

Partition regularity and multiplicatively syndetic sets

by

JONATHAN CHAPMAN (Manchester)

1. Introduction. A system of equations is called *partition regular* if, in any finite colouring of the positive integers $\mathbb{N} = C_1 \cup \dots \cup C_r$, there exists a non-trivial monochromatic solution $\mathbf{x} = (x_1, \dots, x_s)$, meaning that $\mathbf{x} \in C_k^s$ for some k , and $x_i \neq x_j$ for some $i \neq j$. The foundational results in the study of partition regularity are the theorems of Schur [Sch16] and van der Waerden [Wae27]. Schur's theorem states that the equation $x + y = z$ is partition regular, whilst van der Waerden's theorem shows that any finite colouring of \mathbb{N} yields arbitrarily long monochromatic (non-trivial) arithmetic progressions.

The theorems of Schur and van der Waerden are both examples of partition regularity being exhibited by certain linear systems of equations. In particular, these systems are *dilation invariant*, meaning that if $\mathbf{x} = (x_1, \dots, x_s)$ is a solution, then so is $\lambda\mathbf{x} = (\lambda x_1, \dots, \lambda x_s)$ for any $\lambda \in \mathbb{Q}$. In this paper we study the properties of general dilation invariant systems of equations, not just those which are linear. We show that the regularity of such systems is inexorably connected with a special class of sets known as *multiplicatively syndetic sets*.

1.1. Syndeticity. Syndetic sets originate from the study of topological dynamics of semigroups (see [EEN01, HS12]). Given a semigroup (G, \cdot) , a set $S \subseteq G$ is called (*left*)-*syndetic* if there exists a finite set $F \subseteq G$ such that, for each $g \in G$, we have $S \cap (g \cdot F) \neq \emptyset$. Here $g \cdot F := \{gt : t \in F\}$.

The most familiar notion of syndeticity arises in the additive setting where $(G, \cdot) = (\mathbb{N}, +)$. In this case a syndetic subset S is called *additively syndetic* and is just an infinite set with 'bounded gaps'. That is, S is additively syndetic if and only if $S = \{a_1, a_2, \dots\}$ for some infinite

2020 *Mathematics Subject Classification*: Primary 11B30; Secondary 05D10.

Key words and phrases: arithmetic combinatorics, arithmetic Ramsey theory, partition regularity, multiplicatively syndetic sets.

Received 21 April 2019; revised 19 January 2020.

Published online 19 June 2020.

sequence $a_1 < a_2 < \dots$ such that the gaps $|a_{n+1} - a_n|$ are uniformly bounded.

In this paper, we study syndetic sets in the multiplicative semigroup (\mathbb{N}, \cdot) .

DEFINITION (Multiplicatively syndetic set). Let $F \subset \mathbb{N}$ be a non-empty finite set. We say that $S \subseteq \mathbb{N}$ is a *multiplicatively F -syndetic set* if, for every $a \in \mathbb{N}$, we have $S \cap (a \cdot F) \neq \emptyset$.

Multiplicatively syndetic sets have a number of interesting properties. Graham, Witsenhausen, and Spencer [GWS77] observed that multiplicatively syndetic sets have positive density. Much later, Bergelson [Ber10, Lemma 5.11] used methods from ultrafilter theory to show that multiplicatively syndetic sets are *additively central* ⁽¹⁾ (which implies that they have positive density).

The fact that multiplicatively syndetic sets have positive density plays a significant role in the work of Chow, Lindqvist and Prendiville [CLP18]. They demonstrate how multiplicatively syndetic sets can be used to obtain partition regularity results for non-linear equations via an ‘induction on colours’ argument. Their work shows that a sufficient condition for a dilation invariant equation to be partition regular is that it has a solution inside ⁽²⁾ every multiplicatively syndetic set. Our first main theorem is a converse of this result.

THEOREM 1.1 (Partition regularity is equivalent to syndetic solubility). *Let \mathcal{E} be a dilation invariant finite system of equations. Then \mathcal{E} is partition regular if and only if \mathcal{E} has a non-trivial solution inside every multiplicatively syndetic set.*

As an immediate corollary to this theorem, we obtain the following *dilation invariant consistency theorem*.

COROLLARY 1.2 (Dilation invariant consistency theorem). *Let $\mathcal{E}_1, \dots, \mathcal{E}_s$ be s dilation invariant partition regular finite systems of equations. Then in any finite colouring $\mathbb{N} = C_1 \cup \dots \cup C_r$ there exists a colour class C_t such that each \mathcal{E}_i has a solution inside C_t .*

1.2. Brauer configurations. Van der Waerden [Wae27] proved that, for all $r, k \in \mathbb{N}$, there exists a (minimal) $W(r, k) \in \mathbb{N}$ such that, in any r -colouring of $\{1, \dots, W(r, k)\}$, there exists a monochromatic arithmetic progression of length k . Obtaining good bounds for $W(r, k)$ is a notoriously

⁽¹⁾ A subset of \mathbb{N} is called *additively central* if it is a member of a minimal idempotent ultrafilter on $(\mathbb{N}, +)$ (see [Ber10, Definition 5.8]).

⁽²⁾ A system \mathcal{E} is said to have a solution in a set S if there exists a solution to \mathcal{E} with each entry in S .

difficult problem. Over 60 years after van der Waerden’s original paper, Shelah [She88] obtained the first primitive recursive bounds. The best bounds currently known are due to Gowers [Gow01] who obtained the bound

$$(1.1) \quad W(r, k) \leq 2^{2^{r \cdot 2^{2^{k+9}}}}.$$

In §4 we consider a variation of van der Waerden’s theorem concerning configurations of the form

$$\{x, d, x + d, x + 2d\}.$$

These are arithmetic progressions of length 3 along with their common difference. Brauer [Bra28] was the first to establish the partition regularity of these configurations, and so we refer to them as *Brauer configurations (of length 3)*. We also call the corresponding Rado numbers the *(r -colour) Brauer numbers*. Specifically, we define $B(r) \in \mathbb{N}$ to be the smallest positive integer such that every r -colouring of the interval $\{1, \dots, B(r)\}$ yields a monochromatic set of the form $\{x, d, x + d, x + 2d\}$.

To show that Brauer configurations (of length 3) are partition regular, Theorem 1.1 informs us that it is sufficient to prove that all multiplicatively syndetic sets contain such configurations. Our next result establishes a quantitative version of Brauer’s theorem for multiplicatively syndetic sets.

THEOREM 1.3. *There exists a positive absolute constant $c > 0$ such that the following is true. Let $S \subseteq \mathbb{N}$ be a multiplicatively F -syndetic set, for some non-empty finite set $F \subset \mathbb{N}$. Let M denote the largest element of F . If $N \geq 3$ satisfies*

$$M \leq \exp(c\sqrt{\log \log N}),$$

then there exist $d, x \in \mathbb{N}$ such that $\{x, d, x + d, x + 2d\} \subseteq S \cap \{1, \dots, N\}$.

This theorem is analogous to Green and Tao’s result [GT09, Theorem 1.1] that sets $A \subseteq \{1, \dots, N\}$ which lack 4-term arithmetic progressions have size

$$|A| \leq N \exp(-c\sqrt{\log \log N}).$$

We deduce Theorem 1.3 from a more general density result (Theorem 4.1), which concerns dense sets $A \subseteq \{1, \dots, N\}$ lacking arithmetic progressions of length 3 with common difference lying in a given multiplicatively syndetic set S . This density result is proven in §4 by combining the methods of Green and Tao [GT09] with a ‘multiplicatively syndetic induction on colours’ argument of Chow, Lindqvist, and Prendiville [CLP18].

Brauer’s theorem may be proved by iteratively applying van der Waerden’s theorem. As indicated by Cwalina and Schoen [CS17], the best bounds one can obtain for the Brauer numbers by incorporating Gowers’ bound (1.1) into this argument are of the form

$$B(r) \leq \text{tow}((5 + o(1))r).$$

Here $\text{tow}(n)$ denotes an exponential tower of 2's of height n . Explicitly, we take $\text{tow}(1) := 2$ and for all $n \geq 2$ define

$$\text{tow}(n) := 2^{\text{tow}(n-1)}.$$

By incorporating Theorem 1.3 into an induction on colours argument, we obtain an asymptotic improvement on this bound.

THEOREM 1.4 (Tower bound for $B(r)$). *For each $r \in \mathbb{N}$,*

$$(1.2) \quad B(r) \leq \text{tow}((1 + o(1))r).$$

In general, for a partition regular system \mathcal{E} of equations, one can define the r -colour Rado number $R_{\mathcal{E}}(r)$ to be the smallest $N \in \mathbb{N}$ such that every r -colouring of $\{1, \dots, N\}$ yields a monochromatic solution to \mathcal{E} . Cwalina and Schoen [CS17, Theorem 1.5] proved that if \mathcal{E} is a partition regular homogeneous linear equation of the form

$$a_1x_1 + \dots + a_sx_s = 0,$$

where $a_1, \dots, a_s \in \mathbb{Z} \setminus \{0\}$, then

$$R_{\mathcal{E}}(r) \ll_{\mathcal{E}} 2^{O_{\mathcal{E}}(r^4 \log r)}.$$

The improvements obtained by Cwalina and Schoen for single equations ultimately derive from the fact that single linear equations are controlled by the U^2 norm (see §4 for a definition of the U^s norms), and so they can be analysed with (linear) Fourier analysis. However, Brauer configurations of length 3 are controlled by the U^3 norm and therefore require methods from quadratic Fourier analysis. In [CP20] we use higher order Fourier analysis to improve on Theorem 1.4 by obtaining a double exponential bound of the form

$$B(r) \leq \exp \exp(r^C).$$

More generally, we show that a bound of the above form holds for Brauer configurations of any length k (with constant C depending on k).

Notation. The positive integers are denoted by \mathbb{N} . Given $X \geq 1$, we let $[X] := \{n \in \mathbb{N} : 1 \leq n \leq X\} = \{1, \dots, \lfloor X \rfloor\}$.

Let f and g be positive valued functions. We write $f \ll g$, or $g \gg f$, or $f = O(g)$, if there exists a positive constant C such that $f(x) \leq Cg(x)$ for all x . If we require the constant C to depend on some parameters $\lambda_1, \dots, \lambda_k$, then we write $f \ll_{\lambda_1, \dots, \lambda_k} g$ or $f = O_{\lambda_1, \dots, \lambda_k}(g)$.

The letters c and C are typically used to denote absolute constants, whose values may change from line to line. We usually write c to denote a small constant $0 < c < 1$, whereas C usually denotes a large constant $C > 1$.

2. Multiplicative syndeticity and partition regularity. We begin by formally introducing the concepts of partition regularity and multiplicative syndeticity mentioned in the introduction. After establishing the basic properties of multiplicatively syndetic sets, we prove Theorem 1.1 and Corollary 1.2.

2.1. Systems of equations. We consider finite systems \mathcal{E} of polynomial equations in $s \in \mathbb{N}$ variables of the form

$$(2.1) \quad \begin{aligned} p_1(t_1, \dots, t_s) &= 0, \\ &\vdots \\ p_k(t_1, \dots, t_s) &= 0, \end{aligned}$$

where each $p_i \in \mathbb{Q}[t_1, t_2, \dots, t_s]$ is a polynomial in $\{t_i\}_{i=1}^s$. In this paper we only consider systems of finitely many equations, each with finitely many variables. For related results concerning the regularity of infinite systems, see [BHLS15, HLS03].

We usually refer to such a system \mathcal{E} of polynomial equations simply as a *system of equations*. We call $\mathbf{x} \in \mathbb{N}^s$ a *solution* to the system \mathcal{E} if $p_i(\mathbf{x}) = 0$ for all i . A solution $\mathbf{x} = (x_1, \dots, x_s)$ is *non-trivial* if the entries of \mathbf{x} are not all equal. Given a set $S \subseteq \mathbb{Q}$, we say that \mathcal{E} has a [non-trivial] solution in S if there is a [non-trivial] solution $\mathbf{x} = (x_1, \dots, x_s) \in \mathbb{N}^s$ to \mathcal{E} with $x_i \in S$ for all i .

A system \mathcal{E} is called *dilation invariant* if whenever $\mathbf{x} = (x_1, \dots, x_s)$ is a solution to \mathcal{E} , then $\lambda\mathbf{x} = (\lambda x_1, \dots, \lambda x_s)$ is also a solution for every $\lambda \in \mathbb{Q}$. For the majority of this paper, we restrict our attention to dilation invariant systems of polynomial equations. However it should be noted that most of the results we prove in this section apply to any dilation invariant system of equations and not just those consisting of polynomial equations.

2.2. Partition regularity. As mentioned in the introduction, the partition regularity of equations is a well-studied topic in Ramsey theory. Recall that an r -colouring of a set X is a partition $X = C_1 \cup \dots \cup C_r$ of X into r colour classes C_i . Equivalently, an r -colouring can be defined by a function $\chi : X \rightarrow A$ for some set $A = \{a_1, \dots, a_r\}$ with $|A| = r$ (usually we take $A = [r]$). These two characterisations can be seen to be equivalent by taking $\chi^{-1}(a_i) = C_i$. A subset $Y \subseteq X$ is called χ -monochromatic if χ is constant on Y , or equivalently if $Y \subseteq C_i$ for some colour class C_i .

DEFINITION (Partition regularity). Let $S \subseteq \mathbb{Q}$ be a non-empty set and let \mathcal{E} be a system of equations with coefficients in \mathbb{Q} . Let $r \in \mathbb{N}$. We say that \mathcal{E} is (*kernel*) r -regular over S if, for each r -colouring $\chi : \mathbb{Q} \rightarrow [r]$, there exists a χ -monochromatic non-trivial solution \mathbf{x} to \mathcal{E} with entries in S . We

call such an \mathbf{x} a χ -*monochromatic (non-trivial) solution* to A . We say that \mathcal{E} is (*kernel*) *partition regular over S* if \mathcal{E} is r -regular over S for every $r \in \mathbb{N}$.

In practice, when one shows that a given system \mathcal{E} is r -regular, the proof actually yields a number $R_{\mathcal{E}}(r)$ (known as the r *colour Rado number* for \mathcal{E}) such that \mathcal{E} is r -regular over the finite interval $[R_{\mathcal{E}}(r)]$. This is certainly the case whenever one obtains a quantitative regularity result, as in [CS17, Gow01, Sch16, Wae27]. By assuming (some form of) the axiom of choice, one can show that if \mathcal{E} is r -regular, then such an $R_{\mathcal{E}}(r)$ necessarily exists. This result is known as the compactness principle:

COMPACTNESS PRINCIPLE. *Let \mathcal{E} be a finite system of equations in finitely many variables. Let $A \subseteq \mathbb{N}$ and let $r \in \mathbb{N}$. Then \mathcal{E} is r -regular over A if and only if there exists a finite set $F \subseteq A$ such that \mathcal{E} is r -regular over F .*

Proof. See [GRS90, Theorem 4]. ■

REMARK. For the rest of this section, we assume (some form of) the axiom of choice in order to make use of the compactness principle. This assumption is not required for any of the remaining sections.

2.3. Multiplicatively thick sets. The compactness principle informs us that a system \mathcal{E} of equations is partition regular if and only if, for each $r \in \mathbb{N}$, we can find a finite set $F_r \subset \mathbb{N}$ such that \mathcal{E} is r -regular over F_r . Thus, a sufficient condition for \mathcal{E} to be partition regular over a set $A \subseteq \mathbb{N}$ would be that $F_r \subseteq A$ for all $r \in \mathbb{N}$. The problem with this condition is that it is quite possible that the only set which could satisfy this property is $A = \mathbb{N}$. If \mathcal{E} is a dilation invariant system of equations, then we can relax this condition to the requirement that, for each $r \in \mathbb{N}$, we can find $t_r \in \mathbb{N}$ such that $t_r \cdot F_r \subseteq A$. This motivates the following definition.

DEFINITION (Multiplicatively thick set). Let $T \subseteq \mathbb{N}$. We say that T is *multiplicatively thick* if, for each finite set $F \subset \mathbb{N}$, there exists $t \in \mathbb{N}$ such that $t \cdot F \subseteq T$.

PROPOSITION 2.1 (Regularity over thick sets). *Let \mathcal{E} be a dilation invariant system of equations. Let $r \in \mathbb{N}$. Then the following are all equivalent:*

- (I) \mathcal{E} is r -regular;
- (II) \mathcal{E} is r -regular over every multiplicatively thick set;
- (III) \mathcal{E} is r -regular over some multiplicatively thick set.

Proof. The implications (II) \Rightarrow (III) and (III) \Rightarrow (I) are immediate. It only remains to establish (I) \Rightarrow (II).

Suppose \mathcal{E} is r -regular. By compactness, we can find a finite set $F \subset \mathbb{N}$ such that \mathcal{E} is r -regular over F . Now let $T \subseteq \mathbb{N}$ be a multiplicatively thick set. We can then find $t \in \mathbb{N}$ such that $t \cdot F \subseteq T$. Now suppose $\chi : T \rightarrow [r]$ is an r -colouring of T . Define a new r -colouring $\tilde{\chi} : F \rightarrow [r]$ of F by $\tilde{\chi}(x) = \chi(tx)$.

Since \mathcal{E} is r -regular over F , we can find a $\tilde{\chi}$ -monochromatic solution \mathbf{x} to \mathcal{E} in F . By dilation invariance, $t\mathbf{x}$ is a χ -monochromatic solution to \mathcal{E} in T . Thus \mathcal{E} is r -regular over T . ■

2.4. Multiplicatively syndetic sets. We have now reduced regularity over \mathbb{N} to regularity over a multiplicatively thick set. The utility of Proposition 2.1 is demonstrated in the following argument. Suppose that we have a dilation invariant system \mathcal{E} of equations and an integer $r > 1$ such that \mathcal{E} is $(r - 1)$ -regular. We would like to use this to test whether \mathcal{E} is r -regular. Suppose that we have an r -colouring $\mathbb{N} = C_1 \cup \dots \cup C_r$. Informally, if one of the colour classes C_j is ‘small’, we would expect, by $(r - 1)$ -regularity, to find a monochromatic solution in a colour class C_i with $i \neq j$.

To make this rigorous, suppose that we have a colour class C_j whose complement $\mathbb{N} \setminus C_j$ is multiplicatively thick. The remaining $r - 1$ colour classes induce an $(r - 1)$ -colouring on $\mathbb{N} \setminus C_j$. By Proposition 2.1, since \mathcal{E} is $(r - 1)$ -regular, we can find a monochromatic solution to \mathcal{E} inside $\mathbb{N} \setminus C_j$.

This shows that if the dilation invariant system \mathcal{E} is $(r - 1)$ -regular but not r -regular, then there is an r -colouring $\mathbb{N} = C_1 \cup \dots \cup C_r$ without non-trivial monochromatic solutions to \mathcal{E} such that each complement $\mathbb{N} \setminus C_i$ is not multiplicatively thick. Observe that $\mathbb{N} \setminus C_i$ is not multiplicatively thick if and only if there exists a finite set $F \subset \mathbb{N}$ such that, for every $n \in \mathbb{N}$, we have $(n \cdot F) \cap C_i \neq \emptyset$. This motivates the following definition.

DEFINITION (Multiplicatively syndetic set). Let $S \subseteq \mathbb{N}$. Let $F \subset \mathbb{N}$ be a non-empty finite set. We say that S is *multiplicatively F -syndetic* if, for each $n \in \mathbb{N}$, we can find some $t \in F$ such that $nt \in S$. Equivalently, for every $n \in \mathbb{N}$, we have $(n \cdot F) \cap S \neq \emptyset$. We call $S \subseteq \mathbb{N}$ *multiplicatively syndetic* if S is multiplicatively F -syndetic for some non-empty finite set $F \subset \mathbb{N}$.

REMARK. Chow, Lindqvist, and Prendiville [CLP18] define an *M -homogeneous set* to be a set which intersects every homogeneous arithmetic progression $x \cdot [M]$ of length M for every $x \in \mathbb{N}$. We observe that an M -homogeneous set is exactly the same as a multiplicatively $[M]$ -syndetic set.

As mentioned previously, multiplicatively syndetic sets can be equivalently defined in terms of multiplicatively thick sets.

PROPOSITION 2.2. *Let $S \subseteq \mathbb{N}$. Then the following are all equivalent:*

- (I) S is multiplicatively syndetic;
- (II) for every multiplicatively thick set T , we have $S \cap T \neq \emptyset$;
- (III) $\mathbb{N} \setminus S$ is not multiplicatively thick.

Proof. (I) \Rightarrow (II): Suppose S is multiplicatively F -syndetic for some $F \subset \mathbb{N}$, and suppose $T \subseteq \mathbb{N}$ is a multiplicatively thick set. This means that we can

find $t_T \in \mathbb{N}$ such that $t_T \cdot F \subseteq T$. Since S is multiplicatively F -syndetic, we have $(t_T \cdot F) \cap S \neq \emptyset$. In particular, $S \cap T \neq \emptyset$.

(II) \Rightarrow (III): Follows from the fact that S and $\mathbb{N} \setminus S$ are disjoint.

(III) \Rightarrow (I): Since $\mathbb{N} \setminus S$ is not multiplicatively thick, we can find a non-empty finite set $F \subset \mathbb{N}$ such that $t \cdot F \not\subseteq \mathbb{N} \setminus S$ for every $t \in \mathbb{N}$. This implies that S is multiplicatively F -syndetic. ■

In Proposition 2.1 we showed that a dilation invariant system of equations is r -regular if and only if it is r -regular over all multiplicatively thick sets. This is a consequence of the ‘largeness’ of multiplicatively thick sets. We now prove a similar result for multiplicatively syndetic sets. To do this, we identify such sets with finite colourings in the following manner.

DEFINITION (Encoding function). Let $S \subseteq \mathbb{N}$ be a multiplicatively F -syndetic set for some non-empty finite $F \subset \mathbb{N}$. The *encoding function for (S, F)* is the function $\tau_{S;F} : \mathbb{N} \rightarrow F$ defined by

$$\tau_{S;F}(n) := \min\{t \in F : nt \in S\}.$$

Note that the assertion that $\tau_{S;F}$ is a well-defined total function is equivalent to the statement that S is multiplicatively F -syndetic.

The encoding function $\tau_{S;F}$ defines a finite colouring of \mathbb{N} . Moreover, if a set A is monochromatic with respect to this colouring, then there exists $t \in F$ such that $t \cdot A \subseteq S$. This observation leads to the following result.

PROPOSITION 2.3 (Syndetic sets contain PR configurations). *Let $A \subseteq \mathbb{N}$. Let $S \subseteq \mathbb{N}$ be a multiplicatively F -syndetic set for some non-empty finite set $F \subset \mathbb{N}$. Let \mathcal{E} be a dilation invariant system of equations, and let $r \in \mathbb{N}$. If \mathcal{E} is $(|F| \cdot r)$ -regular over A , then \mathcal{E} is r -regular over $S \cap (F \cdot A)$.*

Proof. Suppose $\chi : S \cap (F \cdot A) \rightarrow [r]$ is an r -colouring. Let $\tau = \tau_{S;F}$. Now let $\tilde{\chi} : A \rightarrow F \times [r]$ be the product colouring given by

$$\tilde{\chi}(n) := (\tau(n), \chi(n\tau(n))).$$

Since \mathcal{E} is $(|F| \cdot r)$ -regular over A , we can find a $\tilde{\chi}$ -monochromatic solution \mathbf{a} to \mathcal{E} whose entries a_i all lie in A . From the definition of $\tilde{\chi}$, we can find $t \in F$ such that $\tau(a_i) = t$ for each entry a_i . From the dilation invariance of \mathcal{E} , we deduce that $\mathbf{x} := t\mathbf{a}$ is a χ -monochromatic solution to \mathcal{E} whose entries $x_i = ta_i$ all lie in $S \cap (F \cdot A)$. ■

This proposition immediately gives the following corollary.

COROLLARY 2.4. *Let \mathcal{E} be a dilation invariant system of equations. Then the following are all equivalent:*

- (I) \mathcal{E} is partition regular (over \mathbb{N});
- (II) \mathcal{E} is partition regular over every multiplicatively syndetic set;
- (III) \mathcal{E} is partition regular over some multiplicatively syndetic set $S \subseteq \mathbb{N}$.

We have thus shown that partition regularity over \mathbb{N} is equivalent to partition regularity over a particular multiplicatively syndetic set. Our goal now is to prove Theorem 1.1 and thus show that partition regularity over \mathbb{N} is actually equivalent to 1-regularity over every multiplicatively syndetic set.

Recall that our motivation for introducing multiplicatively syndetic sets came from considering colourings in which some of the colour classes were not multiplicatively thick. This leads to the following induction argument first developed in [CLP18] to establish partition regularity of certain non-linear dilation invariant equations.

LEMMA 2.5 (‘Induction on colours’ schema). *Let \mathcal{E} be a dilation invariant system of equations. If \mathcal{E} is r -regular (for some $r \in \mathbb{N}$), then there exists a finite set $F = F(\mathcal{E}, r) \subset \mathbb{N}$ so that the following holds. If $\mathbb{N} = C_1 \cup \dots \cup C_{r+1}$ is an $(r+1)$ -colouring which lacks monochromatic solutions to \mathcal{E} , then each colour class C_i must be a multiplicatively F -syndetic set.*

Proof. Since \mathcal{E} is r -regular, the compactness principle allows us to find a non-empty finite set $F = F(\mathcal{E}, r) \subset \mathbb{N}$ such that \mathcal{E} is r -regular over F . By dilation invariance, in any colouring χ of \mathbb{N} , if there exists a set of the form $x \cdot F$ (with $x \in \mathbb{N}$) which receives at most r distinct colours, then there exists a χ -monochromatic solution to \mathcal{E} in $x \cdot F$. By contraposition, if $\mathbb{N} = C_1 \cup \dots \cup C_{r+1}$ is an $(r+1)$ -colouring which lacks monochromatic solutions to \mathcal{E} , then each colour class C_i is a multiplicatively F -syndetic set. ■

This lemma shows that when we are trying to prove that a given dilation invariant system \mathcal{E} is partition regular, we only need to consider colourings in which all of the colour classes are multiplicatively syndetic. Combining this with Corollary 2.4 allows us to prove Theorem 1.1.

Proof of Theorem 1.1. If \mathcal{E} is partition regular, then Corollary 2.4 implies that \mathcal{E} is partition regular over every multiplicatively syndetic set. In particular, \mathcal{E} has a non-trivial solution inside every multiplicatively syndetic set.

Conversely, suppose \mathcal{E} is not partition regular. If \mathcal{E} is not 1-regular, then \mathcal{E} has no non-trivial solutions in the multiplicatively syndetic set \mathbb{N} . Suppose then that \mathcal{E} is 1-regular. By Lemma 2.5, there exists a finite colouring of \mathbb{N} with no monochromatic non-trivial solutions to \mathcal{E} and with each colour class being a multiplicatively syndetic set. Therefore each colour class is a multiplicatively syndetic set which has no non-trivial solutions to \mathcal{E} . ■

This result therefore reduces the task of establishing r -regularity over \mathbb{N} for every $r \in \mathbb{N}$ to establishing solubility in every multiplicatively syndetic set. Whilst this may not immediately appear to be helpful, we can obtain Corollary 1.2 very easily from this new approach.

Proof of Corollary 1.2. For each $k \in [s]$, let m_k denote the number of variables appearing in the equations defining the system \mathcal{E}_k . We can therefore define a dilation invariant system \mathcal{E} in $m = m_1 + \dots + m_s$ variables whose solutions are precisely tuples of the form $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(s)})$, where $\mathbf{x}^{(k)} \in \mathbb{Q}^{m_k}$ is a solution to \mathcal{E}_k . Since each \mathcal{E}_i is partition regular, it follows from Theorem 1.1 that \mathcal{E} is partition regular. This implies the desired result. ■

REMARK. When each \mathcal{E}_i is a partition regular linear homogeneous equation, the above result is an immediate consequence of Rado's Criterion [Rad33, Satz IV]. Our proof shows that it is not necessary to utilise such a strong result.

2.5. Multiplicatively piecewise syndetic sets. By using encoding functions, one can show that for a non-empty finite set $F \subset \mathbb{N}$, a set $S \subseteq \mathbb{N}$ is multiplicatively F -syndetic if and only if

$$\mathbb{N} = \bigcup_{t \in F} t^{-1}S,$$

where $t^{-1}S := \{n \in \mathbb{N} : nt \in S\}$. Our proof of Proposition 2.3 used this fact to show that regularity over \mathbb{N} can be 'lifted' to regularity over a multiplicatively syndetic set. However, we proved in Proposition 2.1 that regularity over \mathbb{N} is equivalent to regularity over a multiplicatively thick set. This motivates the introduction of the following weaker form of syndeticity.

DEFINITION (Multiplicatively piecewise syndetic set). Let $F \subset \mathbb{N}$ be a non-empty finite set, and let $S \subseteq \mathbb{N}$. We say that S is *multiplicatively piecewise F -syndetic* if the set $\bigcup_{t \in F} (t^{-1}S)$ is multiplicatively thick. We call $S \subseteq \mathbb{N}$ (*multiplicatively*) *piecewise syndetic* if S is multiplicatively piecewise F -syndetic for some $F \subset \mathbb{N}$.

Another way to view multiplicatively piecewise syndetic sets is through the following 'partial encoding' formulation. Given a non-empty finite set $F \subset \mathbb{N}$ and a set $S \subseteq \mathbb{N}$, define a partial function ⁽³⁾ $\tau_{S;F} : \mathbb{N} \dashrightarrow F$ by

$$(2.2) \quad \tau_{S;F}(n) := \min\{t \in F : nt \in S\},$$

for all $n \in \mathbb{N}$ for which the above quantity is defined. We refer to this partial function as the (*partial*) *encoding function for (S, F)* . By the *domain of $\tau_{S;F}$* we mean the set of all $n \in \mathbb{N}$ for which (2.2) is defined.

We remarked earlier that S is multiplicatively F -syndetic if and only if $\tau_{S;F}$ is a total function, meaning that $\tau_{S;F}(n)$ is defined for all $n \in \mathbb{N}$. Similarly, we see that S is multiplicatively piecewise F -syndetic if and only if the domain of $\tau_{S;F}$ is multiplicatively thick.

⁽³⁾ We use the partial function notation $f : A \dashrightarrow B$ to mean that f only defines a function on a (possibly empty) subset of A .

This technique of identifying a multiplicatively syndetic set with its encoding function was the key idea in the proof of Proposition 2.3. A similar argument can be used to obtain the following analogous result.

PROPOSITION 2.6 (Piecewise syndetic sets contain PR configurations). *Let $S \subseteq \mathbb{N}$ be a multiplicatively piecewise F -syndetic set for some non-empty finite set $F \subset \mathbb{N}$. Let \mathcal{E} be a dilation invariant system of equations, and let $r \in \mathbb{N}$. If \mathcal{E} is $(|F| \cdot r)$ -regular over \mathbb{N} , then \mathcal{E} is r -regular over S .*

Proof. Suppose $\chi : S \rightarrow [r]$ is an r -colouring of S . Since S is multiplicatively piecewise syndetic, the set $T := \cup_{t \in F} (t^{-1} \cdot S)$ is multiplicatively thick. Let $\tau : T \rightarrow F$ denote the encoding function given by

$$\tau(n) := \min\{t \in F : nt \in S\}.$$

Now let $\tilde{\chi} : T \rightarrow F \times [r]$ be the product colouring given by

$$\tilde{\chi}(n) := (\tau(n), \chi(n\tau(n))).$$

By Proposition 2.1, \mathcal{E} is $(|F| \cdot r)$ -regular over T . Thus, we can find a $\tilde{\chi}$ -monochromatic solution \mathbf{x} to \mathcal{E} such that every entry of \mathbf{x} is in T . From the definition of $\tilde{\chi}$, we can find $t \in F$ such that $\tau(x_i) = t$ for every entry x_i of \mathbf{x} . The dilation invariance of \mathcal{E} then shows that $\mathbf{y} := t\mathbf{x}$ is a χ -monochromatic solution to \mathcal{E} whose entries $y_i = tx_i$ all lie in S . ■

We end this section by synthesising all of our results relating partition regularity with solubility in multiplicatively syndetic sets into the following theorem.

THEOREM 2.7 (Summary of results). *Suppose \mathcal{E} is a dilation invariant system of equations. Then the following are all equivalent:*

- (I) \mathcal{E} is partition regular (over \mathbb{N});
- (II) \mathcal{E} is partition regular over every multiplicatively thick set;
- (III) \mathcal{E} is partition regular over every multiplicatively piecewise syndetic set;
- (IV) \mathcal{E} is partition regular over every multiplicatively syndetic set;
- (V) \mathcal{E} has a solution in every multiplicatively piecewise syndetic set;
- (VI) \mathcal{E} has a solution in every multiplicatively syndetic set.

Proof. Proposition 2.1 establishes the equivalence (I) \Leftrightarrow (II). Proposition 2.3 and Theorem 1.1 together show that (I) \Leftrightarrow (IV) \Leftrightarrow (VI). Similarly, we deduce from Proposition 2.6 that (I) \Leftrightarrow (III). Since syndeticity implies piecewise syndeticity, we see that (V) \Rightarrow (VI). Finally, since solubility is equivalent to 1-regularity, we observe that (III) \Rightarrow (V). ■

3. Density of multiplicatively syndetic sets. In Ramsey theory, there are multiple concepts of ‘largeness’. The most familiar of these is the notion of (asymptotic) density.

DEFINITION (Asymptotic density). Let $A \subseteq \mathbb{N}$. The *upper (asymptotic) density* $\bar{d}(A)$ of A is defined by

$$\bar{d}(A) := \limsup_{N \rightarrow \infty} \frac{|A \cap [N]|}{N}.$$

The *lower (asymptotic) density* $\underline{d}(A)$ of A is defined by

$$\underline{d}(A) := \liminf_{N \rightarrow \infty} \frac{|A \cap [N]|}{N}.$$

The *natural (asymptotic) density* $d(A)$ of A is defined by

$$d(A) := \lim_{N \rightarrow \infty} \frac{|A \cap [N]|}{N},$$

whenever the above limit exists, which occurs if and only if $\bar{d}(A) = \underline{d}(A)$.

One can generalise this definition to obtain a notion of asymptotic density for general cancellative, left amenable semigroups (see [BG18] for further details). The above definition comes from the case where the semigroup is $(\mathbb{N}, +)$. As such, asymptotic density is a form of ‘additive largeness’.

Recent research has led to the surprising discovery that multiplicatively large sets, such as multiplicatively syndetic sets, are additively large. Bergelson [Ber10, Lemma 5.11] proved that multiplicatively syndetic sets are additively central, which implies that they have positive upper asymptotic density. However, due to the infinitary nature of central sets, no explicit bounds on the density of multiplicatively syndetic sets can be extracted from this result.

In their work on the partition regularity of non-linear equations, Chow, Lindqvist, and Prendiville [CLP18, Lemma 4.2] independently proved that multiplicatively syndetic sets have positive (lower) asymptotic density. They obtained the following quantitative result (for the case $F = [M]$).

LEMMA 3.1. *Let $F \subset \mathbb{N}$ be a non-empty finite set, and let M denote the largest element of F . Then for any $N \in \mathbb{N}$ and any multiplicatively F -syndetic set $S \subseteq [N]$, we have*

$$(3.1) \quad |S \cap [N]| \geq \frac{1}{|F|} \left\lfloor \frac{N}{M} \right\rfloor.$$

Proof. Define an encoding function $\tau : [N/M] \rightarrow F$ for S by

$$\tau(x) := \min\{t \in F : tx \in S\}.$$

By the pigeonhole principle, there exists $t \in F$ with $|\tau^{-1}(t)| \geq \frac{1}{|F|} \lfloor [N/M] \rfloor$. Thus,

$$|S| \geq |\{tx : x \in \tau^{-1}(t)\}| \geq \frac{1}{|F|} \left\lfloor \frac{N}{M} \right\rfloor. \blacksquare$$

In fact, the density of multiplicatively syndetic sets had been studied much earlier. In 1977 Graham, Witsenhausen, and Spencer [GWS77] determined the maximum asymptotic density for sets lacking linear configurations of the form $\{a_1x, \dots, a_sx\}$. Taking complements enables one to determine the minimum density of a multiplicatively F -syndetic set for $F = \{a_1, \dots, a_s\}$. After this reformulation, their result is as follows.

THEOREM 3.2 ([GWS77, Theorem 2]). *Let $F \subset \mathbb{N}$ be a non-empty finite set. Let P_F be the set of primes dividing elements of F . Let $S(P_F)$ denote the set of all P_F -smooth numbers, meaning that $x \in S(P_F)$ if and only if every prime factor of x lies in P_F . Write $S(P_F) = \{d_1 < d_2 < \dots\}$, where d_k is the k th smallest element of $S(P_F)$. For each $k \in \mathbb{N}$, let*

$$g_F(k) := \min\{|X \cap \{d_1, \dots, d_k\}| : X \subseteq \mathbb{N} \text{ is multiplicatively } F\text{-syndetic}\}.$$

Finally, let $K(F) = \{k \in \mathbb{N} : g_F(k) = g_F(k-1)\}$. Then for any multiplicatively F -syndetic set $S \subseteq \mathbb{N}$, we have the sharp bound

$$\underline{d}(S) \geq \delta_{\min}(F) := 1 - \prod_{p \in S(P_F)} (1 - p^{-1}) \sum_{k \in K(F)} d_k^{-1}.$$

Graham, Witsenhausen, and Spencer remark that there are difficulties in evaluating $\delta_{\min}(F)$ due to the complicated nature of the set $K(F)$. In particular, they could not obtain an explicit evaluation for $\delta_{\min}(F)$ in the case where $F = \{1, 2, 3\}$. Erdős and Graham [EG80] subsequently conjectured that $\delta_{\min}(\{1, 2, 3\})$ is irrational. This conjecture remains open (see [CEG02] for further details and developments related to this problem).

In the case where $F = \{1, p, p^2, \dots, p^{k-1}\}$ for some prime p and some $k \in \mathbb{N}$, Graham, Witsenhausen, and Spencer observed that $\delta_{\min}(F) = \frac{p+1}{p^k-1}$. We now consider $F = \{1, a, a^2, \dots, a^{k-1}\}$, where $a \in \mathbb{N}$ need not be prime, and explicitly construct a multiplicatively F -syndetic set of minimum density.

DEFINITION (Multiplicity function). Let $a \geq 2$ be a positive integer. Define the a -multiplicity function $\nu_a : \mathbb{N} \rightarrow \mathbb{N} \cup \{0\}$ by

$$\nu_a(n) := \max\{k \in \mathbb{N} \cup \{0\} : a^k \mid n\}.$$

LEMMA 3.3 (The set $S(a, k)$). *Let $a, k \in \mathbb{N} \setminus \{1\}$, and let*

$$F = \{1, a, a^2, \dots, a^{k-1}\}.$$

Let $S(a, k) \subseteq \mathbb{N}$ be defined by

$$(3.2) \quad S(a, k) := \{n \in \mathbb{N} : \nu_a(n) \equiv k-1 \pmod{k}\}.$$

Then $S(a, k)$ is a multiplicatively F -syndetic set and has natural density

$$\underline{d}(S(a, k)) = \frac{a-1}{a^k-1}.$$

Proof. By noting that $\nu_a(an) = \nu_a(n) + 1$, we see that, for any $n \in \mathbb{N}$, the set $\nu_a(n \cdot F)$ is a complete residue system modulo k . Thus, $S(a, k)$ is a multiplicatively F -syndetic set.

It only remains to check that $S(a, k)$ achieves the required density bound. For each $m \in \mathbb{N}$, let $A_m := \{n \in \mathbb{N} : \nu_a(n) = km - 1\}$. Hence,

$$S(a, k) = \bigcup_{m \in \mathbb{N}} A_m.$$

Observe that $n \in A_m$ if and only if $n \equiv a^{km-1}b \pmod{a^{km}}$ for some $b \in \{1, \dots, a-1\}$. We therefore deduce that A_m has natural density

$$d(A_m) = \frac{a-1}{a^{km}}.$$

For each $r \in \mathbb{N}$, let $B_r = \bigcup_{m=1}^r A_m$. Since A_i and A_j are disjoint for all $i \neq j$, we deduce from the finite additivity of natural density that

$$d(B_r) = \sum_{m=1}^r d(A_m) = (a-1) \sum_{m=1}^r a^{-km} = (a-1) \frac{1-a^{-rk}}{a^k-1}.$$

By noting that $B_r \subseteq S(a, k)$, we deduce that $\underline{d}(S(a, k)) \geq d(B_r)$ for all $r \in \mathbb{N}$. Taking $r \rightarrow \infty$ gives the lower bound $\underline{d}(S(a, k)) \geq (a-1)/(a^k-1)$.

We now compute an upper bound. First observe that

$$S(a, k) \setminus B_r \subseteq a^{kr-1} \cdot \mathbb{N}$$

for all $r \in \mathbb{N}$. Thus,

$$\bar{d}(S(a, k)) \leq \bar{d}(B_r) + \bar{d}(a^{kr-1} \cdot \mathbb{N}) = d(B_r) + \frac{1}{a^{kr-1}}$$

for all $r \in \mathbb{N}$. Taking $r \rightarrow \infty$ gives the desired upper bound. ■

We now show that $S(a, k)$ has minimal density.

THEOREM 3.4 (Minimal $\{1, a, \dots, a^{k-1}\}$ -syndetic set). *Let $a, k \in \mathbb{N} \setminus \{1\}$, and let $F = \{1, a, a^2, \dots, a^{k-1}\}$. Let $S(a, k)$ be the set defined in (3.2). Let $N \in \mathbb{N}$. If $X \subseteq \mathbb{N}$ is a multiplicatively F -syndetic set, then*

$$|X \cap [N]| \geq |S(a, k) \cap [N]|.$$

Proof. We may assume that $N \geq a^{k-1}$, since otherwise $S(a, k) \cap [N] = \emptyset$ and the result is vacuously true. Let $X \subseteq \mathbb{N}$ be a multiplicatively F -syndetic set. Let $m \in S(a, k) \cap [N]$. Since every element of $S(a, k)$ is divisible by a^{k-1} , we deduce that $(a^{-(k-1)}m) \cdot F \subseteq [N]$. As X is multiplicatively F -syndetic, we can find some $t(m) \in F$ such that $a^{-(k-1)}mt(m) \in X$. We can therefore define a function $g : S(a, k) \cap [N] \rightarrow X$ by $g(n) = a^{-(k-1)}nt(n)$. To complete the proof it is sufficient to show that g is injective.

Suppose that $n, n' \in S(a, k)$ satisfy $g(n) = g(n')$. Thus, $nt(n) = n't(n')$. Applying ν_a to this equation and then reducing modulo k gives

$$\nu_a(t(n)) \equiv \nu_a(t(n')) \pmod{k}.$$

Now observe that the function which maps $t \in F$ to the residue class of $\nu_a(t)$ modulo k is injective. This shows that $t(n) = t(n')$, hence $n = n'$. ■

4. A syndetic density increment strategy. Brauer [Bra28] established the following common generalisation of Schur's theorem and van der Waerden's theorem.

BRAUER'S THEOREM. *For all $k, r \in \mathbb{N}$, there exists $N_0 = N_0(k, r) \in \mathbb{N}$ such that, in any r -colouring $[N_0] = C_1 \cup \dots \cup C_r$ of $[N_0] = \{1, \dots, N_0\}$, there is a colour class C_i containing a set of the form*

$$(4.1) \quad \{x, d, x + d, x + 2d, \dots, x + (k - 1)d\}.$$

In this section we study the case where $k = 3$. This corresponds to configurations of the form

$$(4.2) \quad \{x, d, x + d, x + 2d\}.$$

We refer to such configurations as *Brauer configurations (of length 3)*. These are 3-term arithmetic progressions along with their common difference. Alternatively, one can view Brauer configurations as solutions $\{x, y, z, d\}$ to the dilation invariant system

$$\begin{aligned} x - 2y + z &= 0, \\ x - y + d &= 0. \end{aligned}$$

As shown by Theorem 1.1, Brauer's theorem is equivalent to the assertion that all multiplicatively syndetic sets contain Brauer configurations. In this section, we derive quantitative bounds on the minimal $N \in \mathbb{N}$ for which the set $S \cap [N]$ must contain a Brauer configuration, for a given multiplicatively syndetic set S . The main theorem of this section is as follows.

THEOREM 4.1 (Syndetic Brauer). *Let $M \geq 2$ be a positive integer. Let $A \subseteq [N]$ be such that $|A| \geq \delta N$ for some $0 < \delta \leq 1/2$. Let $S \subseteq \mathbb{N}$ be a multiplicatively $[M]$ -syndetic set. If there does not exist a 3-term arithmetic progression in A with common difference in S , then*

$$(4.3) \quad \log \log N \ll \log(\delta^{-1}) \log(M/\delta).$$

REMARK. We impose the restriction $\delta \leq 1/2$ to ensure that δ is bounded away from 1. One could replace $1/2$ by any positive quantity strictly less than 1 at the cost of increasing the implicit constant in (4.3).

In [GT09], Green and Tao used a density increment strategy over quadratic factors to obtain new bounds for the sizes of subsets of $[N]$ lacking 4-term arithmetic progressions. In this section we combine their methods with the induction on colours argument (Lemma 2.5) of Chow, Lindqvist, and Prendiville [CLP18] to prove Theorem 4.1.

By incorporating the density bounds obtained in §3, we can use Theorem 4.1 to prove Theorem 1.3.

Proof of Theorem 1.3 given Theorem 4.1. Let $c > 0$ be a small positive constant to be specified later, and assume that $M, N \in \mathbb{N}$ satisfy

$$2 \leq M \leq \exp(c\sqrt{\log \log N}).$$

Let $S \subseteq \mathbb{N}$ be a multiplicatively F -syndetic set for some non-empty $F \subseteq [M]$. By taking c sufficiently small, we may assume that $N \geq M$. This implies that the density of $S \cap [N]$ in $[N]$ is positive. Moreover, by Lemma 3.1, we can choose a subset $S' \subseteq S$ such that the density δ of $S' \cap [N]$ in $[N]$ satisfies both $0 < \delta \leq 1/2$ and $\delta^{-1} \ll M^2$. This gives

$$\log(\delta^{-1}) \log(M/\delta) \ll \log^2 M.$$

Hence, by choosing c to be sufficiently small, we can ensure that (4.3) does not hold. Thus, by taking $A = S' \cap [N]$, we conclude from Theorem 4.1 that $S \cap [N]$ contains a Brauer configuration. ■

For each $r \in \mathbb{N}$, define the r -colour Brauer number $B(r)$ to be the minimum $N \in \mathbb{N}$ such that any r -colouring of $[N]$ yields a monochromatic Brauer configuration. In other words, $B(r)$ is the minimal value that $N_0(3, r)$ may take in the statement of Brauer's theorem. We can use Theorem 4.1 to obtain the following recursive bound for these numbers.

THEOREM 4.2 (Recursive bound for $B(r)$). *For each $r \in \mathbb{N}$, we have*

$$(4.4) \quad B(r+1) \leq 2^{B(r)O(\log(r+1))}.$$

By some computation (which is given in the Appendix), this theorem leads to the tower type bound in Theorem 1.4.

Proof of Theorem 4.2 given Theorem 4.1. Let $r \in \mathbb{N}$. Let $M := B(r)$ and let $\delta := (r+1)^{-1}$. Note that $M \geq B(1) = 3$. Suppose $N \in \mathbb{N}$ is such that $N < B(r+1)$. Therefore, we have an $(r+1)$ -colouring $[N] = C_1 \cup \dots \cup C_{r+1}$ with no monochromatic sets of the form (4.2). By dilation invariance and our choice of M , it follows that there cannot exist a set of the form $x \cdot [M] \subseteq [N]$ which is r -coloured. This implies that $C_i \cup (\mathbb{N} \setminus [N])$ is multiplicatively $[M]$ -syndetic for all $i \in [r]$.

Without loss of generality, assume that C_1 is the largest colour class. By the pigeonhole principle, $|C_1| \geq \delta N$. Hence, by taking $A = C_1$ and

$S = C_1 \cup (\mathbb{N} \setminus [N])$ in the statement of Theorem 4.1, we deduce that

$$\log \log B(r+1) \ll \log(r+1) \log((r+1)B(r)).$$

By noting that $B(r) \geq r+1$ (since one can r -colour $[r]$ so that each element has a unique colour), this gives

$$\log \log B(r+1) \ll \log(r+1) \log B(r).$$

Exponentiating twice then gives (4.4). ■

4.1. Norms. Given a non-empty finite set A and a function $f : A \rightarrow \mathbb{C}$, we define the expectation $\mathbb{E}_A(f)$ of f over A by

$$\mathbb{E}_A(f) = \mathbb{E}_{x \in A} f(x) := \frac{1}{|A|} \sum_{x \in A} f(x).$$

The functions we encounter in this section are usually defined on $[N]$ or $\mathbb{Z}/p\mathbb{Z}$, where $N \in \mathbb{N}$, and $p \in \mathbb{N}$ is a prime. When $p > N$, it is convenient to consider $[N]$ as a subset of the field $\mathbb{Z}/p\mathbb{Z}$ by reducing modulo p . Given a function $f : \mathbb{Z} \rightarrow \mathbb{C}$ which is supported on $[N]$, we can then consider f as a function defined on $\mathbb{Z}/p\mathbb{Z}$ by taking $f(x) = 0$ for all $x \in (\mathbb{Z}/p\mathbb{Z}) \setminus [N]$.

We make use of two different types of norms. The standard L^p norms are used to measure the overall size of a function, whilst the Gowers uniformity U^s norms (introduced in [Gow01, Lemma 3.9]) measure the degree to which a function exhibits non-uniformity.

DEFINITION (L^p norms). Let A be a set and let $f : A \rightarrow \mathbb{C}$ be a finitely supported function. The L^1 norm $\|f\|_{L^1(A)}$ of f is defined by

$$\|f\|_{L^1(A)} := \mathbb{E}_{x \in A} |f(x)|.$$

The L^∞ norm $\|f\|_{L^\infty(A)}$ of f is defined by

$$\|f\|_{L^\infty(A)} := \max_{x \in A} |f(x)|.$$

We say that f is *1-bounded* (on A) if $\|f\|_{L^\infty(A)} \leq 1$.

DEFINITION (U^s norms). Let $f : \mathbb{Z}/p\mathbb{Z} \rightarrow \mathbb{C}$. For each $s \geq 1$, the U^s norm ⁽⁴⁾ $\|f\|_{U^s(\mathbb{Z}/p\mathbb{Z})}$ of f is defined by

$$(4.5) \quad \|f\|_{U^s(\mathbb{Z}/p\mathbb{Z})} := \left(\mathbb{E}_{x, h_1, \dots, h_s \in \mathbb{Z}/p\mathbb{Z}} \Delta_{h_1, \dots, h_s} f(x) \right)^{1/2^s},$$

where the *difference operators* Δ_{h_1, \dots, h_s} are defined by

$$\Delta_h f(x) := f(x) \overline{f(x+h)} \quad \text{and} \quad \Delta_{h_1, \dots, h_s} f := \Delta_{h_1} \cdots \Delta_{h_s} f.$$

⁽⁴⁾ For $s \geq 2$ the U^s norms are indeed norms (see [TV06, Chapter 11]), but the U^1 ‘norm’ is only a seminorm.

The Gowers uniformity norms can also be defined recursively. By expanding and rearranging (4.5), we observe that

$$(4.6) \quad \|f\|_{U^{s+1}}^{2^{s+1}} = \mathbb{E}_{h \in \mathbb{Z}/p\mathbb{Z}} \|\Delta_h f\|_{U^s}^{2^s}.$$

4.2. Counting Brauer configurations. To prove Theorem 4.2, we use an induction on colours argument similar to [CLP18, §4].

For the remainder of this section, we let N denote a positive integer and consider Brauer configurations in the interval $[N]$. It is useful to embed $[N]$ in an abelian group which is not much larger than $[N]$. We therefore let p denote a prime ⁽⁵⁾ satisfying $3N < p < 6N$, and embed $[N]$ in the group $\mathbb{Z}/p\mathbb{Z}$ by reducing modulo p . We also let $M \in \mathbb{N}$ be a positive integer with $M > 1$ so that we may consider multiplicatively $[M]$ -syndetic sets.

Let $S \subseteq \mathbb{Z}/p\mathbb{Z}$ and let $f_1, f_2, f_3 : \mathbb{Z}/p\mathbb{Z} \rightarrow \mathbb{C}$. The counting functional A_S we use is defined by

$$A_S(f_1, f_2, f_3) := \mathbb{E}_{x \in \mathbb{Z}/p\mathbb{Z}} \mathbb{E}_{d \in S} f_1(x) f_2(x+d) f_3(x+2d).$$

For brevity, we write $A_S(f) := A_S(f, f, f)$.

LEMMA 4.3 (Counting Brauer configurations with common difference in S). *Let $N \in \mathbb{N}$ with $N \geq 3$. If $S \subseteq [N/3]$ is non-empty, then*

$$A_S(1_{[N]}) > 1/18.$$

Proof. Since $S \subseteq [N/3]$, for all $d \in S$ we have $N - 2d \geq N/3$. Combining this with the bound $p < 6N$ gives

$$A_S(1_{[N]}) = \frac{1}{p} \mathbb{E}_{d \in S} (N - 2d) \geq \frac{N}{3p} > \frac{1}{18}. \quad \blacksquare$$

We now use the two different types of norms introduced earlier to control the size of A_S . The simplest way to bound A_S is by using the L^1 norm.

LEMMA 4.4 (L^1 control for A_S). *Let $S \subseteq [N]$ and let $f, g : \mathbb{Z}/p\mathbb{Z} \rightarrow \mathbb{C}$ be 1-bounded functions. Then*

$$(4.7) \quad |A_S(f) - A_S(g)| \leq 3\|f - g\|_{L^1(\mathbb{Z}/p\mathbb{Z})}.$$

Proof. First let $k \in \{1, 2, 3\}$ and let $f_1, f_2, f_3 : \mathbb{Z}/p\mathbb{Z} \rightarrow \mathbb{C}$ be functions such that f_i is 1-bounded for all $i \neq k$. By a change of variables, we see that

$$\begin{aligned} |A_S(f_1, f_2, f_3)| &\leq \mathbb{E}_{d \in S} \mathbb{E}_{x \in \mathbb{Z}/p\mathbb{Z}} |f_1(x) f_2(x+d) f_3(x+2d)| \\ &\leq \mathbb{E}_{d \in S} \mathbb{E}_{x \in \mathbb{Z}/p\mathbb{Z}} |f_k(x + (k-1)d) 1_S(d)| \\ &= \left(\mathbb{E}_{y \in \mathbb{Z}/p\mathbb{Z}} |f_k(y)| \right) \left(\mathbb{E}_{d \in S} 1_S(d) \right). \end{aligned}$$

⁽⁵⁾ Such a prime exists by Bertrand's postulate.

We therefore deduce that

$$(4.8) \quad |A_S(f_1, f_2, f_3)| \leq \|f_k\|_{L^1(\mathbb{Z}/p\mathbb{Z})}$$

for all $k \in \{1, 2, 3\}$.

Now observe that, by multilinearity, we have the telescoping identity

$$(4.9) \quad A_S(f) - A_S(g) = A_S(f - g, f, f) + A_S(g, f - g, f) \\ + A_S(g, g, f - g).$$

Applying the triangle inequality and using (4.8) gives (4.7). ■

In addition to A_S , for functions $f_1, f_2, f_3 : \mathbb{Z}/p\mathbb{Z} \rightarrow \mathbb{C}$, we introduce the auxiliary counting functional AP_3 given by

$$AP_3(f_1, f_2, f_3) := \mathbb{E}_{x, d \in \mathbb{Z}/p\mathbb{Z}} f_1(x) f_2(x + d) f_3(x + 2d).$$

Since Brauer configurations contain 3-term arithmetic progressions, it is perhaps unsurprising that the uniformity of A_S is related to the uniformity of AP_3 . Indeed, the original motivation for the introduction of the U^s norms in [Gow01, §3] was the observation that they control counting functionals for arithmetic progressions. This result is referred to in the literature as a *generalised von Neumann theorem*. In the case of 3-term arithmetic progressions, the result is as follows.

LEMMA 4.5 (Generalised von Neumann theorem). *Let $f_1, f_2, f_3 : \mathbb{Z}/p\mathbb{Z} \rightarrow \mathbb{C}$ be 1-bounded functions. Then*

$$(4.10) \quad |AP_3(f_1, f_2, f_3)| \leq \min_{1 \leq k \leq 3} \|f_k\|_{U^2(\mathbb{Z}/p\mathbb{Z})}.$$

Proof. This follows from two applications of the Cauchy–Schwarz inequality. For the full details, see [TV06, Lemma 11.4]. ■

We now prove an analogous result for the A_S functional.

LEMMA 4.6 (Generalised von Neumann theorem for A_S). *Let $S \subseteq [N]$ be a non-empty set, and let $f_1, f_2, f_3 : \mathbb{Z}/p\mathbb{Z} \rightarrow \mathbb{C}$ be 1-bounded functions. Then*

$$|A_S(f_1, f_2, f_3)| \leq \frac{p^{1/2}}{|S|^{1/2}} \min_{1 \leq k \leq 3} \|f_k\|_{U^3(\mathbb{Z}/p\mathbb{Z})}.$$

Proof. Observe that we can rewrite $A_S(f_1, f_2, f_3)$ as

$$A_S(f_1, f_2, f_3) = \frac{p}{|S|} \mathbb{E}_{x, d \in \mathbb{Z}/p\mathbb{Z}} f_1(x) f_2(x + d) f_3(x + 2d) 1_S(d).$$

By applying the Cauchy–Schwarz inequality (with respect to the d variable),

we see that the quantity $|A_S(f_1, f_2, f_3)|^2$ is bounded above by

$$\begin{aligned}
& \left(\mathbb{E}_{d \in \mathbb{Z}/p\mathbb{Z}} \frac{p^2}{|S|^2} 1_S(d) \right) \left(\mathbb{E}_{d \in \mathbb{Z}/p\mathbb{Z}} \left| \mathbb{E}_{x, d \in \mathbb{Z}/p\mathbb{Z}} f_1(x) f_2(x+d) f_3(x+2d) \right|^2 \right) \\
&= \frac{p}{|S|} \mathbb{E}_{d, x, x' \in \mathbb{Z}/p\mathbb{Z}} f_1(x) f_1(x') f_2(x+d) f_2(x'+d) f_3(x+2d) f_3(x'+2d) \\
&= \frac{p}{|S|} \mathbb{E}_{d, h, x \in \mathbb{Z}/p\mathbb{Z}} \Delta_h f_1(x) \Delta_h f_2(x+d) \Delta_h f_3(x+2d) \\
&= \frac{p}{|S|} \mathbb{E}_{h \in \mathbb{Z}/p\mathbb{Z}} \text{AP}_3(\Delta_h f_1, \Delta_h f_2, \Delta_h f_3).
\end{aligned}$$

Note that the penultimate equality above follows from a change of variables of the form $h = x' - x$. Now let $k \in \{1, 2, 3\}$. Using Lemma 4.5 and (4.6) along with an application of Hölder's inequality gives

$$\begin{aligned}
|A_S(f_1, f_2, f_3)|^2 &\leq \frac{p}{|S|} \mathbb{E}_{h \in \mathbb{Z}/p\mathbb{Z}} \|\Delta_h f_k\|_{U^2(\mathbb{Z}/p\mathbb{Z})} \\
&\leq \frac{p}{|S|} \left(\mathbb{E}_{h \in \mathbb{Z}/p\mathbb{Z}} 1^{4/3} \right)^{3/4} \left(\mathbb{E}_{h \in \mathbb{Z}/p\mathbb{Z}} \|\Delta_h f_k\|_{U^2(\mathbb{Z}/p\mathbb{Z})}^4 \right)^{1/4} \\
&= \frac{p}{|S|} \|f_k\|_{U^3(\mathbb{Z}/p\mathbb{Z})}^2. \blacksquare
\end{aligned}$$

LEMMA 4.7 (U^3 controls A_S). *Let $S \subseteq \mathbb{N}$ be a multiplicatively $[M]$ -syndetic set, and let $f, g : \mathbb{Z}/p\mathbb{Z} \rightarrow [0, 1]$. If $N \geq 18M^2$, then*

$$(4.11) \quad |A_{S \cap [N/3]}(f) - A_{S \cap [N/3]}(g)| \leq 18M \|f - g\|_{U^3(\mathbb{Z}/p\mathbb{Z})}.$$

Proof. Note that as f and g are non-negative, the difference $f - g$ is a 1-bounded function. Applying the previous lemma and the triangle inequality to the telescoping identity (4.9) then gives

$$|A_{S \cap [N/3]}(f) - A_{S \cap [N/3]}(g)| \leq \frac{3p^{1/2}}{|S \cap [N/3]|^{1/2}} \|f - g\|_{U^3(\mathbb{Z}/p\mathbb{Z})}.$$

Combining Lemma 3.1 with the assumption $N \geq 18M^2$ gives

$$|S \cap [N/3]| \geq \frac{N}{6M^2} > \frac{p}{36M^2}.$$

This implies the desired bound (4.11). \blacksquare

We now proceed to prove Theorem 4.1 using a density increment strategy. This argument combines Green and Tao's quadratic Fourier-analytic methods [GT09] for finding sets lacking arithmetic progressions of length 4 with the techniques used by Chow, Lindqvist, and Prendiville [CLP18, Lemma 7.1] to obtain a 'homogeneous' generalisation of Sárközy's theorem [Sár78].

The original density increment strategy of Roth [Rot53] was used to show that subsets of $[N]$ which lack arithmetic progressions of length 3 have size $o(N)$. This method was subsequently modified by Gowers [Gow98] to

prove an analogous result for arithmetic progressions of length 4, and then further generalised to progressions of arbitrary length [Gow01]. The argument proceeds as follows. Let $\delta_0 > 0$. Suppose $A \subseteq [N]$ lacks arithmetic progressions of length 3 and satisfies $|A| = \alpha N$ for some $\alpha \geq \delta_0$. Then, provided that N is ‘not too small’, meaning that $N > C(\delta_0)$ for some positive constant $C(\delta_0)$ depending only on δ_0 , we can find an arithmetic progression $P \subseteq [N]$ of length $N' := |P| \geq F(N, \delta_0)$ on which A has a density increment

$$(4.12) \quad \alpha' := \frac{|A \cap P|}{|P|} \geq \alpha + c(\delta_0).$$

Here $c(\delta_0) > 0$ is a constant depending only on δ_0 , and F is an explicit positive function such that, for any fixed $\delta > 0$, $F(N, \delta) \rightarrow \infty$ as $N \rightarrow \infty$.

We can then apply an affine transformation of the form $x \mapsto ax + b$ to injectively map $A \cap P$ into $[N']$ with image $A' \subseteq [N']$. Since arithmetic progressions are translation-dilation invariant, A' also lacks arithmetic progressions of length 3 and satisfies $|A'| = \alpha' N' > \delta_0 N'$. We can then iterate this argument. Since the density is increasing by at least $c(\delta_0)$ after each iteration, this process must eventually terminate. We can then procure an upper bound for the size of the original N in terms of $C(\delta_0)$, $c(\delta_0)$ and $F(\cdot, \delta_0)$.

An important aspect of this method is that it uses the translation-dilation invariance of arithmetic progressions. However, more general configurations, such as Brauer configurations, are not translation invariant. This is emphasised by the fact that the odd numbers have density $1/2$ and yet they do not contain any Brauer configurations. A density analogue of Brauer’s theorem is therefore impossible, and so this argument cannot help us prove Brauer’s theorem.

The major insight of Chow, Lindqvist, and Prendiville [CLP18, §7] is that one can separate such configurations into a ‘translation invariant part’ and a ‘non-translation-invariant part’. For instance, if $\{x, d, x + d, x + 2d\}$ is a Brauer configuration, then so is $\{x + h, d, (x + d) + h, (x + 2d) + h\}$ for any $h \in \mathbb{N}$. Brauer configurations therefore consist of a translation-invariant part $\{x, x + d, x + 2d\}$ and a non-translation-invariant part $\{d\}$.

This allows us to modify the density increment strategy of Roth and Gowers to prove Brauer’s theorem. Instead of studying a single set A lacking Brauer configurations, we study a pair of sets A and S with the following properties:

- (i) (Density) $A \subseteq [N]$ satisfies $|A| \geq \delta N$.
- (ii) (Syndeticity) $S \subseteq \mathbb{N}$ is a multiplicatively $[M]$ -syndetic set.
- (iii) (Brauer free) There is no arithmetic progression of length 3 in A with common difference in $S \cap [N/3]$.

As in the original density increment argument, we show that, provided N is ‘not too small’, we can find a long arithmetic progression $P \subseteq [N]$ on which we have a density increment of the form (4.12). As before, we can apply an affine transformation to obtain a new set $A' \subseteq [N']$ with increased density. We can also obtain a new multiplicatively $[M]$ -syndetic set $S' = d^{-1}S$, where d is the common difference of the progression P .

Recall that the translation invariant part $\{x, x + d, x + 2d\}$ of a Brauer configuration is required to come from the dense set A , whilst the non-translation-invariant part $\{d\}$ comes from the multiplicatively syndetic set S . Thus we have obtained new sets $A', S' \subseteq [N']$ satisfying (i)–(iii). Iterating this procedure as in the Gowers–Roth argument allows us to prove Theorem 4.1.

THEOREM 4.8 (Density increment for Brauer configurations). *There exists a constant $C_0 > 1$ such that the following is true. Let $A \subseteq [N]$ be such that $|A| \geq \delta N$ for some $\delta > 0$, and let $S \subseteq \mathbb{N}$ be a multiplicatively $[M]$ -syndetic set. Suppose that there do not exist $x \in A$ and $d \in S \cap [N/3]$ such that $\{x, x + d, x + 2d\} \subseteq A$. If*

$$(4.13) \quad N > \exp(C_0 \delta^{-C_0} M^{C_0}),$$

then there exist positive constants $C, c > 0$ and an arithmetic progression P in $[N]$ satisfying

$$|P| \gg N^{c\delta^C M^{-C}}$$

such that we have the density increment

$$\frac{|A \cap P|}{|P|} \geq (1 + c)\delta.$$

REMARK. The bound (4.13) is needed to ensure that N is not too small to satisfy the conclusion of the theorem. Moreover, by taking $C_0 \geq 2$, we can assume that $N \geq 18M^2$. This allows us to make use of Lemma 4.7.

Proof of Theorem 4.1 given Theorem 4.8. Let $C_1 \geq C_0$ be a large positive parameter (which does not depend on M or δ) to be specified later. We use the following iteration algorithm. After the i th iteration, we have an integer $N_i \in \mathbb{N}$, a real number $\delta_i \geq \delta > 0$, an infinite set $S_i \subseteq \mathbb{N}$, and a finite set $A_i \subseteq [N_i]$ satisfying the following three properties:

- (I) $|A_i| \geq \delta_i |N_i|$;
- (II) S_i is a multiplicatively $[M]$ -syndetic set;
- (III) there is no 3-term arithmetic progression in A_i with common difference in $S_i \cap [N_i/3]$.

We begin by defining the initial variables $N_0 := N$, $A_0 := A$, $S_0 := S$, $\delta_0 := \delta$. The iteration step of the algorithm proceeds as follows. If after the

i th iteration we have

$$(4.14) \quad N_i \leq \exp(C_1 \delta^{-C_1} M^{C_1}),$$

then the algorithm terminates. If not, then we can apply Theorem 4.8 with $N = N_i$ and $f = 1_{A_i}$ to obtain an arithmetic progression

$$P_i = \{a_i, a_i + d_i, \dots, a_i + (|P_i| - 1)d_i\} \subseteq [N_i]$$

which satisfies the length bound

$$(4.15) \quad |P_i| \gg N_i^{c\delta^C M^{-C}},$$

and provides the density increment

$$(4.16) \quad \frac{|A_i \cap P_i|}{|P_i|} \geq (1 + c) \delta_i.$$

Moreover, by partitioning P_i into two shorter progressions if necessary, we can ensure that $d_i |P_i| \leq N_i$, provided that C_1 is sufficiently large. We then take

$$\begin{aligned} N_{i+1} &:= |P_i|, \\ A_{i+1} &:= \{x \in [N_{i+1}] : a_i + (x - 1)d_i \in A_i \cap P_i\}, \\ S_{i+1} &:= d_i^{-1} S_i = \{x \in \mathbb{N} : d_i x \in S_i\}, \\ \delta_{i+1} &:= \frac{|A_{i+1}|}{|N_{i+1}|}. \end{aligned}$$

We now claim that $(N_{i+1}, A_{i+1}, S_{i+1}, \delta_{i+1})$ satisfy (I)–(III). Property (I) follows immediately from our choice of δ_{i+1} . Property (II) follows from the fact that $x \cdot [M]$ intersects $d_i^{-1} S_i$ if and only if $(d_i x) \cdot [M]$ intersects S_i . Finally, notice that if A_{i+1} contains a 3-term arithmetic progression with common difference q , then A_i contains a 3-term arithmetic progression with common difference $d_i q$. Since $d_i |P_i| \leq N_i$, we conclude that $(N_{i+1}, A_{i+1}, S_{i+1}, \delta_{i+1})$ satisfies property (III).

We have therefore shown that our algorithm may continue with the new variables $(N_{i+1}, A_{i+1}, S_{i+1}, \delta_{i+1})$. Moreover, after applying the iteration process t times, we see from (4.16) that the density δ_t satisfies $\delta_t \geq (1 + c)^t \delta$. Since $\delta_t \leq 1$ for all t , we conclude that the algorithm must terminate after T steps for some $T \ll \log(\delta^{-1})$.

We deduce that (4.14) must hold for $i = T$. By (4.15), we see that

$$N_T \geq N^{O(c^T \delta^{CT} M^{-CT})}.$$

Combining these two bounds for N_T gives

$$c^T \delta^{CT} M^{-CT} \log N \ll C_1 \delta^{-C_1} M^{C_1}.$$

Rearranging and taking logarithms gives

$$\log \log N \ll \log C_1 + T \log(c^{-1}) + (C_1 + CT) \log(M/\delta) \ll \log(\delta^{-1}) \log(M/\delta).$$

We have therefore established (4.3), as required. ■

4.3. Quadratic Fourier analysis for Brauer configurations. The goal of the rest of this section is to prove Theorem 4.8. We achieve this by adapting the methods used by Green and Tao [GT09] to study subsets of $[N]$ which lack arithmetic progressions of length 4.

The objective of their argument is to show that if a subset $A \subseteq [N]$ of density α lacks arithmetic progressions of length 4, then there exists a long arithmetic progression $P \subseteq [N]$ upon which A achieves a density increment $|A \cap P| \geq (\alpha + c(\alpha))|P|$. Gowers' argument yields an increment of the form $c(\alpha) \gg \alpha^C$. The key insight of Green and Tao is that one can obtain a larger density increment if one first shows that A has a density increment on a 'quadratic Bohr set'. A linearisation procedure can then be applied to obtain a long arithmetic progression P' which provides a density increment $c(\alpha) \gg \alpha$.

REMARK. In this subsection we make use of several results from [GT09]. The statements of these theorems contain a number of technical terms from quadratic Fourier analysis. We assume the reader is familiar with the notation used in [GT09, §§3–5] and employ this henceforth without further explanation.

The first theorem we need is [GT09, Theorem 5.6].

THEOREM 4.9 (Quadratic Koopman–von Neumann theorem). *Let $\varepsilon > 0$ and let $f : \mathbb{Z}/p\mathbb{Z} \rightarrow [-1, 1]$. Suppose $K \in \mathbb{N}$ satisfies $K \geq C\varepsilon^{-C}$ for some absolute constant $C > 0$. Then there exists a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ in $\mathbb{Z}/p\mathbb{Z}$ of complexity at most $(O(\varepsilon^{-C}), O(\varepsilon^{-C}))$ and resolution K such that*

$$(4.17) \quad \|f - \mathbb{E}(f \mid \mathcal{B}_2 \vee \mathcal{B}_{\text{triv}})\|_{U^3(\mathbb{Z}/p\mathbb{Z})} \leq \varepsilon.$$

This theorem allows us to approximate (in the U^3 norm) a 1-bounded function f with a more 'structured' function $g := \mathbb{E}(f \mid \mathcal{B}_2 \vee \mathcal{B}_{\text{triv}})$. In the proof of Theorem 4.8 we take $f = 1_A$, where $A \subseteq [N]$ is a dense subset of $[N]$ which lacks arithmetic progressions of length 3 with common difference in $S \cap [N/3]$, for a given multiplicatively syndetic set S . Our goal is to obtain a density increment on a quadratic factor for this f . To do this, we first show that it is sufficient to obtain a density increment with respect to the approximation g .

COROLLARY 4.10 (Brauer configurations on a quadratic factor). *Let $f : \mathbb{Z}/p\mathbb{Z} \rightarrow [0, 1]$ be a 1-bounded non-negative function supported on $[N]$. Let $\delta > 0$. Suppose that $N \geq 18M^2$ and*

$$(4.18) \quad |A_{S \cap [N/3]}(f) - A_{S \cap [N/3]}(\delta 1_{[N]})| \geq \delta^3/18.$$

Then there exists a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ in $\mathbb{Z}/p\mathbb{Z}$ of complexity at most $(O(\delta^{-C}M^C), O(\delta^{-C}M^C))$ and resolution $O(\delta^{-C}M^C)$ such that

$$(4.19) \quad |A_{S \cap [N/3]}(g) - A_{S \cap [N/3]}(\delta 1_{[N]})| \geq \delta^3/36,$$

where $g := \mathbb{E}(f \mid \mathcal{B}_2 \vee \mathcal{B}_{\text{triv}})$.

Proof. Let $\varepsilon = \delta^3 M^{-1}/648$. By Theorem 4.9, for some absolute constant $C > 0$, we have a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ in $\mathbb{Z}/p\mathbb{Z}$ of complexity at most $(O(\delta^{-C} M^C), O(\delta^{-C} M^C))$ and resolution $O(\delta^{-C} M^C)$ such that (4.17) holds. From Lemma 4.7 and (4.17) we deduce

$$|A_{S \cap [N/3]}(f) - A_{S \cap [N/3]}(g)| \leq 18M\varepsilon = \delta^3/36.$$

An application of the triangle inequality to (4.18) then gives (4.19). ■

We now follow the approach of Green and Tao [GT09, Corollary 5.8] by replacing f with $\mathbb{E}(f \mid \mathcal{B}_2 \vee \mathcal{B}_{\text{triv}})$ to obtain a density increment on a quadratic factor.

COROLLARY 4.11 (Density increment on a quadratic Bohr set). *There exists a constant $\tilde{C}_0 \geq 2$ such that the following is true. Let $S \subseteq \mathbb{N}$ be a multiplicatively $[M]$ -syndetic set, and let $f : \mathbb{Z}/p\mathbb{Z} \rightarrow [0, 1]$ be supported on $[N]$. Suppose $\mathbb{E}_{[N]}(f) \geq \delta$ for some $\delta > 0$. Suppose further that conditions (4.13) and (4.18) both hold, for some $C_0 \geq \tilde{C}_0$. Then there exists a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ in $\mathbb{Z}/p\mathbb{Z}$ of complexity at most $(O(\delta^{-C} M^C), O(\delta^{-C} M^C))$ and resolution $O(\delta^{-C} M^C)$, and an atom B of the factor $\mathcal{B}_2 \vee \mathcal{B}_{\text{triv}}$ with density $|B|/p \gg \exp(-O(\delta^{-C} M^C))$ which is contained in $[N]$ and is such that*

$$(4.20) \quad \mathbb{E}_B(f) \geq (1 + c)\delta$$

for some absolute constant $c > 0$.

Proof. Let $(\mathcal{B}_1, \mathcal{B}_2)$ be the quadratic factor obtained from Corollary 4.10. Let $g = \mathbb{E}(f \mid \mathcal{B}_2 \vee \mathcal{B}_{\text{triv}})$. Note that $[N] \in \mathcal{B}_{\text{triv}}$, and so $[N] \in \mathcal{B}_2 \vee \mathcal{B}_{\text{triv}}$. Since g is constant on atoms of $\mathcal{B}_2 \vee \mathcal{B}_{\text{triv}}$ and f is supported on $[N]$, also g is supported on $[N]$. This implies that $\mathbb{E}_B(f) = \mathbb{E}_B(g)$ for any $B \in \mathcal{B}_2 \vee \mathcal{B}_{\text{triv}}$. Thus, it is sufficient to establish (4.20) with g in place of f .

Let $\eta > 0$ be a small constant (to be chosen later) and define

$$\Omega := \{x \in [N] : g(x) \geq (1 + \eta)\delta\}.$$

Since g is constant on atoms of $\mathcal{B}_2 \vee \mathcal{B}_{\text{triv}}$, we deduce that Ω can be partitioned into atoms of $\mathcal{B}_2 \vee \mathcal{B}_{\text{triv}}$. We can therefore finish the proof if we can show that one of these atoms B satisfies

$$(4.21) \quad |B| \gg \exp(-O(\delta^{-C} M^C))p.$$

Recall that the factor $\mathcal{B}_2 \vee \mathcal{B}_{\text{triv}}$ has complexity and resolution $O(\delta^{-C} M^C)$, and so contains at most $\exp(O(\delta^{-C} M^C))$ atoms. Thus the pigeonhole principle implies that (4.21) holds if we can obtain a bound of the form

$$(4.22) \quad |\Omega| \gg \delta^3 p.$$

Define the function $h : \mathbb{Z}/p\mathbb{Z} \rightarrow \mathbb{C}$ by

$$h(x) = 1_{[N] \setminus \Omega}(x)g(x).$$

Thus $\|h\|_{L^\infty(\mathbb{Z}/p\mathbb{Z})} < (1 + \eta)\delta$. Taking $\eta \leq 1$ and applying Lemma 4.4 gives

$$|\Lambda_{S \cap [N/3]}(h) - \Lambda_{S \cap [N/3]}(\delta 1_{[N]})| \leq 12\delta^2 \|h - \delta 1_{[N]}\|_{L^1(\mathbb{Z}/p\mathbb{Z})}.$$

From the fact that g is a 1-bounded function, we have

$$\|h - \delta 1_{[N]}\|_{L^1(\mathbb{Z}/p\mathbb{Z})} \leq \|g - \delta 1_{[N]}\|_{L^1(\mathbb{Z}/p\mathbb{Z})} + |\Omega|/p.$$

Since h is a 1-bounded function, Lemma 4.4 also gives

$$|\Lambda_{S \cap [N/3]}(g) - \Lambda_{S \cap [N/3]}(h)| \leq 3|\Omega|/p.$$

Combining these three bounds and using the triangle inequality in (4.19) gives

$$(4.23) \quad \delta^2 \|g - \delta 1_{[N]}\|_{L^1(\mathbb{Z}/p\mathbb{Z})} + |\Omega|/p \gg \delta^3.$$

Recall that $\mathbb{E}_{[N]}(g) = \mathbb{E}_{[N]}(f) \geq \delta$. Thus,

$$\begin{aligned} \|g - \delta 1_{[N]}\|_{L^1(\mathbb{Z}/p\mathbb{Z})} &\leq \|g - \delta 1_{[N]}\|_{L^1(\mathbb{Z}/p\mathbb{Z})} + \frac{\mathbb{E}(g - \delta 1_{[N]})}{\mathbb{Z}/p\mathbb{Z}} \\ &= 2\|(g - \delta 1_{[N]})_+\|_{L^1(\mathbb{Z}/p\mathbb{Z})} \ll |\Omega|/p + \delta\eta. \end{aligned}$$

If η is sufficiently small (relative to the implicit constants), we can then substitute this bound into (4.23) to obtain (4.22). ■

To complete the proof of Theorem 4.8 it only remains to convert this density increment on a quadratic Bohr set into a density increment on an arithmetic progression. This is accomplished by implementing the following ‘linearisation’ procedure of Green and Tao [GT09, Proposition 6.2].

THEOREM 4.12 (Linearisation of quadratic Bohr sets). *Suppose $(\mathcal{B}_1, \mathcal{B}_2)$ is a quadratic factor in $\mathbb{Z}/p\mathbb{Z}$ of complexity at most (d_1, d_2) and resolution K for some $K \in \mathbb{N}$. Let B_2 be an atom of \mathcal{B}_2 . Then for all $N \in \mathbb{N}$, there is a partition of $B_2 \cap [N]$ as a union of $\ll d_2^{O(d_2)} N^{1-c/(d_1+1)(d_2+1)^3}$ disjoint arithmetic progressions in $\mathbb{Z}/p\mathbb{Z}$.*

Proof of Theorem 4.8. Let $\tilde{C}_0 \geq 2$ be the positive constant appearing in the statement of Corollary 4.11. Let A and S be as defined in the statement of Theorem 4.8. Suppose N satisfies (4.13) for some $C_0 \geq \tilde{C}_0$. Let $f = 1_A$. By Corollary 4.11 there is a quadratic factor $(\mathcal{B}_1, \mathcal{B}_2)$ in $\mathbb{Z}/p\mathbb{Z}$ of complexity at most $(O(\delta^{-C} M^C), O(\delta^{-C} M^C))$ and resolution $O(\delta^{-C} M^C)$, and an atom $B \subseteq [N]$ of $\mathcal{B}_2 \vee \mathcal{B}_{\text{triv}}$ with density $|B|/p \geq \exp(-O(\delta^{-C} M^C))$ such that $E_B(f) \geq (1 + c_0)\delta$. By Theorem 4.12, we may write B as a union of $\exp(O(\delta^{-C} M^C)) N^{1-c\delta^C M^{-C}}$ arithmetic progressions. By an application of the pigeonhole principle (see [GT09, Lemma 6.1]), we deduce that one of these progressions P satisfies

$$|P| \geq \exp(-O(\delta^{-C} M^C)) N^{c\delta^C M^{-C}} \quad \text{and} \quad \frac{|A \cap P|}{|P|} = \mathbb{E}_P f \geq \left(1 + \frac{c_0}{2}\right)\delta.$$

Notice that the only property of the parameter C_0 appearing in (4.13) that we have used is that $C_0 \geq \tilde{C}_0$. We may therefore take C_0 to be sufficiently large so that $|P| \gg N^{c'\delta^{C'}M^{-c'}}$ for some absolute constants $C', c' > 0$. ■

Appendix. Obtaining a tower bound. In this section, we show how Theorem 4.2 implies Theorem 1.4. By Theorem 4.2, there is a positive constant $\tilde{C} > 0$ such that

$$B(r+1) \leq 2^{B(r)^{\tilde{C} \log(r+1)}}$$

for all $r \in \mathbb{N}$.

Given $n \in \mathbb{N}$ and $a_1, \dots, a_n \in [2, \infty)$, define the tower function

$$T_n(a_1, a_2, \dots, a_n) := a_1^{a_2^{\dots^{a_n}}}$$

Let $K \geq 1$ be a large positive constant, to be chosen later. We can now introduce the auxiliary function $F : \mathbb{N} \rightarrow \mathbb{R}$ by

$$F(r) := T_{r+1}(2, \dots, 2, Kr^2).$$

Thus,

$$F(1) = 2^K, \quad F(2) = 2^{2^{4K}}, \quad F(3) = 2^{2^{2^{9K}}}, \quad F(4) = 2^{2^{2^{2^{16K}}}}, \quad \dots$$

Our goal is to show that if K is sufficiently large relative to \tilde{C} , then

$$B(r) \leq F(r)$$

for all $r \in \mathbb{N}$. To demonstrate why this is enough to prove Theorem 1.4, we first need to investigate the growth of tower functions.

LEMMA A.1 (Towers dominate cubes). *For all $r \geq 5$ we have*

$$(A.1) \quad r^3 \leq \text{tow}(r-1).$$

Proof. We first observe that (A.1) holds for $r = 5$. Suppose then that $r > 5$ and assume the induction hypothesis that

$$(r-1)^3 \leq \text{tow}(r-2).$$

Note that since $r > 5$, we have

$$\frac{r^3}{(r-1)^3} = \left(1 + \frac{1}{r-1}\right)^3 < \frac{125}{64} < 2.$$

This gives

$$r^3 \leq 2(r-1)^3 \leq 2 \cdot \text{tow}(r-2).$$

Using the elementary fact that $2k \leq 2^k$ for all $k \in \mathbb{N}$, we deduce

$$r^3 \leq 2^{\text{tow}(r-2)} = \text{tow}(r-1).$$

The desired result now follows by induction. ■

This lemma enables us to bound F above by a tower function.

COROLLARY A.2 (Tower bound for F). *For all $r \in \mathbb{N}$,*

$$(A.2) \quad F(r) \leq \text{tow}((1 + o(1))r).$$

Proof. Recall that F is an exponential tower of height $r + 1$, with Kr^2 as the ‘top’ term, and with the remaining terms in the tower equal to 2. By the previous lemma, when r is sufficiently large, we have

$$Kr^2 \leq \text{tow}(\lceil K^{1/3}r^{2/3} \rceil).$$

By adding the heights of the towers, we deduce that

$$F(r) \leq \text{tow}(r + \lceil K^{1/3}r^{2/3} \rceil)$$

for all sufficiently large r . This gives (A.2). ■

We require the following elementary result concerning manipulations of exponentials.

LEMMA A.3. *For all $a, b, k \geq 2$,*

$$(A.3) \quad a^b k \leq a^{b+k} \leq a^{bk}.$$

This gives the following bound on the growth of F .

COROLLARY A.4 (Tower growth of F). *For all $r \in \mathbb{N}$,*

$$F(r)^r \leq \log_2 F(r + 1).$$

Proof. Since $K \geq 1$, the case $r = 1$ can be verified by inspection. Suppose then that $r \geq 2$. By iteratively applying Lemma A.3, we deduce that

$$\begin{aligned} \log_2 F(r + 1) &= T_{r+1}(2, \dots, 2, K(r + 1)^2) \\ &\geq T_r(2, \dots, 2, 2^{Kr^2} \cdot r) \\ &\geq T_{r-1}(2, \dots, 2, 2^{2^{Kr^2}} \cdot r) \\ &\vdots \\ &\geq T_2(2, T_r(2, \dots, 2, K^{r^2}) \cdot r) = F(r)^r. \quad \blacksquare \end{aligned}$$

Proof of Theorem 1.4. By Corollary A.2, it is sufficient to show that

$$(A.4) \quad B(r) \leq F(r)$$

for all $r \in \mathbb{N}$. Since $\tilde{C} \log(n + 1) = o(n)$, we can choose $n_0 \in \mathbb{N}$ with $n_0 \geq 5$ such that

$$\tilde{C} \log(n + 1) \leq n$$

for $n \geq n_0$. By taking K sufficiently large, we can assume that (A.4) holds for $r \leq n_0$. Suppose then that $r > n_0$ and assume the induction hypothesis

$$B(r - 1) \leq F(r - 1).$$

By Theorem 4.2 and the fact that $r > n_0$, we have

$$\log_2 B(r) \leq B(r-1)^{\tilde{C} \log r} \leq B(r-1)^{r-1}.$$

Now by the induction hypothesis and Corollary A.4, we conclude that

$$\log_2 B(r) \leq F(r-1)^{r-1} \leq \log_2 F(r).$$

This establishes the induction step. ■

Acknowledgements. The author would like to thank Sean Prendiville for his constant support and encouragement, and for his helpful comments on an earlier draft of this paper. We also thank the anonymous referee for their comments and suggestions on an earlier version of this paper.

References

- [BHLS15] B. Barber, N. Hindman, I. Leader and D. Strauss, *Partition regularity without the columns property*, Proc. Amer. Math. Soc. 143 (2015), 3387–3399.
- [Ber10] V. Bergelson, *Ultrafilters, IP sets, dynamics, and combinatorial number theory*, in: Ultrafilters across Mathematics, Contemp. Math. 530, Amer. Math. Soc., Providence, RI, 2010, 23–47.
- [BG18] V. Bergelson and D. Glasscock, *Multiplicative richness of additively large sets in \mathbb{Z}^d* , J. Algebra 503 (2018), 67–103.
- [Bra28] A. Brauer, *Über Sequenzen von Potenzresten*, Sitzungsber. Preuss. Akad. Wiss. 1928, 9–16.
- [CP20] J. Chapman and S. Prendiville, *On the Ramsey number of the Brauer configuration*, Bull. London Math. Soc. 52 (2020), 316–334.
- [CLP18] S. Chow, S. Lindqvist and S. Prendiville, *Rado’s criterion over squares and higher powers*, J. Eur. Math. Soc., to appear; arXiv:1806.05002v2 (2020).
- [CEG02] F. Chung, P. Erdős and R. Graham, *On sparse sets hitting linear forms*, in: Number Theory for the Millennium, I (Urbana, IL, 2000), A K Peters, Natick, MA, 2002, 257–272.
- [CS17] K. Cwalina and T. Schoen, *Tight bounds on additive Ramsey-type numbers*, J. London Math. Soc. 96 (2017), 601–620.
- [EEN01] D. B. Ellis, R. Ellis and M. Nerurkar, *The topological dynamics of semigroup actions*, Trans. Amer. Math. Soc. 353 (2001), 1279–1320.
- [EG80] P. Erdős and R. L. Graham, *Old and New Problems and Results in Combinatorial Number Theory*, Monogr. Enseign. Math. 28, L’Enseign. Math. and Univ. de Genève, Genève, 1980.
- [Gow98] W. T. Gowers, *A new proof of Szemerédi’s theorem for progressions of length four*, Geom. Funct. Anal. 8 (1998), 529–551.
- [Gow01] W. T. Gowers, *A new proof of Szemerédi’s theorem*, Geom. Funct. Anal. 11 (2001), 465–588.
- [GRS90] R. L. Graham, B. L. Rothschild and J. H. Spencer, *Ramsey Theory*, 2nd ed., Wiley, New York, 1990.
- [GWS77] R. L. Graham, H. S. Witsenhausen, and J. H. Spencer, *On extremal density theorems for linear forms*, in: Number Theory and Algebra, Academic Press, New York, 1977, 103–109.

- [GT09] 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 Univ. Press, Cambridge, 2009, 180–204.
- [HLS03] N. Hindman, I. Leader and D. Strauss, *Open problems in partition regularity*, *Combin. Probab. Comput.* 12 (2003), 571–583.
- [HS12] N. Hindman and D. Strauss, *Algebra in the Stone–Čech Compactification: Theory and Applications*, 2nd ed., de Gruyter, Berlin, 2012.
- [Rad33] R. Rado, *Studien zur Kombinatorik*, *Math. Z.* 36 (1933), 424–480.
- [Rot53] K. F. Roth, *On certain sets of integers*, *J. London Math. Soc.* 28 (1953), 104–109.
- [Sár78] A. Sárközy, *On difference sets of sequences of integers. I*, *Acta Math. Acad. Sci. Hungar.* 31 (1978), 125–149.
- [Sch16] I. Schur, *Über die Kongruenz $x^m + y^m = z^m \pmod{p}$* , *Jahresber. Deutsch. Math. Verein.* 25 (1916), 114–117.
- [She88] S. Shelah, *Primitive recursive bounds for van der Waerden numbers*, *J. Amer. Math. Soc.* 1 (1988), 683–697.
- [TV06] T. Tao and V. Vu, *Additive Combinatorics*, Cambridge Univ. Press, Cambridge, 2006.
- [Wae27] B. L. van der Waerden, *Beweis einer Baudetschen Vermutung*, *Nieuw Arch. Wisk.* 15 (1927), 212–216.

Jonathan Chapman
Department of Mathematics
University of Manchester
Oxford Road
Manchester, M13 9PL, UK
E-mail: jonathan.chapman@manchester.ac.uk