COMBINATORICA
Bolyai Society – Springer-Verlag

COMBINATORICA **21** (2) (2001) 171–197

VERTEX COVERS BY EDGE DISJOINT CLIQUES

TOM BOHMAN*, ALAN FRIEZE†, MIKLÓS RUSZINKÓ‡,
LUBOS THOMA§

Dedicated to the memory of Paul Erdős

Received October 31, 1999

Let H be a simple graph having no isolated vertices. An (H,k)-*vertex-cover* of a simple graph $G=(V,E)$ is a collection $H_1,\dots,H_r$ of subgraphs of G satisfying

1. $H_i \cong H$, for all $i=1,\dots,r$,
2. $\cup_{i=1}^r V(H_i)=V$,
3. $E(H_i)\cap E(H_j)=\emptyset$, for all $i\neq j$, and
4. each $v\in V$ is in at most k of the H_i.

We consider the existence of such vertex covers when H is a complete graph, $K_t, t\geq 3$, in the context of extremal and random graphs.

1. Introduction

Let H be a simple graph having no isolated vertices. For the purposes of this discussion we say that the simple graph $G=(V,E)$ has property $\mathcal{C}_{H,k}$ if there is a collection $H_1,\dots,H_r$ of subgraphs of G satisfying

P1. $H_i \cong H$, for all $i=1,\dots,r$,
P2. $\cup_{i=1}^r V(H_i)=V$,
P3. $E(H_i)\cap E(H_j)=\emptyset$, for all $i\neq j$, and

Mathematics Subject Classification (2000): 05C80, 05C35

* Supported in part by NSF grant DMS-9627408.
† Supported in part by NSF grant CCR-9530974.
‡ Supported in part by OTKA Grants T 030059 and T 29074, FKFP 0607/1999 and by the Bolyai Foundation.
§ Supported in part by NSF grant DMS-9970622.

P4. each $v \in V$ is in at most k of the H_i.

We call the family $\{H_1,\dots,H_r\}$ an (H,k)-vertex-cover of G. Thus when $k=1$ we ask for the existence of a partition of V into *vertex disjoint* copies of H i.e. the existence of an H-*factor.* In this case we assume the necessary divisibility condition, i.e. that $|V(H)|$ divides $|V|$. We study this property when G is a random graph and also when G is extremal w.r.t. minimum degree. We will mainly focus on the case where H is a complete graph K_t and denote our property by $\mathcal{C}_{t,k}$.

Random graphs

The precise threshold for the occurrence of $\mathcal{C}_{2,1}$ i.e. the existence of a perfect matching was found by Erdős and Rényi [8] as part of a series of papers which laid the foundations of the theory of random graphs. The precise threshold for the occurrence of $\mathcal{C}_{3,1}$ i.e. the existence of a vertex partition into triangles remains as one of the most challenging problems in this area (see, for example, the Appendix by Erdős to the monograph by Alon and Spencer [1]).

The thresholds for H-factors have been studied for example by Ruciński [17], by Alon and Yuster [3] and by Krivelevich [13]. For a graph H, let

$$m_1(H) = \max\left(\frac{|E(H')|}{|V(H')|-1}\right)$$

where the maximum is taken over all subgraphs H' of the graph H with at least two vertices. Ruciński showed that the probability $p(n)=O(n^{-1/m_1(H)})$ is a sharp threshold for the property $\mathcal{C}_{H,1}$ for any graph H such that $m_1(H) > \delta(H)$ where $\delta(H)$ stands, as usual, for the minimum degree of the graph H [17]. Note that complete graphs do not satisfy this condition, and therefore the first interesting open case is $H = K_3$. Alon and Yuster showed that $p(n)=O(n^{-1/m_1(H)})$ is a sharp threshold for the property $\mathcal{C}_{H,1}$ for a more general class of graphs that does not contain the complete graphs [3]. In [13], Krivelevich showed that the probability $p(n) = O(n^{-3/5})$ is enough for the random graph to have a K_3-factor **whp**[1] and, in general, if $p(n) = O(n^{-2t/(t-1)(t+2)})$ then the random graph $G_{n,p}$ contains a K_t-factor **whp** (provided t divides n).

An obvious necessary condition for the existence of a (K_t,k)-vertex-cover is that every vertex be incident with at least one copy of K_t.

[1] A sequence of events $\mathcal{E}_n$ occurs *with high probability,* **whp**, if $\mathbf{Pr}(\mathcal{E}_n)=1-o(1)$.

Theorem 1. *Let* $m=\binom{n}{2}((t-1)!(\log n+c_n))^{1/\binom{t}{2}}n^{-2/t}$. *Then*

$$\lim_{n\to\infty}\mathbf{Pr}(G_{n,m}\text{ contains a }(K_t,2)\text{-vertex-cover})=\begin{cases}0 & c_n\to-\infty\\ e^{-e^{-c}} & c_n\to c\\ 1 & c_n\to\infty\end{cases}$$

(Here, $G_{n,m}$ stands for the probability space over the set of all graphs on n vertices and with m edges endowed with the uniform probability measure.) We will prove this as a consequence of the slightly stronger hitting time version. We consider the graph process $G_m=([n],E_m), m=0,1,\dots,\binom{n}{2}$, where $E_0=\emptyset$ and G_m is obtained from G_{m-1} by choosing e_m randomly from $\binom{[n]}{2}\setminus E_{m-1}$ and putting $E_m=E_{m-1}\cup\{e_m\}$. We define two *hitting times*:

$$\tau_1=\tau_1(t)=\min\{m:\ \text{Every } v\in[n] \text{ is contained in a copy of } K_t \text{ in } G_m\},$$
$$\tau_2=\tau_2(t)=\min\{m:\ G_m \text{ contains a } (K_t,2)\text{-vertex-cover}\}.$$

Theorem 2. *For every fixed* $t\geq 3$,

$$\lim_{n\to\infty}\mathbf{Pr}(\tau_1=\tau_2)=1.$$

Moreover, there exists **whp** *a* $(K_t,2)$*-vertex-cover of* G_{τ_2} *containing* $(1+o(1))\frac{n}{t}$ *copies of* K_t.

Remark 1. In fact, our proof of Theorem 2 implies that G_{τ_2} possesses **whp** a $(K_t,2)$-vertex-cover containing at most $\left(\frac{1}{t}+\frac{1}{(\log n)^{1/t}}\right)n$ copies of K_t.

Remark 2. Theorem 2 lends weight to the common conjecture that the threshold for a K_t-factor is m of Theorem 1.

We prove Theorem 2 in Section 2 and show how Theorem 1 follows from Theorem 2 in Section 3.

Extremal graphs

For a graph G on n vertices what is the smallest minimum degree that insures G has $\mathcal{C}_{t,k}$? For $t\geq 3$ and $k\geq 2$ let

$$f(n,t,k)=\max\{d:\exists G \text{ such that } \delta(G)=d, |V(G)|=n \text{ and } G\notin\mathcal{C}_{t,k}\}.$$

We will assume that n is large with respect to t, but k can be arbitrarily large. The smallest minimum degree that guarantees a K_t-factor (this would be, up to divisibility considerations, $f(n,t,1)+1$) was established in the following deep theorem of Hajnal and Szemerédi [10].

Theorem 3 (Hajnal, Szemerédi). *If* $|V(G)|=n$ *and* $\delta(G)\geq(1-\frac{1}{t})n$ *then* G *contains* $\lfloor n/t \rfloor$ *vertex-disjoint copies of* K_t.

Our central result in this section is the following:

Theorem 4. *Let* $t\geq 3$, $k\geq 2$, $n\geq 6t^2-4t$ *and*

$$n=q[(t-1)k+1]+r \text{ where } 1\leq r\leq (t-1)k+1.$$

Then

$$n-qk-\left\lceil \frac{r}{t-1}\right\rceil \leq f(n,t,k)\leq n-qk-\left\lceil \frac{r}{t-1}\right\rceil+1\,.$$

Note that it follows from Theorem 4 that

$$f(n,t,k)=\left\lfloor \frac{[(t-2)k+1]n}{(t-1)k+1}\right\rfloor+c \qquad (1)$$

where $c\in\{0,1,2\}$. It is tempting to believe that $f(n,t,k)$ equals the lower bound given in Theorem 4. This is not the case in general.

Theorem 5. *Let* $n\geq 6$ *and* $k\geq (n-1)/2$.

$$f(n,3,k)=\left\lceil \frac{n}{2}\right\rceil .$$

Note that the value of $f(n,3,k)$ given in Theorem 5 equals the lower bound in Theorem 4 for n even, but equals the upper bound for n odd. (Here $q=0$ and $r=n$).

For H a simple graph with no isolated vertices and G an arbitrary graph an (H,∞)-*vertex-cover* of G is a collection $H_1,\dots,H_r$ of subgraphs of G satisfying P1, P2 and P3. Thus, G has an (H,∞)-vertex-cover if and only if there exists a k such that G has a (H,k)-vertex-cover. To motivate our results on (H,∞)-vertex-covers, we recall the following well-known extension of Theorem 3. Given an arbitrary graph H, Komlós, Sárközy and Szemerédi [15] showed that there is a constant c (depending only on the graph H) such that if $\delta(G)\geq\left(1-\frac{1}{\chi(H)}\right)n$ for a graph G on n vertices, then there is a union of vertex-disjoint copies of H covering all but at most c vertices of G. Weakening the condition on $\delta(G)$ we show in the following theorem the existence of (H,∞)-vertex-covers for graphs H having the property that there is a vertex u of H such that $\chi(H\setminus\{u\})=\chi(H)-1\geq 3$.

Theorem 6. *Let* H *be a graph such that* $\chi(H)\geq 4$ *and such that there is a vertex* u *of* H *with the property that* $\chi(H\setminus\{u\})=\chi(H)-1$. *Then for every* $\epsilon>0$ *and every graph* G *on* n *vertices, if* $\delta(G)\geq\left(1-\frac{1}{\chi(H)-1}+\epsilon\right)n$, *then* G *has an* (H,∞)*-vertex-cover provided* n *is large enough.*

Theorems 4, 5 and 6 are proved in Section 4.

2. Proof of Theorem 2

In this section we will use the following Chernoff bounds on the tails of the binomial random variable $B(n,p)$. For $0 \le \epsilon \le 1$ and $\theta > 0$

$$\mathbf{Pr}(B(n,p) \le (1-\epsilon)np) \le e^{-\epsilon^2 np/2} \tag{2}$$

$$\mathbf{Pr}(B(n,p) \ge (1+\epsilon)np) \le e^{-\epsilon^2 np/3} \tag{3}$$

$$\mathbf{Pr}(B(n,p) \ge \theta np) \le (e/\theta)^{\theta np} \tag{4}$$

All Lemmas introduced in this section will be proven in the subsections that follow.

Let $t \ge 3$ be fixed. We construct a $(K_t, 2)$-vertex-cover in G_m by dividing our graph process into 3 phases and using edges from different phases for different purposes. Before describing the phases, we make some preliminary definitions and the observation that we may restrict our attention to G_m where m lies in a small interval. Let $\alpha, \beta > 0$ be constants such that

$$\beta^{\binom{t}{2}} > 19/20 \text{ and } \alpha + \beta < 1,$$

and let

$$m_a = \alpha \binom{n}{2} \left((t-1)! \log n\right)^{1/\binom{t}{2}} n^{-2/t}, \text{ and}$$
$$m_b = \beta \binom{n}{2} \left((t-1)! \log n\right)^{1/\binom{t}{2}} n^{-2/t}.$$

Furthermore, for $i = 0, 1$ let

$$m_i = \binom{n}{2} \left((t-1)!(\log n - (1-2i)\log\log n)\right)^{1/\binom{t}{2}} n^{-2/t}.$$

Lemma 1.

$$\mathbf{Pr}(\tau_1 \notin [m_0, m_1]) = o(1) .$$

We will use the term 'a collection of K_t's' in the graph G, for a family $\mathcal{A} \subseteq \binom{V(G)}{t}$ such that $G[S]$ is complete for all $S \in \mathcal{A}$. For such a collection $\mathcal{A}$ we set

$$V(\mathcal{A}) = \bigcup_{S \in \mathcal{A}} S \quad \text{and} \quad E(\mathcal{A}) = \bigcup_{S \in \mathcal{A}} \binom{S}{2}, \tag{5}$$

say $\mathcal{A}$ 'covers' a vertex v if $v \in V(\mathcal{A})$, and say $\mathcal{A}$ 'covers' a set of vertices T if $T \subseteq V(\mathcal{A})$.

We are now ready to describe the 3 phases. In the first phase we simply choose m_a edges uniformly at random, producing the graph $G^1 = ([n], E^1)$. Thus,

$$G^1 = G_{n,m_a}.$$

In the second phase we form the graph $G^2=([n],E^2)$ by choosing m_b edges uniformly at random. This is done independently of phase 1 and without knowledge of which edges were placed in phase 1. Thus,

$$G^2 = G_{n,m_b},$$

and a particular edge may appear in both G^1 and G^2. Let $F=E^1\cup E^2$ and $m_{-1}=|F|$. The third phase is the graph process $H_i=([n],F_i), i=m_{-1},\dots,m_1$ where $F_{m_{-1}}=F$ and F_{i+1} is the union of F_i and the set containing a single edge chosen uniformly at random from $\binom{n}{2}\setminus F_i$. In other words, in the third phase we start with the collection of edges generated in phases 1 and 2 and then add new edges one at time until m_1 edges have been placed. Note that for $m_a+m_b\le i\le m_1$ the graphs G_i and H_i are identically distributed.

We henceforth assume

$$m_a + m_b \le \tau_1 \le m_1. \tag{6}$$

We will show that

$$(6) \;\Rightarrow\; \textbf{whp } G_{\tau_1} \text{ has a } (K_t,2)\text{-vertex-cover.} \tag{7}$$

We stress that we do not condition on the value τ_1 in any way (i.e. we work with the probability space described above); rather, we give an argument that depends only on the properties of the graphs G^1 and G^2, which are subgraphs of G_{τ_1}, the properties of G_{m_1}, which contains G_{τ_1}, and the fact that every vertex in G_{τ_1} is contained in a copy of K_t. In fact, what follows actually shows (with only trivial modifications) that **whp** every graph G_m in the sequence $G_{m_a+m_b},\dots,G_{m_1}$ has a $(K_t,2)$-vertex-cover that covers all vertices that are contained in a copy of K_t. Clearly, Theorem 2 follows from (7) and Lemma 1.

How do we construct the $(K_t,2)$-vertex-cover? We first use the phase one edges to greedily cover as many vertices as possible with vertex disjoint K_t's. Let Ξ be an arbitrary maximal collection of vertex disjoint K_t's in G^1, $X\subseteq[n]$ be the set of vertices not covered by Ξ, and

$$r = \left\lceil \frac{n}{(\log n)^{1/t}} \right\rceil .$$

Remark 3. We can easily randomize this choice of K_t's so that X is a random $|X|$-subset of $[n]$. This will be used in the proof of Lemma 4.

Lemma 2. *Let $G=G_{n,m_a}$.*

$$\mathbf{Pr}\left(\exists R\subset[n] \text{ such that } |R|=r \text{ and } G[R] \text{ contains no } K_t\text{'s}\right)=o(1).$$

It follows from Lemma 2 that **whp**

$$|X| \le r. \tag{8}$$

In other words, after using only a small fraction of the edges in G_{τ_1}, only $o(n)$ vertices remain to be covered. We will use the phase 2 edges (as well as a handful of the phase 1 and phase 3 edges) to form a vertex disjoint collection of K_t's that covers X but does not use any edge in $E(\Xi)$.

Before describing the vertex disjoint collection of K_t's that covers X, we make further definitions and preliminary observations. Our first observation concerns the random graph process G_{m_1} alone. Let $\nu_3=4$, $\nu_4=3$ and $\nu_i=2$ for $i=5,6,\ldots$. We define a *cluster* to be a collection $\mathcal{C}=\{S_1,\ldots,S_l\}$ of K_t's in G_{m_1} such that $l\le 2\nu_t$

$$\begin{aligned} \kappa_i \ge 1 \quad &\text{for} \quad i=2,\ldots,l \\ \kappa_i = t \quad \Rightarrow \quad \kappa_{i-1}&=1 \quad \wedge \quad |S_i\cap S_{i-1}|\ge 2 \\ \text{and} \quad &|\{i:\kappa_i\ne 1\}| = \nu_t \end{aligned}$$

where

$$\kappa_i = \left| S_i \cap \left(\bigcup_{j=1}^{i-1} S_j \right) \right| \qquad \text{for} \qquad i=2,\ldots,l.$$

Note the order of the K_t's in a cluster is important: we think of a cluster as being 'built' one K_t at a time. Roughly speaking, a cluster is a very small collection of K_t's that have many or large pairwise intersections.

Lemma 3.

$$\mathbf{Pr}(G_{m_1} \textit{ contains a cluster}) = o(1).$$

We now turn our attention to the graph G^2. For $v\in[n]$ let Υ_v be the collection of K_t's in G^2 that contain v; to be precise,

$$\Upsilon_v = \left\{ S \in \binom{[n]}{t} : v\in S \text{ and } \binom{S}{2} \subseteq E^2 \right\}.$$

Since Υ_v depends only on the graph G^2 while X is small and depends only on the graph G^1, it is usually the case that no $V(\Upsilon_v)$ contains many members of X— see (5) for explanation of notation. To make this statement precise, we let

$$q = \left\lceil \frac{\log n}{\log\log\log n} \right\rceil.$$

Lemma 4.

$$\mathbf{Pr}(\exists v \in [n] \textit{ such that } |V(\Upsilon_v) \cap X| > q) = o(1).$$

We say that

$$v \in [n] \text{ is } \textit{large} \text{ if } |\Upsilon_v| \geq \tfrac{\log n}{20}, \text{ and}$$
$$v \in [n] \text{ is } \textit{small} \text{ if } |\Upsilon_v| < \tfrac{\log n}{20}.$$

With high probability the small vertices are, with respect to connections via K_t's, far apart. To make this statement precise, we define a *chain* to be a pair u, v of distinct small vertices and a collection $S_1, S_2, S_3, S_4 \in \binom{[n]}{t}$ of (not necessarily distinct) sets such that $u \in S_1$, $v \in S_4$,

$$S_1 \cap S_2, S_2 \cap S_3, S_3 \cap S_4 \neq \emptyset, \qquad \text{and} \qquad \binom{S_i}{2} \subseteq E(G_{m_1}) \text{ for } i = 1, 2, 3, 4.$$

Lemma 5.

$$\mathbf{Pr}(G_{m_1} \textit{ contains a chain}) = o(1).$$

We also show that **whp** every K_t containing a small vertex intersects every other K_t in at most one vertex. A *link* is a small vertex $u \in [n]$ and distinct $S_1, S_2 \in \binom{[n]}{t}$ such that $u \in S_1$, $|S_1 \cap S_2| \geq 2$, and $\binom{S_1}{2}, \binom{S_2}{2} \subseteq E(G_{m_1})$.

Lemma 6.

$$\mathbf{Pr}(G_{m_1} \textit{ contains a link}) = o(1).$$

Finally, let

$$X_1 = \{v \in X : v \text{ is small}\},$$
$$X_2 = \{v \in X : v \text{ is large}\}, \text{ and}$$
$$\Phi = \left\{ S \in \tbinom{[n]}{t} : \tbinom{S}{2} \subseteq E(G_{m_1}) \text{ and } S \cap X_1 \neq \emptyset \right\}.$$

We are now prepared to describe the remainder of the $(K_t, 2)$-vertex-cover. We henceforth assume (kUz8),

(9) G_{m_1} does not contain a cluster,

(10) $\forall v \in [n] \quad |V(\Upsilon_v) \cap X| \leq q$,

(11) G_{m_1} does not contain a chain,

(12) G_{m_1} does not contain a link,

and that n is sufficiently large (in a sense that is made clear below). We will show that there exist collections Ξ_1 and Ξ_2 of vertex disjoint K_t's in G_m such that $\Xi_1 \cup \Xi_2$ covers $X_1 \cup X_2$ and

(13) $V(\Xi_1) \cap V(\Xi_2) = \emptyset$ and $E(\Xi) \cap E(\Xi_1 \cup \Xi_2) = \emptyset$.

If follows from Lemmas 1, 2, 3, 4, 5 and 6 that (13) implies Theorem 2.

We cover X_1 in a rather crude way. Let Ξ_1 be an arbitrary collection of K_t's in G_{τ_1} that covers X_1. Note that the collection Ξ_1 uses edges from all 3 phases and that we make use of the fact that every vertex is contained in some K_t in G_{τ_1} when forming Ξ_1. By (11), Ξ_1 is vertex disjoint.

We cover

$$X_2' := X_2 \setminus V(\Xi_1)$$

in a more sophisticated way: we apply the Lovász Local Lemma. We first 'trim' the Υ_v's. For $v \in X_2'$ let Υ_v' be the collection of sets in $S \in \Upsilon_v$ such that

$$S \cap X = \{v\}$$
$$T \in \binom{[n]}{t} \wedge \binom{T}{2} \subseteq E(G_{m_1}) \Rightarrow |S \cap T| \leq 1, \qquad \text{and} \tag{14}$$
$$S \cap V(\Phi) \subseteq \{v\}.$$

In words, we get Υ_v' from Υ_v by throwing away those sets in Υ_v that contain an element of X other than v, intersect another K_t in more than one vertex, or contain a vertex of a K_t that contains a small vertex. By (10) there are at most q sets in Υ_v that contain an element of X other than v. We will show:

(15) There are $\leq \binom{2\nu_t t}{t}$ sets in Υ_v that intersect another K_t in ≥ 2 vertices.

By (11) at most one set in Υ_v intersects $V(\Phi)$. Therefore, we may choose $\Theta_v \subseteq \Upsilon_v'$ such that

$$|\Theta_v| = \left\lceil \frac{\log n}{21} \right\rceil \qquad \text{for all} \qquad v \in X_2'. \tag{16}$$

Proof of (15). Let $\hat{\Upsilon}_v$ denote the collection of K_t's in Υ_v which intersect another K_t in more than one vertex. Let $B = V(\hat{\Upsilon}_v)$. We construct copies $X_1, X_2, \ldots, X_l$ of K_t in G_{m_1} as follows: Suppose we have constructed $X_1, X_2, \ldots, X_k$. Either (i) $B \subseteq V_k = V(X_1 \cup X_2 \cup \cdots \cup X_k)$ or (ii) $B \not\subseteq V_k$. In case (ii) choose $X_{k+1} \in \Upsilon_v$ which is not contained in V_k. If $|X_{k+1} \cap V_k| = 1$ then choose X_{k+2} where $|X_{k+2} \cap X_{k+1}| \geq 2$. If this process continues for ν_t iterations we will have produced a cluster. Thus $l \leq 2\nu_t$ and $|B| \leq 2t\nu_t$, which implies (15).

Now, consider the probability space in which each $v \in X_2'$ chooses $S_v \in \Theta_v$ uniformly at random and independently of the other vertices. For $u \neq v \in X_2', S \in \Theta_u$ and $T \in \Theta_v$ such that $S \cap T \neq \emptyset$ let $A_{u,v,S,T}$ be the event that $S_u = S$ and $S_v = T$. These are the 'bad' events in our application of the Lovász Local Lemma. Clearly,

$$\mathbf{Pr}(A_{u,v,S,T}) = \frac{1}{|\Theta_v||\Theta_u|} \leq \left(\frac{21}{\log n}\right)^2 =: p. \tag{17}$$

Events A_{u_1,u_2,S_1,S_2} and A_{v_1,v_2,T_1,T_2} are dependent if and only if

$$\{u_1,u_2\}\cap\{v_1,v_2\}\neq\emptyset.$$

Thus, the degree in the dependency graph is bounded above by

$$\begin{aligned} d &:= 2\max_{u\in X_2'}\sum_{S\in\Theta_u}\sum_{v\in X_2'}|\{T\in\Theta_v : S\cap T\neq\emptyset\}| \\ &\leq 2\max_{u\in X_2'}\sum_{w\in V(\Theta_u)}|\Upsilon_w\cap X| \\ &\leq 2tq\left\lceil\frac{\log n}{21}\right\rceil \qquad \text{by (10)} \\ &\leq \frac{t(\log n)^2}{10\log\log\log n}. \end{aligned} \tag{18}$$

It follows from (17) and (18) that

$$pd\leq\frac{45t}{\log\log\log n}=o(1).$$

Thus, for n sufficiently large, it follows from the Lovász Local Lemma that there exists a vertex disjoint collection Ξ_2 of K_t's in G^2 that covers X_2' but covers no vertex in $V(\Xi_1)$.

It remains to show that

$$E(\Xi)\cap E(\Xi_1\cup\Xi_2)=\emptyset.$$

This is an immediate consequence of (12) and (14). We have established (13) and completed the proof. ∎

2.1. Proof of Lemma 1

Let $p_i=m_i/\binom{n}{2}$ for $i=0,1$. We use a result of Ruciński [17] and Spencer [18]. We quote Theorem 3.22(i) of Janson, Łuczak and Ruciński [12], after specialising to K_t: Let $\mathcal{C}_t$ be the property that every vertex of a graph is contained in a copy of K_t.

Theorem 7. *Let* $p=((t-1)!(\log n+c_n))^{1/\binom{t}{2}}n^{-2/t}$. *Then*

$$\lim_{n\to\infty}\mathbf{Pr}(G_{n,p}\in\mathcal{C}_t)=\begin{cases}0 & c_n\to-\infty\\ e^{-e^{-c}} & c_n\to c\\ 1 & c_n\to+\infty\end{cases}$$

The lemma follows immediately. ∎

2.2. Proof of Lemma 2

Let $p_a = m_a/\binom{n}{2}$ and consider the random graph $G = G_{n,p_a}$. For $S \in \binom{[n]}{t}$ let B_S be the event that the induced graph $G[S]$ is complete. For R a fixed subset of $[n]$ such that

$$|R| = r = \left\lceil \frac{n}{(\log n)^{1/t}} \right\rceil$$

let the random variable X_R be the number of copies of K_t contained in R. We clearly have

$$\begin{aligned}
\mu &:= \mathbf{E}[X_R] \\
&= \sum_{S \in \binom{R}{t}} \mathbf{Pr}(B_S) \\
&= \binom{r}{t} p_a^{\binom{t}{2}} \\
&= \binom{r}{t} \frac{\alpha^{\binom{t}{2}}(t-1)!\log n}{n^{t-1}} \\
&= \frac{r^t}{t!}(1 + O(1/r)) \frac{\alpha^{\binom{t}{2}}(t-1)!}{n^{t-1}} \log n \\
&= \Omega(n)
\end{aligned}$$

We apply Janson's inequality (again, we follow the notation of [1]) to show that $\mathbf{Pr}(X_R = 0)$ is small. In order to do so, we must bound the parameter Δ.

$$\begin{aligned}
\Delta &= \sum_{S,T \in \binom{R}{t} : 2 \le |S \cap T| \le t-1} \mathbf{Pr}(B_S \wedge B_T) \\
&= \binom{r}{t} \sum_{i=2}^{t-1} \binom{t}{i} \binom{r-t}{t-i} p_a^{2\binom{t}{2} - \binom{i}{2}} \\
&= \sum_{i=2}^{t-1} O\left(n^{2t - i - \frac{2}{t}\left(2\binom{t}{2} - \binom{i}{2}\right) + o(1)}\right) \\
&= \sum_{i=2}^{t-1} O\left(n^{2 + \frac{i(i-1)}{t} - i + o(1)}\right) \\
&= O\left(n^{2/t + o(1)}\right).
\end{aligned}$$

Thus, Janson's inequality gives

$$\mathbf{Pr}(X_R = 0) \le e^{-c_1 n}$$

where c_1 is a positive constant. Applying the first moment method, we have

$$\begin{aligned}\mathbf{Pr}\left(\bigvee_{R\in\binom{[n]}{r}}\{X_R=0\}\right) &\le \binom{n}{r}e^{-c_1 n}\\ &\le \left(\frac{ne}{r}\right)^r e^{-c_1 n}\\ &= \exp\left\{r\left(1+\frac{\log\log n}{t}\right)-c_1 n\right\}\\ &= o(1)\end{aligned}$$

Since this event is monotone, the same holds for G_{n,m_a}.

2.3. Proof of Lemma 3

Let $\mathcal{C}=\{S_1,\dots,S_l\}$ be a fixed collection of K_t's in K_n such that $l\le 2\nu_t$

$$\kappa_i \ge 1 \qquad \text{for} \qquad i=2,\dots,l$$

$$\kappa_i = t \quad \Rightarrow \quad \kappa_{i-1}=1 \quad \wedge \quad |S_i\cap S_{i-1}|\ge 2 \tag{19}$$

$$\text{and} \qquad |\{i:\kappa_i\ne 1\}|=\nu_t \tag{20}$$

where

$$\kappa_i = \left|S_i\cap\left(\bigcup_{j=1}^{i-1}S_j\right)\right| \qquad \text{for} \qquad i=2,\dots,l.$$

Let $a=|V(\mathcal{C})|$ and $b=|E(\mathcal{C})|$.

Claim 1.

$$a-\frac{2b}{t}<-\frac{1}{t}$$

Proof. We observe this difference as we 'build' the collection $\mathcal{C}$ one K_t at a time. For $j=1,\dots,l$ let $\mathcal{C}_j=\{S_1,\dots,S_j\}$, $a_j=|V(\mathcal{C}_j)|$, $b_j=|E(\mathcal{C}_j)|$ and $d_j=a_j-2b_j/t$. Note that

$$d_1=1,$$

and

$$d_{i+1}-d_i\le(t-\kappa_{i+1})-\frac{2}{t}\left(\binom{t}{2}-\binom{\kappa_{i+1}}{2}\right)=(\kappa_{i+1}-1)\left(\frac{\kappa_{i+1}}{t}-1\right). \tag{21}$$

Thus

$$\begin{aligned} &\kappa_{i+1} = 1 \quad \Rightarrow \quad d_{i+1} - d_i = 0 \\ (22) \qquad \text{and} \quad &2 \le \kappa_{i+1} \le t-1 \quad \Rightarrow \quad d_{i+1} - d_i \le \tfrac{2}{t} - 1. \end{aligned}$$

Furthermore, it follows from (19) that

$$(23) \qquad \kappa_{i+1} = t \quad \Rightarrow \quad b_{i+1} \ge b_i + t - 2 \quad \Rightarrow \quad d_{i+1} - d_i \le -\frac{2(t-2)}{t}.$$

Since (by (22) and (23)) the difference $a_i - 2b_i/t$ decreases by at least $1-2/t$ whenever $\kappa_{i+1} \neq 1$, it follows from (20) that $a - 2b/t = d_l < -1/t$. ∎

Let $\mathcal{E}_i$ be the event that there exists a cluster in G_{m_1} with a vertex set of cardinality i, and let b_i be the minimum number of edges in a cluster on i vertices. With $p_{m_1} = m_1/\binom{n}{2}$ we have

$$\begin{aligned} \mathbf{Pr}(\mathcal{E}_i) &\le \binom{n}{i} 2^{\binom{i}{t}} p_{m_1}^{b_i} \\ &= O\left(n^{i - \frac{2b_i}{t} + o(1)}\right) \\ &= O(n^{-\frac{1}{t} + o(1)}). \end{aligned}$$

The lemma then follows from the fact that the cardinality of the vertex set of a cluster is at most $2\nu_t t$, a constant depending only on t.

2.4. Proof of Lemma 4

We first argue that **whp**

$$(24) \qquad |\Upsilon_v| \le 4 \log n \qquad \text{for all } v \in [n].$$

We can calculate in G_{n,p_b} where $p_b = m_b/N$, $N = \binom{n}{2}$ and then use monotonicity to translate the result to G^2. It follows from (15) that **whp** after removing $O(1)$ K_t's from Υ_v we have a collection $\tilde{\Upsilon}_v$ of K_t's which have no vertex in common but v. So in G_{n,p_b}

$$\begin{aligned} \mathbf{Pr}(|\tilde{\Upsilon}_v| \ge \kappa = 3.9 \log n) &\le \frac{\binom{n-1}{t-1}^{\kappa}}{\kappa!} p_b^{\kappa\binom{t}{2}} \le \frac{(\log n)^{\kappa}}{\kappa!} \le (e/3.9)^{3.9 \log n} \\ &= o(n^{-3/2}). \end{aligned}$$

This verifies (24).

Now fix a vertex v. Then $|V(\Upsilon_v)| < 4t\log n$ and $|X| \le r$. Also, X and $V(\Upsilon_v)$ are chosen independently. It follows from this and Remark 3 that

$$\begin{aligned}\mathbf{Pr}(|V(\Upsilon_v)\cap X| \ge q) &\le \frac{\binom{4t\log n}{q}\binom{n-q}{r-q}}{\binom{n}{r}}\\ &\le \left(\frac{4ter\log n}{qn}\right)^q\\ &\le \left(\frac{4te\log\log\log n\log n}{(\log n)^{(t+1)/t}}\right)^{\log n/\log\log\log n}\\ &= O(n^{-A})\end{aligned}$$

for any constant $A>0$.

There are n choices for v and the lemma follows.

2.5. Proof of Lemmas 5 and 6

Let

$$p = ((t-1)!\log n)^{1/\binom{t}{2}}\, n^{-2/t} \qquad \text{and} \qquad p_{m_1} = \frac{m_1}{\binom{n}{2}}.$$

The main work of this section is the following claim.

Claim 2. *Let $H=(A,B)$ be a fixed graph whose vertex set A is a subset of $[n]$, and let $x,y\in A$ be distinct fixed vertices. If $b:=|B|$ and $a:=|A|\le 4t$ then*

1. $\mathbf{Pr}((x \text{ is small }) \wedge (H\subseteq G_{m_1})) = O(p_{m_1}^b n^{-3/4})$
2. $\mathbf{Pr}((x \text{ and } y \text{ are small}) \wedge (H\subseteq G_{m_1})) = O(p_{m_1}^b n^{-3/2})$

Proof. We only prove 2; the proof of 1 is both similar and easier. Let $\mathcal{R}_x$ be the event that x is small, $\mathcal{R}_y$ be the event that y is small, and let $\mathcal{R}_H$ be the event $B\subseteq E(G_{m_1})$. Furthermore, let

$$N_x = \{v\in[n] : x\sim_{G^2} v\}\setminus A \qquad \text{and} \qquad N_y = \{v\in[n] : y\sim_{G^2} v\}\setminus(A\cup N_x),$$

G_x be the induced graph $G^2[N_x]$, and $G_y = G^2[N_y]$. Finally, let $\epsilon>0$ be a constant such that

$$\beta+\epsilon<1 \text{ and } (\beta-\epsilon)^{\binom{t}{2}} \ge \frac{3}{4}+\frac{1}{20}(1+\log 20). \tag{25}$$

Case 1. $t=3$

We condition on the event that N_x and N_y are of nearly the expected size. Let $\mathcal{R}_1$ be the event that

$$(\beta-\epsilon)np \le |N_x|, |N_y| \le (\beta+\epsilon)np, \tag{26}$$

and $\mathcal{R}_2$ be the event that

$$|E(G_x)|, |E(G_y)| \le \frac{\log n}{20}. \tag{27}$$

We have

$$\mathbf{Pr}(\mathcal{R}_H \wedge \mathcal{R}_x \wedge \mathcal{R}_y) \le \mathbf{Pr}(\mathcal{R}_2 | \mathcal{R}_1 \wedge \mathcal{R}_H)\mathbf{Pr}(\mathcal{R}_H) + \mathbf{Pr}(\overline{\mathcal{R}_1}). \tag{28}$$

Now the Chernoff bounds show that in $G_{n,p_{m_1}}$ we have

$$\mathbf{Pr}(\bar{\mathcal{R}}_1) = O(\exp\{-n^{1-\frac{2}{t}+o(1)}\}), \tag{29}$$

and we can inflate this by $O(n)$ to show the same for G_{m_1}.

Then, where $N = \binom{n}{2}$

$$\begin{aligned} \mathbf{Pr}(\mathcal{R}_H) &\le \binom{\binom{a}{2}}{b}\binom{N-b}{m_1-b}\Big/\binom{N}{m_1} \\ &= O(p_{m_1}^b). \end{aligned} \tag{30}$$

To bound $\mathbf{Pr}(\mathcal{R}_2)$ we condition on $N_x = S, N_y = T$ satisfying (26), where S, T are fixed subsets of $[n]$. Now let $\hat{\mathcal{R}}_2$ denote the event

$$|E(S)|, |E(T)| \le \frac{\log n}{20}.$$

We show that for $\gamma \ge \beta - \epsilon$, in $G_{n,\gamma p}$ we have

$$\mathbf{Pr}_{\gamma p}(\hat{\mathcal{R}}_2) = O(n^{-3/2}). \tag{31}$$

The monotonicity of $\hat{\mathcal{R}}_2$ plus the concentration of the number of edges of $G_{n,\gamma p}$ around $\gamma N p$ then allows us to assert (31) for G^2. Indeed, then

$$O(n^{-3/2}) = \mathbf{Pr}_{\gamma p}(\hat{\mathcal{R}}_2) = \sum_m \binom{N}{m}(\gamma p)^m(1-\gamma p)^{N-m}\mathbf{Pr}_m(\hat{\mathcal{R}}_2)$$

and so taking $\beta - \epsilon \le \gamma$ we see that if $\mathbf{Pr}_{m_1}(\hat{\mathcal{R}}_2) \ge An^{-3/2}$ then $\mathbf{Pr}_{\gamma p}(\hat{\mathcal{R}}_2) \ge An^{-3/2}/2$.

The random variable $X=|E(G_x)|$ (in $G_{n,\gamma p}$) is a binomial random variable $B(s,p)$ where $s=\binom{|S|}{2}$, having mean μ where

$$(\beta-\epsilon)^3\log n<\mu<(\beta+\epsilon)^3\log n.$$

So,

$$\begin{aligned}\mathbf{Pr}_{\gamma p}\left(X\leq\frac{\log n}{20}\right)&\leq\sum_{l=0}^{\lfloor\frac{\log n}{20}\rfloor}\binom{s}{l}(\gamma p)^l(1-\gamma p)^{s-l}\\&\leq(1+o(1))\sum_{l=0}^{\lfloor\frac{\log n}{20}\rfloor}e^{-\mu}\frac{\mu^l}{l!}\\&\leq 2e^{-\mu}\frac{\mu^{\lfloor\frac{\log n}{20}\rfloor}}{\lfloor\frac{\log n}{20}\rfloor!}\\&\leq 3\exp\left\{-\log n\left((\beta-\epsilon)^3-\frac{1}{20}(1+\log 20)\right)\right\}\\&\leq 3n^{-3/4}\end{aligned}$$

We apply the same argument to $|E(G_y)|$ (adding the appropriate conditioning on the number of edges within N_y). The proof now follows from (28)–(31).

Case 2. $t\geq 4$

We bound $\mathbf{Pr}(\mathcal{R}_x\wedge\mathcal{R}_y\wedge\mathcal{R}_H)$ by conditioning on the event that the neighborhoods of x and y are of nearly the expected size and have nearly the expected number of edges. Let $\mathcal{R}_3$ is the event that

$$\begin{gathered}(\beta-\epsilon)pn\leq|N_x|,|N_y|\leq(\beta+\epsilon)pn,\\(\beta-\epsilon)p\binom{|N_x|}{2}\leq|E(G_x)|\leq(\beta+\epsilon)p\binom{|N_x|}{2},\ \text{and}\\(\beta-\epsilon)p\binom{|N_y|}{2}\leq|E(G_y)|\leq(\beta+\epsilon)p\binom{|N_y|}{2}.\end{gathered}$$

Let $\mathcal{R}_4$ be the event that both G_x and G_y contain fewer than $\frac{\log n}{20}$ copies of K_{t-1}. We now bound the probability of $\mathcal{R}_x\wedge\mathcal{R}_y\wedge\mathcal{R}_H$ as follows:

$$\begin{aligned}\mathbf{Pr}(\mathcal{R}_x\wedge\mathcal{R}_y\wedge\mathcal{R}_H)&\leq\mathbf{Pr}(\mathcal{R}_4|\mathcal{R}_H\wedge\mathcal{R}_3)\mathbf{Pr}(\mathcal{R}_H)+\mathbf{Pr}(\overline{\mathcal{R}_3})\\&\leq\mathbf{Pr}(\mathcal{R}_4|\mathcal{R}_H\wedge\mathcal{R}_3)O(p_{m_1}^b)+O(\exp\{-n^{1-\frac{2}{t}+o(1)}\}).\end{aligned}\tag{32}$$

We bound $\mathbf{Pr}(\mathcal{R}_4|\mathcal{R}_H \wedge \mathcal{R}_3)$ by an application of the Poisson approximation on the number of K_t's in the random graph $G_{n,m}$ given by Theorem 6.1 of [9, page 68]. We let n' and m' be integers satisfying

$$(\beta-\epsilon)pn \le n' \le (\beta+\epsilon)pn, \text{ and} \tag{33}$$

$$(\beta-\epsilon)p\binom{n'}{2} \le m' \le (\beta+\epsilon)p\binom{n'}{2}, \tag{34}$$

and condition on the event that $|N_x|=n'$ and $|E(G_x)|=m'$. Note that under this conditioning G_x can be viewed as the random graph $G_{n',m'}$. Following the notation of [9], we have

$$\frac{1}{2}(n')^{2-\frac{2}{t-2}}\omega_1 \le m' \le \frac{1}{2}(n')^{2-\frac{2}{t-2}}\omega_2$$

where

$$\omega_1 = (\beta-\epsilon)^{\frac{t}{t-2}}\left((t-1)!\log n\right)^{1/\binom{t-1}{2}}$$

and

$$\omega_2 = (\beta+\epsilon)^{\frac{t}{t-2}}\left((t-1)!\log n\right)^{1/\binom{t-1}{2}}.$$

Let $X=X_{K_t}$ be the number of copies of K_t in $G_{n',m'}$. The expected number of such K_t's, $\lambda:=\mathbf{E}[X]$, is then bounded as follows:

$$(\beta-\epsilon)^{\binom{t}{2}}\log n \le \lambda \le (\beta+\epsilon)^{\binom{t}{2}}\log n.$$

It then follows from Theorem 6.1 of [9] that

$$\begin{aligned}
\mathbf{Pr}\left(X \le \frac{\log n}{20}\right) &\le (1+o(1))\sum_{k=0}^{\lfloor\frac{\log n}{20}\rfloor} e^{-\lambda}\frac{\lambda^k}{k!} \\
&\le 2e^{-\lambda}\frac{\lambda^{\lfloor\frac{\log n}{20}\rfloor}}{\lfloor\frac{\log n}{20}\rfloor!} \\
&\le 2e^{-\lambda}\left(\frac{20e\lambda}{\log n}\right)^{\frac{\log n}{20}} \\
&\le 2\exp\left\{-(\beta-\epsilon)^{\binom{t}{2}}\log n\right\}(20e)^{\frac{\log n}{20}} \\
&= 2\exp\left\{-\log n\left((\beta-\epsilon)^{\binom{t}{2}} - \frac{1}{20}(1+\log 20)\right)\right\} \\
&\le 2n^{-3/4}
\end{aligned}$$

With (32) this completes the proof. ∎

Proof of Lemma 5. Let $\mathcal{S}_1$ be the event that there is a chain in G_{m_1}. For a fixed collection $\mathcal{A}$ of K_t's in K_n and distinct $u, v \in [n]$ which define a possible chain, it follows from an argument along the line of the proof of Claim 1 that

$$|V(\mathcal{A})| \leq 1 + \frac{2|E(\mathcal{A})|}{t}$$

and it follows from Claim 2 that

$$\mathbf{Pr}\left((u \text{ and } v \text{ are small }) \wedge E(\mathcal{A}) \subseteq E(G_{m_1})\right) \leq O(p_{m_1}^{|E(\mathcal{A})|} n^{-3/2}).$$

Applying the first moment method we have

$$\begin{aligned}
\mathbf{Pr}(\mathcal{S}_1) &\leq \binom{n}{2} \sum_{i=t}^{4t-3} \binom{n-2}{i-2} 2^{\binom{i}{t}} O(p_{m_1}^{\frac{(i-1)t}{2}} n^{-3/2}) \\
&\leq \sum_{i=t}^{4t-3} O(n^{i - \frac{2}{t}\frac{(i-1)t}{2} - \frac{3}{2} + o(1)}) \\
&\leq \sum_{i=t}^{4t-3} O(n^{-\frac{1}{2} + o(1)}) \\
&= o(1)
\end{aligned}$$

∎

Proof of Lemma 6. Let $\mathcal{S}_2$ be the event that there is a link in G_{m_1}. For fixed $S, T \in \binom{[n]}{t}$ such that $|S \cap T| \geq 2$ and $x \in S \cup T$ it follows from Claim 2 that

$$\mathbf{Pr}\left((x \text{ is small }) \wedge \binom{S}{2} \cup \binom{T}{2} \subseteq E(G_{m_1})\right) = O(p_{m_1}^{\binom{t}{2} - \binom{|S \cap T|}{2}} n^{-3/4}).$$

Applying the first moment method we have

$$\begin{aligned}
\mathbf{Pr}(\mathcal{S}_2) &\leq n \binom{n-1}{t-1} \sum_{i=2}^{t-1} \binom{t}{i} \binom{n-t}{t-i} O(p_{m_1}^{2\binom{t}{2} - \binom{i}{2}} n^{-3/4}) \\
&\leq \sum_{i=2}^{t-1} O(n^{2t - i - 2(t-1) + \frac{2}{t}\binom{i}{2} - \frac{3}{4} + o(1)}) \\
&\leq \sum_{i=2}^{t-1} O(n^{\frac{5}{4} - i + \frac{i(i-1)}{t} + o(1)}) \\
&= o(1)
\end{aligned}$$

∎

3. Proof of Theorem 1.

In view of Theorem 2 we need only prove that

$$\lim_{n\to\infty} \mathbf{Pr}(G_{n,m} \in \mathcal{C}_t) = \begin{cases} 0 & c_n \to -\infty \\ e^{-e^{-c}} & c_n \to c \\ 1 & c_n \to \infty \end{cases} \tag{35}$$

Using Theorem 2 of Łuczak [16] we can derive (35) directly from Theorem 7. ∎

4. Proofs of Theorems 4–6

We prove Theorem 4 via an application of the following theorem of Hajnal and Szemerédi. For $k \le n$ the *Turán graph* $T_k(n)$ is the complete k-partite graph on n vertices where the parts in the vertex partition have cardinalities

$$\left\lfloor \frac{n}{k} \right\rfloor, \left\lfloor \frac{n+1}{k} \right\rfloor, \ldots, \left\lfloor \frac{n+k-1}{k} \right\rfloor.$$

In other words, the parts in the partition are as near as possible to being equal (i.e. the partition is a so-called *equipartition*). Below we use the following theorem proved by Hajnal and Szemerédi (cf. Theorem 3).

Theorem 8 (Hajnal, Szemerédi). *If G is a graph on n vertices having maximum degree $\Delta(G) = \Delta$ then*

$$G \subseteq T_{\Delta+1}(n).$$

For a graph G, let $\overline{G}$ be the complement of G. It is easy to see that Theorem 8 is equivalent to

Theorem 9. *If G is a graph on n vertices having minimum degree $\delta(G) = \delta$ then*

$$\overline{T_{n-\delta}(n)} \subseteq G.$$

Proof of Theorem 4. We establish the lower bound by example. Consider the complete t-partite graph on n vertices having parts $V_1, \ldots, V_t$ such that $|V_1| = q$ and

$$|V_2|, \ldots, |V_t| \in \left\{ lq + \left\lceil \tfrac{r}{t-1} \right\rceil, lq + \left\lfloor \tfrac{r}{t-1} \right\rfloor \right\}.$$

If $q = 0$ then G contains no t-clique and therefore has no (K_t, l)-vertex-cover. If $q > 0$ then, by the definition of r, there exists V_i such that $|V_i| > ql$, and G has no (K_t, l)-vertex-cover.

Suppose G is a graph on n vertices having

$$\delta(G) \geq n - ql - \left\lceil \frac{r}{t-1} \right\rceil + 2.$$

Let

$$s = ql + \left\lceil \frac{r}{t-1} \right\rceil - 2.$$

It follows from Theorem 9 that $\overline{T_s(n)} \subseteq G$. In words, there exists an equipartition $V(G) = V_1 \cup \ldots \cup V_s$ such that the induced graph $G[V_i]$ is complete for $i = 1, \ldots s$. We will show that the collection of cliques $G[V_1], \ldots, G[V_s]$ can be transformed into a (K_t, l)-vertex-cover.

Claim 3.

$$t - 1 \leq |V_i| \leq t \text{ for } i = 1, \ldots, s.$$

Proof. We merely observe that $s(t-1) < n$ while $st \geq n$.

$$\begin{aligned}\left[ql + \left\lceil \frac{r}{t-1} \right\rceil - 2\right](t-1) &\leq ql(t-1) + \left(\frac{r}{t-1} + 1\right)(t-1) - 2(t-1) \\ &\leq ql(t-1) + r - (t-1) \\ &< n.\end{aligned}$$

On the other hand,

$$\begin{aligned}\left[ql + \left\lceil \frac{r}{t-1} \right\rceil - 2\right] t &\geq \left[ql + \frac{r}{t-1} - 2\right] t \\ &= n + q(l-1) + \frac{r}{t-1} - 2t.\end{aligned} \tag{36}$$

Now, since $n \geq 6t^2 - 4t$, at least one of the following holds:

- $r \geq 2t(t-1)$
- $q \geq 2t$
- $q(t-1)l \geq 4t(t-1)$.

In any of these situations, the expression in (36) is greater than or equal to n. ∎

If follows from Claim 3 that we may assume that for some m we have $|V_1|=\ldots=|V_m|=t-1$ and $|V_{m+1}|=\ldots=|V_s|=t$.

Claim 4.

$$m<(l-1)(q+1).$$

Proof. Since $V_1,\ldots,V_s$ is a partition, we must have $(t-1)m+t(s-m)=n$. However,

$$\begin{aligned}(t-1)(l-1)(q+1)+t\left[ql+\left\lceil\frac{r}{t-1}\right\rceil-2-(l-1)(q+1)\right]\\
&=q[(t-1)l+1]+t\left\lceil\frac{r}{t-1}\right\rceil+1-l-2t\\
&\le q[(t-1)l+1]+t\left(\frac{r}{t-1}+\frac{t-2}{t-1}\right)+1-l-2t\\
&\le n+\frac{1}{t-1}+t\frac{t-2}{t-1}+1-2t\\
&=n-t\\
&<n\end{aligned}$$

∎

We transform $G[V_1],\ldots,G[V_s]$ into a (K_t,l)-vertex-cover by expanding the clique V_i by one vertex for $i=1,\ldots,m$. To be precise, we will show that there exist $x_1,\ldots,x_m\in V(G)$ such that

1. $x_i\sim v \quad \forall v\in V_i$,
2. $|\{x_i : x_i=v\}|\le l-1 \quad \forall v\in V(G)$,
3. $x_i\in V_j\Rightarrow x_j\notin V_i$,
4. $x_i\notin V_i$.

Note that the third condition must be included to prevent two of the expanded cliques from containing a common edge. For $i=1,\ldots,m$ let

$$A_i=\{v\in V(G)\setminus V_i : v\sim u \quad \forall u\in V_i\}$$

Claim 5. *$|A_i|\ge q+t$ for $i=1,\ldots m$.*

Proof. Since, for $v\in V_i$,

$$\begin{aligned}|\{x\in V(G)\setminus V_i : x\not\sim v\}|&\le n-1-\delta(G)\\
&\le ql+\left\lceil\frac{r}{t-1}\right\rceil-3,\end{aligned}$$

we have

$$\begin{aligned}|\{x \in V(G) \setminus V_i : \exists v \in V_i \text{ such that } x \not\sim v\}| \\ &\le (t-1)\left[ql + \left\lceil \frac{r}{t-1} \right\rceil - 3\right] \\ &\le ql(t-1) + (t-1)\left(\frac{r}{t-1} + \frac{t-2}{t-1}\right) - 3(t-1) \\ &= ql(t-1) + r - 2t + 1.\end{aligned}$$

Therefore

$$\begin{aligned}|A_i| &= |V(G) \setminus V_i| - |\{x \in V(G) \setminus V_i : \exists v \in V_i \text{ such that } x \not\sim v\}| \\ &\ge n - (t-1) - [ql(t-1) + r - 2t + 1] \\ &= q + t\end{aligned}$$

∎

Now, we choose the x_i's one at a time in an order $x_1 = x_{i_1}, x_{i_2}, \dots, x_{i_m}$ as follows. Suppose $x_{i_1}, \dots, x_{i_k}$ have been chosen.

(37) $$\text{If } x_{i_k} \in V_j \text{ and } j \notin \{i_1, \dots, i_k\} \text{ then } j = i_{k+1}.$$

Otherwise i_{k+1} is chosen arbitrarily from $\{j : 1 \le j \le m\} \setminus \{i_1, \dots, i_k\}$. In other words, we chose the x_i's in an order such that at most one x_i falls in V_j before x_j is chosen. For $k = 1, \dots, m$ let

$$U_k = \{v \in V(G) : |\{1 \le j < k : x_{i_j} = v\}| = l - 1\}.$$

In words, U_k is the set of vertices that satisfy condition 2 with equality after $x_{i_1}, \dots, x_{i_{k-1}}$ have been determined. Thus, we must have $x_{i_k} \notin U_k$. By Claim 4

(38) $$|U_k| \le \left\lfloor \frac{m-1}{l-1} \right\rfloor < q + 1.$$

For $k = 1, \dots, m$ let

$$R_k = \bigcup_{1 \le j < k : x_{i_j} \in V_k} V_{i_j}.$$

(Note that the union here is over zero or one set only). By condition 3 we must have $x_{i_k} \notin R_k$. By the construction of the ordering given in (37),

(39) $$|R_k| \le t - 1.$$

An arbitrary $x_{i_k} \in (A_{i_k} \setminus U_k) \setminus R_k$ satisfies 1, 2, and 3. By (38), (39) and Claim 5 such an element exists. ∎

Proof of Theorem 6. Let $\epsilon > 0$ and let G be a graph on n vertices with $\delta(G) = \delta \geq (1 - \frac{1}{\chi(H)-1} + \epsilon)n$. We show that any collection of edge disjoint copies of H that does not cover $V(G)$ can be extended to cover at least one new vertex. To be precise, we show that if a family $\mathcal{F} = \{\Gamma_1, \ldots, \Gamma_m\}$ of copies of H in G and a vertex $v \in V(G)$ satisfy

$$\begin{aligned} &m < n, \\ &\Gamma_i = (V(\Gamma_i), E(\Gamma_i)) \text{ are copies of } H \text{ in } G \text{ for all } i, \\ &E(\Gamma_i) \cap E(\Gamma_j) = \emptyset \text{ for all } i \neq j, \end{aligned} \tag{40}$$

and

$$v \notin \cup_{i=1}^m V(\Gamma_i) \ ,$$

then there exists a family $\mathcal{F}' = \{\Upsilon_1, \ldots, \Upsilon_l\}$ such that for all i $\Upsilon_i = (V(\Upsilon_i), E(\Upsilon_i))$ are copies of H in G

$$E(\Upsilon_i) \cap E(\Upsilon_j) = \emptyset \text{ for all } i \neq j \tag{41}$$

and

$$\cup_{i=1}^l V(\Upsilon_i) \supseteq \Big(\bigcup_{i=1}^m V(\Gamma_i)\Big) \cup \{v\} \ .$$

Note that we include the possibility of $m = 0$. Clearly, an inductive argument based on (40) and (41) above implies the theorem. Further, we may assume $m < n$ in (40). Suppose, on the contrary, that we have a family $\mathcal{F}^* = \{\Gamma_1, \ldots, \Gamma_m\}, m \geq n$, constructed inductively by (40) and (41) such that it does not cover all vertices. However, by the inductive construction of $\mathcal{F}^*$ every vertex is already in some copy of H included in the family $\mathcal{F}^*$. A contradiction.

To proceed with the proof we need to establish some notational conventions. Let u be the vertex of H such that $\chi(H \setminus \{u\}) = \chi(H) - 1 =: \chi - 1$. Set $H' = H \setminus \{u\}$, $h = |V(H)|$, and $e_H = |E(H)|$. For $\mathcal{F}$ and a vertex v as in (40), let N_v be the set of neighbors of v, $d_v = |N_v|$ and $F = \cup_{i=1}^m E(\Gamma_i)$. Our analysis will focus on the consideration of the subgraphs $L = G[N_v]$ and $L' = (N_v, E(L) \setminus F)$. We extend $\mathcal{F}$ to $\mathcal{F}'$ by simply finding a copy of H which contains v but no edges in F. Clearly, if there exists a copy of H' in L', then this H' together with v gives a copy of H that extends $\mathcal{F}$. (Note H' is a subgraph of $L = G[N_v]$).

We have for $|E(L)|\geq\frac{d_v}{2}\Big(\delta-(n-d_v)\Big)$. Since $\delta\geq\Big(\frac{\chi-2}{\chi-1}+\epsilon\Big)n$ is equivalent to $\delta-n\geq-\frac{1}{\chi-2}\delta+\epsilon n\frac{\chi-1}{\chi-2}$, we get

$$\begin{aligned}|E(L)| &\geq \frac{d_v}{2}\Big(\delta-(n-d_v)\Big)\\ &\geq \frac{d_v}{2}\Big(d_v-\frac{1}{\chi-2}\delta+\epsilon n\frac{\chi-1}{\chi-2}\Big)\\ &\geq \frac{d_v^2}{2}\cdot\frac{\chi-3}{\chi-2}+\epsilon n\frac{d_v}{2}\cdot\frac{\chi-1}{\chi-2}\,.\end{aligned}$$

Since we are assuming that $|\mathcal{F}|<n$, we have

$$|F\cap E(L)|\leq|F|\leq e_H n\ ,$$

and it follows

$$\begin{aligned}|E(L')| &= |E(L)|-|F\cap E(L)|\\ &\geq \frac{d_v^2}{2}\cdot\frac{\chi-3}{\chi-2}+\epsilon n\cdot\frac{d_v}{2}\cdot\frac{\chi-1}{\chi-2}-e_H n\\ &\geq \binom{d_v}{2}\cdot\frac{\chi-3}{\chi-2}+\frac{1}{2}\epsilon\binom{d_v}{2}\frac{\chi-1}{\chi-2}\\ &\quad+\Big(\frac{1}{2}\epsilon\binom{d_v}{2}\frac{\chi-1}{\chi-2}+\frac{d_v}{2}\cdot\frac{\chi-3}{\chi-2}+\epsilon\frac{d_v}{2}\cdot\frac{\chi-1}{\chi-2}-e_H n\Big)\,.\end{aligned}$$

Letting $\epsilon'=\frac{1}{2}\cdot\frac{\chi-1}{\chi-2}\cdot\epsilon$ and d_v be large enough (i.e. n large enough), we conclude that

$$\frac{1}{2}\epsilon\binom{d_v}{2}\frac{\chi-1}{\chi-2}+\frac{d_v}{2}\cdot\frac{\chi-3}{\chi-2}+\epsilon\frac{d_v}{2}\cdot\frac{\chi-1}{\chi-2}-e_H n\geq 0$$

and thus, $|E(L')|\geq\Big(\frac{\chi-3}{\chi-2}+\epsilon'\Big)\binom{d_v}{2}$. By the Erdős - Stone theorem (see e.g. [6]) there exists a copy of H' in L'. Taking this copy of H' together with v and edges needed gives us a new copy of H by which we extend $\mathcal{F}$ to $\mathcal{F}'$.

Proof of Theorem 5. We are going to determine the exact value of $f(n,3,k), k\geq\frac{n-1}{2}$ and $n\geq 6$. First, note that in any (K_3,∞)-vertex-cover of a graph G on n vertices no vertex lies in more than $\frac{n-1}{2}$ copies of K_3. In order to get a tight result we assume G is a graph on n vertices with $\delta(G)\geq\lceil n/2\rceil+1$. Let $\mathcal{F}=\{\Gamma_1,\dots,\Gamma_m\}$ and v be as in (40) with $H=K_3$. We use the notation introduced in the proof of Theorem 6. Unlike in the proof of Theorem 6, in order to get a tight result it does not suffice to simply add a new K_3 to $\mathcal{F}$. Our argument includes consideration of several different kinds of modifications of $\mathcal{F}$.

It follows from our minimal degree condition that

$$d_L(x) \geq 2, \quad \text{for all } \; x \in N_v. \tag{42}$$

If there is an edge in L not contained in $F = \cup_{i=1}^m E(\Gamma_i)$ then this edge together with v gives an extension of $\mathcal{F}$ that contains v, and therefore we can assume

$$E(L) \subset F. \tag{43}$$

It follows from (42) and (43) that $|F \cap E(L)| \geq d_v = |N_v|$, and therefore

$$3|\mathcal{F}_3| + |\mathcal{F}_2| \geq d_v \geq \frac{n}{2} + 1\ , \tag{44}$$

where $\mathcal{F}_j = \{\Gamma \in \mathcal{F} : |V(\Gamma) \cap V(L)| = j\}, j = 2,3$. Since $H = K_3$, to simplify the description we identify $\Gamma \in \mathcal{F}$ with its vertex set, i.e. $\Gamma = \{x_1, x_2, x_3\}$. Consider $\Gamma_A = \{x_1, x_2, y\} \in \mathcal{F}_2$ with $x_1, x_2 \in N_v$ and $y \in V(G) \setminus (N_v \cup \{v\})$. If there exists $\Gamma_B \in \mathcal{F}, \Gamma_B \neq \Gamma_A$, such that $y \in \Gamma_B$ then $(\mathcal{F} \setminus \{\Gamma_A\}) \cup \{\{x_1, x_2, v\}\}$ is an extension of $\mathcal{F}$ containing v. Therefore, we can assume

$$|\mathcal{F}_2| \leq |V(G) \setminus (N_v \cup \{v\})| \leq \frac{n}{2} - 2\ , \tag{45}$$

because otherwise there exists a pair $\Gamma_A, \Gamma_B \in \mathcal{F}, \Gamma_A = \{x_1, x_2, y\}, \Gamma_B = \{z_1, z_2, y\}$ as above. It follows from (44) and (45) that $|\mathcal{F}_3| \geq 1$. Now, consider $\Gamma_A \in \mathcal{F}_3$. If there exists $\Gamma_B \in \mathcal{F}$ such that $\Gamma_A \cap \Gamma_B = \{x\}$ then $(\mathcal{F} \cup \{\Gamma_A \setminus \{x\} \cup \{v\}\}) \setminus \{\Gamma_A\}$ is an extension of $\mathcal{F}$ containing v. So, we can henceforth assume

$$\Gamma_A \in \mathcal{F}_3, \Gamma_B \in \mathcal{F} \Longrightarrow \Gamma_A \cap \Gamma_B = \emptyset. \tag{46}$$

Once again, we consider $\Gamma_A = \{x_1, x_2, x_3\} \in \mathcal{F}_3$. Since $d_G(x_i) \geq n/2 + 1 > 3$ (here we use our assumption on n) there exists $u \in V \setminus \{v, x_1, x_2, x_3\}$ and $a \neq b \in \{1,2,3\}$ such that u is adjacent to both x_a and x_b. Let $c \in \{1,2,3\} \setminus \{a, b\}$ and set

$$\mathcal{F}' = \mathcal{F} \setminus \{\Gamma_A\} \cup \{\{x_a, x_b, u\}, \{x_a, x_c, v\}\}.$$

By (46) the family $\mathcal{F}'$ is edge-disjoint and covers v.

In order to prove the lower bound on $f(n,3,k)$ we consider the following two graphs. If $n = 2m$, H_n^e is the complete bipartite graph on the vertex set $Z_1 \cup Z_2, |Z_1| = |Z_2| = m$. In the case $n = 2m+1$, H_n^o consists of the edges of the complete bipartite graph on the vertex set $Z_1 \cup Z_2, |Z_1| = m+1, |Z_2| = m$. Moreover, if $|Z_1|$ is even, H_n^o contains edges of a perfect matching of Z_1 and in the case $|Z_1|$ is odd, H_n^o contains edges of a maximal matching, say M, of Z_1 together with a single edge $\{x,y\}$ where x is the vertex of Z_1 which does not belong to M and y is any vertex of $Z_1 \setminus \{x\}$. Clearly, $\delta(H_n^e) = \lceil n/2 \rceil$ and $\delta(H_n^o) = \lceil n/2 \rceil$. Further, neither of H_n^e and H_n^o contains a (K_3, ∞)-vertex-cover because H_n^e does not contain any copy of K_3 and H_n^o contains only at most $\lceil (n+1)/4 \rceil$ copies of K_3. ∎

Acknowledgement. We thank the referees for their careful reading of the manuscript which has led to several improvements in the presentation.

References

[1] N. Alon, J. H. Spencer: *The Probabilistic Method*, John Wiley and Sons, 1991.
[2] N. Alon, R. Yuster: Almost H-factors in dense graphs, *Graphs and Combinatorics,* **8** (1992), 95–102.
[3] N. Alon, R. Yuster: Threshold functions for H-factors, *Combinatorics, Probability and Computing,* **2** (1993), 137–144.
[4] N. Alon, R. Yuster: H-factors in dense graphs, *J. Combin. Theory,* Ser. B, **66** (1996), 269–282.
[5] B. Bollobás: *Random graphs*, Academic press, 1985.
[6] B. Bollobás: *Extremal Graph Theory*, LMS Monographs, No. 11, Academic Press 1978.
[7] F. Chung, R. Graham: *Erdős on graphs—his legacy of unsolved problems,* A. K. Peters, 1998.
[8] P. Erdős and A. Rényi: On the existence of a factor of degree one of a connected random graph, *Acta Math. Acad. Sci. Hungar.,* **17** (1966) 359–368.
[9] A. M. Frieze: On small subgraphs of random graphs, Proceedings of Random Graphs '89, (A. M. Frieze and T. Łuczak Eds.) John Wiley and Sons, 1992, 67–90.
[10] A. Hajnal, E. Szemerédi: Proof of a conjecture of Erdős, *Combin. Theory and Appl., II,* Colloq. Math. Soc. J. Bolyai **4**, North-Holland, 1970, 601–623.
[11] S. Janson: Poisson approximation for large deviations, *Random Structures and Algorithms,* **1** (1990), 221–229.
[12] S. Janson, T. Łuczak and A. Ruciński: *Random Graphs*, John Wiley and Sons, 2000.
[13] M. Krivelevich: Triangle factors in random graphs, *Combinatorics, Probability and Computing,* **6** (1997), 337–347.
[14] D. E. Knuth, R. Motwani and B. G. Pittel: Stable husbands, *Random Structures and Algorithms,* **1** (1990), 1–14.
[15] J. Komlós, G. N. Sárközy, E. Szemerédi: Proof of the Alon–Yuster conjecture, in preparation.
[16] T. Łuczak: On the equivalence of two basic models of random graphs, In: Proceedings of Random Graphs '87 (M. Karoński, J. Jaworski and A. Ruciński, Eds.) John Wiley and Sons, 1990, 151–157.
[17] A. Ruciński: Matching and covering the vertices of a random graph by copies of a given graph, *Discrete Math.,* **105** (1992), 185–197.

[18] J. Spencer: Threshold functions for extension statements, *J. Combin. Theory,* Ser. A **53** (1990) 286–305.

[19] N. C. Wormald: The perturbation method and triangle-free random graphs, *Random Structures and Algorithms*, **9** (1996), 253–270.

Tom Bohman

Department of Mathematical Sciences,
Carnegie Mellon University
Pittsburgh, PA 15213, USA
tbohman@andrew.cmu.edu

Alan Frieze

Department of Mathematical Sciences,
Carnegie Mellon University
Pittsburgh, PA 15213, USA
alan@random.math.cmu.edu

Miklós Ruszinkó

Department of Mathematical Sciences,
Carnegie Mellon University
Pittsburgh, PA 15213, USA
Permanent Address:
Computer and Automation
Research Institute
of the Hungarian Academy of Sciences,
Budapest, P.O.Box 63,
Hungary-1518.
ruszinko@lutra.sztaki.hu

Lubos Thoma

Department of Mathematical Sciences,
Carnegie Mellon University
Pittsburgh, PA 15213, USA
thoma@qwes.math.cmu.edu