

The cover time of random digraphs

Colin Cooper* Alan Frieze†

September 29, 2009

Abstract

We study the cover time of a random walk on the random digraph $D_{n,p}$ when $np = d \log n$, $d > 1$. We prove that **whp** the cover time is asymptotic to $d \log(d/(d-1)) \cdot n \log n$. To obtain this result we prove that **whp** the stationary distribution of any vertex v is given asymptotically by $\pi_v \sim \deg^-(v)/m$ where $\deg^-(v)$ is the in-degree of v and m is the number of edges of $D_{n,p}$. We note that if $d = d(n) \rightarrow \infty$ with n , the stationary distribution is asymptotically uniform.

1 Introduction

Let $D = (V, E)$ be a strongly connected digraph with $|V| = n$, and $|E| = m$. For a simple random walk $\mathcal{W}_v = (\mathcal{W}_v(t), t = 0, 1, \dots)$ on D starting at $v \in V$, let C_v be the expected time taken to visit every vertex of D . The *cover time* C_D of D is defined as $C_D = \max_{v \in V} C_v$.

For connected undirected graphs, the cover time is well understood, and has been extensively studied. It is an old result of Aleliunas, Karp, Lipton, Lovász and Rackoff [2] that $C_G \leq 2m(n-1)$. It was shown by Feige [9], [10], that for any connected graph G , the cover time satisfies $(1 - o(1))n \log n \leq C_G \leq (1 + o(1))\frac{4}{27}n^3$, where $\log n$ is the natural logarithm. An example of a graph achieving the lower bound is the complete graph K_n which 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 $2n/3$ has cover time asymptotic to the upper bound of $(4/27)n^3$.

For directed graphs cover time is less well understood, and there are strongly connected digraphs with cover time exponential in n . An example of this is the digraph consisting 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.

a directed cycle $(1, 2, \dots, n, 1)$, and edges $(j, 1)$, from vertices $j = 2, \dots, n - 1$. Starting from vertex 1, the expected time for a random walk to reach vertex n is 2^{n-2} .

In a sequence of papers we investigated the cover time of various classes of random graphs. The main results of these papers can be summarized as follows:

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

In this paper we turn our attention to random directed graphs. Let $D_{n,p}$ be the random digraph with vertex set $V = [n]$ where each possible directed edge (i, j) , $i \neq j$ is included with probability p . It is known that if $np = d \log n = \log n + \gamma$ where $\gamma = (d - 1) \log n \rightarrow \infty$ then $D_{n,p}$ is strongly connected **whp**. If γ as defined tends to $-\infty$ then **whp** $D_{n,p}$ is not strongly connected. (We do not have a direct reference to this result. It is easy to show that if $np = \log n - \gamma$ where $\gamma \rightarrow \infty$ then there are vertices of in-degree zero. On the other hand, if $np = \log n + \gamma$ where $\gamma \rightarrow \infty$ then [11] shows that the random digraph is Hamiltonian and hence strongly connected). Strong connectivity for $np = \log n + \gamma$ where $\gamma \rightarrow \infty$ also follows from the proof of Lemma 6. It is also easy to prove directly.

We determine the cover time of $D_{n,p}$ for values of p at or above the threshold for strong connectivity.

Theorem 1. *Let $np = d \log n$ where $d \log n = \log n + \gamma$ and $\gamma = (d - 1) \log n \rightarrow \infty$. Then **whp***

$$C_{D_{n,p}} \sim d \log \left(\frac{d}{d-1} \right) n \log n.$$

Note that if $d = d(n) \rightarrow \infty$ with n , then we have $C_{D_{n,p}} \sim n \log n$.

The method we use to find the cover time of $D_{n,p}$ requires us to know the stationary distribution of the random walk. For an undirected graph G , the stationary distribution is $\pi_v = \deg(v)/2m$, where $\deg(v)$ denotes degree of vertex v , and m is the number of edges in G . For a digraph D , let $\deg^-(v)$ denote in-degree of vertex v , $\deg^+(v)$ denote out-degree, and let m be the number of edges in D . For strongly connected digraphs in which each vertex v has in-degree equal to out-degree ($\deg^-(v) = \deg^+(v)$), then $\pi_v = \deg^+(v)/m$. For general digraphs, however, there is no simple formula for the stationary distribution, and the main

technical task of this paper is to find good estimates for π_v in the case of $D_{n,p}$. We summarize our result concerning the stationary distribution in the following theorem.

We summarize our result concerning the stationary distribution in Theorem 2 below. For a given vertex v , define a quantity $\zeta^*(v)$, which characterizes the in-neighbour w of v with minimum out-degree as follows:

$$\zeta^*(v) = \min_{\deg^+(w)} \left(\frac{\deg^-(w)}{\deg^+(w)} : w \in N^-(v) \right). \quad (1)$$

Theorem 2. *Let $np = d \log n$ where $d \log n = \log n + \gamma$. Let $n \rightarrow \infty$ and assume $\gamma \rightarrow \infty$. Then **whp**, for all $v \in V$,*

$$\pi_v = \frac{(1 + o(1))}{m} (\deg^-(v) + \zeta^*(v)),$$

where $\deg^-(v)$ is the in-degree of v and m is the number of edges of $D_{n,p}$, and $\epsilon_v \in [0, 1]$.

We note the following special cases.

Remark 1. *We prove in Lemma 12 that **whp** $\zeta^*(v) = o(\deg^-(v))$ for almost all vertices v . For these vertices, the $\zeta^*(v)$ term can be absorbed into the error term of π_v .*

Remark 2. *If $\gamma = \omega(\log \log n)$ then **whp** $\zeta^*(v) = o(\deg^-(v))$ for all vertices v , and in particular when $d = 1 + \delta$, $\delta > 0$ then the minimum out-degree is $\Omega(\log n)$. In these cases, $\pi_v \sim \deg^-(v)/m$.*

Remark 3. *If $d = d(n) \rightarrow \infty$ with n , **whp** the stationary distribution of $D_{n,p}$ is $\pi_v \sim 1/n$.*

Our analysis is based on Lemma 3 which is given in Section 2. This lemma is proved in its current state in [7]. It provides a good estimate of the probability that a walk starting at vertex u does not visit another vertex v in steps $T, T + 1, \dots, t$ where T is a mixing time. In order to apply the lemma we need an estimate of the stationary distribution π_v of the random walk for all $v \in V$. We also need an estimate of T , and of a parameter $R_T(1)$ which is the expected number of returns made to v during the first T steps by a walk starting from v .

As previously remarked, the main problem is to estimate the stationary distribution, and we will briefly describe our technique for doing this. Once we know the stationary distribution, it is basically plain sailing. We establish that the conditions for Lemma 3 hold, and using this lemma we apply the methods used in previous papers such as [7].

We approximate the stationary distribution $\boldsymbol{\pi}$ using the expression $\boldsymbol{\pi} = \boldsymbol{\pi} P^k$. To do this we need the value of P^k , where P is the transition matrix. By a suitable choice of k we find we can bound

$$P_x^{(k)}(y) = \mathbf{Pr}(\mathcal{W}_x(k) = y)$$

from above and below by values independent of x , i.e. $P_x^{(k)}(y) \sim \theta_y = \frac{\deg^-(y)}{m}$, where the value for θ_y is verified in Section 5. Here $\deg^-(y)$ is the in-degree of y and m is the number of edges of $D_{n,p}$. We next give a brief intuitive picture of how this is done.

Let $k = 2\ell + 1$. We grow an out-branching T_X from x to depth ℓ , and an in-branching T_Y to y of depth ℓ , and join the top level of these branchings by random edges. By careful choice of ℓ these branchings are almost tree-like. Suppose that vertex w lies in the ℓ -th level of the out-branching T_X . For each path $P_w = (x = v_0, v_1, \dots, v_\ell = w)$ from x to w , the probability that \mathcal{W}_x follows this path for the first ℓ steps is $1/(\deg^+(v_0)\dots\deg^+(v_{\ell-1}))$. We estimate the sum of these probabilities over all vertices w in the ℓ -th level of the out-branching T_X as a random variable. This is reasonably straightforward. We need to do this for T_Y as well and this is more delicate. We approximate the required sum of probabilities from above and below by random variables W_t^1, W_t^0 that are defined iteratively. This is Lemma 5, described in Section 4. In addition, the upper bound analysis requires us to deal with error terms that can be ignored for the lower bound.

The structure of the paper is therefore as follows: Section 2 describes Lemma 3. Section 3 provides a technical lemma about the degree sequence of $D_{n,p}$. Section 4 describes Lemma 5. Section 5 deals with estimating the stationary distribution and forms the main body of this paper. Section 6 is short and gives a bound on the mixing time, using the results of Section 5. We then estimate the cover time asymptotically in Section 7.

1.1 Chernoff Bounds

The following inequalities will be used for the sum $Z = Z_1 + Z_2 + \dots + Z_N$ of independent random variables $0 \leq Z_i \leq 1$, $i = 1, 2, \dots, N$ with $\mathbf{E}(Z_1 + Z_2 + \dots + Z_N) = N\mu$:

$$\mathbf{Pr}(|Z - N\mu| \geq \epsilon N\mu) \leq 2e^{-\epsilon^2 N\mu/3}. \quad (2)$$

$$\mathbf{Pr}(Z \geq \alpha N\mu) \leq (e/\alpha)^{\alpha N\mu}. \quad (3)$$

2 Main Lemma

In this section D denotes a fixed strongly connected digraph with n vertices. 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 D is ergodic with stationary distribution π .

Let

$$d(t) = \max_{u,x \in V} |P_u^{(t)}(x) - \pi_x|,$$

and let T be a positive integer such that for $t \geq T$

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

Fix two vertices u, v . 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 $t = 0, 1, \dots$. Let

$$R_T(z) = \sum_{j=0}^{T-1} r_j z^j \quad (5)$$

and

$$\lambda = \frac{1}{KT} \quad (6)$$

for a sufficiently large constant K .

For $t \geq T$ let $\mathbf{A}_v(t)$ be the event that \mathcal{W}_u does not visit v in steps $T, T+1, \dots, t$.

Lemma 3. *Suppose that*

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

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

(b) $T^2\pi_v = o(1)$ and $T\pi_v = \Omega(n^{-2})$ for all $v \in V$.

There exists

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

where $R_T(1)$ is from (5), such that for all $v \in V$ and $t \geq T$,

$$\mathbf{Pr}(\mathbf{A}_v(t)) = \frac{(1 + O(T\pi_v))}{(1 + p_v)^t} + O(T^2\pi_v e^{-\lambda t/2}). \quad (8)$$

A similar lemma was first proved in [5]. It was amended slightly for [7] and this latter version is the one quoted above.

3 Degree Sequence

We state a simple lemma concerning the degree sequence of $D_{n,p}$. The lemma can be proven by the use of the first and second moment method (see [4] for similar calculations).

Let $np = d \log n$ and let

$$\Delta_0 = 30np. \quad (9)$$

Lemma 4. Assume that $np = d \log n$ where $1 < d = O(1)$ and $(d - 1) \log n \rightarrow \infty$. Let

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

denote the expected number of vertices v with $\deg^-(v) = k$. Let $D(k)$ denote the actual number and

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

The degree sequence has the following properties.

(a) If $d - 1 \geq (\log n)^{-1/3}$ then **whp**

$$K_1 = \emptyset, \quad \min\{k \in K_2\} \geq (\log n)^{1/2}, \quad |K_2| = O(\log \log n).$$

(b) For all degrees $k \in [1, \Delta_0]$ the following hold **whp**.

$$\begin{aligned} k \in K_0, & \quad D(k) = 0, \\ k \in K_1, & \quad D(k) \leq (\log \log n)^2, \\ k \in K_2, & \quad D(k) \leq (\log n)^4, \\ k \in K_3, & \quad \frac{\overline{D}(k)}{2} \leq D(k) \leq 2\overline{D}(k). \end{aligned}$$

(c) Let $k^* = (d - 1) \log n$. Let $\gamma_d = (d - 1) \log(d/(d - 1))$. **Whp** there are at least $n^{\gamma_d} / (3\pi \log n (d(d - 1))^{1/2})$ vertices v with $\deg^-(v) = k^*$ and $\deg^+(v) = d \log n$.

(d) Let \mathcal{D} be the event

$$\{\exists v \in V : \deg^+(v) > \Delta_0 \text{ or } \deg^-(v) > \Delta_0\}, \quad (10)$$

then

$$\Pr(\mathcal{D}) \leq n^{-10} e^{-10np}. \quad (11)$$

Note. From now until Section 5.3, we will work under

$$\textbf{Assumption 1 :} \quad d \geq 2. \quad (12)$$

(Note that 2 is somewhat arbitrary here. Any constant larger than 1 will suffice.)

Under Assumption 1, there is a constant $c > 0$ and an interval

$$I = [cnp, Cnp], \quad (13)$$

where $C = 30$, such that if $\nu \in [n - n^{67}, n]$ then

$$\Pr(\text{Bin}(\nu, p) \in I) = 1 - o(n^{-1}). \quad (14)$$

Let \mathcal{E}_S be the event that the in-degree and out-degree of all vertices in $S \subseteq V$ are in the interval I . Then for any $S \subseteq V$ we have

$$\Pr(\mathcal{E}_S) = 1 - o(1). \quad (15)$$

4 A Useful Lemma

The *weight* of a path P of length ℓ from vertex z to y is the probability that the random walk \mathcal{W}_z reaches y in exactly ℓ steps by following P . The following lemma is used to estimate the stationary distribution of the random walk. It gives a good approximation to the weight of (the paths in) an in-tree T_Y of height ℓ rooted at a vertex y . The precise construction of the tree T_Y is described at the beginning of Section 5.1 below.

We introduce a random variable $W_t = W_{t, \sigma}$ to model the weight of T_Y . Here $t \leq \ell$ is a non-negative integer and σ is an arbitrary sequence σ_i of positive integers in

$$I_0 = [n - n^{67}, n].$$

We note that σ is just a convenient notational device for saying that we have some freedom in choosing the number of trials in each of a collection of binomial random variables.

Initially, when $t = 0$, σ is empty. For given t , the entries in σ list the values of σ_i used in the recurrence (16) given below, along with all entries of the σ_i from previous applications of the recurrence. The exact values of σ_i or the length of σ are not relevant to our construction. The value of σ_i can differ at each application of (16).

Lemma 5. *Let $W_{0, \sigma} = 1$ for all σ . For $t \geq 1$, let B_t be an independent copy of $\min\{\text{Bin}(\sigma_0, p), \Delta_0\}$, where $\sigma_0 \in I_0$ and Δ_0 is given by (9). For $i = 1, \dots, B_t$, let A_i be an independent copy of W_{t-1, σ_i} , and let $D_{i,t}$ be an independent copy of $1 + \text{Bin}(\sigma_i, p)$ where $\sigma_i \in I_0$ for $1 \leq i \leq B_t$. If $B_t = 0$ let $W_t = 0$ and if $B_t \geq 1$ let*

$$W_t = W_{t, \sigma} = \sum_{i=1}^{B_t} \frac{A_i}{D_{i,t}}. \quad (16)$$

Suppose that $M > 1$, $Mt = o(np)$ and $np = o(n^{33})$. Then for $1 \leq |\lambda| \leq M$,

$$\mathbf{E}(e^{\lambda W_t}) \leq \exp \left\{ \lambda + \frac{5M|\lambda|t}{np} \right\}. \quad (17)$$

Proof We proceed by induction on t . The claim is true for $t = 0$. Let $q = 1 - p$ and let $t \geq 1$, then

$$\mathbf{E}(e^{\lambda W_t}) \leq \sum_{k=0}^{\sigma_0} \binom{\sigma_0}{k} p^k q^{\sigma_0-k} \prod_{i=1}^k \mathbf{E}(e^{\lambda A_i/D_{i,t}}) + \mathbf{Pr}(Bin(\sigma_0, p) \geq \Delta_0) \prod_{i=1}^{\Delta_0} \mathbf{E}(e^{\lambda A_i/D_{i,t}}) \quad (18)$$

where

$$\mathbf{E}(e^{\lambda A_i/D_{i,t}}) = \sum_{l=0}^{\sigma_i} \binom{\sigma_i}{l} p^l q^{\sigma_i-l} \mathbf{E}(e^{\lambda A_i/(l+1)}).$$

Now

$$\begin{aligned} \sum_{l=0}^{\sigma_i} \frac{1}{l+1} \binom{\sigma_i}{l} p^l q^{\sigma_i-l} &= \frac{1}{(\sigma_i+1)p} \sum_{l=0}^{\sigma_i} \binom{\sigma_i+1}{l+1} p^{l+1} q^{\sigma_i-l} = \\ &= \frac{1}{(\sigma_i+1)p} (1 - (1-p)^{\sigma_i+1}) = \frac{1 + O(n^{-.33})}{np}. \end{aligned}$$

and

$$\begin{aligned} \sum_{l=0}^{\sigma_i} \frac{1}{(l+1)^2} \binom{\sigma_i}{l} p^l q^{\sigma_i-l} &\leq 2 \sum_{l=0}^{\sigma_i} \frac{1}{(l+1)(l+2)} \binom{\sigma_i}{l} p^l q^{\sigma_i-l} \\ &= \frac{2}{(\sigma_i+1)(\sigma_i+2)p^2} \sum_{l=0}^{\sigma_i} \binom{\sigma_i+2}{l+2} p^{l+2} q^{\sigma_i-l} = \\ &= \frac{2}{(\sigma_i+1)(\sigma_i+2)p^2} (1 - (1-p)^{\sigma_i+2} - (\sigma_i+2)p(1-p)^{\sigma_i+1}) = \frac{2 + O(n^{-.33})}{n^2 p^2}. \end{aligned}$$

Applying induction, and using $e^x \leq 1 + x + x^2$ for $|x| \leq 1$ we see that,

$$\begin{aligned} \mathbf{E}(e^{\lambda A_i/D_{i,t}}) &\leq \sum_{l=0}^{\sigma_i} \binom{\sigma_i}{l} p^l q^{\sigma_i-l} \exp \left\{ \frac{\lambda}{l+1} + \frac{5M|\lambda|(t-1)}{(l+1)np} \right\} \\ &\leq \sum_{l=0}^{\sigma_i} \binom{\sigma_i}{l} p^l q^{\sigma_i-l} \left(1 + \frac{1}{l+1} \left(\lambda + \frac{5M(t-1)|\lambda|}{np} \right) + \frac{2\lambda^2}{(l+1)^2} \right) \\ &\leq 1 + \frac{\lambda(1 + O(n^{-.33}))}{np} + \frac{5M(t-1)|\lambda|(1 + O(n^{-.33}))}{n^2 p^2} + \frac{4\lambda^2(1 + O(n^{-.33}))}{n^2 p^2} \\ &\leq 1 + \frac{\lambda}{np} + \frac{|\lambda|M(5t-2)}{n^2 p^2} + O\left(\frac{\lambda}{n^{1.33}p}\right). \end{aligned} \quad (19)$$

We observe then that $\mathbf{E}(e^{\lambda A_i/D_{i,t}}) = 1 + o(1)$ and so

$$\mathbf{Pr}(Bin(\sigma_0, p) \geq \Delta_0) \prod_{i=1}^{\Delta_0} \mathbf{E}(e^{\lambda A_i/D_{i,t}}) \leq e^{-10np} (1 + o(1))^{\Delta_0} \leq e^{-5np}. \quad (20)$$

Plugging (19) and (20) into (18) we get

$$\begin{aligned}
\mathbf{E}(e^{\lambda W_t}) &\leq \sum_{k=0}^{\sigma_0} \binom{\sigma_0}{k} p^k q^{\sigma_0-k} \left(1 + \frac{\lambda}{np} + \frac{|\lambda|M(5t-2)}{n^2 p^2} + O\left(\frac{\lambda}{n^{1.33}p}\right)\right)^k + e^{-5np} \\
&= \left(1 + \frac{\lambda}{n} + \frac{M(5t-2)|\lambda|}{n^2 p} + O\left(\frac{\lambda}{n^{1.33}}\right)\right)^{\sigma_0} + e^{-5np} \\
&\leq \exp\left\{\lambda + \frac{|\lambda|M(5t-1)}{np}\right\} (1 + e^{|\lambda|-5np}) \\
&\leq \exp\left\{\lambda + \frac{5M|\lambda|t}{np}\right\}
\end{aligned}$$

which is (17). \square

We will apply Lemma 5 as follows. From (17) we have that $\mathbf{E}(e^{MW_t}) \leq e^{M+\varsigma M}$ and $\mathbf{E}(e^{-MW_t}) \leq e^{-M+\varsigma M}$ for $\varsigma = \frac{5t}{cnp}$. So, if $S_k = X_1 + \dots + X_k$ is the sum of k independent copies of W_t , then

$$\Pr(S_k \leq a) \leq e^{Ma} \mathbf{E}(e^{-MS_k}) \leq e^{-kM(1-\varsigma)+Ma}, \quad (21)$$

$$\Pr(S_k \geq b) \leq e^{-Mb} \mathbf{E}(e^{MS_k}) \leq e^{kM(1+\varsigma)-Mb}. \quad (22)$$

The recursive computation of W_t as in (16) gives rise to a randomly weighted tree that we denote by T_{W_t} . If $t = 0$ then T_{W_t} is a single vertex. Otherwise, we have a root of degree B_t with children T_{A_i} for $i = 1, 2, \dots, B_t$. The weight of the root ρ of T_{W_t} is one. The weight of the i th child v_i of ρ will be $1/D_{i,t}$ from (16). The children of v_i will have weights $1/D_{j,t-1}$ and so on. I.e. creation of a vertex involves instantiation of a random variable $D = 1 + \text{Bin}(\sigma, p)$ for some $\sigma \in I_0$ and the weight of the vertex is then $1/D$.

If P is a path from a leaf to the root of T_{W_t} then its weight is defined to be the product of the weights of the vertices on the path. With this definition, the random variable W_t is equal to the sum of the weights of the leaf to root paths.

Furthermore, inequality (17) remains true if we evaluate W_t in a top down approach, revealing σ as is needed i.e level by level in the tree T_{W_t} .

5 Estimating the Stationary Distribution

We approximate the stationary distribution π as follows: Iterating the equation $\pi = \pi P$, k times gives $\pi = \pi P^k$. For fixed y this gives $\pi_y = \sum_{x \in V} \pi_x P_x^{(k)}(y)$. By bounding $P_x^{(k)}(y)$ from above and below by values independent of x , i.e. $P_x^{(k)}(y) \sim \theta_y$ we obtain $\pi_y \sim \theta_y$.

5.1 Lower Bound on the stationary distribution

To bound $P_x^{(k)}(y)$ from below, we use walks between x and y consisting of simple directed (x, y) -paths of length k . Let $np = d \log n$. The proof consists of four cases, depending on the value of d .

Case O: $1 + o(1) \leq d \leq 2$

Case I: $d \geq 2$ and $d = O(1)$ or $d \rightarrow \infty$ and $d \leq n^{3/10}$.

Case II: $n^{3/10} \leq d \leq n^{3/5}$.

Case III: $d \geq n^{3/5}$.

Until Section 5.3, we work under Assumption 1 that $d \geq 2$. Case O is treated separately, in Section 5.3. Most of the work in Lemma 6 below is in the proof of Case I. The proof of Lemma 6 uses Lemma 5. To apply this lemma, we need a careful construction of an in-tree rooted at vertex y .

Under Assumption 1, the term $\zeta(v)$ of Theorem 2 gets absorbed into the ϵ_v error term.

Lemma 6. *There exists an integer $s = O(\log_{np} n)$ such that **whp** for all $x, y \in V$,*

$$P_x^{(s)}(y) \geq (1 - o(1)) \frac{\text{deg}^-(y)}{m} \text{ for all } v \in V,$$

where m is the number of edges in $D_{n,p}$.

Proof

Case I: Let

$$\ell = \left\lfloor \frac{2}{3} \log_{np} n \right\rfloor. \tag{23}$$

For a vertex v let $N^-(v)$ be the set of in-neighbours of v and for a set S , let $N^-(S) = \bigcup_{v \in S} N^-(v)$. Define $N^+(v), N^+(S)$ similarly with respect to out-neighbours.

Construction of T_Y . For fixed $y \in V$, we build a tree T_Y rooted at y , a truncated breadth-first fashion. T_Y has level sets $Y_i, i = 0, 1, \dots, \ell$, and vertex set $Y = \bigcup_{i=0}^{\ell} Y_i$.

Define $Y_0 = \{y\}$. We next describe how to construct Y_{i+1} from Y_i for $i \geq 0$. We process the vertices of Y_i in increasing label order $v_1, v_2, \dots, v_{|Y_i|}$. Initially $\widehat{N}_i = \emptyset$. For $v_k \in Y_i$ let

$$I(v_k) = V \setminus (Y_0 \cup \dots \cup Y_i \cup \widehat{N}_i).$$

When examining the in-neighbours of v_k we retain the first (at most) Δ_0 in-neighbours in $I(v_k)$ in the natural label order. Let $N^*(v_k)$ be this set of retained vertices and update $\widehat{N}_i := \widehat{N}_i \cup N^*(v_k)$. As $\Delta_0^\ell = O(n^{2/3+o(1)})$ we have that $|Y_0 \cup Y_1 \dots \cup Y_\ell| = o(n^{0.67})$, and the in-degree of v_k in $I(v_k)$ fits the definition of B_t in Lemma 5.

The construction above means that for $w \in Y_{i+1}$ there is a unique edge (w, v) in T_Y to some $v = v_k$, in Y_i . Let $S(w) \subseteq Y(i)$ be the vertices $\{v_1, \dots, v_k\}$. Let $Y^*(w) = Y_0 \cup Y_1 \cup \dots \cup Y_{i-1} \cup S(w)$. Then the only edge from w to $Y^*(w)$ is (w, v_k) . The distribution of the other out-edges of w is unknown. Because $|Y^*(w)| = o(n^{0.67})$ the distribution of out-edges of w fits the definition of $D_{i,t}$ in Lemma 5.

Construction of T_X . Given Y and $x \in V$, ($x = y$ is allowed) we define $X_0 = \{x\}$, X_1, \dots, X_ℓ where $X_{i+1} = N^+(X_i) \setminus (Y \cup X_0 \cup \dots \cup X_i)$ for $0 \leq i < \ell$.

Let $X = \bigcup_{i=0}^{\ell} X_i$ and let T_X denote the BFS tree constructed in this manner. If $w \in X_{i+1}$ is the out-neighbour of more than one vertex of X_i , we only keep the edge (z, w) with z as small as possible for T_X . Let $X = \{x_0 = x, x_1, \dots, x_N\}$ where x_i is the i -th vertex added to T_X .

For $u \in X_i$ let P_u denote the path of length i from x to u in T_X and

$$\alpha_{i,u} = \prod_{\substack{w \in P_u \\ w \neq u}} \frac{1}{\deg^+(w)} \leq \Pr(\mathcal{W}_x(i) = u).$$

Similarly, for $v \in Y_i$ let Q_v denote the path from v to y in T_Y and

$$\beta_{i,v} = \prod_{\substack{w \in Q_v \\ w \neq y}} \frac{1}{\deg^+(w)} \leq \Pr(\mathcal{W}_v(i) = y). \quad (24)$$

Given $D_{n,p}$ we have

$$P_x^{(2\ell+1)}(y) \geq Z = Z(x, y) = \sum_{\substack{u \in X_\ell \\ v \in Y_\ell}} \alpha_{\ell,u} \beta_{\ell,v} \frac{1_{uv}}{\deg^+(u)} \quad (25)$$

where 1_{uv} is the indicator for the existence of the edge (u, v) and we take $\frac{1_{uv}}{\deg^+(u)} = 0$ if $\deg^+(u) = 0$.

Let $C = C(x, y)$ denote X_i, Y_i , $0 \leq i \leq \ell$ and the collection of edge sets $(X_{i-1} : X_i)$, $1 \leq i \leq \ell$ in T_X (resp. $(Y_i : Y_{i-1})$, $1 \leq i \leq \ell$ in T_Y).

Now, at this point we know that $u \in X_\ell$ does not have neighbours in $T_Y \setminus Y_\ell$. Other possible edges are unconditioned. So the distribution of $\deg^+(u)$ is $\text{Bin}(\nu, p)$ where $n - n^{.67} \leq \nu \leq n - 1$. Thus

$$\mathbf{E} \left(\frac{1_{uv}}{\deg^+(u)} \middle| C \right) = \sum_{k=1}^{\nu} \binom{\nu}{k} p^k (1-p)^{\nu-k} \frac{k}{\nu} \frac{1}{k} = \frac{1}{n} (1 + O(n^{-.33})). \quad (26)$$

We therefore have,

$$\mathbf{E}(Z | C) = \frac{1 + o(1)}{n} \left(\sum_{u \in X_\ell} \alpha_{\ell,u} \right) \left(\sum_{v \in Y_\ell} \beta_{\ell,v} \right). \quad (27)$$

We will show next that for some $\epsilon_X, \epsilon_Y = o(1)$ and \mathcal{E} as in (15),

$$\Pr\left(\sum_{u \in X_\ell} \alpha_{\ell,u} < 1 - \epsilon_X \mid \mathcal{E}_{X \setminus X_\ell}\right) = o(n^{-10}), \quad (28)$$

$$\Pr\left(\sum_{v \in Y_\ell} \beta_{\ell,v} < (1 - \epsilon_Y) \frac{\deg^-(y)}{np}\right) = o(n^{-1}). \quad (29)$$

Note that there is one inequality (28) needed for each $x, y \in V$ and one inequality (29) needed for each $y \in V$. For a given y , the latter inequality is independent of x .

Proof of (29): For each $v_i \in Y_i$ (the i th level of T_Y) and $j < i$ with path $P = (v_i, v_{i-1}, \dots, v_j)$ in T_Y from v_i to $v_j \in Y_j$, let

$$\gamma_{i,j;v_i,v_j} = \prod_{k=j+1}^i \frac{1}{\deg^+(v_k)}.$$

Given $i < j \leq \ell$ and $v \in Y_i$ we let

$$Y_{v,j} = \{w \in Y_j : \text{the path from } w \text{ to } y \text{ in } T_Y \text{ goes through } v\}. \quad (30)$$

Let

$$\sum_{v \in Y_\ell} \beta_{\ell,v} = \sum_{w \in Y_\ell} \gamma_{\ell,0;w,y} = \sum_{v \in Y_1} \frac{S_v}{\deg^+(v)}, \quad (31)$$

where $Y_{v,\ell}$ is defined in (30), and

$$S_v = \sum_{w \in Y_{v,\ell}} \gamma_{\ell,1;w,v}.$$

Recall that $np = d \log n$, $d > 1$. Let $I = [cnp, Cnp]$ (see (13)). If $\deg^-(y) \in I$, we say y is *centered*. Let \mathcal{E}_y be the event that y is centered. Then from (14), $\Pr(-\mathcal{E}_y) = o(n^{-1})$.

Let $\zeta = 1/\log \log \log n$. A vertex is *normal* if at most $\zeta_0 = \lceil 4/(\zeta^3 d) \rceil$ of its in-neighbours have out-degrees which are not in the range $[(1 - \zeta)np, (1 + \zeta)np]$. Let $\mathcal{N}(y)$ be the event y is normal. If y is centered, the probability y is not normal is at most

$$\sum_{s=cnp}^{Cnp} \binom{s}{\zeta_0} (2e^{-\zeta^2 np/3})^{\zeta_0} = O(n^{-\Omega(\log \log \log n)}).$$

Thus

$$\Pr(-(\mathcal{N}(y), \mathcal{E}_y)) = o(n^{-1}). \quad (32)$$

Let $d_y^- = \deg^-(y)$. If y is normal and centered, and $\epsilon_Y \geq 2\zeta$ then, referring to (31),

$$\Pr \left(\sum_{v \in Y_\ell} \beta_{\ell,v} \leq (1 - \epsilon_Y) \frac{d_y^-}{np} \right) \leq \Pr \left(\sum_{j=1}^{d_y^- - \zeta_0} \frac{S_j}{(1 + \zeta)np} \leq (1 - \epsilon_Y) \frac{d_y^-}{np} \right) \quad (33)$$

$$\leq \Pr \left(\sum_{j=1}^{d_y^- - \zeta_0} S_j \leq (1 - \epsilon_Y/2)d_y^- \right). \quad (34)$$

From the construction of T_Y , and the construction of T_{W_ℓ} given following Lemma 5, we see that there will be some sequence σ_v such that the S_v are distributed as independent copies of $A_v = W_{\ell-1, \sigma_v}$. We can then use (21) to estimate the probability (34).

For d constant, let $\epsilon_Y = 2\zeta$ and $M = 10/(\epsilon_Y)$. In this case $\frac{M\ell}{np} = O\left(\frac{\log \log \log n}{\log \log n}\right) = o(1)$ and Lemma 5 is applicable. For $np = \omega \log n$, $\omega \rightarrow \infty$ let $\epsilon_Y = 1/\omega^{1/3}$ and $M = 1$. In this case $\frac{M\ell}{np} = \frac{1}{\omega^{1/3} \log \log n} = o(1)$ and Lemma 5 is again applicable.

We now complete the proof of (29).

$$\begin{aligned} \Pr \left(\sum_{v \in Y_\ell} \beta_{\ell,v} \leq (1 - \epsilon_Y) \frac{d_y^-}{np} \right) &\leq \Pr \left(\sum_{j=1}^{d_y^- - \zeta_0} A_j \leq \left(1 - \frac{\epsilon_Y}{2}\right) d_y^- \right) + \Pr(\neg(\mathcal{N}(y), \mathcal{E}_y)) \\ &\leq \max_{d_y^- \in I} \left(e^{-M((d_y^- - \zeta_0)(1 - \frac{5(\ell-1)}{np}) - (1 - \epsilon_Y/2)d_y^-)} \right) + o(n^{-1}) \\ &\leq e^{-M\epsilon_Y \log n/3} + o(n^{-1}) \\ &= o(n^{-1}). \end{aligned} \quad (35)$$

Proof of (28): For $u \in X_\ell$ let $P_u = P_{u:\ell} = (u_0 = x, u_1, \dots, u_\ell = u)$ denote the path from x to u in T_X . Let $\deg_X^+(v) = |N^+(v) \cap X|$ denote the out-degree of v in T_X . For the related walk on the digraph T_X , starting at x ; X_ℓ is reached with probability $\Phi = 1$ in exactly ℓ steps, after which the walk halts. Thus

$$1 = \Phi = \sum_{u \in X_\ell} \prod_{\substack{v \in P_u \\ v \neq u}} \frac{1}{\deg_X^+(v)} \geq \sum_{u \in X_\ell} \alpha_{\ell,u}. \quad (36)$$

There is an assumption that $\deg_X^+(v) > 0$ for $v \in X$ and this will be justified **whp** below, see (39).

For $x_j \in X$ let $f(x_j) = |N^+(x_j) \cap (Y \cup \{x_0, x_1, \dots, x_{j-1}\})|$ so that $\deg^+(v) = \deg_X^+(v) + f(v)$.

Now

$$\begin{aligned}
\Phi &= \sum_{u \in X_\ell} \prod_{v \in P_u} \frac{1}{\deg^+(v) - f(v)} \\
&= \sum_{u \in X_\ell} \left(\prod_{v \in P_u} \frac{1}{\deg^+(v)} \right) \left(\prod_{v \in P_u} \frac{1}{1 - f(v)/\deg^+(v)} \right) \\
&= \sum_{u \in X_\ell} \alpha_{\ell,u} \left(\prod_{v \in P_u} \frac{1}{1 - f(v)/\deg^+(v)} \right)
\end{aligned}$$

Now if

$$\prod_{v \in P_u} \frac{1}{1 - f(v)/\deg^+(v)} \leq 1 + h \quad \forall u \in X_\ell, \quad (37)$$

then $\sum_{u \in X_\ell} \alpha_{\ell,u} = 1 - o(1)$ provided $h = o(1)$. Under the assumption $f(v)/\deg^+(v) < 1$ we have

$$\frac{1}{1 - f(v)/\deg^+(v)} \leq \exp \left(\frac{f(v)}{\deg^+(v) - f(v)} \right). \quad (38)$$

For $u \in X$ we can bound $\sum_{v \in P_u} f(v)$ by the binomial $\text{Bin}(N_X, p)$ where $N_X = |Y| + |X|$. Assuming $\mathcal{E}_{X \cup Y}$, see (14), we have $N_X = n^{2/3+o(1)}$. Recall that $np \leq n^{3/10}$ (Case I). Using the Chernoff bound (3), we have that

$$\begin{aligned}
\Pr \left(\sum_{v \in P_u} f(v) \geq \frac{np}{\omega} \middle| \mathcal{E}_{(X \setminus X_\ell) \cup (Y \setminus Y_\ell)} \right) &\leq \Pr \left(\text{Bin}(n^{2/3+o(1)}, p) \geq \frac{np}{\omega} \middle| \mathcal{E}_{(X \setminus X_\ell) \cup (Y \setminus Y_\ell)} \right) \\
&\leq (1 + o(1)) \Pr \left(\text{Bin}(n^{2/3+o(1)}, p) \geq \frac{np}{\omega} \right) \\
&\leq n^{-\frac{1}{100} \frac{np}{\omega}}.
\end{aligned}$$

So with $\omega = o(np)$, with conditional probability at least $1 - n^{-np/(100\omega)}$,

$$\text{(i)} \quad \frac{f(v)}{\deg^+(v)} \leq \frac{1}{c\omega} \quad \text{for } v \in P_u. \quad (39)$$

$$\text{(ii)} \quad \sum_{v \in P_u} \frac{f(v)}{\deg^+(v) - f(v)} \leq \frac{1}{c\omega - 1}.$$

Going back to (38) we see from (ii) that with (conditional) probability at least $1 - n^{-np/(100\omega)}$

$$\prod_{v \in P_u} \frac{1}{1 - f(v)/\deg^+(v)} \leq \exp \left(\frac{1}{c\omega - 1} \right) = 1 + O \left(\frac{1}{\omega} \right). \quad (40)$$

There are at most n trees and n paths per tree and so (37), with $h = O(1/\omega) = o(1)$, follows from (40). This completes the proof of (28).

Evaluation of (25): Let $\mathcal{C} = \mathcal{E}_{X \setminus X_\ell} \cap \mathcal{E}_{Y \setminus Y_\ell}$ so that $\Pr(C \in \mathcal{C}) = 1 - o(1)$ (see (14)). In order to apply (27) we next establish the concentration of Z given $C \in \mathcal{C}$. If $u \in X_\ell$ and $C \in \mathcal{C}$ then $|Y_\ell| = n^{2/3+o(1)}$ and

$$\Pr(|N^+(u) \cap Y_\ell| \geq 100) \leq \Pr(\text{Bin}(n^{2/3+o(1)}, p) \geq 100) \leq n^{-3}. \quad (41)$$

We write $Z = \sum_{u \in X_\ell} Z_u$ where

$$Z_u = \frac{\alpha_{\ell,u}}{\deg^+(u)} \sum_{v \in (N^+(u) \cap Y_\ell)} \beta_{\ell,v}.$$

Conditional on $|N^+(u) \cap Y_\ell| \leq 100$ we have $Z_u \leq 100/(cnp)^{2\ell+1} = B$, say. Let $\widehat{Z}_u = Z_u/B$, then for $u \in X_\ell$ the \widehat{Z}_u are independent random variables, and $0 \leq \widehat{Z}_u \leq 1$. Let $\widehat{Z} = \sum_{u \in X_\ell} \widehat{Z}_u$. Referring to (26), (28), (29) we note that if $C \in \mathcal{C}$ then

$$\mathbf{E}(\widehat{Z} \mid C) \geq \frac{1}{B} \frac{c}{n} (1 - \epsilon_X)(1 - \epsilon_Y)(1 - O(n^{-.33})) - O(|X_\ell|n^{-3}) = n^{1/3-o(1)},$$

where the n^{-3} is from (41).

Let $\widehat{\mu} = \mathbf{E}(\widehat{Z} \mid C)$. It follows from (2) that if $0 \leq \theta \leq 1$,

$$\Pr(|\widehat{Z} - \widehat{\mu}| \geq \theta \widehat{\mu} \mid C) \leq 2e^{-\theta^2 \widehat{\mu}/3}.$$

With $\theta = 4(np/\widehat{\mu})^{1/2}$ we find that,

$$\Pr(|\widehat{Z} - \widehat{\mu}| \geq 4(np\widehat{\mu})^{1/2} \mid C) = o(n^{-4}),$$

and hence that

$$\Pr(|Z - \mathbf{E}(Z \mid C)| \geq 4B(np\widehat{\mu})^{1/2} \mid C) = o(n^{-4}).$$

Now $(4B\sqrt{np\widehat{\mu}}) = O(n^{-7/6+o(1)})$, and so this implies that

$$\Pr\left(|Z - \mathbf{E}(Z \mid C)| = O\left(\frac{1}{n^{7/6+o(1)}}\right) \mid C\right) = 1 - o(n^{-4}). \quad (42)$$

It then follows from (27), (28) and (29) that

$$\Pr\left(Z \leq (1 - o(1)) \frac{\deg^-(y)}{m} \mid C\right) = o(n^{-4}). \quad (43)$$

So,

$$\Pr\left(\exists x, y : P_x^{(2\ell+1)}(y) \geq (1 - o(1)) \frac{\deg^-(y)}{m} \mid \mathcal{E}_V\right) = O(n^{-2}). \quad (44)$$

Since $\Pr(\mathcal{E}_V) = 1 - o(1)$, this completes the analysis for case I.

Case II: Fix $x, y \in V$. A sequence of events \mathcal{E}_n is said to occur *quite surely* **qs** if $\Pr(\mathcal{E}_n) = 1 - O(n^{-K})$ for any constant $K > 0$. The vertex x will have $\sim np$ out-neighbours X_+ and y will **qs** have $\sim np$ in-neighbours Y_- and **qs** $|X_+ \cap Y_-| \sim np^2$. Given this, the Chernoff bound shows that **qs** there are at least $(1 - o(1))n^2p^3$ paths of length three joining x to y . Therefore, for any $x, y \in V$,

$$P_x^{(3)}(y) \geq (1 - o(1)) \frac{n^2p^3}{(np)^3} = \frac{1 - o(1)}{n}$$

and we can proceed as we did for the previous case from (44).

Case III: $\omega \geq n^{3/5}$. We use a similar argument to the previous case. We now use the fact that **qs** there are at least $\sim np^2$ paths of length two from x to y .

This completes the proof of Lemma 6. □

Lemma 7. *Whp for all $y \in V$,*

$$\pi_y \geq (1 - o(1)) \frac{\deg^-(y)}{m}.$$

Proof It follows from Lemma 6 that **whp**, that for any $y \in V$, and $s = 2\ell + 1$ we have

$$\pi_y = \sum_{x \in V} \pi_x P_x^{(s)}(y) \geq (1 - o(1)) \frac{\deg^-(y)}{m} \sum_{x \in V} \pi_x = (1 - o(1)) \frac{\deg^-(y)}{m}. \quad (45)$$

□

5.2 Upper Bound on the stationary distribution

Lemma 7 above proves that the expression in Theorem 2 is a lower bound on the stationary distribution. As $\sum \pi_y = 1$, this can be used to derive an upper bound of $\pi_y \leq (1 + o(1)) \frac{\deg^-(y)}{m}$ for all but $o(n)$ vertices y . To extend this upper bound to *all* $y \in V$ is the subject of the following lemma.

Lemma 8. *There exists an integer $s = O(\log_{np} n)$ such that **whp** for all $x, y \in V$,*

$$P_x^{(s)}(y) \leq (1 + o(1)) \frac{\deg^-(y)}{m},$$

where m is the number of edges of $D_{n,p}$.

The proof of Lemma 8 is split into two cases. Each case requires a sequence of lemmas. As before, most of the work is in proving the first case. We use the notation $\alpha \ll \beta$ to mean that α/β is small enough so that any implied inequalities hold.

Case I: $np \leq n^\delta$ where $0 < \delta \ll \eta \ll 1$ are positive constants.

Case II: $np \geq n^\delta$.

Proof of Case I.

Let

$$\Lambda = \log_{np} n.$$

We will use the following values:

$$\begin{aligned} \ell_0 &= (1 + \eta)\Lambda, & \ell_1 &= (1 - 10\eta)\Lambda, & \ell_2 &= 11\eta\Lambda, \\ \ell_3 &= (1 - \eta/10)\Lambda, & \ell_4 &= \eta\Lambda/20, & \ell_5 &= 9\eta\Lambda/10. \end{aligned}$$

The plan once again is to estimate $P_x^{\ell_0+1}(y)$ using breadth-first trees T_X, T_Y . This time it is easier to grow T_X to a depth ℓ_1 and T_Y to a relatively small depth ℓ_2 . With this choice, Lemma 9 below, implies that $|Y|$ will contain no more than $|Y|$ edges **whp**. This reduces the complexity of the argument.

Lemma 9. *Suppose that $np \leq n^\delta$ where $\delta \ll 1$. For all $S \subseteq V$, $|S| \leq s_0 = \frac{1}{2} \log_{np} n$, S contains at most $|S|$ edges **whp**.*

Proof The expected number of sets S with more than $|S|$ edges can be bounded by

$$\sum_{s=3}^{s_0} \binom{n}{s} \binom{s^2}{s+1} p^{s+1} \leq \sum_{s=3}^{s_0} (e^2 np)^s \text{sep} = o(1).$$

□

We fix x, y and grow T_X from x to a depth ℓ_1 , and T_Y into y to a depth ℓ_2 . The definition of T_X will change, but we will retain the notation. Fortunately, dealing with T_X becomes trivial. The need for parameters ℓ_3, ℓ_4, ℓ_5 indicates that things are more complicated for T_Y .

Define $Y_0 = \{y\}, Y_1, \dots, Y_{\ell_2}$ where $Y_{i+1} = N^-(Y_i) \setminus (Y_0 \cup \dots \cup Y_i)$ for $0 \leq i < \ell_2$.

We form the BFS tree T_Y from $Y_0, Y_1, \dots, Y_{\ell_2}$ as we did with the Y_i 's in the proof of Lemma 6 (i.e. if $w \in Y_{i-1}$ is the in-neighbour of more than one vertex of Y_i , we only keep the edge (z, w) with z as small as possible for T_Y). Our upper bound construction of T_Y is different, in that we do not truncate in-neighbourhoods at Δ_0 . The probability that there is a neighbourhood sufficiently large to truncate is $o(n^{-10})$. If there is no truncation, the construction is the same as for the lower bound.

Now let X_i , $0 \leq i \leq \ell_3$ be the set of all vertices that are reachable from x by a walk of length i . These sets may be larger than the X_i of Lemma 6 as we allow them to overlap, and they are not disjoint from Y . For $0 \leq \ell \leq \ell_3$, let

$$X_\ell^* = X_\ell^*(x) = \bigcup_{i=0}^{\ell} X_i.$$

For $1 \leq i \leq \ell_1$ let

$$\tilde{X}_i = \{a \in X_i : \exists b \in X_j \text{ and } j \leq i \text{ such that } (a, b) \text{ is an edge}\}$$

and let $\tilde{X} = \bigcup_{i=0}^{\ell_1} \tilde{X}_i$.

Lemma 10.

(a) Let

$$\mathcal{L}_a(\ell) = \{\forall z \in X^* : z \text{ has } \leq 100/\eta \text{ in-neighbours in } X_\ell^*\},$$

then $\Pr(\neg \mathcal{L}_a(\ell_3)) = O(n^{-9})$.

(b) Let

$$\mathcal{L}_b = \left\{ \forall 1 \leq i \leq \ell_1 : |\tilde{X}_i| \leq n^{2\delta-10\eta} \Delta_0^i + \log^2 n \right\},$$

then $\Pr(\neg \mathcal{L}_b) = O(n^{-9})$.

Proof

(a) Let $z \in X_{\ell_3}^*$. Let ζ be the number of in-neighbours of z in $\bigcup_{i=1}^{\ell_3} X_i$. Suppose that we build T_{X^*} as in Section 5.1, by growing a breadth-first out-tree from x to depth ℓ_3 . In this way we expose only one in-neighbour of z . Thus ζ is dominated by $1 + \text{Bin}(|X^*|, p) \leq 1 + \text{Bin}(\Delta_0^{\ell_3}, p) + \text{Bin}(\max\{0, |X^*| - \Delta_0^{\ell_3}\}, p)$. We apply (11) to deal with the second binomial. Hence if $r + 1 = 100/\eta$,

$$\Pr(\zeta \geq r + 1) \leq (1 + o(1)) \Delta_0^{r\ell_3} p^r + n^{-10} e^{-10np} \leq 2n^{r(\delta-\eta/10)} + n^{-10} e^{-10np} \leq n^{-9}.$$

Part (a) of the lemma follows.

(b) We claim that $|\tilde{X}_{i+1}|$ is dominated by $|\tilde{X}_1| + |\tilde{X}_2| + \dots + |\tilde{X}_i| + \text{Bin}(\Delta_0^{2i+2}, p) + n1_{\mathcal{D}}$ where \mathcal{D} is defined prior to (11). The sum $|\tilde{X}_1| + |\tilde{X}_2| + \dots + |\tilde{X}_i|$ accounts for vertices that are in X_{i+1} and also in \tilde{X}_j for some $j \leq i$. For a vertex $z \in X_{i+1}$ not so far counted, there are two possibilities. (i) $z \notin \bigcup_{k=0}^i X_k$ or (ii) there exists $j \leq i$ such that $z \in X_j$ and $z \notin \bigcup_{k=j+1}^i X_k$. In the first case the number of edges from z to $\bigcup_{k=0}^{i+1} X_k$ is dominated by $\text{Bin}(|X_0| + \dots + |X_{i+1}|, p)$. In the second case the number of edges from z to $\bigcup_{k=j+1}^{i+1} X_k$ is dominated by $\text{Bin}(|X_{j+1}| + \dots + |X_{i+1}|, p)$. We bound the number of vertices counted by (i) and (ii) by the number of edges described in (i) and (ii) and this is dominated by $\text{Bin}(\Delta_0^{i+1}(1 + \Delta_0 + \dots + \Delta_0^{i+1}), p) + n1_{\mathcal{D}}$. The result follows from Chernoff bounds and induction on i . Thus with probability $1 - o(n^{-10})$ we have

$$|\tilde{X}_{i+1}| \leq n^{2\delta-10\eta} (\Delta_0 + \Delta_0^2 + \dots + \Delta_0^i) + i \log^2 n + n^{\delta-10\eta} \Delta_0^{i+1} + \log^2 n.$$

□

We now consider the weight of various types of walks of length $\ell_0 + 1$ from x to y . Some of these are simple directed paths in the BFS trees, of the type considered in the lower bound, and some use back edges of the BFS trees or contain cycles etc.

For $u \in X_{\ell_1}$ we let

$$\alpha_{u, \ell_1} = \Pr(\mathcal{W}_x(\ell_1) = u)$$

where of course

$$\sum_{u \in X} \alpha_{u, \ell_1} = 1. \quad (46)$$

We also define the $\beta_{i,v}$ as we did in (24). We will estimate

$$P_x^{\ell_0+1}(y) \leq Z_x^{\ell_0+1}(y) + Q_x^{\ell_0+1}(y) + R_x^{\ell_0+1}(y) \quad (47)$$

where the definitions of the terms on the right hand side are as follows:

- $Z_x^{\ell_0+1}(y)$ is given by,

$$Z_x^{\ell_0+1}(y) = \sum_{\substack{u \in X_{\ell_1} \\ v \in Y_{\ell_2} \setminus X}} \alpha_{\ell_1, u} \beta_{\ell_2, v} \frac{1_{uv}}{\deg^+(u)}, \quad (48)$$

is the probability that $\mathcal{W}_x(\ell_0+1) = y$ and the (ℓ_1+1) th edge (u, v) is such that $v \in Y_{\ell_2} \setminus X$ and the last ℓ_2 steps of the walk use edges of the tree T_Y .

These are the simplest walks to describe and they make up almost all the walk probability. They go through T_X and then monotonically level by level through T_Y .

- $Q_x^{\ell_0+1}(y)$ is the probability that $\mathcal{W}_x(\ell_0+1) = y$ and the (ℓ_1+1) th edge (u, v) is such that $v \in Y_{\ell_2} \cap X$ and the last ℓ_2 steps of the walk use edges of the tree T_Y .
- $R_x^{\ell_0+1}(y)$ is the probability that $\mathcal{W}_x(\ell_0+1) = y$ and the last ℓ_2 steps of the walk use an edge which is not part of the tree T_Y .

Upper bound for $Z_x^{\ell_0+1}(y)$. We proceed in a similar manner to the proof of (29), and use the same definition of central and normal, and similar values of the parameters. Thus $\zeta = 1/\log \log \log n$ and $\zeta_0 = \lceil 4/(\zeta^3 d) \rceil$ in the definition of normality. When $np = d \log n$, d constant, we let $\epsilon_Y = 4\zeta$ and $M = 10/\epsilon_Y$. When $\frac{M\ell}{np} = O\left(\frac{\log \log \log n}{\log \log n}\right) = o(1)$ as in the lower bound proof. For $np = \omega \log n$, $\omega \rightarrow \infty$ let $\epsilon_Y = 1/\omega^{1/3}$ and $M = 1$. In this case $\frac{M\ell}{np} = \frac{1}{\omega^{1/3} \log \log n} = o(1)$.

Let \mathcal{E}_y^* be the event that y and all of its in-neighbours are centered. A simple calculation shows that $\Pr(\mathcal{E}_y^*) = 1 - o(n^{-1})$.

Arguing in a similar manner as for (33) to (35) we get

$$\Pr\left(\sum_{v \in Y_{\ell_2}} \beta_{\ell_2, v} \geq (1 + \epsilon_Y) \frac{d_y^-}{np}\right)$$

$$\begin{aligned}
&\leq \Pr \left(\sum_{v \in Y_{\ell_2}} \beta_{\ell_2, v} \geq (1 + \epsilon_Y) \frac{d_y^-}{np} \middle| (\mathcal{E}_y^*, \mathcal{N}(y)) \right) + \Pr(\neg(\mathcal{E}_y^*, \mathcal{N}(y))) \\
&\leq \Pr \left(\sum_{i=1}^{d_y^- - \zeta_0} \frac{A_i}{(1 - \zeta)np} + \sum_{i=d_y^- - \zeta_0 + 1}^{d_y^-} \frac{A_i}{cnp} \geq (1 + \epsilon_Y) \frac{d_y^-}{np} \right) + o(n^{-1}) \\
&\leq \Pr \left(\sum_{i=1}^{d_y^- - \zeta_0} A_i \geq (1 + \epsilon_Y/2)d_y^- \right) + \Pr \left(\sum_{i=1}^{\zeta_0} A_i \geq c\epsilon_Y d_y^-/3 \right) + o(n^{-1}) \tag{49} \\
&= \Pr \left(\sum_{i=1}^{d_y^- - \zeta_0} A_i \geq (1 + \epsilon_Y/2)d_y^- \right) + o(n^{-1}) \tag{50} \\
&\leq \max_{d_y^- \in I} \exp \left\{ M \left((d_y^- - \zeta_0) \left(1 - \frac{5(\ell - 1)}{np} \right) - (1 + \epsilon_Y/2)d_y^- \right) \right\} + o(n^{-1}) \\
&\leq e^{-M\epsilon_Y \log n/3} + o(n^{-1}) \\
&= o(n^{-1}).
\end{aligned}$$

To go from (49) to (50) we use (22). Let $k = \zeta_0$, $\varsigma = o(1)$, $b = c\epsilon_Y d_y^-/3$ and $M = K/\epsilon_Y$ for some large constant K . Thus $k(1 + \varsigma) \leq c\epsilon_Y d_y^-/6$, and

$$\Pr \left(\sum_{i=1}^{\zeta_0} A_i \geq c\epsilon_Y d_y^-/3 \right) = o(n^{-1}).$$

Using the Hoeffding inequality, we see as in (42), (43) that

$$\Pr \left(\exists x, y : Z(x, y) \geq (1 + o(1)) \frac{d^-(y)}{m} \right) = o(1). \tag{51}$$

In computing the expectation of Z we observe that some of the vertices in the top level of T_X are previously inspected in our construction and now $\mathbf{E}(1_{uv}/d^+(u)) \leq 1/n$.

Upper bound for $Q_x^{\ell_0+1}(y)$. From Lemma 10 we see that we can assume

$$|\tilde{X}_{\ell_1}| \leq N = n^{1-19\eta}.$$

Suppose that $W = (x = u_{\ell_1}, \dots, u_1, u_0, v_0, v_1, \dots, y = v_{\ell_2})$ is a walk that contributes to $Q_x^{\ell_0+1}(y)$. As $u_0 \in X_{\ell_1}$, and $v_0 \in X_i \cap Y$ for some $0 \leq i \leq \ell_1$, we deduce that $v_0 \in N^+(\tilde{X}_{\ell_1})$. Next let $1 \leq \ell(W) = \min \{j : v_j \notin X\}$. We put $\ell(W) = \ell_2 + 1$ if $v_0, v_1, \dots, v_{\ell_2} \in X$. We estimate the contribution to $Q_x^{\ell_0+1}(y)$ from paths W for which $\ell(W) = j$ for some fixed j .

Let \mathcal{P}_j be the set of paths length j we can grow rooted in \tilde{X}_{ℓ_1} . Given $|\tilde{X}_{\ell_1}| = N$ the number of such paths at most $|\mathcal{P}_j| \leq N\Delta_0^j$. After we have grown X and \mathcal{P}_j , we can imagine that we

grow T_Y from y , *only as far as level* $\ell_2 - j$. Let $Y'_{\ell_2-j} = Y_{\ell_2-j} \setminus X$. We have $|Y'_{\ell_2-j}| \leq \Delta_0^{\ell_2-j}$ and so bearing in mind the event \mathcal{D} , given by (10), there are at most (a conditioned on \mathcal{D}) $\text{Bin}(N\Delta_0^j|Y_{\ell_2-j}|, p)$ number of paths in \mathcal{P}_j which terminate at y . This binomial is **qs** at most $n^{o(1)}$.

After summing over $j \leq \ell_2$ and using $\mathcal{L}_a(\ell_1)$ we see that the number of paths that start at x , reach \tilde{X}_{ℓ_1} after ℓ_1 steps and finish with a path of length ℓ_2 through T_Y that uses an edge from X to $(Y \setminus X)$ is at most $n^{o(1)}(100/\eta)^{\ell_2} = n^{o(1)}$. Furthermore, $\mathcal{L}_a(\ell_1)$ also implies that the number of vertices in \tilde{X}_{ℓ_1} that are the start of a path of length at most ℓ_2 to y , and that use only edges in X , is at most $(100/\eta)^{\ell_2} = n^{o(1)}$ ($y \in X$ for this case).

This restricts the number of choices for v_0 . This also bounds the number of walks contributing to $Q_x^{\ell_0+1}(y)$ to $n^{o(1)}$ since for such walks (i) there is a unique path from v_0 to y in T_Y and (ii) there are at most $(100/\eta)^{\ell_1} = n^{o(1)}$ walks that can use v_0 . Thus

$$\Pr \left(Q_x^{\ell_0+1}(y) \geq \frac{n^{o(1)}}{(c \log n)^{\ell_0}} = n^{-1-\eta+o(1)} \right) = o(n^{-2}). \quad (52)$$

Upper bound for $R_x^{\ell_0+1}(y)$. Observe that Lemma 9 implies that **whp** the set Y contains at most $|Y|$ edges. For, if Y contained more than $|Y| + 1$ edges then it would contain two distinct cycles C_1, C_2 . In which case, C_1, C_2 and the shortest undirected path joining them in Y would form a set S which satisfies the conditions of Lemma 9.

Thus there is at most one edge $e = (u_y, v_y)$ contained in Y that is not part of T_Y . We must have $u_y \in Y_i, v_y \in Y_j$ for some $0 \neq j \geq i - 1$. Otherwise $(u_y, v_y) \in T_Y$.

Suppose first that $j \geq \ell_5$. Then conditional on the occurrence of $\mathcal{L}_a(\ell_3), \mathcal{E}_V$ (see (15) for the definition of \mathcal{E}_V) there are certainly less than $\Delta_0^{\frac{2}{5} \log_{np} n} \times (100/\eta)^{\ell_3}$ walks contributing to $R_x^{\ell_0+1}(y)$. The first term in this factor is the number of walks of length at most $\frac{2}{5} \log_{np} n$ that start in X^* and end at v_y and the second factor bounds the number of walks of length at most ℓ_3 that start at x and end at the starting point of such a walk. Thus under these circumstances,

$$R_x^{\ell_0+1}(y) \leq \frac{\Delta_0^{\frac{2}{5} \log_{np} n} \times (100/\eta)^{\ell_3}}{(cnp)^{\ell_0}} \leq n^{-1-\eta/2}. \quad (53)$$

Remark 4. *This proves that if x, y are such that there are no extra edges within ℓ_5 of y and if $\mathcal{L}_a(\ell_3), \mathcal{L}_b, \mathcal{E}_V$ occur, then with probability $1 - o(n^{-2})$, $P_x^{\ell_0+1}(y) = (1 + o(1)) \frac{\text{deg}^-(y)}{m}$.*

Let ℓ_y be the distance from v_y to y if v_y exists, and $\ell_y = 0$ otherwise. Let $\ell_6 = \ell_0 + 1 - \ell_y$, if v_y exists then from Remark 4,

$$\Pr \left(P_x^{\ell_6}(v_y) = (1 + o(1)) \frac{\text{deg}^-(v_y)}{m} \mid \mathcal{L}_a(\ell_3), \mathcal{L}_b, \mathcal{E}_V \right) = 1 - o(n^{-2}).$$

This is because $\ell_6 \geq (1 + \frac{\eta}{20}) \log_{np} n$ and we can repeat the arguments above with v_y replacing y and with a smaller value of η . This assumes that the high probability event \mathcal{L}_9 described in Lemma 9 occurs. In which case there is no edge (u_{v_y}, v_{v_y}) to deal with and we are in the case described in Remark 4.

We write

$$R_x^{\ell_0+1}(y) \leq A_1$$

where A_1 is the probability that \mathcal{W}_x goes to y through v_y and $A_1 = 0$ if v_y does not exist. Also,

$$\Pr(\mathcal{W}_{v_y}(\ell_y) = y) = O(1/np). \quad (54)$$

So assuming v_y exists

$$\begin{aligned} A_1 &= P_x^{\ell_6}(v_y) \Pr(\mathcal{W}_{v_y}(\ell_y) = y) \\ &\leq (1 + o(1)) \frac{\Delta_0}{m} \Pr(\mathcal{W}_{v_y}(\ell_y) = y) \\ &\leq \frac{A_2}{n^2 p} = o\left(\frac{\deg^-(y)}{m}\right) \end{aligned}$$

for some constant $A_2 > 0$.

Hence,

$$\Pr(R_x^{\ell_0+1}(y) \leq A_2/n^2 p \mid \mathcal{L}_a(\ell_3), \mathcal{L}_b, \mathcal{E}_V, j < \ell_5) = 1 - o(n^{-2})$$

and in conjunction with (53)

$$\Pr(\exists x, y : R_x^{\ell_0+1}(y) > A_2/n^2 p \mid \mathcal{L}_a(\ell_3), \mathcal{L}_b, \mathcal{E}_V) = o(1). \quad (55)$$

Removing conditioning, it follows from equations (51), (52), (53) and (55) that **whp** for all $x, y \in V$,

$$P_x^{\ell_0+1}(y) \leq (1 + o(1)) \frac{\deg^-(y)}{m}.$$

This completes the proof for Case I of Lemma 8.

Case II: $np \geq n^\delta$.

We can deal with this case by using a concentration inequality (56) from Kim and Vu [13]: Let $\Upsilon = (W, E)$ be a hypergraph where $e \in E$ implies that $|e| \leq s$. Let

$$Z = \sum_{e \in E} w_e \prod_{i \in e} z_i$$

where the $w_e, e \in E$ are positive reals and the $z_i, i \in W$ are independent random variables taking values in $[0, 1]$. For $A \subseteq W, |A| \leq s$ let

$$Z_A = \sum_{\substack{e \in E \\ e \supseteq A}} w_e \prod_{i \in e \setminus A} z_i.$$

Let $M_A = \mathbf{E}(Z_A)$ and $M_j(Z) = \max_{A, |A| \geq j} M_A$ for $j \geq 0$. There exist positive constants a and b such that for any $\lambda > 0$,

$$\Pr(|Z - \mathbf{E}(Z)| \geq a\lambda^s \sqrt{M_0 M_1}) \leq b|W|^{s-1} e^{-\lambda}. \quad (56)$$

For us, W will be the set of edges of \vec{K}_n the complete digraph on n vertices. E will be the set of sets of edges in walks of length $s = \lceil 2/\delta \rceil$ between two fixed vertices x and y in \vec{K}_n , and $w_e = 1$. Z will be the number of walks that are in $D_{n,p}$. In which case we have

$$\begin{aligned} (n)_{s-1} p^s &\leq \mathbf{E}(Z) \leq (1 + o(1)) n^{s-1} p^s \\ M_j &\leq (1 + o(1)) n^{s-j-1} p^{s-j} \leq (1 + o(1)) \mathbf{E}(Z) / np \quad \text{for } j \geq 1. \end{aligned}$$

So $M_0 = \mathbf{E}(Z)$ and applying (56) with $\lambda = (\log n)^2$ we see that for any x, y we have

$$\Pr(Z \geq \mathbf{E}(Z) + O(\mathbf{E}(Z) n^{-\delta/2} (\log n)^2)) \leq O(n^{-3}).$$

Thus **whp**

$$P_x^s(y) \leq (1 + o(1)) \frac{n^{s-1} p^s}{((1 - \epsilon_1) np)^s} = \frac{1 + o(1)}{n} \quad \forall x, y \in V.$$

This completes the proof of Lemma 8. □

We can now give an upper bound complementing that in Lemma 7, thus completing the proof of Theorem 2 in the case where Assumption 1 holds.

Lemma 11. *Whp for all $y \in V$,*

$$\pi_y \leq (1 + o(1)) \frac{\deg^-(y)}{m}.$$

Proof It follows from Lemma 8 that **whp**, for any $y \in V$, and $s = \ell_0 + 1$ that

$$\pi_y = \sum_{x \in V} \pi_x P_x^{(s)}(y) \leq (1 + o(1)) \frac{\deg^-(y)}{m} \sum_{x \in V} \pi_x = (1 + o(1)) \frac{\deg^-(y)}{m}. \quad (57)$$

□

5.3 Removing Assumption 1

We will assume now that

$$1 + o(1) \leq d \leq 2.$$

We need to prove some additional lemmas. Let a vertex be *small* if it has in-degree or out-degree at most $\log n/20$ and *large* otherwise. Let *weak distance* refer to distance in the underlying graph of $D_{n,p}$.

Lemma 12.

- (a) **Whp** there are fewer than $n^{1/5}$ small vertices.
- (b) **Whp** every pair of small vertices are at weak distance at least $\ell_{10} = \frac{\log n}{10 \log \log n}$ apart.
- (c) **Whp** there does not exist a vertex v with $\max \{ \deg^+(v), \deg^-(v) \} \leq \log n / 20$.
- (d) Re-call the definition of $\varsigma^*(v)$ from Theorem 2. **Whp** for all vertices y ,

$$\sum_{z \in N^-(y)} \frac{\deg^-(z)}{\deg^+(z)} = (1 + o(1))(\deg^-(y) + \varsigma^*(v)).$$

Proof

- (a) The expected number of small vertices is at most

$$n \sum_{k=0}^{\log n / 20} \binom{n-1}{k} p^k q^{n-1-k} = O(n^{.1998}).$$

Part (a) now follows from the Markov inequality.

- (b) The expected number of pairs of small vertices at distance ℓ_{10} or less is at most

$$n^2 \sum_{k=0}^{\ell_{10}} 2^k n^k p^{k+1} \left(2 \sum_{l=0}^{\log n / 20} \binom{n-1}{l} p^l q^{n-1-l} \right)^2 =$$

$$O(n \ell_{10} (2d \log n)^{\ell_{10}+1} (20ed)^{\log n / 10} n^{-2d}) = O(n \cdot n^{1/10+o(1)} \cdot n^{1/2} \cdot n^{-2}) = o(1).$$

- (c) The expected number of vertices with small out- and in-degree is $O(n^{1-2 \times .8002}) = o(1)$.

- (d) For $1 \leq k \leq \Delta_0$ let

$$\lambda_k = \begin{cases} 1 & 1 \leq k \leq \frac{\log n}{(\log \log n)^4} \\ (\log \log n)^4 & \frac{\log n}{(\log \log n)^4} \leq k \leq \Delta_0 \end{cases}.$$

Let $\epsilon = \frac{1}{\log \log n}$. The probability there exists a vertex of in-degree $k \in [1, \Delta_0]$ with λ_k in-neighbours of degree outside $(1 \pm \epsilon)np$, is bounded by

$$\sum_{k=1}^{\Delta_0} n \binom{n-1}{k} p^k q^{n-1-k} \binom{k}{\lambda_k} (4e^{-\epsilon^2 np / 3})^{\lambda_k} \leq \sum_{k=1}^{\Delta_0} 2n^{1-d} \left(\frac{nep}{k} \cdot 2 \cdot n^{-\epsilon^2 d \lambda_k / (4k)} \right)^k = o(1).$$

Now assume that there are fewer than λ_k neighbours of v of degree outside $(1 \pm \epsilon)np$. Then, assuming at most one neighbour w of v is small,

$$\sum_{z \in N^-(y) \setminus \{w\}} \frac{\deg^-(z)}{\deg^+(z)} = \begin{cases} (1 + O(\epsilon))k & 1 \leq k \leq \frac{\log n}{(\log \log n)^4} \\ (1 + O(\epsilon))(k - \lambda_k) + O(\lambda_k) & \frac{\log n}{(\log \log n)^4} \leq k \leq \Delta_0 \end{cases}.$$

This completes the proof of the lemma. \square

Let *weak* distance refer to distance in the underlying graph of $D_{n,p}$, and let a cycle in the underlying graph be called a *weak* cycle.

Lemma 13. *Whp there does not exist a small vertex that is within weak distance ℓ_{10} of a weak cycle C of length at most ℓ_{10} .*

Proof Let v, C be such a pair. Let $|C| = i$ and j be the weak distance of v from C . The probability that such a pair exists is at most

$$\sum_{i=3}^{\ell_{10}} (2np)^i \sum_{j=0}^{\ell_{10}} (2np)^j \sum_{l=0}^{\log n/20} 2 \binom{n-1}{l} p^l q^{n-1-l} \\ = O(n^{1/10+o(1)} \cdot n^{1/10+o(1)} \cdot n^{-4/5+o(1)}) = o(1).$$

\square

5.3.1 Lower bound on steady state

For this we proceed as in Section 5.1 but initially restrict our analysis to large x, y . Also, we do not include the small vertices when creating the X_i, Y_i . Our previous analysis holds up with $c = 1/20$ in the use of Lemma 5. Furthermore, avoiding the $\leq n^{1/5}$ small vertices (see Lemma 12(a)) is easily incorporated because in the proof we have allowed for the avoidance of $n^{2/3+o(1)}$ vertices from $\bigcup_i X_i$ etc.

In this way, we show for all large x, y that **whp**,

$$P_x^{(2\ell+1)}(y) \geq (1 - o(1)) \frac{\deg^-(y)}{m}. \quad (58)$$

If x is small, then **whp** it will only have large out-neighbours (see Lemma 12(b)) and so if y is large then

$$P_x^{(2\ell+2)}(y) = \frac{1}{\deg^+(x)} \sum_{z \in N^+(x)} P_z^{(2\ell+1)}(y) \geq (1 - o(1)) \frac{\deg^-(y)}{m}. \quad (59)$$

A similar argument deals with small y and x arbitrary i.e.

$$P_x^{(2\ell+2)}(y) = \sum_{z \in N^-(y)} \frac{P_x^{(2\ell+1)}(z)}{\deg^+(z)} \geq (1 - o(1)) \sum_{z \in N^-(y)} \frac{\deg^-(z)}{m} \frac{1}{\deg^+(z)} \geq (1 - o(1)) \frac{\deg^-(y)}{m}. \quad (60)$$

We have used Lemma 12(d) to justify the last inequality.

We thus obtain a proof that for all x, y , **whp**

$$P_x^{(2\ell+2)}(y) \geq (1 - o(1)) \frac{\deg^-(y)}{m}.$$

(We apply the argument of (60) to (58) to replace $\ell + 1$ by $\ell + 2$).

The extra term $\varsigma^*(y)$ arises as follows:

$$P_x^{(2\ell+3)}(y) = \sum_{z \in N^-(y)} \frac{P_x^{(2\ell+3)}(z)}{\deg^+(z)} \geq (1 - o(1)) \sum_{z \in N^-(y)} \frac{\deg^-(z)}{m \deg^+(z)} \geq (1 - o(1)) \frac{\deg^-(y) + \varsigma^*(y)}{m}.$$

Now we can proceed as in (45).

5.3.2 Upper bound on steady state

Returning to the proof of Section 5.2 we claim that Assumption 1 is only used in (49), (52), (53) and (54). We now explain how these parts of the proof alter when Assumption 1 is removed.

In (52) and (53) we used $(cnp)^{\ell_0}$ as a lower bound on the product of out-degrees on a path of length ℓ_0 . Using Lemmas 12 and 13, and noting that $\ell_0/\ell_{10} < 11$, we see that after dropping Assumption 1 we can replace this lower bound by $(cnp)^{\ell_0-11}$ and then the proof continues unchanged.

The proof as is works perfectly well if we assume that y is large and if it has no small in-neighbours and there is no small vertex in Y . We call such a vertex y *ordinary*. There is flexibility in choosing η and using this we can assume that

$$P_x^{(\ell_1+i)}(y) \leq (1 + o(1)) \frac{\deg^-(y)}{m} \tag{61}$$

for all $1 \leq i \leq \ell_2$ and ordinary y .

If y is small then from Lemmas 12 and 13 we can assume that all of its in-neighbours are ordinary. This is under the assumption that $2\ell_2 < \ell_{10}$ e.g. if $\eta \leq 1/250$. So in this case we can use Lemma 12(d) and obtain for $1 \leq i \leq \ell_2$ that

$$P_x^{(\ell_1+i)}(y) = \sum_{\xi \in N^-(y)} \frac{P_x^{(\ell_1+i-1)}(\xi)}{\deg^+(\xi)} \leq \frac{1 + o(1)}{m} \sum_{\xi \in N^-(y)} \frac{\deg^-(\xi)}{\deg^+(\xi)} = (1 + o(1)) \frac{\deg^-(y)}{m}.$$

Suppose now that y is large and that there is a small vertex $z \in Y$. We can assume from Lemma 13 that Y does not contain any edge not in T_Y . Either $z = w \in N^-(y)$ or, if not, let w be the in-neighbour of y on the path from z to y in T_Y .

Suppose first that $z \neq w$. The possible existence of vertices of small out-degree is allowed for in the random variable $D_{i,t}$ of Lemma 5, so the results of that lemma hold unaltered for w . In the case where $z = w$ then

$$\begin{aligned} P_x^{(\ell_0+1)}(y) &\leq \frac{1+o(1)}{m} \left(\frac{\deg^-(w)}{\deg^+(w)} + \sum_{u \in N^-(y) \setminus w} \frac{\deg^-(u)}{\deg^+(u)} \right) \\ &= \frac{(1+o(1))}{m} (\deg^-(y) + \varsigma^*(y)). \end{aligned}$$

We have now completed the proof without Assumption 1.

6 Upper Bound on Mixing time

We next show that the mixing time T as defined in (4) satisfies

$$T = o(\ell \log n) = o((\log n)^2) \tag{62}$$

where ℓ is given by (23).

Define

$$\bar{d}(t) = \max_{x, x' \in V} |P_x^{(t)} - P_{x'}^{(t)}| \tag{63}$$

to be the maximum over x, x' of the variation distance between $P_x^{(t)}$ and $P_{x'}^{(t)}$. Equation (44) implies that

$$\bar{d}(2\ell + 1) = o(1). \tag{64}$$

Lemma 20 of Chapter 2 of Aldous and Fill [1] proves that

$$\bar{d}(s+t) \leq \bar{d}(s)\bar{d}(t) \text{ and } \max_x |P_x^{(t)} - \pi_x| \leq \bar{d}(t)$$

and so (62) follows immediately from (64).

7 The Cover Time

We see immediately from (62) that Condition (b) of Lemma 3 is satisfied. The proof that Condition (a) of Lemma 3 is satisfied, is as follows. We show below that **whp** for all $v \in V$

$$R_T(1) = 1 + o(1). \tag{65}$$

Let $\lambda = 1/KT$ as in (6). The value of $T = o(\log^2 n)$ is given by (62). For $|z| \leq 1 + \lambda$, we have

$$R_T(z) \geq 1 - \sum_{t=1}^T r_t |z|^t \geq 1 - (1 + \lambda)^T \sum_{t=1}^T r_t = 1 - o(1),$$

and thus for $v \in V$, the value of p_v in (7) is given by

$$p_v = (1 + o(1)) \frac{\deg^-(v)}{m}. \quad (66)$$

Proof of (65): If $d \geq (\log n)^2$, then the minimum out-degree of $D_{n,p}$ is $\Omega(d \log n)$. In which case we have for any x, y

$$\Pr(\mathcal{W}_v(t) = y \mid \mathcal{W}_v(t-1) = x) = O\left(\frac{1}{d \log n}\right). \quad (67)$$

The expected number of returns to $v \in V$ by \mathcal{W}_v is therefore $O(T/d \log n)$.

If $d \leq (\log n)^2$ then a first moment calculation shows that **whp** for every vertex $v \in V$, there is at most one edge from a vertex in $N^+(v)$ to $\{v\} \cup N^+(v)$ or from a vertex in $N^+(N^+(v))$ to $\{v\} \cup N^+(v) \cup N^+(N^+(v))$. Similarly, there is **whp** at most one vertex within distance 10 of v that has in-degree or out-degree less than cnp (see (13)). Thus with probability $1 - O(1/(\log n)^2)$, $x = \mathcal{W}_v(3)$ satisfies $\text{dist}(x, v) \geq 3$ and then the probability of a return to v is $O(T/(\log n)^3) = o(1/\log n)$.

7.1 Upper Bound on the Cover Time

For $np = d \log n$, d constant, let $t_0 = (1 + \epsilon)d \log\left(\frac{d}{d-1}\right) n \log n$. For $np = d \log n$, $d = d(n) \rightarrow \infty$ let $t_0 = (1 + \epsilon)n \log n$. Here $\epsilon \rightarrow 0$ sufficiently slowly so that all of the inequalities claimed below are valid.

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

$$C_u = \mathbf{E}(T_D(u)) = \sum_{t>0} \Pr(T_D(u) \geq t), \quad (68)$$

$$\Pr(T_D(u) \geq t) = \Pr(T_D(u) > t-1) = \Pr(U_{t-1} > 0) \leq \min\{1, \mathbf{E}(U_{t-1})\}. \quad (69)$$

Recall that $\mathbf{A}_v(t)$ denotes the event that $\mathcal{W}_u(t)$ did not visit vertex v in the interval $[T, t]$. It follows from (68), (69) that for any $t \geq T$,

$$C_u \leq t + 1 + \sum_{s \geq t} \mathbf{E}(U_s) \leq t + 1 + \sum_v \sum_{s \geq t} \Pr(\mathbf{A}_v(s)). \quad (70)$$

Assume first that $d(n) \rightarrow \infty$. If $T = o(s)$ then (8) of Lemma 3 implies

$$\Pr(\mathbf{A}_v(s)) \leq (1 + o(1)) \exp \left\{ -\frac{(1 - o(1))s}{n} \right\} + O(e^{-\Omega(s/T)}). \quad (71)$$

Plugging (71) into (70) we get

$$\begin{aligned} C_u &\leq t_0 + 1 + 2n \sum_{s \geq t_0} \left(\exp \left\{ -\frac{(1 - o(1))s}{n} \right\} + O(e^{-\Omega(s/T)}) \right) \\ &\leq t_0 + 1 + 3n^2 \exp \left\{ -\frac{(1 - o(1))t_0}{n} \right\} + O(nTe^{-\Omega(t_0/T)}) \\ &= (1 + o(1))t_0. \end{aligned} \quad (72)$$

Now assume that d is constant. For $v \in V$ we have

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

and

$$\pi_v \geq (1 - o(1)) \frac{\deg^-(v)}{m}.$$

In place of (72) we use the bounds on the number of vertices of degree k given in Lemma 4. Thus

$$C_u \leq t_0 + 1 + o(1) + \sum_{i=0}^3 S_i \quad (73)$$

where

$$\begin{aligned} S_i &= \sum_{k \in K_i} D(k) \sum_{s \geq t_0} \exp \left\{ -\frac{(1 - o(1))ks}{m} \right\} \\ &\leq m \sum_{k \in K_i} \frac{D(k)}{k} e^{-(1 - o(1))kt_0/m} \\ &\leq m \sum_{k \in K_i} \frac{D(k)}{k} \left(\frac{d-1}{d} \right)^{(1+\epsilon/2)k}. \end{aligned}$$

For the main term,

$$\begin{aligned} S_3 &\leq \frac{4m}{n^{d-1}} \sum_{k=cnp}^{Cnp} \left(\frac{nep}{k} \right)^k \left(\frac{d-1}{d} \right)^{(1+\epsilon/2)k} \\ &\leq 4m (Cnp) e^{-\epsilon cnp/2d} \\ &= o(t_0). \end{aligned} \quad (74)$$

We have used the fact that $(nep(d-1))/(kd)^k$ is maximized at $k = np(d-1)/d$, and $m = dn \log n(1 + o(1))$ **whp**.

Continuing we get

$$\begin{aligned}
S_1 &\leq m \sum_{k \in K_1} \frac{D(k)}{k} \left(\frac{d-1}{d} \right)^{(1+\epsilon/2)k} \\
&\leq m \sum_{k=1}^{15} \frac{(\log \log n)^2}{k} \left(\frac{d-1}{d} \right)^{(1+\epsilon/2)k} \\
&= o(m)
\end{aligned} \tag{75}$$

since $D(k) \leq (\log \log n)^2$ and either (i) $d-1 \geq (\log n)^{-1/3}$ and $K_1 = \emptyset$ or (ii) $d-1 < (\log n)^{-1/3}$ (see Lemma 4(a)). Similarly

$$\begin{aligned}
S_2 &\leq m \sum_{k \in K_2} \frac{D(k)}{k} \left(\frac{d-1}{d} \right)^{(1+\epsilon/2)k} \\
&\leq m \sum_{k=16}^{\Delta_0} \frac{(\log n)^4}{k} \left(\frac{d-1}{d} \right)^{(1+\epsilon/2)k} \\
&= o(m)
\end{aligned} \tag{76}$$

since $D(k) \leq (\log n)^4$ and either (i) $d-1 \geq (\log n)^{-1/3}$ and $\min\{k \in K_2\} \geq (\log n)^{1/2}$ or (ii) $d-1 < (\log n)^{-1/3}$ (see Lemma 4(a)).

We now see that $C_u \leq t_0 + o(t_0)$ follows from (73)–(76).

7.2 Lower Bound on the Cover Time

For $np = d \log n$, $d = O(1)$, let $t_1 = (1 - \epsilon)d \log \left(\frac{d}{d-1} \right) n \log n$. For $np = d \log n$, $d = d(n) \rightarrow \infty$ let $t_1 = (1 - \epsilon)n \log n$. Here $\epsilon \rightarrow 0$ sufficiently slowly so that all inequalities claimed below are valid.

Case 1: $np \leq n^\delta$ where $0 < \delta \ll \eta$ is a positive constant.

Let $k^* = (d-1) \log n$, and let $\hat{V} = \{v : \deg^-(v) = k^* \text{ and } \deg^+(v) = d \log n\}$. **Whp** the size $|\hat{V}| \sim n^* = \frac{n^{\gamma_d}}{2\pi \log n (d(d-1))^{1/2}}$ (see Lemma 4(c)). Let us first work under Assumption 1 of Lemma 4. In this case $\gamma_d = (d-1) \ln(d/(d-1)) = \Omega(1)$ and we write $n^* = n^{\gamma_d - o(1)}$. The maximum degree in D is $\Delta_0 = O(np)$ and so V^* contains a sub-set V_1^* of size $n^{\gamma_d - o(1)}$ such that $v, w \in V_1^*$ and $x \in V$ implies

$$dist(x, v) + dist(x, w) > (\log n)^{2/3}. \tag{77}$$

$$dist(y, x) + dist(x, y) > 2(\log n)^{2/3}, \text{ for } y = v, w. \tag{78}$$

Each $v \in V_1^*$ has $\pi_v \sim \frac{d-1}{dn}$ and so we can choose a subset V^{**} of size $\geq n^{\gamma d - o(1)}$ such that if $v_1, v_2 \in V^{**}$ then

$$|\pi_{v_1} - \pi_{v_2}| \leq \frac{1}{n(\log n)^{10}}. \quad (79)$$

Indeed, suppose that $\pi_v \in \left[\frac{d-1}{2dn}, \frac{2(d-1)}{dn}\right]$ for $v \in \hat{V}$. Divide this interval into $(\log n)^{10}$ equal sized sub-intervals and then use the pigeon-hole principle.

We can also assume that if $v \in V^{**}$ then there is no edge joining two vertices within out-distance $\leq (\log \log n)^2$ of v . This removes only $n^{o(1)}$ vertices **whp**.

Now choose $u \notin V^{**}$ and let V^\dagger denote the set of vertices in V^{**} that have not been visited by \mathcal{W}_u by time t_1 . Then $\mathbf{E}(|V^\dagger|) \rightarrow \infty$, as the following calculation shows;

$$\mathbf{E}(|V^\dagger|) \geq n^{\gamma d - o(1)} \left(\exp \left\{ -\frac{(1 + o(1))k^*t_1}{m} \right\} - o(e^{-\Omega(t_1/T)}) \right) - T,$$

where the last term accounts for possible visits before time T .

We now drop Assumption 1 and assume that $d = 1 + \frac{\gamma}{\log n}$ where $\gamma \rightarrow \infty$. Equations (77), (78) now hold for all $v, w \in \hat{V}$. This follows from a calculation similar to that given in Lemma 12(b). The size of V^{**} is at least $n^*/(\log n)^{10}$ and we can again write

$$\begin{aligned} \mathbf{E}(|V^\dagger|) &\geq \frac{n^*}{(\log n)^{10}} \left(\exp \left\{ -\frac{(1 + o(1))k^*t_1}{m} \right\} - o(e^{-\Omega(t_1/T)}) \right) - T \\ &\rightarrow \infty. \end{aligned}$$

As in previous papers we can use the Chebyshev inequality in to show that $V^\dagger \neq \emptyset$ **whp**, thus completing the proof of Theorem 1. This will follow if we can prove that

$$\mathbf{Var}(|V^\dagger|) = o(\mathbf{E}(|V^\dagger|)^2) + O(|\hat{V}|^2 n^{-2}) = o(\mathbf{E}(|V^\dagger|)^2).$$

To establish this inequality, we will show that if $v, w \in V^\dagger$ then

$$\mathbf{Pr}(\mathbf{A}_v(t_1) \cap \mathbf{A}_w(t_1)) \leq (1 + o(1))\mathbf{Pr}(\mathbf{A}_v(t_1))\mathbf{Pr}(\mathbf{A}_w(t_1)). \quad (80)$$

To prove this, we identify vertices v, w into a ‘‘supernode’’ σ to obtain a digraph D_σ with $n - 1$ vertices. In this digraph σ has in-degree $\deg^-(v) + \deg^-(w) = 2k^*$.

The stationary distribution of D_σ .

The arguments we used in Section 5 remain valid in D_σ , and thus

$$\pi_\sigma \sim (1 - o(1)) \frac{2k^*}{m}.$$

However, we need to be more precise. Let π^* denote the vector of steady states in D_σ and for a vertex x of D_σ let

$$\hat{\pi}_x = \begin{cases} \pi_x & x \neq \sigma \\ \pi_v + \pi_w & x = \sigma \end{cases}.$$

We will prove for all $x \in V(D_\sigma)$, that

$$|\pi_x^* - \hat{\pi}_x| = O\left(\frac{1}{n(\log n)^8}\right). \quad (81)$$

Proof of (81).

Let $\xi = \hat{\pi} - \pi^*$ be the difference between $\hat{\pi}$ and π^* . Let P^* be the transition matrix of the walk on D_σ , then

$$P^*(x, y) = \begin{cases} P(x, y) & x, y \neq \sigma \\ (P(v, y) + P(w, y))/2 & x = \sigma \\ P(x, v) + P(x, w) & y = \sigma \end{cases}.$$

Let ξ' be the transpose of ξ . It follows from the steady state equations that

$$(\xi' P^*)_x = \begin{cases} \hat{\pi}_x - \pi_x^* & x \notin N^+(\{v, w\}) \\ \hat{\pi}_x - \pi_x^* + \frac{\pi_w - \pi_v}{2} P(v, x) & x \in N^+(v) \\ \hat{\pi}_x - \pi_x^* + \frac{\pi_v - \pi_w}{2} P(w, x) & x \in N^+(w) \end{cases}.$$

We rewrite this as

$$\xi'(I - P^*) = \eta' \quad (82)$$

where $\eta_x = 0$ for $x \notin N^+(\{v, w\})$ and $|\eta_x| \leq |\pi_v - \pi_w|/2$ otherwise.

Multiplying (82) on the right by $M = \sum_{t=0}^{T-1} (P^*)^t$ we have

$$\xi'(I - P^*)M = \xi'(I - (P^*)^T) = \eta'M. \quad (83)$$

But

$$(P^*)^T = \Pi + E \quad (84)$$

where Π is the $(n-1) \times (n-1)$ matrix with each row equal to $(\pi^*)'$. It follows from Section 6 that each entry of E has absolute value bounded by n^{-3} .

Now write $\xi = \alpha\pi^* + \zeta$ where $\zeta \perp \pi^*$. It follows from (83) and $(\pi^*)'P^* = (\pi^*)'$ that

$$(\alpha\pi^* + \zeta)'(I - (P^*)^T) = \zeta'(I - (P^*)^T) = \zeta'(I - \Pi - E) = \eta'M.$$

Now $\xi'\Pi = 0$ implies that $\zeta'\Pi = -\alpha(\pi^*)'\Pi = -\alpha(\pi^*)'$ and so

$$\zeta'(I - E) = \eta'M - \alpha(\pi^*)'.$$

As $\zeta \perp \pi^*$ this implies that

$$\zeta'(I - E)\zeta = \eta'M\zeta. \quad (85)$$

Note that

$$|\eta'M\zeta| \leq \sum_{t=0}^{T-1} |\eta'(P^*)^t\zeta| \leq T|\eta||\zeta|. \quad (86)$$

Now

$$|\zeta'(I - E)\zeta| \geq |\zeta|^2 - |\zeta'E\zeta| \geq |\zeta|^2 - n^{-3} \left(\sum_{i=1}^{n-1} |\zeta_i| \right)^2 \geq |\zeta|^2(1 - n^{-2}). \quad (87)$$

It follows from (85), (86) and (87) that

$$|\zeta|^2(1 - n^{-2}) \leq T|\eta||\zeta|$$

and so using (79) we find that

$$|\zeta| = O\left(\frac{1}{n(\log n)^8}\right). \quad (88)$$

Now let $\mathbf{1}$ denote the $(n - 1)$ -vector of 1's. Then

$$0 = \mathbf{1} - \mathbf{1} = (\hat{\pi} - \pi^*)'\mathbf{1} = \xi'\mathbf{1} = \alpha + \zeta'\mathbf{1}.$$

Using (88) this gives

$$|\alpha| \leq |\mathbf{1}||\zeta| = O\left(\frac{1}{n^{1/2}(\log n)^8}\right).$$

Now $\xi_x = \alpha\pi_x^* + \zeta_x$ for all x and so

$$\xi_x^2 \leq 2\alpha^2(\pi_x^*)^2 + 2\zeta_x^2 = O\left(\frac{1}{n(\log n)^{16}} \cdot \frac{1}{n^2} + \frac{1}{n^2(\log n)^{16}}\right) = O\left(\frac{1}{n^2(\log n)^{16}}\right).$$

This completes the proof of (81). \square

Proof of (80).

For $v \in V^{**}$, we first tighten (65) to

$$R_v = 1 + o(1/(\log n)^2). \quad (89)$$

Indeed (77) and (78) imply that for $t \leq (\log n)^{2/3}$ the walk \mathcal{W}_v will be at distance $\geq 2(\log n)^{2/3} - t$ from v . Then once the walk is at a vertex w within $\leq (\log n)^{2/3}$ of v its chance of getting closer is only $O(1/\log n)$. This being true with at most one exception for a vertex of low out-degree. The probability that there is a time t such that \mathcal{W}_v is within $\leq (\log n)^{2/3}$ of v and it makes 10 steps closer to v in the next 100 steps is $O(T(\log n)^{-9}) = O((\log n)^{-7})$. This implies (89).

Similarly,

$$R_\sigma = 1 + o(1/(\log n)^2). \quad (90)$$

Using the suffix \mathbf{Pr}_σ to denote probabilities related to random walks in D_σ , it follows that

$$\begin{aligned} \mathbf{Pr}_\sigma(\mathbf{A}_\sigma(t_1)) &\leq \exp\left\{-\frac{(1 + O(T\pi_\sigma^*))\pi_\sigma^*t_1}{m}\right\} - o(e^{-\Omega(t_1/T)}) \\ &\leq \exp\left\{-\frac{(1 + o(1/\log n))(\pi_v + \pi_w)t_1}{m}\right\} - o(e^{-\Omega(t_1/T)}) \\ &= (1 + o(1))\mathbf{Pr}(\mathbf{A}_v(t_1))\mathbf{Pr}(\mathbf{A}_w(t_1)). \end{aligned}$$

But, using rapid mixing in D_σ ,

$$\begin{aligned}\Pr_\sigma(\mathbf{A}_\sigma(t_1)) &= \sum_{x \neq \sigma} P_{\sigma,u}^T(x) \Pr_\sigma(\mathcal{W}_x(t) \neq \sigma, T \leq t \leq t_1) \\ &= \sum_{x \neq \sigma} \left((1 + o(1)) \frac{\deg^-(x)}{m} \right) \Pr_\sigma(\mathcal{W}_x(t) \neq \sigma, T \leq t \leq t_1)\end{aligned}$$

On the other hand,

$$\begin{aligned}\Pr(\mathbf{A}_v(t_1) \cap \mathbf{A}_w(t_1)) &= \sum_{x \neq v,w} P_u^T(x) \Pr(\mathcal{W}_x(t) \neq v, w, T \leq t \leq t_1) \\ &= \sum_{x \neq v,w} \left((1 + o(1)) \frac{\deg^-(x)}{m} \right) \Pr(\mathcal{W}_x(t) \neq v, w, T \leq t \leq t_1)\end{aligned}$$

But,

$$\Pr_\sigma(\mathcal{W}_x(t) \neq \sigma, T \leq t \leq t_1) = \Pr(\mathcal{W}_x(t) \neq v, w, T \leq t \leq t_1)$$

because random walks from x that do not meet v, w or σ have the same measure in both digraphs.

It follows that

$$\Pr(\mathbf{A}_v(t_1) \cap \mathbf{A}_w(t_1)) - \Pr_\sigma(\mathbf{A}_\sigma(t_1)) \leq O(n^{-2}) + o(1)(\Pr(\mathbf{A}_v(t_1) \cap \mathbf{A}_w(t_1)) + \Pr_\sigma(\mathbf{A}_\sigma(t_1))).$$

This implies that

$$\begin{aligned}\Pr(\mathbf{A}_v(t_1) \cap \mathbf{A}_w(t_1)) &\leq (1 + o(1))\Pr_\sigma(\mathbf{A}_\sigma(t_1)) + O(n^{-2}) \leq \\ &\quad (1 + o(1))\Pr(\mathbf{A}_v(t_1))\Pr(\mathbf{A}_w(t_1)) + O(n^{-2}),\end{aligned}$$

which completes the proof of (80). \square

Case 2: $np \geq n^\delta$.

In this range we take $t_1 = (1 - \epsilon)n \log n$ and let V^{**} be the set of vertices of degree $\lfloor np \rfloor$. A simple second moment calculation shows that **whp** we have $|V^{**}| = n^{1/2+o(1)}$. We then choose ϵ so that $\mathbf{E}(|V^\dagger|) \geq n^{1/4}$. It is then only a matter of verifying (80). We define D_σ as in the previous case and now use (56) to show that **whp** we have $\pi_\sigma \geq (2 - o(1))/n$. The details are as in the previous case.

This completes the proof of Theorem 1. \square

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] N. Alon and J. Spencer, *The Probabilistic Method*, Second Edition, Wiley-Interscience, (2000).
- [4] C. Cooper and A. M. Frieze, The cover time of sparse random graphs, *Proceedings of the 14th ACM-SIAM Symposium on Discrete Algorithms*, (2003) 140-147.
- [5] C. Cooper and A. M. Frieze, The cover time of random regular graphs, *SIAM Journal on Discrete Mathematics*, 18 (2005) 728-740.
- [6] C. Cooper and A. M. Frieze, The cover time of the preferential attachment graph, to appear in *Journal of Combinatorial Theory Series B*, 97(2) 269-290 (2007).
- [7] C. Cooper and A. M. Frieze, The cover time of the giant component of $G_{n,p}$. *Random Structures and Algorithms*, 32, 401-439 J. Wiley (2008).
- [8] C. Cooper and A. M. Frieze, Corrigendum: The cover time of the giant component of a random graph, *Random structures and algorithms* 34 (2009) 300-304 .
- [9] U. Feige, A tight upper bound for the cover time of random walks on graphs, *Random Structures and Algorithms*, 6 (1995) 51-54.
- [10] U. Feige, A tight lower bound for the cover time of random walks on graphs, *Random Structures and Algorithms*, 6 (1995) 433-438.
- [11] A.M. Frieze, An algorithm for finding Hamilton cycles in random digraphs, *Journal of Algorithms* 9 (1988) 181-204.
- [12] W. Hoeffding, Probability inequalities for sums of bounded random variables, *Journal of the American Statistical Association* 58 (1963) 13-30.
- [13] J. H. Kim and V. Vu, *Concentration of multivariate polynomials and its applications*. *Combinatorica* 20 (2000) 417-434.