The cover time of the giant component of a random graph

Colin Cooper[*] Alan Frieze[†]

July 23, 2008

Abstract

We study the cover time of a random walk on the largest component of the random graph $G_{n,p}$. We determine its value up to a factor $1 + o(1)$ whenever $np = c > 1$, $c = O(\ln n)$. In particular we show that the cover time is not monotone for $c = \Theta(\ln n)$. We also determine the cover time of the k-cores, $k \geq 2$.

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 *vertex cover time* C_G of G is defined as $C_G = \max_{v \in V} C_v$. The (vertex) cover time of connected graphs has been extensively studied. It is a classic result of Aleliunas, Karp, Lipton, Lovász and Rackoff [2] that $C_G \leq 2m(n-1)$. It was shown by Feige [13], [14], that for any connected graph G, the cover time satisfies $(1 - o(1))n \ln n \leq C_G \leq (1 + o(1))\frac{4}{27}n^3$. As an example of a graph achieving the lower bound, the complete graph K_n has cover time determined by the Coupon Collector problem. The *lollipop* graph consisting of a path of length $n/3$ joined to a clique of size $2n/3$ gives the asymptotic upper bound for the cover time.

We say that a sequence of events $\mathcal{E}_n$ occurs *with high probability*, **whp**, if $\lim_{n\to\infty} \mathbf{Pr}(\mathcal{E}_n) = 1$.

In an earlier paper [7] we studied the cover time of the random graph $G_{n,p}$ when $np = d \ln n$ where d is (asymptotic to some fixed) constant and $(d-1)\ln n \to \infty$. This sharpened a result

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

[†]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA15213, USA, Research supported in part by NSF Grant ccf0502793

of Jonasson, who proved in [18] that when the expected average degree $(n-1)p$ grows faster than $\ln n$, **whp**a random graph has asymptotically the same cover time as the complete graph K_n, whereas, when $np = \Omega(\ln n)$ this is not the case. The paper [7] established the following result: (The notation $A_n \sim B_n$ means that $\lim_{n\to\infty} A_n/B_n = 1$)

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

$$C_G \sim \left(d \ln\left(\frac{d}{d-1}\right)\right) n \ln n.$$

The function $f(d) = d\ln(d/(d-1))$ is monotone decreasing for $d > 1$, tending to 1 as $d \to \infty$ in agreement with [18]. As $d \to 1^+$, $f(d)$ is unbounded and the cover time is no longer of order $n \ln n$. As an example,

$$np = \ln n + e^{o(\ln\ln n)} \text{ implies } C_G \sim (n\ln n)\ln\ln n. \tag{1}$$

The threshold for connectivity in $G_{n,p}$ is at $np = \ln n$ and thus Theorem 1 gives the cover time once the graph is connected **whp**. At $np = \ln n$, the number of isolated vertices is approximately Poisson distributed with parameter 1. Below $np = \ln n$ the graph is disconnected **whp**. In this paper we discuss the cover time of the largest component of $G_{n,p}$. In particular we let $c = np$ and assume that that $c-1$ is at least a positive constant and that $c \leq \ln n + o(\ln n)$. Thus our assumption implies that **whp** the largest component is a giant i.e. of linear size.

We have avoided the *phase transition* where $c = 1 + o(1)$. Determining the asymptotic cover time in this range is an interesting open question.

Let $\boldsymbol{C}_1(G) = (V_1(G), E_1(G))$ denote the largest component of $G \in G_{n,p}$. When $np = c > 1$, c constant, $\boldsymbol{C}_1(G)$ is **whp** a unique component with $\sim nx$ vertices [12], where x is the solution in $(0,1)$ of $x = 1 - e^{-cx}$.

Theorem 2. *Let $np = c$, and let x denote the solution in $(0,1)$ of $x = 1 - e^{-cx}$. If C_{C_1} denotes the cover time of the giant component of $G_{n,p}$. then* **whp**

(a) *If $1 + \Omega(1) \leq c \leq \ln n/\omega$, where $\omega \to \infty$. then*

$$C_{C_1} \sim t_1^* = \frac{cx(2-x)}{4(cx - \ln c)} n(\ln n)^2.$$

(b) *Suppose that $c = \beta \ln n$ where $\beta = \alpha + \delta$, $0 < \alpha < 1$ is a constant and $\delta \to 0$. Then*

$$C_{C_1} \sim t_1^* = \gamma n (\ln n)^2$$

where

$$\gamma = \gamma(\alpha) = \max\left\{\alpha\ell(1-\alpha\ell) : \ \ell \text{ is a positive integer}\right\}. \tag{2}$$

(c) *Suppose that* $c = (1 + \delta)\ln n$ *where* $\delta = o(1)$. *Let* $\delta^+ = \max\{-\delta, 0\}$, *then*

$$C_{C_1} \sim t_1^* = n \ln n(\ln\ln n + \delta^+ \ln n).$$

Note that $x \to 1$ as $c \to \infty$ and so the expression for t_1^* in (a) tends to $\frac{1}{4}n(\ln n)^2$. If $\alpha \to 0$ then $\gamma(\alpha) \to 1/4$ and so (a),(b) are consistent. When $\alpha \in (1/3, 1)$ the cover time is asymptotically $\sim \alpha(1-\alpha)n(\ln n)^2$. Thus choosing $\alpha = 1 - \delta^+$, we see that (b),(c) are consistent for suitable δ^+. Finally, the formula in (c) is consistent with the value given in (1).

In case (b), let ℓ_γ be the value(s) of ℓ attaining the maximum γ. For $k \geq 1$, when α lies in the interval $[1/(2k+1), 1/(2k-1)]$ then $\ell_\gamma = k$. For $k \geq 2$, when $\alpha = 1/(2k-1)$, both k and $k-1$ are solutions to ℓ_γ. The function $\alpha k(1-\alpha k)$ has a maximum at $\alpha = 1/2k$, and thus the cover time is not monotone in the interval $[1/(2k+1), 1/(2k-1)]$. Curiously, $\gamma(1/2k) = 1/4$ for all $k = \ell_\gamma = 1, 2, ...$, and thus the (asymptotic) cover time given in (a) for $c \to \infty$ re-occurs at every $\alpha = 1/2k$. Thus the formula in case (b) has curious properties and can be attributed to the need to cover all vertices whose distance from the 2-core of $G_{n,p}$ is $\ell = 1, 2, \ldots,$. There is a trade-off between the number of such vertices and the distance ℓ. Given α, the expression $\alpha\ell(1-\alpha\ell)n(\ln n)^2$ is the time required to cover all such vertices. As α increases, the number of each type of vertex decreases, but the expected time to cover each of them increases and so the cover time does not always decrease monotonically.

The edge cover time of a graph G is defined similarly to the vertex cover time. It is the expected time to cover all edges. It is certainly bounded below by the vertex cover time; and the star is an example of a graph which achieves the $n \ln n$ lower bound for edge cover time. The upper bound is not so clear, but Zuckerman [22] shows that the $2mn$ bound of [2] is sufficient.

It will be seen in the proofs that for $c = o(\ln n)$ or $c \sim \alpha \ln n$, $(\alpha < 1)$ the vertices of degree 1 are the last to be covered **whp**; whereas when $c \sim \ln n$ vertices of degree 1 vie with other vertices of low degree to be last to be covered. Up to some (unknown) value of $c \leq \ln n$, **whp** all edges of the 2-core are covered before the last vertex of degree 1. In this case the time to cover all vertices of degree 1 precisely determines the edge cover time of the giant. In general we have

Theorem 3. *Under the same conditions as those of Theorem 2,* **whp** *the edge cover time of the giant component of* $G_{n,p}$ *is asymptotically equal to the vertex cover time.*

We also establish the cover time of the 2-core, and any non-empty k-cores, $k \geq 3$. For $k \geq 2$, the k-core is defined as the (possibly empty) sub-graph of $\boldsymbol{C}_1$ obtained by recursively removing any vertices of degree at most $k-1$. Thus the k-core is the largest sub-graph of $\boldsymbol{C}_1(G)$ with minimum degree k.

Theorem 4. *Let* C_{C_2} *denote the cover time of the 2-core of* $G_{n,p}$, *then*

(a) *If* $1 < c \leq \ln n/\omega$*, where* $\omega \to \infty$*, then*

$$C_{C_2} \sim t_2^* = \frac{cx^2}{16(cx - \ln c)} n(\ln n)^2.$$

(b) *Suppose that* $c = \beta \ln n$ *where* $\beta = \alpha + \delta$, $0 < \alpha < 1$ *is a constant and* $\delta \to 0$*. Then*

$$C_{C_2} \sim t_2^* = \gamma n(\ln n)^2,$$

where

$$\gamma = \max\left(\frac{\alpha \lfloor \ell/2 \rfloor \lceil \ell/2 \rceil}{\ell}(1 - \alpha(\ell - 1)) : \ \ell \text{ is a positive integer, } 2 \leq \ell \leq (\lfloor \alpha^{-1} \rfloor + 1) \right). \tag{3}$$

(c) *Suppose that* $c = (1 + \delta) \ln n$ *where* $\delta = o(1)$*. Let* $\delta^+ = \max\{-\delta, 0\}$*, then*

$$C_{C_2} \sim t_2^* = n \ln n(\ln \ln n + \frac{\delta^+}{2} \ln n).$$

Whereas the results for the 2-core have marked similarities with those for the giant component, the results for the κ-cores, $\kappa \geq 3$, are rather different.

Theorem 5.

$$f_k(c, \xi) = 1 - e^{-c\xi} \left(1 + c\xi + \cdots + \frac{(c\xi)^{k-2}}{(k-2)!} \right).$$

For $k \geq 3$ *let* c_k *be the smallest* c *for which* $\xi = f_k(c, \xi)$ *has a solution* $x_k > 0$*, and for* $c > c_k$*, let* x_k *be the largest solution in* $(0, 1)$ *of* $\xi = f_k(c, \xi)$*. Then for* $c > c_k$*,* c *constant and* $np = c$*,* **whp**

$$C_{C_k} \sim t_k^* = \left(\frac{k-1}{k(k-2)} c x_k^2 \right) n \ln n.$$

1.1 Structure of the paper

The structure of the paper is as follows. In Section 2 we prove a general lemma, the *first visit time lemma*, on which our results are based. We use generating functions to give a good estimate of the probability that a random walk starting at a vertex u, has not visited vertex v after t steps. Given that this walk is rapidly mixing, the probability of no visit to v depends mainly on the local structure at v. In particular, it depends on R_v, the expected number of returns that the random walk $\mathcal{W}_v$ makes to v within a *short* amount of time. Most of the technical work in the paper involves estimating upper and lower bounds for R_v for various types of vertex. Armed with information about the values R_v it is relatively easy to estimate the expected number of unvisited vertices at time t and we can obtain an upper bound for the cover time directly. A lower bound is obtained by applying the Chebyshev inequality to the number of unvisited vertices at a given time.

Section 2 discusses the first visit time lemma. Section 3 establishes the cover time of the giant component for all the ranges of Theorem 2 and also proves Theorem 3. We feel that it is better to do all cases of Theorem 2 together, rather than ask the reader to re-do the proof three times, albeit with small twists to the argument. In Section 4 we give outline proofs of Theorems 4, 5.

2 Estimating first visit probabilities

2.1 Convergence of the random walk

In this section G denotes a fixed connected graph with n vertices and m edges. A random walk $\mathcal{W}_u$ is started from a vertex u. 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 that the random walk $\mathcal{W}_u$ on G is ergodic i.e. G is not bipartite. Thus, the random walk $\mathcal{W}_u$ has the steady state distribution π, where $\pi_v = d(v)/(2m)$. Here $d(v)$ is the degree of vertex v.

2.2 Generating function formulation

We use the approach of [8], [9]. We have found some simplifications in the arguments given there and so we will give a detailed proof.

Let $d(t) = \max_{u,x\in V} |P_u^{(t)}(x) - \pi_x|$, and let T be such that, for $t \geq T$

$$\max_{u,x\in V} |P_u^{(t)}(x) - \pi_x| \leq n^{-3}. \tag{4}$$

It follows from e.g. Aldous and Fill [1] that $d(s+t) \leq 2d(s)d(t)$ and so for $k \geq 1$,

$$\max_{u,x\in V} |P_u^{(kT)}(x) - \pi_x| \leq \frac{2^{k-1}}{n^{3k}}. \tag{5}$$

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

$$H(z) = \sum_{t=T}^{\infty} h_t z^t \tag{6}$$

generate h_t for $t \geq T$. This changes the definition of $H(z)$ from that used in [8], [9] where we included the coefficients $h_0, h_1, \ldots, h_{T-1}$ in the definition of $H(z)$ which gave rise to technical problems.

Next, considering the walk $\mathcal{W}_v$, starting at v, let $r_t = \mathbf{Pr}(\mathcal{W}_v(t) = v)$ be the probability that this walk returns to v at step $t = 0, 1, ...$ Let

$$R(z) = \sum_{t=0}^{\infty} r_t z^t$$

generate r_t. Our definition of return involves $r_0 = 1$.

For $t \geq T$ let $f_t = f_t(u \to v)$ be the probability that the first visit of the walk $\mathcal{W}_u$ to v in the period $[T, T+1, \ldots]$ occurs at step t. Let

$$F(z) = \sum_{t=T}^{\infty} f_t z^t$$

generate f_t. Then we have

$$H(z) = F(z)R(z). \tag{7}$$

Finally, for $R(z)$ let

$$R_T(z) = \sum_{j=0}^{T-1} r_j z^j. \tag{8}$$

2.3 First visit time lemma: Single vertex v

Let

$$\lambda = \frac{1}{KT} \tag{9}$$

for some sufficiently large constant K.

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. An almost identical lemma was first proved in [8].

Lemma 6. *Suppose that*

(a) *For some constant $\theta > 0$, we have*

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

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

There exists

$$p_v = \frac{\pi_v}{R_T(1)(1+O(T\pi_v))}, \tag{10}$$

where $R_T(1)$ *is from (8), such that for all* $t \geq T$*,*

$$f_t(u{\rightarrow}v) = (1+O(T\pi_v))\frac{p_v}{(1+p_v)^{t+1}} + o(e^{-\lambda t/2}). \tag{11}$$

Proof Write

$$R(z) = R_T(z) + \widehat{R}_T(z) + \frac{\pi_v z^T}{1-z}, \tag{12}$$

where $R_T(z)$ is given by (8) and

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

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

$$H(z) = \widehat{H}_T(z) + \frac{\pi_v z^T}{1-z}. \tag{13}$$

Equation (5) implies that the radii of convergence of both $\widehat{R}_T$ and $\widehat{H}_T$ exceed $1+2\lambda$. Moreover, for $Z = H, R$ and $|z| \leq 1+\lambda$,

$$|\widehat{Z}(z)| = o(n^{-2}). \tag{14}$$

Using (12), (13) we rewrite $F(z) = H(z)/R(z)$ from (7) as $F(z) = B(z)/A(z)$ where

$$A(z) = \pi_v z^T + (1-z)(R_T(z) + \widehat{R}_T(z)), \tag{15}$$

$$B(z) = \pi_v z^T + (1-z)\widehat{H}_T(z). \tag{16}$$

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

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

Let $z = 1 + \beta\pi_v$, where $\beta = O(1)$. Since $T\pi_v = o(1)$ we have

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

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

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

It follows that $A(z)$ has a real zero at z_0, where

$$z_0 = 1 + \frac{\pi_v}{R_T(1)(1+O(T\pi_v))} = 1 + p_v, \tag{17}$$

say. We also see that

$$A'(z_0) = -R_T(1)(1 + O(T\pi_v)) \neq 0 \tag{18}$$

and thus z_0 is a simple zero (see e.g. [6] p193). The value of $B(z)$ at z_0 is

$$B(z_0) = \pi_v \left(1 + O(T\pi_v)\right) \neq 0. \tag{19}$$

Thus,

$$\frac{B(z_0)}{A'(z_0)} = -(1 + O(T\pi_v))p_v. \tag{20}$$

Thus (see e.g. [6] p195) the principal part of the Laurent expansion of $F(z)$ at z_0 is

$$f(z) = \frac{B(z_0)/A'(z_0)}{z - z_0}. \tag{21}$$

To approximate the coefficients of the generating function $F(z)$, we now use a standard technique for the asymptotic expansion of power series (see e.g.[21] Th 5.2.1).

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

Let $a_t = [z^t]g(z)$, then (see e.g.[6] p143), $a_t = g^{(t)}(0)/t!$. By the Cauchy Inequality (see e.g. [6] p130) we see that $|g^{(t)}(0)| \leq Mt!/(1+\lambda)^t$ and thus

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

As $[z^t]F(z) = [z^t]f(z) + [z^t]g(z)$ and $[z^t]1/(z - z_0) = -1/z_0^{t+1}$ we have

$$[z^t]F(z) = \frac{-B(z_0)/A'(z_0)}{z_0^{t+1}} + O(T\pi_v e^{-t\lambda/2}). \tag{22}$$

Thus, we obtain

$$[z^t]F(z) = (1 + O(T\pi_v))\frac{p_v}{(1+p_v)^{t+1}} + O(T\pi_v e^{-t\lambda/2}),$$

which completes the proof of (11).

Now $M = \max_{z \in C_\lambda} |g(z)| \leq \max |f(z)| + \max |F(z)| = O(T\pi_v) + \max |F(z)|$, where $F(z) = B(z)/A(z)$. On C_λ we have, using (14)-(16),

$$|F(z)| \leq \frac{O(\pi_v)}{\lambda |R_T(z)| - O(T\pi_v)} = O(T\pi_v).$$

We now prove that z_0 is the only zero of $A(z)$ inside the circle C_λ and this implies that $F(z) - f(z)$ is analytic inside C_λ. We use Rouché's Theorem (see e.g. [6]), 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 $\phi(z), \gamma(z)$ be given by $\phi(z) = (1-z)R_T(z)$ and $\gamma(z) = \pi_v z^T + (1-z)\widehat{R}_T(z)$.

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

As $\phi(z)+\gamma(z) = A(z)$ we conclude that $A(z)$ has only one zero inside the circle C_λ. This is the simple zero at z_0. □

Corollary 7. *For $t \geq T$ let $\boldsymbol{A}_t(v)$ be the event that $\mathcal{W}_u$ does not visit v in steps $T, T+1, \ldots, t$. Then, under the assumptions of Lemma 6,*

$$\mathbf{Pr}(\boldsymbol{A}_t(v)) = \frac{(1+O(T\pi_v))}{(1+p_v)^t} + o(e^{-\lambda t/2}).$$

Proof We use Lemma 6 and

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

and note that $T^2\pi_v = o(1)$. □

For the rest of the paper u, v will not be fixed and so it is appropriate to replace the notation $R_T(1)$ by something dependent on v. We use R_v.

3 Cover time of the giant component

In this section we prove the three parts of Theorem 2.

3.1 Typical graphs in $G_{n,p}$

The giant component $\boldsymbol{C}_1$ of G consists of a 2-core $\boldsymbol{C}_2$ and a *mantle* $\boldsymbol{M}$ of edges $E(\boldsymbol{C}_1) \setminus E(\boldsymbol{C}_2)$ consisting of pendant sub-trees. Whp $\boldsymbol{C}_2(G)$ consists of a giant 2-connected block $\boldsymbol{B}_2$, and a few small unicyclic sub-graphs $\boldsymbol{U}_2$ ($O(1)$ edges in expectation) each joined to $\boldsymbol{B}_2$ at a cut vertex. These pendant sub-trees and unicyclic sub-graphs are often treated in the same way in our proofs; we use the term *pendicle* to denote either of them.

The following technical lemma, proved in Appendix A, shows that if $c > 1$ then $cx - \ln c > 0$, so that the values of t^* in Theorems 2, 4 are well defined.

Lemma 8. **(a)** *For $c > 1$ the equation $x = 1 - e^{-cx}$ has a unique solution $x \in (0, 1)$. This solution x satisfies $ce^{-cx} < 1$ and $cx - \ln c > \ln(2 - 1/c)$.*

(b) *Let $\theta(c) = cx(2 - x)/(cx - \ln c)$, then $\theta(c)$ is monotone decreasing for $c > 1$ and $\lim_{c\to 1^+} \theta(c) = 4$, $\lim_{c\to\infty} \theta(c) = 1$.*

As previously remarked, the theorems of Section 1 have several cases which require slightly different definitions. The cases (a)-(d) below refer to Theorem 2.

Definitions

For the proof we split Case (a) into 2 sub-cases: In Case (a1), we have $\omega > (\ln \ln n)^4$ and in Case (a2) we see $\omega \le (\ln \ln n)^4$. Let

$$\sigma_1 = \begin{cases} 1 & \textit{Case (a1)} \\ \frac{\ln n}{\omega(\ln \ln n)^{10}} & \textit{Case (a2)} \\ \frac{\ln n}{(\ln \ln n)^{10}} & \textit{Case (b)} \\ \frac{\ln n}{100} & \textit{Case (c)} \end{cases}. \tag{23}$$

We will say that a vertex of degree at most σ_1 is *small*, and a vertex of degree greater than σ_1 is *large*.

A path P is *s-attached to the 2-core $\boldsymbol{C}_2(G)$* if at least s internal vertices of the path have at least σ_1 edges to vertices of the 2-core other than P. A vertex v has an *s-attached k-neighbourhood* if all paths of length k starting at vertex v are s-attached. We define a particular value $s = s_0$ by

$$s_0 = \begin{cases} 64 \ln \ln n & \textit{Case (a1)} \\ 64 & \textit{Case (a2)} \end{cases}.$$

The k-neighbourhood of w is the set of vertices at distance at most k from w. Define $k = k_0$ by

$$k_0 = A_0 \ln \ln n \qquad \textit{Case (a)}.$$

Here $A_0 = A_0(c)$ is a sufficiently large constant.

Let

$$L = \begin{cases} \frac{\ln n + 2s_0 \ln \ln n + cs_0}{cx - \ln c} & \textit{Case (a1)} \\ 2\omega & \textit{Case (a2)} \\ \lfloor \alpha^{-1} \rfloor + 1 & \textit{Case (b)} \\ 2 & \textit{Case (c)} \end{cases}. \tag{24}$$

Special vertices

To get a precise lower bound on the cover time we construct a set of vertices which are hard to cover. For a vertex v in a pendant sub-tree T of the mantle there is a unique path vPw from v joining the 2-core at the root vertex w of T, which in this context we denote by $w(v)$. A vertex v is *special* if the following properties hold:

S1.

$$d(v) = \begin{cases} 1 & \textit{Cases } (a), (b) \\ \leq \ln \ln n & \textit{Case } (c) \end{cases}.$$

S2. **Cases (a),(b)** only. The distance from v to its closest vertex $w(v)$ in $\boldsymbol{C}_2$ is

$$\ell_0 = \begin{cases} \lceil \ln n/(2(cx - \ln c)) \rceil & \textit{Case } (a1) \\ \lceil \omega/2 \rceil & \textit{Case } (a2) \\ \mathrm{argmax}_\ell \{\alpha\ell(1-\alpha\ell)\} & \textit{Case } (b) \end{cases}.$$

S3. **Cases (a),(b)** only. If $w = w(v)$ is the root of the sub-tree containing v, then $w(v)$ has no neighbours outside the 2-core other than the vertex x on the path $vPxw$ from v to w. In Case (b), we also require that w is a large vertex.

S4. **Case (a)** only. The k_0-neighbourhood $N_{k_0}(C_2, w)$ of $w = w(v)$ in the 2-core is a tree (contains no induced cycles) and is an s_0-attached k_0-neighbourhood for w.

Typical properties

A random graph $G \in G_{n,p}$ ($np = c$, $c > 1$) is *typical* if it has the properties listed below. We first explain our notation. Some properties are only valid in certain cases of Theorem 2, others in all cases. Thus **P0** is valid in all cases and whereas **P3**a will only be used in Case (a) of Theorem 2.

P0.

$$\begin{array}{lll} |V(\boldsymbol{C}_1)| \sim xn & |E(\boldsymbol{C}_1)| \sim cx(2-x)n/2 & \\ |V(\boldsymbol{C}_2)| \sim (x - cx + cx^2)n & |E(\boldsymbol{C}_2)| \sim cx^2 n/2 & \\ |V(\boldsymbol{C}_k)| \sim f_k(c, x_k)n & |E(\boldsymbol{C}_k)| \sim cx_k n/2, & k \geq 3, c > c_k \end{array}$$

Furthermore

$$\mathbf{Pr}(||V(\boldsymbol{C}_1)| - xn| \geq n^{3/4}) \leq e^{-n^{1/4}}.$$

P1. (i) The maximum size of a pendicle of $\boldsymbol{C}_1$ is $\Theta(\ln n)$ and altogether, there are $O(\ln n)$ vertices of $\boldsymbol{C}_2$ in unicyclic pendicles.
(ii) The maximum degree of $\boldsymbol{C}_1$ is $\Delta = O(\ln n)$.

P2. The *conductance* $\Phi(\boldsymbol{C}_1) = \Omega(1/\ln n)$, where

$$\Phi = \min_{\pi(S) \le 1/2} \frac{|E(S : \overline{S})|}{d(S)}.$$

Here $d(S) = \sum_{v \in S} d(v)$, $\pi(S) = \frac{d(S)}{d(\boldsymbol{C}_1)}$, and $E(S : \overline{S})$ is the edge set between S and $V \setminus S$ in the graph induced by $\boldsymbol{C}_1$.

P3a. (i) For $10x^{-1}s_0 \le \ell < L$, there are at most n_ℓ paths in $\boldsymbol{C}_1$ of length ℓ that are not s_0-attached, where

$$n_\ell = \begin{cases} cn(ce^{-cx})^\ell (\ell e^c)^{s_0} (\ln n)^2 & \textit{Case (a1)} \\ (2ce^{-c})^\ell n^{1+2s_0/\omega} & \textit{Case (a2)} \end{cases}.$$

(ii) All paths in $\boldsymbol{C}_1$ of length at least L are s_0-attached.

P3b. (i) A connected sub-graph of size $\ln \ln n$ contains at most $\Lambda = \lfloor \alpha^{-1} \rfloor$ *small* vertices.

(ii) Let n_ℓ be the number of paths in $\boldsymbol{C}_1$ of length ℓ in which all vertices except at most one are small. Then

$$n_\ell \le \begin{cases} n^{1-\alpha\ell+o(1)} & \ell \le \Lambda \\ 0 & \ell > \Lambda \end{cases}.$$

(iii) All vertices of the mantle are small.

(iv) No cycle of size at most $(\ln \ln n)^2$ contains a small vertex.

P3c. A connected sub-graph of size $\ln \ln n$ contains at most one small vertex and no cycle of length $\le \ln \ln n$ contains a small vertex.

P4. There are $O((\ln n)^{5k_0})$ vertices within distance $2k_0$ of cycles of size at most $2k_0$ in G.

P5a. (i) A path of length at most k_0^3 has at most one shortcut.
(ii) A path of length k_0^2 with one endpoint in a cycle C of length at most $2k_0$, and edges disjoint from C, is $3s_0$-attached.
(iii) All paths joining two disjoint cycles of length at most $2k_0$ are $3s_0$-attached.

P5b. (i) Two cycles of length at most $\ln \ln n$ are at distance at least $\ln \ln n$ from each other.
(ii) There are at most $(\ln n)^4$ triangles.

P6a. Any cycle $C = xPyQx$ such that xPy, xQy are of length at least k_0 has internally disjoint s_0-attached sub-paths $xP'z_1, xQ'z_2$ where $z_1 \ne z_2$.

P7a. There are $\Theta(n^{1/2-o(1)})$ special vertices v with pair-wise disjoint neighborhoods $N_{k_0}(C_2, w(v))$.

P7b. There is a set of $\Omega(n^{1-\alpha\ell_0-o(1)})$ special vertices such that if $v_1, v_2 \in S$ then every path from v_1 to v_2 contains at least two large vertices.

P7c. There are $\Theta\left(\frac{n^{\delta^+}((1-\delta^+)\ln n)^k}{k!}\right)$ vertices of degree k for $1 \le k \le \ln\ln n$.

The following lemma is proved in Appendix B.

Lemma 9. *If $np = c$, $1 + \Omega(1) < c \le \ln n + (\ln\ln n)^{1/2}$, then $G_{n,p}$ is typical,* **whp**.

3.2 Mixing time

It follows from Jerrum and Sinclair [17] that

$$|P_u^{(t)}(x) - \pi_x| \le (\pi_x/\pi_u)^{1/2}(1 - \Phi^2/2)^t. \tag{25}$$

In our case, we have $\pi_x/\pi_u = O(\ln n)$ and $\Phi = \Omega(1/\ln n)$ and so we can take

$$T = (\ln n)^3 \ln\ln n. \tag{26}$$

and satisfy all the conditions of Lemma 6. The extra factor $\ln\ln n$ will come in useful in the proof of Lemma 14.

We remark that there is a technical point here. The result of [17] 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 (26) still holds. Using a lazy walk doubles the cover time, as it asymptotically doubles the value of R_v, the value of π_v being unchanged. Otherwise, it has a negligible effect on the analysis and we will ignore it for the rest of the paper and continue as though there are no lazy steps.

3.3 R_v for typical graphs

We assume the random walk takes place on the giant component $\boldsymbol{C}_1$ of a typical graph G. Recall that R_v is the expected number of returns made to vertex v during time $0 \le t \le T$ by a random walk $\mathcal{W}_v(\boldsymbol{C}_1)$ starting at v.

Given a vertex v, let M_v be the union of $\boldsymbol{B}_2$ (the giant 2-connected block) and all paths from v to $\boldsymbol{B}_2$. If v is outside $\boldsymbol{B}_2$ in a pendicle, then there will be one or more such paths, but usually one and **whp** at most two. Let $\widehat{R}_v$ be the expected number of returns to v in T steps, of a random walk $\mathcal{W}_v(M_v)$ walking on M_v.

Lemma 10. *Let $d(v)$ be the degree of v in $\boldsymbol{C}_1$ and let $b(v)$ be the degree of v in M_v, then*

$$R_v \le \frac{d(v)}{b(v)}\widehat{R}_v.$$

Proof Let $\xi = b(v)/d(v)$. Let $R_0 \geq R_v$ be the expected number of returns to v in T steps, allowing the walk to enter the parts of the pendicle at v which are not in M_v, but ignoring any time lost in walking them. Then $R_0 = \frac{1}{\nu}\widehat{R}_v$. This is because every visit to v in the walk on M_v gives rise, in expectation, to $1\xi + 2(1-\xi)\xi + 3(1-\xi)^2\xi + ... = 1/\xi$ visits from neighbours not in M_v. □

Let v be a vertex of the giant component. We construct a sub-graph Γ_v of M_v rooted at v as follows: Let L be given by (24), and for $k = 1, 2, \ldots, L$, let

$$S_k(v) = \{v\} \cup N(v) \cup ... \cup N_k(v), \tag{27}$$

where $N_i(v)$ is the set of vertices at distance i from v in M_v. We now prune in a breadth first manner starting from v.

As previously remarked, Theorem 2 has three cases (a),(b),(c), each of which requires slightly different treatment. In the subsequent proofs we often emphasize the case(s) under consideration.

(**Case (a) of Theorem 2.**) Recall from Property **P3**a(2) that all paths of length at least L are s_0-attached. If $u \in S_k(v)$ and all paths from u to v in $S_k(v)$ are s_0 attached, delete all edges from u to $N_{k+1}(v)$. We do this with one caveat: we delete the edge uw only if this does not create a (w, v)-path which is not s_0-attached. This means, that if u is in a cycle of M_v, some edges incident with u may be pruned and others not.

(**Case (b),(c) of Theorem 2.**) We prune from $u \neq v$ if u is a large vertex. The caveat above becomes: Do not delete edge uw if there is no large vertex $x \neq v$ on the (w, v)-path.

After the pruning is completed, at some level $k \leq L$, we have a connected sub-graph Γ_v rooted at v. If we have (partially) pruned at u, we say u is a *boundary vertex*. Denote the set of boundary vertices by Γ_v°. Define $r(v)$ the *radius* of Γ_v as

$$r(v) = \max_{w \in \Gamma_v^\circ} \text{dist}(v, w). \tag{28}$$

We next note some properties of Γ_v.

Lemma 11.
(i) Γ_v is a tree or contains a unique cycle.

*(ii) (***Case (a) of Theorem 2.***) Let $a \in \Gamma_v^\circ$. If Γ_v is a tree there is a unique (a, v)-path which is s_0-attached. If Γ_v contains a cycle, all paths from a to v contain an $s_0/4$-attached sub-path.*

*(iii) (***Case (b),(c) of Theorem 2.***) The boundary Γ_v° consists of large vertices and non-boundary vertices, with the possible exception of v, are small.*

Proof of (i):
(**Case (a) of Theorem 2.**) **P6**a implies that any such cycle has size at most $2k_0$. **P5**a(2)

implies that any such cycle is at distance at most k_0^2 from v. Then **P5**a(1) implies there is at most one such cycle.
(**Case (b)(c) of Theorem 2.**) This follows from **P5**b(i) (resp. **P3**c).

Proof of (ii): If the (a, v)-path is not unique, there is a sub-graph $vPyCzQa$, where C is a cycle, (and possibly P, Q are empty). In this case, either at least one of vPz, yQa is an $s_0/4$-attached path or *both* branches zC_1y, zC_2y of C are internally $s_0/4$-attached. □

Before proceeding we note some standard results on random walks on a path, which we often use to obtain bounds on R_v (see e.g. Feller [15] p314).

For an unbiased random walk on $(0, 1, ..., k)$ starting at vertex 1 and with absorbing states $0, k$, $\mathbf{Pr}($ absorption at $0) = 1 - 1/k$, $\mathbf{Pr}($ absorption at $k) = 1/k$. Thus if 0 is reflecting, the expected number of visits to 0 by the walk before reaching k is k.

Now suppose that we have a path of length a and $d-1$ paths of length $b \geq a$ all with a common vertex O and otherwise vertex disjoint. A walk is started at O and ends when it reaches an endpoint of a path different from O. If ρ is the probability of return to O before absorption then

$$\rho = 1 - \frac{1}{d}\left(\frac{1}{a} + \frac{d-1}{b}\right) = 1 - \frac{b + (d-1)a}{dab}.$$

Consequently, the expected number of returns to O before absorption is

$$\frac{1}{1-\rho} = \frac{dab}{a + (d-1)b}. \tag{29}$$

For a biased random walk on $(0, 1, ..., k)$, starting at vertex 1, with absorbing states $0, k$, and with transition probabilities at vertices $(1, ..., k-1)$ of $q = \mathbf{Pr}(\text{move left})$, $p = \mathbf{Pr}(\text{move right})$; then

$$\mathbf{Pr}(\text{absorption at } k) = \frac{(q/p) - 1}{(q/p)^k - 1}. \tag{30}$$

Lemma 12. *Let R_v^* be the expected number of returns to v in a random walk on Γ_v where the vertices in Γ_v° are made into absorbing states. Then*

$$\widehat{R}_v = R_v^*(1 + o(1)).$$

Proof
(**Case (a) of Theorem 2.**) First suppose that Γ_v is a tree. Given that a walk on Γ_v has reached $w \in \Gamma_v^\circ$ the expected number of returns to v is at most $TR_v^* p_1$ where

$$p_1 = \max_{w \in \Gamma_v^\circ} \mathbf{Pr}(\text{ the walk } \mathcal{W}_w \text{ reaches } v \text{ before returning to } \Gamma_v^\circ).$$

We can get an upper bound on p_1 as follows: By the construction of Γ_v the particle has to pass through at least s_0 vertices of degree at least $\sigma_1 + 2$ (see (23)) in order to reach the root

v. Thus p_1 is at most the probability of absorption at vertex s_0 in a biased random walk on $(0, 1, ..., s_0)$, starting at vertex 1, with absorbing states $0, s_0$, and with transition probabilities at vertices $(1, ..., s-1)$ of $q = \frac{\sigma_1+1}{\sigma_1+2}$, $p = \frac{1}{\sigma_1+2}$. Thus

$$p_1 \leq \mathbf{Pr}(\text{absorption at } 0) = \frac{\sigma_1}{(\sigma_1+1)^{s_0}-1} < 1/(\ln n)^{10}. \tag{31}$$

It follows that the expected number of returns to v from w within T steps is $O(R_v^* T/(\ln n)^{10})$. The proof when Γ_v contains a cycle is similar, except that s_0 is replaced by $\lfloor s_0/4 \rfloor$.

(**Case (b) of Theorem 2.**) The set Γ_v° consists entirely of large vertices, i.e. vertices of degree greater than σ_1 (see (23)). Let $w \in \Gamma_v^\circ$. By **P3**b(i,ii) all but $\Lambda = O(1)$ of its neighbours in M_v are large, whereas in Γ_v, w has at most 2 neighbours. Assume the particle arrives at w for the first time at some fixed step t. With probability $1 - O(1/\sigma_1)$, after the next step $\mathcal{W}_v$ will be at a large neighbour w_1 of w not in Γ_v. Arguing similarly, we see that with probability $1 - O(1/\sigma_1)$, after $\Lambda + 7$ steps we are at distance six large vertices from w in a direction away from v; and these vertices have the property that the particle either moves directly towards or directly away from v when traversing them. We call this second event a success. The expected number of failures before a success is $1/(1-O(1/\sigma_1))$, each failure incurring at most R_v^* expected returns to v.

By the usual comparison with a biased walk on a path, once we are at a vertex x distance 6 large vertices away from Γ_v°, the walk will either return to a vertex x' (equivalent to x), at distance 6 from Γ_v° or arrive at Γ_v° with probability $O(1/\sigma_1^5)$.

Thus the probability that the particle returns to v at some step up to T after a visit to Γ_v° is at most

$$O(1/\sigma_1) + O(T/\sigma_1^5), \tag{32}$$

and the expected number of returns to v is at most

$$R_v^*/(1-O(1/\sigma_1)) + R_v^* O(T/\sigma_1^5) = R_v^*(1+o(1)).$$

(**Case (c) of Theorem 2.**) The proof is similar to (b) above. □

We next note a property of random walks on undirected graphs which follows from result on electrical networks (see e.g. Doyle and Snell [11]). Let v be a given vertex in a graph G and S a set of vertices disjoint from v. Let $p(G)$, *the escape probability*, be the probability that, starting at v, the walk reaches S before returning to v. For an unbiased random walk,

$$p(G) = \frac{1}{d(v) R_{EFF}},$$

where $R_{EFF} = R_{EFF}(G)$ is the effective resistance of G. We assume each edge of G has resistance 1. In the notation of this paradigm, deleting an edge corresponds to increasing the resistance of that edge to infinity. Thus by Raleigh's Monotonicity Law, if edges are deleted from G to form a sub-graph G' then $R_{EFF}(G') \geq R_{EFF}(G)$. Provided we do not delete any edges incident with v, it follows that $p(G') \leq p(G)$. However $p(G) = 1 - \rho$, where ρ is the probability that the walk returns to v before absorption at S, and hence $\rho' \geq \rho$. Thus $R_{v,S}$, the expected number of returns to v before absorption at S satisfies

$$R_{v,S} = \frac{1}{1-\rho} \leq \frac{1}{1-\rho'} = R'_{v,S}. \tag{33}$$

Finally we note that sub-dividing edges increases effective resistance, as an edge of resistance 1 whose resistance is increased to 2, is equivalent to two edges of resistance 1 in series. This allows us to increase the length of paths to the boundary when considering upper bounds.

Lemma 13.

(a) *(**Case (a),(b) of Theorem 2.**) If v doesn't lie on a cycle of Γ_v then* $R_v \leq (1+o(1))\frac{d(v)}{b(v)}r(v)$.

(b) *(**Case (a) of Theorem 2.**) If v is on a cycle of Γ_v then $R_v \leq (1+o(1))\frac{d(v)}{b(v)-1}r(v)$.*

(c) *(**Case (b) of Theorem 2.**) If v is on a cycle of Γ_v then $R_v = 1 + o(1)$.*

(d) *(**Case (c) of Theorem 2.**) $R_v = 1 + o(1)$ for all $v \in V$.*

(e) *(**Case (a),(b) of Theorem 2.**) If v is a special vertex then $R_v \geq \ell_0 - O(1)$.*

Proof

It follows from Lemmas 10 and 12 that it suffices to consider a random walk on Γ_v with absorbing states at Γ_v°.

(a) Assume first that Γ_v is a tree. There is a path from each edge incident with v to the boundary. If any of these paths P is of length less than $r(v)$ insert edges in Γ'_v to increase the length of P to $r(v)$. The remarks prior to this lemma show that $R'_v \geq R_v$ where R'_v is the expected number of returns to v on Γ'_v. For walks on Γ'_v we have $R'_v = r(v)$.

If Γ_v contains a cycle C disjoint from v, we can delete (one of) the furthest cycle edges from v. If this creates a vertex w of degree one then delete w and repeat until we have formed a tree sub-graph Γ'_v without increasing $r(v)$. The process of deletion will finish before we reach a neighbour of v. The previous argument for case (a), when Γ_v is a tree, is now valid.

(b) If Γ_v contains a cycle C through v then we delete a cycle as in Case (a) and prune vertices of degree one. If the pruning process ends before we reach a neighbour of v then we bound R_v as in Case (a). If however, the deletion of a cycle edge (x, y) results in an induced path from v, ending at x, then we treat this path as not part of M_v in the argument of Lemma 10. This reduces $b(v)$ to $b(v) - 1$ and explains the bound.

(c) In this case v is a large vertex and Γ_v consists of a star centered at v together with an edge joining two neighbours of v. Thus $R_v^* = 1$ and the bound follows.

(d) $R_v \geq 1$ and as argued in Lemma 12 above, on reaching a boundary vertex w, the probability that $\mathcal{W}_v$ returns to v in T steps is only $o(1)$. If v is small, then Γ_v is a star and we are done. If Γ_v contains a cycle C, then v is large and C is a triangle. We delete one edge to form a star giving an upper bound. If v is large and Γ_v is a tree, then there is at most one small neighbour w on a path of length $L = 2$ to the boundary. The probability of visiting w before reaching the boundary is $O(1/\ln n)$.

(e) Let vPw be the unique path in the mantle of length ℓ_0 joining v to $w = w(v)$. Let $\mathcal{T}(\ell_0)$ be the tree in $\boldsymbol{C}_1$ rooted at w containing the vertex v. In this tree vertices v, w have degree 1. In order to prove that $R_v \geq \ell_0 - O(1)$ we need to prove that **whp** the walk $\mathcal{W}_v$ has reached w by time T. This is because the expected number of returns to v before $\mathcal{W}_v$ reaches w is precisely ℓ_0.

Label the vertices on vPw as $(v = v_0, ..., v_{\ell_0} = w)$. In general let $\mathcal{T}(k)$ be the sub-tree of $\mathcal{T}(\ell_0)$ containing $(v_0, ..., v_k)$ obtained by pruning $\mathcal{T}(\ell_0)$ at vertex v_k, which now becomes a leaf. Let $m(k)$ be the number of edges of $\mathcal{T}(k)$. For a walk restricted to $\mathcal{T}(k)$ define a random variable $\boldsymbol{H}(i,k)$ as the number of steps before v_k is first visited by a walk starting at v_i. Let $H(i,k) = \mathbf{E}(\boldsymbol{H}(i,k))$, the *access time*. For walks on a path there is a standard method of calculating $H(i,k)$ which extends easily to trees. We have $H(i,k) = H(i,k-1) + H(k-1,k)$. Also $H(k,k) = 1 + H(k-1,k)$, where $H(k,k)$ has the meaning of the expected time to a first return in $\mathcal{T}(k)$ which can only be via v_{k-1}. Thus we have $H(i,k) = H(i,k-1) + H(k,k) - 1$. By the ergodic theorem for $\mathcal{T}(k)$, $H(k,k) = 2m(k)/d(k) = 2m(k)$. As v is special, v is a leaf in $\boldsymbol{C}_1$, and thus

$$H(0,\ell_0) = (2m(\ell_0) - 1) + (2m(\ell_0 - 1) - 1) + \cdots + 1 \leq 2\ell_0 m(\ell_0) = O((\ln n)^2),$$

where by **P1**, $m(\ell_0) = O(\ln n)$. It follows that

$$\mathbf{Pr}(\boldsymbol{H}(v,w) \geq T) = O\left(\frac{1}{\ln n}\right),$$

and finally,

$$R_v \geq \ell_0 \left(1 - O\left(\frac{1}{\ln n}\right)\right).$$

□

3.4 The conditions of Lemma 6

Lemma 14. *For $|z| \leq 1 + \lambda$, there exists a constant $\theta > 0$ such that $|R_T(z)| \geq \theta$.*

Proof As before, let Γ_v° be the set of absorbing states of Γ_v. For walks in Γ_v, starting at v, let $\beta(z) = \sum_{t=1}^T \beta_t z^t$ where β_t is the probability of a first return to v at time $t \leq T$. Let $\alpha(z) = 1/(1-\beta(z))$, and write $\alpha(z) = \sum_{t=0}^\infty \alpha_t z^t$, so that α_t is the probability that our walk is at v at time t. We shall prove below that the radius of convergence of $\alpha(z)$ is at least $1 + \Omega(1/L^2)$.

We can write

$$\begin{aligned} R_T(z) &= \alpha(z) + Q(z) \\ &= \frac{1}{1-\beta(z)} + Q(z), \end{aligned} \tag{34}$$

where $Q(z) = Q_1(z) + Q_2(z)$, and

$$\begin{aligned} Q_1(z) &= \sum_{t=s_0+1}^{T} (r_t - \alpha_t) z^t \\ Q_2(z) &= -\sum_{t=T+1}^{\infty} \alpha_t z^t. \end{aligned}$$

Here we have used the fact that $\alpha_t = r_t$ for $0 \leq t \leq s_0$, as all paths from the boundary to v contain at least s_0 vertices of degree at least $\sigma_1 + 2$. We note that $Q(0) = 0$, $\alpha(0) = 1$ and $\beta(0) = 0$.

We claim that the expression (34) is well defined for $|z| \leq 1 + \lambda$. We will show below that

$$|Q_2(z)| = o(1) \tag{35}$$

for $|z| \leq 1 + 2\lambda$ and thus the radius of convergence of $Q_2(z)$ (and hence $\alpha(z)$) is greater than $1 + \lambda$. This will imply that $|\beta(z)| < 1$ for $|z| \leq 1 + \lambda$. For suppose there exists z_0 such that $|\beta(z_0)| \geq 1$. Then $\beta(|z_0|) \geq |\beta(z_0)| \geq 1$ and we can assume (by scaling) that $\beta(|z_0|) = 1$. We have $\beta(0) < 1$ and so we can assume that $\beta(|z|) < 1$ for $0 \leq |z| < |z_0|$. But as ρ approaches 1 from below, (34) is valid for $z = \rho|z_0|$ and then $|R_T(\rho|z_0|)| \to \infty$, contradiction.

Recall that $\lambda = 1/KT$. Clearly $\beta(1) \leq 1$ and so for $|z| \leq 1 + \lambda$

$$\beta(|z|) \leq \beta(1+\lambda) \leq \beta(1)(1+\lambda)^T \leq e^{1/K}.$$

Using $|1/(1-\beta(z))| \geq 1/(1+\beta(|z|))$ we obtain

$$|R_T(z)| \geq \frac{1}{1+\beta(|z|)} - |Q(z)| \geq \frac{1}{1+e^{1/K}} - |Q(z)|. \tag{36}$$

We now prove that $|Q(z)| = o(1)$ for $|z| \leq 1 + \lambda$ and the lemma will follow.

Turning our attention first to $Q_1(z)$, we note that $r_t - \alpha_t$ is at most the probability of a return to v within time T, after a visit to Γ_v° for a walk on $\boldsymbol{C}_1$. The following results hold both for

Γ_v a tree, and Γ_v containing a cycle. Considering the various cases of Theorem 2, we see that from (31) (resp. paragraph preceding (32))

$$|Q_1(z)| \leq (1+\lambda)^T Q_1(1) \leq (1+\lambda)^T T/(\ln n)^{5-o(1)} = o(1). \tag{37}$$

We next turn our attention to $Q_2(z)$. The proof of (35) given below holds for all cases of Theorem 2.

Let σ_t be the probability that the walk on Γ_v has not been absorbed by step t. Then $\sigma_t \geq \alpha_t$, and so

$$|Q_2(z)| \leq \sum_{t=T+1}^{\infty} \sigma_t |z|^t,$$

Assume first that Γ_v is a tree. We estimate an upper bound for σ_t as follows: Consider an unbiased random walk $X_0^{(b)}, X_1^{(b)}, \ldots$ starting at $|b| < a \leq L$ on the finite line $(-a, -a+1, ..., 0, 1, ..., a)$, with absorbing states $-a, a$.

$X_m^{(0)}$ is the sum of m independent ± 1 random variables. So the central limit theorem implies that there exists a constant $c > 0$ such that

$$\mathbf{Pr}(X_{ca^2}^{(0)} \geq a \text{ or } X_{ca^2}^{(0)} \leq -a) \geq 1 - e^{-1/2}.$$

Consequently, for any b with $|b| < a$,

$$\mathbf{Pr}(|X_{2ca^2}^{(b)}| \geq a) \geq 1 - e^{-1}. \tag{38}$$

Hence, for $t > 0$,

$$\sigma_t = \mathbf{Pr}(|X_\tau^{(0)}| < a,\ \tau = 0, 1, \ldots, t) \leq e^{-\lfloor t/(2ca^2) \rfloor}. \tag{39}$$

Thus the radius of convergence of $Q_2(z)$ is at least $e^{1/(3ca^2)}$. As $a \leq L$, $e^{1/(3ca^2)} \gg 1 + 2\lambda$ and for $|z| \leq 1 + 2\lambda$,

$$|Q_2(z)| \leq \sum_{t=T+1}^{\infty} e^{2\lambda t - \lfloor t/(2ca^2) \rfloor} = o(1).$$

This lower bounds the radius of convergence of $\alpha(z)$, proves (35) and then (37), (35) and (36) complete the proof of the case where Γ_v is a tree.

We now turn to the case where Γ_v contains a unique cycle C. The place where we have used the fact that Γ_v is a tree is in (39) which relies on (38). Let x be the furthest vertex of C from v in Γ_v. This is the only possible place where the random walk is more likely to get closer to v at the next step. We can see this by considering the breadth first construction of Γ_v. Thus we can compare our walk with random walk on $[-a, a]$ where there is a unique value $d < a$ such that only at $\pm d$ is the walk more likely to move towards the origin and even then this probability is at most $2/3$.

From (38) we see that

$$\mathbf{Pr}(\exists \tau \le ca^2 :\ |X_\tau^{(b)}| = d) \ge 1 - e^{-1/2}.$$

The probability the particle walks from e.g. d to a without returning to the cycle is at least $1/3(a-d)$. Thus

$$\mathbf{Pr}(\exists \tau \le ca^2 :\ |X_{\tau+a-d}^{(b)}| = a) \ge (1 - e^{-1/2})/3a,$$

and

$$\sigma_t = \mathbf{Pr}(|X_\tau^{(0)}| < a,\ \tau = 0, 1, \ldots, t) \le (1 - (1 - e^{-1/2})/3a)^{\lfloor t/(2ca^2) \rfloor} \le e^{-t/(20cL^3)}.$$

The radius of convergence of $Q_2(z)$ is therefore at least $1 + \frac{1}{25cL^3} > 1 + 2\lambda$, assuming that K (defined in (9)) is sufficiently large. Finally, if $z \in C_\lambda$ then

$$|Q_2(z)| \le \sum_{t=T+1}^{\infty} e^{(\lambda - 1/(20cL^3))t} \le \frac{e^{-T/(25cL^3)}}{1 - e^{-1/(25cL^3)}} = o(1)$$

using the extra factor $\ln \ln n$ in the definition of T (see (26)). □

3.5 Upper bound on cover time

Let $t_1 = t_1^*(1+\epsilon)$, where $\epsilon \to 0$ sufficiently slowly that any subsequently claimed inequalities are valid. An upper bound of $t_1(1+o(1))$ for the cover time of $\boldsymbol{C}_1$ is established in Lemma 16 (below), which we prove after some preliminary steps.

Let $T_G(u)$ be the time taken by the random walk $\mathcal{W}_u$ to visit every vertex of a connected graph G. 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:

$$C_u = \mathbf{E}(T_G(u)) = \sum_{t>0} \mathbf{Pr}(T_G(u) \ge t), \tag{40}$$

$$\mathbf{Pr}(T_G(u) \ge t) = \mathbf{Pr}(T_G(u) > t-1) = \mathbf{Pr}(U_{t-1} > 0) \le \min\{1, \mathbf{E}(U_{t-1})\}. \tag{41}$$

As in Corollary 7, let $\boldsymbol{A}_v(t)$, $t \ge T$ be the event that $\mathcal{W}_u(t)$ has not visited v in the interval $[T, t]$. It follows from (40), (41) that for all $t \ge T$,

$$C_u \le t + 1 + \sum_{s \ge t} \mathbf{E}(U_s) \le t + 1 + \sum_v \sum_{s \ge t} \mathbf{Pr}(\boldsymbol{A}_s(v)). \tag{42}$$

Recall from (10) that $p_v = (1+o(1))d(v)/(2mR_v)$, where m is the number of edges and $R_v = R_T(1)$. In Section 3.4 we established that condition (a) of Lemma 6 holds, and in Section 3.3 we derived bounds for the parameter R_v. Using **P0**, **P1**(ii) and T given by (26), it is easily checked that condition (b) of Lemma 6 holds. Thus by Corollary 7, the probability that $\mathcal{W}_u$ has not visited v during $[T, t]$ is given by

$$\begin{aligned} \mathbf{Pr}(\boldsymbol{A}_t(v)) &= (1+o(1))e^{-tp_v} + o(e^{-\lambda t/2}) && (43) \\ &= (1+o(1))e^{-tp_v}, && (44) \end{aligned}$$

where from (9), $\lambda = 1/KT$ and $p_v/\lambda = O(T\pi_v) = o(1)$ by Lemma 6(b). From Lemma 13 we see that $R_v \leq (1+o(1))L$ where L is given by (24). This implies that for any set $S \subseteq V$

$$\sum_{v\in S}\sum_{s\geq t}\mathbf{Pr}(\boldsymbol{A}_s(v)) = O\left(mL\right)\sum_{v\in S}\mathbf{Pr}(\boldsymbol{A}_t(v)). \quad (45)$$

We note in all cases of Theorem 2, that

$$mL = O(t_1^*/\ln\ln n). \quad (46)$$

We adopt the notation $\overset{O}{\leq}$ to stand for $\leq O()$ (resp. $\overset{\Omega}{\geq}$ to stand for $\geq \Omega()$) and thus avoid large unsightly brackets.

Lemma 15. *Let* $F(k) = k(cx - \ln c) + \frac{(1+\epsilon)(\ln n)^2}{4k(cx-\ln c)}$*, then for* $c > 1$

$$\sum_{k=s_0}^{L} e^{-F(k)} \overset{O}{\leq} n^{-(1+\epsilon)^{1/2}}\sqrt{\ln n}.$$

Proof The function $F(k)$ is minimized at $k^* = \frac{(1+\epsilon)^{1/2}\ln n}{2(cx-\ln c)}$. For $k = (1+\delta)k^*$, $F(k) = (1+\epsilon)^{1/2}(1+\delta^2/(2(1+\delta)))\ln n$ and thus for $\delta \leq 1$,

$$e^{-F(k)} \leq n^{-(1+\epsilon)^{1/2}} e^{-\frac{(1+\epsilon)^{1/2}\ln n}{4}\delta^2}.$$

Thus

$$\begin{aligned} \sum_{k=s_0}^{L} e^{-F(k)} &\overset{O}{\leq} \int_{s_0}^{L} e^{-F(k)}\,dk \\ &\overset{O}{\leq} n^{-(1+\epsilon)^{1/2}} k^* \int_{-\infty}^{\infty} e^{-\frac{(1+\epsilon)^{1/2}\ln n}{4}\delta^2}\,d\delta \\ &\overset{O}{\leq} n^{-(1+\epsilon)^{1/2}} \frac{\sqrt{\ln n}}{cx-\ln c}. \end{aligned}$$

The lemma now follows from Lemma 8(a), as $cx - \ln c > \ln(2-1/c)$ for $c > 1$. □

We now prove the upper bounds on cover time for the various cases given in Theorem 2.

Lemma 16. *$C_u \le (1+o(1))t_1$ for all $u \in V(\boldsymbol{C}_1)$.*

Proof Let

$$\begin{aligned} V_1 &= \{v : v \text{ doesn't lie on a cycle of } \Gamma_v\} \\ V_2 &= \{v : v \text{ lies on a cycle of } \Gamma_v\} \end{aligned}$$

Going back to (45) and (46) it will suffice to show

$$\sum_{v \in V_i} \mathbf{Pr}(\boldsymbol{A}_{t_1}(v)) = o(1), \qquad i = 1, 2. \tag{47}$$

(**Case (a(1)) of Theorem 2.**) We first consider V_1, and evaluate $\sum_{v \in V_1} \mathbf{Pr}(\mathcal{A}_{t_1}(v))$. Let $r(v)$ be the radius of Γ_v (see (28)). Lemma 13(a) gives $R_v \le (1+o(1))d(v)r(v)/b(v)$. Thus, using (43)-(44)

$$\begin{aligned} \mathbf{Pr}(\mathcal{A}_{t_1}(v)) &\le 2\exp\left\{-t_1 \frac{d(v)}{2m} \frac{b(v)}{(1+o(1))d(v)r(v)}\right\} & (48) \\ &\le 2\exp\left\{-\frac{(1+\epsilon/2)(\ln n)^2}{4r(v)(cx - \ln c)}\right\}. & (49) \end{aligned}$$

Let $D_\ell = \{v \in V_1, r(v) = \ell\}$ then $|D_\ell| \le n_{\ell-1} = O(n_\ell)$. It follows from **P3**a, that for $\ell \ge 10s_0/x$, and from Lemma 15 that for $t \ge t_1$

$$\begin{aligned} \sum_{v \in V_1} \mathbf{Pr}(\mathcal{A}_t(v)) &\le \sum_{\ell=1}^{L} |D_\ell| \exp\left\{-\frac{(1+\epsilon/2)(\ln n)^2}{4\ell(cx - \ln c)}\right\} & (50) \\ &\le \sum_{\ell=1}^{10s_0/x} n\exp\left\{-\frac{(1+\epsilon/2)(\ln n)^2}{4\ell(cx - \ln c)}\right\} \\ &+ \sum_{\ell=10s_0/x}^{L} cn(\ln n)^2 (Le^c)^{s_0} \exp\left\{-\ell(cx - \ln c) - \frac{(1+\epsilon/2)(\ln n)^2}{4\ell(cx - \ln c)}\right\} & (51) \\ &\overset{O}{\le} n^{-100} + n^{1-(1+\epsilon/2)^{1/2}} (Le^c)^{s_0} (\ln n)^{5/2} \\ &= o(1). & (52) \end{aligned}$$

We next consider the case $v \in V_2$. v belongs to a cycle of Γ_v. By **P6**a any such cycle is of length at most $2k_0$ and by **P4** there are $O((\ln n)^{5k_0})$ such vertices and by Lemma 13(b), $R_v \le \frac{d(v)}{b(v)-1} L(1+o(1))$ where $b(v) \ge 2$, Consequently,

$$\sum_{v \in V_2} \mathbf{Pr}(\mathcal{A}_t(v)) \overset{O}{\le} (\ln n)^{5k_0} e^{-t/(3mL)} = o(1). \tag{53}$$

(**Case (a(2)) of Theorem 2.**) As in Case (a(1)), for V_1,

$$\begin{aligned}
\sum_{v\in V_1}\mathbf{Pr}(\mathcal{A}_{t_1}(v)) &\leq \sum_{\ell=1}^{L}|D_\ell|\exp\left\{-\frac{(1+\epsilon/2)(\ln n)^2}{4\ell(cx-\ln c)}\right\}\\
&\leq \sum_{\ell=1}^{10s_0/x} n\exp\left\{-\frac{(1+\epsilon/2)(\ln n)^2}{4\ell(cx-\ln c)}\right\}\\
&\qquad + \sum_{\ell=10s_0/x}^{L}(2c)^\ell n^{1-\ell/\omega-\omega/(4\ell)+2s_0/\omega-\epsilon\omega/(8\ell)} && (54)\\
&\leq n^{-100}+\sum_{\ell=10s_0/x}^{L}\exp\left\{-\left(\frac{\epsilon\omega}{8\ell}-O\left(\frac{1}{\omega}\right)\right)\ln n\right\} && (55)\\
&= o(1). && (56)
\end{aligned}$$

The case $v\in V_2$ is dealt with as it was in the previous case.

(**Case (b) of Theorem 2.**) Using (48) we obtain

$$\begin{aligned}
\sum_{v\in V_1}\mathbf{Pr}(\mathcal{A}_{t_1}(v)) &\leq \sum_{\ell=1}^{L}|D_\ell|\exp\{-(1+\epsilon/2)\gamma\ln n/(\alpha\ell)\} && (57)\\
&\leq \sum_{\ell=1}^{L} n^{1-\alpha\ell+o(1)-(1+\epsilon/2)\gamma/(\alpha\ell)} && (58)\\
&= n^{o(1)-\epsilon\gamma/2\alpha L} && (59)\\
&= o(1).
\end{aligned}$$

Line (59) follows from $1-\alpha\ell-\gamma/\alpha\ell\leq 0$ when γ is given by (2).

In the case $v\in V_2$, we see that v is large and Γ_v is a star plus a single edge. Furthermore, $R_v=1+o(1)$, (see Lemma 13(c)) and so

$$\sum_{v\in V_3}\mathbf{Pr}(\mathcal{A}_{t_1}(v)) \overset{O}{\leq} (\ln n)^4 e^{-(1-o(1))\sigma_1 t_1/(2m)} = o(1).$$

(**Case (c) of Theorem 2.**) Recall that now $R_v=1+o(1)$ for all $v\in V$ (see Lemma 13(d)). Thus from (42)-(45) and **P7**c,

$$\begin{aligned}
C_u-(t_1+1) &\overset{O}{\leq} n^{\delta^+}\sum_{k\geq 1}\frac{(\ln n)^k}{k!}\sum_{s\geq t_1}\exp\left\{-\frac{ks}{(1+o(1))n\ln n}\right\} && (60)\\
&\overset{O}{\leq} n\ln n\sum_{k\geq 1}\frac{1}{k!}\exp\left\{-k\epsilon(\ln\ln n+\delta^+\ln n)/2\right\}\\
&= o(t_1). && (61)
\end{aligned}$$

$\square$

This completes the proofs of the upper bound on cover time for the cases given in Theorem 2.

3.6 Lower bound on cover time

Let $t_0 = t_1^*(1-\epsilon)$ where $\epsilon \to 0$ sufficiently slowly that any subsequently claimed inequalities are valid. We prove that at time t_0, the probability that the set S of special vertices is covered by the walk $\mathcal{W}_u$ tends to zero, (see Section 3.1 for the definition of special). Hence $T_1(u) > t_0$ **whp** which implies that $C_G \geq t_1^* - o(t_1^*)$.

Let $\eta(x, y)$ denote the probability that $\mathcal{W}_x$ visits y within the first T steps. Let

$$\Sigma_x = \left\{ y : \max\{\eta(x,y), \eta(y,x)\} \geq \frac{1}{(\ln n)^{10}} \right\}.$$

We will prove in Section 3.6.1 below that **whp**

$$|\Sigma_x| \leq (\ln n)^{15} \qquad \text{for all } x \in V. \tag{62}$$

Given this we prove below that in special cases (a), (b), (c), we can choose a sufficiently large subset $S \subseteq S_1$ satisfying

$$\eta(v, v') \leq 1/(\ln n)^{10} \qquad \text{for all } v, v' \in S. \tag{63}$$

Let $X = \sum_{v \in S} Y_v$ denote the subset of S which is unvisited in $[T, t_0]$. It follows from Corollary 7 that

$$\mathbf{E}(X) = \sum_{v \in S} \mathbf{Pr}(\boldsymbol{A}_{t_0}(v)) = \sum_v (1 + O(T\pi_v))e^{-t_0 p_v} + o(\sqrt{n}e^{-\lambda t_0/2}), \tag{64}$$

Case (a): By **P7**a, there is a set S_1 of $\Theta(n^{1/2-o(1)})$ special vertices v with pairwise disjoint neighbourhoods $N_{k_0}(C_2, w(v))$. Given this we use (62) to choose a set $S \subseteq S_1$ of size $\Theta(n^{1/2-o(1)})$ satisfying (63). The value of p_v in (64) is given by

$$p_v \leq \begin{cases} \frac{(2+o(1))(cx - \ln c)}{2m \ln n} & \textit{Case } (a(1)) \\ \frac{2+o(1)}{n \ln n} & \textit{Case } (a(2)) \end{cases}.$$

Thus for some constant $A_1 > 0$,

$$\mathbf{E}(X) \geq n^{A_1 \epsilon}. \tag{65}$$

Case (b): By **P7**b there is a set S of $\Theta(n^{1-\alpha\ell_0-o(1)})$ special vertices such that $\Gamma_v, \Gamma_{v'}$ are disjoint for $v, v' \in S$. We assume that (62), (63) hold and use (64) with

$$p_v \leq \frac{1+o(1)}{\alpha \ell_0 n \ln n}.$$

Since $\gamma = \alpha\ell_0(1-\alpha\ell_0)$ this gives

$$\mathbf{E}(X) \geq \Theta(n^{1-\alpha\ell_0-\gamma/(\alpha\ell_0)+\epsilon\gamma/(2\alpha\ell_0)}) = \Theta(n^{\epsilon\gamma/(2\alpha\ell_0)})$$

and (65) holds here too.

Case (c): Assume for the moment that

$$\delta^+ \ln n \notin [0, (\ln\ln n)^2].$$

It follows from **P3**c that $\Gamma_v, \Gamma_{v'}$ are disjoint for any pair of special vertices v, v'. Now let

$$k = \begin{cases} 1 & \delta^+ \ln n > (\ln\ln n)^2 \\ \ln\ln n & \delta^+ = 0 \end{cases}.$$

Then, by **P7**c,

$$\mathbf{E}(X) \overset{\Omega}{\geq} n^{\delta^+} \frac{(\ln n)^k}{k!} e^{-k(1-\epsilon/2)(\delta^+ \ln n + \ln\ln n)} \gg (\ln n)^{100},$$

which is sufficiently large to allow the deletion of the set defined in (62).

Having bounded $\mathbf{E}(X)$ from below in all cases, we continue by estimating the second moment of X.

Fix $v, v' \in S$. We will show that

$$\mathbf{Pr}(\mathcal{A}_{t_0}(v) \wedge \mathcal{A}_{t_0}(v')) = (1+o(1))\mathbf{Pr}(\mathcal{A}_{t_0}(v))\mathbf{Pr}(\mathcal{A}_{t_0}(v')). \tag{66}$$

It then follows that

$$\mathbf{E}(X^2) = (1+o(1))\mathbf{E}(X)^2. \tag{67}$$

Using the Chebyshev inequality we see that

$$\mathbf{Pr}(X < \mathbf{E}(X/2)) = o(1)$$

and thus **whp** at least $\mathbf{E}(X)/2 - T > 0$ vertices of S are unvisited at t_0.

We define a new graph G_ψ by identifying v, v' and replacing them with a new node ψ. The conductance of G_ψ can easily be seen to be $\Omega(1/\ln n)$.

Walks in G_ψ can be mapped to walks in G in a natural way. If the walk is not at ψ then it chooses its successor with the same probability. This includes neighbours of v, v', since they are non-adjacent in v. When at ψ, with probability 1/2 it moves to a neighbour of v and with probability 1/2 it moves to a neighbour of v'.

Let R_v^* be the expected number of returns to v in time $n^{1/2}$ steps. Let ρ_v denote the probability of a return to v within T steps. Define $R_{v'}^*, R_\psi^*, \rho_{v'}, \rho_\psi$ similarly. We claim that

C1

$$R_v^* = R_v + O(n^{1/2}\pi_v).$$

C2

$$(1-\rho_v)\sum_{k=1}^{n^{1/4}} k\rho_v^{k-1} \le R_v^* \le \sum_{k=1}^{\infty} k\rho_v^{k-1}(1-\rho_v) + O(n^{1/2}\pi_v).$$

C3

$$\rho_\psi = \frac{\rho_v + \rho_{v'}}{2} + O(1/(\ln n)^{10}).$$

C1 comes from the fact that after time T the probability of being at v is close to π_v and so the expected number of visits to v in the time interval $[T, n^{1/2}]$ is $O(n^{1/2}\pi_v)$.

The LHS of **C2** is the probability of $0 \le k \le n^{1/4}$ returns in time $n^{1/2}$ followed by no return within T steps. The RHS is the expected number of quick returns altogether and the $O(n^{1/2}\pi_v)$ term accounts for returns that take longer than T.

The first term in **C3** accounts for choosing a neighbour of v with probability 1/2 etc. and the error term uses (63) to bound the probability of visits from v to v' within time T etc.

C1 and **C2** can be seen to hold for v' and ψ too.

Now $R_v^* = O(\ln n)$ implies that $\rho_v = 1 - \Omega(1/\ln n)$, otherwise the first inequality in **C2** would fail. It then follows from **C1** and **C2** that

$$R_v = \frac{1}{1-\rho_v} + O(n^{1/2}\pi_v).$$

Applying **C3** we then see that

$$\frac{2}{R_\psi} = \frac{1}{R_v} + \frac{1}{R_{v'}} + O(1/(\ln n)^{10}). \tag{68}$$

So, with $\mathbf{Pr}_\psi$ referring to probability in the space of random walks on G_ψ,

$$\begin{aligned} \mathbf{Pr}_\psi(\mathcal{A}_{t_0}(\psi)) &= (1+o(1))\exp\left\{-\frac{t_0\pi_\psi}{(1+O(T\pi_\psi))R_\psi}\right\} \\ &= (1+o(1))\exp\left\{-\frac{t_0\pi_v}{R_v}\right\}\exp\left\{-\frac{t_0\pi_{v'}}{R_{v'}}\right\} \\ &= (1+o(1))\mathbf{Pr}(\mathcal{A}_{t_0}(v))\mathbf{Pr}(\mathcal{A}_{t_0}(v')). \end{aligned} \tag{69}$$

But, using rapid mixing in G_ψ,

$$\begin{aligned}
\mathbf{Pr}_\psi(\mathcal{A}_{t_0}(\psi)) &= \sum_{x\neq\psi} P^{T_\psi}_{\psi,u}(x)\mathbf{Pr}_\psi(\mathcal{W}_x(t-T_\psi)\neq\psi,\ T_\psi\leq t\leq t_0) \\
&= \sum_{x\neq\psi}\left(\frac{deg(x)}{2m}+O(n^{-3})\right)\mathbf{Pr}_\psi(\mathcal{W}_x(t-T_\psi)\neq\psi,\ T_\psi\leq t\leq t_0) \\
&= \sum_{x\neq v,w}\left(P^{T_\psi}_u(x)+O(n^{-3})\right)\mathbf{Pr}(\mathcal{W}_x(t-T_\psi)\neq v,v',\ T_\psi\leq t\leq t_0) \qquad (70)\\
&= \mathbf{Pr}(\mathcal{W}_u(t)\neq v,v',\ T_\psi\leq t\leq t_0)+O(n^{-3}) \\
&= \mathbf{Pr}(\mathcal{A}_{t_0}(v)\wedge\mathcal{A}_{t_0}(v'))+O(n^{-3}). \qquad (71)
\end{aligned}$$

Equation (70) follows because there is a natural measure preserving map ϕ between walks in G that start at $x\neq v,v'$ and avoid v,v' and walks in G_ψ that avoid ψ.

We are left with Case (c) and $\delta^+\ln n\in[0,(\ln\ln n)^2]$. Our current argument shows that for every start vertex u, there are **whp** at least $(n^{\delta^+}\ln n)/2$ vertices of degree 1 that will not be visited in the time interval $[T,t_0]$. However, we must also consider the possibility that these vertices have been visited before time T. The problem we have is that $(n^{\delta^+}\ln n)/2$ may be smaller than T. But, suppose that u is distributed as a particle in steady state distribution of a random walk on $\boldsymbol{C}_1$. Then the probability that the walk visits a vertex of degree 1 during $[0,T]$ is $O(Tn^{\delta^+-1})=o(1)$. We deduce that almost every u is such that a walk starting from u, **whp** avoids all vertices of degree 1 during the interval $[0,T]$. For any such u, there will **whp** be unvisited vertices of degree 1 by time t_0.

This completes the proof of Theorem 2.

3.6.1 Proof of (62)

Let Z_x be the number of vertices visited by $\mathcal{W}_x$ in the first T steps. Then

$$T\geq\mathbf{E}(Z_x)=\sum_{y\in V}\eta(x,y). \qquad (72)$$

For $\epsilon>0$, let

$$A_\epsilon(x)=\{y\in V:\ \eta(x,y)\geq\epsilon\}.$$

It follows from (72) that

$$|A_\epsilon(x)|\leq\frac{T}{\epsilon}. \qquad (73)$$

Similarly, for $\epsilon>0$ let

$$B_\epsilon(x)=\{y\in V:\ \eta(y,x)\geq\epsilon\}.$$

By stationarity, for fixed t,

$$\sum_{y \in V} \pi_y \mathbf{Pr}(\mathcal{W}_y(t) = x) = \pi_x.$$

Thus

$$\begin{aligned}
T\pi_x &= \sum_{1 \le t \le T} \sum_{y \in V} \pi_y P_y^{(t)}(x) \\
&= \sum_{y \in V} \pi_y \sum_{1 \le t \le T} P_y^{(t)}(x) \\
&\ge \sum_{y \in V} \pi_y \eta(y, x) \\
&\ge \sum_{y \in B_x(\epsilon)} \pi_y \eta(y, x) \\
&\ge \pi_{min} \epsilon |B_x(\epsilon)|.
\end{aligned}$$

where $\pi_{min} = \min \{\pi_y : y \in V\}$.

Consequently,

$$|B_\epsilon(x)| \le \frac{T\pi_x}{\epsilon \pi_{min}}. \tag{74}$$

Now it follows from (73) and (74) that for all $x \in V$,

$$\begin{aligned}
|\left\{y : \eta(x, y) \ge 1/(\ln n)^{10}\right\}| &= O((\ln n)^{13} \ln \ln n) \\
|\left\{y : \eta(y, x) \ge 1/(\ln n)^{10}\right\}| &= O((\ln n)^{14} \ln \ln n)
\end{aligned}$$

Equation (62) follows immediately.

3.7 The edge cover time of the giant component

We will limit ourselves to a brief exposition of the proofs of Theorem 3 and the remaining theorems, as their analysis is very similar to that for the cover time of the Giant. We identify certain vertices as special by direct analogy with the definitions used for the case of the Giant component (see Section 3).

For an edge $e = \{u, v\}$, the probability $\mathbf{Pr}(\mathcal{A}_t(e))$ that e has not been visited by step t can be found by subdividing e with a vertex w of degree 2. That is, $\mathcal{A}_t(e)$ occurs in $\boldsymbol{C}_1$ iff $\mathcal{A}_t(w)$ occurs in the modified graph.

In the expressions below, $\Delta = O(\log n)$ denotes the maximum degree of $G_{n,p}$.

Case (a(1)) of Theorem 3.
The probability that some edge of the 2-core has not been covered by time $t_0 = (1-\epsilon)t_1^*$ is at most (see (51))

$$\sum_{\ell=1}^{10s_0/x} m\exp\left\{-\frac{(1-3\epsilon/2)(\ln n)^2}{2\ell(cx-\ln c)}\right\}+ \\ \sum_{\ell=10s_0/x}^{L} c\Delta n(\ln n)^2(Le^c)^{s_0}\exp\left\{-\ell(cx-\ln c)-\frac{(1-3\epsilon/2)(\ln n)^2}{2\ell(cx-\ln c)}\right\}=o(1).$$

In the above sum we are estimating each $\mathbf{Pr}(\boldsymbol{A}_{t_0}(e))$ by considering a random walk where e alone is split. This accounts for the m replacing n in the first sum.

We next consider edges of the mantle. Every edge of an arborescence must have been covered by the time the last vertex of degree 1 of that arborescence is covered. Conversely, the last vertex of degree 1 of that arborescence is covered at exactly the same step as the unique edge incident with it.

Case (a(2)) of Theorem 3.
The proof is similar to case (a(1)). The probability that some edge of the 2-core has not been covered by time $t_0 = (1-\epsilon)t_1^*$ is at most (see (54))

$$\sum_{\ell=1}^{10s_0/x} m\exp\left\{-\frac{(1-3\epsilon/2)(\ln n)^2}{2\ell(cx-\ln c)}\right\}+\sum_{\ell=10s_0/x}^{L}\Delta(2c)^\ell n^{1-\ell/\omega-\omega/(2\ell)+2s_0/\omega}e^{3\epsilon\omega\ln n/(2\ell)} = o(1).$$

Case (b) of Theorem 3.
The probability that some edge of the 2-core has not been covered by time $t_0 = (1-\epsilon)t_1^*$ is at most (see (58))

$$\sum_{\ell=1}^{L}\Delta\left(n^{1-\alpha\ell+o(1)-2(1+\epsilon/3)\gamma/(\alpha\ell)}+n^{1-\alpha(\ell-1)+o(1)-\sigma_1(1+\epsilon/3)\gamma/(\alpha\ell)}\right)=o(1).$$

Case (c) of Theorem 3.
Comparing with (60) we see that, where $\hat{C}_u$ denotes the expected time to visit all edges of the 2-core,

$$\hat{C}_u-(t_1+1)\overset{O}{\leq}\Delta n^{\delta^+}\sum_{k\geq 2}\frac{(\ln n)^k}{k!}\sum_{s\geq t_1}\exp\left\{-\frac{ks}{(1+o(1))n\ln n}\right\}=o(t_1).$$

This completes the proof of Theorem 3.

4 Cover time of the k-cores, $k \geq 2$

4.1 Cover time of 2-core

We make some minor amendments to the analysis used in the proof of Theorem 2, which we now explain.

Proof of Theorem 4(a).

(i) During the construction of Γ_v in Section 3.3, we require only that paths from v to Γ_v° are $\lceil s_0/2 \rceil$-attached. Case (a) of Lemma 12 is still true with this change i.e. $p_1 \leq 1/(\ln n)^6$ and so $\widehat{R}_v = R_v^*(1+o(1))$.

(ii) We distinguish between vertices of the giant 2-connected block $\boldsymbol{B}_2$ of $\boldsymbol{C}_2$, and vertices in $\boldsymbol{U}$, the pendicles consisting of unicyclic components rooted at $\boldsymbol{B}_2$. We define two sets

$$\begin{aligned} V_1 &= \{v \in \boldsymbol{B}_2 : v \text{ does not lie on a cycle of } \Gamma_v\} \\ V_2 &= \boldsymbol{C}_2 \setminus V_1. \end{aligned}$$

(iii) Suppose that $v \in V_1$ has degree $d = d(v) \geq 2$. The we can find d internally disjoint paths from v to the boundary of Γ_v. Let α be the minimum and β be the maximum lengths of these paths. By deleting edges not on these paths and sub-dividing edges, we may ensure that there is one path of length α and there are $d-1$ paths of length β. This construction only increases the expected number of returns to v. Thus

$$R_v \leq (1+o(1))\frac{d\alpha\beta}{\alpha+(d-1)\beta} \leq (1+o(1))\frac{d\alpha\beta}{\alpha+\beta}. \tag{75}$$

This follows from Lemma 10 and (29) and $b(v) = d(v)$.

(iv) For $v \in V_1$, let $\ell(v) = \alpha + \beta$ and let $D_\ell = \{v \in V_1, \ell(v) = \ell\}$. Thus $|D_\ell| \leq \ell n_\ell$ and $R(v) \leq (1+o(1))d(v)\ell/4$.

Recall that $t_2^* = \frac{cx^2}{16(cx-\ln c)} n(\ln n)^2$, and let $t_2 = t_2^*(1+\epsilon)$. Now $2m \sim cx^2 n$ (see **P0**) and thus,

$$\mathbf{Pr}(\mathcal{A}_{t_2}(v)) \leq 2\exp\left\{-t_2 \frac{d(v)}{2m}\frac{1-o(1)}{R_v}\right\} \leq 2\exp\left\{-\frac{(1+\epsilon/2)(\ln n)^2}{4\ell(cx-\ln c)}\right\}.$$

That t_2 is an upper bound for the expected time to cover V_1 now follows from arguments similar to (50)-(53).

(v) For $v \in V_2$, we have $R_v^* \leq d(v)L$ and so

$$\mathbf{Pr}(\mathcal{A}_{t_2}(v)) \leq 2\exp\left\{-\frac{(1+\epsilon/2)(\ln n)^2}{8L(cx-\ln c)}\right\} = n^{-\delta},$$

for some $\delta > 0$ whereas, by **P4**, $|V_2| = O((\ln n)^{5k_0})$.

For the lower bound, we say that a vertex v of $\boldsymbol{C}_2$ is special, if $d(v) = 2$, v is at the center of a path of length ℓ_0 (or $\ell_0 + 1$ if ℓ_0 is even) attached to the 2-core only at its endpoints, and properties **S2**-**S4** hold. $R_v \geq \ell_0/2 - O(1)$ (see proof of Lemma 13(e)) and $\pi_v \sim 2/(cx^2n)$ and thus $p_v \leq (4 + o(1))/(\ell_0 cx^2 n)$. The proof then follows the structure of the proof of the lower bound in Theorem 2(a).

Proof of Theorem 4(b).

If v is a large vertex of V_1, then by arguments similar to Lemma 13 (c) we find that $R_v = 1 + o(1)$. Let V_1' be the set of small vertices of V_1, and let $v \in V_1'$. When we prune Γ_v as in (iii) above, we find v is in a path PvQ of length $\ell = \ell(v)$, $\ell \geq 2$, in which all of the $\ell - 1$ internal vertices are small. As in (75), we have $R_v^* \leq (d(v)/2)(2\lfloor \ell/2 \rfloor \lceil \ell/2 \rceil / \ell)$. For such a v,

$$\mathbf{Pr}(\mathcal{A}_{t_2}(v)) \leq 2 \exp\left\{ -\frac{(1+\epsilon/2)\gamma \ln n}{\alpha} \frac{\ell}{\lfloor \ell/2 \rfloor \lceil \ell/2 \rceil} \right\}.$$

Let $D_\ell = \{v \in V_1' : \ell(v) = \ell\}$. By **P3**b(ii) there are at most $n_{\ell-1} = n^{1-\alpha(\ell-1)+o(1)}$ such paths, and so

$$\begin{aligned}
\sum_{v \in V_1'} \mathbf{Pr}(\mathcal{A}_{t_2}(v)) &\leq 2 \sum_{\ell=2}^{L} \ell |D_\ell| \exp\left\{ -\frac{(1+\epsilon/2)\ell\gamma \ln n}{\alpha \lfloor \ell/2 \rfloor \lceil \ell/2 \rceil} \right\} \\
&\leq \sum_{\ell=2}^{L} n^{1-\alpha(\ell-1)+o(1)-(1+\epsilon/3)\gamma\ell/(\alpha\lfloor \ell/2 \rfloor \lceil \ell/2 \rceil)} \\
&= o(1),
\end{aligned}$$

provided we choose γ as in (3).

Proof of Theorem 4(c). This is a straightforward imitation of (60)-(61).

4.2 Cover time of the k-core, $k \geq 3$

For $\boldsymbol{C}_k, k \geq 3$, the likely presence of many large induced trees of degree k will determine both upper and lower bounds for the cover time, by applying the methods of [8], [9].

We first summarize some properties of the k-core. Let $f(x) = f_k(c, x)$ where

$$f_k(c, x) = 1 - e^{-cx} \left(1 + cx + \cdots + \frac{(cx)^{k-2}}{(k-2)!} \right).$$

The threshold for the appearance of the k-core is the minimum value c_k, of c such that a positive solution of $x = f(x)$ exists. For $c > c_k$, let x_k be the largest solution in $(0, 1)$ of $x = f(x)$.

Then $|E(\boldsymbol{C}_k)| \sim ncx_k^2/2$, $|V(\boldsymbol{C}_k)| \sim nf_{k-1}(c, x_k)$ and n_k, the number of vertices of degree k satisfies $n_k \sim n(cx_k)^k e^{-cx_k}/k!$. These estimates are accurate to within $O(n^{2/3}(\ln n)^{7/3})$ **whp** (see e.g. [10], [20]).

Let $v \in \boldsymbol{C}_k$, and G_v be the sub-graph of $\boldsymbol{C}_k$ rooted at v, of depth r. We say that v is *locally regular*, if G_v is a tree, and all vertices of G_v have degree k in $\boldsymbol{C}_k$. Following [8], [9] it is these locally regular vertices v which have the largest value of $R_v \sim (k-1)/(k-2)$, and **whp** there are n^δ of such vertices are still not covered at $T_k^*(1-\epsilon)$, for suitable $\delta, \epsilon \to 0$. Thus, to prove Theorem 5 it will be enough to establish the existence of a large set of such locally regular vertices **whp**.

Let Y count locally regular sub-graphs of the k-core, and v be locally regular. Working outwards from the root vertex v, with $l_0 = 1$, let l_i be the number of vertices at level i of G_v. For $i \geq 1$, $l_i = k(k-1)^{i-1}$. Let $r = \ln\ln n$ and $a = l_1 + \cdots + l_r = k((k-1)^r - 1)/(k-2) = (\ln n)^{O(1)}$. Thus $|V(G_v)| = a + 1$, and $(a+1)k = 2a + (k-1)l_r$.

We model the k-core via a configuration model. Condition on the degree sequence of $\boldsymbol{C}_k$ and assume that it has $\nu \sim nf_{k-1}(c, x)$ vertices and $\mu \sim cx^2 n/2$ edges. Let W be a set of size 2μ and let $W_1, \ldots, W_\nu$ partition W where $|W_i|$ equals the degree of the ith vertex of $\boldsymbol{C}_k$. A random partition F of W into μ pairs defines a random multigraph $\gamma(F)$ in the usual way, Bollobás [5]. Within this model,

$$\begin{aligned}\mathbf{E}(Y) &= \binom{n_k}{a+1}\binom{a+1}{l_0\ l_1\ l_2 \cdots l_r} k^{l_1} l_1! \cdots k^{l_r} l_r! \\ &\quad \times \binom{2\mu - (a+1)k}{(k-1)l_r}((k-1)l_r)! \frac{\Phi(2\mu - (2a + 2(k-1)l_r))}{\Phi(2\mu)} \\ &= (n_k)_{(a+1)} \left(\frac{k}{2}\right)^a 2^{(a+1)k} \frac{(\mu)_{(a+1)k-a}}{(2\mu)_{(a+1)k}}\end{aligned}$$

where $\Phi(2m) = (2m)!/(m!2^m)$, and

$$\frac{(\mu)_{(a+1)k-a}}{(2\mu)_{(a+1)k}} \sim \frac{1}{2^{(a+1)k}\mu^a}.$$

So

$$\begin{aligned}\mathbf{E}(Y) &\sim (n_k)^{a+1}\left(\frac{k}{cx_k^2 n}\right)^a \\ &\sim n\,\frac{(cx_k)^k}{k!} e^{-cx_k} \left(\frac{e^{-cx_k}}{x_k}\frac{(cx_k)^{k-1}}{(k-1)!}\right)^a \\ &= n^{1-o(1)}.\end{aligned}$$

We prove concentration of Y by a standard martingale argument. We can construct $\gamma(F)$ by taking a random permutation of W and pairing up adjacent elements. If we swap two elements of the permutation then we affect two edges $\gamma(F)$ and then we can change Y by at most $\omega = O(k^r) = O((\ln n)^{O(1)})$ since there are at most this many vertices at distance r from any fixed edge. It follows that

$$\mathbf{Pr}(|Y - \mathbf{E}(Y)| \geq A\omega\sqrt{n \ln n}) \leq \exp\left(-\frac{A^2\omega^2 n \ln n}{2\omega^2 cn}\right) = O\left(n^{-A^2/2c}\right).$$

This immediately implies the same order of concentration for the k-core of $G_{n,p}$ itself.

Upper bound for cover time. Recall the definition of $V_k(v)$ in (27), which defines a subgraph T_v, the local neighbourhood of v to depth L. For simplicity we choose $L = \ln \ln n$. If v is locally regular then $R_v = (1 + o(1))(k-1)/k-2)$. For any other tree T_v, we can prune to a locally regular sub-tree. This increases R_v to $(1 + o(1))(k-1)/k-2)$ by Raleigh's Monotonicity Theorem. Thus for vertices v with T_v a tree we have

$$p_v \leq \frac{k}{cx_k^2 n}\frac{k-2}{k-1}(1 + o(1)).$$

Recalling that $t_k^* = cx_k^2(k-1)/(k(k-2))n \ln n$, we see that

$$ne^{-p_v t_k^*(1+\epsilon)} = o(1/\ln n),$$

for $\epsilon = o(1)$ but not too small. Thus from (45), $t_k^*(1 + o(1))$ is an upper bound for the time taken to cover all locally tree-like vertices where T_v is a tree.

For local neighbourhoods T_v with a cycle, pruning an edge could introduce two vertices u, u' of degree $k-1$. In the worst case u, u' are now of degree 2. However every other vertex on a path from v to the boundary has degree at least 3, so $R_v = O(\ln \ln n)$ by comparison with a biased walk on a path (see Lemma 12). As there are $O((\ln n)^{\ln \ln n})$ vertices v for which T_v contains a cycle (see **P4**), and

$$O((\ln n)^{\ln \ln n})e^{-\frac{k}{cx_k^2 n}\frac{t_k^*}{O(\ln \ln n)}} = o(1/\ln n),$$

it follows from (45) that t_k^* is also an upper bound for the time to cover all locally non-tree-like vertices.

Lower bound for cover time. There are at most $O(k^{2r} \ln n)$ locally regular vertices within distance $2r$ of a given locally regular vertex v. Let S be a maximal set of locally regular vertices at distance at least $2r + 1$ from each other. Choosing $r = \ln \ln n$, by the previous discussion we see that $|S| \times O(k^{2r} \ln n) \geq n^{1-o(1)}$ i.e. $|S| = n^{1-o(1)}$. The lower bound now follows from the arguments already used previously.

Final Remark It is well known that if a k-core, $\boldsymbol{C}_k, k \geq 1$ has ν vertices and μ edges, then it has the same distribution as a random graph with this number of vertices and edges, but conditioned to have minimum degree at least k. Thus with some extra effort we could have couched our results within this model. This would have lengthened the paper. We do not claim to have checked the calculations and we do not claim that we can deduce results for the conditional model from what we have proved. We may be misguided, but we do not forsee any significant difficulties in carrying out such a project.

References

[1] D. Aldous and J. Fill, Reversible Markov Chains and Random Walks on Graphs, http://stat-www.berkeley.edu/pub/users/aldous/RWG/book.html.

[2] 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.

[3] I. Benjamini, G. Kozma and N. Wormald, The mixing time of the giant component of a random graph, to appear.

[4] B. Bollobás, *Random Graphs,* 2nd edition, CUP (2001).

[5] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European Journal on Combinatorics* 1 (1980) 311-316.

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

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

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

[9] C. Cooper and A. M. Frieze, The cover time of the preferential attachment graph, *Journal of Combinatorial Theory Series B*, 97 (2007) 269-290.

[10] C. Cooper, The cores of random hypergraphs with a given degree sequence, *Random Structures and Algorithms*, 25 (2004) 353-375.

[11] P. G. Doyle and J. L. Snell *Random Walks and Electrical Networks.* Carus Mathematical Monograph 22, AMA (1984).

[12] P. Erdős and A. Rényi, The evolution of random graphs, *Publ. Math. Inst. Hungar. Acad. Sci.* 5 (1960) 17-61.

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

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

[15] W. Feller, *An Introduction to Probability Theory, Volume I,* (Second edition) Wiley (1960).

[16] N. Fountoulakis and B. Reed, The evolution of the mixing rate, to appear.

[17] 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

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

[19] B. Pittel, On tree census and the giant component in sparse random graphs, *Random Structures and Algorithms*, 1.3 (1990) 311-342.

[20] B. Pittel, J. Spencer and N. Wormald, Sudden emergence of a giant k-core in a random graph, *Journal of Combinatorial Theory Series B*, 67 (1996) 111-151.

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

[22] D. Zuckerman, On the time to traverse all edges of a graph, *Information Processing Letters*, 38 (1991) 335-337.

A Proof of Lemma 8

(a) For any c the equation $x = 1 - e^{-cx}$ has the solution $x = 0$. Using $x = 1 - e^{-cx}$, provided $0 < x < 1$, we obtain

$$\begin{aligned} c &= \frac{1}{x}\ln\frac{1}{1-x} \\ &= 1 + \frac{x}{2} + \cdots + \frac{x^j}{j+1} + \cdots. \end{aligned} \tag{76}$$

The RHS of (76) increases monotonically from 1 to ∞ as $x \uparrow 1$. Thus for any $c > 1$ the equation $x = 1 - e^{-cx}$ has a unique solution in $(0, 1)$. This solution satisfies

$$\begin{aligned} ce^{-cx} &= \frac{1-x}{x}\ln\frac{1}{1-x} \\ &= \frac{1-x}{x}\left(x + \frac{x^2}{2} + \cdots + \frac{x^j}{j} + \cdots\right) \\ &= 1 - \frac{x}{2} - \frac{x^2}{6} - \cdots - \frac{x^j}{j(j+1)} - \cdots \qquad (77) \\ &< 1. \qquad (78) \end{aligned}$$

Thus

$$\ln(ce^{-cx}) = \ln c - cx < \ln(1 - x/2). \tag{79}$$

From (76) we see that

$$\begin{aligned} c &= 1 + x\left(\frac{1}{2} + \frac{x}{3} + \cdots + \frac{x^{j-1}}{j+1} + \cdots\right) \\ &\leq 1 + x\left(\frac{1}{2} + c - 1\right), \end{aligned}$$

and we see that $x \geq 2(c-1)/(2c-1)$. Combining this with (79) gives $cx - \ln c > \ln(2 - 1/c)$.

(b) We write $\theta(c) = (2 - x)/(1 - g(c))$ where $g(c) = (\ln c)/cx$. Since $x \to 1$ as $c \to \infty$ we see immediately $g(c) \to 0$ and $\theta(c) \to 1$ as $c \to \infty$. We next show that $g(c)$ is monotone decreasing with increasing c and this will show that $\theta(c)$ is monotone decreasing.

From $1 - x - e^{-cx} = 0$ we find that $dx/dc = x/(e^{cx} - c)$. Thus

$$\begin{aligned} \frac{dg(c)}{dc} &= \frac{1}{c^2 x}\left(1 - \ln c\left(1 + \frac{c}{x}\frac{dx}{dc}\right)\right) \\ &= -\frac{1}{c^2 x(1 - ce^{-cx})}(\ln c + ce^{-cx} - 1). \end{aligned}$$

Let $h(c) = \ln c + ce^{-cx} - 1$. We claim that for $c > 1$, $h(c)$ is monotone increasing and hence positive, and so (see (78)) $dg/dc < 0$. From (77) we have

$$\frac{dh(x)}{dx} = \frac{1}{c}\frac{dc}{dx} - \left(\frac{1}{2} + \frac{x}{3} + \cdots + \frac{x^{j-1}}{j+1} + \cdots\right),$$

and from (76)

$$\frac{1}{c}\frac{dc}{dx} = \frac{\frac{1}{2} + \frac{2x}{3} + \cdots + \frac{jx^{j-1}}{j+1} + \cdots}{1 + \frac{x}{2} + \cdots + \frac{x^j}{j+1} + \cdots}.$$

The result we require is equivalent to

$$\frac{1}{2} + \frac{2x}{3} + \cdots + \frac{(j+1)x^j}{j+2} + \cdots > \left(1 + \frac{x}{2} + \cdots + \frac{x^j}{j+1} + \cdots\right) \times \left(\frac{1}{2} + \frac{x}{3} + \cdots + \frac{x^j}{j+2} + \cdots\right).$$

The coefficient of x^k on the LHS is $(k+1)/(k+2)$ and on the RHS it is

$$\sum_{j=0}^{k} \frac{1}{j+1} \cdot \frac{1}{k-j+2},$$

which is the sum of $k+1$ terms of which the first is $1/(k+2)$ and the others are strictly less than this value.

Finally, it follows from (77) that $cx - \ln c = x/2 + O(x^2)$ as $x \to 0$ and so
$\theta(c) = c(2-x)/(1/2 + O(x))$ as $x \to 0$ giving that $\theta(c) \to 1/4$ as $c \to 1$. □

B Proof of Lemma 9

Properties **P0**,**P1**: Proofs of our assertions can be found for example in [12], [19] and [20].

Property **P2**: This has been proved by Benjamini, Kozma and Wormald [3] and Fountoulakis and Reed [16].

Property **P3**a: (i) Assume that $10x^{-1}s_0 \leq \ell \leq L$.

Case (a1): $\sigma_1 = 1, s_0 = 60 \ln\ln n, L = \frac{\ln n + 2s_0 \ln\ln n + cs_0}{cx - \ln c}$.
The expected number of paths P of length ℓ which are not s_0-attached is at most

$$\binom{n}{\ell+1}(\ell+1)!\left(O(e^{-n^{1/4}}) + (1+o(1))p^\ell \sum_{t \leq s_0} \binom{\ell}{t} x^t (1-x)^{\ell-t}\right). \tag{80}$$

Explanation We choose a set S of $\ell+1$ vertices and order them $s_0, s_1, \ldots, s_\ell$. We then expose the edges of $V \setminus S$. The $O(e^{-n^{1/4}})$ term is the probability that there is no giant of size ξn ($\xi = x + O(n^{-1/4})$) produced. We now expose the edges from S to $V \setminus S$. The probability that some fixed set $T \subseteq S \setminus \{s_0\}$ of size t contains the only vertices adjacent to at least two vertices of the exposed giant K is

$$(1-(1-p)^{\xi n})^t (1-p)^{\xi n(\ell-t)} = x^t(1-x)^{\ell-t}(1+O(n^{-1/4}\ln n)).$$

We then multiply by p^ℓ for the probability of the existence of the path, which places the sub-path induced by T in $\boldsymbol{C}_2$.

The expression in (80) can be bounded by

$$2cn(ce^{-cx})^\ell \left(\frac{\ell e x}{s_0(1-x)}\right)^{s_0} \leq 2cn(ce^{-cx})^\ell (\ell e^c)^{s_0} = O(n_\ell/(\ln n)^2) \tag{81}$$

and the claim follows from the Markov inequality.

Case (a2): $\sigma_1 = \frac{\ln n}{\omega(\ln\ln n)^{10}}, s_0 = 50, L = 2\omega$.
We replace (80) by

$$\binom{n}{\ell+1}(\ell+1)!\left(O(e^{-n^{1/4}}) + (1+o(1))p^\ell \sum_{t \leq s_0} \binom{\ell}{t} \left(\sum_{i=0}^{\sigma_1} \binom{\xi n}{i} p^i (1-p)^{\xi n - i}\right)^{\ell-t}\right)$$
$$\leq n(2c)^\ell (c^{\sigma_1} e^{-cx})^{\ell - s_0} \tag{82}$$

and the claim follows from the Markov inequality.

(ii) If $\ell = L$ then the middle term of (81) and the RHS of (82) are both $o(1/\ln n)$ and so **whp** all paths of length L are s_0-attached.

Property **P3**b: $\sigma_1 = \frac{\ln n}{(\ln\ln n)^{10}}, \Lambda = \lfloor \alpha^{-1} \rfloor$.
(i) The probability that there exists S, $|S| \leq \ln\ln n$ such that S induces a connected sub-graph and also contains $\Lambda + 1$ small vertices is at most

$$\sum_{k=\Lambda+1}^{\ln\ln n} \binom{n}{k} k^{k-2} p^{k-1} \binom{k}{\Lambda+1} \left(\sum_{i=0}^{\sigma_1} \binom{n - \ln\ln n}{i} p^i (1-p)^{n-\ln\ln n - i} \right)^{\Lambda+1} \leq$$
$$n(\ln\ln n)^{O(\ln\ln n)} n^{-(\alpha - o(1))(\Lambda+1)} = o(1).$$

(ii) The expected number of paths with all but at most one vertex small can be bounded by

$$\binom{n}{\ell+1} (\ell+1)! p^\ell (\ell+1) \left(\sum_{i=0}^{\sigma_1} \binom{n - \ln\ln n}{i} p^i (1-p)^{n-\ln\ln n-i} \right)^\ell$$
$$\leq n(\ell+1) c^\ell n^{-(\alpha-o(1))\ell} = n^{1-\alpha\ell+o(1)}. \quad (83)$$

The Markov inequality proves the bound is valid **whp** and when $\ell > \Lambda$ the RHS of (83) is $o(1)$.

(iii) The expected number of trees of size $k = O(\ln n)$ in the mantle can be bounded by

$$\binom{n}{k} k^{k-2} p^{k-1} \left(O(e^{-n^{1/4}}) + (k-1)(1-p)^{k(xn-k)} \right) \leq kc^{k-1} n^{1-(\alpha-o(1))k}$$

and this is $o(1)$ for $k \geq \Lambda + 1$.

(iv) The expected number of cycles of size $\leq (\ln\ln n)^2$ which contain a small vertex is

$$\sum_{k=3}^{(\ln\ln n)^2} n^k p^k n^{-\alpha+o(1)} = o(1).$$

Property **P3**c. The proof follows that of **P3**b.

Property **P4**: $k_0 = A_0 \ln\ln n$.
Whp the maximum degree is $O(\ln n)$ and the expected number of cycles of length at most $2k_0$ is $O(k_0 c^{2k_0})$. Thus with probability $1 - O(1/\ln n)$ there are at most $O(k_0 c^{2k_0} \ln n)$ such cycles and most $O(k_0^2 c^{2k_0} (\ln n)^{2k_0+2})$ vertices on or within distance $2k_0$ of such cycles.

Property **P5**a (i) The expected number of paths of length at most k_0^3, with at least 2 shortcuts is at most

$$\sum_{\ell=4}^{k_0^3} \binom{n}{\ell} \ell! \binom{\ell}{2}^2 p^{\ell+1} = O\left(\frac{c^{k_0^3+1} k_0^{12}}{n} \right) = o(1).$$

(ii) The expected number η of paths contradicting **P5**a(2) satisfies

$$\begin{aligned}\eta &\leq 2\binom{n}{k_0^2}(k_0^2)!p^{k_0^2}\binom{k_0^2}{3s_0}(1-x+O(n^{-1/4}))^{k_0^2-3s_0}\sum_{k=3}^{2k_0}k\binom{n}{k}k!p^k \\ &= 5k_0(ce^{-cx})^{k_0^2}c^{2k_0+1}\left(\frac{k_0^2e^{1+cx}}{3s_0}\right)^{3s_0} \\ &\leq 5k_0e^{-k_0^2x/2}c^{2k_0+1}\left(\frac{k_0^2e^{1+cx}}{3s_0}\right)^{3s_0} \qquad \textit{after using } (77) \\ &= o(1).\end{aligned}$$

Part (iii) follows directly from (a), (b).

Property **P5**b. (i) The probability that there is a pair of small cycles that are close together is at most the probability that there is a set of $\leq 3k$ vertices spanning $\geq k+1$ edges where $3 \leq k \leq \ln\ln n$. And so this is at most

$$\sum_{k=3}^{3\ln\ln n}\binom{n}{k}\binom{\binom{k}{2}}{k+1}p^{k+1} \leq n^{-1}(\ln n)^{3\ln\ln n+6} = o(1).$$

(ii) The expected number of triangles is $O((\ln n)^3)$.

Property **P6**a: The expected number of cycles violating the condition is at most

$$\begin{aligned}&2\sum_{\ell_1,\ell_2\geq k_0}n^{\ell_1+\ell_2}p^{\ell_1+\ell_2}\binom{\ell_1}{s_0}x^{s_0}(1-x)^{\ell_1-s_0}\binom{\ell_2}{s_0}x^{s_0}(1-x)^{\ell_2-s_0} \\ &\leq 2\sum_{\ell_1,\ell_2\geq k_0}\left(\frac{\ell_1 ex}{s_0(1-x)}\right)^{s_0}(ce^{-cx})^{\ell_1}\left(\frac{\ell_2 ex}{s_0}\right)^{s_0(1-x)}(ce^{-cx})^{\ell_2} \\ &= 2\left(\frac{ex}{s_0(1-x)}\right)^{2s_0}\left(\sum_{\ell\geq k_0}\ell^{s_0}(ce^{-cx})^{\ell}\right)^2.\end{aligned} \tag{84}$$

Case a(1):
If $u_\ell = \ell^{s_0}(ce^{-cx})^\ell$ then

$$u_{\ell+1}/u_\ell \leq \exp\{s_0/k_0 - (cx-\ln c)\} \leq \exp\{-(cx-\ln c)/2\} < 1.$$

So, if $\zeta = \frac{k_0ex(ce^{-cx})^{k_0/s_0}}{s_0(1-x)}$ then the RHS of (84) is $O(\zeta^{2s_0})$.

Suppose first that $x \leq 1/2$. Then we assume that A_0 is large enough so that if $B_0 = A_0/60$ then $\zeta \leq 2B_0e(ce^{-cx})^{B_0} < 1/2$ and then we have the RHS of (84) equal to $o(1)$.

If $x \geq 1/2$ then $\zeta = \frac{k_0e^{1+cx}x(ce^{-cx})^{k_0/s_0}}{s_0} \leq B_0e^{c+1}(ce^{-c/2})^{B_0} \leq 1/2$ for large enough A_0 and we also have the RHS of (84) equal to $o(1)$.

Case a(2): The RHS of (84) is at most

$$3e^{100c}(A_0 \ln\ln n)^{100} e^{-2ck_0(1-o(1))} = o(1).$$

Property **P7**a

Let S_1 be the set of vertices satisfying **S1**, **S2**, **S3**. Let S_2 be the set of vertices $v \in S_1$ with a path rooted to the 2-core at $w(v)$ which extends to a path of length k_0 in the 2-core, which is not s_0 attached. Let S be the set of special vertices, then using **P4** we see that $|S| = |S_1| - |S_2| - O((\ln n)^{3k_0+2})$. Let $m = nx + O(n^{3/4})$. Thus

$$\begin{aligned} &\mathbf{E}(|S_1|) = o(1)+ \\ &\binom{n}{\ell_0+1}(\ell_0+1)!p^{\ell_0}(1-p)^{n-2}(1-p)^{m(\ell_0-1)+\binom{\ell_0}{2}}(1-p)^{n-m}\sum_{i\geq 2}\binom{m}{i}p^i(1-p)^{m-i} \qquad (85)\\ &= \Theta(n(ce^{-cx})^{\ell_0}e^{-c(2-x)}) \\ &= \Theta(n^{1/2-o(1)}). \end{aligned}$$

Explanation of (85): We choose a vertex v and an ordered set of vertices $A = (v, v_1, v_2, \ldots, v_{\ell_0+1} = w(v))$. We then expose the edges of $H = G - A$ to obtain a graph distributed as $G_{n-\ell_0-1,p}$. With probability $1 - O(e^{-n^{1/4}})$, G will have a giant component $\boldsymbol{C}_1'$ of size $m \sim xn$. The $(1-p)^{n-2}$ is for **S1**. The final vertex $w(v)$ has no edges to the small components of H (for **S3**), or to the vertices of A. The vertex $w(v)$ is in the 2-core, so has at least 2 edges to $\boldsymbol{C}_1'$ the giant of H.

If we repeat the calculation with disjoint pairs of ℓ_0-sets we find that $\mathbf{E}(|S_1|^2) \sim (\mathbf{E}(S_1|)^2$. Thus the Chebyshev inequality can be used to show that **whp** $|S_1| \sim \mathbf{E}(|S_1|) = \Theta(n^{1/2-o(1)})$.

Coming now to S_2, we have

Case (a1):

$$\begin{aligned} \mathbf{E}(|S_2|) \;\leq\; &\binom{n}{\ell_0+1}(\ell_0+1)!p^{\ell_0}(1-p)^{n-2}(1-p)^{m(\ell_0-1)+\binom{\ell_0}{2}} &(86)\\ &\times(1-p)^{n-m}\sum_{i\geq 2}\binom{m}{i}p^i(1-p)^{m-i} &(87)\\ &\times\binom{n-\ell_0}{k_0}k_0!p^{k_0}\sum_{t<s_0}\binom{k_0}{t}x^t(1-x)^{k_0-t}. &(88) \end{aligned}$$

Explanation of (86)–(88): This is similar to (85). We first choose two ordered sets A, B of size ℓ_0+1, k_0 respectively, and examine the giant component $\boldsymbol{C}_1'$ of $H = G - A - B$. The lines (86), (87) establish the properties **S1**, **S2** and **S3** and attach at least two edges from $w(v)$ to

$\boldsymbol{C}_1'$. The line (88) considers the path wQx_{k_0} rooted at $w(v)$ on the vertices of $B = (x_1, ..., x_{k_0})$. The last term upper bounds the probability that such a path is not s_0-attached. We see that

$$\mathbf{E}(|S_2|) \overset{O}{\leq} \mathbf{E}(|S_1|) \left(\frac{k_0 ex}{s_0}\right)^{s_0} (ce^{-cx})^{k_0} = O\left(\frac{\mathbf{E}(|S_1|)}{(\ln n)^2}\right)$$

Case (a2): Here we replace (88) by

$$\binom{n-\ell_0}{k_0} k_0! p^{k_0} \sum_{t<s_0} \binom{k_0}{t} \left(\sum_{i=0}^{\sigma_1-1} \binom{\xi n}{i} p^i (1-p)^{\xi n - i}\right)^{k_0-t} \leq (2c)^{k_0} n^{-k_0/(2\omega)} = o(1/(\ln n)^2)$$

Applying the Markov inequality, with probability $1 - O(1/\ln n)$ we have that $|S_2| \leq |S_1|/\ln n$.

We combine this with our bounds on $|S_1|$ and **P4** to obtain that **whp** $|S| = \Theta(n^{1/2-o(1)})$.

Finally, we note that the expected number of pairs v_1, v_2 of vertices of degree one, both at distance ℓ_0 from $\boldsymbol{C}_2$ such that $w(v_1)$, $w(v_2)$ are at most $2k_0$ apart is bounded by

$$O(n^2 e^{-n^{1/4}})) + \left(n^{\ell_0+1} p^{\ell_0} (1-p)^{nx\ell_0(1+O(n^{3/4}))}\right)^2 \sum_{i=0}^{2k_0-1} n^i p^{i+1} = (\ln n)^{O(\ln\ln n)}.$$

Thus **whp** there are $(\ln n)^{O(\ln\ln n)}$ such pairs. Removing them from S yields the required set of $\Theta(n^{1/2-o(1)})$ special vertices satisfying **P7**a.

Property **P7**b. Let S_1 be the set of vertices satisfying **S1**, **S2**, **S3**. Then if $m = |\boldsymbol{C}_2| = n - O(n^{-1/4})$,

$$\mathbf{E}(|S_1|) = O(e^{-n^{1/4}}) + \binom{n}{\ell_0+1} (\ell_0+1)! p^{\ell_0} (1-p)^{n-2} (1-p)^{m(\ell_0-1)+\binom{\ell_0}{2}} (1-p)^{n-m} \sum_{i \geq \sigma_1} \binom{m}{i} p^i (1-p)^{m-i} = n^{1-\ell_0\alpha-o(1)}.$$

We use Chebyshev to show that $|S_1| \sim \mathbf{E}(|S_1|)$ **whp**. Since the maximum degree of $G_{n,p}$ is $\leq 10\ln n$ **whp**, each $v \in S_1$ is within distance $\ln\ln n$ of $\leq (10\ln n)^{\ln\ln n} = n^{o(1)}$ other members of S_1 and we can finish the argument via **P3**b(i).

Property **P7**c. This involves a straightforward second moment calculation. $\square$