Diffusion limited aggregation in the layers model

Colin Cooper[*]
Department of Informatics
King's College
London WC2B 4BG
England

Alan Frieze[†]
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh PA15213
U.S.A.

December 12, 2023

Abstract

1 Introduction

Diffusion limited aggregation. In the classical model of Diffusion Limited Aggregation (DLA), introduced by Witten and Sander [10], [11], the process begins with a single particle cluster placed at the origin of a space, and then, one-at-a-time, particles make a random walk "from infinity" until they collide with, and stick to, the existing cluster. The process is particularly natural in Euclidean space with particles making Brownian motion, or on d–dimensional lattices. Simulations of DLA in two dimensions show tree-like figures with long branches, and Kesten [6] proved that for lattices, if $d = 2$ the length of these arms is at most order $n^{2/3}$ and if $d > 2$, at most $n^{2/d}$. A similar process, Reaction Limited Aggregation (RLA), differs in that particles stick with probability less than one on each collision. Indicative publications on RLA include Ball et al [2] or Meakin and Family [9].

A distinct but related process, Internal Diffusion Limited Aggregation (IDLA), was introduced by Diaconis and Fulton [3], as a protocol for recursively building a random aggregate of particles. In IDLA particles are added to the source vertex of an infinite graph, and make

*Research supported at the University of Hamburg, by a Mercator fellowship from DFG Project 491453517
†Research supported in part by NSF Grant DMS1952285

a random walk (over occupied vertices) until they visit an unoccupied vertex at which point they halt. Thus the first particle occupies the source, and subsequent particles stick to the outside of the component rooted at the source. Hence IDLA is the opposite of DLA, in that the particles walk in the occupied vertices until they find a space, rather than walking until they collide with the occupied component. The focus in IDLA has been on the limiting shape of the component formed by the occupied vertices. The formative work by Lawler et al [7], proved that, on d-dimensional lattices the limiting shape approaches a Euclidean ball; a result subsequently refined in [1], [5] and [8], amongst others.

As pointed out by Kesten in Remark 3 of [6], the concept of releasing a particle "from infinity" is somewhat informal. In [6] this is circumvented by using the limiting distribution of the first visit to the boundary of the current component from far away, conditional on making such a visit. An alternative approach is to attach a source vertex to a large but finite graph with a designated sink vertex. This analogue of DLA for finite graphs, was previously studied by Frieze and Pegden [4] for the Boolean lattice $\mathcal{B} = \{0,1\}^n$. In [4] the process evolves at discrete time steps $t = 0, 1, \ldots$; each of which has an associated cluster C_t. The initial cluster C_0 consists of a single vertex $\mathbf{0} = (0, \ldots, 0) \in \mathcal{B}$, the sink vertex. The cluster C_t is produced from C_{t-1} by choosing a random decreasing walk ρ_t from the source vertex $\mathbf{1} = (1, \ldots, 1)$. Let v be the last vertex of the initial segment of ρ_t which is disjoint from C_{t-1}, and set $C_t = C_{t-1} \cup \{v\}$. The process terminates at the first time t_f when $\mathbf{1} \in C_{t_f}$. The current paper continues this analysis of DLA on finite graphs by considering trees, and other graphs with a layered structure.

The layers model. Let $S_0, S_1, \ldots, S_k, S_{k+1}$ be disjoint sets of vertices, where S_i is connected to S_{i+1} by a complete bipartite graph, with edges directed from level i to level $i+1$, to form a layered graph G. The set S_i, at distance i from the source, forms layer i of G, at level i. We assume S_0 contains a single vertex v, the *source vertex* of the particles, and S_{k+1} contains a single vertex z the *sink vertex* of the particle cluster. Initially z is the only occupied vertex. The sink vertex z is an artificial vertex added to ensure connectivity of C_t, and to maintain compatibility with the hypercube model of [4].

Let $N_i = |S_i|$ be the size of set S_i. For $i = 1, \ldots, k$, the sizes of S_i, are either taken to be the same size, $N_i = n/k$ (the equal layers model), growing geometrically so that $N_i = d^i$ (the growing layers model).

To simplify notation we assume that in the equal layers model G is an $(n+1)$–vertex graph with layer sizes $N_i = n/k$ for $i = 1, .., k$; and that in the growing layers model, $d^k = n$ and thus G is a $(d^{k+1} - 1)/(d-1)$–vertex graph. Similarly for trees of height k and branching factor d, we assume that the final level is of size $d^k = n$.

DLA in the layers model. Initially, at step $t = 0$ all vertices of G are unoccupied. A step t consists of placing an active particle on the (unoccupied) source vertex and moving the particle forward level by level according to a random walk ρ_t from source to sink until it halts. To move forward from level $i < k$, the active particle currently at vertex u in level i chooses a random neighbour w in level $i + 1$. If the neighbour w is unoccupied the particle moves to w. If the neighbour w is occupied, the particle halts at vertex u in level i, occupies that vertex and adds a directed edge (u, w) to the edge induced component C_t rooted at the sink. As the sink z is occupied at the start, a particle reaching level $i + 1 = k$ halts there.

At any step, the component C_t is a rooted tree with edges directed towards the root, the sink z. The process stops at a step t_f, *the finish time*, when the source vertex v at level zero is occupied by a halted particle. In which case there is a directed path connecting the source v to the sink z, all of whose vertices are occupied by halted particles.

On deletion of the sink z, the digraph $D_t = C_t \backslash \{z\}$ consists of arborescences rooted at the occupied vertices in level k. Let $v = u_0 u_1 \cdots u_k u_{k+1} = z$ be the *connecting path*, the path connecting source and sink at the end of the process. Perhaps the most surprising thing is that in the digraph D_{t_f}, with high probability the arborescence rooted at u_k containing the connecting path is precisely the connecting path $v = u_0 u_1 \cdots u_k$, and contains no other vertices.

Theorem 1. *Let $v = u_0 u_1 \cdots u_k u_{k+1} = z$ be the path, connecting source and sink at the end of the process. Let $w = u_k$ be the vertex at level k on this path. The following results hold with high probability.*

1. *Equal layers model. For $i = 1, ..., k$ let $N_i = n/k$.*
   Let
   $$T_f = [(k+1)!(n/k)^k]^{1/(k+1)} = n^{k/(k+1)}\gamma_k$$
   where $\gamma_k = O(1)$ and $\gamma_k \to 1/e$ as $k \to \infty$.
   Provided $1 \le k \le \sqrt{(\log n)/(\log\log n)}$, the finish time t_f satisfies $T_f/\omega \le t_f \le \omega T_f$, where ω is any function tending to ∞.

2. *Trees of branching factor d, and the growing layers model. For $i = 0, ..., k$ let $N_i = d^i$.*
   Let
   $$T_f = \sqrt{k}\, d^{k+3/2-\sqrt{2k+2}}.$$
   (a) *Provided $d \to \infty$, $k \to \infty$, (and $k \ll d$ for trees) the finish time t_f satisfies $T_f/(\omega d^{O(1/\sqrt{k})}) \le t_f \le \omega T_f$.*
   (The extra term $d^{O(1/\sqrt{k})}$ is there to deal with some rounding error in j^*.)

   (b) *At t_f, levels $i = 1, \ldots, k - \lceil \sqrt{2k+2} - 1 \rceil$ contain a single occupied vertex, the vertex u_i of the the connecting path.*

3. *On termination, let $v = u_0 u_1, \cdots u_k z$ be the path connecting source v and sink z in C_{t_f}, and let $D_{t_f} = C_{t_f} \backslash \{z\}$. In either model, with high probability, the vertices $\{u_1, \cdots, u_k\}$ have in-degree one in C_{t_f}. Thus the arborescence rooted at u_k in D_{t_f} is exactly this connecting path.*

The proof of Theorem 1 also gives an upper bound on the finish time of DLA for trees of branching factor at least two.

Corollary 2. *Let*

$$T = d^{k+2-\sqrt{2k+2}+(1/2)\log k}.$$

For $d \geq 2$, the expected finish time of DLA on a tree of branching factor d is $O(T)$.

Notation. We use $A_n \sim B_n$ to denote $A_n = (1 + o(1))B_n$ and thus $\lim_{n\to\infty} A_n/B_n = 1$. We use ω to denote a quantity which tends to infinity with n more slowly than any other variables in the given expression. The expression $f(n) \ll g(n)$ indicates $f(n) = o(g(n))$. The inequality $A \lesssim B$ which stands for $A \leq (1 + o(1))B$ is used to unclutter notation in some places.

2 Bounds on occupancy in the layers model

For $t \geq 0$, let $L_i(t)$ be the number of particles halted in level i at the end of step t. Thus $L_i(0) = 0$ for all $i \leq k$, and $t = \sum_{i=0}^{k} L_i(t)$. We refer to L_i as the occupancy of level i. Note that $L_k(t) \leq t$, and $L_i(t) \leq \min(t, N_i)$.

The first step is to prove the following proposition.

Proposition 3. *Let*

$$\mu_{k-j}(t) = \frac{1}{N_k N_{k-1} \cdots N_{k-j+1}} \frac{t^{j+1}}{(j+1)!}. \tag{1}$$

Assume $t < t_f$, and thus the process has not stopped. Provided $\mu_{k-\ell}(t) \to \infty$ for all $\ell \leq j$, then $L_{k-j}(t) \sim \mu_{k-j}(t)$ in all models.

Let $\mathcal{H}(t) = (L_0(t), L_1(t), \ldots, L_k(t))$ be the history of the process up to and including step t.

$$\mathbb{E}\left(L_i(t+1) \mid \mathcal{H}(t)\right) = L_i(t) + \frac{L_{i+1}(t)}{N_{i+1}} \prod_{j=0}^{i} \left(1 - \frac{L_j(t)}{N_j}\right), \qquad i < k, \tag{2}$$

$$\mathbb{E}\left(L_k(t+1) \mid \mathcal{H}(t)\right) = L_k(t) + \prod_{j=0}^{k} \left(1 - \frac{L_j(t)}{N_j}\right). \tag{3}$$

Note that (3) follows from (2) as $L_{k+1}(t)/N_{k+1} = 1$ for all t, and that if $L_0(t) = 1$, the above recurrences give $L_i(t+1) = L_i(t)$.

To simplify matters for small t, we can condition on the event that $L_k(s) = s$ for $s \leq \sqrt{N_k/\omega}$, an event of probability $1-1/\omega$, and use the fact that the evolution of $(L_0(t), L_1(t), \ldots, L_k(t))$ is Markovian. Whatever happens, $L_i(t)$ is monotone non-decreasing in t.

2.1 Upper bound on occupancy at step t

The underlying random walk ρ_t, from source v to sink z defines a (v, z)-path given by $v = u_0 u_1, ..., u_k z = v_{k+1}$. Particle t follows this walk until halting at a vertex u_i, where u_{i+1} is the first occupied vertex encountered on the path.

Let $B_i(t)$ denote the occupied (blocked) vertices in level i at time t. We define an *upper-blocked process* which we use to upper bound $L_i(t)$. This process gives rise to sets $\widehat{B}_i(t) \supseteq B_i(t)$ and random variables $\widehat{L}_i(t) = |\widehat{B}_i(t)| \geq L_i(t)$. For *every* vertex u_j, $j \leq k$ on the walk ρ_t, if u_{j+1} is occupied add a vertex to $\widehat{B}_j(t)$ as follows. If $u_j \notin \widehat{B}_j(t)$ add u_j to $\widehat{B}_j(t+1)$. If $u_j \in \widehat{B}_j(t)$ add some other $u'_j \in S_j \setminus \widehat{B}_j(t)$ to $\widehat{B}_j(t+1)$.

In particular if particle $t+1$ halts at vertex u_i in the DLA process, then either u_i is added to both $B_i(t+1)$ and $\widehat{B}_i(t+1)$, or u_i is already a member of $\widehat{B}_i(t)$. In either case $B_i(t) \subseteq \widehat{B}_i(t)$ for all i and $t \geq 0$. It follows (among other things) that $L_i(t) \leq \widehat{L}_i(t)$, and $\widehat{L}_k(t) = t$. The next lemma gives w.h.p. bounds for $\widehat{L}_i(t)$.

Lemma 4. *Let $\mu_{k-j}(t)$ as be given by* (1)*.*

1. *The expectations $\mathbb{E}\,\widehat{L}_i(t)$ satisfy the recurrence*

$$\mathbb{E}\,(\widehat{L}_i(t+1) \mid \widehat{H}(t)) = \widehat{L}_i(t) + \frac{\widehat{L}_{i+1}(t)}{N_{i+1}}. \tag{4}$$

2. *If $j^2/t = o(1)$ then $\mathbb{E}\,\widehat{L}_{k-j}(t) \sim \mu_{k-j}(t)$.*

3. *If $t \geq t_{k-j}(\omega)$ as defined in* (12) *then w.h.p. $\widehat{L}_{k-j}(t) \sim \mu_{k-j}(t)$.*

Proof. Equation (4) follows because the upper bound process increases the size of $\widehat{B}_j(t)$ whenever the walk ρ_t contains a vertex of $\widehat{B}_{j+1}(t)$, this being true at all levels $j = 0, ..., k$. We have

$$\mathbb{E}\,(L_i(t+1) \mid \mathcal{H}(t)) \leq L_i(t) + \frac{L_{i+1}(t)}{N_{i+1}} \leq \widehat{L}_i(t) + \frac{\widehat{L}_{i+1}(t)}{N_{i+1}} = \mathbb{E}\,(\widehat{L}_{i+1}(t+1) \mid \widehat{\mathcal{H}}(t)).$$

Iterating (4) backwards for $0 \le s \le t$, and using $\widehat{L}_i(0) = 0$, gives

$$\mathbb{E}\,\widehat{L}_i(t) = \frac{1}{N_{i+1}} \sum_{s=0}^{t-1} \mathbb{E}\,\widehat{L}_{i+1}(s). \tag{5}$$

We claim for $j \ge 0$ that

$$\mathbb{1}_{\{t\ge j\}}\,\frac{(t-j)^{j+1}}{(j+1)!} \le (N_k N_{k-1}\cdots N_{k-j+1})\,\mathbb{E}\,\widehat{L}_{k-j}(t) \le \frac{t^{j+1}}{(j+1)!}. \tag{6}$$

For given t, the induction is backwards on $k-j$ from $j=0$. Now $N_{k+1}=1$ and $\widehat{L}_{k+1}(t)=1$ which implies that $\widehat{L}_k(s) = s$ for $0 \le s \le t$. So, (6) is true when $j=0$ and the first non-trivial case is $j=1$. From (5) we see that

$$\mathbb{E}\,\widehat{L}_{k-1}(t) = \frac{1}{N_k}\sum_{s=0}^{t-1} s, \tag{7}$$

which illustrates how (6) arises from bounding this sum.

For the general induction put $i = k-(j+1)$, and multiply (5) by $M_{j-1} = N_k N_{k-1}\cdots N_{k-j+1}$. Insert (6) with $i+1 = k-j$ into this, to give

$$\frac{1}{N_{k-j}}\sum_{s=j}^{t-1}\frac{(s-j)^{j+1}}{(j+1)!} \le M_{j-1}\,\mathbb{E}\,\widehat{L}_{k-(j+1)}(t) \le \frac{1}{N_{k-j}}\sum_{s=1}^{t-1}\frac{s^{j+1}}{(j+1)!} \tag{8}$$

By comparison of the sum with the related integral we have that

$$\frac{(t-1)^{m+1}}{m+1} \le 1^m + 2^m + \cdots + (t-1)^m \le \frac{t^{m+1}}{m+1}. \tag{9}$$

Use (9) in (8) with $m = j+1$, giving

$$\frac{\mathbb{1}_{\{t\ge j+1\}}}{N_{k-j}}\,\frac{(t-(j+1))^{j+2}}{(j+2)!} \le M_j\,\mathbb{E}\,\widehat{L}_{k-(j+1)}(t) \le \frac{1}{N_{k-j}}\frac{t^{j+2}}{(j+2)!},$$

which completes the induction for (6). Moreover, provided $j^2/t = o(1)$,

$$\mathbb{E}\,\widehat{L}_{k-j}(t) = \frac{1}{N_k N_{k-1}\cdots N_{k-j+1}}\frac{t^{j+1}}{(j+1)!}\left(1 - O\left(j^2/t\right)\right) = \mu_{k-j}(t)(1+o(1)). \tag{10}$$

This completes the proof of Lemma 4.(1) and Lemma 4.(2) . Note that $t_{k-j}(\omega) \gg (n/k)^{1/k} \gg k^2$ for $j \ge 1$, ensuring that $j^2/t = o(1)$ for j, t of interest.

2.2 Concentration of $\widehat{L}_i(t)$ for sufficiently large t.

We now prove Lemma 4.(3) .

By definition, $\widehat{L}_k(t) = t = \mu_k(t)$, establishing Lemma 4.(3) for $j = 0$.

We next consider $\widehat{L}_{k-j}(t)$ for $1 \leq j \leq k-1$. The random variable $\widehat{L}_{k-j}(t)$ is obtained by choosing a random vertex $u \in S_{k-j}$ irrespective of the current occupancy of u, and choosing a random neighbour $w \in S_{k-j+1}$. If w is occupied then $\widehat{L}_{k-j}(t+1) = \widehat{L}_{k-j}(t) + 1$. Thus, $\widehat{L}_{k-j}(t+1) = \widehat{L}_{k-j}(t) + Q_{k-j}(t)$ where $\mathbb{P}(Q_{k-j}(t) = 1) = (\widehat{L}_{k-j+1}(t)/N_{k-j})$ independently of any previous outcomes.

In particular, $\mathbb{E}\, Q_{k-1}(t) = t/N_k$, and by equation (7), $\mathbb{E}\, \widehat{L}_{k-1}(t) = t(t-1)/(2N_k) \sim \mu_{k-1}(t)$. By Hoeffding's Inequality,

$$\mathbb{P}\left(\widehat{L}_{k-1}(t) \notin \mu_{k-1}(t)\left(1 \pm \frac{1}{\omega}\right)\right) \leq 2\exp\left\{-\frac{(1-o(1))t^2}{6\omega^2 N_k}\right\} \leq 2e^{-\omega} \tag{11}$$

if $t \geq 2\omega^{3/2}\sqrt{2N_k}$.

For $j \geq 0$ we proceed as follows: let $t_1 = t_1(k-j)$ be such that $\mu_{k-j}(t_1) = 1$. Thus $t_1(k) = 1$ and

$$t_1(k-j) = [(j+1)! N_k N_{k-1} \cdots N_{k-j+1}]^{1/(j+1)}.$$

The fact that $t_1(k-j+1) \ll t_1(k-j)$ so that $\mu_{k-j+1}(t_1(k-j))$ is sufficiently large, is a model dependent calculation given in subsequent sections. The statement of Lemma 4 assumes the truth of this.

Let $t_k(\omega) = 1$, and for $j \geq 1$, let $t_{k-j}(\omega) = (4\omega^3)^{1/(j+1)} t_1(k-j)$ so that

$$t_{k-j}(\omega) = (4\omega^3)^{1/(j+1)}[(j+1)! N_k N_{k-1} \cdots N_{k-j+1}]^{1/(j+1)}, \tag{12}$$

and $\mu_{k-j}(t_{k-j}(\omega)) = 4\omega^3$.

Let $\mathcal{E}_j$ denote the event that $\widehat{L}_{k-j}(t) \in \mu_{k-j}(t)(1 \pm 1/\omega)$ for $t \geq t_{k-j}(\omega)$. We claim that $\mathbb{P}(\mathcal{E}_j) \geq 1 - 2je^{-\omega}$ and we verify this by induction.

For $j = 0$, $\mathbb{P}(\mathcal{E}_0) = 1$, and for $j = 1$, $\mathbb{P}(\mathcal{E}_1) \geq 1 - 2e^{-\omega}$ follows from equation (11). For $j \geq 2$, by Hoeffding's inequality, we have that for $t \geq t_{k-j}(\omega)$,

$$\mathbb{P}(\neg\mathcal{E}_j \mid \mathcal{E}_{j-1}) \leq 2\exp\left\{-\frac{\mu_{k-j}(t)}{3\omega^2}(1-\delta)\right\} \leq 2\exp\left\{-\frac{\mu(t_{k-j}(\omega))}{3\omega^2}(1-o(1))\right\} \leq 2e^{-\omega},$$

where $\delta = O(1/\omega + j^2/t)$, includes the correction from (10).

We complete the induction, on the assumption that $\log k \ll \omega$, via

$$\mathbb{P}(\neg\mathcal{E}_j) \leq \mathbb{P}(\neg\mathcal{E}_j \mid \mathcal{E}_{j-1}) + \mathbb{P}(\neg\mathcal{E}_{j-1}) \leq 2e^{-\omega} + 2(j-1)e^{-\omega} = 2je^{-\omega} = o(1). \tag{13}$$

$\square$

2.3 Lower bound on occupancy at step t

A lower bound $\widetilde{L}_i(t)$ on $L_i(t)$ can be found as follows. We consider a process in which a particle halts at level i whenever it chooses an out neighbour in $\widehat{B}_{i+1}(t)$. Let $L_j^*(t)$ bound $\widehat{L}_j(t)$ from above w.h.p. Let $\widetilde{L}_i(t)$ be obtained by replacing L by L^* in in the bracketed terms on the RHS of (2)–(3). Then $\widetilde{L} \leq L \leq \widehat{L}$ w.h.p., and we have the following recurrences.

$$\mathbb{E}\,\widetilde{L}_i(t+1) = \mathbb{E}\,\widetilde{L}_i(t) + \frac{\mathbb{E}\,\widetilde{L}_{i+1}(t)}{N_{i+1}} \prod_{j=0}^{i} \left(1 - \frac{L_j^*(t)}{N_j}\right), \tag{14}$$

$$\mathbb{E}\,\widetilde{L}_k(t+1) = \mathbb{E}\,\widetilde{L}_k(t) + \prod_{j=0}^{k} \left(1 - \frac{L_j^*(t)}{N_j}\right). \tag{15}$$

The solution to these recurrences is given in Section 3.2 for the equal layers model, and in Section 4 for the growing layers model.

3 Analysis of DLA in the equal layers model

The first subsection below shows that provided k is not too large, then w.h.p. there exists a large gap in the number of steps between the time for which $\mathbb{E}\,\widehat{L}_{k-j} = 1$ with all lower values zero, and the time for which $\mathbb{E}\,\widehat{L}_{k-(j+1)} = 1$. This gap allows $\widetilde{L}_{k-j}$ to increase and become concentrated, whilst at the same time all lower values remain zero w.h.p. This allows us to conclude that provided $\mu_i(t)$ is sufficiently large, $L_i(t)$ is concentrated around $\mu_i(t)$. We use this to conclude that $L_0(t) = 0$ until $L_1(t) \to \infty$ suitably fast, and thus estimate the finish time of the DLA process in the equal layers model.

3.1 Evolution of the state vector $\widehat{L}$ in the equal layers model.

Let $\widehat{L} = (\widehat{L}_0, \widehat{L}_1, \ldots, \widehat{L}_k)$ be the state vector of the upper-blocked process. The entries in $\widehat{L}$ are non-negative integers, and if $\widehat{L}_i = 0$, then $\widehat{L}_{i-1} = 0$. Let

$$T_f = \left[(k+1)! N_1 \cdots N_k\right]^{1/(k+1)}, \tag{16}$$

where $N_1 = \cdots = N_k = n/k$. It will be shown below that T_f is asymptotically the expected finish time. The following argument for $t \leq \omega T_f$ proves there is a large enough gap $t'' - t$ between $\mu_{k-j}(t) = 1$ and $\mu_{k-(j+1)}(t'') = 1$ for $\widehat{L}_{k-j}(t'')$ to be concentrated.

Define $\beta = \beta_k = N_k/\omega T_f$. As

$$[(k+1)!]^{1/(k+1)} = k\Theta(1 + k^{3/2k}) = k\Theta(1),$$

then

$$\beta = \frac{N_k}{\omega T_f} = \frac{n/k}{\omega[(k+1)!(n/k)^k]^{1/(k+1)}} = \Theta(1)\frac{1}{\omega k}\left(\frac{n}{k}\right)^{1/(k+1)}. \tag{17}$$

We can assume that $\beta \geq (\log n)^{k/2}$. This will be true if $\omega \leq \log n$ since we assume that $k \leq \sqrt{(\log n)/\log\log n}$.

We have for any $t \leq \omega T_f$, using that $N_{k-j} = N_k = n/k$ in the equal layers model, that

$$\mu_{k-(j+1)}(t) = \frac{t}{(j+2)N_{k-j}} \cdot \mu_{k-j}(t) \leq \frac{1}{j\beta}\mu_{k-j}(t), \tag{18}$$

and similarly

$$\mu_{k-(j+\ell)}(t) = \frac{t}{(j+\ell+1)N_{k-\ell-1}} \cdot \mu_{k-(j+\ell)-1}(t) \leq \frac{1}{(j\beta)^\ell}\mu_{k-j}(t), \qquad \text{for } 1 \leq \ell \leq k-j. \tag{19}$$

Consider $\mathbb{E}\,\widehat{L}_{k-j}(t)$. When $\mu_{k-j}(t) \sim 1$, then $\mathbb{E}\,\widehat{L}_{k-j}(t) \sim 1$ and the Markov inequality implies that w.h.p. $\widehat{L}_{k-j}(t) \in I_\omega = [0, 1, ..., \omega]$. From (18)–(19), when $\mu_{k-j}(t) \sim 1$,

$$\mu_{k-(j+\ell)}(t) \leq \frac{1+o(1)}{(j\beta)^\ell}, \qquad \text{for } 1 \leq \ell \leq k-j, \tag{20}$$

and thus w.h.p. $\widehat{L}_0 = 0, \widehat{L}_1 = 0, \ldots, \widehat{L}_{k-(j+1)} = 0$.

By arguments similar to that for (18), (19), we see that

$$\mu_{k-j+\ell}(t) \geq \mu_{k-j}(t)\; j(j-1)\cdots(j-\ell+1)\,\beta^\ell \qquad \text{for } \ell \leq j.$$

So if $\mu_{k-j}(t) \sim 1$ then for $\ell \leq j$, $\mu_{k-j+\ell}(t) \to \infty$ and $\widehat{L}_{k-j+\ell}$ is equal to $(1+o(1))\mathbb{E}\,\widehat{L}_{k-j+\ell}$ w.h.p. (As already remarked at the end of Section 2.1, $k^2/t = o(1)$ for t of interest here and we can invoke Lemma 4.) Thus at times t such that $\mu_{k-j}(t) \sim 1$ (implying that $t \leq \omega T_f$), w.h.p., the state vector $\widehat{L}$ is such that $\mu_{k-\ell} \to \infty$ for $\ell \leq j-1$, and

$$\widehat{L}(t) = (0, \ldots, 0,\ \widehat{L}_{k-j} \in I_\omega,\ (1+o(1))\mu_{k-j+1},\ (1+o(1))\mu_{k-j+2}, \ldots,\ (1+o(1))\mu_k). \tag{21}$$

Define $t' > t$ so that $\mathbb{E}\,\widehat{L}_{k-j}(t') \sim \mu_{k-j}(t') = \omega$ where $\omega \to \infty$ but $\omega \ll \beta$. It follows from (1) that $t' \sim \omega^{1/(j+1)}t$. Using $T_f = \gamma_k n^{k/(k+1)}$ from Proposition 5.1, for $j < k$ we have $t' \leq \omega T_f$ which follows from $t' = \Theta(1)\omega^{1/(j+1)}N_k^{j/(j+1)} \ll \omega T_f$ for ω growing sufficiently slowly,

$$\mu_{k-(j+1)}(t') = \frac{t'}{(j+2)N_k}\omega \leq \frac{\omega}{\beta}, \qquad \mu_{k-j}(t') = \omega, \qquad \mu_{k-j+1}(t') = \frac{(j+1)N_k}{t'}\omega \geq \omega\beta.$$

Choosing $\log\log n \leq \omega \ll \beta$, we have that w.h.p. $\widehat{L}_{k-(j+1)}(t') = 0$ as before and $\mu_{k-j+1} \geq \omega \log n$. But now

$$\widehat{L}(t') = (0, ..., 0,\ (1+o(1))\mu_{k-j},\ (1+o(1))\mu_{k-j+1}, ...,\ (1+o(1))\mu_k). \tag{22}$$

This condition persists until t'' such that $\mu_{k-(j+1)}(t'') \sim 1$. Equating $\mu_{k-j}(t) \sim 1 \sim \mu_{k-(j+1)}(t'')$ and using $t \le \omega T_f$ gives

$$(t'')^{j+2} \sim t^{j+1}(j+2)N_k \quad \Longrightarrow \quad t'' \ge t\left(\frac{N_k}{t}\right)^{1/(j+2)} \ge \ t\beta^{1/k} \ge t(\log n)^{1/2},$$

by our choice of β. We conclude that $t'' \ge (\log n)^{1/2} t$, so the claimed gap exists.

3.2 Lower bound on occupancy in the equal layers model.

Let $L_j^*(t) = \omega \mu_j(t)$. By (21), (22) of the previous section, this is a w.h.p. upper bound on $\widehat{L}_j(t)$ for $t \le \omega T_f$. Referring to (14)–(15), we claim that

$$\prod_{j=0}^{i}\left(1 - \frac{L_j^*(t)}{N_j}\right) \ge 1 - \sum_{j=0}^{i}\left(\frac{L_j^*(t)}{N_j}\right) \ge 1 - \frac{iL_i^*(t)}{N_i} = 1 - o(1). \tag{23}$$

Note first that as $T_f = O(kn^{k/(k+1)})$ and $N_k = n/k$, then provided $k \le \sqrt{\log n/\log\log n}$ and ω grows sufficiently slowly,

$$\frac{k\omega^2 T_f}{N_k} = O\left(\frac{\omega^2 k^{2+1/(k+1)}}{n^{1/(k+1)}}\right) = o(1); \tag{24}$$

and as $N_{k-j+1} = N_k$ and $\mu_{k-j}(t)/\mu_{k-j+1}(t) = t/(j+1)N_{k-j+1} = o(1)$ by (24), it follows that

$$\mu_{k-j}(t)/N_{k-j} \le t/N_k. \tag{25}$$

Thus (23) follows from (24), (25) and

$$\frac{jL_j^*(t)}{N_j} = \frac{j\omega\mu_j(\omega T_f)}{N_j} \le \frac{k\omega\mu_k(\omega T_f)}{N_k} \le \frac{k\omega^2 T_f}{N_k} = o(1). \tag{26}$$

Let $t_i(\omega)$ be given by (12). For those $i \le k$, and $t \ge t_i(\omega)$, so that $\mu_i(t) \to \infty$ suitably fast, we obtain that $\mathbb{E}\,\widetilde{L}_i(t) \sim \mu_i(t) \sim \mathbb{E}\,\widehat{L}_i(t)$ and thus w.h.p.,

$$\mathbb{E}\,L_{k-j}(t+1) \sim \mathbb{E}\,\widehat{L}_{k-j}(t+1) \sim \frac{t^{j+1}}{(j+1)!}\frac{1}{N_k N_{k-1}\cdots N_{k-j+1}}. \tag{27}$$

As $L_i(t) \le \widehat{L}_i(t)$ by construction, and w.h.p. $\widehat{L}_i(t) \le (1+o(1))\mathbb{E}\,\widehat{L}_i(t)$, this implies $L_i(t) \le (1+o(1))\mu_i(t)$ w.h.p., if $\mu_i(t) \to \infty$.

At step $t+1$, equation (23) implies that, in the DLA-process, particle $t+1$ arrives at level i with probability $(1-o(1))$. It halts at this level with probability $L_{i+1}(t)/N_{i+1} = o(1)$, (see (2), (25)), and thus w.h.p. all but $t(1-o(1))$ particles arrive at level k, the $o(1)$ term being given in (26).

It is an induction on $\ell = k-1, \dots, i$ in that order, that provided $t \geq t_\ell(\omega)$, then $\mu_\ell(t) \to \infty$ sufficiently fast, the values $L_k(t), L_{k-1}(t), \dots, L_i(t)$ are concentrated in the lower tail, according to the arguments in Section 2.2. Thus $L_i(t) \sim \mu_{i+1}(t)$ w.h.p..

In the case that $\mathbb{E}\,\widehat{L}_i(t) = o(1)$ then a fortiori $L_i(t) = 0$ w.h.p., and if $\mathbb{E}\,\widehat{L}_i(t) = O(1)$ then w.h.p. $L_i(t) \leq \omega\mu_i(t)$. We can thus argue that a gap theorem similar to that in Section 3.1 holds for the state vector $L = (L_0, L_1, \dots, L_k)$.

In particular, and most importantly, let $t_1 = t_1(0)$ be such that $\mu_0(t_1) = 1$, and thus $t_1 \sim T_f$. Let $t' = t_1/\omega$. Then w.h.p., $L_0(t') = O(1/\omega)$ and $\mathbb{E}\,L_1(t') \geq \beta/\omega$ where β is given in (17). It follows that for $t \geq t'$, $L_1(t)$ is concentrated around $\mu_1(t)$.

This completes the proof of Proposition 3 for the equal layers model.

3.3 Finish time of DLA in the equal layers model.

Proposition 5. *For $k \geq 2$, let G be an equal layers graph with level sizes $N_i = n/k$, $i = 1, ..., k$. Let T_f be given by (16), then*

1. $T_f = n^{k/(k+1)} e^{-1}(1 + O(\log k/k))$
2. *With high probability, the finish time t_f of the DLA process in G satisfies $T_f/\omega \leq t_f \leq \omega T_f$, where $\omega \to \infty$ arbitrarily slowly.*

Proof. Note that

$$\frac{(k+1)!}{k^k} \sim \frac{1}{k^k}\sqrt{2\pi}e^{-(k+1)}k^{k+3/2}(1+1/k)^{k+3/2} = e^{-(k+1)}\ \Theta_k,$$

where $\Theta_k \sim (e^{1-1/k+O(1/k^2)}\sqrt{2\pi}k^{3/2})$. Thus

$$T_f = [(k+1)!N_1 \cdots N_k]^{1/(k+1)} = \left(\frac{(k+1)!}{k^k}n^k\right)^{1/(k+1)} = n^{\frac{k}{k+1}}e^{-1}(\Theta_k)^{1/(k+1)}, \qquad (28)$$

where $(\Theta_k)^{1/k+1} = e^{O(\log k/k)}$ is bounded for $k \geq 2$ and tends to one as $k \to \infty$.

The choice of T_f is such that $\mathbb{E}\,\widehat{L}_0(T_f) \sim 1$, indeed from (10)

$$\mathbb{E}\,\widehat{L}_0(T_f) = \ 1 + O\left(\frac{k^2}{T_f}\right) = \ 1 + O\left(\frac{k^2}{n^{k/(k+1)}}\right), \qquad (29)$$

which is $1 + o(1)$, provided $k = o(\sqrt{n})$ say. Moreover $\mathbb{E}\,\widehat{L}_0(T_f/\omega) \sim 1/\omega^{k+1}$ proving that $t_f = \Omega(T_f)$ w.h.p.

The value of $\mathbb{E}\,\widehat{L}_{k-j}(t)$ is monotone decreasing with increasing j. This follows from (18) because $T_f = o(N_k)$, making β large, see (28). We next investigate the concentration of $L_1(T_f)$, which continues the discussion in the last paragraph of the previous section.

$$\mathbb{E}\,L_1(T_f) \sim \mu_1(T_f) \sim \frac{(T_f)^k}{k!(n/k)^{k-1}} = \frac{1}{k!}[(k+1)!)^k(n/k)]^{\frac{1}{k+1}} = \Theta(n^{1/(k+1)}k^{-3/2(k+1)}). \tag{30}$$

For concentration of $\widehat{L}_1(T_f)$ we require $\mathbb{E}\,L_1(T_f) \to \infty$, which occurs whenever $k = o(\log n)$.

At this point in the process, even if $L_0(T_f) = 0$, the expected waiting time for a particle to hit $L_1(T_f)$ is $O(n^{k/(k+1)}/k) = O(T_f)$, so w.h.p. this will occur by time ωT_f, completing the proof of Proposition 5, and hence Theorem 1.1. □

3.4 Existence of a unique connecting path component.

We regard occupied vertices as coloured either red or blue, with all occupied vertices in level k coloured blue.

A particle halts at vertex u in level i, if it chooses an edge uw to an occupied neighbour w in level $i+1$. We consider this edge uw as being directed from u to w in the component rooted at the sink. If u is the first in-neighbour of w then u is coloured blue. If however w already has an in-neighbour u', then u, u' and all other in-neighbours are (re)coloured red. At any step, the red vertices in a level are those with siblings, and the blue ones are the unique in-neighbour of some vertex in the next level. The process halts when there is a directed path of occupied vertices $v = u_0 u_1 \cdots u_k$ from the source to level k.

Lemma 6. *With high probability, the path $v = u_0 u_1 \cdots u_k = w$ from the source to level k is blue and thus the in-arborescence of halted particles rooted at $w = u_k$ is exactly this (v, w)-path.*

Proof. As before, let B_i be the set of occupied vertices in level i, where $L_i = |B_i|$. As each $u \in B_i$ has a unique out-neighbour, the subset $\mathrm{Out}(B_i)$ of B_{i+1} with at least one in-neighbour has size at most L_i.

Let $\mathbb{1}_{\{k-j,s\}}$ be the indicator that particle s halts in level $k-j$ and is coloured red due to a pre-existing sibling. In this case particle s has chosen an out-neighbour in the existing set $\mathrm{Out}(B_{k-j}) \subseteq B_{k-j+1}$, and thus

$$\mathbb{E}\ \mathbb{1}_{\{k-j,s\}} \le \frac{\mathbb{E}\,\widehat{L}_{k-j}(s)}{(n/k)} \sim \frac{\mu_{k-j}(s)}{(n/k)}.$$

If $Z_{k-j}(t)$ is the number of red vertices in level $k-j$, then

$$\mathbb{E}\, Z_{k-j}(t) \lesssim 2\sum_{s=1}^{t} \frac{\mu_{k-j}(s)}{(n/k)} = \frac{2}{(n/k)}\sum_{s=1}^{t} \frac{s^{j+1}}{(j+1)!(n/k)^j}$$

$$\sim 2\frac{t^{j+2}}{(j+2)!(n/k)^{j+1}} = 2\mu_{k-(j+1)}(t). \quad (31)$$

The factor 2 covers the case where the pre-existing sibling was blue but is recoloured red.

Define a random variable $\widehat{Z}_{k-j}(t) = Z_{k-j}(t) + Q_{k-j}(t) \geq Z_{k-j}(t)$, where Q_{k-j} is Bernoulii $Be(p_{k-j})$, with $p_{k-j} = (\widehat{L}_{k-j} - |\mathrm{Out}(B_{k-j})|)/(n/k)$. As $\mu_i \to \infty$ at t_f for all $i \geq 1$, for $i \geq 2$, $\widehat{Z}_i(t_f) = (1+o(1))\mathbb{E}\,\widehat{Z}_i(t_f)$ w.h.p., since $Z_i(t_f) \leq \widehat{Z}_i(t_f) \leq \widehat{L}_i(t_f)$. On the other hand, by definition $\mathbb{E}\, L_0(T_f) = 1$, so by the Markov Inequality $Z_1(t_f) \leq \widehat{Z}_1(t_f) \leq \omega$.

Consider the colour of the path $v = u_0u_1u_2\cdots u_k$ at $t \leq t_f$. Vertex v is blue as it is the first neighbour of u_1. Let $R_i(s)$ be the red vertices in level i at step s. Vertex u_1 was chosen u.a.r. so, as $\mu_0(T_f) \sim 1$ (see (29)), $\mu_1(T_f) = \Theta(1)n^{1/(k+1)}$ (see (30)), and $t_f \geq T_f/\omega$,

$$\mathbb{P}(u_1 \in R_1(t)) = \frac{Z_1(t)}{L_1(t)} \leq \frac{\omega}{\mu_1(t)} \leq \frac{\omega^2}{n^{1/(k+1)}}. \quad (32)$$

Let s_1 be the step at which $u_1 = u_1(s_1)$ was occupied by a halted particle. This happened because s_1 chose an occupied out-neighbour $u_2 = u_2(s_2)$, where $s_2 < s_1$. If u_2 is red at step t, then either it became red before step s_1, or at some later step. As u_2 was chosen u.a.r. by the particle s_1 at u_1 from among the $L_2(s_1)$ occupied vertices in level 2,

$$\mathbb{P}(u_2 \in R_2(s_1)) = \frac{Z_2(s_1)}{L_2(s_1)} \lesssim 2\omega\frac{\mu_1(s_1)}{\mu_2(s_1)} = \frac{2\omega s_1}{k(n/k)} = \frac{2\omega s_1}{n}.$$

The factor of 2 is from (31) and ω covers lack of concentration of $\widehat{Z}_2$ or L_2, if s_1 is early. If u_3 is the chosen out-neighbour of u_2,

$$\mathbb{P}(u_2 \in R_2(t)\setminus R_2(s_1)) \leq \frac{1}{(n/k)}\sum_{\tau=s_1+1}^{t} \mathbb{E}\ \mathbb{1}_{\{u_3 \text{ chosen at } \tau\}} = \frac{t-s_1}{(n/k)}.$$

Thus

$$\mathbb{P}(u_2 \in R_2(t)) \leq \frac{\omega k t}{n},$$

and similarly for $u_3, \ldots, u_{k-1}$. Thus

$$\mathbb{P}((v, u_k)\text{–path is blue}) \geq \prod_{i=1}^{k-1}\left(1 - \frac{\omega k t}{n}\right) = 1 - O\left(\frac{\omega k^2 t}{n}\right),$$

where $t \le t_f \le \omega n^{k/(k+1)}$. Provided $\omega k^2/n^{1/(k+1)} = o(1)$, which holds for $3k \log k \le n$, the path from the source to vertex $w = u_k$ level k is blue, and thus the in-arborescence root at w is exactly this path. □

4 Analysis of DLA in the growing layers model

Proposition 7. *Let G be a growing layers graph with level sizes $N_i = d^i$, for $i = 0, ..., k$. Let*

$$T_f = \sqrt{k}\ d^{k+3/2-\sqrt{2k+2}}. \tag{33}$$

Let $j = \max\{i : (2k+2) \ge i(i+1)\}$, and let $j^* = \sqrt{2k+2} - 1$ be the solution to $2k+2 = i(i+1)$.

1. *Provided $d \to \infty$, $k \to \infty$, and $j = j^*$ the finish time t_f satisfies $T_f/\omega \le t_f \le \omega T_f$, where $\omega \to \infty$ arbitrarily slowly.*
2. *If $j \ne j^*$ the error in finish time t_f above introduced by ignoring rounding is $d^{-O(1/\sqrt{k})}$, and thus $T_f d^{-O(1/\sqrt{k})}/\omega \le t_f \le \omega T_f$.*
3. *At t_f, levels $i = 1, \dots, k - \lceil \sqrt{2k+2} - 1 \rceil$ contain a single occupied vertex, the vertex u_i of the the connecting path.*

Proof. The size N_i of layer i is defined as d^i. It follows that the product of the set sizes in the denominator of $\mu_{k-j}(t)$ in (1) is given by

$$N_k N_{k-1} \cdots N_{k-j+1} = d^k d^{k-1} \cdots d^{k-j+1} = d^{kj-j(j-1)/2},$$

and thus (1) becomes

$$\mu_{k-j}(t) = \frac{t^{j+1}}{(j+1)!\ d^{kj-j(j-1)/2}}. \tag{34}$$

Assume d is sufficiently large. The upper bound $\mathbb{E}\,\widehat{L}_{k-j}$ is obtained in Section 2. However, a problem can arise in the growing layers model in the upper bound calculations. The value of μ_{k-j} can decrease with increasing j and then (anomalously) increase again. This is presumably because the recurrence used to establish it assumes $\mu_{k-\ell} \to \infty$ for all $\ell < j$, which is not the case.

Let t be such that $\mu_{k-j}(t) = 1$, and let $t_1 = t_1(k-j)$ be $\lceil t \rceil$, so that $\mu_{k-j}(t) \sim 1$ at step t_1. From (34),

$$\mu_{k-j}(t_1) = \frac{t_1^{j+1}}{(j+1)!\ d^{kj-j(j-1)/2}} \sim 1 \quad \Longrightarrow \quad t_1 \sim [(j+1)!]^{1/(j+1)}\ d^{\frac{2kj-j(j-1)}{2(j+1)}}. \tag{35}$$

From (1)

$$\frac{\mu_{k-j+1}(t)}{\mu_{k-j}(t)} = \frac{(j+1)d^{k-j+1}}{t},$$

so setting $\mu_{k-j}(t_1(k-j)) \sim 1$ gives

$$\begin{aligned} \mu_{k-j+1}(t_1) &\sim \frac{j+1}{[(j+1)!]^{1/j+1}} \; d^{k-j+1-\frac{2kj-j(j-1)}{2(j+1)}} \\ &\sim \frac{e}{(2\pi(j+1))^{1/2(j+1)}} \; d^{\frac{2k+2-j(j+1)}{2(j+1)}}. \end{aligned} \tag{36}$$

The leading term on the RHS is bounded, and the exponent of d on the RHS is positive provided $2k+2 > j(j+1)$, which ensures that $\mathbb{E}\,\widehat{L}_{k-j+1}(t)$ is sufficiently large close to $t_1(k-j)$.

What value of $k-j$ maximizes the step $t_1 = t_1(k-j)$ at which $\mu_{k-j}(t) \sim 1$? Write the exponent of d on the RHS of (35) as $f(j)/2$ where

$$f(j) = \frac{2kj - j(j-1)}{(j+1)} = (2k+2) - j - \frac{2k+2}{j+1}.$$

The maximum of $f(j)$ occurs when $(j^*+1)^2 = (2k+2)$, giving $f(j^*) = (j^*)^2$. The (not necessarily integer) value of j^*, $k-j^*$ and $d^{f(j^*)/2}$ are

$$j^* = \sqrt{2k+2} - 1, \qquad k - j^* = k + 1 - \sqrt{2k+2}, \qquad d^{f(j^*)/2} = d^{k+3/2-\sqrt{2k+2}}. \tag{37}$$

Integer solutions to $(j^*+1)^2 = (2k+2)$ exist for some values of k. Let a be a positive integer, and put $j^*+1 = 2a$ so that $k = 2a^2-1$; this holds for example for $k = 1, 5, 17, 31$. In the case that j^* is not integer, we address the rounding error at the end of this section.

Ignoring rounding for now, we evaluate $t_1 = t_1(k-j)$, at $j = j^*$ where $f(j) = j^2$, to find

$$\begin{aligned} t_1 &\sim [(j+1)!]^{1/(j+1)} \; d^{f(j)/2} \\ &\sim e^{-1}(\sqrt{2\pi})^{1/(j+1)} \; (j+1)^{1+1/2(j+1)} \;\; d^{j^2/2} \\ &= C_{k-j} \; \sqrt{2k+2} \;\; d^{k+3/2-\sqrt{2k+2}}, \end{aligned} \tag{38}$$

where $C_{k-j} = O(1)$. Note that $t_1 = \Theta(T_f)$. From (34),

$$\frac{\mu_{k-(j+1)}(t)}{\mu_{k-j}(t)} = \frac{t}{(j+2)d^{k-j}}$$

so that at t_1, for some C, $C' = O(1)$,

$$\mu_{k-j^*}(t_1) \sim 1, \qquad \mu_{k-(j^*-1)}(t_1) = Cd^{1/2}, \qquad \mu_{k-(j^*+1)}(t_1) = C'd^{1/2}.$$

At first this seems confusing, as one would expect to have $\mu_{k-(j^*-1)}(t_1) = Cd^{1/2} = o(1)$ here. But assuming $d^{1/2} \to \infty$, $\mathbb{E}\,\widehat{L}_{k-j^*+1} = C''d^{1/2}$, which is the last value at which the condition $\mu_{k-j+1} \to \infty$ is valid in the recurrence from $k-j+1$ to $k-j$; and is where the assumption in Proposition 3 breaks down.

Note that at t_1/ω, $\mu_{k-j^*}(t_1/\omega) = O(\omega^{-(j^*+1)})$ from the definition of $t_1(k-j^*)$. Thus w.h.p. $\widehat{L}_{k-j^*}(t_1/\omega) = o(1)$ in the upper process. Consequently all levels $k-\ell$, $\ell \geq j^*$, are zero, w.h.p. in both the upper and the DLA process. In which case, as occupancy is monotone non-decreasing, this holds in the state vector $L(t)$ of the DLA process, for all $t \leq t_1/\omega$ and all $\ell \geq j^*$.

For $\ell < j^*$, as in the equal layers model we will use (23) to obtain (27). To do this we check conditions (24) and (25) for $t \leq t_f$. Provided $k \to \infty$, but $k = o((\log n/\log\log n)^2)$,

$$\frac{k\omega^2 T_f}{N_k} = \frac{\omega^2 k^{3/2} d^{k+3/2-\sqrt{2k+2}}}{d^k} = \omega^2 k^{3/2} d^{3/2-\sqrt{2k+2}} = o(1); \tag{39}$$

and as $N_{k-j} = d^{k-j}$, and $j^*+1 = \sqrt{2k+2}$,

$$\frac{\mu_{k-\ell}(t)}{\mu_{k-\ell+1}(t)} = \frac{t}{(\ell+1)d^{k-\ell+1}} \leq \omega\sqrt{k}d^{1/2-\sqrt{2k+2}+\ell} \leq \omega\sqrt{k}d^{1/2-\sqrt{2k+2}+j^*-1} = \omega\sqrt{k}d^{-3/2},$$

provided $t \leq \omega T_f$. Thus provided $d > k\omega^2$,

$$\frac{\mu_{k-\ell}(t)}{N_{k-\ell}} \leq \frac{\omega\sqrt{k}}{d^{3/2}}\frac{\mu_{k-\ell+1}(t)}{N_{k-\ell}} = \frac{\omega\sqrt{k}}{d^{1/2}}\frac{\mu_{k-\ell+1}(t)}{N_{k-\ell+1}} \leq \frac{\mu_{k-\ell+1}(t)}{N_{k-\ell+1}}.$$

As $\mu_k(t) = t$, it follows that

$$\mu_{k-j}(t)/N_{k-j} \leq t/N_k, \tag{40}$$

and thus (26) holds. As before, let $t_i(\omega)$ be given by (12). For those $i \leq k$, and $t \geq t_i(\omega)$, so that $\mu_i(t) \to \infty$ suitably fast, we obtain that $\mathbb{E}\,\widetilde{L}_i(t) \sim \mu_i(t) \sim \mathbb{E}\,\widehat{L}_i(t)$, and we have $L_{k-\ell}(t) \sim \widehat{L}_{k-\ell}(t) \sim \mu_{k-\ell}(t)$.

Let t_0 be the first step at which $L_{k-j^*}(t) = 1$. Then either $t_0 \leq t_1$, or, as the probability a particle halts at level $k-j^*$ is

$$\phi = \frac{L_{k-(j^*-1)}(t_1)}{N_{k-(j^*-1)}} \sim \frac{Cd^{1/2}}{d^{k-j^*+1}} = \frac{C}{d^{k+3/2-\sqrt{2k+2}}};$$

the probability this does not occur in a further t_1 steps is, see (38),

$$(1-\phi)^{t_1} \leq e^{-t_1\phi} = e^{-(1+o(1))C\sqrt{2k+2}} = o(1),$$

assuming $k \to \infty$. As $\widehat{L}_{k-j^*}(\max(t_0, t_1)) \leq \omega$, w.h.p.; we assume that $L_{k-j^*}(\max(t_0, t_1)) \leq \omega$, and for $\ell > j^*$, $L_{k-\ell} = 0$.

Let u be the vertex in level $k-j^*$ containing the unique particle halted at t_0. Construct a path back from u to the source as follows. Wait until a particle halts at w_{k-j^*-1} in level $k-j^*-1$ by choosing edge $w_{k-j^*-1}u$. The expected time for this is d^{k-j^*}. In a further expected time d^{k-j^*-1}, the path will extend backwards, as a particle will halt in level $k-j^*-2$ by choosing edge to w_{k-j^*-1} etc. Thus in a further

$$T = d^{k-j^*} + d^{k-j^*-1} + \cdots + d = d^{k-j^*}\left(\frac{1-1/d^{k-j^*}}{1-1/d}\right) = \Theta(d^{k-j^*}) = \Theta(d^{k+1-\sqrt{2k+2}})$$

expected steps there will be a path $vw_1 \cdots w_{k-j^*-1}u$ of halted particles extending from the source v to vertex u thus stopping the DLA process (if it has not already halted). This path should be unique, as the expected time for it to branch backwards at any level i is $d^i \gg d^{i-1}$ if $d \to \infty$.

On the other hand the time to create another halted particle in level $k-j^*$ has expected value

$$\frac{1}{\phi} = \Theta(d^{k+3/2-\sqrt{2k+2}}) = \Theta(Td^{1/2}).$$

On the assumption that $d \to \infty$, in at most $\omega(t_1 + T)$ steps the process has halted w.h.p., as claimed.

The effect of rounding error in j^*. In the case that j^* is not integer, we require the maximum j such that $2k+2 > j(j+1)$. Clearly $j = \lfloor j^* \rfloor$ satisfies $2k+2 > j(j+1)$, but what about $j = \lceil j^* \rceil$? Put $j = j^* + \varepsilon$. Further analysis, not given here, shows that the condition $2k+2 \geq j(j+1)$, is satisfied by $j = \lceil j^* \rceil$ up to some $\varepsilon \in (1/2, 1)$.

Let $j = \max\{i : (2k+2) \geq i(i+1)\}$ and suppose that $j = \lfloor j^* \rfloor$ so that $j^* = j + \varepsilon$. Let $T_M = \Theta(t_1(k-j))$ be given by

$$T_M = \sqrt{k} d^{\frac{1}{2}(2k+2-j-(2k+2)/(j+1))},$$

be a revised estimate of the order of the halting time, where $T_M \leq T_f$ as j^* maximizes T_f. As $j^* = \sqrt{2k+2} - 1$, $T_f = \sqrt{k} d^{(j^{*2}/2)}$, see (37), and $2k+2-j^*-(j^*)^2 = \sqrt{2k+2}$,

$$\begin{aligned}
\frac{T_M}{T_f} &= d^{\frac{1}{2}\left((2k+2)-j^*+\varepsilon-\frac{2k+2}{j^*+1-\varepsilon}-(j^*)^2\right)} \\
&= d^{\frac{1}{2}\left(\sqrt{2k+2}-\frac{\sqrt{2k+2}}{1-\varepsilon/(\sqrt{2k+2})}+\varepsilon\right)} \\
&= d^{-\frac{\varepsilon^2}{2\sqrt{2k+2}}(1+O(1/\sqrt{k}))}.
\end{aligned}$$

Choosing $j = \lceil j^* \rceil = j^* + \varepsilon$ gives the same result. Thus the effect of rounding is to alter T_f by $\Theta(1)d^{-O(1/\sqrt{k})}$. □

Existence of a unique connecting path component. Finally, we prove that the arborescence rooted at level k containing the connecting path from source to sink, consists uniquely of that path. The proof is similar to Lemma 6 for the equal layers model. At t_f, w.h.p. there is a unique path from level $k - j^*$ to level zero, so that level $k - j^* + 1$ plays the role of level one. By analogy with (30),

$$\mathbb{P}(u_{k-j^*+1} \in R_{k-j^*+1}(t)) = \frac{Z_{k-j^*+1}(t)}{L_{k-j^*+1}(t)} \leq \frac{\omega}{\mu_{k-j^*+1}(t)} \leq O\left(\frac{\omega^2}{d^{1/2}}\right),$$

where we used an earlier result that $\mu_{k-(j^*-1)}(t_1) = Cd^{1/2}$.

Thus as $t_f \leq \omega t_1(j^*)$, where $j^* = \sqrt{2k+2} - 1$ and $t_1(k - j^*)$ is given by (38)

$$\mathbb{P}((v, u_k)\text{–path is blue}) \geq \left(1 - \frac{\omega^2}{d^{1/2}}\right) \prod_{j=1}^{j^*-2} \left(1 - \frac{\omega t_f}{d^{k-j}}\right) = 1 - O\left(\frac{\omega^2}{d^{1/2}}\right) - O\left(\frac{\omega\sqrt{k}}{d^{3/2}}\right).$$

5 DLA on trees with large branching factor

Let G be a labelled tree with branching factor d and final level k, such that $d^k = n$. As before, the source of the particles is the unique vertex v at level zero, an artificial sink vertex z at level $k+1$ is attached to the vertices at level k, and S_i is the set of vertices in level i.

Let $j^* = \sqrt{2k+2} - 1$, and let T_f be given by

$$T_f = \sqrt{k}\, d^{k+3/2-\sqrt{2k+2}}. \tag{41}$$

The average number of particles arriving at a vertex x at level $k - j^*$ by step T_f is order $\sqrt{kd}$, and the distance j^* to level k is less than this. However, the average number of particles passing through a vertex w at level $k - j^* + 1$ by step T_f is

$$\frac{T_f}{d^{k-j}} = \sqrt{\frac{k}{d}} = o(1), \tag{42}$$

provided $k \ll d$, which we assume to be true. Thus the event that j^* particles have arrived at w is an upper tail event. There should be few vertices w with a path to level k consisting of j^* such halted particles. We assume henceforth that $j \leq j^* - 1$, and look for such paths.

Let $X_w = X_w(k-j, \ell, t)$ be the indicator for the event that exactly ℓ particles reach vertex w in level $k - j$ at or before step t. On the assumption that all levels $i < k - j$ are unoccupied at t,

$$\mathbb{E}\, X_w = \binom{t}{\ell} \left(\frac{1}{d^{k-j}}\right)^{\ell} \left(1 - \frac{1}{d^{k-j}}\right)^{t-\ell}. \tag{43}$$

Let $X=\sum_{w\in S_{k-j}} X_w$ be the number of vertices at level $k-j$ which have exactly $\ell=j+1$ particles reaching them. Thus $\mathbb{E}\,X=d^{k-j}\mathbb{E}\,X_w$, and

$$\begin{aligned}\mathbb{E}\,X(X-1) =&(d^{k-j})(d^{k-j}-1)\binom{t}{2\ell}\binom{2\ell}{\ell}\left(\frac{1}{d^{k-j}}\right)^{\ell}\left(\frac{1}{d^{k-j}-1}\right)^{\ell}\left(1-\frac{2}{d^{k-j}}\right)^{t-2\ell}\\ =&\left(1+O\left(\frac{\ell}{d^{k-j}}\right)\right)(\mathbb{E}\,X)^2.\end{aligned}$$

Thus, as $X\le t_f/\ell$ deterministically,

$$\mathbf{Var}X=\mathbb{E}\,X+O\left(\frac{\ell}{d^{k-j}}\right)(\mathbb{E}\,X)^2=(1+O(\sqrt{k/d})\,\mathbb{E}\,X=(1+o(1))\mathbb{E}\,X. \tag{44}$$

Let $\ell=j+1$. The probability the particles arriving at w form a path $w=w_0w_1w_2\cdots w_j=u$ of halted particles from level $k-j$ to level k is obtained as follows. Let $1\le s_0\le s_1\le s_2\le\cdots\le s_j\le t$ be the steps (and particle labels) of the halted particles forming this path. Thus s_0 reaches $w_j=u$ an event of probability $1/d^j$, s_1 reaches w_{j-1} and picks u as a neighbour, an event also of probability $1/d^j$ and so on, and finally s_j picks w_1 as a neighbour and halts at w, an event of probability $1/d$. Thus

$$\mathbb{P}(\text{path } w_0w_1\cdots w_j\mid (j+1)\text{ arrivals at }w)=\frac{1}{d^j}\frac{1}{dd^2\cdots d^j}.$$

We call such a path *exact* if the halted particles on the path are the only particles in the sub-tree from level $k-j$ to level k rooted at w. Let Y_w be the number of exact paths from w to level k at step t. As there are d^j choices for u given w, we have

$$\begin{aligned}\mathbb{E}\,Y_w =&d^j\frac{1}{d^j}\frac{1}{dd^2\cdots d^j}\binom{t}{j+1}\left(\frac{1}{d^{k-j}}\right)^{j+1}\left(1-\frac{1}{d^{k-j}}\right)^{t-(j+1)}\\ =&(1-O(j^2/t))\,\frac{1}{d^k\cdots d^{k-j}}\frac{t^{j+1}}{(j+1)!}e^{-t/d^{k-j}} \qquad (45)\\ \sim&\frac{t^{j+1}}{(j+1)!}\frac{1}{d^k\cdots d^{k-j}}=\frac{\mu_{k-j}(t)}{d^{k-j}}, \qquad (46)\end{aligned}$$

where $\mu_{k-j}(t)$ is given by (34), and we used the assumption (42) that $j\le j^*-1$. Let $Y(k-j,j+1,t)=\sum Y_w$ be the number of exact paths from level $k-j$ to k at step t. There are d^{k-j} choices for w, so $\mathbb{E}\,Y\sim\mu_{k-j}(t)$.

The value $t_1(k-j^*)$ from (38) satisfies $t_1=\Theta(T_f)$ where T_f is from (33). Using (36) and (37) we obtain that

$$\mu_{k-j^*+1}(T_f)=\Theta(d^{1/2}),$$

so $\mathbb{E}\, Y(k-j^*+1, j+1, t_1(k-j^*) \sim \mu_{k-j^*+1}(t_1) = \Theta(d^{1/2})$.

Comparing (43) and (46), we see that $\mathbb{E}\, Y = \mathbb{E}\, X/d^{j(j+1)/2}$, and thus $\mathbb{E}\, X = \Theta(d^{1/2+j(j+1)/2})$. By the Chebychev inequality, w.h.p. $X \sim \mathbb{E}\, X(k-j, j+1, t)$, see (44). The events of exact path construction at vertices w, w' of level $k-j$ are independent, given exactly $j+1$ particles arrive at each of w, w'. As $\mathbb{E}\, Y(k-j, j+1, t) = \Theta(d^{1/2})$ where $d \to \infty$, then, for some small $\varepsilon > 0$, w.h.p. $(1-\varepsilon)\mathbb{E}\, Y \le Y \le (1+\varepsilon)\mathbb{E}\, Y$. Conditional on this, $Y = \Theta(d^{1/2})$.

Next consider that $\ell \ge j+2$ particles reach vertex w, so that the path is not exact, as at least one other halted particle occupies the sub-tree rooted at w. Let $Q(k-j, \ell, t)$ be the number of these events. Then, provided (42) holds,

$$\sum_{\ell \ge j+2} \mathbb{E}\, Q(k-j, \ell, t) = \Theta(1)\frac{t}{d^{k-j}}\mathbb{E}\, Y = O(1)\sqrt{\frac{k}{d}}\, \mathbb{E}\, Y = o(\mathbb{E}\, Y).$$

From Section 4, at time $t_1(k-j^*)/\omega$, $\mu_{k-j^*}(t_1(k-j^*)/\omega) = (1/\omega)^{j^*+1}$ and we assume levels $1, ..., k-j^*$ are empty. At t_1 there are $\Theta(d^{1/2})$ paths from $k-j$ to level k, of which $o(d^{1/2})$ are not exact. Either one of these paths has extended back to some x in $k-j^*$ or level $k-j^*$ is currently unoccupied. In the latter case, any particles which passed via x in between times t_1/ω and t_1 must have chosen some out neighbour of x other than w. The probability a given particle extends some exact path from level $k-j^*+1$ to $k-j^*$ is order $d^{1/2}/d^{k-j^*+1}$. Let $T(\text{hit})$ be the time for the first such extension to happen. Then

$$\mathbb{E}\, T(\text{hit}) = d^{k-j^*+1/2} = d^{k+3/2-\sqrt{2k+2}} = O(T_f).$$

By analogy with the analysis in Section 4, w.h.p. this selected exact path of occupied vertices from $k-j^*$ to k will extend back to the source v, in

$$T = O(d^{k-j^*}) = O(d^{k+1-\sqrt{2k+2}}) = O\left(\frac{\mathbb{E}\, T(\text{hit})}{\sqrt{d}}\right)$$

expected steps; whereas the expected wait for a second exact path to extend is $\mathbb{E}\, T(\text{hit})$. This completes the proof of Theorem 1 for trees.

5.1 An upper bound on finish time for trees with any $d \ge 2$

Proposition 8. *Let*

$$T = d^{k+3/2-\sqrt{2k+2}+1/2\log_d k}.$$

Then $\omega d^{1/2} T$ is a w.h.p. upper bound on the finish time of DLA in any tree with branching factor $d \ge 2$, and height k.

Let j be such that $d^{k-j} \geq T$ so that $e^{-T/d^{k-j}} = \Theta(1)$ in the calculation of $\mathbb{E}\, Y$ (see (45)). Let h be integer such that

$$k - j = k - j^* + h \geq k + 1 - \sqrt{2k+2} + h \geq k + 3/2 - \sqrt{2k+2} + (1/2)\log_d k$$

so that $h = \lceil 1/2 + 1/2 \log_d k \rceil$. In Section 5 we assumed that $k \ll d$ so that $h = 1$. At $T \sim t_1(k - j^*)$ from the growing layers model

$$\frac{\mu_{k-j^*+h}}{\mu_{k-j^*}} = \Theta(1) 2^{h/2} d^{h^2/2}.$$

This ratio is $\omega(1)$, as either $d \to \infty$ or if d is constant, then $\log_d k = \log_d \log_d n$. Either some path of halted particles already extends to a level i where $i < k - j^* + h$, or all vertices in these levels are unoccupied, and w.h.p. there exists a (not necessarily exact) path of length $j^* - h + 1$. Pick the longest path $P(w)$, where w in level $k - j^* + \ell$ is the final vertex on this path. It takes at most $\Theta(d^{k-j^*+h})$ expected steps to extend $P(w)$ back to the source. Failing this either the path from w to v is now blocked internally in this time by some longer path, or some other disjoint path P' has become longer. Either way, w.h.p. the process has halted in a further $\omega d^{k-j*+h} = \omega d^{1/2} T$ steps.

References

[1] A. Asselah and A. Gaudilliere. Sub-logarithmic fluctuations for internal DLA. The Annals of Probability, 41(3), 1160-1179 (2013).

[2] R. C. Ball, D. A. Weitz, T. A. Witten, and F. Leyvraz. Universal kinetics in reaction-limited aggregation, Phys. Rev. Lett. 58, 274, (1987).

[3] P. Diaconis and W. Fulton. A growth model, a game, an algebra, lagrange inversion, and characteristic classes. Rend. Sem. Mat. Univ. Pol. Torino, 49(1):95–119, (1991).

[4] A. M. Frieze and W. Pegden. Diffusion limited aggregation on the Boolean lattice, Annals of Applied Probability Vol. 28, No. 6 , 3528–3557, (2018).

[5] D. Jerison, L. Levine, and S. Sheffield. Logarithmic fluctuations for Internal DLA. Journal of the American Mathematical Society Vol. 25(1), pp. 271-301 (2012)

[6] H. Kesten. How long are the arms in DLA? Journal of Physics A: Mathematical and General, 20(1), L29–L33, (1987).

[7] G. F. Lawler, M. Bramson and D. Griffeath. Internal Diffusion Limited Aggregation, Ann. Probab. 20 (4) 2117–2140, (1992).

[8] G. F. Lawler. Subdiffusive fluctuations for internal diffusion limited aggregation. Ann. Probab., 23(1):71–86, (1995).

[9] P. Meakin and F. Family. Structure and dynamics of reaction-limited aggregation, Phys. Rev. A 36, 5498(R), (1987).

[10] T. A. Witten and L. M. Sander. Diffusion-limited aggregation, a kinetic critical phenomenon, Physics Review Letters, Vol. 47, No. 19, 1400–1403, (1981).

[11] T. A. Witten and L. M. Sander. Diffusion-limited aggregation, Phys. Rev. B 27, 5686–5697, (1983).