The cover time of sparse random graphs.

Colin Cooper * Alan Frieze†

February 25, 2003

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 with $V = [n]$, and let $|E| = m$. 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 [6], [7]), 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 [10] that **whp**

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

(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 $\omega = (c-1)\log n \to \infty$. This condition on ω ensures that **whp** $G_{n,p}$ is connected, (see Erdős and Rényi [5]).

*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.

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

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

Our method of proof follows an idea from our study of random walks on web graphs [4]. It can easily be adapted to the case of a random r-regular graph G_r on vertex set v where r is sufficiently large.

Theorem 2. **Whp** $C_{G_r} = (1 + O(r^{-1}))n \log n$.

In Section 2 we give some properties that hold **whp** in $G_{n,p}$. In Section 3 we show that graphs with these properties have cover time given by Theorem 1.

2 Properties of $G_{n,p}$

Let δ, Δ denote minimum and maximum degree, and let $d(u, v)$ denote distance between the vertices u, v of the graph G.
Let $np = c \log n$ where $c > 1$. **Whp** $G \in G_{n,p}$ has the structural properties **P0**–**P7** given below. We say that a graph G with these properties is *typical*. The proof of the following lemma is given in the Appendix.

Lemma 1. *Let $p = \frac{c \log n}{n}$ where $\omega = (c-1)\log n \to \infty$ 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

$$\delta(G) \geq \begin{cases} 1 & c \leq 1 + e^{-50} \\ \alpha \log n & c > 1 + e^{-50} \end{cases}$$

where $\alpha = \alpha^*/2$ and $\alpha^* > e^{-60}$ satisfies $c - 1 = \alpha^* \log(ce/\alpha^*)$.
P2: There are at most $n^{1/3}$ *small* vertices (i.e of degree at most $\log n/20$) 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 $H = G - L$. For $S \subseteq V - L$ let $e_H(S, \overline{S})$ be the number of edges of H with one end in S and the other in $\overline{S} = V - (L \cup S)$.

For all $H \subseteq G$ such that $\delta(H) \geq 1$, and for all $S \subseteq V - L$, $|S| \leq n/2$

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

P4: Let $\overline{D}(k) = n\binom{n-1}{k} p^k (1-p)^{n-1-k}$ denote the expected size of $D(k)$ in $G_{n,p}$.
Let $D(k)$ be the number of vertices of degree k in G. Define

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

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

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

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

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

Let $X = \{v : \delta_v \leq \alpha \log n\}$ where δ_v is the minimum degree of a neighbour of v, *excluding neighbours of degree one.* Then

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

P7: The minimum distance between two *small* cycles of length $\leq \frac{\log n}{10\log\log n}$ is at least $\frac{\log n}{\log\log n}$ and the minimum distance between a small vertex and a small cycle is at least $\frac{\log n}{10\log\log n}$.

3 The cover time of a typical graph

In this section G denotes a fixed graph with vertex set v which satisfies **P0**–**P7** 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 $W_{u,H}$. For an unbiased random walk on a connected graph H with $m(H)$ edges, $\pi_{u,H}(v) = \frac{d_H(v)}{2m(H)}$ where $d_H(v)$ denotes degree in G.
Our definition of typical does not rule out G being bipartite, even though $G_{n,p}$ is non-bipartite **whp** for these values of p. In which case there is no steady state distribution. We therefore assume that in such a case, at each step, the random walk does nothing with probability 1/2 and only moves to an adjacent vertex with probability 1/2. We double the expected time to cover the vertices, but the asymptotic number of non-trivial steps remains the same.
Let $H(v) = G - \{v\}$ if v is not a neighbour of a vertex w of degree 1, and let $H(v) = G - \{v, w\}$ if v has a neighbour w of degree 1. Note that **P2** rules out a neighbour 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*

$$|P_{u,H}^{(t)}(x) - \pi_{u,H}(x)| = O(n^{-10}). \tag{1}$$

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

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

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 [9] that

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

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 [9] assumes that the walk is *lazy*, and only makes a move to a neighbour with probability 1/2 at any step. This halves the conductance but still (2) 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. $\square$
For $v \neq u \in V$, let $\boldsymbol{A}_t(v)$ be the event that $W_{u,G}(t)$ has not visited v.

Lemma 3. *Let δ_v be as defined in* **P6**. *If $t > 2\tau_0$ then*

$$\mathbf{Pr}(\boldsymbol{A}_t(v)) \leq \left(1 - \left(\frac{\delta_v - 1}{\delta_v} - O\left(\frac{1}{\log n}\right)\right)\frac{d(v)}{2m}\right)^{t-2\tau_0} \mathbf{Pr}(\boldsymbol{A}_{2\tau_0}(v))$$

$$\mathbf{Pr}(\boldsymbol{A}_t(v)) \geq \left(1 - \left(\frac{\delta_v}{\delta_v - 1} + O\left(\frac{1}{\log n}\right)\right)\frac{d(v)}{2m}\right)^{t-2\tau_0} \mathbf{Pr}(\boldsymbol{A}_{2\tau_0}(v))$$

Proof Fix $w \neq v$ and $y \in N_H(v)$. Let $\mathcal{W}_k(y)$ denote the set of walks of length $2\tau_0$ in $H(v)$ which start at w, finish at y and which *leave* a vertex in the neighbourhood $N_H(v)$ exactly k times. Let $\mathcal{W}_k = \bigcup_y \mathcal{W}_k(y)$ and let $W = (w_0, w_1, \ldots, w_{2\tau_0}) \in \mathcal{W}_k(y)$. Let

$$\rho_W = \frac{\mathbf{Pr}(X_{w,G}(s) = w_s, s = 0, 1, \ldots, 2\tau_0)}{\mathbf{Pr}(X_{w,H}(s) = w_s, s = 0, 1, \ldots, 2\tau_0)}. \tag{3}$$

Then

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

This is because

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

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

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

where $p_k = \mathbf{Pr}(W_{w,H}(2\tau_0) \in \mathcal{W}_k)$. We will show later that

$$p_0 + p_1 \geq 1 - O((\log n)^{-1}) \tag{4}$$

which immediately implies that

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

Now fix y and write

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

Let $p_{k,y}$ denote the probability of the event that $W_{w,H}(2\tau_0)$ leaves a vertex of $N_H(v)$ k times, conditional on $X_{w,H}(2\tau_0) = y$. Thus

$$p_{k,y} = \frac{\mathbf{Pr}(W_{w,H} \in \mathcal{W}_k(y))}{\mathbf{Pr}(X_{w,H}(2\tau_0) = y)},$$

and

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

We will show later that

$$p_{0,y} + p_{1,y} \geq 1 - O((\log n)^{-1}) \tag{5}$$

and so

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

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

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

Therefore

$$\begin{aligned}
&\mathbf{Pr}(\overline{\boldsymbol{A}_t(v)} \mid \boldsymbol{A}_{t-1}(v)) \leq \\
&\left(\frac{\delta_v}{\delta_v - 1} + O\left(\frac{1}{\log n}\right)\right) \sum_{x \in N_H(v)} P^{(2\tau_0)}_{w,H(v)}(x)\frac{1}{d(x)} = \\
&\left(\frac{\delta_v}{\delta_v - 1} + O\left(\frac{1}{\log n}\right)\right) \times \\
&\quad \sum_{x \in N_H(v)} \left(\frac{d(x) - 1}{2m - 2d(v)} + O\left(\frac{1}{n^{10}}\right)\right)\frac{1}{d(x)} \geq \\
&\left(\frac{\delta_v}{\delta_v - 1} + O\left(\frac{1}{\log n}\right)\right) \times \\
&\quad \left(\frac{d(v)}{2m} - \frac{1}{2m - 2d(v)} \sum_{x \in N_H(v)} \frac{1}{d(x)}\right) = \\
&\quad \left(\frac{\delta_v}{\delta_v - 1} + O\left(\frac{1}{\log n}\right)\right)\frac{d(v)}{2m}.
\end{aligned}$$

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

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

and the lemma follows immediately.

Proof of (4,5). Clearly, we only need to prove (5) and so fix $y\in N_H(v)$.

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 $\mathbf{Pr}(W)=\mathbf{Pr}(W_{a,H}(t)=W)$. Then for $x\in V(H)$ we have

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

and with W_3 equal to the reversal of W_2,

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

It follows that the variation distance between 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 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 $O(n^{-9})$.

Let $\mathcal{E}_i,\ i=1,3$ be the event that W_i contains more than one vertex of $N_H(v)$. We will prove that

$$\mathbf{Pr}(\mathcal{E}_i)=O(1/\log n) \tag{6}$$

which will imply (5).

To prove (6), we 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 $\bar{a}\in J=[1,\frac{\log n}{(\log\log n)^2}]$ we have

$$\mathbf{Pr}(Z_{t+1}=a+1\mid Z_t=a)\ge 1-\frac{20}{\log n},\qquad a\in I\backslash\bar{a}.$$

Let t_0 be the first time that W_1 visits $N_H(v)$. We will assume the worst case, namely that $w\in N_H(v)$ and thus $Z_0=1$. It follows then from **P2** and **P7** that

$$\mathbf{Pr}(Z_2=3\mid Z_0=1)\ge\left(1-\frac{20}{\log n}\right)^3.$$

If $Z_2=3$ and there is a subsequent return to $N_H(v)$ then there exists $\tau\le\tau_0$ such that $Z_{\tau+2}\le Z_{\tau+1}\le Z_\tau$. But now **P2** and **P7** imply

$$\mathbf{Pr}(\exists\tau\le\tau_0:\ Z_{\tau+2}\le Z_{\tau+1}\le Z_\tau)=O\left(\frac{\tau_0}{(\log n)^2}\right)$$

and this proves (6). □

3.1 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 $u\in V$

$$C_u\le t_0+o(m). \tag{7}$$

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:

$$\begin{aligned}
\mathbf{Pr}(T_G(u)>t)&=\mathbf{Pr}(U_t>0)\le\min\{1,\mathbf{E}\,U_t\}, && (8)\\
C_u=\mathbf{E}\,T_G(u)&=\sum_{t>0}\mathbf{Pr}(T_G(u)\ge t) && (9)
\end{aligned}$$

It follows from (8,9) that for all t

$$C_u\le t+\sum_{s>t}\mathbf{E}\,U_s=t+\sum_{v\in V}\sum_{s>t}\mathbf{Pr}(\boldsymbol{A}_s(v)). \tag{10}$$

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

$$\begin{aligned}
&\mathbf{Pr}(\boldsymbol{A}_s(v))\le\\
&\left(1-\left(\frac{\delta_v-1}{\delta_v}-\frac{A_1}{\log n}\right)\frac{d(v)}{2m}\right)^{s-2\tau_0}\mathbf{Pr}(\boldsymbol{A}_{2\tau_0}(v))
\end{aligned}$$

In particular, if $\delta_v\ge\alpha\log n$ where α is as in **P1**, then

$$\mathbf{Pr}(\boldsymbol{A}_s(v))\le\exp\left(-\frac{sd(v)}{2m}\left(1-\frac{A_2}{\log n}\right)\right).$$

We distinguish unvisited vertices by type according to **P4**. Thus

$$\mathbf{E}\,U_s\le T_0(s)+T_1(s)+T_2(s)+T_3(s) \tag{11}$$

where

$$T_0(s) = 2\sum_{k=1}^{n-1} n\binom{n-1}{k}p^k(1-p)^{n-k-1}e^{-\frac{sk}{2m}\left(1-\frac{A_2}{\log n}\right)},$$

and for $i = 1, 2$

$$T_i(s) = \sum_{k\in K_i} D(k)e^{-\frac{tk}{2m}\left(1-\frac{A}{\log n}\right)},$$

and for $i = 3$,

$$T_3(s) = \sum_{v\in X}\left(1-\left(\frac{\delta_v-1}{\delta_v}-O\left(\frac{1}{\log n}\right)\right)\frac{d(v)}{2m}\right)^{s-2\tau_0} \leq 2\sum_{v\in X}\exp\left\{-\left(\frac{\delta_v-1}{\delta_v}-\frac{A_1}{\log n}\right)\frac{sd(v)}{2m}\right\}.$$

Now for $\gamma > 0$,

$$\sum_{s=t_0+1}^{\infty} e^{-\gamma s} \leq \gamma^{-1}e^{-\gamma t_0}. \quad (12)$$

Let $\gamma = \frac{k}{2m}\left(1-\frac{A_2}{\log n}\right)$ and let $\lambda = t_0\gamma/k$. Using (12) we get

$$\begin{aligned}\sum_{s=t_0+1}^{\infty} T_0(s) \leq\ & 5m\sum_{k=1}^{n-1}\frac{n}{k}\binom{n-1}{k}p^k(1-p)^{n-k-1}e^{-k\lambda}\\ &\leq 10\frac{m}{p}e^{\lambda}\sum_{k=1}^{n-1}\binom{n}{k+1}p^{k+1}(1-p)^{n-k-1}e^{-(k+1)\lambda}\\ &< 11\frac{m}{p}\frac{c}{c-1}\left(1-p+pe^{-\lambda}\right)^n\\ &\leq 11\frac{me^{A_2(c-1)}}{(c-1)\log n}\\ &= o(m). \quad (13)\end{aligned}$$

Continuing we get

$$\sum_{s=t_0+1}^{\infty} T_1(s) \leq A_3m\sum_{k\in K_1}\frac{(\log\log n)^2}{k}e^{-k\lambda} = o(m) \quad (14)$$

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

$$\sum_{s=t_0+1}^{\infty} T_2(s) \leq A_4m\sum_{k\in K_2}\frac{(\log n)^4}{k}e^{-k\lambda} = o(m) \quad (15)$$

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 $e^{\lambda} \geq (1-o(1))(\log n)^{1/3}$.

$$\begin{aligned}\sum_{s=t_0+1}^{\infty} T_3(s) &\leq \sum_{s=t_0+1}^{\infty}\sum_{v\in X}\exp\left\{-\frac{sd(v)}{3m}\right\}\\ &\leq \sum_{v\in X}\frac{3m}{d(v)}\exp\left\{-\frac{t_0d(v)}{3m}\right\}\\ &\leq \sum_{v\in X}\frac{60m}{\log n}\exp\left\{-\frac{t_0\log n}{60m}\right\} \quad \text{by } \mathbf{P2}\\ &\leq \sum_{v\in X}\frac{60m}{\log n}\left(\frac{c-1}{c}\right)^{\log n/30}\\ &= o(m) \quad (16)\end{aligned}$$

since either (i) $c \geq 1+e^{-50}$ and $X = \emptyset$ or (ii) $c < 1+e^{-50}$, in which case we use $(c-1)/c \leq e^{-50}$.
As $C_G = \max_{u\in V} C_u$, the upper bound on C_G now follows from (7), (11), (13), (14), (15), (16) and (10) with $t = t_0$. □

3.2 The lower bound on cover time

For any vertex u, we can find a set of vertices, S, such that at time $t_1 = t_0(1-\epsilon)$, $\epsilon \to 0$, 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

$$\epsilon = \frac{10}{(c-1)\log c/(c-1)}\frac{\log\log n}{\log n} = o(1)$$

and

$$\delta = \frac{(\log n)^3}{|S^*|}.$$

Note that

$$\overline{D}(k^*) = \Omega\left(\frac{n^{(c-1)\ln(c/(c-1))}}{\sqrt{(c-1)\log n}}\right) = \Omega((\log n)^a) \quad (17)$$

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 $A_t(v,w)$ be the event that W has not visited v or w by step t. Let $Q \subseteq S$ be given by

$$\begin{aligned}Q = \{v \in S^* : &\ \mathbf{Pr}(A_{2\tau_0}(v)) < 1-\delta, \text{ or}\\ &\mathbf{Pr}(A_{2\tau_0}(v,w)) < (1-\delta)^2, \text{ for some } w \in S\}. \quad (18)\end{aligned}$$

By time $2\tau_0$, W has visited at most $2\tau_0+1$ vertices, so

$$\sum_{v\in V}\mathbf{Pr}(\overline{A}_{2\tau_0}(v)) \leq 2\tau_0+1$$

and
$$\sum_{v,w\in V}\mathbf{Pr}(\overline{A}_{2\tau_0}(v,w)) \leq (2\tau_0+1)^2.$$

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

Therefore, if $S = S^* \setminus Q$,
$$|S| \geq \frac{\overline{D}(k^*)}{3}.$$

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

$$\mathbf{E}\,|S(t)| \geq \sum_{v\in S}\left(1-\left(1+\frac{A_5}{\log n}\right)\frac{k^*}{2m}\right)^{t-2\tau_0}\mathbf{Pr}(\boldsymbol{A}_{2\tau_0}(v)).$$

Setting $t = t_1$ we have

$$\begin{aligned}\mathbf{E}\,|S(t_1)| &= \Omega\left(\frac{n^{(c-1)\log c/(c-1)}}{\sqrt{(c-1)\log n}}\exp\left(-\frac{k^*}{2m}t_1\right)\right)\\ &= \Omega\left(\frac{n^{\epsilon(c-1)\log c/(c-1)}}{\sqrt{(c-1)\log n}}\right)\\ &= \Omega((\log n)^9).\end{aligned}\tag{19}$$

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 $2m - \zeta(v,w)$ where $\zeta(v,w) = 2d(v) + 2d(w) = 4k^*$. It follows from a minor alteration to the proof of Lemma 3 that for $v, w \in S$

$$\begin{aligned}\mathbf{E}\,(Y_{v,t_1}Y_{w,t_1}) &\leq \left(1-\left(1+\frac{A_6}{\log n}\right)\frac{2k^*}{2m}\right)^{t_1-2\tau_0}\\ &\leq (1+o(1))\mathbf{E}\,Y_{v,t_1}\mathbf{E}\,Y_{w,t_1}.\end{aligned}\tag{20}$$

It follows from (19) and (20) that

$$\mathbf{Pr}(S(t_1)\neq 0) \geq \frac{(\mathbf{E}\,|S(t_1)|)^2}{\mathbf{E}\,|S(t_1)|^2} = \left(\frac{\mathbf{E}|S(t_1)|(|S(t_1)|-1)}{(\mathbf{E}|S(t_1)|)^2}+\frac{1}{\mathbf{E}\,|S_{t_1}|}\right)^{-1} = 1-o(1).$$

□

3.3 Random Regular Graphs

We give a very brief treatment. G_r is a strong expander **whp** and so the mixing time of a walk on $H(v), H(v,w)$ is $O(\log n)$. The expressions in Lemma 3 become

$$\begin{aligned}\mathbf{Pr}(\boldsymbol{A}_t(v)) &=\\ &= \left(1-(1+O(r^{-1}))\frac{1}{n}\right)^{t-2\tau_r}\mathbf{Pr}(\boldsymbol{A}_{\tau_r}(v))\\ &\sim \exp\left(-(1+O(r^{-1}))\frac{t}{n}\right)\mathbf{Pr}(\boldsymbol{A}_{2\tau_r}(v)).\end{aligned}$$

where $\tau_r = O(\log n)$. This gives the upper bound as before and for the lower bound we only need to choose a large set S of vertices, at least distance 2 from each other, for which $\mathbf{Pr}(\boldsymbol{A}_{2\tau_r}(v))$ is small. This is straightforward. □

Acknowledgement. We thank Johan Jonasson for pointing out an 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] C. Cooper and A.M. Frieze, Crawling on web graphs, to appear in Proceedings of STOC 2002.

[5] P. Erdős and A. Rényi, On random graphs I, *Publ. Math. Debrecen* 6 (1959) 290-297.

[6] U. Feige, A tight upper bound for the cover time of random walks on graphs, *Random structures and algorithms* 6 (1995) 51-54.

[7] U. Feige, A tight lower bound for the cover time of random walks on graphs, *Random structures and algorithms* 6 (1995) 433-438.

[8] S. Janson, T. Łuczak and A. Ruciński, *Random graphs*, Wiley (2000).

[9] 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

[10] J. Jonasson, On the cover time of random walks on random graphs, *Combinatorics, Probability and Computing*, 7 (1998), 265-279.

4 Appendix: Typical graph properties

A proof of **P0** can be found in Bollobás [2] or Janson, Łuczak and Ruciński [8].
P1 is a straightforward application of the first and second moment method. The degree sequence of random graphs is studied in detail in [2].
A proof of **P2** can be found in [3].
P3: **Case of** $1 \le s = |S| \le n/(c\log n)$**.**
We first prove that **whp** $e_G(S,S) \le s\log\log n$. Now $\mathbf{E}\; e_G(S,S) = \binom{s}{2}p$ and thus for fixed s,

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

By property **P0** and the definition of H, both G and H contain no isolated vertices and hence $d(S) > 0$. We write $e(S,\overline{S})/d(S) = 1 - 2e(S,S)/d(S)$. 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 $\overline{S}$. By property **P1** the set S_1 induces no edges, and the neighbours of vertices of S_1 are distinct. Thus

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

Case of $n/c\log n \le s \le n/2$**.**
The expected value of $e_H(S,\overline{S})$ is at least $\mu = s(n-s-4)p$. Thus for fixed s,

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

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

$$\begin{aligned}
&\mathbf{Pr}(\exists S : d_H(S) \ge \frac{3}{2}\mathbf{E}\; d_H(S))\\
&\le \binom{n}{s}e^{-\frac{c}{25}s\log n}\\
&\le \exp\left(-s\left(\frac{c}{25}\log n - \log\frac{ne}{s}\right)\right)\\
&= o(n^{-2}).
\end{aligned}$$

Thus

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

P4a: First observe that

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

Then

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

Similarly,

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

A simple calculation for the range of values of p in question, gives

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

Thus

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

Applying the Chebychef inequality we see that

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

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

$$\mathbf{Pr}(\exists k\in K_3 :\; D(k)\le\frac{1}{2}\overline{D}(k)\text{ or }D(k)\ge 2\overline{D}(k)) = = O\left(\frac{c}{\log n}\right).$$

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

$$\frac{\overline{D}(k+1)}{\overline{D}(k)} \sim \frac{c \log n}{k+1} \qquad \text{when } k = O(\log n). \qquad (21)$$

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

$$\overline{D}(k) = An^{1-c} \left(\frac{ce \log n}{k} \right)^k \frac{1}{k^{1/2}}. \qquad (22)$$

Suppose there exists $k \in K_1 \cup K_2$ such that $k < (\log n)^{1/2}$. It follows from (22) that $c - 1 < (\log n)^{-1/3}$. Now suppose that $k \in K_2$ implies $k \geq (\log n)^{1/2}$. Observe from (22) 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 (21) that $|K_2| = O(c^{2/3} \log \log n)$.
P5: This follows immediately from Chernoff bounds.
P6: From (17) 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 $|B^*| \leq |\{(v,w) \in (V^*)^2 : \ dist(v,w) \leq 2\}|$. Therefore

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

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]. $\Box$