

Explosive growth for a constrained Hastings–Levitov aggregation model

Nathanaël Berestycki
 University of Vienna
 nathanael.berestycki@univie.ac.at

Vittoria Silvestri
 University of Rome La Sapienza
 silvestri@mat.uniroma1.it

Abstract

We consider a constrained version of the HL(0) Hastings–Levitov model of aggregation in the complex plane, in which particles can only attach to the part of the cluster that has already been grown. Although one might expect that this gives rise to a non-trivial limiting shape, we prove that the cluster grows explosively: in the upper half plane, the aggregate accumulates infinite diameter as soon as it reaches positive capacity. More precisely, we show that after nt particles of (half-plane) capacity $1/(2n)$ have attached, the diameter of the shape is highly concentrated around $\sqrt{t \log n}$. This illustrates a new instability phenomenon for the growth of single trees/fjords in unconstrained HL(0).

1 Introduction

We study the growth of an aggregate on the upper half-plane, described by the composition of random conformal maps, which is a modification of the well known Hastings–Levitov HL(0) aggregation model. Let us succinctly describe it. Each map represents the arrival of a new particle, and the randomness comes from the attachment locations on the boundary of the aggregate.

Denote by $\mathbb{H} = \{z \in \mathbb{C} : \Im(z) > 0\}$ the upper half-plane. For $n \geq 1$ let $F : \mathbb{H} \rightarrow \mathbb{C}$ be the slit map

$$F(z) = \sqrt{z^2 - 1/n},$$

which grows a vertical slit of length $1/\sqrt{n}$ and half-plane capacity $1/(2n)$ at 0 (see Figure 1 below).

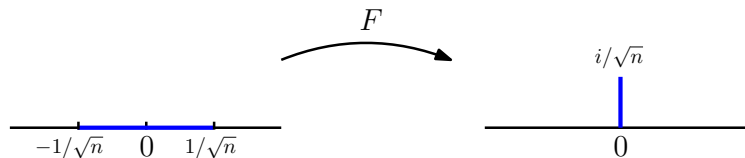


Figure 1: The map F growing a slit of height $1/\sqrt{n}$ at the origin.

Then, if $x \in \mathbb{R}$, the map

$$F_x(z) = x + \sqrt{(z-x)^2 - 1/n}$$

grows a vertical slit of height $1/\sqrt{n}$ at x . We use this to define a growth model as follows. Let $\Phi_0(z) = z$, and denote¹ by $\mathcal{K}_1 = \overline{\mathbb{H} \setminus \Phi_1(\mathbb{H})}$ the closure of the complement of $\Phi_1(\mathbb{H})$ in \mathbb{H} , so that

¹For a set $\mathcal{K} \subseteq \mathbb{H}$ and a map $f : \mathbb{H} \rightarrow \mathbb{C}$, we write $f(\mathcal{K}) = \{f(z) : z \in \mathcal{K}\}$.

\mathcal{K}_1 is a compact set consisting of a vertical slit at 0. Denote by $\Gamma_1(z) = \Phi_1^{-1}(z)$ the associated mapping-out function. Then Γ_1 maps the slit \mathcal{K}_1 to the interval $[-L_1, R_1]$ on the real line, with $L_1 = R_1 = 1/\sqrt{n}$. To the initial cluster \mathcal{K}_1 we attach particles as follows. Suppose inductively that for $k \geq 1$ we have defined the map Φ_k that grows a random compact set \mathcal{K}_k , which we think of as the cluster up to the k^{th} arrival. Let $\Gamma_k = \Phi_k^{-1}$ denote the associated mapping out function, and assume that $\Gamma_k(\mathcal{K}_k)$ is an interval on the real line, call it $[-L_k, R_k]$. Then choose a uniformly random point x_{k+1} on $[-L_k, R_k]$, and set

$$\Phi_{k+1} = \Phi_k \circ F_{x_k}, \quad \mathcal{K}_{k+1} = \overline{\mathbb{H} \setminus \Phi_{k+1}(\mathbb{H})}. \quad (1.1)$$

Note that $\mathcal{K}_{k+1} = \mathcal{K}_k \cup P_{k+1}$ where $P_{k+1} = \overline{\Phi_k([x_{k+1}, x_{k+1} + i/\sqrt{n}])}$ represents the $(k+1)^{\text{th}}$ particle. Unfolding the recursion, we see that this growth mechanism corresponds to setting

$$\Phi_k(z) = F_{x_1} \circ F_{x_2} \circ \dots \circ F_{x_k}(z),$$

for $x_1, x_2 \dots x_k$ random points with $x_1 = 0$ and $x_{j+1} \sim \text{Uniform}[-L_j, R_j]$ for $j \geq 1$. We refer to x_k as the attachment location of the k^{th} particle P_k . Note that $L_k, R_k \geq 0$ for all $k \geq 1$ by construction.

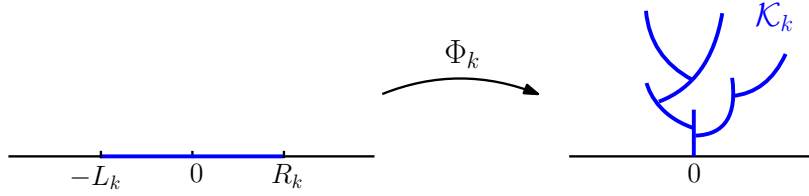


Figure 2: The map Φ_k maps the interval $[-L_k, R_k]$ to the cluster \mathcal{K}_k with k particles.

The above model is a variant of the well-known Hastings–Levitov model $\text{HL}(0)$, where the growth usually takes place outside the unit disc $\mathbb{D} = \{|z| \leq 1\}$, rather than on the upper half-plane \mathbb{H} [5, 14, 15] (see however [3] by Berger, Procaccia and Turner, where this model is introduced in the upper half-plane). We recall that the Hastings–Levitov model was originally proposed in [5] as a continuum version of the celebrated Diffusion Limited Aggregation (DLA) model on the lattice, which ought to be more amenable to analysis thanks to its formulation in terms of conformal maps.

We call this new model **constrained Hastings–Levitov**, as the particles are only allowed to attach on the boundary of the grown region, rather than everywhere on the real line. In other words, the growth of the cluster is restricted to a single tree. See Figure 3 for a simulation of constrained $\text{HL}(0)$ outside the unit disc.

1.1 Main results

The question is to describe the shape of the cluster \mathcal{K}_n for large n . This is implicitly obtained by solving the Loewner equation

$$\begin{cases} \dot{f}_t^{(n)}(z) = \partial_z f_t^{(n)}(z) \int_{\mathbb{R}} \frac{1}{x-z} \mu_t^{(n)}(dx) \\ f_0^{(n)}(z) = z \end{cases} \quad (1.2)$$

driven by the random measure

$$\mu_t^{(n)}(dx) = \sum_{k \geq 0} 1\left(t \in \left[\frac{k}{n}, \frac{k+1}{n}\right)\right) \delta_{x_k}(dx), \quad (1.3)$$

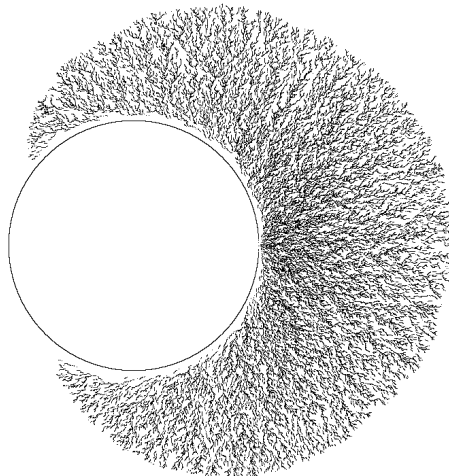


Figure 3: Restricted HL(0) on the outside of the unit disc. To (approximately) quote Magritte, this is not a shape theorem.

where δ_x denotes the Dirac's delta measure centered at $x \in \mathbb{R}$, and x_1, x_2, \dots denote the attachment locations of subsequent particles on the real line. Note that, by definition of the model, x_k is uniformly distributed in the interval $[-L_{k-1}, R_{k-1}]$, so we start by describing the asymptotic behaviour of this random interval for large n .

To state the main result, write $\lambda_k = L_k + R_k$ for the length of the interval after k arrival. As the half-plane capacity of each particle is $1/(2n)$ and half-plane capacity is additive [11], we obtain a half-plane capacity of order $O(1)$ when $k = tn$ particles are added to the aggregate. Based on this simple observation we typically expect a nontrivial shape in this regime; this is in particular what happens for unconstrained HL(0), and seems consistent with the simulations in Figure 3. However this turns out to not be the case: indeed our results show that after nt particles the diameter diverges like $\sqrt{t \log n}$. To prove this, the key is to understand the size of the interval in which particles are allowed to attach after we map out the cluster. This is specified by the result below:

Theorem 1.1. *For any $\varepsilon > 0$ and $T > 0$ we have*

$$\mathbb{P} \left(\sup_{t \in [(\log n)^9/n, T]} \left| \frac{R_{nt}}{\sqrt{t \log n}} - \frac{1}{2} \right| \leq \varepsilon \right) \geq 1 - e^{-c(\log n)^2}$$

for n large enough ($n \geq n_0(\varepsilon, T)$). Here $c = c(\varepsilon, T)$ is a constant which may depend on ε and T but not on n . Moreover, the same holds with L_{nt} in place of R_{nt} . In particular

$$\mathbb{P} \left(\sup_{t \in [(\log n)^9/n, T]} \left| \frac{\lambda_{nt}}{\sqrt{t \log n}} - 1 \right| \leq \varepsilon \right) \geq 1 - 2e^{-c(\log n)^2} \quad (1.4)$$

for n large enough.

Note in particular that after nt arrivals, even though the capacity of the cluster is of unit order, the length of the interval tends to infinity!

1.2 Geometric consequences and conjectures

We use Theorem 1.1 to deduce geometric information on the cluster \mathcal{K}_{nt} , showing that it has diverging diameter in n .

Theorem 1.2. *Let $\text{Diam}(\mathcal{K}_{nt}) = \sup\{|x - y| : x, y \in \mathcal{K}_{nt}\}$ denote the diameter of \mathcal{K}_{nt} . Then*

$$\frac{\text{Diam}(\mathcal{K}_{nt})}{\sqrt{t \log n}} \rightarrow 1$$

in probability. Furthermore,

$$\max_{z \in \mathcal{K}_{nt}} \{\Im(z)\} \leq \sqrt{2t},$$

almost surely.

The above theorem says that the shape of the cluster is very elongated, being of order $O(1)$ in height and of order approximately $\sqrt{t \log n}$ in width.

It is tempting to conjecture that the cluster shape (or better its complement in the upper half-plane) is more precisely described by the following Loewner–Kufarev evolution: denote by μ_t the uniform measure on the interval $[-\frac{\sqrt{t \log n}}{2}, \frac{\sqrt{t \log n}}{2}]$, and let f_t be the deterministic map solution to the Loewner equation

$$\begin{cases} \dot{f}_t(z) = \partial_z f_t(z) \int_{\mathbb{R}} \frac{1}{x - z} \mu_t(dx) \\ f_0(z) = z. \end{cases} \quad (1.5)$$

Note that both μ_t and f_t depend on n , although we did not make this explicit in the notation.

Recall the conformal map Φ_k , defined in (1.1), which maps the upper half plane \mathbb{H} to $H_k = \mathbb{H} \setminus \mathcal{K}_k$ and grows the restricted Hastings–Levitov cluster \mathcal{K}_k up to the k^{th} particle. Then Φ_k can be described by a Loewner–Kufarev evolution driven by the random measure $\mu_t^{(n)}$ defined in (1.3). It is a consequence of Theorem 1.1 that $\mu_t^{(n)}$ is close, in a suitable sense, to the deterministic measure μ_t defined above. Namely, one can show that for any $\varepsilon, t > 0$ and any function $\varphi \in C_b^{1,1}([0, t] \times \mathbb{R})$ there exists a constant $C = C(t, \|\varphi\|_{\infty}, \|\dot{\varphi}\|_{\infty})$ such that

$$\mathbb{P} \left(\left| \int_0^t \int_{\mathbb{R}} \varphi(s, x) \mu_s^{(n)}(dx) - \int_0^t \int_{\mathbb{R}} \varphi(s, x) \mu_s(dx) \right| \leq C\varepsilon \text{ eventually in } n \right) = 1. \quad (1.6)$$

It is therefore natural to guess that Φ_{nt} is close to the conformal map f_t (we note that a related continuity statement of the Loewner–Kufarev equation was obtained by [1], [6] and [7]). Indeed, using (1.6) together with a uniform continuity of the Loewner–Kufarev evolution, one can prove the following statement: for any $\varepsilon, t > 0$ and compact set $K \subset \mathbb{H}$ there exists a constant $C(t, K)$ such that

$$\mathbb{P} \left(\sup_{z \in K} |\Phi_{nt}(z) - f_t(z)| \leq C(t, K) \varepsilon \text{ eventually in } n \right) = 1. \quad (1.7)$$

However, the conformal map f_t , although deterministic, still depends on n and is in the pointwise limit as $n \rightarrow \infty$ close to the trivial identity map, so this statement in itself does not convey useful information about the geometry of the cluster: it simply tells us that Φ_{nt} is close to the identity map for sufficiently large n .

In fact, formulating a precise statement on how close Φ_{nt} needs to be to f_t in order to capture meaningful geometric information on \mathcal{K}_{nt} is in itself not entirely trivial. At the very least one would need to prove that the size of the difference between Φ_{nt} and f_t is smaller than the size of the discrepancy between f_t and the identity map: that is, one would need to show that

$$|\Phi_{nt}(z) - f_t(z)| = o(|f_t(z) - z|),$$

say uniformly on a compact. In other words, (1.7) needs to be formulated quantitatively. As a necessary first step, one would therefore need to approximate $|f_t(z) - z|$. We believe its order of magnitude can be guessed from Theorem 1.2: indeed, since the half-plane capacity of \mathcal{K}_{nt} is t and the diameter is approximately $\sqrt{t \log n}$, then (using for instance the geometric interpretation of half-plane capacity, see for instance [10]), it is natural to guess that the typical height of the cluster is of order $\sqrt{t/\log n}$ for a fixed time t . Therefore we conjecture that

$$|f_t(z) - z| \asymp \sqrt{t/\log n} \tag{1.8}$$

for a fixed t and a fixed z , where the implied constants do not depend on $t > 0$ and z in some given compact set.

Simulations. To complement this discussion we include some simulations of our restricted HL(0) aggregation model on the upper half plane. These simulations do not show the entire cluster but only an envelope of the cluster: roughly speaking they show the conformal image of $\{z \in \mathbb{H} : \Im(z) = \varepsilon\}$ under the conformal map Φ_{nt} sending \mathbb{H} to $\mathbb{H} \setminus \mathcal{K}_{nt}$, for some small $\varepsilon > 0$, and for various values of n and t ; in fact they are parametrised instead by N (the number of particles) and the individual size (i.e., length) which is called d in these pictures (or equivalently, half-plane capacity $d^2/2$). The relation between these parameters is as follows:

$$\begin{aligned} d^2 &= 1/n, \\ N &= t/d^2, \end{aligned}$$

where N is the number of particles.

We also include for these same simulations the image of the cluster under the conformal map Φ_k^{-1} , which is the interval $[-L_k, R_k]$ of the real line, described by Theorem 1.1 and crucial to the arguments of this paper. The simulations also show the actual locations of where particle successively attach.

1.3 Aggregation outside the unit disc

We have so far focused for convenience and ease on the situation in the upper-half plane, but it is in fact more standard to consider HL(0) as an aggregation model outside the unit disc (the recent paper [3] by Berger, Procaccia and Turner being the exception). The model is defined similarly by replacing the upper-half plane \mathbb{H} with the unit disc \mathbb{D} , and the elementary maps describing the slits are obtained from F by conjugating with respect to a conformal map from $\mathbb{C} \setminus \mathbb{D}$ to \mathbb{H} . See e.g. [14] for details.

In order to keep the paper short and simple, we decided not to include results related to this case, although it is not hard to see that our techniques apply to this setting as well, at least as long as the two arms of the cluster remain macroscopically far apart (i.e., they do not shield the unit disc almost entirely). In particular, with relatively little effort it should be possible to show the following result: for any $0 < \delta < 1$, let τ_δ denote the first $k \geq 1$ such that the k th particle attaches at a position whose argument is in $[\pi(1 - \delta), \pi(1 + \delta)]$. Then, for any $0 < \delta < 1$, as $n \rightarrow \infty$,

$$\tau_\delta \sim \frac{4\pi^2 n(1 - \delta)^2}{\log n} \tag{1.9}$$

in probability, where $X_n \sim Y_n$ in probability as $n \rightarrow \infty$ means $X_n/Y_n \rightarrow 1$ in probability. Observe that the surprising phenomenon here is that $\tau_\delta = o(n)$.

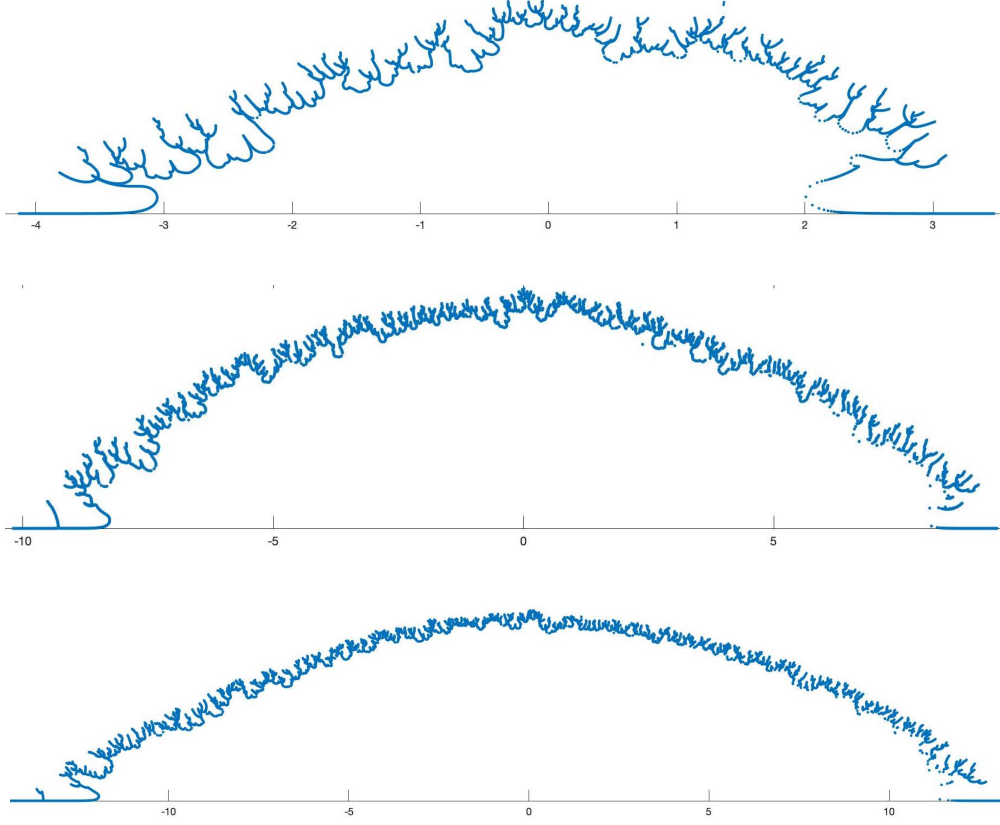


Figure 4: Simulations of cluster envelope at $\varepsilon = 10^{-5}$ with $N = 1000, 5000$ and 10000 with $d = 0.1$. This corresponds to $n = 10^2$ and $t = 10, 50$ and 100 respectively.

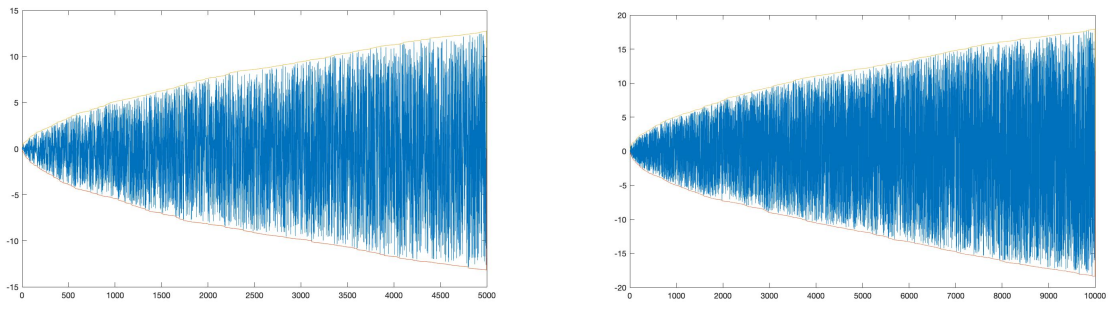


Figure 5: Simulations of the allowed interval with $N = 5000$ and 10000 (second and third simulations in Figure 4).

1.4 Heuristic relation to instability in unconstrained Hastings–Levitov

Compared to the unconstrained model (especially the disc one discussed in Section 1.3), the only difference with our constrained model is that particles can only attach to the tree generated by the first particle. In other words, the constrained model describes the growth of a *single* tree. From this perspective, the explosive growth demonstrated in Theorem 1.1 shows a fundamental instability property which seems not to have been noticed before: left to its own devices, a single tree or fjord grows explosively. The existence of a shape theorem for regular (i.e., unconstrained) HL(0) appears from this point of view nothing short of miraculous: essentially, this suggests that multiple fjords and subtrees are “canceling each other out” from the point of view of conformal maps.

We believe that this effect becomes less pronounced when particle sizes are rescaled according to the derivative of the conformal map at the location where particles attach – this is for instance the situation of the HL(α), for $\alpha > 0$ (see e.g. [8], [13] and [12] for a definition and some fine properties of the model).

It has been proposed (and indeed proved for $\alpha < 1$ in [13] for the unconstrained model) that compensating the size of the particles by a power of the derivative of the conformal map has an overall identical effect as keeping the size fixed but reweighing the law of the attachment by a power of the harmonic measure. In the constrained model however it is conceivable that these two models behave differently for every $\alpha > 0$. Furthermore, in either version it is entirely unclear whether the constrained models remain explosive for positive values of α . Understanding this might shed some light on the expected transition between the smooth and turbulent phases of HL(α) at $\alpha = 1$ predicted in the original paper of Hastings and Levitov [5], and so ultimately on the fundamental problem to show the emergence of fractality in models related to DLA (see, e.g., [9]).

1.5 Sketch of proof of Theorem 1.1 and heuristics

The main argument for proving Theorem 1.1 will consist in the following estimate. Call $r(t) = R_{nt}/\sqrt{\log n}$. We wish to show that $r(t) \asymp \sqrt{t}$. Roughly speaking we show that

$$\frac{dr}{dt} \approx \frac{1}{8r(t)}. \quad (1.10)$$

This comes from a dyadic decomposition where we consider the contribution to the position of the front R_{nt} of particles arriving at time $k \in [nt, n(t+h)]$ in a dyadic interval $I_j = [R_k - 2^{j+1}, R_k - 2^j]$ of length 2^j at distance 2^j from the right front R_k . Here one should think that 2^j is much smaller than n , but at least $1/\sqrt{n}$; with the minimal distance $1/\sqrt{n}$ determined by the behaviour of the conformal map which uniformises a single particle. We show that the contribution to the push of the front of scale j during this interval is, roughly,

$$\Delta_j \approx nh \times \frac{|I_j|}{2r\sqrt{\log n}} \times \frac{\log 2}{n2^{j+1}} \approx \frac{h \log 2}{4r\sqrt{\log n}}$$

where the first term nh is the total number of particles arriving during the interval, the second term corresponds to the probability for a particle to fall in I_j , and the last term gives the average push caused by such a particle. Since the right hand side does not depend on j , we see that all scales contribute equally. This is the reason beyond the explosive growth.

Summed over all scales, this gives

$$R_{(t+h)n} - R_{tn} \approx \sum_j \Delta_j \approx \frac{h\sqrt{\log n}}{8r(t)}$$

where the constant 8 comes from the fact the smallest dyadic scale starts at distance $1/\sqrt{n}$ from the tip of the interval. Dividing by $\sqrt{\log n}$ to express the left hand side in terms of r , this suggests that $r(t)$ satisfies the differential equation (1.10).

Equivalently, starting from very small $r(0)$, we show that the time (measured in units of $1/n$) for r to double is roughly r^2 (that is, we must add roughly $r^2 n$ particles). More precisely, starting from an initial value $r(0) = r_n$, the solution to this ODE is

$$r(t) = (1/2)\sqrt{t + r_n^2/4},$$

so that, for arbitrary $\delta > 0$, the time it takes to reach $r_n(1 + \delta)$ is

$$t = 4r_n^2((1 + \delta)^2 - 1) = 4r_n^2(2\delta + \delta^2) \approx 8\delta r_n^2.$$

2 Proof of Theorem 1.1

Fix $t \in (0, T]$ and $\varepsilon > 0$ as in the statement of Theorem 1.1. Let $\delta > 0$ denote a constant to be chosen later, possibly depending on ε . Recall that for $1 \leq k \leq nT$ we denote by λ_k the length of the interval of the cluster under the mapping out function Γ_k , so that $\lambda_k = L_k + R_k$. Define

$$l(n) = \frac{(\log n)^3}{\sqrt{n}},$$

and introduce the stopping times

$$\begin{aligned} T_1 &= \inf\{i \geq 1 : \lambda_i \geq l(n)\}, \\ T_k &= \inf\{i \geq T_{k-1} : \lambda_i \geq (1 + \delta)\lambda_{T_{k-1}}\}, \quad k \geq 2. \end{aligned}$$

We start by showing that, once the length of the interval exceeds $l(n)$, the time it takes for it to grow by a factor $1 + \delta$ cannot be too small.

Proposition 2.1 (Lower bound on doubling time). *For any $k \geq 2$ it holds*

$$\mathbb{P}\left(T_k - T_{k-1} \geq \left(\frac{2\delta}{1 + \varepsilon}\right) \frac{n}{\log n} \lambda_{T_{k-1}}^2 \middle| \mathcal{F}_{T_{k-1}}\right) \geq 1 - e^{-c(\log n)^2}$$

for n large enough. Here $c = c(\varepsilon, \delta)$ is a constant which may depend on ε and δ but not on n nor k .

Proof. Write $m_k = \left(\frac{2\delta}{1 + \varepsilon}\right) \frac{n}{\log n} \lambda_{T_{k-1}}^2$ for brevity. Then

$$\begin{aligned} \mathbb{P}\left(T_k - T_{k-1} < m_k \middle| \mathcal{F}_{T_{k-1}}\right) &\leq \mathbb{P}(\lambda_{T_{k-1} + m_k} \geq (1 + \delta)\lambda_{T_{k-1}}) \\ &\leq \mathbb{P}\left(R_{T_{k-1} + m_k} \geq R_{T_{k-1}} + \frac{\delta}{2}\lambda_{T_{k-1}} \middle| \mathcal{F}_{T_{k-1}}\right) \\ &\quad + \mathbb{P}\left(L_{T_{k-1} + m_k} \geq L_{T_{k-1}} + \frac{\delta}{2}\lambda_{T_{k-1}} \middle| \mathcal{F}_{T_{k-1}}\right). \end{aligned}$$

We bound the first term in the right hand side. The second term can be bounded using exactly the same argument.

The idea is to decompose over dyadic scales $[2^j, 2^{j+1})$ corresponding to the distance between the right end of the interval and the position of the particle that attaches. Define

$$\begin{aligned} j_{min} &= \inf\left\{j \in \mathbb{Z} : 2^j \geq \frac{1}{\sqrt{n}}\right\}, \\ j_{max} &= \inf\left\{j \in \mathbb{Z} : 2^j \geq \sqrt{T \log n}\right\}. \end{aligned}$$

For $i \geq 1$ and $j_{\min} \leq j \leq j_{\max}$, let

$$\begin{aligned} I_{i,\min} &= (R_i - 2^{j_{\min}}, R_i], \\ I_{i,j} &= (R_i - 2^{j+1}, R_i - 2^j]. \end{aligned}$$

These intervals actually depend on time and are random, but they have fixed width $|I_{i,j}| = 2^j$, and a given particle arriving at time $i+1$ attaches in one and only interval $I_{i,j}$ for some j .

Let us describe the effect that a new arrival x_{i+1} at time $i+1$ has on the interval $[-L_i, R_i]$.

Lemma 2.2. *Suppose a new particle attaches at time i at position $x \in I_{i,j}$ with $j \geq j_{\min}$, so that x is at least distance $1/\sqrt{n}$ from R_i . Then we have deterministically:*

$$R_{i+1} - R_i \leq \frac{1}{2n(R_i - x)}$$

and

$$R_{i+1} - R_i \geq \frac{1}{2n(R_i - x)} - \frac{C}{(R_i - x)^3 n^2},$$

for C absolute constant.

Proof. Observe that R_{i+1} is obtained from R_i by applying the map

$$w \mapsto x + \sqrt{(w - x)^2 + 1/n}$$

so that

$$R_{i+1} = x + \sqrt{z^2 + 1/n}, \text{ with } z = R_i - x.$$

By convexity of the square root function,

$$\begin{aligned} R_{i+1} &\leq x + \sqrt{z^2 + 1/n} \\ &\leq x + z \left(1 + \frac{1}{2z^2 n}\right) = x + z + \frac{1}{2zn} \\ &= R_i + \frac{1}{2zn} = R_i + \frac{1}{n(R_i - x)}, \end{aligned}$$

which gives the desired upper bound. On the other hand, for the lower bound, by a Taylor expansion of $\sqrt{1+x}$ near zero,

$$R_{i+1} \geq R_i + 1/(2nz) - C/(n^2 z^3)$$

(which is valid since $z \geq 1/\sqrt{n}$ by assumption on x). Since $z = R_i - x$ we get the desired inequality. \square

Lemma 2.3. *In the same setting as above the following hold:*

$$\begin{aligned} \mathbb{E}(R_{i+1} - R_i | x_{i+1} \in I_{i,j}) &\leq \frac{\log 2}{n2^{j+1}}, \\ \mathbb{E}(R_{i+1} - R_i | x_{i+1} \in I_{i,j}) &\geq \frac{\log 2}{n2^{j+1}} - \frac{C}{n^2 2^{3j}}, \\ \text{Var}(R_{i+1} - R_i | x_{i+1} \in I_{i,j}) &\leq \frac{C}{n^2 2^{2j}} \end{aligned}$$

with C absolute constant that may change from line to line.

Proof. For the first upper bound, we use Lemma 2.2 and note that, conditionally on $x_{i+1} \in I_{i,j}$, x_{i+1} is uniformly distributed over $I_{i,j}$, so that

$$\mathbb{E}(R_{i+1} - R_i | x_{i+1} \in I_{i,j}) \leq \int_{2^j}^{2^{j+1}} \frac{dx}{2nx2^j} = \frac{\log 2}{n2^{j+1}},$$

as desired. For the lower bound, we proceed similarly and note that

$$\begin{aligned} \mathbb{E}(R_{i+1} - R_i | x_{i+1} \in I_{i,j}) &\geq \int_{R_i - 2^{j+1}}^{R_i - 2^j} \frac{1}{2(R_i - x)} \frac{dx}{n2^j} - \int_{R_i - 2^{j+1}}^{R_i - 2^j} \frac{C}{(R_i - x)^3} \frac{dx}{n2^{2j}} \\ &\geq \frac{\log 2}{n2^{j+1}} - \frac{C}{n2^{2j}} \int_{2^j}^{2^{j+1}} \frac{du}{u^3} \geq \frac{\log 2}{n2^{j+1}} - \frac{C}{n^2 2^{3j}} \end{aligned}$$

as desired. Finally,

$$\begin{aligned} \text{Var}(R_{i+1} - R_i | x_{i+1} \in I_{i,j}) &\leq \mathbb{E}((R_{i+1} - R_i)^2 | x_{i+1} \in I_{i,j}) \\ &\leq \int_{2^j}^{2^{j+1}} \frac{1}{n^2 x^2} \frac{dx}{2^{j+1}} \leq \frac{C}{n^2 2^{2j}}, \end{aligned}$$

as desired. \square

Let us now go back to the proof of Proposition 2.1. For $i \in [T_{k-1} + 1, T_{k-1} + m_k]$ let x_i denote the particles attachment locations, and for $j_{\min} \leq j \leq j_{\max}$, let

$$\Delta_j = \sum_{i=T_{k-1}}^{T_{k-1}+m_k-1} 1_{\{x_{i+1} \in I_{i,j}\}} (R_{i+1} - R_i)$$

denote the push to the right front caused by particles arriving in $I_{i,j}$ during the phase $[T_{k-1}, T_{k-1} + m_k]$ (recall that the actual positions of the intervals $I_{i,j}$ change at each step, but their length is fixed). Note that

$$\begin{aligned} \mathbb{E}(\Delta_j | \mathcal{F}_{T_{k-1}}) &= \sum_{i=T_{k-1}}^{T_{k-1}+m_k-1} \mathbb{E}(1_{\{x_{i+1} \in I_{i,j}\}} (R_{i+1} - R_i) | \mathcal{F}_{T_{k-1}}) \\ &= \sum_{i=T_{k-1}}^{T_{k-1}+m_k-1} \mathbb{E}(R_{i+1} - R_i | x_{i+1} \in I_{i,j}) \mathbb{P}(x_{i+1} \in I_{i,j}) \\ &\leq \sum_{i=T_{k-1}}^{T_{k-1}+m_k-1} \frac{\log 2}{n2^{j+1}} \frac{|I_{i,j}|}{\lambda_{T_{k-1}}} = \frac{\log 2}{2n\lambda_{T_{k-1}}} m_k = \left(\frac{\delta}{1+\varepsilon} \right) \frac{\lambda_{T_{k-1}}}{\log_2 n}. \end{aligned}$$

Lemma 2.4. For $k \geq 2$, $\delta > 0$ and any $\varepsilon \in (0, 1]$ it holds

$$\mathbb{P} \left(\sum_{j=j_{\min}}^{j_{\max}} \Delta_j < \frac{\delta}{2} \left(1 - \frac{\varepsilon}{8}\right) \lambda_{T_{k-1}} \middle| \mathcal{F}_{T_{k-1}} \right) \geq 1 - e^{-c(\log n)^2}$$

for n large enough ($n \geq n_0(\varepsilon, T)$), where $c = c(\delta, \varepsilon)$ is a constant that may depend on δ and ε .

Proof. For $j_{\min} \leq j \leq j_{\max}$ set

$$D_j := \Delta_j - \left(\frac{\delta}{1+\varepsilon} \right) \frac{\lambda_{T_{k-1}}}{\log_2 n}.$$

In words, D_j is obtained by subtracting from Δ_j an upper-bound on the conditional expectation of the push for each particle using Lemma 2.3. Therefore, for each $j_{\min} \leq j \leq j_{\max}$, D_j is the terminal value at time m_k of a supermartingale. Furthermore, its total quadratic variation Q_j at time m_k (i.e. the sum of the conditional variances of the increments) is bounded by

$$\begin{aligned} Q_j &\leq \sum_{i=T_{k-1}}^{T_{k-1}+m_k-1} \mathbb{E}((R_{i+1} - R_i)^2 | x_{i+1} \in I_{i,j}) \mathbb{P}(x_{i+1} \in I_{i,j} | \mathcal{F}_i) \\ &\leq \sum_{i=T_{k-1}}^{T_{k-1}+m_k-1} \frac{C}{n^2 2^{2j}} \frac{|I_{i,j}|}{\lambda_{T_{k-1}}} = \frac{C}{n^2 2^{2j}} \frac{2^j}{\lambda_{T_{k-1}}} m_k \\ &= C \left(\frac{2\delta}{1+\varepsilon} \right) \frac{\lambda_{T_{k-1}}}{2^j n \log n} =: b, \end{aligned}$$

where the last equality defines b . Note also that each jump is bounded by $K = \frac{1}{n^{2j}}$. Applying Freedman's martingale inequality (Proposition 2.1 in [4]) we deduce that, with

$$a = \frac{\delta \lambda_{T_{k-1}}}{\log_2 n} \left(\frac{1}{1+\varepsilon/2} - \frac{1}{1+\varepsilon} \right)$$

(note that $Ka \asymp b$ as $n \rightarrow \infty$),

$$\begin{aligned} \mathbb{P} \left(\Delta_j \geq \frac{\delta \lambda_{T_{k-1}}}{\log_2 n} \frac{1}{(1+\varepsilon/2)} \middle| \mathcal{F}_{T_{k-1}} \right) &= \mathbb{P}(D_j \geq a, Q_j \leq b | \mathcal{F}_{T_{k-1}}) \\ &\leq \exp \left(-\frac{a^2}{Ka+b} \right) \leq \exp \left(-c \frac{n 2^j \lambda_{T_{k-1}}}{\log n} \right) \leq e^{-c(\log n)^2} \end{aligned}$$

since $2^j \geq 1/\sqrt{n}$ and $\lambda_{T_{k-1}} \geq \lambda_{T_1} \geq l(n)$. Here c is a constant that may depend on δ and ε . This bound can be summed over all scales, to get that

$$\begin{aligned} \mathbb{P} \left(\exists j \in [j_{\min}, j_{\max}] : \Delta_j \geq \frac{\delta \lambda_{T_{k-1}}}{(\log_2 n)(1+\varepsilon/2)} \middle| \mathcal{F}_{T_{k-1}} \right) \\ \leq (j_{\max} - j_{\min}) \mathbb{P} \left(\Delta_j \geq \frac{\delta \lambda_{T_{k-1}}}{(\log_2 n)(1+\varepsilon/2)} \middle| \mathcal{F}_{T_{k-1}} \right) \\ \leq (\log_2 n) e^{-c(\log n)^2} \leq e^{-c'(\log n)^2} \end{aligned}$$

for n large enough. We used that $j_{\max} - j_{\min} \leq \frac{1}{2}(\log_2 n)(1+\varepsilon/4)$ for n large enough ($n \geq n_0(\varepsilon, T)$). It follows that, except on an event of probability at most $\exp(-c(\log n)^2)$, it holds

$$\sum_{j=j_{\min}}^{j_{\max}} \Delta_j \leq \sum_{j=j_{\min}}^{j_{\max}} \frac{\delta \lambda_{T_{k-1}}}{(\log_2 n)(1+\varepsilon/2)} \leq \frac{\delta}{2} \left(1 - \frac{\varepsilon}{8} \right) \lambda_{T_{k-1}} \quad (2.1)$$

for n large enough, as long as $\varepsilon \leq 1$, which we assumed. \square

It remains to bound the displacement caused by particles arriving closer than $1/\sqrt{n}$ from the right front, which is encoded in

$$\Delta_{\min} = \sum_{i=T_{k-1}}^{T_{k-1}+m_k-1} 1_{\{x_{i+1} \in I_{i,\min}\}} (R_{i+1} - R_i).$$

Lemma 2.5. For $k \geq 2$ and any $\varepsilon, \delta > 0$ it holds

$$\mathbb{P}\left(\Delta_{\min} < \left(\frac{6\delta}{1+\varepsilon}\right) \frac{\lambda_{T_{k-1}}}{\log n} \middle| \mathcal{F}_{T_{k-1}}\right) \geq 1 - e^{-c(\log n)^2}$$

for n large enough, where c is a constant that may depend on δ and ε .

Proof. This can be bounded crudely: indeed, note that if $x_{i+1} \in I_{i,\min}$ then

$$R_{i+1} - R_i \leq 1/\sqrt{n}$$

so

$$\Delta_{\min} \leq \frac{1}{\sqrt{n}} N_{\min}, \text{ where } N_{\min} = \#\{T_{k-1} + 1 \leq i \leq T_{k-1} + m_k : x_i \in I_{i,\min}\}.$$

Note that N_{\min} is dominated by a sum of independent indicator random variables where the probability of success is at most

$$p_{\min} = \frac{2}{\sqrt{n}\lambda_{T_{k-1}}},$$

since $\lambda_i \geq \lambda_{T_{k-1}}$ for all $i \geq T_{k-1}$ and $2^{j_{\min}-1} < 1/\sqrt{n}$. Hence

$$\mathbb{E}(N_{\min}) \leq m_k p_{\min} = \left(\frac{4\delta}{1+\varepsilon}\right) \frac{\sqrt{n}}{\log n} \lambda_{T_{k-1}}, \quad (2.2)$$

We conclude the proof of the lemma with the elementary fact on binomial random variables (which can also be deduced from Friedman's inequality) that for some $c > 0$,

$$\mathbb{P}(N_{\min} \geq (3/2)m_k p_{\min}) \leq \exp(-cm_k p_{\min}),$$

which implies that

$$\begin{aligned} \mathbb{P}\left(\Delta_{\min} \geq \frac{3}{2} \frac{m_k p_{\min}}{\sqrt{n}} \middle| \mathcal{F}_{T_{k-1}}\right) &\leq \mathbb{P}(N_{\min} \geq (3/2)m_k p_{\min}) \\ &\leq \exp(-cm_k p_{\min}) \\ &\leq \exp(-c(\log n)^2) \end{aligned}$$

as $\lambda_{T_{k-1}} \geq l(n) = (\log n)^3/\sqrt{n}$ since $k \geq 2$ by assumption. This, together with (2.2), concludes the proof. \square

To conclude the proof of Proposition 2.1 we note that

$$R_{T_{k-1}+m_k} - R_{T_{k-1}} = \sum_{i=T_{k-1}}^{T_{k-1}+m_k-1} (R_{i+1} - R_i) = \sum_{j=j_{\min}}^{j_{\max}} \Delta_j + \Delta_{\min}.$$

from which, putting Lemmas 2.4 and 2.5 together, for $k \geq 2$ we obtain

$$\begin{aligned} \mathbb{P}\left(R_{T_{k-1}+m_k} - R_{T_{k-1}} \geq \frac{\delta}{2} \lambda_{T_{k-1}} \middle| \mathcal{F}_{T_{k-1}}\right) &\leq \mathbb{P}\left(\sum_{j=j_{\min}}^{j_{\max}} \Delta_j + \Delta_{\min} \geq \frac{\delta}{2} \lambda_{T_{k-1}} \middle| \mathcal{F}_{T_{k-1}}\right) \\ &\leq \mathbb{P}\left(\sum_{j=j_{\min}}^{j_{\max}} \Delta_j \geq \frac{\delta}{2} \left(1 - \frac{\varepsilon}{8}\right) \lambda_{T_{k-1}} \middle| \mathcal{F}_{T_{k-1}}\right) + \mathbb{P}\left(\Delta_{\min} \geq \frac{\delta\varepsilon}{16} \lambda_{T_{k-1}} \middle| \mathcal{F}_{T_{k-1}}\right) \\ &\leq 2e^{-c(\log n)^2} \end{aligned}$$

for n large enough ($n \geq n_0(\varepsilon, T)$). Similarly one shows that

$$\mathbb{P}\left(L_{T_{k-1}+m_k} - L_{T_{k-1}} \geq \frac{\delta}{2}\lambda_{T_{k-1}} \mid \mathcal{F}_{T_{k-1}}\right) \leq 2e^{-c(\log n)^2},$$

and so

$$\mathbb{P}(T_k - T_{k-1} < m_k \mid \mathcal{F}_{T_{k-1}}) \leq 4e^{-c(\log n)^2}$$

for n large enough. \square

To derive the upper bound on the growth, for any $t \in (0, T]$ define

$$S(t) = \inf\{i \geq 1 : \lambda_i \geq \sqrt{t \log n}\}.$$

Note that $S(t)$ also depends on n , but we choose not to highlight this in the notation. We will use Proposition 2.1 to show that for any $\varepsilon > 0$

$$\mathbb{P}\left(\inf_{t \in (0, T]} \frac{S(t)}{nt} < \frac{1}{(1+\varepsilon)^5}\right) \leq e^{-c(\log n)^2} \quad (2.3)$$

for n large enough ($n \geq n_0(\varepsilon, T)$). To see this, note that for any $k \geq 2$

$$\lambda_{T_k} \leq (1+\delta)\lambda_{T_{k-1}} + \frac{2}{\sqrt{n}} \leq (1+\delta)^{k-1}\lambda_{T_1} + \frac{2(k-1)}{\sqrt{n}},$$

since the length of the interval increases by at most $1/\sqrt{n}$ in each direction upon a new arrival. Now using that $\lambda_{T_1} \leq l(n) + 2/\sqrt{n}$ we get

$$\lambda_{T_k} \leq (1+\delta)^{k-1} \left(l(n) + \frac{2}{\sqrt{n}} \right) + \frac{2(k-1)}{\sqrt{n}} \leq (1+\delta)^{k-1} \frac{(\log n)^3}{\sqrt{n}} (1+\varepsilon),$$

where the last inequality holds as long as

$$2k \leq \varepsilon(\log n)^3. \quad (2.4)$$

It follows that if we let $K(t)$ denote the largest integer such that

$$(1+\delta)^{K(t)-1} < \frac{\sqrt{nt}}{(\log n)^{5/2}(1+\varepsilon)}, \quad (2.5)$$

then $\lambda_{T_{K(t)}} < \sqrt{t \log n}$, which forces $S(t) \geq T_{K(t)}$. Note that such $K(t)$ satisfies (2.4) for n large enough ($n \geq n_0(\varepsilon, \delta, T)$). Thus

$$\begin{aligned} \mathbb{P}\left(\inf_{t \in (0, T]} \frac{S(t)}{nt} < \frac{1}{(1+\varepsilon)^5}\right) &\leq \mathbb{P}\left(\inf_{t \in (0, T]} \frac{T_{K(t)}}{nt} < \frac{1}{(1+\varepsilon)^5}\right) \\ &\leq \mathbb{P}\left(\inf_{t \in (0, T]} \frac{1}{nt} \sum_{k=2}^{K(t)} (T_k - T_{k-1}) < \frac{1}{(1+\varepsilon)^5}\right). \end{aligned}$$

Now, using that $\lambda_{T_k} \geq (1+\delta)^{k-1}l(n)$ for all $k \geq 2$ by definition of the stopping times T_k 's, it is easy to check that

$$\frac{1}{(1+\varepsilon)^2(1+\delta)^3} \leq \frac{1}{nt} \sum_{k=2}^{K(t)} m_k$$

for all $t \in (0, T]$. Taking $\delta = \varepsilon$ we get

$$\begin{aligned}
\mathbb{P}\left(\inf_{t \in (0, T]} \frac{S(t)}{nt} < \frac{1}{(1 + \varepsilon)^5}\right) &\leq \mathbb{P}\left(\exists t \in (0, T] : \sum_{k=2}^{K(t)} (T_k - T_{k-1}) < \sum_{k=2}^{K(t)} m_k\right) \\
&= \mathbb{P}\left(\exists 2 \leq K \leq K(T) : \sum_{k=2}^K (T_k - T_{k-1}) < \sum_{k=2}^K m_k\right) \\
&\leq \sum_{K=2}^{K(T)} \mathbb{P}(\exists 2 \leq k \leq K : T_k - T_{k-1} < m_k) \\
&\leq \sum_{K=2}^{K(T)} \sum_{k=2}^K \mathbb{E}(\mathbb{P}(T_k - T_{k-1} < m_k | \mathcal{F}_{T_{k-1}})) \\
&\leq (K(T))^2 e^{-c(\log n)^2} \leq e^{-c'(\log n)^2}
\end{aligned}$$

for n large enough (depending on ε and T). Here we have used Proposition 2.1 in the fourth inequality, and (2.5) in the last one.

We now turn to showing that the doubling time cannot be too large.

To start with, we need to control the time T_1 it takes to get to minimal length $l(n) = (\log n)^3 / \sqrt{n}$. It is easy to check that, while $\lambda_i \leq l(n)$, one has

$$R_{i+1} - R_i \geq \frac{1}{\sqrt{n}(\log n)^4}$$

for n large enough (n larger than an absolute constant): simply use that the minimal push to the right is obtained when the new particle lands at the left end $x_{i+1} = -L_i$. It follows that $\lambda_{i+1} - \lambda_i \geq 1/(\sqrt{n}(\log n)^4)$, from which we deduce that, deterministically, $T_1 \leq (\log n)^7$.

Now, once the interval has reached minimal length $l(n)$, it cannot grow too slowly, as shown in the next proposition.

Proposition 2.6 (Upper bound on doubling time). *For any $k \geq 2$ it holds*

$$\mathbb{P}\left(T_k - T_{k-1} < 2\delta(1 + \varepsilon) \frac{n}{\log n} \lambda_{T_{k-1}}^2 \middle| \mathcal{F}_{T_{k-1}}\right) \geq 1 - e^{-c(\log n)^2}$$

for n large enough ($n \geq n_0(\varepsilon, T)$). Here $c = c(\delta, \varepsilon)$ is a constant which may depend on ε and δ but not on n nor k .

Proof. This can be proved reasoning as in Proposition 2.1. Write

$$n_k = 2\delta(1 + \varepsilon) \frac{n}{\log n} \lambda_{T_{k-1}}^2$$

for brevity, and define

$$\tilde{\Delta}_j = \sum_{i=T_{k-1}}^{T_{k-1} + n_k - 1} 1_{\{x_{i+1} \in I_{i,j}\}} (R_{i+1} - R_i)$$

for $\tilde{j}_{min} \leq j \leq \tilde{j}_{max}$, where \tilde{j}_{max} is as before, and \tilde{j}_{min} is chosen, depending on ε , so that

$$\left(\frac{\log 2}{2n} - \frac{C}{n^2 2^{2j}}\right) \geq \frac{\log n}{2n} \frac{1 + \varepsilon/2}{1 + \varepsilon},$$

where C is the absolute constant in Lemma 2.2. This ensures that

$$\mathbb{E}(\tilde{\Delta}_j | \mathcal{F}_{T_{k-1}}) \geq \delta(1 + \varepsilon/2) \frac{\lambda_{T_{k-1}}}{\log_2 n}.$$

Thus

$$-\tilde{\Delta}_j + \delta(1 + \varepsilon/2) \frac{\lambda_{T_{k-1}}}{\log_2 n}$$

is the terminal value at time n_k of a supermartingale, and one again concludes by Freedman's inequality that

$$\mathbb{P}\left(\tilde{\Delta}_j < \frac{\delta \lambda_{T_{k-1}}}{\log_2 n} \middle| \mathcal{F}_{T_{k-1}}\right) \leq e^{-c(\log n)^2}$$

for n large enough, depending on ε and T , and a constant $c = c(\delta, \varepsilon)$. This can be summed over all scales $\tilde{j}_{min} \leq j \leq \tilde{j}_{max}$ to get that, since

$$R_{T_{k-1}+n_k} - R_{T_{k-1}} = \sum_{i=T_{k-1}}^{T_{k-1}+n_k-1} (R_{i+1} - R_i) \geq \sum_{j=\tilde{j}_{min}}^{\tilde{j}_{max}} \Delta_j,$$

for $k \geq 2$ it holds

$$\begin{aligned} \mathbb{P}\left(R_{T_{k-1}+n_k} - R_{T_{k-1}} < \frac{\delta}{2} \lambda_{T_{k-1}} \middle| \mathcal{F}_{T_{k-1}}\right) &\leq \mathbb{P}\left(\sum_{j=\tilde{j}_{min}}^{\tilde{j}_{max}} \tilde{\Delta}_j < \frac{\delta}{2} \lambda_{T_{k-1}} \middle| \mathcal{F}_{T_{k-1}}\right) \\ &\leq e^{-c(\log n)^2} \end{aligned}$$

for n large enough. The result follows by observing that the same bound holds for $L_{T_{k-1}+n_k} - L_{T_{k-1}}$. \square

We now discuss how the above result allows to conclude that the interval cannot grow too slowly, once it has reached minimal length $l(n)$. We show that, with $t_0(n) = (\log n)^8/n$,

$$\mathbb{P}\left(\sup_{t \in [t_0(n), T]} \frac{S(t)}{nt} > (1 + \varepsilon)^5\right) \leq e^{-c(\log n)^2} \quad (2.6)$$

for $n \geq n_0(\varepsilon, T)$. For $t \in [t_0(n), T]$ let $\tilde{K}(t)$ denote the smallest integer such that

$$(1 + \delta)^{\tilde{K}(t)-1} > \frac{\sqrt{nt}}{(\log n)^{5/2}},$$

so that

$$(1 + \delta)^{\tilde{K}(t)-2} \leq \frac{\sqrt{nt}}{(\log n)^{5/2}}. \quad (2.7)$$

Then $\lambda_{T_{\tilde{K}(t)}} \geq \sqrt{t \log n}$, which forces $S(t) \leq T_{\tilde{K}(t)}$. Moreover, since $\lambda_{T_k} \leq (1 + \delta)\lambda_{T_{k-1}} + 2/\sqrt{n}$ for all $k \geq 2$, it is easily checked that, by the choice of $\tilde{K}(t)$,

$$\frac{1}{nt} \sum_{k=2}^{\tilde{K}(t)} n_k \leq (1 + \varepsilon)^2 (1 + \delta)^2$$

for all $t \in (0, T]$ and n large enough, depending only on δ, ε . Thus, taking again $\delta = \varepsilon$, we have that

$$\begin{aligned}
\mathbb{P}\left(\sup_{t \in [t_0(n), T]} \frac{S(t)}{nt} > (1 + \varepsilon)^5\right) &\leq \mathbb{P}\left(\sup_{t \in [t_0(n), T]} \frac{T_{\tilde{K}(t)}}{nt} > (1 + \varepsilon)^5\right) \\
&\leq \mathbb{P}\left(\sup_{t \in [t_0(n), T]} \frac{T_1}{nt} > \varepsilon\right) + \mathbb{P}\left(\sup_{t \in [t_0(n), T]} \frac{1}{nt} \sum_{k=2}^{\tilde{K}(t)} (T_k - T_{k-1}) > (1 + \varepsilon)^4\right) \\
&\leq \mathbb{P}(T_1 > \varepsilon(\log n)^8) + \mathbb{P}\left(\exists t \in [t_0(n), T] : \sum_{k=2}^{\tilde{K}(t)} (T_k - T_{k-1}) > \sum_{k=2}^{\tilde{K}(t)} n_k\right) \\
&\leq \mathbb{P}\left(\exists 2 \leq K \leq \tilde{K}(T) : \sum_{k=2}^K (T_k - T_{k-1}) > \sum_{k=2}^K n_k\right) \\
&\leq \sum_{K=2}^{\tilde{K}(T)} \mathbb{P}(\exists 2 \leq k \leq K : T_k - T_{k-1} > n_k) \\
&\leq \sum_{K=2}^{\tilde{K}(T)} \sum_{k=2}^K \mathbb{E}(\mathbb{P}(T_k - T_{k-1} > n_k | \mathcal{F}_{T_{k-1}})) \\
&\leq (\tilde{K}(T))^2 e^{-c(\log n)^2} \leq e^{-c'(\log n)^2}
\end{aligned}$$

for n large enough ($n \geq n_0(\varepsilon, T)$). The fourth inequality above holds for n large enough so that $\varepsilon(\log n)^8 \geq (\log n)^7$, which makes $\mathbb{P}(T_1 > \varepsilon(\log n)^8)$ vanish. For the last one we have used (2.7).

Putting together the upper and lower bounds (2.3) and (2.6) for $S(t)$ we gather that for any $\varepsilon \in (0, 1]$ and $T > 0$ it holds

$$\mathbb{P}\left(\frac{nt}{(1 + \varepsilon)^5} \leq S(t) \leq nt(1 + \varepsilon)^5 \quad \forall t \in [t_0(n), T]\right) \geq 1 - e^{-c(\log n)^2} \quad (2.8)$$

for n large enough (depending on ε, T). To deduce that a similar concentration result holds for the length λ_{nt} we use that

$$\sqrt{t \log n} \leq S(t) \leq \sqrt{t \log n} + \frac{2}{\sqrt{n}}$$

for all $t \in (0, T]$. This tells us that, on the event in (2.8),

$$\lambda_{nt(1+\varepsilon)^5} \geq \lambda_{S(t)} \geq \sqrt{t \log n} \quad \forall t \in [t_0(n), T],$$

which implies

$$\lambda_{nt} \geq \sqrt{\frac{t \log n}{(1 + \varepsilon)^5}} \geq (1 - 3\varepsilon)\sqrt{t \log n} \quad \forall t \in [t_0(n)(1 + \varepsilon)^5, T]$$

for ε small enough ($\varepsilon \leq \varepsilon_0$ with ε_0 absolute constant).

For the upper bound note that, again on the event in (2.8),

$$\lambda_{nt(1-\varepsilon)^5} \leq \lambda_{S(t)} \leq \sqrt{t \log n} + \frac{2}{\sqrt{n}},$$

which implies

$$\lambda_{nt} \leq \sqrt{\frac{t \log n}{(1 - \varepsilon)^5}} + \frac{2}{\sqrt{n}} \leq (1 + 3\varepsilon)\sqrt{t \log n} \quad \forall t \in [t_0(n)(1 - \varepsilon)^5, T]$$

for ε small enough ($\varepsilon \leq \varepsilon_0$ with ε_0 absolute constant). This shows that for any $\varepsilon > 0$ small enough, and any $T > 0$, there exists $n_0(\varepsilon, T)$ such that

$$\mathbb{P} \left((1 - 3\varepsilon)\sqrt{t \log n} \leq \lambda_{nt} \leq (1 + 3\varepsilon)\sqrt{t \log n} \quad \forall t \in \left[\frac{(\log n)^9}{n}, T \right] \right) \geq 1 - e^{-c(\log n)^2} \quad (2.9)$$

for all $n \geq n_0(\varepsilon, T)$, which concludes the proof of (1.4).

It remains to deduce that R_{nt} and L_{nt} are both concentrated around $\frac{1}{2}\sqrt{t \log n}$. We show that there exists an absolute constant ε_0 such that for all $\varepsilon \in (0, \varepsilon_0)$ it holds

$$\mathbb{P} \left(\frac{\sqrt{t \log n}}{2(1 + \varepsilon)^6} \leq R_{nt} \leq \frac{1}{2}\sqrt{t \log n}(1 + \varepsilon)^6 \quad \forall t \in \left[\frac{(\log n)^9}{n}, T \right] \right) \geq 1 - e^{-c(\log n)^2}$$

for n large enough, and c constant independent of n . To this end, recall the definition of $K(t)$ from (2.5), and note that for any $t \in (t_0(n), T]$, restricting to the high probability event $\{T_k - T_{k-1} \geq m_k \text{ for all } 2 \leq k \leq K(t)\}$, we have

$$T_{K(t)} \geq \sum_{k=2}^{K(t)} (T_k - T_{k-1}) \geq \sum_{k=2}^{K(t)} m_k \geq \frac{nt}{(1 + \varepsilon)^5}.$$

It follows that

$$R_{\frac{nt}{(1+\varepsilon)^5}} \leq R_{T_{K(t)}} = R_{T_1} + \sum_{k=2}^{K(t)} (R_{T_k} - R_{T_{k-1}})$$

for all $t \in (t_0(n), T]$. Restricting further to the event $\{T_k - T_{k-1} \leq n_k \text{ for all } 2 \leq k \leq \max\{K(T), \bar{K}(T)\}\}$ we have that $R_{T_k} - R_{T_{k-1}} \leq R_{T_{k-1}+n_k} - R_{T_{k-1}}$. It can be shown, using exactly the same argument as the one of Proposition 2.1, that

$$\mathbb{P} \left(R_{T_{k-1}+n_k} - R_{T_{k-1}} \leq \frac{\delta}{2}(1 + \varepsilon)^2 \lambda_{T_{k-1}} \text{ for all } 2 \leq k \leq K(T) \right) \geq 1 - e^{-c(\log n)^2}$$

for n large enough ($n \geq n_0(\varepsilon, T)$). Thus, with probability exceeding $1 - 3e^{-c(\log n)^2}$, we have that

$$R_{\frac{nt}{(1+\varepsilon)^5}} \leq \lambda_{T_1} + \frac{\delta}{2}(1 + \varepsilon)^2 \sum_{k=2}^{K(t)} \lambda_{T_{k-1}} \leq \frac{(1 + \varepsilon)^3}{2} \sqrt{t \log n}$$

for all $t \in (t_0(n), T]$, where the last inequality can be checked using the definition of $K(t)$. Rearranging, this gives

$$\mathbb{P} \left(R_{nt} \leq \frac{(1 + \varepsilon)^6}{2} \sqrt{t \log n} \quad \forall t \in (t_0(n), T] \right) \geq 1 - 2e^{-c(\log n)^2}.$$

Similarly one shows that

$$\mathbb{P} \left(L_{nt} \leq \frac{(1 + \varepsilon)^6}{2} \sqrt{t \log n} \quad \forall t \in (t_0(n), T] \right) \geq 1 - 2e^{-c(\log n)^2}.$$

Combining these with the lower bound in (2.9) we also deduce that, with probability exceeding $1 - 5e^{-c(\log n)^2}$,

$$R_{nt} = \lambda_{nt} - L_{nt} \geq \sqrt{t \log n} \left(\frac{1}{(1 + \varepsilon)^3} - \frac{1}{2}(1 + \varepsilon)^6 \right) \geq \frac{1}{2(1 + \varepsilon)^6} \sqrt{t \log n}$$

for all $t \in [(\log n)^9/n, T]$, where the last inequality holds as long as $\varepsilon \leq \varepsilon_0$ for some absolute constant $\varepsilon_0 > 0$. The same holds for L_{nt} , which concludes the proof of Theorem 1.1.

3 Proof of Theorem 1.2

We first prove that the height of the cluster is deterministically bounded. This comes from the simple observation that the half-plane capacity of \mathcal{K}_{nt} is deterministically given by t and the following simple lemma: if K is a compact \mathbb{H} -hull such that $x + iy \in K$ for some $x \in \mathbb{R}, y > 0$ then

$$\text{hcap}(K) \geq y^2/2.$$

This can be proved using a simple reflection argument about the line $\{\Re(z) = x\}$ and monotonicity of the half-plane capacity; we leave the details to the reader (this is implicit in [11], and explicitly proved in the note [10], see also Proposition 6.1 of [2] for a closely related statement).

As a consequence, since $\text{hcap}(\mathcal{K}_{nt}) = t$ then we deduce that

$$\max_{z \in \mathcal{K}_{nt}} \{\Im(z)\} \leq \sqrt{2t},$$

as desired.

Now let us turn to the diameter. The argument is based on estimating the harmonic measure of \mathcal{K}_{nt} viewed from infinity (which is conformally invariant) and relating this to the diameter, given that the cluster is necessarily very flat by the above. We start with the lower bound on the diameter which is slightly easier.

Fix $t > 0$. Let $L = (1 - \varepsilon)\sqrt{t \log n}$ and denote by $K = \mathcal{K}_{nt}$ for ease of notations. Let $z = iy$ with $y > 0$ very large. Suppose for contradiction that $\text{Diam}(K) \leq L$, so K could fit in a rectangle \mathcal{R} of height $\sqrt{2t}$ and width L , in the upper half plane and with the bottom side on the real line, and which we can assume to be centered without loss of generality. See Figure 6. Then applying the conformal map g_K , the hull K is by definition mapped to the interval $[-L_t, R_t]$. Let $\tilde{\mathcal{R}} = g_K(\mathcal{R} \setminus K)$. If we then apply the conformal map $g_{\tilde{\mathcal{R}}}$, then by monotonicity, we have

$$R_t \leq r, L_t \leq \ell$$

where $-\ell$ and r respectively are the extreme points of $g_{\tilde{\mathcal{R}}}(\partial\tilde{\mathcal{R}} \cap \mathbb{H})$. (See Figure 6). In fact by symmetry note that $\ell = r$. Furthermore we claim that

$$\ell \leq L/2 + C, \tag{3.1}$$

where C is a constant depending only on t . From here the contradiction (and thus the lower bound) follows, since we know that $\min(R_t, L_t) \geq (1 - \varepsilon/2)(1/2)\sqrt{t \log n}$ with probability converging to 1 as $n \rightarrow \infty$.

This can be seen directly because the conformal map $g_{\tilde{\mathcal{R}}}$ is known explicitly (it is a Schwarz–Christoffel map, since $D = \mathbb{H} \setminus \mathcal{R}$ is a polygon), but it can be seen more robustly as follows. Fix $z = iy$ with $y > 0$ very large. Let τ_D denote the first time at which a Brownian motion B leaves the domain D . Let τ denote the first time that the imaginary part of B reaches $\sqrt{2t}$. Let E be the event that $|\Re(B_\tau)| \leq L/2$.

$$\mathbb{P}^{iy}(B_{\tau_D} \in \mathcal{R}) = \mathbb{P}^{iy}(E) + \mathbb{P}^z(B_{\tau_D} \in \mathcal{R}; E^c).$$

Now observe that $\mathbb{P}^z(E)$ may be computed through the integral over the interval $[-L/2, L/2]$ of a Cauchy distribution with scale parameter $y' = y - \sqrt{2t}$: hence after a change of variables,

$$\mathbb{P}^{iy}(E) = \int_{-L/(2y')}^{L/(2y')} \frac{1}{\pi(x^2 + 1)} dx \sim \frac{L}{\pi y}$$

as $y \rightarrow \infty$ (and $t > 0$ fixed), where $a_n \sim b_n$ means that the ratio $a_n/b_n \rightarrow 1$. Now let us turn to E^c , and observe that if $\tau_D \in \mathcal{R}$ and E^c both hold, then necessarily Brownian motion hits one

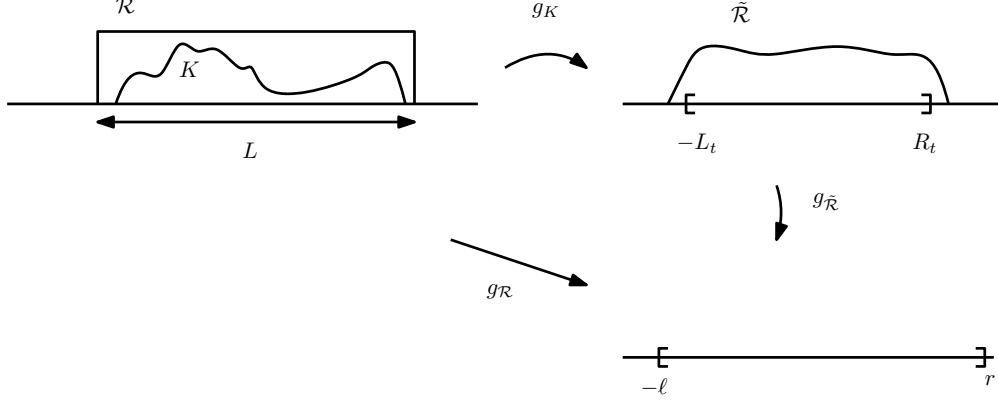


Figure 6: Composition of conformal maps $g_{\mathcal{R}} = g_{\hat{\mathcal{R}}} \circ g_K$.

of the vertical sides of the rectangle \mathcal{R} before the real line. In words, if I_1, I_2 are the vertical sides, and $D^i = \mathbb{H} \setminus I_i$, then

$$(\{B_{\tau_D} \in \mathcal{R}\} \cap E^c) \subset \{B_{\tau_{D^1}} \in I_1\} \cup \{B_{\tau_{D^2}} \in I_2\}$$

Hence by symmetry between I_1 and I_2 ,

$$\mathbb{P}^{iy}(B_{\tau_D} \in \mathcal{R}; E^c) \leq 2\mathbb{P}^{iy}(B_{\tau_{D^1}} \in I_1) \sim 2\frac{\text{cap}(I_1)}{y}$$

Altogether, since $\text{cap}(I_1) = 2\sqrt{2t}/\pi$, we see that

$$\lim_{y \rightarrow \infty} y\mathbb{P}^{iy}(B_{\tau_D} \in \mathcal{R}) \leq L/\pi + 4\sqrt{2t}/\pi. \quad (3.2)$$

To get the desired inequality on ℓ we simply argue by conformal invariance:

$$\mathbb{P}^{iy}(B_{\tau_D} \in \mathcal{R}) = \mathbb{P}^{g_{\mathcal{R}}(iy)}(B_{\tau_{\mathbb{H}}} \in [-\ell, \ell]) \sim \frac{2\ell}{\pi y} \quad (3.3)$$

by a similar integration of the Cauchy distribution and the fact that $g_{\mathcal{R}}(iy) \sim iy$. Multiplying by y and letting $y \rightarrow \infty$, and comparing (3.3) with (3.2), we get

$$\frac{2\ell}{\pi} \leq \frac{L}{\pi} + 4\sqrt{2t}/\pi,$$

which proves (3.1) with $C = \sqrt{32t}$. This proves the desired lower bound on $\text{Diam}(\mathcal{K}_{nt})$.

Now let us turn to the corresponding upper bound. We again argue by contradiction. Let $L = (1 + \varepsilon)\sqrt{t \log n}$. Suppose $\text{Diam}(K) \geq L$. Then there exists $a, b \in K$ with $|a - b| \geq L$. Since K is connected there is a curve in K connecting a to b (and so its maximal height is at most $\sqrt{2t}$). We argue similarly as above. Let I_1, I_2 be the vertical segments connecting a and b respectively to the real line; call x_1 and x_2 the respective base points on \mathbb{R} of I_1 and I_2 . By translation we can assume $x_1 = -x_2$ and by the triangle inequality $|x_1 - x_2| \geq L - O(1)$. Let $E = \{\tau_H \leq \tau_{\mathbb{H}}\} = \{B_{\tau_{\mathbb{H}}} \in K\}$.

$$\mathbb{P}^{iy}(B_{\tau_{\mathbb{H}}} \in [x_1, x_2]) = \mathbb{P}^{iy}(B_{\tau_{\mathbb{H}}} \in [x_1, x_2]; E) + \mathbb{P}(B_{\tau_{\mathbb{H}}} \in [x_1, x_2]; E^c).$$

For the same topological reasons as above, if $B_{\tau_{\mathbb{H}}} \in [x_1, x_2]$ and E^c both hold, it must be the case that the Brownian motion B touched either I_1 or I_2 before the real line (indeed, to enter

the interval $[x_1, x_2]$ the Brownian curve cannot avoid I_1 and I_2 without touching the curve connecting a and b , but then E holds). Therefore,

$$\mathbb{P}^{iy}(B_{\tau_{\mathbb{H}}} \in [x_1, x_2]) \leq \mathbb{P}^{iy}(E) + \mathbb{P}^{iy}(B_{\tau_{D_1}} \in I_1) + \mathbb{P}^{iy}(B_{\tau_{D_2}} \in I_2).$$

Now we multiply by y and let $y \rightarrow \infty$. Using the same argument as above, we see that the left hand side converges to $|x_1 - x_2|/\pi = L/\pi - O(1)$. The right hand side has three terms. The first, by conformal invariance of Brownian motion, converges to $|L_t - R_t|/\pi$. The second and third respectively converge to $\text{cap}(I_1)$ and $\text{cap}(I_2)$, both of which are $O(1)$. Consequently, we obtain

$$L \leq |L_t - R_t| + O(1).$$

By Theorem 1.1, the probability of the above tends to zero.

Acknowledgements. This work took place while N.B. was visiting professor at the University of Rome La Sapienza, whose hospitality is gratefully acknowledged.

The authors also thank James Norris for some useful and inspiring conversations, both in relation with this project and over the years. We also thank Alan Sola for a number of comments on a draft of this article.

NB's work was supported by a University of Vienna start-up grant, and FWF grant P33083 on "Scaling limits in random conformal geometry".

References

- [1] Robert O Bauer. Discrete Löwner evolution. In *Annales de la Faculté des sciences de Toulouse: Mathématiques*, volume 12, pages 433–451, 2003.
- [2] Nathaniel Berestycki and James Norris. Lectures on Schramm–Loewner Evolution. *Lecture notes, available on the webpages of the authors*, 2014.
- [3] Noam Berger, Eviatar B Procaccia, and Amanda Turner. Growth of stationary Hastings–Levitov. *arXiv preprint arXiv:2008.05792*, 2020.
- [4] David A Freedman. On tail probabilities for martingales. *Ann. Probab.*, 3(1):100–118, 1975.
- [5] Matthew B Hastings and Leonid S Levitov. Laplacian growth as one-dimensional turbulence. *Physica D: Nonlinear Phenomena*, 116(1):244–252, 1998.
- [6] Fredrik Johansson and Alan Sola. Rescaled Lévy–Loewner hulls and random growth. *Bulletin des sciences mathématiques*, 133(3):238–256, 2009.
- [7] Fredrik Johansson Viklund, Alan Sola, and Amanda Turner. Scaling limits of anisotropic Hastings–Levitov clusters. *Ann. Inst. Henri Poincaré Probab. Stat.*, 48(1):235–257, 2012.
- [8] Fredrik Johansson Viklund, Alan Sola, and Amanda Turner. Small-particle limits in a regularized Laplacian random growth model. *Comm. Math. Phys.*, 334(1):331–366, 2015.
- [9] Harry Kesten. How long are the arms in DLA? *Journal of Physics A: Mathematical and General*, 20(1):L29, 1987.
- [10] Steven Lalley, Gregory Lawler, and Hariharan Narayanan. Geometric interpretation of half-plane capacity. *Electronic Communications in Probability*, 14:566–571, 2009.
- [11] Gregory F Lawler. *Conformally invariant processes in the plane*. Number 114. American Mathematical Soc., 2008.
- [12] James Norris, Vittoria Silvestri, and Amanda Turner. Scaling limits for planar aggregation with subcritical fluctuations. *arXiv preprint arXiv:1902.01376*, 2019.

- [13] James Norris, Vittoria Silvestri, and Amanda Turner. Stability of regularized Hastings–Levitov aggregation in the subcritical regime. *arXiv preprint arXiv:2105.09185*, 2021.
- [14] James Norris and Amanda Turner. Hastings-Levitov aggregation in the small-particle limit. *Comm. Math. Phys.*, 316(3):809–841, 2012.
- [15] Vittoria Silvestri. Fluctuation results for Hastings–Levitov planar growth. *Probability Theory and Related Fields*, 167(1-2):417–460, 2017.