Looking for vertex number one

Alan Frieze*, Wesley Pegden†
Department of Mathematical Sciences,
Carnegie Mellon University,
Pittsburgh PA 15213,
U.S.A.

January 29, 2016

Abstract

Given an instance of the preferential attachment graph $G_n = ([n], E_n)$, we would like to find vertex 1, using only 'local' information about the graph; that is, by exploring the neighborhoods of small sets of vertices. Borgs et. al gave an an algorithm which runs in time $O(\log^4 n)$, which is local in the sense that at each step, it needs only to search the neighborhood of a set of vertices of size $O(\log^4 n)$. We give an algorithm to find vertex 1, which w.h.p. runs in time $O(\omega \log n)$ and which is local in the strongest sense of operating only on neighborhoods of single vertices. Here $\omega = \omega(n)$ is any function that goes to infinity with n.

1 Introduction

The Preferential Attachment Graph G_n was first discussed by Barabási and Albert [2] and then rigorously analysed by Bollobás, Riordan, Spencer and Tusnády [3]. It is perhaps the simplest model of a natural process that produces a graph with a power law degree sequence.

The Preferential Attachment Graph can be viewed as a sequence of random graphs $G_1, G_2, \ldots, G_n$ where G_{t+1} is obtained from G_t as follows: Given G_t, we add vertex $t+1$ and m random edges $\{e_i = (t+1, u_i) : 1 \leq i \leq m\}$ incident with vertex $t+1$. Here the constant m is a parameter of the model. The vertices u_i are not chosen uniformly from V_t, instead they are chosen with probabilities proportional to their degrees. This tends to generate some very high degree vertices,

*Research supported in part by NSF grant CCF1013110
†Research supported in part by NSF grant DMS

compared with what one would expect in Erdős-Rényi models with the same edge-density. We refer to $u_1, u_2, \ldots, u_m$ as the *left* choices of vertex $t+1$. We also say that $t+1$ is a *right* neighbor of u_i for $i = 1, 2, \ldots, m$.

We consider the problem of searching through the preferential attachment graph looking for vertex number 1, using only local information. This was addressed by Borgs, Brautbar, Chayes, Khanna and Lucier [5] in the context of the Preferential Attachment Graph $G_n = (V_n, E_n)$. Here $V_n = [n] = \{1, 2, \ldots, n\}$. They present the following local algorithm that searches for vertex 1, in a graph which may be too large to hold in memory in its entirety.

1: Initialize a list $\mathcal{L}$ to contain an arbitrary node u in the graph.

2: **while** $\mathcal{L}$ does not contain node 1 **do**

3: Add a node of maximum degree in $N(\mathcal{L})$ to $\mathcal{L}$ **od**;

4: return $\mathcal{L}$

Here for vertex set $\mathcal{L}$, we let $N(\mathcal{L}) = \{w \notin \mathcal{L} :\ \exists v \in \mathcal{L}\ s.t.\ \{v, w\} \in E_n\}$.

They show that w.h.p. the algorithm succeeds in reaching vertex 1 in $O(\log^4 n)$ steps. (We assume that an algorithm can recognize vertex 1 when it is reached.) In [5], they also show how a local algorithm to find vertex 1 can be used to give local algorithms for some other problems. We also note that Brautbar and Kearns [6] considered local algorithms in a more general context. There the algorithm is allowed to jump to random vertices as well as crawl around the graph in the search for vertices of high degree and high clustering coefficient.

We should note that, as the maximum degree in G_n is $n^{1/2-o(1)}$ w.h.p., one cannot hope to have a polylog(n) time algorithm if we have to check the degrees of the neighbors as we progress. Thus the algorithm above operates on the assumption that we can find the highest-degree neighbor of a vertex in $O(1)$ time. This would be the case, for example, if the neighborhood of a vertex is stored as a linked-list which is sorted by degrees. In the same situation, we can also determine the K highest degree neighbors of a vertex in constant time for any constant K, and in the present manuscript we assume such a constant-time step is possible. In particular, in this setting, each of steps 2-7 of the following *Degree Climbing Algorithm* takes constant time.

We let $d_n(v)$ denote the degree of vertex $v \in V_n$.

Algorithm DCA**:**

The algorithm generates a sequence of vertices $v_1, v_2, \ldots,$ until vertex 1 is reached.

Step 1 Carry out a random walk on G until it is mixed; i.e., until the variation distance between the current vertex and the steady state is $o(1)$. We let v_1 be the terminal vertex of the walk. (See Remark 1.1 for comments on this step.)

Step 2 $t \leftarrow 1$.

Step 3 **repeat**

Step 4 Let $C_t = \{w_1, w_2, \dots, w_{m/2}\}$ be the $m/2$ neighbors of v_t of largest degree.
(In the case of ties for the $m/2$th largest degree, vertices will be placed randomly in C_t in order to make $|C_t| = m/2$. Also m is large here and we could replace $m/2$ by $\lfloor m/2 \rfloor$ if m is odd without affecting the analysis by very much.)

Step 5 Choose v_{t+1} randomly from C_t.

Step 6 $t \leftarrow t + 1$.

Step 7 **until** $d_n(v_t) \geq \frac{n^{1/2}}{\log^{1/100} n}$ (**SUCCESS**) or $t > 2\omega \log n$ (**FAILURE**), where $\omega \to \infty$ is arbitrary.

Step 8 Assuming Success, starting from v_T, where T is the value of t at this point, do a random walk on the vertices of degree at least $\frac{n^{1/2}}{\log^{1/20} n}$ until vertex 1 is reached.

Remark 1.1. *It is known that w.h.p. the mixing time of a random walk on G_n is $O(\log n)$, see Mihail, Papadimitriou and Saberi [11]. So we can assume that the distribution of v_1 is close to the steady state $\pi_v = \frac{d_n(v)}{2mn}$.*

Note that Algorithm DCA is a local algorithm in a strong sense: the algorithm only requires access to the current vertex and its neighborhood. (Unlike the algorithm from [5], it does not need access to the neighborhood of the entire set $P_t = \{v_1, \dots v_t\}$ of vertices visited so far.) Our main result is the following:

Theorem 1.2. *If m is sufficiently large then then w.h.p. Algorithm* DCA *finds vertex 1 in G_n in $O(\omega \log n)$ time.*

DCA is thus currently the fastest as well as the "most local" algorithm to find vertex 1. We conjecture that the factor ω in the running time is unnecessary.

Conjecture 1.3. *Algorithm finds vertex 1 in G_n in $O(\log n)$ time, w.h.p.*

We note that w.h.p. the diameter of G_n is $\sim \frac{\log n}{\log \log n}$ and so we cannot expect to improve the execution time much below $O(\log n)$.

The bulk of our proof consists of showing that the execution of Steps 2–7 requires only time $O(\omega \log n)$ w.h.p. for any $\omega = \omega(n) \to \infty$. This analysis requires a careful accounting of conditional probabilities. This is facilitated by the conditional model of the preferential attachment graph due to Bollobás and Riordan [4]. One contribution of our paper is to recast their model in terms of sums of independent copies of the rate one exponential random variables; this will be essential to our analysis.

Outline of the paper

In Section 2 we reformulate the construction of Bollobás and Riordan [4] in terms of sums of independent copies of the exponential random variable of rate one.

Section 3 is the heart of the paper. The aim is to show that if v_t is not too small, then the ratio v_{t+1}/v_t is bounded above by 3/4 in expectation. We deduce from this that w.h.p. the main loop, Steps 2–7, only takes $O(\omega \log n)$ rounds. The idea is to determine a degree bound Δ such that many of v_t's left neighbors have degree at least Δ, while only few of v_t's right neighbors have degree at least Δ. In this way, v_{t+1} is likely to be significantly smaller than v_t.

Once we find a vertex v_T of high enough degree, then we know that w.h.p. v_T is not very large and lies in a small connected subgraph of vertices of high degree that contains vertex one. Then a simple argument based on the worst-case covertime of a graph suffices to show that only $o(\log n)$ more steps are required.

Our proofs will use various parameters. For convenience, we collect here in table form a dictionary of some notations, giving a brief (and imprecise) description of the role each plays in our proof, for later reference.

	Definition	**Role in proof**
$\omega :=$	$O(\log \log n)$	An arbitrarily chosen slowly growing fucntion.
$\lambda_0 :=$	$\dfrac{1}{\log^{40/m} n}$	A (usually valid) lower bound on random variables η_i (cf. Section 2.1).
$n_1 :=$	$\log^{1/100} n$	W.h.p. the main loop never visits $v \leq n_1$.
$P_t :=$	$\{v_1, \dots v_t\}$	The set of vertices visited up to time t.
$\Psi :=$	$(\log \log n)^{10}$	Vertices $v > \Psi v_t$ will not be important in the search for v_{t+1}.
$L :=$	$m^{1/5}$	A large constant, significantly smaller than m.

Notation: We write $A_n \sim B_n$ if $A_n = (1 + o(1))B_n$ as $n \to \infty$. We write $\alpha \lesssim \beta$ in place of $\alpha \leq o(1) + (1 + o(1))\beta$.

2 Preliminaries

2.1 A different model of the preferential attachement graph

Bollobás and Riordan [4] gave an ingenious construction equivalent to the preferential attachment graph model. We choose $x_1, x_2, \ldots, x_{2mn}$ independently and uniformly from $[0,1]$. We then let $\{\ell_i, r_i\} = \{x_{2i-1}, x_{2i}\}$ where $\ell_i < r_i$ for $i = 1, 2, \ldots, mn$. We then sort the r_i in increasing order $R_1 < R_2 < \cdots < R_{mn}$ and let $R_0 = 0$. We then let

$$W_j = R_{mj} \text{ and } w_j = W_j - W_{j-1} \text{ and } I_j = (W_{j-1}, W_j]$$

for $j = 1, 2, \ldots, n$. Given this we can define G_n as follows: It has vertex set $V_n = [n]$ and an edge $\{x, y\}$, $x \leq y$ for each pair ℓ_i, r_i, where $\ell_i \in I_x$ and $r_i \in I_y$.

We recast the construction of Bollobás and Riordan as follows: we can generate the sequence $R_1, R_2, \ldots, R_{mn}$ by letting

$$R_i = \left(\frac{\Upsilon_i}{\Upsilon_{mn+1}}\right)^{1/2}, \tag{1}$$

where $\Upsilon_0 = 0$ and

$$\Upsilon_N = \xi_1 + \xi_2 + \cdots + \xi_N \text{ for } N \geq 1$$

and where $\xi_1, \xi_2, \ldots, \xi_{mn+1}$ are independent exponential rate one random variables i.e. $\mathbf{Pr}(\xi_i \geq x) = e^{-x}$ for all i. This is because $r_1^2, r_2^2, \ldots, r_{mn}^2$ are independent and uniform in $[0,1]$ (as they are each chosen as the maximum of two uniform points) and the order statistics of N independent uniform $[0,1]$ random variables can be expressed as the ratios $\Upsilon_i/\Upsilon_{N+1}$ for $1 \leq i \leq N$.

We refer to the distribution of Υ_N as $\mathrm{ERL}(N)$, as it is known in the literature as the Erlang distribution.

2.2 Important properties

The advantage of our modification of the variant of the Bollobàs and Riordan construction is that if we define

$$\eta_i := \xi_{(i-1)m+1} + \xi_{(i-1)m+2} + \cdots + \xi_{im},$$

then η_i is closely related to the size of I_i. It can then be used to estimate the degree of vertex i. This will simplify the analysis since η_i is simply a sum of exponentials.

In this section, we make this claim (along with other more obscure asymptotic properties of this model) precise. In particular, we let $\mathcal{E}$ denote the event that the following properties hold for G_n. In the appendix, we prove that G_n has all these properties w.h.p.

(P1) For $\Upsilon_{k,\ell} = \Upsilon_k - \Upsilon_\ell$, we have

$$\Upsilon_{k,\ell} \in (k-\ell)\left[1 \pm \frac{L\theta_{k,\ell}^{1/2}}{3(k-\ell)^{1/2}}\right]$$

for $(k,\ell) = (mn+1, 0)$ or

$$\frac{k-\ell}{m} \in \{\omega, \omega+1, \ldots, n\} \text{ and } k-l \geq \begin{cases} 1 & l=0 \\ \log^2 n & k \geq \log^{30} n, l>0 \\ \log^{1/300} n & 0<l<k<\log^{30} n. \end{cases}$$

Here, where $n_0 = \frac{\lambda_0^2 n}{\omega \log^2 n}$, $\lambda_0 = \frac{1}{\log^{20/m} n}$,

$$\theta_{k,\ell} = \begin{cases} \log k & \omega \leq l < k \leq \log^{30} n \\ k^{1/2} & \omega \leq k \leq n^{2/5}, l=0 \\ (k-\ell)^{1/2} & \log^{30} n < k \leq n^{2/5} \\ \frac{(k-\ell)^{3/2}\log n}{n^{1/2}} & n^{2/5} < k \leq n_0. \\ \frac{n}{\omega^{3/2}\log^2 n} & n_0 < k. \end{cases}$$

(P2) $W_i \in \left(\frac{i}{n}\right)^{1/2}\left[1 \pm \frac{L\theta_i^{1/2}}{i^{1/2}}\right] \sim \left(\frac{i}{n}\right)^{1/2}$ for $\omega \leq i \leq n$.

(P3) $w_i \in \frac{\eta_i}{2m(in)^{1/2}}\left[1 \pm \frac{2L\theta_i^{1/2}}{m^{1/2} i^{1/2}}\right] \sim \frac{\eta_i}{2m(in)^{1/2}}$ for $\omega \leq i \leq n$.

(P4) $\lambda_0 \leq \eta_i \leq 40m \log\log n$ for $i \in [\log^{30} n]$.

(P5) $\eta_i \leq \log n$ for $i \in [n]$.

Some properties give asymptotics for intermediate quantities in the Bollobas/Riordan model (e.g., **(P2)**, **(P3)**), while the rest give worst-case bounds on parameters in various ranges for i. The very technical **(P1)** is just giving constraints on the gaps between the points Υ_k in the Bollobas/Riordan model.

2.3 Inequalities

We will use the following inequalities from Hoeffding [9] at several points in the paper. Let $Z = Z_1 + Z_2 + \ldots + Z_N$ be the sum of independent $[0,1]$ random variables and suppose that $\mu = \mathbf{E}(Z)$. Then if $\alpha > 1$ we have

$$\mathbf{Pr}(Z \geq (1+\alpha)\mu) \leq \exp\left\{-\frac{\alpha^2\mu}{2+\alpha/3}\right\} \leq \begin{cases} \exp\left\{-\frac{\alpha^2\mu}{3}\right\}, & \alpha \leq 1. \\ \exp\left\{-\frac{\alpha\mu}{3}\right\}, & \alpha > 1. \end{cases} \tag{2}$$

$$\mathbf{Pr}(Z \geq \beta\mu) \leq e^{-\mu}\left(\frac{e}{\beta}\right)^{\beta\mu}, \ \beta > 1 \tag{3}$$

$$\mathbf{Pr}(Z \leq (1-\alpha)\mu) \leq \exp\left\{-\frac{\alpha^2\mu}{2}\right\}, \quad 0 \leq \alpha \leq 1. \tag{4}$$

Our main use for these inequalities is to get a bound on vertex degrees, see Section 2.4.

In addition to these concentration inequalities, we use various inequalities bounding the tails of the random variable η. We note that the probability density $\phi(x)$ of the sum η of m independent exponential rate one random variables is given by

$$\phi(x) := \frac{x^{m-1}e^{-x}}{(m-1)!}.$$

That is,

$$\mathbf{Pr}(a \leq \eta \leq b) = \int_a^b \phi(y)dy. \tag{5}$$

The equation (5) is a standard result, which can be verified by induction on m (for example, see exercise 4.14.10 of Grimmett and Stirzaker [8]). Although we will frequently need to bound the probability (5), this integral cannot be evaluated exactly in general, and thus we will often use simple bounds on $\phi(\eta)$. We summarise what we need in the following lemma:

Lemma 2.1.

(a)

$$\mathbf{Pr}(\eta \leq xm) \leq m(xe^{1-x})^m \quad \text{for } x \leq 1 - \frac{1}{m}. \tag{6}$$

(b)

$$\mathbf{Pr}(\eta \leq x) \leq (1-e^{-x})^m \leq x^m. \tag{7}$$

(c)

$$\mathbf{Pr}(\eta \geq \beta m) \leq \left(\frac{e\beta}{e^\beta}\right)^m \leq e^{-3m/10} \text{ for } \beta \geq 2.$$

(d)

$$\mathbf{Pr}(\eta \geq (1+\alpha)m) \leq e^{-\alpha^2 m/3} \text{ for } 0 < \alpha < 1.$$

(e)

$$\mathbf{Pr}(\eta \leq (1-\alpha)m) \leq e^{-\alpha^2 m/2} \text{ for } 0 < \alpha < 1.$$

Proof. (a) $\phi(\eta)$ is maximized at $\eta = m-1$. Taking $\phi(mx)$ ($x \leq 1 - 1/m$) as an upper bound on $\phi(y)$ for $y \in [0, mx]$ and $m! \geq (m/e)^m$ in (5) gives us (6).

(b) Writing $\eta = \xi_1 + \xi_2 + \cdots + \xi_m$ we have $\mathbf{Pr}(\eta \leq x) \leq \prod_{i=1}^m \mathbf{Pr}(\xi_i \leq x)$.

(c) If $\eta = \xi_1 + \xi_2 + \cdots + \xi_N$, then with $\lambda = (\beta - 1)/\beta$,

$$\mathbf{Pr}(\eta \geq \beta m) = \mathbf{Pr}(e^{\lambda\eta} \geq e^{\lambda\beta m}) \leq e^{-\lambda\beta m}\,\mathbf{E}(e^{\lambda\eta}) = e^{-\lambda\beta m}\prod_{i=1}^{m}\mathbf{E}(e^{\lambda\xi_i}) = $$
$$e^{-\lambda\beta m}(1-\lambda)^{-m} = (\beta e^{-(\beta-1)})^m. \quad (8)$$

(d) Putting $\beta = 1 + \alpha$ into (8) we see that

$$\mathbf{Pr}(\eta \geq (1+\alpha)m) \leq ((1+\alpha)e^{-\alpha})^m \leq e^{-\alpha^2\alpha m/3}.$$

(e) With $\lambda = \alpha/(1-\alpha)$ we now have

$$\mathbf{Pr}(\eta \leq (1-\alpha)m) = \mathbf{Pr}(e^{-\lambda\eta} \geq e^{-\lambda(1-\alpha)m}) \leq e^{\lambda(1-\alpha)m}\,\mathbf{E}(e^{-\lambda\eta}) = e^{\lambda(1-\alpha)m}\prod_{i=1}^{m}\mathbf{E}(e^{-\lambda\xi_i}) = $$
$$e^{\lambda(1-\alpha)m}(1+\lambda)^{-m} = ((1-\alpha)e^{\alpha})^m \leq e^{-\alpha^2 m/2}.$$

□

2.4 Properties of the degree sequence

We will use the following properties of the degree sequence throughout: let

$$\zeta(i) = \left(\frac{n}{i}\right)^{1/2}\left(1 - \left(\frac{i}{n}\right)^{1/2} - \frac{5L\log\log n}{\omega^{3/4}\log n}\right) \quad (9)$$
$$\zeta^+(i) = \left(\frac{n}{i}\right)^{1/2}\left(1 - \left(\frac{i}{n}\right)^{1/2} + \frac{5L\log\log n}{\omega^{3/4}\log n}\right). \quad (10)$$

Note that

$$\zeta(i) \sim \zeta^+(i) \qquad \text{if } i \leq n\left(1 - \frac{2\log\log n}{\log n}\right). \quad (11)$$
$$\zeta(i) \sim \left(\frac{n}{i}\right)^{1/2} \qquad \text{if } i = o(n). \quad (12)$$

Also, let $\bar{d}_n(i)$ denote the expected value of $d_n(i)$ in G_n.

Lemma 2.2.

(a) If $\mathcal{E}$ occurs then $\bar{d}_n - m \in [\eta_i\zeta(i), \eta_i\zeta^+(i)]$.

(b) $\mathbf{Pr}(d_n(i) - m \leq (1-\alpha)\eta_i\zeta(i)) \leq e^{-\alpha^2\eta_i\zeta(i)/2}$ for $0 \leq \alpha \leq 1$.

(c) $\mathbf{Pr}(d_n(i) - m \geq (1+\alpha)\eta_i\zeta^+(i)) \leq e^{-\alpha^2\eta_i\zeta^+(i)/3}$ *for* $0 \leq \alpha \leq 1$.

(d) $\mathbf{Pr}(d_n(i) - m \geq \beta\eta_i\zeta^+(i)) \leq (e/\beta)^{\beta\eta_i\zeta^+(i)}$ *for* $\beta \geq 2$.

(e) *W.h.p.* $\eta_i \geq \lambda_0$ *and* $\omega \leq i \leq n^{1/2}$ *implies that* $d_n(i) \sim \eta_i \left(\frac{n}{i}\right)^{1/2}$.

(f) *W.h.p.* $\omega \leq i \leq \log^{30} n$ *implies that* $d_n(i) \sim \eta_i \left(\frac{n}{i}\right)^{1/2}$.

(g) *W.h.p.* $\omega \leq i \leq n^{1/2}$ *implies* $d_n(i) \lesssim \max\{1, \eta_i\} \left(\frac{n}{i}\right)^{1/2}$.

(h) *W.h.p.* $n^{1/2} \leq i \leq n$ *implies* $d_n(i) \leq n^{1/3}$.

(i) *W.h.p.* $1 \leq i \leq \log^{1/49} n$ *implies that* $d_n(i) \geq \frac{n^{1/2}}{\log^{1/20} n}$.

(j) *W.h.p.* $d_n(i) \geq \frac{n}{\log^{1/20} n}$ *implies* $i \leq \log^{1/9} n$.

Proof. We defer the proof, which is straightforward but tedious, to the appendix. □

Remark 2.3. We will for the rest of the paper condition on the occurrence of $\mathcal{E}$. All probabilities include this conditioning. We will omit the conditioning in the text in order to simplify expressions.

3 Analysis of the main loop

Since the variation distance after Step 1 is $o(1)$, it suffices to prove Theorem 1.2 under the assumption that we begin Step 2, with v_1 chosen randomly, exactly according to the stationary distribution.

The main loop consists of Steps 2–7. Let $v_0 = 1$ and $v_1, v_2, \ldots, v_s$ for $s \geq 1$ be the sequence of vertices followed by the algorithm up to time s. Let $\rho_t = v_{t+1}/v_t$, and define T_1, T_2 by

$$T_1 = \min\left\{t : v_t \leq \log^{30} n\right\} \text{ and } T_2 = T_1 + 30\omega\log_{4/3}\log n \text{ and } T_0 = \min\left\{2\omega\log_{4/3} n, T_2\right\}. \quad (13)$$

We will prove, see Lemma 3.2, that

$$\mathbf{E}(\rho_t) \leq \frac{3}{4} \text{ for } 1 \leq t \leq T_0. \quad (14)$$

Recalling that T is the time when Step 8 begins, we note that if $T < t \leq T_0$ then this statement is meaningless. So, we will keep to the following notational convention: if X_t is some quantity that depends on $t \leq T$ and $t > T$ then $X_t = X_T$.

Now, roughly speaking, if $r = 2\log_{4/3} n$ and μ is the number of steps in the main loop, then we would hope to have

$$\mathbf{Pr}(\mu \geq r) \leq \mathbf{Pr}\left(\rho_0\rho_1\cdots\rho_r \geq \frac{1}{n}\right) \leq n\,\mathbf{E}(\rho_0\rho_1\cdots\rho_r) \leq \frac{1}{n}$$

and so w.h.p. the algorithm will complete the main loop within $2\log_{4/3} n$ steps. Unfortunately, we cannot justify the last inequality, seeing as the ρ_t are not independent. I.e. we cannot replace $\mathbf{E}(\rho_0\rho_1\cdots\rho_r)$ by $\prod_{i=0}^r \mathbf{E}(\rho_i)$. We proceed instead as in the next lemma.

Lemma 3.1. *Assuming* (14) *we have the w.h.p.* DCA *completes the main loop in at most* T_0 *steps.*

Proof. We let s_0 denote the number of vertices visited by the main loop, and then define $Z_s = \rho_0\rho_1\cdots\rho_s$ for $s \leq s_0$, and $Z_s = \rho_0\rho_1\cdots\rho_{s_0}(\frac{3}{4})^{s-s_0}$ for $s > s_0$.

Suppose first that $T_1 > \omega\log_{4/3} n$. Now (14) and Jensen's inequality implies that for $s \geq 1$,

$$\begin{aligned}\mathbf{E}(\log(Z_s)) = \sum_{i=0}^{\min(s,s_0)} \mathbf{E}(\log(\rho_i)) + \sum_{\min(s,s_0)+1}^{s} \log\tfrac{3}{4} \\ \leq \sum_{i=0}^{\min(s,s_0)} \log\mathbf{E}(\rho_i) + \sum_{\min(s,s_0)+1}^{s} \log\tfrac{3}{4} \leq s\log(3/4). \quad (15)\end{aligned}$$

Now

$$\log(Z_s) \geq (s-s_0)\log(3/4) - \log n \geq s\log(3/4) - \log n \quad (16)$$

since $\rho_1\rho_2\ldots\rho_{s_0} \geq 1/n$.

Now let

$$\alpha = \mathbf{Pr}(\log(Z_s) \leq (1-\beta)s\log(3/4))$$

where α, β are to be determined. Then, (15), (16) imply that

$$(1-\alpha)(1-\beta)s\log(3/4) + \alpha(s\log(3/4) - \log n) \leq \mathbf{E}(\log(Z_s)) \leq s\log(3/4). \quad (17)$$

Equation (17) then implies that

$$\alpha \geq \frac{\beta s\log(4/3)}{\beta s\log(4/3) + \log n}. \quad (18)$$

Now putting $s = \omega\log_{4/3} n$ and $\beta = 1/2$ we see that (18) becomes

$$\alpha \geq 1 - \frac{2}{\omega+2} = 1 - o(1).$$

So w.h.p. after at most $\omega\log_{4/3} n$ steps, we will have exited the main loop.

Suppose now that $T_1 \leq \omega\log_{4/3} n$. Using the argument that gave us (18) we obtain

$$T - T_1 \leq \omega\log_{4/3}\log^{30} n \ w.h.p.$$

$\square$

To prove Lemma 3.2, we will use a method of deferred decisions, exposing various parameters of G_n as we proceed. At time t, we will consider all random variables in the model from Section 2.1 as being exposed if they have affected the algorithm's trajectory thus far, and condition on their particular evaluation. To reduce the conditioning necessary, we will actually analyze a modified algorithm, **NARROW-DCA**(τ), and then later show that the trajectory of **NARROW-DCA**(τ) is the same as that of the **DCA** algorithm, w.h.p., when identical sources of randomness are used.

NARROW-DCA(τ) is the same as the DCA algorithm, except that for the first τ rounds of the algorithm, a modified version of **Step 4** is used:
Modified Step 4 Let

$$C_t = \{w_1, w_2, \ldots, w_{m/2}\}$$

be the $m/2$ neighbors of v_t of largest degree from $\{1, \ldots, \Psi v_t\}$ where $\Psi := (\log\log n)^{10}$.

For rounds $\tau+1, \tau+2, \ldots$, the behavior of NARROW-DCA(τ) is the same as for DCA.

Notice that **NARROW-DCA** "cheats" by using the indices of the vertices, which we do not actually expect to be able to use. Nevertheless, we will see later that w.h.p., for $\tau = 2\omega \log_{4/3} n$, the path of this algorithm is the same as for the DCA algorithm, justifying its role in our analysis.

3.1 Analyzing one step

Our analysis of one step of the main loop consists of the following lemma:

Lemma 3.2. *Let ρ_t be the ratio of v_{t+1}/v_t which appears in a run of the algorithm* **NARROW-DCA**(t)*. Then for all $t \le T_0$ (see (13)), we have that*

$$\mathbf{E}(\rho_t) \le \frac{3}{4} \quad \text{and} \quad \mathbf{Pr}(\rho_t \ge \Psi) \le \frac{1}{\log^2 n}. \tag{19}$$

The first statement ensures that **NARROW-DCA**(t) makes progress in expectation in the tth jump. The second part of this statement implies by induction that for any $t \le \omega \log n$, the behavior of **NARROW-DCA**(t) is identical to the behavior of the DCA algorithm for the first t steps. Thus together these statements give (14).

To prove Lemma 3.2, we will prove a stronger statement which is conditioned on the *history* of the algorithm at time t. The history $\mathcal{H}_t$ of the process at the *end* of step t consists of

(H1) The sequence $v_1, v_2, \ldots, v_t$.

(H2) The left-choices $\lambda(v_s, 1), \lambda(v_s, 2), \ldots, \lambda(v_s, m), 1 \le s < t$ and the corresponding left neighbors $N_L(v_s) = \{u_{1,s}, u_{2,s}, \ldots, u_{m,s}\}$.

(H3) The lists $u'_{1,s}, u'_{2,s}, \ldots, u'_{r,s}$ of all vertices $u'_{k,s}$ which have the property that (i) $v_s \in N_L(u'_{k,s})$ and (ii) $u'_{k,s} \le \Psi v_s$ for $1 \le s < t$. (It is important to notice that $s < t$ here.)

(H4) The values η_{v_i} and the intervals I_{v_i} for $i = 1, 2, \ldots, t$.

(H5) The values η_w and the intervals I_w and the degrees $\deg(w)$, for $w \in \bigcup_{i=1}^{t} N(v_i)$.

Here,

$$N(v) = N_L(v) \cup N_R(v) \text{ where } N_R(v) = \{w \leq \Psi v : v \in N_L(w)\}.$$

We note that at any step t, and for a fixed random sequence used in the **NARROW-DCA**(t) algorithm, $\mathcal{H}_t$ contains all random variables which have determined the behavior of the algorithm so far, in the sense that if we modify any random variables from the random graph model described in Section 2.1 while preserving all values in the history, then the trajectory of the algorithm will not change. We write H_t to refer to a particular evaluation of the history (so that we will be conditioning on events of the form $\mathcal{H}_t = H_t$).

Structure of the proof

The essential structure of our proof of Lemma 3.2 is as follows:

Part 1 We will define the notion of a *typical* history H_t.

Part 2 We will prove that for $t \leq T_0$ and any typical history H_t, random variables η_v which are not explicitly exposed in H_t are essentially unconditioned by the event $\mathcal{H}_t = H_t$ (Lemma 3.3).

Part 3 We will prove by induction that $\mathcal{H}_t$ is typical w.h.p., for $t \leq T_0$.

Part 4 We will use **Part 2** and **Part 3** to prove that for $t \leq T_0$ that where $\mid H_t$ is short for $\mid \mathcal{H}_t = H_t$,

$$\mathbf{E}(\rho_t \mid H_t) \leq \frac{2}{3} + \frac{21\eta_{v_t}}{mL} + \frac{L^3}{m^2} \quad \text{and} \quad \mathbf{Pr}\left(\rho_t \geq \Psi \mid H_t\right) \leq \frac{1}{\log^2 n} \tag{20}$$

by using using nearly unconditioned distributions of random variables which are not revealed in H_t to estimate the probabilities of various events. (Note that η_{v_t} in (20) is simply a real number determined by H_t.)

Part 5 We will also prove for $t \leq T_0$ that

$$\mathbf{E}(\eta_{v_{t+1}} \mid H_t) \leq 4m. \tag{21}$$

Now the expected value statement in (19) follows from (21) and the first part of (20), by removing the conditioning on H_t.

Part 1

Let P_t denote the sequence of vertices $v_1, v_2, \ldots, v_t$ determined by the history H_t. We now define the notion of a *typical* history H_t. For this purpose, we consider the reordered values $0 \le \lambda_1^{(t)} < \lambda_2^{(t)} < \cdots < \lambda_{N(t)}^{(t)}$ where

$$\Lambda_0^{(t)} = \left\{ \lambda_1^{(t)}, \lambda_2^{(t)}, \ldots, \lambda_{N(t)}^{(t)} \right\} = \left\{ \lambda(v_s, i) : 1 \le s \le t,\, 1 \le i \le m \right\}.$$

Given this we define $v = v_j^{(t)}$ to be the index such that $\lambda_j^{(t)} \in I_v$ and then let

$$V_L^{(t)} = \left\{ v_j^{(t)} : 1 \le j \le N(t) \right\}.$$

We also define

$$V_R^{(t)} = \{ v : v \in N_R(P_t) \}$$

Now let us reorder

$$V^{(t)} = \left\{ x_1^{(t)} < x_2^{(t)} < \cdots < x_{M(t)}^{(t)} \right\} = V_L^{(t)} \cup V_R^{(t)}.$$

We define the extreme points $x_0^{(t)} = 0$ and $x_{M(t)+1}^{(t)} = n + 1$ and define

$$X_j^{(t)} = [x_{j-1}^{(t)} + 1, x_j^{(t)} - 1] \quad \text{and} \quad X^{(t)} = \bigcup_{j=1}^{M(t)+1} X_j^{(t)} = [n] \setminus V^{(t)} \quad \text{and} \quad N_j^{(t)} = |X_j^{(t)}|,$$

$$U_j^{(t)} = [W_{x_{j-1}^{(t)}+1}, W_{x_j^{(t)}-1}] \quad \text{and} \quad U^{(t)} = \bigcup_{j=1}^{N} U_j^{(t)} \quad \text{and} \quad L_j^{(t)} = |U_j^{(t)}|.$$

A typical history $H_t, t \le T_0$ is now one with the following properties:

(S1) There do not exist $s_1, s_2 \le t$ such that either (i) $s_1 \le t - 2$ and v_{s_1} and v_{s_2} are neighbors or (ii) $s_1 \le t - 3$ and there exists a vertex w such that $w \in N(v_{s_1}) \cap N(v_{s_2})$. (We say that the path is *self-avoiding*.)

(S2) The points of $\Lambda^{(t)}$ are *well-separated*, in the following sense:

$$|x_j^{(t)} - x_{j-1}^{(t)}| \ge \begin{cases} \log^2 n & x_{j-1}^{(t)} \ge \log^{30} n. \\ \log^{1/400} n & \text{Otherwise.} \end{cases} \tag{22}$$

We observe that

(T1) If H_t is typical then v_{j+1} is chosen from $X^{(j)}$ for all $j < t$.

(T2) Each $U_j^{(t)}$ is the union of intervals $I_v, v \in X_j^{(t)}$.

Part 2

We prove the following:

Lemma 3.3. *For any vertex $v \in X^{(t)}$, any interval $R \subseteq \mathbb{R}$, and any typical history H_t, we have that $v \notin P_t \cup N(P_t)$ implies*

$$\mathbf{Pr}\left(\eta_v \in R \mid H_t\right) \sim \mathbf{Pr}\left(\mathrm{ERL}(m) \in R\right). \tag{23}$$

The following lemma is the starting point for the proof of Lemma 3.3.

Lemma 3.4. *Let $j \in [M(t)+1]$, let H_t be any typical history, and let X' be the value of $X_j^{(t)}$ in H_t. Then the distribution of the random variables $\eta_v, v \in X'$ conditioned on $\mathcal{H}_t = H_t$ is equivalent to the distribution of the random variables η_v, $v \in X'$ conditioned only on the relationship $\sum_{v \in X'} \eta_v = A_1^2 - A_0^2$, where A_1, A_0 are the values of $W_{x_j^{(t)}-1}$ and $W_{x_{j-1}^{(t)}+1}$, respectively, in H_t.*

Proof. Suppose we fix everything except for $\eta_v, v \in X'$. By everything we mean every other η_w and all of the $\lambda(v,i)$ and the random bits we use to make our choices in **Step 5** of DCA; we let H_t be the corresponding history. Suppose now that we replace $\eta_v, v \in X'$ with $\eta'_v, v \in X'$ without changing the sum $\sum_{v \in X'} \eta_v$. Then $W_{x_{j-1}^{(t)}+1}$ remains the same, as it depends only on η_v for $v \notin X'$, and thus $W_{x_j^{(t)}-1}$ remains the same as well, since the difference $A_1^2 - A_0^2$ is unchanged.

In particular, this implies that H_t remains a valid history. We confirm this by induction. Suppose that $H_s, s < t$ remains valid. We first note that because the $\lambda(v_s,i)$ are unchanged, none of $v_s's$ left neighbors are in $X_j^{(t)}$. Also, $N_R(v_s)$ and the vertex degrees for $w \in N_R(v_s)$ will not be affected by the change, even if $v_s < \min X_j^{(t)}$. So H_{s+1} will be unchanged, completing the induction. □

We are now ready to prove Lemma 3.3.

Proof of Lemma 3.3. Suppose that $v \in X' = X_j^{(t)}$, then $M = N_j^{(t)} \geq \zeta_n \to \infty$. We now use Lemma 3.4 to write

$$\mathbf{Pr}(\eta_v \leq x \mid H_t) = \mathbf{Pr}\left(\eta_v \leq x \,\middle|\, \sum_{w \in X'} \eta_w = A_1^2 - A_0^2,\right)$$

where A_1 and A_0 are the values of $W_{x_j^{(t)}-1}$ and $W_{x_{j-1)+1}^{(t)}}$, respectively, in H_t, so that $A_1 - A_0$ is the value of $L_j^{(t)}$ in H_t.

Now from **(P1)** we have that $A := A_1^2 - A_2^2 \in [(1-\varepsilon)mM, (1+\varepsilon)mM]$ for $M = |X'|$ w.h.p., for any $\varepsilon > 0$. Thus we fix any $\mu \in [(1-\varepsilon)mM, (1+\varepsilon)mM]$ and show that

$$\mathbf{Pr}\left(\eta_v \leq x \,\middle|\, \sum_{w \in X'} \eta_w = \mu\right) = (1 + O(\varepsilon))\,\mathbf{Pr}\left(\mathrm{ERL}(m) \leq x\right).$$

The lemma follows since ε is arbitrary.

We write

$$\begin{aligned}
&\mathbf{Pr}\left(\eta_v \le x \,\middle|\, \sum_{w\in X'} \eta_w = \mu\right) \\
&= \int_{\eta=0}^{x} \frac{\eta^{m-1}e^{-\eta}}{(m-1)!} \cdot \frac{(\mu-\eta)^{(M-1)m-1}e^{-(\mu-\eta)}}{((M-1)m-1)!} \cdot \frac{(Mm-1)!}{\mu^{Mm-1}e^{-\mu}} d\eta \\
&= \int_{\eta=0}^{x} \frac{\eta^{m-1}e^{-\eta}}{(m-1)!} \cdot \frac{\left(1-\frac{\eta}{\mu}\right)^{(M-1)m-1} e^{\eta}\prod_{i=1}^{m}(Mm-i)}{\mu^m} d\eta \\
&= \int_{\eta=0}^{x} \frac{\eta^{m-1}e^{-\eta}}{(m-1)!} \cdot \exp\left\{\eta - ((M-1)m-1)\left(\frac{\eta}{\mu} + O\left(\frac{\eta^2}{\mu^2}\right)\right)\right\} \cdot \left(1+O\left(\frac{m}{M}\right)\right) d\eta \\
&= (1+O(\varepsilon)) \int_{\eta=0}^{x} \frac{\eta^{m-1}e^{-\eta}}{(m-1)!} d\eta.
\end{aligned}$$

Here we used that H_t typical implies that $M \ge \log^{1/400} n \to \infty$. □

Part 3

In the next section we will need a lower bound on v_{t+1}. Let

$$\phi_v = \begin{cases} \frac{1}{\log^3 n} & v \ge \log^{30} n \\ \frac{1}{(\log\log n)^3}. & v < \log^{30} n. \end{cases}$$

Lemma 3.5. *W.h.p. $\rho_t \ge \phi_{v_t}$ for $1 \le t \le T_0$.*

Proof. The values of $\lambda(v_t, i), i = 1, 2, \ldots, m$ are unconditioned by H_t, see **(H2)**. It then follows from **(P2)** that if $v_t \ge \log^{30} n$ then

$$\mathbf{Pr}(v_{t+1} \le \phi_{v_t} v_t \mid H_t) \lesssim m\frac{W_{\phi_{v_t}}}{W_{v_t}} \lesssim m\phi_{v_t}^{1/2} = \frac{m}{\log^{3/2} n}. \tag{24}$$

There are $O(\omega \log n)$ choices for t and so this deals with $v_t \ge \log^{30} n$.

Now there are $O(\log\log n)$ choices of $t \in [T_1, T_0]$ for which $v_t \le \log^{30} n$. In this case we can replace the RHS of (24) by $1/(\log\log n)^{3/2}$. □

We will also need to bound the size of $N_R(v_t)$ for all t.

Lemma 3.6. *W.h.p., for all $t \le T_0$,*

$$|N_R(v_t)| \le \begin{cases} \log^3 n & v_t \ge \log^{30} n. \\ (\log\log n)^{20} & v_t \le \log^{30} n. \end{cases}$$

Proof. The size of $N_R(v), v = v_t$ is stochastically bounded by $Bin(\Psi v, \eta_v/v)$. This is because if $w \in N_R(v)$ then $w \leq \Psi v$. Also, for any such w, the probability that it has v as a left neighbor is at most $mw_v/W_w \lesssim \eta_v/(vw)^{1/2} \leq \eta_v/v$. This uses property **(S1)** to see that the values of $\lambda(w,i), i = 1,2,\dots,m$ are unconditioned by H_t. Thus, if $\theta_v = \log^3 n$ if $v \geq \log^{30} n$ and equal to $(\log\log n)^{20}$ otherwise,

$$\mathbf{Pr}(|N_R(v)| \geq \theta_v \mid H_t) \leq \binom{\Psi v}{\theta_v}\left(\frac{\eta_v}{v}\right)^{\theta_v} \leq \left(\frac{e\Psi\eta_v}{\theta_v}\right)^{\theta_v}. \tag{25}$$

If $v \geq \log^{30} n$ then the RHS of (25) is at most $(e/\log n)^{\log^3 n}$ which is clearly small enough to handle T possible values for t. If $v < \log^{30} n$ then the RHS of (25) is at most $(40e/(\log\log n)^9)^{(\log\log n)^{20}}$ which is small enough to handle $O(\omega \log\log n)$ possible values for t such that $v < \log^{30} n$. $\square$

Continuing **Part 3**, we now show that the DCA walk doesn't contain cycles.

Lemma 3.7. *W.h.p. the path $P_t, t \leq T_0$ is self avoiding.*

Proof. We proceed by induction and assume that the claim of the lemma is valid up to time $t-1$. Now consider the choice of v_t.

Case 1: There is an edge $v_s v_t$ where $s \leq t-2$:
(a): $v_t \in N_L(v_s) \cap N_L(v_{t-1})$.
We bound the probability of this (conditional on $\mathcal{E}, H_t$) asymptotically by

$$\sum_{s\in[t-2]}\sum_{v\in N_L(v_s)} \frac{mw_v}{W_{v_{t-1}}} \lesssim \sum_{s\in[t-2]}\sum_{v\in N_L(v_s)} \frac{\eta_v}{2(vv_{t-1})^{1/2}}. \tag{26}$$

Here, and throughout the proof of Case 1, v denotes a possibility for v_t and $mw_v/W_{v_{t-1}}$ bounds the probability that v_{t-1} chooses v. Remember that these choices are still uniform, given the history.

We split the sum in (26) as

$$\sum_{\substack{s\in[t-2]\\ v_s>\log^{30} n}}\sum_{v\in N_L(v_s)} \frac{\eta_v}{2(vv_{t-1})^{1/2}} + \sum_{\substack{s\in[t-2]\\ v_s\leq\log^{30} n}}\sum_{v\in N_L(v_s)} \frac{\eta_v}{2(vv_{t-1})^{1/2}}.$$

Consider the first sum. There are less than t choices for s; m choices for v and $\eta_v \leq \log n$. Now $v \in N_L(v_s)$ and Lemma 3.5 implies that $v \geq \log^{27} n$. So we can bound the first sum by

$$(\#\ of\ s)\cdot(\#\ of\ v)\cdot(\max\eta_v)\cdot\frac{1}{v^{1/2}} \leq T_0\cdot m\cdot\log n\cdot\frac{1}{\log^{27/2} n} = o\left(\frac{1}{\log^{11} n}\right).$$

Summing this estimate over $t \leq T_0$ gives $o(1)$.

For the second sum, we bound the number of choices of s by $O(\omega\log\log n)$ and η_v by $O(\log\log n)$, since $v \leq v_s$. We use the fact (see Section 3.2) that $v_{t-1} \geq \log^{1/100} n$. So we can therefore bound the second sum by

$$(\#\ of\ s)\cdot(\#\ of\ v)\cdot(\max\eta_v)\cdot\frac{1}{v_{t-1}^{1/2}}\leq_b \omega\log\log n\cdot m\cdot\log\log n\cdot\frac{1}{\log^{1/200}n}=o\left(\frac{1}{\log^{1/300}n}\right).$$

(We use $A\leq_b B$ in place of $A=O(B)$.)
There are $O(\omega\log\log n)$ choices for $T_0\geq t>s\geq T_1$ and so we can sum this estimate over choices of t.

(b): $v_t\in N_L(v_s)\cap N_R(v_{t-1})$.
Using $v_t\in N_R(v_{t-1})$, we bound the probability of this asymptotically by

$$\sum_{\substack{s\in[t-2]\\ v_s>\log^{30}n}}\sum_{v\in N_L(v_s)}\frac{\eta_{v_{t-1}}}{2(vv_{t-1})^{1/2}}+\sum_{\substack{s\in[t-2]\\ v_s\leq\log^{30}n}}\sum_{v\in N_L(v_s)}\frac{\eta_{v_{t-1}}}{2(vv_{t-1})^{1/2}}.$$

For the first sum we use the argument of Case (a) without any change, except for bounding $\eta_{v_{t-1}}$ by $\log n$ as opposed to bounding η_v by the same. This gives a bound

$$(\#\ of\ s)\cdot(\#\ of\ v)\cdot(\max\eta_{v_{t-1}})\cdot\frac{1}{v^{1/2}}\leq_b T_0\cdot m\cdot\log n\cdot\frac{1}{\log^{27/2}n}=o\left(\frac{1}{\log^{11}n}\right).$$

This is small enough to inflate by the number of choices for t.

For the second sum we split into two cases: (i) $v_{t-1}\geq\log^{30}n$ and (ii) $v_{t-1}<\log^{30}n$. This enables us to control $\eta_{v_{t-1}}$. For the first case we obtain

$$(\#\ of\ s)\cdot(\#\ of\ v)\cdot(\max\eta_{v_{t-1}})\cdot\frac{1}{v_{t-1}^{1/2}}\leq\omega\log\log n\cdot m\cdot\log n\cdot\frac{1}{\log^{15}n}=o\left(\frac{1}{\log^{13}n}\right).$$

The RHS is small enough to handle the $O(\omega\log n)$ choices for t.

For the second case we obtain

$$(\#\ of\ s)\cdot(\#\ of\ v)\cdot(\max\eta_{v_{t-1}})\cdot\frac{1}{v_{t-1}^{1/2}}\leq\omega\log\log n\cdot m\cdot\log\log n\cdot\frac{1}{\log^{1/200}n}=o\left(\frac{1}{\log^{1/300}n}\right).$$

The RHS is small enough to handle the $O(\omega\log\log n)$ choices for t.

(c): $v_t\in N_R(v_s)\cap N_L(v_{t-1})$.
Using $v_t\in N_L(v_{t-1})$, we bound the probability of this asymptotically by

$$\sum_{\substack{s\in[t-2]\\ v_s>\log^{29}n}}\sum_{v\in N_R(v_s)}\frac{\eta_v}{2(vv_{t-1})^{1/2}}+\sum_{\substack{s\in[t-2]\\ v_s\leq\log^{29}n}}\sum_{v\in N_R(v_s)}\frac{\eta_v}{2(vv_{t-1})^{1/2}}.$$

For the first sum we use $v\geq v_s$ and the argument of Case (a) without change, but notice we split over $v_s>\log^{29}n$ or not here. This gives a bound of

$$(\#\ of\ s)\cdot(\#\ of\ v)\cdot(\max\eta_v)\cdot\frac{1}{v^{1/2}}\leq T_0\cdot m\cdot\log n\cdot\frac{1}{\log^{29/2}n}=o\left(\frac{1}{\log^{12}n}\right).$$

For the second sum we use $v \leq \Psi v_s$ to bound v by $\log^{30} n$. We also use Lemma 3.6 to bound the number of choices of v by $(\log\log n)^{20}$. This gives a bound of

$$(\#\ of\ s)\cdot(\#\ of\ v)\cdot(\max \eta_v)\cdot\frac{1}{v_{t-1}^{1/2}} \leq_b \omega \log\log n\cdot(\log\log n)^{20}\cdot\log\log n\cdot\frac{1}{\log^{1/200} n} = o\left(\frac{1}{\log^{1/300} n}\right).$$

(d): $v_t \in N_R(v_s) \cap N_R(v_{t-1})$.
Using $v_t \in N_R(v_{t-1})$, we bound the probability of this asymptotically by

$$\sum_{\substack{s\in[t-2]\\ v_s>\log^{30} n}} \sum_{v\in N_R(v_s)} \frac{\eta_{v_{t-1}}}{2(vv_{t-1})^{1/2}} + \sum_{\substack{s\in[t-2]\\ v_s\leq\log^{30} n}} \sum_{v\in N_R(v_s)} \frac{\eta_{v_{t-1}}}{2(vv_{t-1})^{1/2}}.$$

For the first sum we use $v \geq v_s$ and Lemma 3.6 to bound the number of choices for v and then we have a bound of

$$(\#\ of\ s)\cdot(\#\ of\ v)\cdot(\max \eta_{v_{t-1}})\cdot\frac{1}{v^{1/2}} \leq T_0\cdot\log^3 n\cdot\log n\cdot\frac{1}{\log^{15} n} = o\left(\frac{1}{\log^9 n}\right).$$

For the second sum we split into two cases: (i) $v_{t-1} \geq \log^{30} n$ and (ii) $v_{t-1} < \log^{30} n$. This enables us to control $\eta_{v_{t-1}}$. We also use Lemma 3.6 to bound the number of choices for v in each case. Thus in the first case we have the bound

$$(\#\ of\ s)\cdot(\#\ of\ v)\cdot(\max \eta_{v_{t-1}})\cdot\frac{1}{v_{t-1}^{1/2}} \leq_b \omega \log\log n\cdot\log^3 n\cdot\log n\cdot\frac{1}{\log^{15} n} = o\left(\frac{1}{\log^{10} n}\right).$$

In the second case we have

$$(\#\ of\ s)\cdot(\#\ of\ v)\cdot(\max \eta_{v_{t-1}})\cdot\frac{1}{v_{t-1}^{1/2}} \leq_b \omega \log\log n\cdot(\log\log n)^{20}\cdot\log\log n\cdot\frac{1}{\log^{1/200} n} = o\left(\frac{1}{\log^{1/300} n}\right).$$

Case 2: There is a path v_s, v, v_t where $s < t$.
The calculations that we have done for Case 1 carry through unchanged. We just replace v_{t-1} by v_t throughout the calculation and treat v as an arbitrary vertex as opposed to a choice of v_t. □

The $x_j^{(t)}$ are separated

We now prove that w.h.p. points λ_i are well-separated. Let

$$J_1 = \left\{j : v_j \geq \log^{30} n\right\}.$$

Lemma 3.8. *Equation (22) holds w.h.p. for all $t \leq T_0$.*

Proof. We consider cases.
Case 1: $x_{j-1}^{(t)}, x_j^{(t)} \in V_R^{(t)}$.
For this we write

$$\zeta_{v,w} = \begin{cases} \log^2 n & \min\{v,w\} \geq \log^{30} n. \\ \log^{1/300} n & \text{Otherwise.} \end{cases}.$$

$$
\begin{aligned}
&\mathbf{Pr}(\exists 1 \le s \le t, v \in N_R(v_s), w \in N_R(v_t) : |v-w| \le \zeta_{v,w} \mid \mathcal{E}, H_t) \le \\
&\lesssim \sum_{1\le s\le t\le T_0} \sum_{\substack{v\in N_R(v_s), w\in N_R(v_t)\\ |v-w|\le\zeta_{v,w}}} \frac{\eta_{v_s}\eta_{v_t}}{(v_s v_t v w)^{1/2}} \\
&\le \sum_{1\le s\le t\le T_0} \sum_{v\in N_R(v_s)} \frac{\zeta_{v,w}\eta_{v_s}\eta_{v_t}}{(v-\zeta_{v,w})(v_s v_t)^{1/2}} \qquad (27)\\
&\le 2 \sum_{1\le s\le t\le T_0} \frac{\zeta^*_{s,t}\eta_{v_s}\eta_{v_t}|N_R(v_s)|}{(v_s v_t)^{1/2}}. \qquad (28)
\end{aligned}
$$

Here $\zeta^*_{s,t}$ will be a bound on the possible value of $\zeta_{v,w}$ in (27).

Case 1a: $\max\{v_s, v_t\} \ge \log^{29} n$**:**
In this case $\zeta^*_{s,t} \le \log^2 n$ and we can bound the summand of (28) by

$$\zeta^*_{s,t} \cdot \log^2 n \cdot \log^3 n \cdot \frac{1}{\log^{29/2} n} = \frac{1}{\log^{15/2} n}.$$

Multiplying by a bound T_0^2 on the number of summands gives a bound of $o(1)$. Here, and in the next case, we use Lemma 3.6 to bound $|N_R(v_s)|$.

Case 1b: $\max\{v_s, v_t\} < \log^{29} n$**:**
Here we have $\max\{v, w\} \le \Psi \log^{29} n \le \log^{30} n$. In this case we can bound the summand of (28) by

$$\log^{1/300} n \cdot (40m \log\log n)^2 \cdot (\log\log n)^{20} \cdot \frac{1}{\log^{1/100} n} = o\left(\frac{(\log\log n)^5}{\log^{1/200}}\right).$$

We only have to inflate this by $(T_0 - T_1)^2 = O((\omega \log\log n)^2)$. This completes the case where $x^{(t)}_{j-1}, x^{(t)}_j \in R^{(t)}$.

Case 2: $x^{(t)}_{j-1}, x^{(t)}_j \in V^{(t)}_L$
We first show that the gaps $\lambda_j - \lambda_{j-1}$ are large. Define

$$\beta_1 = \frac{\log^{15/2} n}{n^{1/2}} \text{ and } \beta_2 = \frac{\log^{1/300} n}{n^{1/2}}.$$

and

$$\varepsilon_j = \begin{cases} \beta_1 & \lambda_j = \lambda(v_t, i), v_t \in J_1. \\ \beta_2 & \textit{otherwise.} \end{cases}$$

and

$$\sigma_1 = \frac{1}{\log^{15/2} n} \text{ and } \sigma_2 = \frac{1}{\log^{1/200} n}.$$

We drop the superscript t for the rest of the lemma.

Claim 3.9.

$$\mathbf{Pr}(\exists \lambda_j \in \Lambda_0 : \lambda_{j-1} > \lambda_j - \varepsilon_j \mid H_t) = o(1).$$

Proof of Claim 3.9. This follows from the fact that

$$\mathbf{Pr}(\exists j : \lambda_{j-1} > \lambda_j - \varepsilon_j) \leq o(1) + (1+o(1))(m^2T_1^2\sigma_1 + m^2(T_0-T_1)(T_1\sigma_1 + (T_0-T_1)\sigma_2)).$$

We have fewer than $m^2T_1^2$ choices for $s = \tau(j-1), t = \tau(j) \in J_1$. Assume first that $s < t$. Given such a choice, we have that w.h.p. $W_{v_t} \gtrsim \log^{15} n/n^{1/2}$ by **(P2)**. Now λ_j will have been chosen uniformly from 0 to $\approx W_{v_t}$ and so the probability it lies in $[\lambda_{j-1}, \lambda_{j-1}+\varepsilon_j]$ is at most $\approx \beta_1/W_{v_s}$, which explains the term $m^2T_1^2\sigma_1$. If $s > t$ then we repeat the above argument with $[\lambda_{j-1}, \lambda_{j-1}+\varepsilon_1]$ replaced by $[\lambda_j - \varepsilon_1, \lambda_j]$

The term $m^2(T_0-T_1)T_1\sigma_1$ arises in the same way with $j-1 \in J_1, j \notin J_1$ or vice-versa.

The term $m^2(T_0-T_1)^2\sigma_2$ arises from the case where $j-1, j \notin J_1$. Here we can only assume that $W_j \gtrsim \log^{1/200} n/n^{1/2}$. This follows from **(P2)**, **(P4)** and Lemma 2.2 and the fact that we exit the main loop when we see a vertex of degree at least $n^{1/2}/\log^{1/100} n$. Assuming that $s < t$ we see that the probability that λ_j lies in $[\lambda_{j-1}, \lambda_{j-1}+\varepsilon_2]$ is at most $\beta_2/W_{v_t} \sim \beta_2/(\log^{1/200} n/n^{1/2}) = o(1)$. $\square$

Given the Claim and **(P4)**, **(P5)** we have that w.h.p.

$$W_{v_j^{(t)}-1} - W_{v_{j-1}^{(t)}+1} \geq \begin{cases} \beta_1 - \frac{\log n}{n^{1/2}} \geq \frac{1}{2}\beta_1 & j \in J_1. \\ \beta_2 - \frac{40m\log\log n}{n^{1/2}} \geq \frac{1}{2}\beta_2 & j \notin J_1 \end{cases} \tag{29}$$

Now,

$$W_{v_j^{(t)}-1} - W_{v_{j-1}^{(t)}+1} = \left(\frac{\Upsilon_{m(\tau(j)-1)}}{\Upsilon_{mn+1}}\right)^{1/2} - \left(\frac{\Upsilon_{m(\tau(j-1)+1)}}{\Upsilon_{mn+1}}\right)^{1/2} = \frac{\Upsilon_{m(\tau(j)-1)} - \Upsilon_{m(\tau(j-1)+1)}}{\Upsilon_{mn+1}^{1/2}(\Upsilon_{m(\tau(j)-1)}^{1/2} + \Upsilon_{m(\tau(j-1)+1)}^{1/2})}.$$

Or,

$$\sum_{u\in X_j^{(t)}} \eta_u = (W_{v_j^{(t)}-1} - W_{v_{j-1}^{(t)}+1})\Upsilon_{mn+1}^{1/2}(\Upsilon_{m(\tau(j)-1)}^{1/2} + \Upsilon_{m(\tau(j-1)+1)}^{1/2}) \geq \begin{cases} \beta_1 n^{1/2} & j \in J_1. \\ \beta_2 n^{1/2} & j \notin J_1 \end{cases}.$$

It follows that w.h.p.

$$|x_{j-1}^{(t)} - x_j^{(t)}| = |X_j^{(t)}| \geq \begin{cases} \frac{\beta_1 n^{1/2}}{\log n} & j \in J_1. \\ \frac{\beta_2 n^{1/2}}{40m\log\log n} & j \notin J_1 \end{cases}.$$

Case 3: $x_j^{(t)} \in V_L^{(t)}, x_{j-1}^{(t)} \in V_R^{(t)}$
Let $\theta_v = \beta_1, v \geq \log^{30} n$ and $\theta_v = \beta_2$ otherwise. We write

$$\begin{aligned} \mathbf{Pr}(\exists s < t, v, k : v \in N_R(v_s), \lambda(v_t, k) \in I_v \pm \theta_v \mid \mathcal{E}, H_t) &\leq \sum_{s,t,v,k} \frac{\eta_{v_s}}{(vv_s)^{1/2}} \cdot \frac{w_v + 2\theta_v}{W_{v_t}} \\ &\lesssim \sum_{s,t,v,k} \left(\frac{\eta_{v_s}\eta_v}{v(v_sv_t)^{1/2}} + \frac{2n^{1/2}\theta_v\eta_{v_s}}{v(v_sv_t)^{1/2}}\right) \end{aligned} \tag{30}$$

We bound the sum in the RHS of (30) as follows: If $\max\{v, v_s\} \geq \log^{30} n$ then we bound the first sum by

$$(\#s,t,k)\cdot \log^2 n \cdot \sum_{v=1}^{n} \frac{1}{v}\cdot \frac{1}{\log^{15} n} \leq_b (\omega^2 \log^2 n)\cdot \log^2 n \cdot \log n \cdot \frac{1}{\log^{15} n} = o(1).$$

We bound the second sum by

$$(\#s,t,k)\cdot 2\log^{15/2} n \cdot \log n \cdot \sum_{v=1}^{n} \frac{1}{v}\cdot \frac{1}{\log^{15} n} \leq (\omega^2 \log^2 n)\cdot \log^{15/2} n \cdot \log n \cdot \log n \cdot \frac{1}{\log^{15} n} = o(1).$$

When $\max\{v, v_s, v_t\} < \log^{30} n$ we bound the first sum by

$$(\#s,t,k)\cdot (40\log\log n)^2 \cdot \sum_{v=1}^{\log^{30} n} \frac{1}{v}\cdot \frac{1}{\log^{1/200} n} \leq_b (\omega \log\log n)^2 \cdot (\log\log n)^2 \cdot \log\log n \cdot \frac{1}{\log^{1/200} n} = o(1). \quad (31)$$

We bound the second sum by

$$(\#s,t,k)\cdot \log^{1/300} n \cdot 40\log\log n \cdot \sum_{v=1}^{\log^{30} n} \frac{1}{v}\cdot \frac{1}{\log^{1/200} n} \leq_b (\omega \log\log n)^2 \cdot \log^{1/300} n \cdot \log\log n \cdot \frac{1}{\log^{1/200} n} = o(1). \quad (32)$$

Finally, if $\max\{v, v_s\} < \log^{30} n \leq \log^{30} n$ then we have to replace $(\#s,t,k)$ in (31), (32) by $O(\omega^2 \log n \log\log n)$. But this is compensated by a factor $1/v_t^{1/2} \leq 1/\log^{15} n$.

It follows that (29) holds w.h.p. and the proof continues as for Case 2.

$\square$

Part 4

We now assume $t \leq T_0$. We begin by showing that DCA only uncovers a small part of the distribution of the η's.

Let $\Xi_t = P_t \cup N(P_t)$ and

$$S_{t,j} = \sum_{v\in \Xi_t}^{t} w_v.$$

Lemma 3.10. *W.h.p., $S_{t,j} = o(W_j)$ for $\log^{1/100} n \leq j$ and $1 \leq t \leq T_0$.*

Proof. Assume first that $j \geq \log^{30} n$. It follows from **(P2)**, **(P3)**, **(P5)** and Lemma 3.6 that w.h.p.

$$S_{t,j} \leq T_0 \times \frac{\max \eta_{v_s}}{n^{1/2}} \times (m + \max |N_R(v_s)|) \lesssim \frac{T_0 \log n(m + \log^3 n)}{2mn^{1/2}} = O\left(\frac{\omega \log^5 n}{n^{1/2}}\right).$$

$$W_j \geq (1 - o(1)) \left(\frac{j}{n}\right)^{1/2} = \Omega\left(\frac{\log^{10} n}{n^{1/2}}\right).$$

This completes this case. Now assume that $j \leq \log^{30} n$. **(P2)**, **(P3)**, **(P4)** and Lemma 3.6 that w.h.p.

$$S_{t,j} \lesssim 40m \log\log n \times \frac{\omega \log_{4/3} \log n}{2mn^{1/2}} \times (m + (\log\log n)^{20}) \ll W_j \sim \left(\frac{j}{n}\right)^{1/2}$$

for $\log^{1/100} n \leq j \leq \log^{30} n$. □

Dealing with left neighbors

The calculation of the ratio ρ_t takes contributions from two cases: where v_{t+1} is a left-neighbor of v_t, and where v_{t+1} is a right-neighbor of v_t.

Lemma 3.11.

$$\mathbf{E}(\rho_t \mathbf{1}_{v_{t+1}<v_t} \mid H_t) \leq \frac{2}{3}.$$

Proof. Let D denote the $(m/2)$th largest degree of a vertex in $N_R(v_t)$. We write

$$\begin{aligned}
\mathbf{E}(\rho_t \mathbf{1}_{v_{t+1}<v_t} \mid H_t) &= \sum_d \mathbf{E}(\rho_t \mathbf{1}_{v_{t+1}<v_t} \mid H_t \mid D = d) \, \mathbf{Pr}(D = d) \\
&\leq \sum_d \mathbf{E}\left(\frac{\zeta_d}{v_t}\right) \mathbf{Pr}(D = d) \\
&= \mathbf{E}\left(\frac{\zeta_D}{v_t}\right),
\end{aligned}$$

where ζ_d is the index of the smallest degree left neighbor of v_t that has degree at least d. We let $\zeta_d = 0$ if there are no such left neighbors. We now couple ζ with a random variable that is independent of the algorithm and can be used in its place.

Going back to Section 2.1 let us associate ℓ_k for $k \geq \omega$ with an index μ_k chosen uniformly from $[\lfloor k/m \rfloor]$. In this way, vertex $i \geq \omega$ is associated with m uniformly chosen vertices $a_{i,1}, a_{i,2}, \ldots, a_{i,m}$ in $[i-1]$. Furthermore, we can couple these choices so that if $N_L(i) = \{b_{i,1}, b_{i,2}, \ldots, b_{i,m}\}$ then given we have (i) $\mathbf{Pr}(b_{i,j} \leq a_{i,j}) \geq 1 - o(1)$ and (ii) $b_{i,j} \leq 2a_{i,j}$ for all i, j. This because $\mathbf{Pr}(b_{i,j} \leq k) \sim W_k/W_i \sim (k/i)^{1/2}$ (giving (i)) and $(k/i)^{1/2} \geq k/i$ and $(1 - o(1))(k/i)^{1/2} \geq k/2i$ (giving (ii)).

So now let μ be the index of the uniform choice associated with the largest degree left neighbor of v_t that has degree at least D. Thus

$$\mathbf{E}\left(\frac{\zeta_D}{v_t}\right) \lesssim \mathbf{E}\left(\frac{\mu}{v_t}\right) = \frac{1}{2} + o(1) \leq \frac{2}{3}.$$

$\square$

Dealing with right neighbors

It will be more difficult to consider the contribution of right-neighbors. In preparation, for $\lambda_0 \leq \gamma \leq 1 - 1/m$ we define

$$\Delta_\gamma^i := m + \gamma m \zeta(i)$$

where $\zeta(i), \zeta^+(i)$ are defined in (9), (10) respectively. We note that $\eta_i \zeta(i)$ is a lower bound for the expected degree of vertex $i, i \geq \omega$, see Lemma 2.2(a). Note also that $\eta_i \zeta^+(i)$ is an upper bound for the expected degree of vertex $i, i \geq \omega$.

The parameter Δ_γ^i is a degree threshold. For a suitable parameter γ, we wish it to be known to Analyzer that there should be many left-neighbors but few right-neighbors which have degree greater than Δ_γ^i. We define

$$\gamma_i^* = \max\left\{\gamma : \left|\left\{j \in N_L(i) : d_n(j) \geq \Delta_\gamma^i\right\}\right| \geq m/2\right\}.$$

$\Delta_{\gamma_{v_t}^*}^{v_t}$ is a lower bound on the degree needed for vertex $j > v_t$ to be considered by DCA as the next vertex; thus we proceed by analyzing the distribution of $\gamma_{v_t}^*$. We first derive upper bounds for $\mathbf{Pr}(\gamma_{v_t}^* \leq \gamma \mid H_t)$.

Lemma 3.12. *There exists $c_1 > 0$ such that*

$$\mathbf{Pr}(\gamma_{v_t}^* \leq \gamma \mid H_t) \lesssim (\gamma^{1/2} e^{1-\gamma^{1/2}})^{c_1 m^2} + m e^{-c_1 \gamma^{1/2} m},\ 0 \leq \gamma \leq \frac{1}{8}. \tag{33}$$

$$\mathbf{Pr}(\gamma_{v_t}^* \leq \gamma \mid H_t) \lesssim (\gamma^{1/2} e^{1-\gamma^{1/2}})^{c_1 m^2} + m e^{-c_1/\gamma^{1/2}},\ 0 \leq \gamma \leq \frac{1}{8}. \tag{34}$$

$$\mathbf{Pr}(\gamma_{v_t}^* \leq \tfrac{5}{4} \mid H_t) \lesssim e^{-c_1 m}. \tag{35}$$

$$\mathbf{Pr}(\gamma_{v_t}^* \geq \gamma \mid H_t) \lesssim \gamma^{-c_1 m}, \quad \gamma \geq 10^5. \tag{36}$$

Proof. For $j < v_t$, we define events $\mathcal{A}_j = \left\{\eta_j \leq \gamma^{1/2} m\right\}$ and $\mathcal{D}_j = \left\{d_n(j) \leq \Delta_\gamma^{v_t}\right\}$. We need to estimate $\mathbf{Pr}\left(\bigcap_{j \in S} \mathcal{D}_j\right)$ for subsets $S \subseteq N_L(v_t)$ of size $m/2$. We write

$$\bigcap_{j \in S} \mathcal{D}_j \subseteq \bigcap_{j \in S} (\mathcal{A}_j \cup (\bar{\mathcal{A}}_j \cap \mathcal{D}_j)) \subseteq \bigcap_{j \in S} \mathcal{A}_j \cup \bigcup_{j \in S} (\bar{\mathcal{A}}_j \cap \mathcal{D}_j). \tag{37}$$

Now, using inequality (6) and equation (23), we see that if $0 \leq \gamma \leq 1/8$ then for $j < v_t$,

$$\mathbf{Pr}(\eta_j \leq \gamma^{1/2} m \mid H_t) \lesssim m(\gamma^{1/2} e^{1-\gamma^{1/2}})^m. \tag{38}$$

The RHS of (38) includes a factor of $1+o(1)$ due to conditioning on $\mathcal{E}, H_t$.

So,

$$\mathbf{Pr}\left(\bigcap_{j\in S}\mathcal{A}_j \;\middle|\; H_t\right) \lesssim (m(\gamma^{1/2}e^{1-\gamma^{1/2}})^m)^{|S|}. \tag{39}$$

Furthermore, because $j \in N_L(v_t)$ implies that $i \geq j$ and hence $\zeta(v_t) \leq \zeta(j)$,

$$\begin{aligned}\mathbf{Pr}((d_n(j) \leq \Delta_\gamma^i) \wedge \bar{\mathcal{A}}_j \mid H_t) &\leq \mathbf{Pr}\left(d_n(j) - m \leq \gamma m\zeta(j) \mid H_t\right)\mathbf{Pr}(\eta_j > \gamma^{1/2}m|H_t)\\ &\lesssim \exp\left\{-\frac{(1-\gamma^{1/2})^2}{2}\gamma^{1/2}m\zeta(j)\right\}.\end{aligned} \tag{40}$$

Explanation of (40)**:** We remark first that the conditioning on $\mathcal{E}, H_t$ only adds a $(1+o(1))$ factor to the upper bound on our probability estimate. We now apply Lemma 2.2(b) with $1-\alpha = \gamma$ and $\eta_j \geq \gamma^{1/2}m$.

From (37) (summing over all $m/2$ subsets of $N_L(v_t)$) and (40) (summing over $N_L(v_t)$) we obtain

$$\begin{aligned}\mathbf{Pr}(\gamma_{v_t}^* \leq \gamma \mid H_t) = \mathbf{Pr}(|\left\{j \in N_L(v_t) : d_n(j) \leq \Delta_\gamma^i\right\}| \geq m/2| \mid H_t) \lesssim\\ 2^m\left((m(\gamma^{1/2}e^{1-\gamma^{1/2}})^m)\right)^{m/2} + m\exp\left\{-\frac{(1-\gamma^{1/2})^2}{2}\gamma^{1/2}m\zeta(v_t)\right\}.\end{aligned} \tag{41}$$

We observe that $j \in N_L(v_t)$ implies that $d_n(j) \geq m+1$. So,

$$m+1 < \Delta_\gamma^i \text{ implies } \zeta(v_t) > \frac{1}{m\gamma}. \tag{42}$$

Using (42) in (41) verifies (34), after bounding $2^m\left((m(\gamma^{1/2}e^{1-\gamma^{1/2}})^m)\right)^{m/2}$ by $(\gamma^{1/2}e^{1-\gamma^{1/2}})^{c_1m^2}$.

From (37) and (40),

$$\begin{aligned}\mathbf{Pr}(\gamma_{v_t}^* \leq \gamma \mid H_t) \lesssim (m(\gamma^{1/2}e^{1-\gamma^{1/2}})^m)^{m/2}+\\ +\mathbf{Pr}(\mathcal{A}_1 \mid H_t) + \mathbf{Pr}(\bar{\mathcal{A}}_1 \mid H_t)^{-1}m\exp\left\{-\frac{(1-\gamma^{1/2})^2}{2}\gamma^{1/2}m\zeta\left(\frac{9n}{10}\right)\right\}\end{aligned} \tag{43}$$

Here

$$\mathcal{A}_1 = \left\{\left|N_L(v_t) \cap \left[\frac{9n}{10}\right]\right| \geq \frac{m}{4}\right\}.$$

Explanation of (43)**:** The first term is from (39). If $\bar{\mathcal{A}}_1$ holds then v_t has at least one left neighbor $j \leq 9n/10$. The final term comes from using (40) and $\zeta(j) \geq \zeta(9n/10)$. The factor $\mathbf{Pr}(\bar{\mathcal{A}}_1)^{-1}$ handles the conditioning on $\bar{\mathcal{A}}_1$. The factor m is the union bound for choices of j.

Now $\left|N_L(v_t) \cap \left[\frac{9n}{10}\right]\right|$ is dominated by the binomial $Bin(m, n/10)$ and so $\mathbf{Pr}(\mathcal{A}_1) \leq e^{-d_1m}$. Now $\zeta(9n/10) \geq 1/20$ and plugging these facts into (40) yields (33). Here we have absorbed the e^{-d_1m} term into $me^{-c_1\gamma^{1/2}m}$ and we will do so again below.

We continue with the proof of (35). For $j \in N_L(v_t)$, we observe that if $d_n(j) \leq \Delta_\gamma^{v_t}$ and $\gamma \leq \frac{5}{4}$ then

$$d_n(j) - m \leq \frac{5m}{4}\zeta(v_t).$$

We now estimate the probability that a uniform random choice of $j \in N_L(v_t)$ (for fixed H_t, which determines v_t) has certain properties.

We first observe that

$$\mathbf{Pr}\left(j \geq \frac{3i}{5} \middle| H_t\right) \lesssim \left(1 - \frac{W_{3i/5}}{W_{v_t}}\right) \sim \left(1 - \left(\frac{3}{5}\right)^{1/2}\right) < \frac{2}{5}. \tag{44}$$

(For this we used **(P2)**.)

Now (6) implies that

$$\mathbf{Pr}(\eta_j \leq 0.99m \mid H_t) \leq e^{-d_2 m}. \tag{45}$$

Moreover, for $\eta_j > 0.99m$ and $j < 3v_t/5$, we have

$$\frac{\zeta(j)}{\zeta(v_t)} = \left(\frac{v_t}{j}\right)^{1/2} \left(\frac{1 - \left(\frac{j}{n}\right)^{1/2} - \varepsilon}{1 - \left(\frac{v_t}{n}\right)^{1/2} - \varepsilon}\right) \geq \left(\frac{v_t}{j}\right)^{1/2}$$

where

$$\varepsilon = \frac{5L \log\log n}{\omega^{3/4} \log n}. \tag{46}$$

Thus we have

$$\mathbf{E}(d_n(j) - m \mid H_t) \geq \eta_j \zeta(j) \gtrsim 0.99m \left(\frac{5}{3}\right)^{1/2} \zeta(v_t).$$

Now $0.99 \times (5/3)^{1/2} = 1.278.. > 1.01 \times 5/4$ and so

$$\mathbf{Pr}\left(d_n(j) - m \leq \frac{5\zeta(v_t)}{4} \;\middle|\; H_t\right) \leq \mathbf{Pr}\left(d_n(j) - m \leq \frac{\zeta(j)\eta_j}{1.01} \;\middle|\; H_t\right) \leq e^{-d_3 \eta_j \zeta(j)} \leq e^{-d_4 m} \tag{47}$$

using Lemma 2.2(b). It follows from (44) and (45) and (47) that

$$\mathbf{Pr}\left(\gamma_{v_t}^* < \frac{5}{4} \;\middle|\; H_t\right) \leq \mathbf{Pr}\left(Bin\left(m, e^{-d_2 m} + e^{-d_4 m} + \frac{2}{5}\right) \geq \frac{m}{2}\right) \leq e^{-d_3 m}.$$

This completes the proof of (35).

To deal with (36) we observe that if $d_n(j) \geq \Delta_\gamma^{v_t}$ and $\gamma \geq 10^5$ then

$$j \in N_L(v_t) \text{ and } j \leq \frac{v_t}{\gamma^{1/2}} \quad \text{or} \quad \eta_j \geq \gamma^{1/2} m \left(\frac{j}{v_t}\right)^{1/2} \geq \gamma^{1/4} m \quad \text{or} \quad d_n(j) - m \geq \gamma^{3/4} \eta_j \zeta(j).$$

But

$$\mathbf{Pr}\left(j \in N_L(v_t) \text{ and } j \leq \frac{v_t}{\gamma^{1/2}} \;\middle|\; H_t\right) \lesssim \frac{W_{v_t/\gamma^{1/2}}}{W_{v_t}} \lesssim \frac{1}{\gamma^{1/4}}. \tag{48}$$

And, using **(P3)** and $\gamma \geq 10^5$,

$$\begin{aligned}\mathbf{Pr}\left(\eta_j \geq \gamma^{1/4} m \;\middle|\; H_t\right) &\lesssim \sum_{l=2v_t/\gamma}^{v_t} \frac{w_l}{W_{v_t}} \int_{\eta_l=\gamma^{1/4}m}^{\infty} \frac{\eta_l^m e^{-\eta_l}}{(m-1)!} d\eta_l \\ &\lesssim \sum_{l=2v_t/\gamma}^{v_t} \frac{e^{-\gamma^{1/4}m}}{2(v_t l)^{1/2}} \\ &\lesssim e^{-\gamma^{1/4}m}.\end{aligned} \tag{49}$$

Lastly, using (44), (45) and Lemma 2.2(d) and $\zeta^+(j) \lesssim \zeta(j)$ for $j \leq 3n/5$ we have

$$\mathbf{Pr}\left(d_n(j) - m \geq \gamma^{3/4}\eta_j\zeta(j) \;\mid\; H_t\right) \leq \frac{2}{5} + e^{-d_2 m} + e^{-\gamma^{3/4}\eta} \leq 0.41. \tag{50}$$

It follows from (48), (49) and (50) that

$$\mathbf{Pr}\left(\gamma_{v_t}^* \geq \gamma \mid H_t\right) \lesssim \mathbf{Pr}\left(Bin\left(m, (1+o(1))\left(\frac{1}{\gamma^{1/4}} + e^{-\gamma^{1/4}m} + 0.41\right)\right) \geq \frac{m}{2}\right) \lesssim e^{-d_3 m}.$$

This completes the proof of the lemma. □

Corollary 3.13. *W.h.p.* $\gamma_{v_s}^* \geq 1/(\log\log n)^2$ *for* $s = 1, 2, \ldots, T = O(\log n)$.

Proof. The value of $\gamma_{v_s}^*$ is determined when v_s is first visited and in this case we can apply Lemma 3.12. In which case the result follows directly from (34). □

We now have a handle on the distribution of $\gamma_{v_t}^*$. We now put bounds on the expected number of $j > v_t$ that can be considered to be a candidate for v_{t+1}, conditioned on the value of $\gamma_{v_t}^*$. In particular, we let

$$D_\gamma^{v_t} = \left\{j > v_t : d_n(j) \geq \Delta_\gamma^{v_t}\right\}.$$

We will bound the size of D_γ^i by dividing D_γ^i into many parts bounding each part; in particular, $\kappa \in \mathbb{N}$ we let

$$J_\gamma^{i,\kappa} = \begin{cases} \left[i, \frac{i}{\gamma^2}\right] \cap D_\gamma^i & \kappa = 0. \\ \left[\frac{i}{\gamma^2}\left(1 + \frac{\kappa-1}{L}\right), \frac{i}{\gamma^2}\left(1 + \frac{\kappa}{L}\right)\right] \cap D_\gamma^i & 1 \leq \kappa \leq \frac{2n\gamma^2 L}{i}. \end{cases} \tag{51}$$

Note that $J_\gamma^{i,0} = \emptyset$ if $\gamma \geq 1$.

Finally, we let

$$r_\gamma^{i,\kappa} := |J_\gamma^{i,\kappa}| \quad \text{and} \quad r_\gamma^i := \sum_{\kappa \geq 0} r_\gamma^{i,\kappa} \qquad \text{and} \quad s_\gamma^i := \sum_{\kappa \geq 0} \sum_{j \in J_\gamma^{i,\kappa}} j \leq \frac{i}{\gamma^2} \sum_{\kappa \geq 0} \left(1 + \frac{\kappa+1}{L}\right) r_\gamma^{i,\kappa}. \tag{52}$$

Remark 3.14. *We have that* $\mathbf{E}\left(\frac{2}{m}\frac{1}{v_t}s_\gamma^{v_t} \mid H_t\right)$ *is an upper bound on the expectation of the ratio* $\rho_t = \frac{v_{t+1}}{v_t}$, *conditioned on the event that* $v_{t+1} > v_t$, *since each right neighbor whose index is included in the sum* $s_\gamma^{v_t}$ *has probability of at most* $\frac{2}{m}$ *of being chosen by the algorithm.*

Lemma 3.15. *If* $v_t \leq n\left(1 - \frac{3L^3}{m}\right)$ *then*

$$\mathbf{E}(r_\gamma^{v_t} \mid H_t) \leq \frac{\eta_{v_t}}{\gamma L}\left(L \max\{0, (1-\gamma)\} + 7 + 10Le^{-c_2\gamma L}\right). \tag{53}$$

$$\mathbf{E}\left(\frac{s_\gamma^{v_t}}{v_t} \mid H_t\right) \leq \frac{\eta_{v_t}}{\gamma^3 L}\left(L \max\{0, (1-\gamma)\} + 13 + 100Le^{-c_2\gamma L}\right). \tag{54}$$

Moreover,

$$\textit{If } v_t \leq n/5 \textit{ and } \kappa \geq (\log\log n)^4 \textit{ and } \gamma \geq 1/(\log\log n)^2 \textit{ then } \mathbf{Pr}(r_\gamma^{v_t,\kappa} > 0) \leq \frac{1}{\log^2 n}. \tag{55}$$

Note that (55) implies the second inequality in (19).

Proof. Recall from Lemma 3.10 that w.h.p.,

$$S_{T,j} = o(W_j) \text{ for } j \geq \log^{1/100} n. \tag{56}$$

We write

$$\begin{aligned}
&\mathbf{E}(r_\gamma^{v_t,\kappa} \mid H_t) \\
&\lesssim \sum_{j \in J_\gamma^{v_t,\kappa}} \frac{mw_{v_t}}{W_j - S_{t,j}} \int_{\eta_j=0}^{\infty} \frac{\eta_j^{m-1}e^{-\eta_j}}{(m-1)!} \mathbf{Pr}\left(d_n(j) \geq \Delta_\gamma^{v_t} \mid H_t, \eta_j\right) d\eta_j &(57)\\
&\lesssim \eta_{v_t} \sum_{j \in J_\gamma^{v_t,\kappa}} \frac{1}{2(v_t j)^{1/2}} \int_{\eta_j=0}^{\infty} \frac{\eta_j^{m-1}e^{-\eta_j}}{(m-1)!} \mathbf{Pr}\left(d_n(j) \geq \Delta_\gamma^{v_t} \mid H_t, \eta_j\right) d\eta_j. &(58)
\end{aligned}$$

Explanation of (57) **and** (58)**:** We sum over the relevant j and fix η_j. We multiply by the density of η_j and integrate. Using (56) we see that

$$\frac{mw_{v_t}}{W_j - S_{t,j}} \sim \frac{mw_{v_t}}{W_j} \sim \frac{\eta_{v_t}}{2(v_t j)^{1/2}}.$$

This is asymptotically equal to the expected number of times j chooses v_t as a neighbor.

Thus

$$\mathbf{E}(r_\gamma^{v_t,\kappa} \mid H_t) \lesssim \sum_{j \in J_\gamma^{v_t,\kappa}} \frac{\eta_{v_t}}{2(v_t j)^{1/2}} I_j \tag{59}$$

where

$$I_j = \int_{\eta_j=0}^{\infty} \frac{\eta_j^{m-1}e^{-\eta_j}}{(m-1)!} \mathbf{Pr}\left(d_n(j) \geq \Delta_\gamma^{v_t} \mid H_t, \eta_j\right) d\eta_j \leq 1.$$

If m is large then

$$\mathbf{E}\left(r_\gamma^{v_t,0} \mid H_t\right) \lesssim \begin{cases} 0 & \gamma \geq 1. \\ \sum_{j \in J_0^{v_t}(\gamma)} \frac{\eta_{v_t}}{2(v_t j)^{1/2}} \lesssim \eta_{v_t} \frac{1-\gamma}{\gamma} & \gamma < 1. \end{cases} \tag{60}$$

Continuing, for $\kappa \geq 1$, we write:

$$I_j \lesssim A_1 + A_2 + A_3, \tag{61}$$

where

$$A_1 = \int_{\eta_j=0}^{(1-1/L)\gamma m} \frac{\eta_j^{m-1} e^{-\eta_j}}{(m-1)!} \mathbf{Pr}\left(d_n(j) \geq \Delta_\gamma^{v_t} \mid \eta_j\right) d\eta_j.$$
$$A_2 = \int_{\eta_j=(1-1/L)m}^{(1+1/L)m} \frac{\eta_j^{m-1} e^{-\eta_j}}{(m-1)!} \mathbf{Pr}\left(d_n(j) \geq \Delta_\gamma^{v_t} \mid \eta_j\right) d\eta_j$$
$$A_3 = \int_{\eta_j=(1+1/L)m}^{\infty} \frac{\eta_j^{m-1} e^{-\eta_j}}{(m-1)!} \mathbf{Pr}\left(d_n(j) \geq \Delta_\gamma^{v_t} \mid \eta_j\right) d\eta_j.$$

and then write $r_\gamma^{v_t,\kappa} = r_\gamma^{v_t,\kappa,1} + r_\gamma^{v_t,\kappa,2} + r_\gamma^{v_t,\kappa,3}$. Here $r_\gamma^{v_t,\kappa,l}$ is equal to the RHS of (59) with I_j replaced by A_l. The implicit $(1+o(1))$ factor in (61) arises from replacing $\mathbf{Pr}\left(d_n(j) \geq \Delta_\gamma^{v_t} \mid \eta_j, H_t\right)$ by $\mathbf{Pr}\left(d_n(j) \geq \Delta_\gamma^{v_t} \mid \eta_j\right)$ in the integrals, i.e., ignoring the conditioning due to H_t. Since $j > v_t$, the only effect of H_t is on W_j through w_{v_t}. Here we have that w.h.p. $W_j \sim \left(\frac{v_t}{n}\right)^{1/2}$ and $w_{v_t} \sim \frac{\eta_{v_t}}{2m(v_t n)^{1/2}} = O\left(\frac{\log n}{2m(v_t n)^{1/2}}\right) = o(W_{v_t})$.

Case 1: $n_1 \leq v_t < n/5$:
Note that in this case

$$\zeta(v_t) \geq \frac{1}{2}\left(\frac{n}{v_t}\right)^{1/2} \geq 1. \tag{62}$$

In the following we use Lemma 2.1 to estimate the integrals over η_j. We observe that

$$\begin{aligned} &\mathbf{E}(r_\gamma^{v_t,\kappa,1} \mid H_t) \\ &\lesssim \sum_{j \in J_\gamma^{v_t,\kappa}} \frac{\eta_{v_t}}{2(v_t j)^{1/2}} \int_{\eta_j=0}^{\left(1-\frac{1}{L}\right)m} \frac{\eta_j^{m-1} e^{-\eta_j}}{(m-1)!} \mathbf{Pr}\left(d_n(j) \geq \Delta_\gamma^{v_t} \mid H_t, \eta_j\right) d\eta_j, \\ &\lesssim \sum_{j \in J_\gamma^{v_t,\kappa}} \frac{\eta_{v_t}}{2(v_t j)^{1/2}} \int_{\eta_j=0}^{\left(1-\frac{1}{L}\right)m} \frac{\eta_j^{m-1} e^{-\eta_j}}{(m-1)!} \cdot \left(\begin{cases} 1 & 1 \leq \kappa \leq 10L, \\ (e(L/\kappa)^{1/2})^{d_0 \gamma m \zeta(v_t)} & \kappa > 10L \end{cases}\right) d\eta_j, \end{aligned} \tag{63}$$

Explanation of (63)**:** We remark first that the conditioning on H_t only adds a $(1+o(1))$ factor to the upper bound on our probability estimate. We will use Lemma 2.2 to bound the probability that degrees are large. Now with our bound on v_t and within the range of integration, the ratio of $\Delta_\gamma^{v_t} - m$ to the mean of $d_n(j) - m$ is

$$\frac{\Delta_\gamma^{v_t} - m}{\zeta^+(j)} = \frac{\gamma m \left(\frac{n}{v_t}\right)^{1/2}\left(1 - \left(\frac{v_t}{n}\right)^{1/2} - o(1)\right)}{\eta_j \left(\frac{n}{j}\right)^{1/2}\left(1 - \left(\frac{j}{n}\right)^{1/2} + o(1)\right)} \gtrsim \frac{\gamma m}{\eta_j}\left(\frac{j}{v_t}\right)^{1/2} \geq$$

$$\frac{L}{L-1}\left(1+\frac{\kappa-1}{L}\right)^{1/2} \geq \left(\frac{\kappa}{L}\right)^{1/2} \text{ when } \kappa \geq 10L. \quad (64)$$

We then use (11) and Lemma 2.2(d) with $\beta = (\kappa/L)^{1/2}$.

Continuing, we observe that

$$\left(1+\frac{\kappa}{L}\right)^{1/2} - \left(1+\frac{\kappa-1}{L}\right)^{1/2} \leq \frac{1}{2L} \tag{65}$$

and so

$$\begin{aligned}
\mathbf{E}(r_\gamma^{v_t,\kappa,1} \mid H_t) &\leq \frac{\eta_{v_t} e^{-m/(2L^2)}}{\gamma {v_t}^{1/2}} \left(\left(\left(1+\frac{\kappa+1}{L}\right) v_t \right)^{1/2} - \left(\left(1+\frac{\kappa-1}{L}\right) v_t \right)^{1/2} \right) \times \\
&\qquad \left(\begin{cases} 1 & 1 \leq \kappa \leq 10L, \\ (e(L/\kappa)^{1/2})^{d_0\gamma m \zeta(v_t)} & \kappa > 10L, \end{cases} \right) \\
&\leq \frac{\eta_{v_t} e^{-m/(2L^2)}}{\gamma L} \cdot \left(\begin{cases} 1 & 1 \leq \kappa \leq 10L, \\ (e(L/\kappa)^{1/2})^{d_0\gamma m \zeta(v_t)} & \kappa > 10L. \end{cases} \right)
\end{aligned} \tag{66}$$

Continuing, it follows from (65) that

$$\sum_{j \in J_\gamma^{v_t,\kappa}} \frac{1}{j^{1/2}} \leq \frac{{v_t}^{1/2}}{\gamma L}. \tag{67}$$

$$\begin{aligned}
&\mathbf{E}(r_\gamma^{v_t,\kappa,2} \mid H_t) \\
&\leq \sum_{j \in J_\gamma^{v_t,\kappa}} \frac{\eta_{v_t}}{2(v_t j)^{1/2}} \int_{\eta_j=(1-1/L)m}^{(1+1/L)m} \frac{\eta_j^{m-1} e^{-\eta_j}}{(m-1)!} \mathbf{Pr}\left(d_n(j) \geq \Delta_\gamma^{v_t} \mid H_t, \eta_j\right) d\eta_j \\
&\leq \sum_{j \in J_\gamma^{v_t,\kappa}} \frac{\eta_{v_t}}{2(v_t j)^{1/2}} \times \begin{cases} 1 & \kappa \leq 3, \\ \exp\left\{-\frac{d_1\gamma m \zeta(v_t)}{L^2}\right\} & 4 \leq \kappa \leq 10L, \\ (e(L/\kappa)^{1/2})^{d_0\gamma m \zeta(v_t)} & \kappa > 10L, \end{cases} \qquad (68) \\
&\leq \frac{\eta_{v_t}}{\gamma L} \times \begin{cases} 1 & 1 \leq \kappa \leq 3, \\ e^{-d_1\gamma m \zeta(v_t)/L^2} & 4 \leq \kappa \leq 10L, \\ (e(L/\kappa)^{1/2})^{d_0\gamma m \zeta(v_t)} & \kappa > 10L. \end{cases} \qquad (69)
\end{aligned}$$

where we have used (67).

Explanation for (68)**:** We proceed in a similar manner to (64) and use

$$\frac{\Delta_\gamma^{v_t} - m}{\zeta^+(j)} = \frac{\gamma m \left(\frac{n}{v_t}\right)^{1/2} \left(1 - \left(\frac{v_t}{n}\right)^{1/2} - o(1)\right)}{\eta_j \left(\frac{n}{j}\right)^{1/2} \left(1 - \left(\frac{j}{n}\right)^{1/2} + \varepsilon\right)} \geq 1 + \frac{1}{L} \text{ if } \kappa \geq 4, \eta_j \geq \left(1 - \frac{1}{L}\right) m.$$

Then we use Lemma 2.2(c),(d).

Continuing,

$$\mathbf{E}(r_\gamma^{v_t,\kappa,3} \mid H_t) \le \sum_{j \in J_\gamma^{v_t,\kappa}} \frac{\eta_{v_t}}{2(v_t j)^{1/2}} \int_{\eta_j=(1+1/L)m}^{\infty} \frac{\eta_j^{m-1} e^{-\eta_j}}{(m-1)!} \mathbf{Pr}\left(d_n(j) \ge \Delta_\gamma^{v_t} \mid H_t, \eta_j\right) d\eta_j \tag{70}$$

We bound the integral in (70) by something independent of j and then as above, there is a factor $\eta_{v_t}/\gamma L$ arising from the sum over j.

For all $1 \le \kappa \le 80L+1$, we simply use the bound

$$\int_{\eta_j=m\left(1+\frac{1}{L}\right)}^{\infty} \frac{\eta_j^{m-1} e^{-\eta_j}}{(m-1)!} d\eta_j \le \exp\left\{-\frac{d_2 m}{L^2}\right\}. \tag{71}$$

For $\kappa \ge 80L+2$, we split the integral from (70) into pieces B_1^κ, B_2^κ (whose definition depends on κ), which we will bound individually.

In particular, we use

$$\begin{aligned} B_1^\kappa &= \int_{\eta_j=m\left(1+\frac{\kappa-1}{L}\right)^{1/4}}^{\infty} \frac{\eta_j^{m-1} e^{-\eta_j}}{(m-1)!} \mathbf{Pr}\left(d_n(j) \ge \Delta_\gamma^{v_t} \mid H_t, \eta_j\right) d\eta_j \\ &\le \int_{\eta_j=m\left(1+\frac{\kappa-1}{L}\right)^{1/4}}^{\infty} \frac{\eta_j^{m-1} e^{-\eta_j}}{(m-1)!} d\eta_j \\ &\le e^{-d_3 m(\kappa/L)^{1/4}} \end{aligned} \tag{72}$$

and

$$\begin{aligned} B_2^\kappa = &\int_{\eta_j=m\left(1+\frac{1}{L}\right)}^{m\left(1+\frac{\kappa-1}{L}\right)^{1/4}} \frac{\eta_j^{m-1} e^{-\eta_j}}{(m-1)!} \mathbf{Pr}\left(d_n(j) \ge \Delta_\gamma^{v_t} \mid H_t, \eta_j\right) d\eta_j \\ &\le \mathbf{Pr}\left(d_n(j) \ge \Delta_\gamma^{v_t} \,\middle|\, H_t, \eta_j \le m\left(1+\frac{\kappa-1}{L}\right)^{1/4}\right) \le \left(\frac{eL^{1/4}}{\kappa^{1/4}}\right)^{d_4 \gamma m \zeta(v_t)} \end{aligned} \tag{73}$$

to bound the integral in (70) by $B_1^\kappa + B_2^\kappa$ for all $\kappa \ge 80L+2$.

Therefore, gathering the many terms together (and using that $\kappa \le \frac{2nL\gamma^2}{v_t}$ from (51)) and relying on m large to allow crude upper bounding, we see that

$$\begin{aligned} \frac{\gamma L}{\eta_{v_t}} \mathbf{E}(r_\gamma^{v_t} \mid H_t) &\lesssim L \max\{0, (1-\gamma)\} \,[\text{from (60)}] + 10Le^{-m/(2L^2)} [\text{from (66)}] \\ &+ 10Le^{-d_1 \gamma m \zeta(v_t)/L^2} [\text{from (69)}] + \left(2 + e^{-m/(2L^2)}\right) \sum_{\kappa=10L}^{2nL\gamma^2/i} \left(\frac{eL^{1/2}}{\kappa^{1/2}}\right)^{d_0 \gamma m \zeta(v_t)} [\text{from (66) and (69)}] \\ &+ 6[\text{from (69)}] + 100L \exp\left\{-\frac{d_2 m}{L^2}\right\} [\text{from (71)}] \end{aligned}$$

$$+\sum_{\kappa=80L+2}^{2nL\gamma^2/i}\left(e^{-d_3m(\kappa/L)^{1/4}}+\left(\frac{eL^{1/4}}{\kappa^{1/4}}\right)^{d_4\gamma m\zeta(v_t)}\right)\text{[from (72) and (73)].} \quad (74)$$

We first observe that if $\frac{n}{v_t} < \frac{10}{\gamma^2}$ then the summations $\kappa = 10L, \ldots, 2nL\gamma^2/v_t$ etc. above are empty. For larger n/v_t we can therefore assume that $\gamma m(n/v_t)^{1/2} \geq m$ which implies (see (62)) that $\gamma m\zeta(v_t) \geq m/2$ and then we can assume that

$$\sum_{\kappa=10L}^{2nL\gamma^2/v_t}\left(\frac{eL^{1/2}}{\kappa^{1/2}}\right)^{d_0\gamma m\zeta(v_t)} \leq \frac{1}{1000} \text{ and } \sum_{\kappa=80L+1}^{2nL\gamma^2/v_t}\left(\frac{eL^{1/4}}{\kappa^{1/4}}\right)^{d_4\gamma m\zeta(v_t)} \leq \frac{1}{1000}. \quad (75)$$

Plugging these estimates into (74) and making some simplifications, we obtain (53).

Going back to (52) we have

$$\frac{\gamma^3 L}{\eta_{v_t}}\mathbf{E}\left(\frac{2s_\gamma^{v_t}}{mi}\;\middle|\; H_t\right) \leq L\max\{0,(1-\gamma)\}$$
$$+200Le^{-m/(2L^2)}+100Le^{-d_1\gamma m\zeta(v_t)/L^2}+\left(2+e^{-m/(2L^2)}\right)\sum_{\kappa=10L}^{2nL\gamma^2/v_t}\frac{2\kappa}{L}\left(\frac{eL^{1/2}}{\kappa^{1/2}}\right)^{d_0\gamma m\zeta(v_t)}$$
$$+12+10^4L\exp\left\{-\frac{d_2m}{L^2}\right\}+\sum_{\kappa=80L+2}^{2nL\gamma^2/v_t}\frac{2\kappa}{L}\left(e^{-d_3m(\kappa/L)^{1/4}}+\left(\frac{eL^{1/4}}{\kappa^{1/4}}\right)^{d_4\gamma m\zeta(v_t)}\right).$$

Making similar estimates to what we did for (75) gives us (54).

We obtain (55) from **(P5)**, (66), (69), (72) and (73). Indeed, if $J_\gamma^{v_t,\kappa} \neq \emptyset$ then from its definition we must have $v_t \leq \frac{2L\gamma^2 n}{\kappa-1}$. Together with $v_t \leq n/5$ we obtain that $\zeta(v_t) \geq \frac{\kappa^{1/2}}{2L^{1/2}\gamma}$. Thus, in this case,

$$\left(\frac{eL^{1/2}}{\kappa^{1/2}}\right)^{d_0\gamma m\zeta(v_t)} \leq \left(\frac{eL^{1/2}}{(\log\log n)^2}\right)^{d_0m(\log\log n)^2/2L^{1/2}} = o\left(\frac{1}{\log^{10} n}\right). \quad (76)$$

This deals with the probabilities in (66) and (69). For (69) we rely m large to to show that $e^{-d_3m(\kappa/L)^{1/4}} = o(1/\log^{10} n)$. Equation (72) is dealt with in a similar manner to (66). Here we have $\left(\frac{eL^{1/4}}{\kappa^{1/4}}\right)^{d_4\gamma m\zeta(v_t)}$ which is the square root of (76).

Case 2: $n/5 \leq v_t \leq n\left(1-\frac{3L^3}{m}\right)$:
The upper bound on v_t implies that

$$m\zeta(v_t) \geq L^3.$$

Using the same definitions of $r_\gamma^{v_t,\kappa,l}, l = 1,2,3$ as above:

$$\sum_{\kappa\geq 1}\mathbf{E}(r_\gamma^{v_t,\kappa,1} \mid H_t) \leq \sum_{\kappa\geq 1}\sum_{j\in J_\gamma^{v_t,\kappa}}\frac{\eta_{v_t}}{2(v_tj)^{1/2}}\int_{\eta_j=0}^{\left(1-\frac{1}{L}\right)m}\frac{\eta_j^{m-1}e^{-\eta_j}}{(m-1)!}d\eta_j,$$

$$
\begin{aligned}
&\le \sum_{\kappa\ge 1}\sum_{j\in J_\gamma^{v_t,\kappa}} \frac{\eta_{v_t}}{2(v_t j)^{1/2}} e^{-m/(2L^2)}, \qquad \text{from Lemma 2.1(e),}\\
&\le \frac{\eta_{v_t}}{\gamma}\left(\frac{n}{v_t}\right)^{1/2} e^{-m/(2L^2)}\\
&\le \frac{5^{1/2}\eta_{v_t}}{\gamma} e^{-m/(2L^2)}.
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{\kappa\ge 1} \mathbf{E}(r_\gamma^{v_t,\kappa,2} \mid H_t)\\
&\lesssim \sum_{\kappa\ge 1}\sum_{j\in J_\gamma^{v_t,\kappa}} \frac{\eta_{v_t}}{2(v_t j)^{1/2}} \int_{\eta_j=(1-1/L)m}^{(1+1/L)m} \frac{\eta_j^{m-1}e^{-\eta_j}}{(m-1)!}\, \mathbf{Pr}\left(d_n(j)\ge \Delta_\gamma^{v_t} \mid \eta_j\right) d\eta_j\\
&\lesssim \sum_{\kappa\ge 1}\sum_{j\in J_\gamma^{v_t,\kappa}} \frac{\eta_{v_t}}{2(v_t j)^{1/2}} \int_{\eta_j=(1-1/L)m}^{(1+1/L)m} \frac{\eta_j^{m-1}e^{-\eta_j}}{(m-1)!} \cdot \begin{cases} 1 & \kappa\le 3,\\ \exp\left\{-\frac{d_5\gamma m\zeta(v_t)}{L^2}\right\} & 4\le\kappa\le 10L,\\ (e(L/\kappa)^{1/2})^{d_6\gamma m\zeta(v_t)} & \kappa>10L, \end{cases}\\
&\le \sum_{\kappa\ge 1}\frac{\eta_{v_t}}{\gamma L}\begin{cases} 2 & \kappa\le 3,\\ 4e^{-d_5\gamma L} & 4\le\kappa\le 10L,\\ \left(\frac{eL^{1/2}}{\kappa^{1/2}}\right)^{d_6\gamma L^3} & \kappa>10L, \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
\sum_{\kappa\ge 1} \mathbf{E}(r_\gamma^{v_t,\kappa,3} \mid H_t) &\le \sum_{\kappa\ge 1}\sum_{j\in J_\gamma^{v_t,\kappa}} \frac{\eta_{v_t}}{2(v_t j)^{1/2}} \int_{\eta_j=(1+1/L)m}^{\infty} \frac{\eta_j^{m-1}e^{-\eta_j}}{(m-1)!}\\
&\le \sum_{\kappa\ge 1}\sum_{j\in J_\gamma^{v_t,\kappa}} \frac{\eta_{v_t}}{2(v_t j)^{1/2}} e^{-m/(3L^2)}, \qquad \text{from Lemma 2.1(d),}\\
&\le \frac{\eta_{v_t}}{\gamma} e^{-m/(3L^2)} \left(\frac{n}{v_t}\right)^{1/2}\\
&\le \frac{5^{1/2}\eta_{v_t}}{\gamma} e^{-m/(3L^2)}.
\end{aligned}
$$

The above upper bounds are small enough to give the lemma in this case, without trouble. □

We are now in a position to prove (20). We confirmed the second part of the statement (20) above, using (55), so only the first part remains. The first part follows immediately from Lemma 3.11 and the following, by addition:

Lemma 3.16.

$$
\mathbf{E}(\rho_t \mathbf{1}_{v_{t+1}\ge v_t} \mid H_t) \le \frac{21\eta_{v_t}}{mL} + \frac{L^3}{m^2}.
$$

Proof. We consider cases.
Case 1: $n_1 \le v_t \le n\left(1-\frac{3L^3}{m}\right)$**:** Then,

$$\mathbf{E}(\rho_t \mathbf{1}_{v_{t+1}\ge v_t} \mid H_t) \le I_1 + I_2 + I_3 + I_4,$$

where

$$\begin{aligned}
I_1 &= \int_{\gamma=1/8}^{5/4} \mathbf{E}(\rho_t \mathbf{1}_{v_{t+1}\ge v_t} \mid \gamma^*_{v_t} = \gamma) d\,\mathbf{Pr}(\gamma^*_{v_t} \le \gamma), \\
&\le \int_{\gamma=1/8}^{5/4} \left(\frac{2\times 8^3}{mL} \times \eta_{v_t} \times \left(\frac{7L}{8} + 13 + 100Le^{-c_2\gamma L}\right)\right) d\,\mathbf{Pr}(\gamma^*_{v_t} \le \gamma), \\
&\qquad\qquad\qquad\qquad\qquad\qquad \text{by Remark 3.14, Lemma 3.15,} \\
&\le \frac{1000\eta_{v_t}}{m} \int_{\gamma=1/8}^{5/4} d\,\mathbf{Pr}(\gamma^*_{v_t} \le \gamma), \\
&\le \eta_{v_t} e^{-c_1 m}, \qquad \text{from (35).} && (77) \\
I_2 &= \int_{\gamma=5/4}^{10000} \left(\frac{2}{m\gamma^3 L} \times \eta_{v_t} \times (13 + 100Le^{-c_2\gamma L})\right) d\,\mathbf{Pr}(\gamma^*_{v_t} \le \gamma), \\
&\le \frac{20\eta_{v_t}}{mL}. && (78) \\
I_3 &= \int_{\gamma=10000}^{\infty} \left(\frac{2}{m\gamma^3 L} \times \eta_{v_t} \times (13 + 100Le^{-c_2\gamma L})\right) d\,\mathbf{Pr}(\gamma^*_{v_t} \le \gamma) \\
&\le \frac{27\eta_{v_t}}{10^{15}Lm} \int_{\gamma=100}^{\infty} \gamma^{-cm} d\gamma, \qquad \text{from (36),} \\
&= \frac{27\eta_{v_t}}{10^{15}Lm} \times \frac{1}{10^{2(cm-1)}(cm-1)}. && (79) \\
I_4 &= \int_{\gamma=0}^{1/8} \mathbf{E}(\rho_t \mathbf{1}_{v_{t+1}\ge v_t} \mid \gamma^*_{v_t} = \gamma) d\,\mathbf{Pr}(\gamma^*_{v_t} \le \gamma) \\
&\le e^{-d_0 m^{1/2}}, && (80)
\end{aligned}$$

To obtain the term $e^{-d_0 L}$ in (80) we use (33) and (34) to obtain

$$\begin{aligned}
&\max\left\{\gamma \in [0,1/8] : \gamma^{-2}\,\mathbf{Pr}(\gamma^* \le \gamma)\right\} \le \\
&\max\left\{\gamma \in [0,1/8] : (\gamma^{1/2-4/(c_1m^2)} e^{1-\gamma^{1/2}})^{c_1m^2}\right\} + \\
&\qquad\qquad\qquad\qquad \max\left\{\gamma \in [0,1/8] : m\gamma^{-2}\min\left\{e^{-c_1m\gamma^{1/2}}, e^{-c_1/\gamma^{1/2}}\right\}\right\} \\
&\le (8^{4/(c_1m^2)-1/2} e^{1-1/8^{1/2}})^{c_1m^2} + m^2 e^{-c_1 m^{1/2}}.
\end{aligned}$$

The first case of the lemma now follows from (77), (78), (79) and (80).

Case 2: $v_t > n\left(1-\frac{3L^3}{m}\right)$**:**
We observe first that $n \le v_t\left(1+\frac{4L^3}{m}\right)$. Then we let $Z = d_n(v_t) - m$ be the number of right

neighbors of v_t. Furthermore,

$$\mathbf{E}(Z \mid H_t) \lesssim \sum_{j=v_t+1}^{n} \frac{w_i}{W_j} \lesssim \sum_{j=v_t+1}^{v_t\left(1+\frac{4L^3}{m}\right)} \frac{\eta_{v_t}}{2(v_t j)^{1/2}} \lesssim \eta_{v_t} \left(\left(1 + \frac{4L^3}{m}\right)^{1/2} - 1 \right) \leq \eta_{v_t} \frac{4L^3}{m}. \tag{81}$$

Case 2a: $\eta_{v_t} \geq 1/L^{1/2}$**.**
We use (81) and Lemma 2.2(d) to prove

$$\mathbf{Pr}\left(Z \geq \frac{\eta_{v_t}}{L} \;\middle|\; H_t \right) \leq e^{-d_1 \eta_{v_t} L} \leq e^{-d_1 L^{1/2}}. \tag{82}$$

Then we can write

$$\mathbf{E}(\rho_t \mathbf{1}_{v_{t+1} \geq v_t} \mid H_t) \leq \left(1 + \frac{4L^3}{m}\right) \times e^{-d_1 L^{1/2}} + \frac{2\eta_{v_t}}{Lm} \leq \frac{3\eta_{v_t}}{Lm}.$$

Explanation: ρ_t will be at most $\left(1 + \frac{4L^3}{m}\right)$ if the unlikely event in (82) occurs. Failing this, the chance that $\rho_t > 1$ is at most $\frac{2Z}{m} \leq \frac{2\eta_{v_t}}{Lm}$.

Case 2b: $\eta_{v_t} < 1/L^{1/2}$**.**
It follows from (81) that $\mathbf{E}(Z \mid H_t) \lesssim 4L^{5/2}/m$. It then follows from Lemma 2.2(d) that

$$\mathbf{Pr}\left(Z \geq \frac{L^3}{3m} \middle| H_t \right) \leq e^{-d_2 L^{1/2}}.$$

We then have

$$E(\rho_t \mathbf{1}_{v_{t+1} \geq v_t} \mid H_t) \leq \left(1 + \frac{4L^3}{m}\right) \times e^{-d_2 L^{1/2}} + \frac{2L^3}{3m^2} \leq \frac{L^3}{3m^2}. \quad \square$$

Part 5

We now prove (21). To do this, we will obtain a recurrence for $\mathbf{E}(\eta_{v_{t+1}} \mid H_t)$, and, at the end, obtain the bound $4m$ by averaging over the possible histories H_t.

We begin by writing

$$\mathbf{E}(\eta_{v_{t+1}} \mid H_t) = \mathbf{E}(\eta_{v_{t+1}} \mathbf{1}_{v_{t+1} < v_t} \mid H_t) + \mathbf{E}(\eta_{v_{t+1}} \mathbf{1}_{v_{t+1} > v_t} \mid H_t) \tag{83}$$

We consider each term in (83) separately. For the **first term**, since

$$\eta_{v_{t+1}} \mathbf{1}_{v_{t+1} < v_t} \leq \max\left\{\eta_l :\ 1 \leq l < v_t,\, l \in N_L(v_t)\right\} \mathbf{1}_{v_{t+1} < v_t} \leq \max\left\{\eta_l :\ 1 \leq l < v_t,\, l \in N_L(v_t)\right\},$$

we have that

$$\mathbf{E}(\eta_{v_{t+1}} \mathbf{1}_{v_{t+1} < v_t} \mid H_t) \leq \mathbf{E}\left(\max\left\{\eta_l :\ 1 \leq l < v_t,\, l \in N_L(v_t)\right\} \mid H_t\right)$$

$$
\begin{aligned}
&= \int_{\eta=0}^{\infty} \mathbf{Pr}(\max\{\eta_l :\ 1 \le l < v_t,\, l \in N_L(v_t)\} \ge \eta \mid H_t) d\eta \\
&= \int_{\eta=0}^{\infty} \mathbf{Pr}(\exists 1 \le l < v_t,\, l \in N_L(v_t) :\ \eta_l \ge \eta \mid H_t) d\eta \\
&\le \int_{\eta=0}^{\infty} \sum_{l=1}^{v_t-1} \mathbf{Pr}(l \in N_L(v_t) \text{ and } \eta_l \ge \eta \mid H_t) d\eta \\
&\lesssim \int_{\eta=0}^{\infty} \sum_{l=1}^{v_t-1} \frac{w_l}{W_{v_t}} \int_{\eta_l=\eta}^{\infty} \frac{\eta_l^{m-1} e^{-\eta_l}}{(m-1)!} d\eta_l d\eta \\
&\lesssim \sum_{l=1}^{v_t-1} \int_{\eta=0}^{\infty} \int_{\eta_l=\eta}^{\infty} \frac{\eta_l}{2m(lv_t)^{1/2}} \frac{\eta_l^{m-1} e^{-\eta_l}}{(m-1)!} d\eta_l d\eta \\
&\lesssim 2m + (1+o(1)) \int_{\eta=2m}^{\infty} \int_{\eta_l=\eta}^{\infty} \frac{\eta_l^{m} e^{-\eta_l}}{(m-1)!} d\eta_l d\eta \\
&\lesssim 2m + \int_{\eta=2m}^{\infty} 4e^{-3\eta/10} d\eta, \text{ from Lemma 2.1(c),} \\
&\lesssim 2m + 20e^{-3m/5} \\
&\le 3m. \qquad (84)
\end{aligned}
$$

We now bound the **second term** of (83). We consider two cases, according to properties of the history H_t (which determines v_t and η_{v_t}).

Case 1: H_t is such that $v_t \le \left(1 - \frac{1}{\omega^{1/2}}\right) n$.
In this case, we have that

$$
\begin{aligned}
&\mathbf{E}(\eta_{v_{t+1}} \mathbf{1}_{v_{t+1} > v_t} \mid H_t) \\
&\quad \le \mathbf{E}\left(\max\left\{\eta_l :\ v_t < l \le n,\, v_{t+1} \in N_L(l), d_n(l) \ge \Delta_{\gamma^*_{v_t}}^{v_t}\right\} \mathbf{1}_{v_{t+1} > v_t} \mid H_t\right) \\
&\quad\quad \le \mathbf{E}\left(\max\left\{\eta_l :\ v_t < l \le n,\, v_{t+1} \in N_L(l), d_n(l) \ge \Delta_{\gamma^*_{v_t}}^{v_t}\right\} \mid H_t\right).
\end{aligned}
$$

So we have that

$$
\begin{aligned}
&\mathbf{E}(\eta_{v_{t+1}} \mathbf{1}_{v_{t+1} > v_t} \mid H_t) \\
&\le \mathbf{E}\left(\max\left\{\eta_l :\ v_t < l \le n,\, v_{t+1} \in N_L(l), d_n(l) \ge \Delta_{\gamma^*_{v_t}}^{v_t}\right\} \mid H_t\right) \\
&= \int_{\eta=0}^{\infty} \mathbf{Pr}(\max\left\{\eta_l :\ v_t < l \le n,\, v_t \in N_L(l), d_n(l) \ge \Delta_{\gamma^*_{v_t}}^{v_t}\right\} \ge \eta \mid H_t) d\eta \\
&= \int_{\eta=0}^{\infty} \mathbf{Pr}\left(\exists v_t < l \le n,\, v_t \in N_L(l) :\ \eta_l \ge \eta, d_n(l) \ge \Delta_{\gamma^*_{v_t}}^{v_t} \mid H_t\right) d\eta \\
&\le \sum_{l=v_t+1}^{n} \int_{\eta=0}^{\infty} \mathbf{Pr}\left((v_t \in N_L(l)) \wedge (\eta_l \ge \eta) \wedge (d_n(l) \ge \Delta_{\gamma^*_{v_t}}^{v_t}) \mid H_t, \eta_l\right) d\eta \\
&\lesssim \sum_{l=v_t+1}^{n} \frac{\eta_{v_t}}{2(lv_t)^{1/2}} \int_{\eta=0}^{\infty} \int_{\eta_l=\eta}^{\infty} \frac{\eta_l^{m-1} e^{-\eta_l}}{(m-1)!} \mathbf{Pr}\left(d_n(l) \ge \Delta_{\gamma^*_{v_t}}^{v_t} \mid H_t, \eta_l\right) d\eta_l d\eta. \qquad (85)
\end{aligned}
$$

Recall that in the final two lines, v_t and η_{v_t} are not random variables, but are the actual values of these random variables in the history H_t, so this is a deterministic upper bound on $\mathbf{E}(\eta_{v_{t+1}}\mathbf{1}_{v_{t+1}>v_t} \mid H_t)$.

We split the sum in the RHS of (85) into $E_1+E_2+E_3+E_4$ according to the ranges of l and η, and bound each separately. The first part consists of

$$E_1 = \sum_{l=v_t+1}^{n} \frac{\eta_{v_t}\mathbf{1}_{l\le 4m^2v_t/(\gamma^*_{v_t})^2}}{2(lv_t)^{1/2}} \int_{\eta=0}^{2m}\int_{\eta_l=\eta}^{\infty} \frac{\eta_l^{m-1}e^{-\eta_l}}{(m-1)!}\,\mathbf{Pr}\left(d_n(l)\ge \Delta^{v_t}_{\gamma^*_{v_t}} \mid H_t,\eta_l\right) d\eta_l d\eta.$$

Even though v_t and η_{v_t} are constants (determined by H_t), we caution that $\gamma^*_{v_t}$ and so also E_1 are random variables.

Observe that we have that

$$\int_{\eta_l=\eta}^{\infty} \frac{\eta_l^{m-1}e^{-\eta_l}}{(m-1)!}\,\mathbf{Pr}\left(d_n(l)\ge \Delta^{v_t}_{\gamma^*_{v_t}} \mid H_t,\eta_l\right) d\eta_l \le \int_{\eta_l=\eta}^{\infty} \frac{\eta_l^{m-1}e^{-\eta_l}}{(m-1)!} d\eta_l \le 1$$

which allows us to write

$$E_1\mathbf{1}_{\gamma^*_{v_t}\le 5/4} \le \mathbf{1}_{\gamma^*_{v_t}\le 5/4} \sum_{l=v_t+1}^{n} \frac{2m\eta_{v_t}\mathbf{1}_{l\le 4m^2v_t/(\gamma^*_{v_t})^2}}{2(lv_t)^{1/2}} \le \frac{5m^2\eta_{v_t}}{\gamma^*_{v_t}}\mathbf{1}_{\gamma^*_{v_t}\le 5/4}. \tag{86}$$

We will use this expression when we take the expectation over $\gamma^*_{v_t}\le 5/4$.

We also have that

$$\int_{\eta_l=\eta}^{\infty} \frac{\eta_l^{m-1}e^{-\eta_l}}{(m-1)!}\,\mathbf{Pr}\left((d_n(l)\ge \Delta^{v_t}_{\gamma^*}) \wedge (\gamma^*_{v_t} > 5/4) \mid H_t,\eta_l\right) d\eta_l \le I_1+I_2+I_3, \tag{87}$$

where

$$\begin{aligned}
I_1 &= \int_{\eta_l=\eta}^{7m/8} \frac{\eta_l^{m-1}e^{-\eta_l}}{(m-1)!} d\eta_l \le m\left(\frac{7}{8}e^{1/8}\right)^m \le e^{-d_0 m}, \quad \text{from Lemma 2.1(a)},\\
I_2 &= \int_{\eta_l=9m/8}^{\infty} \frac{\eta_l^{m-1}e^{-\eta_l}}{(m-1)!} d\eta_l \le e^{-d_1 m}, \quad \text{from Lemma 2.1(d)}.\\
I_3 &= \int_{\eta_l=7m/8}^{9m/8} \frac{\eta_l^{m-1}e^{-\eta_l}}{(m-1)!}\,\mathbf{Pr}\left(\left(d_n(l)-m\ge \gamma^*_{v_t}m\zeta(v_t)\right)\wedge(\gamma^*_{v_t}>5/4)\middle| H_t,\eta_l\right) d\eta_l,\\
&\le \int_{\eta_l=7m/8}^{9m/8} \frac{\eta_l^{m-1}e^{-\eta_l}}{(m-1)!}\,\mathbf{Pr}\left(d_n(l)-m\ge \frac{5}{4}m\zeta(v_t)\middle| H_t,\eta_l\right) d\eta_l.
\end{aligned} \tag{88}$$

We bound I_3 with two subcases:
Subcase 1a: $\zeta(l)>0$.

$$I_3 \le \int_{\eta_l=7m/8}^{9m/8} \frac{\eta_l^{m-1}e^{-\eta_l}}{(m-1)!}\,\mathbf{Pr}\left(d_n(l)-m\ge \frac{10\zeta(v_t)}{9\zeta(l)}\eta_l\zeta(l)\middle| H_t,\eta_l\right) d\eta_l \ \ [\text{since } m\ge \frac{8}{9}\eta_l],$$

$$\leq \int_{\eta_l=7m/8}^{9m/8} \frac{\eta_l^{m-1}e^{-\eta_l}}{(m-1)!} \begin{cases} \exp\left\{-\frac{(10\zeta(v_t)-9\zeta(l))^2\eta_l}{81\zeta(l)}\right\} & 10\zeta(v_t) \leq 18\zeta(l) \\ \exp\left\{-\frac{(10\zeta(v_t)-9\zeta(l))\eta_l}{27}\right\} & 10\zeta(v_t) > 18\zeta(l) \end{cases} \quad \text{[from (2) and } \ell \geq v_t+1\text{]},$$

$$\leq \int_{\eta_l=7m/8}^{9m/8} \frac{\eta_l^{m-1}e^{-\eta_l}}{(m-1)!} \begin{cases} \exp\left\{-\frac{7m\zeta(v_t)}{648}\right\} & 10\zeta(v_t) \leq 18\zeta(l) \\ \exp\left\{-\frac{7m\zeta(v_t)}{216}\right\} & 10\zeta(v_t) > 18\zeta(l) \end{cases}$$

$$\leq e^{-m\zeta(v_t)/100}. \tag{89}$$

Subcase 1b: $\zeta(l) \leq 0$.

In this case, we go back to (88) and use $\zeta^+(l)$ in place of $\zeta(l)$, see (10).

$$I_3 \leq \int_{\eta_l=7m/8}^{9m/8} \frac{\eta_l^{m-1}e^{-\eta_l}}{(m-1)!} \mathbf{Pr}\left(d_n(l) - m \geq \frac{10\zeta(v_t)}{9\zeta^+(l)}\eta_l\zeta^+(l) \middle| H_t, \eta_l\right) d\eta_l, \tag{90}$$

For ε as in (46) we see that $\zeta(l) \leq 0$ implies that $l \geq n(1-\varepsilon)^2$. In which case

$$\zeta^+(l) \leq \frac{2\varepsilon}{1-\varepsilon} \leq 3\varepsilon. \tag{91}$$

On the other hand, $v_t \leq 1 - \frac{1}{\omega^{1/2}}$ implies that

$$\zeta(v_t) \geq \frac{1}{\omega^{1/2}} - 2\varepsilon \geq \frac{1}{2\omega^{1/2}}. \tag{92}$$

Comparing (91) and (92), we see that $\zeta(v_t) \gg \zeta^+(l)$. From this and (3) with $\beta = \frac{10\zeta(v_t)}{9\zeta^+(l)} \geq \frac{1}{6\varepsilon\omega^{1/2}}$ we deduce that

$$\mathbf{Pr}\left(d_n(l) - m \geq \frac{10\zeta(v_t)}{9\zeta^+(l)}\eta_l\zeta^+(l) \middle| H_t, \eta_l\right) \leq (6\varepsilon\omega^{1/2})^{10\eta_l\zeta(v_t)} \leq (6\varepsilon\omega^{1/2})^{35m\zeta(v_t)/36}.$$

Plugging this estimate into (90) we obtain something stronger than (89), finishing Subcase 1b and giving that $I_3 \leq e^{-m\zeta(v_t)/100}$ in all cases.

Having bounded the three terms in (87), we then have that

$$\begin{aligned} E_1\mathbf{1}_{\gamma^*_{v_t}>5/4} &\leq \sum_{l=v_t+1}^{n} \frac{\eta_{v_t}\mathbf{1}_{l\leq 4m^2v_t/(\gamma^*_{v_t})^2}}{(lv_t)^{1/2}}\left(e^{-d_2m} + e^{-m\zeta(v_t)/100}\right) \\ &\leq \eta_{v_t}\left(e^{-d_2m}\frac{5m}{\gamma^*_{v_t}} + \frac{e^{-m\zeta(v_t)/100}}{v_t^{1/2}}\sum_{l=v_t+1}^{n}\frac{1}{l^{1/2}}\right) \\ &\leq \eta_{v_t}\left(4me^{-d_2m} + e^{-m\zeta(v_t)/100}\cdot\frac{(n+1)^{1/2}-(v_t+1)^{1/2}}{v_t^{1/2}}\right) \\ &\leq \eta_{v_t}(4me^{-d_2m} + 2\zeta(v_t)e^{-m\zeta(v_t)/100}) \\ &\leq \eta_{v_t}\left(4me^{-d_2m} + \frac{200}{m}\right). \end{aligned} \tag{93}$$

It follows from (86) and (93) that

$$E_1 \le \eta_{v_t}\left(\frac{5m^2}{\gamma^*_{v_t}}\mathbf{1}_{\gamma^*_{v_t}\le 5/4} + \left(4me^{-d_2 m} + \frac{200}{m}\right)\right).$$

We continue with the other parts of the RHS of (85):

$$\begin{aligned}
E_2 &= \sum_{l=v_t+1}^{n} \frac{\eta_{v_t}\mathbf{1}_{l\le 4m^2 v_t/(\gamma^*_{v_t})^2}}{2(lv_t)^{1/2}} \int_{\eta=2m}^{\infty}\int_{\eta_l=\eta}^{\infty} \frac{\eta_l^{m-1}e^{-\eta_l}}{(m-1)!}\mathbf{Pr}(d_n(l)\ge \Delta^{v_t}_{\gamma^*} \mid H_t,\eta_l)d\eta_l d\eta \\
&\le \sum_{l=v_t+1}^{4m^2 v_t/(\gamma^*_{v_t})^2} \frac{\eta_{v_t}}{2(lv_t)^{1/2}} \int_{\eta=2m}^{\infty}\int_{\eta_l=\eta}^{\infty} \frac{\eta_l^{m-1}e^{-\eta_l}}{(m-1)!}d\eta_l d\eta \\
&\le \sum_{l=v_t+1}^{4m^2 v_t/(\gamma^*_{v_t})^2} \frac{\eta_{v_t}}{2(lv_t)^{1/2}} \int_{\eta=2m}^{\infty}\left(\frac{e\eta/m}{e^{\eta/m}}\right)^m d\eta \quad \text{from Lemma 2.1(c),} \\
&\le \sum_{l=v_t+1}^{4m^2 v_t/(\gamma^*_{v_t})^2} \frac{\eta_{v_t}}{2(lv_t)^{1/2}} \times m\int_{x=2}^{\infty} e^{-3mx/10}dx \\
&= \sum_{l=v_t+1}^{4m^2 v_t/(\gamma^*_{v_t})^2} \frac{10\eta_{v_t}e^{-3m/5}}{6(lv_t)^{1/2}} \\
&\le \frac{e^{-d_3 m}\eta_{v_t}}{\gamma^*_{v_t}}.
\end{aligned} \tag{94}$$

Note that we aborbed an $O(m)$ factor into the expression in (94). This is valid because m is large. We continue to do this where possible.

$$\begin{aligned}
E_3 &= \sum_{l=v_t+1}^{n} \frac{\eta_{v_t}\mathbf{1}_{l>4m^2 v_t/(\gamma^*_{v_t})^2}}{2(lv_t)^{1/2}} \int_{\eta=0}^{\gamma^*_{v_t}(l/v_t)^{1/2}}\int_{\eta_l=\eta}^{\infty} \frac{\eta_l^{m-1}e^{-\eta_l}}{(m-1)!}\mathbf{Pr}(d_n(l)\ge \Delta^{v_t}_{\gamma^*_{v_t}} \mid H_t)d\eta_l d\eta \\
&\le \sum_{l=4m^2 v_t/(\gamma^*_{v_t})^2}^{n} \frac{\eta_{v_t}}{2(lv_t)^{1/2}} \int_{\eta=0}^{\gamma^*_{v_t}(l/v_t)^{1/2}}\int_{\eta_l=\gamma^*_{v_t}(l/v_t)^{1/2}}^{\infty} \frac{\eta_l^{m-1}e^{-\eta_l}}{(m-1)!}d\eta_l d\eta \\
&\le \sum_{l=4m^2 v_t/(\gamma^*_{v_t})^2}^{n} \frac{\eta_{v_t}}{2(lv_t)^{1/2}} \int_{\eta=0}^{\gamma^*_{v_t}(l/v_t)^{1/2}} \exp\left\{-\frac{3\gamma^*_{v_t}(l/v_t)^{1/2}}{10}\right\}d\eta \quad \text{from Lemma 2.1(c),} \\
&\le \sum_{l=4m^2 v_t/(\gamma^*_{v_t})^2}^{n} \frac{\eta_{v_t}}{2(lv_t)^{1/2}} \exp\left\{-\frac{3\gamma^*_{v_t}(l/v_t)^{1/2}}{10}\right\}, \\
&\le \frac{\eta_{v_t}\gamma^*_{v_t}}{v_t}\int_{x=4m^2 v_t/(\gamma^*_{v_t})^2}^{n} \exp\left\{-\frac{3\gamma^*_{v_t}(x/v_t)^{1/2}}{10}\right\}dx, \\
&\le \frac{\eta_{v_t}\gamma^*_{v_t}}{v_t} \times \frac{8v_t}{(\gamma^*_{v_t})^2}\int_{y=m}^{\infty} ye^{-3y/5}dy,
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\eta_{v_t} e^{-d_4 m}}{\gamma^*_{v_t}}. \\
E_4 &= \sum_{l=i}^{n} \frac{\eta_{v_t} \mathbf{1}_{l > 4m^2 v_t/(\gamma^*_{v_t})^2}}{2(lv_t)^{1/2}} \int_{\eta=\gamma^*_{v_t}(l/v_t)^{1/2}}^{\infty} \int_{\eta_l=\eta}^{\infty} \frac{\eta_l^{m-1} e^{-\eta_l}}{(m-1)!} \mathbf{Pr}(d_n(l) \geq \Delta^{v_t}_{\gamma^*_{v_t}} \mid H_t, \eta_l) d\eta_l d\eta \\
&\leq \sum_{l=4m^2 v_t/(\gamma^*_{v_t})^2}^{n} \frac{\eta_{v_t}}{2(lv_t)^{1/2}} \int_{\eta=\gamma^*_{v_t}(l/v_t)^{1/2}}^{\infty} \int_{\eta_l=\eta}^{\infty} \frac{\eta_l^{m-1} e^{-\eta_l}}{(m-1)!} d\eta_l d\eta \\
&\leq \sum_{l=4m^2 v_t/(\gamma^*_{v_t})^2}^{n} \frac{\eta_{v_t}}{2(lv_t)^{1/2}} \int_{\eta=\gamma^*_{v_t}(l/v_t)^{1/2}}^{\infty} e^{-3\eta/10} d\eta \\
&\leq \sum_{l=4m^2 v_t/(\gamma^*_{v_t})^2}^{n} \frac{5\eta_{v_t}}{3(lv_t)^{1/2}} \exp\left\{ -\frac{3\gamma^*_{v_t}(l/v_t)^{1/2}}{10} \right\}, \\
&\leq \frac{2\eta_{v_t}}{i^{1/2}} \int_{x=4m^2 v_t/(\gamma^*_{v_t})^2}^{\infty} x^{-1/2} \exp\left\{ -\frac{3\gamma^*_{v_t}(x/j)^{1/2}}{10} \right\} dx, \\
&= \frac{4\eta_{v_t}}{\gamma^*_{v_t}} \int_{y=2m}^{\infty} e^{-3y/10} dy, \\
&\leq \frac{\eta_{v_t} e^{-d_5 m}}{\gamma^*_{v_t}}.
\end{aligned}
$$

Thus,

$$
\mathbf{E}(\eta_{v_{t+1}} \mathbf{1}_{v_{t+1} > v_t} \mid H_t) \leq E_1 + E_2 + E_3 + E_4 \leq \begin{cases} \left(\frac{5m^2}{\gamma^*_{v_t}} + \frac{e^{-d_6 m}}{\gamma^*_{v_t}} \right) \eta_{v_t} \leq \frac{7m^2}{\gamma^*_{v_t}} \eta_{v_t} & \gamma^*_{v_t} \leq 5/4 \\ \left(\frac{e^{-d_7 m}}{\gamma^*_{v_t}} + \frac{200}{m} \right) \eta_{v_t} & \gamma^*_{v_t} > 5/4. \end{cases}
$$

We now integrate with respect to the value of $\gamma^*_{v_t}$. (Note that $\gamma^*_{v_t}$ is actually a discrete random variable, so that $\mathbf{Pr}(\gamma^*_{v_t} \leq \gamma \mid H_t)$ is discontinuous, but one can view this as a Riemann-Stieltjes integral. We write $\mathbf{Pr}^\dagger(\gamma^*_{v_t} \leq \gamma)$ below in place of $\mathbf{Pr}(\gamma^*_{v_t} \leq \gamma \mid H_t)$.) Using Lemma 3.12 we see that if m is large then integrating over γ,

$$
\begin{aligned}
&\mathbf{E}(\eta_{v_{t+1}} \mathbf{1}_{v_{t+1} > v_t} \mid H_t) \\
&\leq \eta_{v_t} \left(\int_{\gamma=0}^{5/4} \frac{7m^2}{\gamma} d\, \mathbf{Pr}^\dagger(\gamma^*_{v_t} \leq \gamma) + \int_{\gamma=5/4}^{\infty} \frac{e^{-d_7 m}}{\gamma} d\, \mathbf{Pr}^\dagger(\gamma^*_{v_t} \leq \gamma) + \frac{200}{m} \right) \\
&\leq \eta_{v_t} \left(\int_{\gamma=0}^{5/4} \frac{7m^2}{\gamma} d\, \mathbf{Pr}^\dagger(\gamma^*_{v_t} \leq \gamma) + e^{-d_8 m} + \frac{200}{m} \right) \\
&= \eta_{v_t} \left(\int_{\gamma=0}^{1/m} \frac{7m^2}{\gamma} d\, \mathbf{Pr}^\dagger(\gamma^*_{v_t} \leq \gamma) + \int_{\gamma=1/m^{1/2}}^{5/4} \frac{7m^2}{\gamma} d\, \mathbf{Pr}^\dagger(\gamma^*_{v_t} \leq \gamma) + e^{-d_8 m} + \frac{200}{m} \right) \\
&\leq \eta_{v_t} \left(\left[\frac{7m^2}{\gamma} \mathbf{Pr}^\dagger(\gamma^*_{v_t} \leq \gamma) \right]_0^{1/m} + \int_{\gamma=0}^{1/m^{1/2}} \frac{7m^2}{\gamma^2} \mathbf{Pr}^\dagger(\gamma^*_{v_t} \leq \gamma) d\gamma \right.
\end{aligned}
$$

$$
\begin{aligned}
&\qquad\qquad + \int_{\gamma=1/m}^{5/4} \frac{7m^2}{\gamma} d\,\mathbf{Pr}^\dagger(\gamma_{v_t}^* \le \gamma) + e^{-d_8 m} + \frac{200}{m}\Bigg) \\
&\le \eta_{v_t} \Bigg(e^{-d_9 m^{1/2}} + \int_{\gamma=0}^{1/m} \frac{7m^2}{\gamma^2}\,\mathbf{Pr}^\dagger(\gamma_{v_t}^* \le \gamma) d\gamma + \\
&\qquad \int_{\gamma=1/m}^{5/4} \frac{7m^3}{\gamma} d\,\mathbf{Pr}^\dagger(\gamma_{v_t}^* \le \gamma) + e^{-d_8 m} + \frac{200}{m}\Bigg), \qquad \text{from (34)}, \\
&\le \eta_{v_t} \left(e^{-d_9 m^{1/2}} + e^{-d_{10} m^{1/2}} + 7m^4\,\mathbf{Pr}^\dagger\left(\frac{1}{m} \le \gamma_{v_t}^* \le 5/4\right) + e^{-d_8 m} + \frac{200}{m}\right), \\
&\le \eta_{v_t} \left(e^{-d_9 m^{1/4}} + e^{-d_{10} m^{1/2}} + 7m^4 e^{-c_1 m} + e^{-d_8 m} + \frac{200}{m}\right), \qquad \text{from (35)} \\
&\le \eta_{v_t} \left(e^{-d_{12} m^{1/4}} + \frac{200}{m}\right).
\end{aligned} \tag{95}
$$

Combining (84) and (95) via (83), we have that

$$
\mathbf{E}(\eta_{v_{t+1}} \mid H_t) \le 3m + \left(e^{-cm^{1/4}} + \frac{200}{m}\right)\eta_{v_t}. \tag{96}
$$

This completes Case 1. Case 2 is much shorter.

Case 2: H_t is such that $v_t > \left(1 - \frac{1}{\omega^{1/2}}\right) n$.

$$
\begin{aligned}
\mathbf{E}(\eta_{v_{t+1}} \mid H_t) &\lesssim \sum_{l=v_t+1}^{n} \frac{\eta_{v_t}}{2(lv_t)^{1/2}} \int_{\eta=0}^{\infty} \int_{\eta_l=\eta}^{\infty} \frac{\eta_l^{m-1} e^{-\eta_l}}{(m-1)!} d\eta_l d\eta \\
&\sim \sum_{l=v_t+1}^{n} \frac{\eta_{v_t}}{2(lv_t)^{1/2}} \int_{\eta=0}^{\infty} e^{-\eta} \sum_{i=1}^{m} \frac{\eta^{m-i}}{(m-i)!} d\eta \\
&\sim \sum_{l=v_t+1}^{n} \frac{m\eta_{v_t}}{2(lv_t)^{1/2}} \\
&\lesssim m\eta_{v_t} \frac{(n+1)^{1/2} - (v_t+1)^{1/2}}{(v_t+1)^{1/2}} \\
&\le \frac{m\eta_{v_t}}{\omega^{1/2}}.
\end{aligned}
$$

This completes Case 2. In particular, for sufficiently large n we see that for any typical H_t (i.e., in both Case 1 and Case 2), the bound from (96) is valid. Putting

$$
\mathcal{E}_t = \mathcal{E} \cap \{H_t \text{ is typical}\}
$$

we deduce from (96) that

$$
\mathbf{E}(\eta_{v_{t+1}} \mid \mathcal{E}_t) \le 3m + \left(e^{-cm^{1/4}} + \frac{200}{m}\right)\mathbf{E}(\eta_{v_t} \mid \mathcal{E}_t) \tag{97}
$$

$$\lesssim 3m + \left(e^{-cm^{1/4}} + \frac{200}{m}\right) \mathbf{E}(\eta_{v_t} \mid \mathcal{E}_{t-1}). \tag{98}$$

We obtain (98) from (97) because $\mathcal{E}_t \subseteq \mathcal{E}_{t-1}$ and so

$$\mathbf{E}(\eta_{v_t} \mid \mathcal{E}_{t-1}) \geq \mathbf{E}(\eta_{v_t} \mid \mathcal{E}_t)\, \mathbf{Pr}(\mathcal{E}_t \mid \mathcal{E}_{t-1}) \sim \mathbf{E}(\eta_{v_t} \mid \mathcal{E}_t).$$

Because m is large, (21) will follow using (98) as a recurrence once we have shown that

$$\mathbf{E}(\eta_{v_1}) \leq 3m. \tag{99}$$

Here we will use the assumption that v_1 is chosen exacty according to the stationary distribution for a simple random walk on G_n. In particular, we have

$$\mathbf{Pr}(\eta_{v_1} \geq \eta) \leq \mathbf{E}\left(\sum_{i=1}^{n} \frac{d_n(i)}{2mn} 1_{\eta_{v_1} \geq \eta}\right),$$

and Lemma 2.2 implies that if $\eta \geq 2m$

$$\mathbf{E}((d_n(i) - m) 1_{\eta_{v_1} \geq \eta}) \lesssim \left(\frac{n}{i}\right)^{1/2} \int_{\eta=2m}^{\infty} \frac{\eta^m e^{-\eta}}{(m-1)!} d\eta \lesssim \left(\frac{n}{i}\right)^{1/2} \times 4me^{-\eta/2}.$$

Furthermore,

$$\mathbf{E}(m 1_{\eta_{v_t} \geq \eta}) = m\, \mathbf{Pr}(\eta_{v_1} \geq \eta) \leq 5me^{-\eta/2}.$$

So, if $\eta \geq 2m$, then

$$\mathbf{Pr}(\eta_{v_1} \geq \eta) \lesssim \frac{9e^{-\eta/2}}{2n^{1/2}} \sum_{i=1}^{n} \frac{1}{i^{1/2}} \leq 10e^{-\eta/2}.$$

Therefore,

$$\mathbf{E}(\eta_{v_1}) \leq 2m + \int_{\eta=2m}^{\infty} \mathbf{Pr}(\eta_{v_1} \geq \eta) d\eta \leq 2m + 10 \int_{\eta=2m}^{\infty} e^{-\eta/2} d\eta$$

and (99) follows. □

3.2 Exiting the main loop

In summary, it follows that w.h.p. DCA reaches Step 7 in $O(\omega \log n)$ time. Also, at this time $v_T \leq \log^{1/49} n$. This follows from Lemma 2.2(g), 2.2(h) and **(P4)**. Furthermore, this justifies using n_1 as a lower bound on vertices visited during the main loop. The random walk of Step 8 will w.h.p. take place on $[\log^{1/9} n]$. This follows from Lemma 2.2(j). Vertex 1 will be in the same component as v_t in the subgraph of G_n induced by vertices of degree at least $\frac{n^{1/2}}{\log^{1/20} n}$. This is because there is a path from v_T to vertex 1 through vertices in $[v_T]$ only and furthermore it follows from Lemma 2.2(i) that w.h.p. every vertex on this path has degree at least $\frac{n^{1/2}}{\log^{1/20} n}$. The expected time to visit all vertices of a graph with ν vertices is $O(\nu^3)$, see for example Aleliunas, Karp, Lipton, Lovász and Rackoff [1]. Consequently, vertex 1 will be reached in a further $O((\log^{1/9} n)^3) = o(\log n)$ steps w.h.p, completing the proof of Theorem 1.2. □

4 Concluding remarks

We have described an algorithm that finds a distinguished vertex quickly and which is local in a strong sense. There are some natural questions that are left unanswered:

- Can the running time be improved from $O(\omega \log n)$ to $O(\log n)$?
- Can we get polylog expected running time for DCA if $m = 2$?
- Can we extend the analysis to other more general models of web graphs e.g. Cooper and Frieze [7]. In this case, we would not be able to use the model described in Section 2.

As a final observation, the algorithm DCA could be used to find the vertex of largest degree: if we replace Step 8 by "Do the random walk for $\log n$ steps and output the vertex of largest degree encountered" then w.h.p. this will produce a vertex of highest degree. This is because $\log n$ will be enough time to visit all vertices $v \leq \log^{1/39} n$, where the maximum degree vertex lies.

References

[1] R. Aleliunas, R.M. Karp, R.J. Lipton, L. Lovász and C. Rackoff, Random Walks, Universal Traversal Sequences, and the Complexity of Maze Problems. *Proceedings of the 20th Annual IEEE Symposium on Foundations of Computer Science* (1979) 218-223.

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

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

[4] B. Bollobás and O. Riordan, The Diameter of a Scale-Free Random Graph, *Combinatorica* 24 (2004) 5-34.

[5] C. Borgs, M. Brautbar, J. Chayes, S. Khanna and B. Lucier, The Power of Local Information in Social Networks, http://arxiv.org/abs/1212.0884.

[6] M. Brautbar and M. Kearns, Local Algorithms for Finding Intersting Individuals in Large Networks, in *Innovations in Computer Science (ITCS)* (2010) 188-199.

[7] C. Cooper and A.M. Frieze, On a general model of web graphs, *Random Structures and Algorithms* 22 (2003) 311-335.

[8] G. Grimmett and D. Stirzaker, Probability and Random Processes, Third Edition, Oxford University Press, Oxford UK, 2001.

[9] W. Hoeffding, Probability inequalities for sums of bounded random variables, *Journal of the American Statistical Association* 58 (1963) 13-30.

[10] M. Jerrum and A. Sinclair. The Markov chain Monte Carlo method: an approach to approximate counting and integration. In *Approximation Algorithms for NP-hard Problems.* (D. Hochbaum ed.) PWS (1996) 482-520.

[11] M. Mihail, C. Papadimitriou and A. Sabeeri, On Certain Connectivity Properties of the Internet Topology, *Journal of Computer and System Sciences* 72 (2006) 239251.

A Proofs of properties (P1)–(P5)

In this section we give proofs of **(P1)–(P5)**, which we list here for convenience.

(P1) For $\Upsilon_{k,\ell} = \Upsilon_k - \Upsilon_\ell$, we have

$$\Upsilon_{k,\ell} \in (k-\ell)\left[1 \pm \frac{L\theta_{k,\ell}^{1/2}}{3(k-\ell)^{1/2}}\right]$$

for $(k,\ell) = (mn+1, 0)$ or

$$\frac{k-\ell}{m} \in \{\omega, \omega+1, \ldots, n\} \text{ and } k-l \geq \begin{cases} 1 & l=0 \\ \log^2 n & k \geq \log^{30} n, l > 0 \\ \log^{1/300} n & 0 < l < k < \log^{30} n. \end{cases}$$

Here $n_0 = \frac{\lambda_0^2 n}{\omega \log^2 n}$, $\lambda_0 = \frac{1}{\log^{20/m} n}$,

$$\theta_{k,\ell} = \begin{cases} \log k & \omega \leq l < k \leq \log^{30} n, l > 1 \\ k^{1/2} & \omega \leq k \leq n^{2/5}, l = 0 \\ (k-\ell)^{1/2} & \log^{30} n < k \leq n^{2/5} \\ \frac{(k-\ell)^{3/2}\log n}{n^{1/2}} & n^{2/5} < k \leq n_0. \\ \frac{n}{\omega^{3/2}\log^2 n} & n_0 < k. \end{cases}$$

(P2) $W_i \in \left(\frac{i}{n}\right)^{1/2}\left[1 \pm \frac{L\theta_i^{1/2}}{i^{1/2}}\right] \sim \left(\frac{i}{n}\right)^{1/2}$ for $\omega \leq i \leq n$.

(P3) $w_i \sim \frac{\eta_i}{2m(in)^{1/2}}$ for $\omega \leq i \leq n$.

(P4) $\lambda_0 \leq \eta_i \leq 40m \log\log n$ for $i \in [\log^{30} n]$.

(P5) $\eta_i \leq \log n$ for $i \in [n]$.

Proof of (P1)

Applying Lemma 2.1(d),(e) to (1) for $i \geq 1$ we see that

$$\begin{aligned}
&\mathbf{Pr}(\neg\mathbf{(P1)}) \\
&\leq 2\sum_{k=\omega}^{n} \exp\left\{-\frac{L^2\theta_{k,0}}{27}\right\} + 2\sum_{k-\ell=\log^{1/300} n}^{n} \exp\left\{-\frac{L^2\theta_{k,\ell}}{27}\right\} + 2\exp\left\{-\frac{L^2\theta_{mn+1,0}}{27}\right\} \\
&= 2\sum_{k=\omega}^{n^{2/5}} \exp\left\{-\frac{L^2k^{1/2}}{27}\right\} + 2\sum_{k=n^{2/5}+1}^{n_0} \exp\left\{-\frac{L^2k^{3/2}\log n}{27n^{1/2}}\right\} + 2\sum_{k=n_0+1}^{n+1} \exp\left\{-\frac{L^2n}{27\omega^{3/2}\log^2 n}\right\} \\
&+ 2\sum_{k-\ell=\log^{1/300} n}^{\log^{30} n} \exp\left\{-\frac{L^2\log k}{27}\right\} + 2\sum_{k-\ell=\log^{30} n}^{n^{2/5}} \exp\left\{-\frac{L^2(k-\ell)^{1/2}}{27}\right\} + \\
&2\sum_{k-\ell=n^{2/5}+1}^{n_0} \exp\left\{-\frac{L^2(k-\ell)^{3/2}\log n}{27n^{1/2}}\right\} + 2\sum_{k-\ell=n_0+1}^{n} \exp\left\{-\frac{L^2n}{27\omega^{3/2}\log^2 n}\right\} \\
&= o(1).
\end{aligned}$$

Proof of (P2)

For this we use

$$W_i = \left(\frac{\Upsilon_{mi}}{\Upsilon_{mn+1}}\right)^{1/2}.$$

Then,

$$W_i \notin \left(\frac{i}{n}\right)^{1/2}\left[1 \pm \frac{L\theta_i^{1/2}}{i^{1/2}}\right]$$

implies that either

$$\Upsilon_{mn+1} \notin (mn+1)\left[1 \pm \frac{L\theta_i^{1/2}}{3(mn+1)^{1/2}}\right] \text{ or } \Upsilon_{mi} \notin mi\left[1 \pm \frac{L\theta_i^{1/2}}{3i^{1/2}}\right].$$

These events are ruled out w.h.p. by **(P1)**.

Proof of (P3)

We use $(1+x)^{1/2} \leq 1 + \frac{x}{2}$ for $0 \leq |x| \leq 1$. Then,

$$w_i = \left(\frac{\Upsilon_{mi}}{\Upsilon_{mn+1}}\right)^{1/2} - \left(\frac{\Upsilon_{m(i-1)}}{\Upsilon_{mn+1}}\right)^{1/2}$$

$$= \left(\frac{\Upsilon_{m(i-1)}}{\Upsilon_{mn+1}}\right)^{1/2} \left(\left(1+\frac{\eta_i}{\Upsilon_{m(i-1)}}\right)^{1/2} - 1\right)$$
$$\leq \frac{\left(m(i-1)\left(1+\frac{L\theta_i^{1/2}}{3m^{1/2}(i-1)^{1/2}}\right)\right)^{1/2}}{\left((mn+1)\left(1-\frac{L\theta_i^{1/2}}{3(mn+1)^{1/2}}\right)\right)^{1/2}} \frac{\eta_i}{2m(i-1)\left(1-\frac{L\theta_i^{1/2}}{3m^{1/2}(i-1)^{1/2}}\right)}$$
$$\leq \frac{\eta_i}{2m(in)^{1/2}}\left(1+\frac{2L\theta_i^{1/2}}{m^{1/2}i^{1/2}}\right).$$

A similar calculation gives

$$w_i \geq \frac{\eta_i}{2m(in)^{1/2}}\left(1-\frac{2L\theta_i^{1/2}}{m^{1/2}i^{1/2}}\right).$$

Proof of (P4)

The upper bound follows from Lemma 2.1(c). For the lower bound, we observe by (7) that the expected number of $i \leq \log^{30} n$ with $\eta_i \leq \lambda_0$ is at most $\log^{30} n \times \lambda_0^m = o(1)$.

Proof of (P5)

This follows from Lemma 2.1(c).

B Proof of Lemma 2.2

We restate the lemma for convenience.

Lemma B.1.

(a) If $\mathcal{E}$ occurs then $\bar{d}_n - m \in [\eta_i\zeta(i), \eta_i\zeta^+(i)]$.

(b) $\mathbf{Pr}(d_n(i) - m \leq (1-\alpha)m\zeta(j)) \leq e^{-\alpha^2\eta_i\zeta(i)/2}$ for $0 \leq \alpha \leq 1$.

(c) $\mathbf{Pr}(d_n(i) - m \geq (1+\alpha)m\zeta^+(j)) \leq e^{-\alpha^2\eta_i\zeta^+(i)/3}$ for $0 \leq \alpha \leq 1$.

(d) $\mathbf{Pr}(d_n(i) - m \geq \beta m\zeta^+(j)) \leq (e/\beta)^{\beta\eta_i\zeta^+(i)}$ for $\beta \geq 2$.

(e) W.h.p. $\eta_i \geq \lambda_0$ and $\omega \leq i \leq n^{1/2}$ implies that $d_n(i) \sim \eta_i \left(\frac{n}{i}\right)^{1/2}$.

(f) W.h.p. $\omega \leq i \leq \log^{30} n$ implies that $d_n(i) \sim \eta_i \left(\frac{n}{i}\right)^{1/2}$.

(g) *W.h.p.* $\omega \leq i \leq n^{1/2}$ *implies that* $d_n(i) \lesssim \max\{1, \eta_i\} \left(\frac{n}{i}\right)^{1/2}$.

(h) *W.h.p.* $n^{1/2} \leq i \leq n$ *implies* $d_n(i) \leq n^{1/3}$.

(i) *W.h.p.* $1 \leq i \leq \log^{1/49} n$ *implies that* $d_n(i) \geq \frac{n^{1/2}}{\log^{1/20} n}$.

(j) *W.h.p.* $d_n(i) \geq \frac{n}{\log^{1/20} n}$ *implies* $i \leq \log^{1/9} n$.

Proof. (a) Suppose that we fix the values for $W_1, W_2, \ldots, W_n$. Then the degree $d_n(i)$ of vertex i can be expressed

$$d_n(i) = m + \sum_{j=i}^{n} \sum_{k=1}^{m} \zeta_{j,k}$$

where the $\zeta_{j,k}$ are independent Bernouilli random variables such that

$$\mathbf{Pr}(\zeta_{j,k} = 1) \in \left[\frac{w_i}{W_j}, \frac{w_i}{W_{j-1}}\right].$$

So, putting

$$\bar{d}_n(i) = \mathbf{E}(d_n(i))$$

we have

$$mw_i \sum_{j=i}^{n} \frac{1}{W_j} \leq \bar{d}_n(i) - m \leq mw_i \sum_{j=i-1}^{n} \frac{1}{W_j}.$$

Now assuming that **(P2)**, we have for $\omega \leq i \leq n$,

$$\sum_{j=i}^{n} \frac{1}{W_j} \geq \sum_{j=i}^{n} \left(\frac{n}{j}\right)^{1/2} \left(1 - \frac{2L\theta_j^{1/2}}{j^{1/2}}\right).$$

But

$$\begin{aligned}
\sum_{j=\omega}^{n} \frac{\theta_j^{1/2}}{j} &\leq \sum_{j=\omega}^{n^{2/5}} \frac{1}{j^{3/4}} + \sum_{j=n^{2/5}+1}^{n_0} \frac{\log^{1/2} n}{n^{1/4} j^{1/4}} + \sum_{j=n_0+1}^{n} \frac{n^{1/2}}{j\omega^{3/4} \log n} \\
&\leq 4n^{1/10} + \frac{4n^{1/2}}{3\omega^{3/4} \log n} + \frac{3n^{1/2} \log\log n}{\omega^{3/4} \log n} \\
&\leq \frac{4n^{1/2} \log\log n}{\omega^{3/4} \log n}.
\end{aligned}$$

It follows that

$$\begin{aligned}
\bar{d}_n(i) &\geq m + mw_i n^{1/2} \left(2(n^{1/2} - (i+1)^{1/2}) - \frac{9Ln^{1/2} \log\log n}{\omega^{3/4} \log n}\right) \\
&\geq m + \eta_i \left(\frac{n}{i}\right)^{1/2} \left(1 - \left(\frac{i}{n}\right)^{1/2} - \frac{1}{n^{1/2} i^{1/2}} - \frac{9L \log\log n}{2\omega^{3/4} \log n}\right),
\end{aligned}$$

after using **(P3)**.

A similar calculation gives a similar upper bound for $\bar{d}_n(i)$ and this proves that

$$i \geq \omega \text{ implies that } \bar{d}_n(i) \in m + \eta_i \left(\frac{n}{i}\right)^{1/2} \left[1 - \left(\frac{i}{n}\right)^{1/2} \pm \frac{5L \log\log n}{\omega^{3/4} \log n}\right].$$

It follows from (2) and (4) that

$$\mathbf{Pr}\left(d_n(i) - m \leq (1-\alpha)\eta_i \zeta(i) \middle| \, \eta_i\right) \leq \exp\left\{-\frac{L^2 \eta_i n^{1/2}}{4 i^{1/2} \omega^{1/2}}\right\}.$$

$$\mathbf{Pr}\left(d_n(i) - m \geq (1+\alpha)\eta_i \zeta(i) \middle| \, \eta_i\right) \leq \exp\left\{-\frac{L^2 \eta_i n^{1/2}}{4 i^{1/2} \omega^{1/2}}\right\}.$$

(a) For $\eta_i \geq \lambda_0$ and $\omega \leq i \leq n_0$ we have

$$\exp\left\{-\frac{L^2 \eta_i n^{1/2}}{4 i^{1/2} \omega^{1/2}}\right\} \leq e^{-L^2 \log n/4}.$$

(b) This follows from (a) and (4).

(c) This follows from (a) and (2).

(d) This follows from (a) and (3).

(e) This follows from (a), (b), (c) and (12).

(f) This follows from (e) and **(P4)**.

(g) This follows from (c) and (12).

(h) The degree of $i \geq n^{1/2}$ is stochastically dominated by the degree of $n^{1/2}$. Also, the probability that $d_n(n^{1/2})$ exceeds the stated upper bound is $o(1/n)$. So (h) follows from (g).

(i) For $\omega \leq i \leq \log^{1/49} n$, this follows from (f) and **(P4)**. For $1 \leq i < \omega$ we can use (b) with $\eta_i \geq \lambda_0$ and $\alpha = n^{-1/10}$.

(j) This follows from (e), (f) and (g) and **(P4)**. □