Europ. J. Combinatorics (1985) **6**, 317–326

Near Perfect Coverings in Graphs and Hypergraphs

P. Frankl and V. Rödl

Suppose we are given a bipartite graph with vertex set X, Y, $|X| = n$, $|Y| = N$, each point in $X(Y)$ has degree $D(d)$ fixed, respectively, moreover, each pair of points $x, x' \in X$ has at most $D/(\log n)^3$ (common) neighbours. Let $t(X, Y)$ denote the minimum number of vertices of Y needed to cover all vertices of X. We prove (Theorem 1.1) that $t(X, Y)d/n$ tends to 1 as n tends to infinity.

This result has many applications:

Theorem [5]. Suppose $k > r > 1$ are fixed, $n \to \infty$. Then there exists a collection of $(1+o(1)) \times \binom{n}{r}/\binom{k}{r}$ k-subsets of an n-set so that each r-subset is contained in at least one member of the collection.

Analogues and strengthenings of this result are deduced. E.g. for vector spaces, orthogonal or simplectic geometries, random collections of k-sets with constant probabilities, etc.

Theorem 3.3. Suppose $\mathcal{G}$ is a graph on v vertices and e edges and $\mathcal{R}$ is a random graph on n vertices and edge probability $e/\binom{v}{2}$. Then there exists a collection of $(1+o(1))\binom{n}{2}/\binom{v}{2}$ induced subgraphs of $\mathcal{R}$ on v vertices, isomorphic to $\mathcal{G}$ and such that each edge (non-edge) of $\mathcal{R}$ is covered by an edge (non-edge) of a graph in the collection.

1. Introduction

Most problems of graph theory and combinatorics can be formulated as a packing or covering problem for hypergraphs.

Let us recall that a hypergraph $\mathcal{H}$ is a collection of nonempty subsets—called edges—of a set X. One defines the *packing number* $\nu(\mathcal{H})$ as the maximum number of pairwise disjoint edges of $\mathcal{H}$. The *point covering* number $t(\mathcal{H})$ is the minimum number t so that there exist t edges of $\mathcal{H}$ whose union is the whole set X. The hypergraph $\mathcal{H}$ is said to be *d-uniform*, if all edges of $\mathcal{H}$ consist of d vertices. Then, clearly, $\nu(\mathcal{H}) \leq n/d \leq t(\mathcal{H})$ holds. Moreover, equality on one side implies equality on the other. A family of edges of $\mathcal{H}$ showing $\nu(\mathcal{H}) = n/d$ is called a *perfect packing* (and is necessarily a *perfect covering* as well). The existence problem of perfect packings is a very difficult one. For example, if $X = \{(i,j) = 1 \leq i < j \leq 111\}$ and $\mathcal{H} = \{\{(i,j): i, j \in E, |E| = 11, E \subset \{1, \ldots, 111\}\}$, then the existence of a perfect packing is equivalent to the existence of a projective plane of order 10.

A hypergraph can be also considered as a bipartite graph with underlying set X, $\mathcal{H}$ where $x \in X$ and $H \in \mathcal{H}$ form an edge iff $x \in H$ holds. Then $\mathcal{H}$ is d-uniform iff the bipartite graph is d-regular on the side of $\mathcal{H}$.

Looking at this bipartite graph from the other side we obtain the *dual hypergraph* whose vertex set is $\mathcal{H}$, edge set is X, and the incidence relations are unchanged.

The point covering number of the dual hypergraph is the *edge-covering number*, $\tau(\mathcal{H})$ of the original, that is, min τ such that there exists $T \subset X$, $|T| = \tau$ and $T \cap H \neq \varnothing$ holds for all $H \in \mathcal{H}$. It is easy to see that $\nu(\mathcal{H}) \leq \tau(\mathcal{H})$ holds, moreover, if $\mathcal{H}$ is d-uniform then $\tau(\mathcal{H}) \leq d\nu(\mathcal{H})$.

A special case of a theorem of Lovász [L] says that if $\mathcal{H}$ is d-uniform, *d'-regular* (i.e. every point is contained in d' edges) then

$$t(\mathcal{H}) \leq \frac{n}{d}(1 + \log d').$$

For $Y \subset X$ let us define $\mathcal{H}(Y) = \{H \in \mathcal{H}: Y \subset H\}$, $\deg(Y) = |\mathcal{H}(Y)|$.

0195-6698/85/040317+10 \$02.00/0

In this paper we apply the probabilistic approach of [5] to show that for a wide class of hypergraphs $t(\mathcal{H}) \leq n(1+o(1))/d$ holds, that is *near perfect* coverings exist.

THEOREM 1.1. *Suppose $\varepsilon > 0$ is arbitrary, $\mathcal{H}$ is a d-uniform hypergraph on X, $|X| = n$, $a > 3$ is a real number. There exists a positive real $\delta = \delta(\varepsilon)$ such that if for some D one has $(1-\delta)D < \deg(x) < (1+\delta)D$ for all $x \in X$ and $\deg(\{x, y\}) < D/(\log n)^a$ holds for all distinct $x, y \in X$, then, for all $n > n_0(\delta)$,*

$$t(\mathcal{H}) \leq n(1+\varepsilon)/d \text{ holds.}$$

REMARK. Clearly $t(\mathcal{H}) \leq n(1+\varepsilon)/d$ implies $\nu(\mathcal{H}) \geq n(1-d\varepsilon)/d$.

A few applications of this theorem are described in the next sections. The proof of the theorem is presented in Section 5.

2. PACKING OF GRAPHS BY SUBGRAPHS

Let $\mathcal{G}(n, p)$ denote the random graph with edge probability p, that is, each edge is present in $\mathcal{G}(n, p)$ with independent probability p.

Let $\mathcal{A}$ be a fixed graph with v vertices and e edges. For an arbitrary graph $\mathcal{G}$, $|\mathcal{G}|$ denotes the number of its edges and $\pi(\mathcal{G}, \mathcal{A})$ the packing number of $\mathcal{G}$ with respect to $\mathcal{A}$, that is the maximum number of pairwise edge disjoint copies of $\mathcal{A}$ in $\mathcal{G}$.

Similarly $\pi_i(\mathcal{G}, \mathcal{A})$ is the induced packing number of $\mathcal{G}$, that is the maximum number of pairwise edge disjoint induced copies of $\mathcal{A}$ in $\mathcal{G}$.

Theorem 1.1 has the following immediate corollaries.

COROLLARY 2.1 [6]. *Suppose $\mathcal{A}$ is a fixed graph, K_n is the complete graph on n vertices. Then for an arbitrary positive ε and $n > n_0(\varepsilon)$ we have $\pi(K_n, \mathcal{A}) > (1-\varepsilon)\binom{n}{2}/|\mathcal{A}|$.*

In fact Wilson proved more, he showed that $\pi(K_n, \mathcal{A}) = \binom{n}{2}/|\mathcal{A}|$ provided $n > n_0(\varepsilon)$, $v(v-1) \mid n$ and $d \mid v-1$ hold, where d is the greatest common divisor of the degrees of $\mathcal{A}$ and v is the number of vertices of $\mathcal{A}$.

COROLLARY 2.2 [2]. *For an arbitrary fixed integers k*

$$\pi(K_n, K_k) = (1+o(1))\binom{n}{2}\bigg/\binom{k}{2}.$$

COROLLARY 2.3 [1]. *Suppose $\mathcal{A}$ is a fixed graph, p is a fixed real, $0 < p < 1$ then for n tending to infinity we have*

$$\pi(G(n, p), \mathcal{A}) = (1-o(1))\binom{n}{2}p\bigg/|\mathcal{A}|.$$

Recall that $\chi(\mathcal{A})$ denotes the chromatic number of $\mathcal{A}$ and $K(n_1, n_2, \ldots, n_r)$ is the complete r-chromatic graph, i.e. it has vertex set $X_1 \cup X_2 \cup \cdots \cup X_r$, $|X_i| = n_i$, $X_i \cap X_j = 0$ and (x_i, x_j) is an edge iff $x_i \in X_i$, $x_j \in X_j$ and $i \neq j$. Suppose that $n_1 \leq n_2 \leq \cdots \leq n_r$.

THEOREM 2.4. *Suppose $\mathcal{A}$ is a fixed graph, $\chi(\mathcal{A}) \leq r$, $\varepsilon > 0$ is a real number. Then there exists a real δ such that for $n_1 > (1-\delta)n_i$ $(2 \leq i \leq r)$ and all $n_1 > n_0(\mathcal{A}, \varepsilon)$ one has*

$$\pi(K(n_1, n_2, \ldots, n_r), \mathcal{A}) > (1-\varepsilon)\frac{|K(n_1, n_2, \ldots, n_r)|}{|\mathcal{A}|}.$$

3. Packing of Hypergraphs by Subhypergraphs

Suppose $\mathcal{H}$ *and* $\mathcal{B}$ are d-uniform hypergraphs. The packing number of $\mathcal{H}$ with respect to $\mathcal{B}$, $\pi(\mathcal{H}, \mathcal{B})$ is defined in complete analogy with the graph case, it is the maximum number of pairwise edge-disjoint copies of $\mathcal{B}$ in $\mathcal{H}$.

We denote by $\mathcal{H}_d(n, p)$ the random d-uniform hypergraph with each edge having probability p for being chosen.

Also, K_n^d denotes the complete d-uniform hypergraph on n vertices.

Theorem 3.1. *Suppose $\mathcal{B}$ is a fixed d-uniform hypergraph and n tends to infinity. Then*

$$\pi(K_n^d, \mathcal{B}) = (1-o(1))\binom{n}{d}\Big/|\mathcal{B}|$$

holds.

Corollary 3.2 [5]. *Suppose k is fixed, $n \to \infty$. Then*

$$\pi(K_n^d, K_k^d) = (1-o(1))\binom{n}{d}\Big/\binom{k}{d}.$$

Recall that a (d, k, n)-design is a family $\mathcal{S}$ of k-subsets of the n-set X, such that every d-subset of X is contained in exactly one member of $\mathcal{S}$. That is, it corresponds to a perfect packing of K_n^d by K_k^d. The existence problem of (d, k, n)-designs is a hopelessly difficult one, e.g. no such design is known for $d \geq 6$. Wilson [7] showed that for $d = 2$ and fixed k the trivial necessary conditions $(\binom{k}{2} | \binom{n}{2}$, $(k-1)|(n-1))$ are sufficient for $n > n_0(k)$. His theorem gives an estimation of c/n for $o(1)$ in Corollary 3.2. Corollary 3.2 shows that near-designs exist always, as it was conjectured by Erdös and Hanani [2].

Suppose $r \geq 2$ is an integer and the edges of a d-uniform hypergraph $\mathcal{B}$ are coloured by r colours, i.e. partitioned into r subhypergraphs $\mathcal{B}_1, \ldots, \mathcal{B}_r$.

If $\mathcal{H} = \mathcal{H}_1 \cup \cdots \cup \mathcal{H}_r$ is an r-coloured hypergraph then we can define the coloured packing number $\pi_c(\mathcal{H}, \mathcal{B})$ as the maximum number of edge disjoint copies of $\mathcal{B}$ in $\mathcal{H}$ with the additional property that if an edge $B \in \mathcal{B}$ has colour i in $\mathcal{B}$ then it has colour i in $\mathcal{H}$ as well.

Define the rational numbers b_i by $b_i = |\mathcal{B}_i|/|\mathcal{B}|$. Note that $\sum b_i = 1$. Define $\mathcal{H}_d(n, p, (b_1 b_2, \ldots, b_r))$ a random partition $\mathcal{H}_1 \cup \cdots \cup \mathcal{H}_r$ of the edges of $\mathcal{H}_d(n, p)$ where the probability for each edge $H \in \mathcal{H}_d(n, p)$ of the event $H \in \mathcal{H}_i$ is b_i.

Theorem 3.3. *Suppose $0 < p < 1$ and an r-coloured d-uniform hypergraph $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ is fixed. Then*

$$\pi_c(\mathcal{H}_d(n, p, (b_1, \ldots, b_r)), \mathcal{B}) = (1-o(1))\binom{n}{d}p\Big/|\mathcal{B}|$$

holds.

An interesting special case of this theorem is $p = 1$, $\mathcal{B} = K_k^d$, $r = 2$. Then Theorem 3.3 asserts that $\mathcal{H}_d(n, b_1)$ can be packed near perfectly by copies of $\mathcal{B}_1$ in a way that the complements of $\mathcal{B}_1$ (i.e. $\mathcal{B}_2$) form a near perfect packing of the complement of $\mathcal{H}_d(n, b_1)$.

4. Packing of Geometric Configurations

Suppose K is a finite field $|K|=q$ and V is an n-dimensional vector space over K. Then $\binom{V}{d}$ is the collection of all d-dimensional subspaces of V,

$$\left|\binom{V}{d}\right|=\begin{bmatrix} n \\ d \end{bmatrix}=\prod_{i=0}^{d-1}\frac{q^n-q^i}{q^d-q^i}.$$

In 1972 Ray-Chaudhuri [4] raised the problem of the existence of designs among subspaces. A family $\mathcal{W}\subset\binom{V}{r}$ is called a (n, r, t)-design or simply t-design if for all $U\in\binom{V}{t}$ there is exactly one $W\in\mathcal{W}$ with $U<W$. Clearly, this implies

$$|\mathcal{W}|=\begin{bmatrix} n \\ t \end{bmatrix}\Big/\begin{bmatrix} r \\ t \end{bmatrix}.$$

Let us note that 1-designs exist if and only if r divides n but no general existence theorems for t-designs with $t\geqslant 2$ are known.

Here we show that 'almost' t-designs exist:

THEOREM 4.1. *Suppose integers r, t are given, $r>t>0$. Then for $n\to\infty$ there exists*

$$\mathcal{W}\subset\binom{V}{r},\qquad |\mathcal{W}|=(1-o(1))\begin{bmatrix} n \\ t \end{bmatrix}\Big/\begin{bmatrix} r \\ t \end{bmatrix}$$

such that $\dim(W\cap W')<t$ *holds for all distinct $W, W'\in\mathcal{W}$.*

Suppose V is endowed with a nondegenerate scalar product f: $V\times V\to K$ (i.e. $\dim W+\dim W^{\perp}=n$ for all subspaces $W<V$), f may be orthogonal, simplectic (in this case n is even) or Hermitian (then $|K|$ is a square).

A subspace $W<V$ is called totally singular if $f(w, w')=0$ for all $w, w'\in W$. Denote by $\begin{bmatrix} n \\ t \end{bmatrix}_f$ the number of t-dimensional totally singular subspaces of V.

THEOREM 4.2. *Given $r>t\geqslant 0$, fixed and $n\to\infty$, then there exists a collection $\mathcal{W}$ of r-dimensional totally singular subspaces of V such that* $\dim(W\cap W')<t$ *for all distinct*

$$W, W'\in\mathcal{W}\quad\textit{and}\quad |\mathcal{W}|=(1-o(1))\begin{bmatrix} n \\ t \end{bmatrix}_f\Big/\begin{bmatrix} r \\ t \end{bmatrix}.$$

5. Proof of Theorem 1.1

We will need the following simple inequality:

PROPOSITION 5.1. *Suppose $Y_1, Y_2, \ldots, Y_b$ are totally independent identically distributed random variables with $p(Y_i=1)=p$, $p(Y_i=0)=1-p$, $s\geqslant 1$, $t=\lceil esbp\rceil$. Then we have*

$$p\left(\sum_{i=1}^{b} Y_i\geqslant t\right)<2\cdot s^{-t}. \tag{5.0}$$

PROOF. The probability on the LHS is

$$\sum_{j\geqslant t}\binom{b}{j}p^j(1-p)^{b-j}<\sum_{j\geqslant t}\binom{b}{j}p^j.$$

In this sum the ratio of consecutive terms is

$$\binom{b}{j+1}p^{j+1}\Big/\binom{b}{j}p^j<\frac{bp}{j+1}<\tfrac{1}{2},\quad \text{for } j\geqslant t.$$

Thus the sum is bounded by the double of the first term. On the other hand, using $t! > (t/\mathrm{e})^t$, we have

$$\binom{b}{t} p^t < \frac{(bp)^t \mathrm{e}^t}{t^t} \leq s^{-t},$$

and (5.0) follows.

THE MAIN LEMMA

NOTATION. We shall write $A \sim_\rho B$ to denote that

$$1-\rho \leq A/B \leq 1+\rho.$$

LEMMA 5.2. *Suppose $\varepsilon > 0$ is given and $\mathcal{F}$ is d-uniform hypergraph on X, $|X| = n$, $|\mathcal{F}| = m$ such that the following two properties are satisfied for all $x, y \in X$:*
(i) *$|\mathcal{F}(x)| \sim_\rho D$ ($\rho = \rho(\varepsilon)$ is a sufficiently small positive real)*
(ii) *$|\mathcal{F}(x, y)| < D/(\log n)^a$, $a > 3$.*
Then for $n > n_0(\varepsilon, \rho)$ there exists $\mathcal{R} \subset \mathcal{F}$ such that

$$|\mathcal{R}| \underset{2\rho}{\sim} \varepsilon n/d$$

$$\left|\bigcup \mathcal{R}\right| \underset{4\rho}{\sim} n(1-\mathrm{e}^{-\varepsilon}),$$

and $\tilde{\mathcal{F}} = \{F \in \mathcal{F}\colon F \cap R = \varnothing$ for all $R \in \mathcal{R}\}$ with vertex set $X - \bigcup \mathcal{R}$ satisfies (i) and (ii) for $\tilde{\rho} = 6\rho$, $\tilde{D} = e^{-(d-1)\varepsilon} D$, $\tilde{a} > a - o(1)$.

PROOF. Note that $|\mathcal{F}|d \sim_\rho Dn$ and let $\mathcal{R}$ be the random hypergraph which we obtain by choosing each edge of $\mathcal{F}$ with independent probability ε/D. Then the expected number of edges of $\mathcal{R}$ is $E(|\mathcal{R}|) = \varepsilon|\mathcal{F}|/D \sim_\rho \varepsilon n/d$ and with probability close to 1 exponentially in n $|\mathcal{R}| \sim_{2\rho} \varepsilon n/d$ holds.

Next we show that with probability tending to 1 $|\bigcup \mathcal{R}| \sim \varepsilon n$ holds. For a given $x \in X$ we have

$$p\left(x \in \bigcup \mathcal{R}\right) = 1 - \left(1 - \frac{\varepsilon}{D}\right)^{|\mathcal{F}(x)|} \underset{\rho}{\sim} 1 - e^{-\varepsilon}$$

Thus $E(|\bigcup \mathcal{R}|) \sim_\rho n(1-\mathrm{e}^{-\varepsilon})$. Let Z_i be the random variable defined by

$$Z_i = \begin{cases} 1 & \text{if } x_i \in \bigcup \mathcal{R} \\ 0 & \text{if } x_i \notin \bigcup \mathcal{R} \end{cases}, \quad X = \{x_1, \ldots, x_n\}.$$

Then $|\bigcup \mathcal{R}| = \sum Z_i$. We break up this sum into $t = n/(\log n)^{a/3}$ parts, i.e. suppose $\{1, 2, \ldots, n\}$ is partitioned into $I_1 \cup \cdots \cup I_t$, $|I_j - (\log n)^{a/3}| \leq 1$.

Next we estimate $\sum_{i \in I_j} Z_i$. Let I be one of I_j, $1 \leq j \leq t$. Note that for $i, i' \in I$

$$|\mathcal{F}(x_i) \cap \mathcal{F}(x_{i'})| = |\mathcal{F}(x_i, x_{i'})| < D/(\log n)^a$$

holds. Consequently,

$$\sum_{i,i' \in I} |\mathcal{F}(x_i) \cap \mathcal{F}(x_{i'})| < \binom{|I|}{2} D/(\log n)^a < D/2(\log n)^{a/3}. \tag{5.1}$$

Let us set $\mathcal{Y} = \bigcup_{x_i \in I} \mathcal{F}(x_i)$ and partition $\mathcal{Y}$ into $\mathcal{Y}_1 \cup \mathcal{Y}_2 \cup \cdots \cup \mathcal{Y}_d$ where $\mathcal{Y}_l$ consists of those elements of $\mathcal{Y}$ which appear exactly l times in the union. Now (5.1) implies

$$\sum_{2 \leq l \leq d} l|\mathcal{Y}_l| < D/(\log n)^{a/3} \tag{5.2}$$

and thus

$$|\mathcal{F}(x_i) \cap \mathcal{Y}_1| \underset{2\rho}{\sim} D \quad \text{for every } i \in I. \tag{5.3}$$

Let us define a new random variable Z_i^* by

$$Z_i^* = \begin{cases} 1, & \text{if } x_i \in R \in \mathcal{R} \text{ for some } R \in \mathcal{Y}_1, \\ 0, & \text{otherwise.} \end{cases}$$

We have $\sum_{i \in I} (Z_i - Z_i^*) \leq \sum_{2 \leq l \leq d} l|Y_l \cap \mathcal{R}|$ and moreover, clearly $Z_i^* \leq Z_i$ holds. The reason for considering Z_i^* is that the Z_i^* are completely independent for $i \in I$. Thus,

$$\sum Z_i^* \underset{3\rho}{\sim} |I|(1-e^{-\varepsilon})$$

holds with probability close to 1 exponentially in $|I|$, i.e., greater than $1 - 1/n^b$ for $n > n_0$ and some $b > 1$.

This implies $\sum_{i=1}^n Z_i^* \sim_{3\rho} n(1-e^{-\varepsilon})$ holds with probability greater than $1 - 1/n^{(b-1)}$.

On the other hand let $\mathcal{U}$ be the union of all $\mathcal{Y}_2 \cup \cdots \cup \mathcal{Y}_d$ for I running through $I_1, \ldots, I_t$. In view of (5.2), we have

$$|\mathcal{U}| < nD/(\log n)^{2a/3}.$$

Thus, in view of (5.0)

$$|\mathcal{U} \cap \mathcal{R}| < n/(\log n)^{a/3}$$

holds with probability greater than $1 - 1/n^b$, i.e., with this probability $\sum (Z_i - Z_i^*) < dn/(\log n)^{a/3}$ holds, yielding

$$\left|\bigcup \mathcal{R}\right| \underset{4\rho}{\sim} n(1-e^{-\varepsilon}) \quad \text{with probability} > 1 - 2/\, n^{(b-1)}.$$

Now we are coming to the most difficult part of the lemma, namely we want to show that the remainder, $\tilde{F} = \{F \in \mathcal{F}\colon F \cap R = \emptyset \text{ for all } R \in \mathcal{R}\}$ with vertex set $\tilde{X} = X - \bigcup \mathcal{R}$ satisfies (i) (and (ii)).

Let us consider an arbitrary vertex $x \in X - \bigcup \mathcal{R}$. This means that none of the edges from $\mathcal{F}(x)$ was chosen into $\mathcal{R}$.

Suppose $F \in \mathcal{F}(x)$. When does F remain, i.e., $F \in \tilde{\mathcal{F}}(x)$? Clearly, it is equivalent to $y \in \tilde{X}$ for all $y \in F$, i.e. to the event that none of the edges $\mathcal{F}(F, x) = \bigcup_{y \in F - \{x\}} \mathcal{F}(y)$ was chosen.

Clearly we have

$$(d-1)D - \binom{d-1}{2} \frac{D}{(\log n)^a} \underset{\rho}{\leq} |\mathcal{F}(F, x)| \leq (d-1)D. \tag{5.4}$$

Consequently, we have

$$p(F \in \tilde{\mathcal{F}}) = \left(1 - \frac{\varepsilon}{D}\right)^{|\mathcal{F}(F,x)|} \underset{\rho}{\sim} e^{-(d-1)\varepsilon}. \tag{5.5}$$

Thus the expected number of edges in $\tilde{F}(x)$ satisfies

$$E(|\tilde{\mathcal{F}}(x)|) \underset{2\rho}{\sim} e^{-(d-1)\varepsilon} D.$$

However, we need to prove $|\tilde{\mathcal{F}}(x)| \sim e^{-(d-1)\varepsilon} D$ with high probability. This would easily follow if the events $F \in \tilde{\mathcal{F}}$, $F' \in \tilde{\mathcal{F}}$ were independent, however, it is not the case.

To circumvent this difficulty we shall partition $\mathcal{F}(x)$ into stars.

Recall that a star in $\mathcal{F}(x)$ is a sub-family $\{F_1, F_2, \ldots, F_t\}$ satisfying $F_i \cap F_{i'} = \{x\}$ for $i \neq i'$. Let $t = [(\log n)^{a/3}]$ and suppose that we have already defined the pairwise disjoint stars $\mathcal{F}^{(1)}, \mathcal{F}^{(2)}, \ldots, \mathcal{F}^{(s)}$ in $\mathcal{F}(x)$. We want to continue by finding a star $\mathcal{F}^{(s+1)}$ in $\mathcal{G} = \mathcal{F}(x) - \bigcup_{i=1}^s \mathcal{F}^{(i)}$ with $|\mathcal{F}^{(s+1)}| = t$. Suppose it is impossible, i.e., there exist

$F_1, F_2, \ldots, F_l \in \mathscr{G}$, forming a star that for all $F \in \mathscr{G}\{F_1, F_2, \ldots, F_l, F\}$ is not a star. That is, the $(d-1)l$-element set $G = (F_1 \cup \cdots \cup F_l) - \{x\}$ meets all $F \in \mathscr{G}$. Thus $\mathscr{G} \subset \bigcup_{g \in G} \mathscr{F}(x, g)$, yielding

$$|\mathscr{G}| \leq (d-1)l \frac{D}{(\log n)^a} < d \frac{D}{(\log n)^{2a/3}}.$$

Now let us define the random variable $Y(F)$ by

$$Y(F) = \begin{cases} 1, & \text{if } F \in \tilde{\mathscr{F}}, \\ 0, & \text{if } F \notin \tilde{\mathscr{F}}. \end{cases}$$

Then

$$\sum_{F \in (\mathscr{F}(x) - \mathscr{G})} Y(F) \leq |\tilde{\mathscr{F}}(x)| \leq \sum_{F \in (\mathscr{F}(x) - \mathscr{G})} Y(F) + dD/(\log n)^{2a/3}. \tag{5.6}$$

The variables $Y(F)$ are still not independent, however, their dependence is very limited for $F, F' \in \mathscr{F}^{(i)}$. Namely, their correlation is caused only by sets $H \in \mathscr{F}$ such that there exist $y, y' \in H$ with $y \in F$, $y' \in F'$.

Thus for a given star $\mathscr{F}^{(i)}$ the number of such 'bad' H counted with multiplicity is upperbounded by

$$\binom{t}{2}(d-1)^2 \frac{D}{(\log n)^a} \leq (d-1)^2 D/2(\log n)^{a/3}.$$

Let $\mathscr{H}^{(i)}$ be their collection.

Set $\mathscr{H} = \bigcup_i \mathscr{H}^{(i)}$ and let $\mu(H)$ denote the multiplicity of H. Thus we have

$$\sum_{H \in \mathscr{H}} \mu(H) < (d-1)^2 D^2/(\log n)^{2a/3}. \tag{5.7}$$

We claim that

$$\mu(H) < \binom{d}{2} D/(\log n)^a.$$

In fact, denote by $\mathscr{N}(H)$ the collection

$$\mathscr{N}(H) = \{F \in \mathscr{F}(x) \colon (F - \{x\}) \cap H \neq \emptyset\}.$$

Clearly $|\mathscr{N}(H)| < dD/(\log n)^a$ holds.

On the other hand for $F, F' \in (\mathscr{N}(H) \cap \mathscr{F}^{(i)})$ $(F - \{x\}) \cap H \neq (F' - \{x\}) \cap H$ (since $F \cap F' = \{x\}$). Thus each $F \in \mathscr{N}(H)$ adds at most $d-1$ to the multiplicity of H. This implies the claim.

Let us define again auxiliary random variables. Suppose $F \in \mathscr{F}^{(i)}$, we set

$$Y(F)^* = \begin{cases} 1, & \text{if } F \cap R = \emptyset \text{ for all } R \in (\mathscr{R} - \mathscr{H}^{(i)}), \\ 0, & \text{otherwise}. \end{cases} \tag{5.8}$$

Clearly $Y(F)^* \geq Y(F)$, the $Y(F)^*$ are totally independent for $F \in \mathscr{F}^{(i)}$. Also $\sum_{F \in \mathscr{F}(x)} (Y(F)^* - Y(F)) \leq \sum_{H \in \mathscr{H} \cap \mathscr{R}} \mu(H)$.

Let us define $\mathscr{H}_i = \{H \in \mathscr{H} \colon \mu(H) = i\}$ and $l = \max_{H \in \mathscr{H}} \mu(H)$, Thus

$$l < \binom{d}{2} D/(\log n)^a$$

holds.

To estimate the RHS of (5.8) we write

$$\sum_{H \in \mathscr{H} \cap \mathscr{R}} \mu(H) = \sum_{1 \leq i \leq l} i|\mathscr{H}_i \cap \mathscr{R}| = \sum_{1 \leq i \leq l} |(\mathscr{H}_i \cup \cdots \cup \mathscr{H}_l) \cap \mathscr{R}|. \tag{5.9}$$

In view of (5.7) we have

$$|\mathcal{H}_i \cup \cdots \cup \mathcal{H}_l| < \frac{(d-1)^2 D^2}{i(\log n)^{2a/3}}.$$

Using $\varepsilon \leqslant 1$, the independence of the choice of $R \in \mathcal{R}$ and Proposition 5.1 we infer

$$|(\mathcal{H}_i \cup \cdots \cup \mathcal{H}_l) \cap \mathcal{R}| < \frac{e(\log n)^{a/3}(d-1)^2 D\varepsilon}{i(\log n)^{2a/3}} < \frac{e(d-1)^2 D}{i(\log n)^{a/3}}. \tag{5.10}$$

holds with probability at least 1-2 $\cdot \log n^{-(\log n)^{a/3}} > 1 - 1/n^2 l$ if $n > n_0(\varepsilon, a, d)$.

Thus (5.10) holds simultaneously for all $1 \leqslant i \leqslant l < \binom{d}{2} D/(\log n)^a$ with probability at least $1 - 1/n^2$ for $n > n_0$. Substituting (5.10) into (5.9) we infer

$$\sum_{H \in \mathcal{H} \cap \mathcal{R}} \mu(H) < \left(\sum_{1 \leqslant i \leqslant l} \frac{1}{i} \right) \frac{e(d-1)^2}{(\log n)^{a/3}} D < \frac{3(d-1)^2 D}{(\log n)^{(a-3)/3}}$$

holds with probability at least $1 - 1/n^2$.

Using (5.5) we see that $p(Y(F)^* = 1) \sim_{3\rho} \mathrm{e}^{-(d-1)\varepsilon}$ for $n > n_0(\rho, \varepsilon)$. Thus

$$\sum_{F \in \mathcal{F}^{(i)}} Y(F)^* \underset{4\rho}{\sim} \mathrm{e}^{-(d-1)\varepsilon} |\mathcal{F}^{(i)}|$$

holds with probability greater than $1 - 1/n^2$. Combining this with the upper bound for (5.8) and with (5.6) we infer that

$$|\tilde{\mathcal{F}}(x)| \underset{6\rho}{\sim} \mathrm{e}^{-(d-1)\varepsilon} D$$

holds with probability at least $1 - 2/n^2$.

Thus with probability greater than $1 - 2/n^2$ the same holds simultaneously for all $\chi \in \tilde{X}$. Thus (i) holds with $\tilde{D} = D\,\mathrm{e}^{-(d-1)\varepsilon}$, $\tilde{\rho} = 6\rho$. Since $\tilde{D}$ did not decrease drastically, (ii) is automatic with $\tilde{a} = a - o(1)$ for n sufficiently large.

Proof of Theorem 1.1. Let $\mathcal{H}$ be a uniform hypergraph satisfying the assumptions of Theorem 1.1. Setting $\mathcal{F} = \mathcal{H}$ and applying Lemma 5.2 we obtain $\mathcal{R} = \mathcal{R}_1$ and $\tilde{\mathcal{F}} = \mathcal{F}_1$ with the properties from Lemma 5.2. We repeat this and apply again Lemma 5.2 (now to $\mathcal{F}_1$) to obtain $\mathcal{R}_2$ and $\mathcal{F}_2$. Repeating this procedure t-times (t large— this will be specified later) we obtain a sequence $\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_t$ and $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_t$ satisfying $|\mathcal{R}_{j+1}| \sim_{\varepsilon/3} \varepsilon n/d\, \mathrm{e}^{-j\varepsilon}$. Hence, after t steps we covered all but $\mathrm{e}^{-\varepsilon t} n$ points with $\sum_{j=0}^{t-1} \varepsilon n/d\, \mathrm{e}^{-j\varepsilon}$ edges. Pick for each of these uncovered points one edge containing this point. Thus we get

$$\sum_{j=0}^{t-1} \varepsilon \frac{n}{d} \mathrm{e}^{-j\varepsilon} + \mathrm{e}^{-\varepsilon t} n \leqslant (1+\varepsilon) \frac{n}{d}, \quad \text{for } t \geqslant t_0(\varepsilon, d),$$

edges covering all points.

In fact, to be more exact, given ε one chooses first t, $t = t(\varepsilon)$, to ensure the above inequality. Then choose ρ so small that Lemma 5.2 will work even in the t the steps, i.e. with $\rho' = 6^{t-1}\rho$, also $2\rho' < \varepsilon/3$.

6. Proof of Theorems 2.4 and 3.3

In order to prove the theorems from Sections 2, 3 and 4 we need to verify the assumptions of Theorem 1.1. This is usually easy and amounts always to the same sort of considerations. Therefore we present here the two 'most difficult' cases only.

PROOF OF THEOREM 2.4. Consider a $|\mathscr{A}|$-uniform hypergraph H the vertices of which are the edges of $K(n_1, n_2, \ldots, n_r)$ and with $\mathscr{A}' \subset K(n_1, n_2, \ldots, n_r)$ forming an edge iff $\mathscr{A}'$ is isomorphic to $\mathscr{A}$. Then clearly conditions of Theorem 1.1 are satisfied. Applying Theorem 1.1 we get the required family of copies of $\mathscr{A}$.

PROOF OF THEOREM 3.3. Let l be the number of vertices of $\mathscr{B}$ and let V be the vertex set of $\mathscr{B}$. We may suppose that

$$r = \binom{l}{d}, \qquad \mathscr{B} = \binom{V}{d} \quad \text{and} \quad |\mathscr{B}_1| = |\mathscr{B}_2| = \cdots = |\mathscr{B}_r| = 1$$

for otherwise we subdivide each $\mathscr{B}_i$ and $[V]^d - \mathscr{B}$ into singletons and solving the new problem we solve the required as well. Thus it suffices to prove that

$$\pi_c(H_d(n, p, \underbrace{(1/r, 1/r, \ldots, 1/r)}_{r=\binom{l}{d} \text{ times}}), \mathscr{B}) = (1 - o(1))\binom{n}{d}p \Big/ \binom{l}{d}$$

where $\mathscr{B}$ is K_l^d with edges coloured by colours $1, 2, \ldots, \binom{l}{d}$.

Consider now the hypergraph H the vertex set of which equals to the edge set of $H_d(n, p)$. Edges of H are formed by copies of $\mathscr{B}$ in $H_d(n, p)$ i.e. an $\binom{l}{d}$-tuple of edges of $H_d(n, p)$ forms an edge of H if it is the edge set of some complete graph K' and moreover if there exists an isomorphism $\varphi: \mathscr{B} \to K'$ which preserves colours. It is now a matter of routine to verify that H satisfies the assumptions of Theorem 1.1.

7. COVERING RANDOM GRAPHS WITH NONCONSTANT PROBABILITIES

We can extend previous results about random graphs and derive corresponding results where edges are chosen with nonconstant probabilities.

THEOREM 7.1. *Let $\varepsilon > 0$ and $p = n^{\varepsilon - 1/2}$, then*

$$\pi(\mathscr{G}(n, p), K_3) = (1 - o(1))\frac{n^{3/2+\varepsilon}}{6}$$

and more generally

THEOREM 7.2. *Let $\varepsilon > 0$ and $p = n^{\varepsilon - 2/k+1}$, then*

$$\pi(\mathscr{G}(n, p), K_k) = (1 - o(1))\binom{n}{2}p \Big/ \binom{k}{2}.$$

We are omitting the proof of Theorem 7.2 and present a simple proof of the weaker Theorem 7.1 only.

PROOF OF THEOREM 7.1. This is rather easy. It is sufficient to show that the number of triangles containing a given edge does not depend essentially on the choice of the edge. To estimate this, call a pair x, y of vertices of $\mathscr{G}(n, p)$ bad if it is contained in more than $(1+\delta)n^{2\varepsilon}$ or less than $(1-\delta)n^{2\varepsilon}$ triangles, respectively. Thus

$$\begin{aligned}\operatorname{Prob}[\exists \text{ bad } x, y \text{ in } \mathscr{G}(n, p)] &< \binom{n}{2} \sum_{|i - n^{2\varepsilon}| > \delta n^{2\varepsilon}} \binom{n-2}{i}\left(\frac{n^{2\varepsilon}}{n}\right)^i \left(1 - \frac{n^{2\varepsilon}}{n}\right)^{n-2-i} \\ &< n^2 c_\delta^{-n^{2\varepsilon}} = o(1)\end{aligned}$$

and therefore the assumptions of Theorem 1.1 are fulfilled.

References

1. M. Ajtai, J. Komlós, V. Rödl and E. Szemerédi, *On coverings of random graphs*, *Commentationes Mathematicae Universitatis Carolinae* **23**, 1 (1982), 193–198.
2. P. Erdös, H. Hanani, *On a limit theorem in combinatorial analysis*, *Publ. Math. Debrecen* **10** (1963), 10–13.
3. L. Lovász, *On the ratio of optimal integral and fractional covers*, *Discrete Math.* **13** (1975), 383–390.
4. D. K. Ray-Chaudhuri, *Unsolved problems*, Hypergraph Seminar (Proc. First Working Seminar, Ohio, State University of Columbus, Ohio, 1972) *Lecture Notes in Mathematics*, vol. 411 (C. Berge and D. K. Ray-Chaudhuri, eds) 1974, pp. 278–287.
5. V. Rödl, *On a packing and covering problem*, *Europ. J. Combinatorics* **6** (1985), 69–78.
6. R. M. Wilson, *Decomposition of complete graphs into subgraphs isomorphic to a given graph*, Proc. 5th British Combinatorial Conf. (Aberdeen) 1975), *Congressus Numeratium XV*, Utilitatis Math. Publ. Inc., Winnipeg, 1976, pp. 647–659.
7. R. M. Wilson, *An existence theory for pairwise balanced designs*, III, *J. Combin. Theory* **18** (1975), 71–79.

Received 2 August 1984 and in revised form 11 February 1985

P. Frankl
CNRS, Paris, France

and

V. Rödl
FJFI, CVUT, Prague, Czechoslovakia