Transference for the Erdős–Ko–Rado theorem

József Balogh, Béla Bollobás, and Bhargav Narayanan

Abstract. For natural numbers $n, r \in \mathbb{N}$ with $n \geq r$, the Kneser graph $K(n,r)$ is the graph on the family of r-element subsets of $\{1,\dots,n\}$ in which two sets are adjacent if and only if they are disjoint. Delete the edges of $K(n,r)$ with some probability, independently of each other: is the independence number of this random graph equal to the independence number of the Kneser graph itself? We answer this question affirmatively as long as r/n is bounded away from $1/2$, even when the probability of retaining an edge of the Kneser graph is quite small. This gives us a random analogue of the Erdős–Ko–Rado theorem since an independent set in the Kneser graph is the same as a uniform intersecting family. To prove our main result, we give some new estimates for the number of disjoint pairs in a family in terms of its distance from an intersecting family; these might be of independent interest.

1. Introduction

Over the past twenty years, a great deal of work has gone into proving 'sparse random' analogues of classical extremal results in combinatorics. Some of the early highlights include a version of Mantel's theorem for random graphs proved by Babai, Simonovits, and Spencer [1], the Ramsey theoretic results of Rödl and Ruciński [21, 22], and a random analogue of Szemerédi's theorem due to Kohayakawa, Łuczak and Rödl [16]. Very general transference theorems have since been proved by Conlon and Gowers [7], Schacht [25], Balogh, Morris and Samotij [3] and Saxton and Thomason [24]. The surveys of Łuczak [20] and Rödl and Schacht [23] provide a detailed account of such results.

In this paper, we shall be interested in proving such a transference result for a central result in extremal set theory, the Erdős–Ko–Rado theorem. A family of sets $\mathcal{A}$ is said to be *intersecting* if $A \cap B \neq \varnothing$ for all $A, B \in \mathcal{A}$. Writing $X^{(r)}$ for the family of all r-element subsets of a set X and $[n]$ for the set $\{1, 2, ..., n\}$, a classical result of Erdős, Ko and Rado [9] asserts that if $n > 2r$ and $\mathcal{A} \subset [n]^{(r)}$ is intersecting, then $|\mathcal{A}| \leq \binom{n-1}{r-1}$ with equality if and only if $\mathcal{A}$ is a *star*. As is customary, we define the *star centred at* $x \in [n]$ to be the family of all the r-element subsets of $[n]$ containing x and we call an intersecting family *trivial* if it is contained in a star.

Date: 12 October 2014.

2010 *Mathematics Subject Classification.* Primary 05D05; Secondary 05C80, 05D40.

If $\mathcal{A} \subset [n]^{(r)}$ is intersecting and has cardinality comparable to that of a star, must $\mathcal{A}$ necessarily resemble a star? Such questions about the 'stability' of the Erdős–Ko–Rado theorem have received a great deal of attention. Perhaps the earliest stability result about the Erdős–Ko–Rado theorem was proved by Hilton and Milner [11] who determined how large a uniform intersecting family can be if one insists that the family is nontrivial. Furthering this line of research, Friedgut [10], Dinur and Friedgut [8], and Keevash and Mubayi [14] have shown that every 'large' uniform intersecting family is essentially trivial. Finally, let us mention that Balogh, Das, Delcourt, Liu and Sharifzadeh [2] have recently shown, amongst other things, that almost all r-uniform intersecting families are trivial when $r < (n - 8\log n)/3$.

As stated earlier, our aim in this note to prove a transference result for the Erdős–Ko–Rado theorem. The notion of stability we shall consider here was introduced by Bollobás, Narayanan and Raigorodskii [6] (see also [4]). To present this notion of stability, it will be helpful to consider $[n]^{(r)}$ in a different incarnation, as the *Kneser graph* $K(n,r)$. The Kneser graph $K(n,r)$ is the graph on $[n]^{(r)}$ where two vertices, i.e., r-element subsets of $[n]$, are adjacent if and only if they are disjoint. We shall freely switch between these two incarnations of $[n]^{(r)}$.

Observe that a family $\mathcal{A} \subset [n]^{(r)}$ is intersecting if and only if $\mathcal{A}$ induces an independent set in $K(n,r)$. Writing $\alpha(G)$ for the size of the largest independent set of a graph G, the Erdős–Ko–Rado theorem asserts that $\alpha(K(n,r)) = \binom{n-1}{r-1}$ when $n > 2r$. Let $K_p(n,r)$ denote the random subgraph of $K(n,r)$ obtained by retaining each edge of $K(n,r)$ independently with probability p. Bollobás, Narayanan and Raigorodskii [6] asked the following natural question: is $\alpha(K_p(n,r)) = \binom{n-1}{r-1}$? They proved, when $r = r(n) = o(n^{1/3})$, that the answer to this question is in the affirmative even after practically all the edges of the Kneser graph have been deleted. More precisely, they showed that in this range, there exists a (very small) critical probability $p_c(n,r)$ with the following property: as $n \to \infty$, if $p/p_c > 1$, then with high probability, $\alpha(K_p(n,r)) = \binom{n-1}{r-1}$ and the only independent sets of this size in $K_p(n,r)$ are stars, while if $p/p_c < 1$, then $\alpha(K_p(n,r)) > \binom{n-1}{r-1}$ with high probability.

Bollobás, Narayanan and Raigorodskii also asked what happens for larger values of r, and conjectured in particular that as long as r/n is bounded away from $1/2$, such a random analogue of the Erdős–Ko–Rado theorem should continue to hold for $K_p(n,r)$ for some p bounded away from 1. In this note, we shall prove this conjecture and a bit more.

Theorem 1.1. *For every $\varepsilon > 0$, there exist constants $c = c(\varepsilon) > 0$ and $c' = c'(\varepsilon) > 0$ with $c < c'$ such that for all $n, r \in \mathbb{N}$ with $r \leq (1/2 - \varepsilon)n$,*

$$\mathbb{P}\left(\alpha(K_p(n,r)) = \binom{n-1}{r-1}\right) \to \begin{cases} 1 & \text{if } p \geq \binom{n-1}{r-1}^{-c}, \text{ and} \\ 0 & \text{if } p \leq \binom{n-1}{r-1}^{-c'} \end{cases}$$

as $n \to \infty$. *In particular, with high probability,* $\alpha(K_{1/2}(n,r)) = \binom{n-1}{r-1}$.

All the work in proving Theorem 1.1 is in showing that $c(\varepsilon)$ exists; as we shall see, the existence of $c'(\varepsilon)$ follows from a simple second moment calculation.

Let us briefly describe some of the ideas that go into the proof of Theorem 1.1. We shall prove two results which, taken together, show that a large family $\mathcal{A} \subset [n]^{(r)}$ without a large intersecting subfamily must necessarily contain many pairs of disjoint sets, or in other words, must induce many edges in $K(n,r)$; we do this in Section 3. We put together the pieces and give the proof of Theorem 1.1 in Section 4. In Section 5, we briefly describe some approaches to improving the dependence of $c(\varepsilon)$ on ε in Theorem 1.1. We conclude with some discussion in Section 6.

2. Preliminaries

Henceforth, a 'family' will be a uniform family on $[n]$ unless we specify otherwise. To ease the notational burden, we adopt the following notational convention: when n and r are clear from the context, we write $\mathbf{V} = \binom{n}{r}$, $\mathbf{N} = \binom{n-1}{r-1}$, $\mathbf{M} = \binom{n-r-1}{r-1}$ and $\mathbf{R} = \binom{2r}{r}$.

We need a few results from extremal set theory, some classical and some more recent. The first result that we will need, due to Hilton and Milner [11], bounds the cardinality of a nontrivial uniform intersecting family. Writing $\mathcal{A}_x$ for the subfamily of a family $\mathcal{A}$ that consists of those sets containing x, we have the following.

Theorem 2.1. *Let* $n, r \in \mathbb{N}$ *and suppose that* $n > 2r$. *If* $\mathcal{A} \subset [n]^{(r)}$ *is an intersecting family with* $|\mathcal{A}| \geq \mathbf{N} - \mathbf{M} + 2$, *then there exists an* $x \in [n]$ *such that* $\mathcal{A} = \mathcal{A}_x$. □

The next result we shall require, due to Friedgut [10], is a quantitative extension of the Hilton–Milner theorem which says that any sufficiently large uniform intersecting family must resemble a star.

Theorem 2.2. *For every* $\varepsilon > 0$, *there exists a* $C = C(\varepsilon) > 0$ *such that for all* $n, r \in \mathbb{N}$ *with* $\varepsilon n \leq r \leq (1/2 - \varepsilon)n$, *the following holds: if* $\mathcal{A} \subset [n]^{(r)}$ *is an intersecting family and* $|\mathcal{A}| = \mathbf{N} - k$, *then there exists an* $x \in [n]$ *for which* $|\mathcal{A}_x| \geq \mathbf{N} - Ck$. □

We will also need the following well-known inequality for cross-intersecting families due to the second author [5].

Theorem 2.3. *Let* $(A_1, B_1), \dots, (A_m, B_m)$ *be pairs of disjoint* r*-element sets such that* $A_i \cap B_j \neq \varnothing$ *for* $i, j \in [m]$ *whenever* $i \neq j$. *Then* $m \leq \mathbf{R}$. □

Finally, we shall require a theorem of Kruskal [17] and Katona [12]. For a family $\mathcal{A} \subset [n]^{(r)}$, its shadow in $[n]^{(k)}$, denoted $\partial^{(k)}\mathcal{A}$, is the family of those k-sets contained in some member of $\mathcal{A}$. For $x \in \mathbb{R}$ and $r \in \mathbb{N}$, we define the *generalised binomial coefficient*

$\binom{x}{r}$ by setting

$$\binom{x}{r} = \frac{x(x-1)\dots(x-r+1)}{r!}.$$

The following formulation of the Kruskal–Katona theorem is due to Lovász [19].

Theorem 2.4. *Let $n, r, k \in \mathbb{N}$ and suppose that $k \le r \le n$. If the cardinality of $\mathcal{A} \subset [n]^{(r)}$ is $\binom{x}{r}$ for some real number $x \ge r$, then $|\partial^{(k)}\mathcal{A}| \ge \binom{x}{k}$.* □

To avoid clutter, we omit floors and ceilings when they are not crucial. We use the standard $o(1)$ notation to denote any function that tends to zero as n tends to infinity; the variable tending to infinity will always be n unless we explicitly specify otherwise.

3. The number of disjoint pairs

Given a family $\mathcal{A}$, we write $e(\mathcal{A})$ for the number of disjoint pairs of sets in $\mathcal{A}$; equivalently, $e(\mathcal{A})$ is the number of edges in the subgraph of the Kneser graph induced by $\mathcal{A}$. In this section, we give some bounds for $e(\mathcal{A})$.

We denote by $\mathcal{A}^*$ the largest intersecting subfamily of a family $\mathcal{A}$; if this subfamily is not unique, we take any subfamily of maximum cardinality. We write $\ell(\mathcal{A}) = |\mathcal{A}| - |\mathcal{A}^*|$ for the difference between the cardinality of $\mathcal{A}$ and the largest intersecting subfamily of $\mathcal{A}$.

Trivially, we have $e(\mathcal{A}) \ge \ell(\mathcal{A})$. Our first lemma says that we can do much better than this trivial bound when $\ell(\mathcal{A})$ is large.

Lemma 3.1. *Let $n, r \in \mathbb{N}$. For any $\mathcal{A} \subset [n]^{(r)}$,*

$$e(\mathcal{A}) \ge \frac{\ell(\mathcal{A})^2}{2\mathbf{R}}.$$

Proof. To prove this lemma, we need the notion of an induced matching. An *induced matching of size m* in a graph G is a set of $2m$ vertices inducing a subgraph consisting of m independent edges; equivalently, we refer to these m edges as an induced matching of size m. The *induced-matching number* of G, in notation, $m(G)$, is the maximal size of an induced matching in G.

Proposition 3.2. *Let $G = (V, E)$ be a graph with $m(G) = m \ge 1$. Then*

$$|E| \ge \frac{k^2}{4m},$$

where $k = |V| - \alpha(G)$.

Proof. Let us choose $X = \{x_1, \dots, x_m\}$ and $Y = \{y_1, \dots, y_m\}$ so that the edges $x_1y_1, \dots, x_my_m$ constitute an induced matching. Let $Z = \Gamma(X \cup Y)$ be the set of neighbours of the vertices in $X \cup Y$; thus $X \cup Y \subset Z$. Since $m(G) = m$, the set

$V(G) \setminus Z$ is independent, so $|Z| \geq k$. Since some vertex in $X \cup Y$ has at least $|Z|/2m$ neighbours, we conclude that $\Delta(G) \geq |Z|/2m \geq k/2m$ where $\Delta(G)$ is the maximum degree of G.

Now define a sequence of graphs $G = G_0 \supset G_1 \supset \cdots \supset G_k$ and a sequence of vertices $x_0, x_1, \ldots, x_k$ by taking x_i to be a vertex of G_i of maximal degree and G_{i+1} to be the graph obtained from G_i by deleting x_i. We know from our earlier arguments that $\Delta(G_i) \geq (k-i)/2m$, so $|E| \geq \sum_{i=0}^{k} \Delta(G_i) \geq k^2/4m$. □

To apply the previous proposition, we need the following corollary of Theorem 2.3 the proof of which is implicit in [2]; we include the short proof here for completeness.

Proposition 3.3. *For $n \geq 2r$, the induced-matching number of $K(n,r)$ is*

$$m(K(n,r)) = \binom{2r-1}{r-1} = \frac{\mathbf{R}}{2}.$$

Proof. Let $A_1B_1, \ldots, A_mB_m$ be an induced matching in $K(n,r)$. For $m+1 \leq i \leq 2m$, we set $A_i = B_{i-m}$ and $B_i = A_{i-m}$. We apply Theorem 2.3 to $(A_1, B_1), \ldots, (A_{2m}, B_{2m})$ and conclude that $2m \leq \mathbf{R}$. The $\mathbf{R}/2$ partitions of $[2r]$ into disjoint r-sets form an induced matching, so $m(K(n,r)) = \mathbf{R}/2$, as claimed. □

The lemma follows by applying Proposition 3.2 to $G_{\mathcal{A}}$, the subgraph of the Kneser graph $K(n,r)$ induced by $\mathcal{A}$. □

Note that Lemma 3.1 is only effective when $\ell(\mathcal{A}) \geq 2\mathbf{R}$. The next, somewhat technical, lemma complements Lemma 3.1 by giving a better bound when $\ell(\mathcal{A})$ is small provided the size of $\mathcal{A}$ is large.

Lemma 3.4. *For every $\varepsilon, \eta > 0$, there exist constants $\delta = \delta(\varepsilon, \eta) > 0$ and $C = C(\varepsilon) > 0$ with the following property: for all $n, r \in \mathbb{N}$ with $\varepsilon n \leq r \leq (1/2 - \varepsilon)n$, we have*

$$e(\mathcal{A}) \geq \ell(\mathcal{A})^{1+\delta} - C\ell(\mathcal{A})$$

for any family $\mathcal{A} \subset [n]^{(r)}$ with $|\mathcal{A}| = \mathbf{N}$ and $\ell(\mathcal{A}) \leq \mathbf{N}^{1-\eta}$.

To clarify, the $C(\varepsilon)$ in the statement of the lemma above is the same as the $C(\varepsilon)$ guaranteed by Theorem 2.2.

Proof of Lemma 3.4. First, let us note that since we always have $e(\mathcal{A}) \geq \ell(\mathcal{A})$, it suffices to prove the lemma under the assumption that n is sufficiently large.

Let $\ell = \ell(\mathcal{A})$. We start by observing that most of $\mathcal{A}$ must be contained in a star. Indeed, as before, let $\mathcal{A}^*$ denote the largest intersecting subfamily of $\mathcal{A}$; by definition, $|\mathcal{A}^*| = \mathbf{N} - \ell$. Since we have assumed that $\varepsilon n \leq r \leq (1/2 - \varepsilon)n$, we may assume, by Theorem 2.2, that $|\mathcal{A}_n^*| \geq \mathbf{N} - C\ell$, where $C = C(\varepsilon)$ is as guaranteed by Theorem 2.2. Hence, $|\mathcal{A}_n| \geq |\mathcal{A}_n^*| \geq \mathbf{N} - C\ell$.

We also know that $|\mathcal{A}_n| \le |\mathcal{A}^*| \le \mathbf{N} - \ell$; let $\mathcal{B}$ be a subset of $\mathcal{A} \setminus \mathcal{A}_n$ of cardinality exactly ℓ. We shall bound $e(\mathcal{A})$ by counting the number of edges between $\mathcal{B}$ and $\mathcal{A}_n$ in $K(n,r)$.

Let us define

$$\mathcal{A}' = \{A \setminus \{n\} : A \in \mathcal{A}_n\} \subset [n-1]^{(r-1)}$$

and

$$\mathcal{B}' = \{[n-1] \setminus B : B \in \mathcal{B}\} \subset [n-1]^{(n-r-1)}.$$

Clearly, to count the number of edges between $\mathcal{A}_n$ and $\mathcal{B}$ in $K(n,r)$, it suffices to count the number of pairs (A', B') in $\mathcal{A}' \times \mathcal{B}'$ with $A' \subset B'$. This quantity is obviously bounded below by the number of sets $A' \in \mathcal{A}'$ contained in at least one $B' \in \mathcal{B}'$.

Since $\mathcal{A}' \subset [n-1]^{(r-1)}$ and $|\mathcal{A}'| \ge \mathbf{N} - C\ell$, the number of sets $A' \in \mathcal{A}'$ contained in some $B' \in \mathcal{B}'$ is at least $|\partial^{(r-1)}\mathcal{B}'| - C\ell$. Consequently,

$$e(\mathcal{A}) \ge |\partial^{(r-1)}\mathcal{B}'| - C\ell.$$

We shall show that there exists a $\delta = \delta(\varepsilon, \eta) > 0$ such that, under the conditions of the lemma, $|\partial^{(r-1)}\mathcal{B}'| \ge \ell^{1+\delta}$ for all sufficiently large $n \in \mathbb{N}$. We deduce the existence of such a δ from Theorem 2.4, the Kruskal–Katona theorem. We may assume that

$$\ell = |\mathcal{B}'| = \binom{x}{n-r-1}$$

for some real number $x \ge n - r - 1$. It follows from Theorem 2.4 that

$$|\partial^{(r-1)}\mathcal{B}'| \ge \binom{x}{r-1}.$$

Let us put $r = (1/2 - \beta)n$ and $x = \vartheta n$. We now calculate, ignoring error terms that are $o(1)$, what values β and ϑ can take. We know that $\varepsilon \le \beta \le 1/2 - \varepsilon$. Since $x \ge n - r - 1$, we also know that $\vartheta \ge 1/2 + \beta$. On the other hand, since

$$\binom{\vartheta n}{(1/2+\beta)n} = \ell \le \mathbf{N}^{1-\eta} = \binom{n-1}{r-1}^{1-\eta} \le \binom{n}{r}^{1-\eta} = \binom{n}{(1/2-\beta)n}^{1-\eta},$$

it follows from Stirling's approximation for the factorial function that there exists some $\delta'(\varepsilon, \eta) > 0$ such that $\vartheta \le 1 - \delta'$.

Hence, it suffices to check that there exists a $\delta = \delta(\varepsilon, \eta) > 0$ for which the inequality

$$\binom{\vartheta n}{(1/2-\beta)n} \ge \binom{\vartheta n}{(1/2+\beta)n}^{1+\delta}$$

holds for all $\beta \in [\varepsilon, 1/2 - \varepsilon]$ and $\vartheta \in [1/2 + \beta, 1 - \delta']$ as long as n is sufficiently large. This is easily checked using Stirling's formula. □

4. Proof of the main result

Armed with Lemmas 3.1 and 3.4, we are now ready to prove Theorem 1.1.

Proof of Theorem 1.1. Let us fix $\varepsilon > 0$ and assume that $r \leq (1/2 - \varepsilon)n$. Clearly, it is enough to prove Theorem 1.1 for all sufficiently small ε; it will be convenient to assume that $\varepsilon < 1/10$. As mentioned earlier, Bollobás, Narayanan and Raigorodskii have proved Theorem 1.1 in a much stronger form when $r = o(n^{1/3})$. So to avoid having to distinguish too many cases, we shall assume that r grows with n; for concreteness, let us suppose that $r \geq n^{1/4}$. A consequence of these assumptions is that in this range, $\mathbf{V}$, $\mathbf{N}$ and $\mathbf{M}$ all grow much faster than any polynomial in n.

First, let Y denote the (random) number of independent sets $\mathcal{A} \subset [n]^{(r)}$ in $K_p(n,r)$ with $|\mathcal{A}| = \mathbf{N}+1$ and $\ell(\mathcal{A}) = 1$; in other words, independent sets of size $\mathbf{N}+1$ containing an entire star. We begin by showing that there exists a $c' = c'(\varepsilon)$ such that if $p \leq \mathbf{N}^{-c'}$, then $Y > 0$ with high probability. Clearly,

$$\mathbb{E}[Y] = \binom{n}{1}\binom{\mathbf{V}-\mathbf{N}}{1}(1-p)^{\mathbf{M}}.$$

Note that if $r \leq (1/2 - \varepsilon)n$, then we may choose a suitably small $c' = c'(\varepsilon)$ such that $\mathbf{M} \geq \mathbf{N}^{c'}$. It follows that if c' is sufficiently small, then

$$\mathbb{E}[Y] \geq n(\mathbf{V}-\mathbf{N})\exp\left(-(p+p^2)\mathbf{M}\right) \geq (e+o(1))n(\mathbf{V}-\mathbf{N}),$$

so $\mathbb{E}[Y] \to \infty$ when $p \leq \mathbf{N}^{-c'}$.

Therefore, to show that $Y > 0$ with high probability, it suffices to show that $\mathrm{Var}[Y] = o(\mathbb{E}[Y]^2)$ or equivalently, that $\mathbb{E}[(Y)_2] = (1+o(1))\mathbb{E}[Y]^2$, where $\mathbb{E}[(Y)_2] = \mathbb{E}[Y(Y-1)]$ is the second factorial moment of Y.

Writing $\mathcal{S}_x$ for the star centred at x, we note that

$$\mathbb{E}[(Y)_2] = \sum_{a,b,A,B} \mathbb{P}(\mathcal{S}_a \cup \{A\} \text{ and } \mathcal{S}_b \cup \{B\} \text{ are independent}),$$

the sum being over ordered 4-tuples (a,b,A,B) with $a,b \in [n]$, $A \in [n]^{(r)} \setminus \mathcal{S}_a$ and $B \in [n]^{(r)} \setminus \mathcal{S}_b$ such that $(a,A) \neq (b,B)$. Now, observe that

$$\begin{aligned}\sum_{a\neq b} \mathbb{P}(\mathcal{S}_a \cup \{A\} \text{ and } \mathcal{S}_b \cup \{B\} \text{ are independent}) &\leq (n^2)(\mathbf{V}-\mathbf{N})^2(1-p)^{(2-o(1))\mathbf{M}} \\ &= (1+o(1))\mathbb{E}[Y]^2,\end{aligned}$$

and

$$\begin{aligned}\sum_{a=b,A\neq B} \mathbb{P}(\mathcal{S}_a \cup \{A\} \text{ and } \mathcal{S}_b \cup \{B\} \text{ are independent}) &\leq n(\mathbf{V}-\mathbf{N})^2(1-p)^{2\mathbf{M}} \\ &= o(\mathbb{E}[Y]^2).\end{aligned}$$

By Chebyshev's inequality, we conclude that $Y > 0$ with high probability, so the independence number of $K_p(n,r)$ is at least $\mathbf{N}+1$ with high probability if $p \le \mathbf{N}^{-c'}$.

Next, for each $\ell \ge 1$, let X_ℓ denote the (random) number of independent sets $\mathcal{A} \subset [n]^{(r)}$ in $K_p(n,r)$ with $|\mathcal{A}| = \mathbf{N}$ and $\ell(\mathcal{A}) = \ell$. To complete the proof of Theorem 1.1, it clearly suffices to show that for some $c = c(\varepsilon) > 0$, all of the X_ℓ are zero with high probability provided $p \ge \mathbf{N}^{-c}$. We shall prove this by distinguishing three cases depending on which of Theorem 2.1, Lemma 3.1 and Lemma 3.4 is to be used.

Let $C = C(\varepsilon)$ be as in Theorem 2.2. Note that since $r \le (1/2 - \varepsilon)n$, it is easy to check using Stirling's approximation that we can choose positive constants $c_m = c_m(\varepsilon)$ and $c_r = c_r(\varepsilon)$ such that $\mathbf{M} \ge \mathbf{N}^{c_m}$ and $\mathbf{R} \le \mathbf{N}^{1-c_r}$.

We now set $L_m = \mathbf{N}^{c_m}/2$ and $L_r = \mathbf{N}^{1-c_r/4}$ and distinguish the following three cases.

Case 1: $\ell \le L_m$ **.** Let $\mathcal{A} \subset [n]^{(r)}$ be a family of cardinality $\mathbf{N}$ with $\ell(\mathcal{A}) = \ell$. Since

$$\ell \le L_m = \mathbf{N}^{c_m}/2 \le \mathbf{M} - 2,$$

we see that $\mathcal{A}^*$, the largest intersecting subfamily of $\mathcal{A}$, satisfies

$$|\mathcal{A}^*| = \mathbf{N} - \ell \ge \mathbf{N} - \mathbf{M} + 2.$$

It follows from Theorem 2.1 that there is an $x \in [n]$ for which $\mathcal{A}^*$ is contained in the star centred at x. Consider the ℓ sets in $\mathcal{A} \setminus \mathcal{A}^*$. Any such set is disjoint from exactly $\mathbf{M}$ members of the star centred at x and hence from at least $\mathbf{M} - \ell$ members of $\mathcal{A}^*$. This tells us that $e(\mathcal{A}) \ge \ell(\mathbf{M} - \ell)$. Since $\ell \le \mathbf{M}/2$, it follows that

$$\begin{aligned}
\mathbb{E}[X_\ell] &\le n\binom{\mathbf{N}}{\ell}\binom{\mathbf{V}}{\ell}(1-p)^{\ell(\mathbf{M}-\ell)} \\
&\le n\binom{2^n}{\ell}^2 \exp(-p\ell\mathbf{M}/2) \\
&\le \exp(2n\ell - p\ell\mathbf{M}/2).
\end{aligned}$$

Hence, if $c \le c_m/2$ so that $p \ge \mathbf{N}^{-c_m/2}$, it is clear that

$$\sum_{\ell=1}^{L_m} \mathbb{E}[X_\ell] \le \sum_{\ell=1}^{L_m} \exp\left(2n\ell - \frac{\ell\mathbf{N}^{c_m/2}}{2}\right) = o(1).$$

So with high probability, for each $1 \le \ell \le L_m$, the random variable X_ℓ is zero.

Case 2: $\ell \ge L_r$**.** Again, let $\mathcal{A} \subset [n]^{(r)}$ be a family of cardinality $\mathbf{N}$ with $\ell(\mathcal{A}) = \ell$. We know from Lemma 3.1 that

$$e(\mathcal{A}) \ge \frac{\ell^2}{2\mathbf{R}} \ge \frac{\mathbf{N}^{2-c_r/2}}{2\mathbf{N}^{1-c_r}} = \frac{\mathbf{N}^{1+c_r/2}}{2}.$$

So it follows that

$$\sum_{l \ge L_r} \mathbb{E}[X_\ell] \le \binom{\mathbf{V}}{\mathbf{N}} \exp\left(-p\frac{\mathbf{N}^{1+c_r/2}}{2}\right) \le \exp\left(n\mathbf{N} - p\frac{\mathbf{N}^{1+c_r/2}}{2}\right).$$

Hence, if $c \leq c_r/4$ so that $p \geq \mathbf{N}^{-c_r/4}$, we have

$$\sum_{l \geq L_r} \mathbb{E}[X_\ell] \leq \exp\left(n\mathbf{N} - \frac{\mathbf{N}^{1+c_r/4}}{2}\right) = o(1).$$

So once again, with high probability, the sum $\sum_{\ell \geq L_r} X_\ell$ is zero.

Before we proceed further, let us first show that that we may now assume without loss of generality that $r \geq \varepsilon n$. This is because one can check that the arguments in Cases 1 and 2 together prove Theorem 1.1 when $r \leq \varepsilon n$ for all sufficiently small ε. It is easy to check using Stirling's formula that if ε is sufficiently small, indeed, if $\varepsilon < 1/10$ for example, then it is possible to choose positive constants $c'_m(\varepsilon)$ and $c'_r(\varepsilon)$ so that for all $r \leq \varepsilon n$, we have $\mathbf{M} \geq \mathbf{N}^{c'_m}$, $\mathbf{R} \leq \mathbf{N}^{1-c'_r}$ and $\mathbf{N}^{c'_m}/2 \geq \mathbf{N}^{1-c'_r/4}$. So the arguments above yield a proof of Theorem 1.1 when $r \leq \varepsilon n$. Therefore, in the following, we assume that $r \geq \varepsilon n$.

Case 3: $L_m \leq \ell \leq L_r$**.** As before, consider any family $\mathcal{A} \subset [n]^{(r)}$ of cardinality $\mathbf{N}$ with $\ell(\mathcal{A}) = \ell$. First note that since $\varepsilon n \leq r \leq (1/2 - \varepsilon n)$ and $\ell \leq L_r = \mathbf{N}^{1-c_r/4}$ where c_r is a constant depending only on ε, by Lemma 3.4, there exists a $\delta = \delta(\varepsilon)$ such that

$$e(\mathcal{A}) \geq \ell^{1+\delta} - C\ell.$$

Since $\ell \geq L_m = \mathbf{N}^{c_m}/2$, it follows that

$$e(\mathcal{A}) \geq \ell^{1+\delta} - C\ell \geq \ell^{1+\delta/2}$$

for all sufficiently large n.

Next, consider $\mathcal{A}^*$, the largest intersecting subfamily of $\mathcal{A}$, which has cardinality $\mathbf{N} - \ell$. We know from Theorem 2.2 that there exists an $x \in [n]$ such that $|\mathcal{A}^*_x| \geq \mathbf{N} - C\ell$, so $|\mathcal{A}_x| \geq \mathbf{N} - C\ell$. It is then easy to see that

$$\begin{aligned}\mathbb{E}[X_\ell] &\leq n\binom{\mathbf{N}}{C\ell}\binom{\mathbf{V}}{C\ell}(1-p)^{\ell^{1+\delta/2}} \\ &\leq \exp\big(\ell\big(2Cn - p\ell^{\delta/2}\big)\big).\end{aligned}$$

Hence, if $c \leq c_m\delta/4$ so that $p \geq \mathbf{N}^{-c_m\delta/4}$, it follows that

$$\sum_{\ell=L_m}^{L_r} \mathbb{E}[X_\ell] \leq \sum_{\ell=L_m}^{L_r} \exp\big(\ell\big(2Cn - \mathbf{N}^{c_m\delta/4}/2\big)\big) = o(1),$$

so with high probability, for each $L_m \leq \ell \leq L_r$, the random variable X_ℓ is zero.

Putting the different parts of our argument together, we find that if $0 < \varepsilon < 1/10$,

$$c = c(\varepsilon) = \min\left(\frac{c_m(\varepsilon)}{2}, \frac{c'_m(\varepsilon)}{2}, \frac{c_r(\varepsilon)}{4}, \frac{c'_r(\varepsilon)}{4}, \frac{c_m(\varepsilon)\delta(\varepsilon)}{2}\right)$$

and $p \geq \mathbf{N}^{-c}$, then for all $r = r(n) \leq (1/2 - \varepsilon)n$, we have

$$\mathbb{P}\left(\alpha(K_p(n,r)) = \binom{n-1}{r-1}\right) \to 1$$

as $n \to \infty$. This completes the proof of Theorem 1.1. □

5. Refinements

We briefly discuss how one might tighten up the arguments in Theorem 1.1 so as to improve the dependence of $c(\varepsilon)$ on ε in the result. However, since it seems unlikely to us that these methods will be sufficient to determine the precise critical threshold at which Theorem 1.1 ceases to hold, we shall keep the discussion in this section largely informal.

5.1. **Containers for sparse sets in the Kneser graph.** The first approach we sketch involves using ideas from the theory of 'graph containers' to count large sparse sets in the Kneser graph more efficiently.

The theory of graph containers was originally developed to efficiently count the number of independent sets in a graph satisfying some kind of 'supersaturation' condition. The basic principle used to construct containers for graphs can be traced back to the work of Kleitman and Winston [15]. A great deal of work has since gone into refining and generalising their ideas, culminating in the results of Balogh, Morris and Samotij [3] and Saxton and Thomason [24]; these papers also give a detailed account of the history behind these ideas and we refer the interested reader there for details about how the general methodology was developed. Here we shall content ourselves with a brief discussion of how these ideas might be used to improve the dependence of $c(\varepsilon)$ on ε in Theorem 1.1.

Let us write $Y_m = Y_m(n,r)$ for the number of families $\mathcal{A} \subset [n]^{(r)}$ with $|\mathcal{A}| = \mathbf{N}$ and $e(\mathcal{A}) = m$. Clearly, to show that $\alpha(K_p(n,r)) = \mathbf{N}$ with high probability, it suffices to show that $\sum_{m \geq 1} Y_m (1-p)^m = o(1)$. Hence, it would be useful to have good estimates for Y_m. We shall derive some bounds for Y_m; see Theorem 5.2 below. These bounds are not strong enough (especially for small values of m) to prove Theorem 1.1. However, note that in our proof of Theorem 1.1, we use the somewhat cavalier bound of $\binom{\mathbf{V}}{\mathbf{N}}$ for the number of families $\mathcal{A}$ of size $\mathbf{N}$ for which $\ell(\mathcal{A})$ is equal to some prescribed value (in Case 2 of the proof); we can instead use Theorem 5.2 to count more efficiently.

To prove an effective container theorem, one needs to first establish a suitable supersaturation property. Lovász [18] determined the second largest eigenvalue of the Kneser graph; by combining Lovász's result with the expander mixing lemma, Balogh, Das, Delcourt, Liu and Sharifzadeh [2] proved the following supersaturation theorem for the Kneser graph.

Proposition 5.1. *Let $n, r, k \in \mathbb{N}$ and suppose that $n > 2r$ and $k \leq \mathbf{V} - \mathbf{N}$. If $\mathcal{A} \subset [n]^{(r)}$ has cardinality $\mathbf{N} + k$, then $e(\mathcal{A}) \geq k\mathbf{M}/2$.* □

Using Proposition 5.1, we prove the following container theorem for the Kneser graph.

Theorem 5.2. *For every $\varepsilon > 0$, there exists a $\hat{C} = \hat{C}(\varepsilon) > 0$ such that for every $\beta > 0$ and all $n, r, m \in \mathbb{N}$ with $\varepsilon n \leq r \leq (1/2 - \varepsilon)n$, the following holds: writing*

$$k_1 = \hat{C}\left(\frac{\mathbf{N}}{\beta\mathbf{M}} + \left(\frac{m\mathbf{N}}{\beta\mathbf{M}}\right)^{1/2}\right)$$

and

$$k_2 = k_1 + \hat{C}\beta\mathbf{N},$$

there exist, for $1 \leq i \leq \sum_{j=0}^{k_1} \binom{\mathbf{V}}{j}$, families $\mathcal{B}_i \subset [n]^{(r)}$ each of cardinality at most $\mathbf{N} + k_2$ with the property that each $\mathcal{A} \subset [n]^{(r)}$ with $e(\mathcal{A}) \leq m$ is contained in one of these families.

The advantage of this formulation of Theorem 5.2 in terms of k_1, k_2 and β is that we can apply the theorem with a value of $\beta > 0$ suitably chosen for the application at hand.

It is easy to check from Theorem 5.2 that $Y_m = Y_m(n, r)$, the number of families $\mathcal{A} \subset [n]^{(r)}$ with $|\mathcal{A}| = \mathbf{N}$ and $e(\mathcal{A}) = m$, satisfies

$$\begin{aligned} Y_m(n,r) &\leq \left(\sum_{j=0}^{k_1}\binom{\mathbf{V}}{j}\right)\binom{\mathbf{N}+k_2}{\mathbf{N}} = 2\binom{\mathbf{V}}{k_1}\binom{\mathbf{N}+k_2}{k_2} \leq 2\binom{\mathbf{V}}{k_1}\binom{\mathbf{V}}{k_2} \\ &\leq 2\exp\left(\hat{C}n\left(\beta\mathbf{N} + \frac{2\mathbf{N}}{\beta\mathbf{M}} + \left(\frac{4m\mathbf{N}}{\beta\mathbf{M}}\right)^{1/2}\right)\right) \end{aligned}$$

for all $\beta > 0$ such that $k_1 < \mathbf{V}/3$. We can then optimise this bound by choosing β depending on how large m is in comparison to $\mathbf{M}$ and $\mathbf{N}$. For example, when $m \geq \mathbf{N}/\mathbf{M}^{1/2}$, we can take $\beta = (m/\mathbf{N}\mathbf{M})^{1/3}$ and easily check that $Y_m(n, r) \leq \exp(10\hat{C}n(m\mathbf{N}^2/\mathbf{M})^{1/3})$. The reader may check that this estimate for Y_m when combined with the Hilton–Milner theorem is sufficient to prove Theorem 1.1 when r/n is bounded above by and away from ϑ, where $\vartheta \approx 0.362$ is, writing $H(x) = -x\log x - (1-x)\log(1-x)$, the unique real solution to the equation

$$3(1-\vartheta)H\left(\frac{\vartheta}{1-\vartheta}\right) = 2H(\vartheta)$$

in the interval $(0, 1)$.

Proof of Theorem 5.2. We start by proving a lemma whose proof is loosely based on the methods of Saxton and Thomason [24]. Before we state the lemma, let us have

some notation. Given a graph $G = (V, E)$ and $U \subset V(G)$, we write

$$\mu(U) = \frac{|E(G[U])|}{|V|};$$

in other words, $\mu(U)$ is the number of edges induced by U divided by the number of vertices of G. Also, we write $\mathcal{P}(X)$ for the collection of all subsets of a set X.

Lemma 5.3. *Let $G = (V, E)$ be a graph with average degree d and maximum degree Δ. For every $a \geq 0$ and $b > 0$, there is a map $\mathcal{C} : \mathcal{P}(V) \to \mathcal{P}(V)$ with the following property: for every $U \subset V$ with $\mu(U) \leq a$, there is a subset $T \subset V$ such that*

(1) $T \subset U \subset \mathcal{C}(T)$,
(2) $|T| \leq 2|V|(a/bd)^{1/2} + |V|/bd$, *and*
(3) $\mu(\mathcal{C}(T)) \leq 2\Delta(a/bd)^{1/2} + \Delta/bd + bd$.

Proof. We shall describe an algorithm that constructs T given U. The algorithm will also construct $\mathcal{C}(T)$ in parallel; it will be clear from the algorithm that $\mathcal{C}(T)$ is entirely determined by T and in no way depends on U.

Fix a linear ordering of the vertex set V of G. If u and v are adjacent and u precedes v in our ordering, we call v a *forward neighbour* of u and u a *backward neighbour* of v. For a vertex $v \in V$, we write $F(v)$ for the set of its forward neighbours.

We begin by setting $T = \varnothing$ and $A = V$. We shall iterate through V in the order we have fixed and add vertices to T and remove vertices from A as we go along; at any stage, we write $\Gamma(T)$ to denote the set of those vertices which, at that stage, have k or more backward neighbours in T where k is the least integer strictly greater than $(abd)^{1/2}$.

As we iterate through the vertices of V in order, we do the following when considering a vertex v.

(1) If $v \in \Gamma(T)$, we remove v from A; if it is also the case that $v \in U$, then we add v to T.
(2) If $v \notin \Gamma(T)$, we consider the size of $S = F(v) \setminus \Gamma(T)$.
    (a) If $|S| \geq bd$, we remove v from A; if it is also the case that $v \in U$, then we add v to T.
    (b) If $|S| < bd$, we do nothing.

The algorithm outputs T and A when it terminates; we then set $\mathcal{C}(T) = A \cup T$. It is clear from the algorithm that $\mathcal{C}(T)$ is uniquely determined by T and that $T \subset U \subset \mathcal{C}(T)$.

We first show that $|T| \leq 2|V|(a/bd)^{1/2} + |V|/bd$. Consider the partition $T = T_1 \cup T_2$ where T_1 consists of those vertices which were added to T on account of condition (1) and T_2 of those vertices which were added to T when considering condition (2a). The upper bound for $|T|$ follows from the following two claims.

Claim 5.4. $|T_1| \leq |E(G[U])|/k$.

Proof. Clearly, each vertex of T_1 has at least k backward neighbours in $T \subset U$. Hence, $k|T_1| \leq |E(G[U])|$. □

Claim 5.5. $|T_2| \leq k|V|/bd$.

Proof. Let us mark all the edges from v to $F(v) \setminus \Gamma(T)$ when a vertex v gets added to T on account of condition (2a). The number of marked edges is clearly at least $bd|T_2|$. On the other hand, by the definition of $\Gamma(T)$, each vertex is joined to at most k of its backward neighbours by a marked edge. Hence, $bd|T_2| \leq k|V|$. □

Consequently, since $(abd)^{1/2} < k \leq (abd)^{1/2} + 1$, we have

$$\begin{aligned} |T| = |T_1| + |T_2| &\leq \frac{a|V|}{k} + \frac{k|V|}{bd} \\ &\leq \frac{a|V|}{(abd)^{1/2}} + \frac{((abd)^{1/2}+1)|V|}{bd} \leq 2|V|\Big(\frac{a}{bd}\Big)^{1/2} + \frac{|V|}{bd}. \end{aligned}$$

It remains to show that $\mu(\mathcal{C}) \leq 2\Delta(a/bd)^{1/2} + \Delta/bd + bd$. To see this, recall that $\mathcal{C}(T) = A \cup T$ and notice that

$$|E(G[\mathcal{C}(T)])| \leq \Delta|T| + |E(G[A])| \leq \Delta|T| + bd|V|.$$

To see the last inequality, i.e., $|E(G[A])| \leq bd|V|$, note that a vertex v is removed from A by our algorithm unless we have $|F(v) \setminus \Gamma(T)| < bd$ at the stage where we consider v. Since each member of $\Gamma(T)$ is (eventually) removed from A, we see that each vertex of A has at most bd forward neighbours in A and the inequality follows. The claimed bound for $\mu(\mathcal{C})$ then follows from our previously established upper bound for $|T|$. □

To prove Theorem 5.2, we now combine Lemma 5.3 with Proposition 5.1. First note that the Kneser graph $K(n,r)$ has $\mathbf{V} = n\mathbf{N}/r$ vertices and is $(n-r)\mathbf{M}/r$ regular.

Let us take $\hat{C}(\varepsilon) = 20/\varepsilon^2$. It is easy to check that given $\beta > 0$ and a family $\mathcal{A} \subset [n]^{(r)}$ with $e(\mathcal{A}) \leq m$, we can apply Lemma 5.3 with $a = m/\mathbf{V}$ and $b = \beta$ to get families $\mathcal{T} \subset [n]^{(r)}$ and $\mathcal{C}(\mathcal{T}) \subset [n]^{(r)}$ such that $\mathcal{T} \subset \mathcal{A} \subset \mathcal{C}(\mathcal{T})$, $|\mathcal{T}| \leq k_1$ and $e(\mathcal{C}(\mathcal{T})) \leq k_2\mathbf{M}/2$. Hence, by Proposition 5.1, we see that $|\mathcal{C}(\mathcal{T})| \leq \mathbf{N} + k_2$. The theorem then follows by taking the families $\mathcal{C}(\mathcal{T})$ for every $\mathcal{T} \subset [n]^{(r)}$ with $|\mathcal{T}| \leq k_1$. □

5.2. **Stability for the Kruskal–Katona theorem.** An important ingredient in our proof of Theorem 1.1 is Lemma 3.4 which gives a uniform lower bound, using Theorem 2.2 and the Kruskal–Katona theorem, for $e(\mathcal{A})$ in terms of $\ell(\mathcal{A})$ when the size of $\mathcal{A}$ is large.

However, there is a price to be paid for proving such a uniform bound: the bound is quite poor for most families to which the lemma can be applied. Indeed, the families which are extremal for the argument in the proof of Lemma 3.4 must possess a great

deal of structure. Instead of the Kruskal–Katona theorem, one should be able to use a stability version of the Kruskal–Katona theorem, as proved by Keevash [13] for example, to prove a more general result that accounts for the structure of the family under consideration.

6. Conclusion

Several problems related to the question considered here remain. First of all, it would be good to determine the largest possible value of $c(\varepsilon)$ with which Theorem 1.1 holds. It is likely that one needs new ideas to resolve this problem.

Second, one would also like to know what happens when r is very close to $n/2$. Perhaps most interesting is the case when $n = 2r + 1$; one would like to know the values of p for which we have $\alpha(K_p(2r+1, r)) = \binom{2r}{r-1}$ with high probability. A simple calculation shows that $p = 3/4$ is the threshold at which we are likely to find a star and an r-set not in the star all the edges between which are missing in $K_p(2r+1, r)$ which suggests that the critical threshold should be $3/4$. However, it would even be interesting to show that $\alpha(K_p(2r+1, r)) = \binom{2r}{r-1}$ with high probability for, say, all $p \geq 0.999$.

Acknowledgements

The first author was partially supported by a Simons fellowship, NSF CAREER grant DMS-0745185, an Arnold O. Beckman Research Award (UIUC Campus Research Board 13039) and Marie Curie grant FP7-PEOPLE-2012-IIF 327763. The second author would like to acknowledge support from EU MULTIPLEX grant 317532 and NSF grant DMS-1301614.

Some of the research in this paper was carried out while the authors were visitors at the Alfréd Rényi Institute of Mathematics. This research was continued while the third author was a visitor at the University of Memphis. The authors are grateful for the hospitality of the Rényi Institute and the third author is additionally grateful for the hospitality of the University of Memphis.

We would also like to thank Andrew Thomason for some helpful discussions about graph containers.

References

1. L. Babai, M. Simonovits, and J. Spencer, *Extremal subgraphs of random graphs*, J. Graph Theory **14** (1990), 599–622. 1
2. J. Balogh, S. Das, M. Delcourt, H. Liu, and M. Sharifzadeh, *The typical structure of intersecting families of discrete structures*, J. Combin. Theory Ser. A, To appear. 2, 5, 10

3. J. Balogh, R. Morris, and W. Samotij, *Independent sets in hypergraphs*, J. Amer. Math. Soc., To appear. 1, 10
4. L. I. Bogolyubskiy, A. S. Gusev, M. M. Pyaderkin, and A. M. Raigorodskii, *The independence numbers and the chromatic numbers of random subgraphs of some distance graphs*, Mat. Sb., To appear. 2
5. B. Bollobás, *On generalized graphs*, Acta Math. Acad. Sci. Hungar **16** (1965), 447–452. 3
6. B. Bollobás, B. Narayanan, and A. Raigorodskii, *On the stability of the Erdős-Ko-Rado theorem*, J. Combin. Theory Ser. A, To appear. 2
7. D. Conlon and T. Gowers, *Combinatorial theorems in sparse random sets*, Ann. of Math., To appear. 1
8. I. Dinur and E. Friedgut, *Intersecting families are essentially contained in juntas*, Combin. Probab. Comput. **18** (2009), 107–122. 2
9. P. Erdős, C. Ko, and R. Rado, *Intersection theorems for systems of finite sets*, Quart. J. Math. Oxford **12** (1961), 313–320. 1
10. E. Friedgut, *On the measure of intersecting families, uniqueness and stability*, Combinatorica **28** (2008), 503–528. 2, 3
11. A. J. W. Hilton and E. C. Milner, *Some intersection theorems for systems of finite sets*, Quart. J. Math. Oxford **18** (1967), 369–384. 2, 3
12. G. Katona, *A theorem of finite sets*, Theory of graphs (Proc. Colloq., Tihany, 1966), Academic Press, New York, 1968, pp. 187–207. 3
13. P. Keevash, *Shadows and intersections*: *stability and new proofs*, Adv. Math. **218** (2008), no. 5, 1685–1703. 14
14. P. Keevash and D. Mubayi, *Set systems without a simplex or a cluster*, Combinatorica **30** (2010), 175–200. 2
15. D. J. Kleitman and K. J. Winston, *On the number of graphs without* 4*-cycles*, Discrete Math. **41** (1982), 167–172. 10
16. Y. Kohayakawa, T. Łuczak, and V. Rödl, *Arithmetic progressions of length three in subsets of a random set*, Acta Arith. **75** (1996), 133–163. 1
17. J. B. Kruskal, *The number of simplices in a complex*, Mathematical optimization techniques, Univ. of California Press, Berkeley, Calif., 1963, pp. 251–278. 3
18. L. Lovász, *On the Shannon capacity of a graph*, IEEE Trans. Inform. Theory **25** (1979), 1–7. 10
19. ———, *Combinatorial problems and exercises*, 2nd ed., AMS Chelsea Publishing, Providence, RI, 2007. 4
20. T. Łuczak, *Randomness and regularity*, Proceedings of the International Congress of Mathematicians, Eur. Math. Soc., Zürich, 2006, pp. 899–909. 1
21. V. Rödl and A. Ruciński, *Threshold functions for Ramsey properties*, J. Amer. Math. Soc. **8** (1995), 917–942. 1

22. _____, *Rado partition theorem for random subsets of integers*, Proc. London Math. Soc. **74** (1997), 481–502. 1
23. V. Rödl and M. Schacht, *Extremal results in random graphs*, Erdős Centennial, Springer Berlin Heidelberg, 2013, pp. 535–583. 1
24. D. Saxton and A. Thomason, *Hypergraph containers*, Invent. Math., To appear. 1, 10, 11
25. M. Schacht, *Extremal results for random discrete structures*, Ann. of Math., To appear. 1

Department of Mathematics, University of Illinois, 1409 W. Green Street, Urbana IL 61801, USA, *and* Bolyai Institute, University of Szeged, 6720 Szeged, Hungary

Email address: jobal@math.uiuc.edu

Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK, *and* Department of Mathematical Sciences, University of Memphis, Memphis TN 38152, USA, *and* London Institute for Mathematical Sciences, 35a South St., Mayfair, London W1K 2XF, UK

Email address: b.bollobas@dpmms.cam.ac.uk

Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK

Email address: b.p.narayanan@dpmms.cam.ac.uk