On the Radial Growth of Ballistic Aggregation and Other Aggregation Models

Tillmann Bosch[1] · **Steffen Winter**[2]

Abstract
For a class of aggregation models on the integer lattice $\mathbb{Z}^d$, $d \geq 2$, in which clusters are formed by particles arriving one after the other and sticking irreversibly where they first hit the cluster, including the classical model of diffusion-limited aggregation (DLA), we study the growth of the clusters. We observe that a method of Kesten used to obtain an almost sure upper bound on the radial growth in the DLA model generalizes to a large class of such models. We use it in particular to prove such a bound for the so-called ballistic model, in which the arriving particles travel along straight lines. Our bound implies that the fractal dimension of ballistic aggregation clusters in $\mathbb{Z}^2$ is 2, which proves a long standing conjecture in the physics literature.

Keywords Diffusion-limited aggregation · Incremental aggregation · Radial growth · Fractal dimension · Ballistic aggregation

MSC Classification Primary: 82B24 · 60J10; Secondary: 60D05 · 28A80

1 Introduction

Consider a process of cluster formation on the hypercubic lattice $\mathbb{Z}^d$, $d \geq 2$, which starts at time one with a single particle placed at the origin. Then at each time step $n = 2, 3, \ldots$, a particle arrives and is added to the cluster by placing it at a site neighbouring the existing cluster. The position of the new particle is chosen at random independently of all previous choices and according to some distribution which only depends on the existing cluster at that time. We will call such aggregation models *incremental aggregation* in the sequel. Many such
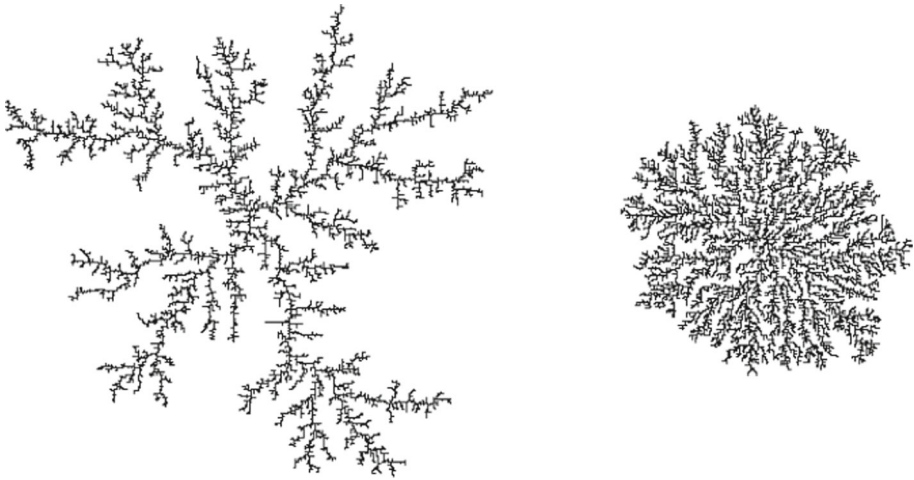
✉ Steffen Winter
steffen.winter@kit.edu

Tillmann Bosch
tillmannbosch3@gmail.com

1   Langstraße 43, 68169 Mannheim, Germany

2   Institute of Stochastics, Karlsruhe Institute of Technology, Englerstr. 2, 76131 Karlsruhe, Germany
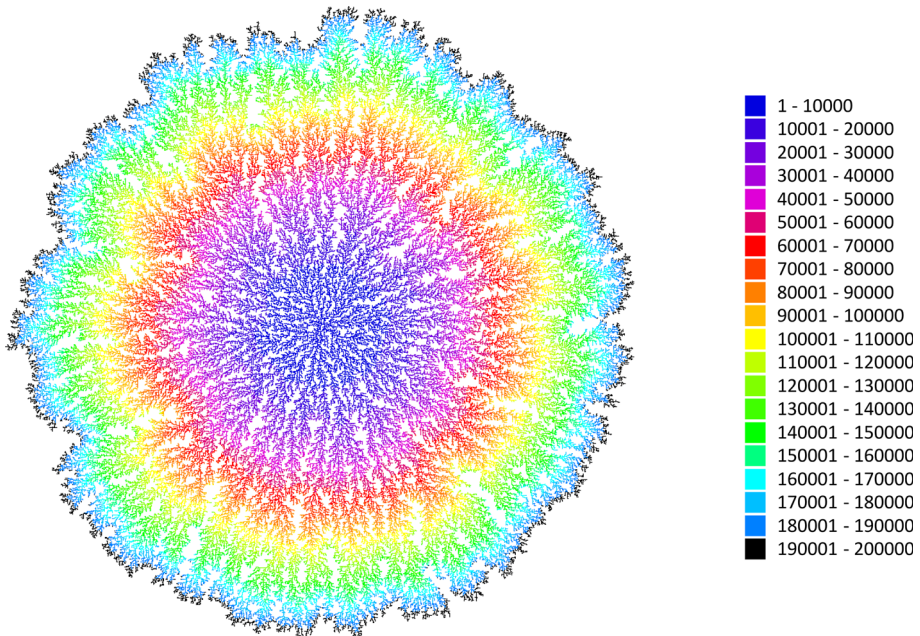
Fig. 1 Typical realization of an aggregation cluster of DLA (left) and the ballistic aggregation model (right) of size $n = 10,000$

models are known and have been studied intensively in the physics literature. They differ by the choice of the distributions governing the selection of the locations for the arriving particles.

The most popular among these models is certainly *diffusion-limited aggregation (DLA)*, suggested by Witten and Sander [1] in 1981 as a model for metal-particle aggregation. Particles arrive from 'far away' and perform symmetric random walks until they hit the cluster for the first time. The distribution specifying the resulting cluster formation mechanism is known as the harmonic measure. The formed clusters have a dendrite-like structure and show a fractal behaviour, see Fig. 1 (left). Another even older model is the *Eden growth model* due to Eden [2], in which the position of the next particle is chosen uniformly out of all sites neighboring the current cluster. The clusters formed in this model tend to be very dense and appear ball like with a rough boundary. Another variation is *diffusion-limited annihilation* due to Meakin and Deutch [3], nowadays better known as *internal DLA*. Here the particles start at the origin, i.e. within the cluster, perform symmetric random walks and are attached where they exit the cluster. As in the Eden model, the formed clusters tend to be ball-like with a rough boundary.

Another popular model in the physics literature is the *Vold-Sutherland model* or *ballistic aggregation*, in which arriving particles travel along straight lines (ballistic trajectories) and stick where they hit the cluster. It can be traced back to the work of the colloid chemists Marjorie J. Vold and David N. Sutherland, see e.g. [4, 5]. It has been suggested for situations in which molecules move in a low density vapour such that thermal interactions can be disregarded. The clusters formed by this model are denser than the ones formed by DLA with branches reconnecting more frequently, but leaving much more space unoccupied as those of the Eden model, see Figs. 1 (right) and 2.

All these models and many variants have been studied intensively in physics and have found numerous applications, we refer to the surveys and books [5–9] for further details and references. Most results have been obtained either from simulations or from intriguing but non-rigorous theoretical considerations. Some of these models have also inspired mathematical research and a number of results have been obtained. However, the progress up to now is

■	1 - 10000
■	10001 - 20000
■	20001 - 30000
■	30001 - 40000
■	40001 - 50000
■	50001 - 60000
■	60001 - 70000
■	70001 - 80000
■	80001 - 90000
■	90001 - 100000
■	100001 - 110000
■	110001 - 120000
■	120001 - 130000
■	130001 - 140000
■	140001 - 150000
■	150001 - 160000
■	160001 - 170000
■	170001 - 180000
■	180001 - 190000
■	190001 - 200000

Fig. 2 Realization of a ballistic cluster of size $n = 200{,}000$, in which the arrival times of the particles are encoded by color. It is apparent that the cluster grows rather uniformly in all directions and that late particles cannot get close to the origin (Color figure online)

far behind what physicists have discovered. We remark that there are also versions of these models (so-called *off-lattice models*) in which clusters are grown in the continuum and not restricted to any lattice, most notably Hastings–Levitov models [10], see also in the above references. In many respects off-lattice models show a similar behaviour as their lattice counterparts and some rigorous results have been obtained, see e.g. [11, 12]. However, the precise relation between lattice and off-lattice models remains a challenging open problem. As the latter do not fit our framework here, we will concentrate on lattice models in the sequel.

One of the most basic questions (to which we will mainly restrict our attention in this paper) is the speed of growth of the radius (or diameter) of the cluster as the number of particles tends to ∞. Regarding DLA, some rigorous almost sure bounds are due to Kesten [13, 14] (with some improvements for dimension 3 in [15] and recently in [16]), which we will recall in Sect. 3 below. They imply in particular that DLA clusters are not full-dimensional (in a discrete sense) and thus exhibit some fractal behaviour. For internal DLA (and the Eden model), the growth rate is $n^{1/d}$ in $\mathbb{Z}^d$, as the number of particles n tends to ∞ and thus the clusters are full-dimensional. For internal DLA even some shape theorems have been established showing that asymptotically the clusters are balls [17]. For the Eden model the existence of a limit shape is also known, due to some relation of this model to first passage percolation, cf. [18] and the references therein. The limit shape is conjectured to deviate from a Euclidean ball. Up to now this has only been established in very high dimensions. For further results regarding DLA and internal DLA, we refer to the survey [19] (and the references therein), in which results are also reviewed for graphs other than the hypercubic lattices $\mathbb{Z}^d$. (It is obvious that the models may be transferred to any infinite connected graph.) In contrast to the situation for the other mentioned models, it seems that the

ballistic model has not received much attention in the maths literature so far. It is mentioned (but not analyzed very far) in [20], in which also a number of new aggregation models are proposed and simulated. A closely related class of models is *ballistic deposition*, in which random surfaces are grown by particles raining down onto a surface and for which a number of mathematical results are available [21–23], see Remark 5.9 for a discussion of possible connections to ballistic aggregation.

In this article we make a first attempt to investigate rigorously the growth behaviour of the ballistic model. To this end, we start from a more general viewpoint and set up a natural class of models (incremental aggregation), which includes the models mentioned above (and many more). We revisit the argument of Kesten that he used in [13] to obtain an upper bound on the growth rates of DLA in $\mathbb{Z}^2$ and observe that the same argument may be applied to any incremental aggregation model. In a nutshell, the 'method' reduces the problem of finding a bound for the radial growth to the much easier task to establish an upper bound on the local mass of the distributions determining the model. We do not claim much originality here, as once the class of incremental aggregation models is set up it is fairly easy to see that Kesten's argument works in general. Then we apply this new method to the ballistic model. First we give a rigorous definition of the ballistic model using some notions from geometric probability. Then we state some bound on the local mass for the ballistic measures (i.e., the distributions defining the ballistic model) and finally we apply Kesten's method to obtain the desired growth rates. It turns out that the method has power in particular in dimension $d = 2$. For the ballistic model in $\mathbb{Z}^2$ we establish that the growth exponent is $1/2$ (i.e., the fractal dimension is 2), which is the value conjectured in the physics literature. For $d \geq 3$, our results imply that the fractal dimension is at least 2 (while the conjectured value is d).

In order to illustrate our results, we have also simulated aggregation clusters of the ballistic model (as we define it here) and DLA for comparison, see Figs. 1 and 2. The Python code is freely available, cf. [24], together with more pictures of large simulated clusters.

The paper is organized as follows. In Sect. 2, we introduce incremental aggregation and show that this class of models includes DLA and the Eden model. In Sect. 3 we provide some definitions regarding radial growth and state some simple facts about growth exponents and fractal dimensions of our models. We also recall the results of Kesten on the radial growth of DLA. In Sect. 4, we state Kesten's method (Theorem 4.1) for general incremental aggregation and demonstrate how it can be applied to recover Kesten's growth bounds for DLA and the known bounds for the Eden model. In Sect. 5, we define the ballistic model and apply Kesten's method to obtain the mentioned results on the growth of these models. Finally, in Sect. 6 we provide a proof of our main tool, Kesten's method.

2 Incremental Aggregation

We set up a framework in which all the models can be treated in a unified way. Let $\mathcal{P}_f^d$ denote the family of finite subsets of $\mathbb{Z}^d$, i.e.

$$\mathcal{P}_f^d := \{A \subset \mathbb{Z}^d \mid A \text{ is finite}\},$$

which we equip with the discrete σ-algebra. We consider the *nearest neighbor graph* on $\mathbb{Z}^d$, i.e., the graph $(\mathbb{Z}^d, E)$ with vertex set $\mathbb{Z}^d$ and edge set $E := \{\{x, y\} \subset \mathbb{Z}^d : \|x - y\| = 1\}$. (Here and throughout $\| \cdot \|$ denotes the Euclidean norm.) A set $A \in \mathcal{P}_f^d$ is called *connected*, if the subgraph of $(\mathbb{Z}^d, E)$ generated by A is. For $A \in \mathcal{P}_f^d$, the *(outer) boundary* of A is the set

$$\partial A := \{y \in \mathbb{Z}^d \setminus A : \exists x \in A \text{ such that } \{x, y\} \in E\}.$$

Throughout let $(\Omega, \mathcal{A}, \mathbb{P})$ be some suitable probability space. For $A \in \mathcal{P}_f^d \setminus \{\emptyset\}$, a *random point in A* is a measurable mapping $y_A : \Omega \to \mathbb{Z}^d$ with $\mathbb{P}(y_A \in A) = 1$. Denote by $\mathcal{D}(A)$ the family of all probability measures on A, i.e., of all possible distributions of a random point in A. Moreover, a *random finite set* is a measurable mapping $F : \Omega \to \mathcal{P}_f^d$.

Definition 2.1 Let $\mathcal{M} := (\mu_A)_{A \in \mathcal{P}_f^d}$ be a family of distributions s.t. $\mu_A \in \mathcal{D}(A)$ for each $A \in \mathcal{P}_f^d$. A sequence $(F_n)_{n \in \mathbb{N}}$ of random finite sets $F_n \subset \mathbb{Z}^d$ is called *incremental aggregation (with distribution family $\mathcal{M}$)*, if it satisfies the following conditions:

(i) $F_1 := \{y_1\}$, where $y_1 := 0 \in \mathbb{Z}^d$;
(ii) for any $n \in \mathbb{N}$, $F_{n+1} := F_n \cup \{y_{n+1}\}$, where y_{n+1} is a random point in $\mathbb{Z}^d$ whose conditional distribution given F_n is $\mu_{\partial F_n}$, i.e.,

$$\mathbb{P}(y_{n+1} = y \mid F_n = A) := \mu_{\partial A}(y) \qquad \text{for any } A \in \mathcal{P}_f^d \text{ and } y \in \mathbb{Z}^d.$$

F_n is called *cluster* or *aggregate at time n*, and $F_\infty := \bigcup_{n \in \mathbb{N}} F_n$ the *infinite cluster*. Observe that

$$0 \in F_1 \subset F_2 \subset F_3 \subset \cdots \subset F_\infty \subseteq \mathbb{Z}^d.$$

Moreover, for any $n \in \mathbb{N}$, almost surely F_n is connected and $\#F_n = n$. It is easy to see that $(F_n)_n$ is a Markov chain, in particular, F_{n+1} depends on F_n, but not on the order, in which the points have been added to F_n. Different aggregation models arise now by choosing different families $\mathcal{M}$ of distributions, in particular the Eden model and DLA are included in the framework:

Example 2.2 (Eden growth model) For each $A \in \mathcal{P}_f^d$, let μ_A be the uniform distribution on A. Then incremental aggregation with distribution family $\mathcal{M} = (\mu_A)_{A \in \mathcal{P}_f^d}$ is known as *Eden growth model*. Here, given $F_n = A \in \mathcal{P}_f^d$, each point y in the outer boundary ∂A of A has exactly the same chance (namely $1/\#\partial A$) of being the next point y_{n+1} added to the cluster.

Example 2.3 (Diffusion limited aggregation (DLA)) Let $(S_n^x)_{n \in \mathbb{N}_0}$ be a symmetric random walk on $\mathbb{Z}^d$ started at $x \in \mathbb{Z}^d$, i.e., $\mathbb{P}(S_0^x = x) = 1$ and for each $n \in \mathbb{N}$ and each $z \in \mathbb{Z}^d$

$$\mathbb{P}(S_n^x = y \mid S_{n-1}^x = z) = \frac{1}{2d}, \text{ for all neighbors } y \text{ of } z.$$

For $A \subset \mathbb{Z}^d$, let $T_A^x := \min\{n \in \mathbb{N} : S_n^x \in A\}$ be the *hitting time of A*. It is well-known that in $\mathbb{Z}^2$, a random walk started at x is *recurrent*, i.e. $\mathbb{P}(T_{\{x\}}^x < \infty) = 1$, and in $\mathbb{Z}^d$, $d \geq 3$ it is *transient*, i.e. $\mathbb{P}(T_{\{x\}}^x < \infty) < 1$. For $A \in \mathcal{P}_f^d$ and any point $x \in \mathbb{Z}^d \setminus A$, define a distribution on A by

$$H_A^x(y) := \mathbb{P}\left(S_{T_A^x}^x = y \mid T_A^x < \infty\right), \qquad y \in A.$$

It is a nontrivial but well-established fact that by setting

$$h_A(y) := \lim_{\|x\| \to \infty} H_A^x(y), \qquad y \in A$$

a probability measure h_A is defined on A, called the *harmonic measure* of A. *DLA* is now defined to be incremental aggregation with distribution family $(h_A)_{A \in \mathcal{P}_f^d}$. So in the DLA model particles perform random walks started at infinity and stick where they first hit the existing cluster. This produces sparse dendrite-like structures as in Fig. 1 (left).

Below in Sect. 5 we will discuss another class of models called ballistic aggregation. In these models the particles travel along straight lines instead of random walk paths. In order to introduce them properly, we will need some concepts from geometric probability.

3 Growth of Arms

Our aim is to study the speed of growth of aggregation clusters in incremental aggregation models. For any finite set $A \subset \mathbb{Z}^d$ with $0 \in A$, let

$$\text{rad}(A) := \max_{x \in A} \|x\| \tag{3.1}$$

denote its *radius*. (Recall that $\| \cdot \|$ denotes the Euclidean norm.) Observe that

$$\text{rad}(A) = \inf\{r > 0 : A \subset B(0, r)\},$$

justifying the terminology. Denote by $\text{diam}(A)$ the *diameter* of A and note that

$$\frac{1}{2}\,\text{diam}(A) \leq \text{rad}(A) \leq \text{diam}(A). \tag{3.2}$$

Instead of the radius one could similarly work with the diameter in the sequel, which might be more natural in any more general setting. Since we are only interested in incremental aggregation, which is started at 0, we stick to the radius for historic reasons. The following simple observation bounding the radius of a finite set $A \subset \mathbb{Z}^d$ in terms of its cardinality $\#A$ turns out to be very useful.

Lemma 3.1 *Let $A \subset \mathbb{Z}^d$ be a finite, connected set with $0 \in A$. Then*

$$\frac{1}{2}\left((\#A)^{1/d} - 1\right) \leq \text{rad}(A) \leq \#A. \tag{3.3}$$

Proof Let $r := \text{rad}(A)$. Then all points of $A \subset \mathbb{Z}^d$ are contained in the ball $B(0, r) \subset [-r, r]^d$. Hence $\#A \leq (2r + 1)^d$, implying the first inequality in (3.3). For the second one note that since $0 \in A$ and A is connected, A must contain a path P from 0 to some point $y \in A$ with $\|y\| = r$. Such a path contains at least r points and thus $\#A \geq \#P \geq r$, showing the second inequality in (3.3).   □

Note that the constant $\frac{1}{2}$ in (3.3) could be improved (and the -1 could be omitted for the price that the first inequality would then hold only for sufficiently large radii) but we do not need such improvements in the sequel. Note also that for the first inequality the assumption that A is connected and contains 0 is not needed.

Let $(A_n)_{n \in \mathbb{N}}$ be an increasing sequence of finite subsets of $\mathbb{Z}^d$. We are looking for a suitable exponent α (the *'growth rate'* of the sequence) and a suitable constant $c > 0$ such that

$$\text{rad}(A_n) \sim c \cdot (\#A_n)^{\alpha}, \quad \text{as } n \to \infty.$$

(Here $a_n \sim b_n$ means that the quotient a_n/b_n converges to 1 as $n \to \infty$.) Since such exponent α might not exist in general, we define the *lower and upper growth rate* of the sequence (A_n) by

$$\underline{\alpha}_f := \underline{\alpha}_f((A_n)) := \liminf_{n \to \infty} \frac{\log(\text{rad}(A_n))}{\log(\#A_n)} \quad \text{and} \quad \overline{\alpha}_f := \limsup_{n \to \infty} \frac{\log(\text{rad}(A_n))}{\log(\#A_n)}.$$

Alternatively, the *lower and upper fractal dimension* are defined by

$$\underline{\delta}_f := \liminf_{n \to \infty} \frac{\log(\#A_n)}{\log(\mathrm{rad}(A_n))} \quad \text{and} \quad \overline{\delta}_f := \limsup_{n \to \infty} \frac{\log(\#A_n)}{\log(\mathrm{rad}(A_n))}.$$

It is immediately clear from the definitions that fractal dimension and growth rate are directly connected. In general one has $\underline{\delta}_f = 1/\overline{\alpha}_f$ and $\overline{\delta}_f = 1/\underline{\alpha}_f$. This means in particular that an upper bound on the growth rate implies a lower bound on the fractal dimension and vice versa.

$\#A_n$ may be interpreted as the *volume* of the cluster A_n. (We may identify $\#A_n$ with the union $\Box A_n := \bigcup_{y \in A_n} C_y$, where $C_y := y + [-\frac{1}{2}, \frac{1}{2}]^d$ is the unit cube centered at y. Then the volume of $\Box A_n$ is exactly $\#A_n$.) Moreover, $\mathrm{rad}(A_n)$ may be interpreted as the diameter, cf. (3.2), which justifies calling the exponents $\underline{\delta}_f$ and $\overline{\delta}_f$ fractal dimensions.

The trivial bounds on the radius stated in Lemma 3.1 imply immediately some general bounds on the fractal dimensions and growth rates.

Proposition 3.2 *Let* $(A_n)_{n \in \mathbb{N}}$ *be an increasing sequence of finite, connected subsets of* $\mathbb{Z}^d$ *with* $0 \in A_1$. *Then*

$$1 \le \underline{\delta}_f \le \overline{\delta}_f \le d \quad \text{and hence} \quad 1/d \le \underline{\alpha}_f \le \overline{\alpha}_f \le 1.$$

Proof The lower bound 1 for $\underline{\delta}_f$ follows by applying the right hand side inequality in (3.3) to A_n and taking the lim inf. The upper bound for $\overline{\delta}_f$ follows similarly by applying the left hand side inequality in (3.3) to A_n and taking the lim sup. $\qquad \square$

In incremental aggregation models, the sequence $(F_n)_{n \in \mathbb{N}}$ consists of random sets and therefore, for each $n \in \mathbb{N}$, $\mathrm{rad}(F_n)$ is a nonnegative random variable (while $\#F_n = n$ is constant almost surely). Hence the fractal dimensions and growth rates are random variables, too. Observe that they satisfy almost surely the bounds stated in Proposition 3.2.

Let now $(F_n)_{n \in \mathbb{N}}$ be DLA in $\mathbb{Z}^d$, $d \ge 2$ as defined in Example 2.3. The following celebrated result due to Kesten provides an almost sure upper bound on the radii of DLA clusters. For $d = 3$, a log-term appears which was improved slightly by Lawler [15] and later again by Benjamini and Yadin [16].

Theorem 3.3 (cf. [13, 14, 16]) *For DLA* $(F_n)_n$ *in* $\mathbb{Z}^d$ *there exists a constant* $c > 0$ *(depending only on* d*) such that almost surely for* n *sufficiently large*

$$\mathrm{rad}(F_n) \le \begin{cases} c\, n^{2/3}, & \text{if } d = 2, \\ c\, n^{1/2}(\log n)^{1/2}, & \text{if } d = 3, \\ c\, n^{2/(d+1)}, & \text{if } d \ge 4. \end{cases}$$

Theorem 3.3 implies the following lower bounds on the fractal dimension.

Corollary 3.4 *For DLA* $(F_n)_{n \in \mathbb{N}}$ *in* $\mathbb{Z}^d$, $d \ge 2$, *one has almost surely*

$$\underline{\delta}_f \ge \frac{d+1}{2} \quad \text{and} \quad \overline{\alpha}_f \le \frac{2}{d+1}.$$

Proof This follows immediately by inserting the estimate in Theorem 3.3 into the definition of the lower fractal dimension (and the upper growth exponent). For $d = 2$, for instance, this yields almost surely

$$\underline{\delta}_f = \liminf_{n \to \infty} \frac{\log n}{\log(\mathrm{rad}(F_n))} \ge \lim_{n \to \infty} \frac{\log n}{\log(c n^{2/3})} = \frac{3}{2}.$$

$\qquad \square$

Simulations of the model suggest that the above bounds are not sharp. In $\mathbb{Z}^2$ they suggest that the fractal dimension exists and equals $\delta_f = \frac{5}{3} \approx 1.66$, which is strictly larger than the rigorous bound stated above. For general $d \geq 2$, it is conjectured, cf. e.g. [5], that for DLA in $\mathbb{Z}^d$ the dimension is given by

$$\delta_f = \frac{d^2 + 1}{d + 1}.$$

4 DLA and Kesten's Method

In his paper [13], Kesten used a certain strategy of proof for his bounds for the DLA model in $\mathbb{Z}^2$, which is also described in Lawler's book, see [15, §2.6], and generalized to arbitrary graphs with bounded degree in [16]. It turns out that this method can be adapted to provide some bounds for any incremental aggregation model. The next theorem below describes this general method. Its proof will be discussed later in Sect. 6. Let $\mathcal{P}_\partial^d$ be the family of outer boundaries of aggregation clusters in $\mathbb{Z}^d$, that is, let

$$\mathcal{P}_\partial^d := \left\{ B \in \mathcal{P}_f^d : \exists A \in \mathcal{P}_f^d \text{ such that } A \text{ is connected, } 0 \in A \text{ and } \partial A = B \right\}.$$

Recall that in any incremental aggregation model, at any time step n the next point y_{n+1} to be attached is chosen from the outer boundary ∂F_n of the current cluster F_n. Therefore, it is in fact sufficient to specify the distributions μ_B for all $B \in \mathcal{P}_\partial^d$ in order to determine an incremental aggregation model completely. As a consequence, it is sufficient to impose conditions on the distribution family only for the distributions of outer boundary sets as in the statement below. Other sets are not relevant for the model.

Theorem 4.1 (Kesten's method) *Let $\mathcal{M} = (\mu_A)_{A \in \mathcal{P}_f^d}$ be some family of distributions such that $\mu_A \in \mathcal{D}(A)$ for each $A \in \mathcal{P}_f^d$. Suppose there exist some positive constants q and C such that for all $r > 1$, any $B \in \mathcal{P}_\partial^d$ with radius at least r and any $z \in B$,*

$$\mu_B(z) \leq C\,r^{-q}. \tag{4.1}$$

Then there is a constant c, such that incremental aggregation $\mathcal{F} = (F_n)_{n \in \mathbb{N}}$ with distribution family $\mathcal{M}$ satisfies almost surely

$$\mathrm{rad}(F_n) \leq c\,n^{1/(q+1)}$$

for n sufficiently large.

The method is rather crude. Roughly it says that if for all relevant distributions in the family $\mathcal{M}$ the mass concentrated on single vertices is not too large compared to the radius of the corresponding cluster (implying that the probability mass must be spread out over a significant number of vertices), then some bound on the radial growth follows. This is plausible. The more spread out distributions are, the more options there are for the next particle to be placed. As only few of the possible locations are extremal in the sense that they lead to radial growth of the cluster, this limits the speed of growth.

In [16], the Assumption (4.1) is called a φ-radius Beurling estimate in the context of the DLA model (for the function $\varphi(r) = cr^{-q}$). Our proof of Theorem 4.1 in Sect. 6 below follows essentially the proof of Lawler [15, § 2.6] but one might also want to compare it with [16, Lemma 2.4 and Theorem 2.6]. The important observation is that the relevant arguments in these proofs do not only apply to DLA but to any incremental aggregation.

To illustrate Kesten's method, we briefly describe how it is applied to obtain the known bounds in $\mathbb{Z}^2$ for DLA and the Eden model. In the next section, we will use it to study ballistic aggregation. For the DLA model in $\mathbb{Z}^2$, one can use the following well-known estimate for harmonic measures.

Proposition 4.2 (cf. e.g. [15, Proposition 2.5.2]) *For any $r > 1$, any connected set $A \subset \mathcal{P}_f^d$ with $0 \in A$ and radius at least r, and any $z \in A$, one has*

$$h_A(z) \leq \begin{cases} C \, r^{-1/2}, & \text{if } d = 2, \\ C \, (\log r)^{1/2} r^{-1}, & \text{if } d = 3, \\ C \, r^{-1}, & \text{if } d \geq 4. \end{cases}$$

Observe that outer boundaries need not be connected in $(\mathbb{Z}^d, E)$ such that the above estimates are not directly applicable. However, since any path started outside a set A and ending in A will first hit ∂A before hitting A (see e.g. [25, Lemma 2.23] or [26, Lemma 2.1] for a proof), we have

$$h_{\partial A} = h_{A \cup \partial A} \quad \text{for any } A \in \mathcal{P}_f^d.$$

Hence the above estimates hold also for any boundary set $B \in \mathcal{P}_\partial^d$. Applying Theorem 4.1 to DLA in $\mathbb{Z}^2$, for which, by Proposition 4.2, the hypothesis is satisfied for $q = 1/2$, we conclude the existence of a constant $c > 0$ such that almost surely $\text{rad}(F_n) \leq c \, n^{3/2}$ for n sufficiently large. This is the bound stated in Theorem 3.3 above.

Remark 4.3 For DLA in $\mathbb{Z}^3$, the bound for the harmonic measure stated in Proposition 4.2 implies that the hypothesis of Theorem 4.1 is satisfied for any $q < 1$. Thus, this theorem yields for any $p > \frac{1}{2}$ the existence of a constant c_p such that a.s. the bound $\text{rad}(F_n) \leq c_p \, n^p$ holds for n sufficiently large. This is enough to conclude (by letting $p \to 1/2$) that a.s. $\underline{\delta}_f \geq 2$, as stated in Corollary 3.4, but it is not enough to get the logarithmic correction for the bound on the radius stated in Theorem 3.3 for the case $d = 3$. For this a slight refinement of Kesten's method is necessary, see e.g. [16].

For $d \geq 4$, Theorem 4.1 yields the existence of some $c > 0$ such that a.s. the bound $\text{rad}(F_n) \leq c \, n^{1/2}$ holds for n sufficiently large. This is not as good as the bound stated in Theorem 3.3 and shows the rather poor performance of the method in higher dimensions. It is known that the exponent $q = 1$ in Proposition 4.2 is optimal. There are sets A (e.g. (discrete) line segments) for which the harmonic measure has atoms of this order, cf. e.g. [15, §2.4] for details. Hence, Theorem 4.1 cannot provide a better bound for the growth rate in this case.

For the Eden growth model in $\mathbb{Z}^d$ (as defined in Example 2.2 above), we have the following estimate for the defining family of distributions (μ_A):

Lemma 4.4 *For any $r > 1$, any boundary set $B \in \mathcal{P}_\partial^d$ with $\text{rad}(B) \geq r$ and any $z \in B$,*

$$\mu_B(z) \leq \sqrt{d}(2(d-1))^{-1} r^{-1}.$$

Proof Denote by $e_1, \ldots, e_d$ the coordinate directions. For $B \in \mathcal{P}_\partial^d$ there is a finite connected set $A \in \mathcal{P}_f^d$ with $0 \in A$ and $\text{rad}(A) \geq r - 1$, such that $\partial A = B$. Because of its radius A must contain a point y with $\|y\| \geq r - 1$. Hence, there is a coordinate direction e_i such that the i-th coordinate of y satisfies $|y_i| \geq (r-1)/\sqrt{d}$. Without loss of generality we can assume $i = 1$ and $y_1 > 0$. Since y_1 is an integer, we have $y_1 \geq \lfloor (r-1)/\sqrt{d} \rfloor + 1$ (where $\lfloor x \rfloor$ denotes the integer part of $x > 0$) such that $y_1 + 1 \geq (r-1)/\sqrt{d} + 1 \geq r/\sqrt{d}$.

Since A is connected, also its projection onto the e_1-axis must be connected. Hence, for each $k = 0, \ldots, y_1$ there is a vertex z_k in A whose first coordinate is k. If one moves from z_k in direction $\pm e_i$, $i = 2, \ldots, d$ to the first vertex outside A, then this new vertex will be in $\partial A = B$. It is clear that all the $2(d-1)(y_1+1)$ vertices of B reached in this way are distinct, which implies that

$$\#B \geq 2(d-1)(y_1+1) \geq \frac{2(d-1)}{\sqrt{d}} r.$$

Hence, $\mu_B(z) = (\#A)^{-1} \leq \sqrt{d}(2(d-1))^{-1}r^{-1}$ as asserted.   □

Applying now Kesten's method to the Eden growth model $(F_n)_n$ in $\mathbb{Z}^d$ using the estimate provided in Lemma 4.4 (for the exponent $q = 1$), we infer that there is a constant $c > 0$, such that almost surely

$$\operatorname{rad}(F_n) \leq c\, n^{1/2}$$

for n sufficiently large. This implies $\delta_f \geq 2$ for the lower fractal growth dimension of the Eden model in $\mathbb{Z}^d$. For $d = 2$, by Proposition 3.2, we therefore recover

$$\delta_f = 2.$$

It is also well-known that $\delta_f = d$ for the Eden model in $\mathbb{Z}^d$, $d \geq 3$, but Kesten's method as stated in Theorem 4.1 is not capable of providing such a result, as this would require an estimate of the form (4.1) with $q = d - 1$, which is simply not true. Indeed, for any $r \in \mathbb{N}$ there is a cluster A such that for $B = \partial A$, $\operatorname{rad}(B) = r$ and $\mu_B(z) \geq (2(d-1))^{-1}r^{-1}$. (Take e.g. $A := \{(k, 0, \ldots, 0) \in \mathbb{Z}^d : k \in \{0, 1, \ldots, r-1\}$. Then $\#B = 2 + 2(d-1)r \leq 2(d-1)r$ and so $\mu_B(z) \geq (2(d-1))^{-1}r^{-1}$.) This precludes an estimate of the form (4.1) to hold for any $q > 1$.

The hypothesis in Kesten's method is rather restrictive. Requiring some deterministic estimate to be satisfied for all outer boundaries and at all locations z leads to an exponent q that is often too small in order to provide a good bound for the radial growth. Although the methods is able to provide lower bounds on radial growth in any dimension d, it seems that the method is strong only in $\mathbb{Z}^2$.

It seems plausible, that it should be enough to have a bound on $\mu_{\partial A}(z)$ available for outer boundaries of 'typical' aggregation clusters A (or for 'most of them'). However, this will require further investigation and is not covered by Theorem 4.1 above. Another approach (which leads in fact to the growth bounds for DLA stated above) are estimates in terms of the volume of the clusters rather than its radius, which will be a topic of further investigation.

5 Ballistic Aggregation

We now introduce the ballistic model rigorously and discuss its growth properties. Ballistic aggregation in $\mathbb{Z}^d$ will be defined as incremental aggregation $(F_n)_{n \in \mathbb{N}}$ with a suitable distribution family $\mathcal{M} = (b_A)_{A \in \mathcal{P}_f^d}$. In order to determine the model all we have to do is to choose the family $\mathcal{M}$, i.e., to fix a distribution b_A on each finite set $A \in \mathcal{P}_f^d$. Roughly, for each $z \in A$, $b_A(z)$ will be determined by the probability that a 'directed random line through A hits z first'. In order to define this properly, we recall some ideas from stochastic geometry, in particular the concept of an *isotropic random line*. In what follows, we will view $\mathbb{Z}^d$ also as a subset of $\mathbb{R}^d$ embedded in the natural way.

Let $A(d, 1)$ be the space of lines in $\mathbb{R}^d$ (i.e., the affine Grassmannian of 1-flats). We equip it with the usual hit-or-miss topology and the associated Borel σ-algebra $\mathcal{A}(d, 1)$, see e.g. [27, Chap. 13] for details. For any compact set $K \subset \mathbb{R}^d$ let

$$[K] := \{L \in A(d, 1) : L \cap K \neq \emptyset\}.$$

Then $\mathcal{A}(d, 1) = \sigma(\{[K] : K \in \mathcal{K}^d\})$, where $\mathcal{K}^d$ denotes the family of all *convex bodies*, that is, compact, convex sets in $\mathbb{R}^d$.

It is well-known, cf. e.g. [27, Theorem 13.2.12], that there is a unique Haar measure μ_1 on $A(d, 1)$ (i.e. a Euclidean motion-invariant Radon measure) normalized such that

$$\mu_1([B_d]) = \kappa_{d-1}.$$

Here B_n is the unit ball in $\mathbb{R}^n$ and $\kappa_n := \lambda_n(B_n)$ denotes its volume. We point out that from a physical viewpoint μ_1 is the most natural choice for modelling the trajectory of an arriving particle. Due to its invariance properties, it does not prefer any directions or locations. (In the physics literature, the precise choice of the underlying line measure in ballistic models is not always disclosed. In some cases, the directions of the sampled lines are concentrated on the directions of the coordinate axes, also Meakin [4] uses a line measure different to μ_1, see also Remark 5.2 below.

Observe that, by the Crofton formula, cf. e.g. [27, Theorem 5.1.1], for any convex body $K \in \mathcal{K}^d$,

$$\mu_1([K]) = \int_{A(d,1)} \mathbf{1}\{K \cap L \neq \emptyset\}\mu_1(dL) \tag{5.1}$$
$$= \int_{A(d,1)} V_0(K \cap L)\mu_1(dL) = \alpha_d\, V_{d-1}(K),$$

where the constant is given by $\alpha_d := \frac{2(d-1)!\kappa_{d-1}}{d!\kappa_d} = \frac{2\kappa_{d-1}}{d\kappa_d}$ and $V_j(K)$ is the intrinsic volume of K of degree j. If K has nonempty interior, then $V_{d-1}(K)$ is half the surface area of K. This means, the measure of lines hitting a given convex body K is up to some universal constant given by the surface area of K. Following [27, Definition 8.4.2], for any convex body $K \in \mathcal{K}^d$ with $V_{d-1}(K) > 0$, an *isotropic random line through K* is defined to be a random variable $L : \Omega \to A(d, 1)$ with distribution given by

$$\mathbb{P}(L \in \mathcal{A}) := \mathbb{P}^K(\mathcal{A}) := \frac{\mu_1(\mathcal{A} \cap [K])}{\mu_1([K])}, \qquad \mathcal{A} \in \mathcal{A}(d, 1).$$

This definition extends without any problem to arbitrary compact sets $K \subset \mathbb{R}^d$, provided that $\mu_1([K]) > 0$. However, the convenient interpretation (5.1) of μ_1 in terms of V_{d-1} (or the surface area) is no longer valid if K is not convex. Observe that for any compact $K \subset \mathbb{R}^d$, one has $\mu_1([K]) \leq \mu_1([\text{conv}(K)]) < \infty$, where $\text{conv}(K)$ denotes the convex hull of K. Thus, for any compact set $K \subset \mathbb{R}^d$ with $\mu_1([K]) > 0$, the distribution $\mathbb{P}^K$ is well-defined and so is an *isotropic random line through K*. In $\mathbb{R}^2$ one has the following additional property, which turns out to simplify the analysis significantly: for any connected, compact set $C \subset \mathbb{R}^2$,

$$[C] = [\text{conv}(C)].$$

Therefore, $\mathbb{P}^C = \mathbb{P}^{\text{conv}(C)}$ (provided $V_1(\text{conv}(C)) > 0$, i.e., C not a singleton). Thus in dimension 2 one can work with convex hulls instead of the original sets. Unfortunately, this is not possible in any higher dimension.

For any finite set $A \in \mathcal{P}_f^d$ denote by $\square A$ the union of grid boxes centered in A, that is, let

$$\square A := \bigcup_{z \in A} C_z \qquad \text{where } C_z = \left[-\frac{1}{2}, \frac{1}{2} \right]^d + z.$$

Then, for any $d \geq 2$ and $A \in \mathcal{P}_f^d \setminus \{\emptyset\}$, the distribution $\mathbb{P}^{\square A}$ of an isotropic random line through $\square A$ is well-defined. The idea for the definition of the ballistic measure b_A is now as follows: We identify A with $\square A$ and generate an isotropic random line L through $\square A$. We choose randomly one of the two directions on L. Then for any $z \in A$, we let $b_A(z)$ be the probability that, when traveling along L in the chosen direction, the square C_z is the first square in $\square A$ visited by L. More formally, for each $L \in A(d, 1)$ let us denote by $v_L \in \mathbb{S}^{d-1}$ one of the two directions spanning L. (We choose v_L such that the mapping $L \mapsto v_L$ is measurable.) For $A \in \mathcal{P}_f^d$ and $L \in A(d, 1)$, let

$$L^A := \{z \in A : C_z \cap L \neq \emptyset\},$$

that is, L^A is the set of those points in A whose associated boxes are intersected by L. Traversing L in direction v_L induces an order in L^A. (For certain lines the order in which the boxes are visited is not well-defined, namely if L hits a box first at an intersection point of two or several boxes not visited before. In such a case the order can be made unique by fixing some order of the boxes C_z in $z \in \mathbb{Z}^d$, such as the one induced by the lexicographic order in $\mathbb{Z}^d$. However, for μ_1-almost all lines L the order in L^A is well-defined without this.) Denote by $\min(L^A)$ and $\max(L^A)$ the first and last point in L^A 'visited' by L.

Definition 5.1 (*Ballistic measure*) Let $A \in \mathcal{P}_f^d \setminus \{\emptyset\}$. For $z \in \mathbb{Z}^d$ and $L \in A(d, 1)$ define

$$b_A(z, L) := \begin{cases} 1, & \text{if } L^A = \{z\}, \\ \frac{1}{2}, & \text{if } \#(L^A) > 1 \text{ and } z \in \{\min(L^A), \max(L^A)\}, \\ 0, & \text{otherwise.} \end{cases}$$

Then a probability measure b_A on A is well-defined by setting

$$b_A(z) := \int_{A(d,1)} b_A(z, L) \mathbb{P}^{\square A}(dL), \qquad z \in \mathbb{Z}^d.$$

We call b_A the *ballistic measure* on A.

Observe that $b_A(z) = 0$ for $z \notin A$ and $\sum_{z \in A} b_A(z) = 1$. Hence, b_A is indeed a probability measure which is concentrated on A, that is, $b_A \in \mathcal{D}(A)$. Note also that, for any $z \in A$,

$$b_A(z) = \int_{A(d,1)} b_A(z, L) \mathbb{P}^{\square A}(dL) \leq \int_{A(d,1)} \mathbf{1}\{L \in [C_z]\} \mathbb{P}^{\square A}(dL) = \mathbb{P}^{\square A}([C_z]). \quad (5.2)$$

This gives a first upper bound of the probability mass of b_A at single vertices. *Ballistic aggregation (BA) on $\mathbb{Z}^d$* is now defined to be incremental aggregation with distribution family $(b_A)_{A \in \mathcal{P}_f^d}$, where b_A is the ballistic measure on A.

Remark 5.2 It is a natural question to ask at this point, how an isotropic random line (IRL) through a given set K may be sampled and how ballistic clusters can be generated according to this model. A useful fact to note is that the invariant line measure μ_1 permits rejection sampling, that is, given sets $K \subset M \subset \mathbb{R}^d$ with $\mu_1([K]) > 0$ (and thus $\mu_1([M]) > 0$), one

can generate an IRL through the larger set M and if it intersects K, then it is an IRL through K. More formally, if L_K is an IRL through K and L_M one through M, then

$$\mathbb{P}(L_K \in \mathcal{A}) = \mathbb{P}(L_M \in \mathcal{A} | L_M \in [K]), \quad \mathcal{A} \in \mathcal{A}(d, 1),$$

see e.g. [27, Lemma 8.4.1]. This property of μ_1 allows to choose a ball B large enough to contain the current cluster $\square A$ and generate IRLs through B until one hits the cluster $\square A$. An IRL through a ball B of radius $r > 0$ centred at 0 may be generated as follows. Fix a line $L \in G(d, 1)$ containing the origin and let $B' := B \cap L^\perp$. Choose a random point X uniformly in B' and independently a random rotation ϑ uniformly in SO_d, i.e., according to the unique Haar measure ν on SO_d with $\nu(SO_d) = 1$. Then $L_B = \vartheta(L + X)$ is an IRL through B, cf. e.g. [27, p. 340]. For $d = 2$, this amounts to choosing a distance p uniformly in $(0, r)$ and an angle α uniformly in $[0, 2\pi)$. Then the random line $L_{p,\alpha} := \{x \in \mathbb{R}^2 : \langle x, n_\alpha \rangle = p\}$ is an IRL through B. Here n_α denotes the unit normal having angle α with the e_1-axis. (We point out that other 'natural' generation mechanisms of a random line through a ball, e.g., choosing two random points in the ball or on its boundary or choosing a random point on the boundary and a random angle, do not necessarily produce an IRL, see also Bertrand's paradoxon in the literature.) For $d = 3$ (and similarly in higher dimensions), a uniform random rotation ϑ may be generated by choosing three random points X_1, X_2, X_3 uniformly in the ball B, computing an orthonormal basis Z_1, Z_2, Z_3 using the Gram-Schmidt algorithm and replacing Z_3 by $-Z_3$, if necessary, to obtain a positively oriented ONB. Then the random rotation ϑ defined by $\vartheta(e_i) = Z_i, i = 1, 2, 3$ is uniform in SO_3.

With this knowledge how to sample IRLs in balls, it is now easy to generate ballistic clusters. To add a particle to an existing cluster, it is enough to produce IRLs in a ball containing the cluster, reject them until the cluster is hit for the first time, choose uniformly one of the two directions on the generated line and find the first point neighboring the current cluster, when traversing the line in this direction. □

For the ballistic model in $\mathbb{Z}^d$, $d \geq 2$, we have the following main result concerning the radial growth of clusters, which parallels the one obtained by Kesten for DLA.

Theorem 5.3 *Let $d \in \mathbb{N}$, $d \geq 2$ and $(F_n)_{n \in \mathbb{N}}$ be ballistic aggregation in $\mathbb{Z}^d$. There exists a constant $c > 0$ such that almost surely for n sufficiently large*

$$\mathrm{rad}(F_n) \leq c \, n^{1/2}.$$

Combining this bound with the trivial lower bound provided by Propositon 3.2, we conclude that in $\mathbb{Z}^2$ the fractal dimension δ_f of BA clusters exists and equals 2. This confirms a long standing conjecture in the physics literature, cf. [4, 5, 28, 29].

Corollary 5.4 *For the ballistic model in $\mathbb{Z}^2$, one has $\delta_f = 2$ almost surely.*

Proof On the one hand, by Propositon 3.2, $\overline{\delta}_f \leq 2$ almost surely. On the other hand, Theorem 5.3 implies

$$\underline{\delta}_f = \liminf_{n \to \infty} \frac{\log n}{\log(\mathrm{rad}(F_n))} \geq \lim_{n \to \infty} \frac{\log n}{\log(c n^{1/2})} = 2.$$

□

Remark 5.5 We can even get a slightly stronger conclusion from the above theorem regarding the asymptotic size of BA clusters. The ballistic model in $\mathbb{Z}^2$ satisfies almost surely

$$\liminf_{n \to \infty} \frac{\#F_n}{(\mathrm{rad}(F_n))^2} \geq \liminf_{n \to \infty} \frac{n}{c^2 n} = c^{-2} > 0.$$

This may be interpreted as saying that asymptotically BA clusters cover a positive portion of the plane. (More precisely, the quotient of the area covered by the cluster $\square F_n$ (given by $\# F_n$) and the area of the smallest ball containing $\square F_n$ ($\approx \pi \cdot \mathrm{rad}(F_n)^2$) is asymptotically bounded from below, as $n \to \infty$.)

For $d \geq 3$, the conclusion regarding the fractal dimension of the BA model is not quite as strong as for $d = 2$.

Corollary 5.6 *For the ballistic model in $\mathbb{Z}^d$, $d \geq 3$, one has $\underline{\delta}_f \geq 2$ almost surely.*

As in the case $d = 2$, this can be directly deduced from Theorem 5.3. Unfortunately, this lower bound for the fractal dimension is far from the conjectured value $\delta_f = d$ in this case.

The proof of Theorem 5.3 will be based on Kesten's method (Theorem 4.1 above). In order to apply it, a suitable estimate for the local growth of the ballistic measures b_A is required, which we state now.

Proposition 5.7 *For any $d \in \mathbb{N}$, $d \geq 2$, there exists a constant $C_d > 0$ such that, for any $r \geq 1$, any connected set $A \in \mathcal{P}_f^d$ with $0 \in A$ and $\mathrm{rad}(A) \geq r$, and any $z \in A$,*

$$b_A(z) \leq C_d \, r^{-1}.$$

For $d = 2$, $C_2 = 2$ is a suitable constant.

Before we provide a proof of Proposition 5.7, we first clarify that Theorem 5.3 easily follows from this statement.

Proof of Theorem 5.3 By Proposition 5.7, the hypothesis in Kesten's method (Theorem 4.1) is satisfied for the exponent $q = 1$ for the ballistic model $(F_n)_{n \in \mathbb{N}}$ on $\mathbb{Z}^d$ for any $d \geq 2$. Hence, from this theorem, the stated bounds on the radial growth of this model (with the exponent $1/(1 + q) = 1/2$) follow at once.

$\square$

Observe that any improvement on the exponent $q = 1$ in Proposition 5.7 (i.e., any larger q) would allow to deduce from Kesten's method a better bound for the radial growth. Unfortunately, the exponent $q = 1$ in Proposition 5.7 turns out to be optimal in any dimension $d \geq 2$, as the following remark clarifies.

Remark 5.8 The exponent $q = 1$ in Proposition 5.7 is optimal (largest possible) in any space dimension $d \geq 2$. Consider for any $d \in \mathbb{Z}$, $d \geq 2$ and any integer $r \geq 1$ the set

$$A = A_r := \left\{ (s, 0, \ldots, 0) \in \mathbb{Z}^d : s \in \{0, \ldots, r\} \right\}.$$

Observe that $\square A$ is a union of $r + 1$ unit size boxes placed side by side in a row. Hence $\square A$ is convex and $\mathrm{rad}(A) = r$. Moreover, noting that for any directed line L hitting the cube C_0 at all, C_0 is always the first or last cube of $\square A$ visited by L, we obtain for the ballistic measure at $z = 0$ (and similarly at $z = (r, 0, \ldots, 0)$)

$$b_A(z) = \int_{A(d,1)} b_A(z, L) \mathbb{P}^{\square A}(dL) \geq \frac{1}{2} \int_{A(d,1)} \mathbf{1}\{L \in [C_z]\} \mathbb{P}^{\square A}(dL) = \frac{1}{2} \mathbb{P}^{\square A}([C_z])$$

$$= \frac{1}{2} \frac{\mu_1([C_z])}{\mu_1([\square A])} = \frac{1}{2} \frac{V_{d-1}(C_z)}{V_{d-1}(A_r)} = \frac{1}{2} \cdot \frac{d}{1 + (d-1)(r+1)} = \frac{1}{2} \cdot \frac{d}{(d-1)r + d}.$$

For any $r \geq d$, the denominator in the last expression is bounded from above by $d\,r$ and therefore we obtain $b_A(z) \geq 1/2 \cdot r^{-1}$. Hence a bound as in Proposition 5.7 valid for the ballistic measures b_A of all finite sets $A \in \mathcal{P}_f^d$ cannot be true for any $q > 1$.

Proof of Proposition 5.7 We first discuss the case $d = 2$, for which our argument is much simpler than for the general case. Fix $r \geq 1$. Let $A \in \mathcal{P}_f^2$ be a connected set with $0 \in A$ and $\text{rad}(A) \geq r$. Note that the connectedness implies $[\Box A] = [\text{conv}(\Box A)]$. The set $\Box A$ contains the square C_0 and another unit square C_y with $y \in \mathbb{Z}^2$ and $\|y\| \geq r$. Hence $\text{conv}(\Box A)$ contains $\text{conv}(C_0 \cup C_y)$ and thus a rectangle R_r with sidelengths r and 1. (Choose one side of R_r parallel to the vector y.) Using (5.2) and the properties of the measure μ_1, we conclude that, for any $z \in A$,

$$b_A(z) \leq \mathbb{P}^{\Box A}([C_z]) = \frac{\mu_1([C_z])}{\mu_1([\text{conv}(\Box A)])} \leq \frac{\mu_1([C_0])}{\mu_1([R_r])} = \frac{V_1(C_0)}{V_1(R_r)} = \frac{2}{1+r} \leq 2r^{-1}.$$

This shows the claimed estimate for $d = 2$ and also that $C_2 = 2$ is a suitable constant in this case.

Now let us turn to the case $d \geq 3$. We start similarly as for $d = 2$. Let $r \geq 1$ and let $A \in \mathcal{P}_f^d$ be a connected set with $0 \in A$ and $\text{rad}(A) \geq r$. As before, the set $\Box A$ contains the square C_0 and another unit square C_y with $y \in \mathbb{Z}^d$ and $\|y\| \geq r$, but now we can not simply replace $\Box A$ by its convex hull, since the latter set may be hit by significantly many more lines than $\Box A$.

Since A is connected, it must contain a path from 0 to y, that is, there are $m \in \mathbb{N}$ and points $p_0 := 0, p_1, \ldots, p_{m-1}, p_m := y \in A$ such that $\|p_i - p_{i-1}\| = 1$ for $i = 1, \ldots, m$. Let $\Gamma := \{p_0, \ldots, p_m\}$ and let G denote (the graph of) the shortest curve connecting these points in the given order. Notice that G consists of axis-parallel segments of length 1. For $F \subset \mathbb{R}^d$ and $\varepsilon > 0$ let

$$F_{\oplus \varepsilon} := \left\{ x \in \mathbb{R}^d : \inf_{y \in F} \|x - y\| \leq \varepsilon \right\}$$

denote the ε-parallel set of F. Observe that $G_{\oplus 1/2} \subset \Box \Gamma$. (Indeed, any point in $G_{\oplus 1/2}$ is contained in some ball $B(x, \frac{1}{2})$ with $x \in G$. If $x \in \Gamma$, then $B(x, \frac{1}{2}) \subset C_x$ and if x is on the segment $S(p_k, p_{k+1})$ connecting p_k and p_{k+1}, then obviously $B(x, \frac{1}{2}) \subset C_{p_k} \cup C_{p_{k+1}}$.) This yields

$$\mu_1([G_{\oplus 1/2}]) \leq \mu_1([\Box \Gamma]) \leq \mu_1([\Box A]).$$

We claim now that there is a constant $\widetilde{C}_d$ (independent of G and r) such that

$$\mu_1([G_{\oplus 1/2}]) \geq \widetilde{C}_d r. \tag{5.3}$$

Using (5.2), the properties of the measure μ_1 and (5.1), we conclude from (5.3) that, for any $z \in A$,

$$b_A(z) \leq \mathbb{P}^{\Box A}([C_z]) = \frac{\mu_1([C_z])}{\mu_1([\Box A])} \leq \frac{\mu_1([C_0])}{\mu_1([G_{\oplus 1/2}])} \leq \frac{\alpha_d V_{d-1}(C_0)}{\widetilde{C}_d r} = \frac{\alpha_d d}{\widetilde{C}_d} r^{-1}.$$

This shows the estimate stated in Proposition 5.7 (for the constant $C_d := \alpha_d d / \widetilde{C}_d$).

It remains to provide a proof of (5.3). The rough idea here is to compare the measure of lines through the tube $G_{\oplus 1/2}$ with the measure of lines through the tube $S_{\oplus 1/2}$, where $S := S(0, y)$ is the segment connecting 0 and y, and to observe that the latter is of order r. Observe that any hyperplane hitting S does also hit the curve G, since both curves have the same endpoints. (Indeed, if the hyperplane contains an endpoint of S then it obviously intersects also G. Otherwise the hyperplane contains an inner point of S and separates the common endpoints of S and G. But then the curve G must also intersect the hyperplane when passing from one side to the other.) This will be useful for decomposing $\mu_1([G_{\oplus 1/2}])$.

Let $A(d, d-1)$ denote the set of all $(d-1)$-flats and μ_{d-1} the unique Euclidean motion invariant measure on $A(d, d-1)$ such that $\mu_{d-1}([B^d]) = 2$, see [27, Ch. 13.2] for details. For $H \in A(d, d-1)$ denote by $A(H, 1)$ the set of lines in H and let μ_1^H be the invariant line measure on $A(H, 1)$ (which is the image measure of μ_1 on $\mathbb{R}^{d-1}$ under some isometry from $\mathbb{R}^{d-1}$ to H). Recall that for $d \geq 3$ the measure of any set $\mathcal{A}$ of lines may be determined by computing its line measure in hyperplanes and then integrating over all hyperplanes. More precisely, for any $\mathcal{A} \in \mathcal{A}(d, 1)$,

$$\mu_1(\mathcal{A}) = \int_{A(d,d-1)} \mu_1^H(\mathcal{A})\mu_{d-1}(dH),$$

see e.g. [27, Thm. 7.1.2 and the subsequent Remark], and in particular [27, Lemma 13.2.2] for the measurability of the mapping $H \mapsto \mu_1^H(\mathcal{A})$.

Let $s := y/\|y\|$ be the direction of the segment S. For any $H \in A(d, d-1)$, denote by $v_H \in \mathbb{S}^{d-1}$ some unit normal of H and by p_H the corresponding (signed) distance to 0 (such that $H = \{x \in \mathbb{R}^d : \langle x, v_H \rangle = p_H\}$). Note that v_H is uniquely determined for a.a. H if we require additionally $\langle s, v_H \rangle \geq 0$. For a lower bound, we do not need to compute the line measure in all hyperplanes. We restrict our attention to the following set. Fix some $\delta > 0$ and let

$$\mathcal{H} := \{H \in A(d, d-1) : \langle v_H, s \rangle \geq \delta \text{ and } H \cap S \neq \emptyset\}.$$

Observe that, for each $H \in \mathcal{H}$, we have $G \cap H \neq \emptyset$, i.e. there is some point $x \in G \cap H$, implying that $G_{\oplus 1/2}$ contains the closed ball $B(x, \frac{1}{2})$ with radius $\frac{1}{2}$ and center x. Hence, by (5.1),

$$\mu_1^H([G_{\oplus 1/2}]) \geq \mu_1^H([B(x, 1/2)]) = \alpha_{d-1} V_{d-2}(B^{d-1}(1/2)) =: \widetilde{c}_d,$$

where $B^k(t)$ denotes a ball in $\mathbb{R}^k$ with radius t (centered at 0). The most important thing to note here is that $\widetilde{c}_d$ is some positive constant independent of H and r. Therefore, we infer

$$\mu_1([G_{\oplus 1/2}]) = \int_{A(d,d-1)} \mu_1^H([G_{\oplus 1/2}])\,\mu_{d-1}(dH)$$

$$\geq \int_{\mathcal{H}} \mu_1^H([G_{\oplus 1/2}])\,\mu_{d-1}(dH) \geq \widetilde{c}_d\,\mu_{d-1}(\mathcal{H}). \tag{5.4}$$

Employing [27, Theorem 13.2.12], we conclude that

$$\mu_{d-1}(\mathcal{H}) = \int_{G(d,d-1)} \int_{F^\perp} \mathbf{1}_{\mathcal{H}}(F + z)\lambda_1(dz)\nu_{d-1}(dF)$$

$$= \int_{G(d,d-1)} \mathbf{1}\{\langle s, v_F \rangle \geq \delta\} \int_{F^\perp} \mathbf{1}\{(F + z) \cap S \neq \emptyset\}\lambda_1(dz)\nu_{d-1}(dF),$$

where $G(d, d-1)$ denotes the Grassmannian of hyperplanes in $\mathbb{R}^d$ and ν_{d-1} the invariant measure on $G(d, d-1)$. Moreover, as before v_F denotes the unit normal of F such that $\langle s, v_F \rangle \geq 0$ and λ_1 is the 1-dim. Lebesgue measure on $F^\perp$. Now observe that the inner integral in the last term is bounded from below by $r\delta$. Indeed, we have $(F + z) \cap S \neq \emptyset$ if and only if $\langle y, v_F \rangle \geq \langle z, v_F \rangle \geq 0$. Moreover, if the first indicator is 1, then $\langle y, v_F \rangle = \|y\|\langle s, v_F \rangle \geq r\delta$. Hence $\lambda_1(\{z \in F^\perp : (F + z) \cap S \neq \emptyset\}) \geq \lambda_1(\{z \in F^\perp : r\delta \geq \langle z, v_F \rangle \geq 0\}) = r\delta$ as claimed.

It follows that

$$\mu_{d-1}(\mathcal{H}) \geq r\delta \cdot \nu_{d-1}(\{F \in G(d, d-1) : \langle v_F, s \rangle \geq \delta\}) =: \hat{c}_d\, r, \tag{5.5}$$

where the constant $\hat{c}_d$ is positive and independent of r. Now the claim (5.3) follows (for the constant $\widetilde{C}_d := \widetilde{c}_d \, \hat{c}_d$) by combining (5.4) and (5.5). □

Remark 5.9 In ballistic deposition [5, 9], a cluster is grown in $\mathbb{Z}^d$ by letting particles rain down onto a $d-1$-dimensional initial surface (e.g., $\mathbb{Z}^{d-1} \times \{0\}$ or a subset of it). As in ballistic aggregation, the particles are attached to the existing surface where they first touch it, but in contrast to our ballistic model, the particles are assumed to arrive all from the same direction (vertical from above). In this situation, not the radius but the height $h_t(z)$ of the cluster at time t above a site $z \in \mathbb{Z}^{d-1}$ is the variable of interest, for which a strong law of large numbers was obtained in [21]. It implies in particular that the height grows linearly in t (essentially, $t^{-1}h_t(z)$ converges a.s. to a constant, as $t \to \infty$), see also [22, Thm. 2.3]. This means specifically that the grown clusters in ballistic deposition are full-dimensional, just as conjectured for ballistic aggregation. The analysis of such models in the mentioned references goes far beyond this. In [21] a kind of shape theorem is proved for clusters grown from a quite arbitrary initial surface, and in [22] also fluctuations of the heights are studied, i.e. the roughness of the surfaces. Recently, modifications of these models (in $\mathbb{Z}^2$) with elongated particles of random height and different sticking rules have been considered in [23]. It is an interesting open question, whether some of the methods and ideas can be transferred from deposition to aggregation models.

Note that our ballistic model (in which particles have no preferred directions) can easily be modified in such a way that certain directions are preferred. Particles may even be constrained to arrive from a fixed single direction. We believe that our results on radial growth are not very sensitive to such modifications (which is also backed by observations in simulations [5, 28]), and that our dimension bounds can rather easily be extended to such situations.

6 Proof of Kesten's Method

The first step towards a proof of Theorem 4.1 is an equivalent reformulation of the conclusion in a more convenient form. Let $(F_n)_{n \in \mathbb{N}}$ be some incremental aggregation. Recall the definition of the radius $\mathrm{rad}(F_n)$ from (3.1). We define the two random functions

$$R : \mathbb{N} \to [0, \infty), \quad n \mapsto \mathrm{rad}(F_n)$$

and

$$T : [0, \infty) \to \mathbb{N}, \quad r \mapsto \min\{j \in \mathbb{N} \mid R(j) \geq r\}.$$

Clearly, $T(r)$ is the first time step at which the growing cluster has radius at least r. It is easy to see that almost surely the functions R and T are non-decreasing. Moreover, one has almost surely the relations $T(R(n)) \leq n$ for each $n \in \mathbb{N}$, and $R(T(r)) \geq r$ for each $r \in [0, \infty)$. Note also that almost surely

$$R(n) \to \infty \text{ as } n \to \infty \quad \text{and} \quad T(r) \to \infty \text{ as } r \to \infty.$$

The conclusion in Theorem 4.1 is an almost sure upper bound on the random function R. Our aim is to reformulate this upper bound on R equivalently as a lower bound on the waiting time T. The following statement provides this link. Recall that a function $h : [0, \infty) \to [0, \infty)$ is called multiplicative if and only if $h(xy) = h(x)h(y)$ for all $x, y \geq 0$. Note that multiplicativity implies $h(0) = 0$ and $h(1) = 1$. (The functions $x \mapsto x^p$, $p > 0$ are generic examples for such h.)

Lemma 6.1 *Let $a > 0$ and $h : [0, \infty) \to [0, \infty)$ be a bijective, multiplicative and increasing function. For $c > 0$ define the events*

$$A_c := \{\omega \in \Omega \mid \exists N \in \mathbb{N} : R(\omega)(n) \leq c\, h(n) \text{ for all } n \geq N\}$$

and

$$D_c := \{\omega \in \Omega \mid \exists r_0 \in \mathbb{N} : T(\omega)(ar) \geq c\, h^{-1}(r) \text{ for all } r \geq r_0\}.$$

Then the following assertions are equivalent:

(i) There exists a constant $c > 0$ such that $\mathbb{P}(A_c) = 1$.
(ii) There exists a constant $\bar{c} > 0$ such that $\mathbb{P}(D_{\bar{c}}) = 1$.

Proof Fix some $c > 0$ and let $\omega \in A_c$. Choose $N \in \mathbb{N}$ such that $R(n) \leq c\, h(n)$ holds for all $n \geq N$. (Here and in the sequel N, R and T depend on ω, which we suppress in the notation.) Since h is bijective and multiplicative, the last inequalities are equivalent to $\tilde{c}\, h^{-1}(R(n)) \leq n$ being satisfied for all $n \geq N$, where $\tilde{c} := h^{-1}(1/c)$. Since $T(s) \to \infty$ as $s \to \infty$, we can find $r_0 \in \mathbb{N}$ large enough such that $T(ar_0) > N$. Obviously this implies $T(ar) \geq T(ar_0) > N$ for all $r \geq r_0$, since T is increasing. Hence we infer that $\tilde{c}h^{-1}(R(T(ar))) \leq T(ar)$ holds for each $r \geq r_0$. Observe now that h^{-1} is increasing and multiplicative since h is. Taking into account that $R(T(ar)) \geq ar$, we conclude that for each $r \geq r_0$

$$T(ar) \geq \tilde{c}h^{-1}(ar) = \tilde{c}h^{-1}(a) \cdot h^{-1}(r) =: \bar{c} \cdot h^{-1}(r).$$

Hence $\omega \in D_{\bar{c}}$. It is now important to note that $\bar{c} = h^{-1}(a/c)$ does not depend on ω (but only on h, a and c). Therefore we have proved that $A_c \subseteq D_{\bar{c}}$.

With a completely analogous argument one can show that $D_{\bar{c}} \subseteq A_c$. Hence $A_c = D_{\bar{c}}$ for any $c > 0$ (and $\bar{c} = h^{-1}(a/c)$). In particular, if $\mathbb{P}(A_c) = 1$ for some $c > 0$, then $\mathbb{P}(D_{\bar{c}}) = 1$. Similarly, if $\mathbb{P}(D_{\hat{c}}) = 1$ for some $\hat{c} > 0$, then $c := a/h(\hat{c}) > 0$, $\bar{c} = h^{-1}(a/c) = \hat{c}$ and therefore $\mathbb{P}(A_c = 1) = 1$. This completes the proof. □

Assume now that the hypothesis of Theorem 4.1 is satisfied for some $q > 0$. Consider for any $c > 0$ the events

$$A_c := \left\{\omega \in \Omega \mid \exists N \in \mathbb{N} : R(\omega)(n) \leq c\, n^{\frac{1}{1+q}} \text{ for all } n \geq N\right\}$$

and

$$D_c := \left\{\omega \in \Omega \mid \exists r_0 \in \mathbb{N} : T(\omega)(2r) \geq c r^{1+q} \text{ for all } r \geq r_0\right\}.$$

It is clear that the conclusion of Theorem 4.1 may be formulated as follows in terms of the events A_c: *There exists a constant $c > 0$ such that $\mathbb{P}(A_c) = 1$.*

Note that A_c and D_c are the same as the corresponding events defined in Lemma 6.1 for the choices $a = 2$ and

$$h : [0, \infty) \to [0, \infty), \quad x \mapsto x^{\frac{1}{1+q}}.$$

Clearly, h is bijective, multiplicative and increasing as required. Hence, by Lemma 6.1, the conclusion of Theorem 4.1 holds if and only if there exists a constant $\beta > 0$ such that

$$\mathbb{P}(D_\beta) = 1. \tag{6.1}$$

Therefore, in order to prove Theorem 4.1 it clearly suffices to prove that its hypothesis implies the existence of some constant $\beta > 0$ such that (6.1) holds.

In order to achieve this we introduce some more notation. For $n \in \mathbb{N}$ we write $F_n = \{y_1, \ldots, y_n\}$, where y_j denotes the (random) point added to the cluster at time $j \in \mathbb{N}$. Fix some $\beta > 0$, which will be determined later on. For $r \in \mathbb{N}$ let $\widetilde{m}_r := \beta r^{1+q}$ and define

$$V_r := \{\omega \in \Omega \mid T(\omega)(2r) < \widetilde{m}_r\}.$$

Let $S_r := \{x \in \mathbb{Z}^d \mid r \le |x| < r + 1\}$ be the *discrete sphere* in $\mathbb{Z}^d$ of radius r centered at the origin. (It is easy to see that any connected set $A \in \mathbb{Z}^d$ with $0 \in A$ and $\mathrm{rad}(A) \ge r$ will contain at least one point of S_r, since there will be a path in A connecting 0 to some point at distance at least r which can not avoid intersecting the set S_r.) Further we define the set of all possible (self-avoiding) paths in $\mathbb{Z}^d$ of length r (i.e., passing r edges) with starting point in S_r by

$$Z_r := \big\{\mathbf{z} := (z_0, \ldots, z_r) \in (\mathbb{Z}^d)^r \mid z_0 \in S_r, \{z_i, z_{i+1}\} \in E \text{ for } i = 0, \ldots, r - 1$$
$$\text{and } z_i \ne z_j \text{ for } i \ne j\big\}. \tag{6.2}$$

For any $\mathbf{z} \in Z_r$ we define the event

$$W_r(\mathbf{z}) := \big\{\omega \in \Omega \mid \exists j_0 < \cdots < j_r \le \widetilde{m}_r \text{ such that } y_{j_i}(\omega) = z_i \text{ for } i = 0, \ldots, r\big\}$$

and the union of these events by

$$W_r := \bigcup_{\mathbf{z} \in Z_r} W_r(\mathbf{z}).$$

Note that this is a finite union, as for any $r \in \mathbb{N}$ there are only finitely many paths in Z_r. Note also that W_r as well as V_r depend (via $\widetilde{m}_r$) on the choice of β, which is suppressed in the notation but which will become important later. We claim that for each $r \in \mathbb{N}$

$$V_r \subseteq W_r. \tag{6.3}$$

Indeed, if $\omega \in V_r$, then $T(\omega)(2r) < \widetilde{m}_r$. For $m_r := \max\{j \in \mathbb{N} \mid j \le \widetilde{m}_r\}$ this implies $T(\omega)(2r) \le m_r$, and hence $\mathrm{rad}(F_{m_r}(\omega)) \ge 2r$. Thus $F_{m_r}(\omega)$ contains a point x with $\|x\| \ge 2r$. Since $F_{m_r}(\omega)$ is connected and contains the origin, there exists a path $\mathbf{p} = (y_{i_1}, \ldots, y_{i_\ell})$ in $F_{m_r}(\omega)$ connecting 0 and x (i.e., $y_{i_0} = 0$ and $y_{i_\ell} = x$) with the additional property that $i_0 < i_1 \ldots < i_\ell$, that is, the points of $\mathbf{p}$ have been added to the cluster in the right order. (Not every path between 0 and x might satisfy this, but such a path must exist, since otherwise the cluster would be disconnected at some time.) $\mathbf{p}$ contains a point z_0 of S_r and thus a subpath $\mathbf{p}'$ connecting z_0 to x. Since $\|z_0 - x\| > r - 1$, $\mathbf{p}'$ has length at least r. Hence, by taking the first r steps of $\mathbf{p}'$, we find some path $\mathbf{z}$ of length exactly r, that is, some $\mathbf{z} \in Z_r$ in $F_{m_r}(\omega)$. This means $\omega \in W_n(\mathbf{z}) \subset W_n$ which proves the set inclusion (6.3).

The following estimate will allow to complete the proof of (6.1).

Lemma 6.2 *There is some $\beta > 0$ and constants $c_1 > 0$, $q \in (0, 1)$ and $r_0 \in \mathbb{N}$ (which depend on β) such that for all $r \in \mathbb{N}$ with $r \ge r_0$*

$$\mathbb{P}(W_r) \le c_1 r^{d-1} q^r.$$

Before we prove Lemma 6.2 let us discuss how it is used to complete the proof of (6.1). Let us fix $\beta > 0$ as required for applying the lemma. Then, clearly, since $q < 1$, we have

$$\sum_{r=1}^{\infty} \mathbb{P}(W_r) \le r_0 + \sum_{r \ge r_0} \mathbb{P}(W_r) \le r_0 + \sum_{r \ge r_0} c_1 r^{d-1} q^r < \infty$$

and thus, by (6.3),

$$\sum_{r=1}^{\infty} \mathbb{P}(V_r) < \infty.$$

Hence, by the Borel–Cantelli lemma, $\mathbb{P}(\limsup_{r\to\infty} V_r) = 0$. Since

$$(\limsup_{r\to\infty} V_r)^c = \{\omega \in \Omega \mid \exists r_0 \in \mathbb{N} \text{ s.t. } \omega \in V_r^c \text{ for all } r > r_0\}$$

$$= \{\omega \in \Omega \mid \exists r_0 \in \mathbb{N} \text{ s.t. } T(\omega)(2r) \geq \beta r^{1+q} \text{ for all } r \geq r_0\} = D_\beta$$

we conclude that $\mathbb{P}(D_\beta) = 1$. This completes the proof of (6.1) and thus of Kesten's method (up to a proof of Lemma 6.2).

In order to prepare the proof of Lemma 6.2, we introduce some more notation and provide a few auxiliary results. For any $r \in \mathbb{N}$, any path $\mathbf{z} = (z_0, \dots, z_r) \in Z_r$ (cf. (6.2)) and any $i \in \{0, \dots, r\}$, we define random variables

$$\tau_i^{\mathbf{z}} : \Omega \to \mathbb{N}_\infty, \quad \tau_i^{\mathbf{z}}(\omega) = \begin{cases} j, & \text{if } j \in \mathbb{N} \text{ and } y_j(\omega) = z_i, \\ \infty, & \text{if } z_i \notin F_\infty(\omega), \end{cases}$$

that is, $\tau_i^{\mathbf{z}}$ is either the time point j at which z_i is added to the cluster or ∞ if this never happens. Further, for $i \in \{1, \dots, r\}$, we define waiting times

$$\sigma_i^{\mathbf{z}} : \Omega \to \mathbb{N}_\infty, \quad \omega \mapsto \begin{cases} \tau_i^{\mathbf{z}}(\omega) - \tau_{i-1}^{\mathbf{z}}(\omega), & \text{if } \tau_{i-1}^{\mathbf{z}}(\omega) < \tau_i^{\mathbf{z}}(\omega) < \infty, \\ \infty, & \text{else,} \end{cases}$$

so $\sigma_i^{\mathbf{z}}$ is the waiting time between adding z_{i-1} and z_i to the cluster, if both are added and z_{i-1} is added before z_i. We also consider the event

$$U_{\mathbf{z}} := \{\tau_0^{\mathbf{z}} < \cdots < \tau_r^{\mathbf{z}} < \infty\}$$

that all the vertices of the path $\mathbf{z}$ are added to the cluster, and in the right order (i.e., z_ℓ before $z_{\ell+1}$). Note that $\mathbb{P}(U_{\mathbf{z}}) > 0$. We will now prove that the random variable $\sigma_i^{\mathbf{z}}$ conditioned on $U_{\mathbf{z}}$ always dominates a geometrically distributed random variable with parameter

$$p_r := Cr^{-q}, \tag{6.4}$$

where the constants C and q are the ones from the hypothesis in Theorem 4.1. In particular, they are the same for all $i \in \{0, \dots, r-1\}$ and independent of $r \in \mathbb{N}$.

Lemma 6.3 *For any $r \in \mathbb{N}$ and $i = 1, \dots, r$, we have*

$$\mathbb{P}(\sigma_i^{\mathbf{z}} > k | U_{\mathbf{z}}) \geq \mathbb{P}(Y > k)$$

for any $k \in \mathbb{N}$, where Y is geometrically distributed with parameter p_r as in (6.4).

Proof Fix $i \in \{1, \dots, r\}$. Let $U_{\mathbf{z}}^{i,j} := U_{\mathbf{z}} \cap \{\tau_{i-1}^{\mathbf{z}} = j\}$, $j \in \mathbb{N}$ and let $k \in \mathbb{N}$. It suffices to show that for any j such that $\mathbb{P}(U_{\mathbf{z}}^{i,j}) > 0$

$$\mathbb{P}(\sigma_i^{\mathbf{z}} > k | U_{\mathbf{z}}^{i,j}) \geq \mathbb{P}(Y > k). \tag{6.5}$$

Indeed, since $U_{\mathbf{z}} = \bigcup_{j \in \mathbb{N}} U_{\mathbf{z}}^{i,j}$ is a disjoint decomposition, we infer that

$$\mathbb{P}(\sigma_i^{\mathbf{z}} > k | U_{\mathbf{z}}) = \sum_{j \in \mathbb{N}} \mathbb{P}(\{\sigma_i^{\mathbf{z}} > k\} \cap \{\tau_i^{\mathbf{z}} = j\} | U_{\mathbf{z}})$$

$$= \sum_{j \in \mathbb{N}} \frac{\mathbb{P}(\{\sigma_i^{\mathbf{z}} > k\} \cap U_{\mathbf{z}}^{i,j})}{\mathbb{P}(U_{\mathbf{z}})} = \sum_{j : \mathbb{P}(U_{\mathbf{z}}^{i,j}) > 0} \frac{\mathbb{P}(\{\sigma_i^{\mathbf{z}} > k\} \cap U_{\mathbf{z}}^{i,j})}{\mathbb{P}(U_{\mathbf{z}})}$$

$$= \sum_{j : \mathbb{P}(U_{\mathbf{z}}^{i,j}) > 0} \mathbb{P}(\sigma_i^{\mathbf{z}} > k | U_{\mathbf{z}}^{i,j}) \frac{\mathbb{P}(U_{\mathbf{z}}^{i,j})}{\mathbb{P}(U_{\mathbf{z}})}.$$

Now, by (6.5), the first probability in each summand can be bounded from below by $\mathbb{P}(Y > k)$ and taking this factor out of the summation, the remaining summands sum to 1. Hence, we obtain $\mathbb{P}(\sigma_i^{\mathbf{z}} > k | U_{\mathbf{z}}) \geq \mathbb{P}(Y > k)$ as asserted in the lemma.

For a proof of (6.5), fix $j \in \mathbb{N}$ and define for any $\ell \in \mathbb{N}$ the event $A_\ell := \{\omega \in \Omega : y_{j+\ell}(\omega) \neq z_i\}$. Set $A_0 := U_{\mathbf{z}}^{i,j}$. Then, by the multiplication formula of conditional probabilities, we have for any $k \in \mathbb{N}$

$$\mathbb{P}(\sigma_i^{\mathbf{z}} > k | U_{\mathbf{z}}^{i,j}) = \mathbb{P}\left(\bigcap_{\ell=1}^{k} A_\ell | A_0\right) = \prod_{m=1}^{k} \mathbb{P}\left(A_m | \bigcap_{\ell=0}^{m-1} A_\ell\right).$$

Now observe that the condition in the m-th factor implies, that at time $j + m$ the cluster F_{j+m} has radius at least r and contains z_{i-1} but not z_i. (And thus, due to the choice of $\mathbf{z}$, z_i is in the outer boundary of F_{j+m}.) Hence, by the hypothesis (4.1) in Kesten's method (Theorem 4.1), we get

$$\mathbb{P}\left(A_m | \bigcap_{\ell=0}^{m-1} A_\ell\right) \geq 1 - p_r.$$

We conclude that $\mathbb{P}(\sigma_i^{\mathbf{z}} > k | U_{\mathbf{z}}^{i,j}) \geq (1 - p_r)^k = \mathbb{P}(Y > k)$ holds for any $k \in \mathbb{N}$, proving (6.5). □

Since we want to sum geometric random variables in a moment, we also provide the following standard bound, which can be found e.g. in [15, Lemma 2.6.2].

Lemma 6.4 *Let $n \in \mathbb{N}$, $n \geq 2$ and let $Y_1, \ldots, Y_n$ be independent geometrically distributed random variables with parameter $0 < p < \frac{1}{2}$. Then for every $a \geq 2p$,*

$$\mathbb{P}\left(\sum_{i=1}^{n} Y_i \leq a \frac{n}{p}\right) \leq (2ae^2)^n.$$

Note that the stated bound has power only if $2ae^2 < 1$, which is possible only if $p < (4e^2)^{-1}$. Based on this estimate we obtain the following upper bound for the waiting time $\tau_r^{\mathbf{z}}$ until the path $\mathbf{z} \in Z_r$ is completely added to the cluster conditioned on the event that it is added.

Lemma 6.5 *For any $\beta > 0$ there is some $r_0 \in \mathbb{N}$ such that for all $r \in \mathbb{N}$ with $r \geq r_0$ and all $\mathbf{z} \in Z_r$*

$$\mathbb{P}(W_r(\mathbf{z})) \leq (2\beta Ce^2)^r,$$

where C is the constant in the hypothesis of Theorem 4.1.

Proof Fix some $\beta > 0$ and set $a := \beta C$. Choose $r_0 \in \mathbb{N}$ large enough such that $p_{r_0} < \min\{\frac{1}{2}, \frac{a}{2}\}$. Note that this implies $a = \beta C \geq 2p_r$ for all $r \geq r_0$.

Observe that $W_r(\mathbf{z}) \subset \{\tau_r^{\mathbf{z}} \leq \tilde{m}_r\} \cap U_{\mathbf{z}}$ and thus

$$\mathbb{P}(W_r(\mathbf{z})) \leq \mathbb{P}(W_r(\mathbf{z})|U_{\mathbf{z}}) \leq \mathbb{P}(\tau_r^{\mathbf{z}} \leq \tilde{m}_r|U_{\mathbf{z}}).$$

Note that under the condition $U_{\mathbf{z}}$, $\tau_r^{\mathbf{z}}$ may be represented as $\tau_r^{\mathbf{z}} = \tau_0^{\mathbf{z}} + \sum_{i=1}^{r} \sigma_i^{\mathbf{z}}$. Hence, for any $t > 0$, we obtain

$$\mathbb{P}(\tau_r^{\mathbf{z}} \leq t|U_{\mathbf{z}}) \leq \mathbb{P}(\tau_r^{\mathbf{z}} - \tau_0^{\mathbf{z}} \leq t|U_{\mathbf{z}}) = \mathbb{P}\left(\sum_{i=1}^{r} \sigma_i^{\mathbf{z}} \leq t|U_{\mathbf{z}}\right).$$

Now recall from Lemma 6.3 that, conditioned on $U_{\mathbf{z}}$, each of the random variables $\sigma_i^{\mathbf{z}}$ dominates a geometric random variable Y with parameter p_r. Moreover, given $\tau_{i-1}^{\mathbf{z}}$, the randomness in the variable $\sigma_i^{\mathbf{z}}$ will only depend on the subsequent steps of the construction. Hence their sum dominates a sum $\sum_{i=1}^{r} Y_i$ of independent geometric random variables Y_i with parameter p_r. That is, for any $t > 0$

$$\mathbb{P}(\tau_r^{\mathbf{z}} \leq t|U_{\mathbf{z}}) \leq \mathbb{P}\left(\sum_{i=1}^{r} Y_i \leq t\right).$$

Specializing now to $t = \tilde{m}_r$ and recalling from the definition of $\tilde{m}_r$ and (6.4) that $\tilde{m}_r = \beta r^{1+q} = \beta C \frac{r}{p_r} = a\frac{r}{p_r}$, we can apply Lemma 6.4 with $n = r$, $p = p_r$ and $a = \beta C$. (Note that the hypothesis $a \geq 2p_r$ is satisfied for $r \geq r_0$.)

We obtain that for any $r \in \mathbb{N}$ with $r \geq r_0$,

$$\mathbb{P}(W_r(\mathbf{z})) \leq \mathbb{P}\left(\tau_r^{\mathbf{z}} \leq \tilde{m}_r|U_{\mathbf{z}}\right) \leq \mathbb{P}\left(\tau_r^{\mathbf{z}} \leq \frac{ar}{p_r}|U_{\mathbf{z}}\right) \leq (2\beta Ce^2)^r,$$

which completes the proof. $\square$

Now we have all the ingredients to provide a proof of Lemma 6.2.

Proof of Lemma 6.2 Choose $\beta > 0$ small enough that $4d\beta Ce^2 < 1$. By definition of W_r and Lemma 6.5, there exists $r_0 \in \mathbb{N}$ such that, for any $r \geq r_0$,

$$\mathbb{P}(W_r) \leq \sum_{\mathbf{z} \in Z_r} \mathbb{P}(W_r(\mathbf{z})) \leq \#Z_r \cdot (2\beta Ce^2)^r.$$

Recall that the paths of Z_r start in S_r and observe that there is some constant c_1 such that $\#S_r \leq c_1 r^{d-1}$. Hence there are at most $c_1 r^{d-1}$ starting points for paths in Z_r. Moreover, since each point in $\mathbb{Z}^d$ has $2d$ neighbors and paths in Z_r have length r, there are at most $(2d)^r$ possible paths with a given starting point. Hence $\#Z_r \leq c_1 r^{d-1}(2d)^r$. We infer that for any $r \geq r_0$

$$\mathbb{P}(W_r) \leq c_1 r^{d-1}(2d)^r \cdot (2\beta Ce^2)^r = c_1 r^{d-1} q^r,$$

where $q := 4d\beta Ce^2 < 1$, due to the choice of β, and $c_1 > 0$. Hence we have found some $\beta > 0$ (and suitable constants c_1, q and r_0) such that the asserted bound in Lemma 6.2 holds for all $r \in \mathbb{N}$ with $r \geq r_0$. $\square$

Data Availability The computer code and datasets generated during the current study are solely for the purpose of illustration and are available in the GitHub repository, see [24].

Declarations

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Witten, T.A., Sander, L.M.: Diffusion-limited aggregation, a kinetic critical phenomenon. Phys. Rev. Lett. **47**, 1400–1403 (1981). https://doi.org/10.1103/PhysRevLett.47.1400
2. Eden, M.: A two-dimensional growth process. In: Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, vol. IV, pp. 223–239. University of California Press, Berkeley (1961)
3. Meakin, P., Deutch, J.M.: The formation of surfaces by diffusion limited annihilation. J. Chem. Phys. **85**(4), 2320–2325 (1986). https://doi.org/10.1063/1.451129
4. Meakin, P.: The Vold-Sutherland and Eden models of cluster formation. J. Colloid Interface Sci. **96**(2), 415–424 (1983). https://doi.org/10.1016/0021-9797(83)90044-9
5. Vicsek, T.: Fractal Growth Phenomena. World Scientific, Teaneck (1989). https://doi.org/10.1142/0511
6. Stanley, H.E., Ostrowsky, N. (eds.): On Growth and Form. NATO Advanced Science Institutes Series E: Applied Sciences, vol. 100. Martinus Nijhoff Publishers, Dordrecht (1986)
7. Jullien, R., Botet, R.: Aggregation and Fractal Aggregates. World Scientific, Teaneck (1987)
8. Kolb, M.: Structure formation by aggregation: models and applications. In: Kramer, B. (ed.) Advances in Solid State Physics, pp. 381–389. Springer, Berlin (2001). https://doi.org/10.1007/3-540-44946-9_31
9. Barabási, A.-L., Stanley, H.E.: Fractal Concepts in Surface Growth. Cambridge University Press, Cambridge (1995). https://doi.org/10.1017/CBO9780511599798
10. Hastings, M.B., Levitov, L.S.: Laplacian growth as one-dimensional turbulence. Physica D **116**(1), 244–252 (1998). https://doi.org/10.1016/S0167-2789(97)00244-3
11. Silvestri, V.: Fluctuation results for Hastings-Levitov planar growth. Probab. Theory Related Fields **167**(1–2), 417–460 (2017). https://doi.org/10.1007/s00440-015-0688-7
12. Berger, N., Procaccia, E.B., Turner, A.: Growth of stationary Hastings-Levitov. Ann. Appl. Probab. **32**(5), 3331–3360 (2022). https://doi.org/10.1214/21-AAP1761
13. Kesten, H.: How long are the arms in DLA? J. Phys. A **20**(1), 29–33 (1987). https://doi.org/10.1088/0305-4470/20/1/007
14. Kesten, H.: Upper bounds for the growth rate of DLA. Physica A **168**(1), 529–535 (1990). https://doi.org/10.1016/0378-4371(90)90405-H
15. Lawler, G.F.: Intersections of Random Walks. Probability and Its Applications. Birkhäuser, Boston (1991)
16. Benjamini, I., Yadin, A.: Upper bounds on the growth rate of diffusion limited aggregation (2017). arXiv:1705.06095
17. Lawler, G.F., Bramson, M., Griffeath, D.: Internal diffusion limited aggregation. Ann. Probab. **20**(4), 2117–2140 (1992)

18. Auffinger, A., Damron, M., Hanson, J.: 50 Years of First-Passage Percolation. University Lecture Series, vol. 68. American Mathematical Society, Providence (2017). https://doi.org/10.1090/ulect/068

19. Sava-Huss, E.: From fractals in external DLA to internal DLA on fractals. In: Freiberg, U., Hambly, B., Hinz, M., Winter, S. (eds.) Fractal Geometry and Stochastics VI. Progress in Probability, vol. 76, pp. 273–298. Birkhäuser/Springer, Cham (2021). https://doi.org/10.1007/978-3-030-59649-1_12

20. Molchanov, I.S.: Diffusion-limited aggregation with jumps and flights. J. Stat. Comput. Simul. **64**(4), 357–381 (1999). https://doi.org/10.1080/00949659908811985

21. Seppäläinen, T.: Strong law of large numbers for the interface in ballistic deposition. Ann. l'Inst. Henri Poincare B Probab. Stat. **36**(6), 691–736 (2000). https://doi.org/10.1016/S0246-0203(00)00137-0

22. Penrose, M.D.: Growth and roughness of the interface for ballistic deposition. J. Stat. Phys. **131**(2), 247–268 (2008). https://doi.org/10.1007/s10955-008-9507-1

23. Comets, F., Dalmau, J., Saglietti, S.: Scaling limit of the heavy tailed ballistic deposition model with p-sticking. Ann. Probab. **51**(5), 1870–1931 (2023). https://doi.org/10.1214/23-aop1635

24. Bosch, T.: Incremental-aggregations. https://github.com/tillmanntristanbosch/incremental-aggregations (2021). Accessed 23rd August 2023

25. Kesten, H.: Aspects of first passage percolation. In: Hennequin, P.L. (ed.) École d'Été de Probabilités de Saint Flour XIV - 1984, pp. 125–264. Springer, Berlin (1986)

26. Deuschel, J.-D., Pisztora, A.: Surface order large deviations for high-density percolation. Probab. Theory Relat. Fields **104**(4), 467–482 (1996). https://doi.org/10.1007/BF01198162

27. Schneider, R., Weil, W.: Stochastic and Integral Geometry. Probability and Its Applications (New York). Springer, Berlin (2008). https://doi.org/10.1007/978-3-540-78859-1

28. Bensimon, D., Shraiman, B., Liang, S.: On the ballistic model of aggregation. Phys. Lett. A **102**(5), 238–240 (1984). https://doi.org/10.1016/0375-9601(84)90701-1

29. Ball, R.C., Witten, T.A.: Causality bound on the density of aggregates. Phys. Rev. A **29**, 2966–2967 (1984). https://doi.org/10.1103/PhysRevA.29.2966