

The cover time of sparse random graphs.

Colin Cooper* Alan Frieze†

May 29, 2005

Abstract

We study the cover time of a random walk on graphs $G \in G_{n,p}$ when $p = \frac{c \log n}{n}, c > 1$. We prove that **whp** the cover time is asymptotic to $c \log \left(\frac{c}{c-1} \right) n \log n$.

1 Introduction

Let $G = (V, E)$ be a connected graph, let $|V| = n$, and $|E| = m$. For $v \in V$ let C_v be the expected time taken for a simple random walk W on G starting at v , to visit every vertex of G . The *cover time* C_G of G is defined as $C_G = \max_{v \in V} C_v$. The cover time of connected graphs has been extensively studied. It is a classic result of Aleliunas, Karp, Lipton, Lovász and Rackoff [1] that $C_G \leq 2m(n-1)$. It is also known (see Feige [7], [8]), that for any connected graph G

$$(1 - o(1))n \log n \leq C_G \leq (1 + o(1))\frac{4}{27}n^3.$$

In this paper we study the cover time of the random graph, $G \in G_{n,p}$. It was shown by Jonasson [11] that **whp**¹

(a) $C_G = (1 + o(1))n \log n$ if $\frac{np}{\log n} \rightarrow \infty$.

(b) If $c > 1$ is constant and $np = c \log n$ then $C_G > (1 + \alpha)n \log n$ for some constant $\alpha = \alpha(c)$.

Thus Jonasson has shown that when the expected average degree $(n-1)p$ grows faster than $\log n$, a random graph has the same cover time **whp** as the complete graph K_n , whose cover time is determined by the Coupon Collector problem. Whereas, when $np = \Omega(\log n)$ this is not the case.

In this paper we sharpen Jonasson's results for the case $np = c \log n$ where $\omega = (c-1) \log n \rightarrow \infty$. This condition on ω ensures that **whp** $G_{n,p}$ is connected, (see Erdős and Rényi [6]).

*Department of Mathematical and Computing Sciences, Goldsmiths College, London SW14 6NW, UK

†Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA15213. Supported in part by NSF grant CCR-9818411.

¹A sequence of events \mathcal{E}_n is said to occur *with high probability* (**whp**) if $\lim_n \rightarrow \infty \Pr(\mathcal{E}_n) = 1$

Theorem 1. *Suppose that $np = c \log n = \log n + \omega$ where $\omega = (c - 1) \log n \rightarrow \infty$ and $c = O(1)$. If $G \in G_{n,p}$, then **whp***

$$C_G \sim c \log \left(\frac{c}{c-1} \right) n \log n.$$

It follows from the results of Broder and Karlin [4] that if $c - 1 = \Omega(1)$ then **whp** $C_G = O(n \log n)$. This is because $G_{n,p}$ is then an expander **whp**. Thus it was known that **whp** $C_G = \Theta(n \log n)$ for this value of p .

Our paper improves on previous results by giving asymptotically exact values for the cover time. When $c - 1 = \Omega(1)$ we give the exact constant of multiplication in $C_G = \Theta(n \log n)$. Also, when $c - 1 = \frac{\omega}{\log n}$ where $\omega = o(\log n)$ we establish the novel result that $C_G \sim n \log n \log \log n$. The challenge is to find a technique which can make an accurate average case analysis of the random walk.

In the next section we give some properties that hold **whp** in $G_{n,p}$. In Section 4 we show that a graph with these properties has a cover time described by Theorem 1.

2 Outline of proof

In Section 3 we describe the properties of $G_{n,p}$ that will be needed for our proof. We call graphs with these properties *typical* and estimate the cover time of typical graphs.

For a walk starting at vertex u and a vertex $v \neq u$ and a time s , we define $\mathcal{A}_s(v)$ to be the event that W_u has not visited v by time s . We then argue that

$$C_u \leq t + \sum_{v \in V} \sum_{s > t} \Pr(\mathcal{A}_s(v)) \tag{1}$$

for any $t > 0$, see equation (14).

We then write

$$\Pr(\mathcal{A}_s(v)) = \prod_{i=1}^s \Pr(\mathcal{A}_i(v) \mid \mathcal{A}_{i-1}(v)) \tag{2}$$

and spend much of the time estimating $\Pr(\mathcal{A}_i(v) \mid \mathcal{A}_{i-1}(v))$ for i not too small.

Heuristically, a random walk of length s under the conditioning $\mathcal{A}_s(v)$ should be much like a random walk on $H = G[V \setminus \{v\}]$. Now the probability of $\bar{\mathcal{A}}_s(v)$ when we *forbid* a visit to v in the first $s - 1$ steps is close to

$$\sum_{w \in N_G(v)} \frac{d_H(w)}{2|E(H)|} \cdot \frac{1}{d_H(w)}.$$

Here, $d_H(w)$ is the degree of w in H and $\frac{d_H(w)}{2|E(H)|}$ is the steady state of a random walk on H . $N_G(v)$ is the set of neighbours of v in G . If H is an expander and s much larger than the mixing time then the sum of these probabilities is the probability that W_u is at a neighbour of v at time $s - 1$. The factor $\frac{1}{\text{degree}_H(w)}$ is then the probability we move to v at the next step. The main part of the analysis in Lemma 3 goes to show that

$$\Pr(\mathcal{A}_i(v) \mid \mathcal{A}_{i-1}(v)) \sim 1 - \frac{d(v)}{2m}$$

for large i . This is then used in (2) to estimate $\Pr(\mathcal{A}_s(v))$ from above and below.

The upper bound follows from (1) after some calculation.

The lower bound uses the Chebyshev inequality and requires the estimate $\Pr(\mathcal{A}_s(v_1) \wedge \mathcal{A}_s(v_2)) \sim \Pr(\mathcal{A}_s(v_1))\Pr(\mathcal{A}_s(v_2))$ for suitably chosen pairs v_1, v_2 .

3 Properties of $G_{n,p}$

Let $d(v)$ be the degree of vertex $v \in V$ and let δ, Δ denote the minimum and maximum degree. Let $\delta_v \geq 2$ be the minimum degree of a neighbour of v , *excluding neighbours of degree one*. Let $\text{dist}(u, v)$ denote the distance between the vertices u, v of the graph G .

Let $c = c(n) > 1$ where $c > 1$. We say that a graph G is *typical* if it has the structural properties **P0–P7** given below. The proof of the following lemma is given in the Appendix.

Lemma 1. *Let $p = \frac{c \log n}{n}$ where $\omega = (c - 1) \log n \rightarrow \infty$ and $c = O(1)$. Then **whp** $G \in G_{n,p}$ is typical.*

P0: G is connected.

P1: $\Delta(G) \leq \Delta_0 = (c + 10) \log n$ and

$$\delta(G) \geq \begin{cases} 1 & c \leq 1 + e^{-500} \\ \alpha \log n & c > 1 + e^{-500} \end{cases}$$

where $\alpha = \alpha^*/2$ and $\alpha^* > e^{-600}$ satisfies $c - 1 = \alpha^* \log(c\alpha^*)$.

P2: Call a vertex *small* if its degree is at most $\log n/20$. There are at most $n^{1/3}$ *small* vertices and no two small vertices are within distance $\leq \frac{\log n}{(\log \log n)^2}$ of each other.

P3: For $L \subseteq V$, $|L| \leq 4$, let $H = G[V \setminus L]$ be the subgraph of G induced by $V \setminus L$. For $S \subseteq V \setminus L$ let $e_H(S, \bar{S})$ be the number of edges of H with one end in S and the other in $\bar{S} = V \setminus (L \cup S)$. For all $H \subseteq G$ such that $\delta(H) \geq 1$, and for all $S \subseteq V \setminus L$, $|S| \leq n/2$,

$$\frac{e_H(S, \bar{S})}{d_H(S)} \geq \frac{1}{6}.$$

P4: Let $D(k)$ be the number of vertices of degree k in G . Let

$$\bar{D}(k) = n \binom{n-1}{k} p^k (1-p)^{n-1-k}$$

denote the expected size of $D(k)$ in $G_{n,p}$.

Define

$$\begin{aligned} K_0 &= \{k \in [1, \Delta_0] : \bar{D}(k) \leq (\log n)^{-2}\}. \\ K_1 &= \{1 \leq k \leq 15 : (\log n)^{-2} \leq \bar{D}(k) \leq \log \log n\}. \\ K_2 &= \{k \in [16, \Delta_0] : (\log n)^{-2} \leq \bar{D}(k) \leq (\log n)^2\}. \\ K_3 &= [1, \Delta_0] \setminus (K_0 \cup K_1 \cup K_2). \end{aligned}$$

P4a: If $k \in K_3$ then $\frac{1}{2}\overline{D}(k) \leq D(k) \leq 2\overline{D}(k)$, and

$$D(k) \begin{cases} = 0 & k \in K_0 \\ \leq (\log \log n)^2 & k \in K_1 \\ \leq (\log n)^4 & k \in K_2 \end{cases}$$

P4b: If $\omega \geq (\log n)^{2/3}$ then $K_1 = \emptyset$ and

$$\min\{k \in K_2\} \geq (\log n)^{1/2} \quad \text{and} \quad |K_2| = O(\log \log n).$$

P5: The number of edges $m = m(G)$ of G satisfies $|m - \frac{1}{2}cn \log n| \leq n^{1/2} \log n$.

P6: Let $k^* = \lceil (c-1) \log n \rceil$, $V^* = \{v : d(v) = k^*\}$ and let $B^* = \{v \in V^* : \text{dist}(v, w) \leq \frac{10 \log n}{(\log \log n)^2} \text{ for some } w \in V^*, w \neq v\}$. Then

$$|V^*| \geq \frac{1}{2}\overline{D}(k^*) \quad \text{and} \quad |B^*| \leq \frac{1}{10}\overline{D}(k^*).$$

Let $X = \{v : \delta_v \leq \alpha \log n\}$. Then

$$|V^* \cap X| \leq \frac{1}{10}\overline{D}(k^*).$$

P7 Call a cycle *short* if its length is at most $\frac{\log n}{10 \log \log n}$. The minimum distance between two short cycles is at least $\frac{\log n}{\log \log n}$ and the minimum distance between a small vertex and a short cycle is at least $\frac{\log n}{10 \log \log n}$.

P8: G has at least one triangle and at least one 5-cycle.

4 The cover time of a typical graph

In this section G denotes a fixed graph with vertex set $[n]$ which satisfies **P0–P8** and u is some arbitrary vertex from which a walk is started. For a subgraph H of G let $W_{u,H}$ denote a random walk on H which starts at vertex u and let $W_{u,H}(t)$ denote the walk generated by the first t steps. Let $X_{u,H}(t)$ be the vertex reached at step t and let $P_{u,H}^{(t)}(v) = \mathbf{Pr}(X_{u,H}(t) = v)$. Let $\pi_{u,H}(v)$ be the steady state probability of the random walk $W_{u,H}$. Note that properties **P1** and **P8** guarantee that the Markov chain associated with the walk is irreducible and aperiodic and therefore it has a steady state. For an unbiased random walk on a connected graph H with $m(H)$ edges, $\pi_H(v) = \pi_{u,H}(v) = \frac{d_H(v)}{2m(H)}$ where $d_H(v)$ denotes degree in H .

Let $H = H(v) = G[V \setminus \{v\}]$ if v is not a neighbour of a vertex w of degree 1, and let $H(v) = G[V \setminus \{v, w\}]$ if v has a neighbour w of degree 1. (Note that **P2** rules out a vertex having two neighbours of degree 1). For a subgraph H let $N_H(v)$ be the neighbourhood of v in H (i.e $N_H(v) = N_G(v) \cap V(H)$). When $H = G$ we drop the H from the above notation and often drop the u as well.

Lemma 2. *Let G be typical, then there exists a sufficiently large constant $K > 0$ such that if $\tau_0 = K \log n$ then for all $v \in V$, and for all $u, x \in H = H(v)$, after $t \geq \tau_0$ steps*

$$|P_{u,H}^{(t)}(x) - \pi_{u,H}(x)| = O(n^{-10}). \quad (3)$$

Proof The *conductance* Φ of the walk $W_{u,H}$ is defined by

$$\Phi(W_{u,H}) = \min_{\pi(S) \leq 1/2} \frac{e_H(S : \bar{S})}{d_H(S)}.$$

It follows from **P3** that the conductance Φ of the walk $W_{u,H}$ satisfies $\Phi \geq \frac{1}{6}$. Now it follows from Jerrum and Sinclair [10] that

$$|P_{u,H}^{(t)}(x) - \pi_{u,H}(x)| = O\left(n^{1/2} \left(1 - \frac{\Phi^2}{2}\right)^t\right). \quad (4)$$

For sufficiently large K , the RHS above will be $O(n^{-10})$ at τ_0 . We remark that there is a technical point here. The result of [10] assumes that the walk is lazy, and only makes a move to a neighbour with probability $1/2$ at any step. This leaves the steady state distribution unchanged, halves the conductance and (4) remains true. For us it is sufficient simply to keep the walk lazy for $2\tau_0$ steps until it is mixed. This is negligible compared to the cover time. □

For $v \neq u \in V$, let $\mathcal{A}_t(v)$ be the event that $W_{u,G}(t)$ does not visit v .

Lemma 3.

(a) *If $t > 2\tau_0$ and $\delta_v \geq 2$ then*

$$\begin{aligned} \Pr(\mathcal{A}_t(v)) &\leq \left(1 - \left(\left(\frac{\delta_v - 1}{\delta_v}\right)^2 - O\left(\frac{1}{\log n}\right)\right) \frac{d(v)}{2m}\right)^{t-2\tau_0} \Pr(\mathcal{A}_{2\tau_0}(v)) \\ \Pr(\mathcal{A}_t(v)) &\geq \left(1 - \left(\left(\frac{\delta_v}{\delta_v - 1}\right)^2 + O\left(\frac{1}{\log n}\right)\right) \frac{d(v)}{2m}\right)^{t-2\tau_0} \Pr(\mathcal{A}_{2\tau_0}(v)) \end{aligned}$$

(b) *Suppose that $v, v' \in V^* \setminus X$ (see **P6**) and that $\text{dist}(v, v') > \frac{10 \log n}{(\log \log n)^2}$. Then*

$$\Pr(\mathcal{A}_{2\tau_0}(v) \cap \mathcal{A}_{2\tau_0}(v')) \leq \left(1 - \left(1 + O\left(\frac{1}{\log n}\right)\right) \frac{k^*}{m}\right)^{t-2\tau_0}.$$

Proof The main idea of the proof is to show that the variation distance between a random walk of length τ_0 on H and a random walk on G , *conditioned* not to visit v is sufficiently small.

(a) Fix $w \neq v$ and $y \in N_H(v)$. Let $\mathcal{W}_k(y)$ denote the set of walks in $H(v)$ which start at w , finish at y , are of length $2\tau_0$ and which *leave* a vertex in the neighbourhood $N_H(v)$ exactly k times. (Note that the walk can leave $y \in N_H(v)$ without necessarily leaving $N_H(v)$). Let $\mathcal{W}_k = \bigcup_y \mathcal{W}_k(y)$ and let $W = (w_0, w_1, \dots, w_{2\tau_0}) \in \mathcal{W}_k(y)$. Let

$$\rho_W = \frac{\Pr(X_{w,G}(s) = w_s, s = 0, 1, \dots, 2\tau_0)}{\Pr(X_{w,H}(s) = w_s, s = 0, 1, \dots, 2\tau_0)}. \quad (5)$$

Then

$$1 \geq \rho_W \geq \left(\frac{\delta_v - 1}{\delta_v}\right)^k.$$

This is because

$$\frac{\Pr(X_{w,H}(s) = w_s \mid X_{w,H}(s-1) = w_{s-1})}{\Pr(X_{w,G}(s) = w_s \mid X_{w,G}(s-1) = w_{s-1})} = \begin{cases} 1 & w_{s-1} \notin N_G(v) \\ \frac{d_G(w_{s-1})}{d_G(w_{s-1})-1} & w_{s-1} \in N_G(v) \end{cases}$$

If $\mathcal{E} = \{X_{w,G}(\tau) \neq v, 0 \leq \tau \leq 2\tau_0\}$ then

$$\begin{aligned} \Pr(\mathcal{E}) &= \sum_{k \geq 0} \sum_{W \in \mathcal{W}_k} \Pr(W_{w,G}(2\tau_0) = W) \\ &= \sum_{k \geq 0} \sum_{W \in \mathcal{W}_k} \rho_W \Pr(W_{w,H}(2\tau_0) = W) \\ &\geq \sum_{k \geq 0} p_k \left(\frac{\delta_v - 1}{\delta_v} \right)^k \end{aligned}$$

where

$$p_k = \sum_{W \in \mathcal{W}_k} \Pr(W_{w,H}(2\tau_0) = W) = \Pr(W_{w,H}(2\tau_0) \in \mathcal{W}_k).$$

We will show in Lemma 4, below, that

$$p_0 + p_1 + p_2 \geq 1 - O((\log n)^{-1}) \quad (6)$$

which immediately implies that

$$\Pr(\mathcal{E}) \geq p_0 + p_1 \left(1 - \frac{1}{\delta_v}\right) + p_2 \left(1 - \frac{1}{\delta_v}\right)^2 \geq \left(1 - \frac{1}{\delta_v}\right)^2 - O((\log n)^{-1}).$$

Now fix y and write

$$\begin{aligned} \Pr(X_{w,G}(2\tau_0) = y \mid \mathcal{E}) &= \sum_{k \geq 0} \sum_{W \in \mathcal{W}_k(y)} \Pr(W_{w,G}(2\tau_0) = W) \Pr(\mathcal{E})^{-1} \\ &= \sum_{k \geq 0} \sum_{W \in \mathcal{W}_k(y)} \rho_W \Pr(W_{w,H}(2\tau_0) = W) \Pr(\mathcal{E})^{-1}. \end{aligned}$$

Now if

$$\begin{aligned} p_{k,y} &= \frac{\Pr(W_{w,H} \in \mathcal{W}_k(y))}{\Pr(X_{w,H}(2\tau_0) = y)} \\ &= \Pr(W_{w,H}(2\tau_0) \text{ leaves a vertex of } N_H(v) \text{ } k \text{ times} \mid X_{w,H}(2\tau_0) = y) \end{aligned}$$

then

$$\sum_{k \geq 0} p_{k,y} \left(\frac{\delta_v - 1}{\delta_v} \right)^k \leq \frac{\Pr(X_{w,G}(2\tau_0) = y \mid \mathcal{E})}{\Pr(X_{w,H}(2\tau_0) = y)} \leq \Pr(\mathcal{E})^{-1}.$$

We will show in Lemma 4, below, that

$$p_{0,y} + p_{1,y} + p_{2,y} \geq 1 - O((\log n)^{-1}) \quad (7)$$

and so

$$\left(\frac{\delta_v - 1}{\delta_v} \right)^2 - O\left(\frac{1}{\log n} \right) \leq \left| \frac{\Pr(X_{w,G}(2\tau_0) = y \mid \mathcal{E})}{\Pr(X_{w,H}(2\tau_0) = y)} \right| \leq \left(\frac{\delta_v}{\delta_v - 1} \right)^2 + O\left(\frac{1}{\log n} \right).$$

Taking w as $X_{u,G}(t - 2\tau_0 - 1)$, and conditioning on $\mathcal{A}_{t-2\tau_0-1}(v)$, we deduce that

$$\left(\frac{\delta_v - 1}{\delta_v}\right)^2 - O\left(\frac{1}{\log n}\right) \leq \left| \frac{\Pr(X_{u,G}(t-1) = y \mid \mathcal{A}_{t-1}(v))}{\Pr(X_{w,H}(2\tau_0) = y)} \right| \leq \left(\frac{\delta_v}{\delta_v - 1}\right)^2 + O\left(\frac{1}{\log n}\right).$$

Therefore

$$\begin{aligned} \Pr(\mathcal{A}_t(v) \mid \mathcal{A}_{t-1}(v)) &\geq 1 - \left(\left(\frac{\delta_v}{\delta_v - 1}\right)^2 + O\left(\frac{1}{\log n}\right) \right) \sum_{y \in N_H(v)} P_{w,H(v)}^{(2\tau_0)}(y) \frac{1}{d(y)} \\ &= 1 - \left(\left(\frac{\delta_v}{\delta_v - 1}\right)^2 + O\left(\frac{1}{\log n}\right) \right) \sum_{y \in N_H(v)} \left(\frac{d(y) - 1}{2m - 2d(v)} + O\left(\frac{1}{n^{10}}\right) \right) \frac{1}{d(y)} \\ &\geq 1 - \left(\left(\frac{\delta_v}{\delta_v - 1}\right)^2 + O\left(\frac{1}{\log n}\right) \right) \frac{1}{2m - 2d(v)} \left(d(v) - \sum_{y \in N_H(v)} \frac{1}{d(y)} \right) \\ &= 1 - \left(\left(\frac{\delta_v}{\delta_v - 1}\right)^2 + O\left(\frac{1}{\log n}\right) \right) \frac{d(v)}{2m}. \end{aligned}$$

Here we use **P2** to see that $\sum_{y \in N_H(v)} \frac{1}{d(y)} \leq \frac{40d(v)}{\log n}$.

Similarly,

$$\Pr(\mathcal{A}_t(v) \mid \mathcal{A}_{t-1}(v)) \leq 1 - \left(\left(\frac{\delta_v - 1}{\delta_v}\right)^2 - O\left(\frac{1}{\log n}\right) \right) \frac{d(v)}{2m}$$

and the lemma follows immediately. \square

Lemma 4. *Equations (6),(7) are valid.*

Proof Clearly, we only need to prove (7) and so fix $y \in N_H(v)$.

The main idea is to show that a random walk of length $2\tau_0$ from w to y is close in distribution to a random walk of length τ_0 from w to x followed by the reversal of a random walk W_3 of length τ_0 from y to x , where x is chosen from the the steady state of a walk. Each of W_1, W_3 is basically a random walk of length τ_0 and it easy to estimate the number of returns to $N_H(v)$ for such walks.

Let $\mathcal{W}(a, b, t)$ denote the set of walks in H from a to b of length t and for $W \in \mathcal{W}(a, b, t)$ let $\Pr(W) = \Pr(W_{a,H}(t) = W)$. Then for $x \in V(H)$ we have

$$\begin{aligned} \Pr(X_{w,H}(\tau_0) = x \mid X_{w,H}(2\tau_0) = y) &= \sum_{\substack{W_1 \in \mathcal{W}(w,x,\tau_0) \\ W_2 \in \mathcal{W}(x,y,\tau_0)}} \frac{\Pr(W_1)\Pr(W_2)}{\Pr(\mathcal{W}(w,y,2\tau_0))} \\ &= \pi_{x,H}^{-1} \sum_{\substack{W_1 \in \mathcal{W}(w,x,\tau_0) \\ W_2 \in \mathcal{W}(x,y,\tau_0)}} \frac{\Pr(W_1)\pi_{x,H}\Pr(W_2)}{\Pr(\mathcal{W}(w,y,2\tau_0))} \end{aligned}$$

and with W_3 equal to the reversal of W_2 ,

$$\begin{aligned}
&= \pi_{x,H}^{-1} \pi_{y,H} \sum_{\substack{W_1 \in \mathcal{W}(w,x,\tau_0) \\ W_3 \in \mathcal{W}(y,x,\tau_0)}} \frac{\Pr(W_1)\Pr(W_3)}{\Pr(\mathcal{W}(w,y,2\tau_0))} \\
&= \frac{\pi_{x,H}^{-1} \pi_{y,H}}{\Pr(\mathcal{W}(w,y,2\tau_0))} \Pr(\mathcal{W}(w,x,\tau_0)) \Pr(\mathcal{W}(y,x,\tau_0)) \\
&= \frac{\pi_{x,H}^{-1} \pi_{y,H}}{\Pr(\mathcal{W}(w,y,2\tau_0))} (\pi_{x,H} - O(n^{-10}))^2 \\
&= \pi_{x,H} - O(n^{-9} \log n).
\end{aligned}$$

It follows that the variation distance between $X_{w,H}(\tau_0)$ and a vertex chosen from the steady state distribution π_H is $O(n^{-8} \log n)$. Now given $x = X_{w,H}(\tau_0)$, $W_{w,H}(\tau_0)$ is a random walk of length τ_0 from w to x and $W_2 = (x = X_{w,H}(\tau_0), X_{w,H}(\tau_0 + 1), \dots, y = X_{w,H}(2\tau_0))$ is a random walk of length τ_0 from x to y . For $W \in \mathcal{W}(y,x,\tau_0)$ let $\mathbf{Q}(W)$ be the probability that $(y, X_{w,H}(2\tau_0 - 1), \dots, X_{w,H}(\tau_0)) = W$. Then we have

$$\begin{aligned}
\mathbf{Q}(W) &= (1 + O(n^{-8} \log n)) \frac{\pi_{x,H} \Pr(W^{reverse})}{\Pr(\mathcal{W}(x,y,\tau_0))} \\
&= (1 + O(n^{-8} \log n)) \frac{\pi_{y,H} \Pr(W)}{\Pr(\mathcal{W}(x,y,\tau_0))} \\
&= (1 + O(n^{-8} \log n)) \frac{\pi_{y,H} \pi_{x,H} \Pr(W)}{\pi_{x,H} \Pr(\mathcal{W}(x,y,\tau_0))} \\
&= (1 + O(n^{-8} \log n)) \frac{\pi_{y,H} \pi_{x,H} \Pr(W)}{\pi_{y,H} \Pr(\mathcal{W}(y,x,\tau_0))}
\end{aligned}$$

Thus if $W = (w_1, w_2, \dots, w_{\tau_0} = x)$ then

$$\mathbf{Q}(W \mid X_{w,H}(\tau_0) = x) = (1 + O(n^{-8} \log n)) \frac{\Pr(W)}{\Pr(\mathcal{W}(y,x,\tau_0))}$$

and so the distribution of $W_2^{reverse}$ is within variation distance $O(n^{-8} \log n)$ of that of a random walk of length τ_0 from y to a vertex x chosen with distribution π_H .

Thus the variation distance between the distribution of a random walk of length $2\tau_0$ from w to y and that of $W_1, W_3^{reversed}$ is $O(n^{-8} \log n)$ where W_1, W_3 are obtained by (i) choosing x from the steady state distribution and then (ii) choosing a random walk W_1 from w to x and a random walk W_3 from y to x . Furthermore, the variation distance between the distribution of W_1 and a random walk of length τ_0 from w is $O(n^{-9})$. Similarly, the variation distance between distribution of W_3 and a random walk of length τ_0 from y is $O(n^{-9})$.

Now consider W_1 and let Z_t be the distance of $X_{w,H}(t)$ from v . We observe from **P2** and **P7** that except for at most one value $\bar{a} \in J = [1, \frac{\log n}{2(\log \log n)^2}]$ we have

$$\Pr(Z_{t+1} = a + 1 \mid Z_t = a) \geq 1 - \frac{20}{\log n}, \quad a \in I \setminus \bar{a}.$$

and this will enable us to prove

$$\Pr(W_1 \text{ or } W_3 \text{ make a return to } N_H(v)) = O(1/\log n) \tag{8}$$

and this implies (7). (Note that a move from $N_H(v)$ to $N_H(v)$ has to be counted as a return here.)

To prove (8), let t_0 be the first time that W_1 visits $N_H(v)$. We have to estimate the probability that W_1 returns to $N_H(v)$ later on and so we can assume w.l.o.g. that $w \in N_H(v)$ i.e. $Z_0 = 1$.

It follows from **P2** and **P7** that

$$\Pr(Z_i = i + 1, i = 1, \dots, 6 \mid Z_0 = 1) \geq \left(1 - \frac{40}{\log n}\right)^6. \quad (9)$$

To check this consider two possibilities: Let $N^7(v)$ denote the set of vertices within distance 7 or less of v in G .

- (a) $N^7(v)$ does not contain a small vertex. Since there is at most one edge joining two vertices in $N^7(v)$, we see that $\Pr(Z_{i+1} > Z_i) = 1 - \frac{40}{\log n}$ for $i = 1, \dots, 6$ and (9) follows.
- (b) On the other hand, if there is a small vertex x in $N^7(v)$ then with probability $\geq 1 - \frac{20}{\log n}$ the first move from w takes us further away from x and (9) follows as before.

If $Z_3 = 4$ and there is a return to $N_H(v)$ then there exists $\tau \leq \tau_0$ such that $Z_\tau = 4, Z_{\tau+1} = 3$ and $Z_{\tau+2} \leq 3$. If there is no small vertex within distance 4 of v then **P2** and **P7** imply

$$\Pr(\exists \tau \leq \tau_0 : Z_\tau = 4, Z_{\tau+1} = 3, Z_{\tau+2} \leq 3) = O\left(\frac{\tau_0}{(\log n)^2}\right). \quad (10)$$

If there is a *unique* small vertex within distance 4 of v and $Z_6 = 7$ and there is a return to $N_H(v)$ then there exists $\tau \leq \tau_0$ such that $Z_\tau = 7, Z_{\tau+1} = 6$ and $Z_{\tau+2} = 5$ (no short cycles close to v now). We can then argue as in (10) that the probability of this is $O\left(\frac{\tau_0}{(\log n)^2}\right)$. This completes the proof of part (a) of the lemma.

(b) We simply run through the proof as in (a), replacing v by v, v' : $H = H(v, v') = G - \{v, v'\}$, $N_H(v, v') = N_G(v) \cup N_G(v')$. The proof of (7) remains valid because v, v' are far apart. \square

4.1 The upper bound on cover time

From here on, A_1, A_2, \dots are a sequence of unspecified positive constants.

Let $t_0 = \lceil 2m \log \frac{c}{c-1} \rceil$. We now prove for typical graphs, that for any vertex $u \in V$

$$C_u \leq t_0 + o(m). \quad (11)$$

Let $T_G(u)$ be the time taken to visit every vertex of G by the random walk W_u . Let U_t be the number of vertices of G which have not been visited by W_u at step t . We note the following:

$$\Pr(T_G(u) > t) = \Pr(U_t > 0) \leq \min\{1, \mathbf{E} U_t\}, \quad (12)$$

$$C_u = \mathbf{E} T_G(u) = \sum_{t>0} \Pr(T_G(u) > t) \quad (13)$$

It follows from (12,13) that for all t

$$C_u \leq t + \sum_{s>t} \mathbf{E} U_s = t + \sum_{v \in V} \sum_{s>t} \Pr(\mathcal{A}_s(v)). \quad (14)$$

Now, by Lemma 3, for $s > 2\tau_0$,

$$\begin{aligned} \Pr(\mathcal{A}_s(v)) &\leq \left(1 - \left(\left(\frac{\delta_v - 1}{\delta_v}\right)^2 - \frac{A_1}{\log n}\right) \frac{d(v)}{2m}\right)^{s-2\tau_0} \Pr(\mathcal{A}_{2\tau_0}(v)) \\ &\leq \exp\left(-\frac{sd(v)}{2m} \left(1 - \frac{A_2}{\log n}\right)\right), \quad \text{if } \delta_v \geq \alpha \log n \end{aligned}$$

where α is as in **P1**.

Then from **P4**,

$$\begin{aligned} \mathbf{E} U_s &\leq \sum_{i=1}^3 \sum_{k \in K_i} \sum_{\substack{d(v)=k \\ v \notin X}} \Pr(\mathcal{A}_s(v)) + \sum_{v \in X} \Pr(\mathcal{A}_s(v)) \\ &\leq \sum_{i=1}^3 \sum_{k \in K_i} D(k) \exp\left(-\frac{kd(v)}{2m} \left(1 - \frac{A_2}{\log n}\right)\right) + \sum_{v \in X} \Pr(\mathcal{A}_s(v)) \\ &\leq T_3(s) + T_1(s) + T_2(s) + T_X(s) \end{aligned} \tag{15}$$

where

$$\begin{aligned} T_3(s) &= 2 \sum_{k=1}^{n-1} n \binom{n-1}{k} p^k (1-p)^{n-1-k} e^{-\frac{sk}{2m} \left(1 - \frac{A_2}{\log n}\right)}, \\ T_i(s) &= \sum_{k \in K_i} D(k) e^{-\frac{sk}{2m} \left(1 - \frac{A_2}{\log n}\right)}, \quad i = 1, 2 \end{aligned}$$

and

$$\begin{aligned} T_X(s) &= \sum_{v \in X} \left(1 - \left(\left(\frac{\delta_v - 1}{\delta_v}\right)^2 - O\left(\frac{1}{\log n}\right)\right) \frac{d(v)}{2m}\right)^{s-2\tau_0} \\ &\leq 2 \sum_{v \in X} \exp\left\{-\left(\left(\frac{\delta_v - 1}{\delta_v}\right)^2 - \frac{A_3}{\log n}\right) \frac{sd(v)}{2m}\right\}. \end{aligned}$$

Now for $\gamma > 0$,

$$\sum_{s=t_0+1}^{\infty} e^{-\gamma s} = \frac{1}{e^\gamma - 1} e^{-\gamma t_0} \leq \gamma^{-1} e^{-\gamma t_0}. \tag{16}$$

Let $\lambda = \frac{t_0}{2m} \left(1 - \frac{A_2}{\log n}\right)$. Applying (16) we get

$$\begin{aligned}
\sum_{s=t_0+1}^{\infty} T_3(s) &\leq 3m \sum_{k=1}^{n-1} \frac{n}{k} \binom{n-1}{k} p^k (1-p)^{n-k-1} e^{-k\lambda} \\
&\leq 6 \frac{m}{p} e^\lambda \sum_{k=1}^{n-1} \binom{n}{k+1} p^{k+1} (1-p)^{n-k-1} e^{-(k+1)\lambda} \\
&< 7 \frac{m}{p} \frac{c}{c-1} \left(1-p+pe^{-\lambda}\right)^n \\
&\leq 7 \frac{mn}{(c-1)\log n} e^{-np+npe^{-\lambda}} \\
&\leq 8 \frac{me^{2A_2}}{(c-1)\log n} \\
&= o(m).
\end{aligned} \tag{17}$$

We have used the estimation,

$$\begin{aligned}
npe^{-\lambda} &\leq (c \log n) \left(\frac{c-1}{c}\right) \left(1 + \frac{1}{c-1}\right)^{A_2/\log n} \\
&\leq (1 + O(n^{-1}))((c-1)\log n) \left(1 + \frac{2A_2}{(c-1)\log n}\right).
\end{aligned}$$

Note that we have used $(c-1)\log n \rightarrow \infty$ to get the second line.

Continuing we get

$$\begin{aligned}
\sum_{s=t_0+1}^{\infty} T_1(s) &\leq A_4 m \sum_{k \in K_1} \frac{(\log \log n)^2}{k} e^{-k\lambda} \\
&= o(m)
\end{aligned} \tag{18}$$

since either (i) $\omega \geq (\log n)^{2/3}$ and $K_1 = \emptyset$ or (ii) $\omega < (\log n)^{2/3}$ and $e^\lambda \geq (1 - o(1))(\log n)^{1/3}$.

$$\begin{aligned}
\sum_{s=t_0+1}^{\infty} T_2(s) &\leq A_5 m \sum_{k \in K_2} \frac{(\log n)^4}{k} e^{-k\lambda} \\
&= o(m)
\end{aligned} \tag{19}$$

since either (i) $\omega \geq (\log n)^{2/3}$ and $\min\{k \in K_2\} \geq (\log n)^{1/2}$ and $|K_2| = O(\log \log n)$ or (ii) $\omega < (\log n)^{2/3}$ and $e^\lambda \geq (1 - o(1))(\log n)^{1/3}$.

Note now that $\delta_v \geq 2$ and if $v \in X$ (see **P6**) then from **P2** $d(v) \geq \log n/20$. Thus

$$\begin{aligned}
\sum_{s=t_0+1}^{\infty} T_X(s) &\leq \sum_{s=t_0+1}^{\infty} \sum_{v \in X} \exp \left\{ -\frac{sd(v)}{10m} \right\} \\
&\leq \sum_{v \in X} \frac{10m}{d(v)} \exp \left\{ -\frac{t_0 d(v)}{10m} \right\} \\
&\leq \sum_{v \in X} \frac{200m}{\log n} \exp \left\{ -\frac{t_0 \log n}{200m} \right\} && \text{by **P2**} \\
&\leq \sum_{v \in X} \frac{200m}{\log n} \left(\frac{c-1}{c} \right)^{\log n/201} \\
&= o(m)
\end{aligned} \tag{20}$$

since either (i) $c \geq 1 + e^{-500}$ and $X = \emptyset$ or (ii) $c < 1 + e^{-500}$, in which case we use $(c-1)/c \leq e^{-500}$.

As $C_G = \max_{u \in V} C_u$, the upper bound on C_G now follows from (11), (15), (17), (18), (19), (20) and (14) with $t = t_0$. \square

4.2 The lower bound on cover time

For any vertex u , we can find a set of vertices S such that at time $t_1 = t_0(1 - \epsilon)$, $\epsilon \rightarrow 0$, the probability the set S is covered by the walk W_u tends to zero. Hence $T_G(u) > t_1$ **whp** which implies that $C_G \geq (1 - o(1))t_0$.

We construct S as follows. Let k^*, V^*, B^* be as defined in Property **P6**.

Let $S^* = V^* \setminus (B^* \cup X)$ and let

$$\epsilon = \frac{10}{(c-1) \log c / (c-1)} \frac{\log \log n}{\log n} = o(1) \text{ and } \delta = \frac{(\log n)^3}{|S^*|}.$$

Note that

$$\bar{D}(k^*) = \Omega \left(\frac{n^{(c-1) \ln(c/(c-1))}}{\sqrt{(c-1) \log n}} \right) = \Omega((\log n)^a) \tag{21}$$

for any constant $a > 0$. Then **P6** implies that $|S^*| = \Omega((\log n)^a)$ for any constant $a > 0$.

Now for $v, w \neq u$ let $\mathcal{A}_t(v, w)$ be the event that W has not visited v or w by step t .

Let $Q \subseteq S^*$ be given by

$$Q = \{v \in S^* : \Pr(\mathcal{A}_{2\tau_0}(v)) < 1 - \delta, \text{ or } \Pr(\mathcal{A}_{2\tau_0}(v, w)) < (1 - \delta)^2, \text{ for some } w \in S^*\}.$$

Now in time $2\tau_0$, W can visit at most $2\tau_0 + 1$ vertices and so

$$\sum_{v \in V} \Pr(\bar{\mathcal{A}}_{2\tau_0}(v)) \leq 2\tau_0 + 1 \text{ and } \sum_{v, w \in V} \Pr(\bar{\mathcal{A}}_{2\tau_0}(v, w)) \leq \binom{2\tau_0 + 1}{2}.$$

Thus

$$|Q| \leq \frac{2\tau_0 + 1}{\delta} + \frac{2\tau_0(2\tau_0 + 1)}{2(1 - (1 - \delta)^2)} = o(|S^*|).$$

Therefore, if $S = S^* \setminus Q$,

$$|S| \geq \frac{\overline{D}(k^*)}{3}.$$

Let $S(t)$ denote the subset of S which has not been visited by W by time t . Now

$$\mathbf{E} |S(t)| \geq \sum_{v \in S} \left(1 - \left(1 + \frac{A_6}{\log n} \right) \frac{k^*}{2m} \right)^{t-2\tau_0} \Pr(\mathcal{A}_{2\tau_0}(v)).$$

Setting $t = t_1$ we have

$$\begin{aligned} \mathbf{E} |S(t_1)| &= \Omega \left(\frac{n^{(c-1) \log c / (c-1)}}{\sqrt{(c-1) \log n}} \exp \left(-\frac{k^*}{2m} t_1 \right) \right) \\ &= \Omega \left(\frac{n^{\epsilon(c-1) \log c / (c-1)}}{\sqrt{(c-1) \log n}} \right) \\ &= \Omega((\log n)^9). \end{aligned} \tag{22}$$

Let $Y_{v,t}$ be the indicator for the event that $W_u(t)$ has not visited vertex v at time t . As $v, w \in S$ are not adjacent, and have no common neighbours, when we delete v, w the total degree of $H(v, w)$ is $2m - 2d(v) - 2d(w)$, and $d(v) = d(w) = k^*$. It follows from Lemma 3(b) that for $v, w \in S$

$$\begin{aligned} \mathbf{E} (Y_{v,t_1} Y_{w,t_1}) &\leq \left(1 - \left(1 + O \left(\frac{1}{\log n} \right) \right) \frac{k^*}{m} \right)^{t_1-2\tau_0} \\ &\leq (1 + o(1)) \mathbf{E} Y_{v,t_1} \mathbf{E} Y_{w,t_1}. \end{aligned} \tag{23}$$

It follows therefore that

$$\Pr(S(t_1) \neq \emptyset) \geq \frac{(\mathbf{E} |S(t_1)|)^2}{\mathbf{E} |S(t_1)|^2} = \frac{1}{\frac{\mathbf{E} |S_{t_1}| (|S_{t_1}| - 1)}{(\mathbf{E} |S(t_1)|)^2} + (\mathbf{E} |S_{t_1}|)^{-1}} = 1 - o(1)$$

from (22) and (23). □

Acknowledgement We thank Johan Jonasson for pointing out a significant error in earlier draft.

References

- [1] R. Aleliunas, R.M. Karp, R.J. Lipton, L. Lovász and C. Rackoff, Random Walks, Universal Traversal Sequences, and the Complexity of Maze Problems. *Proceedings of the 20th Annual IEEE Symposium on Foundations of Computer Science* (1979) 218-223.
- [2] B. Bollobás, *Random graphs* (Second edition), Cambridge University Press (2001).
- [3] B. Bollobás, T. Fenner and A.M. Frieze, An algorithm for finding Hamilton paths and cycles in random graphs, *Combinatorica* 7 (1987) 327-341.
- [4] A.Z. Broder and A. Karlin, Bounds on the cover time, *Journal of Theoretical Probability* 2 (1989) 101-120.
- [5] C. Cooper and A.M. Frieze, Crawling on web graphs, to appear in *Proceedings of STOC 2002*.

- [6] P. Erdős and A. Rényi, On random graphs I, *Publ. Math. Debrecen* 6 (1959) 290-297.
- [7] U. Feige, A tight upper bound for the cover time of random walks on graphs, *Random Structures and Algorithms* 6 (1995) 51-54.
- [8] U. Feige, A tight lower bound for the cover time of random walks on graphs, *Random Structures and Algorithms* 6 (1995) 433-438.
- [9] S.Janson, T.Luczak and A.Ruciński, *Random graphs*, Wiley (2000).
- [10] M. Jerrum and A. Sinclair. The Markov chain Monte Carlo method: an approach to approximate counting and integration. In *Approximation Algorithms for NP-hard Problems*. (D. Hochbaum ed.) PWS (1996) 482-520
- [11] J. Jonasson, On the cover time of random walks on random graphs, *Combinatorics, Probability and Computing*, 7 (1998), 265-279.

5 Appendix: Typical graph properties

A proof of **P0,P1** can be found in Bollobás [2] or Janson, Luczak and Ruciński [9].
A proof of **P2** can be found in [3].

P3: Case of $1 \leq s = |S| \leq n/(c \log n)$.

We first prove that **whp** $e_G(S, S) \leq s \log \log n$. Now

$$\begin{aligned}
& \Pr(\exists S : e_G(S, S) \geq s \log \log n) \\
& \leq \binom{n}{s} \binom{\binom{s}{2}}{s \log \log n} p^{s \log \log n} \leq \left(\frac{ne}{s}\right)^s \left(\frac{spe}{2 \log \log n}\right)^{s \log \log n} \\
& \leq \exp\left(-s \left(\log \log n \cdot \log\left(\frac{2n \log \log n}{cse \log n}\right) - \log \frac{ne}{s}\right)\right) \\
& = o(n^{-2}).
\end{aligned}$$

By property **P0** and the definition of H , both G and H contain no isolated vertices and hence $d(S) > 0$. We write $e(S, \bar{S})/d(S) = 1 - 2e(S, S)/d(S)$. Partition S into sets S_1 and S_2 , where S_1 are the vertices of S of degree at most $(\log n)/10$. Let T_1 be the neighbour set of S_1 in S_2 and let T_2 be the neighbour set of S_1 in \bar{S} . By property **P1** the set S_1 induces no edges, and the neighbours of vertices of S_1 are distinct. Thus

$$\begin{aligned}
\frac{2e_H(S, S)}{d_H(S)} & \leq \frac{2(|T_1| + |S_2| \log \log n)}{2|T_1| + |T_2| + |S_2|((\log n)/10 - |L|)} \\
& \leq \frac{2 + \log \log n}{(\log n)/10} = o(1).
\end{aligned}$$

Now use

$$d_H(S) = e_H(S, \bar{S}) + 2e_H(S, S).$$

Case of $n/(c \log n) \leq s \leq n/2$.

The expected value of $e_H(S, \bar{S})$ is at least $\mu = s(n - s - 4)p$. Thus from Chernoff bounds, for fixed s ,

$$\begin{aligned} \Pr(\exists S : e_H(S, \bar{S}) \leq \mu/2) &\leq \binom{n}{s} e^{-\frac{c}{9} \frac{s(n-s-4)}{n} \log n} \\ &\leq \exp\left(-s \left(\frac{c}{18} \log n - \log \frac{ne}{s}\right)\right) \\ &= o(n^{-2}). \end{aligned}$$

We note that $\mathbf{E} d_H(S) = 2\binom{s}{2}p + s(n - s - |L|)p$. Thus

$$\begin{aligned} \Pr(\exists S : d_H(S) \geq \frac{3}{2} \mathbf{E} d_H(S)) &\leq \binom{n}{s} e^{-\frac{c}{20} s \log n} \\ &\leq \exp\left(-s \left(\frac{c}{20} \log n - \log \frac{ne}{s}\right)\right) \\ &= o(n^{-2}). \end{aligned}$$

Thus

$$\frac{e_H(S, \bar{S})}{d_H(S)} \geq \frac{\frac{1}{2}s(n-s-4)p}{\frac{3}{2}(2\binom{s}{2}p + s(n-s)p)} \geq \frac{1}{6}.$$

P4a: First observe that

$$\Pr(\exists k \in K_0 : D(k) > 0) \leq \sum_{k \in K_0} \bar{D}(k) = \frac{|K_0|}{(\log n)^2} = O\left(\frac{1}{\log n}\right).$$

Then

$$\Pr(\exists k \in K_1 : D(k) > (\log \log n)^2) \leq \sum_{k \in K_1} \frac{\bar{D}(k)}{(\log \log n)^2} = O\left(\frac{1}{\log \log n}\right).$$

Similarly,

$$\Pr(\exists k \in K_2 : D(k) > (\log n)^4) \leq \sum_{k \in K_2} \frac{\bar{D}(k)}{(\log n)^4} = O\left(\frac{1}{\log n}\right).$$

A simple calculation gives that for our range of values of p

$$\mathbf{E} (D(k)(D(k) - 1)) = \bar{D}(k)^2 \left(1 + O\left(\frac{\log n}{n}\right)\right).$$

Thus

$$\mathbf{Var} D(k) = \bar{D}(k) \left(1 + O\left(\frac{\bar{D}(k) \log n}{n}\right)\right).$$

Applying the Chebychef inequality we see that

$$\Pr(D(k) \leq \frac{1}{2} \bar{D}(k) \text{ or } D(k) \geq 2 \bar{D}(k)) \leq \frac{4 \mathbf{Var} D(k)}{\bar{D}(k)^2} = \frac{4}{\bar{D}(k)} \left(1 + O\left(\frac{\log n}{n}\right)\right).$$

So, as $|K_3| = O(\log n)$,

$$\Pr(\exists k \in K_3 : D(k) \leq \frac{1}{2} \bar{D}(k) \text{ or } D(k) \geq 2 \bar{D}(k)) = O\left(\frac{1}{\log n}\right).$$

P4b: The sequence $(\overline{D}(k), k \geq 0)$ is unimodal and

$$\frac{\overline{D}(k+1)}{\overline{D}(k)} \sim \frac{c \log n}{k+1} \quad \text{when } k = O(\log n). \quad (24)$$

Moreover, for $k \leq \Delta_0$ there is a positive constant $A = A(k)$ such that

$$\overline{D}(k) \sim An^{1-c} \left(\frac{ce \log n}{k} \right)^k \frac{1}{k^{1/2}}. \quad (25)$$

Suppose first that there exists $k \in K_1 \cup K_2$ such that $k < (\log n)^{1/2}$. It follows from (25) that $c-1 < (\log n)^{-1/3}$, for if $c-1 \geq (\log n)^{-1/3}$ and $k < (\log n)^{1/2}$ then $\overline{D}(k) = o((\log n)^{-2})$.

Now suppose that $k \in K_2$ implies $k \geq (\log n)^{1/2}$. Observe from (25) that both $\overline{D}(\lfloor (c-c^{1/3}) \log n \rfloor)$ and $\overline{D}(\lfloor (c+c^{1/3}) \log n \rfloor)$ are much greater than $(\log n)^2$. Thus either $k \leq (c-c^{1/3}) \log n$ or $k \geq (c+c^{1/3}) \log n$. In either case, we see from iterating (24) that $|K_2| = O(\log \log n)$.

P5: This follows immediately from Chernoff bounds.

P6: From (21) we see that $\lceil (c-1) \log n \rceil \in K_3$ for c constant. That $|V^*| \geq \frac{1}{2} \overline{D}(k^*)$ now follows from **P3**. Now $|B^*| \leq |\{(v, w) \in (V^*)^2 : \text{dist}(v, w) \leq d = \frac{10 \log n}{(\log \log n)^2}\}|$. Therefore

$$\mathbf{E} |B^*| \leq \overline{D}(k^*)^2 \sum_{k=1}^d n^k p^{k+1} = o(\overline{D}(k^*)),$$

and the second part of **P6** follows from the Markov inequality. The third part is a similar first moment calculation.

P7: A proof of similar results can be found in [3].

P8: The expected number of triangles tends to infinity and we can use the Chebychef inequality to show that one exists **whp**. The same argument will work for 5-cycles. \square