The cover time of a sparse random intersection graph

Mindaugas Bloznelis* Jerzy Jaworski† Katarzyna Rybarczyk‡

Abstract

Random intersection graphs model real affiliation networks, where each of $n$ actors selects an attribute set from a given collection of $m$ attributes and two actors establish adjacency relation whenever they share a common attribute. In the random intersection graph $G(n, m, p)$ actors select attributes independently at random and with probability $p$. We establish the cover time of the simple random walk on $G(n, m, p)$ at the connectivity threshold and above it. We consider the range of $n, m, p$ where the typical attribute is shared by (stochastically) bounded number of actors.

keywords: random intersection graphs, random walk, cover time, conductance.
MSC: 05C81, 05C80, 05C82, 91D30.

1 Introduction and results

Random walk is a standard tool of data collection in large networks. The cover time, i.e., the expected time needed to visit all vertices of the network, is a fundamental characteristic of a random walk. We study the cover time of the simple random walk in a special class of networks, called affiliation networks, where each actor is prescribed a finite set of attributes and two actors establish adjacency relation whenever they share a common attribute. For example, in the film actor network two actors are adjacent if they have played in the same movie, in the collaboration network two scientists are adjacent if they have co-authored a publication, in the consumer co-purchase network two consumers are adjacent if they have purchased similar products. Many complex networks have such a configuration of attributes prescribed to actors of the network. This structure of the affiliation network with actors and attributes might be hidden (see [9]). However it is plausible that it determines the form of various real life complex networks such as social networks, internet network, WWW.

*Faculty of Mathematics and Informatics, Vilnius University, 03225 Vilnius, Lithuania.
†Faculty of Mathematics and Computer Science, Adam Mickiewicz University, 60-614 Poznań, Poland
‡supported by NCN (National Science Center) grant 2014/13/D/ST1/01175
An important feature of real affiliation networks is the clustering property, i.e., the tendency of nodes to cluster together by forming relatively small groups with a high density of ties within a group [12]. We are interested in whether and how the clustering property affects the random walk performance. For this purpose we study a mathematical model of affiliation network, called random intersection graph, where attributes are prescribed to actors at random. We analyse the random intersection graph model introduced in [11], see also [8]. It was studied in the context of random walks already in [15] however only partial results were obtained.

We establish the first order asymptotics for the cover time of a random intersection graph. Our results are mathematically rigorous. They are inspired by the fundamental study by Cooper and Frieze [5], [6] of the cover time of the Erdős-Rényi random graph, where edges are inserted independently at random.

In this section we define the random graph model $G(n, m, p)$ and formulate our results. We let $n, m \to \infty$ and use the asymptotic notation $o(\cdot), O(\cdot), \Omega(\cdot), \Theta(\cdot), \approx$ explained in [10]. The phrase “with high probability” will mean that the probability of the event under consideration tends to one as $n, m$ tend to infinity.

The random intersection graph $G(n, m, p)$ with vertex set $V$ of size $n$ is defined using an auxiliary set of attributes $W$ of size $m$. Every vertex $v \in V$ is assigned a random set of attributes $W(v) \subset W$ and two vertices $u, v \in V$ become adjacent whenever they share a common attribute, i.e., $W(u) \cap W(v) \neq \emptyset$. We assume, in addition, that events $w \in W(v)$ are independent and have (the same) probability $p$. Noting that every set $\mathcal{V}(w) = \{v : w \in W(v)\} \subset V$ of vertices sharing an attribute $w \in W$ induces a clique in the intersection graph we can represent $G(n, m, p)$ as a union of $m$ randomly located cliques. The sizes $|\mathcal{V}(w)|$, $w \in W$, of these cliques are independent binomial random variables with the common distribution $\text{Bin}(n, p)$.

In this paper we assume that the expected size of the typical clique $E|\mathcal{V}(w)| = np = \Theta(1)$. Furthermore, we focus on the connectivity threshold, which for $G(n, m, p)$ is defined by the relation

$$ mp(1 - e^{-np}) = \ln n, $$

see [14], [17]. Namely, for $a_n = mp(1 - e^{-np}) - \ln n$ tending to $+\infty$ the probability that $G(n, m, p)$ is connected is $1 - o(1)$. For $a_n \to -\infty$ this probability is $o(1)$. Hence we will assume below that $m = O(n \ln n)$. It is convenient to represent the connectivity threshold in terms of the edge density (i.e., the probability that $u, v \in V$ are adjacent in $G(n, m, p)$). We denote by $p_I$ the edge density of $G(n, m, p)$. A simple calculation shows that $p_I = 1 - (1 - p^2)^m = mp^2(1 + O(p^2))$. Furthermore, at the connectivity threshold $\hat{p}_I$ the edge density $p_I$ approximately equals $(1 - e^{-np})^{-1}np(n^{-1} \ln n) =: \hat{p}_I$. Note that the latter expression differs by the factor $(1 - e^{-np})^{-1}np > 1$ from the edge density $n^{-1} \ln n$ that defines the connectivity threshold for the Erdős-Rényi random graph on $n$ vertices, see [10].
We recall that given a connected graph $G$ with the vertex set $V$, $|V| < \infty$, the cover time $C(G) = \max_{u \in V} C_u$, where $C_u$ is the expected number of steps needed by the simple random walk starting from vertex $u$ to visit all the vertices of $G$.

**Theorem 1.** Let $n, m \to +\infty$. Let $p = p(m, n) > 0$ and $c = c(m, n) > 1$ be such that $np = \Theta(1)$, $c = O(1)$ and

$$np(1 - (1 - p)^{n-1}) = c \ln n \quad \text{and} \quad (c - 1) \ln n \to \infty.$$  

Then with high probability the cover time of a random walk on $G(n, m, p)$ is

$$(3) \quad (1 + o(1)) \cdot \ln \left( \frac{np}{\ln \left( \frac{c-1}{c} (e^{np} - 1) + 1 \right)} \right) \cdot \frac{np}{1 - e^{-np}} \cdot cn \ln n.$$  

Note that under assumptions of the theorem $G(n, m, p)$ is just above the connectivity threshold. Indeed, for $np = \Theta(1)$ the quantities $(1 - p)^{n-1}$ and $e^{-np}$ are $O(n^{-1})$ close. Therefore, for the edge density $p_I$ defined by formula (2) we have $p_I = (c + O(n^{-1}))\hat{p}_I$. Furthermore $(c - 1) \ln n \to +\infty$ implies that $p_I > \hat{p}_I$. Moreover, we obtain from (3) that the random intersection graph with edge density $p_I = c\hat{p}_I$, $c > 1$, has the cover time

$$(4) \quad (1 + o(1))\lambda c\hat{p}_In^2,$$

where \( \lambda = \ln \left( \frac{np}{\ln \left( \frac{c-1}{c} (e^{np} - 1) + 1 \right)} \right).$$

It is interesting to compare (3) with the respective cover time of the Erdős - Rényi random graph $G(n, q)$. While comparing two different random graph models (both just above their respective connectivity thresholds) one may set the reference point to be a connectivity threshold. Then (4) is compared with the cover time of $G(n, q)$ for the edge density $q$ being just above the connectivity threshold $\hat{q}_B = n^{-1} \ln n$. The cover time of $G(n, c\hat{q}_B)$, $c > 1$, as $n \to +\infty$ is given by the formula, see (5),

$$(5) \quad (1 + o(1)) \ln (c/(c - 1))c\hat{q}_Bn^2.$$  

Interestingly, $\lambda = \lambda(np, c)$ is always less than $\ln(c/(c - 1))$ (we show this in Sect 4.2 below). Furthermore, we have $\lambda(np, c) \to \ln(c/(c - 1))$ as $np \to 0$. The later fact is not surprising as $G(n, m, p)$ is the union of iid random cliques $\mathcal{V}(w)$, $w \in \mathcal{W}$. For $np \to 0$ the absolute majority of the cliques are either empty or induce a single edge of $G(n, m, p)$. Therefore in this range $G(n, m, p)$ starts looking similar to a union of iid edges. But a union of iid edges represents the Erős-Rényi random graph.

On the other hand, in the case where edge densities of $G(n, m, p)$ and $G(n, q)$ are the same, the cover time of $G(n, m, p)$ is always larger than $G(n, q)$. To see this we set the edge density of $G(n, q)$ to be $q = c\hat{p}_I = c\hat{q}_B$, $c > 1$, where we write for short $\tilde{c} = np(1 - e^{-np})^{-1}c$. By (5), the cover time of $G(n, c\hat{q}_B)$ is

$$(1+ o(1)) \ln(\tilde{c}/(\tilde{c} - 1))c\hat{q}_B = \ln(\tilde{c}/(\tilde{c} - 1))c\hat{p}_I.$$
In Sect. 4.2 below we show that \( \lambda > \ln(\bar{c}/(\bar{c} - 1)) \). Therefore, given the edge density (namely, \( c^1 = \bar{c} \hat{q}_B, c > 1 \)), the cover time of the random intersection graph is always larger than that of the binomial graph. This can be explained by the fact that the abundance of cliques in \( G(n, m, p) \) may slow down the random walk considerably. It is interesting to trace the relation between the ratio \( \lambda/\ln(\bar{c}/(\bar{c} - 1)) \) and the expected size of the typical clique \( np = \mathbb{E}[\mathcal{V}(w)] \). In Figure 1 we present a numerical plot of the function \( np \to \lambda(np, c)/\ln(\bar{c}/(\bar{c} - 1)) \) for \( 0 \leq np \leq 30 \) and \( c = 1.1, c = 2 \) and \( c = 10 \).

From a technical point of view the main reason that makes the cover times so different is that at the connectivity thresholds the degree distributions of \( G(n, m, p) \) and \( G(n, q) \) (and consequently the stationary distributions of respective random walks) differ a lot. We believe that results on the degree distribution of \( G(n, m, p) \) obtained in Sect. 3 below are interesting in its own right as they give an insight into the structure of random intersection graphs.

We conclude this section with the outline of the proof. In the proof we use heuristics developed in works of Cooper and Frieze [5], [6]. First we show that with high probability \( G(n, m, p) \) has certain properties which are listed in Section 2. We call an intersection graph with such properties typical. Some of the proofs of the properties are standard. These are postponed to Appendix. Furthermore, the conductance property that guarantees \( O(\ln n) \) upper bound for the mixing time is shown in a separate paper [19]. In Section 3 we establish the properties of the degree distribution that, in fact, define the cover time asymptotics [3]. Then in Section 4 we study the expected number of returns of a random walk to a given vertex and evaluate the probability of the first visit to a vertex. Here we use a lemma due to Cooper and Frieze (see Corollary 7 in [6]) that relates the time needed to visit a vertex to the return probabilities of the walk that starts from that vertex. We note that in contrast to the Erdős-Rényi graph the typical vertex of \( G(n, m, p) \) may belong to quite a few small cycles. This makes our analysis of the return probabilities and conductance property a bit more involved. Finally, in Section 5, combining the results of Sections 3 and 4 we establish matching upper and lower bounds for the cover time that hold uniformly over the typical graphs.
2 Typical graphs

2.1 Intersection graphs - notation

In order to define needed properties we first introduce some notation. For any graph \( G \) we denote by \( \mathcal{E}(G) \) its edge set. We remark that the intersection graph \( \mathcal{G}(n,m,p) \) with the vertex set \( V \) and attribute set \( W \) is obtained from the bipartite graph \( \mathcal{B}(n,m,p) \) with bipartition \( (V,W) \), where each vertex \( v \in V \) and each attribute \( w \in W \) are linked independently and with probability \( p \). Any two vertices \( u, v \in V \) are adjacent in the intersection graph whenever they share a common neighbor in the bipartite graph. By \( \mathcal{G} \) and \( \mathcal{B} \) we denote realised instances of random graphs \( \mathcal{G}(n,m,p) \) and \( \mathcal{B}(n,m,p) \) and always assume that \( \mathcal{G} \) is defined by \( \mathcal{B} \). For convenience, edges in \( \mathcal{B} \) we call links. Cycles in \( \mathcal{B} \) we call \( \mathcal{B} \)-cycles. It is easy to see that for \( k \geq 3 \) any \( \mathcal{B} \)-cycle \( v_1w_1v_2\ldots w_{k-1}w_kv_1 \) \((v_i \in V, w_i \in W)\) defines the cycle \( v_1v_2\ldots v_kv_1 \) in \( \mathcal{G} \). Furthermore, any induced cycle \( v_1v_2\ldots v_kv_1 \) \((k \geq 3)\) in \( \mathcal{G} \) is defined by some \( \mathcal{B} \)-cycle \( v_1w_1v_2\ldots w_{k-1}w_kv_1 \). Note that in \( \mathcal{G} \) there might be many other cycles besides those defined by \( \mathcal{B} \)-cycles. We denote \( \text{dist}(v,v') \) the distance between \( v \) and \( v' \) in \( \mathcal{G} \). We additionally denote:

- the set of attributes of \( v \in V \), which contribute to at least one edge in \( \mathcal{G} \):
  \[ W'(v) = \{ w \in W(v) : |V_G(w)| \geq 2 \}; \]
- the set of vertices which have attribute \( w \):
  \[ \mathcal{V}(w) = \{ v \in V : w \in W(v) \}; \]

the set of neighbors of a vertex \( v \in V \) and the set of neighbors of \( S \subseteq V \)
\[ \mathcal{N}(v) = \{ v' \in V \setminus \{ v \} : \mathcal{W}(v) \cap \mathcal{W}(v') \neq \emptyset \}, \quad \mathcal{N}(S) = \bigcup_{v \in S} \mathcal{N}(v) \setminus S. \]

Furthermore, for \( i = 0, 1, 2, \ldots \) we denote \( \mathcal{N}_i(v) \) the set of vertices at distance \( i \) from \( v \) in \( \mathcal{G} \). So that \( \mathcal{N}_0(v) = \{ v \} \) and \( |\mathcal{N}(v)| = |\mathcal{N}_1(v)| =: \deg(v) \) is the degree of \( v \in V \). We denote \( D(k) \) the number of vertices of degree \( k \) in \( \mathcal{G} \) and \( D(k,i) \) the number of vertices \( v \) of degree \( k \) such that \( |W'(v)| = i \). Define

\[ \text{SMALL} = \{ v : |W'(v)| \leq 0.1 \ln n \} \quad \text{and} \quad \text{LARGE} = V \setminus \text{SMALL}. \]

Vertices belonging to SMALL (LARGE) we call small (large). We denote
\[ d_0 = mp(1 - (1 - p)^{n-1}), \quad d_1 = nmp^2, \quad \Delta = \max\{4d_0, 12d_1\} \].

Note that \( d_0 = \mathbb{E}|W'(v)| \), \( d_1 \) is the approximate expected degree and \( np = \Theta(1) \) implies
\[ d_0 \asymp d_1 \asymp \Delta \asymp \ln n. \]
In our calculations we will frequently use the fact that
\[ mp(1 - e^{-np}) = d_0(1 + O(n^{-1})) = (1 + O(n^{-1}))c \ln n. \]

By \( \Pr_G \) and \( \mathbb{E}_G \) we denote the conditional probability and expectation given \( G \). By \( c', c'' \) we denote positive constants that can be different in different places. Throughout the proof the inequalities hold for \( n \) large enough. If it does not influence the result, we consequently omit \( \lfloor \cdot \rfloor \) and \( \lceil \cdot \rceil \) for the sake of clarity of presentation.

### 2.2 Typical intersection graphs

**Lemma 2.** Let \( m, n \to +\infty \). Assume that conditions of Theorem 1 hold. Then there exists a constant \( a_* > 0 \) that may depend on the sequences \( \{p(m,n)\} \) and \( \{c(m,n)\} \), but not on \( n, m \) such that with high probability \( G(n,m,p) \) and \( B(n,m,p) \) have properties \( \text{P0–P8} \) listed below.

**Proof.** All the properties, but \( \text{P2, P8} \), are shown in Appendix. \( \text{P8} \) is shown in Section 3. \( \text{P2} \) is proved in the accompanying paper [19]. \( \square \)

- **P0** \( G \) is connected and has at least one odd cycle.
- **P1** \( |E(G)| - \frac{n^2mp^2}{2} \leq n^{1/2} \ln n. \)
- **P2**
  \[
  \min_{S, |S| \leq n^{1/2}} \frac{e(S, \bar{S})}{2e(S, S) + e(S, \bar{S})} > \frac{1}{50},
  \]
  where, for any \( S \subseteq V \), \( e(S, S) \) is the number of edges induced by set \( S \) and \( e(S, \bar{S}) \) is the number of edges between \( S \) and \( V \setminus S \).
- **P3** For all \( v \in V \) we have \( |W'(v)| \leq \Delta \) and \( \deg(v) \leq \Delta. \)
- **P4** Each adjacent pair \( v, v' \in V \) shares at most \( \max\{2np; 4\} \frac{\ln n}{\ln \ln n} \) common neighbors.
- **P5** Every \( v \in V \) has at least \( |W'(v)| - 1 \) neighbors in \( G \). Every \( v \in \text{LARGE} \) has at least \( (\ln n)/11 \) neighbors in \( G \).
- **P6** For every \( v \) and \( 1 \leq i \leq a_* \ln n/\ln \ln n \) each vertex from \( \mathcal{N}_i(v) \) has at most two neighbors in \( \mathcal{N}_{i-1}(v) \).
- **P7** Any two small vertices are at least \( a_* \ln n/\ln \ln n \) links apart. Each small vertex and each \( B \)-cycle of length at most \( a_* \ln n/\ln \ln n \) are at least \( a_* \ln n/\ln \ln n \) links apart. Any two \( B \)-cycles of length at most \( a_* \ln n/\ln \ln n \) are at least \( a_* \ln n/\ln \ln n \) links apart.
P8 Introduce the numbers

\begin{equation}
\bar{D}(k, i) = \frac{k^i}{k!} (mp^{e^{-np}})^i (np)^{k} n^{1-c}, \quad \bar{D}(k) = \sum_{i=1}^{k} \bar{D}(k, i).
\end{equation}

Here \(c\) is from \([2]\) and \(\frac{k^i}{k!}\) denotes Stirling’s number of the second kind. We remark that \(\bar{D}(k), \bar{D}(k, i)\) are approximations to \(\mathbb{E}D(k), \mathbb{E}D(k, i)\), see \([15]\) below. Define

\[
K_1 = \{1 \leq k \leq 20 : \bar{D}(k) \leq \ln \ln n\}; \\
K_2 = \{21 \leq k \leq \Delta : \bar{D}(k) \leq (\ln n)^2\}; \\
K_3 = \{1, 2, \ldots, \Delta\} \setminus (K_1 \cup K_2).
\]

P8a For \(k \in K_1\) we have \(D(k) \leq (\ln \ln n)^2\), for \(k \in K_2\) we have \(D(k) \leq (\ln n)^4\), and for \(k \in K_3\) we have \(\frac{1}{2}D(k) \leq D(k) \leq \frac{3}{2} \bar{D}(k)\).

P8b If \((c - 1) \geq \ln^{-1/3} n\) then \(D(k) = 0\) for all \(k \leq \ln^{1/2} n\).

P8c Define

\[
i_0 = \lceil (c - 1) \ln n \rceil, \quad k_0 = \lceil i_0 \cdot \max\{10 e^{np}(e^{np} - 1), 2\} \rceil, \\
I = \{k : i_0 \leq k \leq k_0\} \text{ and } \bar{D}(k, i_0) \geq i_0^2\}.
\]

Let \(D^*(k, i_0)\) denote the number of vertices \(v \in \mathcal{V}\) such that \(|W'(v)| = i_0, \deg(v) = k,\) and \(v\) is at distance at least \(\ln n/(\ln \ln n)^3\) from any other vertex \(v' \in \mathcal{V}\) with \(|W'(v')| = i_0\). We have \(I \neq \emptyset\) and \(D^*(k, i_0) \geq \bar{D}(k, i_0)/2\) for \(k \in I\).

We call an instance \(\mathcal{G}\) of \(\mathcal{G}(n, m, p)\) typical if it has properties P0-P8.

3 Vertex degrees in \(\mathcal{G}(n, m, p)\)

In this section we study vertex degrees in \(\mathcal{G}(n, m, p)\) and establish the property P8. We consider vertices \(v \in \mathcal{V}\) with degrees at most \(\Delta\) and with \(|W'(v)| \leq \Delta\). Note that \(\Delta \leq c' \ln n\) for some constant \(c' > 0\), see \([3]\).

Notation and auxiliary results. Notation introduced in this section do not extend to other sections. In the proof we use the following relations for \(1 \leq t < i \leq k\) and \(1 \leq h \leq k-i+1\)

\begin{equation}
\binom{k}{i} \frac{i^{k-i}}{2} \geq \binom{k}{i}, \quad \binom{k}{i} k^{2(t-i)} \geq \binom{k}{i} \geq i^{h-1} \binom{k-h+1}{i}, \quad \binom{k}{i} \geq i^{h-1} \binom{k-h}{i-1}.
\end{equation}

The first inequality is shown in \([13]\). The second one is equivalent to \(\frac{k_i}{j_i} \geq k^{-2}, j \geq 1,\) which follows from the fact that \(j \rightarrow \frac{k_j}{j_i} \geq k^{-2}\), decreases, see \([4]\), combined with
\{\frac{k}{k-1}\} = \binom{k}{2}^{-1}. The third and fourth inequalities follow by multiple application of the recursion relation \{n+1\} = r_{n} + \{n\}.

By \(A_i, A'_i\) and \(B_k, B'_k\) we denote subsets of \(W\) and \(V\) of sizes \(i\) and \(k\) respectively. In what follows it is convenient to think of \(B_k\) and \(A_i\) as realised neighborhoods \(N(v) = B_k\) and \(W'(v) = A_i\) of some \(v \in V\) \((B'_k\) and \(A_i\) refer to respective neighborhoods of another vertex \(u \in V\)). We say that \(A_i\) covers \(B_k\) if each node from \(B_k\) is linked to some vertex from \(A_i\) in \(B(n,m,p)\). For \(i \leq k\) we call \(A_i\) a cover of \(B_k\) if \(A_i\) covers \(B_k\) and no proper subset of \(A_i\) covers \(B_k\) (note that \(A_i\) may cover \(B_k\) not being a cover of \(B_k\)). A cover \(A_i\) is an economic cover (e-cover) if there are exactly \(k\) links between \(A_i\) and \(B_k\). The probability that \(A_i\) is an e-cover of \(B_k\) is

\begin{equation}
(1 - p)^{ik-k} \hat{p}_{k,i}, \quad \text{where} \quad \hat{p}_{k,i} = \binom{k}{i} i! p^i.
\end{equation}

For \(A_t \subseteq A_i\) consider a configuration of links between \(A_t\) and \(B_k\) such that \(A_t\) is an e-cover of \(B_k\) and each node belonging to \(A_t\) \(\setminus A_i\) is linked to a single vertex from \(B_k\). We call such a configuration basic \((A_t/B_k\) basic configuration). Denote the event \(A_{t,A_t}B_k\) contains an \(A_t/B_k\) basic configuration as a subgraph and its probability \(\hat{p}_{t,i} = \Pr\{A_{t,A_t}B_k\}\). For \(i \leq k\) we have

\begin{equation}
(1 - p)^{ik-k} \hat{p}_{k,i} \leq \hat{p}_{k,i}^{*} \leq \hat{p}_{k,i} + \delta_{k,i}, \quad \text{where} \quad \delta_{k,i} = \sum_{t=1}^{i-1} \binom{i}{t} \binom{k}{t} t! p^k (kp)^{i-t}.
\end{equation}

The first inequality is obvious and the second one follows by the union bound: \(\binom{i}{t}\) counts e-covers \(A_t \subseteq A_i\) of size \(t\), \(\binom{k}{i} t!\) counts configurations of \(k\) links between \(A_t\) and \(B_k\) that realize e-cover \(A_t\). Furthermore, \(k^{i-t}\) upper bounds the number of ways to link members of \(A_t\) \(\setminus A_i\) to arbitrary vertices of \(B_k\). Note that \(\delta_{k,i}\) is negligible compared to \(\hat{p}_{k,i}\). By (8),

\begin{equation}
\delta_{k,i} \leq \hat{p}_{k,i} \sum_{t=1}^{i-1} \binom{i}{t} (k^3 p)^{i-t} \leq \hat{p}_{k,i} ((1 + k^3 p)^i - 1) \leq c'' \hat{p}_{k,i} ik^3 p.
\end{equation}

Hence, we have uniformly in \(1 \leq i \leq k \leq \Delta\) that

\begin{equation}
\hat{p}_{k,i}^{*} = \hat{p}_{k,i} (1 + O(n^{-1} \ln^4 n)).
\end{equation}

For \(v \in V \setminus B_k\) define the event \(A_{v,A_i}B_k = \{v\}\) as each node from \(A_i\) is linked to \(v\), there are no links between \(A_i\) and \(V \setminus (B_k \cup \{v\})\), and none element of \(W \setminus A_i\) belongs to \(W'(v)\). Its probability

\begin{equation}
\hat{p}_{k,i} = \Pr\{A_{v,A_i}B_k\} = p^{t}(1-p)^{i(n-k-1)}(1-p+p(1-p)^{n-1})^{m-i} = p^{t} e^{-ip} e^{-d_0(1+O(n^{-1} \ln^2 n))}.
\end{equation}
Similarly, for $B_k, B'_k, A_i, A_j$ and $u \neq v$ such that $(B_k \cup B'_k) \cap \{u, v\} = \emptyset$ and $A_i \cap A_j = \emptyset$ the probability that events $\mathcal{A}_{v,A_i,B_k}$ and $\mathcal{A}_{u,A_j,B'_k}$ occur simultaneously
\[
p'_{k,i,j}(0) := \Pr\{\mathcal{A}_{v,A_i,B_k} \cap \mathcal{A}_{u,A_j,B'_k}\} = p^{i+j}(1-p)^{(i+j)(n-k-1)}((1-p)^2 + 2p(1-p)^{n-1})^{m-i-j}.
\]
Note that $B_k$ and $B'_k$ may intersect. Here $((1-p)^2 + 2p(1-p)^{n-1})^m$ is the probability that none element from $\mathcal{W} \setminus (A_i \cup A_j)$ belong to $\mathcal{W}'(v) \cup \mathcal{W}'(u)$. Furthermore, for $|A_i \cap A_j| = r \in \{1,2\}$ and $v \in B'_k$, $u \in B_k$ the probability $\Pr\{\mathcal{A}_{v,A_i,B_k} \cap \mathcal{A}_{u,A_j,B'_k}\}$ is at most
\[
p'_{k,i,j}(r) := p^{i+j}(1-p)^{(i+j-2r)(n-k-1)}((1-p)^2 + 2p(1-p)^{n-1})^{m-i-j+r}.
\]
We remark that $p'_{k,i,j}(r) \leq c'p'_{k,i,j}(0)$, for $r = 1,2$ and
\[
p'_{k,i,j}(0) = p^{i+j}e^{-(i+j)np}e^{-2d_0} = p'_{k,i}p'_{k,i}(1 + O(n^{-1}\ln^2n)).
\]
For $k = 1,2,\ldots,\Delta$ denote $\tilde{i}_k = \min\{1; [k/\ln\ln n]\}$. Note that in a connected $G$ for each vertex $v$ of degree $k \in [1,\ln\ln n]$ we have $|\mathcal{W}'(v)| \geq 1 = i_k$. Denote
\[
D''(k) = \sum_{i=\tilde{i}_k}^{k} D(k,i), \quad X = \sum_{\ln\ln n < \Delta} \sum_{i=1}^{\tilde{i}_k-1} D(k,i), \quad Y = \sum_{k=1}^{\Delta} \sum_{r \geq 1} D(k,k+r).
\]

**Expected values.** Here we show that uniformly in $i \leq k \leq \Delta$
\[
\mathbb{E}D(k,i) = \bar{D}(k,i)(1 + O(n^{-1}\ln^4n)) \quad \text{and} \quad \mathbb{E}D(k) = \bar{D}(k)(1 + O(n^{-1}\ln^4n)).
\]
Furthermore, for some $\varepsilon > 0$ (depending on the sequence $np = \Theta(1)$) we have
\[
\mathbb{E}Y = O(n^{-\varepsilon}) \quad \text{and} \quad \mathbb{E}X = O(\ln^{-10}n).
\]
Let us prove the first part of (15). Let $v \in \mathcal{V}$ and $1 \leq i \leq k$. Given $A_i, B_k \subset \mathcal{V} \setminus \{v\}$, we have $\Pr\{\mathcal{W}'(v) = A_i, \mathcal{N}(v) = B_k\} = p'_{k,i}p'_{k,i}$. By the union rule,
\[
p_{k,i} := \Pr\{|\mathcal{W}'(v)| = i, |\mathcal{N}(v)| = k\} = \binom{n-1}{k} \binom{m}{i} p^*_k, p^*_{k,i} = \left\{\binom{k}{i}\frac{np^k}{k!}(mpe^{-np})^i e^{-d_0}(1 + O(n^{-1}\ln^4n))\right\}.
\]
In the last step we invoked [9], [12], [13] and used the approximations
\[
\binom{n-1}{k} = \frac{n^k}{k!}(1 + O(k^2/n)), \quad \binom{m}{i} = \frac{m^i}{i!}(1 + O(i^2/m)).
\]
Finally, using (17) we evaluate the expectation
\[
\mathbb{E}D(k,i) = np_{k,i} = \bar{D}(k,i)(1 + O(n^{-1}\ln^4n)).
\]
Let us prove the second part of (15). We split

\[ \mathbb{E}D(k) = \sum_{1 \leq i \leq k} \mathbb{E}D(k, i) + R, \quad R = \sum_{r \geq 1} \mathbb{E}D(k, k + r) \]

and show that \( R = \mathbb{E}D(k, k)O(n^{-1} \ln^2 n) \). Given \( r \geq 1, v \in \mathcal{V} \) and \( B_k \subset \mathcal{V} \setminus \{v\} \), any instance \( \mathcal{B} \) favouring the event \( |\mathcal{W}'(v)| = k + r, \mathcal{N}(v) = B_k \) contains an \( A_k/B_k \) basic configuration for some \( A_k \). In addition, there is \( A'_r \subset \mathcal{V} \setminus A_k \) such that \( \mathcal{W}'(v) = A_k \cup A'_r \). Hence, the probability \( p^*_k = \Pr\{|\mathcal{W}'(v)| > k, \mathcal{N}(v) = B_k\} \) is at most

\[ \binom{m}{k} \sum_{r \geq 1} \binom{m-k}{r} p^*_k p'_k k^{-r}(pk)^r. \]

Here \( \binom{m}{k} \binom{m-k}{r} \) counts pairs \( A_k, A'_r \) and \( (pk)^r \) upper bounds the probability that each node from \( A'_r \) is linked to some vertex from \( B_k \). Using (21) we bound

\[ R = n \binom{n-1}{k} p^*_k \leq n \binom{n-1}{k} \binom{m}{k} p^*_k p'_k k^{-r} R', \]

where

\[ R' = \sum_{r \geq 1} \binom{m-k}{r} p'_k k^{-r}(pk)^r \leq \sum_{r \geq 1} (mp^2k)^r = O(n^{-1} \ln^2 n). \]

Here we used \( p'_k k^{-r} \leq p^* k^{-r} p'_k \). From (17), (19), (22), (23) we obtain \( R = O(n^{-1} \ln^2 n) \mathbb{E}D(k, k) \).

Let us prove the first part of (16). (20) combined with the first relation of (15) imply \( \mathbb{E}Y = O(n^{-1} \ln^2 n) \sum_{1 \leq k \leq \Delta} \bar{D}(k, k) \). Here

\[ \sum_{1 \leq k \leq \Delta} \bar{D}(k, k) = n^{1-c} \sum_{k \geq 1} (nmp^2e^{-np})^k / k! \leq n^{1-c} e^{nmp^2e^{-np}}. \]

Invoking \( mpe(1 - e^{-np})(1 + O(n^{-1})) = c \ln n \), we write the right side in the form \( n^{1-c} e^{nmp(e^np - 1)} e^{-c\ln n (1 + O(n^{-1} \ln n))} \). For \( np = \Theta(1) \) this quantity is \( O(n^{1-c}) \) for some \( \varepsilon > 0 \), since the ratio \( np/(e^np - 1) < 1 \) is bounded away from 1. Hence \( \mathbb{E}Y = O(n^{-\varepsilon} \ln^2 n) \).

Let us prove the second bound of (16). By (15),

\[ \mathbb{E}D(k, i) \leq 2\bar{D}(k, i) \leq \binom{k}{i} (ke)^{i-2} \binom{k}{i} (mp^2e^{-np})^i (np)^k n^{-c} \leq f(k, i). \]

Here we used \( \binom{k}{i} \leq \binom{k}{i} i^{-2} \), see (8), and \( \binom{k}{i} \leq (ke/i)^i \). The inequality \( f(k, i + 1)/f(k, i) \geq 1 \) implies that \( i \rightarrow f(k, i) \) increases for \( 1 \leq i < i_k \). Hence,

\[ \sum_{1 \leq i \leq i_k - 1} \mathbb{E}D(k, i) < i_k f(k, i_k). \]
From (19), (26), (27) and (29) we obtain

$$\ln(i_k f(k, i_k)) \leq i_k + i_k \ln k + (k - 2i_k + 1) \ln i_k + i_k \ln(c_2 \ln n) + k \ln c_1 - \ln k! = -\ln k! + k \ln i_k + O(k) = -k \ln \ln \ln n + O(k).$$

Combining the latter bound with (24) we obtain

$$\mathbb{E}X \leq \sum_{k > \ln \ln n} i_k f(k, i_k) \leq \sum_{k > \ln \ln n} e^{-0.5k \ln \ln \ln n} = o(\ln^{-10} n).$$

Finally, we observe that \( P_0, P_3 \) and (16) imply

(25) \( \Pr \{ \forall v \in V \text{ we have } 1 \leq \deg(v) \leq \Delta \text{ and } i_{\deg(v)} \leq |W'(v)| \leq \deg(v) \} = 1 - o(1). \)

**Concentration.** Here we upper bound the variance \( \text{Var}D''(k) \), see (30) below. Given \( \{u, v\} \subset V \), let \( p''_u \) be the probability that \( |W'(u) \cap W'(v)| \geq 3 \) and \( p''_v \) be the probability that \( \deg(u) = \deg(v) = k, |W'(u) \cap W'(v)| \leq 2 \) and \( i_k \leq |W'(u)|, |W'(v)| \leq k \). We have

(26) \( n(n - 1)p''_+ \leq \mathbb{E}(D''(k)(D''(k) - 1)) \leq n(n - 1)(p''_+ + p''_v), \)

where \( p''_v \leq (\binom{m}{3})p^6 \leq c'n^{-3} \ln^3 n \) is negligible. Let us evaluate \( p''_+ \). We split

(27) \( p''_+ = \sum_{0 \leq r \leq 2} \sum_{i_k \leq i, j \leq k} p_{k, i, j}(r), \)

where \( p_{k, i, j}(r) \) stands for the probability of the event

(28) \( \{ |W'(u)| = j, |W'(v)| = i, |W'(u) \cap W'(v)| = r, \deg(u) = \deg(v) = k \}. \)

We show below that uniformly in \( 1 \leq i_k \leq i, j \leq k \leq \Delta \)

(29) \( p_{k, i, j}(r) = p_{k, i}p_{k, j}(1 + O\left(\frac{\ln n}{n}\right)), \quad p_{k, i, j}(r) \leq O(n^{-r+0.1})p_{k, i}p_{k, j}, \quad r = 1, 2. \)

From (19), (26), (27) and (29) we obtain

$$\mathbb{E}(D''(k)(D''(k) - 1)) = (\mathbb{E}D''(k))^2(1 + O(n^{-0.9})) + O(n^{-1} \ln^3 n).$$

Therefore, the variance

(30) \( \text{Var}(D''(k)) = \mathbb{E}(D''(k)(D''(k) - 1)) + \mathbb{E}D''(k) - (\mathbb{E}D''(k))^2 = (\mathbb{E}D''(k))^2O(n^{-0.9}) + O(n^{-1} \ln^3 n) + \mathbb{E}D''(k). \)
Proof of (29). Let \( r = 0 \). For \( u \neq v, B_k, B'_k, A_i, A_j \) such that \((B_k \cup B'_k) \cap \{u, v\} = \emptyset\) and \( A_i \cap A_j = \emptyset \) we have

\[
\Pr\{A_{A_i, B_k} \cap A_{A_j, B'_k} \cap A_{w, A_i, B_k} \cap A_{u, A_j, B'_k}\} = p_{k,i,j}^* p_{k,i,j}^* p_{k,i,j}^* (0).
\]

Summing over \( A_i, A_j \) with \( A_i \cap A_j = \emptyset \) and over (not necessarily distinct) \( B_k, B'_k \) we obtain

\[
p_{k,i,j}(0) = \binom{n-1}{k}^2 \left( \frac{m}{i} \right) \left( \frac{m-i}{j} \right) p_{k,i,j}^* p_{k,i,j}^* p_{k,i,j}^* (0) = p_{k,i,j} (1 + O(n^{-1} \ln^2 n)).
\]

In the last step we used (14) and (17).

Let \( r = 1 \). For \( 1 \leq i, j \leq k \) we split \( p_{k,i,j}(1) = \sum_{0 \leq h \leq k-1} p_{k,i,j}(1, h) \), where \( p_{k,i,j}(1, h) \) is the probability of the event (28) intersected with the event that \( w = A_i \cap A_j \) has \( h \) neighbors in \( V \setminus \{u, v\} \). We have

\[
p_{k,i,j}(1, h) \leq m \binom{m-1}{i-1} \binom{m-i}{j-1} \binom{n-2}{h} \binom{n-h-2}{k-h-1} \cdot p^h p_{k-h-1,i-1}^* p_{k-h-1,j-1}^* p_{k,i,j}^* (1).
\]

The first line counts triplets \( \{w\}, A_i, A_j \) such that \( \{w\} = A_i \cap A_j \) and triplets \( B_h, B_k, B'_h \) such that \( u \in B_k \subset V \setminus \{v\}, v \in B'_k \subset V \setminus \{u\} \) and \( B_h \subset B_k \subset B'_h \). Furthermore, \( p^h \) is the probability that \( u \) is linked to each vertex of \( B_h \), \( p_{k-h-1,i-1}^* \) is the probability that \( A_i \setminus \{w\} \) covers \( B_k \setminus (B_h \cup \{u\}) \), \( p_{k-h-1,j-1}^* \) is the probability that \( A_j \setminus \{w\} \) covers \( B'_k \setminus (B_h \cup \{v\}) \).

For \( h = k-1 \) we put \( p_{0,i}^* := (p(k-1))^* \) so that \( p_{0,i}^* \) upper bounds the probability that \( B_k \setminus \{u\} \) covers \( A_i \setminus \{w\} \) \( \left( B'_k \setminus \{v\} \right) \) covers \( A_j \setminus \{\{w\}\} \).

Now we show that

\[
p_{k-h-1,i-1}^* \leq c^* \binom{k}{i} \binom{m-1}{i-1} \binom{m-i}{j-1} \binom{n-2}{h} \binom{n-h-2}{k-h-1} p^h p_{k-h-1,i-1}^* p_{k-h-1,j-1}^* p_{k,i,j}^* (1).
\]

The second inequality follows from \( \binom{k}{i} \binom{m-1}{i-1} \binom{m-i}{j-1} \binom{n-2}{h} \binom{n-h-2}{k-h-1} \geq \binom{k-h}{i} \binom{m-1}{i-1} \binom{m-i}{j-1} \binom{n-2}{h} \binom{n-h-2}{k-h-1} \).

Let \( h < k - 1 \). Denote \( \tau = (i-1) \land (k-h-1) \). For \( \tau = i-1 \) we obtain from (10), (11) that

\[
p_{k-h-1,i-1}^* = p_{k-h-1,i-1} (1 + o(1)) \leq c^* \binom{k-h}{i-1} (i-1)! p^h p_{k-h-1,i-1}^*.
\]

Now the inequalities \( \binom{k}{i} \geq i^h \binom{k-h}{i} \geq i^h \binom{k-h-1}{i-1} \), see (8), imply (32). For \( \tau = k-h-1 \) we apply (33) and invoke (32), see (8). We obtain

\[
p_{k-h-1,i-1}^* \leq \sum_{t=1}^{\tau} \binom{i-1}{t} t! \tau^{2(\tau-t)} (p\tau)^{i-t-1}.
\]
Then using \( p r^3 = o(1) \) we upper bound the right side by

\[
p^r (p r)^{i-r-1} (i - 1)! \sum_{t=1}^{r} (pr^3)^{r-t} \leq c' p^r (p r)^{i-r-1} (i - 1)!. \]

Furthermore, we multiply the right side by \( \{ k^2 i^r k - 1 \} \), see (14), and use \( pr i \leq 1 \) to get (32). Proof of (32) is complete.

In the next step we invoke (32) in (31) and apply (18). We obtain

\[
\sum_{0 \leq h < k-1} (p k, i, j)(1, h) \leq c' m^{i+j-1} \left( \frac{np}{k!} \right)^2 \left\{ \begin{array}{c} k \\ i \end{array} \right\} \left\{ \begin{array}{c} k \\ j \end{array} \right\} \sum_{0 \leq h < k-1} (p k, i, j)(1, h) \leq c'' p k, i, j S^*_{k, i, j}(h),
\]

where

\[
S^*_{k, i, j}(h) := \frac{((k h + 1)^2)}{h! (np)^{h+2} (ij)^h}, \quad h < k - 1, \quad \text{and} \quad S^*_{k, i, j}(k - 1) := \frac{k!}{(np)^{k+1}} (p k, i, j)^{i+j-2} i^j j^{k-j}. \]

In the last step of (34) we used \( \sum_{0 \leq h < k-1} S^*_{k, i, j}(h) = O(n^{0.1}) \) uniformly in \( i_k \leq i, j \leq k \leq c' \ln n. \)

Furthermore, we have \( \sum_{0 \leq h < k-1} S^*_{k, i, j}(h) = O(n^{0.1}) \) uniformly in \( i_k \leq i, j \leq k \leq c' \ln n. \)

Hence, (34) imply (29) for \( r = 1. \) For \( r = 2 \) the proof of (29) is much the same.

Text below for preprint only: 1. Routine Calculation of \( \sum_{0 \leq h < k-1} S^*_{k, i, j}(h) = O(n^{0.1}). \)

\[
\sum_{0 \leq h < k-1} S^*_{k, i, j}(h) \leq \frac{k^2}{(np)^2} \sum_{0 \leq h < k-1} (\frac{k^2/(ijnp)}{h!})^h \leq k^2 e^{\frac{k^2}{(ijnp)} \left( \frac{h}{(np)^2} \right)} \leq e^{e^{(\ln n)^2}}.
\]

To upper bound \( S^*_{k, i, j}(k - 1) \) we drop \( (i - 1)! \) and \( (j - 1)! \) and estimate

\[
S^*_{k, i, j}(k - 1) \leq k (i j)^2 \frac{k!}{(ij)^k} (p k, i, j)^{i+j-2} \frac{1}{(np)^k} \leq k^3 e^{\frac{k!}{i_k^2 (np)^k}} \leq k^3 \frac{1}{i_k^2 (np)^k} \leq n^{0.1}.
\]

Here \( k \leq (\ln \ln n)^3 \Rightarrow (k/(np))^k \leq (c' k)^k \leq n^{0.1} \) and \( k > (\ln \ln n)^3 \Rightarrow k/i_k^2 \leq (\ln n)^{-1}. \)

End of Routine Calculation.

Text below for preprint only: 2. Bounds for \( p_{k, i, j}(2). \)

Let \( r = 2. \) We split \( p_{k, i, j}(2) = \sum_{0 \leq h < k-1} p_{k, i, j}(2, h), \) where \( p_{k, i, j}(2, h) \) is the probability of the event \( \{ \} \) intersected with the event that \( \{ w, z \} = A_i \cap A_j \) has \( h \) neighbors in \( V \setminus \{ u, v \}, \) i.e., \( |V(z) \cup V(w)| = h + 2. \) We have

\[
\sum_{0 \leq h < k-1} (p k, i, j)^2 \leq \left( \frac{m}{2} \right) \left( \frac{m - 2}{i - 2} \right) \left( \frac{m - i}{j - 2} \right) \left( \frac{n - 2}{h} \right) \left( \frac{n - 2 - h}{k - 1 - h} \right)^2 \cdot (2p)^h p_{k, i, j}(2, h).
\]

\[
\sum_{0 \leq h < k-1} (p k, i, j)^2 \leq \left( \frac{m}{2} \right) \left( \frac{m - 2}{i - 2} \right) \left( \frac{m - i}{j - 2} \right) \left( \frac{n - 2}{h} \right) \left( \frac{n - 2 - h}{k - 1 - h} \right)^2 \cdot (2p)^h p_{k, i, j}(2, h).
\]
Here \( (\begin{array}{cc}
\m & \n-2 \\
\j & \j-2
\end{array}) \) counts triplets \( \{w, z\}, A_i, A_j \) such that \( \{w, z\} = A_i \cap A_j \). \( (\begin{array}{cc}
\v & \v-2 \\
\k & \k-2
\end{array}) \) counts triplets \( B_h, B_k, B_k' \) such that \( u \in B_k \subset V \setminus \{v\}, v \in B_k' \subset V \setminus \{u\} \) and \( B_h \subset B_k \cap B_k' \) is the set of neighbors of \( \{w, z\} \) from \( V \setminus \{u, v\} \). Furthermore, \( (2p)^h \) upper bounds the probability that \( \{w, z\} \) covers \( B_h \). \( p_{k-h-1,i-2} \) is the probability that \( A_i \setminus \{w, z\} \) covers \( B_k \setminus (B_h \cup \{u\}) \), \( p_{k-h-1,i-2}^* \) is the probability that \( A_j \setminus \{w, z\} \) covers \( B_k' \setminus (B_h \cup \{v\}) \).

We only prove the first inequality. Denote \( \tau := (i - 2) \wedge (k - h - 1) \). For \( \tau = i - 2 \) we have, see (10), (11),

\[ p_{k-h-1,i-2}^* \leq c' \left( \begin{array}{c}
\tau \\
i
\end{array} \right) \left( \begin{array}{c}
\tau \\
2
\end{array} \right)! p^{k-h-1} \leq \left( \begin{array}{c}
k \\\ni
\end{array} \right) \left( \begin{array}{c}
2 \\\n2
\end{array} \right)! \tau^{i-k}(pk)^{i-k}. \]

We only prove the first inequality. Denote \( \tau := (i - 2) \wedge (k - h - 1) \). For \( \tau = i - 2 \) we have, see (10), (11),

\[ p_{k-h-1,i-2}^* \leq c' \left( \begin{array}{c}
k \\\ni
\end{array} \right) \left( \begin{array}{c}
2 \\\n2
\end{array} \right)! \tau^{i-k}(pk)^{i-k}. \]

In the last step we used \( \left( \begin{array}{c}
k \\\ni
\end{array} \right) \geq \tau^{i-k} \left( \begin{array}{c}
i \\\ni
\end{array} \right) \), see (8). For \( \tau = k - h - 1 \) we have

\[ p_{k-h-1,i-2}^* \leq \sum_{t=1}^{\tau} \left( \begin{array}{c}
\tau \\
t
\end{array} \right) \left( \begin{array}{c}
\tau \\
2
\end{array} \right)! t^{i-k}(pk)^{i-k} \leq \sum_{t=1}^{\tau} \left( \begin{array}{c}
\tau \\
t
\end{array} \right) \left( \begin{array}{c}
\tau \\
2
\end{array} \right)! \tau^{i-k}(pk)^{i-k}. \]

We multiply the right side by \( \left( \begin{array}{c}
k \\\ni
\end{array} \right) \tau^{i-k} \geq 1 \) and obtain

\[ p_{k-h-1,i-2}^* \leq c' \left( \begin{array}{c}
k \\\ni
\end{array} \right) \left( \begin{array}{c}
2 \\\n2
\end{array} \right)! \tau^{i-k}(pk)^{i-k}. \]

Note that for large \( n \) we have \( p\tau i < 1 \) since \( p\tau i \leq pk^2 = o(1) \).

Invoking (36) and (18) in (35) we obtain

\[ p_{k,i,j}(2, h) \leq c' m^{i+j-2} \left( \frac{np^2 k!}{(k!)^2} \right) \left( \begin{array}{c}
k \\\ni
\end{array} \right) \left( \begin{array}{c}
j \\\nj
\end{array} \right) \left( \begin{array}{c}
k \\\j\cdot
\end{array} \right) \cdot S_{k,i,j}(2, h) \leq c'' p_{k,i} p_{k,j} \bar{S}_{k,i,j}(h), \]

where

\[ \bar{S}_{k,i,j}(h) = \frac{2^h \binom{(k)_{h+1}}{2}}{h!(np)^{h(ij)_{h-1}}}, \quad h < k - 1, \quad \text{and} \quad \bar{S}_{k,i,j}(k - 1) = \frac{2^{k-1} k! k^2 (pk)^{i+j-4} \cdot i-k \cdot j-k}{(np)^{k+1} (i-2)!(j-2)!}. \]
In the last step of (37) we used \( p'_{k,i,j}(2) \leq c' p'_{k,i,j}(0) \leq c'' p'_{k,i,j} \), see (14), and (17). The same argument as above yields \( \sum_{0 \leq k \leq k-1} S_{k,i,j}(\epsilon) = O(n^{0.1}) \) uniformly in \( i_k \leq i, j \leq k \leq c' \ln n \). These bounds together with (37) imply (29) for \( r = 2 \).

End of Bounds for \( p_{k,i,j}(2) \).

End of the preprint only text.

Now we prove P8. In view of (25) it suffices to show that P8 holds for \( D''(k) \).

Proof of P8a. The second part of P8a follows from (15) by Markov’s inequality. To prove the first part we show that

\[
1 - \Pr\left\{ \frac{1}{2} \bar{D}(k) \leq D''(k) \leq \frac{3}{2} \bar{D}(k), \; k \in K_3 \right\} \leq \Pr\{ \cup_{k \in K_3} \mathcal{B}_k \} \leq \sum_{k \in K_3} \Pr\{ \mathcal{B}_k \} = o(1).
\]

Here we denote for short \( \mathcal{B}_k = \{ |D''(k) - \bar{D}(k)| > \bar{D}(k)/2 \} \). The first two inequalities are obvious. To prove the last bound we show that \( \Pr\{ \mathcal{B}_k \} \leq O(n^{-0.9}) + 2/\bar{D}(k) \). From (15), (16) we obtain

\[
|\bar{D}(k) - \mathbb{E}D''(k)| \leq |\bar{D}(k) - \mathbb{E}D(k)| + \mathbb{E}X + \mathbb{E}Y = O(n^{-1} \ln^2 n) \bar{D}(k) + O(\ln^{-9} n).
\]

For \( k \in K_3 \) we have \( 0.9 \leq \mathbb{E}D''(k)/\bar{D}(k) \leq 1.1 \). Now, by Chebyshev’s inequality and (30),

\[
\Pr\{ \mathcal{B}_k \} \leq \Pr\{ |D''(k) - \mathbb{E}D''(k)| \geq \bar{D}(k)/3 \} \leq O(n^{-0.9}) + 1.1/\bar{D}(k).
\]

Proof of P8b. We need to show that \( \Pr\{ \deg(v) > \ln^{1/2} n, \; \forall v \in \mathcal{V} \} = 1 - o(1) \). In view of P5 it suffices to prove that \( p_0 := \Pr\{ \exists v \in \mathcal{V} : |\mathcal{W}'(v)| < 2 \ln^{1/2} n \} = o(1) \). Note that each \( |\mathcal{W}'(v)| \) has binomial distribution with mean \( d_0 = c \ln n \), see (2). By the union bound and Chernoff’s inequality, see (2.6) in [10], we have

\[
p_0 \leq n \Pr\{ |\mathcal{W}'(v)| \leq \frac{2d_0}{\sqrt{n \ln n}} \} \leq n \exp\left\{ -\left( 1 - \frac{2}{\sqrt{n \ln n}} + \frac{2}{\sqrt{n \ln n}} \ln \frac{2}{\sqrt{n \ln n}} \right) d_0 \right\} = o(1).
\]

Proof of P8c. Let us prove that \( I \neq \emptyset \). We begin with showing auxiliary inequality (38), see below. Given \( y > 0 \), \( q > 1 \) and integer \( i > 1 \), let \( r = \lceil i + iqy \rceil \). We have

\[
\sum_{k \geq r} \binom{k}{i} \frac{y^k}{k!} \leq \frac{y^i}{i!} \sum_{k \geq r} \binom{y(i)q-y-i}{k-i} \frac{q}{(k-i)!} \leq \frac{y^i}{i!} \frac{(y/q)^r}{(r-i)!} q \leq \frac{y^i}{i!} \sqrt{2\pi q} (r-i) q^{-1}.
\]

In the first step we use \( \binom{k}{i} \leq \binom{k}{i}^{i-k} \). In the second step we upper bound the series by the geometric series \( 1 + q^{-1} + q^{-2} + \cdots \) using the fact that the ratio of two consecutive terms is at most \( q^{-1} \). The last inequality follows by Stirling’s approximation. Choosing \( q = 2e \) we upper bound the right side by \( y^i/(2i!) \). Combining this bound with the identity \( \sum_{k \geq i} \binom{k}{i} y^k/k! = (e^y - 1)^i/i! \) and inequality \( e^y - 1 > y \) we obtain for any \( r \geq \lceil i + 2ey \rceil \)

\[
\sum_{k=i}^{r} \binom{k}{i} \frac{y^k}{k!} \geq \frac{(e^y - 1)^i}{i!} - \frac{1}{2 y^{i-1}} \geq \frac{1}{2} \frac{(e^y - 1)^i}{i!}.
\]
For $i = i_0$, $r = k_0$ and $y = np$ this inequality implies
\begin{equation}
\sum_{k=0}^{k_0} \bar{D}(k, i_0) \geq \frac{1}{2} \left( mp(1 - e^{-np})i_0 \right)^{i_0} n^{1-c} \geq 1 + O(i_0/n) \left( \frac{c \ln n}{i_0} \right)^{i_0}.
\end{equation}

In the second step we used \((2)\) and Stirling’s approximation. Furthermore, by the assumption \((c - 1) \ln n \to +\infty\), we have for large $n$ that $c/(c - 1) \geq (c \ln n)/i_0 > 2c/(2c - 1)$. We conclude that the right side of \((39)\) grows exponentially in $i_0$. This proves $I \neq \emptyset$.

Denote, for short, $D^*(k) = D(k, i_0) - D^*(k, i_0)$. We show below that
\begin{align}
(40) & \quad p_1 := \Pr\{D(k, i_0) \geq 0.8\bar{D}(k, i_0), \forall k \in I\} = 1 - o(1), \\
(41) & \quad p_2 := \Pr\{D^*(k) < 0.3\bar{D}(k, i_0), \forall k \in I\} = 1 - o(1).
\end{align}

Note that \((40)\) and \((41)\) imply \(P8c\). It remains to prove \((40), (41)\).

Proof of \((40)\). By \((15)\) we have $i_0^2 \leq \bar{D}(k, i_0) \leq 1.1\bar{E}(k, i_0)$, $\forall k \in I$. Combining the union bound and Chebychev’s inequality we obtain
\begin{equation}
1 - p_1 \leq \sum_{k \in I} \Pr\{D(k, i_0) < 0.88 \bar{E}(k, i_0)\} \leq \sum_{k \in I} 70 \frac{\operatorname{Var}(D(k, i_0))}{(\bar{E}(k, i_0))^2} = o(1).
\end{equation}

In the last step we used $\bar{E}(k, i_0) \geq i_0^2/1.1$ and invoked the approximation
\[
\operatorname{Var}(D(k, i_0)) = (\bar{E}(k, i_0))^2 O(n^{-0.9}) + O(n^{-1} \ln n) + \bar{E}(k, i_0),
\]
which is shown using the same argument as in \((26), (29), (30)\) above.

Proof of \((41)\). We show below that $\bar{E}D^*(k) \leq c' D(k, i_0) \ln^{-3} n$. Then combining the union bound and Markov’s inequality we obtain
\begin{equation}
1 - p_2 \leq \sum_{k \in I} \Pr\{D^*(k) \geq 0.3\bar{D}(k, i_0)\} \leq c' |I| \ln^{-3} n = o(1).
\end{equation}

Given $k$ we upper bound $\bar{E}D^*(k)$ by the expected number of vertex pairs $v \neq u$ such that $|\mathcal{W}'(v)| = |\mathcal{W}'(u)| = i_0$, $\deg(v) = k$ and $\operatorname{dist}(u, v) \leq \ln/\ln \ln n$). The pairs with different intersection sizes $|\mathcal{W}'(v) \cap \mathcal{W}'(u)| = r$ will be counted separately.

For $r = 1$ the expected number of pairs is upper bounded by
\begin{equation}
n(n - 1) \cdot \binom{m}{i_0} \binom{n - 2}{k - 1} \binom{m - i_0}{i_0 - 1} \cdot p^{i_0}(1 - p)^{i_0(n-k-1)} p_{k,i_0}^* \\
\cdot (p(1 - p)(1 - (1 - p)^2)^{i_0} \cdot ((1 - p)^2 + 2p(1 - p)^{n-1})^{m-2i_0+1}.
\end{equation}

Here $n(n - 1)$ counts ordered pairs $v \neq u$. $\binom{m}{k-1}(m-i_0)$ counts non intersecting subsets $A_{i_0}, A_{i_0-1} \subset W$ and $B_k \subset V \setminus \{v\}$ with $u \in B_k$ that can realise $\mathcal{W}'(v), \mathcal{W}'(u) \setminus \mathcal{W}'(v)$ and $\mathcal{N}(v)$ respectively. Furthermore, $p_{k,i_0}^* = \Pr\{A_{i_0}, B_k\}$ and $p^{i_0}(1 - p)^{i_0(n-k-1)}$ is the
conclude that for 

$$r(1-p)(1-(1-p)^{n-2})^{i_0-1}$$ is the probability that each element of $A_{i_0-1}$ is linked to $u$, none to $v$ and each has more than one neighbor in $\mathcal{B}$. Finally, $p(1-p)(1-(1-p)^{n-2})^{i_0-1}$ is the probability that no element of $\mathcal{W} \setminus (A_{i_0} \cup A_{i_0-1})$ belongs to $\mathcal{W}'(v) \cup \mathcal{W}'(u)$.

Using (9), (12), see also (13), we show that (44) is at most $\bar{D}(k, i_0)k \frac{d_k^{i_0-1}}{(i_0-1)!} e^{-d_0}(1 + o(1))$. Next we bound

$$\frac{d_k^{i_0-1}}{(i_0-1)!} \leq \frac{d_k^{i_0}}{i_0!}$$

and use Stirling’s approximation to $i_0!$. We have

$$\frac{d_k^{i_0}}{i_0!} e^{-d_0} \leq \left(\frac{d_k}{i_0}\right)^i_0 e^{i_0-d_0} \leq \left(\frac{c}{c-1}\right)^{1+(c-1)\ln n} e^{1-\ln n} \leq c \frac{\ln n}{n} \left(\frac{c}{c-1}\right)^{(c-1)\ln n}.$$  

In the last step we used $c/(c-1) = o(\ln n)$. Furthermore, for $c > 1$, $c = \Theta(1)$ there exists $\varepsilon > 0$ such that $(c/(c-1))^{c-1} < e^{1-\varepsilon}$ uniformly in $n, m$ (because $x \to (1+x^{-1})^x$ increases for $x > 0$ and approaches $e$ as $x \to +\infty$). Hence, the right side is bounded by $n^{-\varepsilon/2}$. We conclude that for $r = 1$ the expected number of pairs is at most $c'' \bar{D}(k, i_0)n^{-\varepsilon/3}$.

For $r = 2$ we similarly upper bound the expected number of pairs by $c'' \bar{D}(k, i_0)n^{-\varepsilon/3}$. Preprint text only: 3. An upper bound for the expected number of pairs for $r = 2$.

For $r = 2$ the expected number of pairs is upper bounded by

We have

$$n(n - 1) \cdot \binom{m}{i_0} \left(\begin{array}{c} m - 2 \\ k \\ i_0 - 2 \end{array}\right) \cdot p^i_0 (1-p)^{i_0(n-k-1)}p_{k,i_0}^* \cdot \left((1-p)^{2} + 2p(1-p)^{n-1}\right)^{m-2i_0+2} \leq \bar{D}(k, i_0)k \frac{d_k^{i_0-2}}{(i_0-2)!} e^{-d_0}(1 + o(1)) \leq c'' \bar{D}(k, i_0) \ln^{-4} n.$$  

End of an upper bound for the expected number of pairs for $r = 2$.

End of the preprint only text.

For $r \geq 3$ the expected number of pairs is at most $n(n - 1)p'' \leq c'n^{-1}\ln^3 n \leq c'' \bar{D}(k, i_0) \ln^{-4} n$, for $k \in I$. For $r = 0$ we consider separately the pairs that are in the distance $\text{dist}(u, v) = t \in \{2, 3, \ldots\}$. For $t = 1$, the expected number of pairs is at most

$$\left(\begin{array}{c} n(n - 1) \\ k \end{array}\right) \sum_{i_0 = 1}^{\min(n-k, m)} \binom{m}{i_0} \binom{m - 2}{k} p^i_0 (1-p)^{i_0(n-k-1)}p_{k,i_0}^* \cdot \left((1-p)^{2} + 2p(1-p)^{n-1}\right)^{m-2i_0} \cdot (mn^{2})^{t-2}k_{i_0}p.$$

Here $n(n - 1)$ counts ordered pairs $v \neq u$, $\binom{n-2}{k}$ counts sets $B_k \subset \mathcal{V} \setminus \{u, v\}$ that realise $\mathcal{N}(v)$, $\binom{m}{i_0}$ counts non-intersecting pairs $A_i, A'_i \subset \mathcal{W}$ that realise $\mathcal{W}'(v)$, $\mathcal{W}'(u)$. Furthermore, $(mn^{2})^{t-2}k_{i_0}p$ upper bounds the number of paths connecting $B_k$ with $A'_i$ and having $2t - 1$ links. Proceeding as in (44), (45) we upper bound (46) by $c'' \bar{D}(k, i_0) \frac{d_k^{i_0}}{i_0!} e^{-d_0} \ln^t n$, where $\frac{d_k^{i_0}}{i_0!} e^{-d_0} \leq n^{-\varepsilon/2}$. Hence (46) is at most $c'' \bar{D}(k, i_0)n^{-\varepsilon/3}$. 17
4 Probability of the first visit of a vertex

Given a connected graph \( G \) on the vertex set \( \mathcal{V} \), let \( \mathcal{W}_u \) denote the simple random walk starting from \( u \in \mathcal{V} \). Denote \( P_u^{(t)}(v) = \Pr\{\mathcal{W}_u(t) = v\} \), where \( \mathcal{W}_u(t) \) denotes the vertex visited at time \( t = 0, 1, 2, \ldots \) (so that \( P_u^{(0)}(u) = 1 \)). Assuming that \( G \) admits a stationary distribution \( \pi = \{\pi_v, v \in \mathcal{V}\} \) (i.e., \( \lim_{t \to +\infty} P_u^{(t)}(v) = \pi_v \), for all \( u, v \in \mathcal{V} \)) we have \( \pi_v = \deg(v)|\mathcal{E}(G)|^{-1} \). Given integer \( T > 0 \) and \( v \in \mathcal{V} \), denote

\[
R_{T,v}(z) = \sum_{j=0}^{T-1} \Pr\{\mathcal{W}_v(j) = v\} z^t, \quad z \in \mathbb{C}
\]

and, for \( t \geq T \), let \( \mathcal{A}_t(v) \) be the event that \( \mathcal{W}_u \) does not visit \( v \) in steps \( T, T + 1, \ldots, t \).

In the following lemma we consider a sequence of connected graphs \( \{G_n\} \), where \( n \) is the number of vertices of \( G_n \). We assume that each graph admits a stationary distribution \( \pi = \pi(n) \). Furthermore, we assume that \( T = T(n) \) is such that, for \( t \geq T \)

\[
\max_{u,v \in \mathcal{V}} \left| P_u^{(t)}(v) - \pi_v \right| \leq n^{-3}.
\]

The following lemma was proved in [6]. It is stated there as Corollary 7.

**Lemma 3.** Suppose that \( T = T(n) \) satisfies (47) and

(i) there exist \( \Theta > 0 \), \( C_0 > 0 \) and \( n_0 > 0 \) such that uniformly in \( n > n_0 \) we have

\[
\min_{|z| \leq 1 + (C_0 T)^{-1}} \left| R_{T,v}(z) \right| \geq \Theta,
\]

(ii) \( T^2 \pi_v = o(1) \) and \( T \pi_v = \Omega(n^{-2}) \).

Then there exists

\[
p_v = \frac{\pi_v}{R_{T,v}(1)(1 + O(T \pi_v))}
\]

such that for all \( t \geq T \)

\[
\Pr\{\mathcal{A}_t(v)\} = \frac{1 + O(T \pi_v)}{(1 + p_v)^t} + o(e^{-t/(2C_0 T)}).
\]

We note that the bounds \( O(T \pi_v) \) and \( o(e^{-t/(2C_0 T)}) \) in (48) and (49) hold uniformly in \( u, v \) and \( n > n_0 \), provided that conditions (i), (ii) hold uniformly in \( v \) and \( n > n_0 \).
4.1 The expected number of returns

Given an instance $\mathcal{G}$ of the random intersection graph $\mathcal{G}(n,m,p)$, we consider the simple random walk $\mathcal{M}_v$ on $\mathcal{G}$ starting from $v \in \mathcal{V}$. Let $r_i = \Pr_{\mathcal{G}} \{ \mathcal{M}_v(i) = v \}$ denote the probability that the walk returns to $v$ at time $i$ (so that $\mathcal{M}_v(0) = v$ implies $r_0 = 1$). We remark that $z \to R_{T,v}(z) = \sum_{i=0}^{T-1} r_i z^i$ is a random function depending on the realised graph $\mathcal{G}$. Furthermore, given a pair of vertices $x, y$ of $\mathcal{G}$ that are in a distance at least 20, let $\mathcal{G}_x$ denote the graph obtained from $\mathcal{G}$ by merging $x$ and $y$. Here $\mathcal{G}_x = \{ x, y \}$ represents the new vertex obtained from the merged pair. We denote $\deg(x) = \deg(x) + \deg(y)$ the degree of $x$. In $\mathcal{G}_x$ we consider the simple random walk $\mathcal{M}_{\mathcal{G}_x}$ starting from $x$. We denote $R_{T,\mathcal{G}_x}(z) = R_{T,x,y}(z) = \sum_{i=0}^{T-1} r_i z^i$, where $r_i = \Pr_{\mathcal{G}} \{ \mathcal{M}_{\mathcal{G}_x}(i) = \mathcal{G}_x \}$. Furthermore, let $\tau_v$ ($\tau_{\mathcal{G}_x}$) denote the time of the first return of $\mathcal{M}_v$ ($\mathcal{M}_{\mathcal{G}_x}$) to $v$ ($\mathcal{G}_x$) in the interval $[1, +\infty)$. In the lemma below we assume that $m, n, p$ satisfy conditions of Theorem 1.

Lemma 4. Let $C_0 > 0$. Assume that $T = T(n,m) \to \infty$ and $T = o(ln^3 n)$. We have with high probability

\begin{align}
(50) & \quad \sup_{|x| \leq 1 + (C_0 T)^{-1}} |R_{T,v}(z)| = 1 + O(ln^{-1} n) \quad \forall \ v \in \mathcal{V}, \\
(51) & \quad \sup_{|x| \leq 1 + (C_0 T)^{-1}} |R_{T,x,y}(z)| = 1 + O(ln^{-1} n) \quad \forall x, y \in \mathcal{V} \text{ with dist}(x,y) \geq 20, \\
(52) & \quad R_{T,v}(1) = 1 + \bar{p}_v + O(ln^{-2} n), \quad \text{where} \quad \bar{p}_v = \Pr_{\mathcal{G}} \{ \tau_v \leq T - 1 \} \approx ln^{-1} n, \\
(53) & \quad R_{T,x,y}(1) = 1 + \bar{p}_{\mathcal{G}_x} + O(ln^{-2} n), \quad \text{where} \quad \bar{p}_{\mathcal{G}_x} = \Pr_{\mathcal{G}} \{ \tau_{\mathcal{G}_x} \leq T - 1 \} \approx ln^{-1} n, \\
(54) & \quad \frac{\deg(x)}{1 + \bar{p}_{\mathcal{G}_x}} = \left( \frac{\deg(x)}{1 + \bar{p}_v} + \frac{\deg(y)}{1 + \bar{p}_y} + O(\deg(x)ln^{-2} n) \right).
\end{align}

Furthermore, $\{50\}, \{52\}$, respectively, $\{51\}, \{53\}, \{54\}$ hold uniformly in $v \in \mathcal{V}$, respectively, uniformly in $x, y \in \mathcal{V}$ satisfying dist($x,y) \geq 20$.

Proof. We establish $\{50\} \{54\}$ for $\mathcal{G}$ having properties P1-P8, see Lemma 2.

Proof of $\{50\}$. We have $\left| R_{T,v}(z) - 1 \right| \leq \left| \sum_{i=1}^{T-1} r_i z^i \right|$ and for $|z| \leq 1 + (C_0 T)^{-1}$

\begin{align}
\sum_{i=1}^{T-1} r_i z^i & \leq \sum_{i=1}^{T-1} r_i |z|^i \leq (1 + (C_0 T)^{-1})^T \tilde{R} \leq e^{C_0^{-1}} \tilde{R}, \quad \tilde{R} := \sum_{i=1}^{T} r_i.
\end{align}

We show below that $\tilde{R} = O(ln^{-1} n) + O(T ln^{-5} n)$ uniformly in $v \in \mathcal{V}$. Note that $\tilde{R}$ is the expected number of returns to $v$ of the random walk $\mathcal{M}_v$ in the time interval $[1, T]$.

We begin with an observation, denoted (O), about random walks on directed graph with the vertex set $\{0, 1, 2, 3\}$, where 3 is an absorbing state. Assume, that the transitional probabilities $p_{0,3} = p_{0,2} = 0$ and $0 < p_{0,5}, p_{0,1}, p_{1,2}, p_{2,1}, p_{2,3} < 1$ are...
fixed. The walk starts at 0 and it is allowed to make \( t \) steps. Then for any \( t \), the expected number of returns to 0 before visiting 2 is maximized if we choose

\[
p_{0,0} = 1, \quad p_{0,1} = p_{5,2} = p_{1,5} = p_{2,5} = 0, \quad p_{1,0} = 1 - p_{1,2}, \quad p_{2,0} = 1 - p_{2,1} - p_{2,3}.
\]

Case (1). Assume that \( N_1(v) \cup \cdots \cup N_7(v) \) contains no small vertices. The random walk \( W'(i) = \min \{7, \text{dist}(v, W_v(i)) \} \) moves along the path of length 7 and has the state space \( \{0, 1, \ldots, 7\} \). Its transitional probabilities satisfy inequalities

\[
(56) \quad p'_{j+1,j} \leq c' / \ln n, \quad p'_{j,j} \leq c' / \ln \ln n, \quad 0 \leq j \leq 6,
\]

where \( c' \) is an absolute constant. Indeed, by \( \text{P6} \), every \( u \in N_{j+1}(v) \) is adjacent to at most two vertices from \( N_j(v) \). Now \( \text{P5} \) implies \( p'_{j+1,j} = O(\ln^{-1} n) \). Furthermore, by \( \text{P4} \), each of these vertices shares with \( u \) at most \( a_s(\ln n) / \ln \ln n \) common neighbors from \( N_{j+1}(v) \). In addition, by \( \text{P7} \), there can be at most 2 vertices in \( N_{j+1}(v) \) adjacent to \( u \) and having no common neighbors with \( u \) located in \( N_i(v) \). Therefore, every \( u \in N_{j+1}(v) \) can have at most \( 2 + 2a_s(\ln n) / \ln \ln n \) neighbors in \( N_{j+1}(v) \) altogether. Now \( \text{P5} \) imply \( p'_{j+1,j+1} = O(1 / \ln n) \), for \( 0 \leq j \leq 5 \). Finally, we obviously have \( p'_{0,1} = 1 \).

The random walk \( W' \) is lazy: it may stay at state \( j > 1 \) for several consecutive steps. Let \( W'' \) be the fast random walk defined by \( W' \) as follows: \( W'' \) only makes a step when \( W' \) changes its state. In the latter case the moves of \( W' \) and \( W'' \) coincide. Its transitional probabilities

\[
(57) \quad p''_{0,1} = p''_{7,6} = 1, \quad p''_{j+1,j} = p'_{j+1,j} / (1 - p'_{j,j}) \quad \text{for} \quad 0 \leq j \leq 5 \quad \text{and} \quad p''_{i,i} = 0 \quad \forall i.
\]

We have \( \bar{R} = R' \leq R'' \), where \( R' \) and \( R'' \) denote the expected numbers of returns to 0 within the first \( T \) steps of respective random walks \( W' \) and \( W'' \). We split \( R'' = R''_1 + R''_2 \), where \( R''_1 \) is the expected number of returns to 0 before the first visit of 7. We have \( R''_1 \leq E_G X \), where \( X \) is the number of backward steps made by \( W'' \) before visiting 7. The inequality \( p''_{j+1,j} \leq c'(1 + o(1)) / \ln n \), see [56], [57], implies

\[
\Pr_G \{X = k\} \leq ((6c' + o(1)) / \ln n)^k, \quad k \geq 0.
\]
Hence $E_G X = O(\ln^{-1} n)$ and we obtain $R''_1 = O(\ln^{-1} n)$. Furthermore, after visiting 7 the random walk $\mathcal{W}''$ moves to 6. Starting from 6 the walk may visit 0 before visiting 7 again, we call such event a success. The probability of success is $O(\ln^{-6} n)$ see, e.g., formula (30) in \cite{9}. The expected number of successes within the first $T$ steps of the random walk is at most $O(T \ln^{-6} n)$. Hence $R'_2$, the expected number of returns to 0 after the first visit of 7, is at most $O(T(1 + R''_1) \ln^{-6} n)$. Here $R''_1$ accounts for the returns to 0 after a success and before visiting 7 again. We conclude that $\bar{R} \leq R'_1 + R''_1 = O(\ln^{-1} n) + O(T \ln^{-6} n)$.

**Case (2).** Assume that $\mathcal{N}_{k+1}(v)$ contains a small vertex, say $\overline{v}$, for some $0 \leq k \leq 5$. Note that, by \cite{7}, there is no other small vertex in $\mathcal{G}$ within the distance $O(\ln \ln n)$ from $v$. Now we define $\mathcal{W}'(i) = \min\{7, \text{dist}(v, \mathcal{W}_v(i))\}$, for $\mathcal{W}_v(i) \neq \overline{v}$, and put $\mathcal{W}'(i) = \overline{k}$, for $\mathcal{W}_v(i) = \overline{v}$. It is a random walk on the state space $S_k = \{0, 1, \ldots, k, \overline{k}, k+1, \ldots, 7\}$. Furthermore, let $\mathcal{W}''$ be the corresponding fast random walk on $S_k$: $\mathcal{W}''$ only makes a step when $\mathcal{W}'$ changes its state and in the latter case the moves of $\mathcal{W}'$ and $\mathcal{W}''$ coincide. Arguing as in (56), (57) we obtain the corresponding inequalities for the transitional probabilities $\overline{p}_{i,j}$ of $\mathcal{W}''$

\begin{equation}
\overline{p}_{k,\overline{k}}, \overline{p}_{r,\overline{k}}, \overline{p}_{j,j-1} \leq c'/\ln n, \quad r = k+1, k+2, \quad 1 \leq j \leq 6.
\end{equation}

Note that $\overline{p}_{k,\overline{k}}, \overline{p}_{r,\overline{k}}, \overline{p}_{j,j+1} > 0$, $0 \leq j \leq 6$. Furthermore, we have $\overline{p}_{k,r}, \overline{p}_{r,\overline{k}} > 0$ whenever $\overline{v}$ has a neighbor in $\mathcal{N}_r(v)$, $r = k+1, k+2$. Moreover, we have $\overline{p}_{\overline{r},\overline{6}} = 1$ and $\overline{p}_{j,j} = 0$ for all $j \in S_k$. Finally, $\overline{p}_{0,1} = 1$ for $k > 0$ and $\overline{p}_{0,1} + \overline{p}_{0,1} = 1$ for $k = 0$. All the other transitional probabilities $\overline{p}_{i,j}$ are zero. From now on we consider the cases $k = 0$ and $k \geq 1$ separately.

Assume that $k = 0$, i.e., $\overline{v} \in \mathcal{N}_1(v)$. Let $\mathcal{W}^*$ be the random walk on $S_0$ starting from 0 and with transitional probabilities $\overline{p}^*_{i,j} = \overline{p}_{i,j}$ for each $(i, j) \in S_0 \times S_0$, but

\begin{equation}
\overline{p}^*_{0,0} = 1, \quad \overline{p}^*_{r,r} = \overline{p}^*_{r,\overline{r}} = 0, \quad \overline{p}^*_{r,0} = \overline{p}_{r,0} + \overline{p}_{r,\overline{r}}, \quad r = 1, 2.
\end{equation}

We have $\bar{R} = R' \leq R'' \leq R^*$, where $R'$, $R''$ and $R^*$ denote the expected numbers of returns to 0 within the first $T$ steps of respective random walks $\mathcal{W}'$, $\mathcal{W}''$ and $\mathcal{W}^*$. The last inequality follows from observation (O). Let us consider the first $T$ steps of $\mathcal{W}^*$. We
split \( R^* = R_2^* + R_3^* \), where \( R_2^* \) (\( R_3^* \)) denotes the expected number of returns to 0 before (after) the first visit to 2. From (58), (59) we easily obtain that \( R_2^* = O((\ln n)^{-1}) \). After visiting 2 the walk \( \mathcal{W}^* \) moves to 3 with probability at least \( 1 - 2e' / \ln n \) and it moves towards 0 with probability at most \( 2e' / \ln n \). In the latter case the random walk will be back at 2 after perhaps visiting 0 and the expected number of visits to 0 before returning to 2 is at most \( 1 + R^*_2 \). Hence, the expected number of returns to 0 after visiting 2 and before visiting 3 is at most \( O((1 + R_2^*) (\ln n)^{-1}) = O((\ln n)^{-1}) \). Next we consider random walk \( \mathcal{W}^* \) restricted to the path \( \{2, 3, \ldots, 7\} \), where 2 and 7 are reflecting states. Assuming that the walk starts at 2 we add the expected number of at most \( O((\ln n)^{-1}) \) visits to 0 after each return to 2. Proceeding as in Case (1) we estimate \( R_3^* \leq O((\ln n)^{-1}) + O(T (\ln n)^{-5}) \).

Assume that \( 1 \leq k \leq 5 \). We follow the movements of \( \mathcal{W}'' \) on the subset \( S^*_k = \{0, 1, \ldots, k, k + 2, \ldots, 7\} \subset S_k \) and only register a move when the walk changes its state in \( S^*_k \). The walk moves along the path \( S^*_k \) and has left and right reflecting states 0 and 7. From (65) we obtain that it moves right (from each state but 7) with probability at least \( 1 - 2e' / \ln n \). Arguing as in Case (1) we show that the expected number of returns to 0 within the first \( T \) steps is \( O(\ln n) + O(T (\ln n)^{-5}) \). Obviously, it is an upper bound for \( \bar{R} \).

Case (3). For \( k = 6 \), i.e., \( \pi \in \mathcal{N}_7(v) \), the set \( \mathcal{N}_1(v) \cup \cdots \cup \mathcal{N}_6(v) \) contains no small vertex. The argument used in Case (1) yields the bound \( \bar{R} = O((\ln n)^{-1}) + O(T (\ln n)^{-5}) \).

Proof of (51). We proceed as in (55). Using the fact that \( \bar{R} := \sum_{i=1}^T \bar{r}_i \) is the expected number of returns to \( \pi \) of the random walk \( \mathcal{W}_\pi \) in the time interval \([1; T]\) we show that \( \bar{R} = O((\ln n)^{-1}) + O(T (\ln n)^{-5}) \).

Note that the vertex sets \( \mathcal{N}_x := \mathcal{N}_1(x) \cup \cdots \cup \mathcal{N}_6(x) \) and \( \mathcal{N}_y := \mathcal{N}_1(y) \cup \cdots \cup \mathcal{N}_6(y) \), defined by \( \mathcal{G} \), do not intersect since \( x \) and \( y \) are at distance at least 20. We paint vertices of \( \mathcal{N}_x \) red and those of \( \mathcal{N}_y \) blue. While \( \mathcal{W}_\pi \) stays in \( \mathcal{N}_x \) (respectively \( \mathcal{N}_y \)) we call \( \mathcal{W}_\pi \) red (respectively blue). The path drawn by red random walk corresponds to that of \( \mathcal{W}_x \) in \( \mathcal{N}_x \) and the path drawn by blue random walk corresponds to that of \( \mathcal{W}_y \) in \( \mathcal{N}_y \) (the walk may change its color after every visit to \( \pi \)). For \( 1 \leq i \leq 9 \) we denote \( \mathcal{N}_i(\pi) = \mathcal{N}_i(x) \cup \mathcal{N}_i(y) \) the set of vertices at distance \( i \) from \( \pi \) in \( \mathcal{G}_\pi \). Note that \( \mathcal{N}_x \cup \mathcal{N}_y \) may contain at most two small vertices, by P7. Therefore, at least one of the sets \( \mathcal{N}_i(\pi), i = 7, 8, 9 \) has no
small vertices. Assume it is $N_7(\nu)$ (the cases $i = 8, 9$ are treated in the same way).

Now we analyse the blue and red walks similarly as in the proof of (50) above. At the moment of the first visit of $N_7(\nu)$ by $\mathcal{W}_\nu$ the expected number of returns to $\nu$ is $O(\ln^{-1} n)$. Indeed, the number of returns is the sum of returns of the red and the blue walks. But the expected number of returns of the red walk before it reaches $N_7(x)$ is the same as that of $\mathcal{W}_x$ in $G$. This number is $O(\ln^{-1} n)$, see Cases (1), (2) above. Similarly, the expected number of returns of the blue walk before it reaches $N_7(y)$ is $O(\ln^{-1} n)$.

After the first visit of $N_7(\nu)$ the random walk $\mathcal{W}_\nu$ stays in a distance at least 7 from $\nu$ until it makes the first move from $N_7(\nu)$. This move can be red or blue. A red (blue) move is successful if continuing from $N_6(x)$ ($N_6(y)$) the red (blue) walk visits $\nu$ before visiting $N_7(\nu)$ again. The probability of success is $O(\ln^{-5} n)$, see Cases (1), (2) above. Note that after a successful visit to $\nu$ and before visiting $N_7(\nu)$ again the walk $\mathcal{W}_\nu$ may return several times to $\nu$, but the expected number of such returns is $O(\ln^{-1} n)$.

Hence, each success occurs with probability $O(\ln^{-5} n)$ and it adds $1 + O(\ln^{-1} n)$ expected number of returns to $\nu$. The expected number of successes in the time interval $[1, T]$ is at most $O(T \ln^{-5} n)$. Therefore $\bar{R} = O(\ln^{-1} n) + O(T \ln^{-5} n)$.

Proof of (52), (53). We only show (52). The proof of (53) is much the same. Let $Z_v(t)$ be the number of returns of $\mathcal{W}_v$ to $v$ in the time interval $[t, T - 1]$. We define $Z_v(t) \equiv 0$ for $t \geq T$. Let $I_{\{\tau_v \leq T - 1\}}$ denote the indicator of the event $\tau_v \leq T - 1$. We have

$$R_{T,v}(1) = 1 + \mathbb{E}_G Z_v(1) \quad \text{and} \quad \mathbb{E}_G Z_v(1) \leq \bar{R} = O(\ln^{-1} n).$$

The last bound is shown in the proof of (50) above. Note that $I_{\{\tau_v \leq T - 1\}} \leq Z_v(1)$ implies $\bar{p}_v \leq \mathbb{E}_G Z_v(1) = O(\ln^{-1} n)$. Furthermore, (P3) implies $\text{Pr}_G(\tau_v = 2) \geq 1/\Delta$ and (6) implies $1/\Delta \ll \ln^{-1} n$. Now from the inequality $\bar{p}_v \geq \text{Pr}_G(\tau_v = 2)$ we obtain $\bar{p}_v \ll \ln^{-1} n$.

Let us prove the first relation of (52). The identity

$$Z_v(1) = I_{\{\tau_v \leq T - 1\}} Z_v(1) = I_{\{\tau_v \leq T - 1\}} (1 + Z_v(\tau_v + 1))$$

implies $\mathbb{E}_G Z_v(1) = \bar{p}_v + \bar{R}$, where $\bar{R} = \mathbb{E}_G I_{\{\tau_v \leq T - 1\}} Z_v(\tau_v + 1) = O(\ln^{-2} n)$. Indeed,

$$\bar{R} = \sum_{i=1}^{T-1} \mathbb{E}_G I_{\{\tau_v = i\}} Z_v(i + 1) = \sum_{i=1}^{T-1} \text{Pr}_G(\tau_v = i) \mathbb{E}_G Z_v(i + 1) \leq \bar{p}_v \mathbb{E}_G Z_v(1) = O(\ln^{-2} n).$$

Here we used $\mathbb{E}_G Z_v(i) \leq \mathbb{E}_G Z_v(1)$, for $i \geq 1$.

Proof of (54). Let $q_x = \text{deg}(x)/\text{deg}(\nu)$ and $q_y = \text{deg}(y)/\text{deg}(\nu)$ denote the probabilities that the first move of $\mathcal{W}_x$ is red and blue respectively. Let $A_x$ denote the event that the first return of $\mathcal{W}_x$ in the time interval $[1, T - 1]$ occurs before the first visit to $N_7(x)$. Similarly we define the events $A_y$ and $A_\nu$. Define the probabilities $\bar{p}_u^* = \text{Pr}_G \{ \tau_u \leq T - 1 \} \cap A_u \}$ for $u = x, y, \nu$. The relations $\text{Pr}_G \{ A_u \} = 1 - O(T \ln^{-5} n)$ (which are shown using the same argument as (50), (51)) imply $\bar{p}_u = \bar{p}_u^* + O(T \ln^{-5} n)$ for $u = x, y, \nu$. Now the identity

23
\[ p'_x = q_x p'_x + q_y p'_y \] implies \[ p_x = q_x p_x + q_y p_y + O(T \ln^{-5} n). \] Combining this identity with relations

\[
(1 + p_u)^{-1} = 1 - p_u + O(p_u^2) = 1 - p_u + O(\ln^{-2} n), \quad u = x, y, \kappa
\]

we obtain

\[
\frac{1}{1 + p_x} = \frac{q_x}{1 + p_x} + \frac{q_y}{1 + p_y} + O(T \ln^{-5} n) + O(\ln^{-2} n).
\]

4.2 Probability of the first visit by time \((3)\)

Here for any typical \(G\) and any \(u, v \in \mathcal{V}\) we evaluate the probability that the simple random walk \(W_u\) starting from \(u \in \mathcal{V}\) does not visit \(v\) after time \(T\) and before time \((3)\). We choose \(T = \Theta(\ln n)\) satisfying (47) and Lemma 3 (ii). By P1, P2, P3, such \(T = \Theta(\ln n)\) exists and it does not depend on particular instance \(G\), see [16].

We write the principal term of (3) in the form \(\lambda mn^2 p^2\) and approximate it by

\[
t_0 = \lambda_0 mn^2 p^2 \quad \text{and} \quad t_1 = \lambda_1 mn^2 p^2.
\]

Here

\[
\lambda = \ln \frac{np}{\ln (a + 1)}, \quad \lambda_1 = \ln \frac{np}{\ln (Aa + 1)}, \quad \lambda_0 = (1 + \varepsilon_n)\lambda
\]

and

\[
A = \exp \left( \frac{10 \ln \ln n}{(c - 1) \ln n} \right), \quad a = \frac{c - 1}{\ln e^{np} - 1}, \quad \varepsilon_n = \frac{\ln \ln n}{\ln n}.
\]

In Fact 5 we collect several observations about \(\lambda\)'s. The proof is given in Appendix B.

**Fact 5.**

(i) \(\lambda > 0\) is bounded away from 0 by a constant.

(ii) \(\lambda \leq 2 \ln \ln n\).

(iii) If \((c - 1) \leq (\ln n)^{-1/3}\) then \(\lambda \geq (\ln \ln n)/4\).

(iv) \(\lambda = \Theta(1 + |\ln(c - 1)|)\).

(v) \(\lambda_1 = (1 + o(1))\lambda\).

(vi) \(\ln(ck/(ck - 1)) < \lambda < \ln(c/(c - 1))\), where \(\kappa = np(1 - e^{-np})^{-1}\).
Note that inequalities $\lambda_1 < \lambda < \lambda_0$ (the first one follows as $A > 1$) and Fact 5 (v) imply
\[ t_1 < \lambda mn^2p^2 < t_0 \quad \text{and} \quad t_1 = (1-o(1))t_0 \leq (1 + |\ln(c-1)|)n \ln n. \]

Now we show that whp we have uniformly in $v, x, y \in V$ with $\text{dist}(x, y) \geq 20$
\begin{align}
(62) \quad \Pr_G \{ \mathcal{A}_{t_1}(v) \} &= e^{-\deg(v)\lambda/(1+\bar{p}_v)}(1 + o(1)) + o(n^{-3}) \\
&\geq e^{-\deg(v)\lambda_i(1 + o(1)) + o(n^{-3})}, \\
(63) \quad \Pr_G \{ \mathcal{A}_{t_1}(x) \cap \mathcal{A}_{t_1}(y) \} &= \Pr_G \{ \mathcal{A}_{t_1}(x) \} \Pr_G \{ \mathcal{A}_{t_1}(y) \} (1 + o(1)) + o(n^{-3}).
\end{align}
We recall that $\bar{p}_v = O(\ln n)$ is defined in (53).

**Proof of (62).** From P1, P3 and (48), (52) we obtain for $T = O(\ln n)$ that
\begin{align}
(64) \quad \pi_v &= \deg(v)/(2|E(G)|) = \deg(v)(mn^2p^2)^{-1}(1 + O(n^{-1/2})) = O(n^{-1}) \\
(65) \quad p_v &= \pi_v(1 + \bar{p}_v)^{-1}(1 + O(\ln^2 n)) = \pi_v(1 + O(\ln n)).
\end{align}
Combining these relations and using P3 and Fact 5(ii) we obtain
\begin{align}
(66) \quad t_1p_v &= \frac{\lambda_i \deg(v)}{1 + \bar{p}_v} (1 + O(\ln^2 n)) = \frac{\lambda_i \deg(v)}{1 + \bar{p}_v} + O\left(\frac{\ln n}{\ln n}\right).
\end{align}
Furthermore, (64), (65) imply
\begin{align}
(67) \quad (1 + p_v)t_1 &= e^{t_1,\text{ln}(1+p_v)} = e^{t_1,p_v+O(t_1,\tilde{e}_v^2)} = e^{t_1,p_v} (1 + O(t_1/n^2)).
\end{align}
Finally, we apply Lemma 3 with $C_0 = 1$ (condition (i) of Lemma 3 holds by (50)) and derive (62) from (49), (66) and (67). Note that in this step we estimate the remainder term of (49), $e^{-t_1/2C_0T} = o(n^{-3})$.

**Proof of (63).** We use the observation of [CooperFrieze GiantPaper] that $\mathcal{A}_{t_1}(x) \cap \mathcal{A}_{t_1}(y) = \mathcal{A}_{t_1}(x)$, where $\mathcal{A}_{t_1}(x)$ is the event that the same random walk $\mathfrak{W}_u$, when considered in $G_x$, does not visit the vertex $x$ of $G_x$ in steps $T, T + 1, \ldots, t_1$. We recall that $G_x$ is obtained from $G$ by merging vertices $x$ and $y$ into one vertex denoted $x$. Therefore, $\Pr_G \{ \mathcal{A}_{t_1}(x) \cap \mathcal{A}_{t_1}(y) \} = \Pr_G \{ \mathcal{A}_{t_1}(x) \}$. Proceeding as in the proof of (62) (see also remark below) we show that
\begin{align}
(68) \quad \Pr_G \{ \mathcal{A}_{t_1}(x) \} = e^{-\lambda_i \deg(x)/(1+\bar{p}_v)}(1 + o(1)) + o(n^{-3}).
\end{align}
Furthermore, combining (62) with (68) and using (54) we obtain
\[ \Pr_G \{ \mathcal{A}_{t_1}(x) \} = \Pr_G \{ \mathcal{A}_{t_1}(x) \} \Pr_G \{ \mathcal{A}_{t_1}(y) \} (1 + o(1)) + o(n^{-3}) \]
thus showing (63). We remark that (68) refers to the random walk in $G_x$ starting at $u$. We note that the expansion property P2 extends to $G_x$ and, therefore, Lemma 3 applies with the same (mixing time) $T$. The only difference is that now we verify condition (i) of in Lemma 3 using (51) instead of (50).

Finally, we note that $\bar{p}_v = O(\ln n)$ implies $\lambda_i/(1 + \bar{p}_v) \geq \lambda$. Thus by (66), (67) we get
\begin{align}
(69) \quad (1 + p_v)^{-t_0} &= (1 + o(1)) e^{-\lambda_i \deg(v)/(1+\bar{p}_v)} \\
&\leq (1 + o(1)) e^{-\lambda \deg(v)}. 
\end{align}
5 Cover Time

In this section we consider the simple random walk \( \mathfrak{W}_u \) on typical \( G \) starting at a vertex \( u \in V \). Given \( G \) and \( u \) we denote \( C_u \) the expected time taken for the walk to visit every vertex of \( G \). We show that \( \text{whp} \) \( t_1 \leq C_u \leq t_0(1 + o(1)) \). In the proof we choose \( T = \Theta(\ln n) \) satisfying conditions (47) and Lemma 3(ii) and use the short-hand notation \( x = mpe^{-\lambda} \), \( y = npe^{-\lambda}, y_1 = npe^{-\lambda_1} \).

5.1 Upper bound

For each \( u \in V \) and all \( t \geq T \) we have, see (42) of [6] that

\[
C_u \leq t + 1 + \sum_v \sum_{s \geq t} \Pr_{G} \{ A_s(v) \} .
\]  

For \( t_0 \) and \( \lambda \) defined in (60), (61), we have by Lemma 3, see also (50),

\[
\sum_v \sum_{s \geq t} \Pr_{G} \{ A_s(v) \} \leq (1 + o(1)) \sum_v \sum_{s \geq t_0} \left( (1 + p_v)^{-s} + o(e^{-s/(2C_0T)}) \right)
\]

\[
= (1 + o(1)) \sum_v \frac{(1 + p_v)^{-t_0}}{1(1 + p_v)^{-1}} + o(n^{-1})
\]

\[
= (1 + o(1)) \sum_v \frac{mn^2p^2}{\deg(v)} e^{-\lambda \deg(v)} + o(n^{-1})
\]

\[
= (1 + o(1)) \sum_k D(k) \frac{mn^2p^2}{k} e^{-\lambda k} + o(n^{-1}).
\]

In the third line we used (64), (65), and (69). We show below that the sums

\[ S_i := \sum_{k \in K_i} D(k) \frac{1}{k} e^{-\lambda k} = o(1), \quad i = 1, 2, 3. \]

These bounds imply \( \sum_v \sum_{s \geq t} \Pr_{G} \{ A_s(v) \} = o(t_0) \). Now (70) yields \( C_u \leq t_0 + 1 + o(t_0) \).

We first estimate \( S_1, S_2 \). For \( (c - 1) \leq (\ln n)^{-1/3} \) we have, by Fact 5(iii) and P8a,

\[ S_1 \leq 20(\ln \ln n)^2 e^{-\frac{1}{4}\ln \ln n} = o(1), \quad S_2 \leq \Delta(\ln n)^4 e^{-\frac{21}{4}\ln \ln n} = o(1). \]

For \( (c - 1) \geq (\ln n)^{-1/3} \) we have, by P8a, P8b and Fact 5(i),

\[ S_1 + S_2 \leq \Delta(\ln n)^4 \frac{1}{(\ln n)^{1/2}} e^{-\lambda (\ln n)^{1/2}} = o(1). \]
Now we estimate $S_3$. By property P8,

$$
(71) \quad S_3 \leq \frac{3}{2} \sum_{k=1}^{\Delta} \bar{D}(k) \frac{1}{k} e^{-\lambda k} \leq \frac{3}{2} n^{1-c} \sum_{i=1}^{\infty} (mpe^{-np})^i \sum_{k=i}^{\infty} \frac{k}{k!} (npe^{-\lambda})^k.
$$

To estimate the inner sum we use the following inequalities shown in Appendix B below.

**Fact 6.** For $y > 0$ we have $\sum_{k=1}^{\infty} \{k\} \frac{\lambda^k}{k!} \leq \frac{e^{y-1}}{y} \leq 4 \frac{(e^{y-1})^i}{i!}$. For $0 < e^y - 1 < 1/2$ we have $\sum_{k=1}^{\infty} \{k\} \frac{\lambda^k}{k!} \leq 3 \frac{1}{2} \frac{(e^{y-1})^i}{i!} \leq 3 \frac{1}{2} \frac{(e^{y-1})^i}{i!} y$.

For $e^y - 1 < 1/2$ we obtain from (71) using Fact 6 that

$$
S_3 \leq 4.5 n^{1-c} \sum_{i=1}^{\infty} \frac{(e^y - 1)^{i+1}}{y(i+1)!} x^i = \frac{4.5}{xy} n^{1-c} \sum_{i=1}^{\infty} \frac{(e^y - 1)x^{i+1}}{(i+1)!} \leq \frac{4.5}{xy} n^{1-c} e^{(e^y - 1)x}.
$$

Furthermore, relations

$$
(72) \quad x(e^{np} - 1) = (1 + O(n^{-1} \ln n)) c \ln n \quad \text{and} \quad \frac{e^{np} - 1}{e^{np} - 1 - c} = \frac{c - 1}{c},
$$

see (2) and (61), imply $n^{1-c} e^{(e^y - 1)x} = 1 + o(1)$. Finally, we establish the bound $S_3 = o(1)$ by showing that $xy \to +\infty$. For small $y > 0$ satisfying $e^y - 1 < 1/2$ we have $2y > (e^y - 1)$. Now (72) implies

$$
2xy > x(e^y - 1) = x(e^{np} - 1)(c - 1)/c = (1 + o(1))(c - 1) \ln n \to +\infty.
$$

For $e^y - 1 \geq 1/2$ we obtain from (71) using the first inequality of Fact 6 that

$$
S_3 \leq 6 n^{1-c} \sum_{i=1}^{\infty} \frac{(e^y - 1)^i}{y(i+1)!} x^i \leq \frac{6}{xy(e^y - 1)} n^{1-c} e^{(e^y - 1)x} = \frac{6}{xy(e^y - 1)} (1 + o(1)).
$$

Now the right side is $O(x^{-1}) = o(1)$, since $y(e^y - 1) \geq (\ln(3/2))/2$ is bounded away from zero. This completes the proof of the upper bound.

### 5.2 Lower bound

Here we define a large set of special vertices $S$ and show that with high probability some vertices from $S$ are visited by $\mathcal{W}_u$ only after time $t_1$, i.e., in the last phase before covering all the vertices.

For $i_0$, $k_0$ and $\bar{D}(k, i_0)$ defined in P8 denote

$$
I_1 = \{k : i_0 \leq k \leq k_0 \text{ and } \bar{D}(k, i_0) > k_0^{-2} e^{k\lambda_1}\}, \quad I_0 = \{i_0, i_0 + 1, \ldots, k_0\} \setminus I_1.
$$
Note that $k_0^{-2}e^{k_1} > i_0^2$ since $\lambda_1 = (1 + o(1)) \lambda$ is bounded away from 0 and $i_0 \approx k_0$. Hence $I_1 \subset I$, where $I$ is from P8. Define the set of special vertices

$$S = \left\{ v \in V : |W'(v)| = i_0, \deg(v) \in I_1, v' \in \arg min_{v' \in I} \text{dist}(v, v') \geq \frac{\ln n}{(\ln \ln n)^3} \right\}$$

and let $X$ be the number of vertices in $S$ that are not visited in steps $T, T + 1, \ldots, t_1$.

Note that $I_1 \subset I$ implies $|S| = \sum_{k \in I_1} D^*(k, i_0)$, see P8.

We show below that $E_GX = \Omega(\ln^9 n)$ and $E_GX^2 - (E_GX)^2 = o((E_GX)^2)$ uniformly over typical $G$ and $u \in V$. These bounds yield $\Pr_G\{2X > E_GX\} \rightarrow 1$, by Chebyshev’s inequality. Hence whp $X = \Omega(\ln^9 n)$. As the number of vertices visited within the first $T$ steps is at most $T = O(\ln n)$ we will find in $S$ at least $X - T = \Omega(\ln^9 n)$ vertices unvisited by the time $t_1$. Thus $C_u \geq t_1$.

Let us prove that $E_GX = \Omega(\ln^9 n)$. It follows from (62) and P8c that

$$E_GX = \sum_{v \in S} \Pr_G\left\{A_{I_1}(v)\right\} = (1 + o(1)) \sum_{k \in I_1} D^*(k, i_0)e^{-k\lambda_1} + o(1)$$

$$\geq \left(\frac{1}{2} + o(1)\right) \sum_{k \in I_1} D(k, i_0)e^{-k\lambda_1} + o(1).$$

We write the sum $\sum_{k \in I_1} D(k, i_0)e^{-k\lambda_1}$ in the form

$$\left(\sum_{k \geq i_0} - \sum_{k \geq k_0} - \sum_{k \leq i_0} \right) D(k, i_0)e^{-k\lambda_1} =: S^* - S_0^* - S_1^*$$

and show that $S^* = \Omega(\ln^9 n)$ and $S_i^* = o(1)$, $i = 0, 1$. The bound $S_1^* \leq |I_1|k_0^{-2} = o(1)$ is obvious. Let us prove that $S_0^* = o(1)$. In the proof we use inequalities

$$(73) \quad \left\{\begin{array}{l}
\frac{k}{i_0} \frac{1}{k!} \leq \frac{(ke)^{i_0}i_0^k}{i_0^{2i_0}k^{k}}e^{k_0} \leq \frac{(ke)^{i_0}i_0}{i_0^{2i_0}k^{k_0}}e^{k_0} \frac{i_0}{k^{k_0}(k-k_0)!} \frac{k^{k_0}}{i_0^{k_0}}.
\end{array}\right.$$\]

The second inequality follows by $k_0 \leq k$. To get the first one we combine the inequalities

$$\left\{\begin{array}{l}
\frac{k}{i_0} \leq \frac{k}{i_0} i_0^{-i_0} \leq (ke)^{i_0} i_0^{k_2i_0} \quad \text{and} \quad k! \geq (k-k_0)!(k/e)^{k_0}
\end{array}\right.$$\]

that follow from (8), $\left(\begin{array}{l}k \\
\end{array}\right) \leq (ke/s)^s$ and $k!/(k-s)! \geq (k/e)^s$ respectively (the last inequality follows by induction on $s$). From (73) we obtain

$$D(k, i_0)e^{-k\lambda_1} \leq \frac{e^{i_0}}{n^{e^{-1}}} \frac{xk_0}{i_0^2} \frac{e^{i_0}y_1}{k_0} \frac{k}{(k-k_0)!}.$$
Note that the first factor $e^{i_0 n^{1-c}} \leq e$. Now summing over $k \geq k_0$ gives

$$S^*_0 \leq e \cdot \left( \frac{xk_0}{i_0^2} \right)^{i_0} \left( \frac{e^{i_0 y_1}}{k_0} \right)^{k_0} e^{i_0 y_1} = e \cdot \left( \frac{x e^{y_1 k_0}}{i_0^2} \right)^{i_0} \left( \frac{e^{i_0 y_1}}{k_0} \right)^{k_0}.$$  

Furthermore, using $e^{y_1} = Aa + 1$, $y_1 \leq Aa$, and the first relation of (72) we upper bound $S^*_0/e$ by

$$(74) \quad \left( (1 + O(n^{-1} \ln n)) c \frac{k_0}{c - 1} Aa + 1 \right)^{i_0} \cdot \left( \frac{e^{i_0} k_0}{Aa} \right)^{k_0}.$$  

Now assume that $c - 1 \leq (\ln \ln n)^2 / \ln n$. In this case our condition $(c - 1) \ln n \to +\infty$ implies $Aa + 1 = O(1)$. Using $i_0 \approx k_0$ we upper bound (74) by

$$(\Theta(1))^{i_0+1} \left( \frac{1}{c - 1} \right)^{i_0} (Aa)^{k_0} = (\Theta(1))^{k_0} A^{k_0} (c - 1)^{k_0 - i_0} = o(1).$$  

In the first step we used $a / (c - 1) = \Theta(1)$. In the last step we used $k_0 - i_0 \approx k_0$ and $A^{k_0} = O(e^{O(\ln \ln n)})$. This shows $S^*_0 = o(1)$.

Next, assume that $c - 1 > (\ln \ln n)^2 / \ln n$. Using $e^s \leq 1 + 2s$ for small $s = 10(\ln \ln n)/(c - 1) \ln n$ we bound $A \leq 1 + 2s$. Now, the inequality $(1 + 2s)/(c - 1)/c \leq 1$, which holds for $c = O(1)$, yields $aA \leq e^{np} - 1$. Furthermore, a crude upper bound $A \leq 3/e$ yields $aA \leq (e^{np} - 1)(c - 1)c^{-1}(3/e)$. Invoking these upper bounds for $aA$ in the first and second factors of (74) and using $(1 + O(n^{-1} \ln n))^{i_0} = 1 + o(1)$ we upper bound (74) by

$$2 \left( \frac{k_0}{i_0 e^{np} - 1} \right)^{i_0} \left( \frac{3 i_0}{k_0 e^{np} - 1} \right)^{k_0} \frac{c - 1}{c} (e^{np} - 1)^{k_0 - i_0} \leq 2 \left( \frac{i_0}{k_0} e^{nc} (e^{np} - 1)^{k_0 - i_0} \right)^{3 k_0} = o(1).$$  

In the first step we used $(c - 1)/c < 1$ and $(e^{np})^{k_0 - 2i_0} \geq 1$. This proves $S^*_0 = o(1)$.

Let us prove that $S^* = \Omega(n^g n)$. By properties of Stirling’s numbers we have

$$S^* = n^{1-c} x^{i_0} \sum_{k \geq i_0} \left( \frac{k!}{i_0!} \right)^{k_0} \left( \frac{1}{k !} \right)^{y_0} = n^{1-c} x^{i_0} \left( \frac{e^{y_1} - 1}{i_0!} \right)^{i_0}.$$  

Furthermore, using Stirling’s approximation to $i_0!$ and invoking the relations

$$n^{1-c} \sim e^{-i_0}, \quad e^{y_1} - 1 = A(e^{np} - 1)(c - 1)/c, \quad x(e^{np} - 1) = mp(1 - e^{-np})$$

we get

$$S^* \sim \frac{1}{\sqrt{i_0}} A^{i_0} \left( \frac{mp(1 - e^{-np} c - 1)}{i_0} \right)^{i_0} \sim \frac{1}{\sqrt{i_0}} A^{i_0} = \Omega(\ln^9 n).$$

Finally, we show that $\mathbb{E}_G X^2 - (\mathbb{E}_G X)^2 = o((\mathbb{E}_G X)^2)$. We have $X(X - 1) = \sum_{i,v} \mathbb{I}_u \mathbb{I}_v$, where $\mathbb{I}_v$ denotes the indicator of event $\mathcal{A}_t(v)$. By (63),

$$\mathbb{E}_G \mathbb{I}_u \mathbb{I}_v = (1 + o(1))\Pr_G \{ \mathcal{A}_t(u) \} \Pr_G \{ \mathcal{A}_t(v) \} + o(n^{-3}).$$

Hence $\mathbb{E}_G X(X - 1) = (1 + o(1))(\mathbb{E}_G X)^2$. Finally, for $\mathbb{E}_G X \to +\infty$ we obtain

$$\mathbb{E}_G X^2 - (\mathbb{E}_G X)^2 = \mathbb{E}_G X(X - 1) + \mathbb{E}_G X - (\mathbb{E}_G X)^2 = o(\mathbb{E}_G X)^2.$$

29
6 Conclusions

We determined the expected cover time of a random walk in random intersection graph \( G(n, m, p) \) above its connectivity threshold. Our results were compared with corresponding results obtained for Erdős–Rényi random graph model. This comparison led us to conclusion that the presence of clustering and specific degree distribution in affiliation networks delay the covering of the network by a random walk (with relation to the random walk on \( G(n, q) \) with corresponding edge density).

We studied the random intersection graph model introduced by Karoński at al. in [11]. However various different random intersection graph models have been studied since, for example, in the context of security of wireless sensor networks [3, 20] or scale free networks [1] (see also [2], [8], and [18] for more models and applications). In the context of obtained results it would be intriguing to study the cover time in other models of random intersection graphs in order to understand more the relation between clustering, degree distribution, and the expected cover time of a random walk in real life networks.

Appendix

A Proof of Lemma 2

Before we proceed to the proof of \( \text{P0-P7} \) we collect several auxiliary results. In the proof we use the following version of Chernoff’s inequality, see (2.6) in [10].

**Lemma 7.** Let \( X \) be a random variable with the binomial distribution and expected value \( \mu \), then for any \( 0 < \varepsilon < 1 \)

\[
\Pr \{ X \leq \varepsilon \mu \} \leq \exp \left( -\psi(\varepsilon)\mu \right), \text{ where } \psi(\varepsilon) = \varepsilon \ln \varepsilon + 1 - \varepsilon.
\]

**Fact 8.** There exists a constant \( a_\star > 0 \) depending on the sequences \( \{p(m,n)\} \) and \( \{c(n,m)\} \) such that with high probability

- (i) any two \( \mathcal{B} \)-cycles of length at most \( a_\star \ln n / \ln \ln n \) are at least \( a_\star \ln n / \ln \ln n \) links apart from each other.
- (ii) any two small vertices are at least \( a_\star \ln n / \ln \ln n \) links apart from each other.
- (iii) any small vertex is at least \( a_\star \ln n / \ln \ln n \ln \ln n \) links apart from any \( \mathcal{B} \)-cycle of length at most \( a_\star \ln n / \ln \ln n \).

**Fact 9.** With high probability there are at most \( \frac{\ln 3}{\ln \ln n} \) \( \mathcal{B} \)-cycles consisting of 4 links.

**Fact 10.** With high probability \( |\mathcal{V}(w)| \leq \frac{\ln n}{\ln \ln n} \max\{2, np\} \) for every \( w \in \mathcal{W} \).
Proof of Fact 8. Given $a_* > 0$ denote $j_0 = [a_* \ln n / \ln \ln n]$.

Proof of (i). The expected number of pairs of $B$-cycles $C_k$, $C_r$ of length $2k$ and $2r$ which are connected by a (shortest) path of length $i \geq 1$ (links) containing $i_1$ internal vertices and $i_2$ internal attributes (those outside $C_k$ and $C_r$) is at most

$$\begin{align*}
(76) \quad \binom{n}{k} \binom{m}{k} p^{2k} \binom{n}{r} \binom{m}{r} p^{2r} \binom{n}{i_1} \binom{m}{i_2} p^i 4k! i_1! i_2! \leq 4n^{k+r+i} m^{k+r+i} p^{2k+2r+i}.
\end{align*}$$

Note that $i = i_1 + i_2 + 1$ and for $i = 1$ we have $i_1 = i_2 = 0$. Assuming that $n \leq m$ and $mp \geq 1$ we see that the sum of $(76)$ over $2 \leq k, r \leq j_0$ and $0 \leq i_1 + i_2 \leq j_0$ is at most

$$\begin{align*}
(77) \quad \sum_{k=2}^{j_0} \sum_{r=2}^{j_0} \sum_{i_1=0}^{j_0} \sum_{i_2=0}^{j_0} 4p(mp)^{2k+2r+i-1} \leq 4j_0^4 p(mp)^{5j_0} = o(1).
\end{align*}$$

To achieve the last bound we choose $a_* > 0$ sufficiently small. We obtain that the expected number of pairs of $B$-cycles of length at most $2j_0$ that are in the distance $d \in [1, j_0]$ is $o(1)$. Hence with high probability we do not observe such a pair.

Next we count pairs of $B$-cycles $C_k$, $C_r$ that share at least one vertex or attribute. For any $B$-cycle $C$ we denote by $V_C$ the set of its vertices and by $W_C$ the set of its attributes. Let $u \in (V_{C_k} \cup W_{C_k}) \setminus (V_{C_k} \cup W_{C_k})$. We can walk from $u$ along $C_r$ in two directions until we reach the set $V_{C_k} \cup W_{C_k}$. In this way we obtain a path belonging to $C_r$ and with endpoints in $V_{C_k} \cup W_{C_k}$. Internal vertices/attributes of the path do not belong to $C_k$. By $i$ we denote the length of the path (number of links). $i_1$ and $i_2$ denote the numbers of internal vertices and attributes of the path ($i = i_1 + i_2 + 1, |i_1 - i_2| \leq 1$). The union of $C_k$ and the path defines an eared $B$-cycle. The expected number of such eared cycles is at most

$$\begin{align*}
(78) \quad \binom{n}{k} \binom{m}{k} p^{2k} \binom{n}{i_1} \binom{m}{i_2} p^i 4k! i_1! i_2! \leq 4kn^{k+i} m^{k+i} p^{2k+i}
\end{align*}$$

Assuming that $n \leq m$ and $mp \geq 1$ we see that the sum of $(78)$ over $2 \leq k \leq j_0$ and $2 \leq i \leq 2j_0$ is at most

$$\begin{align*}
\sum_{k=2}^{j_0} \sum_{i_1=1}^{j_0} \sum_{i_2=1}^{j_0} 4kp(mp)^{2k+i-1} \leq 4j_0^4 p(mp)^{4j_0} = o(1).
\end{align*}$$

The last bound follows by our choice of $a_* > 0$ in $(77)$. We obtain that the expected number of eared $B$-cycles (where the cycle and the ear have at most $j_0$ vertices and $j_0$ attributes each) is $o(1)$. Hence with high probability we do not observe a pair of $C_k$, $C_r$ with $2 \leq k, r \leq j_0$ that share at least one vertex or attribute.

Proof of (ii). Assume that there exist two vertices $v, v' \in \text{SMALL}$, which are $2t$ links apart blue in $B$ for some $2t \leq j_0$. Then blue $B$ contains a path $v w_1 v_1 w_2 \ldots v_{t-1} w_t v'$ of
length $2t$ and the set $(W(v) \cup W(v')) \setminus \{w_1, \ldots, w_t\}$ is of cardinality at most $0.2 \ln n$. The number of possible paths of the form $vw_1v_1w_2 \ldots v_{t-1}w_t v'$ is at most $n^{t+1}m^t$ and the probability that a path of length $2t$ is present in $\mathcal{B}(n, m, p)$ is $p^{2t}$. Furthermore, given the event that the path $vw_1v_1w_2 \ldots v_{t-1}w_t v'$ is present, the cardinality of the set $(W(v) \cup W(v')) \setminus \{w_1, \ldots, w_t\}$ has the binomial distribution $\text{Bin}(m-t, 1-(1-p)^2)$. By Lemma 7, this cardinality is less than $0.2 \ln n$ with probability at most $e^{-1.2 \ln n}$. Indeed, the binomial distribution has the expected value $\mu = 2mp + o(1) > 2c \ln n > 2 \ln n$. Finally, by the union bound the probability that there exist two small vertices within the distance $j_0$ (links) is at most $j_0/2 \sum_{t=1}^{j_0/2} n^{t+1}m^t p^{2t} e^{-1.2 \ln n} \leq j_0 n (np^2)^{j_0/2} e^{-1.2 \ln n} = o(1)$.

Proof of (iii) is a combination of those of (i) and (ii).

Proof of Fact 9. The expected number of $\mathcal{B}$-cycles consisting of 4 links is at most $n^2m^2p^4 = O(\ln^2 n)$. The fact follows by Markov’s inequality.

Proof of Fact 10. Denote for short $s = \max\{2, np\} (\ln n) / \ln \ln n$. By the union bound,

$$\Pr\{\exists w \in W|\mathcal{V}(w)| \geq s\} \leq m \Pr\{|\mathcal{V}(w)| \geq s\} \leq m \left(\frac{n}{s}\right)^p \leq m \left(\frac{enp}{s}\right)^s \leq \exp (\ln m - s \ln s + O(s)) = o(1).$$

Proof of P0-P7.

**P0** Any triple of vertices linked to some $w$ in $\mathcal{B}(n, m, p)$ induce the triangle in $G(n, m, p)$. The degrees of attributes $w \in W$ in $\mathcal{B}(n, m, p)$ are independent Binomial random variables with mean $pn = \Theta(1)$. Therefore, the maximal degree is greater than 2 whp. Hence $\mathcal{G}(n, m, p)$ contains a triangle whp. For the connectivity property we refer to [?].

**P1** Let $X$ be binomial Bin $(n, p)$ random variable. Denote

$$Y_1 = \sum_{w \in W} \left(\left|\mathcal{V}(w)\right| \right)^2, \quad Y_2 = \sum_{w, w'} \left(\left|\mathcal{V}(w) \cap \mathcal{V}(w')\right| \right)^2.$$

We have, by the inclusion-exclusion argument,

$$(79) \quad Y_1 - Y_2 \leq |\mathcal{E}(\mathcal{G}(n, m, p))| \leq Y_1.$$

32
Note that $Y_2$ is the number of $B$–cycles with 4 links. Hence $Y_2 \leq \ln^3 n$ with high probability, by Fact 9. Furthermore, as $|\mathcal{V}(w)|, w \in \mathcal{W}$, are independent copies of $X$, we obtain for $np = \Theta(1)$ that

$$
\mathbb{E}Y_1 = m \frac{\mathbb{E}X(X-1)}{2} = m \frac{n(n-1)p^2}{2} = m \cdot \Theta(1) = \Theta(n \ln n),
$$

$$
\text{Var}Y_1 = m \frac{\mathbb{E}X^2(X-1)^2 - (n(n-1)p^2)^2}{2} = m \cdot \Theta(1) = \Theta(n \ln n).
$$

Now, Chebyshev’s inequality implies $\Pr\{|Y_1 - \mathbb{E}Y_1| > \sqrt{\mathbb{E}Y \ln^{1/4} n}\} = o(1)$. This bound combined with $Y_2 = O(\ln^3 n)$ and (19) shows $P1$.

**P3** Proof of $\text{deg}(v) \leq \Delta$. Given $A \subset \mathcal{W}$ of size $|A| \leq 4mp$, let $Y$ denote the number of vertices $v \in \mathcal{V}$ linked to at least one attribute from $A$. The random variable $Y$ has binomial distribution $\text{Bin} \left(n, 1 - (1 - p)^{|A|}\right)$. Let $Y'$ be a random variable with the distribution $\text{Bin} \left(n, 4d_1/n\right)$. Inequalities $(1 - p)^{|A|} \geq (1 - p)^{4mp} \geq 1 - 4d_1/n$ imply that $Y$ is stochastically dominated by $Y'$, i.e., $\Pr\{Y > s\} \leq \Pr\{Y' > s\}, \forall s > 0$. We have

$$
\Pr\{\text{deg}(v) > 12d_1\} \leq \Pr\{\text{deg}(v) > 12d_1 \mid |\mathcal{W}(v)| \leq 4mp\} + \Pr\{|\mathcal{W}(v)| > 4mp\}
$$

$$
\leq \Pr\{Y' > 12d_1\} + \Pr\{|\mathcal{W}(v)| > 4mp\}
$$

$$
\leq \left(\frac{n}{12d_1}\right) \left(\frac{4d_1}{n}\right)^{\lceil 12d_1 \rceil} + \left(\frac{m}{4mp}\right)^{\lceil 4mp \rceil}
$$

$$
\leq \left(\frac{e}{3}\right)^{\lceil 12d_1 \rceil} + \left(\frac{e}{4}\right)^{\lceil 4mp \rceil} = o(n^{-1}).
$$

Now, by the union bound $\Pr\{\exists v \in \mathcal{V} \text{deg}(v) > 12d_1\} \leq n \Pr\{\text{deg}(v) > 12d_1\} = o(1)$. We conclude that $\max_{v \in \mathcal{V}} \text{deg}(v) > \Delta$ with high probability.

Proof of $|\mathcal{W}(v)| \leq \Delta$. The random variable $|\mathcal{W}(v)|$ has binomial distribution $\text{Bin} \left(m, p_*\right)$ with $p_* = p(1 - (1 - p)^m) = d_0/m$. We have

$$
\Pr\{|\mathcal{W}(v)| > 4d_0\} \leq \left(\frac{m}{4d_0}\right)^{\lceil 4d_0 \rceil} = o(n^{-1}).
$$

Hence, $\Pr\{\exists v \in \mathcal{V} \mid |\mathcal{W}(v)| > 4d_0\} \leq n \Pr\{|\mathcal{W}(v)| > 4d_0\} = o(1)$.

**P4** Any pair of adjacent vertices $v, v' \in \mathcal{V}$ share at most two common attributes (otherwise there were two intersecting $B$–cycles of length 4, the event ruled out by Fact 8(i)). Assume that $v, v'$ share two attributes $w, w'$. In this case all common neighbors of $v, v'$ belong to $\mathcal{V}(w) \cup \mathcal{V}(w')$ (otherwise there were two intersecting $B$–cycles of length at most 6). Now $P4$ follows from Fact 10. Next, assume that $v, v'$ share only one attribute $w$. In this case there might be at most one common neighbor of $v, v'$ outside $\mathcal{V}(w)$ (otherwise there were two intersecting $B$–cycles of length at most 6). Hence the number of common neighbors is at most $|\mathcal{V}(w)| - 2 + 1$ and we obtain $P4$ from Fact 10.
P5 If \( \deg(v) \leq |\mathcal{W}(v)| - 2 \) then we either find \( u \in \mathcal{V} \setminus \{v\} \) linked to at least three different elements of \( \mathcal{W}(v) \) or we find \( u_1, u_2 \in \mathcal{V} \setminus \{v\} \) such that \( u_i \) is linked to at least two elements of \( \mathcal{W}(v) \) for each \( i = 1, 2 \). In both cases there is a pair of short \( B \)-cycles containing \( v \), the event ruled out by Fact 8(i). Hence \( d(v) \geq |\mathcal{W}(v)| - 1 \) with high probability. For a large vertex \( v \), the latter inequality implies \( \deg(v) \geq (\ln n)/11 \).

P6 Assume that we find \( v \in \mathcal{V} \) and \( u_1^*, u_2^*, u_3^* \in N_{i-1}(v) \) and \( u \in N_i(v) \) such that \( v \) is adjacent to each \( u_1^*, u_2^*, u_3^* \). Vertex \( u \) is 2i links apart from \( v \) in \( B \). Furthermore, each \( u_j^* \) is \( 2(i-1) \) links apart from \( v \) and 2 links apart from \( u \), for \( j = 1, 2, 3 \). Therefore we find three distinct shortest paths connecting \( v \) and \( u \) in \( B \) (via \( u_1^*, u_2^* \) and \( u_3^* \)). These paths create at least two short \( B \)-cycles close to \( u \). But, by Fact 8(i), there are no such cycles with high probability.

P7 follows by Fact 8.

B Proof of Facts 5, 6

Proof of Fact 5. Proof of (i). For \( y > x > 0 \) we have \( \ln(1+y) - \ln(1+x) > (y-x) \ln'(1+y) \), since \( x \to \ln'(1+x) \) is decreasing. We apply \( \ln(1+y) - (y-x)(1+y)^{-1} > \ln(1+x) \) to \( x = a \) and \( y = e^{np} - 1 \) and obtain

\[
\ln(1 + a) < np - (1 - e^{-np})/c.
\]

For \( np = \Theta(1) \) and \( 1 < c = O(1) \) we find and absolute constant \( \delta > 0 \) such that \( (1 - e^{-np})/c > \delta \). Hence \( \lambda > \ln(np/(np - \delta)) \) is bounded away from zero.

Proof of (ii). \( (c-1) \ln n \to \infty \) implies \( a > 2 \ln^{-1} n \) for large \( n \). Furthermore, \( \ln(x+1) \geq x/2 \), for \( 0 < x < 2 \), implies \( \ln(a+1) > \ln^{-1} n \). Therefore

\[
\lambda \leq \ln(np/ln n) \leq ln np + \ln \ln n \leq 2 \ln \ln n.
\]

Proof of (iii). Using \( \ln(1 + a) \leq a \) we bound from below

\[
(80) \quad \lambda \geq \ln(np/a) = \ln \frac{cnp}{e^{np} - 1} + \ln((c - 1)^{-1}) = O(1) + \ln((c - 1)^{-1}).
\]

For \( c - 1 \leq \ln^{-1/3} n \) and large \( n \) the right side is at least \( 4^{-1} \ln \ln n \).

Proof of (iv). Our assumptions \( np = \Theta(1) \) and \( 1 < c = O(1) \) imply that the lower bound of \( \ln \frac{cnp}{e^{np} - 1} \), denoted \( b \), is finite. Let us first show that \( 1 + |\ln(c - 1)| = O(\lambda) \). For \( 1 < c < 1.5 \) and \( |\ln(c - 1)| > 2 \max\{-b, 0\} \) this relation follows from (80). Otherwise it follows from (i). Now we show that \( \lambda = O(1 + |\ln(c - 1)|) \). For \( c \leq \min\{1.5, (e^{np} - 1)^{-1}\} \) we have \( a \leq 1 \) and inequality \( \ln(1 + a) \geq a/2 \) implies

\[
\lambda \leq \ln(2np/a) = \ln \frac{2cnp}{e^{np} - 1} + \ln((c - 1)^{-1}) = O(1) + \ln((c - 1)^{-1}).
\]
For \( c > \min\{1.5, (e^{np} - 1)^{-1}\} \) we have \( a = \Theta(1) \). This implies \( \lambda = O(1) \).

Proof of (v). In view of \( \lambda > \lambda_1 \) it suffices to show that \( \lambda \leq \lambda_1 + o(\lambda) \). We firstly assume that \((c - 1) \ln n \leq (\ln \ln n)^2\). By the inequalities \( x(1 - x) \leq \ln(1 + x) \leq x \), we have

\[
\lambda - \lambda_1 = \ln \frac{\ln(aA + 1)}{\ln(a + 1)} \leq \ln \frac{aA}{a(1-a)} = \ln A - \ln(1-a) = \ln A + o(1).
\]

Furthermore,

\[
\ln A = \ln(c-1)^{-\frac{10}{(c-1)\ln n}} + \frac{10\ln((c-1)\ln n)}{(c-1)\ln n} = o(\ln(c-1)^{-1}) + o(1).
\]

In the last step we used \((c - 1) \ln n \to +\infty\). We obtain \( \lambda - \lambda_1 = o(\| \ln(c-1) \| + 1) \). Now (v) follows from (iv).

Next we assume that \((c - 1) \ln n > (\ln \ln n)^2\). In this case \( \ln A \leq 10/\ln \ln n \). By the inequality \( e^x - 1 \leq 2x \), for \( 0 \leq x \leq 1/2 \), we have

\[
A - 1 \leq 2 \ln A \leq 20/\ln \ln n.
\]

Set \( f(x) = \ln \ln(ax + 1) \). Then \( f'(x) = a(\ln(ax + 1))^{-1} (ax + 1)^{-1} \) is a decreasing function, for \( x \geq 0 \). From this fact and the inequalities \( A > 1 \) and \( a/(1 + a) \leq \ln(1 + a) \) we obtain

\[
\lambda - \lambda_1 = \ln \ln(aA + 1) - \ln \ln(a + 1) = f(A) - f(1) \\
\leq f'(1)(A - 1) = \frac{1}{\ln(a + 1)} \frac{a}{1 + a} (A - 1) \leq A - 1 \leq \frac{20}{\ln \ln n} = o(\lambda).
\]

In the last step we used (i).

Proof of (vi). We write for short \( x = np, b = (c - 1)/c, z = e^x \). To prove the first inequality we show that \( x(c\kappa - 1)/(c\kappa) > \ln(b(e^x - 1) + 1) \). To this aim we establish the same inequality, but for the respective derivatives \( \partial/\partial x \). Indeed, the derivative of the left side \( 1 - e^{-x}(1 - b) \) is greater than that of the right side \( (b(e^x - 1) + 1)^{-1} be^x \) because we have \( (1 - (1-b)z^{-1})(b(z - 1) + 1) > bz \), for \( 0 < b < 1 \) and \( z > 1 \).

To prove the second inequality we show that \( x \ln^{-1}(b(e^x - 1) + 1) < b^{-1} \). We have \( xb < \ln(b(e^x - 1) + 1) \) because the respective inequality holds for the derivatives \( \partial/\partial x \).

Proof of Fact 6. We start with auxiliary function \( f_i(t) = t + \sum_{j=1}^{i}(1-e^t)^j j^{-1}, i = 1, 2, \ldots \).

Its derivative \( f'_i(t) = (1-e^t)^i \) satisfies \( f'_i(t)(-1)^i = (e^t - 1)^i > 0 \), for \( t > 0 \). Now \( f_i(0) = 0 \) implies \( f_i(t)(-1)^i > 0 \) for \( t > 0 \) and \( \forall i \). Using this fact we obtain the upper bound

\[
f_i(t)(-1)^i = (e^t - 1)^i - f_{i-1}(t)(-1)^{i-1} \leq (e^t - 1)^i i^{-1}, \quad t > 0, \quad \forall i.
\]
Furthermore, the identity \( \ln(1 + x) = - \sum_{j \geq 1} (-x)^j j^{-1} \) with \( x = e^t - 1 \) implies

\[
(82) \quad f_i(t) = \ln(1 + x) + \sum_{1 \leq j \leq i} (1 - e^t)^j j^{-1} = - \sum_{j \geq i+1} (1 - e^t)^j j^{-1}.
\]

Finally, the well known identity \( \sum_{k=i}^{\infty} \binom{k}{i} \frac{t^k}{k!} = \frac{(e^t - 1)^i}{i!} \) implies

\[
(83) \quad f_i(y) \frac{(-1)^i}{i!} = \int_0^y \frac{(e^t - 1)^i}{i!} dt = \sum_{k \geq i} \binom{k}{i} \frac{y^{k+1}}{(k+1)!}.
\]

The first inequality of Fact 6 is an immediate consequence of (81), (83). The second inequality of Fact 6 follows from (83) combined with (82). In this case we have

\[
f_i(y) \frac{(-1)^i}{i!} = \sum_{j \geq i+1} \frac{(1 - e^y)^j}{j} \frac{(e^y - 1)^{j+1}}{(i+1)!} \left( 1 + (-1)^{i+1} R \right),
\]

where \( R = \sum_{k \geq 1} (1 - e^y)^k (i+1)/(i+1+k) \). Note that \( |e^y - 1| < 1/2 \) implies \( |R| \leq 1/2 \).

References

[1] M. Bloznelis. (2013). Degree and clustering coefficient in sparse random intersection graphs, The Annals of Applied Probability 23, 1254–1289.

[2] M. Bloznelis, E. Godehardt, J. Jaworski, V. Kurauskas, K. Rybarczyk. (2015). Recent Progress in Complex Network Analysis: Models of Random Intersection Graphs. In: Lausen, B., Krolak-Schwerdt, S., Böhmer, M. (eds) Data Science, Learning by Latent Structures, and Knowledge Discovery, pp. 69–78. Springer-Verlag Berlin Heidelberg.

[3] S. R. Blackburn and S. Gerke. (2009). Connectivity of the uniform random intersection graph, Discrete Mathematics 309 (16), 5130–5140,

[4] E.R. Canfield and C. Pomerance. (2002). On the problem of uniqueness for the maximal Stirling number(s) of the second kind, Integers, 2, paper A1.

[5] C. Cooper, A. Frieze. (2007). The cover time of sparse random graphs. Random Structures Algorithms 30, 1–16.

[6] C. Cooper, A. Frieze. (2008). The cover time of the giant component of a random graph. Random Structures Algorithms 32, 401–439.

[7] M. Deijfen and W. Kets. (2009). Random intersection graphs with tunable degree distribution and clustering. Probab. Engry. Inform. Sci., 23: 661–674.
[8] A. Frieze, M. Karoński. *Introduction to Random Graphs*. Cambridge University Press, 2016.

[9] J.-L. Guillaume, M. Latapy (2004). Bipartite structure of all complex networks, *Information Processing Letters*, 5, 215–221.

[10] S. Janson, T. Łuczak, and A. Ruciński, *Random Graphs*, Wiley, New York, 2000.

[11] M. Karoński, E. R. Scheinerman, and K. B. Singer-Cohen. (1999). On random intersection graphs: The subgraph problem. *Combinatorics, Probability and Computing*, 8:131–159.

[12] M. E. J. Newman, S. H. Strogatz, D. J. Watts. (2002). Random graphs with arbitrary degree distributions and their applications. *Physical Review* E 64: 026118.

[13] B. C. Rennie and A. J. Dobson. (1969). On Stirling numbers of the second kind, *Journal of Combinatorial Theory*, 7 (1969), 116–121.

[14] K. Rybarczyk. (2017). The coupling method for inhomogeneous random intersection graphs. *Electron. J. Combin*. 24, no. 2, Paper 2.10.

[15] S. Nikoletseas, Ch. Raptopoulos, and P. G. Spirakis. (2007). Expander Properties and the Cover Time of Random Intersection Graphs. In: Luděk, K. and Antonín, K. (eds.) Mathematical Foundations of Computer Science 2007, Springer Berlin Heidelberg, LNCS 4708, 44–55.

[16] A. Sinclair, M. Jerrum. (1989). Approximate Counting, Uniform Generation and Rapidly Mixing Markov Chains. *Information and Computation*, 82, 93–133.

[17] K. B. Singer-Cohen, *Random intersection graphs, PhD thesis*, Department of Mathematical Sciences, The Johns Hopkins University, 1995.

[18] P. G. Spirakis, S. Nikoletseas, and Ch. Raptopoulos. (2013). A Guided Tour in Random Intersection Graphs. In: Fomin, F. V., Freivalds, R., Kwiatkowska, M., Peleg, D. (eds.) Automata, Languages, and Programming 2013, Springer Berlin Heidelberg, LNCS 7966, 29–35.

[19] K. Rybarczyk, J. Jaworski, M. Bloznelis. A note on expansion properties of random intersection graphs. Manuscript.

[20] O. Yağan and A. M. Makowski. (2012). Zero–one laws for connectivity in random key graphs, *IEEE Transactions on Information Theory* 58(5), 2983–2999.