Vertex covers by edge disjoint cliques

Tom Bohman  Alan Frieze*  Miklós Ruszinkó†
Lubos Thoma‡

Department of Mathematical Sciences,
Carnegie Mellon University.

June 6, 2000

Abstract

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 $H_1, \ldots, H_r$ of subgraphs of $G$ satisfying

1. $H_i \cong H$, for all $i = 1, \ldots, r$,
2. $\bigcup_{i=1}^r V(H_i) = V$,
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 $H_1, \ldots, H_r$ of subgraphs of $G$ satisfying

P1. $H_i \cong H$, for all $i = 1, \ldots, r$,

P2. $\bigcup_{i=1}^r V(H_i) = V$,

P3. $E(H_i) \cap E(H_j) = \emptyset$, for all $i \neq j$, and

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

*Supported in part by NSF grant CCR-9530974.
†Permanent Address: Computer and Automation Research Institute of the Hungarian Academy of Sciences, Budapest, P.O.Box 63, Hungary-1518.
‡Research was partially supported by OTKA Grants T 030059 and T 29074, FKFP 0607/1999 and by the Bolyai Foundation.
§Supported in part by NSF grant DMS-9970622.
We call the family \( \{H_1, \ldots, H_r\} \) an \((H, k)\)-vertex-cover of \( G \). Thus when \( k = 1 \) we ask for the existence of a partition of \( V \) into \textit{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. In the main we will focus on the case where \( H \) is a complete graph \( K_t \) and denote our property by \( \mathcal{C}_{k,t} \).

**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 [7] 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 [15] and by Alon and Yuster [3]. 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. In [15], 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 \( H \). Note that, for example, \( H \) being a complete graph is excluded. Hence, the first interesting case is \( H = K_3 \). In [11], Krivelevich showed that the probability \( p(n) = O(n^{-3/5}) \) is enough for the random graph to have a \( K_3 \)-factor \textbf{whp} and, in general, if \( p(n) = O(n^{-2/(t-1)(t+2)}) \) then the random graph \( G_{n,p} \) contains a \( K_t \)-factor \textbf{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 \( K_t \).

**Theorem 1.** Let \( m = (n)_2((t - 1)! (\log n + c_n))^{1/(t-1)} n^{-2/t} \). Then

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

(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 \( G_m = ([n], E_m), m = 0, 1, \ldots, (n)_2 \), where \( E_0 = \emptyset \) and \( G_m \) is obtained from \( G_{m-1} \) by choosing \( e_m \) randomly from \((n)_2 \setminus E_{m-1}\) and putting

---

\(^{1}\)A sequence of events \( \mathcal{E}_n \) occurs with high probability, \textbf{whp}, if \( \Pr(\mathcal{E}_n) = 1 - o(1) \).
\[ E_m = E_{m-1} \cup \{e_m\} \]. 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 \( t \geq 3 \),

\[
\lim_{n \to \infty} \Pr(\tau_1 = \tau_2) = 1.
\]

Moreover, there exists whp a \((K_t, 2)\)-vertex-cover of \( G_{\tau_2} \) containing \((1 + o(1)) n^t \) copies of \( K_t \).

**Remark 1.** In fact, our proof of Theorem 2 implies that \( G_{\tau_2} \) possesses whp a \((K_t, 2)\)-vertex-cover containing at most \( \left( \frac{1}{t} + \frac{1}{(\log n)^t} \right) n \) 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 \( C_{t,k} \)? For \( t \geq 3 \) and \( k \geq 2 \) let

\[
f(n, t, k) = \max \{d : \exists G \text{ such that } \delta(G) = d, |V(G)| = n \text{ and } G \notin C_{t,k}\}.
\]

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 [9].

**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

\[
n - qk - \left\lfloor \frac{r}{t-1} \right\rfloor \leq f(n, t, k) \leq n - qk - \left\lfloor \frac{r}{t-1} \right\rfloor + 1.
\]

Note that it follows from Theorem 4 that

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

(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 $(H, \infty)$-vertex-cover of $G$ is a collection $H_1, \ldots, H_r$ of subgraphs of $G$ satisfying P1, P2 and P3. Thus, $G$ has an $(H, \infty)$-vertex-cover if and only if there exists a $k$ such that $G$ has a $(H, k)$-vertex-cover. To motivate our results on $(H, \infty)$-vertex-covers, we recall the following well-known extension of Theorem 3. Given an arbitrary graph $H$, Komlős, Sárközy and Szemerédi [13] 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 $(H, \infty)$-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 $(H, \infty)$-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 $0 \leq \epsilon \leq 1$ and $\theta > 0$

$$\Pr(B(n, p) \leq (1 - \epsilon)np) \leq e^{-\epsilon^2np/2} \quad (2)$$

$$\Pr(B(n, p) \geq (1 + \epsilon)np) \leq e^{-\epsilon^2np/3} \quad (3)$$

$$\Pr(B(n, p) \geq \theta np) \leq (e/\theta)^{\theta np} \quad (4)$$

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

Let $t \geq 3$ 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

$$\beta^{(1)} > 19/20 \quad \text{and} \quad \alpha + \beta < 1,$$
and let
\[
m_a = \alpha \left( \frac{n}{2} \right) ((t - 1)! \log n)^{1/2} n^{-2/t}, \quad \text{and}
\]
\[
m_b = \beta \left( \frac{n}{2} \right) ((t - 1)! \log n)^{1/2} n^{-2/t}.
\]
Furthermore, for \( i = 0, 1 \) let
\[
m_i = \left( \frac{n}{2} \right) ((t - 1)! (\log n - (1 - 2i) \log \log n))^{1/2} n^{-2/t}.
\]

**Lemma 1.**
\[
\Pr(\tau_1 \notin [m_0, m_1]) = o(1).
\]

We will use the term ‘a collection of \( K_t \)'s’ in the graph \( G \), for a family \( \mathcal{A} \subseteq \left( V(G) \atop t \right) \) such that \( G[S] \) is complete for all \( S \in \mathcal{A} \). For such a collection \( \mathcal{A} \) we set
\[
V(\mathcal{A}) = \bigcup_{S \in \mathcal{A}} S \quad \text{and} \quad E(\mathcal{A}) = \bigcup_{S \in \mathcal{A}} \left( S \atop 2 \right),
\]
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 \( H_i = ([n], F_i), i = m_{-1}, \ldots, m_1 \) 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 \( \left( \begin{array}{c} n \end{array} \right) \setminus F_i \). 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 \( m_a + m_b \leq i \leq m_1 \) the graphs \( G_i \) and \( H_i \) are identically distributed.

We henceforth assume that
\[
m_a + m_b \leq m \leq m_1
\]
and that every vertex in \( H_m = G_m \) lies in at least one copy of \( K_t \). We will show that
\[
\text{whp } G_m \text{ has a } (K_t, 2)\text{-vertex-cover.} \tag{5}
\]
Theorem 2 follows from (5) 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 \(\Xi\) 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 \(\Xi\), and

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

We can easily randomise 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 \(G = G_{n, m_a}\).

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

It follows from Lemma 2 that whp

\[
|X| \leq r.
\]  

(6)

In other words, after using only a small fraction of the edges in \(G_m\), 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 \(C = \{S_1, \ldots, S_l\}\) of \(K_t\)'s in \(G_{m_1}\) such that \(l \leq 2\nu_i\)

\[
\kappa_i \geq 1 \quad \text{for} \quad i = 2, \ldots, l
\]
\[
\kappa_i = t \quad \Rightarrow \quad \kappa_{i-1} = 1 \quad \land \quad |S_i \cap S_{i-1}| \geq 2
\]
\[
\text{and} \quad |\{i : \kappa_i \neq 1\}| = \nu_i
\]

where

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

Roughly speaking, a cluster is a small collection of \(K_t\)'s that have many or large pairwise intersections.

**Lemma 3.**

\[
\Pr(G_{m_1} \text{ contains a cluster}) = o(1).
\]
We now turn our attention to the graph $G^2$. For $v \in [n]$ let $\Upsilon_v$ be the collection of $K_t$'s in $G^2$ that contain $v$; to be precise,

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

Since $\Upsilon_v$ 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 $X$. To make this statement precise, we let

$$q = \left\lfloor \frac{\log n}{\log \log \log n} \right\rfloor.$$

**Lemma 4.**

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

We say that

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

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 $S_1, S_2, S_3, S_4 \in \left( [n] \atop t \right)$ of (not necessarily distinct) sets such that $u \in S_1$, $v \in S_4$,

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

**Lemma 5.**

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

We also note that no $K_t$ containing a small vertex intersects any other $K_t$ in more than a single vertex. A link is a small vertex $u \in [n]$ and distinct $S_1, S_2 \in \left( [n] \atop t \right)$ 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.**

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

Finally, let

- $X_1 = \{ v \in X : v \text{ is small} \}$,
- $X_2 = \{ v \in X : v \text{ is large} \}$, and
- $\Phi = \left\{ S \in \left( [n] \atop t \right) : \binom{S}{2} \subseteq E(G_{m_1}) \text{ and } S \cap X_1 \neq \emptyset \right\}$.
We are now prepared to describe the remainder of the \((K_1, 2)\)-cover.

We henceforth assume (6),

\[
G_{m_1} \text{ does not contain a cluster,} \quad (7)
\]
\[
\forall v \in [n] \quad |V(Y_v) \cap X| \leq q, \quad (8)
\]
\[
G_{m_1} \text{ does not contain a chain,} \quad (9)
\]
\[
G_{m_1} \text{ does not contain a link,} \quad (10)
\]

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

\[
V(\Xi_1) \cap V(\Xi_2) = \emptyset \quad \text{and} \quad E(\Xi) \cap E(\Xi_1 \cup \Xi_2) = \emptyset. \quad (11)
\]

If follows from Lemmas 1, 2, 3, 4, 5 and 6 that (11) implies Theorem 2.

We cover \(X_1\) in a rather crude way. Let \(\Xi_1\) be an arbitrary collection of \(K_i\)'s in \(G_m\) that covers \(X_1\). Note that the collection \(\Xi_1\) uses edges from all 3 phases and that we make use of the fact that every vertex is contained in some \(K_i\) in \(G_m\) when forming \(\Xi_1\). By (9), \(\Xi_1\) 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 \(Y_v\)'s. For \(v \in X'_2\) let \(\mathcal{Y}_v^t\) be the collection of sets in \(S \in \mathcal{Y}_v\) such that

\[
S \cap X = \{v\}
\]
\[
T \in \binom{[n]}{t} \land \binom{T}{2} \subseteq E(G_{m_1}) \Rightarrow |S \cap T| \leq 1, \quad \text{and} \quad (12)
\]
\[
S \cap V(\Phi) \subseteq \{v\}.
\]

In words, we get \(\mathcal{Y}_v^t\) from \(\mathcal{Y}_v\) by throwing away those sets in \(\mathcal{Y}_v\) that contain an element of \(X\) other than \(v\), intersect another \(K_i\) in more than one vertex, or contain a vertex of a \(K_i\) that contains a small vertex. By (8) there are at most \(q\) sets in \(\mathcal{Y}_v\) that contain an element of \(X\) other than \(v\). We will show

\[
\text{there are } \leq \binom{2^{\text{ad}}}{t} \text{ sets in } \mathcal{Y}_v \text{ that intersect another } K_i \text{ in } \geq 2 \text{ vertices.} \quad (13)
\]

By (9) at most 1 set in \(\mathcal{Y}_v\) intersects \(V(\Phi)\). Therefore, we may choose \(\Theta_v \subseteq \mathcal{Y}_v^t\) such that

\[
|\Theta_v| = \left\lceil \frac{\log n}{21} \right\rceil \quad \text{for all} \quad v \in X'_2. \quad (14)
\]

**Proof of (13)** Let \(\hat{\mathcal{Y}}_v\) denote the collection of \(K_i\)'s in \(\mathcal{Y}_v\) which intersect another \(K_i\) in more than one vertex. Let \(B = V(\hat{\mathcal{Y}}_v)\). We construct copies \(X_1, X_2, \ldots, X_i\)
of $K_2$ in $G_{m_1}$ as follows: Suppose we have constructed $X_1, X_2, \ldots, X_k$. Either (i) $B \subseteq V_k = V(X_1 \cup X_2 \cup \ldots \cup X_k)$ or (ii) $B \not\subseteq V_k$. In case (ii) choose $X_{k+1} \in \mathcal{T}_v$ 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 $\nu_1$ iterations we will have produced a cluster. Thus $l \leq 2\nu_1$ and $|B| \leq 2t\nu_1$, which implies (13).

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,

$$\Pr(A_{u,v,S,T}) = \frac{1}{|\Theta_v||\Theta_u|} \leq \left( \frac{21}{\log n} \right)^2 = p.$$  \hspace{1cm} (15)

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

$$d := 2 \max_{u \in X'_2} \sum_{S \in \Theta_u} \sum_{v \in X'_2} |\{T \in \Theta_v : S \cap T \neq \emptyset\}|$$

$$\leq 2 \max_{u \in X'_2} \sum_{v \in V(\Theta_u)} |\gamma_v \cap X'_2|$$

$$\leq 2tq \left[ \log n \right] \frac{\log n}{21} \quad \text{by (8)}$$

$$\leq \frac{t(\log n)^2}{10 \log \log \log n}.$$  \hspace{1cm} (16)

It follows from (15) and (16) that

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

Thus, for $n$ sufficiently large, it follows from the Lovász Local Lemma that there exists a vertex disjoint collection $\Xi_2$ of $K_2$'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 (10) and (12). We have established (11) and completed the proof. \hfill \Box

### 2.1 Proof of Lemma 1

Let $p_i = m_i/\binom{n}{2}$ for $i = 0, 1$. 

9
We first apply Janson's inequality to show that whp every vertex in $G_{n,p_1}$ is contained in a copy of $K_t$ (we follow the notation of [1, pages 95 and 96]). Let $v$ be a fixed vertex and let $n$ denote the number of copies of $K_t$ in $G$ which are incident with $v$. Next let $S_1, S_2, \ldots, S_{\binom{n-1}{t-1}}$ be an enumeration of the copies of $K_t$ in $K_n$ which contain $v$. Letting $B_j$ be the event $(\binom{s_j}{2}) \subseteq E(G_{n,p_0})$, we have

$$
\mu = \sum_{j=1}^{\binom{n-1}{t-1}} \Pr(B_j) = \binom{n-1}{t-1} p_1^{\binom{s_j}{2}} = (\log n + \log \log n)(1 + O(1/n))
$$

(17)

and

$$
\Delta = \sum_{|S_j \cap S_k| \geq 2} \Pr(B_j \cap B_k)
= \binom{n-1}{t-1} \sum_{r=2}^{t-1} \binom{t-1}{r} p_1^{\binom{s_j}{2} - \binom{s_k}{2}}
= O\left(\sum_{r=2}^{t-1} n^{2t-r-1} \binom{s_j}{2} - \binom{s_k}{2} + o(1)\right)
= O(n^{2t-1-1+o(1)}).
$$

(18)

Then, by Janson's inequality, we have

$$
\Pr(Z = 0) \leq \exp\left\{ -\mu + \frac{1}{1 - \epsilon} \frac{\Delta}{2} \right\}
= \frac{1}{n \log n} \exp\left\{ O(n^{-1+o(1)}) + O(n^{2t-1-1+o(1)}) \right\}
= o(1/n).
$$

(19)

It follows that

$$
\Pr(\exists u \in [n] : u \text{ is not contained in a copy of } K_t \text{ in } G_{n,p_0}) = o(1).
$$

(20)

The event $\{\exists u \in [n] : u \text{ is not contained in a copy of } K_t\}$ is monotone decreasing and so (20) implies that whp every vertex in $[n]$ is contained in a copy of $K_t$ in $G_{n,m_1}$. In other words, $\tau_1 \leq m_1$ whp.

We now turn to the random graph $G_{n,p_0}$ in order to establish our almost sure lower bound on $\tau_1$. For $v \in [n]$ let $Z_v$ be the number of $K_t$'s in $G_{n,p_0}$ that contain $v$, and let $Y$ denote the number of vertices $v$ such that $Z_v = 0$. Since

$$
M = (1 - p_0)^{\binom{n-1}{t-1}} = (1 + o(1)) \frac{\log n}{n}
$$

(21)

is a lower bound on $\Pr(Z_v = 0)$ for each $v \in [n]$, we have

$$
E(Y) \geq (1 + o(1)) \log n.
$$

(22)
We now show that $\text{Var}(Y)$ is small. Indeed,

$$\Pr(Z_1 = Z_2 = 0) \leq \Pr(\mathcal{E}_1) + \Pr(\bar{\mathcal{E}}_2 \bar{\mathcal{E}}_3 | \bar{\mathcal{E}}_1)$$

(23)

where, if $N_i$ is the set of neighbors of $i$ in $G_{n,p_0}$,

$$\mathcal{E}_1 = \left\{ \left(1 - n^{-\frac{1}{2}}\right) n^{p_0} \leq |N_1|, |N_2| \leq 2np_0 \right\} \vee \left\{ |N_1 \cap N_2| \geq n^{-\frac{1}{2}}np_0 \right\}$$

$$\mathcal{E}_2 = \{G_{n,p_0} \text{ contains a copy } H \text{ of } K_{t-1} \text{ such that } H \subseteq N_1\}$$

$$\mathcal{E}_3 = \{G_{n,p_0} \text{ contains a copy } H \text{ of } K_{t-1} \text{ such that } H \subseteq N_2 \setminus N_1\}$$

Applying (2)–(4) we get,

$$\Pr(\mathcal{E}_1) \leq 5 \exp \left\{ -n^{1-\frac{5}{4t} + o(1)} \right\}.$$

Note that

$$\Pr(\bar{\mathcal{E}}_2 \wedge \bar{\mathcal{E}}_3 | N_1, N_2) = \Pr(\bar{\mathcal{E}}_2 | N_1, N_2) \Pr(\bar{\mathcal{E}}_3 | N_1, N_2)$$

because, conditioning on $N_1$ and $N_2$, these events depend on disjoint sets of edges.

Let $W_1$ and $W_2$ be fixed sets that satisfy

$$\left(1 - \frac{1}{n^{\frac{1}{4t}}}\right) n^{p_0} \leq |W_1| \leq 2np_0 \text{ and } \left(1 - \frac{2}{n^{\frac{1}{4t}}}\right) \leq |W_2 \setminus W_1| \leq 2np_0.$$  

It follows from another application of Janson’s inequality that

$$\Pr(\bar{\mathcal{E}}_2 | N_1 = W_1, N_2 \setminus N_1 = W_2), \Pr(\bar{\mathcal{E}}_3 | N_1 = W_1, N_2 \setminus N_1 = W_2)$$

$$\leq \exp \left\{ -\log n + \log \log n + O(n^{-\frac{1}{4t}} + o(1)) + O(n^{-1+\frac{1}{4t}} + o(1)) \right\}.$$  

Therefore,

$$\Pr(Z_1 = Z_2 = 0) = \frac{\log^2 n}{n^2} + o(1),$$

and it follows from (21) that

$$\text{Var}(Y) = o(\log^2 n).$$

It then follows from Chebyshev’s inequality that

$$\Pr(Y = 0) = o(1).$$

(24)

Since the event $\{Y = 0\}$ is monotone increasing, it follows from (24) that

$$\Pr(\text{every vertex in } G_{n,m_0} \text{ is contained in a copy of } K_t) = o(1).$$

In other words, we have shown that whp $\tau_1 > m_0$.  

11
\subsection{Proof of Lemma 2}

Let $p_a = m_a/\binom{n}{3}$ 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

$$\mu := \mathbb{E}[X_R]$$

$$= \sum_{S \in \binom{R}{t}} \mathbb{P}(B_S)$$

$$= \binom{r}{t} p_a^{\binom{t}{2}}$$

$$= \binom{r}{t} \frac{\alpha^2(t-1)! \log n}{n^{t-1}}$$

$$= \frac{r^t}{t!} (1 + O(1/r)) \frac{\alpha^2(t-1)! \log n}{n^{t-1}}$$

$$= \Omega(n)$$

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

$$\Delta = \sum_{S,T \in \binom{R}{t}, 2 \leq |S \cap T| \leq t-1} \mathbb{P}(B_S \wedge B_T)$$

$$= \binom{r}{t} \sum_{i=2}^{t-1} \binom{t}{i} \binom{r-t}{t-i} p_a^{\binom{i}{2} - \binom{i}{2}}$$

$$= \sum_{i=2}^{t-1} O \left( n^{2t-i-\frac{3}{2}\binom{i}{2} - \binom{i}{2} + o(1)} \right)$$

$$= \sum_{i=2}^{t-1} O \left( n^{2^{i-1} - i + o(1)} \right)$$

$$= O \left( n^{2^{t-1} + o(1)} \right).$$

Thus, Janson’s inequality gives

$$\mathbb{P}(X_R = 0) \leq e^{-c_1 n}$$
where \( c_1 \) is a positive constant. Applying the first moment method, we have

\[
\Pr \left( \bigvee_{R \in \binom{[n]}{r}} \{ X_R = 0 \} \right) \leq \binom{n}{r} e^{-c_1 n} \\
\leq \left( \frac{ne^r}{r} \right) e^{-c_1 n} \\
= \exp \left\{ r \left( 1 + \frac{\log \log n}{t} \right) - c_1 n \right\} \\
= o(1)
\]

Since this event is monotone, the same holds for \( G_{n,m} \).

### 2.3 Proof of Lemma 3

Let \( C = \{ S_1, \ldots, S_l \} \) be a fixed collection of \( K_i \)'s in \( K_n \) such that \( l \leq 2\nu_i \)

\[
\kappa_i \geq 1 \quad \text{for} \quad i = 2, \ldots, l \\
\kappa_i = t \quad \Rightarrow \quad \kappa_{i-1} = 1 \land |S_i \cap S_{i-1}| \geq 2 \\
\text{and} \quad |\{i : \kappa_i \neq 1\}| = \nu_i
\]

where

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

Let \( a = |V(C)| \) and \( b = |E(C)| \).

**Claim 7.**

\[
a - \frac{2b}{t} < -\frac{1}{t}
\]

**Proof.** We observe this difference as we ‘build’ the collection \( C \) one \( K_i \) at a time. For \( j = 1, \ldots, l \) let \( C_j = \{ S_1, \ldots, S_j \} \), \( a_j = |V(C_j)| \), \( b_j = |E(C_j)| \) and \( d_j = a_j - 2b_j/t \). Note that

\[
d_1 = 1,
\]

and

\[
d_{i+1} - d_i \leq (t - \kappa_{i+1}) \frac{2}{t} \left( \binom{t}{2} - \binom{\kappa_{i+1}}{2} \right) = (\kappa_{i+1} - 1) \left( \frac{\kappa_{i+1}}{t} - 1 \right).
\]

Thus

\[
\kappa_{i+1} = 1 \quad \Rightarrow \quad d_{i+1} - d_i = 0 \\
\text{and} \quad 2 \leq \kappa_{i+1} \leq t - 1 \quad \Rightarrow \quad d_{i+1} - d_i \leq \frac{2}{t} - 1.
\]
Furthermore, it follows from (25) that
\[ \kappa_{i+1} = t \Rightarrow b_{i+1} \geq b_i + t - 2 \Rightarrow d_{i+1} - d_i \leq -\frac{2(t-2)}{t}. \]
(29)

Since (by (28) and (29)) the difference \( a_i - 2b_i/t \) decreases by at least \( 1 - 2/t \) whenever \( \kappa_{i+1} \neq 1 \), it follows from (26) that \( a - 2b/t = d_i < -1/t \). \( \square \)

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
\[
\Pr(\mathcal{E}_i) \leq \binom{n}{i} 2^{(i)} p_{m_1}^{b_i} = O\left( n^{i - \frac{2i}{t} + o(1)} \right) = O\left( n^{-\frac{1}{t} + o(1)} \right).
\]

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

### 2.4 Proof of Lemma 4

We first argue that whp
\[ |\Upsilon_v| \leq 4 \log n \quad \text{for all } v \in [n]. \]
(30)

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 Lemma 3 and (13) that whp after removing \( O(1) \) \( K_i \)'s from \( \Upsilon_v \) we have a collection \( \tilde{\Upsilon}_v \) of \( K_i \)'s which are disjoint except for there containing \( v \). So in \( G_{n,p_b} \)
\[
\Pr( |\tilde{\Upsilon}_v| \geq \kappa = 3.9 \log n ) \leq \frac{(n-1)^\kappa}{\kappa!} \frac{(n)^{\kappa}}{\kappa!} \leq \frac{(e/3.9)^{3.9\log n}}{n^{-3/2}} = o(n^{-3/2}).
\]

This verifies (30).

Now fix a vertex \( v \). Then \( |V(\Upsilon_v)| < 4t \log n \) and \( |X| \leq r \). Also, \( X \) and \( V(\Upsilon_v) \) are chosen independently. It follows that
\[
\Pr( |V(\Upsilon_v) \cap X| \geq q ) \leq \frac{\binom{4t \log n}{q} \binom{n-q}{r-q}}{\binom{n}{r}} \leq \frac{\binom{4te \log n}{qn}}{q^n} \leq \binom{4te \log \log n \log n}{(\log n)^{(t+1)/t}} \log n/\log \log n = O(n^{-A})
\]

14
for any constant $A > 0$.
There are $n$ choices for $v$ and the lemma follows.

### 2.5 Proof of Lemmas 5 and 6

Let

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

The main work of this section is the following claim.

**Claim 8.** 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 $a := |A| \leq 4t$ then

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

**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

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

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

$$\beta + \epsilon < 1 \text{ and } (\beta - \epsilon)^{\binom{2}{t}} \geq \frac{3}{4} + \frac{1}{20} (1 + \log 20). \tag{31}$$

**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

$$(\beta - \epsilon)np \leq |N_x|, |N_y| \leq (\beta + \epsilon)np, \tag{32}$$

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

$$|E(G_x)|, |E(G_y)| \leq \frac{\log n}{20}. \tag{33}$$

We have

$$\Pr(\mathcal{R}_H \land \mathcal{R}_x \land \mathcal{R}_y) \leq \Pr(\mathcal{R}_2 | \mathcal{R}_1 \land \mathcal{R}_H) \Pr(\mathcal{R}_H) + \Pr(\bar{\mathcal{R}}_1). \tag{34}$$

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

$$\Pr(\bar{\mathcal{R}}_1) = O(\exp\{-n^{1-2/t+o(1)}\}), \tag{35}$$
and we can inflate this by \( O(n) \) to show the same for \( G_{m_1} \).

Then, where \( N = \binom{n}{2} \)

\[
\Pr(\mathcal{R}_H) \leq \binom{\binom{n}{2}}{b} \left( \frac{N - b}{m_1 - b} \right) / \binom{N}{m_1} = O\left(b^m_{m_1}\right),
\]

(36)

To bound \( \Pr(\mathcal{R}_2) \) we condition on \( N_x = S, N_y = T \) satisfying (32), where \( S, T \) are fixed subsets of \( [n] \). Now let \( \hat{\mathcal{R}}_2 \) denote the event

\[
|E(S)|, |E(T)| \leq \frac{\log n}{20}.
\]

We show that for \( \gamma \geq \beta - \epsilon \), in \( G_{n,\gamma p} \) we have

\[
\Pr_{\gamma p}(\hat{\mathcal{R}}_2) = O(n^{-3/2}).
\]

(37)

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 (37) for \( G^2 \). Indeed, then

\[
O(n^{-3/2}) = \Pr_{\gamma p}(\hat{\mathcal{R}}_2) = \sum_m \binom{N}{m} (\gamma p)^m (1 - \gamma p)^{N - m} \Pr_m(\hat{\mathcal{R}}_2)
\]

and so taking \( \beta - \epsilon \leq \gamma \) we see that if \( \Pr_m(\hat{\mathcal{R}}_2) \geq A n^{-3/2} \) then \( \Pr_{\gamma p}(\hat{\mathcal{R}}_2) \geq A n^{-3/2} / 2 \).

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

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

So,

\[
\Pr_{\gamma p} \left( X \leq \frac{\log n}{20} \right) \leq \sum_{l=0}^{\left\lfloor \frac{\log n}{20} \right\rfloor} \binom{s}{l} (\gamma p)^l (1 - \gamma p)^{s-l} \leq (1 + o(1)) \sum_{l=0}^{\left\lfloor \frac{\log n}{20} \right\rfloor} e^{-\mu \frac{l}{l!}} \\
\leq 2 e^{-\mu \frac{\left\lfloor \frac{\log n}{20} \right\rfloor}{\left\lfloor \frac{\log n}{20} \right\rfloor !}} \\
\leq 3 \exp \left\{ - \log n \left( (\beta - \epsilon)^3 - \frac{1}{20 (1 + \log 20)} \right) \right\} \leq 3 n^{-3/4}
\]

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 (34) – (37).
Case 2. $t \geq 4$

We bound $\Pr(\mathcal{R}_x \wedge \mathcal{R}_y \wedge \mathcal{R}_H)$ 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 $\mathcal{R}_3$ be the event that

$$(\beta - \epsilon)pn \leq |N_x|, |N_y| \leq (\beta + \epsilon)pn,$$

$$(\beta - \epsilon)p\left(\frac{|N_x|}{2}\right) \leq |E(G_x)| \leq (\beta + \epsilon)p\left(\frac{|N_x|}{2}\right), \text{ and}$$

$$(\beta - \epsilon)p\left(\frac{|N_y|}{2}\right) \leq |E(G_y)| \leq (\beta + \epsilon)p\left(\frac{|N_y|}{2}\right).$$

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

$$\Pr(\mathcal{R}_x \wedge \mathcal{R}_y \wedge \mathcal{R}_H) \leq \Pr(\mathcal{R}_4|\mathcal{R}_H \wedge \mathcal{R}_3)\Pr(\mathcal{R}_H) + \Pr(\mathcal{R}_3)$$

$$\leq \Pr(\mathcal{R}_4|\mathcal{R}_H \wedge \mathcal{R}_3)O(p_m^b) + O(\exp\{-n^{1-t^2+o(1)}\}). \quad (38)$$

We bound $\Pr(\mathcal{R}_4|\mathcal{R}_H \wedge \mathcal{R}_3)$ 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 [8, page 68]. We let $n'$ and $m'$ be integers satisfying

$$(\beta - \epsilon)pn \leq n' \leq (\beta + \epsilon)pn, \quad (39)$$

$$(\beta - \epsilon)p\left(\frac{n'}{2}\right) \leq m' \leq (\beta + \epsilon)p\left(\frac{n'}{2}\right), \quad (40)$$

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 [8], we have

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

where

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

and

$$\omega_2 = (\beta + \epsilon)^{\frac{t^2}{t-1}} ((t-1)! \log n)^{1/(t^2)}.$$
It then follows from Theorem 6.1 of [8] that

\[
\Pr \left( X \leq \log \frac{n}{20} \right) \leq (1 + o(1)) \sum_{k=0}^{\left\lfloor \log \frac{n}{20} \right\rfloor} e^{-\lambda} \frac{\lambda^k}{k!} \\
\leq 2e^{-\lambda} \frac{\lambda^{\left\lfloor \log \frac{n}{20} \right\rfloor}}{\log n} \\
\leq 2e^{-\lambda} \left( \frac{20\epsilon}{\log n} \right)^{\left\lfloor \log \frac{n}{20} \right\rfloor} \\
\leq 2 \exp \left\{ -(\beta - \epsilon) \left( \frac{\epsilon}{2} \log n \right) \left( 20 \epsilon \right)^{\frac{\log n}{20}} \right\} \\
= 2 \exp \left\{ - \log n \left( (\beta - \epsilon) \left( \frac{\epsilon}{2} \right) - \frac{1}{20} (1 + \log 20) \right) \right\} \\
\leq 2n^{-3/4}
\]

With (38) this completes the proof. □

Proof of Lemma 5. Let \( S_1 \) be the event that there is a chain in \( G_{m_1} \). For a fixed collection \( \mathcal{A} \) of \( K_i \)'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 7 that

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

and it follows from Claim 8 that

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

Applying the first moment method we have

\[
\Pr(S_1) \leq \binom{n}{2}^{4t-3} \sum_{i=t}^{\binom{n-2}{2}} 2^i O\left( p_{m_1}^{\frac{(i-1)i}{2}} n^{-3/2} \right) \\
\leq \sum_{i=t}^{4t-3} O\left( n^{i-\frac{2(i-1)i}{2}-\frac{2}{3}+o(1)} \right) \\
\leq \sum_{i=t}^{4t-3} O\left( n^{-\frac{1}{2}+o(1)} \right) \\
= o(1)
\]

Proof of Lemma 6. Let \( S_2 \) be the event that there is a link in \( G_{m_1} \). For fixed \( S, T \in \binom{[n]}{2} \) such that \( |S \cap T| \geq 2 \) and \( x \in S \cup T \) it follows from Claim 8 that

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

18
Applying the first moment method we have

\[ \Pr(S_2) \leq n \left( \frac{n-1}{t-1} \right) \sum_{i=2}^{t-1} \binom{i}{t} \left( \frac{n-t}{t-i} \right) O\left( \frac{2^t_i - \binom{i}{t}}{n^{3/4}} \right) \]
\[ \leq \sum_{i=2}^{t-1} O\left( n^{2t-i-2(t-1) + \frac{2}{3} \binom{i}{t} \frac{5}{3} + o(1)} \right) \]
\[ \leq \sum_{i=2}^{t-1} O\left( n^{\frac{5}{i} - \frac{i(i+1)}{1} + o(1)} \right) \]
\[ = o(1) \]

\[ \Box \]

3 Proof of Theorem 1.

For a graph \( G \) and a vertex \( v \), we defined prior to (21) \( Z_v(G) = Z_v \) to be the number of \( K_i \)'s in \( G \) that contain \( v \) and \( Y(G) = Y \) to be the number of vertices \( u \) with \( Z_u = 0 \).

In view of Theorem 2 we need only prove that

\[
\lim_{n \to \infty} \Pr(Y(G_{n,m}) = 0) = \begin{cases} 
0 & c_n \to -\infty \\
e^{-e^{-c}} & c_n \to c \\
1 & c_n \to \infty 
\end{cases}
\]

(41)

Using Theorem 2 of Łuczak [14] we can derive (41) from the more easily obtained

\[
\lim_{n \to \infty} \Pr(Y(G_{n,p}) = 0) = \begin{cases} 
0 & c_n \to -\infty \\
e^{-e^{-c}} & c_n \to c \\
1 & c_n \to \infty 
\end{cases}
\]

(42)

where \( p = m/\binom{n}{3} \). Furthermore we need only consider the case \( c_n \to c \) as the others follow by monotonicity. Equation (42) can be proved by showing that \( Y(G_{n,p}) \) is asymptotically Poisson. In particular we need only show that for \( k = O(1) \),

\[
\lim_{n \to \infty} n^{k} \Pr(Z_i(G_{n,p}) = 0, 1 \leq i \leq k) = e^{-ck}
\]

(43)

and then apply e.g. Theorem 20 of Bollobás [5].

Equation (43) follows from

\[
\Pr(Z_i(G_{n,p}) = 0 \mid Z_j(G_{n,p}) = 0, 1 \leq j < i) \sim \frac{e^{-e}}{n}
\]

(44)

for \( 1 \leq i \leq k \).

Using \( N_j \) to denote the neighbourhood of \( j \) in \( G_{n,p} \) we let
• $\nu_1$ denote the number of $K_{i-1}$ in $N_i \setminus \bigcup_{j=1}^{i-1} N_j$.

• $\nu_2$ denote the number of $K_{i-1}$ in $N_i$ which use a vertex of $\bigcup_{j=1}^{i-1} N_j$.

We then let $C_i = \{ Z_j(G_{n,p}) = 0, 1 \leq j < i \}$ and write

$$\Pr(Z_i(G_{n,p}) = 0 \mid C_i) = \Pr(\nu_1 = 0 \mid C_i)(1 - \Pr(\nu_2 \neq 0 \mid \nu_1 = 0, C_i)).$$

Then $\Pr(\nu_1 = 0 \mid C_i) \sim e^{-c}/n$ follows from Janson’s inequality and $\Pr(\nu_2 \neq 0 \mid \nu_1 = 0, C_i) \leq \Pr(\nu_2 \neq 0) = o(1/n)$ follows from the FKG inequality and a first moment calculation. \hfill \Box

4 Proofs of Theorems 4 – 6

We prove Theorem 4 via an application of the following theorem of Hajnal and Szemerédi. For $k \leq n$ 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 7 (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 7 is equivalent to

**Theorem 8.** If $G$ is a graph on $n$ vertices having minimum degree $\delta(G) = \delta$ then

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

Let a $(K_t,l)$-vertex-cover be a $K_t$-vertex-cover in which each vertex appears in at most $l$ copies of $K_t$.

**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

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

\[
\delta(G) \geq n - ql - \left\lfloor \frac{r}{t-1} \right\rfloor + 2.
\]

Let

\[
s = ql + \left\lfloor \frac{r}{t-1} \right\rfloor - 2.
\]

It follows from Theorem 8 that \( \overline{T_s(n)} \subseteq G \). In words, there exists an equipartition \( V(G) = V_1 \cup \cdots \cup V_s \) 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 9.

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

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

\[
\left[ ql + \left\lfloor \frac{r}{t-1} \right\rfloor - 2 \right] (t - 1) \leq ql(t - 1) + \left( \frac{r}{t-1} + 1 \right) (t - 1) - 2(t - 1)
\]

\[
\leq ql(t - 1) + r - (t - 1)
\]

\[
< n.
\]

On the other hand,

\[
\left[ ql + \left\lfloor \frac{r}{t-1} \right\rfloor - 2 \right] t \geq \left[ ql + \frac{r}{t-1} - 2 \right] t
\]

\[
= n + q(l - 1) + \frac{r}{t-1} - 2t.
\]

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 (46) is greater than or equal to \( n \).

If follows from Claim 9 that we may assume that for some \( m \) we have \( |V_1| = \cdots = |V_m| = t - 1 \) and \( |V_{m+1}| = \cdots = |V_s| = t \).
Claim 10.

\[ m < (l - 1)(q + 1). \]

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

\[
(t - 1)(l - 1)(q + 1) + t \left[ ql + \left\lfloor \frac{r}{t - 1} \right\rfloor - 2 - (l - 1)(q + 1) \right]
\]

\[
= q[(t - 1)(l + 1) + t \left( \frac{r}{t - 1} \right) + 1 - l - 2t]
\]

\[
\leq q[(t - 1)(l + 1) + t \left( \frac{r}{t - 1} + \frac{t - 2}{t - 1} \right) + 1 - l - 2t]
\]

\[
\leq n + \frac{1}{t - 1} + t \frac{t - 2}{t - 1} + 1 - 2t
\]

\[
= n - t
\]

\[
< n
\]

\[ \square \]

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. \( |\{x_i : x_i = v\}| \leq l - 1 \quad \forall v \in V(G) \),
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 11.** \( |A_i| \geq q + t \) for \( i = 1, \ldots, m \).

**Proof.** Since, for \( v \in V_i \),

\[
|\{x \in V(G) \setminus V_i : x \neq v\}| \leq n - 1 - \delta(G)
\]

\[
\leq ql + \left\lfloor \frac{r}{t - 1} \right\rfloor - 3,
\]

22
we have

$$|\{x \in V(G) \setminus V_i : \exists v \in V_i \text{ such that } x \neq v\}|$$

$$\leq (t - 1) \left[ ql + \left\lceil \frac{r}{t - 1} \right\rceil - 3 \right]$$

$$\leq ql(t - 1) + (t - 1) \left( \frac{r}{t - 1} + \frac{t - 2}{t - 1} \right) - 3(t - 1)$$

$$= ql(t - 1) + r - 2t + 1.$$

Therefore

$$|A_i| = |V(G) \setminus V_i| - |\{x \in V(G) \setminus V_i : \exists v \in V_i \text{ such that } x \neq v\}|$$

$$\geq n - (t - 1) - [ql(t - 1) + r - 2t + 1]$$

$$= q + t \quad \square$$

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

If $x_{i_k} \in V_j$ and $j \not\in \{i_1, \ldots, i_k\}$ then $j = i_{k+1}$.  \hfill (47)

Otherwise $i_{k+1}$ is chosen arbitrarily from $\{j : 1 \leq j \leq m\} \setminus \{i_1, \ldots, i_k\}$. 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 $k = 1, \ldots, m$ let

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

In words, $U_k$ is the set of vertices that satisfy 2. with equality after $x_{i_1}, \ldots, x_{i_{l-1}}$ have been determined. Thus, we must have $x_{i_k} \not\in U_k$. By Claim 10

$$|U_k| \leq \left\lceil \frac{m - 1}{l - 1} \right\rceil < q + 1. \quad \hfill (48)$$

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

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

(Note that the union here is over zero or one set only). By condition 3. we must have $x_{i_k} \not\in R_k$. By the construction of the ordering given in (47),

$$|R_k| \leq t - 1. \quad \hfill (49)$$

An arbitrary $x_{i_k} \in (A_{i_k} \setminus U_k) \setminus R_k$ satisfies 1, 2, and 3. By (48), (49) and Claim 11 such an element exists. \square
Proof of Theorem 6. Let $\epsilon > 0$ and let $G$ be a graph on $n$ vertices with $\delta(G) = \delta \geq (1 - \frac{1}{\chi(H)^2}) + \epsilon)n$. 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{equation}
m < n, \\
\Gamma_i = (V(\Gamma_i), E(\Gamma_i)) \text{ are copies of } H \text{ in } G \text{ for all } i, (50)\\nE(\Gamma_i) \cap E(\Gamma_j) = \emptyset \text{ for all } i \neq j,
\end{equation}
and
\begin{equation}
v \not\in \bigcup_{i=1}^{m} V(\Gamma_i),
\end{equation}
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 $G$
\begin{equation}
E(\Upsilon_i) \cap E(\Upsilon_j) = \emptyset \text{ for all } i \neq j
\end{equation}
and
\begin{equation}
\bigcup_{i=1}^{l} V(\Upsilon_i) \supseteq \left( \bigcup_{i=1}^{m} V(\Gamma_i) \right) \cup \{v\}.
\end{equation}
Note that we include the possibility of $m = 0$. Clearly, an inductive argument based on (50) and (51) above implies the theorem. Further, we may assume $m < n$ in (50). Suppose, on the contrary, that we have a family $\mathcal{F}' = \{\Gamma_1, \ldots, \Gamma_m\}, m \geq n$, constructed inductively by (50) and (51) 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 $\chi(H \setminus \{u\}) = \chi(H) - 1$. Set $H' = H \setminus \{u\}$, $h = |V(H)|$, and $e_H = |E(H)|$. For $\mathcal{F}$ and a vertex $v$ as in (50), let $N_v$ be the set of neighbors of $v$, $d_v = |N_v|$ and $F = \bigcup_{i=1}^{m} E(\Gamma_i)$. 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 $|E(L)| \geq \frac{d_v}{2} \left( \delta - (n - d_v) \right)$. Since $\delta \geq \left( \frac{\chi - 2}{\chi - 1} + \epsilon \right)n$ is equivalent to $\delta - n \geq -\frac{1}{\chi - 2} \delta + \epsilon n \frac{\chi - 1}{\chi - 2}$, we get
\begin{align*}
|E(L)| & \geq \frac{d_v}{2} \left( \delta - (n - d_v) \right) \\
& \geq \frac{d_v}{2} \left( d_v - \frac{1}{\chi - 2} \delta + \epsilon n \frac{\chi - 1}{\chi - 2} \right) \\
& \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{align*}
Since we are assuming that $|\mathcal{F}| < n$, we have

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

and it follows

$$|E(L')| = |E(L)| - |F \cap E(L)|$$

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

$$\geq \left( \frac{d_v}{2} \right) \cdot \frac{\chi - 3}{\chi - 2} + \frac{1}{2} \epsilon \left( \frac{d_v}{2} \right)^2 \frac{\chi - 1}{\chi - 2}$$

$$+ \left( \frac{1}{2} \epsilon \left( \frac{d_v}{2} \right)^2 \frac{\chi - 1}{\chi - 2} + d_v \cdot \frac{\chi - 3}{\chi - 2} + e \frac{d_v}{2} \cdot \frac{\chi - 1}{\chi - 2} - e_H n \right).$$

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

$$\frac{1}{2} \epsilon \left( \frac{d_v}{2} \right)^2 \frac{\chi - 1}{\chi - 2} + \frac{d_v \cdot \chi - 3}{\chi - 2} + e \frac{d_v}{2} \cdot \frac{\chi - 1}{\chi - 2} - e_H n \geq 0$$

and thus, $|E(L')| \geq \left( \frac{\chi - 3}{\chi - 2} + \epsilon' \right) \left( \frac{d_v}{2} \right)^2$. By the Erdős - Stone theorem 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 $(K_3, \infty)$-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 $\mathcal{F} = \{ \Gamma_1, \ldots, \Gamma_m \}$ and $v$ be as in (50) 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

$$d_L(x) \geq 2, \quad \text{for all} \quad x \in N_v. \quad (52)$$

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

$$E(L) \subset F. \quad (53)$$

It follows from (52) and (53) that $|F \cap E(L)| \geq d_v = |N_v|$, and therefore

$$3|\mathcal{F}_3| + |\mathcal{F}_2| \geq d_v \geq \frac{n}{2} + 1, \quad (54)$$

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
\[ \Gamma_A = \{x_1, x_2, y\} \in \mathcal{F}_2 \text{ with } x_1, x_2 \in V(G) \text{ and } y \in V(G) \setminus (V(G) \cup \{v\}). \] If there exists \( \Gamma_B \in \mathcal{F}, \Gamma_B \not= \Gamma_A, \) such that \( y \in \Gamma_B \) then \( (\mathcal{F} \setminus \{\Gamma_A\}) \cup \{\{x_1, x_2, y\}\} \) is an extension of \( \mathcal{F} \) containing \( v \). Therefore, we can assume

\[ |\mathcal{F}_2| \leq |V(G) \setminus (V(G) \cup \{v\})| \leq \frac{n}{2} - 2, \tag{55} \]

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 (54) and (55) 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

\[ \Gamma_A \in \mathcal{F}_3, \Gamma_B \in \mathcal{F} \implies \Gamma_A \cap \Gamma_B = \emptyset. \tag{56} \]

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 \( a \not= b \in \{1, 2, 3\} \) such that \( u \) is adjacent to both \( x_a \) and \( x_b \). Let \( c = \{1, 2, 3\} \setminus \{a, b\} \) and set

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

By (56) 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 \( n = 2m \), \( H_n^k \) is the complete bipartite graph on the vertex set \( Z_1 \cup Z_2, |Z_1| = |Z_2| = m \). In the case \( n = 2m + 1 \), \( H_n^k \) 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, \( H_n^k \) contains edges of a perfect matching of \( Z_1 \) and in the case \( |Z_1| \) is odd, \( H_n^k \) 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, \( \delta(H_n^k) = \lceil n/2 \rceil \) and \( \delta(H_n^k) = \lfloor n/2 \rfloor \).

Further, neither of \( H_n^k \) and \( H_n^k \) contains a \((K_3, \infty)-\)vertex-cover because \( H_n^k \) does not contain any copy of \( K_3 \) and \( H_n^k \) contains only at most \([(n + 1)/4] \) copies of \( K_3 \).

\[ \Box \]

References


