DEVIATION PROBABILITIES FOR ARITHMETIC PROGRESSIONS AND OTHER REGULAR DISCRETE STRUCTURES

GONZALO FIZ PONTIVEROS, SIMON GRIFFITHS, MATHEUS SECCO, AND ORIOL SERRA

ABSTRACT. Let the random variable $X := e(\mathcal{H}[B])$ count the number of edges of a hypergraph \mathcal{H} induced by a random m element subset B of its vertex set. Focussing on the case that \mathcal{H} satisfies some regularity condition we prove bounds on the probability that X is far from its mean. It is possible to apply these results to discrete structures such as the set of k-term arithmetic progressions in the cyclic group \mathbb{Z}_N . Furthermore, we show that our main theorem is essentially best possible and we deduce results for the case $B \sim B_p$ is generated by including each vertex independently with probability p.

1. INTRODUCTION

The problem of determining how well a random variable X is concentrated around its mean $\mathbb{E}[X]$ has a long history and is of great interest in many areas of mathematics. In the area of Combinatorics this question most frequently arises when X counts the number of occurrences of some substructure. The cases of subgraphs in a random graph and arithmetic progressions in a random set have been studied extensively in recent years, as we discuss below. There is a common framework in which these problems may be viewed: Given a hypergraph \mathcal{H} on $[N] := \{1, \ldots, N\}$ and a distribution on subsets $B \subseteq [N]$ we may ask how well concentrated is

$$X := e(\mathcal{H}[B])$$

the number of edges of \mathcal{H} contained in B.

Work on this problem has focussed on the case that $B \sim B_p$ is a *p*-random subset of [N] (each element included independently with probability *p*). In this case general bounds on deviation probabilities of the random variable $X := e(\mathcal{H}[B_p])$ follow from the famous inequality of Kim and Vu [23]. In the case of large deviations (of the order of the mean) further progress has been made by Janson and Ruciński [21], who determined (under certain conditions) the log probability $\log(\mathbb{P}(X > (1 + \delta)\mathbb{E}[X]))$ up to a factor of order $\log 1/p$, and by Warnke [31, 32] who determined the log probability up to a constant factor (in which the constant may depend on δ). Recently Bhattacharya and Mukherjee [4] have shown how

Acknowledgements: G.F.P. was supported by BGSMath Postdoctoral Grant and the Spanish Research Agency under projects MDM-2014-0445 and MTM2017-82166, S.G. was supported by CNPq bolsa de produtividade em pesquisa (Proc. 310656/2016-8) and FAPERJ Jovem cientista do nosso estado (Proc. 202.713/2018), M.S. was supported as a PhD student by CAPES and O.S. was supported by grants MDM-2014-0445 and MTM2017-82166-P of the Spanish Ministry of Science.

the large deviations framework introduced by Chatterjee and Varadhan [9] (in the context of subgraph counts) may be used to understand when replica symmetry is broken.

We focus on the case $B \sim B_m$ is a uniformly random m element subset of $[N] = \{1, \ldots, N\}$. We prove upper bounds on the probability of deviations of the random variable $e(\mathcal{H}[B_m])$, which are particularly relevant in the case that the hypergraph \mathcal{H} is highly regular. Furthermore, since B_p , the *p*-random subset, may be obtained by averaging over the models B_m where m is selected according to Bin(N, p), we may deduce results in the context of p-random sets.

Given a k-uniform hypergraph \mathcal{H} on vertex set [N] and a subset $B \subseteq [N]$ we set

$$N^{\mathcal{H}}(B) := e(\mathcal{H}[B]),$$

the number of edges of \mathcal{H} contained in the set B. We also define

$$L^{\mathcal{H}}(m) := \mathbb{E}\left[N^{\mathcal{H}}(B_m)\right]$$

the expected value of $N^{\mathcal{H}}(B_m)$ where B_m is a uniformly selected *m* element subset of [N]. Our focus will be on studying

$$D^{\mathcal{H}}(B_m) := N^{\mathcal{H}}(B_m) - L^{\mathcal{H}}(m),$$

the deviation of $N^{\mathcal{H}}(B_m)$ from its mean.

We may now state our bounds on the probability of certain deviations $D^{\mathcal{H}}(B_m)$. We say that a hypergraph \mathcal{H} is r-tuple-regular¹ if all r-tuples of vertices are included in the same number of edges. Note that all hypergraphs are 0-tuple-regular, and a hypergraph is 1tuple-regular if it is regular in the usual sense that each vertex is in the same number of edges. We also note that, by a simple double counting argument, r-tuple regular implies r'-tuple regular for all $0 \leq r' \leq r$.

Theorem 1.1. Let $1 \leq r \leq k$. Let \mathcal{H} be a k-uniform hypergraph on [N]. Suppose that \mathcal{H} is (r-1)-tuple-regular with maximum r-degree Δ_r . Then

$$\mathbb{P}\left(|D^{\mathcal{H}}(B_m)| > a\right) \leqslant N^{O_k(1)} \exp\left(\frac{-\Omega_k(1)a^{2/r}}{m\Delta_r^{2/r}}\right)$$

for all a > 0.

Remark 1. We focus throughout the article on hypergraphs rather than even more general structures such as weighted hypergraphs or polynomials, that were considered by Kim and Vu [23]. In fact, the proof goes through with no major changes for the case of hypergraphs with positively weighted edges. Since a general weighted hypergraph may be written as a difference of two with positive weights, one may deduce results in this setting. Similarly, a bounded degree polynomial may be broken into a finite number of homogeneous polynomials which correspond to weighted hypergraphs.

¹Equivalently, in the language of combinatorial designs, \mathcal{H} is an $r - (N, k, \lambda)$ design, for some λ .

Unfortunately some natural hypergraphs are not precisely *r*-tuple-regular, but just very close to being so. Let $\bar{d}_r = \bar{d}_r(\mathcal{H})$ denote the average degree of *r*-sets in a hypergraph \mathcal{H} . We say that \mathcal{H} is (r, η) -near-regular if every *r*-tuple of vertices is contained in $(1 \pm \eta)\bar{d}_r$ edges. Obviously, a hypergraph which is *r*-tuple-regular is (r, η) -near-regular for all $\eta \ge 0$. In particular, Theorem 1.1 is a special case $(\eta = 0)$ of the following theorem.

Theorem 1.2. Let $1 \leq r \leq k$ and let $\eta \in [0, 3^{-r+1}]$. Let \mathcal{H} be a k-uniform hypergraph on [N]. Suppose that \mathcal{H} is $(r-1, \eta)$ -near-regular with maximum r-degree Δ_r . Then

$$\mathbb{P}\left(|D^{\mathcal{H}}(B_m)| > a\right) \leqslant N^{O_k(1)} \exp\left(\frac{-\Omega_k(1)a^{2/r}}{m\Delta_r^{2/r}}\right)$$

for all

$$a \geq (10k!)^{10^r} e(\mathcal{H}) \left(\frac{\eta m^{k-1}}{N^{k-1}}\right)^{r/(r-1)}$$

Remark 2. In the dense case $(m = \Theta(N))$, the bound given by Theorem 1.2 is best possible (up to the constant implicit in $\Omega_k(1)$) for all $r \ge 0$ and across the whole range $\Theta(1) \le \Delta_r \le \Theta(N)$. Examples are given in Section 6.

Remark 3. One may easily observe by a double counting argument that an $(r-1, \eta)$ -near-regular hypergraph \mathcal{H} is $(r'-1, \eta)$ -near-regular for all $1 \leq r' \leq r$. So one may choose which of the above inequalities to apply. Therefore

$$\mathbb{P}\left(D^{\mathcal{H}}(B_m) > a\right) \leqslant N^{O_k(1)} \min_{1 \leqslant r' \leqslant r} \left\{ \exp\left(\frac{-\Omega_k(1)a^{2/r'}}{m\Delta_{r'}^{2/r'}}\right) \right\} .$$

It is worth remarking that the minimum is not always obtained at the extremes $r' \in \{1, r\}$, see the application to arithmetic progressions in Section 5 for example.

Remark 4. The lower bound condition on a given in Theorem 1.2 is best possible in the weak sense that there exist hypergraphs for which it cannot be significantly improved. However, with extra conditions one may weaken the condition on a. In particular, the condition in Theorem 1.2 may be weakened to

$$a \geq (10k!)^{10^r} e(\mathcal{H}) \left(\frac{\eta m^{k-2}}{N^{k-2}}\right)^{r/(r-2)}$$

if \mathcal{H} is regular and $(r-1, \eta)$ -near-regular, for $r \ge 3$. This result is stated as Proposition 3.5 in Section 3.4.

The existing literature has focussed on the probability of deviations of $e(\mathcal{H}(B_p))$ in the model of a *p*-random subset $B_p \subseteq [N]$, in which each element is included in B_p independently with probability *p*. As usual in probabilistic combinatorics we allow for the possibility that p = p(n) is a function of *n*. Our results apply for a range of moderate deviations, which are smaller than the order of magnitude of the mean.

4 GONZALO FIZ PONTIVEROS, SIMON GRIFFITHS, MATHEUS SECCO, AND ORIOL SERRA

For the range of deviations covered by the theorem below it precisely provides the asymptotic log probability of the corresponding deviation event. For comparison, the bounds given by the Kim-Vu inequality [23] are not so strong, however they often apply for a much larger range of deviations.

Theorem 1.3. Let $k \ge r \ge 2$. Let \mathcal{H}_N be a sequence of k-uniform hypergraphs which are $(r-1,\eta)$ -near-regular with maximum r-degree Δ_r and $V(\mathcal{H}_N) = [N]$. Let δ_N be a sequence satisfying

$$\max\left\{\frac{\Delta_r(N\log N)^{r/2}}{p^{k-r/2}e(\mathcal{H})}, \left(\eta^r p^{k-r}\right)^{1/(r-1)}, \frac{1}{\sqrt{pN}}\right\} \ll \delta_N \ll \left(\frac{p^{k-r}e(\mathcal{H})}{N^r\Delta_r}\right)^{1/(r-1)}$$

,

where $0 \leq p \leq 1$ is bounded away from 1. Then

$$\mathbb{P}\left(|D^{\mathcal{H}_N}(B_p)| > \delta_N L^{\mathcal{H}_N}(B_p)\right) = \exp\left(-(1+o(1))\frac{\delta_N^2 p N}{2k^2(1-p)}\right).$$

Remark 5. As in Remark 4, we may be able extend the range of δ_N if we also know that \mathcal{H} is regular as well as $(r-1,\eta)$ -near-regular. Using Proposition 3.5 in place of Theorem 1.2 one may replace the term $(\eta^r p^{k-r})^{1/(r-1)}$ by $(\eta^r p^{2k-2r})^{1/(r-2)}$ in the lower bound of δ_N in Theorem 1.3.

When Δ_r is much larger than $e(\mathcal{H})/N^r$, the order of magnitude of the average *r*-degree, the range of deviations covered by the above theorem may be disappointing. In particular this theorem does not appear to give new results when applied in the context of subgraph counts in the Erdős-Rényi random graph G(n, p). For results concerning moderate deviations for subgraph counts the reader may consult [12, 13, 14, 22]. There is a large literature dedicated to problems of large deviations of subgraph counts, we encourage the interested reader to consult the survey of Chatterjee [10], the recent article of Harel, Mousset and Samotij [17], and the references therein.

On the other hand the condition on δ_N in Theorem 1.3 simplifies significantly if Δ_r is of the same order as the average *r*-degree. For example, when r = 2 and the hypergraph is regular the condition simplifies to

$$\max\left\{\frac{\log N}{p^{k-1}N}, \frac{1}{\sqrt{pN}}\right\} \ll \delta_N \ll p^{k-2}$$

As an immediate application, consider 3-term arithmetic progressions in the cyclic group \mathbb{Z}/\mathbb{NZ} , for N prime. Let us write $D^3(B_p)$ for the deviation of the 3-progressions count in a p-random subset B_p of \mathbb{Z}/\mathbb{NZ} . Note that the expected number of such arithmetic progressions is $L^3(B_p) = p^3 \binom{N}{2}$.

Theorem 1.4. Let δ_N be a sequence satisfying

$$\max\left\{\frac{\log N}{p^2 N}, \frac{1}{\sqrt{pN}}\right\} \ll \delta_N \ll p.$$

Then,

$$\mathbb{P}\left(D^{3}(B_{p}) > \delta_{N}L^{3}(B_{p})\right) = \exp\left(-(1+o(1))\frac{\delta_{N}^{2}pN}{18(1-p)}\right).$$

Furthermore, the same bounds apply to the corresponding negative deviations.

See Section 5 for a discussion of Theorem 1.4 and analogous results.

We remark that the results above would not hold for hypergraphs without some kind of regularity property. Related results for the non-regular setting, which includes arithmetic progressions in $[N] = \{1, \ldots, N\}$ for example, have recently been obtained by Christoph Koch together with the second and third authors [15].

1.1. Layout of the article. In Section 2 we prepare for the proof of Theorem 1.2 by stating a number of auxiliary results which we require. In particular we state a martingale representation for the deviation $D^{\mathcal{H}}(B_m)$, which may be of independent interest. This section also introduces notation and required inequalities from probability theory.

In Section 3 we complete the proof of Theorem 1.2. This is achieved by a double induction argument which links deviation probabilities to the increments of the martingale and vice-versa in successive steps. In Section 4 we show how Theorem 1.3 may be deduced from Theorem 1.2.

Section 5 illustrates possible applications of Theorem 1.2 to obtain deviation probabilities for a variety of arithmetic structures. These include k-arithmetic progressions, Schur triples, additive energy and, more generally, solutions of linear systems in random sets. We also include a direct proof of the deviation result for 3-term arithmetic progressions which is simpler than the general case (Theorem 1.2). Since many of the same ideas arise, the reader may wish to read this proof before that of Theorem 1.2. See Theorem 5.2, and its proof.

In Section 6 we prove that the bound given in Theorem 1.2 is best possible up to the implicit constant $\Omega_k(1)$. The proof is based on a construction of hypergraphs with certain regularity properties. The details of the construction are given in the appendix.

Finally, in Section 7 we give some concluding remarks and discuss some related open questions.

2. Overview and Auxiliary results

In this section we introduce notation and auxiliary results which we require for the proof of Theorem 1.2. In particular we introduce a martingale representation for $D^{\mathcal{H}}(B_m)$ (see Proposition 2.1) which is the basis of our proof of Theorem 1.2. This representation is similar in spirit to that introduced in [14] in the context of subgraph counts. The section is divided into Section 2.1 in which we introduce notation, Section 2.2 in which we state auxiliary results and Section 2.3 in which we state some inequalities from probability theory. The proofs of the auxiliary results of Section 2.2 are presented in Section 2.4.

2.1. Notation. We will continue to use N for the number of vertices of the hypergraph \mathcal{H} . We set $h := e(\mathcal{H})$.

We will continue to use m for the number of elements in the random subset $B_m \subseteq [N]$. We denote its density by t := m/N. We sometimes consider the set B_i at an earlier step i of the process. We denote its density s := i/N.

Given a vertex x of a hypergraph \mathcal{H} we define

$$\mathcal{H}(x) := \{ f \setminus \{x\} : f \in E(\mathcal{H}), x \in f \}$$

In the case that \mathcal{H} is a k-uniform hypergraph on [N], then $\mathcal{H}(x)$ is a (k-1)-uniform hypergraph on $[N] \setminus \{x\}$.

In addition to $N^{\mathcal{H}}(B_m)$, the number of edges of \mathcal{H} in B_m , we shall also consider partially filled edges (with multiplicity). Let $N_j^{\mathcal{H}}(B_m)$ be the number of *j*-subsets of edges of \mathcal{H} which are contained in B_m . Equivalently,

$$N_j^{\mathcal{H}}(B_m) := \sum_{f \in E(\mathcal{H})} \binom{|f \cap B_m|}{j}.$$

For a k-uniform hypergraph \mathcal{H} we have $N^{\mathcal{H}}(B_m) = N_k^{\mathcal{H}}(B_m)$.

Throughout the article we denote by

$$L_j^{\mathcal{H}}(m) := \mathbb{E}\left[N_j^H(B_m)\right] = h\binom{k}{j}\frac{(m)_j}{(N)_j}, \quad \text{and}$$
(1)

$$D_j^{\mathcal{H}}(B_m) := N_j^{\mathcal{H}}(B_m) - L_j^{\mathcal{H}}(m), \qquad (2)$$

the mean and deviation of $N_i^{\mathcal{H}}(B_m)$ respectively.

In order to define the increments of the key martingale representation, we introduce²

$$X_{\ell}^{\mathcal{H}}(B_i) := N_{\ell}^{\mathcal{H}}(B_i) - \mathbb{E}\left[N_{\ell}^{\mathcal{H}}(B_i) \mid B_{i-1}\right].$$

Since $N_{\ell}^{\mathcal{H}}(B_{i-1})$ is determined by B_{i-1} we observe that

$$X_{\ell}^{\mathcal{H}}(B_i) = A_{\ell}^{\mathcal{H}}(B_i) - \mathbb{E}\left[A_{\ell}^{\mathcal{H}}(B_i) \mid B_{i-1}\right],$$

²Implicit in this notation is the fact that we regard B_i as a set together with the information of the order in which the points were added.

where

$$A_{\ell}^{\mathcal{H}}(B_i) := N_{\ell}^{\mathcal{H}}(B_i) - N_{\ell}^{\mathcal{H}}(B_{i-1})$$

denotes the increase in $N_{\ell}^{\mathcal{H}}(B_i)$ with the addition of the *i*-th element. Both the above expressions for $X_{\ell}^{\mathcal{H}}(B_i)$ will be used during the proof. We observe that, by its definition, the sequence $X_{\ell}^{\mathcal{H}}(B_i)$ is a difference martingale for each ℓ .

When the hypergraph \mathcal{H} is clear from context we drop it from the notation. That is, we write $N_j(B_m), D_j(B_m), A_\ell(B_i), X_\ell(B_i)$, etc.

It will sometimes be convenient to write $x \in a \pm b$ to express that x belongs to the interval [a - b, a + b].

2.2. Auxiliary results. The most significant auxiliary result we state here is the martingale representation for $D_i^{\mathcal{H}}(B_m)$.

Proposition 2.1. Let \mathcal{H} be a k-uniform hypergraph and let $1 \leq j \leq k$. Then

$$D_{j}^{\mathcal{H}}(B_{m}) = \sum_{i=1}^{m} \sum_{\ell=1}^{j} \frac{(N-m)_{\ell}(m-i)_{j-\ell}}{(N-i)_{j}} \binom{k-\ell}{k-j} X_{\ell}^{\mathcal{H}}(B_{i}).$$
(3)

The equation (3) expresses $D_j^{\mathcal{H}}(B_m)$ as a linear combination of difference martingales and it is therefore a martingale.

It will also be useful to note the ways in which η -near-regularity is inherited.

Lemma 2.2. Let \mathcal{H} be a k-uniform (r, η) -near-regular hypergraph with maximum (r + 1)degree Δ_{r+1} , for some $\eta \in [0, 1/3]$. Let $x \in V(\mathcal{H})$. Then

(i) \mathcal{H} is $(r-1,\eta)$ -near-regular

(ii) $\mathcal{H}(x)$ is $(r-1, 3\eta)$ -near-regular with maximum r-degree at most Δ_{r+1} .

A major part of the proof of Theorem 1.2 involves controlling the size of the increments in the martingale representation. In some cases we can control $|X_{\ell}^{\mathcal{H}}(B_i)|$ directly and deterministically.

Lemma 2.3. Let \mathcal{H} be a k-uniform hypergraph on [N] which is (r, η) -near-regular. If $1 \leq \ell \leq r$, then

$$|X_{\ell}^{\mathcal{H}}(B_i)| \leq \frac{2\ell\binom{k}{\ell}\eta s^{\ell-1}h}{N}.$$

Finally, given a B_i -measurable event E let E^- be the B_{i-1} -measurable event that E occurs for some extension $B_{i-1} \cup \{x\}$ of B_{i-1} . Since the conditional probability $\mathbb{P}(E|E^-)$ is between 1/N and 1 we immediately obtain the following.

Lemma 2.4. Let E be a B_i measurable then

$$\mathbb{P}(E) \leqslant \mathbb{P}(E^{-}) \leqslant N\mathbb{P}(E) .$$

2.3. Martingale Inequalities. Our proofs will rely on the Azuma–Hoeffding [1, 18] martingale inequality.

Lemma 2.5 (Azuma–Hoeffding inequality). Let $(S_i)_{i=0}^m$ be a martingale with increments $(X_i)_{i=1}^m$, and let $c_i = ||X_i||_{\infty}$ for each $1 \leq i \leq m$. Then, for each a > 0,

$$\mathbb{P}\left(S_m - S_0 > a\right) \leqslant \exp\left(\frac{-a^2}{2\sum_{i=1}^m c_i^2}\right).$$

Furthermore, the same bound holds for $\mathbb{P}(S_m - S_0 < -a)$.

By considering a "truncation" of the increments in which X_i is set to 0 if it could be larger than c_i with positive probability then one immediately obtains the following straightforward variant. We state the lemma in our context of a sequence of random sets $(B_i)_{i=0}^N$. This process may be defined by taking b_1, \ldots, b_N to be a uniformly random permutation of $\{1, \ldots, N\}$ and setting $B_i = \{b_1, \ldots, b_i\}$ for each $i = 0, \ldots N$.

Lemma 2.6 (Azuma–Hoeffding inequality (a variant)). Let $(S_i)_{i=0}^m$ be a martingale with respect to the natural filtration of the process $(B_i)_{i=0}^N$, let $(X_i)_{i=1}^m$ be the increments of the process and let $(c_i)_{i=1}^N$ be a sequence of real numbers. Then, for each a > 0,

$$\mathbb{P}\left(S_m - S_0 > a\right) \leqslant \exp\left(\frac{-a^2}{2\sum_{i=1}^m c_i^2}\right) + N\sum_{i=1}^m \mathbb{P}\left(|X_i| > c_i\right)$$

Furthermore, the same bound holds for $\mathbb{P}(S_m - S_0 < -a)$.

Proof. We first define for $1 \leq i \leq m$ the "truncation" X_i^* of the increment X_i as

$$X_i^* := X_i 1_{\|X_i\| B_{i-1}\|_{\infty} \leq c_i}$$

Let us define a new process $(S_j^*)_{j=0}^m$ by $S_0^* := S_0$ and for $1 \leq j \leq m$:

$$S_j^* := S_0^* + \sum_{i=1}^j X_i^*.$$

Since the event considered by the indicator function is B_{i-1} -measurable we have $\mathbb{E}[X_i^* | B_{i-1}] = 0$ and so $(S_j^*)_{j=0}^m$ is a martingale with respect to the natural filtration of the process $(B_i)_{i=0}^N$. Note also that the increments of this process satisfy $|X_i^*| \leq c_i$ almost surely. Therefore the Azuma-Hoeffding inequality gives us

$$\mathbb{P}\left(S_m^* - S_0^* > a\right) \leqslant \exp\left(\frac{-a^2}{2\sum_{i=1}^m c_i^2}\right).$$

$$\tag{4}$$

We also observe that by union bound and Lemma 2.4

$$\mathbb{P}\left(S_{m}^{*} \neq S_{m}\right) \leqslant \sum_{i=1}^{m} \mathbb{P}\left(X_{i}^{*} \neq X_{i}\right)$$
$$= \sum_{i=1}^{m} \mathbb{P}\left(\|X_{i}|B_{i-1}\|_{\infty} > c_{i}\right)$$
$$\leqslant N \sum_{i=1}^{m} \mathbb{P}\left(|X_{i}| > c_{i}\right).$$
(5)

Finally we have

$$\mathbb{P}(S_m - S_0 > a) \leqslant \mathbb{P}(S_m^* - S_0^* > a) + \mathbb{P}(S_m^* \neq S_m)$$

and so by (4) and (5) we get the desired result.

2.4. Proofs of auxiliary results. We now prove Proposition 2.1 and Lemmas 2.2 and 2.3.

Proof of Proposition 2.1. Fix the hypergraph \mathcal{H} . We prove the required expression (3) by a double induction over m and j. The base cases j = 0 and m = 0 are trivial.

For the induction step we may assume that (3) holds if j' < j or if j' = j and m' < m. The argument proceeds by focussing on the latest point added. We recall that $X_j(B_m) = A_j(B_m) - \mathbb{E}[A_j(B_m)|B_{m-1}]$ and that $A_j(B_m) = N_j(B_m) - N_{j-1}(B_{m-1})$ counts the increase in $N_j(B_m)$ with the addition of the *m*-th element of B_m . Considering that any such increase must consist of a (j - 1)-subset together with an extra element of the same edge (which is not already present) and each vertex has probability 1/(N - m + 1) to be selected next, we have that

$$\mathbb{E}[A_j(B_m)|B_{m-1}] = \frac{(k-j+1)N_{j-1}(B_{m-1}) - jN_j(B_{m-1})}{N-m+1}$$

We will use this expression to find a suitable expression for $D_j(B_m)$ in terms of the deviations $D_{j-1}(B_{m-1})$, $D_j(B_{m-1})$ and $X_j(B_m)$. The first step will be to express D_j as $N_j - L_j$. To reach the later steps we expand $N_j(B_m)$ as $A_j(B_m) + N_j(B_{m-1})$ and when possible express N_j as $L_j + D_j$ and use the identity

$$\frac{(k-j+1)L_{j-1}(B_{m-1}) - jL_j(B_{m-1})}{N-m+1} = L_j(m) - L_j(m-1).$$

We obtain the following expression for $D_i(B_m)$:

$$D_{j}(B_{m}) = N_{j}(B_{m}) - L_{j}(m)$$

= $A_{j}(B_{m}) + D_{j}(B_{m-1}) - (L_{j}(m) - L_{j}(m-1))$
= $\frac{N - m - j + 1}{N - m + 1} D_{j}(B_{m-1}) + \frac{(k - j + 1)}{N - m + 1} D_{j-1}(B_{m-1}) + X_{j}(B_{m}).$

The required expression (3) now follows immediately from the induction hypothesis by simply checking the coefficient of each $X_{\ell}(B_i)$. This may be verified easily by checking the cases (i) i = m and $\ell = j$, (ii) i = m and $\ell < j$, (iii) i < m and $\ell = j$, and (iv) i < m and $\ell < j$.

In case (i), the coefficients on each side are 1 and in case (ii), the coefficients on each side are 0. In case (iii), the coefficients on both sides are $(N-m)_j/(N-i)_j$. Finally in case (iv), the coefficient on the right is given by

$$\begin{split} &\frac{N-m-j+1}{N-m+1} \cdot \frac{(N-m+1)_{\ell}(m-1-i)_{j-\ell}}{(N-i)_{j}} \cdot \binom{k-\ell}{k-j} \\ &+ \frac{k-j+1}{N-m+1} \cdot \frac{(N-m+1)_{\ell}(m-1-i)_{j-1-\ell}}{(N-i)_{j-1}} \cdot \binom{k-\ell}{k-j+1} \\ &= \frac{(N-m)_{\ell-1}(m-1-i)_{j-1-\ell}}{(N-i)_{j}} \binom{k-\ell}{k-j} \left[(N-m-j+1)(m-i-j+\ell) + (N-i-j+1)(j-\ell) \right] \\ &= \frac{(N-m)_{\ell-1}(m-1-i)_{j-1-\ell}}{(N-i)_{j}} \binom{k-\ell}{k-j} \cdot (N-m-\ell+1)(m-i) \\ &= \frac{(N-m)_{\ell}(m-i)_{j-\ell}}{(N-i)_{j}} \binom{k-\ell}{k-j}, \end{split}$$

which agrees with the coefficient on the left hand side.

Proof of Lemma 2.2. Part (i) is immediate by a double counting argument. This argument shows that the average (r-1)-degree is (N-r+1)/r times the average r-degree, while the maximum (r-1)-degree is at most this multiple of the maximum r-degree, and similarly for the minimum.

For (ii) we observe that $\mathcal{H}(x)$ is (k-1)-uniform with maximum *r*-degree at most Δ_{r+1} . We will prove now that $\mathcal{H}(x)$ is $(r-1, 3\eta)$ -near-regular. Since \mathcal{H} is (r, η) -near-regular, it is also $(1, \eta)$ -near-regular, by (i), and therefore $d_{\mathcal{H}}(x) \ge (1-\eta)hk/N$. And so it follows from a simple double counting argument that the average (r-1)-degree in $\mathcal{H}(x)$ satisfies

$$\bar{d}_{r-1}^{\mathcal{H}(x)} = \frac{(k-1)_{r-1}}{(N-1)_{r-1}} d_{\mathcal{H}}(x) = \bar{d}_{r}^{\mathcal{H}} \cdot \frac{N}{hk} d_{\mathcal{H}}(x) \ge (1-\eta) \bar{d}_{r}^{\mathcal{H}}.$$

Now let $A \subseteq V(\mathcal{H}(x))$ be an (r-1) element set. Since \mathcal{H} is (r, η) -near-regular we have

$$d_{\mathcal{H}(x)}(A) = d_{\mathcal{H}}(A \cup \{x\}) \leqslant (1+\eta)\bar{d}_r^{\mathcal{H}},$$

and so, since $\eta \in [0, 1/3]$ we have

$$d_{\mathcal{H}(x)}(A) \leqslant \frac{1+\eta}{1-\eta} \, \bar{d}_{r-1}^{\mathcal{H}(x)} \leqslant (1+3\eta) \, \bar{d}_{r-1}^{\mathcal{H}(x)} \, .$$

L

A near identical argument gives the lower bound $(1 - \eta)/(1 + \eta) \ge (1 - 3\eta)$ times $\bar{d}_{r-1}^{\mathcal{H}(x)}$, and so completes the proof.

We now prove Lemma 2.3, which bounds the possible value of $|X_{\ell}(B_i)|$ in a k-uniform (r, η) -near-regular hypergraph \mathcal{H} .

Proof of Lemma 2.3. Recall first that $X_{\ell}(B_i) := A_{\ell}(B_i) - \mathbb{E}[A_{\ell}(B_i)|B_{i-1}]$ and that $A_{\ell}(B_i) = N_{\ell}(B_i) - N_{\ell}(B_{i-1})$. If b_i is the last element added to B_i , i.e., $B_i = B_{i-1} \cup \{b_i\}$, then

$$A_{\ell}(B_i) = \sum_{\substack{C \subseteq B_{i-1} \\ |C| = \ell - 1}} d(C \cup \{b_i\}).$$

As $1 \leq \ell \leq r$ and \mathcal{H} is (r, η) -near-regular, we have that \mathcal{H} is also (ℓ, η) -near-regular by part (i) of Lemma 2.2. And so $d(C \cup \{b_i\}) = (1 \pm \eta)\overline{d}_\ell$ for all $(\ell - 1)$ element subsets $C \subseteq B_{i-1}$. It follows that

$$A_{\ell}(B_i) = (1 \pm \eta) \binom{i-1}{\ell-1} \bar{d}_{\ell}$$

Now, as $X_{\ell}(B_i) := A_{\ell}(B_i) - \mathbb{E}[A_{\ell}(B_i)|B_{i-1}]$, it follows that

$$|X_{\ell}(B_i)| \leqslant 2\eta \binom{i-1}{\ell-1} \bar{d}_{\ell}$$

Finally, since $\bar{d}_{\ell} = h\binom{k}{\ell} / \binom{N}{\ell}$, and using the bound $(i-1)_{\ell-1} \leq s^{\ell-1}(N-1)_{\ell-1}$, we obtain

$$|X_{\ell}(B_i)| \leqslant \frac{2\ell\binom{k}{\ell}\eta s^{\ell-1}h}{N},$$

as required.

3. Proof of Theorem 1.2

The proof of Theorem 1.2 is given by induction on r. In fact we prove two series of statements with a joint induction. Since $N_k^{\mathcal{H}}(B_m) = N^{\mathcal{H}}(B_m)$, it is clear that Theorem 1.2 is included in the sequence of statements $P_r, r \ge 1$.

 $\mathbf{P_r}$: For all $k \ge j \ge r$ and $\eta \in [0, 3^{-r+1}]$, for all k-uniform hypergraphs \mathcal{H} on [N] which are $(r-1, \eta)$ -near-regular with maximum r-degree Δ_r , we have

$$\mathbb{P}\left(D_j^{\mathcal{H}}(B_m) > a\right) \leqslant N^{O_k(1)} \exp\left(\frac{-\Omega_k(1)a^{2/r}}{m\Delta_r^{2/r}}\right)$$

for all $0 \le m \le N$ and all $a \ge C_r \eta^{r/(r-1)} h t^{(j-1)r/(r-1)}$, where $C_r = (10k!)^{10^r}$.

The other sequence of statements Q_r will be related to the behaviour of the random variables $X_{\ell}^{\mathcal{H}}(B_i)$ that occur in the martingale representation of $D_{\mathcal{H}}(B_m)$. We define Q_r to be the following statement.

 $\mathbf{Q_r}$: For all $k \ge \ell \ge r+1$ and $\eta \in [0, 3^{-r}]$, for all k-uniform hypergraphs \mathcal{H} on [N] which are (r, η) -near-regular with maximum (r+1)-degree Δ_{r+1} , we have

$$\mathbb{P}\left(|X_{\ell}^{\mathcal{H}}(B_{i})| > \alpha\right) \leqslant N^{O_{k}(1)} \exp\left(\frac{-\Omega_{k}(1)\alpha^{2/r}}{i\Delta_{r+1}^{2/r}}\right)$$

for all $0 \leq i \leq N$ and all $\alpha \geq D_r \eta s^{\ell-1} h/N$, where $D_r = (10k!)^{10^r+10}$.

The base case of the induction, P_1 , is proved in Section 3.1 by a straightforward application of the Azuma–Hoeffding inequality. We complete the proof by showing that P_r implies Q_r and Q_r implies P_{r+1} . These proofs are given in Sections 3.2 and 3.3 respectively.

3.1. The base case $-P_1$. Let us fix $k \ge j \ge 1$, $\eta \in [0, 1]$, and a k-uniform hypergraph \mathcal{H} on [N] with maximum degree Δ .

We use the Azuma–Hoeffding inequality applied to the martingale representation:

$$D_{j}(B_{m}) = \sum_{i=1}^{m} \sum_{\ell=1}^{j} \frac{(N-m)_{\ell}(m-i)_{j-\ell}}{(N-i)_{j}} \cdot {\binom{k-\ell}{k-j}} X_{\ell}(B_{i}).$$

In order to do so we must bound the magnitude of the increment

$$Y_i := \sum_{\ell=1}^{j} \frac{(N-m)_{\ell}(m-i)_{j-\ell}}{(N-i)_j} \cdot \binom{k-\ell}{k-j} X_{\ell}(B_i)$$

of the martingale. We observe that the first fraction is always at most 1, and so the coefficient itself is $O_k(1)$. Recalling that $X_{\ell}(B_i) := A_{\ell}(B_i) - \mathbb{E}[A_{\ell}(B_i)|B_{i-1}]$, and that both $A_{\ell}(B_i)$ and $\mathbb{E}[A_{\ell}(B_i)|B_{i-1}]$ are non-negative we have

$$||X_{\ell}(B_i)||_{\infty} \leq ||A_{\ell}(B_i)||_{\infty} \leq O_k(1) \cdot \Delta \qquad a.s.$$

where the second inequality follows since $A_{\ell}(B_i) = N_{\ell}(B_i) - N_{\ell}(B_{i-1})$ is certainly at most $\binom{k-1}{\ell-1} = O_k(1)$ times $d_{\mathcal{H}}(b_i) \leq \Delta$ (any "new" ℓ -sets must be in edges containing b_i).

Since Y_i consists of a finite number of terms and the coefficients are $O_k(1)$,

$$|Y_i| \leqslant O_k(1) \cdot \Delta \qquad a.s$$

By an application of the Azuma–Hoeffding inequality to $D_j(B_m) = \sum_{i=1}^m Y_i$, we have that

$$\mathbb{P}\left(D_j(B_m) > a\right) \leqslant \exp\left(\frac{-a^2}{2m O_k(1) \Delta^2}\right) = \exp\left(\frac{-\Omega_k(1)a^2}{m\Delta^2}\right).$$

This complete the proof of the base case P_1 .

3.2. **P** implies **Q**. In this section, we shall prove that $P_r \Rightarrow Q_r$. Let us fix $k \ge \ell \ge r+1$, $\eta \in [0, 3^{-r}]$ and a k-uniform hypergraph \mathcal{H} on [N] which is (r, η) -near-regular. Let Δ_{r+1} be the maximum (r+1)-degree of \mathcal{H} . Let us also fix $0 \le i \le N$ and $\alpha \ge D_r \eta s^{\ell-1} h/N$. In order to prove Q_r we must prove that

$$\mathbb{P}\left(|X_{\ell}(B_i)| > \alpha\right) \leqslant N^{O_k(1)} \exp\left(\frac{-\Omega_k(1)\alpha^{2/r}}{i\Delta_{r+1}^{2/r}}\right).$$
(6)

We recall that $X_{\ell}(B_i) = A_{\ell}(B_i) - \mathbb{E}[A_{\ell}(B_i)|B_{i-1}]$. Our proof of (6) is based on the following proposition on the deviation of $A_{\ell}(B_i)$ from its mean. We set

$$\lambda_{\ell}(i) := \mathbb{E}\left[A_{\ell}(B_i)\right] = \frac{\ell\binom{k}{\ell}h(i-1)_{\ell-1}}{(N)_{\ell}}$$

Note that $\lambda_{\ell}(B_i)$ is also equal to $L_{\ell}(i) - L_{\ell}(i-1)$.

Proposition 3.1.

$$\mathbb{P}\left(\left|A_{\ell}(B_{i}) - \lambda_{\ell}(i)\right| > \alpha\right) \leqslant N^{O_{k}(1)} \exp\left(\frac{-\Omega_{k}(1) \alpha^{2/r}}{i\Delta_{r+1}^{2/r}}\right)$$

for all $\alpha \ge D_r \eta s^{\ell-1} h/2N$.

Let us first observe that (6) follows from Proposition 3.1. Since $X_{\ell}(B_i) = A_{\ell}(B_i) - \mathbb{E}[A_{\ell}(B_i)|B_{i-1}]$, the event $|X_{\ell}(B_i)| > \alpha$ of (6) may only occur if either $|A_{\ell}(B_i) - \lambda_{\ell}(i)| > \alpha/2$ or $|\mathbb{E}[A_{\ell}(B_i)|B_{i-1}] - \lambda_{\ell}(i)| > \alpha/2$, by the triangle inequality. The first of these events has probability at most

$$N^{O_k(1)} \exp\left(\frac{-\Omega_k(1) \,\alpha^{2/r}}{i\Delta_{r+1}^{2/r}}\right)$$

by Proposition 3.1. The second may only occur if there exists $x \in [N] \setminus B_{i-1}$ such that $|A_{\ell}(B_{i-1} \cup \{x\}) - \lambda_{\ell}(i)| > \alpha/2$ and so the same bound holds by Lemma 2.4 and Proposition 3.1. This completes the proof of (6).

All that remains is to prove Proposition 3.1. We shall base the proof of Proposition 3.1 on Lemma 3.2, which shows how we may view A_{ℓ} in terms a deviation in a hypergraph $\mathcal{H}(x)$, and Lemma 3.3 which uses P_r to bound the probability of such deviations.

It will be useful to condition on the element $b_i = x$ that is added as the *i*th element of the process. Given that $b_i = x$ the set B_i is distributed as

$$B_i = B_{i-1}^{(x)} \cup \{x\}$$

where $B_{i-1}^{(x)}$ is a uniformly random subset of i-1 elements of $[N] \setminus \{x\}$. We also recall that we define the hypergraph

$$\mathcal{H}(x) := \left\{ f \setminus \{x\} : f \in E(\mathcal{H}), \, x \in f \right\}.$$

The first lemma shows that $A_{\ell}(B_{i-1}^{(x)} \cup \{x\})$ may be expressed precisely in terms of the deviation of $(\ell - 1)$ sets in the hypergraph $\mathcal{H}(x)$

Lemma 3.2. For each $x \in [N]$

$$A_{\ell}(B_{i-1}^{(x)} \cup \{x\}) = D_{\ell-1}^{\mathcal{H}(x)}(B_{i-1}^{(x)}) + \mathbb{E}\left[N_{\ell-1}^{\mathcal{H}(x)}(B_{i-1}^{(x)})\right]$$
$$= D_{\ell-1}^{\mathcal{H}(x)}(B_{i-1}^{(x)}) + \lambda_{\ell}(i) \pm \frac{\eta\ell\binom{k}{\ell}hs^{\ell-1}}{N}$$

We use P_r to provide a bound on the probability that $D_{\ell-1}^{\mathcal{H}(x)}(B_{i-1}^{(x)})$ is large. Lemma 3.3. Let $r+1 \leq \ell \leq k$. Then

$$\mathbb{P}\left(\left|D_{\ell-1}^{\mathcal{H}(x)}(B_{i-1}^{(x)})\right| > \alpha\right) \leqslant N^{O_k(1)} \exp\left(\frac{-\Omega_k(1) \alpha^{2/r}}{i\Delta_{r+1}^{2/r}}\right)$$

for all $\alpha \ge C_r(3\eta)^{r/(r-1)}e(\mathcal{H}(x))s^{(\ell-2)r/(r-1)}$. In particular, the result holds for all $\alpha \ge D_r\eta s^{\ell-1}h/4N$.

Let us see how Proposition 3.1 follows from these lemmas.

Proof of Proposition 3.1. By Lemma 3.2 we have that

$$|A_{\ell}(B_i) - \lambda_{\ell}(i)| \leq |D_{\ell-1}^{\mathcal{H}(x)}(B_{i-1}^{(x)})| + \frac{\eta \ell\binom{k}{\ell} h s^{\ell-1}}{N}.$$

Since the second term on the right is at most $\alpha/2$ (this follows from the condition on α and the fact that $D_r \ge 4\ell \binom{k}{l}$), the event that $|A_\ell(B_i) - \lambda_\ell(i)| > \alpha$ is contained in the event that $|D_{\ell-1}^{\mathcal{H}(x)}(B_{i-1}^{(x)})| > \alpha/2$. The required bound now follows immediately from Lemma 3.3.

All that remains is to prove Lemma 3.2 and Lemma 3.3. We begin with Lemma 3.2.

Proof of Lemma 3.2. Recall that

$$A_{\ell}(B_{i-1}^{(x)} \cup \{x\}) := N_{\ell}(B_{i-1}^{(x)} \cup \{x\}) - N_{\ell}(B_{i-1}^{(x)})$$

which is precisely the number of pairs (S, f) where $S \subseteq B_{i-1}^{(x)}$ is a subset of $\ell - 1$ elements, and f is an edge of \mathcal{H} such that $S \cup \{x\} \subseteq f$. Setting $f^- = f \setminus \{x\}$ we note that this condition is equivalent to the fact that $S \subseteq f^-$. It follows that $A_{\ell}(B_{i-1}^{(x)} \cup \{x\})$ is precisely $N_{\ell-1}^{\mathcal{H}(x)}(B_{i-1}^{(x)})$, and so

$$A_{\ell}(B_{i-1}^{(x)} \cup \{x\}) = D_{\ell-1}^{\mathcal{H}(x)}(B_{i-1}^{(x)}) + \mathbb{E}\left[N_{\ell-1}^{\mathcal{H}(x)}(B_{i-1}^{(x)})\right]$$

by the definition of $D_{\ell-1}^{\mathcal{H}(x)}(B_{i-1}^{(x)})$ as the deviation of $N_{\ell-1}^{\mathcal{H}(x)}(B_{i-1}^{(x)})$ from its mean. All that remains is to prove that

$$\mathbb{E}\left[N_{\ell-1}^{\mathcal{H}(x)}(B_{i-1}^{(x)})\right] = \lambda_{\ell}(i) \pm \frac{\eta \ell\binom{k}{\ell} s^{\ell-1} h}{N}$$

Since \mathcal{H} is (r, η) -near-regular (and so $(1, \eta)$ -near-regular by Lemma 2.2) we have $e(\mathcal{H}(x)) = d_{\mathcal{H}}(x) = (1 \pm \eta)hk/N$. We may also observe that $\mathcal{H}(x)$ is (k-1)-uniform on N-1 vertices and so

$$\mathbb{E}\left[N_{\ell-1}^{\mathcal{H}(x)}(B_{i-1}^{(x)})\right] = \frac{e(\mathcal{H}(x))\binom{k-1}{\ell-1}(i-1)_{\ell-1}}{(N-1)_{\ell-1}} \\ = \frac{\ell\binom{k}{\ell}h(i-1)_{\ell-1}}{(N)_{\ell}} \pm \frac{\eta\ell\binom{k}{\ell}h(i-1)_{\ell-1}}{(N)_{\ell}} \\ = \lambda_{\ell}(i) \pm \frac{\eta\ell\binom{k}{\ell}s^{\ell-1}h}{N},$$

as required.

Proof of Lemma 3.3. We prove the required bound by applying the inequality given by P_r to the hypergraph $\mathcal{H}(x)$. We observe that $\mathcal{H}(x)$ is a (k-1)-uniform hypergraph on N-1 vertices. We may also observe that $\mathcal{H}(x)$ inherits the regularity condition $(r-1, 3\eta)$ -near-regular from \mathcal{H} by Lemma 2.2 and the maximum r-degree of $\mathcal{H}(x)$ is at most Δ_{r+1} . By P_r we have that

$$\mathbb{P}\left(\left|D_{\ell-1}^{\mathcal{H}(x)}(B_{i-1}^{-})\right| > \alpha\right) \leqslant N^{O_k(1)} \exp\left(\frac{-\Omega_k(1) \alpha^{2/r}}{i\Delta_{r+1}^{2/r}}\right)$$

for all $\alpha \ge C_r(3\eta)^{r/(r-1)}e(\mathcal{H}(x))s^{(\ell-2)r/(r-1)}$. This is exactly the result we need. All that remains is to verify that this includes all $\alpha \ge D_r\eta s^{\ell-1}h/4N$.

We have that $D_r \ge 10k \cdot 3^{r/(r-1)}C_r$, $\eta < 1$, $(\ell - 2)r \le (\ell - 1)(r - 1)$ and $e(\mathcal{H}(x)) \le (1 + \eta)hk/N \le 2hk/N$. It follows that

$$\frac{D_r \eta h s^{\ell-1}}{4N} \ge \frac{10C_r (3\eta)^{r/(r-1)} h k s^{(\ell-2)r/(r-1)}}{4N} \ge C_r (3\eta)^{r/(r-1)} e(\mathcal{H}(x)) s^{(\ell-2)r/(r-1)}$$

This confirms that the inequality holds across the whole of the range we claimed.

3.3. **Q implies P.** In this section, we will prove that $Q_r \Rightarrow P_{r+1}$. The main idea of the induction step is using our information about the magnitude of increments combined with Lemma 2.6. Let us fix $k \ge j \ge r+1$, $\eta \in [0, 3^{-r}]$ and a k-uniform (r, η) -near-regular hypergraph \mathcal{H} on [N] with maximum (r+1)-degree Δ_{r+1} . We recall the martingale representation

$$D_j(B_m) = \sum_{i=1}^m Y_i,$$

where

$$Y_{i} = \sum_{\ell=1}^{j} \frac{(N-m)_{\ell}(m-i)_{j-\ell}}{(N-i)_{j}} \binom{k-\ell}{k-j} X_{\ell}(B_{i})$$

We prove now an auxiliary lemma that controls the probability that the increments are large. Lemma 3.4. If $\alpha \ge jk!D_r\eta t^{j-1}h/N$, then

$$\mathbb{P}(|Y_i| > \alpha) \leqslant N^{O_k(1)} \exp\left(\frac{-\Omega_k(1)\alpha^{2/r}}{m\Delta_{r+1}^{2/r}}\right)$$

Proof. We first observe that

$$|Y_i| \leq \sum_{\ell=1}^{j} k! t^{j-\ell} |X_\ell(B_i)|,$$
(7)

since

$$\frac{(N-m)_{\ell}(m-i)_{j-\ell}}{(N-i)_j} \binom{k-\ell}{k-j} \leqslant t^{j-l}k!.$$

By an application of Lemma 2.3 and Q_r , since $\alpha \ge jk!D_r\eta t^{j-1}h/N$ the following bound holds for all $1 \le \ell \le k$:

$$\mathbb{P}\left(|X_{\ell}(B_i)| > \frac{\alpha t^{\ell-j}}{jk!}\right) \leqslant N^{O_k(1)} \exp\left(\frac{-\Omega_k(1)\alpha^{2/r}}{m\Delta_{r+1}^{2/r}}\right).$$
(8)

Finally, using (7), (8) and the union bound, we have

$$\mathbb{P}\left(|Y_i| > \alpha\right) \leqslant N^{O_k(1)} \exp\left(\frac{-\Omega_k(1)\alpha^{2/r}}{m\Delta_{r+1}^{2/r}}\right).$$

We are now ready to prove P_{r+1} . Let $a \ge C_{r+1}\eta^{(r+1)/r}ht^{(j-1)(r+1)/r}$. Choosing $\alpha = a^{r/(r+1)}\Delta_{r+1}^{1/(r+1)}$, we can easily verify that

$$\alpha \geqslant jk! D_r \eta t^{j-1} h/N,$$

using $\Delta_{r+1} \ge h/(N^{r+1})$ and $C_{r+1} \ge k^2(k!)^2 D_r^2$. By an application of the Azuma–Hoeffding inequality (the version given in Lemma 2.6) with $c_i = \alpha$ for every *i*, we have

$$\mathbb{P}(D_j(B_m) > a) \leq \exp\left(\frac{-a^2}{2m\alpha^2}\right) + N\sum_{i=1}^m \mathbb{P}(|Y_i| > \alpha).$$

By Lemma 3.4, we obtain

$$\mathbb{P}\left(D_j(B_m) > a\right) \leqslant \exp\left(\frac{-a^2}{2m\alpha^2}\right) + N^{O_k(1)} \exp\left(\frac{-\Omega_k(1)\alpha^{2/r}}{m\Delta_{r+1}^{2/r}}\right).$$

Since $\alpha = a^{r/(r+1)} \Delta_{r+1}^{1/(r+1)}$, this last inequality gives us

$$\mathbb{P}(D_j(B_m) > a) \leqslant N^{O_k(1)} \exp\left(\frac{-\Omega_k(1) a^{2/(r+1)}}{m \Delta_{r+1}^{2/(r+1)}}\right),$$

which establishes the induction step.

Now that we have established the base case P_1 and the implications $P_r \Rightarrow Q_r$ and $Q_r \Rightarrow P_{r+1}$ for all $r \ge 1$ we have completed the proof of Theorem 1.2.

3.4. A weaker condition if \mathcal{H} is regular and $(r-1, \eta)$ -near-regular. We remarked (Remark 4) in the Introduction that the condition on *a* may be weakened to

$$a \geq (10k!)^{10^r} e(\mathcal{H}) \left(\frac{\eta m^{k-2}}{N^{k-2}}\right)^{r/(r-2)}$$

if \mathcal{H} is regular and $(r-1, \eta)$ -near-regular, for $r \ge 3$.

The proof is essentially identical to that given above. We highlight only the differences.

Let \mathcal{H} be regular and (r, η) -near-regular. In this case the $\pm \frac{\eta \ell \binom{k}{\ell} s^{\ell-1} h}{N}$ term in Lemma 3.2 is not necessary. It follows that in this case Q_r holds for all $\alpha \ge C_r(3\eta)^{r/(r-1)} h s^{(\ell-2)r/(r-1)}$ (by a simple adaptation of the proof of Proposition 3.1).

It is then possible to prove the result of Lemma 3.4 for all $\alpha \ge 10j(k+1)!C_r\eta^{r/(r-1)}ht^{(j-2)r/(r-1)}/N$. Following the rest of the proof of Section 3.3 we obtain the required result for all $a \ge C_{r+1}\eta^{(r+1)/(r-1)}ht^{(j-2)(r+1)/(r-1)}$. Taking j = k and swapping r for r-1 we obtain the claimed result, which we now state as a proposition.

Proposition 3.5. Let $1 \leq r \leq k$ and let $\eta \in [0, 3^{-r+1}]$. Let \mathcal{H} be a k-uniform hypergraph on [N]. Suppose that \mathcal{H} is regular and $(r-1, \eta)$ -near-regular with maximum r-degree Δ_r . Then

$$\mathbb{P}\left(|D^{\mathcal{H}}(B_m)| > a\right) \leqslant N^{O_k(1)} \exp\left(\frac{-\Omega_k(1)a^{2/r}}{m\Delta_r^{2/r}}\right)$$

for all

$$a \geq (10k!)^{10^r} e(\mathcal{H}) \left(\frac{\eta m^{k-2}}{N^{k-2}}\right)^{r/(r-2)}$$

4. Deviations $D^{\mathcal{H}}(B_p)$ – Proof of Theorem 1.3

Deviations $D(B_p)$ in the *p*-model are intimately related to deviations $D(B_m)$ in the *m*-model, via the identity

$$\mathbb{P}\left(D^{\mathcal{H}}(B_p) > a\right) = \sum_{m=0}^{N} b_{N,p}(m) \mathbb{P}\left(N^{\mathcal{H}}(B_m) > p^k h + a\right), \qquad (9)$$

where $b_{N,p}(m) := \mathbb{P}(\operatorname{Bin}(N,p)=m)$. Recall that $L^{\mathcal{H}}(p) = p^k h$ is the expected value of $N^{\mathcal{H}}(B_p)$ and we study the probability that the deviation satisfies $D^{\mathcal{H}}(B_p) > \delta_N L^{\mathcal{H}}(p)$.

Our proof of Theorem 1.3 consists of a lower bound (see Section 4.1) and an upper bound (see Section 4.2). The lower bound is based on on the fact that $\mathbb{P}(N(B_m) > p^k h + a)$ is increasing in m and so

$$\mathbb{P}\left(D(B_p) > \delta_N L(p)\right) \ge B_{N,p}(m_+) \mathbb{P}\left(N(B_{m_+}) > \left(1 + \delta_N\right) L(p)\right)$$
(10)

for all $m_+ \ge 0$, where $B_{N,p}(m) := \mathbb{P}(\operatorname{Bin}(N,p) \ge m)$. For a particular choice of $m_+ = m_+(\delta_N)$ we shall prove that

$$B_{N,p}(m_{+}) \mathbb{P}\left(N(B_{m_{+}}) > (1 + \delta_{N})L(p)\right) = \exp\left(-(1 + o(1))\frac{\delta_{N}^{2}pN}{2k^{2}(1 - p)}\right)$$

which gives the required lower bound.

We need to work harder to prove the upper bound. We must control all contributions to the sum (9). We again use that $\mathbb{P}(N(B_m) > p^k h + a)$ is increasing to observe that

$$\mathbb{P}\left(D(B_p) > \delta_N L(p)\right) \leqslant \mathbb{P}\left(N(B_{m_-}) > \left(1 + \delta_N\right)L(p)\right) + B_{N,p}(m_-)$$

for all $m_{-} \ge 0$. We shall then choose $m_{-} = m_{-}(\delta_{N})$ such that

$$B_{N,p}(m_{-}) = \exp\left(-(1+o(1))\frac{\delta_{N}^{2}pN}{2k^{2}(1-p)}\right)$$

and

$$\mathbb{P}\left(N(B_{m_{-}}) > (1 + \delta_N)L(p)\right) \ll \exp\left(\frac{-\delta_N^2 p N}{2k^2(1-p)}\right) \,.$$

The latter inequality is proved using Theorem 1.2.

Based on the above sketch it is clear that the probabilities $b_{N,p}(m)$ and $B_{N,p}(m)$, related to the binomial distribution, are central to our proof. While more precise estimates, up to a multiplicative factor of 1 + o(1) are known³, the following is sufficient for our purposes. Throughout the section we set q := 1 - p.

Theorem 4.1. Suppose that (x_N) is a sequence such that $1 \ll x_N \ll \sqrt{Npq}$. Then

$$b_{N,p}(\lfloor pN + x_N\sqrt{Npq} \rfloor) = \frac{1}{\sqrt{Npq}} \exp\left(-(1+o(1))\frac{x_N^2}{2}\right)$$

and

$$B_{N,p}(\lfloor pN + x_N\sqrt{Npq} \rfloor) = \exp\left(-(1+o(1))\frac{x_N^2}{2}\right).$$

To see that these bounds do indeed follow from Theorem 1.13 of [14] (for example) simply note that for any sequence $1 \ll x_N \ll \sqrt{Npq}$ we have that:

³see for example the bound in [14], adapted from Bahadur [2]

- (i) the $E(x_N, N)$ expression is at most $o(x_N^2)$ (by comparison with a geometric series), and
- (ii) the multiplicative term satisfies $1/\sqrt{2\pi x_N} = \exp(-\log(2\pi x_N)/2) = \exp(o(x_N^2))$.

Both of the values m_{-} and m_{+} discussed above will be chosen in relation to

$$m_* := (1+\delta_N)^{1/k} pN$$

which is chosen so that $L(m^*) = L(p)$. Let us also define

$$x(m) := \frac{m - pN}{\sqrt{Npq}}$$

in general and, in particular, set $x_* := x(m_*)$.

4.1. Lower Bound. First we choose a sequence f_N such that

$$\max\left\{\eta^{r/(r-1)}p^{(k-1)/(r-1)}N, \frac{p^{r/2-k+1}N^{r/2+1}\Delta_r(\log N)^{r/2}}{h}\right\} \ll f_N \ll \delta_N pN.$$

Note that since

$$\delta_N \gg \max\left\{\frac{\Delta_r (N\log N)^{r/2}}{p^{k-r/2}h}, \left(\eta^r p^{k-r}\right)^{1/(r-1)}\right\}$$

we can take such sequence. We then choose $m_+ = m_* + f_N$ and we also set $x_+ = x(m_+)$.

Now we prove the following two lemmas, which together with (10), will give us the desired lower bound.

Lemma 4.2.

$$\mathbb{P}\left(N(B_{m_+}) > (1+\delta_N)L(p)\right) = 1 - o(1).$$

Proof. Note first that it suffices to prove that

$$\mathbb{P}\left(D(B_{m_+}) \leqslant (1+\delta_N)L(p) - L(m_+)\right) = o(1).$$

Observe now that

$$(1+\delta_N)L(p) - L(m_+) = hp^k \left[(1+\delta_N)p^k - \frac{(m_+)_k}{(N)_k} \right]$$
$$\leqslant h \left[(1+\delta_N)p^k - \frac{(m_*+f_N-k)^k}{N^k} \right]$$
$$\leqslant h \left[(1+\delta_N)p^k - \frac{m_*^k}{N^k} - \frac{m_*^{k-1}f_N}{N^k} \right]$$
$$\leqslant -\frac{hp^{k-1}f_N}{N}.$$

So we have

$$\mathbb{P}\left(D(B_{m_+}) \leqslant (1+\delta_N)L(p) - L(m_+)\right) \leqslant \mathbb{P}\left(D(B_{m_+}) \leqslant -\frac{hp^{k-1}f_N}{N}\right)$$

Let us denote $a = (hp^{k-1}f_N)/N$. Since $f_N \gg \eta^{r/(r-1)}p^{(k-1)/(r-1)}N$, if N is sufficiently large, we can apply Theorem 1.2 and obtain

$$\mathbb{P}\left(D(B_{m_+}) \leqslant -a\right) \leqslant N^{O_k(1)} \exp\left(\frac{-\Omega_k(1)a^{2/r}}{m\Delta_r^{2/r}}\right).$$

Using now that

$$f_N \gg \frac{p^{r/2-k+1}N^{r/2+1}\Delta_r(\log N)^{r/2}}{h}$$

we can easily verify that

$$N^{O_k(1)} \exp\left(\frac{-\Omega_k(1)a^{2/r}}{m\Delta_r^{2/r}}\right) = o(1),$$

which gives us the desired result.

Lemma 4.3.

$$B_{N,p}(m_{+}) = \exp\left(-(1+o(1))\frac{\delta_{N}^{2}pN}{2k^{2}(1-p)}\right).$$

Proof. By Theorem 4.1, we have

$$B_{N,p}(m_+) = \exp\left(-(1+o(1))\frac{x_+^2}{2}\right).$$

Observe now that

$$\frac{x_+}{x_*} = 1 + \frac{f_N}{pN[(1+\delta_N)^{1/k} - 1]}.$$

Since $\delta_N \ll 1$, $(1 + \delta_N)^{1/k} - 1 = \Theta(\delta_N)$ and as $f_N \ll \delta_N pN$, we obtain $x_+ = (1 + o(1))x_*$, which gives us

$$B_{N,p}(m_+) = \exp\left(-(1+o(1))\frac{x_*^2}{2}\right)$$

Finally, note that

$$x_* = \sqrt{\frac{pN}{1-p}} \left((1+\delta_N)^{1/k} - 1 \right)$$

and since $(1 + \delta_N)^{1/k} - 1 = (1 + o(1))\delta_N/k$, we obtain the required result.

4.2. Upper Bound. Our upper bound on $\mathbb{P}\left(D^{\mathcal{H}}(B_p) > \delta_N L^{\mathcal{H}}(p)\right)$ requires us to control all the terms of the sum (9), i.e., all the terms in the sum

$$\mathbb{P}\left(D^{\mathcal{H}}(B_p) > \delta_N L^{\mathcal{H}}(p)\right) = \sum_{m=0}^N b_{N,p}(m) \mathbb{P}\left(N^{\mathcal{H}}(B_m) > p^k h + \delta_N L^{\mathcal{H}}(p)\right).$$

In fact we do not require a very precise analysis. We shall simply break the sum into two parts $m < m_{-}$ and $m \ge m_{-}$ for a value of m_{-} we define below. We bound the terms $m \ge m_{-}$

using only the first probability (the binomial) and the terms $m < m_{-}$ using only the second probability (the deviation in the model B_m).

Let g_N be a sequence satisfying

$$\frac{\eta^{r/(r-1)}p^{(k-1)/(r-1)}N}{h}, \frac{N^{r/2+1}(\log N)^{r/2}\Delta_r}{p^{k-r/2-1}h}, \frac{\delta_N^r N^{r+1}\Delta_r}{p^{k-r-1}h} \ll g_N \ll \delta_N pN$$

It is certainly possible to choose such a sequence by the conditions on δ_N in Theorem 1.3.

We define $m_{-} := m_{*} - g_{N}$ and set $x_{-} = x(m_{-})$. By a calculation similar to that given in (11) we have

$$L(m_{-}) \leq (1+\delta_N)L(p) - \frac{hp^{k-1}g_N}{N}.$$
 (11)

We now bound the two parts of the sum. First, for the part $m \ge m_{-}$ we simply use that the sum of these terms is at most

$$B_{N,p}(m_{-}) = \exp\left(-(1+o(1))\frac{\delta_{N}^{2}pN}{2k^{2}(1-p)}\right)$$

This may be verified by simply following the proof of Lemma 4.3 and using that $g_N \ll \delta_N pN$.

Now, we bound the rest of the sum by $\exp(-x_*^2)$ using the following lemma. This will complete the proof of the upper bound.

Lemma 4.4.

$$\sum_{m=0}^{m_{-}} \mathbb{P}\left(N^{\mathcal{H}}(B_m) > p^k h + \delta_N L_{(p)}^{\mathcal{H}}\right) \leqslant \exp(-x_*^2).$$

Proof. Since $\mathbb{P}\left(N^{\mathcal{H}}(B_m) > p^k h + \delta_N L^{\mathcal{H}}(p)\right)$ is increasing in *m* it suffices to prove that

$$\mathbb{P}\left(N^{\mathcal{H}}(B_{m_{-}}) > p^{k}h + \delta_{N}L^{\mathcal{H}}(p)\right) \leqslant \frac{\exp(-x_{*}^{2})}{N}.$$
(12)

By (11) this event is contained in the event that $D^{\mathcal{H}}(B_{m_{-}}) > hp^{k-1}g_N/N$. Three lower bounds on g_N were given above. The first ensures that we may apply Theorem 1.2 to bound the probability of the deviation $D^{\mathcal{H}}(B_{m_{-}}) > hp^{k-1}g_N/N$. The second and third give that the resulting bound is at most $N^{O_k(1)} \exp(-\omega(\log N))$ and $N^{O_k(1)} \exp(\omega(x_*^2))$ respectively. In particular, for any constant C we have that

$$\mathbb{P}\left(D^{\mathcal{H}}(B_{m_{-}}) > hp^{k-1}g_N/N\right) \leqslant N^{O_k(1)}\exp(-C\log N - x_*^2)$$

for all sufficiently large N. Choosing C to be one larger than the constant of the $O_k(1)$ we obtain (12), and so complete the proof of the lemma.

5. Arithmetic configurations in random sets

In this section we collect some applications of Theorems 1.2 and 1.3 to illustrate its use in obtaining bounds of deviations for the count of arithmetic structures in random sets. To simplify matters we will consider the ambient group to be the cyclic group $\mathbb{Z}/N\mathbb{Z}$ with N prime.

5.1. k-progressions. The initial motivation of this article was to analyze moderate deviations in counting the number of k-term arithmetic progressions (k-progressions for short) in random sets. As it has been mentioned in the Introduction, large deviations for this problem have been intensively studied and they are currently well understood, see [21, 31, 5, 17, 8]. It was also proved by Berkowitz, Sah and Sawhney [6] that the number of k-progressions in a dense binomial random subset of \mathbb{Z}_n does not obey a local central limit theorem, despite obeying a central limit theorem. We consider here moderate deviations for the m-model first. Let $N^k(B_m)$ denote the number of k-progressions in a random subset B_m of the cyclic group $\mathbb{Z}/N\mathbb{Z}$, N prime, and by $D^k(B_m)$ the deviation of $N^k(B_m)$.

Let \mathcal{H}_k be the k-uniform hypergraph with vertex set $\mathbb{Z}/N\mathbb{Z}$ and edges the nontrivial kprogressions $\{x, x+d, \ldots, x+(k-1)d\}$ with $x \in \mathbb{Z}/N\mathbb{Z}$ and $d \in \{1, 2, \ldots, (N-1)/2\}$. Every pair $\{x, y\}$ of elements in $\mathbb{Z}/N\mathbb{Z}$ is contained in $\binom{k}{2}$ edges of \mathcal{H}_k , so that the hypergraph is 2-regular and hence 1-regular. A direct application of Theorem 1.1 for r = 2 gives the following result.

Theorem 5.1. For each a > 0,

$$\mathbb{P}\left(|D^{k}(B_{m})| > a\right) \leqslant N^{c_{1}} \exp\left(-c_{2}\frac{a}{m}\right)$$

for some constants c_1, c_2 which depend only on k.

The expected value of $N^k(B_m)$ in this example is (see (1))

$$L^{k}(m) = {\binom{N}{2}} \frac{(m)_{k}}{(N)_{k}} \sim \frac{m^{k}}{2N^{k-2}}.$$

Bounds on moderate deviations are thus obtained in Theorem 5.1 for $a \ll m^k/N^{k-2}$.

In particular, for $m \log N \ll a \ll m^k/N^{k-2}$ we obtain exponentially small bounds for moderate deviations of the k-progressions count, which apply to random sets of $\mathbb{Z}/N\mathbb{Z}$ of size $m \gg (N^{k-2} \log N)^{1/(k-1)}$. This is slightly above the threshold cardinality for the existence of k-progressions in a random set (see Rödl, Ruciński [25]). One can not expect to obtain such small deviation for smaller sets since, as shown in Rué, Spiegel and Zumalacárregui [27], the count of k-progressions within the threshold window converges to a Poisson distribution.

We note that the term in the exponential in Theorem 1.1 for r = 3 corresponding to this example would be $-\Omega_k(1)a^{2/3}/m$ (as $\Delta_3 = O_k(1)$), while for r = 1 we have $\Delta_1 \sim N/2$ and we would obtain $-\Omega_k(1)a^2/mN$. Both are worse than the one obtained for r = 2 as displayed in the bound of Theorem 5.2. This exemplifies Remark 3 in the Introduction.

By following the proof of Theorem 1.2 in the case k = 3 one can make the $o_k(1)$ and $\Omega_k(1)$ terms in the bound of the Theorem explicit (even if not optimized). We include this proof here as it is simpler than the general one described in Section 3 and may illustrate its main lines.

Theorem 5.2. Let $D^3(B_m)$ denote the deviation in counting 3-progressions in $\mathbb{Z}/N\mathbb{Z}$, with N prime. For every a > 0, we have

$$\mathbb{P}\left(|D^3(B_m)| > a\right) \leqslant (Nm+1) \exp\left(-\frac{a}{9m}\right).$$

Proof. In this case the hypergraph \mathcal{H}_3 defined in the proof of Theorem 5.1 is 2-regular. Hence, we have $D_1(B_m) = D_2(B_m) = 0$ (by Lemma 2.3, with $\eta = 0$) and the martingale representation (3) reduces to

$$D_3^3(B_m) = \sum_{j=1}^m \frac{(N-m)_3}{(N-j)_3} X_3(B_j).$$

We bound the probability that $|X_3(B_i)|$ is large as in the proof of Lemma 3.3 by considering the hypergraph $\mathcal{H}(x)$ which is a 2-uniform 1-regular hypergraph with $\Delta_1(\mathcal{H}_3(x)) =$ $\Delta_2(\mathcal{H}_3) = 3$. It follows from the Azuma–Hoeffding inequality (the argument in Section 3.1) that, for each $\alpha > 0$,

$$\mathbb{P}(X_3(B_j)) > \alpha) < \exp\left(-\frac{\alpha^2}{9j}\right).$$

Therefore, as in the argument in Section 3.3, Lemma 2.6 gives, for each a > 0 and $\alpha > 0$,

$$\mathbb{P}\left(D_3^3(B_m) > a\right) \leqslant \exp\left(-\frac{a^2}{2m\alpha^2}\right) + N\sum_{j=1}^m \exp\left(-\frac{\alpha^2}{9j}\right) \leqslant \exp\left(-\frac{a^2}{2m\alpha^2}\right) + Nm\exp\left(-\frac{\alpha^2}{9m}\right)$$

The result follows by choosing $\alpha = a^{1/2}$.

The result follows by choosing $\alpha = a^{1/2}$.

The bounds on deviations for the counting in the m-model can be transferred to the pbinomial model as described in Section 4. As mentioned in the Introduction, Theorem 1.4 follows from Theorem 1.3 taking into account that the hypergraph \mathcal{H}_3 of 3-progressions in $\mathbb{Z}/N\mathbb{Z}$ is 1-regular. We can analogously derive the following result for k-progressions from Theorem 1.3 using again the 1-regularity of the hypergraph \mathcal{H}_k of k-progressions.

Theorem 5.3. Let $D^k(B_n)$ denote the deviation of the k-progressions count in a p-random subset B_p of $\mathbb{Z}/N\mathbb{Z}$, N prime.

Let δ_N be a sequence satisfying

$$\max\left\{\frac{\log N}{p^{k-1}N^2}, \frac{1}{\sqrt{pN}}\right\} \ll \delta_N \ll p^{k-2}.$$

Then,

$$\mathbb{P}\left(D^k(B_p) > \delta_N L^k(p)\right) = \exp\left(-(1+o(1))\frac{\delta_N^2 p N}{2k^2(1-p)}\right).$$

24 GONZALO FIZ PONTIVEROS, SIMON GRIFFITHS, MATHEUS SECCO, AND ORIOL SERRA

Furthermore, the same bounds apply to the corresponding negative deviations.

As the expected number of k-progressions in B_p is $L^k(p) = p^k {N \choose 2}$, Theorem 5.3 provides exponentially small bounds of moderate deviations for $p \gg N^{-1/(2k-3)}$.

5.2. Schur equation. In addition to the 3-progressions, let us mention another 3-variable case, the Schur equation x + y = z, related to problems on sum-free sets. Threshold cardinalities and large deviations for the Schur equation have also been addressed in the literature, see e.g. [16, 31]. Consider the hypergraph \mathcal{H} with vertex set the nonzero elements of $\mathbb{Z}/N\mathbb{Z}$ and edge set the nontrivial Schur triples $\{x, y, z\}$ with x + y = z, none of them zero. Every nonzero element x in $\mathbb{Z}/N\mathbb{Z}$ belongs to N - 3 triples of the form $\{x, y, x + y\}$ with $y \notin \{0, x, -x\}$ and (N - 3)/2 triples of the form $\{y, x - y, x\}$ with $y \notin \{0, x, x/2\}$. Therefore, \mathcal{H} is 1-regular with $d_1(\mathcal{H}) = 3(N-3)/2$ and maximum 2-degree $\Delta_2(\mathcal{H}) = 3$. Let $D^3(B_m)$ denote the deviation on the count of Schur triples of nonzero elements in a random set B_m of $(\mathbb{Z}/N\mathbb{Z})^*$. A proof analogous to the one of Theorem 5.2 gives, for all a > 0,

$$\mathbb{P}\left(D^3(B_m) > a\right) < N^2 \exp\left(-c\left(\frac{a}{m}\right)\right)$$

for some constant c. The above bound is analogous to the one in Theorem 5.2 and applies to the same ranges of a and m. The bounds on the deviations can be analogously transferred to the p-binomial model to obtain a bound as in Theorem 1.4 for the deviation of Schur triples in this model.

This example exemplifies some minor adjustments which can be made to apply Theorem 1.1 for r = 2 to the count of solutions of general 3-term linear equations of the form ax + by + cz = d in $\mathbb{Z}/N\mathbb{Z}$. The corresponding hypergraphs may fail to be regular or near-regular by the degree of some pairs of elements, depending on the values of a, b, c, but can be regularized without affecting the counts in a substantial way. We discuss the general approach for linear systems in Section 5.4 below.

5.3. Sidon equation. The Sidon equation x + y = z + t has also been treated in the literature, see e.g. [24, 27]. In particular, bounds of the deviation on the number of solutions of the Sidon equation are given in [24, Lemma 5.3] by means of the Kim–Vu polynomial concentration inequality [23]. An analogous analysis can be carried over in our context to obtain bounds on moderate deviations giving more precise results in an appropriate range of the size of random sets. The number of nontrivial quadruples (x, y, z, t) in a set $A \subset \mathbb{Z}/N\mathbb{Z}$ which satisfy the Sidon equation x + y = z + t with $\{x, y\} \neq \{z, t\}$ is also called the Additive Energy of the set (which counts ordered quadruples). For unordered subsets, solutions of the Sidon equation in A are either 3-progressions $\{x, y, z = (x+y)/2\}$, which satisfy x + y = 2z, or quadruples $\{x, y, z, t\}$ of distinct elements.

We consider the 4-uniform hypergraph \mathcal{H}_4 which has vertex set $\mathbb{Z}/N\mathbb{Z}$ (as usual we consider N prime) and edges the Sidon quadruples $\{x, y, z, t\}$ with x+y=z+t. Every pair $\{x, y\}$ is contained in (N-3) quadruples of the form $\{x, y, z, x+y-z\}$ with $z \notin \{x, y, (x+y)/2\}$ and

(N-3) quadruples of the form $\{x, y, z, x+z-y\}$ with $z \notin \{x, y, (x-y)/2\}$, so that \mathcal{H}_4 is 2regular with $d_2(\mathcal{H}_4) = 2(N-3)$. Moreover, each triple $\{x, y, z\}$ which is not in arithmetic progression belongs to three distinct quadruples $\{x, y, z, t\}$ with $t \in \{x+y-z, x+z-y, y+z-x\}$, while triples which are 3-progressions, say z = (x+y)/2, belong to two distinct quadruples. Hence, and $\Delta_3(\mathcal{H}_4) = 3$. By denoting by $D^4(B_m)$ the deviation in the count of Sidon quadruples of paiwise distinct elements, Theorem 1.2 with r = 3 gives, for all a > 0,

$$\mathbb{P}\left(D^4(B_m) > a\right) \leqslant N^{c_1} \exp\left(-c_2 \frac{a^{2/3}}{m}\right),$$

for some constants c_1, c_2 . By bounding the deviation $D^S(B_m)$ on the count of the number of solutions of the Sidon equation by the sum of deviations on the count of quadruples of distinct elements and the count of 3-progressions from Theorem 5.2 we obtain

$$\mathbb{P}(D^S(B_m) > a) \leqslant N^{c_1} \exp\left(-c_S \frac{a^{2/3}}{m}\right),$$

for some constant c_S . The expected number of solutions of the Sidon equation in a random set B_m is

$$L^{S}(m) = \frac{(m)_{3}}{2(N-3)} + \frac{(m)_{4}}{2(N-4)}$$

Therefore, for $a \ll m^4/N$ and $a^{2/3}/m \gg \log N$ we obtain exponentially small bounds for the deviation in the count of solutions of the Sidon equation, which apply to random sets of cardinality $m \gg N^{2/5} (\log N)^{3/5}$. This shows that, for these values of m, the Additive Energy of a random m-set in $\mathbb{Z}/N\mathbb{Z}$ is highly concentrated on its mean value. We observe that in this case the threshold for the appearance of solutions of the Sidon equation is $N^{1/4}$ and our bound starts to be effective only above $N^{2/5}$.

As in the preceding examples, Theorem 1.3 can be applied to transfer the bounds to the p-binomial model. Let $D^4(B_p)$ denote the deviation on the number of solutions to the Sidon equation with pairwise distinct entries in a random set B_p of $\mathbb{Z}/N\mathbb{Z}$ and $L^4(p)$ its mean value. Taking into account that the hypergraph \mathcal{H}_4 is 2-regular and has constant maximum degree $\Delta_3(\mathcal{H}_4) = 3$, Theorem 1.3 with r = 3 gives

$$\mathbb{P}(D^4(B_p) > \delta_N L^4(p)) = \exp\left(-(1+o(1))\frac{\delta_N^2 p N}{32(1-p)}\right),$$

for every sequence δ_N satisfying

$$\max\left\{\frac{1}{p^{5/2}}\left(\frac{\log N}{N}\right)^{3/2}, \frac{1}{\sqrt{pN}}\right\} \ll \delta_N \ll p^{1/2}.$$

The above condition on δ_N gives a meaningful range of applications when $p \gg (\log N/N)^{1/2}$.

5.4. Linear systems. The arithmetic configurations considered in the above examples correspond to solutions of linear systems

$$Ax = 0$$

in some finite field \mathbb{F}_N , N a prime, where A is a $(l \times k)$ matrix with integer entries, $l \leq k-2$, $x \in \mathbb{F}_N^k$. This general setting has been widely addressed in the literature, see e.g. [25, 21, 27]. Conditions are imposed on the matrix A to avoid some degenerate cases, the most natural one is that all $l \times l$ submatrices of A are nonsingular. Then substitution of k-l entries in x gives a unique solution to the equation. It follows that the hypergraph \mathcal{H}_A with vertex set \mathbb{F}_N and edges the k-subsets which are entries of a solution to the linear system has maximum (k-l) degree $\Delta_{k-l}(\mathcal{H}_A) \leq {k \choose k-l}$. One can apply Theorem 1.2 in this context whenever the hypergraph is $(k-l-1,\eta)$ -near regular for an appropriate value of η .

Substitution of k-l-1 entries in x give rise to N solutions to the system. If the equation $x_i = x_j$ is linearly independent with the rows of A for every pair $i, j \in [k]$, then $O_k(1)$ solutions may have repeated entries. Matrices A which satisfy this last condition are called *irredundant* (see [25]), a property that we will also assume. Moreover, $O_k(1)$ substitutions may give rise to the same solution. Therefore, under the above conditions on A, every (k-l-1)-set belongs to $\binom{k}{k-l-1}N - O_k(1)$ edges of the hypergraph. It follows that the hypergraph \mathcal{H}_A is $(k-l-1,\eta)$ -near regular where $\eta = O_k(1/N)$. Under the above stated conditions on A, let $D^{\mathcal{H}_A}(B_m)$ denote the deviation on the count of solutions to the system Ax = 0 with pairwise distinct entries in a random set B_m . Hence Theorem 1.2 with r = k-l gives, for some constant c_k and every sufficiently large N,

$$\mathbb{P}\left(D^{\mathcal{H}_A}(B_m) > a\right) \leqslant N^{O_k(1)} \exp\left(-\frac{\Omega_k(1)a^{2/(k-l)}}{m}\right),\,$$

for all $a > c_k \left(\frac{m^{k-1}}{N^{l+1}}\right)^{(k-l)/(k-l-1)}$. Since the expected value of the number of solutions is of the order m^k/N^l , for

$$\max\left\{m^{(k-l)/2} (\log N)^{(k-l)/2}, \left(\frac{m^{k-1}}{N^{l+1}}\right)^{(k-l)/(k-l-1)}\right\} \ll a \ll \frac{m^k}{N^l}$$

we obtain exponentially small bounds for moderate deviations, which apply to random sets of cardinality $m \gg N^{2l/(k+l)} (\log N)^{(k-l)/(k+l)}$.

Let $D^{\mathcal{H}_A}(B_p)$ denote the deviation on the count of solutions to the system Ax = 0 with pairwise distinct entries in a random set B_p of $\mathbb{Z}/N\mathbb{Z}$ and let $L^{\mathcal{H}_A}(p)$ be its expected value. By using Theorem 1.3 we obtain the corresponding bound for moderate deviations in the p-model

$$\mathbb{P}\left(D^{\mathcal{H}_A}(B_p) > \delta_N L^{\mathcal{H}_A}(p)\right) = \exp\left(-(1+o(1))\frac{\delta_N^2 p N}{2k^2(1-p)}\right).$$

for every sequence δ_N satisfying

$$\max\left\{\frac{1}{p^{(k+l)/2}}\left(\frac{\log N}{N}\right)^{(k-l)/2}, \frac{1}{\sqrt{pN}}\right\} \ll \delta_N \ll p^{l/(k-l-1)}$$

and so this is applicable when $p \gg (\log N/N)^{(k-l-1)/(k+l-1)}$.

6. Lower Bounds

The aim of this section is to prove that the inequality established by Theorem 1.2 is tight up to a constant in the exponent, when $m = \Omega(N)$, for a family of hypergraphs which satisfy the conditions of the theorem. The *density* of a k-uniform hypergraph \mathcal{H} with h edges is $h/\binom{N}{k}$. We provide examples across a range of densities.

Proposition 6.1. Let $0 < \tau < 1/2$. For each $r \ge 1$ and each sequence γ_n satisfying $\Theta_r(1/n) \le \gamma_n \le \Theta_r(1)$ for infinitely many values of N, there exists an N-vertex (r + 1)-uniform hypergraph \mathcal{H} with density $\Theta_r(\gamma_N)$, which is $(r - 1, \eta)$ -near-regular for $\eta = O_r(1/\gamma_N N)$ and with maximum r-degree Δ_r , and for which

$$\mathbb{P}\left(D^{\mathcal{H}}(B_m) > a\right) \geq N^{-O_{r,\tau}(1)} \exp\left(\frac{-O_{r,\tau}(1)a^{2/r}}{m\Delta_r^{2/r}}\right),$$

for all $\max\{h/N, \eta h\} \leq a \leq \Omega_{r,\tau}(1)h$ and all $\tau N \leq m \leq (1-\tau)N$.

In the cases r = 1 and r = 2 particularly simple constructions exist. For r = 1 one may take a regular graph on half the vertices and leave the remaining vertices isolated. For r = 2one may take two disjoint copies of a regular 3-uniform hypergraph. Even for r = 3 it is not difficult to describe a family of examples: Suppose N = 3s and partition the vertex set into 3 equal parts $V = V_1 \cup V_2 \cup V_3$, a set e of four vertices will be an edge of \mathcal{H} if it has 2 vertices in each of two parts.

In general, for $r \ge 4$, we need to work harder to produce the examples. Based on the r = 2and r = 3 cases we see that it is useful to partition the vertex set. Given a partition $V = V_1 \cup \cdots \cup V_\ell$ into ℓ equal parts of size $s := N/\ell$ we will label the vertices $v_{i,j} : i \in [\ell], j \in [s]$. We call \mathcal{H} an ℓ -part hypergraph if it is invariant under permutations of the parts $V_i : i \in [\ell]$. That is, given a permutation π of $[\ell]$ then $\{v_{i_1,j_1}, \ldots, v_{i_k,j_k}\}$ is an edge of \mathcal{H} if and only if $\{v_{\pi(i_1),j_1}, \ldots, v_{\pi(i_k),j_k}\}$ is an edge of \mathcal{H} . For example, the r = 2 case above is a 3-uniform 2-part hypergraph and the r = 3 case is a 4-uniform 3-part hypergraph. For general r the hypergraphs we construct will be (r + 1)-uniform and ℓ -part for some ℓ chosen sufficiently large. (In fact, we take $\ell = 4(r + 1)!$)

We consider the deviation event $D^{\mathcal{H}}(B_m) > a$ relative to an auxiliary event in which the random set B_m is unevenly distributed with respect to the parts V_1, \ldots, V_{ℓ} . Specifically, we

define E_{ε}^{ℓ} to be the event⁴ that

$$|B_m \cap V_1| = (1+2\varepsilon)\frac{m}{\ell} \pm 1.$$

$$|B_m \cap V_i| = (1-\varepsilon)\frac{m}{\ell} \pm 1 \qquad i = 2, 3.$$

$$|B_m \cap V_i| = \frac{m}{\ell} \pm 1 \qquad i = 4, \dots, \ell.$$

Since

$$\mathbb{P}\left(D^{\mathcal{H}}(B_m) > a\right) \ge \mathbb{P}\left(E_{\varepsilon}^{\ell}\right) \mathbb{P}\left(D^{\mathcal{H}}(B_m) > a \mid E_{\varepsilon}^{\ell}\right)$$
(13)

it will suffice to prove lower bounds on $\mathbb{P}(E_{\varepsilon}^{\ell})$ (see Lemma 6.4) and $\mathbb{P}(D^{\mathcal{H}}(B_m) > a | E_{\varepsilon}^{\ell})$ (see Lemma 6.5).

The responsiveness of the deviation $D^{\mathcal{H}}(B_m)$ to the uneven distribution E_{ε}^{ℓ} depends on the hypergraph \mathcal{H} . In particular, it depends on the coefficients of the following polynomial

$$Q^{\mathcal{H}}(x) = \sum_{e \in E(\mathcal{H})} (1+2x)^{e_1} (1-x)^{e_2+e_3}$$

where $e_i := |e \cap V_i|$ for $i \in [\ell]$, which may be defined for any ℓ -part hypergaph \mathcal{H} . Let $c_j^{\mathcal{H}}$ be the coefficient of x^j in $Q^{\mathcal{H}}(x)$.

Given an (r+1)-uniform ℓ -part hypergraph \mathcal{H} , we define \mathcal{H} to be (r, η, γ) -nice if

- (i) \mathcal{H} is $(r-1,\eta)$ -near-regular,
- (ii) the density of \mathcal{H} is between γ/ℓ^2 and $3\gamma/\ell^2$,
- (iii) $\Delta_r \leq \gamma N$, and
- (iv) $c_r^{\mathcal{H}} \ge \gamma N^{r+1}/\ell^{r+1}$.

Proposition 6.1 follows from the following two propositions. The first states the required lower bound for this family of hypergraphs.

Proposition 6.2. If \mathcal{H} is an (r+1)-uniform ℓ -part hypergraph (for $\ell = O_r(1)$) with maximum r-degree Δ_r , which is (r, η, γ) -nice for $\eta = O_r(1/\gamma N)$ then

$$\mathbb{P}\left(D^{\mathcal{H}}(B_m) > a\right) \geqslant N^{-O_{r,\tau}(1)} \exp\left(\frac{-O_{r,\tau}(1)a^{2/r}}{m\Delta_r^{2/r}}\right),$$

for all $\max\{h/N, \eta h\} \leq a \leq \Omega_{r,\tau}(1)h$ and all $\tau N \leq m \leq (1-\tau)N$.

The second proposition states that such hypergraphs exist.

Proposition 6.3. For all $r \ge 2$ there exists $\ell = \ell(r)$ such that for all sequences $10/n \le \gamma_n \le \ell^{-2}$ for infinitely many values of N, there exists an N-vertex (r+1)-uniform hypergraph \mathcal{H} which is ℓ -part and (r, η, γ_N) -nice with $\eta = O_r(1/\gamma_N N)$.

⁴the ± 1 is only included as the values themselves may not be integers

The proof of Proposition 6.3 is technical and unlikely to be of general interest and so will be given in the appendix.

Our proof of Proposition 6.2 is based on (13) and the following two lemmas. The first provides a bound on $\mathbb{P}(E_{\varepsilon}^{\ell})$.

Lemma 6.4. Let $\tau \in (0, 1/2)$ and $\ell \in \mathbb{N}$. Provided $\tau N \leq m \leq (1 - \tau)N$ we have

$$\mathbb{P}\left(E_{\varepsilon}^{\ell}\right) \geqslant N^{-O_{\ell,\tau}(1)}\exp(-O_{\ell,\tau}(\varepsilon^{2}m))$$

for all $0 < \varepsilon \leq \tau/2(1-\tau)$.

The second lemma shows the effect of the uneven distribution (given by E_{ε}^{ℓ}) on $N^{\mathcal{H}}(B_m)$ (or at least its expected value). We show that for a certain value of ε the conditional expectation exceeds the unconditioned expectation by at least 2a.

Lemma 6.5. There are constants $C_1 = C_1(\ell, \tau)$ and $C_2 = C_2(\ell, \tau)$ such that if \mathcal{H} is an (r+1)-uniform (r, η, γ) -nice ℓ -part hypergraph, then for all $\max\{h/N, \eta h\} \leq a \leq h$, we have

$$\mathbb{E}\left[N^{\mathcal{H}}(B_m) \mid E_{\varepsilon}^{\ell}\right] \ge L^{\mathcal{H}}(m) + 2a$$

for all $C_1(a/h)^{1/r} \leq \varepsilon \leq C_2$.

We now deduce Proposition 6.2 from these two lemmas.

Proof of Proposition 6.2. We begin with an easy observation: if a random variable X has mean μ , has $|X| \leq h$ and has $\mathbb{E}[X|E] \geq \mu + 2a$ for some event E then

$$\mathbb{P}\left(X > \mu + a\right) \geqslant \frac{a}{h} \mathbb{P}\left(E\right) \,.$$

This follows immediately using that $X \leq h \mathbb{1}_{X > \mu + a} + \mu + a$ and taking the conditional expectation.

Taking $E = E_{\varepsilon}^{\ell}$ where $\varepsilon = \Theta_{\ell,\tau}((a/h)^{1/r})$ is given by Lemma 6.5 (observe that if $\Omega_{r,\tau}(1)$ is sufficiently small in Proposition 6.2, the interval for ε in Lemma 6.5 is nonempty), and using that $h \leq aN$, it follows from Lemma 6.4 that

$$\mathbb{P}\left(D^{\mathcal{H}}(B_m) > a\right) \geqslant N^{-O_{\ell,\tau}(1)} \exp\left(-O_{\ell,\tau}(\varepsilon^2 m)\right) = N^{-O_{\ell,\tau}(1)} \exp\left(\frac{-O_{\ell,\tau}(1)a^{2/r}m}{h^{2/r}}\right)$$

Since $\ell = O_r(1)$, the dependence of the constants on ℓ is just a dependence on r. Also, conditions (ii) and (iii) in the definition of (r, η, γ) -nice give us that $h = \Omega_\tau(m^r \Delta_r)$ and so the required bound follows immediately.

We now turn to the proofs of Lemmas 6.4 and 6.5. The proof of Lemma 6.4 is a relatively straightforward exercise with Stirling's approximation, we include it for completeness.

Proof of Lemma 6.4. Note first that the conditions imposed on ε ensure that the event E_{ε}^{ℓ} is nonempty. Denoting t = m/N, observe that

$$\mathbb{P}\left(E_{\varepsilon}^{\ell}\right) = \frac{\binom{s}{(1+2\varepsilon)st}\binom{s}{(1-\varepsilon)st}^{2}\binom{s}{st}^{\ell-3}}{\binom{s\ell}{ts\ell}}.$$

Using Stirling's approximation for the factorials, if $k = \Omega(n)$, we have

$$\binom{n}{k} = (1+o(1))\sqrt{\frac{n}{2\pi k(n-k)}} \cdot \frac{n^n}{k^k(n-k)^{n-k}}$$

Note that if k = cn, where $c \in (0, 1)$, we have

$$\binom{n}{k} = (1+o(1))\Theta_c(n^{-1/2}) \cdot \exp(nH(c)).$$

where $H(x) = -x \log x - (1 - x) \log(1 - x)$. Then

$$\mathbb{P}\left(E_{\varepsilon}^{\ell}\right) = (1+o(1))\Theta_{\ell,\tau}(s^{(1-\ell)/2}) \cdot \exp[s(H((1+2\varepsilon)t) + 2H((1-\varepsilon)t) - 3H(t))]$$

Using Taylor series, $s = m/t\ell$ and $\tau \leq t \leq 1 - \tau$, we obtain that

$$\mathbb{P}\left(E_{\varepsilon}^{\ell}\right) \geqslant N^{-O_{\ell,\tau}} \exp(-O_{\ell,\tau}(\varepsilon^2 m)),$$

as desired.

We will need to work a little more to prove Lemma 6.5. In particular we shall use the following lemma about the coefficients of $Q^{\mathcal{H}}(x)$ in the case \mathcal{H} is (r, η, γ) -nice. Recall that $L^{\mathcal{H}}(m) = \mathbb{E}\left[N^{\mathcal{H}}(B_m)\right]$.

Lemma 6.6. Let \mathcal{H} be an (r+1)-uniform ℓ -part hypergraph which is (r, η, γ) -nice and let m = tN, where $\tau \leq t \leq 1 - \tau$. Then:

 $\begin{array}{l} (i) \ c_0^{\mathcal{H}} \geqslant L^{\mathcal{H}}(m)/t^{r+1};\\ (ii) \ |c_j^{\mathcal{H}}| = O_r(\eta h + h/N), \ if \ 1 \leqslant j \leqslant r-1;\\ (iii) \ |c_{r+1}^{\mathcal{H}}| = O_r(\Delta_r N^r). \end{array}$

Proof. (i) For the first part, observe that $c_0^{\mathcal{H}} = h$ and $L^{\mathcal{H}}(m) = h \frac{(m)_{r+1}}{(N)_{r+1}}$. Because

$$\frac{(m)_{r+1}}{(N)_{r+1}} \leqslant \left(\frac{m}{N}\right)^{r+1} = t^{r+1},$$

it follows that $c_0^{\mathcal{H}} \ge L^{\mathcal{H}}(m)/t^{r+1}$.

_	
	- 1
	- 1
	- 1
	١.

(ii) In this part, we will use that \mathcal{H} is $(r-1,\eta)$ -near-regular. We first observe that

$$c_{j}^{\mathcal{H}} = \sum_{e \in E(\mathcal{H})} \sum_{i=0}^{j} (-1)^{i} 2^{j-i} \binom{e_{2} + e_{3}}{i} \binom{e_{1}}{j-i}$$
$$= \sum_{i=0}^{j} (-1)^{i} 2^{j-i} \sum_{e \in E(\mathcal{H})} \binom{e_{2} + e_{3}}{i} \binom{e_{1}}{j-i}.$$

Note that by double counting,

$$\sum_{e \in E(\mathcal{H})} \binom{e_2 + e_3}{i} \binom{e_1}{j-i} = \sum_{\substack{A \subset V_1, |A| = j-i \\ B \subset V_2 \cup V_3, |B| = i}} d_j(A \cup B).$$

Because \mathcal{H} is $(r-1,\eta)$ -near-regular and $1 \leq j \leq r-1$, $d_j(A \cup B) \in (1 \pm \eta)\bar{d}_j^{\mathcal{H}}$, which gives us

$$\sum_{e \in E(\mathcal{H})} \binom{e_2 + e_3}{i} \binom{e_1}{j - i} \in \binom{s}{j - i} \binom{2s}{i} \vec{d}_j^{\mathcal{H}} (1 \pm \eta).$$

Using that $\bar{d}_j^{\mathcal{H}} = O_r(h/N^j)$, we have

$$\sum_{e \in E(\mathcal{H})} \binom{e_2 + e_3}{i} \binom{e_1}{j - i} = \binom{s}{j - i} \binom{2s}{i} \vec{d}_j^{\mathcal{H}} \pm O_r(\eta h).$$

Therefore

$$c_j^{\mathcal{H}} = \sum_{i=0}^{j} (-1)^i 2^{j-i} {s \choose j-i} {2s \choose i} d_j^{\mathcal{H}} \pm O_r(\eta h).$$

The sum

$$\sum_{i=0}^{j} (-1)^i 2^{j-i} \binom{s}{j-i} \binom{2s}{i}$$

is a polynomial in s of degree at most j. However the coefficient of s^{j} in this polynomial is

$$\sum_{i=0}^{j} \frac{(-1)^{i} \cdot 2^{j-i} \cdot 2^{i}}{(j-i)!i!} = \frac{2^{j}}{j!} \sum_{i=0}^{j} (-1)^{i} \binom{j}{i} = 0.$$

Hence the sum is actually a polynomial in s of degree at most j - 1. Using again that $\bar{d}_j^{\mathcal{H}} = O_r(h/N^j)$ and using that $s = O_r(N)$, we obtain that $|c_j^{\mathcal{H}}| = O_r(\eta h + h/N)$, as desired.

(iii) In this last part, as the hypergraph is (r + 1)-uniform, $|c_{r+1}^{\mathcal{H}}| = O_r(h)$ and since $h = O_r(\Delta_r N^r)$, it follows that $|c_{r+1}^{\mathcal{H}}| = O_r(\Delta_r N^r)$.

Armed with Lemma 6.6 we now prove Lemma 6.5 about the conditional expectation $\mathbb{E}\left[N^{\mathcal{H}}(B_m) \mid E_{\varepsilon}^{\ell}\right].$

Proof of Lemma 6.5. Let \mathcal{H} be a (r, η, γ) -nice ℓ -part hypergraph and let $\varepsilon = C_1(a/h)^{1/r}$ for a constant C_1 which we can choose later, as a function of ℓ and τ . We must prove that

$$\mathbb{E}\left[N^{\mathcal{H}}(B_m) \mid E_{\varepsilon}^{\ell}\right] \geqslant L^{\mathcal{H}}(m) + 2a.$$

We first observe that

$$\mathbb{E}\left[N^{\mathcal{H}}(B_m)|E_{\varepsilon}^{\ell}\right] = \sum_{e \in E(\mathcal{H})} \prod_{i=1}^{\ell} \frac{\binom{s-e_i}{|B_m \cap V_i|-e_i}}{\binom{s}{|B_m \cap V_i|}}$$
$$= \sum_{e \in E(\mathcal{H})} \prod_{i=1}^{\ell} \frac{\binom{|B_m \cap V_i|}{e_i}}{\binom{s}{e_i}}.$$

As we are conditioned on the event E_{ε}^{ℓ} , we have $|B_m \cap V_1| = (1+2\varepsilon)st \pm 1$, $|B_m \cap V_2| = |B_m \cap V_3| = (1-\varepsilon)st \pm 1$ and $|B_m \cap V_i| = st \pm 1$ for all $4 \leq i \leq \ell$. Using that $(N)_k = N^k(1\pm O(1/N))$, we obtain

$$\mathbb{E}\left[N^{\mathcal{H}}(B_m)|E_{\varepsilon}^{\ell}\right] = t^{r+1}Q^{\mathcal{H}}(\varepsilon) \pm O_r(h/N)$$

By Lemma 6.6, choosing C_2 sufficiently small, there are positive constants α and β depending on r and τ , such that

$$\mathbb{E}\left[N^{\mathcal{H}}(B_m)|E_{\varepsilon}^{\ell}\right] \ge L^{\mathcal{H}}(m) + \alpha \varepsilon^r \Delta_r N^r - \beta(\eta h + h/N)$$

As $\varepsilon = C_1(a/h)^{1/r}$ and $\Delta_r \ge h/N^r$,
$$\mathbb{E}\left[N^{\mathcal{H}}(B_m)|E_{\varepsilon}^{\ell}\right] \ge L^{\mathcal{H}}(m) + \alpha C_1 a - \beta(\eta h + h/N).$$

Since $a \ge \max\{h/N, \eta h\}$, we can choose $C_1 = (2 + 2\beta)/\alpha$, obtaining

$$\mathbb{E}\left[N^{\mathcal{H}}(B_m)|E_{\varepsilon}^{\ell}\right] \geqslant L^{\mathcal{H}}(m) + 2a,$$

as desired.

7. Concluding remarks and open questions

We showed in Section 6 that Theorem 1.2 is best possible up to the implicit constants, at least when $m = \Theta(N)$. However, for particular cases, such as the 3-term arithmetic progressions in \mathbb{Z}_N we do not believe that the bounds obtained are best possible. A lower bound of the form

$$\mathbb{P}\left(D^{3}(B_{m}) > a\right) \ge \exp\left(\frac{-O(1)a^{2/3}N^{2/3}}{m}\right)$$

may be proved by considering the probability that the interval $\{1, \ldots, \lfloor N/3 \rfloor\}$ contains significantly more points than expected. Specifically $(1 + \varepsilon)N/3$ points where $\varepsilon \approx a^{1/3}/N^{2/3}$. Writing $D^3(B_m)$ for the deviation of number of three term arithmetic progressions, it remains an interesting open problem to determine the value of

$$\log \mathbb{P}\left(D^3(B_m) > a\right)$$

up to a constant factor. We stress that this problem is open even in the dense case $m = \Theta(N)$, for $N \ll a \ll N^2$.

		I
- L		J

$$\mathbb{P}\left(D^{\mathcal{H}}(B_m) > a\right) \leqslant N^{O_k(1)} \exp\left(\frac{-\Omega_k(1)a}{m\Delta_2}\right)$$
$$\mathbb{P}\left(D^{\mathcal{H}}(B_m) > a\right) \leqslant N^{O_k(1)} \exp\left(\frac{-\Omega_k(1)aN^{1/2}}{m^{3/2}\Delta_2}\right)$$

to

$$\mathbb{P}\left(D^{\mathcal{H}}(B_m) > a\right) \leqslant N^{O_k(1)} \exp\left(\frac{-\Omega_k(1)aN}{m^2\Delta_2}\right)$$

may hold in general for regular 3-uniform hypergraphs. This final bound, if true, would be best possible up to the implicit constants.

For simplicity we have stated our open problems for 3-uniform hypergraphs. The analogous problems for k-uniform hypergraphs with one of our regularity properties are also open.

Improvements in results for the *m*-model B_m for m = o(N) would almost certainly allow one to extend the range of deviations covered by Theorem 1.3 for the *p*-model B_p .

References

- K. Azuma, Weighted sums of certain dependent random variables, Tohoku Math. J. 19 (1967), no. 3, 357–367.
- [2] R.R. Bahadur, Some approximations to the binomial distribution function, Ann. Math. Statist. 31 (1960), no. 1, 43–54.
- [3] J. Balogh, R. Morris and W. Samotij, Random sum-free subsets of abelian groups, Israel J. Math. 199 (2014), no. 2, 651–685.
- [4] B.B. Bhattacharya and S. Mukherjee, A Note on Replica Symmetry in Upper Tails of Mean-Field Hypergraphs, arXiv:1812.09841.
- [5] B.B. Bhattacharya, S. Ganguly, X. Shao, and Y. Zhao, Upper tails for arithmetic progressions in a random set, Int. Math. Res. Not. 1 (2020), 167–213.
- [6] R. Berkowitz, A. Sah and M. Sawhney, Number of arithmetic progressions in dense random subsets of Z/nZ, arXiv:1907.11807
- [7] S. Boucheron, G. Lugosi and P. Massart, *Concentration inequalities*, Oxford University Press, Oxford, 2013.
- [8] J. Briët and S. Gopi, Gaussian width bounds with applications to arithmetic progressions in random settings, Int. Math. Res. Not. 22 (2020), 8673–8696.
- [9] S. Chatterjee and S.R.S. Varadhan, The large deviation principle for the Erdös-Rényi random graph, European J. Combin. 32 (2011), no. 7, 1000–1017.
- S. Chatterjee, An introduction to large deviations for random graphs, Bull. Amer. Math. Soc. (N. S.) 53 (2016), no. 4, 617–642.
- [11] D. Dellamonica, Y. Kohayakawa, S.J. Lee, V. Rödl and W. Samotij, The number of B_h-sets of a given cardinality, Proc. Lond. Math. Soc. (3) 116 (2018), no. 3, 629–669.

34 GONZALO FIZ PONTIVEROS, SIMON GRIFFITHS, MATHEUS SECCO, AND ORIOL SERRA

- [12] H. Döring and P. Eichelsbacher, Moderate deviations in a random graph and for the spectrum of Bernoulli random matrices, Electron. J. Probab. 14 (2009), 2636–2656.
- [13] V. Féray, P.L. Méliot and A. Nikeghbali, Mod-φ convergence I: Normality zones and precise deviations, Springer Briefs in Probability and Mathematical Statistics, Springer, 2016.
- [14] C. Goldschmidt, S. Griffiths and A. Scott, Moderate deviations of subgraph counts in the Erdős-Rényi random graphs G(n, m) and G(n, p), Trans. Amer. Math. Soc. **373** (2020), 5517–5585.
- [15] S. Griffiths, C. Koch and M. Secco, Deviation probabilities for arithmetic progressions and irregular discrete structures, arXiv:2012.09280
- [16] R. Graham, V. Rödl and A. Ruciński, On Schur properties of random subsets of integers, J. Number Theory 61 (1996), 388–408.
- [17] M. Harel, F. Mousset and W. Samotij, Upper tails via high moments and entropic stability, arXiv:1904.08212.
- [18] W. Hoeffding, Probability inequalities for sums of bounded random variables, J. Amer. Statist. Assoc. 58 (1963), 13–30.
- [19] S. Janson, K.Oleszkiewicz and A. Ruciński, Upper tails for subgraph counts in random graphs, Israel J. Math. 142 (2004), 61–92.
- [20] S. Janson and A. Ruciński, When are small subgraphs of a random graph normally distributed?, Probab. Theory Related Fields 78 (1988), no. 1, 1–10.
- [21] S. Janson and A. Ruciński, Upper tails for counting objects in randomly induced subhypergraphs and rooted random graphs, Ark. Mat. 49 (2011), 79–96.
- [22] S. Janson and L. Warnke, The lower tail: Poisson approximation revisited, Random Structures Algorithms 48 (2015), no.2, 219–246.
- [23] J.H.Kim and V.H. Vu, Concentration of multivariate polynomials and its applications, Combinatorica 20 (2000), no. 3, 417–434.
- [24] Y. Kohayakawa, S. J. Lee, V. Rödl and Wojciech Samotij, The number of Sidon sets and the maximum size of Sidon sets contained in a sparse random set of integers, Random Structures Algorithms 46 (2015), no. 1, 1–25.
- [25] V. Rödl and A. Ruciński, Rado partition theorem for random subsets of integers Proc. London Math. Soc. 74, (1997), 481–502.
- [26] J. Rué, O. Serra and L. Vena, Counting configuration-free sets in groups, European J. Combin. 66 (2017), 281–307.
- [27] J. Rué, C. Spiegel and A. Zumalacárregui, Threshold functions and Poisson convergence for systems of equations in random sets Math. Z. 288 (2018), no. 1-2, 333–360.
- [28] I.Z. Ruzsa, Solving a linear equation in a set of integers I, Acta Arithmetica LXV.3 (1993), 259–282.
- [29] M. Schacht, Extremal results for random discrete structures, Ann. of Math. (2) 184 (2016), 333–365.
- [30] L. Warnke, On the method of typical bounded differences, Combin. Probab. Comput. 25 (2016), no. 2, 269–299.
- [31] L. Warnke, Upper tails for arithmetic progressions in random subsets, Israel J. Math. 221 (2017) 221– 317.
- [32] L. Warnke, On the missing log in upper tail estimates, J. Combin. Theory Ser. B 140 (2020), 98–146.

8. Appendix: Lower bound construction

In this appendix we construct the hypergraphs whose existence is claimed by Proposition 6.3. We shall recall the necessary definitions and state a proposition (Proposition 8.1) which implies Proposition 6.3. Let us recall that we call a hypergraph \mathcal{H} an ℓ -part hypergraph if its vertex set V = [N]may be partitioned $V = V_1 \cup \cdots \cup V_\ell$ into ℓ equal parts of size $s := N/\ell$ in such a way that when we label the vertices $v_{i,j} : i \in [\ell], j \in [s]$ the hypergraph \mathcal{H} is invariant under permutation of the parts $i \in [\ell]$.

Recall also that we define the polynomial

$$Q^{\mathcal{H}}(x) = \sum_{e \in E(\mathcal{H})} (1+2x)^{e_1} (1-x)^{e_2+e_3}$$

where $e_i := |e \cap V_i|$ for $i \in [\ell]$, and that $c_i^{\mathcal{H}}$ denotes the coefficient of x^j in $Q^{\mathcal{H}}(x)$.

Finally, we defined an (r+1)-uniform ℓ -part hypergraph \mathcal{H} to be (r, η, γ) -nice if

- (i) \mathcal{H} is $(r-1,\eta)$ -near-regular,
- (ii) the density of \mathcal{H} is between γ/ℓ^2 and $3\gamma/\ell^2$,
- (iii) $\Delta_r \leq \gamma N$, and
- (iv) $c_r^{\mathcal{H}} \ge \gamma N^{r+1}/\ell^{r+1}$.

The proposition we must prove states that such hypergraphs exist. Specifically, for all $r \ge 2$ there exists $\ell = \ell(r)$ such that for all sequences $\Theta(1/n) \le \gamma_n \le \ell^{-2}$ there are infinitely many (r+1)-uniform ℓ -part hypergraphs \mathcal{H} which are (r, η, γ_N) -nice with $\eta = O_r(1/\gamma_N N)$.

To this end, let us fix $r \ge 2$. We shall work with ℓ throughout the definition without fixing its value. It will be clear at the end of the proof that if ℓ is taken sufficiently large then for all sufficiently large multiples $N = s\ell$ of ℓ the construction satisfies all 4 conditions above. With this in mind we may always assume that $s = N/\ell$ is sufficiently large.

We work with a fixed value of $\gamma \in [10\ell^2/s, 1/2]$ and the example we provide will contain $\approx 2\gamma N^{r+1}/(r+1)!\ell^2$ edges. This gives us the claimed range of densities.

We now state more precisely the type of construction we shall give. In particular we shall fix the choice of ℓ as $\ell = 4(r+1)!$.

Proposition 8.1. Let $r \ge 2$, and fix $\ell = 4(r+1)!$. There exists a constant C = C(r) such that for all $0 \le \gamma \le 1/2$ and all $s \ge 10\ell^2/\gamma$ there exists an ℓ -part (r+1)-uniform hypergraph \mathcal{H} on $s\ell$ vertices which is $(r, C/\gamma s, \gamma)$ -nice.

It is easily checked that Proposition 6.3 follows from this proposition.

A major challenge in the proof of Proposition 8.1 is to ensure that the hypergraph \mathcal{H} we construct is $(r - 1, C/\gamma s)$ -near-regular. This is more difficult than one might imagine. We shall define \mathcal{H} with enough symmetry that any two sets with the same intersection pattern with the parts have (essentially) the same degree in \mathcal{H} . This reduces the number of degrees we need to check to a finite number (the number of partitions of the number r - 1). In order to formalise these details, and define \mathcal{H} , we introduce the concept of type.

8.1. Types and one-type hypergraphs. We continue to use the partition $V = V_1 \cup \ldots V_\ell$ into ℓ parts defined above. Given a set e of elements of V consider the vector $(e_1, \ldots, e_\ell) :=$ $(|e \cap V_1|, |e \cap V_2|, \dots, |e \cap V_\ell|)$. We define $\mathbf{x}(e)$, the *type* of e, to be the vector obtained from (e_1, \dots, e_ℓ) by placing its entries in decreasing order and removing the 0s. For example



Figure 1. An edge of type (2,1,1) and an edge of type (1,1,1,1)

Note that the set of possible types of edges in our (r+1)-uniform hypergraph \mathcal{H} is \mathcal{P}_{r+1} , the set of partitions of r+1. We now define certain families of one-type hypergraphs. Given a type $\mathbf{x} \in \mathcal{P}_{r+1}$ we may define $\mathcal{H}_{\mathbf{x}}$ to be the hypergraph consisting of all edges of type \mathbf{x} .

We shall also consider sparser subhypergraphs of $\mathcal{H}_{\mathbf{x}}$. We remark that each vertex $v \in V$ may be viewed as a pair (i_v, j_v) where $i_v \in [\ell]$ and $j_v \in [s]$. Given $\alpha \in (0, 1)$ we say that a sequence $(i_1, j_1), \ldots, (i_{r+1}, j_{r+1})$ of elements of V is α -good if

$$j_1 + \dots + j_{r+1} \in \{1, \dots, \lfloor \alpha s \rfloor\} \pmod{s}$$

We may define $\mathcal{H}_{\mathbf{x}}^{\alpha}$ to be the hypergraph obtained from $\mathcal{H}_{\mathbf{x}}$ by keeping those edges $e = \{(i_1, j_1), \ldots, (i_{r+1}, j_{r+1})\}$ which are α -good. It will also be useful to define $\mathcal{H}_{\mathbf{x}}^{\alpha}$ to be the equivalent hypergraph, but in which multi-sets are also permitted. That is $\mathcal{H}_{\mathbf{x}}^{\alpha}$ contains those multi-sets of r + 1 vertices which have type \mathbf{x} and are α -good.

The idea of this sparsification is that roughly α proportion of the edges of $\mathcal{H}_{\mathbf{x}}$ remain. In fact the same is true for degrees and higher degrees, up to the *r*-degree which will satisfy $\Delta_r = O(\alpha s)$, where the constant in the $O(\cdot)$ may depend on *r* and ℓ .

8.2. Multi-type hypergraphs and degrees: We construct our hypergraph \mathcal{H} as a union of one-type hypergraphs $\mathcal{H}^{\alpha}_{\mathbf{x}}$ defined above. In fact we do not need to consider all types. We consider only the type (r+1) and types of the form $(\mathbf{x}, 1, 1)$ where $\mathbf{x} \in \mathcal{P}_{r-1}$. For each vector $\mathbf{x} \in \mathcal{P}_{r-1}$ we will define a vector $\mathbf{x}^+ \in \mathcal{P}_{r+1}$ that "extends" \mathbf{x} . For $\mathbf{x} = (r-1)$ we define $\mathbf{x}^+ = (r+1)$, and all other $\mathbf{x} \in \mathcal{P}_{r-1}$ we define $\mathbf{x}^+ = (\mathbf{x}, 1, 1)$. Given $\boldsymbol{\alpha} = (\alpha_{\mathbf{x}} : \mathbf{x} \in \mathcal{P}_{r-1})$, which assigns a value $\alpha_{\mathbf{x}} \in (0, 1)$ to each partition \mathbf{x} of r-1, we define the hypergraph

$$\mathcal{H}^{oldsymbol{lpha}} := \bigcup_{\mathbf{x}\in\mathcal{P}_{r-1}}\mathcal{H}^{lpha_x}_{\mathbf{x}^+}.$$

To re-iterate, the hypergraph \mathcal{H}^{α} is a union of hypergraphs $\mathcal{H}^{\alpha}_{\mathbf{y}}$ where the type \mathbf{y} is either (r+1) or of the form $(\mathbf{x}, 1, 1)$ for some $\mathbf{x} \in \mathcal{P}_{r-1}$. We may think of the $\alpha_{\mathbf{x}}$ as "weights"

which are associated with the types \mathbf{x}^+ . We may also define $\bar{\mathcal{H}}^{\alpha}$ to be the equivalent in which multi-sets are permitted.

We now study degrees of vertices and sets in \mathcal{H}^{α} . These degrees depend on the weights $\alpha_{\mathbf{x}}$ in a predictable manner, and we shall choose the weights to ensure \mathcal{H}^{α} is $(r-1, C/\gamma s)$ near-regular. It will be extremely helpful to us that the degree of an (r-1) set A of type $\mathbf{x} \in \mathcal{P}_{r-1}$ depends strongly on $\alpha_{\mathbf{x}}$ and weakly on all of the other weights $\alpha_{\mathbf{y}}$. Let us first
study the degree of the type \mathbf{x} abstractly, and then relate this to the degree of A itself.

Given $\mathbf{x} \in \mathcal{P}_{r-1}$ let |x| denote the number of entries in the sequence \mathbf{x} . A prototypical set of type \mathbf{x} is the set $A^0_{\mathbf{x}}$ which contains the elements $(1, 1), \ldots, (1, x_1)$ of V_1 the elements $(2, 1), \ldots, (2, x_2)$ of V_2 and so on up to $(a, 1), \ldots, (a, x_a)$, where a := |x|. The abstract degree of \mathbf{x} will be a slight modification of the degree of $A^0_{\mathbf{x}}$ in \mathcal{H}^{α} in that we will work with the multi-set version $\bar{\mathcal{H}}^{\alpha}$. We define the abstract degree $d^{\alpha}_{\mathbf{x}}$ of type \mathbf{x} to be the number of pairs $u, u' \in V$ such that $A^0_{\mathbf{x}} \cup \{u, u'\} \in \bar{\mathcal{H}}^{\alpha}$.

With the following two lemmas we prove that the degrees of sets A of type \mathbf{x} are very well approximated by $d_{\mathbf{x}}^{\alpha}$ and that $d_{\mathbf{x}}^{\alpha}$ itself is very well controlled in terms of the weight $\alpha_{\mathbf{x}}$.

Lemma 8.2. Let $\mathbf{x} \in \mathcal{P}_{r-1} \setminus \{(r-1)\}$ and let $A \subseteq V$ be a subset of r-1 vertices of type \mathbf{x} . Let $\boldsymbol{\alpha} = (\alpha_{\mathbf{x}} : \mathbf{x} \in \mathcal{P}_{r-1})$ and let α' be the largest value of $\alpha_{\mathbf{y}}$ for $\mathbf{y} \in \mathcal{P}_{r-1} \setminus \{(r-1)\}$. Then the degree of A in $\mathcal{H}^{\boldsymbol{\alpha}}$ differs from $d_{\mathbf{x}}^{\boldsymbol{\alpha}}$ by at most $O(\alpha'n)$.

Proof. By the symmetry of the definition of $\overline{\mathcal{H}}^{\alpha}$, the degree of A in $\overline{\mathcal{H}}^{\alpha}$ is precisely $d_{\mathbf{x}}^{\alpha}$. The degree of A in \mathcal{H}^{α} may be slightly smaller as we no longer count multi-sets. It therefore suffices to bound the number of α' -good multi-sets of r + 1 vertices which contain A. There are r choices of which element to repeat (one of the r - 1 elements of A or the new rth element) and in each case only $\alpha_{\mathbf{x}}n \leq \alpha'n$ choices of the new rth element such that the resulting multi-set is $\alpha_{\mathbf{x}}$ -good.

Lemma 8.3. Let $\mathbf{x} \in \mathcal{P}_{r-1}$, and let $|\mathbf{x}| = a \ge 2$. Let $\boldsymbol{\alpha} = (\alpha_{\mathbf{x}} : \mathbf{x} \in \mathcal{P}_{r-1})$ and let $\alpha(a)$ be the largest value of $\alpha_{\mathbf{y}}$ for $\mathbf{y} \in \mathcal{P}_{r-1}$ with $2 \le |\mathbf{y}| \le a - 1$. Then the abstract degree $d_{\mathbf{x}}^{\boldsymbol{\alpha}}$ of \mathbf{x} satisfies

$$\binom{\ell-a}{2}\alpha_{\mathbf{x}}s^{2} - O(s) \leqslant d_{\mathbf{x}}^{\boldsymbol{\alpha}} \leqslant \binom{\ell-a}{2}\alpha_{\mathbf{x}}s^{2} + r^{4}\ell\alpha(a)s^{2}$$

Proof. The abstract degree $d_{\mathbf{x}}^{\alpha}$ of \mathbf{x} is the degree of $A_{\mathbf{x}}^{0}$ in $\overline{\mathcal{H}}^{\alpha}$. Since $\overline{\mathcal{H}}^{\alpha}$ contains $\mathcal{H}_{\mathbf{x}}^{\alpha_{\mathbf{x}}}$ we may prove the lower bound simply by observing that $A_{\mathbf{x}}^{0}$ is contained in $\binom{\ell-a}{2}\alpha_{\mathbf{x}}s^{2} - O(s)$ edges of $\mathcal{H}_{\mathbf{x}}^{\alpha_{\mathbf{x}}}$. To see this note that the edges of $\mathcal{H}_{\mathbf{x}}^{\alpha_{\mathbf{x}}}$ are of type $\mathbf{x}^{+} = (\mathbf{x}, 1, 1)$ so we may extend from $A_{\mathbf{x}}^{0}$ to an edge of type \mathbf{x}^{+} by adding one vertex to each of two new parts of $V_{1} \cup \cdots \cup V_{\ell}$, there are $\binom{\ell-a}{2}$ choices of the two other parts, *s* choices of an element in one, and $|\alpha_{\mathbf{x}}s| = \alpha_{\mathbf{x}}s \pm 1$ choices of an element in the second such that the resulting set is $\alpha_{\mathbf{x}}$ -good.

To prove the upper bound we must consider all the other $\overline{\mathcal{H}}_{\mathbf{y}}^{\alpha_{\mathbf{y}}}$ subhypergraphs of $\overline{\mathcal{H}}^{\alpha}$. In fact, if \mathbf{y} is not closely related to \mathbf{x} then $A_{\mathbf{x}}^{0}$ will not be included in any edges of type \mathbf{y}^{+} (for example an edge of type (3, 1) is not contained in any of type (2, 2, 1, 1)). In fact it

is necessary that $x_i \leq y_i$ for all i = 1, ..., a, and $x_i \leq y_i^+$ for all i = 1, ..., a. This last condition implies that $|\mathbf{y}| \geq a - 2$. The cases that $\mathbf{y} \in \mathcal{P}_{r-1}$ and $x_i \leq y_i^+$ for all i = 1, ..., a correspond to just three possible cases:

- (i) $x_a = 1$ and there is a unique j < a such that $y_j = x_j + 1$, and $y_i = x_i$ for $i \in [a-1] \setminus \{j\}$.
- (ii) $x_{a-1} = x_a = 1$ and there is a unique j < a such that $y_j = x_j + 2$, and $y_i = x_i$ for $i \in [a-2] \setminus \{j\}$.
- (iii) $x_{a-1} = x_a = 1$ and there is a pair j < j' < a such that $y_j = x_j + 1$, $y_{j'} = x_{j'} + 1$ and $y_i = x_i$ for $i \in [a-2] \setminus \{j, j'\}$.

It is easily verified that the number of ways to extend $A^0_{\mathbf{x}}$ to an $\alpha_{\mathbf{y}}$ -good edge of type \mathbf{y} is at most $r\ell \alpha' s^2$ in case (i), at most $r\alpha' s^2$ in case (ii) and at most $r^2 \alpha' s^2$ in case (iii). Since each case corresponds to a bounded number of sequences \mathbf{y} (certainly at most r^2 in each case) we deduce the upper bound.

In fact, there are some particular types \mathbf{x} which we shall analyse particularly carefully so it is useful to count the degree even more carefully in these cases.

Lemma 8.4. If $\mathbf{x} = (r-1)$ then every (r-1)-tuple of type \mathbf{x} has degree

$$\frac{1}{2} \lfloor \alpha_{\mathbf{x}} s^2 \rfloor + O(s)$$

If $|\mathbf{x}| = 2$ then every (r-1)-tuple of type \mathbf{x} has degree

$$\binom{\ell-2}{2}\alpha_{\mathbf{x}}s^2 + O(s)\,.$$

Proof. First, for $\mathbf{x} = (r-1)$, an (r-1)-tuple is of type \mathbf{x} if it consists of r-1 vertices in the same part. Without loss of generality consider $(1, j_1), \ldots, (1, j_{r-1})$. The only edges of \mathcal{H}^{α} which contain such a tuple are the edges of type $\mathbf{x}^+ = (r+1)$, which consist of r+1 elements in the same part. There are clearly $\binom{s-r+1}{2}$ ways to extend to an (r+1) set of this type. The number of choices which produce an $\alpha_{\mathbf{x}}$ -good edge is between

$$\frac{1}{2}s\lfloor\alpha_{\mathbf{x}}s\rfloor - r\alpha_{\mathbf{x}}s$$
 and $\frac{1}{2}s\alpha_{\mathbf{x}}s$.

Since both these expressions are of the form

$$\frac{1}{2}s\lfloor\alpha_{\mathbf{x}}s\rfloor + O(s)$$

this completes the proof of the first part.

Now consider \mathbf{x} with $|\mathbf{x}| = 2$. This means that $\mathbf{x} = (r-1-i, i)$ for some $1 \leq i \leq \lfloor (r-1)/2 \rfloor$. It may be easily checked in these cases that an (r-1)-tuple of type \mathbf{x} is not contained in any \mathbf{y}^+ for $\mathbf{y} \neq \mathbf{x}$, and so the degree of this tuple in \mathcal{H}^{α} is the degree of this tuple in $\mathcal{H}^{\alpha_{\mathbf{x}}}_{\mathbf{x}}$. To complete the proof it suffices to observe that this degree is $\binom{\ell-2}{2}\alpha_{\mathbf{x}}s^2 + O(s)$. This follows easily from the fact that there are $\binom{\ell-2}{2}$ choices of the pair of parts of the last two vertices, s choices for a vertex in the first part, and finally $\lfloor \alpha_{\mathbf{x}} s \rfloor = \alpha_{\mathbf{x}} s \pm 1$ choices of the second vertex, in such a way that an $\alpha_{\mathbf{x}}$ -good edge is created.

8.3. Choice of the weights – completing the proof of Proposition 8.1. As we have discussed above we shall define the hypergraph \mathcal{H} as \mathcal{H}^{α} for a particular choice of the vector of weights α . Our aim is to select the weights such that all (r-1)-tuples have degree approximately γs^2 . To be more precise, all such degrees will be $\gamma s^2 \pm Cs$. This will be sufficient to ensure that the hypergraph \mathcal{H} satisfies conditions (i),(ii) and (iii) of the $(r, C/s, \gamma)$ -nice definition. Once this has been achieved we shall verify that condition (iv) also holds for \mathcal{H} .

Lemma 8.5. If the constant C is chosen sufficiently large then there is a choice of the weights $\boldsymbol{\alpha} = (\alpha_{\mathbf{x}} : \mathbf{x} \in \mathcal{P}_{r-1})$ such that:

(a) The degree of each (r-1)-set is $\gamma s^2 \pm Cs$ (b) $\alpha_{(r-1)} = 2\gamma$ (c) The weights satisfy $0 \leq \alpha_{\mathbf{x}} \leq 4\gamma \ell^{-2}$ for all $\mathbf{x} \in \mathcal{P}_{r-1} \setminus \{(r-1)\}$.

Proof. We shall prove this statement by induction. For each $1 \leq j \leq r-1$ consider the statement that there exists a weighting such that (a) is satisfied for sets whose type **x** has $|\mathbf{x}| \leq j$, (b) and (c) are satisfied, and $\alpha_{\mathbf{x}} = 0$ for all $|\mathbf{x}| > j$.

The base case (j = 1) of this statement is proved by taking $\alpha_{(r-1)} = 2\gamma$. It follows from Lemma 8.4 that the degree of a set of type (r-1) is

$$\frac{1}{2}\alpha_{(r-1)}s^2 + O(s) = \gamma s^2 + O(s),$$

as required.

We shall also treat j = 2 as a base case. We extend the choice above, and so $\alpha_{(r-1)} = 2\gamma$. Note that a set of type **x** with $|\mathbf{x}| = 2$ will not be contained in a part V_i , and so will not be contained in any of the edges of $\mathcal{H}_{(r-1)}^{\alpha_{(r-1)}}$. We define $\alpha_{\mathbf{x}} = \gamma/\binom{\ell-2}{2}$. And so, again by Lemma 8.4, we have that a set of type **x** has degree

$$\binom{\ell-2}{2}\alpha_{\mathbf{x}}s^2 + O(s) = \gamma s^2 + O(s).$$

For the induction step, suppose the weighting has been defined for $||\mathbf{x}|| \leq j - 1$ we shall show how to extend the weighting for $\alpha_{\mathbf{x}}$ with $|\mathbf{x}| = j$. Fix \mathbf{x} with $|\mathbf{x}| = j$. We first focus on the abstract degree $d_{\mathbf{x}}^{\alpha}$ of \mathbf{x} . We use Lemma 8.3. Recall that in Lemma 8.3 $\alpha(a)$ is defined to be $\max\{\alpha_{\mathbf{y}}: 2 \leq |\mathbf{y}| \leq a - 1\}$. By the induction hypothesis we may assume that $\alpha(j) \leq 4\gamma \ell^{-2}$. It follows that the contribution to $d_{\mathbf{x}}^{\alpha}$ from the weights $\alpha_{\mathbf{y}}$ with $|\mathbf{y}| \leq j - 1$ is at most we have that the contribution of all edges of type \mathbf{y}^+ where $|\mathbf{y}| < |\mathbf{x}|$ is at most

$$r^4 \ell \alpha(j) s^2 \leqslant 4\gamma \ell^{-1} r^4 s^2 \leqslant \gamma s^2.$$

Let $0 \leq \gamma_{\mathbf{x}} \leq \gamma$ be such that $\gamma_{\mathbf{x}}s^2$ is the remaining degree necessary, i.e., γs^2 minus the contributions of earlier **y**. We may define

$$\alpha_{\mathbf{x}} = \frac{\gamma_{\mathbf{x}}}{\binom{\ell - |\mathbf{x}|}{2}}.$$

We observe that $\alpha_{\mathbf{x}}$ are at most $4\gamma\ell^{-2}$, and that the abstract degree $d_{\mathbf{x}}^{\alpha}$ is $\gamma s^2 + O(s)$.

Once the weights $\alpha_{\mathbf{x}}$ have been chosen for all $|\mathbf{x}| = j$ we have by Lemma 8.2 (applied with $\alpha' \leq 4\gamma \ell^{-2}$) that all (r-1) tuples have degree

$$\gamma s^2 + O(s)$$
,

as required.

By taking $\mathcal{H} = \mathcal{H}^{\alpha}$, where the weights α are given by Lemma 8.5, we have produced a hypergraph which satisfies properties (i), and (ii) from the definition of $(r, C/s, \gamma)$ -nice. In fact \mathcal{H} also satisfies condition (iii). To see this we consider two cases. For an *r*-set *A* contained in a part V_i the only edges containing *A* are those of type (r + 1), it follows that the degree of *A* is at most $\alpha_{(r-1)}s = 2\gamma s \leq \gamma N$. For any other *r*-set *A* it is contained in edges of at most *r* of the hypergraphs $\mathcal{H}^{\alpha_x}_{\mathbf{x}^+}$ which make up \mathcal{H} . In each case to extend *A* to an edge requires a choice of a part (at most ℓ possibilities) and a choice of a vertex to complete an $\alpha_{\mathbf{x}}$ -good edge (at most $\alpha_{\mathbf{x}}s \leq 4\gamma s/\ell^2$ possibilities). It follows that the degree of *A* is at most $4r\gamma s/\ell \leq \gamma N$.

All that remains is to prove that

$$c_r^{\mathcal{H}} \geqslant \gamma N^{r+1} / \ell^{r+1}$$
.

This is achieved by this final lemma.

Lemma 8.6. Let $\mathcal{H} = \mathcal{H}^{\alpha}$ for some sequence of weights $\boldsymbol{\alpha} = (\alpha_{\mathbf{x}} : \mathbf{x} \in \mathcal{P}_{r-1})$ with $\alpha_{(r-1)} = 2\gamma$ and

$$0 \leqslant \alpha_{\mathbf{x}} \leqslant 4\gamma \ell^{-2}$$

for all $\mathbf{x} \in \mathcal{P}_{r-1} \setminus \{(r-1)\}$. Then

$$c_r^{\mathcal{H}} \ge \frac{\gamma s!}{r!} = \frac{\gamma N^{r+1}}{r!\ell^{r+1}}.$$

Proof. The hypergraph $\mathcal{H} = \mathcal{H}^{\alpha}$ is a union of hypergraphs

$$\bigcup_{\mathbf{x}\in\mathcal{P}_{r-1}}\mathcal{H}_{\mathbf{x}^+}^{\alpha_x}$$

Let us write $c_r^{\mathbf{x}}$ for the coefficient of x^r in $Q^{\mathcal{H}_{\mathbf{x}^+}^{\alpha_x}}(x)$. By linearity we have that

$$c_r^{\mathcal{H}} = \sum_{\mathbf{x}\in\mathcal{P}_{r-1}} c_r^{\mathbf{x}}$$

We shall prove that

(a)
$$c_r^{(r-1)} \ge \frac{r2^r \gamma s^{r+1}}{(r+1)!}$$
,
(b) $|c_r^{\mathbf{x}}| \le \frac{2^{r+2} \gamma s^{r+1}}{\ell}$ if $|\mathbf{x}| = 2$,
(c) $c_r^{\mathbf{x}} = 0$ if $||x|| \ge 3$.

The required result will now follow from linearity. We recall that $\ell = 4(r+1)!$ and that there are at most r/2 types **x** with |x| = 2. It follows that

$$c_r^{\mathcal{H}} \ge \frac{r2^r \gamma s^{r+1}}{(r+1)!} - \frac{r2^{r+1} \gamma s^{r+1}}{\ell}$$
$$= \frac{r2^r \gamma s^{r+1} \left(\ell - 2(r+1)!\right)}{\ell(r+1)!}$$
$$\ge \frac{r2^{r-1} \gamma s^{r+1}}{(r+1)!}$$
$$\ge \frac{\gamma s^{r+1}}{r!},$$

as required.

We now prove (a), (b), and (c) stated above.

We begin with (a). Since $\alpha_{(r-1)} = 2\gamma$ the hypergraph $\mathcal{H}^{2\gamma}_{(r+1)}$ in question contains (2γ) good edges of type (r+1), that is, (2γ) -good edges contained in one of the parts V_i . The coefficient of ε^n in the corresponding polynomial

$$Q^{\mathcal{H}^{2\gamma}_{(r+1)}}(x) = \sum_{e \in E(\mathcal{H})} (1+2x)^{e_1} (1-x)^{e_2+e_3}$$

 $satisfies^5$

$$c_r^{(r-1)} \ge \frac{7}{4} \gamma \binom{s}{r+1} \left((r+1)2^r + 2(r+1)(-1)^r \right) \\\ge \frac{r2^r \gamma s^{r+1}}{(r+1)!},$$

as required.

We now prove (b). A type \mathbf{x} with $|\mathbf{x}| = 2$ is of the form $\mathbf{x} = (j, r - j - 1)$ for some $(r-1)/2 \leq j \leq r-2$. The corresponding \mathbf{x}^+ is $\mathbf{x}^+ = (j, r-j-1, 1, 1)$. And the corresponding hypergraph consists of $\alpha_{\mathbf{x}}$ -good edges of type \mathbf{x}^+ . A contribution to the coefficient of ε^r is only made by edges which use all 3 of the parts V_1, V_2, V_3 and one other part. We obtain

⁵We replace by 2 by 7/4 as we may lose something when we take the integer part. The reader may easily verify the details.

that

$$\begin{aligned} |c_r^{\mathbf{x}}| &\leq 2\alpha_{\mathbf{x}}(\ell-3) \binom{s}{j} \binom{s}{r-j-1} s^2 \left(2^j + 2^{r-1-j} + 2\right) \\ &\leq \frac{8\gamma \ell s^{r+1} 2^{r-1}}{\ell^2 j! (r-1-j)!} \\ &\leq \frac{2^{r+2} \gamma s^{r+1}}{\ell} \,, \end{aligned}$$

as required.

Finally, we note that part (c) follows immediately from the fact that if $|\mathbf{x}| \ge 3$ then at most r-1 elements of any edge of the corresponding hypergraph lie in the union of $V_1 \cup V_2 \cup V_3$ and so the coefficient of ε^r is 0.

Gonzalo Fiz Pontiveros, Department of Mathematics, Universitat Politècnica de Catalunya, Barcelona, Spain.

Email address: gonzalo.fiz@upc.edu

SIMON GRIFFITHS, DEPARTAMENTO DE MATEMÁTICA, PUC-RIO, RUA MARQUÊS DE SÃO VICENTE 225, GÁVEA, RIO DE JANEIRO 22451-900, BRAZIL

Email address: simon@mat.puc-rio.br

MATHEUS SECCO, DEPARTAMENTO DE MATEMÁTICA, PUC-RIO, RUA MARQUÊS DE SÃO VICENTE 225, GÁVEA, RIO DE JANEIRO 22451-900, BRAZIL

Email address: matheussecco@mat.puc-rio.br

Oriol Serra, Department of Mathematics, Universitat Politècnica de Catalunya, and Barcelona Graduate School of Mathematics, Barcelona, Spain.

Email address: oriol.serra@upc.edu