Erdös-Ko-Rado for Random Hypergraphs: Asymptotics and Stability

Combinatorics, Probability and Computing (2017) 26, 406–422.
doi:10.1017/S0963548316000420

Erdős–Ko–Rado for Random Hypergraphs: Asymptotics and Stability

MARCELO M. GAUY[1†], HIỆP HÀN[2‡] and IGOR C. OLIVEIRA[3§]

[1]Institute of Theoretical Computer Science, ETH Zürich, 8092 Zürich, Switzerland
[2]Instituto de Matemáticas, Pontificia Universidad Católica de Valparaíso, Blanco Viel 596, Cerro Barón, Valparaíso, Chile
[3]Faculty of Mathematics and Physics, Charles University in Prague, Sokolovska 83, 186 75 Prague 8, Czech Republic

Received 6 September 2014; revised 15 May 2016

We investigate the asymptotic version of the Erdős–Ko–Rado theorem for the random k-uniform hypergraph $\mathcal{H}^k(n,p)$. For $2 \leqslant k(n) \leqslant n/2$, let $N = \binom{n}{k}$ and $D = \binom{n-k}{k}$. We show that with probability tending to 1 as $n \to \infty$, the largest intersecting subhypergraph of $\mathcal{H}^k(n,p)$ has size

$$(1+o(1))p\frac{k}{n}N$$

for any

$$p \gg \frac{n}{k}\ln^2\left(\frac{n}{k}\right)D^{-1}.$$

This lower bound on p is asymptotically best possible for $k = \Theta(n)$. For this range of k and p, we are able to show stability as well.

A different behaviour occurs when $k = o(n)$. In this case, the lower bound on p is almost optimal. Further, for the small interval $D^{-1} \ll p \leqslant (n/k)^{1-\varepsilon}D^{-1}$, the largest intersecting subhypergraph of $\mathcal{H}^k(n,p)$ has size $\Theta(\ln(pD)ND^{-1})$, provided that $k \gg \sqrt{n\ln n}$.

Together with previous work of Balogh, Bohman and Mubayi, these results settle the asymptotic size of the largest intersecting family in $\mathcal{H}^k(n,p)$, for essentially all values of p and k.

2010 *Mathematics subject classification*: Primary 05D05
Secondary 05D40

† A significant fraction of this work was completed while the author was at the University of São Paulo. Supported by CNPq grants 132238/2012-8 and 248952/2013-7.

‡ Supported by the Millennium Nucleus Information and Coordination in Networks ICM/FIC RC130003, the FONDECYT Iniciación grant 11150913, FAPESP (Proc. 2010/16526-3 and 2103/03447-6), BEPE (2013/11353-1) and PROBRAL CAPES/DAAD Proc. 430/15.

§ Parts of this work were completed while the author was at Columbia University and as a visitor at the University of São Paulo. Supported in part by CNPq grant 200252/2015-1.

1. Introduction

The Erdős–Ko–Rado theorem [11] is a cornerstone in extremal combinatorics. Let $[n]$ denote the set $\{1,2,\dots,n\}$, and let $\binom{[n]}{k}$ denote the set of all k-element subsets of $[n]$. A family of k-element sets $\mathcal{F} \subset \binom{[n]}{k}$ is called a k-uniform hypergraph on the vertex set $[n]$, and such a hypergraph is called *intersecting* if $A \cap B \neq \emptyset$ holds for every edge $A, B \in \mathcal{F}$. The Erdős–Ko–Rado theorem then states that for $2 \leqslant k \leqslant n/2$, an intersecting family $\mathcal{F} \subset \binom{[n]}{k}$ must satisfy $|\mathcal{F}| \leqslant \frac{k}{n}\binom{n}{k}$. This is best possible, as seen by the *principal* hypergraphs $\mathcal{F}_i$, which consist of all edges containing the fixed element $i \in [n]$.

We investigate a random analogue of the Erdős–Ko–Rado theorem in which the ambient space $\binom{[n]}{k}$ in the theorem is replaced by a random space. Random analogues of extremal results have been studied extensively in the past few decades, and we refer to [6, 8, 23, 24] for the history of this line of research and recent breakthroughs.

The ambient random space we will work with is $\mathcal{H}^k(n,p)$, the binomial random k-uniform hypergraph on the vertex set $[n]$ in which each edge $e \in \binom{[n]}{k}$ is included in $\mathcal{H}^k(n,p)$ independently with probability p. Further, for a k-uniform $\mathcal{H}$, let $i(\mathcal{H})$ denote the size of the largest intersecting subhypergraph of $\mathcal{H}$, that is,

$$i(\mathcal{H}) = \max\{|\mathcal{F}| : \mathcal{F} \subset H \text{ and } \mathcal{F} \text{ is intersecting}\}.$$

In this notation the Erdős–Ko–Rado theorem states that

$$i(\mathcal{H}^k(n,1)) = i\left(\binom{[n]}{k}\right) = \frac{k}{n}\binom{n}{k}.$$

Notation. All asymptotic limits in this paper are taken as $n \to \infty$. If we write $a(n) \ll b(n)$ or $a(n) = o(b(n))$, it means that $a(n)/b(n) \to 0$. In particular, the notation $o(1)$ represents a function that goes to 0 as $n \to \infty$, as usual. For simplicity, we omit floor and ceiling functions, whenever they are not essential. We say that a sequence of events $\mathcal{E}_n$ holds asymptotically almost surely if $\Pr[\mathcal{E}_n] \to 1$ as $n \to \infty$. We use $\ln^d c$ to denote $(\ln c)^d$.

We will be interested in $i(\mathcal{H}^k(n,p))$ for $k = k(n)$ and all $p = p(n) \in (0,1)$. This question was investigated by Balogh, Bohman and Mubayi [3], which obtained very precise results for the size and structure of the largest intersecting family in $\mathcal{H}^k(n,p)$, for $k \leqslant n^{1/2-o(1)}$. For larger k, they obtained asymptotic tight bounds on $i(\mathcal{H}^k(n,p))$, but only for rather large values of p. In general, their result highly depends on the range of k and p, and hence it is slightly cumbersome to state. Therefore, we will only partially discuss it here, and refer to [3] for detailed information. Their result concerning the large range of k is given below in Proposition 1.1.

Proposition 1.1 (Proposition 1.3 in [3]). *Let $\delta = \delta(n) > 0$ and $N = \binom{n}{k}$. If $\ln n \ll k < (1-\delta)n/2$ and $p \gg (1/\delta)((\ln n)/k)^{1/2}$, then almost surely, as $n \to \infty$,*

$$i(\mathcal{H}^k(n,p)) = (1+o(1))p(k/n)N.$$

https://doi.org/10.1017/S0963548316000420 Published online by Cambridge University Press

In other words, for this range of p, the expected size of the intersection of a principal family $\mathcal{F}_i$ with $\mathcal{H}^k(n,p)$ is very close to the size of a maximum intersecting subfamily of $\mathcal{H}^k(n,p)$. We extend this result, and provide an almost complete description of $i(\mathcal{H}^k(n,p))$ as follows.

Theorem 1.2. *For all $0 < \varepsilon < 1$ there exists a constant $C > 0$ for which the following holds. Let $p = p(n) \in (0,1)$, $k = k(n)$, where $2 \leqslant k \leqslant n/2$, $N = \binom{n}{k}$, and $D = \binom{n-k}{k}$. Then almost surely, as $n \to \infty$,*

$$\text{(1)} \qquad i(\mathcal{H}^k(n,p)) = (1 \pm \varepsilon)pN \qquad \text{if } N^{-1} \ll p \ll D^{-1},$$

$$\text{(2)} \qquad i(\mathcal{H}^k(n,p)) \geqslant (1-\varepsilon)\frac{N}{D}\ln(pD) \qquad \text{if } D^{-1} \ll p \leqslant (n/k)D^{-1} \text{ and } k \gg \sqrt{n \ln n},$$

$$\text{(3)} \qquad i(\mathcal{H}^k(n,p)) \leqslant C\frac{N}{D}\ln(pD) \qquad \text{if } D^{-1} \ll p \leqslant (n/k)^{1-\varepsilon}D^{-1},$$

$$\text{(4)} \qquad i(\mathcal{H}^k(n,p)) = (1 \pm \varepsilon)p\frac{k}{n}N \qquad \text{if } p \geqslant C(n/k)\ln^2(n/k)D^{-1}.$$

The first bound follows from a standard deletion argument, and we state it here for completeness. Note also that $i(\mathcal{H}^k(n,p))$ is monotone in p. Hence, in the range of p around $(n/k)D^{-1}$ not mentioned in the theorem, we have $i(\mathcal{H}^k(n,p)) = O((N/D)\ln^2(n/k))$ due to (4).

If k is linear in n, the bounds in (1) and (4) determine $i(\mathcal{H}^k(n,p))$ asymptotically for essentially all p. Here, we have a change of behaviour around D^{-1}. Roughly speaking, for p below D^{-1}, essentially all of $\mathcal{H}^k(n,p)$ is intersecting. Beyond that point, that is, for $p \gg D^{-1}$, the largest intersecting subhypergraph of $\mathcal{H}^k(n,p)$ has size very close to the size of the intersection of a principal hypergraph with $\mathcal{H}^k(n,p)$. Observe that cases (2) and (3) are trivial for $k = \Theta(n)$.

For $k = o(n)$ there is a rather short range of p where $i(\mathcal{H}^k(n,p))$ reveals a 'flat' behaviour. Indeed, the upper bound (3) shows that $i(\mathcal{H}^k(n,p))$ grows slowly with p, since it appears only in the ln-term. The corresponding lower bound in (2) shows that for $k \geqslant n^{1/2+o(1)}$ this bound is tight up to a multiplicative constant. We provide no lower bound for the range $k < n^{1/2-o(1)}$ here, as in this case the result of Balogh, Bohman and Mubayi is more satisfactory. Again, we refer to [3] for further information.

Although the 'flat range' phenomenon might come as a surprise, it has been observed elsewhere. Indeed, in the dense case, that is, for $p = 1$, and for $k = o(n)$, the size of the largest intersecting family is vanishing compared to the ambient space, that is,

$$i\left(\binom{[n]}{k}\right) = \frac{k}{n}\binom{n}{k} = o\left(\binom{n}{k}\right).$$

For these so called 'degenerate' problems, the random analogues typically reveal such an intermediate flat behaviour, as observed for example in [19, 20].

The question of the range of p for which the largest intersecting family $\mathcal{F} \subset \mathcal{H}^k(n,p)$ is indeed the projection of a principal family has been successfully addressed in [3] for $k < n^{1/2-o(1)}$. For larger k, which we are mainly interested in, the problem seems to be more complicated, and has only been studied recently in [14], for constant p. We make no contribution to this question here. However, as well as the bounds on $i(\mathcal{H}^k(n,p))$, we are able to show stability for $k = \Theta(n)$ in the same range for p as in case (4) in Theorem 1.2.

Theorem 1.3. *For every $\beta > 0$ and $\varepsilon > 0$ there exist constants $\delta > 0$ and $C > 0$ for which the following holds. For any $\beta n < k(n) < (1/2-\beta)n$ and $p \geqslant C \cdot D^{-1}$, asymptotically almost surely stability holds, that is, for every intersecting family $\mathcal{F} \subset \mathcal{H}^k(n,p)$ of size $|\mathcal{F}| \geqslant (1-\delta)p(k/n)N$, there is an element $i \in [n]$ that is contained in all but at most $\varepsilon p(k/n)N$ elements of $\mathcal{F}$.*

In the dense case, that is, for $p = 1$, the result was proved by Friedgut [12]. Indeed, the proof of Theorem 1.3 relies on the result of Friedgut and on a removal lemma for the Kneser graph due to Friedgut and Regev [13].

Further results and organization. In proving Theorems 1.2 and 1.3, it will be convenient for us to work with the Kneser graph $K(n,k)$. The vertex set of this graph is $\binom{[n]}{k}$, and two k-element sets form an edge if and only if they are disjoint. Hence $K(n,k)$ is a $\binom{n-k}{k}$-regular graph on $\binom{n}{k}$ vertices, and a hypergraph $\mathcal{F} \subset \binom{[n]}{k}$ is intersecting if and only if $\mathcal{F}$ is an independent set in $K(n,k)$. Further, let $K(n,k,p)$ denote the subgraph of $K(n,k)$ induced on the random *vertex* set obtained by including each vertex from $\binom{[n]}{k}$ independently with probability p. Due to the correspondence, all bounds on intersecting subgraphs of $\mathcal{H}^k(n,p)$ will follow from corresponding bounds on the size of largest independent sets in $K(n,k,p)$.

Using this translation, Theorem 1.2 follows from a more general scheme which relies on the technical Proposition 2.4 and Lemma 2.1, to be introduced in the next section. Further, for Theorem 1.3 we will need Lemma 2.3, which together with Lemma 2.1 will be proved in Section 3. Based on these results, we will give the proofs of Theorems 1.2 and 1.3 in Section 2.

In general, the proof scheme based on Proposition 2.4 and Lemma 2.1 can be used to bound the size of the largest independent sets in random subgraphs of any D-regular graph G (in fact a sequence of graphs). Here, by random subgraph we mean the graph induced on a binomial random subset of the vertex set. This application yields asymptotically sharp bounds if G has an independent set of size (close to) $-\lambda_{\min}|V(G)|/(D-\lambda_{\min})$. Indeed, Theorem 1.2 shows such an application to the Kneser graph, and there are many other graphs for which this applies. We refer to [2], for example, for a list of such graphs which include the weak product of the complete graph, line graphs of regular graphs which contain a perfect matching, Paley graphs, some strongly regular graphs, and appropriate classes of random regular graphs (see Section 5.1 of [2]).

The proof of Proposition 2.4 will be given in Section 4. It is based on a description of all independent sets in locally dense graphs. This idea can be traced back to the work of Kleitman and Winston [18], and has been exploited in various contexts since their work. Though similar proofs have been given elsewhere, none of them seems to fit into our context fully. This also applies to the powerful extension of the ideas of Kleitman and Winston to hypergraphs due to Balogh, Morris and Samotij in [6] (see also [23]), which only partially suits our needs.

2. Proofs of Theorems 1.2 and 1.3

As mentioned before, the proofs of the main theorems rely on Proposition 2.4. A central notion employed in this proposition which applies to $K(n,k)$ is as follows.

Definition. Given $\lambda \in (0,1]$, $\gamma \in (0,1]$, and a graph G on N vertices, we say that G is (λ,γ)-*supersaturated* if, for any subset $S \subseteq V(G)$ with $|S| \geqslant \lambda N$, we have

$$e(S) \geqslant \gamma \left(\frac{|S|}{N}\right)^2 e(G).$$

In addition, let $\lambda = \lambda(n) > 0$ and $\gamma = \gamma(n) > 0$. A sequence $\{G_n\}_{n\in\mathbb{N}}$ is called $(\lambda(n),\gamma(n))$-*supersaturated* if G_n is $(\lambda(n),\gamma(n))$-*supersaturated* for each $n \in \mathbb{N}$.

Hence, in a (λ,γ)-supersaturated graph G each set S of size at least λN spans many edges. Indeed, up to the multiplicative factor γ, S spans as many edges as expected from a random subset of $V(G)$ of the same size.

Using an extension of Hoffman's spectral bound [16], one can relate supersaturation to the eigenvalues of a graph. We refer to Section 3 for the proof.

Lemma 2.1. *Let G be a D-regular graph on N vertices, and let $\lambda_{\min}$ denote the smallest eigenvalue of the adjacency matrix of G. Then every set $S \subset V(G)$ satisfies*

$$e(S) \geqslant \left(\frac{\lambda_{\min}}{D}\frac{N}{|S|} + \frac{D-\lambda_{\min}}{D}\right)\left(\frac{|S|}{N}\right)^2 e(G).$$

As the eigenvalues of the Kneser graph are known due to Lovász [21], we immediately conclude the following supersaturation for the Kneser graph.

Lemma 2.2. *Let $2 \leqslant k \leqslant n/2$ and $\tau = \tau(n) > 0$. Then $K(n,k)$ is $((1+\tau)k/n, \tau/(1+\tau))$-supersaturated.*

Proof. The Kneser graph $K(n,k)$ has degree $D = \binom{n-k}{k}$, and the smallest eigenvalue of $K(n,k)$ is given by (see [21])

$$\lambda_{\min} = -\binom{n-k-1}{k-1} = -\frac{k}{n-k}D.$$

Let $S \subset \binom{[n]}{k}$ be of size at least $(1+\tau)\frac{k}{n}N$, with $N = \binom{n}{k}$. Lemma 2.1 implies that

$$e(S) \geqslant \left(-\frac{n}{(n-k)(1+\tau)} + \frac{n}{n-k}\right)\left(\frac{|S|}{N}\right)^2 e(G),$$

and the claim follows. $\square$

Beyond the notion of supersaturation needed for the proof of Theorem 1.2, we will rely on the following notion of robust stability in the proof of Theorem 1.3 (see also [22]).

Definition. Let $\lambda, \varepsilon, \delta > 0$. Let G be a graph on N vertices, and let $\mathcal{B}(G) \subseteq \mathcal{P}(V(G))$ be a family of sets. We say that G is $(\lambda, \mathcal{B}(G))$-stable with respect to (ε, δ) if, for every $S \subseteq V(G)$ with $|S| \geqslant (1-\delta)\lambda N$, we have either

- $e(S) \geqslant \delta(|S|/N)^2 \cdot e(G)$, or

- $|S \setminus B| \leqslant \varepsilon\lambda N$, for some $B \in \mathcal{B}(G)$.

In addition, let $\lambda = \lambda(n) > 0$, $\{G_n\}_{n\in\mathbb{N}}$ be a sequence of graphs, and let $\mathcal{B} = \{\mathcal{B}_n\}_{n\in\mathbb{N}}$ with $\mathcal{B}_n \subset \mathcal{P}(V(G_n))$ be a sequence of families of sets. We say that $\{G_n\}_{n\in\mathbb{N}}$ is $(\lambda,\mathcal{B})$-stable if for any $\varepsilon > 0$ there exists $\delta > 0$ and $n_0 \in \mathbb{N}$ such that, for all $n \geqslant n_0$, the graph G_n is $(\lambda(n),\mathcal{B}_n)$-stable with respect to (ε,δ).

It is instructive to think of $\mathcal{B}(G)$ as the family of largest independent sets in G, and of λN as the size of each $B \in \mathcal{B}$. The first part of the definition roughly says that if G is robustly stable, then any vertex set S whose size is close to the size of a largest independent set in G must either contain many edges, or be close to a largest independent set in structure.

The Kneser graph satisfies robust stability for k linear in n, as stated in the next lemma. It is a direct consequence of the corresponding stability result proved by Friedgut [12], and the removal lemma proved by Friedgut and Regev [13]. Again, we refer to Section 3 for the details of the proof. In the following, let $\mathcal{F}_i \subset \binom{[n]}{k}$ denote the principal hypergraph centred at i, that is, the hypergraph consisting of all k-element subsets of $[n]$ containing $i \in [n]$.

Lemma 2.3. *Let $\beta > 0$ and $k = k(n)$, where $\beta n \leqslant k \leqslant (1/2-\beta)n$, and let $G_n = K(n,k)$. Further, let $\mathcal{B}_n(G_n) = \{\mathcal{F}_i \mid i \in [n]\} \subset \mathcal{P}(V(G_n))$, and set $\mathcal{B} = \{\mathcal{B}_n\}_{n\in\mathbb{N}}$. Then $G = \{G_n\}_{n\in\mathbb{N}}$ is $(k/n,\mathcal{B})$-stable.*

With supersaturation and robust stability defined, we are now ready to state our main technical result. Given a graph H, we use $\alpha(H)$ to denote the size of the largest independent set in H. Also, for a finite set V, we let V_p be a random subset of V obtained by selecting each element $v \in V$ independently with probability p.

Proposition 2.4. *Let $\lambda = \lambda(n)$ and $\gamma = \gamma(n)$ be $(0,1)$-valued functions, and let $G = \{G_n\}_{n\in\mathbb{N}}$ be a family of graphs, where each G_n has $N = N(n)$ vertices (with $\lim_{n\to\infty} N(n) = \infty$) and average degree $D = D(n)$. For any constant $0 < \varepsilon < 1$ there exist constants $C = C(\varepsilon) > 0$ and $\delta = \delta(\varepsilon) > 0$ such that for any probability sequence $p = p(n) \in (0,1]$, the following holds. For a random spanning subgraph $H_n = G_n[V_p]$, where $V = V(G_n)$, we have the following.*

(i) *If $N^{-1} \ll p \ll D^{-1}$, then $\alpha(H_n) = (1 \pm \varepsilon)pN$ asymptotically almost surely.*

(ii) *If G is (λ,γ)-supersaturated and $9D^{-1} \leqslant p \leqslant \lambda^{\varepsilon}(\lambda\gamma D)^{-1}$, then*

$$\mathbb{P}\left(\alpha(H_n) > \frac{4N}{\varepsilon\gamma D}\ln(pD)\right) \leqslant \exp\left\{-\frac{N}{\gamma D}\ln(pD)\right\}.$$

(iii) *If G is (λ,γ)-supersaturated and $p \geqslant C(\lambda\gamma D)^{-1}\ln^2(e/\lambda)$, then*

$$\mathbb{P}(\alpha(H_n) \geqslant (1+\varepsilon)\lambda pN) \leqslant \exp\{-\varepsilon^2 p\lambda N/24\}.$$

(iv) *If G is $(\lambda,\mathcal{B})$-stable and $p \geqslant C(\lambda D)^{-1}\ln^2(e/\lambda)$, then with probability at least*

$$1 - \exp(-\delta^2\lambda pN/2),$$

the following holds: every independent set I in H_n of size at least $(1-\delta)\lambda pN$ satisfies $|I \setminus B| \leqslant \varepsilon\lambda pN$ for some $B \in \mathcal{B}_n$.

In addition, the following result will be needed for the lower bound (2) in Theorem 1.2. It is Shearer's extension [25] of a result due to Ajtai, Komlós and Szemerédi [1].

Proposition 2.5 ([1, 25]). *Let $G = \{G_n\}_{n\in\mathbb{N}}$ be a sequence of graphs on $N = N(n)$ vertices with average degree at most $D = D(n) > 1$. If each G_n is triangle-free, then G_n contains an independent set of size $N(D\ln D - D + 1)/(D-1)^2 \geqslant N(-1+\ln D)/D$.*

Finally, we shall repeatedly use Chernoff's bound for binomial random variables, which we state here for reference (see [17, Theorem 2.1]).

Lemma 2.6. *Given integers $m, s > 0$ and $\zeta \in [0,1]$, we have*

$$\mathbb{P}(\mathrm{Bin}(m,\zeta) \geqslant m\zeta + s) \leqslant e^{-s^2/(2\zeta m + s/3)}, \tag{2.1}$$

$$\mathbb{P}(\mathrm{Bin}(m,\zeta) \leqslant m\zeta - s) \leqslant e^{-s^2/(2\zeta m)}. \tag{2.2}$$

We are now ready to present the proofs of Theorems 1.2 and 1.3.

Proof of Theorem 1.2. Given $0 < \varepsilon < 1$, apply Proposition 2.4 with $\varepsilon/4$ in order to obtain a corresponding constant C_1. Let $C = \max\{32/\varepsilon^2, 32C_1/\varepsilon\}$. Further, let $k = k(n)$, and $G_n = K(n,k)$. Recall that G_n is a D-regular graph on N vertices, with $D = D(n) = \binom{n-k}{k}$ and $N = N(n) = \binom{n}{k}$. Let $H_n = G_n[V_p]$, where $V = V(G_n)$, and V_p is the set obtained by including each vertex of V independently with probability p. We apply Proposition 2.4 to $\{G_n\}_{n\in\mathbb{N}}$, with functions $N(n)$ and $D(n)$ as defined above. Bound (1) of Theorem 1.2 follows immediately from case (i) of Proposition 2.4.

For bounds (3) and (4) of Theorem 1.2, note that by Lemma 2.2 applied with $\tau = \varepsilon/2$, we know that G_n is (λ,γ)-supersaturated, with $\lambda \leqslant (1+\varepsilon/2)k/n$ and $\gamma = \varepsilon/4$. Thus we can apply Proposition 2.4 in both cases. We start with bound (3) of Theorem 1.2. Assume that $k = o(n)$, since for k linear in n this range of p is trivial. By part (ii) of Proposition 2.4 applied with $\varepsilon_1 = \varepsilon/2$, we derive that for $9D^{-1} \leqslant p \leqslant (n/k)^{1-\varepsilon/2}(\varepsilon D)^{-1}$, which contains our interval for p in case (3), we have

$$i(\mathcal{H}^k(n,p)) < \frac{8}{\varepsilon_1^2}\frac{N}{D}\ln(pD) \leqslant C\frac{N}{D}\ln(pD)$$

with probability at least $(1-\exp(-4N\ln(pD)/(\varepsilon D)))$. As $p \gg D^{-1}$, this probability tends to one as n goes to infinity, which gives the upper bound in case (3).

Next we show bound (4) of Theorem 1.2. The lower bound follows by considering the sub-hypergraph of $\mathcal{H}^k(n,p)$ consisting of all hyperedges containing, say, the element n. Using the Chernoff bound (Lemma 2.6), we have with high probability that this (intersecting) subhypergraph has size at least $(1-\varepsilon)p(k/n)N$. For the upper bound, we apply bound (iii) of Proposition 2.4 with $\varepsilon/4$ and λ, γ as chosen above. Then, by the choice of C, we have

$$i(\mathcal{H}^k(n,p)) \leqslant (1+\varepsilon)\frac{k}{n}pN \quad \text{for } p \geqslant C(n/k)D^{-1}\ln^2(n/k) \geqslant C_1(\lambda\gamma D)^{-1}\ln^2(e/\lambda),$$

and the claim follows.

Finally, we prove bound (2) of Theorem 1.2. Observe that this range of p is non-trivial only if $k \ll n$. By Chernoff's bound, almost surely H_n has at least $(1-\varepsilon/32)pN$ vertices. Further, $\mathbb{E}[e(H_n)] = p^2ND/2$, and it is not hard to see that $\mathrm{Var}[e(H_n)] \leqslant 2p^3N^2D + p^2ND$. By Chebyshev's inequality, we derive

$$\mathbb{P}\big(|e(H_n) - \mathbb{E}(e(H_n))| \geqslant \varepsilon p^2ND/32\big) \leqslant \frac{32^2\mathrm{Var}(e(H_n))}{\varepsilon^2 p^4 N^2 D^2},$$

which goes to zero by the choice of p.

Claim 2.7. *For $(n \ln n)^{1/2} \ll k \ll n$ and $p \leqslant (n/k)D^{-1}$, asymptotically almost surely the number of triangles in H_n is at most $\varepsilon pN/32$.*

Proof. The expected number of triangles in H_n is at most

$$p^3\binom{n}{k}\binom{n-k}{k}\binom{n-2k}{k}.$$

Using Markov's inequality and $p \leqslant (n/k)D^{-1}$, the claim follows if we can show that

$$(n/k)^2\binom{n-2k}{k} \ll \binom{n-k}{k}.$$

Indeed,

$$\binom{n-k}{k}\binom{n-2k}{k}^{-1} = \frac{(n-k)\cdots(n-2k+1)}{(n-2k)\cdots(n-3k+1)} \geqslant \left(\frac{n-k}{n-2k}\right)^k \geqslant \left(1+\frac{k}{n}\right)^k,$$

and using $(1+x) \geqslant \exp\{x-x^2\}$ for $0<x<1$, together with our assumption $(n\ln n)^{1/2} \ll k \ll n$, we obtain

$$\binom{n-k}{k}\binom{n-2k}{k}^{-1} \geqslant \exp\{k^2/n - k^3/n^2\} \gg n^2 \geqslant (n/k)^2,$$

which completes the proof of the claim. □

Hence, by removing at most $\varepsilon pN/32$ vertices, we obtain a triangle-free graph with at least $(1-\varepsilon/16)pN$ vertices, and no more than $(1/2+\varepsilon/32)p^2ND$ edges. Consequently, this graph has average degree at most $(1+\varepsilon/4)pD$, and due to Proposition 2.5, it contains an independent set of size

$$\frac{(1-\varepsilon/16)pN}{(1+\varepsilon/4)pD}\big(\ln((1+\varepsilon/4)pD)-1\big) \geqslant (1-\varepsilon)\frac{N}{D}\ln pD.$$

This completes the proof. □

Proof of Theorem 1.3. Let $\beta > 0$ be fixed, and $\beta n \leqslant k \leqslant (1/2-\beta)n$. Again, let G_n denote the Kneser graph $K(n,k)$. Set $\lambda = k/n$, and for a given n, let $\mathcal{B}_n$ be the set of all principal hypergraphs $\mathcal{F}_i$, for $i = 1,\ldots,n$. By Lemma 2.3, the family $G = \{G_n\}$ is $(\lambda,\mathcal{B})$-stable, where $\mathcal{B} = \{\mathcal{B}_n\}_{n\in\mathbb{N}}$. For a given $\varepsilon > 0$, we apply Proposition 2.4 in order to obtain constants C' and $\delta > 0$. Since $k = \Theta(n)$, it is possible to choose an appropriate constant C such that δ and C satisfy the conclusion of the theorem, which completes the proof. □

https://doi.org/10.1017/S0963548316000420 Published online by Cambridge University Press

3. Proofs of Lemmas 2.1 and 2.3

As mentioned before, the proof of Lemma 2.1 is a straightforward extension of Hoffman's bound [16].

Proof of Lemma 2.1. Given a D-regular G with N vertices and smallest eigenvalue $\lambda_{\min}$, we need to show that for every non-empty $S \subset V(G)$,

$$e_S = e(S) \geqslant \left(\frac{\lambda_{\min}}{D}\frac{N}{|S|} + \frac{D-\lambda_{\min}}{D}\right)\left(\frac{|S|}{N}\right)^2 e(G).$$

Let M denote the adjacency matrix of G. For $x, y \in \mathbb{R}^N$, let $\langle x, y\rangle = \sum_{i=1}^N x_i y_i$. Also, let v_S be the $0/1$-characteristic vector of S. First note that $\langle v_S, M v_S\rangle = 2e_S$. Since M is a symmetric real matrix, it is diagonalizable by an orthonormal basis. Let $u_1, \dots, u_N$ be normalized eigenvectors of M with corresponding eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_N = \lambda_{\min}$, respectively. Since G is a D-regular graph, we have $u_1 = (1/\sqrt{N}, \dots, 1/\sqrt{N})$ and $\lambda_1 = D$. Let $v_S = \sum_{i=1}^N a_i u_i$ be the expansion of v_S by eigenvectors. We have

$$2e_S = \langle v_S, M v_S\rangle = \sum_{i=1}^N \lambda_i a_i^2 \geqslant \lambda_1 a_1^2 + \lambda_{\min}\sum_{i=2}^N a_i^2.$$

Now observe that $a_1 = \langle v_S, u_1\rangle = |S|/\sqrt{N}$. In addition, $|S| = \langle v_S, v_S\rangle = \sum_{i=1}^N a_i^2$. Therefore,

$$\begin{aligned}
2e_S &\geqslant D\frac{|S|^2}{N} + \lambda_{\min}\left(|S| - \frac{|S|^2}{N}\right)\\
&= |S|\left(\lambda_{\min} + \frac{|S|}{N}(D-\lambda_{\min})\right)\\
&= \left(\frac{|S|}{N}\right)^2 2e(G)\left(\frac{\lambda_{\min}}{D}\frac{N}{|S|} + \left(1 - \frac{\lambda_{\min}}{D}\right)\right),
\end{aligned}$$

and the lemma follows. □

We now proceed to show robust stability for the Kneser graph for $k = \Omega(n)$. The proof is a direct consequence of stability due to Friedgut [12] and a removal lemma for the Kneser graph due to Friedgut and Regev [13], which we state next.

Proposition 3.1 (Friedgut [12]). *Given $\beta > 0$, let $k = k(n)$ be a sequence of integers satisfying $\beta n \leqslant k \leqslant (1/2 - \beta)n$. For all $\varepsilon > 0$ there exists $\delta > 0$ and n_0 such that, for all $n \geqslant n_0$, the following holds. If $\mathcal{F} \subseteq \binom{[n]}{k}$ is an intersecting family of size at least*

$$(1-\delta)\binom{n-1}{k-1},$$

then there is $i \in [n]$ such that

$$|\mathcal{F} \setminus \mathcal{F}_i| \leqslant \varepsilon\binom{n-1}{k-1}.$$

https://doi.org/10.1017/S0963548316000420 Published online by Cambridge University Press

Proposition 3.2 (Friedgut and Regev [13]). *Given $\beta > 0$, let $k = k(n)$ be a sequence of integers satisfying $\beta n \leqslant k \leqslant (1/2-\beta)n$. Moreover, let $N = \binom{n}{k}$ and $D = \binom{n-k}{k}$. For all $\varepsilon > 0$ there exists $\delta > 0$ and n_0 such that, for all $n \geqslant n_0$, the following holds. Every family $\mathcal{F} \subseteq \binom{[n]}{k}$ which spans at most $\delta|\mathcal{F}|^2(D/N)$ non-intersecting pairs can be made intersecting by removing at most $\varepsilon\binom{n-1}{k-1}$ elements from $\mathcal{F}$.*

Proof of Lemma 2.3. Given any $\varepsilon > 0$, first let $\varepsilon_2 = \varepsilon/2$, and apply Proposition 3.1 to get a corresponding $\delta_2 = \delta_2(\varepsilon_2) > 0$. Now set $\varepsilon_1 = \min(\varepsilon/2, \delta_2/2)$, and this time use Proposition 3.2 in order to obtain an appropriate $\delta_1 = \delta_1(\varepsilon_1) > 0$. Finally, set $\delta = \min(\delta_1, \delta_2/2) = \delta(\varepsilon) > 0$.

It follows that for any family $\mathcal{F}$ with

$$|\mathcal{F}| \geqslant (1-\delta)\binom{n-1}{k-1}$$

and

$$e(\mathcal{F}) \leqslant \delta(|\mathcal{F}|/N)^2(ND/2) \leqslant \delta_1|\mathcal{F}|^2(D/N)$$

there exists an intersecting family $\mathcal{F}' \subseteq \mathcal{F}$ obtained from $\mathcal{F}$ by removing at most $\varepsilon_1\binom{n-1}{k-1}$ of its elements such that

$$|\mathcal{F}'| \geqslant (1-\delta-\varepsilon_1)\binom{n-1}{k-1} \geqslant (1-\delta_2)\binom{n-1}{k-1}.$$

In addition, Proposition 3.1 implies that for some $i \in [n]$, we have

$$|\mathcal{F}' \setminus \mathcal{F}_i| \leqslant \varepsilon_2\binom{n-1}{k-1}.$$

Therefore,

$$|\mathcal{F} \setminus \mathcal{F}_i| \leqslant |\mathcal{F} \setminus \mathcal{F}'| + |\mathcal{F}' \setminus \mathcal{F}_i| \leqslant \varepsilon_1\binom{n-1}{k-1} + \varepsilon_2\binom{n-1}{k-1} \leqslant \varepsilon\binom{n-1}{k-1},$$

which completes the proof. □

4. Proof of Proposition 2.4

We begin with the proof of a simple structural result for independent sets in graphs (Lemma 4.1). For a given graph G, let $\mathcal{I}_G(t)$ denote the set of independent sets of G of size exactly t, and let $\mathcal{I}_G$ denote the set of all independent sets in G.

Lemma 4.1. *Let G be a graph on N vertices, and let $\gamma > 0$ be an arbitrary real number. In addition, let $0 < \ell < t$ be integers. Then, for every independent set $I \subset V(G)$ of size at least t, there is a sequence of vertices $x_1, \ldots, x_\ell \in I$ and a sequence of subsets $V(G) \supseteq X_1 \supseteq \cdots \supseteq X_\ell$ depending only on $x_1, \ldots, x_\ell$ such that:*

- *$x_1, \ldots, x_i \notin X_i$ for all $i \leqslant \ell$,*
- *$I \setminus \{x_1, \ldots, x_i\} \subset X_i$ for all $i \leqslant \ell$.*

Moreover, we have either

(i) $|X_i| \leqslant \left(1 - 2\gamma\frac{e(G)}{N^2}\right)|X_{i-1}|$ *for all $1 \leqslant i \leqslant \ell$, or*

(ii) $e(G[X_i]) < \gamma \frac{|X_i|^2}{N^2} e(G)$ *for some* $1 \leqslant i \leqslant \ell$.

Proof. Fix an independent set I of size at least t. We need to define the required sequences $x_1, \ldots, x_\ell$ and $X_1, \ldots, X_\ell$. Assume that we have already chosen elements $x_1, \ldots, x_{i-1} \in I$ and sets $V(G) = X_0 \supset X_1 \supset \cdots \supset X_{i-1}$ satisfying the conditions of our result. Observe that initially no element has been selected, and for convenience we set $X_0 = V(G)$.

Consider an ordering $(v_1, \ldots, v_{|X_{i-1}|})$ of the vertices in X_{i-1} which satisfies

$$|N(v_i) \cap \{v_{i+1}, \ldots, v_{|X_{i-1}|}\}| \geqslant |N(v_j) \cap \{v_{i+1}, \ldots, v_{|X_{i-1}|}\}|$$

for all $i < |X_{i-1}|$ and all $j > i$. Such an ordering clearly exists, since one can repeatedly choose (and remove) the vertex with highest degree in the remaining graph. In this case we say that this is a *max-ordering* of the elements in X_{i-1}.

Let j be the smallest index such that the vertex v_j in the max-ordering of X_{i-1} is contained in I. Such an index must exist, since $I \setminus \{x_1, \ldots, x_{i-1}\} \subseteq X_{i-1}$ and $i - 1 < \ell < t \leqslant |I|$. We define $x_i = v_j$, and set $S = X_{i-1} \setminus \{v_1, \ldots, v_j\}$.

If $\deg(v_j, S) < 2\gamma|S|e(G)/N^2$ then we let $X_i = S$. Note that, due to the max-ordering and the definition of v_j, every vertex $v \in X_i = \{v_{j+1}, \ldots, v_{|X_{i-1}|}\}$ satisfies $\deg(v, X_i) \leqslant \deg(v_j, X_i)$.This implies that the number of edges in X_i satisfies $e(X_i) < \gamma|X_i|^2 e(G)/N^2$. Otherwise, that is, for the case $\deg(v_j, S) \geqslant 2\gamma|S|e(G)/N^2$, we let $X_i = S \setminus N(v_j)$. Then,

$$|X_i| \leqslant |S| - \deg(v_j, S) = \left(1 - 2\gamma\frac{e(G)}{N^2}\right)|S| \leqslant \left(1 - 2\gamma\frac{e(G)}{N^2}\right)|X_{i-1}|.$$

Finally, observe that it follows from the definition of v_j that we always have $I \setminus \{x_1, \ldots, x_i\} \subset X_i$, which completes the proof. □

From this lemma we immediately deduce the following corollaries.

Corollary 4.2. *Let $G = (V, E)$ be a fixed (λ, γ)-supersaturated graph on N vertices with average degree D, where $\lambda, \gamma > 0$. Let $t \geqslant 1$, and let ℓ be an integer such that $0 < \ell < t$. Finally, set*

$$\nu = \nu(\ell) = \max\left\{\left(1 - \gamma\frac{D}{N}\right)^\ell, \lambda\right\}.$$

Then, for every independent set $I \in \mathcal{I}_G(t)$, there exists a subset $L \subset I$ of size ℓ and a set $P(L)$, depending only on L, of size at most νN such that $I \setminus L \subset P(L) \subset V(G)$. Further, we have $L \cap P(L) = \emptyset$. In particular, it follows that

$$|\mathcal{I}_G(t)| \leqslant \binom{N}{\ell}\binom{\nu N}{t - \ell}.$$

Proof. Given $I \in \mathcal{I}_G(t)$, we apply Lemma 4.1 to obtain a sequence of vertices $x_1, \ldots, x_\ell$ and sets $V = X_0, X_1, \ldots, X_\ell$, as stated. Now set $L = \{x_1, \ldots, x_\ell\}$ and $P(L) = X_\ell$, and observe that $I \setminus L \subset P(L)$ and $L \cap P(L) = \emptyset$.

If

$$|X_i| \leqslant \left(1 - 2\gamma \frac{e(G)}{N^2}\right)|X_{i-1}| \quad \text{for all } i \leqslant \ell,$$

then

$$|P(L)| \leqslant \left(1 - 2\gamma \frac{e(G)}{N^2}\right)^{\ell} N.$$

In other words,

$$|P(L)| \leqslant \left(1 - \gamma \frac{D}{N}\right)^{\ell} N.$$

On the other hand, if

$$e(X_i) < \gamma \frac{|X_i|^2}{N^2} e(G) \quad \text{for some } i \leqslant \ell,$$

then $|P(L)| \leqslant \lambda N$, since by assumption G is (λ,γ)-supersaturated. Altogether, it follows that $|P(L)| \leqslant \nu N$, which completes the proof. □

Corollary 4.3. *Let $\lambda, \varepsilon, \delta > 0$ and let $G = (V,E)$ be a graph on N vertices which is $(\lambda, \mathcal{B})$-stable with respect to (ε, δ). Let*

$$t > \ell \geqslant \ln\left(\frac{1}{(1-\delta)\lambda}\right) \frac{N^2}{2\delta e(G)}.$$

Then, for every independent set $I \in \mathcal{I}_G(t)$, there exists a subset $L \subset I$ of size ℓ and a set $P(L) \subset V(G)$ depending only on L such that $I \setminus L \subset P(L)$ and $L \cap P(L) = \emptyset$. Furthermore, either

- $|P(L)| \leqslant (1-\delta)\lambda N$, *or*
- $|P(L) \setminus B| \leqslant \varepsilon \lambda N$ *for some* $B \in \mathcal{B}$.

Proof. We apply Lemma 4.1 to G using $\gamma = \delta$ in order to obtain a sequence of vertices $x_1, \ldots, x_\ell$ and subsets $X_1, \ldots, X_\ell$ with the desired properties. Let $L = \{x_1, \ldots, x_\ell\}$. If

$$|X_i| \leqslant \left(1 - 2\delta \frac{e(G)}{N^2}\right)|X_{i-1}| \quad \text{for all } i \leqslant \ell,$$

then set $P(L) = X_\ell$. Using our assumption on ℓ, we get

$$|P(L)| \leqslant \left(1 - 2\delta \frac{e(G)}{N^2}\right)^{\ell} N \leqslant (1-\delta)\lambda N.$$

Otherwise, pick the smallest index $j \leqslant \ell$ such that

$$e(G[X_j]) < \delta \frac{|X_j|^2}{N^2} e(G),$$

and let $P(L) = X_j$. Again, if $|P(L)| \leqslant (1-\delta)\lambda N$ we are done. On the other hand, if this condition does not hold, we deduce from the $(\lambda, \mathcal{B})$-stability of G that there exists some $B \in \mathcal{B}$ for which $|P(L) \setminus B| \leqslant \varepsilon \lambda N$, which completes the proof. □

https://doi.org/10.1017/S0963548316000420 Published online by Cambridge University Press

Now that we have all the necessary machinery, we proceed with the proof of Proposition 2.4.

Proof of Proposition 2.4. Let $\lambda = \lambda(n)$, $\gamma = \gamma(n)$, and let $G = \{G_n\}_{n\in\mathbb{N}}$ be a sequence of graphs. For a given $0 < \varepsilon < 1$, let $C_1 = 800/\varepsilon^3$. For the proof of case (iv) in Proposition 2.4, suppose that G is $(\lambda,\mathcal{B})$-stable for some $\mathcal{B}$. Then for $\varepsilon' = \varepsilon/2$, there is a constant $\delta' > 0$ and $n_1 \in \mathbb{N}$ such that, for all $n \geqslant n_1$, the graph G_n is $(\lambda,\mathcal{B}_n)$-stable with respect to (ε',δ'). We choose $\delta = \min\{\delta'/4, \varepsilon'/16, 1/16\}$ and $C_2 = 100/\delta^4$. Finally, set $C = \max\{C_1, C_2\}$, and let $n_0 \geqslant n_1$ be sufficiently large.

We proceed with the proof of case (i) of Proposition 2.4. Assume that $N^{-1} \ll p \ll D^{-1}$. Using the Chernoff bound (Lemma 2.6), we have almost surely $|V_p| = (1 \pm \varepsilon/2)pN$, which proves the upper bound. Further, we have $\mathbb{E}(e(H_n)) = \frac{1}{2}NDp^2$, and by Markov's inequality a.a.s. $e(H_n) \leqslant \varepsilon pN/2$ holds. By deleting at most this number of vertices from H_n, we obtain an independent set of size at least $(1-\varepsilon)pN$, which proves the lower bound.

For part (ii), assume that $9D^{-1} \leqslant p \leqslant \lambda^{\varepsilon}(\lambda\gamma D)^{-1}$. Further, let

$$\ell = (1+\varepsilon)\frac{N}{\gamma D}\ln(pD) > 0 \quad \text{and} \quad t = \frac{4N}{\varepsilon\gamma D}\ln(pD).$$

Let X be the random variable counting the number of independent sets of size exactly t in H_n, that is, $X = |\mathcal{I}_{H_n}(t)|$. By the choice of our parameters, Corollary 4.2 applies, and we obtain

$$\mathbb{E}[X] \leqslant \binom{N}{\ell}\binom{\nu(\ell)N}{t-\ell}p^t,$$

where

$$\nu(\ell) = \max\left\{\left(1-\gamma\frac{D}{N}\right)^{\ell}, \lambda\right\}.$$

Using

$$\binom{n}{k} \leqslant \left(\frac{en}{k}\right)^k$$

and the choice of ℓ and t, we get

$$\binom{N}{\ell} \leqslant \left(\frac{e\gamma D}{\ln(pD)}\right)^{\ell} \quad \text{and} \quad \binom{\nu(\ell)N}{t-\ell} \leqslant \left(\frac{e\nu(\ell)\gamma D}{\ln(pD)}\right)^{t-\ell}.$$

Combining the two inequalities, and noting that our choice of C guarantees that $\ell \leqslant \varepsilon t/2$, we get

$$\mathbb{E}[X] \leqslant \left(\frac{e\gamma pD\nu^{1-\varepsilon/2}}{\ln(pD)}\right)^t.$$

If $\nu(\ell) = \lambda$, we have $\gamma pD\nu^{1-\varepsilon/2} \leqslant \lambda^{\varepsilon/2} \leqslant 1$, since $p \leqslant \lambda^{\varepsilon}(\lambda\gamma D)^{-1}$. On the other hand, if

$$\nu(\ell) \leqslant e^{-\ell\gamma(D/N)} \leqslant (pD)^{-1-\varepsilon},$$

we have

$$\gamma pD\nu^{1-\varepsilon/2} \leqslant \gamma(pD)^{-\varepsilon/2+\varepsilon^2/2} \leqslant \gamma \leqslant 1$$

since $\varepsilon < 1$. Hence, $\mathbb{E}(X) \leqslant (e/\ln(pD))^t$, and the claim follows from Markov's inequality.

https://doi.org/10.1017/S0963548316000420 Published online by Cambridge University Press

For part (iii), assume that $p \geqslant C(\lambda\gamma D)^{-1}\ln^2(e/\lambda)$. Let

$$t = (1+\varepsilon)p\lambda N \quad \text{and} \quad \ell = \frac{N}{\gamma D}\ln\left(\frac{e}{\lambda}\right).$$

We need to upper-bound the following probability:

$$q = \mathbb{P}[\exists I \subset V_p,\, |I| = t,\, I \text{ is an independent set in } G_n].$$

It follows from Corollary 4.3 that for any $I \in \mathcal{I}_{G_n}(t)$, there exist $L \subset I$ of size ℓ and $P(L)$ such that $I \setminus L \subset P(L) \subset V$. Therefore,

$$q \leqslant \sum_L \mathbb{P}[L \subset V_p \text{ and } |V_p \cap P(L)| \geqslant t - \ell],$$

where the sum is taken over all subsets $L \in \binom{V}{\ell}$ that correspond to some independent as given by Corollary 4.3. Using the fact that L and $P(L)$ are disjoint, we obtain

$$q \leqslant \sum_L \mathbb{P}[L \subset V_p] \cdot \mathbb{P}[|V_p \cap P(L)| \geqslant t - \ell]. \tag{4.1}$$

In addition, by our choice of ℓ, it follows that $\nu(\ell) = \lambda$. Therefore, for any such L, we have $|P(L)| \leqslant \nu(\ell)N \leqslant \lambda N$. Further, the choice of ℓ and p implies that $\ell \leqslant (\varepsilon/2)p\lambda N$. Hence with $X = |V_p \cap P(L)|$, we have due to the Chernoff bound that

$$\mathbb{P}(X \geqslant t - \ell) \leqslant \mathbb{P}\left(X \geqslant p|P(L)| + \frac{\varepsilon p\lambda N}{2}\right) \leqslant \exp\left(-\frac{\varepsilon^2 p\lambda N}{12}\right).$$

From (4.1) and

$$\binom{N}{\ell} \leqslant \left(\frac{eN}{\ell}\right)^{\ell},$$

it follows that

$$q \leqslant \left(\frac{eNp}{\ell}\right)^{\ell} \exp\left(-\frac{\varepsilon^2 p\lambda N}{12}\right) = \exp\left(\ell \cdot \ln\left(\frac{eNp}{\ell}\right) - \frac{\varepsilon^2 p\lambda N}{12}\right).$$

Recall that we want to prove that $q \leqslant \exp(-\varepsilon^2 p\lambda N/24)$. With the choice $\ell = (N/\gamma D)\ln(e/\lambda)$ it is now sufficient to show that

$$\frac{1}{\gamma D}\ln(e/\lambda)\ln\left(\frac{ep\gamma D}{\ln(e/\lambda)}\right) \leqslant \frac{\varepsilon^2 p\lambda}{24},$$

or equivalently

$$\frac{24}{\varepsilon^2\gamma\lambda D}\ln(e/\lambda) \leqslant \frac{p}{\ln(ep\gamma D/(\ln(e/\lambda)))}.$$

As the left-hand side is independent of p and the right-hand side is increasing in p, it is sufficient to show the inequality for the endpoint $p = C(\lambda\gamma D)^{-1}\ln^2(e/\lambda)$. In this case the inequality follows from

$$24/\varepsilon^2 \leqslant C\ln(e/\lambda)/\ln\left(\frac{eC}{\lambda}\ln(e/\lambda)\right).$$

Note that $\ln((eC/\lambda)\ln(e/\lambda)) > \ln(e/\lambda) + \ln C$, since $eC/\lambda > \ln(e/\lambda)$. Therefore the bound follows from $48/\varepsilon^2 \leqslant C\ln(e/\lambda)/(\ln(e/\lambda)+\ln C)$, or equivalently

$$\frac{48}{\varepsilon^2} \leqslant \frac{C}{1+\ln(C)/\ln(e/\lambda)}.$$

Since the right-hand side is decreasing in λ, it is sufficient to verify for $\lambda = 1$, which is immediate from the choice of C_1 and C.

For part (iv), let $p \geqslant C(\lambda D)^{-1}\ln^2(e/\lambda)$. Further, let

$$\mathcal{T} = \{I \in \mathcal{I}_{G_n} \colon |I| > (1-\delta)\lambda pN \text{ and } |I \setminus B| > \varepsilon\lambda pN \text{ for all } B \in \mathcal{B}_n\}.$$

Our task is to upper-bound the value of

$$q_{\mathcal{T}} = \mathbb{P}(\text{there is an independent set } I \subset V_p \text{ with } I \in \mathcal{T}).$$

Recall our choice of ε', δ', and n_0, and that G_n is $(\lambda, \mathcal{B}_n)$-stable with respect to (ε', δ') for every $n \geqslant n_0$. We apply Corollary 4.3 with ε', δ', $t = (1-4\delta)\lambda pN$, and

$$\ell = \frac{N}{\delta D}\ln\frac{e}{\lambda} \leqslant \delta\lambda pN.$$

Note that this is a valid choice of ℓ, since

$$\frac{N}{\delta D}\ln\frac{e}{\lambda} \geqslant \ln\left(\frac{1}{(1-\delta)\lambda}\right)\frac{N^2}{2\delta e(G)}.$$

This implies that for every $I \in \mathcal{T}$ there is some $L = L(I) \subset I$ of size ℓ and some $P(L) \subset V(G_n)$, depending only on L and disjoint from L, such that $I \setminus L \subset P(L)$. Hence, if there is an $I \subset V_p$ with $I \in \mathcal{T}$, then there is an L of size ℓ with

(A) $L \subset V_p$, and

(B) $|P(L) \cap V_p| \geqslant (1-\delta)p\lambda N - \ell \geqslant (1-2\delta)p\lambda N$ and $|(P(L)\setminus B) \cap V_p| > \varepsilon\lambda pN - \ell \geqslant \frac{3}{4}\varepsilon\lambda pN$ for all $B \in \mathcal{B}_n$, since $\delta \leqslant \varepsilon'/16 = \varepsilon/32$.

Let $q_{P(L)}$ be the probability that event (B) holds for the random set V_p. As L and $P(L)$ are disjoint, we have

$$q_{\mathcal{T}} \leqslant \sum_L \mathbb{P}[L \subset V_p] \cdot q_{P(L)}, \tag{4.2}$$

where the sum ranges over all $L \in \binom{V}{\ell}$ corresponding to some I as given by Corollary 4.3.

From Corollary 4.3 and the chosen parameters, either $|P(L)| \leqslant (1-\delta')\lambda N$ or $|P(L)\setminus B| \leqslant \varepsilon'\lambda N$ for some $B \in \mathcal{B}_n$. Consider each of the cases separately. If

$$|P(L)| \leqslant (1-\delta')\lambda N \leqslant (1-4\delta)\lambda N$$

then Chernoff's bound (Lemma 2.6) yields

$$\mathbb{P}(|P(L) \cap V_p| \geqslant (1-2\delta)p\lambda N) \leqslant \exp\{-\delta^2\lambda pN\}.$$

Similarly, if $|P(L)\setminus B| \leqslant \varepsilon'\lambda N = \varepsilon\lambda N/2$ for some $B \in \mathcal{B}_n$, then, together with $\delta \leqslant \varepsilon/32$, we have

$$\mathbb{P}\left(|(P(L)\setminus B) \cap V_p| > \frac{3}{4}\varepsilon\lambda pN\right) \leqslant \exp\{-\varepsilon\lambda pN/48\} \leqslant \exp\{-\delta^2\lambda pN\}.$$

Consequently, for every set L as above we have $q_{P(L)} \leqslant \exp\{-\delta^2\lambda pN\}$.

Hence (4.2) combined with

$$\binom{N}{\ell} \leqslant \left(\frac{eN}{\ell}\right)^{\ell}$$

and the choice of $\ell = (N/(\delta D))\ln(e/\lambda)$ yields

$$q_{\mathcal{T}} \leqslant \left(\frac{epN}{\ell}\right)^{\ell} \exp\{-\delta^2\lambda pN\} \leqslant \exp\left\{\ell \ln\left(\frac{e\delta pD}{\ln(e/\lambda)}\right) - \delta^2\lambda pN\right\}.$$

To complete the proof it suffices therefore to show that

$$\ell \ln\left(\frac{e\delta pD}{\ln(e/\lambda)}\right) < \delta^2\lambda pN/2,$$

or equivalently

$$\frac{2\ln(e/\lambda)}{\lambda\delta^3 D} < \frac{p}{\ln(e\delta pD/(\ln(e/\lambda)))}.$$

As the left-hand side does not depend on p, and the right-hand side is monotone increasing in p, it is sufficient to verify this inequality for the endpoint $p = C(\lambda D)^{-1}\ln^2(e/\lambda)$. In this case, and noting that $e\delta C/\lambda > \ln(e/\lambda)$ due to our choice of C_2 and C, the claim follows from

$$\frac{2}{\delta^3} < \frac{C\ln(e/\lambda)}{\ln((e\delta C/\lambda)\ln(e/\lambda))} < \frac{C\ln(e/\lambda)}{2\ln((e\delta C/\lambda))} = \frac{C}{2+2\ln(\delta C)/\ln(e/\lambda)}.$$

As the right-hand side is decreasing in λ, it is sufficient to verify for $\lambda = 1$ which, however, is immediate from the choice of C and C_2. This completes the proof. □

5. Concluding remarks

While this work has been under review, there has been a vivid interest in questions related to random versions of the Erdős–Ko–Rado theorem (see [4, 5, 7, 9, 10, 14, 15]). In particular, as well as the results of Balogh, Bohman and Mubayi [3], the question concerning the structure of the largest intersecting family in the random setting has been addressed in [5, 14, 15] for various ranges of k and p. Moreover, an extension of the robust stability result for intersecting families, Lemma 2.3, has been considered in [9], implying that Theorem 1.3 can be extended to a larger range of k. We refer to these papers for further information.

Acknowledgement

The authors are grateful to Yoshiharu Kohayakawa for mentioning the problem and for several helpful comments. Many thanks go to Oded Regev and Ehud Friedgut for sharing a preliminary version of their paper. Finally, we thank the anonymous referee for carefully reading the paper and for many suggestions that improved its readability.

References

[1] Ajtai, M., Komlós, J. and Szemerédi, E. (1980) A note on Ramsey numbers, *J. Combin. Theory Ser. A* **29** 354–360.

[2] Alon, N., Dinur, I., Friedgut, E. and Sudakov, B. (2004) Graph products, Fourier analysis and spectral techniques. *Geom. Funct. Anal.* **14** 913–940.

[3] Balogh, J., Bohman, T. and Mubayi, D. (2009) Erdős–Ko–Rado in random hypergraphs. *Combin. Probab. Comput.* **18** 629–646.

[4] Balogh, J., Bollobás, B. and Narayanan B. Transference for the Erdős–Ko–Rado theorem. *Forum of Math. Sigma*, to appear https://doi.org/10.1017/fms.2015.21.

[5] Balogh, J., Das, S., Delcourt, M., Liu, H. and Sharifzadeh, M. (2015) Intersecting families of discrete structures are typically trivial. *J. Combin. Theory Ser. A* **132** 224–245.

[6] Balogh, J., Morris, R. and Samotij, W. (2015) Independent sets in hypergraphs. *J. Amer. Math. Soc.* **28** 669–709.

[7] Bollobás, B., Narayanan, B. and Raigorodskii, A. On the stability of the Erdős–Ko–Rado theorem. *J. Combin. Theory Ser. A*, to appear http://dx.doi.org/10.1016/j.jcta.2015.08.002.

[8] Conlon, D. and Gowers, W. Combinatorial theorems in sparse random sets. Submitted DOI 10.4007/annals.2016.184.2.2.

[9] Das, S. and Tran, T. Removal and stability for Erdős–Ko–Rado. *SIAM J. Discrete Math.*, to appear DOI:10.1137/15M105149X.

[10] Devlin, P. and Kahn, J. On stability in the Erdős–Ko–Rado theorem. Submitted DOI:10.1137/15M1012992.

[11] Erdős, P., Ko, C. and Rado, R. (1961) Intersection theorems for systems of finite sets. *Quart. J. Math.* **12** 313–320.

[12] Friedgut, E. (2008) On the measure of intersecting families, uniqueness and stability. *Combinatorica* **28** 503–528.

[13] Friedgut, E. and Regev, O. Manuscript.

[14] Hamm, A. and Kahn, J. On Erdős–Ko–Rado for random hypergraphs I. Submitted.

[15] Hamm, A. and Kahn, J. On Erdős–Ko–Rado for random hypergraphs II. Submitted.

[16] Hoffman, A. (1970) On eigenvalues and colorings of graphs. In *Graph Theory and its Applications (Proc. Advanced Sem., Math. Research Center, Univ. of Wisconsin)*, pp. 79–91.

[17] Janson, S., Łuczak, T. and Ruciński, A. (2000) *Random Graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience.

[18] Kleitman, D. and Winston, K. (1982) On the number of graphs without 4-cycles. *Discrete Math.* **41** 167–172.

[19] Kohayakawa, Y., Kreuter, B. and Steger, A. (1998) An extremal problem for random graphs and the number of graphs with large even-girth. *Combinatorica* **18** 101–120.

[20] Kohayakawa, Y., Lee, S. J., Rödl, V. and Samotij, W. (2015) The number of Sidon sets and the maximum size of Sidon sets contained in a sparse random set of integers. *Random Struct. Alg.* **46** 1–25.

[21] Lovász, L. (1979) On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory* **25** 1–7.

[22] Samotij, W. (2014) Stability results for random discrete structures. *Random Struct. Alg.* **44** 269–289.

[23] Saxton, D. and Thomason, A. (2015) Hypergraph containers. *Invent. Math.* **201** 925–992.

[24] Schacht, M. Extremal results for random discrete structures. Submitted DOI 10.4007/annals.2016.184.2.1.

[25] Shearer, J. (1983) A note on the independence number of triangle-free graphs. *Discrete Math.* **46** 83–87.