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.

October 19, 2024

Abstract

In the classical model of Diffusion Limited Aggregation (DLA), introduced by Witten and Sander, the process begins with a single particle cluster placed at the origin of a space. Then, one at a time, particles make a random walk from infinity until they halt by colliding with the existing cluster.

We consider an analogous version of this process on large but finite graphs with a designated source and sink vertex. Initially the cluster of halted particles contains a single particle at the sink vertex. Starting one at a time from the source, each particle makes a random walk in the direction of the sink vertex. The particle halts at the last unoccupied vertex before the walk enters the cluster for the first time, thus increasing the size of the cluster. This continues until the source vertex becomes occupied, at which point the process ends.

We study this DLA process on several classes of layered graphs, including Cayley trees of branching factor at least two with a sink vertex attached to the leaves. We determine the finish time of the process for the given classes of graphs and show that the subcomponent of the final cluster linking source to sink is essentially a unique path.

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

1 Introduction

Diffusion limited aggregation. In the classical model of Diffusion Limited Aggregation (DLA), introduced by Witten and Sander [12], [13], 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 the d–dimensional lattice $\mathbb{Z}^d$. Simulations of DLA in two dimensions show tree-like figures with long branches. For $\mathbb{Z}^d$, Kesten [8] proved that when the cluster size is N, the length of these arms is almost surely upper bounded by $CN^{2/3}$, when $d = 2$, and by $C_d N^{2/d}$ when $d \geq 3$.

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 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. Although the DLA and IDLA processes differ, the point in common is that they both describe the evolution of a unique cluster by adhesion to the cluster boundary. As with DLA, 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 [9], proved that, on d-dimensional lattices the limiting shape approaches a Euclidean ball; a result subsequently refined in [1], [7] and [10], amongst others.

DLA has been proposed as a model of physical processes in systems as diverse as coagulated aerosols [12] and urban growth [2], a common factor being the dendritic shape of the cluster so obtained (see [5] for illustrative examples). However the main and an original motivation for the DLA process was as a model of dendritic growth in dielectric breakdown and lightning formation. Niemeyer, Pietronero, and Weismann, [11], introduced a dielectric breakdown model which considers DLA in the presence of an electric field which biasses the particles to move in a given direction (e.g. downward). In established models of lightning formation, paths of negatively charged particles (leaders) descend downward from the base of the cloud layer, and paths of positively charged particles percolate upwards from the ground towards them, inducing a lightning strike on meeting. The DLA process was seen as a first approximation of this process, which is itself finite in extent with a source at the top (cloud layer) and a sink (the earth) at the bottom.

We consider a version of the DLA process on large but finite graphs with a designated source and sink vertex. DLA on finite graphs was previously studied for complete binary trees by Hastings and Halsey [6], and for the Boolean lattice (hypercube) by Frieze and Pegden [4]. The current paper continues this analysis of DLA on finite layered graphs, of which the binary tree and hypercube are typical examples.

The layers model. Let G be a finite graph $G = (V, E)$, whose vertices can be partitioned into sets $S_0, S_1, \ldots, S_k, S_{k+1}$ to form a *layered structure* in which all edges are between layer S_i and S_{i+1} $(i = 0, 1, ..., k)$. The sets S_0, S_{k+1} are of size one with $S_0 = \{v\}$ and $S_{k+1} = \{z\}$. The vertex v is the *source vertex* and the vertex z is the *sink vertex*. In certain cases the source and sink may be attached as extra (*artificial*) vertices to an existing graph G to complete the layered structure.

Examples of layered graphs with a symmetric structure include the following: The Boolean lattice (hypercube) $\mathcal{B} = \{0, 1\}^n$, with $v = \mathbf{1}$, $z = \mathbf{0}$. Two dimensional square grids with source the top left hand corner and sink the bottom right hand corner, thus inducing a diagonal orientation. Finite Cayley trees with source the root and an artificial sink. Layered multipartite graphs with attached source and sink.

Graphs in the layers model have a top (the source) and a bottom (the sink). Particles are constrained to move downwards. As such they can be seen as simple models of particles percolating downward through some porous medium.

DLA in the layers model. Initially at step $t = 0$, only the sink vertex z is occupied, and the occupied component is $C_0 = \{z\}$. At a given step t, let ρ_t be a random walk on the underlying graph starting at the source v and moving forward level by level to the sink z. A particle placed on the source vertex (assumed unoccupied) follows the walk ρ_t until it encounters an occupied vertex. Let the path followed by the walk ρ_t be $v = x_0 x_1 x_2, ..., x_k x_{k+1} = z$, and let x_i, $(i \geq 1)$ be the first occupied vertex on the walk. The particle halts at position x_{i-1} and permanently occupies that vertex. The component C_t is formed by adding the vertex x_{i-1} and directed edge $x_{i-1}x_i$ to the component C_{t-1} thus extending the directed tree formed by the occupied vertices and rooted the sink z. As the sink is occupied from the start, any particle which reaches level k automatically halts there. The process ends at a step t_f, *the finish time*, when the source vertex v first becomes occupied by a halted particle. Thus t_f is the final number of particles occupying the graph (not including the sink).

The models of this paper. Let $N_i = |S_i|$ be the size of the set S_i, the i–th layer of the graph G. We regard all edges as directed from the source towards the sink. We consider two models.

- The *Cayley tree* $G(d, k)$ with branching factor d and height k. For convenience we take the size of the last level to be $d^k = n$. Here $d \geq 2$ is fixed or tends to infinity with n sufficiently slowly, so that $k = \log n / \log d$ also tends to infinity with n. The source v is the root vertex at level zero of the tree and the sink z is an artificial vertex connected to all vertices in the final layer S_k of the tree. Excluding the sink, $G(d, k)$ is a $(d^{k+1} - 1)/(d - 1)$–vertex graph.

- The *multipartite layers model.* The sets S_0, S_{k+1} have size one, the layers S_i have size N_i. For $i = 0, .., k$ there is a complete bipartite graph between S_i and S_{i+1}. Here $k \geq 1$ is fixed or tends to infinity with n.

  The *equal (multipartite) layers model* is a graph $M_E(n, k)$ in which the sets of S_i $(i = 1, \ldots, k)$ are the same size $N_i = n/k$, where $N_i \to \infty$. Excluding the source and sink, the equal layers model G is an n–vertex graph. The parameter k can either be a fixed integer $k \geq 1$, or a function of n. In the extremes the graph has one layer, S_1, of n vertices $(k = 1)$, or is a path $(k = n)$.

  The *growing layers model* is a graph $M_G(d, k)$ in which the sets S_i grow geometrically in size with parameter d. As $|S_0| = 1$, then $|S_i| = N_i = d^i$, and we take $d^k = n$. Here $d, k \to \infty$ with n. Excluding the sink, the growing layers model G is a $(d^{k+1}-1)/(d-1)$–vertex graph.

Analysis of the DLA process is in terms of n, which is, up to a constant multiple, the number of vertices in the graph. The model is probabilistic and corrections arising from the exceptional events are estimated as a function of n, even if this is not always made explicit. The parameter k determines the number of levels in the graph, and d the growth rate (if any) of the levels. The value of k or d can be constant in some models or it can tend to infinity as a function of n within some bounds. As $d^k = n$ in the growing layers model, they are related by $k = \log n / \log d$.

Notation. We say a sequence of events $\mathcal{E}_n$, $n \geq 0$ occurs with high probability (w.h.p.) if $\mathbb{P}(\mathcal{E}_n) = 1 - \varepsilon_n$ where $\varepsilon_n = o_n(1)$ for some function $o_n(1) \to 0$ with n, and thus $\lim_{n\to\infty} \mathbb{P}(\mathcal{E}_n) = 1$. Typically we write $o_n(1) = o(1)$ and similarly for other asymptotic notation such as $O(\cdot), \Theta(\cdot), \Omega(\cdot)$. 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 $\omega = \omega_n$ in two ways; either to denote any quantity ω_n which tends to infinity with n but suitably slowly as required in the given proof context, or as a fixed divergent quantity whose value is stated explicitly. The expression $f(n) \ll g(n)$ indicates $f(n)/g(n) = o_n(1)$. The notation $f \to \infty$ indicates that $f = f(n)$ grows unboundedly with increasing n. All results claimed are for sufficiently large n.

Results for the multipartite layers model. At any step t, the occupied component C_t is a tree with edges directed downwards towards the sink z. At the finish time t_f the source becomes occupied, and the component C_{t_f} contains a directed *connecting path* from the source v to the sink z, all of whose vertices contain halted particles.

On deletion of the sink vertex, the digraph $D_t = C_t \backslash \{z\}$ consists of a directed forest with components rooted at the occupied vertices of level k. 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. In the multipartite layers model w.h.p. the tree component rooted at vertex u_k containing the connecting path is precisely the path $v = u_0 u_1 \cdots u_k$. Thus this path grew back to the source without gaining any off-path neighbours due to other particles colliding with it.

Theorem 1. *Let $\omega \to \infty$ slowly with n. The following results hold with high probability in the multipartite layers model.*

(1) Equal layers model, $M_E(n,k)$. For $i = 1, ..., k$ let $N_i = n/k$. Let

$$T_f = [(k+1)!(n/k)^k]^{1/(k+1)} = \gamma_k \, n^{k/(k+1)}$$

where $\gamma_k > 0$ constant, and $\gamma_k \to 1/e$ as $k \to \infty$.

Provided $2 \le k \le \sqrt{\log n}/\log\log n$, the finish time t_f satisfies $T_f/\omega \le t_f \le \omega T_f$.

(2) Growing layers model, $M_G(d,k)$. For $i = 0, ..., k$, let $N_i = d^i$, where $d^k = n$ and thus $k = \log n/\log d$. Let

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

(a) Provided $d \to \infty$, $k \to \infty$ with n, and $k \le \log^2 d$, the finish time t_f satisfies $T_f/\omega \le t_f \le \omega T_f$.

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

(3) Let $v = u_0 u_1 \cdots u_k z$ be the path connecting source v and sink z at t_f, and let $D_{t_f} = C_{t_f} \setminus \{z\}$ be the occupied component with the sink deleted. In either model, with high probability, the vertices $\{u_1, \cdots, u_k\}$ have in-degree one in C_{t_f}. Thus the tree rooted at u_k in D_{t_f} is exactly this path.

Results for Cayley trees. For a Cayley tree $G = G(k,d)$, the connecting path is by definition the unique path from source to sink. We give values for the finish time both for $d \to \infty$, and for d finite. By an indicative argument Hastings and Halsey [6] derived an estimate of $\sqrt{2k}\, 2^{k-\sqrt{2k}}$ for the expected finish time of DLA on the binary tree $G(2,k)$ of height k. We confirm their estimate is of the correct order of magnitude, and give results for $G(d,k)$ for any constant d.

Theorem 2. *With high probability the finish time t_f of DLA on a Cayley tree of branching factor d and height k satisfies $T_f/\omega \le t_f \le \omega T_f$, where T_f is as given below, and where $\omega \to \infty$ slowly with n.*

(1) If $d \to \infty$, $k \to \infty$, and $d \ll k$, then $T_f = \sqrt{k}\; d^{k+3/2-\sqrt{2k+2}}$.

(2) If $d \ge 2$ constant, then $T_f = \sqrt{k}\, d^{k-\sqrt{2k}}$.

Comparison of occupancy. How to compare the results of DLA on different models with each other? One possibility is to define the *packing ratio* $\rho = N/|V|$ of a process as the number of particles N at the finish divided by the number of vertices in the graph. Ignoring constants and lower order terms, $\rho = n^{-1/(k+1)}$ for the equal layers model and $\rho = n^{-\sqrt{2/k}}$ for the growing layers model.

How can the packing ratio be used to compare results for DLA on infinite graphs such as $\mathbb{Z}^d$? The sink in $\mathbb{Z}^d$ is the origin $\mathbf{0}$, as this is where the initial particle is located. At any fixed moment the longest arm of the figure formed by the occupied vertices defines a path back to the edge of the current bounding figure; which we take as the source (of particles entering from outside).

For the two-dimensional grid it is a result of Kesten [8], that, when N particles are added, the longest arm of the DLA figure is upper bounded by $C_2 N^{2/3}$, for some constant C_2. A circle of radius $N^{2/3}$ contains order $n = N^{4/3}$ vertices in $\mathbb{Z}^2$, so $N = \Theta(n^{3/4})$. Ignoring constants, this gives $\rho \leq n^{-1/4}$ for $\mathbb{Z}^2$, at any point in the (infinite) process. For $d \geq 3$, the longest arm length in $\mathbb{Z}^d$ is $C_d N^{2/d}$. A figure of radius $R = N^{2/d}$ has order $n = R^d = N^2$ vertices, so $\rho = n^{-1/2}$.

Road map of proofs. The following is an outline description and we often suppress minor details and qualifiers such as w.h.p.

The first step is to derive a solution to the recurrence (1)–(2) for the expected occupancy of the levels in the multipartite layers model. Under certain assumptions, the solution is asymptotic to (3), and the actual occupancy is concentrated around this value. This is particularly true in the equal layers model where the levels fill up in a more or less regular manner as given by (3), allowing us to estimate the finish time.

Things differ somewhat for the growing layers model. Although the higher levels fill in a regular manner, as given by (3), eventually we reach a level (with a well defined index $k-j^*$) where, in expectation, only a single occupancy occurs. The expected waiting time for further occupancy of this level is much longer than the expected waiting time for a unique path of halted particles to grow backwards to the origin, thus terminating the process. This is the content of Theorem 1.(2).(b). On the other hand Theorem 1.(3) (which requires a separate proof) says something stronger. Namely that the connecting path grows back from level k to the source as a unique path, and without gaining any off-path neighbours due to other particles colliding with it.

Because of the similarity in which the layers grow, the results for the growing layers model tell us the likely behaviour of the Cayley tree model. This allow us to construct proofs for the finish time for Cayley trees.

2 Bounds on occupancy in the layers model

Recall that N_i is the size of layer i for $i = 0, 1, ..., k$, where the value of N_i depends on the model in question. 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 $L_0(t) = 0$ for $t < t_f$, $L_k(t) \leq t$, and generally $L_i(t) \leq \min(t, N_i)$.

General formulation of layers occupancy. Let $\mathcal{H}(t) = (L_0(t), L_1(t), \ldots, L_k(t))$ be the occupancy vector of the DLA process at step t. Then,

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

$$\mathbb{E}\left(L_k(t+1) \mid \mathcal{H}(t)\right) = \; L_k(t) + \prod_{j=0}^{k} \left(1 - \frac{L_j(t)}{N_j}\right). \tag{2}$$

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

The following proposition gives the solution to these recurrences under suitable conditions.

Proposition 3. *For $t \geq 0$ let $\mu_k(t) = t$, and for $1 \leq j \leq k-1$, let*

$$\mu_{k-j}(t) = \frac{1}{N_k N_{k-1} \cdots N_{k-j+1}} \frac{t^{j+1}}{(j+1)!}. \tag{3}$$

For $j \geq 0$, suppose there are steps $T_k < \cdots < T_{k-\ell} < \cdots < T_{k-j}$, such that for all $\ell \leq j$, at $T_{k-\ell}$ the value of $\mu_{k-\ell}(T_{k-\ell}) \to \infty$ sufficiently fast, and for all levels $i < k - \ell$ the value of $\mu_i(T_{k-\ell}) \to 0$ sufficiently fast. Then with high probability for all $\ell \leq j$ and $T_{k-\ell} \leq t \leq t_f$ we have $L_{k-\ell}(t) \sim \mu_{k-\ell}(t)$ in either of the multipartite layers models.

The condition for the existence of T_{k-j} is satisfied for $0 \leq j \leq k-1$ in the equal layers model and for $0 \leq j \leq \sqrt{2k+2} - 2$ in the growing layers model.

The proposition describes a gap property, that when level $k - j$ starts to fill and become concentrated, the levels with lower index $i = 0, 1, ..., k - (j + 1)$ remain empty. The proof of Proposition 3 is inductive backwards from level k. The precise growth of the levels, and the values of j which make it work are to be determined. The first steps are common to the equal and growing layers models, and we include them together in this section. For the equal layers model we complete the proof of Proposition 3 in Sections 3.1 and 3.2. The analogous proof for the growing layers model is given in Section 4.

2.1 Upper bound on occupancy at step t

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

The upper-blocked process. Let $B_i(t)$ denote the occupied (blocked) vertices in level i in the DLA process at the end of step 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, $0 \leq j \leq k$ on the walk ρ_{t+1}, if u_{j+1} is occupied ($u_{j+1} \in \widehat{B}_{j+1}(t)$) 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 $u'_j \in S_j \backslash \widehat{B}_j(t)$ to $\widehat{B}_j(t+1)$. If a layer becomes full we continue with the layer above. As we will prove, this contingency will not occur, as with high probability the first layer to become full is the source.

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 that $L_i(t) \leq \widehat{L}_i(t) \leq t$, as ρ_{t+1} can add at most one vertex to $\widehat{B}_i(t)$. Moreover $\widehat{L}_k(t) = t$ deterministically (provided $t \leq N_k$). As a consequence the finish time $t_f(UB)$ of the upper blocked process satisfies $t_f(UB) \leq t_f(DLA)$.

Let $\widehat{\mathcal{H}}(t) = (\widehat{L}_0(t), \widehat{L}_1(t), \ldots, \widehat{L}_k(t))$ be the occupancy vector of the upper-blocked process at step t. For $0 \leq i \leq k$, the expectation $\mathbb{E}\,\widehat{L}_i(t)$ satisfies the recurrence

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

where $\mathcal{E}(t)$ is the event that $\widehat{L}_i(t) < N_i$ and that $\widehat{L}_{i+1}(t) < N_{i+1}$ for $i < k$. As we only propose to analyse the process as long as no level is full, we assume $\mathbb{1}_{\{\mathcal{E}(t)\}} = 1$ forthwith. Equation (4) follows because the upper blocked process increases the size of $\widehat{B}_j(t)$ (if possible) whenever the walk ρ_t contains a vertex of $\widehat{B}_{j+1}(t)$, this being true at all levels $j = 0, ..., k$.

The evolution of $\widehat{\mathcal{H}}(t) = (\widehat{L}_0(t), \widehat{L}_1(t), \ldots, \widehat{L}_k(t))$ is Markovian, and for $t \leq t_f$ we henceforth assume for $i \geq 1$ that $\widehat{L}_i(t) < N_i$ in our calculations. at $t_f \leq \omega T_f$. If so, referring to (1) and (2), we have

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

The next lemma gives w.h.p. bounds for $\widehat{L}_i(t)$ when none of the layers $i = 1, \ldots, k$ are full. With high probability the source is the only layer to become full in either the upper blocked and DLA process at or before t_f. When $L_0(t) = 1$ at $t = t_f$ the DLA process stops anyway.

The proofs of level occupancy are inductive backwards from level k. For a given level $k-j$ we identify two (not necessarily integer) times, $t_1(k-j)$ and $t_{k-j}(\omega)$, defined as follows. For $j \geq 0$, let $t_1(k-j)$ be the solution to $\mu_{k-j}(t) = 1$. Thus as $\mu_k(t) = t$, $t_1(k) = 1$ and

$$t_1(k-j) = [(j+1)! N_k N_{k-1} \cdots N_{k-j+1}]^{1/(j+1)}. \tag{5}$$

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

$$t_{k-j}(\omega) = \left((4\omega^3)[(j+1)! N_k N_{k-1} \cdots N_{k-j+1}]\right)^{1/(j+1)}. \tag{6}$$

The variable $t_1(k-j)$ is used a reference point in many of our calculations, and concentration of $\widehat{L}_{k-j}(t)$ follows for $t \geq t_{k-j}(\omega)$.

Lemma 4. *Let $\mu_i(t)$ be given by* (3)*. Let $\omega = 6 \log n$. Provided $\widehat{L}_i(t) \leq N_i$, the following hold for $i = 0, 1, ..., k$.*

(1) Deterministically $\widehat{L}_k(t) = t$, and for $j \geq 1$, if $j^2/t = o(1)$, then $\mathbb{E}\,\widehat{L}_{k-j}(t) \sim \mu_{k-j}(t)$.

(2) Suppose that $t_1(k-(j-1)) \ll t_1(k-j)$, and that $j^2/t = o(1)$. If $t \geq t_{k-j}(\omega)$ then w.h.p. $\widehat{L}_{k-j}(t) = \mu_{k-j}(t)(1 + O(1/\omega))$.

Note. The fact that $t_1(k-(j-1)) \ll t_1(k-j)$ is to ensure that $\mu_{k-(j-1)}(t_1(k-j))$ is sufficiently large and thus $\widehat{L}_{k-(j-1)}(t_1(k-j))$ is concentrated as $\widehat{L}_{k-j}(t)$ grows, is a model dependent calculation given in Section 3.1 and Section 4 respectively for the equal and growing layers models.

Proof. As $N_{k+1} = 1$ and $\widehat{L}_{k+1}(t) = 1$, this implies that $\widehat{L}_k(s) = s$ for $0 \leq s \leq t$. Moreover at most one vertex can be added to $\widehat{L}_i(t-1)$ at step t, which implies $\widehat{L}_i(t) \leq t$.

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

$$\mathbb{E}\,\widehat{L}_i(t) = \frac{1}{N_{i+1}} \sum_{s=0}^{t-1} \mathbb{E}\,\widehat{L}_{i+1}(s). \tag{7}$$

We claim for $j \geq 0$ that

$$\mathbb{1}_{\{t \geq j\}} \frac{(t-j)^{j+1}}{(j+1)!} \leq (N_k N_{k-1} \cdots N_{k-j+1})\, \mathbb{E}\,\widehat{L}_{k-j}(t) \leq \frac{t^{j+1}}{(j+1)!}. \tag{8}$$

For given t, the induction is backwards on $k-j$ from $j = 0$. When $j = 0$ (8) is true, so the first non-trivial case is $j = 1$. From (7) we see that

$$\mathbb{E}\,\widehat{L}_{k-1}(t) = \frac{1}{N_k} \sum_{s=0}^{t-1} s, \tag{9}$$

which illustrates how (8) arises from bounding this sum.

For the induction at step $i = k - (j+1)$, let

$$M_{j-1} = N_k N_{k-1} \cdots N_{k-j+1}. \tag{10}$$

Multiply (7) by M_{j-1}, and insert the bounds on $M_{j-1}\mathbb{E}\,\widehat{L}_{k-j}(s)$ from (8) (with $i+1 = k-j$) into this, to give

$$\frac{1}{N_{k-j}} \sum_{s=j}^{t-1} \frac{(s-j)^{j+1}}{(j+1)!} \le M_{j-1}\,\mathbb{E}\,\widehat{L}_{k-(j+1)}(t) \le \frac{1}{N_{k-j}} \sum_{s=1}^{t-1} \frac{s^{j+1}}{(j+1)!} \tag{11}$$

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

$$\frac{(t-1)^{m+1}}{m+1} \le 1^m + 2^m + \cdots + (t-1)^m \le \frac{t^{m+1}}{m+1}. \tag{12}$$

Use (12) in (11) 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 (8). Moreover, provided $j^2/t = o(1)$,

$$\mathbb{E}\,\widehat{L}_{k-j}(t) = \frac{1}{N_k N_{k-1} \cdots N_{k-j+1}} \frac{t^{j+1}}{(j+1)!} \left(1 - O\left(j^2/t\right)\right) = \mu_{k-j}(t)(1+o(1)). \tag{13}$$

This completes the proof of Lemma 4.(1). □

We proceed to the proof of Lemma 4.(2). The first thing to check is that, for the values of k given in Theorem 1, for $t \ge t_{k-j}(\omega)$, $j^2/t = o(1)$, allowing us to use Lemma 4.(1).

Lemma 5. *Let $t_1(k-j)$ be the value of t such that $\mu_{k-j}(t) = 1$ as given by (5). The condition $j^2/t = o(1)$ is satisfied at $t_1(k-j)$ in the equal layers model provided $k = o(n^{1/5})$ and in the growing layers model provided $d \to \infty$. This allows us to assume that for $t \ge t_{k-j}(\omega)$, $\mathbb{E}\,\widehat{L}_{k-j}(t) \sim \mu_{k-j}(t)$ in subsequent calculations.*

Proof. For $j \ge 1$, the value of $t_{k-j}(\omega)$ given in (6), satisfies

$$t_{k-j}(\omega) \gg t_1(k-j) = ((j+1)! N_k \cdots N_{k-j+1})^{1/(j+1)} \ge M_{j-1}^{1/(j+1)},$$

see (10). The product M_{j-1} is model specific, having the values $M_{j-1}(E) = (n/k)^j$ (equal layers model) and $M_{j-1}(G) = d^{kj-j(j-1)/2}$ (growing layers model).

In the first case $M_{j-1}(E)^{1/(j+1)} \ge (n/k)^{1/2}$, and in the second $M_{j-1}(G) \ge d^{k/2}$, this minimum being achieved at $j=1$ or $j=k$. Checking $j^2/M_{j-1}^{1/(j+1)}$ we see that the condition $j^2/t = o(1)$ is satisfied at $t_1(k-j)$ in the equal layers model provided $k = o(n^{1/5})$ and in the growing layers model provided either $d \to \infty$ or $k \to \infty$. □

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

We now prove Lemma 4.(2).

Lemma 6. *Let $\mu_{k-j}(t)$ as be given by (3). Let $\omega \geq 2\log N_k + 4\log k$. Let $t_k(\omega) = 1$, and for $1 \leq j \leq k-1$, as given in (6), let*

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

Let $k^* = \sqrt{\log n}/\log\log n$. If $k \leq k^*$, with probability $1 - O(k^2\omega N_k e^{-\omega})$ it holds that for all $j \leq k-1$ and all $t \geq t_{k-j}(\omega)$ that $\widehat{L}_{k-j}(t) \in [\mu_{k-j}(t)(1-1/\omega), \mu_{k-j}(t)(1+1/\omega)]$.

Proof. The proof is inductive and for level $k-j$ it depends on establishing the result for levels $k, k-1, \ldots, k-j+1$. By definition, $\widehat{L}_k(t) = t = \mu_k(t)$. Let $t_k(\omega) = 1$, establishing Lemma 4.3 for $j = 0$. For $j \geq 0$, let $t_1(k-j)$ be the solution to $\mu_{k-j}(t) = 1$ as given by (5).

For $j \geq 1$,

$$t_{k-j}(\omega) = (4\omega^3)^{1/(j+1)}[(j+1)!N_k N_{k-1}\cdots N_{k-j+1}]^{1/(j+1)} = (4\omega^3)^{1/(j+1)} t_1(k-j), \tag{14}$$

so that $\mu_{k-j}(t_{k-j}(\omega)) = 4\omega^3$ where ω is still to be determined.

The random variable $\widehat{L}_{k-j}(t)$ is obtained from $\widehat{L}_{k-j}(t-1)$ by choosing a random vertex $u \in S_{k-j}$ irrespective of the current occupancy of u, and 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 $\mathbb{P}(Q_{k-j}(t) = 1) = (\widehat{L}_{k-j+1}(t)/N_{k-j+1})$ independently of any previous outcomes. In particular, $\mathbb{E}\, Q_{k-1}(t) = t/N_k$, and by equation (9), $\mathbb{E}\,\widehat{L}_{k-1}(t) = t(t-1)/(2N_k) \sim \mu_{k-1}(t)$. Similarly $\mathbb{E}\,\widehat{L}_{k-j}(t) \sim \mu_{k-j}(t)$ again by the Proof of Lemma 4.2.

For $j \geq 1$, and given $t \geq t_{k-j}(\omega)$ let $\mathcal{A}_j(t)$ denote the event that $\widehat{L}_{k-j}(t) \in [\mu_{k-j}(t)(1-1/\omega), \mu_{k-j}(t)(1+1/\omega)]$. By Hoeffding's Inequality,

$$\mathbb{P}(\neg\mathcal{A}_j(t)) \leq 2\exp\left\{-\frac{\mu_{k-j}(t)}{3\omega^2}(1-\delta)\right\} \leq 2\exp\left\{-\frac{\mu(t_{k-j}(\omega))}{4\omega^2}\right\} \leq 2e^{-\omega}, \tag{15}$$

where $\delta = O(1/\omega + j^2/t)$, includes the correction from (13).

Let $\mathcal{E}_j(t)$ be the event that $\mathcal{A}_j(s)$ holds for all $s \in [t_{k-j}(\omega), t]$. For $j = 0$, $\mathbb{P}(\mathcal{E}_0(t)) = 1$, and for $j \geq 1$, given $\mathcal{E}_{j-1}(t)$, we have inductively that

$$\mathbb{P}(\neg\mathcal{E}_j(t) \mid \mathcal{E}_{j-1}(t)) \leq \sum_{s=t_{k-j}(\omega)}^{t} \mathbb{P}(\neg\mathcal{A}_j(s)). \tag{16}$$

Adding over $i \leq j$ we will be able to complete the induction, via

$$\mathbb{P}(\neg \mathcal{E}_j) \leq \mathbb{P}(\neg \mathcal{E}_j \mid \mathcal{E}_{j-1}) + \mathbb{P}(\neg \mathcal{E}_{j-1}) \leq \sum_{i \leq j} \sum_{t \geq t_{k-i}(\omega)} \mathbb{P}(\neg \mathcal{A}_i(t)) \leq cj\omega N_k e^{-\omega}, \tag{17}$$

provided we can establish the bound on the RHS, which we now do. Here c is some absolute constant.

As $\mu_{k-j}(t)$ in (3) is monotone increasing in t and by (15) is bounded above by $2e^{-\omega} = o(1)$, we use the Euler-MacLaurin Theorem to replace the summation over t in (16) by an integral. Thus,

$$\sum_{t \geq t_{k-j}(\omega)} \mathbb{P}(\neg \mathcal{A}_j(t)) \leq 2 \sum_{t \geq t_{k-j}(\omega)} \exp\left\{ -\frac{\mu(t_{k-j}(\omega))}{4\omega^2} \right\} \leq 3 \int_{t \geq t_0} \exp\left\{ -\frac{t^{j+1}}{C} \right\} dt,$$

where, $t_0 = t_{k-j}(\omega)$, $C = [4\omega^2 (j+1)! N_k \cdots N_{k-j+1}]$, and from (14), $t_{k-j}(\omega) = (4\omega^3 C)^{1/j+1}$. Put $t^{j+1}/C = z^2/2$, so that $t = (Cz^2/2)^{1/(j+1)}$, and $dt/dz = (2/(j+1))(C/2)^{1/(j+1)} z^{2/(j+1)-1}$. Let $z_0 = z(t_0)$, then $z_0 = \sqrt{2\omega}$, and as $j \geq 1$, $z^{2/(j+1)-1} \leq 1$. Finally $(C/2)^{1/(j+1)} = O(\omega^{3/2} N_k)$, giving

$$\int_{t \geq t_0} \exp\left\{ -\frac{t^{j+1}}{C} \right\} dt \leq \frac{2}{j+1} \left(\frac{C}{2} \right)^{1/(j+1)} \int_{z \geq z_0} e^{-z^2/2}\, dz \leq O(\omega^{3/2} N_k)\, \frac{1}{\sqrt{\omega}} e^{-\omega},$$

by using a standard bound on the tail of the Normal distribution, $\mathbb{P}(Z \geq z) \leq 1/(\sqrt{2\pi} z) e^{-z^2/2}$.

Lemma 6 follows for all $1 \leq j \leq k$ by adding the RHS of (17) over $j \leq k$ and choosing $\omega = 2 \log N_k + 4 \log k \leq 6 \log n$ $\square$

2.3 Lower bound on occupancy at step t

So far we only have an upper bound on $\mathbb{E}\, L_j(t)$ given by $\mathbb{E}\, L_j(t) \leq \mathbb{E}\, \widehat{L}_j(t)$. We construct a lower bound $\mathbb{E}\, \widetilde{L}_j(t)$ and prove that for large enough t these bounds converge thus giving the asymptotic value of $\mathbb{E}\, L_j(t)$.

For a given level j, a lower bound on $L_j(t)$ can be found as follows. Define a sub-process of DLA which requires that the particle avoids upper-blocked vertices. Thus to reach level j, a particle must avoid choosing neighbours in $\widehat{B}_1, ..., \widehat{B}_j$ at steps $0, ..., j-1$ of its random walk.

Let $L_j^*(t)$ be a w.h.p. upper bound on $\widehat{L}_j(t)$. This will, for example, be obtained from Lemma 6. if $t \geq t_j(\omega)$. Referring to (1), (2), let $\widetilde{L}_i(t)$ be obtained by replacing L by L^* in the bracketed terms on the RHS. This defines a lower bound $\widetilde{L}_j(t)$, such that w.h.p.

$\widetilde{L}_j(t) \le L_j(t) \le \widehat{L}_j(t)$. Let $\widetilde{\mathcal{H}}(t) = (\widetilde{L}_0(t), \widetilde{L}_1(t), \ldots, \widetilde{L}_k(t))$ be the occupancy vector obtained from this lower bound, then we have the following recurrences.

$$\mathbb{E}\,(\widetilde{L}_i(t+1) \mid \widetilde{\mathcal{H}}(t)) = \widetilde{L}_i(t) + \frac{\widetilde{L}_{i+1}(t)}{N_{i+1}} \prod_{j=0}^{i} \left(1 - \frac{L_j^*(t)}{N_j}\right), \tag{18}$$

$$\mathbb{E}\,(\widetilde{L}_k(t+1) \mid \widetilde{\mathcal{H}}(t)) = \widetilde{L}_k(t) + \prod_{j=0}^{k} \left(1 - \frac{L_j^*(t)}{N_j}\right). \tag{19}$$

These recurrences are solved in Section 3.2 for the equal layers model, and in Section 4 for the growing layers model.

The gap property. It is clear that $0 \le \widetilde{L}_{k-j}(t) \le L_{k-j}(t) \le \widehat{L}_{k-j}(t) \le t$, and that these values are monotone non-decreasing. We choose times

$$t^-(k-j) < t_1(k-j) < t^+(k-j) = t_{k-j}(\omega)$$

such that

$$\mu_{k-j}(t^-) = 1/\omega^3, \quad \mu_{k-j}(t_1) = 1, \quad \mu_{k-j}(t^+) = 4\omega^3.$$

Below $t^-(k-j)$, $\widehat{L}_{k-j}(t) = 0$ w.h.p. for all $j = 1, ..., k$. This follows from the definition of $k < \omega$ and the Markov Inequality. Above $t^+(k-j)$ we have that $\mathbb{E}\,\widetilde{L}_{k-j}(t) \sim \mathbb{E}\,\widehat{L}_{k-j}(t)$ and thus $\mathbb{E}\,L_{k-j}(t) \sim \mu_{k-j}(t)$. The main content of Sections 3 and 4 is to prove this via a *gap property* which implies that

$$t^+(k-j) = t_{k-j}(\omega) \ \ll\ t^-(k-(j+1)) \ \ll\ t_1(k-(j+1)),$$

so that w.h.p. concentration for L_{k-j} occurs while $L_{k-(j+1)}$ is still zero.

3 Analysis of DLA in the equal layers model

It is convenient at this point to obtain an asymptotic estimate for the finish time of DLA in the equal layers model. Let $t_1(k-j)$ and $t_{k-j}(\omega)$ be as given by (5) and (6). Thus $T_f = t_1(0)$, the time at which the source has expected occupancy one (in the upper-blocked process). Recalling that $N_1 = \cdots = N_k = n/k$, let

$$T_f = [(k+1)! N_1 \cdots N_k]^{1/(k+1)} = \left(\frac{(k+1)!}{k^k} n^k\right)^{1/(k+1)}. \tag{20}$$

That T_f indeed approximates the finish time will be shown in Section 3.3.

Lemma 7. *For $k \geq 2$, let G be an equal layers graph with level sizes $N_i = n/k$, $i = 1, ..., k$.*

(1) Let $\gamma_k = e^{-1}(1 + O(\log k/k))$, then

$$T_f = n^{k/(k+1)}\gamma_k. \tag{21}$$

(2) For $1 \leq j \leq k$, let be as given in (5), then

$$t_1(k-j) = ((j+1)!)^{1/(j+1)} (n/k)^{j/(j+1)} = \Theta(j)(n/k)^{j/(j+1)},$$

and

$$t_1(k-j) = t_1(k-(j-1)) \cdot \Theta(1)(n/k)^{1/(j(j+1))}.$$

Let $k^* = \sqrt{\log n}/\log\log n$. If $k \leq k^*$, then $t_1(k-j) \gg t_1(k-(j-1))$ in the equal layers model; as claimed in the note below Lemma 4.

Proof. Note that

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

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

$$(\Theta_k)^{1/k+1} = e^{O(\log k/k)} = 1 + O(\log k/k)$$

is bounded for $k \geq 2$ and tends to one as $k \to \infty$. From (20)

$$T_f = \left(\frac{(k+1)!}{k^k}n^k\right)^{1/(k+1)} = n^{\frac{k}{k+1}}e^{-1}\ (1 + O(\log k/k)). \tag{22}$$

The second part follows by direct calculation using $((j+1)!)^{1/(j+1)} = j\Theta(1 + j^{3/(2j)}) = j\Theta(1)$. $\square$

Note that with $t_{k-j}(\omega)$ as given by (6), then $t_{k-j}(\omega) = (4\omega)^{1/(j+1)}t_1(k-j)$, so asymptotics follow from the above lemma.

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

We prove there is a large gap in the number of steps between the time when $\mathbb{E}\,\widehat{L}_{k-j} = 1$ with all lower values zero, and the time when $\mathbb{E}\,\widehat{L}_{k-(j+1)} = 1$. The gap allows $\widehat{L}_{k-j}$ to increase and become concentrated around μ_{k-j}, whilst all values with a lower index $i < k-j$ remain zero. This confirms the inductive assumption stated below (5) in Lemma 6.

We list various assumptions used in this section.

$$1 \leq k \leq k^* = \frac{\sqrt{\log n}}{\log \log n}, \quad \omega = 6 \log n, \quad \beta = \frac{N_k}{\omega T_f} \geq (6 \log n)^{4+k/2}. \tag{23}$$

Let $\widehat{L} = (\widehat{L}_0, \widehat{L}_1, \dots, \widehat{L}_k)$ 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 $\widehat{L}_{i-1} = 0$.

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, as assumed in Lemma 6. In particular w.h.p. at $t_{k-j}(\omega)$, where $t'' \gg t_{k-j}(\omega) > t'$ we have (31).

Define $\beta = \beta_k = N_k/\omega T_f$. From (21),

$$\beta = \frac{N_k}{\omega T_f} = \frac{n}{\omega k} \frac{1}{\gamma_k n^{k/(k+1)}} = \frac{1}{\gamma_k} \frac{1}{\omega k} n^{1/(k+1)}. \tag{24}$$

We assumed that $\beta \geq (6 \log n)^{4+k/2}$. This is true if $\omega = 6 \log n$ since we assume that $k \leq \sqrt{\log n}/\log \log n$.

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

$$\mu_{k-(j+1)}(t) = \frac{t}{(j+2)N_{k-j}} \cdot \mu_{k-j}(t) \leq \frac{1}{(j+1)\beta} \mu_{k-j}(t). \tag{25}$$

For $j + \ell \leq k$, we can iterate this to give

$$\mu_{k-(j+\ell)}(t) \leq \mu_{k-j}(t) \frac{1}{\beta^\ell} \frac{1}{(j+\ell+1)(j+\ell)\cdots(j+2)} \leq \mu_{k-j}(t) \frac{1}{((j+1)\beta)^\ell}. \tag{26}$$

Consider $\mathbb{E}\, \widehat{L}_{k-j}(t)$. At $t \sim t_1(k-j)$ when $\mu_{k-j}(t) \sim 1$, then (see Lemma 5) $\mathbb{E}\, \widehat{L}_{k-j}(t) \sim 1$. The Markov inequality implies that w.h.p. $\widehat{L}_{k-j}(t) \in I_\omega = [0, 1, ..., \omega]$. From (25)–(26),

$$\mu_{k-(j+\ell)}(t) \leq \frac{1 + o(1)}{((j+1)\beta)^\ell}, \qquad \text{for } 1 \leq \ell \leq k - j, \tag{27}$$

and thus w.h.p. $\widehat{L}_0 = 0, \widehat{L}_1 = 0, \dots, \widehat{L}_{k-(j+1)} = 0$. Using (25)–(26), we see that

$$\mu_{k-j+\ell}(t) \geq \mu_{k-j}(t)\, j(j-1)\cdots(j-\ell+1)\, \beta^\ell \qquad \text{for } \ell \leq j. \tag{28}$$

So if $\mu_{k-j}(t) \sim 1$, then for $1 \leq \ell \leq j$, $t \geq t_{k-j+\ell}(\omega)$ and

$$\mu_{k-j+\ell}(t) \geq \beta^\ell \geq (6 \log n)^{4+k/2}. \tag{29}$$

Thus Lemma 6 holds, and w.h.p. $\widehat{L}_{k-j+\ell}$ is equal to $(1+o(1))\mathbb{E}\,\widehat{L}_{k-j+\ell} \sim \mu_{k-j+\ell}(t)$.

In summary, at time t such that $\mu_{k-j}(t) \sim 1$ (implying that $t \le T_f$, see Proposition 7.(2)), w.h.p., the state vector $\widehat{L}$ is such that $\mu_{k-\ell} \to \infty$ for $\ell \le j-1$, and

$$\widehat{L}(t) = (0, \dots, 0,\ \widehat{L}_{k-j} \in I_\omega,\ (1+o(1))\mu_{k-j+1},\ (1+o(1))\mu_{k-j+2}, \dots,\ \mu_k). \tag{30}$$

Let $t = t_1(k-j)$, let $t' = t_{k-j}(\omega) = (4\omega^3)^{1/((j+1)}t_1(k-j)$, so $\mu_{k-j}(t') \sim 4\omega^3$. By Lemma 6 it holds w.h.p. that $\widehat{L}_{k-j}(t') \sim \mu_{k-j}(t')$.

Let $t'' = t_1(k-(j+1))$. By Proposition 7.(2), $t_1(k-(j+1)) = \Theta(1)(n/k)^{1/(j(j+1))}t_1(k-j)$. As $t' = (4\omega^3)^{1/((j+1)}t_1(k-j)$, it can be checked that $t'' \gg t'$. Also by (26)

$$\sum_{\ell \ge 1} \mu_{k-(j+\ell)}(t') = O\left(\frac{\mu_{k-j}(t')}{\beta}\right) = O\left(\frac{\omega^3}{\beta}\right) = O\left(\frac{1}{\log^{1+k/2}}\right).$$

Thus, applying the Markov inequality to the above, at $t' = t_{k-j}(\omega)$, w.h.p.

$$\widehat{L}(t_{k-j}(\omega)) = (0, ..., 0,\ (1+o(1))\mu_{k-j},\ (1+o(1))\mu_{k-j+1}, ...,\ \mu_k), \tag{31}$$

so that $\widehat{L}_{k-j}(t')$ is concentrated and all lower levels are unoccupied, and thus the claimed gap exists. This condition persists w.h.p. until around $t_1(k-(j+1)) = t'' \gg t' = t_{k-j}(\omega)$, when $\mu_{k-(j+1)}(t_1(k-(j+1))) = 1$ at which point $\widehat{L}(t'')$ resembles (30) and the induction continues.

3.2 Lower bound on occupancy in the equal layers model.

We prove that, for $t \ge t_i(\omega)$, we have $\mathbb{E}\,\widetilde{L}_i(t) \sim \mathbb{E}\,\widehat{L}_i(t) \sim \mu_i(t)$. As with the upper bounds, we will have that $\mathbb{E}\,\widetilde{L}_k \gg \mathbb{E}\,\widetilde{L}_{k-j}$ for $j \ge 1$. The first step is to draw a line between them.

Let $t^- = t_1(k-1)/\omega$, so that $\mu_{k-1}(t^-) = 1/\omega^2$. Then w.h.p $\widehat{L}_{k-1}(t^-) = 0$ and as $\widehat{L}_{k-1}(t)$ is monotone non-decreasing, w.h.p. $\widehat{L}_j(t) = 0$ for all $j \le k-1$ and $t \le t^-$. As $\widehat{L}_k(t) = L_k^*(t) = t$ deterministically, this simplifies (19) for $t \le t^-$.

As before let $\beta = N_k/\omega T_f$ where ω, k are given by (23). Provided $k \ge 2$, if $t \ge t^-$, then $t/\beta \gg \omega^3$. Indeed using (24),

$$t^- = \frac{t_1(k-1)}{\omega} = \frac{1}{\omega}\sqrt{\frac{2n}{k}} \qquad \Longrightarrow \qquad \frac{t^-}{\beta} = \frac{\gamma_k \omega k}{n^{1/(k+1)}} \cdot \frac{1}{\omega}\sqrt{\frac{2n}{k}} \ge 2\gamma_k n^{1/6} \gg \omega^3.$$

For $t \ge t^-$, and $1 \le i \le k-1$ we first consider the product term in (18). Recalling that $N_i = N_k = n/k$, we will prove that

$$\prod_{j=0}^{i}\left(1 - \frac{L_j^*(t)}{N_j}\right) \ge 1 - \sum_{j=0}^{i}\left(\frac{L_j^*(t)}{N_j}\right) = 1 - O\left(\frac{t}{\beta N_k}\right) = 1 - o(1). \tag{32}$$

By (26), $\mu_{k-j} \leq \mu_k/\beta^j$. Assume $t \geq t^-$ and that for some $\ell \leq k-1$, $t_\ell(\omega) \leq t < t_{\ell+1}(\omega)$. Apply Lemma 6 for $j \geq \ell+1$ along with (30) and (31).

If $i < \ell$ then $L_j^*(t) = 0$, for $j \leq i$. Next, if $i = \ell$, $L_j^*(t) = 0$ for $j < i$ and $L_i^*(t) \leq 2\mu_i(t_i(\omega)) \leq 8\omega^3$. Finally assume $i \geq \ell+1$. Then (as $t \geq t^-$),

$$\sum_{j=0}^{i} \frac{L_j^*(t)}{N_j} \leq \frac{8\omega^3}{N_k} + 2\sum_{j=\ell+1}^{i} \frac{t}{\beta^{k-j}N_k} \leq \frac{O(1)}{N_k}\left(\omega^3 + \frac{t}{\beta}\right) = O\left(\frac{t}{\beta N_k}\right). \tag{33}$$

Consider now $\mathbb{E}\,\widetilde{L}_k(t)$. For all $t \geq 0$, $\widehat{L}_k(t) = t$, and $\sum_{j<k} \widehat{L}_j(t) = O(t/\beta)$. So from (19)

$$\mathbb{E}\,\widetilde{L}_k(t) = t - \sum_{s=0}^{t} O\left(\frac{s}{N_k}\right) = t - O\left(\frac{t^2}{N_k}\right) = t\left(1 - O\left(\frac{t}{N_k}\right)\right).$$

For $t \geq t_0$ where $t_0 \to \infty$ arbitrarily slowly we have $\widetilde{L}_k(t) \sim t$ w.h.p., initializing an induction for $\mathbb{E}\,\widetilde{L}_i(t)$ using arguments equivalent Section 2.2 for $\mathbb{E}\,\widehat{L}_i(t)$.

At step $t+1$, equation (32) implies that, in the lower bounds on the process, particle $t+1$ arrives at level i with probability $(1-o(1))$. If $i < k$, it halts at this level with probability $\widetilde{L}_{i+1}(t)/N_{i+1}$. Thus

$$\mathbb{E}\,\widetilde{L}_i(t+1) = \mathbb{E}\,\widetilde{L}_i(t) + \left(1 - O\left(\frac{t}{\beta N_k}\right)\right) \cdot \frac{\mathbb{E}\,\widetilde{L}_{i+1}(t)}{N_{i+1}}. \tag{34}$$

Arguing as in (7) on the inductive assumption that $\mathbb{E}\,\widetilde{L}_{i+1}(t) = \mu_{i+1}(t)(1 - O(t/N_k))$, we find

$$\begin{aligned}
\mathbb{E}\,\widetilde{L}_i(t) =& \frac{1}{N_{i+1}} \sum_{s=0}^{t-1} \mathbb{E}\,\widetilde{L}_{i+1}(s)\left(1 - O\left(\frac{s}{\beta N_k}\right)\right) \\
=& \frac{1}{N_{i+1}} \sum_{s=0}^{t-1} \mu_{i+1}(s)\left(1 - O\left(\frac{s}{N_k}\right)\right)\left(1 - O\left(\frac{s}{\beta N_k}\right)\right) \\
=& \frac{1}{N_{i+1}} \sum_{s=0}^{t-1} \frac{s^i}{i!N_k \cdots N_{i+2}}\left(1 - O\left(\frac{s}{N_k}\right)\right) \\
=& \mu_i(t) - O(1)\frac{t\mu_i(t)}{N_k} \quad = \quad \mu_i(t)\left(1 - O\left(\frac{t}{N_k}\right)\right).
\end{aligned}$$

Thus

$$\mathbb{E}\,\widetilde{L}_i(t) \sim \mathbb{E}\,L_i(t) \sim \mathbb{E}\,\widehat{L}_i(t) \sim \mu_i(t) = \frac{t^{i+1}}{(i+1)!N_k \cdots N_{i+1}} \tag{35}$$

as required.

Let $t_i(\omega)$ be given by (14). For those $i \le k$, such that $t \ge t_i(\omega)$, then $\mu_i(t) \to \infty$ suitably fast and the concentration results of Lemma 6 hold. The gaps inherited from the upper bound argument of Section 3.1 are essentially unaltered.

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

3.3 Finish time of DLA in the equal layers model.

A lower bound on the finish time follows from the upper-blocked process, and an upper bound from the lower bound estimates for DLA.

Proposition 8. *For $1 \le k \le k^* = \sqrt{\log n}/\log\log n$, let G be an equal layers graph with level sizes $N_i = n/k$, $i = 1, ..., k$. Let T_f be given by (20). With probability $1 - O(1/\omega')$, the finish time t_f of the DLA process in G satisfies $T_f/\omega' \le t_f \le \omega' T_f$, where $\omega' \to \infty$ arbitrarily slowly.*

Proof. Let $t_1 = t_1(0)$ be such that $\mu_0(t_1) \sim 1$, and thus $t_1 \sim T_f$. Let $t' = t_1/\omega'$, where $\omega' \to \infty$ slowly. Then w.h.p., $\mathbb{E}\,\widehat{L}_0(t') = O(1/(\omega')^{k+1})$, and thus $\mathbb{P}(\widehat{L}_0(t') > 0) = O(1/\omega')$.

We next investigate the concentration of $L_1(T_f)$. By Lemma 7.(1),

$$\mu_1(T_f) = \mu_0(T_f)\frac{N_1(k+2)}{T_f} \ge \frac{n}{\gamma_k n^{k/(k+1)}} = \Theta(1)n^{1/(k+1)} \gg 4\omega^3, \tag{36}$$

where $\omega = 6\log n$. Thus $T_f \gg t_1(\omega)$ and hence Proposition 6 holds for $\widetilde{L}_1(T_f)$. Suppose at T_f that $L_0(T_f) = 0$. By (36) the expected waiting time τ for a particle to hit $L_1(T_f)$ is

$$\tau \le (1 - o(1))\frac{N_1}{\mu_1(T_f)} = \Theta(1)\frac{n}{k\, n^{1/(k+1)}} = \Theta(1)\frac{n^{k/(k+1)}}{k} = \Theta(T_f/k) = O(T_f)$$

By time $\omega' T_f$, w.h.p. $L_0(\omega' T_f) = 1$, completing the proof of Proposition 8, and hence Theorem 1.(1). □

3.4 Existence of a unique connecting path component.

We now prove Theorem 1.(3) for the equal layers model. We must show w.h.p. that at t_f, the finish time, the path of occupied vertices connecting the source to level k (and hence to the sink) has no off-path neighbours in C_{t_f}.

At a given step t, the edge induced component C_t is obtained from C_{t-1} by adding a newly occupied vertex which points to the neighbour in C_{t-1} which halted the particle: Thus 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 C_t rooted at the sink. An arborescence is a rooted tree with all edges directed towards the root vertex. Thus C_t is an arborescence with root z. On deletion of the z, $D_t = C_t \backslash \{z\}$ becomes a directed forest of arborescences each rooted at a vertex in level k.

Let B_i be the set of occupied vertices in level i, where $L_i = |B_i|$. Given that a particle at u chooses a vertex in B_{i+1} as a neighbour, then this neighbour is chosen uniformly at random (u.a.r.) from the set B_{i+1}.

We regard vertices occupied by halted particles as coloured either red or blue, with all occupied vertices in level k coloured blue. 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 choice of w by the particle at u is independent of the colour of w at this step.

The process halts when there is a directed path of occupied vertices $v = u_0 u_1 \cdots u_k z$ from the source to sink. The source vertex v is blue at t_f as it is the first in-neighbour of u_1.

Lemma 9. *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 arborescence of halted particles rooted at $w = u_k$ is exactly this path.*

Proof. As before, let B_i be the set of occupied vertices in level i, and $L_i = |B_i|$. As each $u \in B_i$ has a unique occupied out-neighbour in C, 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, as $|\mathrm{Out}(B_{k-j})| \leq \widehat{L}_{k-j}(s)$,

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

Let $Z_{k-j}(t)$ be the number of red vertices in level $k-j$. We associate the number of possibly-red vertices at each step with a super-process $\widehat{Z}_{k-j}(t) = Z_{k-j}(t) + Q_{k-j}(t)$, where $Q_{k-j}(t)$ is Bernoulii with parameter $(\widehat{L}_{k-j} - |(\mathrm{Out})B_{k-j}|)$. Thus $Z_{k-j}(t) \leq \widehat{Z}_{k-j}(t)$ and where

$$\mathbb{E}\,(\widehat{Z}_{k-j}(t+1) \mid \widehat{\mathcal{H}}(t)) = \widehat{Z}_{k-j}(t) + \frac{\widehat{L}_{k-j}(t)}{N_{k-j}}.$$

The recurrence for $\mathbb{E}\,\widehat{Z}$ mirrors the recurrence (4) for $\widehat{L}_i(t)$ with $k-j$ replacing $i+1$.

Specifically,

$$\mathbb{E}\,\widehat{Z}_{k-j}(t) \sim \frac{1}{(n/k)} \sum_{s=1}^{t} \mu_{k-j}(s) = \frac{1}{(n/k)} \sum_{s=1}^{t} \frac{s^{j+1}}{(j+1)!(n/k)^j}$$
$$\sim \frac{t^{j+2}}{(j+2)!(n/k)^{j+1}} = \mu_{k-(j+1)}(t). \tag{37}$$

We note that $\widehat{Z}_{k-j}(t)$ is the sum of independent indicator variables, and will be concentrated at or after $t_{k-(j+1)}(\omega)$. The number of red vertices is at most $2Z_{k-j}(t) \le 2\widehat{Z}_{k-j}(t)$, where the factor 2 covers the case where the pre-existing sibling was blue but is recoloured red.

Assume $T_f/\omega \le t_f \le \omega T_f$, where $\omega = \omega'$ from Proposition 8 tends slowly to infinity with n. Denote the path connecting v to level k by $v = u_0 u_1 u_2 \cdots u_k$. By definition $v = u_0$ is blue. For $i \ge 1$, let $R_i(s)$ be the red vertices in level i at step s. Let s_i denote the step at which u_i became occupied. Consider next the colour of u_1 at t_f. As $T_f = t_1(0)$, and $t = t_f \ge T_f/\omega \gg t_1(\omega)$ by the gap property, so we have that $L_1(t) \sim \mu_1(t)$. On the other hand $\widehat{Z}_1(t)$ may not be concentrated, but we can assume $\widehat{Z}_1(t) \le \omega' \mu_0(t)$ where ω' is to be determined. If $t = 2\omega T_f$, then $\mu_0(t) = (2\omega)^{k+1}$ and as $k \ge 2$ then $t \ge t_0(\omega)$, so the red subprocess in level one is concentrated. Crudely[1] put $\omega' = \omega^5$. Vertex u_1 was chosen uniformly at random from the occupied vertices in level one by the particle halting at u_0, so,

$$\mathbb{P}(u_1 \in R_1(t)) \le \frac{2\widehat{Z}_1(t)}{L_1(t)} \le \frac{2\omega' \mu_0(t)}{\mu_1(t)} \le \frac{2\omega' t}{(k+1)(n/k)} \le \frac{2\omega' t}{n}. \tag{38}$$

Next consider the colour of u_2. If u_2 is red at step t, then either (i) it became red before step s_1 when it was chosen uniformly at random by u_1, or (ii) it became red at some later step.

In the first case,

$$\mathbb{P}(u_2 \in R_2(s_1)) \le \frac{2\widehat{Z}_2(s_1)}{L_2(s_1)} \le 2\omega' \frac{\mu_1(s_1)}{\mu_2(s_1)} = \frac{2\omega' s_1}{k(n/k)} = \frac{2\omega' s_1}{n}.$$

By the gap theorem, L_2 is already concentrated at s_1, for otherwise $L_1(s_1)$ would be empty. The ω' covers possible lack of concentration of $\widehat{Z}_2$

In the second case, as u_3 is the chosen out-neighbour of u_2, then u_2 will turn red if some particle chooses u_3 after time s_1, and thus

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

[1]Why? The number of red vertices can only increase with t and is at most $8\omega^3$ in expectation at $t_0(\omega)$. The width of the 'end time interval' is ω^2. Apply the Markov inequality ω^2 times.

Thus

$$\mathbb{P}(u_2 \in R_2(t)) \le \frac{2\omega' k t}{n},$$

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

$$\mathbb{P}(\text{path } v u_1 \cdots u_k \text{ is blue}) \ge \prod_{i=1}^{k-1} \left(1 - \frac{2\omega' k t}{n}\right) = 1 - O\left(\frac{\omega' k^2 t}{n}\right),$$

where $t = t_f \le \omega n^{k/(k+1)}$. Provided $\omega' k^2 \omega / n^{1/(k+1)} = o(1)$, which holds for $k \le \sqrt{\log n} / \log\log n$, the path from the source to vertex $w = u_k$ level k is blue, and thus the arborescence rooted at w is exactly this path. □

4 Analysis of DLA in the growing layers model

In the growing layers model, each layer is larger than the previous one by a factor of d. Thus $N_j = d^j$ for $j = 0, 1, ..., k$ and we take $N_k = d^k = n$. Many of the properties of this model such as a gap property and unique connecting path are similar to the equal layers model.

The main, and most striking difference, is that there is a well defined level at a distance about $\sqrt{2k+2}$ from the end at which agglomerative growth stops and from which and a single path grows back towards the source. Moreover at the end, *except for the last ... levels*, the connecting path is the *only occupied vertex* in the layer. This is in contrast to the equal layers model where all levels have significantly occupancy, and even that of the first level at the end is $\Theta(n^{1/(k+1)})$.

Proposition 10. *Let G be a growing layers graph with level sizes $N_i = d^i$, for $i = 0, ..., k$, where $d \to \infty$, $k \to \infty$, and $k \ge \log^2 d$. Let*

$$T_f = \sqrt{k}\, d^{k+3/2-\sqrt{2k+2}}. \tag{39}$$

(1) The finish time t_f of DLA on G satisfies $T_f/\omega \le t_f \le \omega T_f$, where $\omega \to \infty$ slowly.

(2) 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.

Proof. The size N_i of layer i is $N_i = d^i$. It follows that the product of the set sizes in the denominator of $\mu_{k-j}(t)$ in (3) 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 (3) becomes

$$\mu_{k-j}(t) = \frac{t^{j+1}}{(j+1)!\, d^k d^{k-1} \cdots d^{k-j+1}} = \frac{t^{j+1}}{(j+1)!\, d^{kj-j(j-1)/2}}. \tag{40}$$

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 because the recurrence used to establish it assumes $\mu_{k-\ell} \to \infty$ for all $\ell < j$, which is not the case. We next locate where this happens; this is where the unique path back to the source begins.

An important level. We next show the existence of a level $i \sim k+1-\sqrt{2k+2}$, such that the first occupancy of this level effectively determines the finish time of the process.

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 (40),

$$\mu_{k-j}(t_1) = \frac{t_1^{j+1}}{(j+1)!\, d^{kj-j(j-1)/2}} \sim 1 \quad \Longrightarrow \quad t_1 \sim [(j+1)!]^{1/(j+1)}\, d^{\frac{2kj-j(j-1)}{2(j+1)}}. \tag{41}$$

From (3)

$$\frac{\mu_{k-j+1}(t)}{\mu_{k-j}(t)} = \frac{(j+1)d^{k-j+1}}{t}, \tag{42}$$

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

$$\begin{aligned} \mu_{k-j+1}(t_1) &\sim \frac{j+1}{[(j+1)!]^{1/j+1}}\, d^{k-j+1-\frac{2kj-j(j-1)}{2(j+1)}} \\ &\sim \frac{e}{(2\pi(j+1))^{1/2(j+1)}}\, d^{\frac{2k+2-j(j+1)}{2(j+1)}}. \end{aligned} \tag{43}$$

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

$$j^* = \sqrt{2k+2} - 1, \qquad k - j^* = k+1-\sqrt{2k+2}, \qquad d^{f(j^*)/2} = d^{k+3/2-\sqrt{2k+2}}. \tag{44}$$

In the case where j^* is not integer, the rounding error is addressed in the Appendix, where we show that the condition $k \geq \log^2 d$ given in Proposition 10 is sufficient to ignore the effect of rounding on the value of T_f.

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

$$\begin{aligned} t_1 &\sim [(j+1)!]^{1/(j+1)} \; d^{f(j)/2} \\ &\sim e^{-1}(\sqrt{2\pi})^{1/(j+1)} \; (j+1)^{1+1/2(j+1)} \;\; d^{j^2/2} \\ &= C_{k-j} \; \sqrt{2k+2} \;\; d^{k+3/2-\sqrt{2k+2}}, \end{aligned} \tag{45}$$

on inserting the values from (44), and where $C_{k-j} = e^{-1}(1+O(1/j))$. Note that $t_1 = \Theta(T_f)$. From (42),

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

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

$$\mu_{k-j^*}(t_1) \sim 1, \qquad \mu_{k-(j^*-1)}(t_1) = Cd^{1/2}, \qquad \mu_{k-(j^*+1)}(t_1) = C'd^{1/2}. \tag{46}$$

At first this seems confusing, as one might expect to have $\mu_{k-(j^*-1)}(t_1) = o(1)$ by analogy with the equal layers model. Assuming $d^{1/2} \to \infty$, level $k - j^* + 1$ is the last level 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.

Gap property of $\widehat{L}$ and a lower bound on L. Note that from the definition of $t_1(k-j^*)$, at t_1/ω, we have $\mu_{k-j^*}(t_1/\omega) = O(\omega^{-(j^*+1)})$, where $j^* \sim \sqrt{2k} \to \infty$. Thus w.h.p. $\widehat{L}_{k-j^*}(t_1/\omega) = o(1)$ in the upper process. Consequently all levels $i = k-\ell$, $\ell \geq j^*$, have $\widehat{L}_i(t) = 0$, w.h.p., for $t \leq t_1/\omega$.

In what follows we only consider indexes $k-\ell$ where $0 \leq \ell \leq k - j^* + 1$. By (44) we have $k - j^* + 1 = k + 2 - \sqrt{2k+2}$.

We see from (46) that $\mu_{k-j^*+1}(t_1(k-j^*)) = C'd^{1/2}$. Thus although $d \to \infty$ so that $\widehat{L}_{k-j^*+1}$ will be concentrated around $\mu_{k-(j^*+1)}(t_1)$, we cannot expect it to be as strong as in Lemma 6. Fortunately this will not matter as $k-j^*$ is the last level to which we apply the gap argument. For $\ell \leq j^*-1$ at $t_1(k-\ell)$, the value of $\widehat{L}_{k-\ell+1}$ obeys Lemma 6. In particular, it can be checked that $t_1(k-j^*+1) = \Theta(1)\sqrt{k}d^{k+5/2-(3/2)\sqrt{2k+2}}$. Thus using (42), we obtain $\mu_{k-j^*+2}(t_1(k-j^*+1)) = \Theta(d^{(\sqrt{2k+2}+1)/2})$.

Turning to the lower bound $\widetilde{L}$ as given in (18)–(19) we need to prove that (32) holds. The main task is to find a value of β for the growing layers model which we can use in the arguments given in Section 3.2 for the equal layers model. In what follows $\omega \to \infty$ slowly. The value from Lemma 6 is denoted as $\omega' = 6\log n$.

Using (39), (40), and $N_{k-\ell} = d^{k-\ell}$, for $\ell \geq 1$,

$$\begin{aligned} \frac{\mu_{k-\ell}}{N_{k-\ell}} &= \frac{t}{(k-\ell+1)d^{k-\ell}} \frac{\mu_{k-\ell+1}}{N_{k-\ell+1}} \\ &\leq \frac{\mu_{k-\ell+1}}{N_{k-\ell+1}} \frac{\omega\sqrt{k}}{(k+2-\sqrt{2k+2})} \frac{d^{k+3/2-\sqrt{2k+2}}}{d^{k+2-\sqrt{2k+2}}} \\ &\leq \frac{\omega(1+o(1))}{\sqrt{kd}} \frac{\mu_{k-\ell+1}}{N_{k-\ell+1}} \leq \left(\frac{1}{\beta}\right)^{\ell} \frac{t}{N_k}. \end{aligned}$$

Choosing $\beta = (1+o(1)))\sqrt{kd}/\omega$, (33) becomes

$$\sum_{\ell=j^*+1}^{i} \frac{L^*_{k-\ell}}{N_{k-\ell}} \leq \frac{O(\omega'^3)}{N_{k-j^*}} + 2\sum_{\ell=j^*+1}^{i} \left(\frac{1}{\beta}\right)^{\ell} \frac{t}{N_k} = O\left(\frac{t}{\beta^i N_k}\right).$$

It now follows from the proof in Section 3.2 that for those $k-j^*+1 \leq i \leq k$, and $t \geq t_i(\omega')$, then $\mu_i(t) \to \infty$ suitably fast. We hence obtain that $\mathbb{E}\,\widetilde{L}_i(t) \sim \mu_i(t) \sim \mathbb{E}\,\widehat{L}_i(t)$, and so we have $L_{k-\ell}(t) \sim \widehat{L}_{k-\ell}(t) \sim \mu_{k-\ell}(t)$.

4.1 The finish time in the growing layers model.

We now turn to the proof of (39). We show that, w.h.p., a path (of occupied vertices) grows back to the source from the *first vertex to be occupied* in level $k-j^*$, thus halting the process; and moreover this occurs *before a second vertex becomes occupied* in level $k-j^*$.

Let t_0 be the first step at which $L_{k-j^*}(t) = 1$, where w.h.p. $t_0 \geq t_1/\omega$. 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 (45),

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

where $C' \sim C/e$ and we assume $k \to \infty$.

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 $d \to \infty$.

We next give the proof of Proposition 10.(2). The expected number of steps needed to create another halted particle in level $k-j^*$ is

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

Whereas, w.h.p. on the assumption that $d \to \infty$, in at most $(t_1+T)\omega$ steps the process has halted as claimed, before a second vertex can become occupied in level $k-j^*$.

Existence of a unique connecting path. 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 9 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 Section 3.4 equation (38) etc.,

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

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

Thus as $t_f \le \omega t_1(j^*)$, where $j^* = \sqrt{2k+2}-1$ and $t_1(k-j^*)$ is given by (45)

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

□

5 Theorem 2.(1): Trees with large branching factor

Let $G = G(k,d)$ 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. To establish Theorem 2.(1), we need to prove w.h.p. that $T_f/\omega \le t_f \le \omega T_f$ where $\omega \to \infty$ slowly with n.

We borrow several ideas from the growing layers model, starting with j^* and T_f (see (44) and (39)). Let $j^* = \sqrt{2k+2} - 1$ and let T_f be given by

$$T_f = \sqrt{k}\; d^{k+3/2-\sqrt{2k+2}}. \tag{47}$$

By a uniformity argument, the expected number of particles arriving at a given vertex in level $k - j^*$ by step T_f is $T_f/d^{(k-j^*)} = \sqrt{kd}$. However, the expected number of particles arriving at a vertex w at level $k - j^* + 1$ by step T_f is

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

provided $k \ll d$, which we assume to be true. In order for w to be occupied, the sub-tree rooted at w must contain a path from w to level k consisting of j^* occupied vertices. By the Markov inequality the event that $j^* = \Theta(\sqrt{k})$ particles have arrived at w by step T_f, has probability $O(1/\sqrt{d})$. So there should be few vertices in level $k - j^* + 1$ with a path to level k containing j^* halted particles.

Consider t particles percolating downward from the root of an *infinite d-ary tree*. Particle s starts at step s and each particle transitions one edge at each step, so they never collide. Let w be a vertex in level ℓ of this tree, where the root vertex is at level zero. Let u be a vertex at level $\ell + j$ contained in the subtree rooted at w, and $H(w,u)$ the unique path $w = w_0 w_1 ... w_j = u$ from w in level ℓ to u in level $\ell + j$. Let $1 \le s_0 \le s_1 \le s_2 \le \cdots \le s_j \le t$ be particle labels, where particles s_0 and s_1 transition all edges of H, particle s_2 transitions edges $w_0 w_1 ... w_{j-1}$; in general s_i transitions $w_0 w_1 ... w_{j-i+1}$, and s_j transitions $w_0 w_1$. The probability of this is

$$P(s_0, s_1, ..., s_j) = \left(\frac{1}{d^\ell}\right)^{j+1} \frac{1}{d^j} \frac{1}{d d^2 \cdots d^j}.$$

The expected number of $j+1$ tuples of particles which transition as above is $\binom{t}{j+1} P(s_0, s_1, ..., s_j)$. There are d^j vertices u at level j in the subtree of w, and d^ℓ vertices w in level ℓ of the tree, so the number of sequences $Z_{\ell,j+1}(t)$ forming such a path between the levels has expectation

$$\mathbb{E}\, Z_{\ell,j+1}(t) = d^\ell d^j \binom{t}{j+1} P(s_0, s_1, ..., s_j) = \binom{t}{j+1} \frac{1}{d^{\ell+1} \cdots d^{\ell+j}}.$$

For $j \ll t$, with $\ell = k - j$ and $\mu_{k-j}(t)$ as given by (40) for the growing layers model,

$$\mathbb{E}\, Z_{k-j,j+1}(t) \sim \frac{t^{j+1}}{(j+1)!} \frac{1}{d^{k-j+1} \cdots d^k} = \mu_{k-j}(t). \tag{48}$$

Return now to the finite Cayley tree $G(k,d)$. For some vertex in level $k-j^*$ to be occupied at step t there must be some sequence $(s_0,s_1,...,s_j)$ which satisfies the construction given above. Indeed s_0 halts in level k, causing s_1 to halt in level $k-1$ and so on, until s_j halts in level $k-j$.

Let $j=j^*$ and $t=(1-\varepsilon)t_1(k-j^*)$ where $\varepsilon=\omega/\sqrt{2k+2}=\omega/(j^*+1)$, and $\omega<\sqrt{k}$,

$$\mu_{k-j^*}(t)=(1-\varepsilon)^{j^*+1}\le e^{-\omega}=o(1). \tag{49}$$

We conclude that at t' no such sequence exists w.h.p. and levels $0,1,...,k-j^*$ are empty. By (45), $t_1(k-j^*)=CT_f$, for some constant $C>0$, so the finish time $t_f\ge T_f/\omega$.

From (42), $\mu_{k-j+1}(t)=[(j+1)d^{k-j+1}/t]\cdot\mu_{k-j}(t)$, and so

$$\mu_{k-j^*+1}(t')=\frac{\sqrt{2k+2}\,d^{k-j^*+1}}{(1-\varepsilon)CT_f}\mu_{k-j^*}(t_1)=\Theta(\sqrt{d}).$$

Returning briefly to the infinite d-regular tree process, let $Z'(t)$ be the number of vertices w in level $k-j^*+1$ with j^* particles following a path $w=w_1w_2...w_j=u$ in the subtree rooted at w, plus another particle which passes through w, to any of its children. Then

$$\mathbb{E}\,Z'(t)\le\frac{t}{d^{k-j^*+1}}\mathbb{E}\,Z_{k-j^*+1,j^*}(t)=\frac{t}{d^{k-j^*+1}}\mu_{k-j^*+1}(t).$$

If $t=t_1(k-j^*)$, then for $k\ll d$

$$\mathbb{E}\,Z'(t_1)=\Theta(\sqrt{k})=o(\mu_{k-j^*+1}(t_1)).$$

Thus in expectation there are $(1-o(1))\mu_{k-j^*+1}(t_1)$ vertices w in level $k-j^*+1$ are *exact*. They have occupancy j^*, in their subtree and a unique occupied path to level k. The events that two such vertices w,w' have this path property are independent and we conclude that the number of such vertices is concentrated around its mean at $\Theta(\sqrt{d})$.

From now on the proof mirrors that of the growing layers model in Section 4.1. Thus w.h.p. within a most $2t_1$ steps the first occupancy of a vertex in level $k-j^*$ has occured; and a path grows back to the source from this vertex, halting the process before the second occupancy in level $k-j^*$ can occur.

5.1 Theorem 2.(2): DLA on trees with branching factor $d\ge 2$

Proposition 11. *For $d\ge 2$, let $G=G(k,d)$ be the Cayley tree with branching factor d and height k, where $d^k=n$. Let $T=T(G)$ be given by*

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

Let t_f be the finish time of DLA on G. Then w.h.p. $T/\omega\le t_f\le\omega T$.

Proof. The argument leading to (49) in the previous section only assumed $k \to \infty$ so that we could find some $\omega < \sqrt{k}$ and where $\omega \to \infty$. As $d^k = n$ then $k = \log n/(\log d)$ which is is monotone increasing with deceasing d. So $k \to \infty$ as before and the lower bound on t_f follows from (49) on choosing $t' = (1-\varepsilon)t_1(k-j^*)$, and noting that $t' \approx (1-\varepsilon)T/e > T/\omega$.

For the upper bound, let $j = j^* - h$ where $h = \lceil 1/2 + \log_d k \rceil$. In Section 5, for the upper bound, we assumed that $k \ll d$, in which case $h = 1$, which led to an argument about the occupancy of levels $k - j^*$ and $k - j^* + 1$. In this section the assumption $k \ll d$ may no longer hold, for example, when $d = 2$, $k = \log n/\log 2$.

The value of μ_{k-j^*} given by (48), satisfies $\mu_{k-j+1}(t) = [(j+1)d^{k-j+1}/t] \cdot \mu_{k-j}(t)$. Iterating this, and evaluating at $t_1(k-j^*) = (1+O(1/j))e^{-1}\sqrt{2k+2}d^{k-j^*+1/2}$ from (45),

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

This ratio is $\omega(1)$, as either (i) $d \to \infty$, or (ii) d is constant and then $\log_d k = \log_d \log_d n$, implying that $h = \omega(1)$. As before, with $\mathbb{E}\, Z_{k-j^*+h, j^*-h+1}(t) = \mu_{k-j^*+h}(t)$, consider $Z'(t_1)$, where

$$\mathbb{E}\, Z'(t_1(k-j^*)) \leq \frac{t_1}{d^{k-j^*+h}} \mu_{k-j^*+h}(t_1).$$

The ratio $\rho = t_1(k-j^*)/d^{k-j^*+h}$ satisfies

$$\rho = t_1(k-j^*)/d^{k-j^*+h} = \Theta(1)\frac{\sqrt{k}}{d^{h+1/2}} = \left(\frac{k}{d^{2h+1}}\right)^{1/2}.$$

Assume $k \leq d$, then $2h+1 \geq 3$, so $\rho \leq 1/d$. Alternatively, if $d \leq k$, then $h \geq \log_d k$ and $\rho \leq 1/k$. In either case w.h.p. $(1-o(1))\mu_{k-j^*+h}(t)$ vertices in level $k - j^* + h$ are exact at around $t = t_1$.

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 at t_1. In the latter case, w.h.p. there exists $\Theta(e^h d^{h^2/2})$ exact paths from level $k - j^* + h$ to level k. In expectation, it takes at most

$$\Theta(1)\frac{d^{k-j^*+h}}{e^h d^{h^2/2}} \leq T$$

further steps for one of these paths to extend back to the source. The upper bound now follows from the Markov inequality. □

Proof of Theorem 2.(2). If d is constant, $d^k = n$ implies $k = \log n/\log d$. Hence $\sqrt{2k+2} = \sqrt{2k} + \Theta(1)$, and $d^{3/2} = \Theta(1)$. Thus t_1, T are both $\Theta(T_f)$, where T_f is as given.

References

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

[2] M. Batty, P. Longley and S. Fotheringham. Urban growth and form: scaling, fractal geometry, and diffusion-limited aggregation. Environment and Planning A, 21, 1447-1472 (1989).

[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(6), 3528–3557, (2018).

[5] T. C. Halsey. Diffusion limited aggregation: A model for pattern formation. Physics Today 53 (11), 36–41, (2000).

[6] M. B. Hastings and T. C. Halsey. High-dimensional diffusive growth, Europhys. Lett., 55(5), 679—685 (2001).

[7] 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)

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

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

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

[11] L. Niemeyer, L. Pietronero, and H. J. Wiesmann. Fractal dimension of dielectric breakdown. Phys. Rev. Lett., 52(12):1033–1036, (1984).

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

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

Appendix

The effect of rounding error in the growing layers model. We examine conditions on d, k which allow us to effectively ignore the rounding error on j^* in the value of T_f, and show that $(\log d)/\sqrt{k} = O(1)$ suffices. In the case that j^* is not integer, we require the maximum j such that $2k+2 > j(j+1)$; see the exponent of d in (43). 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 (44), 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 $\Theta(1) d^{-O(1/\sqrt{k})}$. Thus $(\log d)/\sqrt{k} = O(1)$ suffices.