

An Approximate Counting Version of the Multidimensional Szemerédi Theorem

Natalie Behague*, Joseph Hyde, Natasha Morrison†, Jonathan A. Noel‡ and Ashna Wright§

Department of Mathematics and Statistics, University of Victoria, Victoria, B.C., Canada.

`{nbehague,josephhyde,nmorrison,noelj,ashnawright}@uvic.ca`

September 15, 2025

Abstract

For any fixed $d \geq 1$ and subset X of \mathbb{N}^d , let $r_X(n)$ be the maximum cardinality of a subset A of $\{1, \dots, n\}^d$ which does not contain a subset of the form $\mathbf{b} + rX$ for $r > 0$ and $\mathbf{b} \in \mathbb{R}^d$. Such a set A is said to be X -free. The Multidimensional Szemerédi Theorem of Furstenberg and Katznelson states that $r_X(n) = o(n^d)$. We show that, for $|X| \geq 3$ and infinitely many $n \in \mathbb{N}$, the number of X -free subsets of $\{1, \dots, n\}^d$ is at most $2^{O(r_X(n))}$. The proof involves using a known multidimensional extension of Behrend's construction to obtain a supersaturation theorem for copies of X in dense subsets of $[n]^d$ for infinitely many values of n and then applying the powerful hypergraph container lemma. Our result generalizes work of Balogh, Liu, and Sharifzadeh on k -AP-free sets and Kim on corner-free sets.

1 Introduction

Szemerédi's Theorem [31] states that the maximum cardinality of a subset of $[n] := \{1, \dots, n\}$ not containing k points in an arithmetic progression is $o(n)$, settling a conjecture of Erdős and Turán [12] from the 1930s. Szemerédi's original proof was purely combinatorial and led to the discovery of his celebrated Regularity Lemma [32]. A few years later, Furstenberg [13] found

*Research supported by a PIMS Postdoctoral Fellowship.

†Research supported by NSERC Discovery Grant RGPIN-2021-02511 and NSERC Early Career Supplement DGECR-2021-00047 and a Start-Up Grant from the University of Victoria.

‡Research supported by NSERC Discovery Grant RGPIN-2021-02460 and NSERC Early Career Supplement DGECR-2021-00024 and a Start-Up Grant from the University of Victoria.

§Research supported by an NSERC CGS-M.

an alternative proof using ergodic theory. While this approach relies on the Axiom of Choice and, therefore, does not yield effective bounds, it has the benefit of being more amenable to generalization. For example, the first proof of the Density Hales–Jewett Theorem was obtained via an extension of Furstenberg’s ideas [15]. Later, another proof of Szemerédi’s Theorem was given by Gowers [16] using Fourier analysis.

Our main result is a “counting version” of the Multidimensional Szemerédi Theorem (Theorem 1.1 below), which was first proven by Furstenberg and Katznelson [14] using ergodic theory. Later, using extensions of Szemerédi’s Regularity Lemma to hypergraphs, Gowers [17], Tao [33] and Rödl, Nagle, Schacht, and Skokan [22, 25] obtained combinatorial proofs of the Multidimensional Szemerédi Theorem which yield effective bounds (albeit, with very extreme dependencies). In order to state our result, we require a few definitions. Given $d \in \mathbb{N}$ and a set $X \subseteq \mathbb{N}^d$, a *copy* of X is a set of the form

$$\mathbf{b} + rX := \{\mathbf{b} + r\mathbf{x} : \mathbf{x} \in X\},$$

where $\mathbf{b} \in \mathbb{R}^d$ and $r \in \mathbb{R}_{\geq 0}$. A copy of X is *non-trivial* if $r \neq 0$. Let $r_X(n)$ denote the cardinality of the largest subset of $[n]^d$ which does not contain a non-trivial copy of X ; such a set is said to be *X-free*. For $d = 1$, we simply write $r_k(n)$ to mean $r_{\{1, \dots, k\}}(n)$. In this language, Szemerédi’s Theorem [31] says that $r_k(n) = o(n)$, for any $k \geq 3$. The Multidimensional Szemerédi Theorem of [14] is as follows.

Theorem 1.1 (Multidimensional Szemerédi Theorem [14]; see also [17, 22, 25, 33]). *If $d \geq 1$ and X is a finite subset of \mathbb{N}^d , then $r_X(n) = o(n^d)$.*

Our focus in this paper is on the related question of counting the number of X -free subsets of $[n]^d$. It is trivial to see that there are at least $2^{r_X(n)}$ such sets. The following well-known (and still open) question of Cameron and Erdős [9] asks whether this bound is approximately correct for k -term arithmetic progressions.

Question 1.2 (Cameron and Erdős [9, Section 4.2]). *Is the number of subsets of $[n]$ not containing a k -term arithmetic progression equal to $2^{(1+o(1))r_k(n)}$?*

While Question 1.2 is still open, it is reasonable to believe that the answer could be yes. There are many results in the literature which assert that, up to a $(1 + o(1))$ factor in the exponent, the number of combinatorial structures satisfying a certain set of constraints is equal to the number of subsets of the largest such structure. See, for example, the results of Kleitman [21] on antichains, Dong, Mani, and Zhao [10] on t -error correcting codes, and Balogh, Das, Delcourt, Liu, and Sharifzadeh [1] on intersecting families. The following result of Balogh, Liu, and Sharifzadeh [2] makes substantial progress towards answering Question 1.2 in the affirmative.

Theorem 1.3 (Balogh, Liu, and Sharifzadeh [2, Theorem 4.2]). *For $k \geq 3$ and for infinitely many $n \in \mathbb{N}$, the number of subsets of $[n]$ with no k -term arithmetic progression is $2^{O(r_k(n))}$.*

Our main result is the following generalization of Theorem 1.3.

Theorem 1.4. *Let d be a positive integer and let $X \subseteq \mathbb{N}^d$ be a finite set such that $|X| \geq 3$. For infinitely many $n \in \mathbb{N}$, the number of X -free subsets of $[n]^d$ is $2^{O(r_X(n))}$.*

The statement of Theorem 1.4 was recently proven by Kim [20] in the special case that $X = \{\mathbf{0}, \mathbf{e}_1, \dots, \mathbf{e}_d\}$, where $\mathbf{e}_1, \dots, \mathbf{e}_d$ are the vectors of the standard basis of \mathbb{R}^d . An X -free set for this particular choice of X is said to be *corner-free*. It is well-known that the specialization of the Multidimensional Szemerédi Theorem to corner-free sets implies the full Multidimensional Szemerédi Theorem; see, e.g., [17, Proof Theorem 10.3]. However, there seems to be no obvious reduction of Theorem 1.4 to the result of Kim [20] due to the fact that the function $r_X(n)$ for a generic set X may have a very different growth rate when compared with the maximum size of a corner-free set.

As in [2] and [20], we will obtain our results by applying the seminal hypergraph container method, which was introduced by Balogh, Morris, and Samotij [3] and, independently, Saxton and Thomason [28]. It is a widely applicable tool which has since been used to prove a myriad of enumeration results; see the survey [4].

In applying the hypergraph container method to prove Theorem 1.4, the high level idea is to show that the X -free subsets of $[n]^d$ can be covered by $2^{o(r_X(n))}$ sets called “containers”, each of which only has $O(r_X(n))$ elements. As in most applications of this method, the key to obtaining these containers is to prove a “supersaturation” result which ensures that any subset of $[n]^d$ with $C \cdot r_X(n)$ elements, for large C , contains many copies of X . Due to our limited understanding of the exact growth rate of $r_X(n)$ (and, in particular, how severely it “fluctuates” as n grows), we are only able to obtain such a supersaturation result on an infinite subset of \mathbb{N} ; this explains why Theorem 1.4 does not apply to all $n \in \mathbb{N}$.

The exact asymptotics of $r_X(n)$ have not been determined, even when X is a 3-term arithmetic progression. In 1946, Behrend [5] presented a construction of large sets that do not contain 3-term arithmetic progressions, giving the following lower bound on $r_3(n)$,

$$r_3(n) \geq \Omega \left(\frac{n}{2^{2\sqrt{2}} \sqrt{\log_2 n} \cdot \log^{1/4}(n)} \right).$$

In 2011, this bound was improved upon by a factor of $\Theta(\sqrt{\log n})$ by Elkin [11]. Shortly after Elkin’s paper, Green and Wolf [18] found a simpler proof of the same bound. The first bound of $r_3(n) = o(n)$ was famously proved by Roth [26] and has been improved many times throughout the years [6–8, 19, 27, 29]. Currently, the best known upper bound is $r_3(n) \leq n \cdot 2^{-O((\log n)^\beta)}$, for some absolute constant $\beta > 0$, proved by Kelley and Meka [19].

The case for $k > 3$ has also been studied extensively. Currently, the best known bounds on $r_k(n)$ are:

$$\frac{Cn \sqrt[a]{\log n}}{2^{a2^{(a-1)/2}} \sqrt[a]{\log n}} \leq r_k(n) \leq \frac{n}{(\log \log n)^{C'}},$$

where $a = \lceil \log k \rceil$ and C, C' are positive constants depending only on k . The lower bound was obtained by O’Bryant [23] by generalizing the ideas of Elkin [11] and Green and Wolf [18] and incorporating them into an argument of Rankin [24], who gave a natural generalization of Behrend’s construction [5]. The upper bound is due to Gowers [16].

The rest of the paper is organized as follows. In Section 2 we prove a supersaturation lemma for all n using a consequence of the prime number theorem. In Section 3, we use the results of Section 2 together with a generalization of Behrend's construction to prove a stronger supersaturation result for an infinite subset of \mathbb{N} . Finally, in Section 4, we will state a hypergraph container lemma and combine it with the supersaturation result from Section 3 to prove Theorem 1.4. Throughout this paper, let d be a positive integer and let X be a finite subset of \mathbb{N}^d such that $|X| \geq 3$. All logarithms in this paper are natural unless otherwise specified.

2 Supersaturation lemmas

As is typical with applications of the hypergraph container method, one of the ingredients used to prove Theorem 1.4 is a supersaturation result. In this section, we will prove some preliminary supersaturation results that provide a lower bound for the number of copies of X in a large set $A \subseteq [n]^d$. The results of this section will be used in Section 3 to prove a stronger supersaturation result for an infinite set of positive integers. For $A \subseteq [n]^d$, let $\Gamma_X(A)$ be the number of copies of X in A . We begin with the following simple lower bound on $\Gamma_X(A)$.

Lemma 2.1. *For any $A \subseteq [n]^d$,*

$$\Gamma_X(A) \geq |A| - r_X(n).$$

Proof. We proceed by induction on $|A|$. If $|A| \leq r_X(n)$, then the right side of the inequality is at most zero so there is nothing to prove. Now, assume $|A| \geq r_X(n) + 1$. Then A must contain at least one copy of X , say X' . For any $x \in X'$, the set $A \setminus \{x\}$ contains at least $|A| - 1 - r_X(n)$ copies of X by the induction hypothesis. These copies, along with X' which is not in $A \setminus \{x\}$, yield $|A| - r_X(n)$ copies of X in A . \square

The bound in Lemma 2.1 is, obviously, rather trivial. However, as we see in Lemma 2.3 below, we can obtain a seemingly stronger bound by applying Lemma 2.1 within copies of $[M]^d$ inside of $[n]^d$, where $1 \leq M \leq n$, and using an “averaging argument.” The copies of $[M]^d$ that we consider are translates of $p \cdot [M]^d$ where p is a small prime; the following consequence of the prime number theorem (see, e.g., [30]) is useful for counting the choices of p . For $\ell \geq 2$, let $\pi(\ell)$ be the number of primes p such that $2 \leq p \leq \ell$.

Proposition 2.2. *There exists $\ell_0 \in \mathbb{N}$ such that $\pi(\ell) \geq \frac{\ell}{2 \log(\ell)}$ for all $\ell \geq \ell_0$.*

Lemma 2.3. *Let M and n be integers such that $2 \leq M \leq n$. If $A \subseteq [n]^d$ such that*

$$|A| > \max\{4\ell_0 d n^{d-1} M, r_X(n)\} \tag{2.4}$$

where ℓ_0 is as in Proposition 2.2, then

$$\Gamma_X(A) \geq \frac{|A|}{11d \log^2(n)} \cdot \frac{n}{M} \left(\frac{|A|}{2n^d} - \frac{r_X(M)}{M^d} \right).$$

Proof. By translating and scaling X if necessary, we may assume that there does not exist any “smaller” non-trivial copy of X in \mathbb{N}^d . Specifically, we can assume that there does not exist an integer $t > 1$ which divides the coordinates of the vector $\mathbf{x} - \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in X$.

Let p be a prime chosen uniformly at random from all primes $p \leq \frac{|A|}{4dn^{d-1}M}$. Next, choose a “base point” \mathbf{b} uniformly at random from the set $\{0, \dots, n - Mp - 1\}^d$ (which is non-empty because $Mp \leq \frac{|A|}{4dn^{d-1}} < n$) and define

$$G_{\mathbf{b}} = \mathbf{b} + p \cdot [M]^d.$$

Let $A_{\mathbf{b}} := A \cap G_{\mathbf{b}}$. Our goal is to bound $\mathbb{E}[\Gamma_X(A_{\mathbf{b}})]$ from below in terms of $|A|$ and from above in terms of $\Gamma_X(A)$; combining these inequalities will give us the desired lower bound on $\Gamma_X(A)$.

Note that $G_{\mathbf{b}}$ is a non-trivial copy of $[M]^d$, so we can apply Lemma 2.1 within $G_{\mathbf{b}}$ to yield

$$\mathbb{E}[\Gamma_X(A_{\mathbf{b}})] \geq \mathbb{E}[|A_{\mathbf{b}}| - r_X(M)] = \mathbb{E}[|A_{\mathbf{b}}|] - r_X(M) \quad (2.5)$$

by linearity of expectation. So, to get a lower bound on $\mathbb{E}[\Gamma_X(A_{\mathbf{b}})]$, it suffices to bound $\mathbb{E}[|A_{\mathbf{b}}|]$. Define $I_p := \{0, \dots, n - Mp - 1\}^d \cap [n]^d$; that is I_p is obtained from $[n]^d$ by removing all points where some coordinate is ‘large’. As there are fewer than $dn^{d-1}Mp$ such points and $p \leq \frac{|A|}{4dn^{d-1}M}$, we have that

$$|A \cap I_p| > |A| - dn^{d-1}Mp \geq |A| - dn^{d-1}M \left(\frac{|A|}{4dn^{d-1}M} \right) = \frac{|A|}{2}$$

for every possible choice of p . Given the choice of p , for any $v \in I_p$ there are exactly M^d choices of base point $\mathbf{b} \in \{0, \dots, n - Mp - 1\}^d$ with the property that $v \in G_{\mathbf{b}}$. Thus, conditioned on the choice of p , the probability that any given $v \in I_p$ is in $G_{\mathbf{b}}$ is at least $\frac{M^d}{(n-Mp)^d} \geq \frac{M^d}{n^d}$. Therefore, $\mathbb{E}[|A_{\mathbf{b}}|] \geq \mathbb{E}[|A \cap I_p \cap G_{\mathbf{b}}|] \geq \frac{|A|}{2} \cdot \frac{M^d}{n^d}$. Substituting this lower bound into (2.5) gives

$$\mathbb{E}[\Gamma_X(A_{\mathbf{b}})] \geq \frac{|A|}{2} \cdot \frac{M^d}{n^d} - r_X(M). \quad (2.6)$$

Next, we will obtain an upper bound on $\mathbb{E}[\Gamma_X(A_{\mathbf{b}})]$ in terms of $\Gamma_X(A)$. To do this, we let Y be an arbitrary fixed non-trivial copy of X in A and bound the probability that $Y \subseteq G_{\mathbf{b}}$ from above. Let $\mathbf{a} \in \mathbb{R}^d$ and $r > 0$ so that $Y = \mathbf{a} + r \cdot X$ and note that, since $|X| \geq 2$, the choice of \mathbf{a} and r is unique. Then, clearly, Y is a subset of $G_{\mathbf{b}}$ if and only if $(\mathbf{a} - \mathbf{b}) + r \cdot X$ is a subset of $p \cdot [M]^d$. Thus, if Y is a subset of $G_{\mathbf{b}}$, then, for any $\mathbf{x}, \mathbf{y} \in X$, every coordinate of

$$(\mathbf{a} - \mathbf{b} + r\mathbf{x}) - (\mathbf{a} - \mathbf{b} + r\mathbf{y}) = r(\mathbf{x} - \mathbf{y})$$

is a multiple of p . By our assumption that X is the smallest copy of itself in \mathbb{N}^d , this implies that r is an integer and p divides r . Thus, for any non-trivial copy $Y = \mathbf{a} + rX$ of X in A , the probability that Y is a subset of $G_{\mathbf{b}}$ is at most the probability that p is a divisor of r , multiplied by the probability that, for an arbitrary $\mathbf{y}_0 \in Y$, the base point \mathbf{b} is chosen to be one of the at most M^d base points for which $\mathbf{y}_0 \in G_{\mathbf{b}}$. Let us now bound these two

probabilities individually. Since, for any such Y , the integer r has at most $\log_2(r) \leq \log_2(n)$ divisors, the probability that p divides r is at most

$$\frac{\log_2(n)}{\pi(|A|/4dn^{d-1}M)} \leq \frac{\log_2(n)8dn^{d-1}M \log(|A|/4dn^{d-1}M)}{|A|} \leq \frac{8\log_2^2(n) \cdot Mdn^{d-1}}{|A|}$$

where the first inequality uses Proposition 2.2 and the fact that $|A| > 4\ell_0 dn^{d-1}M$. After fixing the choice of p , for any $\mathbf{y}_0 \in Y$, the probability that $\mathbf{y} \in G_{\mathbf{b}}$ is at most

$$\frac{M^d}{(n - Mp)^d} \leq \frac{M^d}{\left(n - \frac{|A|}{4dn^{d-1}}\right)^d}.$$

Observe that

$$\left(\frac{M}{n - \frac{|A|}{4dn^{d-1}}}\right)^d = \left(\frac{M}{n}\right)^d \left(1 - \frac{|A|}{4dn^d}\right)^{-d} \leq \left(\frac{M}{n}\right)^d \left(1 - \frac{1}{4d}\right)^{-d} \leq \left(\frac{M}{n}\right)^d \cdot \frac{4}{3}.$$

Therefore, for any non-trivial copy Y of X in A , we have

$$\mathbb{P}[Y \subseteq A_{\mathbf{b}}] \leq \frac{8\log_2^2(n) \cdot Mdn^{d-1}}{|A|} \cdot \left(\frac{M}{n}\right)^d \cdot \frac{4}{3} \leq \frac{11 \cdot \log_2^2(n)M^{d+1}d}{|A|n}.$$

Using this, we get the following upper bound on $\mathbb{E}[\Gamma_X(A_{\mathbf{b}})]$ by linearity of expectation:

$$\mathbb{E}[\Gamma_X(A_{\mathbf{b}})] \leq \Gamma_X(A) \cdot \frac{11 \cdot \log_2^2(n)M^{d+1}d}{|A|n}. \quad (2.7)$$

Putting together (2.6) and (2.7) and solving for $\Gamma_X(A)$ gives,

$$\Gamma_X(A) \geq \frac{n|A|}{11\log_2^2(n)M^{d+1}d} \left(\frac{|A|M^d}{2n^d} - r_X(M)\right) = \frac{|A|}{11d\log_2^2(n)} \cdot \frac{n}{M} \left(\frac{|A|}{2n^d} - \frac{r_X(M)}{M^d}\right),$$

as desired. \square

3 Stronger supersaturation on an infinite set

Our aim in this section is to prove a stronger supersaturation result for infinitely many values of n . We begin by proving the following basic lemma about functions of n with a particular growth rate.

Lemma 3.1. *For $\beta_r, \beta_m > 0$, let $0 < \alpha < \exp(-\beta_r\beta_m/2)$. For $d \geq 1$, suppose that $r : \mathbb{N} \rightarrow \mathbb{N}$ and $m : \mathbb{N} \rightarrow \mathbb{N}$ are such that*

$$r(n) \geq \frac{n^d}{\exp(\beta_r \sqrt{\log(n)})}$$

and

$$n \geq m(n) \geq \frac{n}{\exp(\beta_m \sqrt{\log(n)})}$$

for all $n \geq 1$. Then there exists an infinite subset N of \mathbb{N} such that

$$\frac{r(n)}{n^d} \geq \frac{\alpha \cdot r(m(n))}{m(n)^d}$$

for all $n \in N$.

Proof. Take $\varepsilon > 0$ so that $\alpha = \frac{\exp(\frac{-\beta_r \beta_m}{2})}{(1+\varepsilon)^3}$. By definition of $m(n)$,

$$\begin{aligned} \sqrt{\log(m(n))} - \sqrt{\log(n)} &\geq \sqrt{\log\left(\frac{n}{\exp(\beta_m \sqrt{\log(n)})}\right)} - \sqrt{\log(n)} \\ &= \sqrt{\log(n) - \beta_m \sqrt{\log(n)}} - \sqrt{\log(n)}. \end{aligned} \quad (3.2)$$

Observe that the expression in (3.2) converges to $\frac{-\beta_m}{2}$ as $n \rightarrow \infty$. Therefore, there exists some $n_0 \in \mathbb{N}$ such that, for all $n \geq n_0$,

$$\sqrt{\log(m(n))} - \sqrt{\log(n)} \geq \frac{-\beta_m}{2} - \frac{\log(1+\varepsilon)}{\beta_r}.$$

This implies that for all $n \geq n_0$,

$$\exp\left(\beta_r \left(\sqrt{\log(m(n))} - \sqrt{\log(n)}\right)\right) \geq \frac{\exp\left(\frac{-\beta_r \beta_m}{2}\right)}{(1+\varepsilon)}. \quad (3.3)$$

Define

$$Y := \left\{ \frac{r(n) \exp(\beta_r \sqrt{\log(n)})}{n^d} : n \in \mathbb{N} \right\}.$$

Let Z be the set of all cluster points of Y in the extended reals, $\mathbb{R} \cup \{-\infty, \infty\}$. Note that, by the lower bound on $r(n)$, we have $Z \subseteq [1, \infty]$. Let $z = \sup(Z)$. We now proceed by cases depending on the value of z .

Suppose $z = \infty$. Define

$$N := \left\{ n \in \mathbb{N} : n \geq n_0 \text{ and } \frac{r(n) \exp(\beta_r \sqrt{\log(n)})}{n^d} \geq \frac{r(m) \exp(\beta_r \sqrt{\log(m)})}{m^d}, \forall m \leq n \right\}.$$

As $z = \infty$, it follows that N is an infinite set. As $m(n) \leq n$, for all $n \in N$ we have,

$$\frac{r(n)}{n^d} \geq \frac{r(m(n)) \cdot \exp(\beta_r \sqrt{\log(m(n))})}{m(n)^d \cdot \exp(\beta_r \sqrt{\log(n)})} \geq \frac{r(m(n))}{m(n)^d} \cdot \frac{\exp(\frac{-\beta_m \beta_r}{2})}{(1+\varepsilon)} \geq \alpha \cdot \frac{r(m(n))}{m(n)^d},$$

where the second inequality follows from (3.3). Therefore, N is as desired.

We may now assume that $z < \infty$. Choose n'_0 large enough such that for all $n \geq n'_0$,

$$r(m(n)) \cdot \frac{\exp(\beta_r \sqrt{\log(m(n))})}{m(n)^d} \leq (1 + \varepsilon)z. \quad (3.4)$$

Note that such an n'_0 exists as $m(n) \rightarrow \infty$ as $n \rightarrow \infty$. Define

$$N := \left\{ n \in \mathbb{N} : n \geq \max\{n_0, n'_0\} \text{ and } \frac{r(n) \cdot \exp(\beta_r \sqrt{\log(n)})}{n^d} \geq \frac{z}{(1 + \varepsilon)} \right\}.$$

By definition of z , the set N is infinite. By definition of N , for all $n \in N$,

$$\begin{aligned} \frac{r(n)}{n^d} &\geq \frac{z}{(1 + \varepsilon) \cdot \exp(\beta_r \sqrt{\log(n)})} \stackrel{(3.4)}{\geq} \frac{r(m(n))}{m(n)^d} \cdot \frac{\exp(\beta_r \sqrt{\log(m(n))})}{\exp(\beta_r \sqrt{\log(n)}) \cdot (1 + \varepsilon)^2} \\ &\stackrel{(3.3)}{\geq} \frac{\exp(-\frac{\beta_r \beta_m}{2})}{(1 + \varepsilon)^3} \cdot \frac{r(m(n))}{m(n)^d} = \alpha \cdot \frac{r(m(n))}{m(n)^d}. \end{aligned}$$

Therefore, N is as desired. \square

In proving the main result of this section, we will apply Lemma 3.1 with $r(n) = r_X(n)$. To do so, we require the following standard extension of the construction of Behrend [5] to higher dimensions. For completeness, the proof is included in Appendix A.

Lemma 3.5. *Let $d \geq 1$ and $X \subseteq \mathbb{N}^d$ be fixed. If $|X| \geq 3$, then there exists $c_X > 0$ depending only on X such that*

$$r_X(n) \geq n^d \cdot \exp(-c_X \sqrt{\log n}).$$

In order to apply Lemma 3.1, we require another function $m(n)$. Define, for the rest of this section,

$$m(n) := \frac{n}{\log^{3|X|}(n)} \left(\frac{r_X(n)}{n^d} \right)^{|X|+2}.$$

Using Lemma 3.5, observe that, for sufficiently large n ,

$$\begin{aligned} m(n) &= \frac{n}{\log^{3|X|}(n)} \left(\frac{r_X(n)}{n^d} \right)^{|X|+2} \geq \frac{n}{\log^{3|X|}(n)} \left(\exp(-c_X \sqrt{\log n}) \right)^{|X|+2} \\ &> n \cdot \left(\exp\left(-(5c_X|X|)\sqrt{\log n}\right) \right). \end{aligned} \quad (3.6)$$

Let α be chosen such that $0 < \alpha < \exp\left(-\frac{5c_X^2|X|}{2}\right)$. This choice of α is fixed for the remainder of this paper.

Lemma 3.7. *There exists an infinite subset N of \mathbb{N} such that, if $n \in N$ and $A \subseteq [n]^d$ where $|A| \geq \frac{4}{\alpha} \cdot r_X(n)$, then*

$$\Gamma_X(A) \geq \frac{\log^{3|X|-2}(n)}{3d} \cdot \left(\frac{n^d}{r_X(n)} \right)^{|X|} \cdot n^d.$$

Proof. Let ℓ_0 be as defined in Proposition 2.2. By Lemma 3.5 and (3.6), we may apply Lemma 3.1 with $m(n)$, $r_X(n)$, and α to obtain an infinite subset N' of \mathbb{N} , such that for all $n \in N'$,

$$\frac{r_X(n)}{\alpha \cdot n^d} \geq \frac{r_X(m(n))}{m(n)^d}. \quad (3.8)$$

Note that $\alpha < 1$ and so the condition $|A| \geq \frac{1}{\alpha} \cdot r_X(n)$ implies that $|A| \geq r_X(n)$. For large enough $n \in N'$, we have

$$\begin{aligned} 4\ell_0 d n^{d-1} m(n) &= 4\ell_0 d n^{d-1} \cdot \frac{n}{\log^{3|X|}(n)} \cdot \left(\frac{r_X(n)}{n^d} \right)^{|X|+2} \leq 4\ell_0 d n^d \left(\frac{r_X(n)}{n^d} \right)^2 \\ &= 4\ell_0 d \left(\frac{r_X(n)}{n^d} \right) \cdot r_X(n) < \frac{4}{\alpha} \cdot r_X(n) \leq |A|, \end{aligned}$$

where the penultimate inequality holds for sufficiently large n as $r_X(n) = o(n^d)$. Therefore A satisfies the conditions of Lemma 2.3 with $M = m(n)$. Applying this we obtain

$$\Gamma_X(A) \geq \frac{1}{11d \log^2(n)} \cdot |A| \cdot \frac{n}{m(n)} \cdot \left(\frac{|A|}{2n^d} - \frac{r_X(m(n))}{m(n)^d} \right).$$

Using $|A| \geq \frac{4}{\alpha} \cdot r_X(n)$ and using (3.8) to bound the final term, we get

$$\begin{aligned} \Gamma_X(A) &\geq \frac{1}{11d \log^2(n)} \cdot |A| \cdot \frac{n}{m(n)} \cdot \left(\frac{2r_X(n)}{\alpha \cdot n^d} - \frac{r_X(n)}{\alpha \cdot n^d} \right) \\ &= \frac{1}{11d \log^2(n)} \cdot |A| \cdot \frac{n}{m(n)} \cdot \left(\frac{r_X(n)}{\alpha \cdot n^d} \right). \end{aligned}$$

Further, we may apply the definition of $m(n)$ and the lower bound on $|A|$ to get

$$\begin{aligned} \Gamma_X(A) &\geq \frac{1}{11d \log^2(n)} \cdot |A| \cdot n \cdot \frac{\log^{3|X|}(n)}{n} \cdot \left(\frac{n^d}{r_X(n)} \right)^{|X|+2} \cdot \left(\frac{r_X(n)}{\alpha \cdot n^d} \right) \\ &\geq \frac{1}{11d \log^2(n)} \cdot \frac{4}{\alpha} \cdot r_X(n) \cdot \log^{3|X|}(n) \cdot \left(\frac{n^d}{r_X(n)} \right)^{|X|+2} \cdot \left(\frac{r_X(n)}{\alpha \cdot n^d} \right) \\ &\geq \frac{\log^{3|X|-2}(n)}{3d\alpha^2} \cdot \left(\frac{n^d}{r_X(n)} \right)^{|X|} \cdot n^d. \end{aligned}$$

Hence, as $\alpha < 1$, we have for large enough $n \in N'$

$$\Gamma_X(A) \geq \frac{\log^{3|X|-2}(n)}{3d} \cdot \left(\frac{n^d}{r_X(n)} \right)^{|X|} \cdot n^d. \quad (3.9)$$

Define N to be the set of all $n \in N'$ which are large enough that (3.9) holds. \square

4 Main result

In this section, we prove Theorem 1.4. The proof applies a corollary of a hypergraph container lemma of Saxton and Thomason [28]. To state this lemma, we require some definitions.

Consider an r -uniform hypergraph $\mathcal{H} = (V(\mathcal{H}), E(\mathcal{H}))$, where $V(\mathcal{H})$ is the set of vertices and $E(\mathcal{H})$ is the set of hyperedges. We use $\bar{d}(\mathcal{H})$ to denote the average degree of a vertex in \mathcal{H} ; by a handshaking argument, $\bar{d}(\mathcal{H}) = r|E(\mathcal{H})|/|V(\mathcal{H})|$. For a set $A \subseteq V(\mathcal{H})$, we use $e_{\mathcal{H}}(A)$ to denote the number of hyperedges contained in A . We say $A \subseteq V(\mathcal{H})$ is *independent* if $e_{\mathcal{H}}(A) = 0$. The *co-degree* of a set $A \subseteq V(\mathcal{H})$ is the number of hyperedges of \mathcal{H} that contain A . We use $d_{\mathcal{H}}(A)$, or simply $d(A)$ when the subscript is obvious from context, to denote the co-degree of a set A . We denote the maximum co-degree of all sets of size k in \mathcal{H} as $\Delta_k(\mathcal{H})$; explicitly

$$\Delta_k(\mathcal{H}) := \max\{d_{\mathcal{H}}(A) : |A| = k\}.$$

Finally, we define $\Delta(\mathcal{H}, \tau)$ for any $\tau > 0$ as

$$\Delta(\mathcal{H}, \tau) := 2^{\binom{r}{2}-1} \sum_{j=2}^r 2^{-\binom{j-1}{2}} \frac{\Delta_j(\mathcal{H})}{\tau^{j-1} \bar{d}(\mathcal{H})}.$$

Notice that $\Delta(\mathcal{H}, \tau)$ depends on all maximum co-degrees in \mathcal{H} , except for $\Delta_1(\mathcal{H})$.

Theorem 4.1 (Hypergraph Container Lemma [28]). *Let \mathcal{H} be an r -uniform hypergraph. Suppose that $0 < \varepsilon, \tau < 1/2$ satisfy*

- $\tau < 1/(200 \cdot r \cdot r!^2)$, and
- $\Delta(\mathcal{H}, \tau) \leq \varepsilon/(12r!)$.

Then there exists $c = c(r) \leq 1000 \cdot r \cdot r!^3$ and a collection \mathcal{C} of subsets of $V(\mathcal{H})$ such that the following hold:

- *for every independent set I in \mathcal{H} , there exists $A \in \mathcal{C}$ such that $I \subseteq A$,*
- $\log |\mathcal{C}| \leq c \cdot |V(\mathcal{H})| \cdot \tau \cdot \log(1/\varepsilon) \cdot \log(1/\tau)$,
- $e_{\mathcal{H}}(A) \leq \varepsilon \cdot |E(\mathcal{H})|$ *for every $A \in \mathcal{C}$.*

We will adapt Theorem 4.1 to our problem. In order to do so, we must view our problem from the perspective of independent sets in a hypergraph. Let us consider the $|X|$ -uniform hypergraph \mathcal{G} with $V(\mathcal{G}) = [n]^d$ and hyperedge set consisting of all copies of X in $[n]^d$. Clearly independent sets in \mathcal{G} are precisely X -free subsets of $[n]^d$. Applying Theorem 4.1 to this hypergraph \mathcal{G} , we immediately get the following corollary.

Corollary 4.2. *For a finite set $X \subseteq \mathbb{N}^d$, let \mathcal{G} be the $|X|$ -uniform hypergraph on n^d vertices encoding the set of all copies of X in $[n]^d$. Suppose that there exist $0 < \varepsilon, \tau < 1/2$ such that*

- $\tau < 1/(200 \cdot |X| \cdot |X|!^2)$

- $\Delta(\mathcal{G}, \tau) \leq \varepsilon/(12|X|!)$.

Then there exists $c = c(|X|) \leq 1000 \cdot |X| \cdot |X|!^3$ and a collection \mathcal{C} of subsets of $V(\mathcal{G})$ such that the following holds:

- for every X -free subset $I \subseteq [n]^d$, there exists $A \in \mathcal{C}$ such that $I \subseteq A$,
- $\log |\mathcal{C}| \leq c \cdot n^d \cdot \tau \cdot \log(1/\varepsilon) \cdot \log(1/\tau)$,
- $e_{\mathcal{G}}(A) \leq \varepsilon \cdot |E(\mathcal{G})|$ for every $A \in \mathcal{C}$.

By combining Lemma 3.7 and Corollary 4.2 we can now prove Theorem 1.4, restated below for convenience.

Theorem 1.4. *Let d be a positive integer and let $X \subseteq \mathbb{N}^d$ be a finite set such that $|X| \geq 3$. For infinitely many $n \in \mathbb{N}$, the number of X -free subsets of $[n]^d$ is $2^{O(r_X(n))}$.*

Proof. Define N to be the infinite subset of \mathbb{N} guaranteed by Lemma 3.7. Consider the $|X|$ -uniform hypergraph \mathcal{G} encoding copies of X in $[n]^d$ and choose $\gamma > 0$ so that $\bar{d}(\mathcal{G}) = \frac{|X||E(\mathcal{G})|}{n^d} \geq \gamma n$ for all n sufficiently large. For $n \in N$, define

$$\varepsilon(n) := \frac{1}{3d} \cdot \frac{\log^{3|X|-2}(n)}{n} \cdot \left(\frac{n^d}{r_X(n)} \right)^{|X|},$$

and

$$\tau(n) := \left(\frac{12|X|! \cdot 2^{|X|^2} \cdot |X|^3}{\gamma n \cdot \varepsilon(n)} \right)^{\frac{1}{|X|-1}}.$$

We wish to apply Corollary 4.2 and so we will verify the conditions for \mathcal{G} with $\varepsilon(n)$, and $\tau(n)$. Clearly, $\varepsilon(n), \tau(n) > 0$ for all n . Using Lemma 3.5, observe for sufficiently large n ,

$$\begin{aligned} \varepsilon(n) &= \frac{1}{3d} \cdot \frac{\log^{3|X|-2}(n)}{n} \cdot \left(\frac{n^d}{r_X(n)} \right)^{|X|} \\ &\leq \frac{1}{3d} \cdot \frac{\log^{3|X|-2}(n)}{n} \cdot \left(\frac{n^d}{n^d \cdot \exp(-c_X \sqrt{\log(n)})} \right)^{|X|} \\ &= \frac{1}{n^{1-o(1)}} < \frac{1}{2}. \end{aligned} \tag{4.3}$$

Also, we clearly have $\frac{1}{\varepsilon(n)} = o(1)$ and so $\tau(n) = o(1)$. Thus, for n sufficiently large, we have

$$\tau(n) < 1/(200 \cdot |X| \cdot |X|!^2) < \frac{1}{2}. \tag{4.4}$$

Next, we wish to bound $\Delta(\mathcal{G}, \tau(n))$. By definition,

$$\Delta(\mathcal{G}, \tau(n)) = 2^{\binom{|X|}{2}-1} \sum_{j=2}^{|X|} 2^{-\binom{j-1}{2}} \frac{\Delta_j(\mathcal{G})}{\tau(n)^{j-1} d(\mathcal{G})}.$$

Every term in the above summation is at most $\frac{\Delta_2(\mathcal{G})}{\tau(n)^{|X|-1} d(\mathcal{G})} \leq \frac{|X|^2}{\tau(n)^{|X|-1} \gamma n}$ and so

$$\Delta(\mathcal{G}, \tau(n)) \leq \frac{2^{|X|^2} \cdot |X|^3}{\tau(n)^{|X|-1} \gamma n} = \frac{\varepsilon(n)}{12|X|!} \quad (4.5)$$

where the equality is by definition of $\tau(n)$.

From (4.3), (4.4) and (4.5), we may apply Corollary 4.2 with $\varepsilon = \varepsilon(n)$ and $\tau = \tau(n)$ to get a constant $c = c(X) \leq 1000 \cdot |X| \cdot |X|^3$ and a collection \mathcal{C} of subsets of $V(\mathcal{G})$ with the following properties:

- for every X -free subset $I \subseteq [n]^d$, there exists $S \in \mathcal{C}$ such that $I \subseteq S$,
- $\log |\mathcal{C}| \leq c \cdot n^d \cdot \tau(n) \cdot \log(1/\varepsilon(n)) \cdot \log(1/\tau(n))$,
- for every $A \in \mathcal{C}$, $e_{\mathcal{G}}(A) \leq \varepsilon(n) \cdot |E(\mathcal{G})|$.

By definition, $\log(1/\varepsilon(n)) = O(\log(n))$ and $\log(1/\tau(n)) = O(\log(n))$. We also have $\frac{n^d \log^3(n)}{r_X(n)} = o\left((n \cdot \varepsilon(n))^{\frac{1}{|X|-1}}\right)$ which implies that $\tau(n) = o\left(\frac{r_X(n)}{n^d \log^3(n)}\right)$. Therefore,

$$\log |\mathcal{C}| = O(n^d \tau(n) \log^2(n)) = o(r_X(n)). \quad (4.6)$$

Note that for every container $A \in \mathcal{C}$, the number of copies of X in A is at most $\varepsilon(n) \cdot |E(\mathcal{G})| < \varepsilon(n) \cdot n^{d+1}$. That is,

$$\Gamma_X(A) \leq \varepsilon(n) \cdot n^{d+1} = \frac{\log^{3|X|-2}(n)}{3d} \cdot \left(\frac{n^d}{r_X(n)}\right)^{|X|} \cdot n^d.$$

So, by Lemma 3.7, every container $A \in \mathcal{C}$ must satisfy

$$|A| < \frac{4}{\alpha} \cdot r_X(n) \quad (4.7)$$

where α is a constant depending only on X . Since every X -free subset of $[n]^d$ is contained in some container in \mathcal{C} , we can use (4.6) and (4.7) to conclude that the number of X -free subsets of $[n]^d$ is at most

$$\sum_{A \in \mathcal{C}} 2^{|A|} \leq |\mathcal{C}| \cdot \max_{A \in \mathcal{C}} 2^{|A|} \stackrel{(4.6), (4.7)}{<} 2^{o(r_X(n))} \cdot 2^{\frac{4}{\alpha} \cdot r_X(n)} = 2^{O(r_X(n))},$$

as desired. □

References

- [1] J. Balogh, S. Das, M. Delcourt, H. Liu, and M. Sharifzadeh, *Intersecting families of discrete structures are typically trivial*, J. Combin. Theory Ser. A **132** (2015), 224–245.
- [2] J. Balogh, H. Liu, and M. Sharifzadeh, *The number of subsets of integers with no k -term arithmetic progression*, Int. Math. Res. Not. IMRN (2017), no. 20, 6168–6186.
- [3] J. Balogh, R. Morris, and W. Samotij, *Independent sets in hypergraphs*, J. Amer. Math. Soc. **28** (2015), no. 3, 669–709.
- [4] ———, *The method of hypergraph containers*, Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. IV. Invited lectures, World Sci. Publ., Hackensack, NJ, 2018, pp. 3059–3092.
- [5] F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U.S.A. **32** (1946), 331–332.
- [6] T. F. Bloom, *A quantitative improvement for Roth’s theorem on arithmetic progressions*, J. Lond. Math. Soc. (2) **93** (2016), no. 3, 643–663.
- [7] T. F. Bloom and O. Sisask, *Logarithmic bounds for Roth’s theorem via almost-periodicity*, Discrete Anal. (2019), Paper No. 4, 20.
- [8] ———, *Breaking the logarithmic barrier in Roth’s theorem on arithmetic progressions*, (2021), E-print arXiv:2007.03528v2.
- [9] P. J. Cameron and P. Erdős, *On the number of sets of integers with various properties*, Number theory (Banff, AB, 1988), de Gruyter, Berlin, 1990, pp. 61–79.
- [10] D. Dong, N. Mani, and Y. Zhao, *On the number of error correcting codes*, Combin. Probab. Comput. **32** (2023), no. 5, 819–832.
- [11] M. Elkin, *An improved construction of progression-free sets*, Israel J. Math. **184** (2011), 93–128.
- [12] P. Erdős and P. Turán, *On Some Sequences of Integers*, J. London Math. Soc. **11** (1936), no. 4, 261–264.
- [13] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256.
- [14] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for commuting transformations*, J. Analyse Math. **34** (1978), 275–291 (1979).
- [15] ———, *A density version of the Hales-Jewett theorem*, J. Anal. Math. **57** (1991), 64–119.

- [16] W. T. Gowers, *A new proof of Szemerédi’s theorem*, Geom. Funct. Anal. **11** (2001), no. 3, 465–588.
- [17] ———, *Hypergraph regularity and the multidimensional Szemerédi theorem*, Ann. of Math. (2) **166** (2007), no. 3, 897–946.
- [18] B. Green and J. Wolf, *A note on Elkin’s improvement of Behrend’s construction*, Additive number theory, Springer, New York, 2010, pp. 141–144.
- [19] Z. Kelley and R. Meka, *Strong bounds for 3-progressions*, E-print arXiv:2302.05537v4, 2023.
- [20] Y. Kim, *The number of k -dimensional corner-free subsets of grids*, Electron. J. Combin. **29** (2022), no. 2, Paper No. 2.53, 19.
- [21] D. Kleitman, *On Dedekind’s problem: The number of monotone Boolean functions*, Proc. Amer. Math. Soc. **21** (1969), 677–682.
- [22] B. Nagle, V. Rödl, and M. Schacht, *The counting lemma for regular k -uniform hypergraphs*, Random Structures Algorithms **28** (2006), no. 2, 113–179.
- [23] K. O’Bryant, *Sets of integers that do not contain long arithmetic progressions*, Electron. J. Combin. **18** (2011), no. 1, Paper 59, 15.
- [24] R. A. Rankin, *Sets of integers containing not more than a given number of terms in arithmetical progression*, Proc. Roy. Soc. Edinburgh Sect. A **65** (1960/61), 332–344 (1960/61).
- [25] V. Rödl and J. Skokan, *Regularity lemma for k -uniform hypergraphs*, Random Structures Algorithms **25** (2004), no. 1, 1–42.
- [26] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 104–109.
- [27] T. Sanders, *On Roth’s theorem on progressions*, Ann. of Math. (2) **174** (2011), no. 1, 619–636.
- [28] D. Saxton and A. Thomason, *Hypergraph containers*, Invent. Math. **201** (2015), no. 3, 925–992.
- [29] T. Schoen, *Improved bound in Roth’s theorem on arithmetic progressions*, Adv. Math. **386** (2021), Paper No. 107801, 20.
- [30] A. Selberg, *An elementary proof of the prime-number theorem*, Ann. of Math. (2) **50** (1949), 305–313.
- [31] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245.

- [32] ———, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS, vol. 260, CNRS, Paris, 1978, pp. 399–401.
- [33] T. Tao, *A variant of the hypergraph removal lemma*, J. Combin. Theory Ser. A **113** (2006), no. 7, 1257–1280.

A Behrend-type constructions in every dimension

The purpose of this appendix is to prove the following extension of the classical result of Behrend [5] to arbitrary sets of cardinality at least three in any dimension. The argument in the proof is standard (similar ideas are used in, e.g., [20, Proof of Lemma 3.3] and [17, Proof of Theorem 10.3]), and we only present it for the sake of completeness.

Lemma 3.5. *Let $d \geq 1$ and $X \subseteq \mathbb{N}^d$ be fixed. If $|X| \geq 3$, then there exists $c_X > 0$ depending only on X such that*

$$r_X(n) \geq n^d \cdot \exp(-c_X \sqrt{\log n}).$$

Before proceeding with the proof of Lemma 3.5, it will be useful to deal with the case $d = 1$ separately. This proof follows the standard argument of Behrend [5].

Lemma A.1. *If $X \subseteq \mathbb{N}$ such that $|X| \geq 3$, then there exists $c_X > 0$ depending only on X such that, for all $n \in \mathbb{N}$,*

$$r_X(n) \geq n \cdot \exp(-c_X \sqrt{\log n}).$$

Proof. Note that $r_X(n) \geq r_{X'}(n)$ for any subset X' of X ; thus, we can assume that $|X| = 3$. Without loss of generality, $X = \{a, b, c\}$ for $a < b < c$. Let $q > 0$ be chosen so that

$$(b - a) = q \cdot (c - b).$$

For $n \geq 1$, define $N := \left\lceil \sqrt{\log(n)} \right\rceil$ and $M = \left\lfloor \frac{n^{1/N}}{q+1} \right\rfloor$. For each vector $\mathbf{x} = (x_1, \dots, x_N) \in [M]^N$, the square of the Euclidean norm, namely $x_1^2 + x_2^2 + \dots + x_N^2$, is an integer in $[N, N \cdot M^2]$. So, by the pigeonhole principle, there exists a choice of r so that r^2 is an integer in $[N, N \cdot M^2]$ and S defined to be the intersection of the sphere of radius r with $[M]^d$ satisfies

$$|S| \geq \frac{M^N}{NM^2 - N + 1} \geq \frac{M^{N-2}}{N}.$$

By definition of M and N , this is at least

$$\frac{1}{N} \left(\frac{n^{1/N}}{2(q+1)} \right)^{N-2} = \frac{n}{n^{2/N} N (2(q+1))^{N-2}} \geq n \cdot \exp(-c_X \sqrt{\log(n)}),$$

where the constant c_X depends only on q . We fix S for the rest of the proof.

Next, define $f : \mathbb{Z}^N \rightarrow \mathbb{Z}$ according to

$$f(x_1, \dots, x_N) = \sum_{i=1}^N x_i ((q+1)M)^{i-1}$$

We define $A := f(S)$. Our goal is to show that A is an X -free subset of $[n]$, at which point we will be done since we have already shown that $|S| \geq n \cdot \exp(-c_X \sqrt{\log(n)})$. Let us start by showing that $A \subseteq [n]$. Clearly every element of A is an integer. Note that, for every $(x_1, \dots, x_N) \in [M]^N$, we have that $f(x_1, \dots, x_N)$ is a positive integer which is bounded above by

$$M \cdot \sum_{i=1}^N ((q+1)M)^{i-1} = M \cdot \left(\frac{(q+1)^N M^N - 1}{(q+1)M - 1} \right).$$

If n is chosen large with respect to X , then we may assume that $qM > 1$ and so the above expression is less than $(q+1)^N M^N$ which, by definition of M , is at most n . Note that we can freely assume that n is large with respect to X since, for any fixed n_0 , we can pick c_X large enough that $n^d \exp(-c_X \sqrt{\log(n)}) < 1 \leq r_X(n)$ for all $n < n_0$, and so it suffices to consider $n \geq n_0$. Therefore, $A \subseteq [n]$.

Next, let us show that $|A| = |S|$. For this, we show that f is injective on $[M]^N$. Let $\mathbf{x} = (x_1, \dots, x_N)$ and $\mathbf{y} = (y_1, \dots, y_N)$ be distinct elements of $[M]^N$ and let j be the largest index such that $x_j \neq y_j$. Then, by the triangle inequality,

$$\begin{aligned} |f(\mathbf{x}) - f(\mathbf{y})| &\geq |x_j - y_j|((q+1)M)^{j-1} - \sum_{i=1}^{j-1} |x_i - y_i|((q+1)M)^{i-1} \\ &\geq ((q+1)M)^{j-1} - M \left(\frac{(q+1)^{j-1} M^{j-1} - 1}{(q+1)M - 1} \right) > 0 \end{aligned}$$

which implies that $f(\mathbf{x}) \neq f(\mathbf{y})$. Therefore, $|A| = |S|$.

Finally, we show that A is X -free. If not, then there must exist \mathbf{x}, \mathbf{y} and \mathbf{z} in S such that $f(\mathbf{x}) < f(\mathbf{y}) < f(\mathbf{z})$ and $f(\mathbf{y}) - f(\mathbf{x}) = q(f(\mathbf{z}) - f(\mathbf{y}))$. In other words, $(q+1)f(\mathbf{y}) - f(\mathbf{x}) - qf(\mathbf{z}) = 0$. Since S is a subset of a sphere, S is X -free and so it cannot be the case that $(q+1)\mathbf{y} - \mathbf{x} - q\mathbf{z} = \mathbf{0}$, and so we can let j be the largest index such that $(q+1)y_j - x_j - qz_j \neq 0$. Now, by the triangle inequality,

$$\begin{aligned} &|(q+1)f(\mathbf{y}) - f(\mathbf{x}) - qf(\mathbf{z})| \\ &\geq |(q+1)y_j - x_j - qz_j|((q+1)M)^{j-1} - \sum_{i=1}^{j-1} |(q+1)y_i - x_i - qz_i|((q+1)M)^{i-1} \\ &\geq ((q+1)M)^{j-1} - ((q+1)M - q - 1) \left(\frac{(q+1)^{j-1} M^{j-1} - 1}{(q+1)M - 1} \right) > 0 \end{aligned}$$

which contradicts the assumption that $(q+1)f(\mathbf{y}) - f(\mathbf{x}) - qf(\mathbf{z}) = 0$. Therefore A is X -free and the proof is complete. \square

We now use Lemma A.1 to prove Lemma 3.5.

Proof of Lemma 3.5. If $d = 1$, then we are done by Lemma A.1, and so we assume that $d \geq 2$. Since $r_X(n) \geq r_{X'}(n)$ for any subset X' of X , we may assume that $X = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$. Our goal is to construct, for each $n \geq 1$, a “large” set $S \subseteq [n]^d$ which is X -free. We may assume throughout the proof that n is divisible by $4(d-1)$.

We start by making numerous (valid) assumptions about the vectors \mathbf{x}, \mathbf{y} and \mathbf{z} . Let T be the convex hull of $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ and note that T is a (possibly degenerate) triangle. By relabelling \mathbf{x}, \mathbf{y} and \mathbf{z} if necessary, we can assume that neither of the interior angles of T at \mathbf{x} nor \mathbf{y} are equal to $\pi/2$; this is because any triangle can contain at most one interior angle of $\pi/2$. Note that, in the special case that \mathbf{x}, \mathbf{y} and \mathbf{z} are collinear, T is a degenerate triangle and the interior angles at \mathbf{x} and \mathbf{y} are either 0 or π and so this property holds automatically. Now, by translating X if necessary, we may assume that \mathbf{z} is the zero vector. Let the entries of \mathbf{x} and \mathbf{y} be given by $\mathbf{x} = (x_1, \dots, x_d)$ and $\mathbf{y} = (y_1, \dots, y_d)$. Given that $\mathbf{z} = \mathbf{0}$, the condition that the interior angles of T at \mathbf{x} and \mathbf{y} are not equal to $\pi/2$ can be expressed algebraically as follows:

$$\langle \mathbf{x}, \mathbf{x} - \mathbf{y} \rangle = \sum_{i=1}^d x_i(x_i - y_i) \neq 0, \quad (\text{A.2})$$

$$\langle \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = \sum_{i=1}^d y_i(x_i - y_i) \neq 0, \quad (\text{A.3})$$

where $\langle \cdot, \cdot \rangle$ is the standard inner product on \mathbb{R}^d . Since $\mathbf{x} \neq \mathbf{y}$, by reordering the indices, rescaling X and swapping \mathbf{x} and \mathbf{y} if necessary, we can assume that $x_d - y_d = 1$ and that $|x_i - y_i| \leq 1$ for all $i \in [d]$. Note that the rescaling may put \mathbf{x} and \mathbf{y} in \mathbb{Q}^d . For each $i \in [d-1]$, let $x_i - y_i = r_i/q_i$ such that r_i and $q_i > 0$ are integers with q_i as small as possible. Note that $q := \max_{i \in [d-1]} q_i$ is a constant depending on X .

Next, we define

$$t := \frac{\sum_{i=1}^d y_i(y_i - x_i)}{\sum_{i=1}^d (x_i - y_i)^2}.$$

Note that t is well-defined because $\mathbf{x} \neq \mathbf{y}$, and that $t \in \mathbb{Q}$ because $\mathbf{x}, \mathbf{y} \in \mathbb{Q}^d$. Moreover, t can be regarded as a quantity which depends only on X . Also, (A.3) implies that $t \neq 0$ and (A.2) implies that $t \neq 1$. So, $X' := \{0, t, 1\}$ is a subset of \mathbb{Q} of cardinality 3. Using Lemma A.1, we let S' be a subset of $[n/4]$ which is X' -free and satisfies

$$|S'| \geq (n/4) \cdot e^{-c_{X'} \sqrt{\log(n/4)}}.$$

Note that X' depends on t which, in turn, depends on X , and so $c_{X'}$ is a constant that depends on X . Define $S \subseteq [n]^d$ by

$$S := \left\{ \mathbf{s} = (s_1, \dots, s_d) \in [n]^d : \sum_{i=1}^d \left(s_i - \frac{n}{2} \right) (x_i - y_i) \in S' \right\}.$$

To complete the proof, we need to analyze the cardinality of S and show that it is X -free. We start with its cardinality. For each $i \in [d-1]$, since $|x_i - y_i| \leq 1$, the number of choices for $s_i \in [n]$ such that

$$-\frac{n}{4(d-1)} \leq \left(s_i - \frac{n}{2}\right)(x_i - y_i) \leq 0$$

is at least $\frac{n}{4(d-1)}$. Since $x_i - y_i$ is a rational with denominator bounded by the constant q for all $i \in [d-1]$, the number of such s_i for which $\left(s_i - \frac{n}{2}\right)(x_i - y_i)$ is also an integer is at least $\left\lfloor \frac{n}{4q(d-1)} \right\rfloor$. Therefore, the number of choices for $(s_1, \dots, s_{d-1}) \in [n]^{d-1}$ such that

$$\sum_{i=1}^{d-1} \left(s_i - \frac{n}{2}\right)(x_i - y_i) \in \mathbb{Z} \cap \left[-\frac{n}{4}, 0\right]$$

is $\Omega(n^{d-1})$, where the constant factor depends only on d and q (which, in turn, depends on X). Recalling that $x_d - y_d = 1$ and that $S' \subseteq [n/4]$, we see that, for any such (s_1, \dots, s_{d-1}) and any $a \in S'$, there is a unique choice of $s_d \in [n]$ so that

$$\sum_{i=1}^d \left(s_i - \frac{n}{2}\right)(x_i - y_i) = a.$$

Thus, we get that

$$|S| = \Omega(n^{d-1}|S'|) = n^d \cdot e^{-c_X \sqrt{\log n}}$$

for a sufficiently small constant $c_X > 0$ depending on X , as desired.

Finally, we show that S is X -free. Suppose, to the contrary, that there exists a non-trivial copy of X in S . In other words, there exists $\mathbf{b} \in \mathbb{Q}^d$ and $r \in \mathbb{Q} \cap (0, \infty)$ such that

$$\{\mathbf{b}, \mathbf{b} + r\mathbf{x}, \mathbf{b} + r\mathbf{y}\} \subseteq S.$$

Let $\mathbf{b} = (p_1, \dots, p_d)$ and define

$$\begin{aligned} a &:= \sum_{i=1}^d \left(p_i - \frac{n}{2}\right)(x_i - y_i), \\ b &:= \sum_{i=1}^d \left(p_i + rx_i - \frac{n}{2}\right)(x_i - y_i), \\ c &:= \sum_{i=1}^d \left(p_i + ry_i - \frac{n}{2}\right)(x_i - y_i). \end{aligned}$$

The condition $\{\mathbf{b}, \mathbf{b} + r\mathbf{x}, \mathbf{b} + r\mathbf{y}\} \subseteq S$ is equivalent to $a, b, c \in S'$. We claim that $\{a, b, c\}$ is a non-trivial copy of $\{0, t, 1\}$, which will contradict the definition of S' and complete the proof. Since $r \neq 0$, we have

$$\frac{a - b}{c - b} = \frac{r \sum_{i=1}^d x_i(y_i - x_i)}{r \sum_{i=1}^d (x_i - y_i)^2} = t.$$

Therefore

$$b + (c - b)\{0, t, 1\} = b + \{0, a - b, c - b\} = \{a, b, c\}$$

and so $\{a, b, c\}$ is a copy of $\{0, t, 1\}$. It is non-trivial because $r \neq 0$ and $\mathbf{x} \neq \mathbf{y}$ implies $c - b = r \cdot \sum_{i=1}^d (x_i - y_i)^2$ is non-zero. This contradiction completes the proof. \square