The cover time of the preferential attachment graph

Colin Cooper[*] Alan Frieze[†]

May 2, 2006

Abstract

The preferential attachment graph $G_m(n)$ is a random graph formed by adding a new vertex at each time step, with m edges which point to vertices selected at random with probability proportional to their degree. Thus at time n there are n vertices and mn edges. This process yields a graph which has been proposed as a simple model of the world wide web [2]. In this paper we show that if $m \geq 2$ then **whp** the cover time of a simple random walk on $G_m(n)$ is asymptotic to $\frac{2m}{m-1} n \log n$.

1 Introduction

Let $G = (V, E)$ be a connected graph. A *random walk* $\mathcal{W}_u$, $u \in V$ on the undirected graph $G = (V, E)$ is a Markov chain $X_0 = u, X_1, \ldots, X_t, \ldots \in V$ associated to a particle that moves from vertex to vertex according to the following rule: the probability of a transition from vertex i, of degree $d(i)$, to vertex j is $1/d(i)$ if $\{i, j\} \in E$, and 0 otherwise. For $u \in V$ let C_u be the expected time taken for $\mathcal{W}_u$ to visit every vertex of G. The *cover time* C_G of G is defined as $C_G = \max_{u \in V} C_u$. The cover time of connected graphs has been extensively studied. It is a classic result of Aleliunas, Karp, Lipton, Lovász and Rackoff [1] that $C_G \leq 2|E|(|V| - 1)$. It was shown by Feige [8], [9], that for any connected graph G with $|V| = n$,

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

The lower bound is achieved by (for example) the complete graph K_n, whose cover time is determined by the Coupon Collector problem.

[*]Department of Computer Science, King's College, University of London, London WC2R 2LS, UK

[†]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA15213. Supported in part by NSF grant CCR-0200945.

In a previous paper [6] we studied the cover time of random graphs $G_{n,p}$ when $np = c \log n$ where $c = O(1)$ and $(c-1)\log n \to \infty$. This extended a result of Jonasson, who proved in [12] that when the expected average degree $(n-1)p$ grows faster than $\log n$, **whp** a random graph has the same cover time (asymptoticaly) as the complete graph K_n, whereas, when $np = \Omega(\log n)$ this is not the case.

Theorem 1. *[6] Suppose that $np = c \log n = \log n + \omega$ where $\omega = (c-1)\log n \to \infty$ and $c \geq 1$. If $G \in G_{n,p}$, then* **whp**[1]

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

The notation $A_n \sim B_n$ means that $\lim_{n\to\infty} A_n/B_n = 1$.

In another paper [7] we used a different technique to study the cover time of random regular graphs. We proved the following:

Theorem 2. *Let $r \geq 3$ be constant. Let $\mathcal{G}_r$ denote the set of r-regular graphs with vertex set $V = \{1, 2, \ldots, n\}$. If G is chosen randomly from $\mathcal{G}_r$, then* **whp**

$$C_G \sim \frac{r-1}{r-2}\, n \log n.$$

In this paper we turn our attention to the preferential attachment graph $G_m(n)$ introduced by Barabási and Albert [2] as a simplified model of the WWW. The preferential attachment graph $G_m(n)$ is a random graph formed by adding a new vertex at each time step, with m edges which point to vertices selected at random with probability proportional to their degree. Thus at time n there are n vertices and mn edges. We use the generative model of [3] (see also [4]) and build a graph sequentially as follows:

- At each time step t, we add a vertex v_t, and we add an edge from v_t to some vertex u, where u is chosen at random according to the distribution:

$$\mathbf{Pr}(u = v_i) = \begin{cases} \frac{d_{t-1}(v_i)}{2t-1}, & \text{if } v_i \neq v_t; \\ \frac{1}{2t-1}, & \text{if } v_i = v_t; \end{cases} \qquad (1)$$

  where $d_{t-1}(v)$ denotes the degree of vertex v at the end of time step $t-1$.

- For some constant m, every m steps we contract the most recently added m vertices $v_{m(k-1)+1}, \ldots, v_{mk}$ to form a single vertex $k = 1, 2, \ldots$.

Let $G_m(n)$ denote the random graph at time step mn after n contractions of size m. Thus $G_m(n)$ has n vertices and mn edges and may be a multi-graph. It should be noted that without the vertex contractions, we generate $G_1(mn)$.

[1] A sequence of events $\mathcal{E}_n$ occurs *with high probability* **whp** if $\lim_{n\to\infty} \mathbf{Pr}(\mathcal{E}_n) = 1$.

We will assume for the purposes of this paper that $m \geq 2$ is a constant.

This is a very nice clean model, but we warn the reader that it allows loops and multiple edges, although **whp** there will be relatively few of them.

We prove

Theorem 3. *If $m \geq 2$ then* **whp** *the preferential attachment graph $G = G_m(n)$ satisfies*

$$C_G \sim \frac{2m}{m-1}\, n \log n.$$

2 The first visit time lemma.

2.1 Convergence of the random walk

In this section G denotes a fixed connected graph with n vertices. Let u be some arbitrary vertex from which a walk $\mathcal{W}_u$ is started. Let $\mathcal{W}_u(t)$ be the vertex reached at step t, let P be the matrix of transition probabilities of the walk and let $P_u^{(t)}(v) = \mathbf{Pr}(\mathcal{W}_u(t) = v)$. We assume the random walk $\mathcal{W}_u$ on G is ergodic with steady state distribution π and note that $\pi_v = \frac{d(v)}{2mn}$.

2.2 Generating function formulation

Fix two distinct vertices u, v. Let h_t be the probability $\mathbf{Pr}(\mathcal{W}_u(t) = v) = P_u^{(t)}(v)$, that the walk $\mathcal{W}_u$ visits v at step t. Let $H(s)$ generate h_t.

Similarly, considering the walk $\mathcal{W}_v$, starting at v, let r_t be the probability that this walk returns to v at step $t = 0, 1, ...$ Let $R(s)$ generate r_t. We note that $r_0 = 1$.

Let $f_t(u{\to}v)$ be the probability that the first visit of the walk $\mathcal{W}_u$ to v occurs at step t. Thus $f_0(u{\to}v) = 0$. Let $F(s)$ generate $f_t(u{\to}v)$. Thus

$$H(s) = F(s)R(s). \tag{2}$$

Let T be the smallest positive integer such that

$$\max_{x \in V} |P_u^{(t)}(x) - \pi_x| \leq n^{-3} \qquad \text{for } t \geq T. \tag{3}$$

For $R(s)$ let

$$R_T(s) = \sum_{j=0}^{T-1} r_j s^j. \tag{4}$$

Thus $R_T(s)$ generates the probability of a return to v during steps $0, ..., T-1$ of a walk starting at v. Similarly for $H(s)$, let

$$H_T(s) = \sum_{j=0}^{T-1} h_j s^j. \tag{5}$$

2.3 First visit time: Single vertex v

The following lemma should be viewed in the context that G is an n vertex graph which is part of a sequence of graphs with n growing to infinity. We prove it in greater generality than is needed for the proof of Theorem 3.

Let T be as defined in (3) and

$$\lambda = \frac{1}{K_1 T} \tag{6}$$

for sufficiently large constant K_1.

Lemma 4. *Suppose that for some constant* $0 < \theta < 1$,

(a) $H_T(1) < (1-\theta)R_T(1)$.

(b) $\min_{|s|\leq 1+\lambda} |R_T(s)| \geq \theta$.

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

Let

$$p_v = \frac{\pi_v}{R_T(1)(1+O(T\pi_v))}, \tag{7}$$

$$c_{u,v} = 1 - \frac{H_T(1)}{R_T(1)(1+O(T\pi_v))}, \tag{8}$$

where the values of the $1 + O(T\pi_v)$ *terms are given implicitly in (15), (18) respectively. Then*

$$f_t(u \to v) = c_{u,v} \frac{p_v}{(1+p_v)^{t+1}} + O(R_T(1)e^{-\lambda t/2}) \quad \textit{for all } t \geq T. \tag{9}$$

Proof Write

$$R(s) = R_T(s) + \widehat{R}_T(s) + \frac{\pi_v s^T}{1-s}, \tag{10}$$

where $R_T(s)$ is given by (4) and

$$\widehat{R}_T(s) = \sum_{t\geq T} (r_t - \pi_v)s^t$$

generates the error in using the stationary distribution π_v for r_t when $t \geq T$. Similarly, let

$$H(s) = H_T(s) + \widehat{H}_T(s) + \pi_v \frac{s^T}{1-s}. \tag{11}$$

Note that for $Z = H, R$ and $|s| \leq 1 + o(1)$,

$$|\widehat{Z}(s)| = o(n^{-2}). \tag{12}$$

This is because the variation distance between the stationary and the t-step distribution decreases exponentially with t.

Using (10), (11) we rewrite $F(s) = H(s)/R(s)$ from (2) as $F(s) = B(s)/A(s)$ where

$$\begin{aligned} A(s) &= \pi_v s^T + (1-s)(R_T(s) + \widehat{R}_T(s)), &(13)\\ B(s) &= \pi_v s^T + (1-s)(H_T(s) + \widehat{H}_T(s)). &(14) \end{aligned}$$

For real $s \geq 1$ and $Z = H, R$, we have

$$Z_T(1) \leq Z_T(s) \leq Z_T(1)s^T.$$

Let $s = 1 + \beta\pi_v$, where $\beta > 0$ is constant. Since $T\pi_v = o(1)$ we have

$$Z_T(s) = Z_T(1)(1 + O(T\pi_v)).$$

$T\pi_v = o(1)$ and $T\pi_v = \Omega(n^{-2})$ and $R_T(1) \geq 1$ implies that

$$A(s) = \pi_v(1 - \beta R_T(1)(1 + O(T\pi_v))).$$

It follows that $A(s)$ has a real zero at s_0, where

$$s_0 = 1 + \frac{\pi_v}{R_T(1)(1 + O(T\pi_v))} = 1 + p_v, \tag{15}$$

say. We also see that

$$A'(s_0) = -R_T(1)(1 + O(T\pi_v)) \neq 0 \tag{16}$$

and thus s_0 is a simple zero (see e.g. [5] p193). The value of $B(s)$ at s_0 is

$$B(s_0) = \pi_v \left(1 - \tfrac{H_T(1)}{R_T(1)(1+O(T\pi_v))} + O(T\pi_v)\right) \neq 0. \tag{17}$$

Thus, from (7), (8)

$$\frac{B(s_0)}{A'(s_0)} = -p_v c_{u,v}. \tag{18}$$

Thus (see e.g. [5] p195) the principal part of the Laurent expansion of $F(s)$ at s_0 is

$$f(s) = \frac{B(s_0)/A'(s_0)}{s - s_0}. \tag{19}$$

Note that s is a complex variable in the above equation.

To approximate the coefficients of the generating function $F(s)$, we now use a standard technique for the asymptotic expansion of power series (see e.g.[14] Th 5.2.1).

We prove below that $F(s) = f(s) + g(s)$, where $g(s)$ is analytic in $C_\lambda = \{|s| = 1+\lambda\}$ and that $M = \max_{s\in C_\lambda} |g(s)| = O(R_T(1))$.

Let $a_t = [s^t]g(s)$, then (see e.g.[5] p143), $a_t = g^{(t)}(0)/t!$. By the Cauchy Inequality (see e.g. [5] p130) we have that $|g^{(t)}(0)| \le Mt!/(1+\lambda)^t$ and thus

$$|a_t| \le \frac{M}{(1+\lambda)^t} = O(R_T(1)e^{-t\lambda/2}).$$

As $[s^t]F(s) = [s^t]f(s) + [s^t]g(s)$ and $[s^t]1/(s - s_0) = -1/(s_0)^{t+1}$ we have

$$[s^t]F(s) = \frac{-B(s_0)/A'(s_0)}{s_0^{t+1}} + O(R_T(1)e^{-t\lambda/2}). \tag{20}$$

Thus, we obtain

$$[s^t]F(s) = c_{u,v}\frac{p_v}{(1+p_v)^{t+1}} + O(R_T(1)e^{-t\lambda/2}),$$

which completes the proof of (9).

Now $M = \max_{s\in C_\lambda} |g(s)| \le \max|f(s)| + \max|F(s)| = o(1) + \max|F(s)|$. Furthermore, as $F(s) = B(s)/A(s)$ on C_λ we have that

$$|F(s)| \le \frac{H_T(1)(1+\lambda)^T + O(T\pi_v)}{|R_T(s)| - O(T\pi_v)} \le \frac{R_T(1)e^{1/K_1} + o(1)}{\theta - o(1)} = O(R_T(1)).$$

We now prove that s_0 is the only zero of $A(s)$ inside the circle C_λ. We use Rouché's Theorem (see e.g. [5]), the statement of which is as follows: *Let two functions $\phi(z)$ and $\gamma(z)$ be analytic inside and on a simple closed contour C. Suppose that $|\phi(z)| > |\gamma(z)|$ at each point of C, then $\phi(z)$ and $\phi(z) + \gamma(z)$ have the same number of zeroes, counting multiplicities, inside C.*

Let the functions $\phi(s), \gamma(s)$ be given by $\phi(s) = (1-s)R_T(s)$ and $\gamma(s) = \pi_v s^T + (1-s)\widehat{R}_T(s)$.

$$|\gamma(s)|/|\phi(s)| \le \frac{\pi_v(1+\lambda)^T}{\lambda\theta} + \frac{|\widehat{R}_T(s)|}{\theta} = o(1).$$

As $\phi(s) + \gamma(s) = A(s)$ we conclude that $A(s)$ has only one zero inside the circle C_λ. This is the simple zero at s_0. □

Corollary 5. *Let $\boldsymbol{A}_t(v)$ be the event that $\mathcal{W}_u$ has not visited v by step t. Then under the same conditions as those in Lemma 4, for $t \ge T$,*

$$\mathbf{Pr}(\boldsymbol{A}_t(v)) = \frac{c_{u,v}}{(1+p_v)^t} + O(R_T(1)\lambda^{-1}e^{-\lambda t/2}).$$

Proof We use Lemma 4 and

$$\mathbf{Pr}(\boldsymbol{A}_t(v)) = \sum_{\tau>t} f_\tau(u\to v).$$

□

Note that $R_T(1) = O(1)$ in our applications of this corollary. In any case $R_T(1) \le T$.

As we leave this section we introduce the notation R_v, H_v to replace $R_T(1), H_T(1)$ (which are not attached to v).

3 The random graph $G_m(n)$

In this section we prove some properties of $G_m(n)$. We first derive crude bounds on degrees.

Lemma 6. *For $k \le \ell$, let $d_\ell(k)$ denote the degree of vertex k in $G_m(\ell)$. For sufficiently large n, we have:*

(a)

$$\mathbf{Pr}(\exists (k,\ell), 1 \le k \le \ell \le n : \ d_\ell(k) \ge (\ell/k)^{1/2}(\log n)^3) = O(n^{-3}).$$

(b)

$$\mathbf{Pr}(\exists k \le n^{1/8} : \ d_n(k) \le n^{1/4}) = O(n^{-1/17}).$$

Proof

We consider the model $G_1(N)$, where $1 \le N \le mn$. As discussed in [4], in $G_m(\nu)$, $d_\nu(s)$ has the same distribution as $d_N(m(s-1)+1) + \cdots + d_N(ms)$ in $G_1(N)$ when $N = m\nu$.

Let $D_k = d_N(1) + \cdots + d_N(k)$ be the sum of the degrees of the vertices $v_1, ..., v_k$ in the graph $G_1(N)$, where $D_k \ge 2k$. The following is a slight extension of (3) of [4]:
Assume $A \ge 1, k \ge 1$, then

$$\mathbf{Pr}(|D_k - 2\sqrt{kN}| \ge 3A\sqrt{N \log N}) \le N^{-2A}. \tag{21}$$

We also need (4) from the same paper: Assume $0 \le d < N - k - s$, then

$$\begin{aligned}
\mathbf{Pr}(d_N(k+1) = d+1 \mid D_k - 2k = s) &= (s+d)2^d \frac{(N-k-s)_d}{(2N-2k-s)_{d+1}} && (22)\\
&= \frac{s+d}{2N-2k-s-d} \prod_{i=0}^{d-1} \left(1 - \frac{s+i}{2N-2k-s-i}\right) && (23)\\
&\le \exp\left\{-\frac{d(s+(d-1)/2)}{2N}\right\}. && (24)
\end{aligned}$$

(a): Let $N = \ell m$, and $k \leq N$. We first consider the case $1 \leq k \leq 100(\log n)^3$. In order to consider the degree of vertex 1, we additionally allow $k = 0$ and $\mathbf{Pr}(D_0 = 0) = 1$.
Let $\lambda = 100(N/(k+1))^{1/2}(\log n)^2$ then

$$\begin{aligned}\mathbf{Pr}(d_N(k+1) \geq \lambda) &\leq \sum_{\substack{0 \leq s \leq N-k \\ d+1 \geq \lambda}} \mathbf{Pr}(d_N(k+1) = d+1 \mid D_k - 2k = s) \\ &\leq N^2 \exp\left\{-\frac{2400(\log n)^4}{k+1}\right\} \leq n^{-20}, \end{aligned} \tag{25}$$

after using (24).

For fixed ℓ, and $N = m\ell$, define $k_0 = k_0(N) = N/\log N$. Assume $100(\log n)^3 < k \leq k_0$. We use (21) with $A = 3\log_\ell n$ to argue that

$$\mathbf{Pr}\left(D_k \leq 2\sqrt{kN} - 9\log_\ell n\sqrt{N\log N}\right) \leq n^{-6}. \tag{26}$$

Now

$$\frac{kN}{81(\log_\ell n)^2 N\log N} \geq \frac{100(\log n)^3(\log \ell)^2}{81(\log n)^2(\log \ell + \log m)} > \log n$$

and

$$N \geq k\log N \geq k\log k \geq k\log\log n$$

and so (26) implies

$$\mathbf{Pr}\left(D_k - 2k \leq 3\sqrt{kN}/2\right) \leq n^{-6},$$

and thus $s > 3\sqrt{kN}/2$ **whp**. Arguing as in (25) we deduce that

$$\mathbf{Pr}(d_N(k+1) \geq 10\sqrt{N/k}(\log n)^2) \leq n^{-6} + N^2 \exp\left\{-\frac{(10\sqrt{N/k}(\log n)^2)(3\sqrt{kN}/2)}{2N}\right\} \leq 2n^{-6}. \tag{27}$$

When $k_0 < k \leq N$, let $N' = 2N\log N$ and $n' = \max\{n, N'/m\}$, and now assume that $N'/(2(\log N')^2 \leq k < N < k_0(N')$. We use (26), (27) evaluated at N' together with the fact that $d_N(k)$ is stochastically dominated by $d_{2N\log N}(k)$ to obtain

$$\mathbf{Pr}(d_N(k+1) \geq 10\sqrt{N'/k}(\log n')^2) = O(n'^{-6}).$$

Using the relationship between $G_m(n)$ and $G_1(N)$, part (a) now follows.

(b): Here $N = nm$, and $k \leq mN^{1/8}$. Using (21) with $A = 2$ we have

$$\mathbf{Pr}(D_k - 2k \geq 8\sqrt{kN\log N}) \leq N^{-4}.$$

We then use (23) to write

$$\mathbf{Pr}(d_N(k) \leq N^{1/4}) \leq N^{-4} + \sum_{d=0}^{N^{1/4}} \frac{d + 8\sqrt{kN\log N}}{2N - 2k - 8\sqrt{kN\log N} - d} = O\left(\frac{\sqrt{k}\log n}{n^{1/4}}\right).$$

Summing the RHS of the above inequality over $k \leq mN^{1/8}$ accounts for the possible values of k and completes the proof of the lemma. □

Let
$$\omega = (\log n)^{1/3}. \tag{28}$$
Let a cycle C be *small* if $|C| \leq 2\omega + 1$. Let a vertex v be *locally-tree-like* if the sub-graph G_v induced by the vertices at distance 2ω or less is a tree. Thus a locally-tree-like vertex is at distance at least 2ω from any small cycle.

Lemma 7. Whp *$G_m(n)$ does not contain a set of vertices S such that (i) $|S| \leq 100\omega$, (ii) the sub-graph H induced by S has minimum degree at least 2 and (iii) H contains a vertex $v \geq n^{1/10}$ of degree at least 3 in H.*

Proof Let Z_1 denote the number of sets S described in Lemma 7, and let $s = |S|$. Then

$$\begin{aligned}
\mathbf{E}\,(Z_1) &\leq o(1) + \sum_{3\leq s\leq 100\omega} \sum_{H} \prod_{(v,w)\in E(H)} \frac{(\log n)^3}{(vw)^{1/2}} \qquad (29)\\
&\leq o(1) + \sum_{3\leq s\leq 100\omega} \sum_{H} (\log n)^{3|E(H)|} \prod_{v\in S} v^{-d_H(v)/2}\\
&\leq o(1) + \sum_{3\leq s\leq 100\omega} (1+(\log n)^3)^{\binom{s}{2}} n^{-1/20} \prod_{v\in S} \frac{1}{v}\\
&\leq o(1) + \sum_{3\leq s\leq 100\omega} (1+(\log n)^3)^{\binom{s}{2}} n^{-1/20} H_n^s\\
&\leq o(1) + 100\omega (\log n)^{20000(\log n)^{2/3}} n^{-1/20}\\
&= o(1).
\end{aligned}$$

where $H_n = \sum_{v=1}^n \frac{1}{v}$.

Explanation of (29)**:** Suppose that $1 \leq \alpha < \beta \leq n$. Then

$$\mathbf{Pr}(G_m(n) \text{ contains edge } (\alpha,\beta)) \mid d_\beta(\alpha) \leq (\beta/\alpha)^{1/2}(\log n)^3) \leq \frac{(\log n)^3}{(\alpha\beta)^{1/2}}. \tag{30}$$

This is because when β chooses its neighbours, the probability it chooses α is at most $\frac{m(\log n)^3(\beta/\alpha)^{1/2}}{2m(\beta-1)}$. Here the numerator is a bound on the degree of α in $G_m(\beta-1)$. We are using Lemma 6 here and the $o(1)$ term accounts for the failure of this bound. Furthermore, this remains an upper bound if we condition on the existence of some of the other edges of H. □

This lemma is used to justify the following corollary: A small cycle is *light* if it contains no vertex $v \leq n^{1/10}$ (it has no "heavy" vertices), otherwise it is *heavy*.

Corollary 8. Whp *$G_m(n)$ does not contain a small cycle within 10ω of a light cycle.*

□

We need to deal with the possibility that $G_m(n)$ contains many cycles.

Lemma 9. **Whp** *$G_m(n)$ contains at most $(\log n)^{10\omega}$ vertices or edges on small cycles.*

Proof Let Z be the number of vertices/edges on small cycles in $G_m(n)$ (including parallel edges). Then

$$\begin{aligned}\mathbf{E}(Z) &\leq o(1)+\sum_{k=2}^{2\omega+1} k\sum_{a_1,\ldots,a_k}\prod_{i=1}^{k}\frac{(\log n)^3}{(a_ia_{i+1})^{1/2}} \\ &\leq o(1)+\sum_{k=2}^{2\omega+1} k(\log n)^{3k}H_n^k \\ &= O((\log n)^{9\omega})\end{aligned} \tag{31}$$

and the result follows from the Markov inequality.

Explanation of (31)**:** We sum over the choices $a_1,a_2,\ldots,a_k$ for the vertices of the cycle. The term $(\log n)^3/(a_ia_{i+1})^{1/2}$ bounds the probability of edge (a_i,a_{i+1}) and comes from the RHS of (30). The $o(1)$ term accounts for the probability it is. □

We estimate the number of non-locally-tree-like vertices.

Lemma 10. **Whp** *there are at most $O(n^{1/2+o(1)})$ non-locally-tree-like vertices.*

Proof A non-locally-tree-like vertex v is within ω of a small cycle. So the expectation of the number Z of such vertices satisfies

$$\begin{aligned}\mathbf{E}(Z) &\leq o(1)+\sum_{\substack{0\leq r\leq\omega\\ 3\leq s\leq 2\omega+1\\ 1\leq i\leq s}}\sum_{\substack{a_0,\ldots,a_r\\ b_1\ldots,b_s}}\frac{(\log n)^3}{(a_0b_1)^{1/2}}\prod_{k=1}^{r-1}\frac{(\log n)^3}{(a_ka_{k+1})^{1/2}}\prod_{l=1}^{s}\frac{(\log n)^3}{(b_lb_{l+1})^{1/2}} \\ &= O(n^{1/2+o(1)}).\end{aligned}$$

The result follows from the Markov inequality.

Here $a_0,a_1,\ldots,a_r$ are the choices for the vertices of a path from v to a small cycle. The path ends at b_1 and the cycle is through $b_1,b_2,\ldots,b_s$. □

Lemma 11. **Whp** *there are at most $n(\log n)^{-\omega}$ vertices $v\geq n/2$ which have more than $(\log n)^{11\omega}$ vertices at distance 3ω or less from them.*

Proof For a fixed vertex v, the expected number of paths of length $\leq 3\omega$ and endpoint v is bounded by

$$\sum_{1\leq r\leq 3\omega}\sum_{a_1,\ldots,a_r}\frac{(\log n)^3}{a_r^{1/2}v^{1/2}}\prod_{k=1}^{r-1}\frac{(\log n)^3}{(a_ka_{k+1})^{1/2}}\leq(\log n)^{10\omega}.$$

The result now follows from the Markov inequality. □

Let

$$\omega_0 = \log\log\log n. \tag{32}$$

We say that v is *locally regular* if it is locally tree-like and the first $2\omega_0$ levels of G_v form a tree of depth $2\omega_0$, rooted at v, in which every non-leaf has branching factor m.

For $j \in [n]$ we let $X(j)$ denote the set of neighbours of j in $[j-1]$ i.e. the vertices "chosen" by j, although not including j; recall that loops are allowed in the scale-free construction. We regard X as a function from $[n]$ to the power set of $[n]$ and so X^{-1} is well defined. The constraint that $X^{-1}(i) = \{j\}$, means j is the only vertex $v > i$ that chooses i.

Lemma 12. Whp, *$G_m(n)$ contains at least $n^{1-o(1)}$ locally regular vertices $v \geq n/2$.*

Proof Let $I_k = \left[n\left(1-\frac{1}{2^k}\right), n\left(1-\frac{1}{2^{k+1}}\right)\right)$ for $1 \leq k \leq \omega_0$. Let

$$J_2 = \{j \in I_2 : \; X(j) \subseteq I_1, \; |X(j)| = m \text{ and } X^{-1}(i) = \{j\} \; \text{ for } i \in X(j)\}.$$

We require $|X(j)| = m$ so that there are no parallel edges originating from j.

Then for $2 < k \leq \omega_0$ we let

$$J_k = \{j \in I_k : \; X(j) \subseteq J_{k-1}\,, |X(j)| = m \text{ and } X^{-1}(i) = \{j\} \; \text{ for } i \in X(j)\}.$$

For $j \in I_2$, define $i_{m+1} = j-1$, then

$$\mathbf{Pr}(j \in J_2) =$$

$$\sum_{\{i_1<\cdots<i_m\}\subseteq I_1} \prod_{k=1}^{m} \prod_{\tau=mi_k+1}^{mi_{k+1}} \left(1-\frac{km}{2\tau-1}\right) \prod_{\tau=mj+1}^{mn} \left(1-\frac{m^2}{2\tau-1}\right) \cdot m! \prod_{i=1}^{m} \frac{m}{2mj+2i-1} \tag{33}$$

$$\sim \sum_{\{i_1<\cdots<i_m\}\subseteq I_1} \left(\prod_{k=1}^{m} \frac{i_k}{j}\right)^{m/2} \cdot \frac{j^{m^2/2}}{n^{m^2/2}} \cdot \frac{m!}{(2j)^m}$$

$$\sim \frac{m!}{(2j)^m n^{m^2/2}} \sum_{\{i_1<\cdots<i_m\}\subseteq I_1} \prod_{k=1}^{m} i_k^{m/2}$$

$$\geq (1-O(1/n)) \frac{1}{(2j)^m n^{m^2/2}} \left(\sum_{i \in I_1} i^{m/2}\right)^m \tag{34}$$

$$\geq \frac{|I_1|^m}{2^{m+m^2/2} n^m}. \tag{35}$$

Explanation of (33)-(35)**:** We sum over the choices $i_1 < i_2 < \cdots < i_m$ for $X(j)$. The double product followed by the single product is the probability that the vertices in set $i_1, i_2, \ldots, i_m$ are chosen by j and j alone. The term $m!$ counts the order in which j chooses these vertices and the final product gives the probability that these choices are made.

To see the derivation of (34) we note that for $b_j \geq 0$

$$(b_1 + \cdots + b_t)^m - ({b_1}^2 + \cdots + {b_t}^2)\binom{m}{2}(b_1 + \cdots + b_t)^{m-2} \leq m! \sum_{i_1 < \cdots < i_m} \prod_{k=1}^{m} b_{i_k}.$$

The line (35) follows by putting $i = n/2$ and $j = n$.

So

$$\mathbf{E}\,(|J_2|) \geq \frac{n}{2^{3m+m^2/2}}.$$

We use a martingale argument to prove that $|J_2|$ is concentrated around its mean.

We work in $G_1(mn)$. Let $Y_1, Y_2, \ldots, Y_{mn}$ denote the sequence of choices of edges added. When vertex i chooses its neighbour, it does so according to the model (1), and thus selects one of the existing $2i - 1$ edge-endpoints uar.

Fix $Y_1, Y_2, \ldots, Y_i$, let $Y_i = (i, v)$ and let $\hat{Y}_i = (i, \hat{v})$ denote an alternative choice of edge-endpoint $\hat{v}$ at step i. For each complete outcome $\mathbf{Y} = Y_1, Y_2, \ldots, Y_{i-1}, Y_i, \ldots, Y_{mn}$ we define a corresponding outcome $\hat{\mathbf{Y}} = Y_1, Y_2, \ldots, Y_{i-1}, \hat{Y}_i, \ldots, \hat{Y}_{mn}$. Let $S(i) = \{i\}$. For $j > i$, $\hat{Y}_j$ is obtained from Y_j as follows: If Y_j creates a new edge (j, v) by choosing one of the $|S(j-1)|$ edge-endpoints at v arising from edges with labels in $S(j-1)$, ie. edges generated directly or indirectly from the edge-endpoint of Y_i, then $\hat{Y}_j$ chooses the corresponding edge-endpoint $\hat{v}$ to create edge $(j, \hat{v})$. If this occurs then $S(j) = S(j-1) \cup \{j\}$. In all other cases $\hat{Y}_j = Y_j$ and $S(j) = S(j-1)$.

We consider the martingale $Z_0, Z_1, \ldots, Z_{mn}$ where

$$Z_t = \mathbf{E}\,(|J_2| \mid Y_1, Y_2, \ldots, Y_t) - \mathbf{E}\,(|J_2| \mid Y_1, Y_2, \ldots, Y_{t-1}).$$

The map $\mathbf{Y} \to \hat{\mathbf{Y}}$ is measure preserving. In going from $\mathbf{Y}$ to $\hat{\mathbf{Y}}$, $|J_2|$, changes by at most 2, according to the in-degree of the vertices $v, \hat{v}$.

The Azuma-Hoeffding martingale inequality then implies that

$$\mathbf{Pr}(||J_2| - \mathbf{E}\,(|J_2|)| \geq u) \leq \exp\left\{-\frac{u^2}{2mn}\right\}. \tag{36}$$

It follows that **qs**[2]

$$||J_2| - \mathbf{E}\,(|J_2|)| \leq n^{1/2} \log n. \tag{37}$$

Thus **qs** we have

$$|J_2| \geq \frac{n}{2^{3m+1+m^2/2}} = A_2 n,$$

which defines the constant A_2.

[2]A sequence of events $\mathcal{E}_n$ occurs *quite surely* (**qs**) if $\mathbf{Pr}(\mathcal{E}_n) = 1 - O(n^{-K})$ for any constant $K > 0$.

Repeating the argument given for $\mathbf{Pr}(j \in J_2)$, we see that for $j \in I_3$

$$\begin{aligned}
&\mathbf{Pr}(j \in J_3 \mid J_2) = \\
&\sum_{\{i_1 < \cdots < i_m\} \subseteq J_2} \prod_{k=1}^{m} \prod_{\tau = mi_k+1}^{mi_{k+1}} \left(1 - \frac{km}{2\tau - 1}\right) \prod_{\tau=mj+1}^{mn} \left(1 - \frac{m^2}{2\tau - 1}\right) \cdot m! \prod_{i=1}^{m} \frac{m}{2mj + 2i - 1} \\
&\sim \frac{1}{(2j)^m n^{m^2/2}} \left(\sum_{i \in J_2} i^{m/2}\right)^m \\
&\geq \frac{|J_2|^m}{2^{m+m^2/2} n^m}.
\end{aligned}$$

Thus,

$$\mathbf{E}\left(|J_3| \mid J_2\right) \geq \frac{|J_2|^m}{2^{m+m^2/2} n^m} |I_3|$$

and given J_2, **qs** $|J_3|$ will be concentrated around its mean to within $n^{1/2} \log n$.

Proceeding in this way we find that for $2 \leq k \leq \omega_0$ we have **qs**

$$|J_k| \geq A_k n$$

where for $k \geq 2$,

$$A_{k+1} = \frac{A_k^m}{2^{m+k+3+m^2/2}},$$

and (inductively) $A_k \geq 2^{-10km^k}$. It follows that $|J_k| \geq 2^{-10km^k} n$ and that $|J_{\omega_0}| = n^{1-o(1)}$.

By construction, any locally tree-like vertex of J_{ω_0} is locally regular. The lemma follows from the bound on the number of non locally tree-like vertices in Lemma 10. □

3.1 Mixing time

The *conductance* Φ of the walk $\mathcal{W}_u$ is defined by

$$\Phi = \min_{\pi(S) \leq 1/2} \frac{e(S : \overline{S})}{d(S)}.$$

Mihail, Papadimitriou and Saberi [13] proved that the *conductance* Φ of the walks $\mathcal{W}$ are bounded below by some absolute constant. Now it follows from Jerrum and Sinclair [10] that

$$|P_u^{(t)}(x) - \pi_x| \leq (\pi_x/\pi_u)^{1/2}(1 - \Phi^2/2)^t. \tag{38}$$

For sufficiently large t, 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 halves the conductance but we still have

$$T = O(\log n) \tag{39}$$

in (3). The cover time is doubled. Asymptotically the values R_v are doubled too. Otherwise, it has a negligible effect on the analysis and we will ignore this for the rest of the paper and continue as though there are no lazy steps.

Notice that Lemma 6 implies $\pi_v = O((\log n)^2 n^{-1/2})$ and so together with (39) we see that

$$T\pi_v = o(1) \text{ and } T\pi_v = \Omega(n^{-2}) \tag{40}$$

for all $v \in V$, as required by Lemma 4.

4 Cover time of $G_m(n)$

4.1 Parameters

Recall that the values of ω, ω_0 are given by (28), (32) respectively.

Assume now that $G_m(n)$ has the following properties: (i) there are $n^{1-o(1)}$ locally regular vertices, (ii) $d(s) \geq n^{1/4}$ for $s \leq n^{1/10}$, (iii) no small cycle is within distance 10ω of a light cycle, (iv) there are at most $(\log n)^{10\omega}$ vertices on small cycles and (v) there are at most $n(\log n)^{-\omega}$ vertices $v \geq n/2$ which have more than $(\log n)^{11\omega}$ vertices at distance 3ω or less from them.

Consider first a locally regular vertex v. It was shown in [7] (Lemma 6) that $R_v = \frac{r-1}{r-2} + o(\omega^{-1})$ for a locally-tree-like vertex w of an r-regular graph. We obtain the same result for v by putting $r = m+1$. Note that the degree of v is irrelevant here. It is the branching factor of the rest of the tree G_v that matters.

Lemma 13. *Suppose that v is locally-tree-like. Then*

(a) $R_v \leq \frac{d(v)}{m-1} + o(1)$.

(b) $d(v) \geq m+1$ *implies* $R_v \leq \frac{d(v)(m+m^{-1}-1)}{d(v)(m+m^{-1}-2)-m^{-1}+1} + o(1)$

(c) *If v is locally regular then* $R_v = \frac{m}{m-1} + o(1)$.

Proof We first define an infinite tree T_v^* by taking the tree T_v' defined by the first $\omega+1$ levels of G_v and then rooting a copy of the infinite tree T_m^∞ which has branching factor m from each leaf of T_v'. This construction is modified in the case that v is locally regular. We now let T_v' be made up from the first ω_0 levels. Thus if v is locally regular, T_v^* is an infinite tree with branching factor m, rooted at v.

Let R_v^* be the expected number of visits to v for an infinite random walk $\mathcal{W}_v^*$ on T_v^*, started at v. We argue first that

$$|R_v - R_v^*| = o(1). \tag{41}$$

Let $r_t^* = \mathbf{Pr}(\mathcal{W}_v^*(t) = v)$. Then

$$\begin{aligned} |R_v - R_v^*| &\leq \sum_{t=\omega+1}^{T} r_t + \sum_{t=\omega+1}^{\infty} r_t^* \\ &\leq o(1) + \sum_{t=\omega+1}^{\infty} e^{-\alpha t} \qquad \text{for some constant } \alpha > 0 \qquad (42) \\ &= o(1). \end{aligned}$$

(When v is locally regular, the sums are from $\omega_0 + 1$.)

Explanation of (42): We prove that $\sum_{t=\omega+1}^{T} r_t = o(1)$ via (38); replace r_t by $\pi_v + O(\zeta^t)$ for some constant $\zeta < 1$. For the second sum we project the walk $\mathcal{W}_v^*$ onto $\{0, 1, 2, \ldots, \}$ by letting $\mathcal{X}(t)$ be the distance of $\mathcal{W}_v^*(t)$ from v. The degree of every vertex in T_v^* is at least m and if a vertex has degree exactly m then its immediate descendants have degree at least $m+1$ and so we see that for any positive $\lambda < 1/2$ and $t \geq 0$ we have

$$\begin{aligned} \mathbf{E}\left(e^{-\lambda(\mathcal{X}(2t+2)-\mathcal{X}(2t))} \mid \mathcal{X}(2t)\right) &\leq \frac{m-1}{m+1}e^{-2\lambda} + \frac{2m-1}{m(m+1)} + \frac{1}{m(m+1)}e^{2\lambda} \qquad (43) \\ &\leq \frac{1}{3}e^{-2\lambda} + \frac{1}{2} + \frac{1}{6}e^{2\lambda} \\ &\leq \frac{1}{3}(1 - 2\lambda + 4\lambda^2) + \frac{1}{2} + \frac{1}{6}(1 + 2\lambda + 4\lambda^2) \\ &\leq e^{-\lambda(1-6\lambda)/3}. \qquad (44) \end{aligned}$$

We take $\lambda = 1/12$ and $\alpha = \lambda(1 - 6\lambda)/3 = 1/72$.

Explanation of (43) If $\mathcal{W}_v^*(t) = w$ and the degree of w is m then all of w's neighbours in T_v^* have degree at least $m+1$. The expression on the RHS of (43) gives the exact expectation if either (i) the degree of w is m and all its neighbours have degree $m+1$ or (ii) the degree of w is $m+1$ and all neighbours have degree m. This situation minimizes the expectation, since the higher the degree the more likely it is that $\mathcal{X}$ increases.

It follows from (44) that

$$\begin{aligned}
\mathbf{E}\left(e^{-\lambda\mathcal{X}(2t)}\right) &= \mathbf{E}\left(\prod_{\tau=0}^{t-1} e^{-\lambda(\mathcal{X}(2\tau+2)-\mathcal{X}(2\tau))}\right) \\
&= \mathbf{E}\left(\mathbf{E}\left(e^{-\lambda(\mathcal{X}(2t)-\mathcal{X}(2t-2))} \mid \mathcal{X}(2t-2)\right)\prod_{\tau=0}^{t-2} e^{-\lambda(\mathcal{X}(2\tau+2)-\mathcal{X}(2\tau))}\right) \\
&\le e^{-\alpha}\mathbf{E}\left(\prod_{\tau=0}^{t-2} e^{-\lambda(\mathcal{X}(2\tau+2)-\mathcal{X}(2\tau))}\right) \\
&\le e^{-\alpha t}.
\end{aligned}$$

Thus

$$r_{2t}^* = \mathbf{Pr}(\mathcal{X}(2t)=0) = \mathbf{Pr}(e^{-\mathcal{X}(2t)} \ge 1) \le \mathbf{E}\left(e^{-\mathcal{X}(2t)}\right) \le e^{-\alpha t}$$

and (42) follows.

Let b_w, $w \in T_v^*$ be the branching factor at w i.e. $b_v = d_v$ and $b_w = d_w - 1$ if w is not the root.

Let $\widehat{T}_w$ be the sub-tree of T_v^* rooted at vertex w. (Thus $\widehat{T}_v = T_v^*$). Let ρ_w denote the probability that a random walk on $\widehat{T}_w$ which starts at w ever returns to w. Our aim is to estimate ρ_v and use

$$R_v^* = \frac{1}{1-\rho_v}. \tag{45}$$

Let $C(w)$ denote the children of w in T_v^*. We use the following recurrence: The parameter k counts the number of returns to x, for $x \in C(w)$.

$$\begin{aligned}
\rho_w &= 1 - \frac{1}{b_w}\sum_{x\in C(w)}\sum_{k\ge 0}\left(1-\frac{1}{d_x}\right)\left(\rho_x\left(1-\frac{1}{d_x}\right)\right)^k(1-\rho_x) && (46) \\
&= 1 - \frac{1}{b_w}\sum_{x\in C(w)}\frac{\left(1-\frac{1}{d_x}\right)(1-\rho_x)}{1-\rho_x\left(1-\frac{1}{d_x}\right)} \\
&= 1 - \frac{1}{b_w}\sum_{x\in C(w)}\frac{b_x - b_x\rho_x}{b_x+1-\rho_x b_x} \\
&= \frac{1}{b_w}\sum_{x\in C(w)}\frac{1}{b_x+1-\rho_x b_x}. && (47)
\end{aligned}$$

Explanation of (46)**:** For each $x \in C(w)$, $1/b_w$ gives the probability that the walk moves to x in the first step. The term $1 - 1/d_x$ is the probability that the first step from x is away from w. Then the term $\rho_x(1 - 1/d_x)$ is the probability that the walk returns to x and does not visit w in its first move from x. We sum over the number of times, k, that this happens. The final factor $1 - \rho_x$ is the probability of no return for the $k + 1$th time.

We see immediately that if T_v^* is a regular tree with branching factor $m \geq 2$ then, with $\rho_w = \rho$ for all w,

$$\rho = \frac{1}{m+1-\rho m} \text{ and hence } \rho = \frac{1}{m}$$

and this deals with the locally regular case. (The solution $\rho = 1$, which implies $R_v^* = \infty$ is ruled out by (42) which implies $R_v^* < \infty$).

If w is in the first ω levels let $b_w = b_w^+ + b_w^-$ where b_w^+ is the number of children w' of w in T_v with $w > w'$ i.e. w chose w' in the construction of $G_m(n)$. If w is at a higher level, we take $b_w = b_w^+ = m$ and $b_w^- = 0$.

We will now prove the following by induction on $\omega + 1 - \ell_w$, where $\ell_w \leq \omega + 1$ is the level of w in the tree.:

(a) $b_w = m - 1$ implies $\rho_w \leq \frac{1}{m}$.

(b) $b_w^+ = m, b_w^- \geq 1$ implies $\rho_w \leq \frac{1}{b_w}\left(1 + \frac{b_w - m}{m+m^{-1}-1}\right)$.

(c) $b_w = b_w^+ = m$ implies $\rho_w \leq \frac{1}{m}$.

(d) $b_w^+ = m-1, b_w^- \geq 1$ implies $\rho_w \leq \frac{1}{b_w}\left(\frac{m-1}{m} + \frac{b_w^-}{m+m^{-1}-1}\right)$

The base case will be $\ell_w = \omega + 1$. For which, Case (c) applies and the induction hypothesis holds from the locally regular case.

The lemma follows from this since only cases (b),(c) can apply to the root v, in which case $b_v = d(v)$.

Let us now go through the inductive step. Let us assume these conditions apply to $x \in C(w)$. Then case by case, the following inequalities will hold:

(a) $b_x + 1 - b_x\rho_x \geq m + \frac{1}{m} - 1$.

(b) $b_x + 1 - b_x\rho_x \geq m + (b_x - m)\left(1 - \frac{1}{m+m^{-1}-1}\right) \geq m$.

(c) $b_x + 1 - b_x\rho_x \geq m$.

(d) $b_x + 1 - b_x\rho_x \geq m + \frac{1}{m} - 1 + b_x^-\left(1 - \frac{1}{m+m^{-1}-1}\right) \geq m + \frac{1}{m} - 1$.

Case (a): In this case $b_w = b_w^+$ and only cases (b),(c) are possible for $x \in C(w)$. In which case $b_x + 1 - b_x\rho_x \geq m$ for $x \in C(w)$ and then (47) implies that $\rho_w \leq 1/m$.

Case (b): In $C(w)$ we have $b_w^+ = m$ cases of (b) or (c) and b_w^- cases of (a) or (d). In the first case we have $b_x + 1 - b_x\rho_x \geq m$. In the second case we have $b_x + 1 - b_x\rho_x \geq m + m^{-1} - 1$. Thus

$$\rho_w \leq \frac{1}{b_w}\left(1 + \frac{b_w - m}{m+m^{-1}-1}\right).$$

Case (c): This follows as in Case (a).

Case (d): In $C(w)$ we have $m-1$ cases of (b) or (c) and b_w^- cases of (a) or (d). Thus

$$\rho_w \leq \frac{1}{b_w}\left(\frac{m-1}{m} + \frac{b_w^-}{m+m^{-1}-1}\right)$$

as is to be shown. □

We deal with non-locally-tree like vertices in a somewhat piece-meal fashion: We remind the reader that if G_v is not tree-like, then it consists of a breadth-first tree T_v of depth ω plus extra edges E_v. Each $e \in E_v$ lies in a small cycle σ_e. If one of these cycles is light, then G_v must be a tree plus a single extra edge, see Corollary 8. Otherwise, all the cycles σ_e are heavy. G_v may of course contain other cycles, but these will play no part in the proof.

Lemma 14. *Suppose that either*
(i) G_v contains a unique light cycle C_v, that $v \notin C_v$ and that the shortest path $P = (w_0 = v, w_1, \dots, w_k)$ from v to C_v is such that $\max\{d(w_1), \dots, d(w_k)\} \geq \omega^3$*, or*
(ii) the small cycles of G_v are all heavy cycles. Then

(a) $R_v \leq \frac{d(v)}{m-1} + o(1))$.

(b) $d(v) \geq m+1$ *implies* $R_v \leq \frac{d(v)(m+m^{-1}-1)}{d(v)(m+m^{-1}-2)+m^{-1}-1} + o(1)$

Proof
(a) Let w be the first vertex on the path from v to C_v which has degree at least ω^3. Let G_v' be obtained from G_v by deleting those vertices, other than w, all of whose paths to v in G_v go through w.(By assumption there are one or two paths). Let R_v' be the expected number of returns to v in a random walk of length ω on G_v' where w is an absorbing state. We claim that

$$R_v \leq R_v' + O(\omega^{-2}). \tag{48}$$

Once we verify this, the proof of (a) follows from the proof of Lemma 13 i.e. embed the tree $H'v$ in an infinite tree by rooting a copy of T_m^∞ at each leaf. To verify (48) we couple random walks on G_v, G_v' until w is visited. In the latter the process stops. In the former, we find that when at w, the probability we get closer to v in the next step is at most ω^{-3} and so the expected number of returns from now on is at most $\omega \times \omega^{-3}$ and (48) follows.

(b) Now consider the case where the small cycles of G_v are all heavy. We argue first that a random walk of length ω that starts at v might as well terminate if it reaches a vertex $w \leq n^{1/10}$, $w \neq v$. By the assumptions made at the start of Section 4.1 we can assume $d(w) \geq n^{1/4}$. Now we can assume from Lemma 9 at least $n_0 = n^{1/4} - (\log n)^{10\omega}$ of the T_v edges incident with w are not in any cycle σ_e contained in G_v. But then if a walk arrives at w, it has a more than $\frac{n_0}{n^{1/4}}$ chance of entering a sub-tree T_w of G_v rooted at w for which every vertex is separated from v by w. But then the probability of leaving T_w in ω steps is

$O(\omega(\log n)^{10\omega}/n^{1/4})$ and so once a walk has reached w, the expected number of further returns to v is $o(\omega^{-1})$. We can therefore remove T_w from G_v and then replace an edge (x,w) by an edge (x,w_x) and make all the vertices w_x absorbing. Repeating this argument, we are left with a tree to which we can apply the argument of Lemma 13. $\square$

Note that if $v \in V_B$ then no bound on R_v has been established:

$$V_B = \{v:\ G_v \text{ contains a unique light cycle } C_v \text{ and the path from } v \text{ to } C_v \text{ contains no vertex of degree at least } \omega^3\}$$

However, for these it suffices to prove

Lemma 15. *If $v \in V_B$ then $R_v \leq 2\omega$.*

Proof We write, for some constant $\zeta < 1$,

$$\begin{aligned} R_v &= \sum_{t=1}^{\omega} r_t + \sum_{t=\omega+1}^{T} (\pi_v + O(\zeta^t)) \\ &\leq \omega + o(\omega) \end{aligned}$$

and the lemma follows. $\square$

We remind the reader that in the following lemma, λ is defined in (6) and $R_T(s)$ is defined in (4).

Lemma 16. *There exists a constant $0 < \theta < 1$ such that if $v \in V$ then $|R_T(s)| \geq \theta$ for $|s| \leq 1 + \lambda$.*

Proof Assume first that v is locally tree-like. We write

$$\begin{aligned} R_T(s) &= A(s) + Q(s) \\ &= \frac{1}{1-B(s)} + Q(s). \end{aligned} \tag{49}$$

Here $A(s) = \sum a_t s^t$ where $a_t = r_t^*$ is the probability that the random walk $\mathcal{W}_v^*$ is at v at time t (see Lemma 13 for the definition of $\mathcal{W}_v^*$). $B(s) = \sum b_t s^t$ where b_t is the probability of a first return at time t. Then $Q(s) = Q_1(s) + Q_2(s)$ where

$$\begin{aligned} Q_1(s) &= \sum_{t=\omega+1}^{T} (r_t - a_t) s^t \\ Q_2(s) &= -\sum_{t=T+1}^{\infty} a_t s^t. \end{aligned}$$

Here we have used the fact that $a_t = r_t$ for $0 \le t \le \omega$.

We now justify equation (49). For this we need to show that

$$|B(s)| < 1 \qquad \text{for } |s| \le 1 + \lambda. \tag{50}$$

We note first that, in the notation of Lemma 13, $B(1) = \rho_v < 1$. Then observe that $b_t \le a_t \le e^{-\alpha t}$. The latter inequality is proved in Lemma 13, see (42). Thus the radius of convergence ρ_B of $B(s)$ is at least e^{α}, $B(s)$ is continuous for $0 \le |s| < \rho_B$, $|B(s)| \le B(|s|)$ and $B(1) < 1$. Thus there exists a constant $\epsilon > 0$ such that $B(s) < 1$ for $|s| \le 1 + \epsilon$. We can assume that $\lambda < \epsilon$ and (50) follows. We will use

$$|R_T(s)| \ge \frac{1}{1 + B(|s|)} - |Q(s)| \ge \frac{1}{1 + B(1 + \lambda)} - |Q(s)| \ge \frac{1}{2} - |Q(s)|.$$

The lemma for locally tree-like vertices will follow once we show that $|Q(s)| = o(1)$. But, using (38),

$$\begin{aligned} |Q_1(s)| &\le (1+\lambda)^T \sum_{t=\omega+1}^{T} (\pi_v + e^{-\Phi^2 t/2} + e^{-\alpha t}) = o(1) \\ |Q_2(s)| &\le \sum_{t=T+1}^{\infty} (e^{-\alpha}(1+\lambda))^t = o(1). \end{aligned}$$

For non tree-like vertices we proceed more or less as in Lemma 14. If $v \notin V_B$ then we truncate G_v at vertices of degree more than $n^{1/4}$, add copies of T_m at leaves and then proceed as above.

If $v \in V_B$ let T_v^* be the graph obtained by adding T_m^{∞} to all the leaves of G_v. Thus T_v^* contains a unique cycle $C = (x_1, x_2, \ldots, x_k, x_1)$. We can write an expression equivalent to (49) and then argument rests on showing that $B(1) < 1$ and $a_s \le \zeta^s$ for some $\zeta < 1$. The latter condition can be relaxed to $a_s \le e^{o(s)}\zeta^s$, allowing us to take less care with small s.

$\boldsymbol{B(1) < 1}$: If $m \ge 3$ there is a $\ge 1 - \frac{2}{m}$ probability of the first move of $\mathcal{W}_v^*$ going into an infinite tree rooted at a neighbour of v and then the probability of return to v is bounded below by a positive constant. The same argument is valid for $m = 2$ when $v \notin C$. So assume that $v \in C$ and that T_v^* consists of C plus a tree T_i attached to x_i for $i = 1, 2, \ldots, k$. Here T_i is empty (if degree of x_i is 2) or infinite. Furthermore, T_i empty, implies that T_{i-1}, T_{i+1} are both infinite. Thus the walk $\mathcal{W}_v^*$ has a constant positive probability of moving into an infinite tree within 2 steps and then never returning to v.

$\boldsymbol{a_s \le e^{o(s)}\zeta^s}$: If C is an even cycle then we can couple the distance X_t of $W_v^*(t)$ to v with a random walk on $\{0, 1, 2, \ldots, \}$ as we did in Lemma 13. If C is an odd cycle let w_1, w_2 be the vertices of C which are furthest from v in T_v^*. If $W_v^*(t) \ne w_1, w_2$ then $\mathbf{E}\,(X_{t+2} - X_t) \ge 1/6$ and otherwise $\mathbf{E}\,(X_{t+2} - X_t) \ge 0$. Thus $\mathbf{E}\,(X_{t+4} - X_t) \ge 1/6$ always and we can use Hoeffding's theorem. □

Lemma 17. *If $v \in V$ and its degree $d_n(v) \le (\log n)^2$ then $H_v < CR_v + o(1)$ for some constant $C < 1$.*

Proof As in Section 2.1 let f_t be the probability that $\mathcal{W}_u$ has a first visit to v at time t. As $H(s) = F(s)R(s)$ we have

$$\begin{aligned} H_v &\leq \mathbf{Pr}(\mathcal{W}_u \text{ visits } v \text{ by time } T-1)R_v \\ &= R_v \sum_{t=1}^{T} f_t. \end{aligned}$$

We now estimate $\sum_{t=1}^{T} f_t$, the probability that $\mathcal{W}_u$ visits v by time T. We first observe that (38) implies

$$\sum_{t=\omega+1}^{T} f_t \leq \sum_{t=\omega+1}^{T} (((\log n)^2/m)^{1/2} e^{-\Phi^2 t/2} + \pi_v) = o(1).$$

Thus it suffices to bound $\sum_{t=1}^{\omega} f_t$, the probability that $\mathcal{W}_u$ visits v by time ω.

Let $v_1, v_2, \ldots, v_k$ be the neighbours of v and let w be the first neighbour of v visited by $\mathcal{W}_u$. Then

$$\begin{aligned} \mathbf{Pr}(\mathcal{W}_u \text{ visits } v \text{ by time } \omega) &= \sum_{i=1}^{k} \mathbf{Pr}(\mathcal{W}_u \text{ visits } v \text{ by time } \omega \mid w = v_i)\mathbf{Pr}(w = v_i) \\ &\leq \sum_{i=1}^{k} \mathbf{Pr}(\mathcal{W}_{v_i} \text{ visits } v \text{ by time } \omega)\mathbf{Pr}(w = v_i). \end{aligned}$$

So it suffices to prove the lemma when u is a neighbour of v.

Let the neighbours of u be $u_1, u_2, \ldots, u_d$, $d \geq m$ and $v = u_d$. If u is locally tree-like than we can write

$$\mathbf{Pr}(\mathcal{W}_u \text{ does not visit } v \text{ by time } \omega) \geq \rho \frac{d-1}{d} - o(1) > 0. \tag{51}$$

Here ρ is a lower bound on the probability of not returning to u in ω steps, given that $\mathcal{W}_u(1) \neq v$. We have seen in the previous lemma that this is at least some positive constant.

If $u \notin V_B$ then we truncate H_u as we did in Lemma 14 and argue for (51).

If $u \in V_B$ and there exist neighbours $u_1, \ldots, u_k$ say, which are not on the unique cycle C of H_u then there is a probability k/d that $\mathcal{W}_u^*(1) = u_i$ for some $i \leq k$ and then the probability that $\mathcal{W}_u$ does not return to u_i in ω steps is bounded below by a constant. The final case is where $m = 2$, $d_n(u) = 2$ and u, u_1, v are part of the unique cycle of H_u. But then with probability $1/2$ $\mathcal{W}_u(1) = u_1$ and then with conditional probability at least $1/3$ $x = \mathcal{W}_u(2)$ is not on C and then the probability that $\mathcal{W}_u$ does not return to x in ω steps is bounded below by a constant. □

4.2 Upper bound on cover time

Let $t_0 = \lceil \frac{2m}{m-1} n \log n \rceil$. We prove that **whp**, for $G_m(n)$, for any vertex $u \in V$, $C_u \leq t_0 + o(t_0)$.

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

$$C_u = \mathbf{E}\left(T_G(u)\right) = \sum_{t>0} \mathbf{Pr}(T_G(u) \geq t), \tag{52}$$

$$\mathbf{Pr}(T_G(u) \geq t) = \mathbf{Pr}(T_G(u) > t-1) = \mathbf{Pr}(U_{t-1} > 0) \leq \min\{1, \mathbf{E}\, U_{t-1}\}. \tag{53}$$

It follows from (52), (53) that for all t

$$C_u \leq t + 1 + \sum_{s \geq t} \mathbf{E}\left(U_s\right) = t + 1 + \sum_{v \in V} \sum_{s \geq t} \mathbf{Pr}(\boldsymbol{A}_s(v)) \tag{54}$$

where $\boldsymbol{A}_s(v)$ is defined in Corollary 5.

For vertices v satisfying Corollary 5 we see that

$$\sum_{s \geq t} \mathbf{Pr}(\boldsymbol{A}_s(v)) \leq (1 + O(T\pi_v)) \frac{R_v}{\pi_v} e^{-(1+O(T\pi_v))t\pi_v/R_v} + O(\lambda^{-2} e^{-\lambda t/2}). \tag{55}$$

The second term arises from the sum of the error terms $O(\lambda^{-1} e^{-\lambda s/2})$ for $s \geq t$.

Recall that V_B is the set of vertices v such that G_v contains a unique light cycle C_v and the path from v to C_v contains no vertex of degree at least ω^3.

We write $V = V_1 \cup V_2 \cup V_3$ where $V_1 = (V \setminus V_B) \cap \{d_n(v) \leq (\log n)^2\}$, $V_2 = \{d_n(v) \geq (\log n)^2\}$ and $V_3 = V_B \cap \{d_n(v) \leq (\log n)^2\}$.

Let $t_1 = (1+\epsilon)t_0$ where $\epsilon = n^{-1/3}$ can be assumed by Lemma 6 to satisfy $T\pi_v = o(\epsilon)$ for all $v \in V - V_2$.

If $v \notin V_B$ then by Lemmas 13(a) and 14(a),

$$t_1(1 + O(T\pi_v))\pi_v/R_v \geq \frac{2m}{m-1} n \log n \cdot \frac{d(v)}{2mn} \cdot \frac{m-1}{d(v)} = \log n. \tag{56}$$

Plugging (56) into (54) and using $R_v \leq 5$ (Lemmas 13 and 14) and $\pi_v \geq \frac{1}{2n}$ for all $v \in V \setminus V_B$ we get

$$\sum_{v \in V_1} \sum_{s \geq t_1} \mathbf{Pr}(\boldsymbol{A}_s(v)) \leq 10n. \tag{57}$$

Suppose now that $v \in V_2$ ie. $d_n(v) \geq (\log n)^2$. After a walk of length T there is an $\Omega((\log n)^2/n)$ chance of being at v. Thus for some constant $c > 0$ and $s \geq t_1$, we have

$$\mathbf{Pr}(\boldsymbol{A}_s(v)) \leq \left(1 - \frac{c(\log n)^2}{n}\right)^{\lfloor s/T \rfloor} \leq \exp\left\{-\frac{cs(\log n)^2}{2Tn}\right\}.$$

Thus

$$\begin{aligned}\sum_{v\in V_2}\sum_{s\geq t_1}\mathbf{Pr}(\boldsymbol{A}_s(v)) &\leq n\sum_{s\geq t_1}\exp\left\{-\frac{cs(\log n)^2}{2Tn}\right\}\\ &\leq \frac{3Tn^2}{c(\log n)^2}\exp\left\{-\frac{ct_1(\log n)^2}{2Tn}\right\}=o(1). \qquad (58)\end{aligned}$$

It remains to deal with $v\in V_3$. We first observe that

$$|V_B|\leq(\log n)^{10\omega}\omega^{3\omega}\leq(\omega\log n)^{10\omega}$$

and from Lemma 15 and (55) we have

$$\begin{aligned}\sum_{v\in V_3}\sum_{s\geq t_1}\mathbf{Pr}(\boldsymbol{A}_s(v)) &\leq (\omega\log n)^{10\omega}\left(2n\omega e^{-(1+o(1))t_1\pi_v/(2\omega)}+O(\lambda^{-2}e^{-\lambda t_1/2})\right)\\ &= o(n). \qquad (59)\end{aligned}$$

Thus combining (57) with (58) and (59) gives

$$C_u\leq t_1+O(n)=t_0+o(t_0),$$

completing our proof of the upper bound on cover time.

4.3 Lower bound on cover time

For some 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 $\mathcal{W}_u$ tends to zero. Hence $T_G(u)>t_1$ **whp** which implies that $C_G\geq t_0-o(t_0)$.

We construct S as follows. Let S be some maximal set of locally regular vertices such that the distance between any two elements of S is least$2\omega+1$. Thus $|S|\geq ne^{-e^{O(\omega_0)}}(\log n)^{-11\omega}\geq n(\log n)^{-12\omega}$.

Let $S(t)$ denote the subset of S which has not been visited by $\mathcal{W}_u$ after step t. Now, by Corollary 5, provided $t\geq T$

$$\mathbf{E}\left(|S(t)|\right)\geq(1-o(1))\sum_{v\in S}\left(\frac{c_{u,v}}{(1+p_v)^t}+o(n^{-2})\right).$$

Let u be a fixed vertex of S. Let $v\in S$ and let $H_T(1)$ be given by (5), then (38) implies that

$$H_T(1)\leq\sum_{t=\omega}^{T-1}(\pi_v+e^{-\Phi^2t/2})=o(1). \qquad (60)$$

$R_v \geq 1$ and so $c_{uv} = 1 - o(1)$. Setting $t = t_1 = (1-\epsilon)t_0$ where $\epsilon = 2\omega^{-1}$, we have

$$\begin{aligned}
&\mathbf{E}\,(|S(t_1)|) \geq (1+o(1))|S|e^{-(1-\epsilon)t_0 p_v} \\
&= (1+o(1)) \exp\left\{\log n - 12\omega \log\log n - (1+o(1))(1-\epsilon)\frac{2m}{m-1} n \log n \cdot \frac{m}{2mn} \cdot \frac{m-1}{m}\right\} \\
&\geq n^{1/\omega}. \qquad (61)
\end{aligned}$$

Let $Y_{v,t}$ be the indicator for the event $\boldsymbol{A}_t(v)$. Let $Z = \{v, w\} \subset S$. We will show (below) that that for $v, w \in S$

$$\mathbf{E}\,(Y_{v,t_1} Y_{w,t_1}) = \frac{c_{u,Z}}{(1+p_Z)^{t+2}} + o(n^{-2}), \qquad (62)$$

where $c_{u,Z} \sim 1$ and $p_Z \sim (m-1)/(mn) \sim p_v + p_w$. Thus

$$\mathbf{E}\,(Y_{v,t_1} Y_{w,t_1}) = (1+o(1))\mathbf{E}\,(Y_{v,t_1})\mathbf{E}\,(Y_{w,t_1})$$

which implies

$$\mathbf{E}\,(|S(t_1)|(|S(t_1)|-1)) \sim \mathbf{E}\,(|S(t_1)|)(\mathbf{E}\,(|S(t_1)|)-1). \qquad (63)$$

It follows from (61) and (63), that

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

Proof of (62). Let Γ be obtained from G by merging v, w into a single node Z. This node has degree $2m$.

There is a natural measure preserving mapping from the set of walks in G which start at u and do not visit v or w, to the corresponding set of walks in Γ which do not visit Z. Thus the probability that $\mathcal{W}_u$ does not visit v or w in the first t steps is equal to the probability that a random walk $\widehat{\mathcal{W}}_u$ in Γ which also starts at u does not visit Z in the first t steps.

We apply Lemma 4 to Γ. That $\pi_Z = \frac{1}{n}$ is clear, and $c_{u,Z} = 1 - o(1)$ is argued as in (60). The vertex Z is tree-like up to distance ω in Γ. The derivation of R_Z in Lemma 13(c) is valid. The fact that the root vertex of the corresponding infinite tree has degree $2m$ does not affect the calculation of R_Z^*. □

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] A. Barabási and R. Albert, Emergence of scaling in random networks, *Science* 286 (1999) 509-512.

[3] B. Bollobás and O. Riordan, The diameter of a scale-free random graph, *Combinatorica* 24 (2004) 5-34.

[4] B. Bollobás and O. Riordan and J. Spencer and G. Tusnády, The degree sequence of a scale-free random graph process, *Random Structures and Algorithms* 18 2001 279-290.

[5] J. Brown and R. Churchill, *Complex Variables and Applications,* (Sixth Edition) McGraw-Hill (1996).

[6] C. Cooper and A. M. Frieze, The cover time of sparse random graphs, *Proceedings of the 14th ACM-SIAM Symposium on Discrete Algorithms* (2003) 140-147.

[7] C. Cooper and A. M. Frieze, The cover time of random regular graphs, *SIAM Journal on Discrete Mathematics* 18 (2005) 728-740.

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

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

[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] W. Hoeffding, Inequalities for sums of bounded random variables, *Journal of the American Statistical Association* 27, (1963) 13-30.

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

[13] M. Mihail, C. Papadimitriou, and A. Saberi , On Certain Connectivity Properties of the Internet Topology, *Proceedings of the 44th Symposium on Foundations of Computer Science* (2003) 28-35.

[14] H. Wilf, *Generatingfunctionology,* Academic Press (1990).