The cover time of random geometric graphs

Colin Cooper[*] Alan Frieze[†]

February 17, 2009

Abstract

We study the cover time of random geometric graphs. Let $I(d) = [0,1]^d$ denote the unit torus in d dimensions. Let $D(x,r)$ denote the ball (disc) of radius r. Let Υ_d be the volume of the unit ball $D(0,1)$ in d dimensions. A random geometric graph $G = G(d,r,n)$ in d dimensions is defined as follows: Sample n points V independently and uniformly at random from $I(d)$. For each point x draw a ball $D(x,r)$ of radius r about x. The vertex set $V(G) = V$ and the edge set $E(G) = \{\{v,w\} : w \neq v, w \in D(v,r)\}$. Let $G(d,r,n)$, $d \geq 3$ be a random geometric graph. Let $c > 1$ be constant, and let $r = (c \log n/(\Upsilon_d n))^{1/d}$. Then **whp**

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

1 Introduction

Let $G = (V,E)$ be a connected graph with $|V| = n$ vertices, and $|E| = m$ edges. 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 $(1-o(1))n\log n \leq C_G \leq (1+o(1))\frac{4}{27}n^3$. 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

[*]Department of Computer Science, King's College, University of London, London WC2R 2LS, UK
[†]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA15213, USA. Supported in part by NSF grant CCF0502793.

length $n/3$ joined to a clique of size $2n/3$ gives the asymptotic upper bound for the cover time.

A few words on notation. Results on random graphs are always asymptotic in n, the size of the vertex set. The notation $A_n \sim B_n$ means that $\lim_{n\to\infty} A_n/B_n = 1$, and **whp** (with high probability) means with probability tending to 1 as $n \to \infty$. Poly-log factors are suppressed in $\tilde{O}, \tilde{\Omega}$.

In a series of papers, we have studied the cover time of various models of a random graph. These results can be summarized as follows:

- [8] If $p = c \log n/n$ and $c > 1$ then **whp** $C_{G_{n,p}} \sim c \log\left(\frac{c}{c-1}\right) n \log n$.
- [9] Let $c > 1$ and let x denote the solution in $(0, 1)$ of $x = 1 - e^{-cx}$. Let X_g be the giant component of $G_{n,p}$, $p = c/n$. Then **whp** $C_{X_g} \sim \frac{cx(2-x)}{4(cx-\log c)} n(\log n)^2$.
- [6] Let $G_{n,r}$ denote a random r-regular graph on vertex set $[n]$ with $r \geq 3$ then **whp** $C_{G_{n,r}} \sim \frac{r-1}{r-2} n \log n$.
- [7] Let $G_m(n)$ denote a *preferential attachment graph* of average degree $2m$ then **whp** $C_{G_m} \sim \frac{2m}{m-1} n \log n$.
- [10] Let $D_{n,p}$ denote a random *digraph* with independent edge probability p). If $p = c \log n/n$ and $c > 1$ then **whp** $C_{D_{n,p}} \sim c \log\left(\frac{c}{c-1}\right) n \log n$.

Let I denote the unit interval $[0, 1]$ and let $I(d) = [0, 1]^d$ denote the unit torus in d dimensions. We use the torus for convenience, to avoid boundary effects. Let $D(x, r)$ denote the ball (disc) of radius r, and thus

$$D(x, r) = \left\{ y \in I(d) : \sum_{i=1}^{d} \min\left\{ |x_i - y_i|^2, |x_i - (1 + y_i)|^2 \right\} \leq r^2 \right\}.$$

Let Υ_d be the volume of the unit ball $D(0, 1)$ in d dimensions. Thus

$$\Upsilon_d = (\pi^{d/2})/\Gamma(d/2 + 1) = \begin{cases} \frac{\pi^k}{k!} & d = 2k, \text{ even} \\ \frac{2^d k! \pi^k}{d!} & d = 2k + 1, \text{ odd} \end{cases}$$

A random geometric graph $G = G(d, r, n)$ in d dimensions is defined as follows: Sample n points V independently and uniformly at random from $I(d)$. For each point x draw a ball $D(x, r)$ of radius r about x. The vertex set $V(G) = V$ and the edge set $E(G) = \{\{v, w\} : w \neq v, w \in D(v, r)\}$

Geometric graphs are widely used as models of ad-hoc wireless networks [16], [17], [22] in which each transmitter has transmission radius r and can only communicate with other transmitters within that radius. In the simplest model of a random geometric graph, the n points representing transmitters, are distributed uniformly at random (uar) in the unit square. Any other point v within the circle radius r centered at a transmitter u, is joined to u by an edge.

If $r \geq \sqrt{c \log n/(\pi n)}$, $c > 1$, then the graph $G(d, r, n)$ is connected **whp** [16, 19], and the cover time of G is well defined. Avin and Ercal [3] consider the cover time of geometric graphs in the case $d = 2$, and prove the following theorem.

Theorem 1. *If* $G = G(2, r, n)$ *and* $r^2 > \frac{8 \log n}{n}$ *then* **whp**

$$C_G = \Theta(n \log n).$$

They also indicate that this result can be generalized to $d \geq 3$. In this paper we consider $d \geq 3$ and replace $\Theta(n \log n)$ by an asymptotically correct constant:

Theorem 2. *Let* $G(d, r, n)$, $d \geq 3$ *be a random geometric graph. Let* $c > 1$ *be constant, and let* $r = \left(\frac{c \log n}{\Upsilon_d n}\right)^{1/d}$. *Then* **whp**

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

As c increases, the RHS of (1) is asymptotic to $n \log n$. It will be clear that we can allow $c \to \infty$ in our analysis and obtain this estimate rigorously. We find it convenient however just to deal with the case of c constant.

Structure of the paper

To prove Theorem 2, we establish bounds on the cover time using the method of [6]. In Section 3 we state a general lemma, the first visit time lemma, on which our results are based. In the next few sections we estimate various quantities needed for this lemma. Then in Section 6, we establish upper and lower bounds on the cover time of G.

2 Some properties of G

The degree, $d(v)$, of vertex $v \in V$ is binomially distributed as $Bin(n-1, p)$, where $p = c \log n/n$ is the volume of $D(v, r)$. It is a simple matter to show using the Chebychev inequality that the number of edges m of G satisfies

$$m \sim \frac{1}{2} cn \log n. \tag{2}$$

We next give some crude bounds on vertex degree.

Lemma 3. *For $c > 1$ there exists a constant $c_0 > 0$ such that* **whp**

$$c_0 \log n \leq d(v) \leq \Delta_0 = (c + 10) \log n \qquad \text{for all } v \in V.$$

Proof For c_0 sufficiently small,

$$\begin{aligned}
\mathbf{Pr}(\exists v: \ d(v) \leq c_0 \log n) &\leq n\mathbf{Pr}(Bin(n-1, p) \leq c_0 \log n) \\
&= n \sum_{k=0}^{c_0 \log n} \binom{n-1}{k} p^k (1-p)^{n-1-k} \\
&\leq 2n \left(\frac{ne}{c_0 \log n} \right)^{c_0 \log n} \left(\frac{c \log n}{n} \right)^{c_0 \log n} n^{-c+o(1)} \\
&= n^{-(c-1-c_0 \log(ce/c_0)-o(1))} \\
&= o(1).
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\mathbf{Pr}(\exists v: \ d(v) \geq \Delta_0) &\leq n\mathbf{Pr}(Bin(n-1, c \log n/n) \geq \Delta_0) \\
&\leq n \sum_{k \geq \Delta_0} \binom{n}{k} \left(\frac{c \log n}{n} \right)^k \left(1 - \frac{c \log n}{n} \right)^{n-k} \\
&\leq 2n \binom{n}{\Delta_0} \left(\frac{c \log n}{n} \right)^{\Delta_0} \left(1 - \frac{c \log n}{n} \right)^{n-\Delta_0} \\
&\leq 3n \left(\frac{ne}{\Delta_0} \right)^{\Delta_0} \left(\frac{c \log n}{n} \right)^{\Delta_0} n^{-c} \\
&= 3n^{1-c-(c+10)\log((c+10)/(ce))} \\
&= o(1).
\end{aligned}$$

□

Let $D(k)$ denote the number of vertices v with $d(v) = k$ in G, and let $\overline{D}(k)$ be the expected number. Thus

$$\overline{D}(k) = n\binom{n-1}{k} p^k (1-p)^{n-1-k} \leq \frac{2}{n^{c-1}} \left(\frac{nep}{k} \right)^k.$$

Let

$$I_c = [c_0 \log n, \Delta_0],$$

where $\Delta_0 = (c + 10)\log n$. The previous lemma shows that **whp** all vertex degrees lie in I_c. The next lemma gives bounds the value of $D(k)$ in terms of $\overline{D}(k)$ for various ranges of $k \in I_c$.

Lemma 4. *Let*

$$\begin{aligned} K_0 &= \{k \in I_c : \overline{D}(k) \leq (\log n)^{-2}\}. \\ K_1 &= \{k \in I_c : (\log n)^{-2} \leq \overline{D}(k) \leq (\log n)^2\}. \\ K_2 &= I_c \setminus (K_0 \cup K_1). \end{aligned}$$

Then, **whp**

(a) *For* $k \in K_0$, $D(k) = 0$.
For $k \in K_1$, $D(k) \leq (\log n)^4$, *and*

$$\min\{k \in K_1\} \geq (\log n)^{1/2} \quad \text{and} \quad |K_1| = O(\log\log n).$$

If $k \in K_2$ *then* $\frac{1}{2}\overline{D}(k) \leq D(k) \leq 2\overline{D}(k)$.

(b) *Let* $\gamma_c = (c-1)\log(c/(c-1))$. *Let* $k_1 = (c-1)\log n$, *and let* $S_1 = \{v : d(v) = k_1\}$, *then*

$$|S_1| = (nep/k_1)^{k_1} n^{1-c} = n^{\gamma_c + o(1)}.$$

Proof An identical calculation is made in [8] for the degree sequence of the random graph $G_{n,p}$. □

The remainder of this section is devoted to proving that vertices of G which are close spatially in $I(d) = [0,1]^d$, are close in edge distance in G. To do this we partition $I(d)$ into sub-cubes of various (appropriate) sizes, and examine the structure of the graph within and between the sub-cubes. This partition approach is used in many of the proofs throughout this paper.

Let

$$h_a = \epsilon_a r \text{ and } h_b = L_1 h_a$$

where $\epsilon_a \leq 1/4d$ is a small positive constant. We assume that $\ell_a = 1/h_a$ is an even integer and L_1 is a large odd integer constant which divides ℓ_a, and thus ℓ_a/L_1 is even. We will assume that

$$L_1 \epsilon_a \gg 1 \tag{3}$$

so that a ball of radius r fits well into a cube of side $L_1 h_a$.

The size of L_1 is constrained by the inequalities (6) and (7); where Γ_d is given by (4) and the parameter L of inequality (7) satisfies (5).

We partition $I(d)$ into grids K_a, K_b where K_a, K_b are made up of cubes of side h_a, h_b and K_a is a refinement of K_b. We assume that ϵ_a is small enough so that if x, y are in K_a-cubes that share a $(d-1)$-dimensional face then x, y are adjacent in G.

Let $B_1, B_2, \ldots, B_M$, $M = \Omega(n/\log n)$ be an enumeration of the K_b-cubes. The aim of the next few lemmas is to show that **whp** we can choose a K_a-cube A_i inside each B_i such that each A_i contains $\Omega(\log n)$ members of V (we say A_i is a *heavy* sub-cube). Furthermore, if B_i, B_j share a $(d-1)$-dimensional face, then there is a sequence of $3L_1+1$ heavy K_a-cubes $X_0, X_1, \ldots X_{3L_1}$ such that (i) $X_0 = A_i$ and $X_{3L_1} = A_j$ and (ii) X_i, X_{i+1} share a $(d-1)$-dimensional face for $0 \le i < 3L_1$. L_1 is some sufficiently large constant. In other words, we find a large d-dimensional grid-like structure with vertices represented by the A_i and edges represented by a sequence of $3L_1+1$ heavy cubes. Furthermore, every other vertex will be within $O(1)$ distance of this "grid".

Each K_a-cube is labeled by a d-tuple in $[\ell_a]^d$. Given a K_b-cube B we define a *line* of B to be a set of L_1 K_a-cubes of B, where the labels are constant except for exactly one index. A *slice* of B is a set of L_1^{d-1} K_a-cubes of B, where the labels are constant on exactly one index. A slice is *extreme* if contains a $(d-1)$-dimensional face of B. Given a line of B, its two *ends* are the two K_a-cubes lying in extreme slices.

If we fix a K_a-cube A, then the number of points in V that are in A is distributed as $Bin(n, \alpha \log n/n)$ where $\alpha = c\epsilon_a^d/\Upsilon_d$. A cube is *light* if it contains fewer than $\epsilon_\ell \alpha \log n$ points in V where $\epsilon_\ell \le c_0/2c$ is a small constant, otherwise it is *heavy.* If C is an arbitrary union of K_a-cubes $A_1, A_2, \ldots, A_k$ then $heavy(C) = \{A_i : A_i \textit{ is heavy}\}$.

Given a K_a-cube A, let $K_L(A)$ be a cube of side Lh_a with A at its centre (assuming that L is an odd integer). Consider $K_L(A)$ to be partitioned into L^d K_a-cubes.

Lemma 5. *Suppose that $L = O(1)$. Let*

$$\Gamma_d = 20\Upsilon_d \epsilon_a^{-d}. \tag{4}$$

Then **whp** *there does not exist a K_a-cube A such that $K_L(A)$ contains Γ_d light K_a-cubes*

Proof The number of points of V in a set of t K_a-cubes is distributed as the binomial $Bin(n, t\alpha \log n/n)$. If they are all light then this number is less than $t\alpha \log n/2$. Using Chernoff bounds, we see that the probability that $K_L(A)$ contains at least $t \ge \Gamma_d$ light K_a-cubes is at most

$$\binom{L^d}{t} e^{-t\alpha \log n/8} \le (L^d\, n^{-c\epsilon_a^d/(8\Upsilon_d)})^t \le n^{-2}.$$

There are $O(n)$ K_a-cubes and the claim follows.

□

We use the following result, which is part of Lemma 9.9 of Penrose [19]: Let $B_Z(n) = [n]^d$ and let A be a subset of $B_Z(n)$. We assume a graph structure with vertices $[n]^d$, and where two vertices are adjacent if their Hamming distance is one. The *external vertex boundary* $\partial^+_{B(n)} A$ is the set of vertices in $B_Z(n) \setminus A$ which are adjacent to some $x \in A$.

Lemma 6. *If $A \subseteq [n]^d$ and $|A| \leq 2n^d/3$ then*

$$|\partial^+_{B(n)} A| \geq (2d)^{-1}(1-(2/3)^{1/d})|A|^{(d-1)/d}. \square$$

Fix a cube C that is the union of K_a-cubes and is of side Lh_a, $L = O(1)$. Consider the graph H_C with vertex set $heavy(C)$ and where two vertices of H_C are adjacent if the corresponding cubes share a $(d-1)$-dimensional face. Let $\kappa_1(H_C)$ denote the size of the largest component of H_C. We a somewhat loosely refer to this as "the largest component of C".

Lemma 7. Whp

$$\kappa_1(H_C) \geq L^d - \gamma\Gamma_d^2$$

for some $\gamma = \gamma(d) \geq 1$.

Proof It follows from Lemma 5 that $|V(H_C)| = |heavy(C)| \geq L^d - \Gamma_d$. Let $W_1, W_2, \ldots, W_s$ be the components of H_C. Suppose that some r-subset of components $W_1, ..., W_r$ satisfies $W = \bigcup_{i=1}^r W_i$ where $|W| = w \leq L^d/2$. Lemma 6 implies that W has at least $a_1 w^{(d-1)/d}$ neighbours for some $a_1 = a_1(d)$. As these must all be light this implies that $\Gamma_d \geq a_1 w^{(d-1)/d} \geq a_1 w^{1/2}$ and the claim follows. $\square$

Recall that $h_b = L_1 h_a$ defines the side length of the K_b-cubes. Let A be the centre K_a-cube of a K_b-cube B. We introduce a new quantity L and define $K_L^*(B)$, the L-centre of B, to be $K_L(A)$. If F is an extreme slice of B, then define its L-centre as follows: If X is the centre K_a-cube of F then $K_L^*(F) = F \cap K_L(X)$.

A line Λ of B containing a cube $\hat{A} \in K_L^*(B)$ is *good* if it satisfies the following conditions:

1. All its K_a-cubes are in $\kappa_1(H_B)$.
2. Let A_1, A_2 be the K_a-cubes at the ends of Λ. Let F_i be the extreme slice containing A_i. Let $F_i^* = K_L^*(F_i)$. Let H_i^* be the sub-graph of H_B induced by F_i^*. We require that $A_i \in \kappa_1(H_i^*)$ for $i = 1, 2$.

We say that a K_a-cube $\hat{A}$ is *good* if $\hat{A} \in K_L^* \cap \kappa_1(H_B)$ and if the d lines through $\hat{A}$ are good. A good cube $\hat{A}$ is *useable* if all K_a-cubes within distance 10 of $\hat{A}$ are good.

Lemma 8. *K_L^* contains at least $L^d - (2d)^{11}L\gamma\Gamma_d^2$ useable cubes,* **whp**.

Proof Let B be a K_b-cube. Every useable sub-cube of B lies in $\kappa_1(H_B)$. The number of lines containing a light cube of B is at most $d\Gamma_d$. The number of lines with an end-cube in the L-centre of its face that violate Condition 2 of goodness is at most $2d \times \gamma\Gamma_{d-1}^2$. (We have applied Lemma 7 to the F_i^* of this condition).

Each cube has at most $(2d)^{10}$ cubes within distance 10. So the number of non-useable cubes in the L-centre is at most $L(2d)^{10}(d\Gamma_d + 2d\gamma\Gamma_{d-1}^2)$. $\square$

We make L large enough so that

$$L^d - (2d)^{11} L\gamma\Gamma_d^2 > L^d/2 \tag{5}$$

and then we can apply Lemma 8 to show there are many useable cubes. We also assume that

$$L_1 \geq L^d. \tag{6}$$

Finally, Lemma 10 below, requires the following lower bound on L_1:

$$L_1 \geq 3^d(\Gamma_d + \gamma\Gamma_d^2)/\epsilon_a. \tag{7}$$

Let $B_1, B_2, \ldots, B_M$, $M = \Omega(n/\log n)$ be an enumeration of the K_b-cubes. Our grids define bipartite graphs, this is why we chose $\ell_a, \ell_a/L_1$ even. Thus each cube will have a parity, with neighbouring cubes having different parity. Similarly, the K_a sub-cubes of each B_i have a parity. We choose a useable cube A_i in each B_i, $i = 1, 2, \ldots, M$, where A_i is chosen to have the same parity as B_i.

In what follows, by a path, we mean a sequence of K_a-cubes (the path vertices) with consecutive cubes sharing a $(d-1)$-face.

Lemma 9. *If K_b-cubes B_i, B_j share a $(d-1)$-dimensional face, then there is a path $P(i,j)$ of length $3L_1$ of heavy cubes joining A_i and A_j, where A_i, A_j are useable cubes, as defined above. These paths are pair-wise internally vertex disjoint.*

Proof First define a path $Q(i,j)$ as follows: Start at A_i and follow the line Λ_i of cubes to the face separating B_i and B_j. Suppose this ends in the cube S_i. Do the same for A_j. Now choose a cube S in $\kappa_1(F_i^*)$ that shares a $(d-1)$-face with a cube S' in $\kappa_1(F_j^*)$. This is possible if L is large enough. $Q(i,j)$ consists of Λ_i, then a path from S_i to S inside $\kappa_1(F_i^*)$, then the edge to S', then a path from S' to S_j inside $\kappa_1(F_j^*)$ and finally the line Λ_j in the direction S_j to A_j. The $Q(i,j)$ are internally vertex disjoint by construction and have odd length at most $2L_1 + L^d \leq 3L_1$, from (6). The odd length follows from the fact that the parities of B_i, B_j differ. We now add to $Q(i,j)$ to bring its length up to $3L_1$. Suppose $Q(i,j)$ has length $3L_1 - 2t$ where t is a positive integer. Suppose that $Q(i,j)$ begins with cubes A_i, S_1, S_2. We will replace the edge (S_1, S_2) by a path of length $2t+1$. Suppose that we number the $2d$ coordinate directions as $\pm 1, \pm 2, \ldots, \pm d$. Suppose that $Q(i,j)$ starts in direction $+k$. If $t = 1$ we replace (S_1, S_2) by (S_1, T_1, T_2, S_2) where T_l is the neighbour of S_l in direction $+(k+1)$. Otherwise, let T_l' be the neighbour of T_l in the direction $+(k+1)$. We can now construct a path $S_1, T_1, T_1', T_2', P_h, T_2, S_2$ where P_h is a path of length $2h+1$, $0 \leq h \leq (L_1 - L)/2$. The path P_h goes out from T_2' for h steps in direction $+k$, then makes a step in direction $-(k+1)$ and then returns to T_2 via h steps in direction $-k$. This path will only use heavy cubes, by the definition of useable. This suffices for $2h+4=2t$ provided $t \leq L_1 - L + 2$. For larger t we will have to go out further in the direction $+(k+1)$ and use two or three paths in the direction $+k$. Our choice of directions will maintain disjointness.

$\square$

We now consider points that do not lie in a cube of $\kappa_1(H_B)$ for any B in the K_b dissection of $I(d)$.

Lemma 10. *For $v \in V$, let $C_v^{(b)}$ be the K_b-cube containing v in the K_b dissection of $I(d)$ and let $C_v^{(a)}$ be the K_a-cube containing v in the K_a dissection of $C_v^{(b)}$. Then, there exists a K_b-cube B_i such that*

(a) *v is at G-distance $\leq 3^d(\Gamma_d + \gamma\Gamma_d^2)$ from $\kappa_1(H_{B_i})$.*

(b) *v is at G-distance $O(1)$ from any point $w \in V$ in B_i.*

Proof Let $\ell = 3^d(\Gamma_d + \gamma\Gamma_d^2)$. Fix $v \in V$ lying in the cube $C_v^{(a)}$ and let $P = (v_0 = v, v_1, \dots, v_k)$ be a shortest path in G from v to a point in an H_{B_i}. Here we are assuming that G is connected, which it is **whp**.

We can assume that the K_a-cubes that contain these v_i are distinct, otherwise the path can be shortened. Given (7) we can see that $k \leq \ell$. Indeed, if $\ell < k$ then (7) implies that the sub-path $(v_0, v_1, \dots, v_\ell)$ stays inside the $\leq 3^d$ K_b-cubes that touch $C_v^{(b)}$. It follows from Lemmas 5 and 7 that at least one of K_a-cubes containing one of the $v_1, v_2, \dots, v_\ell$ is in the largest heavy component of its K_b-cube. This verifies $k < \ell$ and also proves part (a). Once we reach the largest component of B_i, we can reach a vertex in A_i in $O(1)$ steps, which proves part (b). $\square$

3 Estimating first visit probabilities

In this section we describe the main analytical tool we use to compute the cover time. We use the approach of [6, 7, 9, 10]. Let G denote a fixed connected graph, and 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 $P_u^{(t)}(v) = \mathbf{Pr}(\mathcal{W}_u(t) = v)$. Let π be the steady state distribution of the random walk $\mathcal{W}_u$. Let $\pi_v = \pi(v)$ denote the stationary distribution of the vertex v. For an unbiased ergodic random walk on a graph G with $m = m(G)$ edges, $\pi_v = \frac{d(v)}{2m}$, where $d(v)$ denotes the degree of v in G.

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

$$\max_{u,x\in V} |P_u^{(t)}(x) - \pi_x| \leq n^{-3}. \tag{8}$$

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

$$\max_{u,x\in V} |P_u^{(kT)}(x) - \pi_x| \leq \frac{2^{k-1}}{n^{3k}}. \tag{9}$$

Now fix $u \neq v \in V$. Next, let $r_t = \mathbf{Pr}(\mathcal{W}_v(t) = v)$ be the probability that this walk returns to v at step t. Let

$$R_T(z) = \sum_{j=0}^{T-1} r_j z^j. \tag{10}$$

For a large constant $K > 0$, let

$$\lambda = \frac{1}{KT}. \tag{11}$$

For $t \geq 0$, let $\mathcal{A}_t(v)$ be the event that $\mathcal{W}_u$ does not visit v in steps $T, T+1, \ldots, t$ The vertex u will have to be implicit in this definition. The following was proved in [9].

Lemma 11. *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* $T\pi_v = \Omega(n^{-2})$.

Let

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

where $R_T(1)$ is from (10).

Then for all $t \geq T$,

$$\mathbf{Pr}(\mathcal{A}_t(v)) = \frac{(1 + O(T\pi_v))}{(1 + p_v)^t} + O(Te^{-\lambda t/2}). \tag{12}$$

The evaluation of $R_T(z)$ at $z = 1$ occurs frequently in our calculations in this paper. For the rest of the paper u, v will not be fixed and it is appropriate to replace the notation $R_T(1)$ by something dependent on v. We use the notation R_v. For $u \neq v$ we let $R_{u,v}$ denote the expected number of visits by $\mathcal{W}_u$ to v up to time T. Our next two tasks are to estimate T and the R_v.

4 Mixing time of the random walk

We need two basic results on mixing times. First let $\lambda_{\max}$ be the second largest eigenvalue of the transition matrix P. Then,

$$|P_u^{(t)}(x) - \pi_x| \le \left(\frac{\pi_x}{\pi_u}\right)^{1/2} \lambda_{\max}^t \le ((c+10)/c_0)^{1/2} \lambda_{\max}^t. \tag{13}$$

See for example, Jerrum and Sinclair [18].

Next, for each $x \neq y \in V$ let $Q(x,y)$ be a *canonical* path from x to y in G. Then, see for example Sinclair [20], we have that

$$\lambda_{\max} \le 1 - \frac{1}{\rho}, \tag{14}$$

where

$$\rho = \max_{e=\{x,y\} \in E(G)} \frac{1}{\pi(x)P(x,y)} \sum_{\gamma_{ab} \ni e} \pi(a)\pi(b)|\gamma_{ab}|, \tag{15}$$

and $|\gamma_{ab}|$ is the length of the canonical path $Q(a,b)$ from a to b.

Consider the K_b-grid of Section 2. It will help to fix a collection of points $x_i \in A_i \cap V$ for $i = 1, 2, \ldots, M$.

Lemma 12.

$$1 - \lambda_{\max} = \tilde{\Omega}\left(n^{-2/d}\right).$$

Proof We first define canonical paths between the x_i. We can in a natural way express $x_i = y(j_1, j_2, \ldots, j_d)$ where $0 \le j_t < 1/h_b$ for $1 \le t \le d$. The path from $y(j_1, j_2, \ldots, j_d)$ to $y(k_1, k_2, \ldots, k_d)$ goes

$$y(j_1, j_2, \ldots, j_d) \leftrightsquigarrow y(j_1+1, j_2, \ldots, j_d) \leftrightsquigarrow \cdots, y(k_1, j_2, \ldots, j_d) \leftrightsquigarrow y(k_1, j_2+1, \ldots, j_d) \leftrightsquigarrow \cdots \leftrightsquigarrow, y(k_1, k_2, \ldots, k_d).$$

The $\leftrightsquigarrow$ represents a path in G that follows the sub-cubes of a $P(i,j)$, choosing one vertex from each K_a-cube as necessary.

Thus we first increase the first component (mod $1/h_b$) until it is k_1 and then do the same for the second and subsequent components. Each such path has length at most $3dL_1/h_b = O(1/r)$. If we fix a grid edge e (really an edge of a path $\leftrightsquigarrow$) joining $y(j_1, \ldots, j_t, \ldots, j_d)$ to $y(j_1, \ldots, j_t+1, \ldots, j_d)$ then the number of paths through e is $O(h_b^{-d-1})$; any such path starts at $y(l_1, \ldots, l_t, j_{t+1}, \ldots, j_d)$ and ends at $y(j_1, \ldots, j_{t-1}, l'_t, \ldots, l'_d)$ for some $l_1, \ldots, l_t, l'_t, \ldots, l_d$.

We obtain canonical paths for every pair of vertices by using Lemma 10 i.e. we connect each point x of V to its closest $x_i = \phi(x)$. Each x_i is chosen by $O(\log n)$ points in this way. Using Chernoff bounds, we bound the number of points in V at G-distance $O(1)$ (Lemma 10) from any x_i. Our path from x to y goes x to $\phi(x)$ to $\phi(y)$ to y. After this we find that each path has length $O(1/r)$ and each edge is in $\tilde{O}\left(1/r^{d+1}\right)$ paths. It follows from (15) that

$$\rho = \tilde{O}\left(n \cdot 1 \cdot r^{-d-1} \cdot n^{-2} \cdot r^{-1}\right) = \tilde{O}\left(n^{2/d}\right)$$

and the lemma follows from (14). □

Applying (13) we see that we can take

$$T = \tilde{O}\left(n^{2/d}\right) \tag{16}$$

when we use Lemma 11.

5 Expected number of returns during the mixing time etc.

Having obtained a good enough bound on T, we next show that **whp** $R_v = 1 + o(1)$ for all $v \in V$, and that the conditions of Lemma 11 hold.

For a set of vertices B, the escape probability $p_{\text{esc}}(a, B)$, is the probability that a random walk leaving a does not return to a before reaching B. This probability is given by

$$p_{\text{esc}}(a, B) = \frac{1}{d(a)\mathrm{R}_{\text{EFF}}(v, B)}, \tag{17}$$

where $\mathrm{R}_{\text{EFF}}(a, B)$ is the *effective resistance* between a and B in an electrical network with all edges having resistance one, see for example Doyle and Snell [11].

Thus $R_v(B)$, the expected number of returns to v before reaching B is given by

$$R_v(B) = \frac{1}{p_{\text{esc}}(v, B)} = d(v)\mathrm{R}_{\text{EFF}}(v, B).$$

Raleigh's Theorem (see e.g. [11]), states that deleting edges increases effective resistance. Thus, provided we do not prune edges incident with vertex v, edge deletion increases $R_v(B)$.

We will identify a set U_v such that $\mathcal{W}_v$ is very likely to enter U_v before returning to v and such that if $x \in U_v$ then $\mathcal{W}_x$ is unlikely to visit v within time T. We deduce from this the $R_v = 1 + o(1)$.

5.1 The probability a walk visits fixed vertex during T

For $x, y \in V$ we let

$$\eta(x,y) = \mathbf{Pr}(\exists\, 1 \le t \le T : \mathcal{W}_x(t) = y). \tag{18}$$

We aim to show that if y is fixed, then $\eta(x,y) = o(1)$ for almost all choices of x.

Given $\epsilon > 0$, let

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

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,

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

It follows that if

$$U_v = \left\{x :\ \eta(x,v) \ge \frac{1}{(\log n)^2}\right\}$$

then **whp**

$$|U_v| = \tilde{O}(T). \tag{19}$$

5.2 The value of R_v

We next prove that

Lemma 13. Whp *$R_v = 1 + O(1/\log n)$ for all $v \in V$.*

Proof Fix $v \in V$ and make v the centre of a K_b-cube C_v of side $L_1 h_a$ and partition $I(d)$ into K_b-cubes with C_v as one of the cubes. Let $\bar{U}_v = V \setminus U_v$.

Let $\mathcal{W}_v$ starting from v, and let p_v be the probability of a first return to v within time T by this walk. Then

$$1 \leq R_v \leq \frac{1}{1-p_v}, \tag{20}$$

and

$$p_v \leq 1 - p_{\text{esc}}(v, \bar{U}_v) + 1/(\log n)^2. \tag{21}$$

Given (21) and (20), it is sufficient to prove for all $v \in V$ that **whp**

$$p_{\text{esc}}(v, \bar{U}_v) = 1 - O(1/\log n). \tag{22}$$

We focus on proving (22). We construct a sub-graph $G_v^* = (V^*, E^*)$ of G. The set $V^* = W_1^* \cup W_2^*$, where W_1^*, W_2^* are defined as follows. Let S denote the set of heavy K_a-cubes that either (i) belong to a path $P(i,j)$ as defined in Section 2 or (ii) are contained in C_v. For each sub-cube in S, we choose an arbitrary subset of vertices of size $\epsilon_\ell \alpha \log n$, $\alpha = c/(\Upsilon_d \epsilon_a^d)$ and place these vertices in W_1^*. (See the definition of *heavy* above Lemma 5 for the values of these constants.) Let W_2^* consist of v and all of its neighbours and any vertex that is in a heavy cube of C_v. Note that (3) implies that v has no neighbours outside of C_v. Let $E^* = E_1^* \cup E_2^*$, where E_1^* consists of those edges (x, y) where x, y are either contained in the same K_a-cube of S, or are in distinct K_a-cubes of S which share a $(d-1)$-dimensional face. The set E_2^* consists of the edges of G that join two vertices of W_2^*.

The degree of v in G^* is the same as its degree in G and G^* is a sub-graph of G. From Raleigh's Theorem, we see that $\mathrm{R}_{\text{EFF}}(v, \bar{U}_v) \leq \mathrm{R}^*_{\text{EFF}}(v, \bar{U}_v)$. So if $p^*_{\text{esc}}(v, \bar{U}_v)$ is the probability that the random walk $\mathcal{W}_v^*$ on G^* visits $\bar{U}_v$ before returning to v then

$$p_{\text{esc}}(v, \bar{U}_v) \geq p^*_{\text{esc}}(v, \bar{U}_v).$$

So to prove (22), it suffices to prove

$$p^*_{\text{esc}}(v, \bar{U}_v) = 1 - O(1/\log n) \tag{23}$$

Define a d-dimensional grid $\tilde{\Gamma}$ with M vertices as follows: There is one vertex of $\tilde{\Gamma}$ for each of the K_b-cubes $B_1, B_2, \ldots, B_M$ in the partition of $I(d)$ and the cube C_v corresponds to the origin of $\tilde{\Gamma}$.

With one small caveat, the random walk $\mathcal{W}_v^*$ on G^* can be coupled with a random walk $\tilde{\mathcal{W}}$ on $\tilde{\Gamma}$. Let $V_i^* = V^* \cap A_i$, for $i = 1, 2, \ldots, M$, where A_i is the useable K_a-cube of B_i chosen according to Lemma 9. When $\mathcal{W}_v^*$ is inside V_i^*, $\tilde{\mathcal{W}}$ will be at the ith vertex of $\tilde{\Gamma}$. If $\mathcal{W}_v^*$ is on a vertex of a path $P(i,j)$ then $\tilde{\Gamma}$ stays at its current vertex. Since the paths $P(i,j)$ are all of the same length, and since the V_i^* are all the same size, the next vertex that $\mathcal{W}_v^*$ visits is equally likely to be any neighbour of the current vertex. The caveat concerns what to do when $\mathcal{W}_v^*$ is inside C_v or equivalently when $\tilde{\mathcal{W}}$ is at the origin. In this case it is not true that the K_b-cube next visited by $\mathcal{W}_v^*$ is equally likely to be any neighbour. This turns out to be unimportant, but it should be borne in mind in the following discussion.

Now focus on the random walk $\tilde{\mathcal{W}}$ on the $[N]^d$, $N = M^{1/d}$, where $M = \Omega(n/\log n)$. By assumption, $\tilde{\mathcal{W}}$ starts at the origin. Let $\mathcal{J}_V^* = \left\{i \in [M] : V_i^* \subseteq \bar{U}_v\right\}$.

We will prove that with probability bounded below by a constant $\gamma > 0$, $\tilde{\mathcal{W}}$ will (for some small constant c_1) visit $\mathcal{J}_V^*$ within $c_1 N^2$, steps, before returning to the origin. We will also show that

$$\mathbf{Pr}(\mathcal{W}_v^* \text{ revisits } v \mid \mathcal{W}_v^* \text{ visits } W_2^*) \leq \epsilon_v = O(1/\log n). \tag{24}$$

This bound includes the probability of a re-visit to v before $\mathcal{W}^*$ first leaves C_v. Thus,

$$1 - p_{\text{esc}}^*(v, \bar{U}_v) \leq \epsilon_v \sum_{i=0}^{\infty} i(1-\gamma)^i < \frac{\epsilon_v}{\gamma^2} = O(1/\log n)$$

and this completes the proof of (23).

Now because $d \geq 3$ there is a positive probability γ' such that we have $\tilde{\mathcal{W}}(t') \neq 0$ for $1 \leq t' \leq t = c_1 N^2$. This is because a random walk $\hat{\mathcal{W}}$ on the infinite d-dimensional lattice is non-recurrent. Thus there is a constant probability $\zeta_d > 0$ that it does not return to the origin. The symmetry of the grid ensures that this remains true, even if there is a non-uniform choice of neighbour at the origin. If c_1 is small, then there is a greater than $1 - \zeta_d/2$ chance that $\hat{\mathcal{W}}$ stays inside the box $[-N/3, N/3]^d$ for the first t steps and this implies that $\tilde{\mathcal{W}}$ does not return to the origin with probability at least $\zeta_d - \zeta_d/2$. Furthermore, if the walk does not return to the origin, then its subsequent behaviour is precisely that of the standard walk, given the neighbour $X_1 = \tilde{\mathcal{W}}(1)$ of the origin that is first chosen. So we can now think of the distribution of $\tilde{\mathcal{W}}$ as that of a standard walk that first goes to X_1 and then with probability at least $\zeta_d/2$ does not return to the origin. Now for a fixed $x \in N^d$ we have

$$\mathbf{Pr}(\mathcal{W}_v^*(t) \in V_x^* \mid \text{ no return, } X_1) = O(\mathbf{Pr}(\mathcal{W}_v^*(t) \in V_x^* \mid X_1)) = O(\mathbf{Pr}(\hat{\mathcal{W}}(t) = x \mid X_1).$$

For a random walk $\hat{\mathcal{W}}$ on the d-dimensional lattice,

$$\mathbf{Pr}(\hat{\mathcal{W}}(t) = x) = O(\mathbf{Pr}(\hat{\mathcal{W}}(t) = 0)) = O(t^{-d/2}).$$

This can be seen by considering the relevant multinomial terms (see e.g. page 329 of Feller [15] for the case $d = 3$). Thus choosing $t = c_1 N^2$, as above, we have

$$\mathbf{Pr}(\tilde{\mathcal{W}}_v(t) \notin \mathcal{J}_V^*) = \tilde{O}\left(|U_v| N^{-d}\right) = \tilde{O}\left(T N^{-d}\right) = O(1/\log n).$$

So

$$\mathbf{Pr}\left(\mathcal{W}_v^*(t) \notin \bigcup_{i \in \mathcal{J}_V^*} V_i^*\right) = \tilde{O}\left(|U_v| N^{-d}\right) = \tilde{O}\left(T N^{-d}\right) = O(1/\log n).$$

Now any constant $\gamma < \gamma'$ will suffice. □

5.3 Proof of (24):

For this we consider the graph H with vertex set equal to the set of heavy K_a-cubes. Two heavy cubes C_1, C_2 are defined to be adjacent in H if the centers of C_1, C_2 are no more than $r_1 = r - 2d^{1/2}h_a$ apart. In which case, $v_i \in C_i \cap V$, $i = 1, 2$ implies that $(v_1, v_2) \in G$.

Lemma 14. *The ball $D(v, r)$ contains $\Upsilon_d \epsilon_a^{-d}(1 - \epsilon_B)$ K_a-cubes, where $0 \leq \epsilon_B \leq 1 - (1 - 2\epsilon_a d^{1/2})^d$.*

Proof The upper bound of $\frac{\Upsilon_d r^d}{(\epsilon_a r)^d}$ is the ratio of the volume of $D(v, r)$ and the volume of a K_a-cube. For the lower bound, consider $D(v, r')$, $r' = r_1$. Every K_a-cube that touches $D(v, r_1)$ is contained entirely in $D(v, r)$. There are at least $\frac{\Upsilon_d r_1^d}{(\epsilon_a r)^d}$ cubes that touch $D(v, r_1)$ and the lower bound follows. □

It follows that the maximum degree in H satisfies

$$\Delta(H) \leq \Upsilon_d \epsilon_a^{-d} = O(1). \tag{25}$$

Lemma 15. Whp, *for every $w \in V$, $D(w, r)$ contains at least one heavy cube.*

Proof Let $c' = (cr_1^d \log n / \Upsilon_d n)^{1/d}$. If ϵ_a is sufficiently small, we will have $c' > 1$ and then **whp** (see proof of Lemma 3) $D(w, r_1)$ contains at least $c_0' \log n$ points of V where $c_0' \geq c_0/2$, for all $w \in V$. There are at most $\Upsilon_d \epsilon_a^{-d}$ cubes contained entirely in $D(w, r)$. So, one of these cubes must contain at least $\frac{c_0 \epsilon_a^d}{2\Upsilon_d} \log n$ points i.e. is heavy. □

The following bound is somewhat crude, but it will suffice.

Lemma 16. Whp *H contains no component with fewer than $\log \log n$ vertices.*

Proof Let $\delta_C = 1_{C \text{ is heavy}}$ for a K_a-cube C. Now let C be a fixed K_a-cube and S a set of K_a-cubes with $|S| = O(\log n)$. If $p = \frac{c\epsilon_a^d \log n}{\Upsilon_d n} = \frac{\alpha \log n}{n}$, and $n' = n - O((\log n)^2)$ then

$$\mathbf{Pr}(C \text{ is light} \mid \delta_s,\ s \in S) \leq \sum_{i=0}^{\epsilon_\ell \alpha \log n} \binom{n'}{i} p^i (1-p)^{n'-i} \leq n^{-\beta}, \tag{26}$$

where $\beta = \alpha(1 - \epsilon_\ell \log(e/\epsilon_\ell)) + o(1)$. (Strictly speaking we should also condition on the event that every K_a-cube has $O(\log n)$ points of V. However, this happens with probability $1 - O(n^{-K})$ where the constant K can be made as large as necessary).

Let $N = 1/\epsilon_a r$ so that the grid K_a has N^d cubes.

Case 1: $k \leq (c-1)\Upsilon_d \epsilon_a^{-d}/2$.

The probability that H contains a component with k vertices is at most

$$o(1) + n(e\epsilon_a^{-d})^k n^{-\beta(\Upsilon_d \epsilon_a^{-d}(1-\epsilon_B)(r_1/r)^d - k)} = o(1).$$

Explanation: The $o(1)$ term on the LHS is the probability that there is a vertex of degree exceeding $(c+10)\log n$. Choose a K_a-cube C and another $k-1$ that make up a connected component of H. There are at most $(e\epsilon_a^{-d})^k$ ways to choose a tree of cubes with root C. Now let x be the centre of the first chosen cube. The term $n^{-\beta(\Upsilon_d \epsilon_a^{-d}(1-\epsilon_B)(r_1/r)^d - k)}$ bounds the probability that all of the other cubes in $D(x, r_1)$ are light. Note that $\beta(\Upsilon_d \epsilon_a^{-d}(1-\epsilon_B)(r_1/r)^d - k) > 1$ if $\epsilon_\ell, \epsilon_B = \epsilon_B(\epsilon_a)$ are sufficiently small.

Case 2: $(c-1)\Upsilon_d \epsilon_a^{-d}/2 \le k \le \log\log n$.

Choose k K_a-cubes $C_1, C_2, \ldots, C_k$ and let $X = \bigcup_{i=1}^k C_i$. Let Y be the subset of $I(d)$ within distance $r_2 = r_1 - d^{1/2}h_a$ of X. Note that if $y \in Y$ then it is within distance r_1 of a centre of a sub-cube of X. We claim that

$$\text{vol}_d(Y) \ge \Upsilon_d(\rho + r_2)^d,$$

where ρ is the solution to $k\epsilon_a^d r^d = \Upsilon_d \rho^d$. This follows from the classical isoperimetric inequality [4], where ρ is the radius of a ball with the same volume as X. Note that

$$\rho = \left(\frac{k\epsilon_a^d r^d}{\Upsilon_d}\right)^{1/d} = \left(\frac{k}{\Upsilon_d}\right)^{1/d} \epsilon_a r \ge \left(\frac{c-1}{2}\right)^{1/d} r.$$

Arguing as in Lemma 14 we see that if $r_3 = r_2 - 2d^{1/2}h_a$ then $Y \setminus X$ contains at least

$$\Upsilon_d \epsilon_a^{-d}\left((1-\epsilon_B)\left(\frac{\rho + r_3}{r}\right)^d - \left(\frac{\rho}{r}\right)^d\right) \ge \Upsilon_d \epsilon_a^{-d}(1-\epsilon_B)(r_3/r)^d \tag{27}$$

K_a-cubes. So, the probability that H contains a component with k vertices is at most

$$n((c+10)\log n)^k n^{-\beta(\Upsilon_d \epsilon_a^{-d}(1-\epsilon_B)(r_3/r)^d)} = o(1).$$

The term $n^{-\beta(\Upsilon_d \epsilon_a^{-d}(1-\epsilon_B)(r_3/r)^d)}$ bounds the probability that all the cubes mentioned in (27) are light. □

Now consider a random walk $\mathcal{W}^*$ on G^* that starts at some vertex $w \in C_v$. Note first that, when at a neighbour of v, there is only an $O(1/\log n)$ chance of returning to v at the next step. Second note, that at any vertex, there is at least the chance $\epsilon_h = \frac{\epsilon_\ell \alpha}{c+10}$ of moving to a heavy cube (Lemma 15). In particular, with probability at least $\epsilon_h - o(1)$ the next vertex after w lies in a heavy cube and is not equal to v. Let P be some path in G^* of length t_0 where $20\Upsilon_d \epsilon_a^{-d} \ll t_0 \ll L_1$. We can use Lemma 16 to argue that P exists.

Then note that with probability $\left(\frac{\epsilon_h}{\Delta(H)}\right)^{t_0}$ the next t_0 steps of $\mathcal{W}^*$ traverse the cubes of P. By Lemma 5 we see that at least one of the cubes on P will be in $\kappa_1(C_v)$. Once we reach such cube there is at least the chance $\epsilon^* = \epsilon_h^{3^d(\Gamma_d+c_b\Gamma_d^2)+L_1^d+1}$ of leaving C_v by going along the path promised by Lemma 10 to a giant component and then going through this giant component and leaving C_v. Thus the chance of returning to v is $O(1/\log n)$ either when starting at v or when returning to C_v.

This completes the proof of (24) and the lemma. □

5.4 Conditions of Lemma 11

It is clear from (16) that Lemma 11(b) holds. To check condition (a) we see that if $|z| \leq 1+\lambda$ then since by Lemma 13 we have $R_v = 1 + o(1)$, we see that

$$\left|\sum_{j=1}^{T-1} r_j z^j\right| \leq (1+\lambda)^T \sum_{j=1}^{T-1} r_t = (1+\lambda)^T (R_v - 1) = o(1).$$

So $|R_T(z)| \geq 1 - \left|\sum_{j=1}^{T-1} r_j z^j\right| = 1 - o(1)$ for $|z| \leq 1 + \lambda$.

6 Cover time

From (12) of Lemma 11 we have that for all $t \geq T$,

$$\mathbf{Pr}(\mathcal{A}_t(v)) = \frac{1+o(1)}{\left(1+\frac{\pi_v}{R_v(1+O(T\pi_v))}\right)^{t+1}} + O(Te^{-\lambda t/2}). \tag{28}$$

An upper bound is obtained as follows: 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:

$$\mathbf{Pr}(T_G(u) > t) = \mathbf{Pr}(U_t > 0) \leq \min\{1, \mathbf{E}U_t\}, \tag{29}$$

$$C_u = \mathbf{E}T_G(u) = \sum_{t>0} \mathbf{Pr}(T_G(u) > t) \tag{30}$$

It follows from (28,29) that for all t

$$C_u \leq t + 1 + \sum_{s>t} \mathbf{E}U_s = t + 1 + \sum_{v\in V}\sum_{s>t} \mathbf{Pr}(\mathcal{A}_s(v)). \tag{31}$$

Let

$$t^* = \left(c \log\left(\frac{c}{c-1}\right)\right) n \log n$$

and

$$t_0 = (1-\delta)t^* \textit{ and } t_1 = (1+\delta)t^*$$

where $\delta = o(1)$ but goes to zero sufficiently slowly that inequalities below are satisfied.

6.1 Upper bound on the cover time

For $v \in V$ we have by Lemmas 11 and 13 that

$$\mathbf{Pr}(\boldsymbol{A}_s(v)) = (1+o(1)) \exp\{-(1+o(1/\log n))\pi_v s\} + O(Te^{-\Omega(s/T)})$$

and we note that

$$\pi_v = \frac{d(v)}{2m}.$$

Then we find, using the **whp** bounds in Lemma 4,

$$C_u \leq t_1 + 1 + S_1 + S_2 + O(nT^2 e^{-\Omega(s/T)}) \tag{32}$$

where

$$\begin{aligned} S_i &= \sum_{k\in K_i} D(k) \sum_{s\geq t_1} \exp\left\{-\frac{(1-o(1))ks}{2m}\right\} \\ &\leq 3m \sum_{k\in K_i} \frac{D(k)}{k} e^{-(1-o(1))kt_1/2m} \\ &\leq 3m \sum_{k\in K_i} \frac{D(k)}{k} \left(\frac{c-1}{c}\right)^{(1+\delta/2)k}. \end{aligned}$$

For the first term,

$$\begin{aligned} S_1 &\leq 3m \sum_{k\in K_1} \frac{(\log n)^4}{k} \left(\frac{c-1}{c}\right)^{(1+\delta/2)k} \\ &= o(t_1) \end{aligned} \tag{33}$$

since $D(k) \leq (\log n)^4$ and $\min\{k \in K_2\} \geq (\log n)^{1/2}$.

Continuing we get

$$
\begin{aligned}
S_2 &\leq \frac{6m}{n^{c-1}} \sum_{k \in I_c} \left(\frac{nep}{k}\right)^k \left(\frac{c-1}{c}\right)^{(1+\delta/2)k} \\
&\leq 6m \sum_{k \in I_c} e^{-\delta k/2c} \\
&= o(t_1), \qquad (34)
\end{aligned}
$$

where we have used the fact that $(nep(c-1))/(kc))^k$ is maximized at $k = np(c-1)/c$, and $\delta k/\log\log n \to \infty$ for $k \in I_c$.

It now follows from (32) – (33) that $C_u \leq t_1 + o(t_1)$.

6.2 Lower bound on the cover time

We can find a vertex u and a set of vertices S_0 such that at time t_0, the probability the set S_0 is covered by the walk $\mathcal{W}_u$ tends to zero. Hence $T_G(u) > t_0$ **whp** which implies that $C_G \geq (1-o(1))t^*$.

We construct S_0 as follows. Let $k_1 = (c-1)\log n$ be as defined in Lemma 4, and let $S_1 = \{v : d(v) = k_1\}$. Let $A = \{(u,v) : \ u \notin S_1, v \in S_1, \eta(u,v) \geq 1/(\log n)^2\}$, where $\eta(u,v)$ is defined in (18). It follows from (19) that **whp** $|A| = \tilde{O}\left(T|S_1|\right)$. By simple counting, we see that there exists $u \notin S_1$ such that $|\{v \in S_1 : \ (u,v) \in A\}| = \tilde{O}\left(T|S_1|/n\right) = o(|S_1|)$. We choose such a u and let $S_2 = \{v \in S_1 : \ (u,v) \notin A\}$. We then take an independent subset S_0 of S_2. Because the maximum degree of G is $O(\log n)$ we can choose $|S_0| = \Omega(|S_2|/\log n)$.

Let $\mathcal{B}_v$ be the event that $\mathcal{W}_u$ does not visit v in the time interval $[1,T]$. Then, by our choice of u, we see that for $v \in S_0$,

$$\mathbf{Pr}(\mathcal{B}_v) \geq 1 - 1/(\log n)^2. \qquad (35)$$

We need to prove that

$$\mathbf{Pr}(\mathcal{A}_t(v) \mid \mathcal{B}_v) = (1+o(1)) \exp\left\{-\frac{(c-1)t_0(1+o(1))}{cnR_v(1+O(T\pi_v))}\right\} = (1+o(1))\mathbf{Pr}(\mathcal{A}_t(v)). \qquad (36)$$

The proof of this requires just a small change to the proof of Lemma 11, which we include as an appendix. (Equality is needed in (36). The reader will observe that $\mathbf{Pr}(\mathcal{A}_t(v) \mid \mathcal{B}_v) \leq (1+o(1))\mathbf{Pr}(\mathcal{A}_t(v))$ follows immediately from (35)).

Then **whp**, if Z_0 is the number of vertices in S_0 that are not visited in time $[1, t_0]$,

$$\mathbf{E}(Z_0) \geq A_1 \frac{n^{(c-1)\log(c/(c-1))}}{\log n} \exp\left\{-(1+o(1))\frac{(c-1)t_0}{cn}\right\} \geq A_2 n^{\frac{1}{2}\delta(c-1)\log(c/(c-1))} \to \infty \qquad (37)$$

for some constants $A_1, A_2 > 0$.

We show next, for all $v, w \in S_0$, that

$$\eta(v,w) = O(1/\log n). \tag{38}$$

We define a new graph G_ψ by identifying v, w and replacing them with a new node ψ. The proof of Lemma 12 can be modified to show that the mixing time T_ψ of G_ψ will satisfy (16). Indeed, we can assume that our choice of x_i's excludes v, w and then v, w can only appear as endpoints of canonical paths. For a path from x to ψ we can then choose one of the already constructed canonical paths from x to v or x to w.

Similarly, the proof of Lemma 13 can be modified to show that

$$p_{\text{esc}}(\psi, \bar{U}_\psi) = 1 - O(1/\log n). \tag{39}$$

Here $\bar{U}_\psi = \bar{U}_v \cap \bar{U}_w$ and the probability is for a random walk in G_ψ starting at ψ. Our modification of Lemma 13 requires a random walk on the d-dimensional lattice, starting at point x (a surrogate for v's cube), to have positive probability of not returning to x or some other fixed vertex y (a surrogate for w's cube) and vice-versa. This is a simple consequence of Polya's classic result.

Then

$$\eta(v,w) \leq 2(1 - p_{\text{esc}}(\psi, U_\psi)) + O(1/(\log n)^2) = O(1/\log n).$$

The term $2(1 - p_{\text{esc}}(\psi, U_\psi))$ arises as follows: Consider the random walk $\mathcal{W}_\psi$, conditioned on moving to a neighbour of v at the first step. Up until the time that this walk returns to ψ it behaves just like $\mathcal{W}_v$. It has a probability $1 - p_{\text{esc}}(\psi, U_\psi)$ of entering U_ψ before returning to ψ and this can be inflated by a factor 2 to account for the initial conditioning. But then $\mathcal{W}_v$ has a probability of at most $2(1 - p_{\text{esc}}(\psi, U_\psi))$ of avoiding U_ψ before reaching v or w and a fortiori it has a probability of at most $2(1 - p_{\text{esc}}(\psi, U_\psi))$ of avoiding U_ψ before reaching w. The term $O(1/(\log n)^2)$ accounts for walks that reach U_ψ before w.

Now for $v, w \neq u$ let

$$\mathcal{A}_t(v,w) = \mathcal{A}_t(v) \wedge \mathcal{A}_t(w) \textit{ and } \mathcal{B}_{v,w} = \mathcal{B}_v \wedge \mathcal{B}_w.$$

Then for $v \neq w$,

$$\mathbf{Pr}(\mathcal{B}_{v,w}) \geq 1 - \mathbf{Pr}(\bar{\mathcal{B}}_v) - \mathbf{Pr}(\bar{\mathcal{B}}_w) \geq 1 - 2/(\log n)^2 \tag{40}$$

and we will show that

$$\mathbf{Pr}(\mathcal{A}_{t_1}(v,w) \mid \mathcal{B}_{v,w}) \leq A_0 \mathbf{Pr}(\mathcal{A}_{t_1}(v)) \mathbf{Pr}(\mathcal{A}_{t_1}(w)). \tag{41}$$

for all $v, w \in S_0$, for some absolute constant A_0 and

$$\mathbf{Pr}(\mathcal{A}_{t_1}(v,w) \mid \mathcal{B}_{v,w}) = (1+o(1))\mathbf{Pr}(\mathcal{A}_{t_1}(v,w)) = (1+o(1))\mathbf{Pr}(\mathcal{A}_{t_1}(v))\mathbf{Pr}(\mathcal{A}_{t_1}(w)) \tag{42}$$

for almost all pairs $(v,w)\in S_0$.

It then follows that

$$\mathbf{E}(Z_0(Z_0-1)) \le (1+o(1))\mathbf{E}(Z_0)^2$$

and so

$$\mathbf{Pr}(Z_0\ne 0) \ge \frac{\mathbf{E}(Z_0)^2}{\mathbf{E}(Z_0^2)} = \frac{1}{\frac{\mathbf{E}(Z_0(Z_0-1))}{\mathbf{E}(Z_0)^2} + (\mathbf{E}Z_0)^{-1}} = 1-o(1)$$

from (37) and (42).

6.2.1 Proof of (41)

We argue that

$$R_\psi \le \frac{R_v+R_w}{2} + O(1/\log n) \tag{43}$$

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. 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 w. Returns to v,w account for the term $\frac{R_v+R_w}{2}$. We must also account for returns to ψ that come from walks from v to w and vice-versa. This can be overestimated by $R_v\eta(v,w)+R_w\eta(w,v)$, giving the $O(1/\log n)$ term.

Putting $\pi_v=\pi_w=\pi_0=\pi_\psi/2$, this implies that

$$\begin{aligned}
\frac{\pi_\psi}{R_\psi}-\frac{\pi_v}{R_v}-\frac{\pi_w}{R_w} &= \frac{\pi_0}{R_\psi R_v R_w}\left(2R_vR_w - R_\psi(R_v+R_w)\right)\\
&\ge \frac{\pi_0}{R_\psi R_v R_w}\left(2R_vR_w - \left(\frac{R_v+R_w}{2}+O(1/\log n)\right)(R_v+R_w)\right)\\
&= \frac{\pi_0}{2R_\psi R_v R_w}((R_v-R_w)^2 + O(1/\log n))\\
&= O\left(\frac{1}{n\log n}\right).
\end{aligned} \tag{44}$$

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

$$\begin{aligned}
\mathbf{Pr}_\psi(\mathcal{A}_{t_0}(\psi)) &= (1+o(1))\exp\left\{-\frac{t_0\pi_\psi}{(1+O(T\pi_\psi))R_\psi}\right\}\\
&= (1+o(1))\exp\left\{-\frac{t_0\pi_v}{R_v}\right\}\exp\left\{-\frac{t_0\pi_w}{R_w}\right\}\exp\left\{O\left(\frac{t_0}{n\log n}\right)\right\}\\
&= O(\mathbf{Pr}(\mathcal{A}_{t_0}(v))\mathbf{Pr}(\mathcal{A}_{t_0}(w))).
\end{aligned} \tag{45}$$

But, using rapid mixing in G_ψ,

$$\begin{aligned}
\mathbf{Pr}_\psi(\mathcal{A}_{t_0}(\psi)) &= \sum_{x\neq\psi} P_{\psi,u}^{T_\psi}(x)\mathbf{Pr}_\psi(\mathcal{W}_x(t-T_\psi)\neq\psi,\ T_\psi\leq t\leq t_0) \\
&= \sum_{x\neq\psi}\left(\frac{d(x)}{2m}+O(n^{-3})\right)\mathbf{Pr}_\psi(\mathcal{W}_x(t-T_\psi)\neq\psi,\ T_\psi\leq t\leq t_0) \\
&= \sum_{x\neq v,w}\left(P_u^{T_\psi}(x)+O(n^{-3})\right)\mathbf{Pr}(\mathcal{W}_x(t-T_\psi)\neq v,w,\ T_\psi\leq t\leq t_0) \qquad (46)\\
&= \mathbf{Pr}(\mathcal{W}_u(t)\neq v,w,\ T_\psi\leq t\leq t_0)+O(n^{-3}) \\
&= \mathbf{Pr}(\mathcal{A}_{t_0}(v,w))+O(n^{-3}). \qquad (47)
\end{aligned}$$

Equation (46) follows because there is a natural measure preserving map ϕ between walks in G that start at $x\neq v,w$ and avoid v,w and walks in G_ψ that avoid ψ. The map ϕ also shows that

$$\mathbf{Pr}(\mathcal{A}_{t_0}(v,w)\wedge\mathcal{B}_{v,w})=\mathbf{Pr}_\psi(\mathcal{A}_{t_0}(\psi)\wedge\mathcal{B}_\psi)=(1+o(1))\mathbf{Pr}_\psi(\mathcal{A}_{t_0}(\psi)\mid\mathcal{B}_\psi). \qquad (48)$$

But the argument for (36) can be used to show that

$$\mathbf{Pr}_\psi(\mathcal{A}_{t_0}(\psi)\mid\mathcal{B}_\psi)=(1+o(1))\mathbf{Pr}_\psi(\mathcal{A}_{t_0}(\psi)). \qquad (49)$$

Equation (41) follows from (45)–(49).

6.3 Proof of (42)

We get this sharpening of (41) whenever we can replace the $O(1/\log n)$ in (43) by $o(1/\log n)$. This replacement can be done whenever we can replace $O(1/\log n)$ in (38) by $o(1/\log n)$. We show that this can be done for almost all pairs $v,w\in S_0$.

There is a very simple argument when c is sufficiently large. The size of S_0 is $n^{\gamma_c+o(1)}$ **whp** where $\gamma_c=(c-1)\log\left(\frac{c}{c-1}\right)$. For any fixed $v\in S_0$ there are at most $T(\log n)^2$ vertices w such that $\eta(v,w)\geq 1/(\log n)^2$. If $\gamma_c>2/d$ then **whp** $T(\log n)^2=\tilde{O}(n^{2/d})=o(|S_0|)$ and (42) holds. For example, if $c\geq 2$ then $\gamma_c\geq\log 2=.69314718>2/d$ for $d\geq 3$.

So now we must consider the case where $1<c\leq 2$. Let A denote the set of *unordered* pairs $v,w\in S_0$ such that either $\eta(v,w)\geq 1/(\log n)^2$ or $\eta(w,v)\geq 1/(\log n)^2$. To prove (42) it is enough to show that

$$\mathbf{Pr}(\eta(v,w)\geq 1/(\log n)^2\mid v,w\in S_0,\ |v-w|\geq r^{1/2})=o(1). \qquad (50)$$

Here, if $v=(v_1,v_2,\ldots,v_d)$ then $|v|=(v_1^2+v_2^2+\cdots+v_d^2)^{1/2}$.

Note that the expected number of pairs $v, w \in S_0$ such that $|v - w| \leq r^{1/2}$ can be bounded by $\tilde{O}\left(\max\left\{n^{2\gamma_c - 1/2 + o(1)}, 1\right\}\right)$. So **whp** there are at most $\log n$ times this quantity. These pairs can therefore be ignored in our verification of (42).

To prove (50) we choose two points v, w for which $|v - w| \geq r^{1/2}$, condition on $v, w \in S_0$ and then bound $\mathbf{Pr}(\eta(v, w))$ from below. We condition on $v, w \in S_0$ by randomly placing k_1 points into each of $D(v, r), D(w, r)$. We then couple part of the remaining construction of G along with the first T steps $(x_1, x_2, ..., x_T)$ of the random walk $\mathcal{W}_v$. We weaken $x_T = w$ to $x_T \in D(w, r)$.

Let $i_v = \max\{i : x_i \in D(v, 2r)\}$. Notice next that for $i > i_v$, x_{i+1} is either (i) chosen from some previously exposed point or (ii) randomly chosen from $D(x_i, r)$ and these latter choices are made independently. Let J_0 be the set of indices $i > i_v$ where a choice is made according to (ii). Suppose now that $i \in J_0$ and $x_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,d})$ and let $y_{i,j} = x_{i,j} - x_{i-1,j}$. The $y_{i,j}$, $j = 1, 2, \ldots, d$ are not independent. Their sum of squares is at most r^2. On the other hand, if $B(x_i)$ is the cube of side $2r/d^{1/2}$ with centre x_i, then $B(x_i)$ is contained in $D(x_i, r)$. Conditional on $x_i \in B(x_{i-1})$, the $y_{i,j}$, $j = 1, 2, \ldots, d$ are independent. Let J_1 be the indices $i \in J_0$ for which $x_i \in B(x_{i-1})$. Then let $I_t = [t] \cap J_1$ for $t \geq 0$.

The size of I_t is $Bin(|J_0 \cap [t]|, q)$ where q is bounded away from 0. Furthermore, to reach w in t steps, we must have $|J_0 \cap [t]| \geq |v - w|/r - O(\log n) \geq r^{-1/2} - O(\log n)$. So, by use of the Chernoff bounds, we can assume that $|I_1| \geq r^{-1/2}q/2$. Now fix I_1 and condition on the values $y_{i,j}, i \notin I_1$ and let $Z_j = \sum_{i \in I_1} y_{i,j}$, $j = 1, 2, \ldots, d$. Now we have $Z_1, Z_2, \ldots, Z_d$ independent. Fix j. Then $Z_j = \sum_{l=1}^{s} \xi_l$ where **whp** $s \geq r^{-1/2}q/2$ and ξ_l is uniform in $[-r/d^{1/2}, r/d^{1/2}]$. As such it is well approximated by a normal distribution. In particular we can use the Berry-Esseen inequality, see for example [12]:

Let $X_1, X_2, \ldots, X_N$ be i.i.d. with $\mathbf{E}(X_i) = 0, \mathbf{E}(X_i^2) = \sigma^2$ and $\mathbf{E}(|X_i|^3) = \rho < \infty$. If $F_N(x)$ is the distribution of $(X_1 + X_2 + \cdots + X_N)/(\sigma\sqrt{N})$ and $\mathcal{N}(x)$ is the standard normal distribution, then

$$|F_N(x) - \mathcal{N}(x)| \leq \frac{3\rho}{\sigma^3\sqrt{N}}.$$

To have $x_t \in D(w, r)$ each Z_j will have to have to take a value in an interval A_j of length at most $2r$. This interval being determined by the values $x_i, i \notin I_t$. It follows from the Berry-Esseen inequality that $\mathbf{Pr}(Z_j \in A_j) = O(t^{-1/2})$. (We have $\sigma = \Omega(r)$ and $\rho = O(r^3)$). Hence, for some constant C,

$$\mathbf{E}(\eta(v, w)) \leq C \sum_{t=r^{-1/2}q/2}^{T} t^{-d/2} = O(r^{1/4})$$

and (50) and (42) follow.

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] C. Avin and G. Ercal, On the cover time and mixing time of random geometric graphs, *Theoretical Computer Science* 380 (2007) 2-22

[4] K. Ball, An elementary introduction to modern convex geometry, in *Flavors of Geometry*, MRI Publication 31, S. Levy Ed. (1997) 1-58.

[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 random regular graphs, *SIAM Journal on Discrete Mathematics*, 18 (2005) 728-740.

[7] C. Cooper and A. M. Frieze, The cover time of the preferential attachment graph. *Journal of Combinatorial Theory Series B*, Series B, 97(2) 269-290 (2007).

[8] C. Cooper and A. M. Frieze, The cover time of sparse random graphs, *Random Structures and Algorithms* 30 (2007) 1-16.

[9] C. Cooper and A. M. Frieze, The cover time of the giant component of a random graph, *Random Structures and Algorithms*, 32, 401-439 (2008).

[10] C. Cooper and A. M. Frieze, The cover time of random digraphs, Proceedings of RANDOM 2007, (11th. International Workshop on Randomized Techniques in Computation)

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

[12] R. Durrett, *Probability: Theory and Examples*, Brooks/Cole 1991.

[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] P. Gupta and P.R. Kumar, Critical power for asymptotic connectivity in wireless networks, *In Stochastic Analysis, Control, Optimization and Applications*, Birkhaüser, Boston, 1998.

[17] P. Gupta and P.R. Kumar, Internets in the sky: the capacity of three dimensional wireless networks. *Communications in Information and Systems*, 1 (2001) 33-50.

[18] 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

[19] M. D. Penrose. *Random Geometric Graphs.* Oxford University Press (2003).

[20] A. Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combinatorics, probability and Computing*, 1 (1992), 351-370.

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

[22] F. Xue and P. R. Kumar, Scaling laws for ad hoc wireless networks: an information theoretic approach. *Foundations and Trends in Networking*, 1(2) (2006) 145-270.

A Proof of (36)

Fix two vertices u, v. For $t \geq T$ let $h_t = \mathbf{Pr}(\mathcal{W}_u(t) = v \mid \mathcal{B}_v)$ be the conditional probability that the walk $\mathcal{W}_u$ visits v at step t. Let

$$H(z) = \sum_{t=T}^{\infty} h_t z^t \tag{51}$$

generate h_t for $t \geq T$.

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, conditional on the occurrence of $\mathcal{B}_v$. Let

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

generate f_t. Then we have

$$H(z) = F(z)R(z). \tag{52}$$

Write

$$R(z) = R_T(z) + \widehat{R}_T(z) + \frac{\pi_v z^T}{1-z}, \tag{53}$$

where $R_T(z)$ is given by (10) 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,

$$H(z) = \widehat{H}_T(z) + \frac{\pi_v z^T}{1-z} \tag{54}$$

where

$$\widehat{H}_T(z) = \sum_{t \geq T} (h_t - \pi_v) z^t.$$

Equation (9) implies that the radii of convergence of both $\widehat{R}_T$ and $\widehat{H}_T$ exceed $1+2\lambda$. (Although the h_t are conditional probabilities, we can still use (9) for $t \geq 2T$). Moreover, $|z| \leq 1+\lambda$,

$$|\widehat{R}_T(z)| = o(n^{-2}). \tag{55}$$

$$|\widehat{H}_T(z)| = o(1). \tag{56}$$

Equation (55) follows directly from the definition of T. This also implies that

$$\sum_{t \geq 2T} (h_t - \pi_v) z^t = o(n^{-2}).$$

Furthermore, $\sum_{t=T}^{2T} \pi_v z^{2T} = O(T\pi_v) = o(1)$ and

$$\sum_{t=T}^{2T} h_t z^t \leq (1+\lambda)^{2T} \sum_{t=T}^{2T} h_t \leq (1+\lambda)^{2T} \mathbf{Pr}(\mathcal{B}_v)^{-1} \max_{u'} R_{u',v} = o(1).$$

Here $\mathbf{Pr}(\mathcal{B}_v)^{-1}$ handles the conditioning on $\mathcal{B}_v$ and maximising over u' is maximising over $\mathcal{W}_u(T-1)$.

Using (53), (54) we rewrite $F(z) = H(z)/R(z)$ from (52) as $F(z) = B(z)/A(z)$ where

$$A(z) = \pi_v z^T + (1-z)(R_T(z) + \widehat{R}_T(z)), \tag{57}$$

$$B(z) = \pi_v z^T + (1-z)\widehat{H}_T(z). \tag{58}$$

For real $z \geq 1$ we have

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

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 (55) and $R_T(1) \geq 1$ imply 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

$$z_0 = 1 + \frac{\pi_v}{R_T(1)(1 + O(T\pi_v))} = 1 + p_v, \tag{59}$$

say. We also see that

$$A'(z_0) = -R_T(1)(1 + O(T\pi_v)) \neq 0 \tag{60}$$

and thus z_0 is a simple zero (see e.g. [5] p193). The value of $B(z)$ at z_0 is

$$B(z_0) = \pi_v(1 + o(1)) \neq 0. \tag{61}$$

We have used (56) here.

Thus,

$$\frac{B(z_0)}{A'(z_0)} = -(1 + O(T\pi_v))p_v. \tag{62}$$

Thus (see e.g. [5] p195) the principal part of the Laurent expansion of $F(z)$ at z_0 is

$$f(z) = \frac{B(z_0)/A'(z_0)}{z - z_0}. \tag{63}$$

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

$$M = \max_{z \in C_\lambda} |g(z)| = O(T\pi_v). \tag{64}$$

Let $a_t = [z^t]g(z)$, then (see e.g.[5] p143), $a_t = g^{(t)}(0)/t!$. By the Cauchy Inequality (see e.g. [5] 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

$$[z^t]F(z) = \frac{-B(z_0)/A'(z_0)}{z_0^{t+1}} + O(T\pi_v e^{-t\lambda/2}). \tag{65}$$

Thus, we obtain

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

Substituting $R_T(1) = 1 + o(1))$ and $\pi_v \sim \frac{c-1}{cn}$ completes the proof of (36).

A.1 Proof of (64)

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 (55)-(58),

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

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. [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 $\phi(z), \gamma(z)$ 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. □