

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. **TOPROVE 0** □

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. **TOPROVE 1** □

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. **TOPROVE 2** □

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 3.2. *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 3.2, observe that, for sufficiently large n ,

$$\begin{aligned} m(n) &= \frac{n}{\log^{3|X|}(n)} \left(\frac{r_X(n)}{n^d} \right)^{|X|+2} \geq \frac{n}{\log^{3|X|}(n)} \left(\exp(-c_X \sqrt{\log n}) \right)^{|X|+2} \\ &> n \cdot \left(\exp \left(-(5c_X |X|) \sqrt{\log n} \right) \right). \end{aligned} \tag{3.3}$$

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 3.4. *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. TOPROVE 3 □

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 3.4 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. TOPROVE 4 □

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 3.2. *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 3.2, 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. **TOPROVE 5**

□

We now use Lemma A.1 to prove Lemma 3.2.

Proof. **TOPROVE 6**

□