Université du Littoral Côte d'Opale

École Doctorale en Sciences Technologie et Santé (ED 585)

LMPA

Thesis defended by **Keenan Penner**

Defended on **July 7, 2025**

In order to become Doctor from Université du Littoral Côte d'Opale

Academic Field **Mathematics**

Speciality **Mathematics and their interactions**

IDLA aggregates and forests with sources in a hyperplane of $\mathbb{Z}^d$

Thesis supervised by	Nicolas Chenavier	Supervisor
	David Coupier	Co-Supervisor
	Arnaud Rousselle	Co-Monitor

Committee members

Referees	Quentin Berger	Professor at Université Sorbonne Paris Nord	
	Ecaterina Sava-Huss	Professor at Universität Innsbruck	
Examiners	Alexandre Gaudillière	Researcher at Institut de Mathématiques de Marseille	
	Régine Marchand	HDR Associate Professor at Institut Élie Cartan de Lorraine	
	Dominique Schneider	Professor at Université du Littoral Côte d'Opale	Committee President
Supervisors	Nicolas Chenavier	Professor at Université du Littoral Côte d'Opale	
	David Coupier	Professor at Institut Mines Télécom Nord Europe	
	Arnaud Rousselle	Associate Professor at Université Bourgogne Europe	

Colophon

Doctoral dissertation entitled "IDLA aggregates and forests with sources in a hyperplane of $\mathbb{Z}^d$", written by Keenan Penner, completed on July 3, 2025, typeset with the document preparation system LaTeX and the yathesis class dedicated to theses prepared in France.

Université du Littoral Côte d'Opale

École Doctorale en Sciences Technologie et Santé (ED 585)

LMPA

Thèse présentée par **Keenan Penner**

Soutenue le **7 juillet 2025**

En vue de l'obtention du grade de docteur de l'Université du Littoral Côte d'Opale

Discipline **Mathématiques**

Spécialité **Mathématiques et leurs interactions**

Agrégats et forêts IDLA avec sources dans un hyperplan de $\mathbb{Z}^d$

Thèse dirigée par	Nicolas Chenavier	Directeur
	David Coupier	Co-Directeur
	Arnaud Rousselle	Co-Encandrant

Composition du jury

Rapporteurs	Quentin Berger	Professeur à l'Université Sorbonne Paris Nord	
	Ecaterina Sava-Huss	Professeure à l'Universität Innsbruck	
Examinateurs	Alexandre Gaudillière	Chargé de Recherche à l'Institut de Mathématiques de Marseille	
	Régine Marchand	MCF HDR à l'Institut Élie Cartan de Lorraine	
	Dominique Schneider	Professeur à l'Université du Littoral Côte d'Opale	Président du Jury
Directeurs de thèse	Nicolas Chenavier	Professeur à l'Université du Littoral Côte d'Opale	
	David Coupier	Professeur à l'Institut Mines Télécom Nord Europe	
	Arnaud Rousselle	MCF à l'Université Bourgogne Europe	

IDLA aggregates and forests with sources in a hyperplane of $\mathbb{Z}^d$

Abstract

Internal Diffusion Limited Aggregation (IDLA) is a process used to construct random aggregates in $\mathbb{Z}^d$. At the start, it was introduced to model certain industrial chemical techniques such as electropolishing. Since then, many variants of this model have been studied, and have led to results describing the overall shape of the aggregate. Such results are referred to as *shape theorems*. This thesis handles several IDLA models in $\mathbb{Z}^d$, with $d \geq 2$. We study various families of random aggregates, whose construction is based on an IDLA protocol with an infinite number of sources. We establish stabilization results for these aggregates as well as shape theorems.
We begin by studying a multi-source IDLA model built by sending a deterministic number of particles from infinitely many sources. We establish properties of stationarity for this model, along with a global upper bound, which roughly describes the shape of the aggregate far away from the origin. This result allows us to then show a stabilization result, which turns out to be crucial in the second part of our study.
The second part is devoted to the study of a multi-source model similar to the previous one, built this time by sending a random number of particles from infinitely many sources, in a random order. We study this model with the goal of building a translation invariant forest, called the directed infinite-volume IDLA forest. Its existence is non-trivial due to issues of consistancy, and forces us to adapt results obtained for the previous model to this one, requiring even to sharpen some of these results. In this second part, we prove a stabilization result for forests, whose proof is based on arguments of percolation, allowing us to prove the existence of the directed infinite-volume IDLA forest.

Keywords: IDLA, random walks, random forests, random graphs, growth models, shape theorems, stabilization, percolation

Agrégats et forêts IDLA avec sources dans un hyperplan de $\mathbb{Z}^d$

Résumé

L'Agrégation Limitée par Diffusion Interne (IDLA) est un processus permettant de construire des agrégats aléatoires dans $\mathbb{Z}^d$. Il a initialement été introduit pour modéliser des problèmes de chimie industrielle tels que l'électro-polissage. Depuis, de nombreuses variantes de ce modèle ont été étudiées, et ont donné lieu à des résultats permettant de décrire la forme des agrégats obtenus. De tels résultats sont qualifiés dans la littérature de *shape theorems*. Cette thèse traite de plusieurs modèles d'agrégation limitée par diffusion interne sur $\mathbb{Z}^d$, avec $d \geq 2$. Nous étudions différentes familles d'agrégats aléatoires, dont la construction se base sur un protocole IDLA comportant une infinité de sources. Nous établissons des résultats de stabilisation pour ces agrégats, ainsi que des *shape theorems*.
Dans un premier temps, nous étudions un modèle d'IDLA multi-sources construit en envoyant un nombre déterministe de particules depuis un nombre infini de sources. Nous établissons des propriétés de stationnarité pour ce modèle, ainsi qu'une borne globale permettant de contrôler de façon grossière l'agrégat loin de l'origine. Ce résultat nous permet ensuite de montrer un résultat de stabilisation, qui se révèle être crucial pour la seconde partie de notre étude.
La seconde partie est devouée à l'étude d'un modèle multi-sources similaire au premier, mais construit cette fois-ci en envoyant un nombre aléatoire de particules depuis un nombre infini de sources, selon un ordre aléatoire. Nous étudions ce modèle dans le but de construire une forêt invariante par translation, appelée forêt IDLA dirigée de volume infini. L'existence de cette forêt est non-triviale à cause de problèmes de consistance, et nous amène à adapter et à affiner les résultats du précédent modèle pour le second. Nous démontrons dans cette seconde partie un résultat de stabilisation pour des forêts, dont la preuve est basée sur un argument de percolation, nous permettant de démontrer l'existence de la forêt IDLA dirigée de volume infini.

Mots clés : IDLA, marches aléatoires, forêts aléatoires, graphes aléatoires, modèles de croissance, *shape theorems*, stabilisation, percolation

LMPA
Maison de la Recherche Blaise Pascal – 50, rue Ferdinand Buisson – CS 80699 – 62228 Calais Cedex – France

Remerciements

Je tiens tout d'abord à remercier mes directeurs de thèse, Nicolas Chenavier, David Coupier et Arnaud Rousselle pour leur aide précieuse tout au long de ces trois dernières années. Merci de m'avoir guidé du début à la fin, merci pour le temps et l'énergie que vous avez consacré à ce travail, et merci pour votre bienveillance et votre gentillesse. J'ai passé d'excellents moments à travailler à vos côtés, toujours dans une ambiance chaleureuse et conviviale, et je vous suis très reconnaissant pour tout l'enrichissement tant scientifique qu'humain que vous m'avez apporté.

Merci aux rapporteurs Ecaterina Sava-Huss et Quentin Berger pour avoir pris le temps de lire en détail ma thèse, et pour leurs remarques pertinentes. Thank you to the reviewers Ecaterina Sava-Huss and Quentin Berger for taking the time to thoroughly read my thesis and for their insightful comments.

Merci également à Alexandre Gaudillière, Régine Marchand et Dominique Schneider qui me font l'honneur d'être membres de ce jury.

Merci à tous mes collègues du LMPA pour la bonne ambiance que vous avez mise au sein du laboratoire. J'ai partagé de très bons moments à vos côtés, notamment pendant les pauses déjeuner, à échanger sur tout et n'importe quoi, souvent avec humour! J'aimerais remercier Isabelle pour son travail et sa grande disponibilité. Merci à toute l'équipe de co-voiturage Lille - Calais pour les bons moments passés en train et en voiture!

Merci également à mes amis doctorants du laboratoire: Pierre, Julie, Vincent, Martin, Rawaa et Manon. J'aimerais commencer par remercier Pierre, mon tout premier co-bureau, qui m'a guidé dans les méandres administratifs des débuts de thèse, et qui m'a donné de précieux conseils tout au long de ma première année. Tu as toujours gardé une bonne humeur (sauf peut-être quand il s'agissait de battre Ponchik aux échecs) et tu as grandement facilité mon insertion au labo, merci!

On cite souvent l'autonomie, la gestion du temps ou encore la communication comme les principales qualités développées pendant un doctorat. Mais au sein du LMPA, j'ai eu l'opportunité d'en développer bien plus. C'est pour cela que j'aimerais remercier Julie, qui m'a fait apprendre ce qu'était l'humilité, en me confrontant à son TD de Stats de L3 sur lequel je ne maîtrisais visiblement plus la notion de test statistique. Merci à Vincent d'avoir développé mon état d'esprit entrepreneurial avec ses divers projets post-thèse. Enfin, merci à Martin et Rawaa de m'avoir aidé à développer mon esprit d'investigation en menant à mes côtés l'enquête ayant pour but de découvrir l'identité secrète de Wen Chao Lu. Malgré un grand travail d'investigation, le mystère reste entier. Martin, je te confie la lourde tâche d'y venir à bout... Je pourrais citer bien plus de moments ludiques à vos côtés, mais je pense que cette liste se suffit à elle-même pour décrire la bonne ambiance qui a régné dans le bureau des doctorants et au sein du LMPA de façon générale.

J'aimerais désormais remercier celles et ceux que j'ai rencontré au cours de ces trois dernières années aux côtés de l'équipe Probas Stats du Laboratoire Paul Painlevé. Même si je n'étais pas membre officiel de l'équipe, j'ai été très bien accueilli dès mon arrivée, et je garderai de très bons souvenirs des moments passés avec toute l'équipe ainsi qu'avec l'équipe du GT. Un grand merci à vous tous! Merci à Barnabé, David D., Laurence, et Raphaël d'avoir souvent animé les discussions du midi! Merci en particulier au duo Christopher–Benoît, puis Benoît–Rafaël, avec qui j'ai beaucoup rigolé, et ce souvent accompagné d'un Americano. Christopher, tu m'as permis de viser la lune, en me confiant les clés du cluster universitaire sur lequel j'ai pu faire tourner une simulation à 1 milliard de particules. Visiblement, la fac ne partageait pas notre esprit visionnaire, et nous a contraint à interrompre la simulation.

Merci aussi à Benoît pour tous les moments de rigolade au cours de ces trois dernières années, et surtout la dernière année dans le bureau 326. C'était un véritable plaisir de partager ce bureau avec toi pendant ma dernière année. J'ai beaucoup ri et beaucoup appris à tes côtés. Les pauses café, accompagnées (trop souvent!) d'un Snickers sont devenues un rituel sacré avec toi, je compte sur Rafaël pour faire perdurer la tradition. Enfin, Rafaël, merci d'avoir toujours porté un immense intérêt pour l'IDLA, en particulier la propriété abélienne, qui, je le sais bien, te tient à cœur. Plus sérieusement, merci pour toutes les discussions que nous avons pu avoir ensemble, je te souhaite plein de bonnes choses pour la poursuite de ta thèse.

Je remercie également les enseignants en mathématiques que j'ai eus au cours de ma scolarité et de mes études supérieures. Je pense notamment à mon professeur de Terminale Serge Seguin. Merci de m'avoir donné goût aux mathématiques!

J'aimerais remercier quelques amis rencontrés au cours de mes études supérieures: Tristan, Laure, Ulysse, Camille, Alix, Nadhir, Yacine et Philippe. C'est toujours un plaisir de vous revoir et de passer du temps à vos côtés! Un immense merci à mes amis du Lycée International, notamment la team *Species*! Les longs appels passés avec vous chaque soir m'ont fait un bien fou, je vous aime fort! Merci à Arnaud pour les raclées sur GeoGuessr, merci à Dylan pour le moment Bing Bong à Londres, merci à Quentin pour le moment 'Bro bro bro' au Camelot, merci à Nico K. pour les moments Soundboard, merci à Nico R. pour les 'DOWN' mémorables, et merci à Charles d'être Charles. Merci en particulier d'avoir été mon quasi-voisin à la Cité U et d'avoir partagé la routine Métro-Boulot-Chakalito avec moi. Merci aussi pour tes précieux conseils en Python! Merci à Côme pour tout ton soutien, merci de m'avoir toujours proposé de faire de nouvelles activités, que ce soit des concerts ou des baskets. Merci à Isabel pour ta gentillesse et ton écoute. Merci à Aurélien, Chloé et Chris pour les verres occasionnels à Paris! Merci à tous mes collocs de la Rue Colbert, j'ai passé d'excellents moments à vos côtés. Merci d'avoir supporté mes moments de folie où j'expliquais ma thèse de façon incompréhensible. Je vous souhaite plein de bonnes choses pour la suite!

Enfin, j'aimerais finir par remercier ma famille. Merci Mamie, Pey et Bob pour tout l'amour que vous m'avez donné. Merci à mes parents de m'avoir soutenu tout le long de mes études. Vous avez toujours été là pour moi, et j'en suis grandement reconnaissant. Thank you Dad for the peaceful weekends spent at home, for the delicious food you'd have me take back to Lille, and for all your generosity. Merci Maman pour ta joie de vivre, ton humour, et toute ta gentillesse. Merci pour tous les bons moments passés avec toi, que je garde précieusement en mémoire et dans mon cœur. Merci Julia-Marie pour tout ce que tu as fait pour moi depuis le début. Merci de partager le même humour que moi et de m'accompagner dans mes délires. Je n'aurais pas pu tomber sur une meilleure grande sœur. I'd like to finish off by thanking some of my favorite Kansans: thank you Grandma and Grandpa for the kindness you've always shown to me. Thank you Dave, Jill, Deb, Jeff, Adam, Nate, Desiree, Braiden and Virg for all the great times spent together in the summer. Special shoutout to Dave, the linear algebra expert, who I'm sure will enjoy every page of this manuscript.

À ma mère.

Table of contents

Acronyms and symbols

IDLA Internal Diffusion Limited Aggregation

uIDLA Uniform Internal Diffusion Limited Aggregation

GFF Gaussian Free Field

RST Radial Spanning Tree

DSF Directed Spanning Forest

PPP Poisson Point Process

$[\![a, b]\!]$ interval of all integers between a and b ($[\![a, b]\!] := [a, b] \cap \mathbb{Z}$)

$\| \cdot \|_2$ euclidian norm

$\| \cdot \|$ uniform norm

$\mathbb{B}(r)$ ball of radius r in $\mathbb{R}^d$

ω_d volume of d-dimensional euclidean unity ball

$\mathcal{H}$ hyperplane of sources ($\mathcal{H} = \{0\} \times \mathbb{Z}^{d-1}$)

$\mathcal{H}_M$ hyperplane of sources of level at most M ($\mathcal{H}_M = \{0\} \times [\![-M, M]\!]^{d-1}$)

$\mathbb{Z}_M$ strip of size M ($\mathbb{Z}_M = \mathbb{Z} \times [\![-M, M]\!]^{d-1}$)

$\mathcal{R}_k$ slab of width $2k$ ($\mathcal{R}_k = [\![-k, k]\!] \times \mathbb{Z}^{d-1}$)

$\oplus$ smash sum operator

$p_{\mathcal{H}}$ projection operator on $\mathcal{H}$

T_k translation operator of vector k

δ_I inner error

δ_O outer error

$\mathcal{T}_n$ IDLA tree of size n

$\mathcal{T}_\infty$ IDLA tree

$\mathcal{F}_n[M]$ IDLA forest up to time n and level M

$\mathcal{F}_n$ infinite-volume directed IDLA forest up to time n

$\mathcal{F}_\infty$ infinite-volume directed IDLA forest

Summary (in French)

Outline of the current chapter

Résultats connus sur l'IDLA et modèles reliés

Motivations

L'agrégation limitée par diffusion interne (IDLA) est un modèle de croissance aléatoire sur $\mathbb{Z}^d$. Ce modèle fournit un protocole permettant de construire une suite d'agrégats de façon récursive, comme suit. À l'étape 0, l'agrégat est supposé vide. Ensuite, à chaque étape, une particule est lancée depuis l'origine, puis tuée dès qu'elle atteint un site en dehors de l'agrégat en cours. La particule tuée se pose alors sur ce site et l'occupe pour toujours. On redémarre alors le protocole en lançant une nouvelle particule depuis l'origine.

L'IDLA a été introduit pour la première fois par Meakin et Deutch en 1986 (voir [39]) afin de modéliser une technique de chimie industrielle appelée *électropolissage*. Contrairement à l'*électroplaquage*, dont le but est de recouvrir une surface métallique d'une fine couche de matériau, le but de l'électropolissage est de polir la surface d'un métal en lui *ôtant* une fine couche de matériau. Ce procédé a lieu grâce à une réaction d'oxydoréduction. Le métal que l'on souhaite polir est alors baigné dans une solution ionique, et est relié à une source d'alimentation électrique. Le métal joue alors le rôle de l'anode, chargée positivement. Une réaction d'oxydoréduction a lieu à la surface du métal, provoquant une dissolution des anions dans la solution ionique. La surface du métal est alors marquée par un grand nombre de trous, formant un agrégat de sites vides. Il devient alors particulièrement intéressant de pouvoir quantifier la régularité de la surface obtenue par électropolissage. C'est d'ailleurs l'une des motivations premières des problèmes en IDLA : on cherche à décrire le plus précisément possible la forme finale de l'agrégat obtenu, ainsi que les éventuelles fluctuations autour de cette forme finale. De tels résultats sont énoncés dans la littérature sous le nom de *shape theorems*.

Construction et propriétés

Diaconis et Fulton sont les premiers à introduire le modèle IDLA dans un contexte mathématique en 1991 dans [19]. Les auteurs commencent par définir une opération, appelée *smash sum*, entre deux ensembles $A,\ B \subset \mathbb{Z}^d$. Cette *smash sum* est notée $A \oplus B$, et fournit un ensemble aléatoire dont le cardinal est égal à $|A| + |B|$. Tout d'abord, si $A \cap B = \emptyset$, alors on définit simplement $A \oplus B = A \cup B$. Supposons désormais que $A \cap B \neq \emptyset$. On écrit alors $A \cap B = \{z_1, \dots,\ z_k\},\ k \geq 1$. Similairement au protocole IDLA décrit au-dessus, le protocole définissant la *smash sum* dans ce cas consiste à lancer des marches aléatoires indépendantes depuis chaque site de $A \cap B$ et d'ajouter à $A \cup B$ le premier site visité par la marche en dehors de $A \cup B$. Plus formellement, la *smash sum* se définit comme suit : On définit $C_0 = A \cup B$. Puis, on définit C_1 comme

$$C_1 := C_0 \cup \{S_{z_1}(\tau_0)\},$$

où S_{z_1} désigne une marche aléatoire (simple, symétrique) émise depuis z_1, et τ_0 est le temps d'arrêt :

$$\tau_0 := \inf\{t \geq 0,\ S(t) \notin C_0\}.$$

De façon générale, pour $1 \leq i \leq k$, on définit C_i :

$$C_i := C_{i-1} \cup \{S_{z_i}(\tau_{i-1})\},$$

où $\tau_{i-1} := \inf\{t \geq 0,\ S(t) \notin C_{i-1}\}$, et S_{z_i} est une marche aléatoire indépendante des précédentes. Enfin, nous posons $A \oplus B := C_k$. En somme, la *smash sum* de A et B est l'agrégat $A \cup B$ auquel on ajoute $|A \cap B|$ points supplémentaires, obtenus en lançant une marche aléatoire à partir de chaque site de $A \cap B$, puis en ajoutant le premier site visité en dehors de l'agrégat actuel.

Propriété abélienne

Il est raisonnable de penser que la loi de l'agrégat aléatoire obtenu à l'issue de cette *smash sum* dépend de l'ordre dans lequel nous avons indexé les éléments de $A \cap B$. Étonnement, ce n'est pas le cas, et cette propriété est appelée *propriété abélienne*. Elle est énoncée en détail ci-dessous.

Proposition 1 : *(Propriété abélienne) Soit $A \subset \mathbb{Z}^d$ borné, et soient $z_1, \dots, z_k \in \mathbb{Z}^d$. La distribution de*

$$((A \oplus \{z_1\}) \oplus \{z_2\}) \oplus \cdots \oplus \{z_k\}$$

ne dépend pas de l'ordre des z_i. C'est-à-dire, pour toute permutation $\sigma \in \mathfrak{S}_k$ de $\{1, \dots, k\}$, on a :

$$((A \oplus \{z_1\}) \oplus \{z_2\}) \oplus \cdots \oplus \{z_k\} \overset{\text{loi}}{=} ((A \oplus \{z_{\sigma(1)}\}) \oplus \{z_{\sigma(2)}\}) \oplus \cdots \oplus \{z_{\sigma(k)}\}.$$

Il est en fait possible de montrer un résultat plus fort que celui énoncé ci-dessus, et montrer qu'il est possible de mettre en pause la trajectoire d'une marche donnée, de lancer d'autres marches entre-temps, et de reprendre la trajectoire de la marche stoppée, sans jamais changer la loi de $A \oplus B$. On précise que le terme de *particule* est employé pour désigner une marche jusqu'au temps de sortie de l'agrégat.

Cette propriété est centrale dans de nombreux résultats présentés dans cette thèse. Elle nous permet de considérer des protocoles dans lesquels des particules sont arrêtées en cours de route, puis redémarrées, sans jamais changer la loi de l'agrégat final. La propriété abélienne se révèle être fondamentale dans de nombreuses preuves concernant l'IDLA, car la plupart des stratégies utilisées dans les protocoles IDLA reposent, d'une manière ou d'une autre, sur cette propriété.

Preuve de la Proposition 1 : Soient $x,\ y \in \mathbb{Z}^d$ et $A \subset \mathbb{Z}^d$. Soient $S^x,\ S^y$ deux marches aléatoires indépendantes simples, symétriques, partant respectivement de x et de y. On définit les temps d'arrêt suivants :

$$T^x := \inf\{n \geq 0,\ S_n^x \notin A\}, \qquad T^y := \inf\{n \geq 0,\ S_n^y \notin A\}.$$

Soient $(M_k^x)_{k\geq 0}$ et $(M_k^y)_{k\geq 0}$ les incréments de S^x et S^y respectivement *après* que ces marches sortent de A. Ceux-ci sont bien définis puisque $T^x < \infty$ et $T^y < \infty$ presque sûrement (on rappelle que $A \subset \mathbb{Z}^d$ est borné). Les incréments M^x et M^y sont tels que pour tout $n \geq 0$:

$$S_n^x = S_{n\wedge T^x}^x + \sum_{k=1}^{n-T^x} M_k^x \mathbb{1}_{n>T^x},$$

$$S_n^y = S_{n\wedge T^y}^y + \sum_{k=1}^{n-T^y} M_k^y \mathbb{1}_{n>T^y}.$$

Désormais, nous définissons les marches aléatoires $\tilde{S}^x$ et $\tilde{S}^y$ comme suit :

$$\tilde{S}_n^x = S_{n\wedge T^x}^x + \sum_{k=1}^{n-T^x} M_k^y \mathbb{1}_{n>T^x},$$

$$\tilde{S}_n^y = S_{n\wedge T^y}^y + \sum_{k=1}^{n-T^y} M_k^x \mathbb{1}_{n>T^y}.$$

La marche $\tilde{S}^x$ suit la même trajectoire que S^x jusqu'à ce qu'elle quitte A, puis suit une trajectoire définie par les incréments de S^y après que celle-ci quitte A. De même, $\tilde{S}^y$ suit la même trajectoire que S^y jusqu'à ce qu'elle quitte A, puis suit une trajectoire définie par les incréments de S^x après que celle-ci quitte A. Maintenant, appelons E l'agrégat aléatoire obtenu (par *smash sum* de A et $\{x\}$ puis $\{y\}$) après avoir lancé la particule d'abord depuis x en utilisant S^x, puis la particule depuis y en utilisant S^y. De manière similaire, nous appelons $\tilde{E}$ l'agrégat obtenu (par *smash sum* de A et $\{y\}$ puis $\{x\}$) après avoir lancé la particule depuis y en utilisant $\tilde{S}^y$ en premier, puis depuis x en utilisant $\tilde{S}^x$.

On peut montrer que, grâce à notre couplage, on a $E \overset{\text{a.s}}{=} \tilde{E}$.

- Si $S^x(T^x) \neq S^y(T^y)$, alors S^x et S^y quittent A par des sites différents, et il en va de même pour $\tilde{S}^x$ et $\tilde{S}^y$. Dans ce cas, il est clair que lancer la particule d'abord depuis x puis depuis y est identique à lancer d'abord depuis y puis depuis x.
- Supposons maintenant que $S^x(T^x) = S^y(T^y)$. Une fois encore, par construction, il en va de même pour $\tilde{S}^x$ et $\tilde{S}^y$. Dans les deux cas, un site aléatoire z_1 est ajouté à l'agrégat.
  - Pour E, on utilise la trajectoire de S^y, qui ajoute le site aléatoire z_2 à l'agrégat.

– Pour $\tilde{E}$, on utilise la trajectoire de $\tilde{S}^x$ pour déterminer le site aléatoire à ajouter. Cependant, avec notre couplage, la trajectoire de $\tilde{S}^x$ après avoir quitté A est la même que la trajectoire de S^y après avoir quitté A. Ainsi, le même site z_2 est ajouté à l'agrégat.

Dans les deux cas, nous avons bien que $E = \tilde{E} = A \cup \{z_1\} \cup \{z_2\}$. Maintenant, comme $\tilde{S}^x$ et $\tilde{S}^y$ sont des marches aléatoires simples symétriques indépendantes issues respectivement de x et y, nous obtenons :

$$(A \oplus \{x\}) \oplus \{y\} \overset{\text{loi}}{=} (A \oplus \{y\}) \oplus \{x\}.$$

□

Notons que la propriété abélienne ne tient qu'*en loi*, et non *presque sûrement*. Il est même *faux* de dire que les agrégats obtenus en permutant l'ordre d'émission sont identiques presque sûrement. Nous fournissons un exemple simple dans $\mathbb{Z}^2$ pour illustrer ce point. Soit $x = (0, 1)$, $y = (0, -1)$ et soit $A = \{x,\ y\}$. Soient S^x et S^y deux marches aléatoires indépendantes, simples, symétriques sur $\mathbb{Z}^2$. Pour cet exemple, supposons que les trajectoires de S^x et S^y soient telles que $S^x = (\downarrow, \rightarrow, \dots)$, et $S^y = (\uparrow, \uparrow, \leftarrow, \dots)$.

Calcul de $(A \oplus \{x\}) \oplus \{y\}$ **:** Commençons par calculer $A \oplus \{x\}$. Puisque $A \cap \{x\} = \{x\}$, on lance une particule depuis x en utilisant la trajectoire de S^x. Le premier site que cette particule visite en dehors de $A \cup \{x\}$ est $(0, 0)$, qui est alors ajouté à l'agrégat. On obtient alors que $A \oplus \{x\} = \{x,\ y,\ (0, 0)\}$.

Maintenant, calculons $(A \oplus \{x\}) \oplus \{y\}$. Puisque $(A \oplus \{x\}) \cap \{y\} = \{y\}$, on lance une particule depuis y en utilisant la trajectoire de S^y. Cette particule commence par visiter $(0, 0)$ (qui est déjà dans l'agrégat), puis visite x (qui est également dans l'agrégat), et visite enfin $(-1, 1)$, qui est ajouté à l'agrégat actuel. On obtient que $(A \oplus \{x\}) \oplus \{y\} = \{x,\ y,\ (0, 0),\ (-1, 1)\}$.

Calcul de $(A \oplus \{y\}) \oplus \{x\}$ **:** On commence par calculer $A \oplus \{y\}$. Puisque $A \cap \{y\} = \{y\}$, on lance une particule depuis y en utilisant la trajectoire de S^y. Le premier site que cette particule visite en dehors de $A \cup \{y\}$ est $(0, 0)$, qui est alors ajouté à l'agrégat. On obtient que $A \oplus \{y\} = \{x,\ y,\ (0, 0)\}$.

Maintenant, calculons $(A \oplus \{y\}) \oplus \{x\}$. Puisque $(A \oplus \{y\}) \cap \{x\} = \{x\}$, on lance une particule depuis x en utilisant la trajectoire de S^x. Cette particule commence par visiter $(0, 0)$ (qui est déjà dans l'agrégat), puis visite $(1, 0)$, qui est ajouté à l'agrégat actuel. On obtient $(A \oplus \{y\}) \oplus \{x\} = \{x,\ y,\ (0, 0),\ (1, 0)\}$.

Au final, nous obtenons que $(A \oplus \{x\}) \oplus \{y\} \neq (A \oplus \{y\}) \oplus \{x\}$. Les figures 1 et 2 illustrent l'exemple précédent. Les agrégats sont représentés par des points bleus, tandis que la marche aléatoire est représentée par un petit point rouge.

Définition du modèle IDLA

Pour $n \geq 0$, nous définissons l'agrégat IDLA $A(n)$ comme la répétition de n *smash sum* entre $\emptyset$ et $\{0\}$:

$$A(n) = \underbrace{((\emptyset \oplus \{0\}) \oplus \{0\}) \oplus \dots \oplus \{0\}}_{n \text{ fois}}.$$

Autrement dit, posons $A(0) = \emptyset$, et pour tout $n \geq 1$, étant donné une réalisation de $A(n-1)$, nous définissons $A(n)$ comme

$$A(n) = A(n-1) \cup \{S(\tau_{n-1})\},$$

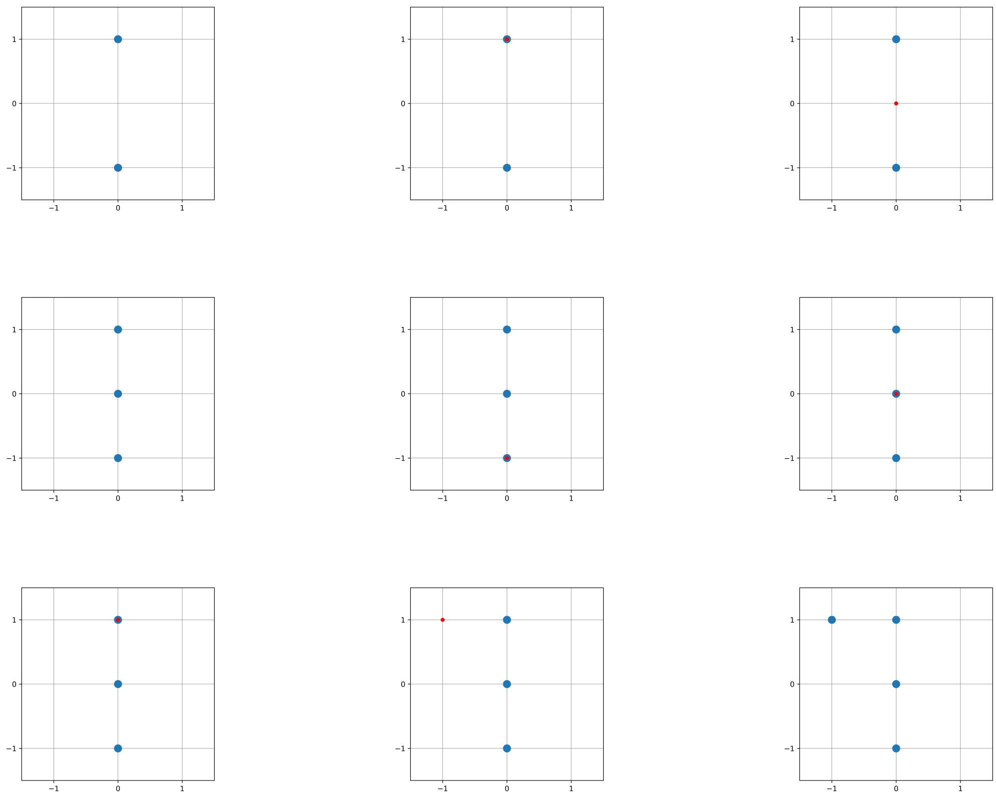


Figure 1: Une réalisation de $(A \oplus \{x\}) \oplus \{y\}$

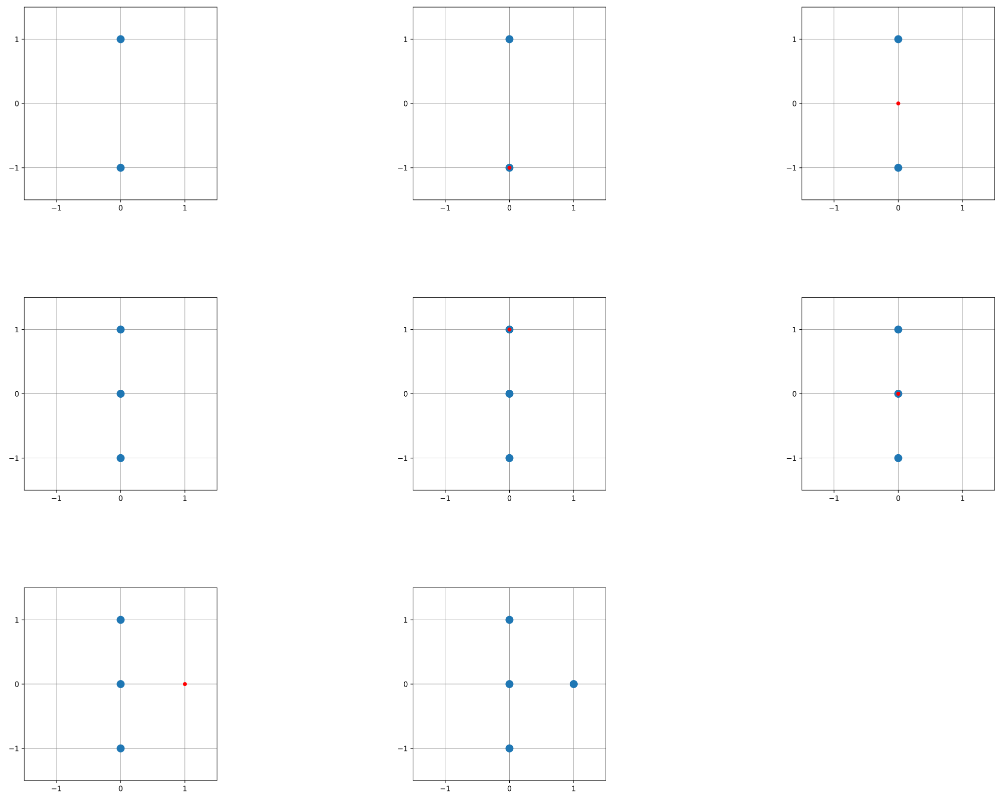
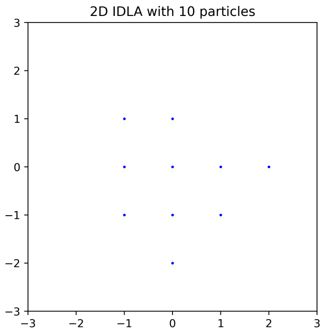


Figure 2: Une réalisation de $(A \oplus \{y\}) \oplus \{x\}$

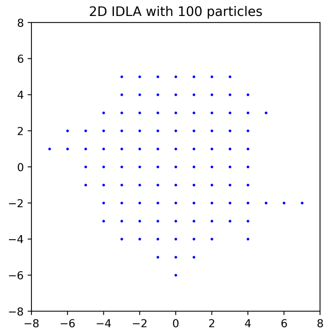
où S désigne une marche aléatoire simple et symétrique sur $\mathbb{Z}^d$, partant de l'origine, indépendante de $A(n-1)$, et où τ_{n-1} est le temps d'arrêt

$$\tau_{n-1} := \inf \{t \geq 0,\ S(t) \notin A(n-1)\}\,.$$

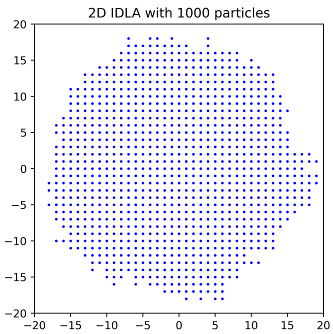
À chaque étape, une nouvelle marche aléatoire, indépendante des précédentes, est alors envoyée depuis l'origine et arrêtée dès qu'elle quitte l'agrégat. Nous donnons quelques réalisations de $A(n)$ avec des valeurs croissantes de n ($n \in \{10, 100, 1000, 10000\}$) à la Figure 3.



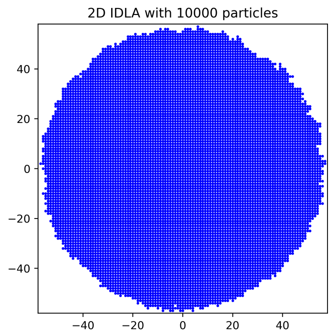
(a) IDLA avec 10 particules



(b) IDLA avec 100 particules



(c) IDLA avec 1000 particules



(d) IDLA avec 10000 particules

Figure 3: Une réalisation du modèle IDLA sur $\mathbb{Z}^2$, avec 10, 100, 1000 et 10000 particules

On pourrait conjecturer à partir de ces simulations qu'en faisant augmenter le nombre n de particules émises, la forme de l'agrégat $A(n)$ se rapproche d'une boule euclidienne. Ceci est en fait vrai, et fait l'objet du premier *shape theorem* concernant le modèle IDLA.

Shape Theorems

Lawler, Bramson et Griffeath sont les premiers à démontrer dans [34] un *shape theorem* en dimension $d \geq 2$ pour de l'IDLA classique. Ils montrent que, convenablement renormalisé, la forme limite de l'agrégat est celle d'une boule euclidienne. Leur résultat est énoncé comme suit :

Théorème 2 : *Soit ω_d le volume de la boule euclidienne de dimension d et de rayon 1 dans $\mathbb{R}^d$.*

Pour tout $\varepsilon > 0$, presque sûrement, pour n suffisamment grand,

$$\mathbb{B}(n(1-\varepsilon)) \subseteq A(\lfloor \omega_d n^d \rfloor) \subseteq \mathbb{B}(n(1+\varepsilon)), \tag{1}$$

où $\mathbb{B}(r) := \{x \in \mathbb{Z}^d,\ \|x\|_2 < r\}$ est la boule euclidienne de rayon r.

Dans toute la suite de ce résumé, afin d'alléger la notation, on écrira systématiquement $\omega_d n^d$ au lieu de $\lfloor \omega_d n^d \rfloor$.

Commentons tout d'abord ce résultat. Au premier abord, il pourrait paraître surprenant d'obtenir comme forme limite une boule *Euclidienne*, étant donné que l'espace sur lequel nous travaillons est l'espace discret $\mathbb{Z}^d$. Cependant, ce résultat est moral, car nous savons qu'une marche aléatoire sur $\mathbb{Z}^d$ convenablement renormalisée converge en loi vers un mouvement Brownien dans $\mathbb{R}^d$. Étant donné l'isotropie du mouvement Brownien, il est cohérent que la forme limite soit celle d'une boule Euclidienne.

La démonstration de (1) utilise des ingrédients classiques que l'on retrouve dans de nombreuses autres démonstrations de *shape theorems* pour divers modèles d'IDLA. En effet, la plupart des preuves des *shape theorems* suivent un schéma similaire : on commence par montrer la borne inférieure en utilisant la propriété abélienne, puis on démontre la borne supérieure en exploitant la borne inférieure. Dans [34], la stratégie pour prouver la borne supérieure consiste à dire ceci : étant donné qu'une proportion significative de particules a déjà été utilisée pour remplir la borne inférieure, il reste alors trop peu de particules pour déborder de la borne supérieure. Il reste alors à contrôler le taux auquel la frontière de l'agrégat se développe, ce que les auteurs montrent par l'étude d'un processus de branchement.

Revenons à la stratégie employée pour montrer la borne inférieure. Si une particule lancée depuis l'origine atteint $\mathbb{B}(n(1-\varepsilon))$ avant de sortir de l'agrégat, on met sa trajectoire en pause, puis on continue à lancer des particules depuis l'origine. Cela donne un agrégat intermédiaire, qui est inclus dans $\mathbb{B}(n(1-\varepsilon))$ par construction. Une fois toutes les particules lancées, on reprend la trajectoire de celles qui avaient été arrêtées à $\mathbb{B}(n(1-\varepsilon))$ jusqu'à ce qu'elles quittent l'agrégat actuel. Grâce à la propriété abélienne, on sait que cet agrégat final a la même loi que $A(\omega_d n^d)$. Les auteurs obtiennent des résultats sur cet agrégat intermédiaire, à partir desquels ils déduisent des propriétés sur $A(\omega_d n^d)$. En particulier, ils utilisent une décomposition astucieuse pour compter les particules qui atteignent $\mathbb{B}(n(1-\varepsilon))$, que nous détaillons brièvement.

Considérons un site $z \in \mathbb{B}(n(1-\varepsilon))$. On définit les variables aléatoires M, N et L comme suit :

N = nombre de *particules* qui visitent z avant de quitter l'agrégat,
M = nombre de *marches* qui visitent z avant de quitter $\mathbb{B}(n)$,
L = nombre de *marches* qui visitent z avant de quitter $\mathbb{B}(n)$, mais après que la particule associée se soit posée.

On a alors l'inégalité suivante :

$$M \leq N + L. \tag{2}$$

En effet, considérons une marche comptée par la variable aléatoire M. Soit sa particule associée est passée par z avant de quitter l'agrégat, auquel cas elle a été comptée par N. Soit la marche est passée par z avant de quitter $\mathbb{B}(n)$ mais après que la particule associée se soit posée. Dans ce cas, elle a été comptée par L. Il s'agit bien d'une inégalité dans (2), et non d'une égalité, car la variable N peut contenir des particules qui ont visité z après avoir quitté $\mathbb{B}(n)$, et ces particules ne seraient pas comptabilisées par M. En écrivant $N \geq M - L$, il devient alors possible d'obtenir un contrôle sur N, qui compte des *particules*, en utilisant les variables M et L, qui comptent toutes les deux des *marches*. Notons que M est totalement indépendant de l'agrégat, il est donc facile à contrôler à l'aide de résultats bien connus sur les marches aléatoires. Le contrôle de la variable L nécessite un peu plus de travail. Tout d'abord, notons que L ne compte que les marches qui atteignent z *après* que la particule associée se soit arrêtée et *avant* de quitter $\mathbb{B}(n)$. On peut écrire L comme une somme d'indicatrices de cet événement, où la somme est prise sur les $\omega_d n^d$ marches. Cependant, puisque L compte des marches qui passent par z uniquement *après* que la particule associée se soit arrêtée, nous pouvons commencer à compter ces marches en fonction du site où chaque particule quitte l'agrégat, ce qui correspond à chaque fois à un site unique de $\mathbb{B}(n)$. Ainsi, au lieu de sommer sur le nombre de marches émises, nous pouvons élargir l'ensemble de sommation en sommant sur tous les sites $y \in \mathbb{B}(n)$ et en considérant des marches émises en y qui visitent z avant de quitter $\mathbb{B}(n)$. Cela nous fournit une variable aléatoire $\tilde{L}$ telle que $L \leq \tilde{L}$ *en loi*. De plus, les marches comptées par $\tilde{L}$ ont l'avantage d'être indépendantes, ce qui signifie que $\tilde{L}$ peut s'écrire comme la somme de variables aléatoires indépendantes, et la rend facile à contrôler. Cette manière astucieuse de compter les particules de N en utilisant M et L est présente dans de nombreuses autres démonstrations de *shape theorems*. Grâce à ce raisonnement, nous pouvons contrôler de façon efficace les particules, qui par nature sont très dépendantes les unes des autres et donc difficiles à étudier, en travaillant avec des sommes de variables indépendantes, qui sont des objets bien plus faciles à étudier.

Une fois la forme limite de l'agrégat prouvée, il est ensuite devenu pertinent d'étudier les fluctuations de l'agrégat autour de sa forme limite. Pour ce faire, on définit l'erreur interne $\delta_I(n)$ et l'erreur externe $\delta_O(n)$ par :

$$n - \delta_I(n) = \sup\left\{r \geq 0,\ \mathbb{B}(r) \subset A(\omega_d n^d)\right\}, \quad n + \delta_0(n) = \inf\left\{r \geq 0,\ A(\omega_d n^d) \subset \mathbb{B}(r)\right\}.$$

Une première borne sur ces deux termes d'erreur est due à Lawler dans [33]. Il obtient des fluctuations d'ordre sous-diffusives, dans le sens où

$$\delta_I(n) = \mathcal{O}\left(n^{1/3}\log^2 n\right),$$

$$\delta_0(n) = \mathcal{O}\left(n^{1/3}\log^4 n\right).$$

Ces bornes ont ensuite été améliorées par Asselah et Gaudillière dans [1] pour obtenir des bornes d'ordre logarithmique. Ils montrent alors que :

$$\delta_I(n) = \mathcal{O}\left(\log n\right),$$

$$\delta_0(n) = \mathcal{O}\left(\log^2 n\right).$$

Peu après, ils améliorent leur borne pour l'erreur externe (voir [4]), tout en améliorant les deux bornes pour des dimensions $d \geq 3$, en obtenant des fluctuations d'ordre sous-logarithmique. Parallèlement, Jerison, Levine et Sheffield obtiennent dans [31] le même résultat en utilisant une approche différente. Ils montrent qu'en dimension $d = 2$,

$$\delta_I(n) = \mathcal{O}\left(\log n\right),$$

$$\delta_O(n) = \mathcal{O}\left(\log n\right).$$

Dès que $d \geq 3$, les fluctuations sont sous-logarithmiques, dans le sens où :

$$\delta_I(n) = \mathcal{O}\left(\sqrt{\log n}\right),$$

$$\delta_O(n) = \mathcal{O}\left(\sqrt{\log n}\right).$$

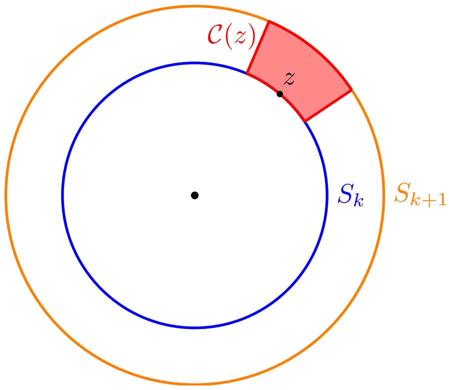


Figure 4: Une représentation des coquilles et des cellules. Les coquilles sont désignées par S_k et S_{k+1}. La cellule de z, hachurée en rouge, est désignée par $\mathcal{C}(z)$. L'exploration par vagues affirme que si un grand nombre de particules sont accumulées à la surface d'une cellule, c'est-à-dire sur $\mathcal{C}(z) \cap S_k$, alors $\mathcal{C}(z)$ sera remplie avec grande probabilité

La différence de fluctuations entre le cas de la dimension 2 et les dimensions supérieures est une conséquence de la récurrence des marches aléatoires lorsque $d = 2$ et de leur caractère transitoire dès que $d \geq 3$. Les estimées concernant la fonction de Green sont donc différentes dans chaque cas, ce qui entraîne des bornes plus précises dans le cas $d \geq 3$.

Les méthodes utilisées pour prouver ces fluctuations sont très similaires à celles que nous avons détaillées concernant la démonstration du premier *shape theorem* de Lawler, Bramson et Griffeath dans [34]. Nous détaillons ici rapidement l'approche d'Asselah et Gaudillière car nous utilisons des méthodes similaires pour établir un *shape theorem* pour notre modèle multi-sources. Dans [4], Asselah et Gaudillière utilisent une méthode d'«exploration par vagues», reposant sur la propriété abélienne. L'espace extérieur à $\mathbb{B}(n)$ est partitionné en coquilles, chaque coquille étant elle-même partitionnée en cellules. Lorsque des particules atteignent la frontière d'une coquille, elles sont arrêtées, et une nouvelle particule est alors lancée depuis l'origine. Cela provoque une accumulation temporaire de particules à la surface des coquilles. Une fois toutes les particules envoyées, les particules mises en pause reprennent leurs trajectoires. Les auteurs montrent alors que si un grand nombre de particules se trouvent accumulées à la surface d'une coquille, alors elles ont de grandes chances de remplir la cellule dans la coquille suivante. Cette «exploration par vagues» peut se voir comme l'analogue d'un barrage retenant de l'eau. Les particules s'accumulant à la frontière de la coquille jouent le rôle de l'eau stockée par le barrage, et une fois qu'il y a suffisamment de particules, on ouvre les vannes du barrage, libérant les particules, qui vont alors remplir la cellule correspondante. Cette stratégie repose une fois de plus sur la propriété abélienne, qui nous donne la possibilité de mettre des particules en pause en cours de route sans jamais changer la loi de l'agrégat final. Une illustration des coquilles et des cellules est donnée par la Figure 4.

Asselah et Gaudillière utilisent également une décomposition similaire à celle de (2) pour contrôler le nombre de particules qui atteignent une coquille. Ils exploitent une propriété d'indépendance négligée par Lawler, Bramson et Griffeath pour obtenir des estimations plus fines de N. À ce jour, ces bornes sont les plus précises connues pour le modèle d'IDLA classique. Asselah et Gaudillière montrent dans [2] que ces bornes sont optimales lorsque $d \geq 3$. En revanche, il n'est pas encore clair que ces bornes soient optimales lorsque $d = 2$.

Modèles similaires

Dans cette section, nous donnons un bref aperçu de variantes et de modèles liés au modèle classique d'IDLA. Nous présentons quelques-uns de ces modèles ainsi que les résultats connus à leur sujet. Nous donnons les grandes lignes de chaque démonstration en restant volontairement vague, car l'idée principale de cette section est avant tout de donner au lecteur une idée globale des ingrédients habituels utilisés pour les démonstrations en IDLA.

IDLA cylindrique

Dans ce paragraphe, nous énonçons plusieurs résultats d'IDLA sur des graphes cylindriques. Pour $N \geq 0$, on définit le graphe cylindrique de $\mathbb{Z}^2$ comme $\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}$. Nous adoptons la notation de [29, 37] et écrivons $Z_N := \mathbb{Z}/N\mathbb{Z}$. On peut alors considérer un protocole d'IDLA sur les graphes cylindriques, où les particules sont lancées uniformément depuis $Z_N \times \{0\}$. On a vu que la forme limite de l'agrégat classique IDLA était une boule euclidienne avec des fluctuations sous-logarithmiques. Dans le cas de l'agrégat IDLA sur $Z_N \times \mathbb{Z}$, il a été démontré que la forme limite est celle d'un rectangle, avec des fluctuations logarithmiques. En effet, Levine et Silvestri montrent dans [37] que pour tout $\gamma > 0,\ m \in \mathbb{N}$, il existe une constante $b_{\gamma,m}$ telle que

$$\mathbb{P}\left(R_{\frac{t}{N}-b_{\gamma,m}\log N} \subset A(t) \subset R_{\frac{t}{N}+b_{\gamma,m}\log N},\ \forall t \leq N^m\right) \geq 1 - N^{-\gamma},$$

où $R_k := \{(x, y) : y \leq k\}$, et $A(t)$ désigne l'agrégat IDLA cylindrique obtenu en lançant t particules uniformément depuis $Z_N \times \{0\}$. Ce résultat vient améliorer celui de [29], dans lequel un résultat similaire avait été montré mais pour des temps $t = N^2$. Ici, Levine et Silvestri étendent ce résultat à des temps longs $t \leq N^m$. Dans le même article, les auteurs s'intéressent également à la question suivante : combien de temps faut-il pour que l'agrégat atteigne une forme rectangulaire, étant donné une configuration initiale quelconque? En d'autres termes, lorsque le processus IDLA ne commence pas à plat, combien de temps lui faut-il pour oublier sa forme initiale ? La réponse à cette question dépend évidemment de la configuration initiale, et les auteurs étudient alors des cas où la position de départ est une «configuration initiale typique», c'est-à-dire une configuration attendue, ou encore, une configuration raisonnablement atteignable pour le modèle d'IDLA cylindrique. Par un contrôle de la hauteur de l'agrégat et un couplage astucieux, ils montrent qu'avec grande probabilité, le processus oublie son état initial en $\mathcal{O}(N^2 \log N)$ étapes. La propriété abélienne est encore une fois centrale dans la démonstration de ce résultat, car elle permet aux auteurs de créer leur *water coupling*, un couplage leur permettant de construire deux agrégats égaux en loi. Ce couplage nécessite que des particules soient lancées, puis gelées avant d'être lancées à nouveau, ce qui, comme nous l'avons vu précédemment, est un raisonnement classique en IDLA.

IDLA sur l'amas de percolation surcritique

Dans ce paragraphe, nous présentons des résultats concernant l'IDLA sur le cluster infini de percolation surcritique dans $\mathbb{Z}^d$. On commence par présenter le modèle de percolation par arêtes sur $\mathbb{Z}^d$. Pour $p \in [0, 1]$, chaque arête de $\mathbb{Z}^d$ est dite *ouverte* (conservée dans le graphe) avec probabilité p et *fermée* (retirée du graphe) avec probabilité $1 - p$, indépendamment des autres arêtes. Lorsque $p = 0$, le graphe obtenu est trivialement réduit à $\{0\}$, tandis que le graphe obtenu lorsque $p = 1$ est $\mathbb{Z}^d$ tout entier. Cependant, pour $p \in\ ]0, 1[$, la configuration de percolation obtenue est nettement moins claire. L'une des interrogations principales sur ces modèles de percolation est l'existence d'un cluster infini. Pour $x \in \mathbb{Z}^d$, on définit le cluster de x comme la réunion de tous les sommets accessibles depuis x en n'utilisant que des arêtes ouvertes du graphe. On note ce cluster par $C(x) := \left\{y \in \mathbb{Z}^d,\ x \leftrightarrow y\right\}$. L'existence d'un cluster infini peut donc s'écrire comme l'événement $\{|C(x)| = \infty\}$, ou encore $\{x \leftrightarrow \infty\}$. Par stationnarité de $\mathbb{Z}^d$, cet événement ne dépend pas de x, et on peut alors définir la probabilité de percolation par

$$\theta_d(p) := \mathbb{P}_p\left(|C(0)| = \infty\right),$$

où $\mathbb{P}_p$ désigne la mesure de percolation de paramètre p. De plus, on définit le paramètre critique de percolation $p_c(d)$ par :

$$p_c(d) := \inf\left\{p \in [0,1],\ \theta_d(p) > 0\right\}.$$

On sait que $p_c(1) = 1$ et $p_c(2) = \frac{1}{2}$ (résultat dû à Kesten). Cependant, en dimension supérieure, les valeurs exactes de $p_c(d)$ ne sont pas encore connues. Une autre question intéressante est la valeur de $\theta_d(p_c(d))$. On sait que $\theta_1(1) = 1$, tandis que $\theta_2\left(\frac{1}{2}\right) = 0$. Dès que $d \geq 3$, le comportement au paramètre critique n'est pas connu.

L'IDLA sur les graphes de percolation surcritique est étudié dans [20]. Les auteurs établissent un *shape theorem* pour les graphes de percolation sur $\mathbb{Z}^d$, dans le cas surcritique $p > p_c(d)$. Plus précisément, soit ω le cluster infini de percolation surcritique sur $\mathbb{Z}^d$ conditionné à contenir l'origine, et soit $b(n) := |\omega \cap \mathbb{B}(n)|$. Notons $\tilde{A}(b(n))$ l'agrégat IDLA sur ω obtenu en lançant $b(n)$ particules depuis l'origine. Alors, pour tout $\varepsilon > 0$ et presque sûrement pour tout ω, pour n suffisamment grand, on a :

$$\mathbb{B}(n(1-\varepsilon)) \cap \omega \subset \tilde{A}(b(n)) \subset \mathbb{B}(n(1+\varepsilon)) \cap \omega.$$

Une fois de plus, la stratégie pour prouver ce *shape theorem* suit le même schéma que précédemment : les auteurs commencent par prouver la borne inférieure puis la borne supérieure, car cette dernière utilise le résultat de la première. Les auteurs fournissent en fait une méthode générale dans laquelle l'existence d'une borne inférieure vérifiant de «bonnes» conditions de régularité fournit automatiquement une technique pour prouver une borne supérieure. Ils prouvent ce résultat pour des graphes généraux sous ces conditions de régularité, et montrent en particulier que le graphe de percolation surcritique sur $\mathbb{Z}^d$ vérifie ces conditions.

Modèles de tas de sable

Dans ce paragraphe, nous définissons un modèle similaire à l'IDLA connu sous le nom de modèle de tas de sable. Ce modèle est souvent lié à l'IDLA car son protocole exhibe lui aussi une propriété abélienne. Le protocole est le suivant : considérons le réseau $\mathbb{Z}^d$, sur lequel chaque site contient un certain nombre entier de grains de sable, formant un tas de sable. Lorsqu'il y a plus de $2d$ grains sur un même site, ce tas de sable s'effondre, donnant à chacun de ses $2d$ voisins exactement un grain de sable. Cette étape est alors répétée tant qu'il reste des sites avec des tas de sable contenant plus de $2d$ grains. Le modèle classique de tas de sable consiste à avoir n, $n \gg 2d$ grains empilés à l'origine et à les laisser s'effondrer. Dans ce cas, si nous appelons S_n l'ensemble de tous les sites visités pendant le processus, Levine et Peres montrent dans [36] que pour tout $\varepsilon > 0$, il existe des constantes positives c_1, c_2 dépendant uniquement de d et c_1', c_2' dépendant uniquement de d et ε telles que pour tout $r > 0$,

$$\mathbb{B}(c_1 r - c_2) \subset S_n \subset \mathbb{B}(c_1' r + c_2'),$$

avec $n = \omega_d r^d$. Tout comme l'IDLA classique, le modèle de tas de sable présente une propriété abélienne : il a été démontré que la configuration finale des tas de sable est indépendante de l'ordre dans lequel les tas s'effondrent.

Une version continue de ce modèle est appelée *divisible sandpile model*, ou tas de sable divisible. Dans ce modèle, au lieu de considérer des valeurs entières pour compter les grains de sable présents sur chaque site, on associe à chaque site une masse donnée, et chaque site déverse une partie de sa masse à ses voisins dès que cette masse est strictement supérieure à 1. Dans ce cas, supposons qu'un site x ait une masse $m > 1$, alors chacun de ses $2d$ voisins reçoit une masse égale à $(m-1)/2d$, laissant x avec une masse égale à 1. Dans ce cas, x devient stable et ne déverse plus de masse. Maintenant, considérons une configuration initiale de masse sur $\mathbb{Z}^d$. Levine et Peres montrent dans [36] que lorsque le nombre de déversements tend vers l'infini, la configuration initiale converge vers une configuration finale, à condition que chaque site instable déverse sa masse une infinité de fois. Ici encore, nous pouvons montrer que cette configuration finale ne dépend pas de l'ordre dans lequel les sites s'effondrent. Levine et Peres montrent un *shape theorem* lorsque la configuration initiale consiste en une masse égale à $\omega_d n^d$ située à l'origine. Ils montrent qu'il existe des constantes $c,\ c'$ dépendant uniquement de d telles que

$$\mathbb{B}(n-c) \subset D(\omega_d n^d) \subset \mathbb{B}(n+c'),$$

où $D(\omega_d n^d)$ est l'ensemble des sites ayant une masse 1 dans la configuration finale du modèle de *divisible sandpile*, commencé avec une masse $\omega_d n^d$ à l'origine.

Modèle rotor-router

Nous présentons un modèle étroitement lié au modèle IDLA classique, appelé modèle *rotor-router*. Ce modèle peut être considéré comme une version déterministe de l'IDLA. Il a été étudié pour la première fois dans [41], initialement sous le nom de «marcheurs eulériens». Le modèle est le suivant. Considérons $\mathbb{Z}^2$, sur lequel nous avons positionné à chaque site un *rotor* pointant dans une direction : nord, est, sud ou ouest. On lance alors une particule depuis l'origine, et à chaque incrément de temps, le rotor sur lequel se situe la particule *active* pivote de 90 degrés dans le sens des aiguilles d'une montre. La particule se déplace alors dans la nouvelle direction indiquée par le rotor, jusqu'à ce qu'elle sorte de l'agrégat. Ce modèle peut être étendu à $\mathbb{Z}^d$, à condition d'avoir un ordre prédéterminé des $2d$ directions à parcourir. Il existe un *shape theorem* pour le modèle rotor-router, similaire au *shape theorem* de l'IDLA classique : quelle que soit la configuration initiale des rotors sur $\mathbb{Z}^d$, l'agrégat obtenu après avoir lancé $\omega_d n^d$ particules converge vers la boule euclidienne de rayon n. Nous pouvons également quantifier les fluctuations autour de cette boule. Plus précisément, si l'on note par $\widetilde{A}(n)$ l'agrégat rotor-router après avoir lancé n particules, alors il existe des constantes positives $c,\ c'$ dépendant uniquement de d telles que

$$\mathbb{B}(n - c\log n) \subset \tilde{A}(\omega_d n^d) \subset \mathbb{B}(n(1 + c' n^{-1/d}\log n)).$$

Nous fournissons une simulation du modèle rotor-router à la Figure 5.

IDLA sur le graphe de Sierpinski

Le modèle IDLA a également été étudié sur une classe spéciale de graphes, notamment sur le graphe de Sierpinski (*Sierpinski gasket*). Commençons par détailler la construction de ce graphe. Soit

$$V_0 = \left\{(0,0),\ (1,0),\ (1/2, \sqrt{3}/2)\right\},$$

et

$$E_0 = \left\{\{(0,0),\ (1,0)\},\ \{(0,0),\ (1/2,\ \sqrt{3}/2)\},\ \{(1,0),\ (1/2,\ \sqrt{3}/2)\}\right\}.$$

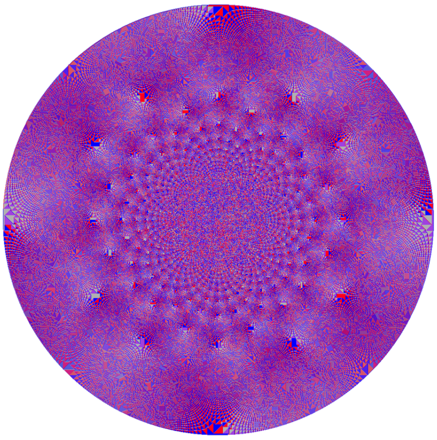


Figure 5: Simulation sur $\mathbb{Z}^2$ du modèle rotor-router avec 1 million de particules. Chaque couleur correspond à une direction d'un rotor. (Simulation : Y. Peres [36])

Pour $n \geq 0$, étant donné (V_n, E_n), on définit de façon récursive

$$V_{n+1} = V_n \cup \{(2^n, 0) + V_n\} \cup \left\{ \left(2^{n-1}, 2^{n-1}\sqrt{3}\right) + V_n \right\},$$

et

$$E_{n+1} = E_n \cup \{(2^n, 0) + E_n\} \cup \left\{ \left(2^{n-1}, 2^{n-1}\sqrt{3}\right) + E_n \right\},$$

où $(x, y) + A := \{(x, y) + a,\ a \in A\}$. Maintenant, soit $V_\infty := \bigcup_{n\geq 0} V_n$ et $E_\infty := \bigcup_{n\geq 0} E_n$. Enfin, on définit $V := V_\infty \cup \{-V_\infty\}$, et $E := E_\infty \cup \{-E_\infty\}$. Le graphe $\mathrm{SG} = (V, E)$ est appelé le *graphe doublement infini de Sierpinski* (doubly infinite Sierpinski gasket graph en anglais). Une illustration de ce graphe est donnée à la Figure 6.

On rappelle maintenant quelques concepts de théorie des graphes. Soit $G = (V, E)$ un graphe. Pour $x,\ y \in V$, on dit que x est connecté à y s'il existe une suite de sommets $x = v_0, v_1, \ldots, v_k = y$ tels que $(v_i, v_{i+1}) \in E$ pour tout $i \geq 0$. Dans ce cas, on note $x \sim y$, et on dit que $(v_0, \ldots, v_k)$ est un chemin de longueur k entre x et y. On peut alors définir la distance de graphe $d_G(x, y)$ comme la longueur du plus court chemin entre x et y. Si un tel chemin n'existe pas, c'est-à-dire si $x \not\sim y$, alors on pose $d_G(x, y) = +\infty$.

Notons $o = (0, 0)$ l'origine du graphe de Sierpinski SG. On pose $\mathcal{S}_0 = \{o\}$. Pour $n \geq 1$, on note $\mathcal{S}_n$ le cluster IDLA obtenu à partir de $\mathcal{S}_{n-1}$ après avoir lancé une particule depuis o et ajouté le premier site qu'elle visite en dehors de $\mathcal{S}_{n-1}$. Chen, Huss, Sava-Huss et Teplyaev montrent dans [13] un *shape theorem* pour l'IDLA sur SG. On note $B(n) := \{x \in V,\ d_G(o, x) < n\}$, et on définit $b_n := |B(n)|$. Alors, pour tout $\varepsilon > 0$, presque sûrement, pour n suffisamment grand, on a :

$$B(n(1-\varepsilon)) \subset \mathcal{S}(b_n) \subset B(n(1+\varepsilon)).$$

Encore une fois, la preuve de ce résultat utilise des ingrédients classiques d'IDLA. On commence par la preuve de la borne inférieure, en raisonnant à l'aide d'une décomposition similaire à celle de (1), et en utilisant des résultats sur les fonctions de Green sur SG. De manière plutôt surprenante, on obtient des bornes pour les fonctions de Green sur SG en utilisant des résultats sur le modèle du tas de sable divisible sur SG. Bien que ce modèle soit strictement déterministe, il partage de nombreuses propriétés avec le modèle IDLA sur SG. Une fois la borne inférieure démontrée, la

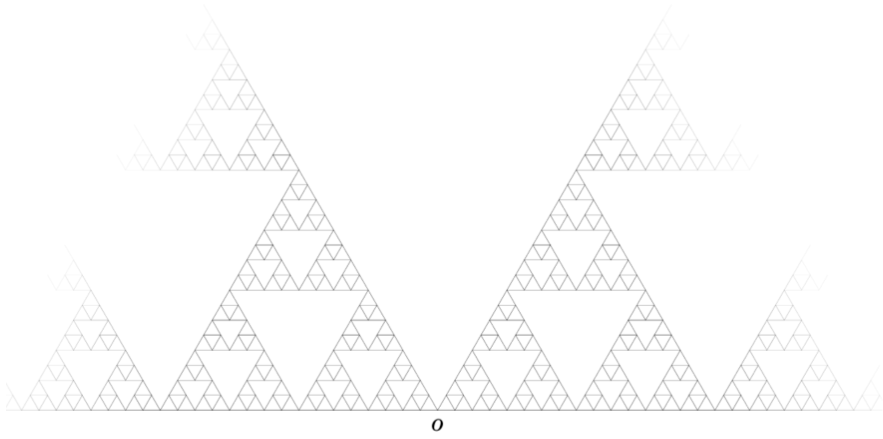


Figure 6: Le graphe doublement infini de Sierpinski. L'origine est notée o. (Figure : E. Sava-Huss)

borne supérieure est obtenue en utilisant un argument déjà connu : après avoir rempli la borne inférieure, il reste trop peu de particules pour déborder de la borne supérieure. Un tel argument avait déjà été utilisé dans [34], par exemple. Cela nécessite un travail supplémentaire sur les fonctions de Green du graphe de Sierpinski SG, et nécessite aussi de s'assurer que SG satisfait de bonnes conditions.

De récents travaux sur ce modèle ont permis de préciser les fluctuations du cluster autour de sa forme limite. Heizmann a montré dans [24] que, presque sûrement, pour tout $\kappa > 0$, il existe une constante $c > 0$ telle que, pour n suffisamment grand, on a :

$$B(n - cn^{1/2}\log(n)^{(1+\kappa)/2\alpha}) \subset \mathcal{S}_n \subset B(n - cn^{1/2+1/2\alpha}\log(n)^{(1-1/\alpha)(1+\kappa)/2\alpha}),$$

où $\alpha = \frac{\log 3}{\log 2}$ est la dimension de Hausdorff du graphe de Sierpinski.

Agrégation limitée par diffusion externe – external DLA

Avant l'introduction de l'IDLA par Meakin et Deutch dans [39], Witten et Sander ont introduit dans [49] le modèle de d'agrégation par diffusion *externe*, appelé aussi *external* DLA. Tout comme l'IDLA, l'*external* DLA est un modèle de croissance aléatoire sur un graphe infini et localement fini G. Ce modèle fournit un protocole pour construire des agrégats $(E_n)_{n\geq 0}$ sur G. Commençons par fixer $o \in G$ l'origine du graphe. On commence avec $E_0 = \{o\}$. Ensuite, à chaque étape, on lance une marche aléatoire «depuis l'infini», qu'on arrête une fois que celle-ci touche la frontière externe de l'agrégat. Le site sur lequel elle s'arrête est alors ajouté à l'agrégat, puis on recommence en envoyant une nouvelle marche aléatoire, indépendante des précédentes, «depuis l'infini». Nous expliquerons plus loin en détail ce que nous entendons par des marches lancées «depuis l'infini»; pour l'instant, le lecteur peut interpréter cela comme une marche lancée très loin du cluster. Contrairement à l'IDLA, les agrégats obtenus par DLA externe sont beaucoup moins bien compris et bien moins réguliers que les clusters habituellement lisses de l'IDLA. Une réalisation de DLA externe sur $\mathbb{Z}^2$ est illustrée à la Figure 7.

Commençons par définir rigoureusement le modèle. Pour cela, nous devons préciser ce que nous entendons par le lancement d'une marche aléatoire «depuis l'infini». Ce concept est difficile à généraliser pour tout graphe G, car il dépend du caractère transient ou récurrent de la marche sur G. Nous commençons par définir quelques notions de marches aléatoires sur un graphe G. Soit S une marche aléatoire sur G, et soit $A \subset G$. On définit le temps d'arrêt $\tau(A)$ comme suit :

$$\tau(A) := \inf\{n \geq 0,\ S_n \in A\}.$$

Pour une marche aléatoire S issue de $x \in \mathbb{Z}^d$, on définit la distribution de premier passage (*hitting distribution*) $H_A(x, y)$ par :

$$H_A(x, y) = \mathbb{P}_x\left(S_{\tau(A)} = y\right),\ \forall y \in A.$$

Nous souhaitons définir μ_A, la mesure harmonique de A *depuis l'infini* en prenant $|x| \to \infty$ dans $H_A(x, \cdot)$. Il faut cependant être prudent, car cette limite n'est pas toujours bien définie et ne définit pas systématiquement une mesure de probabilité sur A. Lorsque la marche aléatoire est récurrente sur G, μ_A est en général bien définie, mais des problèmes surviennent lorsque la marche est transiente. Pour faire simple, nous définissons le modèle d'*external* DLA sur $\mathbb{Z}^d$, où les mesures harmoniques «depuis l'infini» sont plus faciles à définir. Pour cela, nous devons considérer le cas récurrent $d = 2$ et le cas transient $d \geq 3$.

Le cas de $\mathbb{Z}^2$: Dans le cas de $\mathbb{Z}^2$, la marche aléatoire est récurrente, et $\mu_A(y) = \lim_{|x|\to\infty} H_A(x, y)$ est une quantité bien définie, et constitue une mesure de probabilité sur A. Dans ce cas, nous pouvons définir le cluster de DLA externe E_{n+1} comme $E_{n+1} = E_n \cup \{y_{n+1}\}$, où y_{n+1} est choisi selon la distribution $\mu_{\partial E_n}$, où ∂E_n désigne la frontière de E_n.

Le cas $d \geq 3$: Dans le cas de $d \geq 3$, on ne peut pas définir la mesure harmonique de la même manière que ci-dessus, car la marche aléatoire est transiente, et les quantités $\lim_{|x|\to\infty} H_A(x, y)$ sont nulles pour tout $y \in A$. Pour pallier ce problème, nous devons conditionner la marche à toucher l'ensemble A. Dans ce cas, on définit :

$$\mu_A(y) := \lim_{|x|\to\infty} \mathbb{P}_x\left(S_{\tau(A)} = y \mid \tau(A) < \infty\right) = \lim_{|x|\to\infty} \frac{H_A(x, y)}{\sum_{z \in B} H_A(x, z)},$$

qui vérifie $\sum_{y \in A} \mu_A(y) = 1$. Avec cette définition, nous définissons le cluster E_{n+1} en ajoutant à E_n un site y_{n+1} choisi selon $\mu_{\partial E_n}$.

Le cluster infini de DLA externe est alors défini par :

$$E_\infty := \bigcup_{n \geq 0} \uparrow E_n.$$

Nous présentons brièvement quelques résultats et questions concernant le DLA externe. L'une des principales interrogations autour de ce modèle concerne la longueur des « bras » qu'il génère. On définit $r_n := \max\{|x|,\ x \in E_n\}$ comme le rayon de E_n. Dans le cas où $G = \mathbb{Z}^d$, Kesten a démontré dans [32] qu'il existe des constantes $c(d)$ ne dépendant que de d telles que, presque sûrement,

$$\limsup_{n\to\infty} \frac{r_n}{n^{2/3}} \leq c(2), \quad \text{pour } d = 2,$$

$$\limsup_{n\to\infty} \frac{r_n}{n^{2/d}} \leq c(d), \quad \text{pour } d \geq 3.$$

Cette question a depuis été étudiée sur d'autres graphes que $\mathbb{Z}^d$. En particulier, Procaccia et al. (voir [42], [44], [43], [40]) ont étudié ce problème pour du DLA le long de segments, ou dans

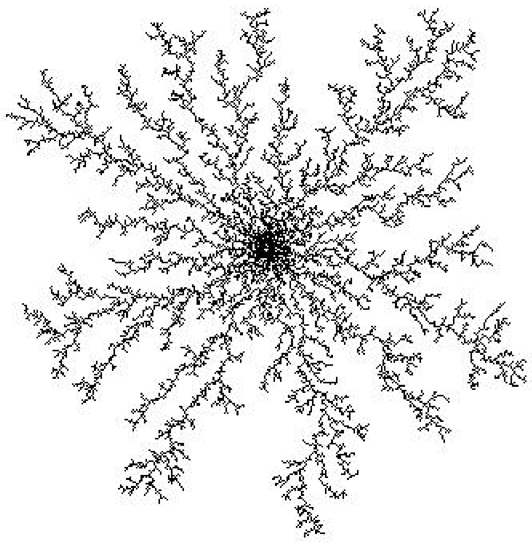


Figure 7: Une réalisation de DLA externe sur $\mathbb{Z}^2$ (Simulation : E. Sava-Huss [46])

le demi-plan supérieur de $\mathbb{Z}^2$. De même, Benjamini et Yadin introduisent dans [9] une méthode robuste qui s'applique à une large classe de graphes, tels que les graphes transitifs de croissance polynomiale, les graphes transitifs de croissance exponentielle, les graphes non moyennables (*non-amenable graph* en anglais), les amas de percolation surcritiques de $\mathbb{Z}^d$, ainsi que les tapis de Sierpinski en grande dimension. En particulier, ils améliorent le taux de croissance de r_n dans le cas de $\mathbb{Z}^3$, passant du taux $n^{2/3}$ à $\sqrt{n \log n}$. (Nous citons ici des résultats provenant de [9], un préprint déposé en 2017 qui, à notre connaissance, n'a pas encore été publié.)

En réalité, on conjecture que ce taux de croissance sur $\mathbb{Z}^d$ devrait être de l'ordre de $n^{1/d}$, mais ce problème reste ouvert depuis l'introduction du modèle. La conjecture s'énonce comme suit :

$$\limsup_{n \to \infty} \frac{\mathbb{E}[r_n]}{n^{1/d}} < +\infty.$$

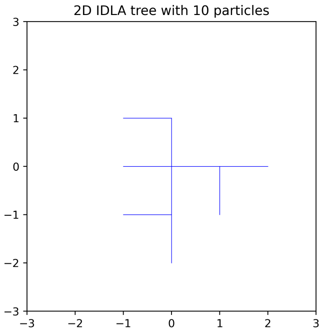
D'autres questions ouvertes concernent le nombre de «bouts» du cluster infini E_∞, ou encore le nombre de sommets occupés à l'intérieur d'une boule centrée autour de l'origine o du graphe. De nombreuses autres questions ouvertes entourent le DLA externe, illustrant à quel point ce modèle est difficile à étudier. Notamment, l'un des outils puissants présents pour l'IDLA mais absents pour le DLA externe est la propriété abélienne. L'absence d'un tel outil pour le DLA externe ainsi que la difficulté de définir l'envoi de marches « depuis l'infini » en font un modèle bien plus ardu à analyser que l'IDLA.

L'arbre IDLA

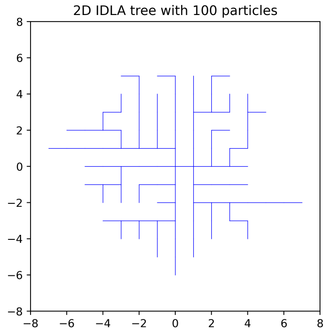
Dans ce paragraphe, nous présentons un arbre aléatoire sur $\mathbb{Z}^d$, enraciné en 0, qui se cache assez naturellement derrière le protocole d'IDLA classique. Cet arbre s'obtient tout simplement en considérant l'arête par laquelle chaque particule sort de l'agrégat $A(n)$. Pour construire cet arbre, on commence par construire une famille d'arbres finis $(\mathcal{T}_n)_{n\geq 1}$. On désigne chaque arbre par $\mathcal{T}_n = (V_n, E_n)$, où V_n désigne son ensemble de sommets et E_n son ensemble d'arêtes. Lorsque $n = 1$, on pose $V_1 = \{0\}$ et $E_1 = \emptyset$. Maintenant, pour tout $n \geq 2$, étant donné une réalisation de $\mathcal{T}_{n-1}$, on pose

$$V_n = V_{n-1} \cup \{S(\tau_{n-1})\}, \qquad E_n = E_{n-1} \cup \{(S(\tau_{n-1} - 1), S(\tau_{n-1}))\},$$

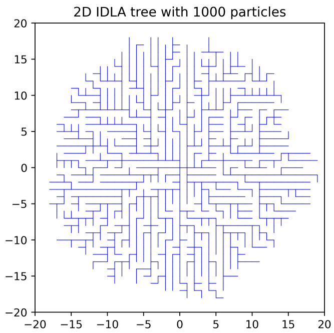
où $\tau_{n-1} = \inf\{t \geq 0,\ S(t) \notin V_{n-1}\}$. L'ensemble des sommets de $\mathcal{T}_n$ est exactement l'agrégat d'IDLA standard $A(n)$, tandis que la nouvelle arête ajoutée à E_{n-1} est l'arête par laquelle la particule est sortie de l'agrégat. Cette construction fournit un arbre, car l'extrémité de chaque nouvelle arête ajoutée forme un nouveau sommet, rendant impossible la formation de cycles. Nous fournissons quelques simulations de $\mathcal{T}_n$ à la Figure 8. Les marches aléatoires utilisées pour ces simulations sont les mêmes que celles utilisées pour les agrégats simulés à la Figure 3. Naturellement, par construction, la famille d'arbres $(\mathcal{T}_n)_{n\geq 1}$ est croissante pour l'inclusion, ce



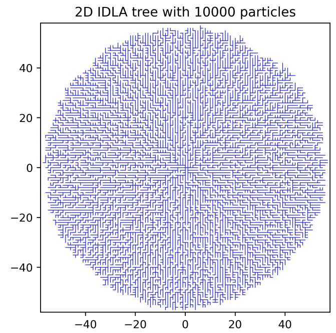
(a) Une réalisation de $\mathcal{T}_{10}$



(b) Une réalisation de $\mathcal{T}_{100}$



(c) Une réalisation de $\mathcal{T}_{1000}$



(d) Une réalisation de $\mathcal{T}_{10000}$

Figure 8: Réalisations de l'arbre IDLA en dimension 2 avec 10, 100, 1000 et 10000 particules

qui nous permet de définir l'arbre IDLA $\mathcal{T}_\infty$ comme suit :

$$\mathcal{T}_\infty := \bigcup_{n\geq 1} \uparrow \mathcal{T}_n \quad \text{p.s.}$$

Notons que, d'après (1), presque sûrement, l'arbre IDLA recouvre entièrement $\mathbb{Z}^d$. Néanmoins, d'autres propriétés concernant $\mathcal{T}_\infty$ restent moins claires. Dans le paragraphe suivant, nous citons quelques questions intéressantes concernant $\mathcal{T}_\infty$ et nous expliquons pourquoi il est difficile d'y répondre.

Absence d'Abélianité

Dans le cas des agrégats IDLA, nous avons vu que la propriété abélienne était un outil puissant pour démontrer de nombreux résultats. On pouvait notamment utiliser cette propriété pour construire des couplages élégants entre les différentes marches aléatoires. Dans le cas où l'on construit un graphe à partir d'un protocole IDLA, en rajoutant l'arête par laquelle la particule sort de l'agrégat, il s'avère que la propriété abélienne est alors fausse. Voici un bref exemple qui démontre ce point. Soit $x = (0, 1)$, $y = (0, 0)$. Posons $A = \{y\}$. Soient $\mathcal{F}_1$ et $\mathcal{F}_2$ les graphes obtenus après avoir lancé une marche aléatoire depuis x puis y, respectivement depuis y puis x. Dans le cas de l'IDLA, comme il ne peut pas se produire de cycle dans le graphe, les graphes obtenus seront systématiquement des forêts, c'est-à-dire des collections d'arbres disjoints.

Dans le cas de $\mathcal{F}_1$, puisque $x \notin A$, le site x est ajouté à l'agrégat et devient une racine pour cette forêt. Quelles que soient les marches S^x et S^y, $\mathcal{F}_1$ possèdera deux racines avec probabilité 1.

Dans le cas de $\mathcal{F}_2$, nous commençons par lancer une particule depuis y. Puisque $y \in A$, cette particule reste en vie et se déplace vers un site voisin. Elle a une probabilité $\frac{1}{4}$ d'aller vers x, auquel cas x serait ajouté à l'agrégat, mais ne constituerait alors *jamais* une racine pour $\mathcal{F}_2$. Ainsi, $\mathcal{F}_2$ ne contient qu'un seul arbre avec probabilité $\frac{1}{4}$, alors que, presque sûrement, $\mathcal{F}_1$ contient 2 arbres.

Les figures 9 et 10 illustrent cet exemple, où $S^x = (\rightarrow, \dots)$ et $S^y = (\uparrow, \leftarrow, \dots)$.

Difficultés et questions ouvertes concernant l'arbre IDLA

Plusieurs questions naturelles apparaissent au sujet de l'arbre IDLA, telles que l'existence de plusieurs branches infinies et, le cas échéant, leur direction asymptotique. Pour un arbre orienté $\mathcal{T} = (V, E)$, enraciné en o, on appelle branche de $\mathcal{T}$ tout chemin issu de o dans $\mathcal{T}$. (Voir la Figure 11 pour l'illustration d'une branche sur l'arbre IDLA $\mathcal{T}_{20000}$).

Nous savons que $\mathcal{T}_\infty$ possède au moins une branche infinie, car chacun de ses sommets est de degré fini. Cependant, il n'est pas du tout clair qu'il existe plusieurs branches infinies. Dans ce cas, sont-elles de nombre fini, ou y en a-t-il une infinité? Dans le cas où il existerait plusieurs branches infinies, nous serions intéressés de savoir si chacune d'elles possède une direction asymptotique, donnée par un certain vecteur $u \in \mathbb{Z}^d$. Une autre question naturelle serait de savoir si les branches de l'arbre IDLA sont tendues. Cela revient à dire que les branches de $\mathcal{T}_\infty$ peuvent être contenues dans un cône étroit, que nous détaillons ici. Pour tout $x \in \mathbb{Z}^d$ et tout $\varepsilon \in [0, \pi[$, on définit le cône $C(x, \varepsilon) := \{y \in \mathbb{Z}^d,\ \theta(x, y) \leq \varepsilon\}$, où θ désigne l'angle (à valeurs dans $[0, \pi]$) entre x et y. Pour tout $x \in \mathbb{Z}^d$, notons $\mathcal{T}_\infty^x \subset \mathcal{T}_\infty$ le sous-arbre enraciné en x, c'est-à-dire l'arbre ayant pour racine x et tous les descendants de x dans $\mathcal{T}$. On dit que $\mathcal{T}$ est *f-straight* si presque sûrement, pour tout $x \in \mathbb{Z}^d$,

$$\mathcal{T}_\infty^x \subset C(x, f(\|x\|)),$$

où f est une fonction positive telle que $f(l) \to 0$ quand $l \to +\infty$. Cette condition sur f impose qu'à mesure où l'on s'éloigne de l'origine, le cône s'affine de plus en plus. Intuitivement, cela traduit que les branches de l'arbre sont relativement droites.

Ces questions restent, à ce jour, sans réponse. La principale difficulté dans l'étude de l'arbre IDLA est son aspect radial. En effet, comme il est enraciné à l'origine, on peut observer sur

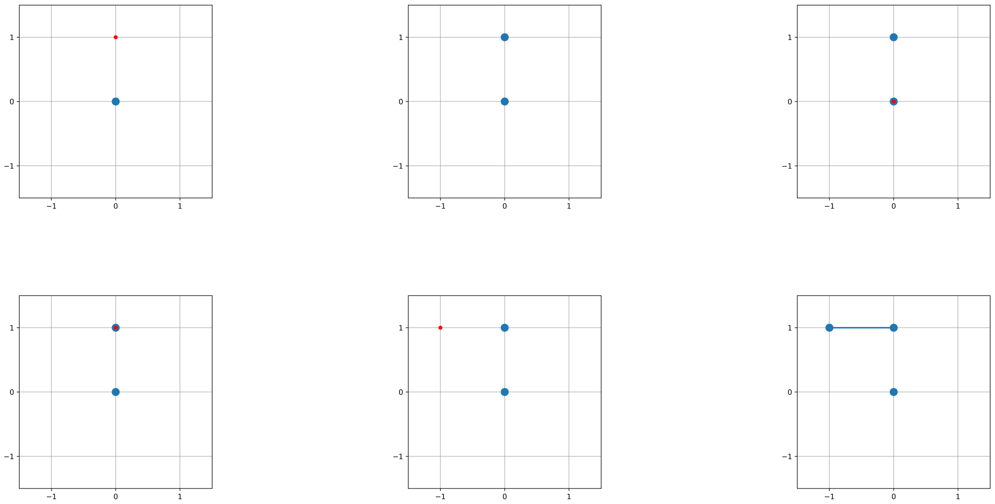


Figure 9: Une réalisation de $\mathcal{F}_1$. La particule est d'abord lancée depuis $x = (0, 1)$, puis depuis $y = (0, 0)$. La forêt $\mathcal{F}_1$ possède deux racines.

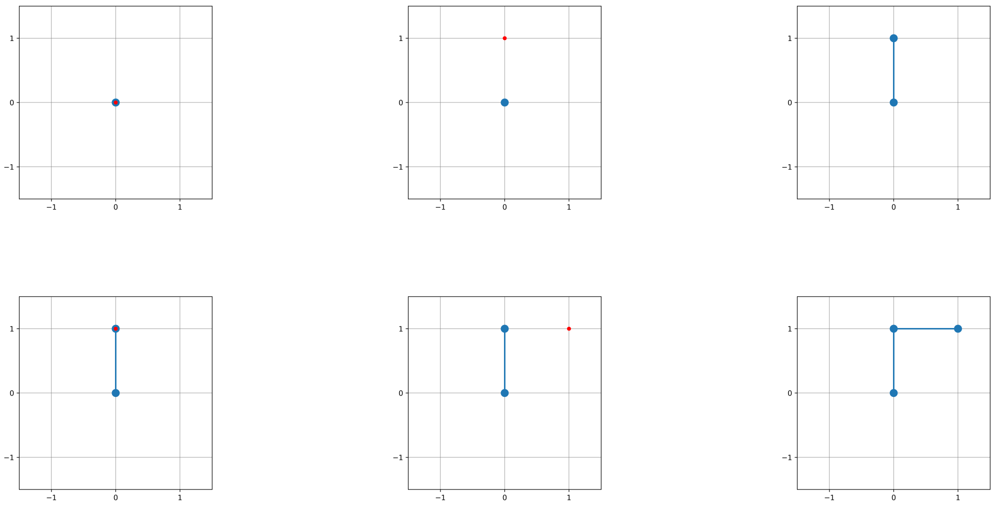


Figure 10: Une réalisation de $\mathcal{F}_2$. La particule est d'abord lancée depuis $y = (0, 0)$, puis depuis $x = (0, 1)$. La forêt $\mathcal{F}_2$ possède une seule racine: il s'agit en fait d'un arbre.

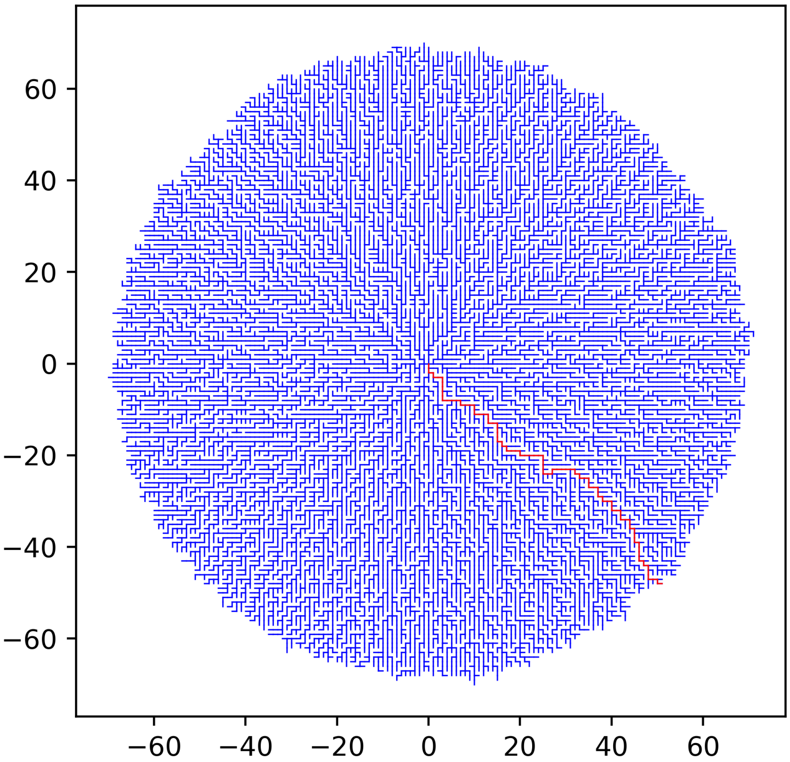


Figure 11: Une réalisation de $\mathcal{T}_{20000}$. Une branche de l'arbre est coloriée en rouge

la Figure 8 que les branches proches de l'origine sont fortement orientées dans cette direction. Cet aspect radial rend impossible l'utilisation d'arguments de stationnarité. De plus, chaque branche est produite par de nombreuses particules, introduisant une forte dépendance dans la construction de $\mathcal{T}_\infty$, et rendant alors son étude très difficile. Enfin, l'absence d'abélianité pour les arbres et les forêts requiert une compréhension plus fine du comportement des particules.

Pour surmonter ces difficultés, les auteurs de [16] ont eu l'idée de construire une forêt ergodique dans $\mathbb{Z}^2$ basée sur un protocole IDLA, appelée forêt IDLA dirigée de volume infini (*infinite-volume directed IDLA forest*), dans l'espoir que cette forêt puisse «approcher» $\mathcal{T}_\infty$. Nous détaillons ce que nous entendons par là ci-dessous. Nous avons mentionné que l'arbre IDLA avait un aspect radial marqué, ce qui le rend très difficile à étudier. Cependant, des simulations de l'arbre IDLA semblent montrer qu'à mesure où l'on s'éloigne de l'origine, l'aspect radial de l'arbre diminue. Par conséquent, nous sommes tentés d'étudier $\mathcal{T}_\infty$ à l'intérieur de la boule $\mathbb{B}(n \cdot e_1, r)$, où $r > 0$ est constant et n grand. Ainsi, $\mathcal{T}_\infty$ restreint à cette boule ne devrait pas présenter d'aspect radial et devrait ressembler davantage à une forêt dirigée par le vecteur $e_1 = (1, 0, \dots, 0)$. Une telle technique est utilisée par Baccelli et Bordenave dans [6] dans le cas du *Radial Spanning Tree* (RST) pour montrer que les branches sont tendues. Dans leur cas, ils approchent le RST par la *Directed Spanning Forest* (DSF). C'est cet argument qui motive la construction de la forêt IDLA de [16].

La construction de cette forêt est non triviale en dimension 2 et nécessite un travail préliminaire considérable sur le protocole IDLA sur lequel elle se base. Par ailleurs, le raisonnement utilisé dans [16] pour construire cette forêt est exclusif à la dimension 2 et ne fonctionne plus pour $\mathbb{Z}^d$ dès que $d \geq 3$. Ceci est expliqué plus en détail en Section 1 du Chapitre 3.

Résumé des résultats

Dans cette section, nous donnons un bref résumé des travaux présentés dans cette thèse. Le Chapitre 1 est un chapitre d'introduction dans lequel on présente (en anglais) les mêmes modèles et résultats sur l'IDLA que dans ce résumé. Le Chapitre 2 de cette thèse est consacré à la démonstration de propriétés concernant un nouveau protocole IDLA en dimension $d \geq 3$, tandis que le Chapitre 3 est dédié à prouver l'existence de la forêt IDLA dirigée en dimensions supérieures. La preuve que nous fournissons dans le Chapitre 3 fonctionne également en dimension 2, et elle généralise ainsi la construction de [16]. On commence par détailler le protocole IDLA sur lequel repose cette construction.

Trois familles d'agrégats IDLA

Dans cette section, nous détaillons le protocole IDLA utilisé pour construire les agrégats nécessaires à la construction de la forêt IDLA dirigée de volume infini. Pour ce faire, nous introduisons trois agrégats aléatoires, notés $A_n[\infty]$, $A_n^*[\infty]$ et $A_n^\dagger[\infty]$, construits en envoyant des particules depuis un ensemble infini de sources dans $\mathbb{Z}^d$. Chaque agrégat est obtenu en prenant la limite (en espace) d'une famille d'agrégats finis $A_n[M]$ (respectivement $A_n^*[M]$ et $A_n^\dagger[M]$).

Construction de $A_n[\infty]$: L'ensemble infini de sources que nous considérons est l'hyperplan $\mathcal{H} := \{0\} \times \mathbb{Z}^{d-1}$ de $\mathbb{Z}^d$, avec $d \geq 2$. Soient $n, M \in \mathbb{N}$. Dans ce qui suit, exactement n particules sont envoyées depuis chaque source de $\mathcal{H}$. On commence par construire une famille d'agrégats $(A_n[M])_{M \geq 0}$ par récurrence, comme suit. Lorsque $M = 0$, $A_n[0]$ est le modèle d'IDLA classique, c'est-à-dire où n particules sont émises depuis l'origine. On appelle *niveau* M l'ensemble des sources dans $\mathcal{H}$ à distance M de l'origine (pour $\|(z_1, \dots, z_d)\| := \max_i |z_i|$). Étant donné une réalisation de $A_n[M-1]$, on lance n particules depuis chaque source de niveau M selon l'ordre lexicographique. Ainsi, $A_n[M]$ est défini comme l'agrégat produit par $A_n[M-1]$ auquel on ajoute les nouveaux sites atteints par les particules lancées au niveau M.

Contrairement à sa forme, le nombre total de sites de l'agrégat $A_n[M]$ est déterministe, et est égal à $\#A_n[M] = n(2M+1)^{d-1}$. De plus, par construction, la séquence d'agrégats $(A_n[M])_{M\geq 0}$ est presque sûrement croissante pour l'inclusion, nous permettant de définir l'agrégat $A_n[\infty]$ comme :

$$A_n[\infty] := \bigcup_{M\geq 0} \uparrow A_n[M] \quad \text{p.s.}$$

Construction de $A_n^*[\infty]$: Tout comme pour $A_n[\infty]$, on commence par construire une famille d'agrégats aléatoires *finis* $(A_n^*[M])_{M\geq 0}$ avant de définir $A_n^*[\infty]$. Contrairement à $A_n[\infty]$, le nombre de particules envoyées depuis chaque source $z \in \mathcal{H}$ sera cette fois-ci aléatoire, donné par une variable aléatoire de Poisson N_z de paramètre n. Soit $(N_z)_{z\in\mathcal{H}}$ une famille de variables aléatoires indépendantes et identiquement distribuées de Poisson de paramètre n. Soit $M \geq 0$. Nous construisons $A_n^*[M]$ en utilisant le même protocole que pour $A_n[M]$, mais au lieu de lancer n particules depuis chaque source z de $\mathcal{H}_M := \{0\} \times [\![-M, M]\!]^{d-1}$, nous lançons N_z particules. On précise que les marches aléatoires sont envoyées indépendamment des N_z. Une fois de plus, la famille $(A_n^*[M])_{M\geq 0}$ reste croissante pour l'inclusion, et on peut ainsi définir $A_n^*[\infty]$ comme :

$$A_n^*[\infty] := \bigcup_{M\geq 0} \uparrow A_n^*[M] \quad \text{p.s.}$$

Construction de $A_n^\dagger[\infty]$: Étant donné que notre objectif est de définir une forêt aléatoire dont la distribution est invariante par translation de vecteurs de $\mathcal{H}$, nous devons introduire un protocole garantissant que chaque source joue le même rôle dans notre construction. Ce sera le cas pour la construction de notre dernier agrégat $A_n^\dagger[\infty]$.

Considérons une famille de processus ponctuels de Poisson (PPP) i.i.d. sur $\mathbb{R}_+$, d'intensité 1, notée $(\mathcal{N}_z)_{z\in\mathcal{H}}$. Le processus $\mathcal{N}_z$ fonctionne comme une *horloge aléatoire*, en fournissant des temps aléatoires correspondant aux temps d'émission depuis z de nos particules. À chaque famille de temps $(\tau_{z,j})_{j\geq 1}$ fournie par $\mathcal{N}_z$, nous associons une famille de marches aléatoires indépendantes $(S_{z,j})$ (également indépendantes des $\mathcal{N}_z$). Ainsi, au temps $\tau_{z,j}$, la j-ème particule est émise depuis la source z, et suit la trajectoire donnée par $S_{z,j}$ avant de quitter l'agrégat (courant). Pour éviter que plusieurs particules ne se déplacent simultanément, on supposera que la trajectoire de chaque particule est réalisée de façon instantanée.

On appelle $A_n^\dagger[M]$ l'agrégat obtenu en utilisant le protocole ci-dessus, en envoyant des particules depuis les sources de $\mathcal{H}_M := \{0\} \times [\![-M, M]\!]^{d-1}$ et sur l'intervalle de temps $[0, n]$. On peut montrer, à l'aide d'un couplage naturel (voir la Section 2.2), que la famille $(A_n^\dagger[M])_{M\geq 0}$ est croissante pour l'inclusion, ce qui nous permet de définir $A_n^\dagger[\infty]$ comme :

$$A_n^\dagger[\infty] := \bigcup_{M\geq 0} \uparrow A_n^\dagger[M] \quad \text{a.s.}$$

Puisque chaque PPP intervenant dans la construction de $A_n^\dagger[\infty]$ a une intensité égale à 1, nous avons $\mathcal{N}_z([0,n]) \overset{\text{loi}}{=} N_z$, où N_z est une variable de Poisson de paramètre n. Par conséquent, les agrégats aléatoires $A_n^*[M]$ et $A_n^\dagger[M]$ ne diffèrent que par l'*ordre* dans lequel leurs particules sont envoyées. Cependant, on sait grâce à la propriété abélienne que cela ne change pas la *loi* de l'agrégat final. Ainsi, pour tout $M \geq 0$,

$$A_n^\dagger[M] \overset{\text{loi}}{=} A_n^*[M]. \tag{3}$$

Puisque les deux familles d'agrégats $(A_n^*[M])_{M\geq 0}$ et $(A_n^\dagger[M])_{M\geq 0}$ sont croissantes pour l'inclusion, on déduit de (3) que $A_n^\dagger[\infty] \overset{\text{loi}}{=} A_n^*[\infty]$. Ce résultat s'avérera particulièrement utile dans toute la suite de cette thèse, car nous pourrons travailler sur l'agrégat $A_n^*[\infty]$ afin de déduire directement des résultats pour $A_n^\dagger[\infty]$, et vice versa. Dans certains cas, il sera plus facile de travailler avec des particules envoyées dans un ordre déterministe, d'où l'étude de $A_n^*[\infty]$, tandis que d'autres fois, il sera bien plus avantageux de travailler avec un ordre aléatoire. C'est le cas par exemple lorsque l'on veut montrer des propriétés de stationnarité pour les agrégats : il est plus facile de montrer des propriétés de stationnarité pour $A_n^\dagger[\infty]$ que pour $A_n^*[\infty]$, car dans le premier cas, les particules sont envoyées sans ordre particulier, tandis que dans le second cas, elles sont envoyées dans un ordre prédéterminé. Ce résultat, qui est une conséquence directe de la propriété abélienne, n'est vrai que dans le cas des *agrégats*, mais ne peut être utilisé dès lors que l'on considère des forêts, pour les raisons expliquées en Section 4.

Résumé du Chapitre 2

Le Chapitre 2 est consacré à l'étude des propriétés de l'agrégat aléatoire $A_n[\infty]$. On généralise plusieurs résultats de [16], démontrés pour la dimension 2, aux dimensions supérieures. Pour rappel, l'agrégat $A_n[\infty]$ s'obtient en lançant exactement n particules depuis chaque site de $\mathcal{H} = \{0\} \times \mathbb{Z}^{d-1}$ suivant l'ordre lexicographique. En suivant cette construction, on commence alors à faire pousser l'agrégat depuis l'origine. Nous montrons un résultat de stabilisation de l'agrégat, qui stipule que presque sûrement, toute particule lancée loin de l'origine ne va pas atteindre des régions proches de l'origine. Plus précisément, on montre le résultat suivant.

Theorem 0.1 (Stabilisation)**.** *Soit $n \geq 0$ et $\alpha > 1$. Presque sûrement, il existe un entier M_0 tel que, pour tout $M \geq M_0$, toute particule contribuant à $A_n[\infty]$ et émise depuis un site de niveau supérieur à M^α ne visite pas la bande $\mathbb{Z}_M = \mathbb{Z} \times [\![-M, M]\!]^{d-1}$.*

Ce résultat de stabilisation de l'agrégat se révèle particulièrement utile pour contrôler la probabilité d'événements faisant intervenir l'agrégat dans des régions proches de l'origine. En effet, si l'on souhaite démontrer une propriété de $A_n[\infty]$ à l'intérieur de la bande $\mathbb{Z}_M$, il suffit alors de considérer l'agrégat *fini* $A_n[M^\alpha]$, qui est beaucoup plus simple à manipuler que $A_n[\infty]$, car construit à l'aide d'un nombre fini de particules.

Ce raisonnement est par exemple nécessaire lorsqu'il s'agit de prouver le *shape theorem* pour $A_n[\infty]$. Nous montrons que presque sûrement, pour n suffisamment grand, la forme limite de l'agrégat $A_n[\infty]$, *proche de l'origine*, est la «dalle» $\mathcal{R}_{n/2} = [\![-n/2, n/2]\!] \times \mathbb{Z}^{d-1}$. Ce *shape theorem* est le résultat principal du Chapitre 2, et est formulé comme suit :

Theorem 0.2. *(Shape theorem) Pour tous entiers $d \geq 3$ et $\alpha \geq 1$, il existe une constante $C = C(d, \alpha) > 0$ telle que, presque sûrement, il existe un entier $N \geq 1$ tel que pour tout $n \geq N$,*

$$\mathcal{R}_{n/2 - C\sqrt{\log n}} \cap \mathbb{Z}_{n^\alpha} \subset A_n[\infty] \cap \mathbb{Z}_{n^\alpha} \subset \mathcal{R}_{n/2 + C\sqrt{\log n}} \cap \mathbb{Z}_{n^\alpha} \ . \tag{4}$$

Commentons brièvement ce résultat. Tout d'abord, la forme limite obtenue est la même que celle dans le Théorème 6.1 de [16], cependant les fluctuations autour de la forme limite diffèrent lors du passage de $d = 2$ à $d \geq 3$, comme c'était le cas pour le *shape theorem* de l'IDLA standard d'Asselah, Gaudillière et Jerison, Levine et Sheffield dans [4], [31]. On obtient ici des fluctuations

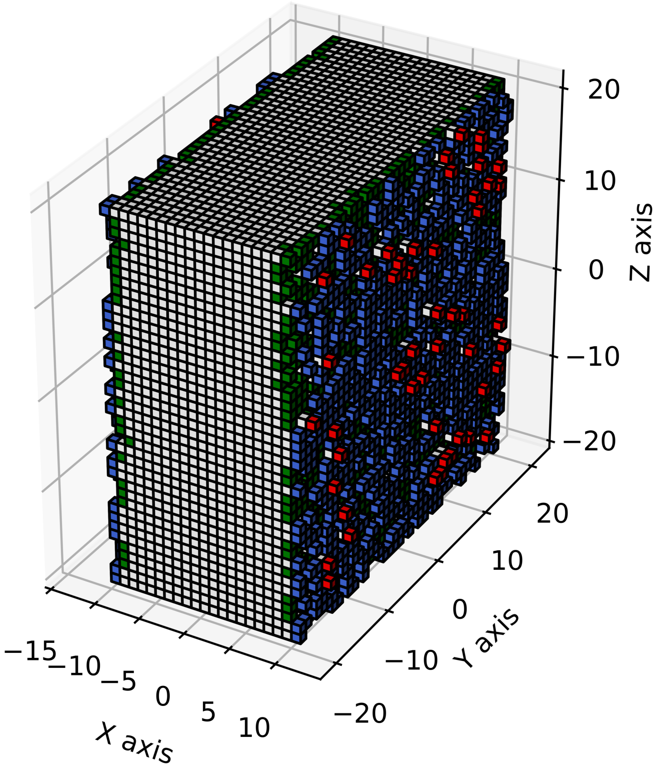


Figure 12: Une simulation de $A_{20}[40] \cap \mathbb{Z}_{20}$. Les points de l'axe des abscisses à la frontière de l'agrégat tels que $|x| = 10$ (resp. $|x| < 10$ et $|x| > 10$) sont représentés en bleu (resp. vert et rouge). Tous les autres points sont représentés en blanc.

autour de $\mathcal{R}_{n/2}$ de l'ordre de $\sqrt{\log n}$ lorsque $d \geq 3$, contrairement aux fluctuations d'ordre $\log n$ lorsque $d = 2$. Cette différence est, une fois de plus, une conséquence du comportement récurrent et transitoire des marches aléatoires lorsque $d = 2$ et $d \geq 3$ respectivement. Ce *shape theorem* affirme qu'au premier ordre, l'agrégat limite $A_n[\infty]$ a une épaisseur en moyenne égale à n, ce qui est cohérent puisque n particules sont lancées par source. Remarquons également que ce résultat ne s'applique pas à l'agrégat $A_n[\infty]$ tout entier, mais à l'agrégat restreint à la bande $\mathbb{Z}_{n^\alpha}$ (éventuellement grande). Il est *faux* que l'agrégat tout entier est contenu dans $\mathcal{R}_{n/2}$, car presque sûrement, il existe une source z pathologique (éventuellement très loin de l'origine) depuis laquelle les n particules émises se sont déplacées dans la direction de $e_1 = (1, 0, \dots, 0)$. Cela implique que presque sûrement, le site $z + n \cdot e_1 \in A_n[\infty]$. Ainsi, pour éviter de tels événements pathologiques, il est nécessaire de se restreindre à une bande proche de l'origine. Une illustration du *shape theorem* est visible à la Figure 12.

Commentons brièvement quelques unes des méthodes utilisées pour établir ce *shape theorem*. Une fois de plus, la propriété abélienne est centrale pour établir notre résultat. Comme c'était le cas dans [4], on utilise ici aussi un argument de coquilles et de cellules pour montrer la borne inférieure. Ce raisonnement nécessite de mettre en pause certaines particules avant de les lancer de nouveau, d'où l'importance de la propriété abélienne. Nous utilisons également une décomposition dans le même esprit que (2) afin de contrôler le nombre de particules par l'intermédiaire de marches aléatoires. Comme il est de coutume dans d'autres preuves, on se sert de la borne inférieure pour ensuite montrer la borne supérieure. Une difficulté supplémentaire de notre modèle par rapport aux autres modèles mentionnés jusqu'à présent est le fait que l'on manipule une infinité de particules. Il nous est donc nécessaire d'établir au préalable le résultat de stabilisation forte, qui nous permet alors de travailler sur l'agrégat restreint à une bande, et donc de se ramener à un nombre *fini* de particules.

On mentionne une autre difficulté du modèle. Celle-ci concerne le passage de la dimension $d = 2$ à $d \geq 3$. On insiste sur le fait que la généralisation des résultats de [16] aux dimensions supérieures est loin d'être immédiate et nécessite beaucoup plus de travail. L'une des principales difficultés provient de la géométrie de l'hyperplan $\mathcal{H}$ duquel sont émises les particules. On explique brièvement les différences qui interviennent lors du passage en dimensions supérieures. La construction de l'agrégat $A_n[M+1]$ à partir de $A_n[M]$ nécessite de lancer des particules supplémentaires à partir des sites du niveau $M+1$, c'est-à-dire à partir des sites de l'ensemble $\mathcal{H}_{M+1} \setminus \mathcal{H}_M$. Le nombre de sites de cet ensemble est égal à

$$(2(M+1)+1)^{d-1} - (2M+1)^{d-1}.$$

Lorsque $d = 2$, cette quantité est égale à 2, et les sites d'émission correspondants sont $(0, M)$ et $(0, -M)$. Cependant, pour $d \geq 3$, cette quantité dépend de M, et est de l'ordre de $(d-1)2^{d-1}M^{d-2}$. La plupart des techniques utilisées en dimension 2 cessent alors d'être valides en dimensions supérieures, car le nombre de particules émises à partir du niveau M explose lorsque $M \to \infty$, ce qui rend difficile le contrôle d'événements impliquant des particules loin de l'origine. Pour pallier ce problème, on fournit une borne globale supérieure pour $A_n[\infty]$, appelée dans cette thèse *global upper bound*, qui permet de contrôler l'épaisseur de l'agrégat pour des régions éloignées de l'origine. La forme choisie est suffisamment «souple» afin de pouvoir traiter les cas pathologiques mentionnés au paragraphe précédent sur le *shape theorem*.

Pour bien définir la borne globale, on commence par définir, pour tout $z \in \mathcal{H}$, la variable aléatoire suivante :

$$X_z(n) := \max\Big\{|z_1'|,\ z' \in A_n[\infty],\ z_i' = z_i \quad \forall i = 2, \ldots, d\Big\}.$$

Cette variable mesure tout simplement la valeur absolue de la première coordonnée du site le plus lointain de $A_n[\infty]$, sur la ligne $\mathbb{Z} \times \{z_2, \ldots, z_d\}$. Pour tout $n,\ M \geq 0$ et tout $\varepsilon > 0$, on définit l'événement suivant :

$$\mathbf{Over}(M, \varepsilon, n) = \bigcup_{l \geq M} \{\exists z \in \mathcal{H} : \|z\| = l,\ X_z(n) > \varepsilon l\}.$$

Cet événement se décrit de façon plus intuitive en définissant le cône $\mathscr{C}_\varepsilon$ suivant :

$$\mathscr{C}_\varepsilon := \bigcup_{l \geq 0} \Big\{z \in \mathbb{Z}^d,\ \|p_{\mathcal{H}}(z)\| = l,\ |z_1| \leq \varepsilon l\Big\},$$

où $p_{\mathcal{H}}$ est l'opérateur de projection orthogonale sur $\mathcal{H}$. Alors, on peut écrire plus simplement

$$\mathbf{Over}(M, \varepsilon, n) = \{A_n[\infty] \cap \mathbb{Z}_M^c \not\subset \mathscr{C}_\varepsilon\}.$$

On choisit de définir le cône $\mathscr{C}_\varepsilon$ de cette façon, car sa forme évasée permet de traiter des cas pathologiques où l'agrégat aurait une épaisseur anormalement élevée. Si l'on s'éloigne suffisamment de l'origine, presque sûrement, un tel phénomène aura lieu, mais au niveau $l \gg 1$, on autorise l'agrégat à être aussi grand que εl, ce qui nous offre suffisamment de marge pour pallier tout phénomène pathologique. On rappelle que l'on s'attend tout de même à avoir, *en moyenne*, un agrégat d'épaisseur n, donc la condition exigée par la borne globale reste assez souple.

La borne globale intervient fréquemment dans cette thèse pour des arguments que l'on appelle *donut method* ou encore *donut argument*. Cette méthode est souvent utilisée afin de contrôler des trajectoires de particules émises loin de l'origine. On explique ici brièvement l'idée conduisant à cette méthode. Quand on se place sur l'événement $\mathbf{Over}(M, \varepsilon, n)^c$, on sait alors que $A_n[\infty] \subset \mathscr{C}_\varepsilon$. Ainsi, toute particule qui sort de $\mathscr{C}_\varepsilon$ sort nécessairement de $A_n[\infty]$. L'idée est alors de découper $\mathscr{C}_\varepsilon$ en anneaux, ou tores (d'où le nom de *donut method*), et de dire qu'une particule émise loin de l'origine et qui visite des régions proches de l'origine a nécessairement traversé un grand nombre de *donuts*. Les dimensions de chaque donut sont bien choisies de sorte que l'on puisse borner sans trop de mal la probabilité qu'une marche traverse un donut donné. On montre alors que traverser chaque donut a un certain *coût* (en probabilité) $0 < c < 1$ pour la particule, et par un argument géométrique, on montre alors que la probabilité de traverser k donuts est bornée par $(1-c)^k$. Cette méthode nous permet, lorsque le nombre de donuts est suffisament grand, d'avoir des décroissances exponentielles pour les probabilités de ce type d'événements.

Résumé du Chapitre 3

Dans le Chapitre 3, on s'intéresse à la construction d'une forêt IDLA dirigée de volume infini. Cette forêt est construite à partir du même protocole que l'agrégat $A_n^\dagger[\infty]$, c'est-à-dire en utilisant un nombre aléatoire de particules, envoyées dans un ordre aléatoire. Le lecteur peut se référer au paragraphe 4 page 23 pour plus de précisions sur la construction de cet agrégat. On cherche à construire une forêt invariante en loi par rapport aux translations de $\mathcal{H}$, ce qui nous conduit à travailler avec le protocole donné par $A_n^\dagger[\infty]$, ce dernier étant construit en utilisant un ordre d'émission donné par un PPP. On commence par détailler la construction de cette forêt.

Construction de $\mathcal{F}_n[M]$

Soit $n,\ M \geq 0$. On peut construire assez naturellement à partir de $A_n^\dagger[M]$ une forêt aléatoire *finie* $\mathcal{F}_n[M]$, simplement en considérant les arêtes de $\mathbb{Z}^d$ par lesquelles les particules de $A_n^\dagger[M]$ quittent l'agrégat (en cours). Soit $(\mathcal{N}_z)_{z\in\mathcal{H}_M}$ la famille de PPP utilisées pour construire $A_n^\dagger[M]$. Soit $\kappa = \sum_{z\in\mathcal{H}_M} \#\mathcal{N}_z([0,n])$ le nombre total de particules émises jusqu'au temps n. On peut indexer les particules émises en fonction de leurs temps de départ $\tau_1 < \tau_2 < \ldots < \tau_\kappa \leq n$. Pour $j \in [\![1,\kappa]\!]$, on désigne par $A[j]$ l'agrégat obtenu au temps τ_j (c'est-à-dire une fois que la j-ème particule a été ajoutée à l'agrégat). On a $A[0] = \emptyset$ et $A[\kappa] = A_n^\dagger[M]$. Pour construire la forêt associée à cet agrégat, nous procédons par récurrence. On commence par poser $\mathcal{F}_n[M,0] = (\emptyset,\emptyset)$. Maintenant, pour $j \in [\![1,\kappa]\!]$, étant donné le graphe aléatoire $\mathcal{F}_n[M,j-1] = (V_{j-1},\ E_{j-1})$, nous obtenons $\mathcal{F}_n[M,j]$ comme suit. Soit z le site nouvellement ajouté à $A[j-1]$ par la j-ième particule.

- Si $z \in \mathcal{H}$ et qu'aucune particule n'a été envoyée auparavant depuis z, alors z est la racine d'un nouvel arbre dans le graphe. On définit alors $\mathcal{F}_n[M,j] = (V_j,\ E_j)$, où
$$V_j = V_{j-1} \cup \{z\}, \qquad E_j = E_{j-1}.$$
- Sinon, soit z' le dernier site de $A[j-1]$ visité par la particule avant d'atteindre z. Alors on définit
$$V_j = V_{j-1} \cup \{z\}, \qquad E_j = E_{j-1} \cup \{(z',z)\}.$$

Enfin, on définit $\mathcal{F}_n[M] := \mathcal{F}_n[M,\kappa]$. Avec cette construction, la forêt $\mathcal{F}_n[M]$ est une union finie d'arbres ayant tous leurs racines dans $\mathcal{H}$, et dont l'ensemble des sommets est $A_n^\dagger[M]$. Afin d'obtenir une forêt stationnaire, une approche naturelle serait de prendre la limite en *espace* (lorsque $M \to \infty$) puis en *temps* ($n \to \infty$) de $\mathcal{F}_n[M]$. Cependant, cette approche ne fonctionne pas, car il y a un problème de consistance (pour l'inclusion) concernant les forêts. Bien que l'on puisse construire un couplage afin d'avoir $A_n^\dagger[M] \subset A_n^\dagger[M']$ lorsque $M \leq M'$, ce même couplage ne garantit pas que $\mathcal{F}_n[M] \subset \mathcal{F}_n[M']$. Ce phénomène peut se constater à l'aide de la Figure 13. Il se peut que certains sommets, présents dans *les deux* forêts, soient atteints par des particules différentes, ce qui peut ajouter une arête à $\mathcal{F}_n[M]$ qui ne sera pas présente dans $\mathcal{F}_n[M']$ (et inversement). Ces discrépances peuvent se produire par un phénomène complexe appelé *chaînes de discrépance* (*chains of changes* en anglais). Pour pallier ce problème, on établit un résultat de stabilisation pour les forêts aléatoires $(\mathcal{F}_n[M])_{M\geq 0}$. Ce résultat est le résultat principal du Chapitre 3, nous permettant alors de définir notre forêt de volume infini. Ce résultat stipule que ces forêts finies, restreintes à une bande près de l'origine, sont consistantes.

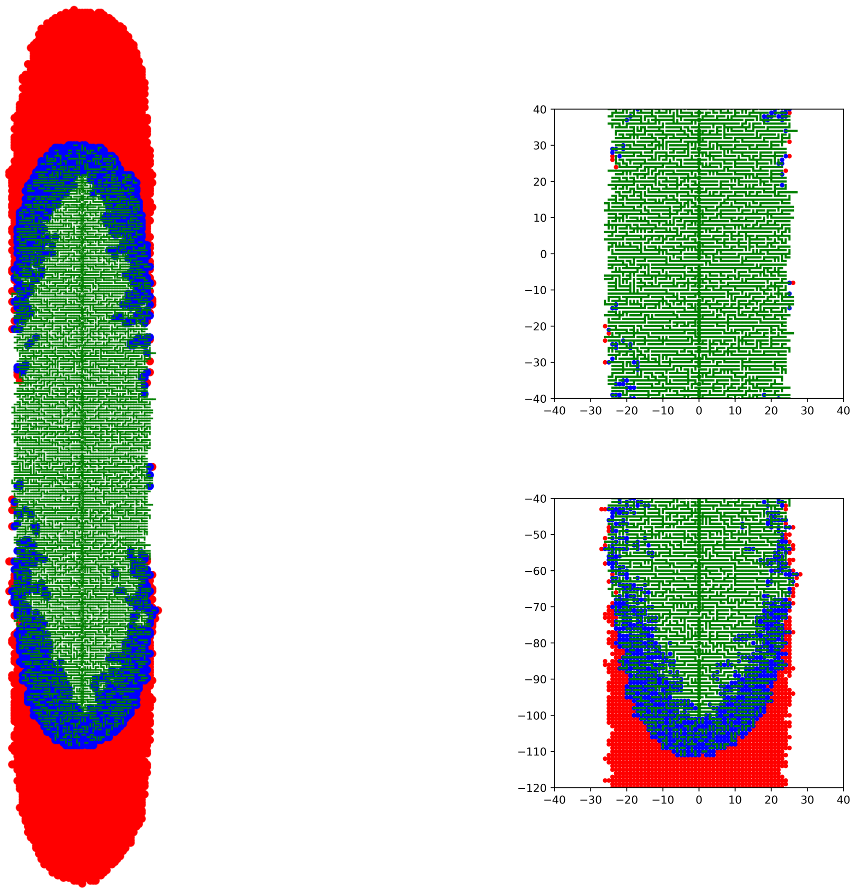


Figure 13: Réalisations de $\mathcal{F}_{50}[100]$ et $\mathcal{F}_{50}[150]$. Les arêtes présentes en vert sont communes aux deux forêts. Les points en vert représentent les racines communes aux deux forêts. Les points en rouge sont des sommets de $A_{50}^\dagger[150] \setminus A_{50}^\dagger[100]$. Les points en bleu sont des sommets communs à $A_{50}^\dagger[150]$ et $A_{50}^\dagger[100]$, mais atteints par différentes particules. Les arêtes correspondant à ces sommets bleus peuvent être différentes dans $\mathcal{F}_{50}[100]$ et dans $\mathcal{F}_{50}[150]$, et résultent de *chains of changes*. Les images à droite sont des zooms de l'image de gauche.

Theorem 0.3 (Stabilisation des forêts)**.** *Pour tout $n \geq 1$ et $K \geq 1$, presque sûrement,*

$$\exists N_0 = N_0(K) \geq 0,\ \forall N_1 \geq N_0,\ \mathcal{F}_n\left[N_1\right] \cap \mathbb{Z}_K = \mathcal{F}_n\left[N_0\right] \cap \mathbb{Z}_K.$$

À l'aide de ce résultat, il est possible de définir la forêt IDLA dirigée de volume infini. Pour cela, on commence par prendre la limite *en espace* des forêts, en définissant $\mathcal{F}_n$ comme

$$\mathcal{F}_n := \bigcup_{K \geq 1} \uparrow \mathcal{F}_n\left[N_0(K)\right] \cap \mathbb{Z}_K, \quad \text{p.s.}$$

Cette union est bien croissante. En effet, presque sûrement, on peut prendre la suite $(N_0(K))_{K\geq 1}$ croissante. Pour tout $K \leq K'$, on a alors :

$$\begin{aligned} \mathcal{F}_n\left[N_0(K)\right] \cap \mathbb{Z}_K &= \mathcal{F}_n\left[N_0(K')\right] \cap \mathbb{Z}_K \\ &\subset \mathcal{F}_n\left[N_0(K')\right] \cap \mathbb{Z}_{K'}. \end{aligned}$$

Il est immédiat que la suite de forêts $(\mathcal{F}_n)_{n\geq 0}$ est croissante pour l'inclusion. Cela découle simplement du fait que $\mathcal{F}_{n+1}$ s'obtient à partir de $\mathcal{F}_n$ en laissant pousser la forêt sur l'intervalle de temps $]n, n+1]$. On peut alors prendre une limite *en temps* afin de définir

$$\mathcal{F}_\infty := \bigcup_{n\geq 1} \uparrow \mathcal{F}_n,\ \text{p.s.}$$

La forêt obtenue est alors appelée forêt IDLA dirigée de volume infini, ou encore *directed infinite-volume IDLA forest.* On montre que cette forêt vérifie les propriétés suivantes :

Proposition 0.4.
1. *Presque sûrement, l'ensemble des sommets de $\mathcal{F}_\infty$, $V(\mathcal{F}_\infty)$, est tel que $V(\mathcal{F}_\infty) = \mathbb{Z}^d$.*
2. *Les distributions de $(\mathcal{F}_n)_{n\geq 0}$ et $\mathcal{F}_\infty$ sont invariantes par translation de vecteur k, $k \in \mathcal{H}$.*
3. *Les distributions de $(\mathcal{F}_n)_{n\geq 0}$ et $\mathcal{F}_\infty$ sont mélangeantes par translation de vecteur k, $k \in \mathcal{H}$.*

On commente brièvement la preuve du Théorème 0.3. Dans [16], en dimension 2, la preuve repose sur une propriété importante de l'agrégat $A_n^\dagger[\infty]$. Chenavier, Coupier et Rousselle montrent que presque sûrement, $A_n^\dagger[\infty]$ est composé de composantes connexes finies. Ils établissent ce résultat en montrant que l'on peut trouver une infinité de droites horizontales vides pour $A_n^\dagger[\infty]$. Cette propriété permet de contourner le problème des *chains of changes* à l'origine des problèmes de consistance des forêts. Pour des dimensions $d \geq 3$, on est tenté d'utiliser le même raisonnement. Il suffirait alors de montrer que l'on peut trouver des hyperplans vides dans $A_n^\dagger[\infty]$. Cependant, un simple argument de percolation permet de montrer qu'il n'est pas possible d'avoir un hyperplan de sites vides dans $A_n^\dagger[\infty]$ (voir Section 1 p.103). Il faut donc trouver une nouvelle stratégie pour stabiliser les forêts.

D'abord, expliquons en détail le phénomène de *chains of changes.* Comme mentionné précédemment, ce phénomène peut provoquer des discrépances entre les forêts $\mathcal{F}_n[M]$ et $\mathcal{F}_n[M']$ lorsque $M < M'$, ce qui entraîne un souci de consistance entre les forêts. Soit $M \geq 0$. Prenons $M' > M$. Par construction, toute particule émise d'un niveau $M < \|i\| \leq M'$ contribue à $A_n^\dagger[M']$ mais pas à $A_n^\dagger[M]$, générant une discrépance entre les deux agrégats et donc entre les deux forêts. Supposons qu'une particule soit émise depuis une source de niveau $M < \|i\| \leq M'$ à l'instant $t_1 \in [0, n[$ et ajoute un site z_1 à l'agrégat $A_n^\dagger[M']$, tandis que $A_n^\dagger[M]$ reste inchangé. Si l'on désigne par $A^\dagger_{t_1^-}[M']$ l'agrégat produit juste avant l'émission de cette particule, on obtient alors :

$$A^\dagger_{t_1}[M'] = A^\dagger_{t_1^-}[M'] \cup \{z_1\} \quad \text{et} \quad A^\dagger_{t_1}[M] = A^\dagger_{t_1^-}[M].$$

En réalité, cette même particule peut être ensuite à l'origine de *plusieurs* discrépances entre les forêts. Si aucune autre particule envoyée depuis les niveaux $\|i\| \leq M$ ne passe par z_1 durant $[t_1, n]$, alors z_1 est une discrépance entre les deux agrégats et donc entre les deux forêts. Sinon, soit t_2 l'instant auquel une particule émise depuis un niveau $\|i\| \leq M$ passe par z_1 durant $[t_1, n]$. Dans ce cas, deux événements se produisent :

- Puisque z_1 appartient déjà à $A^\dagger_{t_2^-}[M']$, la particule ajoute un nouveau site $z_2 \neq z_1$ à $A^\dagger_{t_2^-}[M']$. On a alors : $A^\dagger_{t_2}[M'] = A^\dagger_{t_2^-}[M'] \cup \{z_2\}$.
- Par construction, pour $A^\dagger_{t_2^-}[M]$, le site z_1 est ajouté à l'agrégat, donc $A^\dagger_{t_2}[M] = A^\dagger_{t_2^-}[M] \cup \{z_1\}$.

Remarquons qu'à l'instant t_2, le site z_1 ne constitue plus une discrépance entre les deux agrégats, et c'est désormais z_2 qui est une discrépance. On dit que la divergence a été *relayée* de z_1 à z_2. Bien que z_1 ne soit plus une discrépance entre les deux agrégats, on note cependant que l'arête par laquelle z_1 a été atteint pour $A^\dagger_{t_2}[M]$ peut différer de celle par laquelle il a été atteint pour $A^\dagger_{t_2}[M']$. Ainsi, bien qu'il n'y ait plus de discrépance en z_1 pour les agrégats, les deux forêts peuvent toujours différer au niveau de l'arête menant à z_1. On peut alors imaginer un scénario dans lequel z_2 est relayé en z_3 à un temps (aléatoire) t_3, entraînant potentiellement une autre différence entre les forêts au niveau de l'arête menant à z_2, et ainsi de suite... Un tel phénomène est appelé *chains of changes* et est très difficile à contrôler. On présente ci-dessous l'une des difficultés liées au contrôle des *chains of changes*.

Une stratégie naturelle pour prouver le Théorème 3.1 serait d'utiliser le lemme de Borel-Cantelli, et de montrer que

$$\sum_{N \geq 1} \mathbb{P}\left(\exists N_1 \geq N,\ \mathcal{F}_n[N_1] \cap \mathbb{Z}_K \neq \mathcal{F}_n[N] \cap \mathbb{Z}_K\right) < \infty.$$

Il est naturel de vouloir utiliser une *union bound* et de montrer que

$$\sum_{N \geq 1} \sum_{N_1 \geq N} \mathbb{P}\left(\mathcal{F}_n[N_1] \cap \mathbb{Z}_K \neq \mathcal{F}_n[N] \cap \mathbb{Z}_K\right) < \infty,$$

mais ce raisonnement n'aboutit pas. En effet, bien que l'événement $\{\mathcal{F}_n[N_1] \cap \mathbb{Z}_K \neq \mathcal{F}_n[N] \cap \mathbb{Z}_K\}$ implique l'existence d'une *chain of changes* entre $A^\dagger_n[N_1]$ et $A^\dagger_n[N]$, il est difficile d'obtenir avec cette stratégie des bornes décroissantes en fonction du paramètre N_1. Ce problème vient du fait que nous n'avons *a priori* aucun contrôle sur le niveau à partir duquel cette chaîne est initiée. Une telle chaîne pourrait être émise de n'importe quel niveau à partir de $N+1$, rendant impossible de borner notre probabilité par une quantité dépendant de N_1. Il faut donc trouver une approche différente pour mieux contrôler les *chains of changes*.

On présente dans cette thèse une stratégie inédite pour attaquer ce problème : on transforme le problème des *chains of changes*, qui concerne à la base des trajectoires de marches aléatoires et de particules, en un problème de percolation booléenne. On montre que l'absence de percolation dans un modèle bien choisi entraîne nécessairement la stabilisation des forêts. On s'inspire des techniques utilisées dans le Chapitre 2 sur l'agrégat $A_n[\infty]$ pour obtenir des propriétés similaires sur $A_n^\dagger[\infty]$. Parmi ces résultats, on peut citer une borne globale *affinée*, un *shape theorem*, ainsi qu'un résultat de stabilisation pour l'agrégat. Ces résultats sont nécessaires pour avoir un bon contrôle des rayons intervenant dans notre modèle de percolation booléenne. On montre alors l'absence *presque sûre* de percolation dans notre modèle, ce qui implique le Théorème 0.3.

Perspectives

On présente dans cette section d'éventuelles perspectives de recherche en lien avec les travaux présentés jusqu'ici.

L'arbre IDLA

Une première perspective concerne l'arbre IDLA, ce dernier ayant motivé tout le travail réalisé jusqu'à présent sur les agrégats $A_n[\infty]$ et $A_n^\dagger[\infty]$. Pour rappel, cet arbre est construit en considérant l'arête par laquelle est sortie chaque particule du modèle IDLA classique (voir (4) pour plus de détails concernant la construction de cet arbre). Les questions autour de l'arbre concernent l'existence de plusieurs branches infinies, et le cas échéant, la direction asymptotique de ces branches, ou encore le fait qu'il y en ait dans toutes les directions. Une autre question serait de savoir si les branches de cet arbre sont tendues. Comme nous l'avons mentionné précédemment, l'aspect radial de l'arbre IDLA rend l'objet en question particulièrement difficile à étudier. La stratégie privilégiée serait alors de comparer l'arbre IDLA, restreint à une boule le long d'un des axes et loin de l'origine, à la forêt IDLA construite au Chapitre 3. L'idée derrière cet argument est de dire que loin de l'origine, le caractère radial de l'arbre IDLA semble s'estomper. On serait alors tenté de dire que, restreints à une boule loin de l'origine, ils ont la même distribution. Cette méthode d'approximation de la loi d'un arbre par la loi d'une forêt est classique, utilisée par exemple par Baccelli et Bordenave dans [6] pour comparer la loi du *Radial Spanning Tree* (RST) à celle de la *Directed Spanning Forest* (DSF). Il faudrait alors montrer que la forêt IDLA $\mathcal{F}_\infty$ est un bon candidat pour approximer l'arbre IDLA $\mathcal{T}_\infty$.

Précisons ce que nous entendons par «approcher» l'arbre par la forêt. Soit $r > 0$, on définit $B_n := \mathbb{B}(n \cdot e_1, r)$, où $e_1 = (1, 0, 0, \dots)$. Nous aimerions prouver un résultat de la forme :

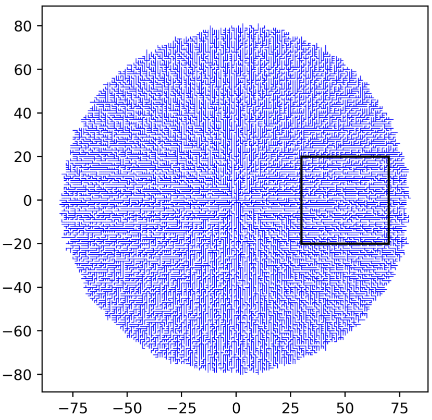
$$\lim_{n\to+\infty} \mathbb{P}\left(\mathcal{T}_\infty \cap B_n \neq \mathcal{F}_\infty \cap B_n\right) = 0.$$

Dans l'approximation de l'arbre par la forêt, il est crucial que le paramètre r dans la définition de B_n soit pris indépendamment de n.

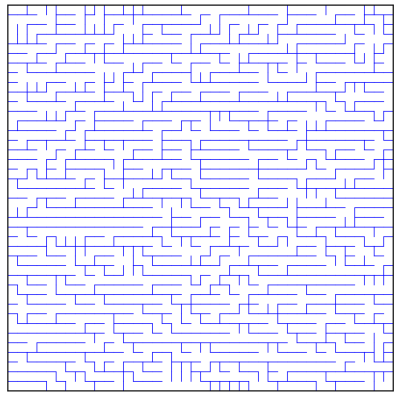
La Figure 14 permet de comparer l'apparence de l'arbre IDLA à celle de la forêt IDLA dirigée. Sur cette figure, on observe une réalisation de l'arbre IDLA dans une fenêtre éloignée de l'origine : il s'agit de la fenêtre $[30, 70] \times [-20, 20]$. On fait de même avec la forêt IDLA, que l'on observe dans la fenêtre $[0, 40] \times [-20, -20]$. On constate que l'arbre IDLA ne présente plus un aspect radial aussi prononcé qu'à l'origine. Les branches que l'on observe dans cette fenêtre sont alors comparables aux branches obtenues pour la forêt dirigée. Notons qu'on aperçoit légèrement l'aspect radial sur le côté gauche de cette simulation, où les branches semblent tout de même se diriger vers l'origine. Sur le côté droit de la simulation, cet aspect est nettement moins prononcé.

La forêt IDLA

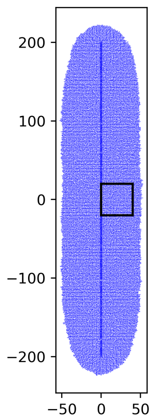
Une autre perspective serait l'étude de la forêt IDLA dirigée de volume infini, dont la construction et l'existence sont démontrées dans le chapitre 3. L'objectif initial derrière la construction de cette forêt était de l'utiliser afin d'approcher l'arbre IDLA, mais certaines questions intéressantes



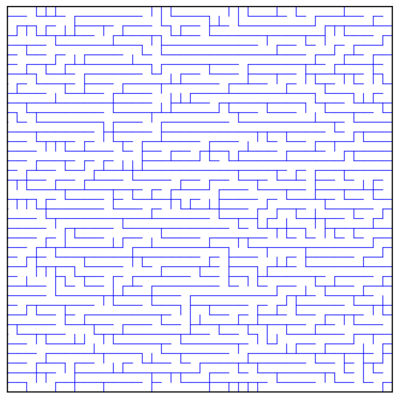
(a) Une réalisation de $\mathcal{T}_{20000}$



(b) La même simulation, observée dans le carré noir



(c) Une réalisation de $\mathcal{F}_{100}[200]$



(d) La même simulation, observée dans le carré noir

Figure 14: Une comparaison de l'arbre IDLA par la forêt IDLA

émergent également autour de cet objet. Nous avons montré que cette forêt était ergodique par rapport aux translations de $\mathcal{H}$. Une question naturelle concerne les arbres de la forêt et notamment leur taille. Sont-ils tous presque sûrement finis? Dans le cas positif, que peut-on dire de la loi de la longueur d'un «arbre typique» (en un sens à préciser) de la forêt?

Un autre résultat intéressant serait de prouver la stationnarité «loin» des sources de $\mathcal{H}$. Pour la raison évidente que la forêt possède toutes ses racines dans $\mathcal{H}$, $\mathcal{F}_\infty$ n'est pas invariant par translations horizontales de T_k, où $k \in \mathbb{Z} \times \{0\}^{d-1}$. Cependant, il serait intéressant d'étudier le comportement en temps long de la forêt, c'est-à-dire loin de $\mathcal{H}$, et de voir si cette forêt présente un comportement stationnaire. Cela impliquerait qu'après un temps long, les branches «oublient» d'où elles viennent et seraient distribuées selon une loi stationnaire. Ces questions sont complexes, et nécessitent sans doute une étude plus fine des propriétés de l'agrégat aléatoire $A_n^\dagger[\infty]$.

Liens avec le GFF

Le *Gaussian Free Field* (GFF), ou champ libre gaussien, est un champ aléatoire très populaire en physique statistique depuis le début des années 70. Ce champ peut s'interpréter comme une généralisation d'un pont Brownien, dans lequel le temps est remplacé par un paramètre multi-dimensionnel. À cause de son comportement très singulier, il est défini comme une distribution aléatoire. Jerison, Levine et Sheffield montrent dans [28] que les fluctuations de l'agrégat IDLA classique, convenablement renormalisées en temps et en espace, convergent vers une variante du GFF. Dès lors, si on définit

$$E_t = \frac{1}{r} \sum_{x \in \mathbb{Z}^2} \left(\mathbb{1}_{x \in A(T(t))} - \mathbb{1}_{x \in \mathbb{B}(r)} \right) \delta_{x/r},$$

où $T(t)$ désigne une variable de Poisson de paramètre t, $A(\cdot)$ l'agrégat IDLA classique, et où $r = \sqrt{t/\pi}$, alors E_t converge au sens des distributions vers une variante du GFF. C'est-à-dire, pour toute famille de fonctions test $\phi_1, \dots, \phi_k$ (d'un espace convenablement choisi), la loi jointe de $(\langle \phi_1, E_t \rangle, \dots, \langle \phi_1, E_t \rangle)$ converge en loi vers $(\langle \phi_1, h \rangle, \dots, \langle \phi_1, h \rangle)$, où h désigne une variante du GFF, appelée *augmented GFF*.

Nous commentons brièvement ce résultat. En lançant $T(t)$ particules, on s'attend à ce que $A(T(t))$ ait, au vu de (1), une forme proche de celle de $\mathbb{B}(r)$. Il faut voir la mesure aléatoire E_t comme la mesure de l'erreur moyenne entre l'agrégat IDLA $A(T(t))$ et sa forme attendue $\mathbb{B}(r)$. Certains de ces écarts par rapport à la moyenne vont s'annuler, et on peut alors montrer que ces fluctuations, convenablement renormalisées, convergent au sens des distributions vers une variante d'un champ gaussien.

Il serait intéressant de voir si un résultat analogue peut s'obtenir concernant les fluctuations de l'agrégat $A_n[\infty]$. Cela semble cependant plus complexe, étant donné que la forme limite de l'agrégat $A_n[\infty]$ est moins bien contrôlée que celle de l'agrégat classique $A(\cdot)$. Notamment, toutes les propriétés d'isotropie de la boule euclidienne ne pourraient pas s'appliquer dans le cas de l'agrégat $A_n[\infty]$. Il faudrait manipuler le pavé $\mathcal{R}_{n/2} = [\![-n/2, n/2]\!] \times \mathbb{Z}^{d-1}$, qui présente de moins bonnes propriétés, ce qui rajouterait une difficulté supplémentaire au problème.

Modèles de compétition

Nous avons présenté jusqu'ici des variantes de l'IDLA dans lesquelles les particules sont toujours envoyées les unes après les autres, sans jamais avoir de particules en vie simultanément. Nous présentons ici un modèle d'IDLA dans lequel les particules présentes sont en compétition les unes entre les autres. Pour définir ce modèle, on se munit de deux familles de PPP, $\left(\mathcal{N}_z^b\right)_{z \in \mathbb{Z}^d}$ et $(\mathcal{N}_z^m)_{z \in \mathbb{Z}^d}$, d'intensités respectives λ et μ. La première famille d'horloges correspond aux temps de naissance (*birth*) des particules, alors que la deuxième famille correspond aux temps de mouvement (*move*) des particules. Ainsi, les particules de ce modèle naissent avec un taux λ et se déplacent avec un taux μ. On se donne une réalisation d'un agrégat $A_{\lambda,\mu}(t) \subset \mathbb{Z}^d$. Les règles de naissance et de mouvement sont les suivantes :

- Des particules naissent en chaque site z selon les horloges données par $\mathcal{N}_z^b$, à condition que z soit dans l'agrégat en cours. Si z n'est pas dans l'agrégat, alors il ne se passe rien.

Autrement dit, une particule ne peut naître sur un site que si ce site est déjà présent dans l'agrégat, et que si aucune autre particule vivante ne l'occupe.

- Une particule présente sur un site z se déplace selon les horloges données par $\mathcal{N}_z^m$. Dans ce cas, la particule choisit un site voisin z' uniformément au hasard parmi ses $2d$ voisins. Si une particule est en vie sur z', alors la particule située en z ne se déplace pas (au moins jusqu'à la prochaine horloge donnée par $\mathcal{N}_z^m$). Par contre, si z' est libre, alors la particule se déplace sur z'. Son prochain mouvement sera donné par la prochaine horloge de $\mathcal{N}_{z'}^m$.
- Si une particule se déplace vers un site z qui ne fait pas partie de l'agrégat, alors la particule meurt, et le site z est alors rajouté à l'agrégat en cours. En particulier, des particules peuvent désormais naître au niveau du site z.

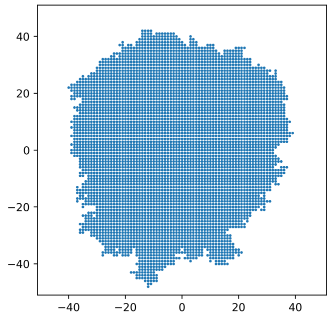
Ainsi, en partant de $A_{\lambda,\mu}(0) := \{0\}$, pour $t > 0$, on définit $A_{\lambda,\mu}(t)$ comme l'agrégat obtenu jusqu'au temps t selon le protocole décrit ci-dessus, en utilisant les horloges de $\left(\mathcal{N}_z^b([0,t])\right)_{z\in\mathbb{Z}^d}$ et de $(\mathcal{N}_z^m([0,t]))_{z\in\mathbb{Z}^d}$.

Nous pensons que ce modèle est un intermédiaire entre le modèle d'*uniform* IDLA, ou uIDLA, et le modèle *Eden*, aussi appelé *modèle de Richardson*. Le protocole de l'uIDLA est quasiment identique au protocole de l'IDLA standard, avec comme seule différence le fait qu'à l'étape $n+1$, la particule émise part d'un site choisi uniformément au hasard dans l'agrégat en cours, au lieu de partir de l'origine. Ce modèle a en particulier été étudié dans [7]. Les auteurs démontrent un *shape theorem* avec comme forme limite la boule euclidienne de rayon n, quand un nombre $|\mathbb{B}(n)|$ de particules est envoyé. Dans le modèle *Eden*, on commence avec un agrégat constitué de $\{0\}$ au temps 0, puis on rajoute un nouveau site x à l'agrégat après un temps exponentiel, dont le taux est proportionnel au nombre de sites voisins de x occupés. Un *shape theorem* existe également pour ce modèle : Richardson montre dans [45] l'existence d'une forme limite non-explicite. Celle-ci est convexe, mais il est conjecturé qu'il ne s'agit a priori pas d'une boule euclidienne.

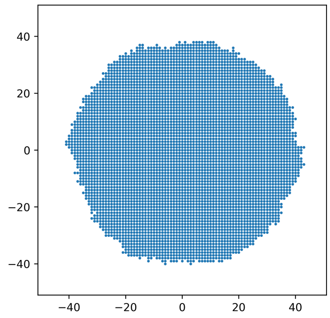
On s'intéresse au comportement de notre modèle en fonction du rapport λ/μ entre les naissances et le mouvement des particules. En particulier, on suspecte que lorsque $\lambda/\mu \ll 1$, le modèle devrait se comporter comme le modèle d'uIDLA. Notre intuition consiste à dire que si $\lambda/\mu \ll 1$, alors les particules naissent à un taux très faible par rapport au taux de mouvement, et le modèle devrait se comporter comme s'il n'y avait qu'une seule particule active à la fois. L'aspect où les particules rentrent en compétition les unes avec les autres s'estompe, et l'on se retrouve alors dans une situation proche de celle du modèle d'uIDLA dans lequel une particule naît uniformément au hasard dans l'agrégat en cours et réalise sa trajectoire jusqu'à en sortir.

Dans le cas où $\lambda/\mu \gg 1$, on suspecte un comportement inverse. Il y aurait alors énormément de particules actives simultanément, et l'agrégat (en cours) aurait alors tous ses sites occupés par des particules actives. Dans ce cas, les particules se bloqueraient toutes entre elles, et seules les particules situées au bord de l'agrégat pourraient alors se déplacer (vers l'extérieur) pour rajouter un nouveau site à l'agrégat. On rappelle que les particules de notre modèle ne peuvent que se déplacer vers des sites inoccupés. On présume alors un comportement similaire à celui de l'*Eden model*.

On fournit une simulation des deux régimes ($\lambda/\mu \gg 1$ et $\lambda/\mu \ll 1$) à la Figure 15. Les deux agrégats sont constitués de 5000 particules. La simulation de gauche correspondrait au régime Eden, tandis que celle de droite correspondrait au régime uIDLA. Dans le cas de l'uIDLA, la forme limite ressemble bien à une boule euclidienne, tandis que la forme limite dans le modèle Eden est beaucoup moins claire. On fournit une autre simulation du modèle Eden à la Figure 16.



(a) Une simulation avec $\lambda = 10^4,\ \mu = 1$



(b) Une simulation avec $\lambda = 1,\ \mu = 10^4$

Figure 15: Une comparaison des deux régimes

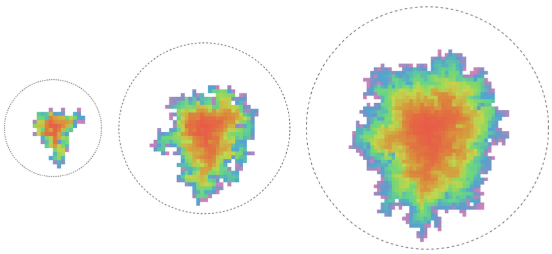


Figure 16: Une simulation du modèle Eden avec 100, 500 et 1500 particules (source: [25])

Chapter 1

Introduction

Outline of the current chapter

1 Known results on IDLA and related models

1.1 Motivations

Internal Diffusion Limited Aggregation (IDLA) is a random growth model on $\mathbb{Z}^d$. It provides a protocol to build aggregates recursively as follows. At step 0, assume the aggregate is empty. Then, at each step, we throw a particle from the origin, and stop it once it visits a site outside of the current aggregate. The particle then occupies that site, and we start over by sending a new particle from the origin.

IDLA was first introduced by Meakin and Deutch in 1986 (see [39]) to model an industrial chemical technique known as *electropolishing*. As opposed to electroplating, whose purpose is to coat objects in a thin layer of metal, the purpose of electropolishing is to polish the surface of a metal by removing a thin layer of material from it. This is achieved through a redox reaction. The metal we wish to polish is dipped inside an electrolyte solution, and connected to a power supply, making it serve as the positively charged anode. An oxidation reaction occurs at the surface of the metal, causing negatively charged anions to dissolve inside the electrolyte solution. This leaves an aggregate of empty sites at the surface of the metal. It then becomes of significant interest to quantify how smooth the surface of a given metal can get through such a process. One of the main motivations when working on IDLA processes is the study of the shape of the aggregate, through results referred to as *shape theorems*. Such results aim to describe the limiting shape of the aggregate while also giving sharp bounds concerning the fluctuations around the limiting shape.

1.2 Construction and properties

Diaconis and Fulton are the first to introduce IDLA in a mathematical context in 1991 in [19]. The authors begin by defining the *smash sum* of two sets $A,\ B \subset \mathbb{Z}^d$, denoted by $A \oplus B$. The smash sum $A \oplus B$ provides a random set, whose cardinality is equal to $|A| + |B|$. First, if $A \cap B = \emptyset$, then let $A \oplus B = A \cup B$. Now, assume $A \cap B \neq \emptyset$. We write $A \cap B = \{z_1, \dots,\ z_k\},\ k \geq 1$. Similarly to the IDLA protocol described above, the protocol behind the smash sum consists in launching independent random walks from each site of $A \cap B$ and adding to $A \cup B$ the first site the walk visits outside of $A \cup B$. Formally, we define the smash sum recursively as follows. Let $C_0 = A \cup B$. We define C_1 as

$$C_1 := C_0 \cup \{S_{z_1}(\tau_0)\},$$

where S_{z_1} denotes a random walk (simple, symmetric) started in z_1, and τ_0 is the following stopping time:

$$\tau_0 := \inf\{t \geq 0,\ S(t) \notin C_0\}.$$

More generally, for $1 \leq i \leq k$, we define C_i as:

$$C_i := C_{i-1} \cup \{S_{z_i}(\tau_{i-1})\},$$

where $\tau_{i-1} := \inf\{t \geq 0,\ S(t) \notin C_{i-1}\}$, and where S_{z_i} denotes a random walk, independent of the previous walks. Finally, we let $A \oplus B := C_k$. Roughly speaking, the smash sum of A and B is the aggregate $A \cup B$ and an additional $|A \cap B|$ points, obtained by launching a random walk from each site of $A \cap B$, and adding the first site it visits outside of the (current) aggregate.

Abelian Property

One may be initially inclined to say that the law of the smash sum depends on the order in which we indexed the elements of $A \cap B$. Surprisingly however, this isn't the case, and this property is referred to as the Abelian Property. It is stated in detail below.

Proposition 1.1 (Abelian Property)**.** *Let $A \subset \mathbb{Z}^d$ be a bounded subset of $\mathbb{Z}^d$, and take $z_1, \dots, z_k \in \mathbb{Z}^d$. The distribution of*

$$((A \oplus \{z_1\}) \oplus \{z_2\}) \oplus \cdots \oplus \{z_k\}$$

does not depend on the order of the z_i's. That is, if we take $\sigma \in \mathfrak{S}_k$ a permutation of $\{1, \dots, k\}$, then:

$$((A \oplus \{z_1\}) \oplus \{z_2\}) \oplus \cdots \oplus \{z_k\} \overset{\text{law}}{=} ((A \oplus \{z_{\sigma(1)}\}) \oplus \{z_{\sigma(2)}\}) \oplus \cdots \oplus \{z_{\sigma(k)}\}.$$

One can actually prove a stronger result than the previous one, and show that it is possible to pause the trajectory of a given walk, throw other walks in the mean time, and resume its trajectory after, without ever changing the law of $A \oplus B$. In all that follows, the term 'particle' will be used to refer to a random walk up to the time it exits the aggregate.

This property is central in many of the results presented in this thesis. It allows us to consider protocols where particles are stopped mid-way, then resumed again, without ever changing the law of the final aggregate. The Abelian Property is in a way the cornerstone of many IDLA proofs, as most of the strategies used for IDLA protocols rely in some way or another on this property.

Proof of Proposition 1.1. Let $x,\ y \in \mathbb{Z}^d$ and $A \subset \mathbb{Z}^d$. Let $S^x,\ S^y$ denote two independent simple, symmetric random walks issued from x and y respectively. We define the following stopping times:

$$T^x := \inf\{n \geq 0,\ S^x_n \notin A\}, \qquad T^y := \inf\{n \geq 0,\ S^y_n \notin A\}.$$

Let $(M^x_k)_{k\geq 0}$ and $(M^y_k)_{k\geq 0}$ denote the increments of S^x and S^y respectively *after* these walks exit A. These are well defined since $T^x < \infty$ and $T^y < \infty$ almost surely, as $A \subset \mathbb{Z}^d$ is bounded. That is, M^x and M^y are such that for any $n \geq 0$:

$$S^x_n = S^x_{n\wedge T^x} + \sum_{k=1}^{n-T^x} M^x_k \mathbb{1}_{n>T^x},$$

$$S^y_n = S^y_{n\wedge T^y} + \sum_{k=1}^{n-T^y} M^y_k \mathbb{1}_{n>T^y}.$$

Now, we define the random walks $\tilde{S}^x$ and $\tilde{S}^y$ as follows:

$$\tilde{S}^x_n = S^x_{n\wedge T^x} + \sum_{k=1}^{n-T^x} M^y_k \mathbb{1}_{n>T^x},$$

$$\tilde{S}^y_n = S^y_{n\wedge T^y} + \sum_{k=1}^{n-T^y} M^x_k \mathbb{1}_{n>T^y}.$$

The walk $\tilde{S}^x$ follows the same trajectory as S^x until it exits A, and then follows a trajectory given by the increments of S^y after it exits A. Similarly, $\tilde{S}^y$ follows the same trajectory as S^y until it exits A, and then follows a trajectory given by the increments of S^x after it exits A. Now, let us call E the random aggregate obtained (by smash sum of A and $\{x\}$ then $\{y\}$) after launching the particle from x using S^x first, and then the particle from y using S^y. Similarly, we call $\tilde{E}$ the aggregate obtained (by smash sum of A and $\{y\}$ then $\{x\}$) after launching the particle from y using $\tilde{S}^y$ first, then from x using $\tilde{S}^x$.

We can show that through our coupling, we have that $E \overset{\text{a.s}}{=} \tilde{E}$. We detail this below:

- If $S^x(T^x) \neq S^y(T^y)$, then S^x and S^y exit A through different sites, and the same goes for $\tilde{S}^x$ and $\tilde{S}^y$. In that case, it is clear that launching from x then y is the same as launching from y then x.
- Now, suppose $S^x(T^x) = S^y(T^y)$. Once again, by construction, the same goes for $\tilde{S}^x$ and $\tilde{S}^y$. In both cases, a random site z_1 is added to the aggregate.
  - For E, we use the trajectory of S^y, which adds the random site z_2 to the aggregate.
  - For $\tilde{E}$, we use the trajectory of $\tilde{S}^x$ to determine the random site to add. However with our coupling, the trajectory of $\tilde{S}^x$ after exiting A is the same as the trajectory of S^y after exiting A. Hence, the same site z_2 is also added to the aggregate.

In both cases, we have that the $E = \tilde{E} = A \cup \{z_1\} \cup \{z_2\}$. Now, since $\tilde{S}^x$ and $\tilde{S}^y$ are independent simple symmetric random walks issued respectively from x and y, we have that

$$(A \oplus \{x\}) \oplus \{y\} \overset{\text{law}}{=} (A \oplus \{y\}) \oplus \{x\}.$$

□

Note that the Abelian Property only holds *in law*, and it is *wrong* to say that the aggregates obtained by switching the order of the z_i's are almost surely the same. We provide a short example in $\mathbb{Z}^2$ to show this. Let $x = (0,1)$, $y = (0,-1)$ and let $A = \{x,\ y\}$. Let S^x and S^y denote two independent simple symmetric random walks on $\mathbb{Z}^2$. For this example, let us assume that the trajectory of S^x and S^y are such that $S^x = (\downarrow, \rightarrow, \dots)$, and $S^y = (\uparrow, \uparrow, \leftarrow, \dots)$.

Computing $(A \oplus \{x\}) \oplus \{y\}$**:** Let us begin by computing $A \oplus \{x\}$. Since $A \cap \{x\} = \{x\}$, we launch a particle from x using the trajectory of S^x. The first site this particle visits outside of $A \cup \{x\}$ is $(0,0)$, which is then added to the aggregate. We get that $A \oplus \{x\} = \{x,\ y,\ (0,0)\}$. Now, let us compute $(A \oplus \{x\}) \oplus \{y\}$. Since $(A \oplus \{x\}) \cap \{y\} = \{y\}$, we launch a particle from y using the trajectory of S^y. This particle begins by visiting $(0,0)$ (which is already in the aggregate), then visits x (also in the aggregate), and visits $(-1,1)$, which is added to the current aggregate. We get that $(A \oplus \{x\}) \oplus \{y\} = \{x,\ y,\ (0,0),\ (-1,1)\}$.

Computing $(A \oplus \{y\}) \oplus \{x\}$**:** We begin by computing $A \oplus \{y\}$. Since $A \cap \{y\} = \{y\}$, we launch a particle from y using the trajectory of S^y. The first site this particle visits outside of $A \cup \{y\}$ is $(0,0)$, which is then added to the aggregate. We get that $A \oplus \{y\} = \{x,\ y,\ (0,0)\}$. Now, let us compute $(A \oplus \{y\}) \oplus \{x\}$. Since $(A \oplus \{y\}) \cap \{x\} = \{x\}$, we launch a particle from x using the trajectory of S^x. This particle begins by visiting $(0,0)$ (which is already in the aggregate), then visits $(1,0)$, which is added to the current aggregate. We get that $(A \oplus \{y\}) \oplus \{x\} = \{x,\ y,\ (0,0),\ (1,0)\}$.

In the end, we obtain that $(A \oplus \{x\}) \oplus \{y\} \neq (A \oplus \{y\}) \oplus \{x\}$. Figures 1.1 and 1.2 illustrate the previous example. The aggregate is represented with blue dots, while the random walk is represented with a small red dot.

Definition of the IDLA model

For $n \geq 0$, we can define the standard IDLA aggregate $A(n)$ as the repetition of n smash sums between $\emptyset$ and $\{0\}$, seen as

$$A(n) = \underbrace{((\emptyset \oplus \{0\}) \oplus \{0\}) \oplus \cdots \oplus \{0\}}_{n \text{ times}}.$$

In other words, we let $A(0) = \emptyset$, and for any $n \geq 1$, given a realization of $A(n-1)$, we define $A(n)$ as

$$A(n) = A(n-1) \cup \{S(\tau_{n-1})\},$$

where S denotes a simple symmetric random walk on $\mathbb{Z}^d$ issued at 0, independent of $A(n-1)$, and τ_{n-1} is the stopping time

$$\tau_{n-1} := \inf \{t \geq 0,\ S(t) \notin A(n-1)\}.$$

At each step, a new random walk, independent of the previous ones, is sent from the origin and stopped once it exits the aggregate. We give a few realizations of $A(n)$ with increasing values of n ($n \in \{10,\ 100,\ 1000, 10000\}$) in Figure 1.3. One could conjecture with these simulations, that as the number of emitted particles n grows larger, the shape of the aggregate $A(n)$ resembles a Euclidean ball. This is in fact true, and constitutes the first shape theorem provided for IDLA.

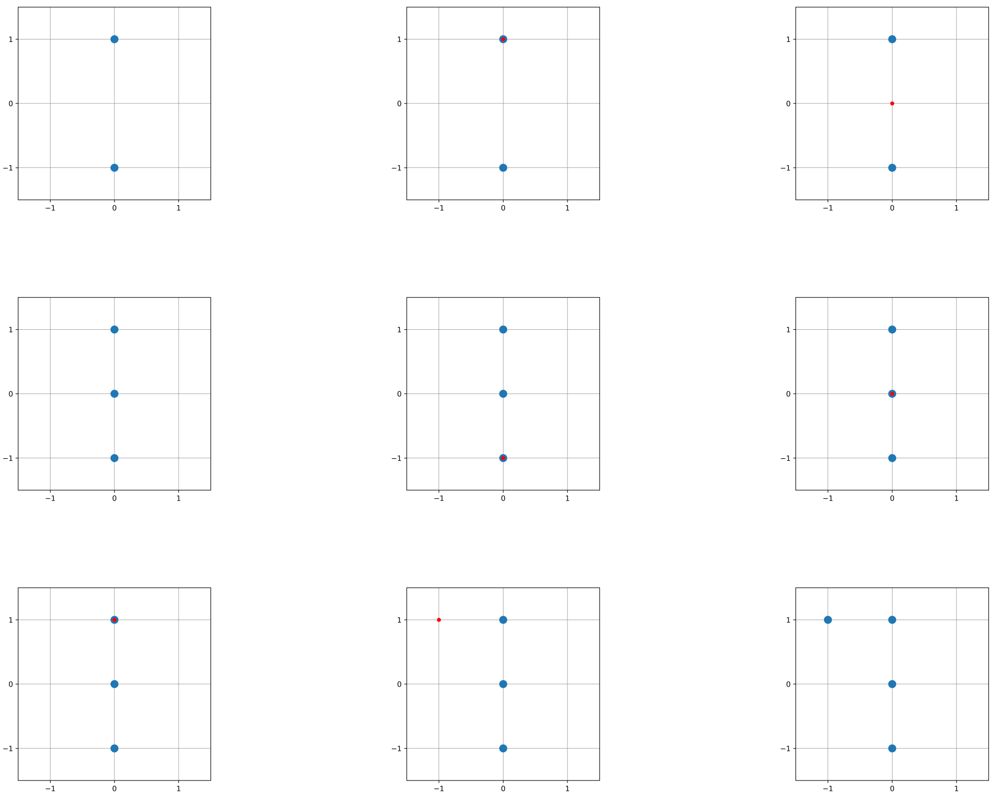


Figure 1.1: A realization of $(A \oplus \{x\}) \oplus \{y\}$

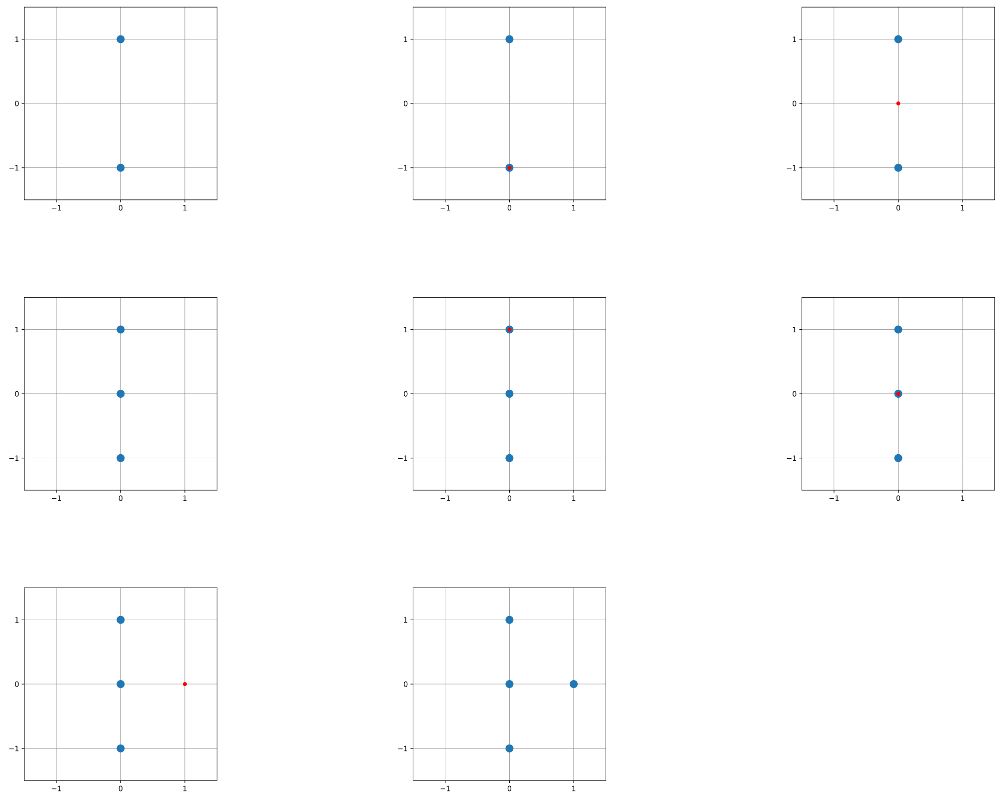
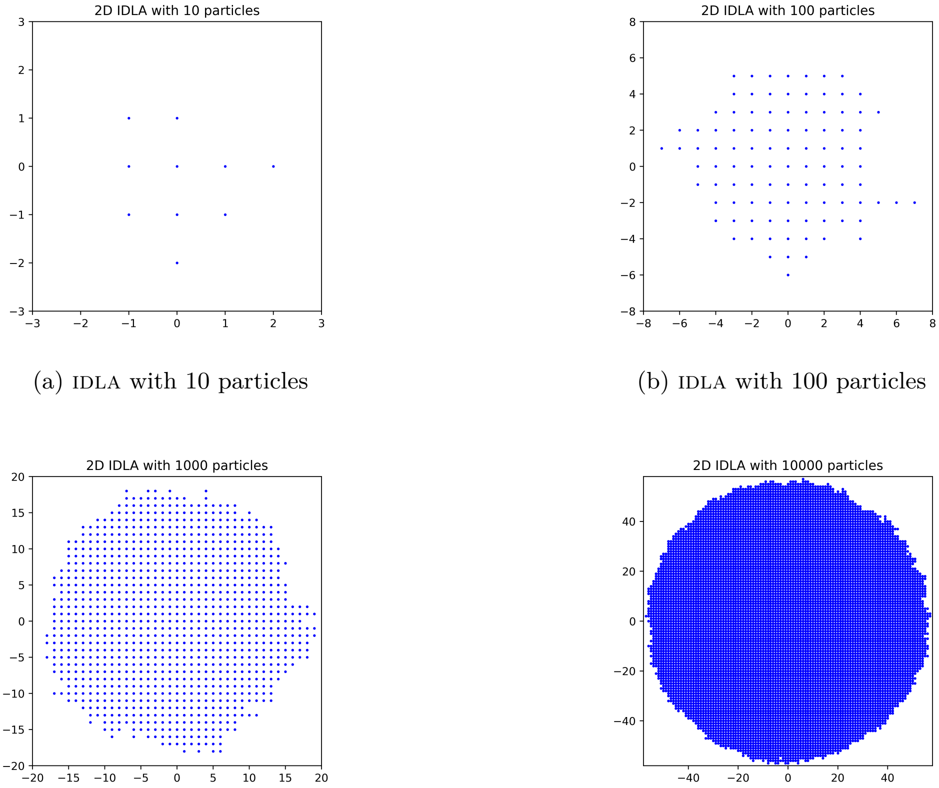


Figure 1.2: A realization of $(A \oplus \{y\}) \oplus \{x\}$



(a) IDLA with 10 particles

(b) IDLA with 100 particles

(c) IDLA with 1000 particles

(d) IDLA with 10000 particles

Figure 1.3: Realizations of 2-dimensional IDLA, with 10, 100, 1000 and 10000 particles

1.3 Shape theorems

Lawler, Bramson and Griffeath are the first to show in [34] a shape theorem in dimensions $d \geq 2$ for classical IDLA. They show that the aggregate, when suitably renormalized, has the same limit shape as a Euclidean ball.

Theorem 1.2. *(Shape theorem) Let ω_d denote the volume of the d-dimensional Euclidean ball of radius 1 in $\mathbb{R}^d$. For any $\varepsilon > 0$, almost surely, for n sufficiently large,*

$$\mathbb{B}(n(1-\varepsilon)) \subseteq A(\lfloor \omega_d n^d \rfloor) \subseteq \mathbb{B}(n(1+\varepsilon)),$$

where $\mathbb{B}(r) := \{x \in \mathbb{Z}^d, \ \|x\|_2 < r\}$ is the Euclidean ball of radius r.

In all that follows, in an effort to lighten notation, we will omit the use of floor functions and write $\omega_d n^d$ rather than $\lfloor \omega_d n^d \rfloor$.

Let us first comment on this result. At first glance, it might seem surprising to obtain a Euclidean ball as the limiting shape, given that the space we are working with is the discrete lattice space $\mathbb{Z}^d$. However, this result is moral, as we know that a random walk on $\mathbb{Z}^d$, properly normalized, converges in distribution to a Brownian motion in $\mathbb{R}^d$. Since Brownian motion is isotropic, it is rational that the limiting shape is that of a Euclidean ball.

The proof of Theorem 1.2 provides classical ingredients found in numerous other proofs of shape theorems for various models of IDLA. Indeed, most proofs of shape theorems tend to follow the same outline: they begin with the proof of the lower bound, which heavily uses the Abelian Property, and then the proof of the upper bound, as the proof of the upper bound tends to use the result from the lower bound. In [34], the strategy to prove the upper bound is to say that a significant proportion of particles have already been used to fill in the lower bound, leaving too few particles to extend over the upper bound. It then remains to control the rate at which the border of the aggregate grows, which the authors show through the study of a branching process.

Let us comment on the strategy employed in the proof of the lower bound. If a particle launched from the origin reaches $\mathbb{B}(n(1-\varepsilon))$ before it exits the aggregate, we essentially pause its trajectory, and continue launching particles from the origin. This gives us a transitional aggregate, which is included in $\mathbb{B}(n(1-\varepsilon))$ by construction. Once all our particles have been launched, we resume the trajectory of the ones we had stopped at $\mathbb{B}(n(1-\varepsilon))$ until they exit the current aggregate. Thanks to the Abelian Property, we know that this final aggregate has the same law as $A(\omega n^d)$. The authors obtain results on the transitional aggregate, from which they deduce properties on $A(\omega n^d)$. They use a clever decomposition to count particles that reach $\mathbb{B}(n(1-\varepsilon))$, which we briefly detail.

Consider a site $z \in \mathbb{B}(n(1-\varepsilon))$ and define the random variables $M,\ N$ and L as follows:

N = number of *particles* that visit z before leaving the aggregate,
M = number of *walks* that visit z before leaving $\mathbb{B}(n)$,
L = number of *walks* that visit z before leaving $\mathbb{B}(n)$, but after the associated particle has settled.

Then we have
$$M \leq N + L. \tag{1}$$
Indeed, consider a walk counted by M. Either its associated particle visits z before exiting the aggregate, in which case it is counted by N, or the walk visits z before leaving $\mathbb{B}(n)$ but after the associated particle settled. In that case, it is counted by L. (1) is an inequality, and not an equality, because N can contain particles that visited z after exiting $\mathbb{B}(n)$, which wouldn't be counted by M. By writing $N \geq M - L$, we can control N, which counts *particles*, using M and L, which both count *walks*. Note that M is completely independent of the aggregate, so is easy to control using known results from random walk theory. To control L, more work is required. Firstly, note that L only counts walks that reach z *after* the associated particle has settled and *before* exiting $\mathbb{B}(n)$. We can write L as a sum of indicators of this event, where the sum is taken over all $\omega_d n^d$ walks. However, since L only counts walks that reach z *after* the associated particle has settled, we can start counting these walks based on the site on which each particle exits the aggregate, which each time is a unique site of $\mathbb{B}(n)$. Hence, rather than summing over the number of emitted walks, we can enlarge the set by summing over all sites of $y \in \mathbb{B}(n)$ and by considering walks that start in y and visit z before exiting $\mathbb{B}(n)$. This gives us a random variable $\tilde{L}$ which is such that $L \leq \tilde{L}$ *in law*. Additionally, the walks counted by $\tilde{L}$ have the advantage of being independent, meaning $\tilde{L}$ can be written as the sum of independent random variables, making it easy to control. This clever way of counting particles of N using M and L is recurrent in many other shape theorem proofs. Through this reasoning, we can control particles, which are by nature highly dependent of each other and thus difficult to study, by working with sums of independent variables, which are very much easier to handle. Using known results on estimates of Green functions, they effectively control these sums of random walks.

Once the limiting shape of the aggregate was proved, it became interesting to study the fluctuations around this shape. To do so, we define the inner error $\delta_I(n)$ and the outer error

$\delta_O(n)$ such that

$$n - \delta_I(n) = \sup\{r \geq 0,\ \mathbb{B}(r) \subset A(\omega_d n^d)\}, \quad n + \delta_0(n) = \inf\{r \geq 0,\ A(\omega_d n^d) \subset \mathbb{B}(r)\}.$$

A first bound on these error terms is due to Lawler in [33], who showed that they were subdiffusive, in the sense that

$$\delta_I(n) = o\left(n^{1/3}\log^2 n\right),$$
$$\delta_0(n) = o\left(n^{1/3}\log^4 n\right).$$

These bounds were then improved by Asselah and Gaudillière in [1] to obtain logarithmic bounds. They show that:

$$\delta_I(n) = \mathcal{O}\left(\log n\right),$$
$$\delta_0(n) = \mathcal{O}\left(\log^2 n\right).$$

Quickly thereafter, they improved their bound for the outer error, (see [4]), while improving both bounds in dimensions $d \geq 3$, obtaining sublogarithmic fluctuations. At around the same time, Jerison, Levine and Sheffield publish in [31] the same result using a different approach. Both show that in dimension $d = 2$,

$$\delta_I(n) = \mathcal{O}\left(\log n\right),$$
$$\delta_O(n) = \mathcal{O}\left(\log n\right).$$

In higher dimensions, the fluctuations are sublogarithmic in the sense that:

$$\delta_I(n) = \mathcal{O}\left(\sqrt{\log n}\right),$$
$$\delta_O(n) = \mathcal{O}\left(\sqrt{\log n}\right).$$

The difference in fluctuations between the 2-dimensional case and higher dimensions stems from the fact that random walks are recurrent when $d = 2$ and transient once $d \geq 3$. Green function estimates are hence different in each case, and result in sharper bounds when $d \geq 3$.

The methods used to prove these fluctuations are very similar to the ones we detailed for the proof of Lawler, Bramson and Griffeath's initial shape theorem in [34]. We briefly detail the approach of Asselah and Gaudillière, as similar methods will be used to establish a shape theorem in the context of our multi-source model. In [4], Asselah and Gaudillière use a method of exploration by waves relying on the Abelian Property. The space outside $\mathbb{B}(n)$ is partitioned in shells, while each shell is partitioned in cells. When particles reach the border of a shell, they are paused, and a new particle is launched from the origin. This causes particles to temporarily accumulate on the surface of shells. Once all particles have been sent, the paused particles resume the trajectory. The authors show that if many particles are at the surface of a shell, they are very likely to fill the corresponding cell. This 'exploration by waves' is analogous to a dam holding water. Particles accumulating at the border of the shell are similar to water being stored by a dam, and once there are enough particles, the floodgates are opened, releasing the particles inside the corresponding cell. This strategy relies once again on the ability to pause particles midway without changing the law of the final aggregate, ensured by the Abelian Property. An illustration of cells and shells is given in Figure 1.4.

Additionally, Asselah and Gaudillière use a decomposition similar to the one in (1) to control the number of particles hitting a shell. They exploit an independence property overlooked by Lawler, Bramson and Griffeath to obtain finer estimates of N. As of today, these bounds are the

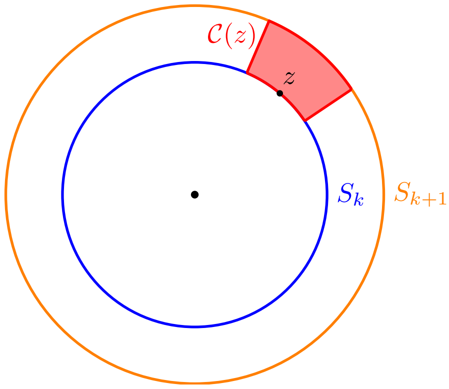


Figure 1.4: A representation of shells and cells. The shells are denoted by S_k and S_{k+1}, and the cell of z is denoted by $\mathcal{C}(z)$ (hatched in red). Exploration by waves claims that if many particles accumulate on the surface of the cell, i.e on $\mathcal{C}(z) \cap S_k$, then the cell $\mathcal{C}(z)$ is likely to be filled.

sharpest known to us for classical IDLA. Asselah and Gaudillière show in [2] that these bounds are actually optimal when $d \geq 3$. However, it has not yet been determined whether these bounds are optimal when $d = 2$.

1.4 Related models

In this section, we give a brief overview of variants and related models of the classical IDLA model. We present a few of these models as well as existing results for them. We purposely stay vague concerning the outline of each proof, as the main idea of this section is rather to give the reader an idea of the usual ingredients used for proofs in IDLA.

Cylindrical IDLA

In this paragraph, we describe some results of IDLA on cylinder graphs. For $N \geq 0$, we define the cylinder graph in $\mathbb{Z}^2$ as $\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}$. We adopt the notation of [29, 37] and write $Z_N := \mathbb{Z}/N\mathbb{Z}$. One can look at IDLA on cylinder graphs, where particles are launched uniformly from $Z_N \times \{0\}$. Just like the classical IDLA aggregate was logarithmically close to a Euclidean ball, the IDLA aggregate on $Z_N \times \mathbb{Z}$ is logarithmically close to a rectangle. Indeed, Levine and Silvestri show in [37] that for any $\gamma > 0,\ m \in \mathbb{N}$, there exists a constant $b_{\gamma,m}$ such that

$$\mathbb{P}\left(R_{\frac{t}{N}-b_{\gamma,m}\log N} \subset A(t) \subset R_{\frac{t}{N}+b_{\gamma,m}\log N},\ \forall t \leq N^m\right) \geq 1 - N^{-\gamma},$$

with $R_k := \{(x, y) : y \leq k\}$, and where $A(t)$ denotes the cylindrical IDLA aggregate obtained after launching t particles uniformly from $Z_N \times \{0\}$. This result is an improvement of the result of [29], where a similar result had been shown but for times $t = N^2$. Here, Levine and Silvestri extend this result to large times, up to $t \leq N^m$. In the same paper, the authors also examine the question of how long it takes for the aggregate to reach a rectangular shape, given it started in some initial configuration. In other words, when the IDLA process does not start from flat, how long does it take for it to forget its initial shape? This obviously depends on the initial configuration, and the authors study cases where the starting position is a 'typical initial' configuration, which is an expected configuration cylindrical IDLA can reach. Through control of the height and a clever coupling for aggregates, they show that with high probability, the process forgets its initial state in $\mathcal{O}(N^2 \log N)$ steps. The Abelian Property is once again central in the proof of this result, as it allows the authors to create their 'water coupling' to build two aggregates equal in law. This water coupling requires particles to be launched, then frozen before being launched again, which we have seen by now is a classical reasoning for IDLA.

IDLA on supercritical percolation graphs

In this paragraph, we present results concerning IDLA on the infinite cluster of supercritical percolation on $\mathbb{Z}^d$. Let us begin by presenting the Bernoulli bond percolation model on $\mathbb{Z}^d$. For $p \in [0,1]$, each edge of $\mathbb{Z}^d$ is *open* (kept in the graph) with probability p and *closed* (removed from the graph) with probability $1-p$, independently of all other edges. When $p = 0$, the graph obtained is trivially reduced to $\{0\}$, while the graph obtained when $p = 1$ is $\mathbb{Z}^d$ entirely. However, for $p \in \,]0, 1[$, the percolation configuration obtained is less clear. One of the main interests for percolation models is the existence of an infinite cluster. For $x \in \mathbb{Z}^d$, we can define the cluster of x as reunion of all reachable vertices from x using open edges of the graph. We write this cluster as $C(x) := \left\{y \in \mathbb{Z}^d,\ x \leftrightarrow y\right\}$. The existence of an infinite cluster can thus be written as the event $\{|C(x)| = \infty\}$, or also $\{x \leftrightarrow \infty\}$. By stationarity of $\mathbb{Z}^d$, this event does not depend on x, and we can define the probability of percolation as

$$\theta_d(p) := \mathbb{P}_p\left(|C(0)| = \infty\right),$$

where $\mathbb{P}_p$ denotes the percolation measure of parameter p. Additionally, define the critical percolation parameter $p_c(d)$ as:

$$\inf\left\{p \in [0,1],\ \theta_d(p) > 0\right\}.$$

We know that $p_c(1) = 1$ and $p_c(2) = \frac{1}{2}$ (due to Kesten). However, in higher dimensions, exact values of $p_c(d)$ are not yet known. Another interesting question is the value of $\theta_d(p_c(d))$. We know that $\theta_1(1) = 1$, while $\theta_2(\frac{1}{2}) = 0$. For $d \geq 3$ however, behavior at the critical parameter is unknown.

IDLA on supercritical percolation graphs is studied in [20]. The authors establish a shape theorem on percolation graphs on $\mathbb{Z}^d$, in the supercritical case $p > p_c(d)$. More precisely, let ω denote the infinite supercritical percolation cluster on $\mathbb{Z}^d$ conditioned on containing the origin, and let $b(n) := |\omega \cap \mathbb{B}(n)|$. Now, denote by $\tilde{A}(b(n))$ the IDLA aggregate on ω generated by launching $b(n)$ particles from the origin. Then, for any $\varepsilon > 0$ and almost surely any ω, given n is large enough, we have:

$$\mathbb{B}(n(1-\varepsilon)) \cap \omega \subset \tilde{A}(b(n)) \subset \mathbb{B}(n(1+\varepsilon)) \cap \omega.$$

Once again, the strategy to prove this shape theorem follows the same outline as before: the authors begin by proving the lower bound and then the upper bound, as the latter uses the former. The authors actually provide a general method in which the existence of a lower bound verifying 'good' regularity conditions automatically provides a technique to prove an upper bound. They prove this result for general graphs under these regularity conditions, and show that in particular, the supercritical percolation graph in $\mathbb{Z}^d$ verifies these conditions.

Sandpile models

In this paragraph, we define a similar model to IDLA known as the sandpile model. This model is often linked to IDLA as its protocol also happens to exude an Abelian Property. The protocol is as follows: consider the lattice $\mathbb{Z}^d$, where each site contains a certain number of grains of sand, making a sandpile. When there are more than $2d$ grains on one site, the stack of sand collapses, granting each of its $2d$ neighbors exactly one grain of sand. This step is repeated as long as sites with stacks of sand with more than $2d$ grains remain. The classical sandpile model consists of having $n,\ n \gg 2d$ grains stacked at the origin and letting them fall. In this case, if we call S_n the set of all sites visited during the process, Levine and Peres show in [36] that for any $\varepsilon > 0$, there exist positive constants $c_1,\ c_2$ depending only on d and $c_1',\ c_2'$ depending only on d and ε such that for any $r > 0$,

$$\mathbb{B}(c_1 r - c_2) \subset S_n \subset \mathbb{B}(c_1' r + c_2'),$$

with $n = \omega_d r^d$. Just like standard IDLA, the sandpile model exhibits an Abelian Property. In this case, it has been shown that the final configuration of the sandpile is independent of the order in which we topple the sites.

A continuous version of this model is the divisible sandpile model. In this model, rather than considering integer-valued grains for each site, we now consider that each site has a given mass, and each site topples part of its mass to its neighbors whenever this mass is strictly greater than 1. In this case, say a site x has mass $m > 1$, then each of its $2d$ neighbors receives a mass equal to $(m-1)/2d$, leaving x with mass equal to 1. In that case, x is stable and no longer topples any mass. Now, consider an initial configuration of mass on $\mathbb{Z}^d$. Levine and Peres show in [36] that when the number of topplings tends to infinity, the initial configuration will converge towards a final configuration, provided each unstable site topples infinitely many times. Here again, we can show that this final configuration does not depend on the order in which the sites topple. They show a shape theorem when the initial configuration consists of a mass $\omega_d n^d$ at the origin, which states that there exist constants $c,\ c'$ depending only on d such that

$$\mathbb{B}(n - c) \subset D(\omega_d n^d) \subset \mathbb{B}(n + c'),$$

where $D(\omega_d n^d)$ is the set of sites of mass 1 in the final configuration of the divisible sandpile model started with mass $\omega_d n^d$ at the origin.

Rotor-router model

We present a model tightly linked with IDLA called the *rotor-router* model. One can consider it as a deterministic version of IDLA. This model was first studied in [41], initially under the name of 'Eulerian walkers'. The model is as follows. Consider $\mathbb{Z}^2$, where we have positioned at each site a rotor pointing towards a direction: north, east, south or west. We launch a particle from the origin, and at each time step, the rotor on the site of the *active* particle is shifted 90 degrees clockwise. The particle then moves in the direction of the newly oriented rotor, until exiting the aggregate. This model can be extended to $\mathbb{Z}^d$, provided that we have a predisposed order of the $2d$ directions to cycle through. A shape theorem for the rotor-router model exists, similar to the shape theorem of classical IDLA: the result states that given any starting configuration of rotors on $\mathbb{Z}^d$, the aggregate obtained when launching $\omega_d n^d$ particles converges to the Euclidean ball of radius n. We also have a result on the fluctuations around this ball. More precisely, let us denote by $\tilde{A}(n)$ the rotor-router aggregate after launching n particles. Then there exist positive constants $c,\ c'$ depending only on d such that

$$\mathbb{B}(n - c\log n) \subset \tilde{A}(\omega_d n^d) \subset \mathbb{B}(n(1 + c' n^{-1/d} \log n)).$$

We provide a simulation of the router rotor model in Figure 1.5.

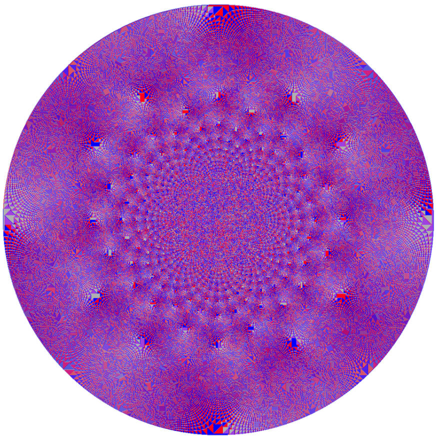


Figure 1.5: Simulation on $\mathbb{Z}^2$ of the rotor-router model with 1 million particles. Each color corresponds to a direction of a router. (Simulation: Y. Peres [36])

IDLA on Sierpinski gaskets

IDLA has also been studied on a special class of graphs, known as Sierpinski gaskets. Let us first detail the construction of this graph. Let

$$V_0 = \left\{(0,0),\ (1,0),\ (1/2, \sqrt{3}/2)\right\},$$

and

$$E_0 = \left\{\{(0,0),\ (1,0)\},\ \{(0,0),\ (1/2,\ \sqrt{3}/2)\},\ \{(1,0),\ (1/2,\ \sqrt{3}/2)\}\right\}.$$

Then, for $n \geq 0$, given (V_n, E_n), define recursively

$$V_{n+1} = V_n \cup \{(2^n, 0) + V_n\} \cup \left\{\left(2^{n-1}, 2^{n-1}\sqrt{3}\right) + V_n\right\},$$

and

$$E_{n+1} = E_n \cup \{(2^n, 0) + E_n\} \cup \left\{\left(2^{n-1}, 2^{n-1}\sqrt{3}\right) + E_n\right\},$$

where $(x, y) + A := \{(x, y) + a,\ a \in A\}$. Now, let $V_\infty := \bigcup_{n\geq 0} V_n$ and $V_\infty := \bigcup_{n\geq 0} V_n$. Finally, define $V := V_\infty \cup \{-V_\infty\}$, and $E := E_\infty \cup \{-E_\infty\}$. The graph SG $= (V, E)$ is called the *doubly infinite Sierpinski gasket graph*. An illustration of this graph is given in Figure 1.6.

We remind the reader about some concepts of graph theory. Consider a graph $G = (V, E)$. For $x,\ y \in V$, we say that x is connected to y if there exists a sequence of vertices $x = v_0, v_1, \ldots, v_k = y$ such that $(v_i, v_{i+1}) \in E$ for all $i \geq 0$. In that case, we write $x \sim y$, and we say that $(v_0, \ldots, v_k)$ is a path of length k between x and y. Now, we can define the graph distance $d_G(x, y)$ as the length of the shortest path between x and y. If no such path exists, that is $x \nsim y$, then we set $d_G(x, y) = +\infty$.

Let $o = (0, 0)$ denote the origin of the Sierpinski gasket graph SG, and let $\mathcal{S}_0 = \{o\}$. For $n \geq 1$, let $\mathcal{S}_n$ denote the IDLA cluster obtained from $\mathcal{S}_{n-1}$ after launching a particle from o, and adding the first site it visits outside of $\mathcal{S}_{n-1}$. Chen, Huss, Sava-Huss and Teplyaev show in [13] a shape theorem for IDLA on SG. Denote by $B(n) := \{x \in V,\ d_{\mathrm{SG}}(o, x) < n\}$, and let $b_n := |B(n)|$.

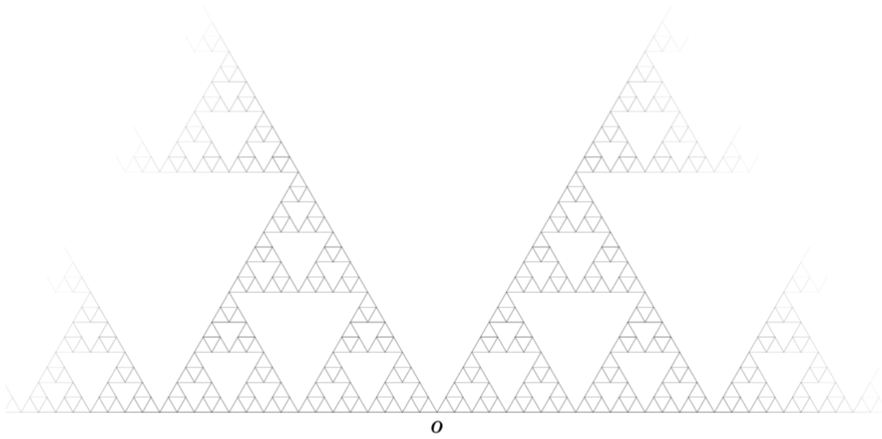


Figure 1.6: The doubly infinite Sierpinski gasket graph. The origin is denoted by o. (Figure: E. Sava-Huss)

Then, for all $\varepsilon > 0$, almost surely, given n is large enough, we have :

$$B(n(1-\varepsilon)) \subset \mathcal{S}(b_n) \subset B(n(1+\varepsilon)).$$

Once again, the proof of this result uses classical ingredients of IDLA. The proof begins with the proof of the inner bound, using a decomposition similar to the one in (1), and using results on Green functions on SG. Surprisingly enough, they obtain bounds for Green functions on SG using known results from the divisible sandpile on SG. Although this model is strictly deterministic, it shares many properties with the IDLA model on SG. Once the inner bound is proved, the outer bound is proved by arguing that only a few particles remain after filling up the inner bound. Such an argument has already been used in [34], for example. This requires additional work on Green functions on the gasket SG, as well as ensuring that SG satisfies satisfactory conditions.

Recent work on this model has led to precision about the fluctuations of the cluster around its limiting shape. Heizmann showed in [24] that almost surely, for any $\kappa > 0$, there exists a constant $c > 0$ such that, given n is large enough, we have:

$$B(n - cn^{1/2}\log(n)^{(1+\kappa)/2\alpha}) \subset \mathcal{S}_n \subset B(n - cn^{1/2+1/2\alpha}\log(n)^{(1-1/\alpha)(1+\kappa)/2\alpha}),$$

where $\alpha = \frac{\log 3}{\log 2}$ is the Hausdorff dimension of the Sierpinski gasket.

External DLA

Before the introduction of IDLA by Meakin and Deutch in [39], Witten and Sander introduced in [49] the model of *external* DLA. Just like IDLA, external DLA is a model used to grow random sets $(E_n)_{n\geq 0}$ on a locally finite and infinite graph G. Fix $o \in G$ as the origin of the graph. We start off with $E_0 = \{o\}$. Then, at each time step, we launch a random walk 'from infinity', and stop it once it hits the outer border of the aggregate. The site on which it stops is then added to the aggregate, and we start over by throwing a new random walk, independent of the previous walks, 'from infinity' again. We will explain in detail what we mean by launching 'from infinity'. For now, the reader can understand this as launching a walk far away from the cluster. As opposed to IDLA, the aggregates obtained by external DLA are far less well understood, and are far less regular than the usual smooth clusters of IDLA. Contrary to IDLA, external DLA does not exhibit an Abelian Property, making it a much harder model to study. A realization of external DLA on $\mathbb{Z}^2$ is shown in Figure 1.7.

Let us properly define the model by detailing what we mean by launching a random walk 'from infinity'. This is a difficult concept to generalize for any graph G, as it depends if the walk is transient or recurrent on G. Let us begin by defining a few concepts of random walks on a graph G. Let S denote a random walk on G, and let $A \subset G$. We define the stopping time $\tau(A)$ as $\tau(A) := \inf\{n \geq 0,\ S_n \in A\}$. Now, for a random walk S starting on some vertex $x \in \mathbb{Z}^d$, we define the hitting distribution $H_A(x, y)$ as

$$H_A(x, y) = \mathbb{P}_x \left(S_{\tau(A)} = y \right), \ \forall y \in A.$$

We wish to define μ_A, the harmonic measure of A *from infinity* by taking $|x| \to \infty$ in $H_A(x, \cdot)$. We must be cautious however, as this limit isn't always well defined, and won't systematically define a probability measure on A. When the random walk is recurrent on G, μ_A is generally well defined, but issues arise when the walk is transient on G. For sake of simplicity, we define external DLA on $\mathbb{Z}^d$, where harmonic measures *from infinity* are easy to define. For this, we must consider the recurrent case $d = 2$ and the transient case $d \geq 3$.

The case of $\mathbb{Z}^2$: In the case of $\mathbb{Z}^2$, the random walk is recurrent, and $\mu_A(y) := \lim_{|x| \to \infty} H_A(x, y)$ is well defined, and defines a probability distribution on A. In that case, we define the external DLA cluster E_{n+1} as $E_{n+1} = E_n \cup \{y_{n+1}\}$, where y_{n+1} is chosen according to the distribution $\mu_{\partial E_n}$, where ∂E_n denotes the border of E_n.

The case $d \geq 3$: In the case of $d \geq 3$, one cannot define the harmonic measure using the same reasoning as above because the random walk is transient, and the quantities $\lim_{|x| \to \infty} H_A(x, y)$ are identically zero for any $y \in A$. To overcome this, we must condition the walk on hitting the set A. In that case, we define

$$\mu_A(y) := \lim_{|x| \to \infty} \mathbb{P}_x \left(S_{\tau(A)} = y \mid \tau(A) < \infty \right) = \lim_{|x| \to \infty} \frac{H_A(x, y)}{\sum_{z \in B} H_A(x, z)},$$

which verifies $\sum_{y \in A} \mu_A(y) = 1$. With this definition, we define the cluster E_{n+1} by adding a site y_{n+1} to E_n, according to the distribution $\mu_{\partial E_n}$.

The infinite external DLA cluster is then defined as

$$E_\infty := \bigcup_{n \geq 0} \uparrow E_n.$$

We briefly present some results and questions for external DLA. One of the main interests is the length of 'arms' in this model. Define $r_n := \max\{|x|,\ x \in E_n\}$ the radius of E_n. In the case where $G = \mathbb{Z}^d$, Kesten showed in [32] that there exist constants $c(d)$ depending only on d such that almost surely,

$$\limsup_{n \to \infty} \frac{r_n}{n^{2/3}} \leq c(2),\ d = 2,$$

$$\limsup_{n \to \infty} \frac{r_n}{n^{2/d}} \leq c(d),\ d \geq 3.$$

This question has since been studied on different graphs. Notably, Procaccia et al. (see [42],

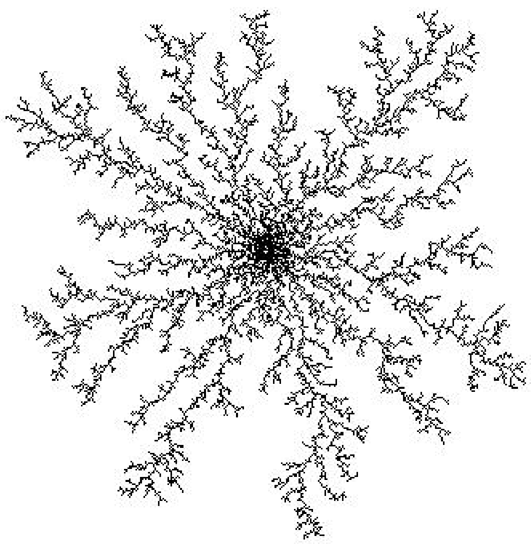


Figure 1.7: A realization of external DLA on $\mathbb{Z}^2$ (Simulation: E. Sava-Huss [46])

[44], [43], [40]) study this question for DLA along long line segments, or in the upper half plane of $\mathbb{Z}^2$. Additionally, Benjamini and Yadin introduce in [9] a robust method that applies to a large class of graphs, such as transitive graphs of polynomial growth, transitive graphs of exponential growth, non-amenable graphs, super-critical percolation clusters on $\mathbb{Z}^d$, and high dimensional pre-Sierpinski carpets. In particular, they improve Kesten's growth rate on $\mathbb{Z}^3$ from $n^{2/3}$ to $\sqrt{n \log n}$. These results are featured in [9], a pre-print from 2017, which, to our knowledge, hasn't been published yet. Readers should be aware that these results haven't been peer-reviewed, so some caution is warranted.

Actually, the growth rate of r_n on $\mathbb{Z}^d$ is conjectured to be of order $n^{1/d}$, and remains an open problem since the introduction of the model. The conjecture is stated as such :

$$\limsup_{n\to\infty} \frac{\mathbb{E}\left[r_n\right]}{n^{1/d}} < +\infty.$$

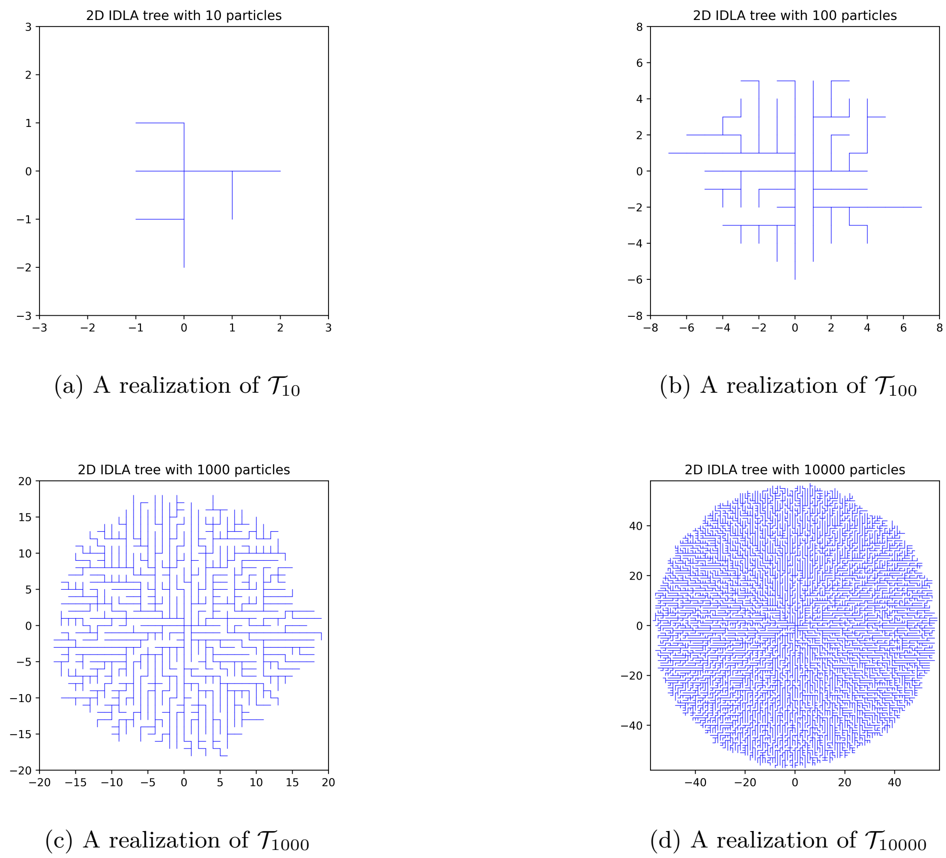
Other open questions deal with the number of ends of the infinite cluster E_∞, or the number of occupied vertices inside a ball centered on the origin of the graph. There are many more open questions surrounding external DLA, showing just how difficult this model is to study. Notably, one of the powerful tools present in IDLA and not in external DLA is the Abelian Property. The absence of such a tool for external DLA, along with the difficulties of properly defining particles 'from infinity' make it a much more difficult model to study than IDLA.

1.5 The IDLA Tree

In this paragraph, we present a random tree on $\mathbb{Z}^d$, rooted in 0, which we can naturally build using the classical IDLA protocol, simply by considering the edge from which each particle exits. We begin by building a family of finite trees $(\mathcal{T}_n)_{n\geq 1}$. We denote each tree by $\mathcal{T}_n = (V_n, E_n)$, where V_n denotes its vertex set and E_n its set of edges. When $n = 1$, we let $V_1 = \{0\}$ and $E_1 = \emptyset$. Now, for any $n \geq 2$, given a realization of $\mathcal{T}_{n-1}$, let

$$V_n = V_{n-1} \cup \{S(\tau_{n-1})\}, \qquad E_n = E_{n-1} \cup \{(S(\tau_{n-1} - 1), S(\tau_{n-1}))\},$$

where $\tau_{n-1} = \inf\{t \geq 0,\ S(t) \notin V_{n-1}\}$. The vertex set of $\mathcal{T}_n$ is exactly the standard IDLA aggregate $A(n)$, while the new edge added to E_{n-1} is the edge from which the particle exits the aggregate. This construction provides a tree, since the extremity of each newly added edge is a new vertex, making it impossible to have loops. We provide a few simulations of $\mathcal{T}_n$ in Figure 1.8. The random walks used for these simulations are the same as the ones used for the aggregates simulated in Figure 1.3. By construction, it comes naturally that the family of trees $(\mathcal{T}_n)_{n\geq 1}$ is



(a) A realization of $\mathcal{T}_{10}$

(b) A realization of $\mathcal{T}_{100}$

(c) A realization of $\mathcal{T}_{1000}$

(d) A realization of $\mathcal{T}_{10000}$

Figure 1.8: Realizations of the 2-dimensional IDLA tree, with 10, 100, 1000 and 10000 particles

increasing with respect to inclusion, allowing us to define the IDLA tree $\mathcal{T}_\infty$ as :

$$\mathcal{T}_\infty := \bigcup_{n\geq 1} \uparrow \mathcal{T}_n \quad \text{a.s.}$$

Note that from Theorem 1.2, almost surely, this tree spans over all of $\mathbb{Z}^d$. Other properties concerning $\mathcal{T}_\infty$ are less clear however. In the following paragraph, we cite a few interesting questions concerning $\mathcal{T}_\infty$, and explain why they are difficult to answer.

Absence of Abelianity

In the case of IDLA aggregates, we saw that the Abelian Property was a powerful tool to prove many results, in particular to come up with *elegant couplings* between random walks. However, when constructing a graph using an IDLA protocol, that is by considering the edge through which the particle exits the aggregate, it turns out that the Abelian Property no longer holds. We provide a brief example to illustrate this point. Let $x = (0, 1)$, $y = (0, 0)$ and let $A = \{y\}$. Let $\mathcal{F}_1$, respectively $\mathcal{F}_2$, denote the graphs obtained after launching a random walk from x then y, respectively from y then x. In the context of IDLA, since no cycles or loops can occur in the graph, the graphs obtained will systematically be forests, that is, collections of disjointed trees.

In the case of $\mathcal{F}_1$, since $x \notin A$, the site x is added to the aggregate and is a root for this forest. Whatever the trajectory of S^x and S^y, $\mathcal{F}_1$ has two roots, with probability 1.

In the case of $\mathcal{F}_2$, we begin by throwing a particle from y. Since $y \in A$, this particle stays alive and moves to a neighboring site. It has probability $\frac{1}{4}$ of going to x, in which case x would be added to the aggregate, but would then *never* constitute a root for $\mathcal{F}_2$. Hence, $\mathcal{F}_2$ contains only a single tree with probability $\frac{1}{4}$, while almost surely, $\mathcal{F}_1$ contains 2 trees. Figures 1.9 and 1.10 illustrate this example, where $S^x = (\rightarrow, \dots)$ and $S^y = (\uparrow, \leftarrow, \dots)$.

Difficulties and open questions about the IDLA tree

When looking at the IDLA tree, some questions naturally come to mind, such as the existence of several infinite branches, and if so, their asymptotic direction. For an oriented tree $\mathcal{T} = (V, E)$ rooted at o, we call a branch of $\mathcal{T}$ any path started in o within $\mathcal{T}$. (See Figure 1.11 for an illustration of a branch in the IDLA tree $\mathcal{T}_{20000}$). We know that at least one of the branches of $\mathcal{T}_\infty$ is infinite, since each vertex is of finite degree. However, it is unclear if there exist several infinite branches. Is there an infinity of infinite branches, or are these finitely numbered? In the case of several infinite branches, we would be interested in knowing if each of them has some asymptotic direction, given by some vector $u \in \mathbb{Z}^d$.

Another natural question is the straightness of the branches. We would be interested in knowing if the branches of $\mathcal{T}_\infty$ could be included inside a narrow cone, which we detail here. For any $x \in \mathbb{Z}^d$ and any $\varepsilon \in [0, \pi[$, we define the cone $C(x, \varepsilon) := \{y \in \mathbb{Z}^d,\ \theta(x, y) \leq \varepsilon\}$, where θ denotes the angle (with values in $[0, \pi]$) between x and y. For any $x \in \mathbb{Z}^d$, let $\mathcal{T}_\infty^x \subset \mathcal{T}_\infty$ denote the subtree rooted at x, that is, the tree rooted at x and containing all descendants of x in $\mathcal{T}$. We say that $\mathcal{T}$ is f-*straight* if almost surely, for all $x \in \mathbb{Z}^d$,

$$\mathcal{T}_\infty^x \subset C(x, f(\|x\|)),$$

where f is a positive function such that $f(l) \to 0$ as $l \to +\infty$. This condition on f implies that as one moves away from the origin, the cone grows more narrow. Intuitively, this should mean that the branches of the tree are relatively straight.

These questions remain, as of today, without an answer. The main difficulty that arises in the study of the IDLA tree is its radial aspect. Indeed, since it is rooted at the origin, one can see on Figure 1.8 that branches near the origin are very much oriented towards the origin. This radial aspect makes it impossible to exploit any stationarity properties. Additionally, each branch is produced by many particles, adding a lot of dependence in the construction of the branches, making it quite difficult to study $\mathcal{T}_\infty$. The absence of Abelianity for trees and forests also requires a deeper understanding of the behavior of particles.

To overcome these difficulties, the authors in [16] construct an ergodic forest in $\mathbb{Z}^2$ based on an IDLA protocol, called the infinite-volume directed IDLA forest, in the hopes that this forest could ‘approach’ $\mathcal{T}_\infty$. We detail what we mean by this below. We mentioned that the IDLA tree

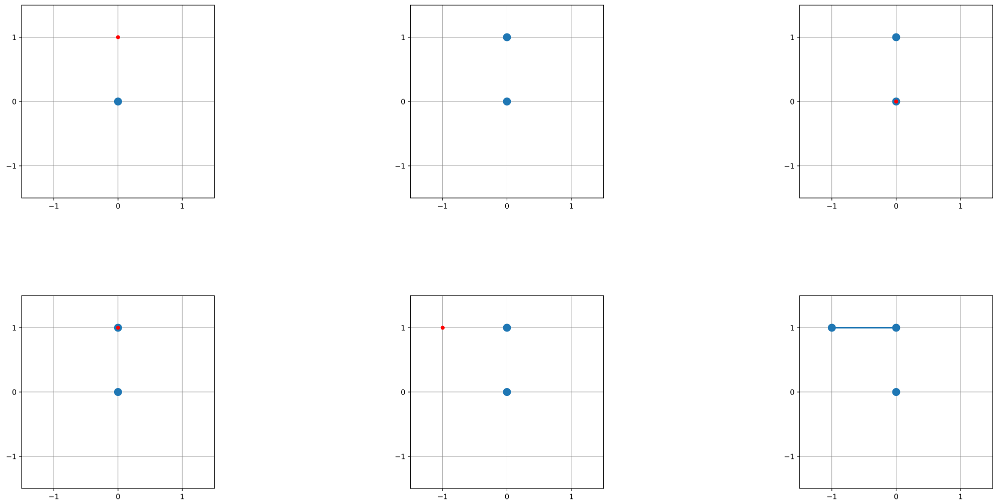


Figure 1.9: A realization of $\mathcal{F}_1$. The particle is first launched from $x = (0, 1)$, then from $y = (0, 0)$. The forest $\mathcal{F}_1$ contains two roots.

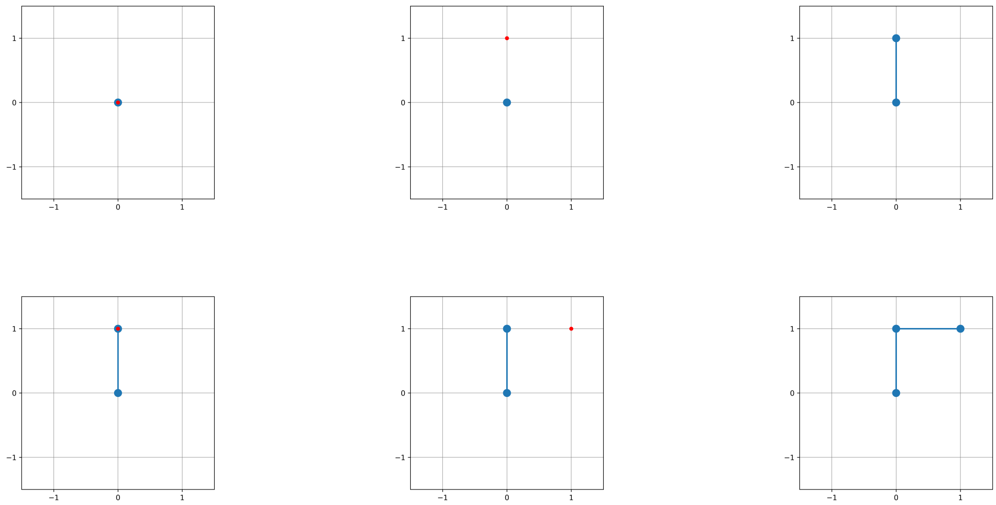


Figure 1.10: A realization of $\mathcal{F}_2$. The particle is first launched from $y = (0, 0)$, then from $x = (0, 1)$. The forest $\mathcal{F}_2$ contains only one root: it is in fact a tree.

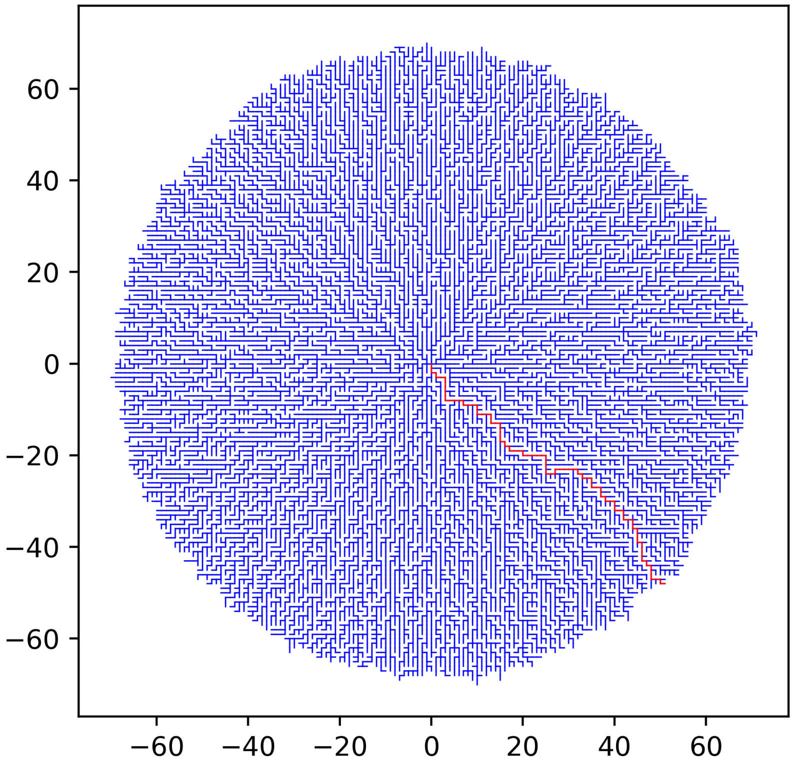


Figure 1.11: A realization of $\mathcal{T}_{20000}$. A branch of the tree is highlighted in red.

had a strong radial aspect to it, making it very difficult to study. However, simulations of the IDLA tree seem to show that as we move away from the origin, the radial aspect of the tree fades away. Therefore, we are tempted to study $\mathcal{T}_\infty$ through the ball $\mathbb{B}(n \cdot e_1, r)$, where $r > 0$ is constant and n is taken very large. Hence, $\mathcal{T}_\infty$ restricted to this ball should not exhibit a radial aspect, and should look more like a forest directed by the vector $e_1 = (1, 0, \dots, 0)$. Such a technique has been successfully used by Baccelli and Bordenave in [6] in the case of the *Radial Spanning Tree* (RST) to show straightness of the branches. In their case, they successfully approximate the RST by the *Directed Spanning Forest* (DSF). This is the motivation behind the construction of the infinite-volume directed IDLA forest of [16].

The construction of this forest is non-trivial in dimension 2, and requires a considerable amount of preliminary work on the IDLA protocol it is based upon. Additionally, the reasoning used in [16] to build this forest is exclusive to $\mathbb{Z}^2$, and fails to work for $\mathbb{Z}^d$ once $d \geq 3$. This is explained in detail in Section 1 of Chapter 3.

2 Outline of the thesis

Chapter 2 of this thesis is devoted to showing properties of a new IDLA protocol in dimension $d \geq 3$, while Chapter 3 is devoted to proving the existence of the *IDLA forest* in higher dimensions. The proof we provide in Chapter 3 also works in dimension 2, generalizing the construction of [16]. Let us first detail the IDLA protocol behind this construction.

2.1 Three families of IDLA aggregates

In this section, we detail the IDLA protocol used to build the aggregates necessary to the construction of the directed infinite-volume IDLA forest. To do so, we introduce three random aggregates, namely $A_n[\infty]$, $A_n^*[\infty]$ and $A_n^\dagger[\infty]$, built by sending particles from an infinite set of sources in $\mathbb{Z}^d$. Each aggregate is obtained by taking the limit (in space) of a family of finite aggregates $A_n[M]$ (respectively $A_n^*[M]$ and $A_n^\dagger[M]$).

Construction of $A_n[\infty]$: The infinite set of sources that we consider is the hyperplane $\mathcal{H} := \{0\} \times \mathbb{Z}^{d-1}$ of $\mathbb{Z}^d$, with $d \geq 2$. Let n, M be non-negative integers. In the sequel, exactly n particles are sent from each source. Let us now build the sequence of aggregates $(A_n[M])_{M\geq 0}$ inductively as follows. When $M = 0$, $A_n[0]$ is the classical IDLA model, i.e. with n particles emitted from the origin. Let us call *level M* the set of sources in $\mathcal{H}$ at distance M from the origin (for $\|(z_1, \dots, z_d)\| := \max_i |z_i|$). Given a realization of $A_n[M-1]$, we throw n particles from each source of level M according to the lexicographical order. So $A_n[M]$ is defined as the aggregate produced by $A_n[M-1]$ and the new sites added by particles launched at level M.

Unlike its shape, the total number of sites in $A_n[M]$ is deterministic, and equals $\#A_n[M] = n(2M+1)^{d-1}$. Besides, by construction, the sequence of aggregates $(A_n[M])_{M\geq 0}$ is almost surely increasing in the sense of inclusion, allowing us to define the limiting aggregate $A_n[\infty]$ as:

$$A_n[\infty] := \bigcup_{M\geq 0} \uparrow A_n[M] \quad \text{a.s.}$$

Construction of $A_n^*[\infty]$: Just like for $A_n[\infty]$, we begin by building a family of *finite* random aggregates $(A_n^*[M])_{M\geq 0}$. Unlike $A_n[\infty]$ however, the number of particles sent from each source z will now be random, given by a Poisson random variable N_z of parameter n. Let $(N_z)_{z\in\mathcal{H}}$ denote a family of i.i.d Poisson random variables of parameter n. Let $M \geq 0$. We build $A_n^*[M]$

using the same protocol as for $A_n[M]$, but instead of throwing n particles from each source z of $\mathcal{H}_M := \{0\} \times [\![-M, M]\!]^{d-1}$, we throw a number given by the N_z's. In this protocol, the random walks are launched independently of the N_z's. Once again, $(A_n^*[M])_{M\geq 0}$ is increasing with respect to inclusion, so we can define $A_n^*[\infty]$ as:

$$A_n^*[\infty] := \bigcup_{M\geq 0} \uparrow A_n^*[M] \quad \text{a.s.}$$

Construction of $A_n^\dagger[\infty]$ **:** Since our aim is to define a random forest whose distribution is invariant w.r.t translation of vectors in $\mathcal{H}$, we need to introduce a protocol ensuring that each source plays the same role in the construction. This will be the case in the construction of $A_n^\dagger[\infty]$.

We consider a family of i.i.d Poisson Point Processes (PPP) on $\mathbb{R}_+$, of intensity 1, which we note by $(\mathcal{N}_z)_{z\in\mathcal{H}}$. These will work as *random clocks* and will be the emission times for our particles. Each process $\mathcal{N}_z$ provides a family of times $(\tau_{z,j})_{j\geq 1}$, with which we can associate a family of independent random walks $(S_{z,j})$ (also independent of the $\mathcal{N}_z$'s). Hence, at time $\tau_{z,j}$, the j-th particle is emitted from the source z, and follows the trajectory given by $S_{z,j}$ before exiting the (current) aggregate. To avoid having multiple particles moving at the same time, we consider that particles realize their trajectories instantaneously.

We call $A_n^\dagger[M]$ the aggregate obtained using the protocol above, by sending particles from sources of $\mathcal{H}_M := \{0\}\times[\![-M, M]\!]^{d-1}$ and on the time interval $[0, n]$. One can show using a natural coupling (see Section 2.2) that $(A_n^\dagger[M])_{M\geq 0}$ is increasing with respect to inclusion, allowing us to define $A_n^\dagger[\infty]$ as:

$$A_n^\dagger[\infty] := \bigcup_{M\geq 0} \uparrow A_n^\dagger[M] \quad \text{a.s.}$$

Since each PPP considered when building $A_n^\dagger[\infty]$ is of intensity 1, we have that $\mathcal{N}_z([0, n]) \overset{\text{law}}{=} N_z$, where N_z is a Poisson variable of parameter n. Therefore, the random aggregates $A_n^*[M]$ and $A_n^\dagger[M]$ only differ in the *order* in which particles are sent. However, we know from the Abelian Property that this does not change the *law* of the final aggregate, thus for all $M \geq 0$,

$$A_n^\dagger[M] \overset{\text{law}}{=} A_n^*[M]. \tag{2}$$

Since both families of aggregates $(A_n^*[M])_{M\geq 0}$ and $(A_n^\dagger[M])_{M\geq 0}$ are increasing, we deduce from (2) that $A_n^\dagger[\infty] \overset{\text{law}}{=} A_n^*[\infty]$. This result will be particularly useful, as we will be able to use $A_n^*[\infty]$ to directly deduce results for $A_n^\dagger[\infty]$ and vice versa. In some cases, it will be easier to work with particles sent in a deterministic order, hence studying $A_n^*[\infty]$, while other times it will be advantageous to work with a random order. This is the case for example when showing stationary invariance of the aggregates: it is much easier to show invariance for $A_n^\dagger[\infty]$ than for $A_n^*[\infty]$, since in the first case, the particles are sent in no particular order, while in the second case particles are sent in a predetermined order. This result, a consequence of the Abelian Property, is only valid for the aggregates, and can not be exploited in the case of forests, for the reasons explained in Section 1.5.

Chapter 2

IDLA with sources in a hyperplane of $\mathbb{Z}^d$

In this chapter, we study the properties of the random aggregate $A_n[\infty]$ (see Section 2.1 for details about its construction). We generalize several results from [16], originally proven in dimension 2, to higher dimensions. We emphasize that the generalization of these results to higher dimensions is far from trivial and requires much more work. Let us begin by explaining some of the challenges of our model. The main difficulty when transitioning from dimension $d = 2$ to $d \geq 3$ arises from the geometry of the hyperplane $\mathcal{H} = \{0\} \times \mathbb{Z}^{d-1}$ from which the particles are emitted. Recall that the construction of $A_n[M+1]$ uses $A_n[M]$, and requires launching additional particles from sites of level $M+1$, that is, from sites in the set $\mathcal{H}_{M+1} \setminus \mathcal{H}_M$. The number of sites in this set is equal to

$$(2(M+1)+1)^{d-1} - (2M+1)^{d-1}.$$

When $d = 2$, this quantity is equal to 2, and the corresponding emission sites are $(0, M)$ and $(0, -M)$. However, for $d \geq 3$, this quantity now depends on M, and is of order $(d-1)2^{d-1}M^{d-2}$. Most of the techniques used in dimension 2 cease to be valid in higher dimensions, because the number of particles emitted from level M explodes as $M \to \infty$, making it difficult to control events involving particles far from the origin.

Another difficulty of our model compared to other IDLA models mentioned so far is the fact that we are dealing with an infinite number of particles. As a reminder, the aggregate $A_n[\infty]$ is built by launching exactly n particles from each site of the *infinite* hyperplane $\mathcal{H} = \{0\} \times \mathbb{Z}^{d-1}$. The presence of an infinite number of particles is a major challenge, and one of the main ways to overcome this is by considering *finite* portions of the aggregate.

Our first result is a *stabilization* result for the aggregate. It states that almost surely, any particle launched far from the origin will not reach regions close to the origin. Essentially, we show that particles of $A_n[\infty]$ emitted above levels M^α, with $\alpha > 1$ and M sufficiently large, do not visit the strip $\mathbb{Z}_M = \mathbb{Z} \times [\![-M, M]\!]^{d-1}$. This stabilization result for the aggregate is particularly useful when it comes to controlling the probability of events involving the aggregate in regions near the origin. Indeed, if we wish to prove a property of $A_n[\infty]$ inside the strip $\mathbb{Z}_M$, thanks to the stabilization result, it is sufficient to work with the *finite* aggregate $A_n[M^\alpha]$ rather than $A_n[\infty]$. This aggregate is much easier to handle than $A_n[\infty]$, as it is constructed using only a *finite* number of particles.

Our second result is a *shape theorem* for $A_n[\infty]$. We show that almost surely, when n is sufficiently large, the limiting shape of the aggregate $A_n[\infty]$, *near the origin*, is the 'slab'

$\mathcal{R}_{n/2} = [\![-n/2, n/2]\!] \times \mathbb{Z}^{d-1}$. The limiting shape obtained is the same as in Theorem 6.1 of [16]. However, the fluctuations around the limiting shape differ between $d = 2$ and $d \geq 3$, as was the case for the shape theorem of Asselah, Gaudillière, and Jerison, Levine, and Sheffield in the context of classical IDLA (see [4], [31]). Here, we obtain fluctuations around $\mathcal{R}_{n/2}$ of order $\sqrt{\log n}$ when $d \geq 3$, in contrast to fluctuations of order $\log n$ when $d = 2$. This difference is, once again, a consequence of the switch between recurrent and transient behavior of random walks when $d = 2$ and $d \geq 3$. Our shape theorem asserts that, at first order, the limiting aggregate $A_n[\infty]$ has a thickness equal to n, which seems consistent since n particles are launched from each source. We also note that this result does not apply to the entire aggregate $A_n[\infty]$, but only to the aggregate inside the strip $\mathbb{Z}_{n^\alpha}$ (with n large and $\alpha \geq 1$). It is *false* that the entire aggregate is contained inside $\mathcal{R}_{n/2}$, because almost surely, there exists a pathological source z (possibly very far from the origin) from which the n emitted particles move in the direction of $e_1 = (1, 0, \ldots, 0)$. This implies that almost surely, the site $z + n \cdot e_1$ belongs to $A_n[\infty]$. Thus, to avoid such pathological events, it is necessary to restrict ourselves to a strip near the origin.

Our final result is a *global upper bound* for the aggregate. Unlike the shape theorem, which describes the shape of the aggregate near the origin, this last result roughly controls the width of the aggregate for regions far from the origin. We show that the aggregate, above a certain level, is almost surely included inside a hypercone. The expanding shape of the cone allows it to effectively handle the pathological events that the shape theorem could not handle. However, due to this, it is not used to obtain a precise shape on the aggregate. Instead, the global upper bound is frequently used in arguments referred to in this thesis as the *donut method* or the *donut argument*. This method is used to control the trajectories of particles emitted far from the origin. We briefly explain the idea behind our method. We know that if $A_n[\infty]$ is included inside the hypercone, this implies that any particle that exits the hypercone necessarily exits $A_n[\infty]$ as well. The idea is to partition the hypercone into rings, or tori (hence the name *donut method*), and to say that a particle emitted far from the origin and visiting regions near the origin must necessarily cross a large number of *donuts*. The dimensions of each donut are carefully chosen so that we can easily bound the probability that a walk crosses a given donut. We then show that crossing a donut has a certain *cost* $0 < c < 1$ (in probability) for each given particle, so using a geometric argument, we show that the probability of a particle crossing k donuts is bounded by $(1-c)^k$. This method therefore allows us, when the number of donuts is sufficiently large, to have exponential decay for such events.

The previous results are featured in a submitted article [15] in collaboration with Nicolas Chenavier, David Coupier and Arnaud Rousselle.

Outline of the current chapter

1 Introduction

The (standard) Internal Diffusion Limited Aggregation (IDLA) is a random growth model $(A_n)_{n\geq 0}$ in $\mathbb{Z}^d$ recursively defined as follows. We start with $A_0 = \emptyset$. At step n, a simple symmetric random walk (independent of everything else) starts from the origin 0, called the *source*, until it exits the current aggregate A_{n-1}, say at some vertex z, which is added to A_{n-1} to get $A_n = A_{n-1} \cup \{z\}$. A first shape theorem was established by Lawler, Bramson and Griffeath in [34]. It asserts that the aggregate A_n (when it is suitably normalized) converges a.s. to an Euclidean ball as the number n of random walks goes to infinity, with fluctuations (w.r.t. the limit shape) which are at most linear. Since then, several papers (by Lawler [33], Asselah and Gaudillière [1, 4, 2] and Jerison, Levine and Sheffield [31, 30, 28]) have improved the bounds for fluctuations which are known to be logarithmic in dimension 2 and sublogarithmic in higher dimensions. Since then, many variants of this model have been considered and corresponding shape theorems have been explored. Let us cite IDLA models on discrete groups with polynomial or exponential growth in [10, 11], on non-amenable graphs in [26], on comb lattices in [5, 27], on cylinder graphs in [29, 37, 48] or on supercritical percolation clusters in [20, 47]. Let us mention that IDLA models with drifted random walks [38] or with uniform starting points [7] have been also studied. The case of multiple sources has been investigated too, see e.g. [16, 35, 21].

In this paper, we aim to extend the shape theorem in dimension $d = 2$ stated in [16] to higher dimensions. As explained below, this generalization is non-trivial and requires new ideas.

The infinite set of sources that we consider is the hyperplane $\mathcal{H} := \{0\} \times \mathbb{Z}^{d-1}$ of $\mathbb{Z}^d$, with $d \geq 3$. A random walk starting from a source of $\mathcal{H}$ and stopped when it exits the current aggregate is called a *particle*. Let n, M be non-negative integers. In the sequel, exactly n particles are sent from each source. Let us now build the sequence of aggregates $(A_n[M])_{M\geq 0}$ inductively as follows. When $M = 0$, $A_n[0]$ is the classical IDLA model, i.e. with n particles emitted from the origin. Let us call *level* M the set of sources in $\mathcal{H}$ at distance M from the origin (for $\|(z_1, \dots, z_d)\| := \max_i |z_i|$). Given a realization of $A_n[M-1]$, we throw n particles from each source of level M according to the lexicographical order. So $A_n[M]$ is defined as the aggregate produced by $A_n[M-1]$ and the new sites added by particles launched at level M.

Unlike its shape, the total number of sites in $A_n[M]$ is deterministic, and equals $\#A_n[M] = n(2M+1)^{d-1}$. Besides, by construction, the sequence of aggregates $(A_n[M])_{M\geq 0}$ is a.s. increasing in the sense of inclusion, allowing us to define the limiting aggregate $A_n[\infty]$ as:

$$A_n[\infty] := \bigcup_{M\geq 0} A_n[M] \quad \text{a.s.}$$

One of our main results is a shape theorem for $A_n[\infty]$. Restricted to the (large) strip $\mathbb{Z}_{n^\alpha} := \mathbb{Z} \times [\![-\lfloor n^\alpha \rfloor, \lfloor n^\alpha \rfloor]\!]^{d-1}$, the aggregate $A_n[\infty]$ looks like a slab with thickness n and sublogarithmic fluctuations as the number of particles n tends to infinity. Let us specify that the slab $\mathcal{R}_x$ is defined as $\mathcal{R}_x := [\![-\lfloor x \rfloor, \lfloor x \rfloor]\!] \times \mathbb{Z}^{d-1}$ for any positive real number x.

Theorem 2.1. *(Shape theorem) For any integers $d \geq 3$ and $\alpha \geq 1$, there exists a constant $C = C(d, \alpha) > 0$ such that, almost surely, there exists an integer $N \geq 1$ such that for any integer $n \geq N$,*

$$\mathcal{R}_{n/2-C\sqrt{\log n}} \cap \mathbb{Z}_{n^\alpha} \subset A_n[\infty] \cap \mathbb{Z}_{n^\alpha} \subset \mathcal{R}_{n/2+C\sqrt{\log n}} \cap \mathbb{Z}_{n^\alpha} \ . \tag{1}$$

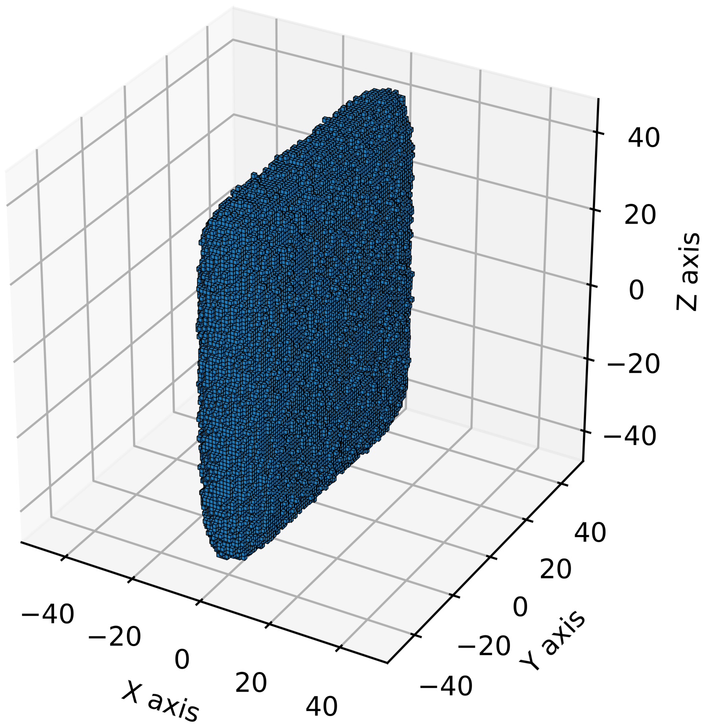


Figure 2.1: A realization of $A_{20}[40]$. Each particle is represented by a cube.

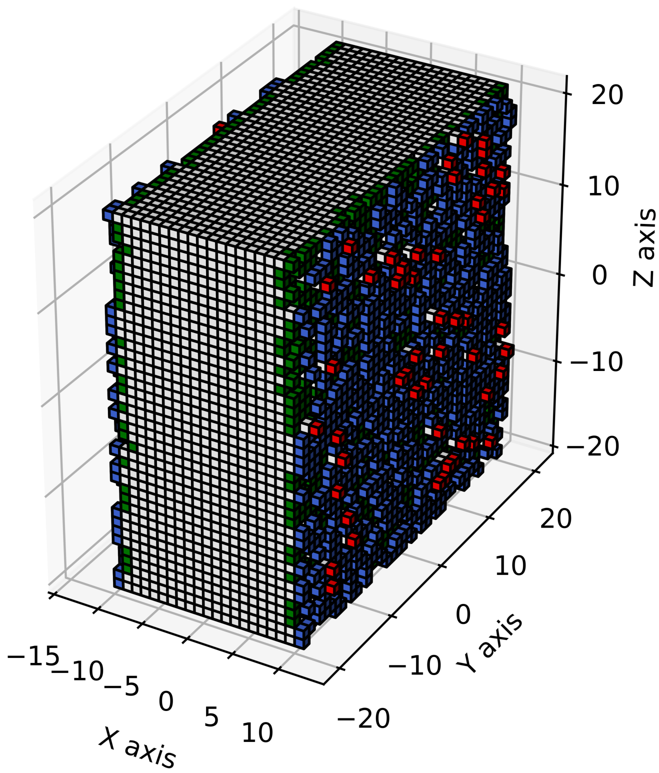


Figure 2.2: A realization of $A_{20}[40] \cap \mathbb{Z}_{20}$. The points with x-coordinate on the border such that $|x| = 10$ (resp. $|x| < 10$ and $|x| > 10$) are colored in blue (resp. green and red). All other points are colored in white.

Let us comment on this shape theorem (see Figures 2.1 and 2.2). It says that at first order, the limiting aggregate $A_n[\infty]$ is of thickness n, which makes sense since n particles are launched per source. Notice that Proposition 2.3 confirms that fact; n is (exactly) the mean thickness of $A_n[\infty]$. Let us also remark that Theorem 2.1 holds for the aggregate $A_n[\infty]$ restricted to the strip $\mathbb{Z}_{n^\alpha}$ (even large). Such a restriction is unavoidable since a.s. there exists some pathological source z (far away from the origin) for which all the n particles always move in the direction of the abscissa, meaning that the site $z+(n,0,\ldots,0)$ belongs to $A_n[\infty]$. Furthermore, Theorem 2.1 specifies the fluctuations of the aggregate $A_n[\infty]$ around its limiting shape $\mathcal{R}_{n/2}$ (both restricted to the strip $\mathbb{Z}_{n^\alpha}$). They are (at most) sublogarithmic while they are (at most) logarithmic in dimension $d = 2$ [16]. This dichotomy between dimension $d = 2$ and higher echoes the results of [4, 30], in which it is proved that the fluctuations for the standard IDLA are also sublogarithmic when $d \geq 3$.

Our second main result is a stabilization result for the aggregate $A_n[\infty]$. It ensures that particles emitted from sources far away from the origin do not reach regions near the origin. It puts forward an independence property between the aggregate $A_n[\infty] \cap \mathbb{Z}_M$ and particles from afar, which, in itself, is a very interesting property. While we will not be using it in this present paper, such a property will be crucial in [14], where we use it to generalize the constructions of IDLA forests from [16] to higher dimensions.

Theorem 2.2. *(Strong stabilization) Let $n \geq 0$ and $\alpha > 1$. A.s. there exists an integer M_0 such that, for any integer $M \geq M_0$, the trajectory of any particle contributing to $A_n[\infty]$ and starting from a level larger than M^α does not visit the strip $\mathbb{Z}_M$.*

In what follows, we assume $\alpha \geq 2$ to be an integer, as picking real values of α requires a heavy use of floor functions. This choice is made simply for the sake of lightening notation.

Theorem 2.2 is an extension of Theorem 3.1 of [16] (concerning the bidimensional case) to dimension $d \geq 3$. As we explain now, this extension is non-trivial and its proof requires a new approach. As in [16], particles contributing to the aggregate $A_n[\infty]$ are sent by successive waves, i.e. from the annuli

$$\mathrm{Ann}(M,j) := \mathcal{H} \cap \big(\mathbb{B}((j+2)M^\alpha) \setminus \mathbb{B}((j+1)M^\alpha)\big),\ j \geq 0,$$

where $\mathbb{B}(\ell)$ denotes the (closed) ball with radius ℓ and centered at the origin (w.r.t. the supremum distance $\|\cdot\|$). When $d = 2$, the hyperplane of sources $\mathcal{H}$ corresponds to the vertical axis and $\mathrm{Ann}(M,j)$ admits only $2M^\alpha$ sources, for any j. When $d \geq 3$, $\#\mathrm{Ann}(M,j)$ depends also on j and increases with j as j^{d-2} (this factor disappears when $d = 2$). The same holds for the number of particles sent during the $(j+1)$-th wave, i.e. from $\mathrm{Ann}(M,j)$. In order to visit the strip $\mathbb{Z}_M$ before stopping, a particle sent during the $(j+1)$-th wave has to travel inside the current aggregate until reaching $\mathbb{Z}_M$. It is more or less likely according to the index j and the thickness of the current aggregate which can then be viewed as a 'random environment' where the particle evolves before stopping. However, there is a certain deterioration of the 'environment' when successive waves are launched. Indeed, if A_{j+1} denotes the aggregate obtained after sending the j-th wave, then particles of the $(j+1)$-th wave contribute to the growth of A_{j+1} into A_{j+2} (i.e. A_{j+2} is thicker than A_j) making easier the travel inside the current aggregate to the strip $\mathbb{Z}_M$ for further particles. Hence, we have to deal with two opposite trends: as j increases, particles of the $(j+1)$-th wave have to travel a longer way to reach $\mathbb{Z}_M$, but this way is more likely since the corresponding aggregate is thicker. In dimension $d = 2$, the number of particles sent at each wave being weak (and constant w.r.t. j), the deterioration phenomenon of the 'environment' is negligible compared to the distance that particles must travel and the stabilization result is not too difficult to obtain in this case (see Section 3.1 of [16]). In dimension

$d \geq 3$, because of the increase of the number of particles sent at each wave, the deterioration of the 'environment' previously mentioned is stronger and the proof used in [16] no longer applies. To address this issue, the idea consists in proving that the aggregate $A_n[\infty]$, beyond some level, is included within a cone centered at the origin (Theorem 2.8). This upper bound presents two advantages. First it is rough enough– the thickness of the cone increases when one moves away from the origin –to take into account the deterioration of the 'environment' phenomenon and pathological sources (previously cited). Second, it is global since it concerns the whole aggregate outside some compact set. This result is referred to as a *rough global upper bound*.

Our paper is organized as follows. In Section 2, we give some properties of $A_n[\infty]$ including invariance (in distribution) w.r.t. translations/symmetries and a mass transport principle. We also recall the so-called *Abelian property* which ensures that the order in which the particles are sent is not important (in distribution) to define $A_n[\infty]$. In Section 3, we discretize our problem into donuts and establish a result (Proposition 2.6) which will be used to derive the rough global upper bound. This upper bound is stated and proved in Section 4. In the last two sections, we prove Theorems 2.1 and 2.2.

2 First properties

2.1 Mass transport property and symmetries

In this section, we state some basic properties satisfied by the random aggregate $A_n[\infty]$. The first property states that given a line $\mathbb{Z} \times \{y\}$, where $y \in \mathbb{Z}^{d-1}$, the average amount of particles that settle on this line is equal to n. Here, we have written $\{y\} := \{(y_2, y_3, \ldots, y_d)\}$ for any $y = (y_2, y_3, \ldots, y_d) \in \mathbb{Z}^{d-1}$. One can interpret this as the following statement: on average, the n particles sent from each source $(0, y)$ settle on the line $\mathbb{Z} \times \{y\}$.

Proposition 2.3. *Let $n \geq 1$. For all $y \in \mathbb{Z}^{d-1}$*

$$\mathbb{E}\left[\#\left(A_n[\infty] \cap (\mathbb{Z} \times \{y\})\right)\right] = n.$$

Just as in Section 2 of [16], a consequence of Proposition 2.3 is a result of *weak stabilization*, which claims that a particle sent far from the origin does not *settle* close to the origin. This result differs from our result of strong stabilization given in Section 5, as the latter shows that a particle sent far from the origin does not *visit* areas close to the origin. Moreover, unlike strong stabilization, weak stabilization does not provide any exploitable bounds, which makes it impossible to use arguments such as the Borel-Cantelli Lemma.

In the following proposition, we claim that the distribution of the random aggregate $A_n[\infty]$ is invariant with respect to translations and symmetries. In what follows, we denote by T_k the translation operator with respect to vector $k \in \mathcal{H}$ and $S_{k'}$ the point reflection operator across a point $k' \in \frac{1}{2}\mathcal{H}$, that is:

$$\forall x \in \mathbb{Z}^d,\ T_k(x) := x + k \quad \text{and} \quad S_{k'}(x) = 2k' - x.$$

For $B \subset \mathbb{Z}^{d-1}$, let

$$T_k B = \{T_k(x),\ x \in B\} \quad \text{and} \quad S_{k'} B = \{S_{k'}(x),\ x \in B\}.$$

Proposition 2.4. *Let $n \geq 0$, $k \in \mathcal{H}$, $k' \in \frac{1}{2}\mathcal{H}$.*

1. *The distribution of* $A_n[\infty]$ *is invariant with respect to* T_k*, i.e.* $T_k A_n[\infty] \overset{\text{law}}{=} A_n[\infty]$.

2. *The distribution of* $A_n[\infty]$ *is invariant with respect to* $S_{k'/2}$*, i.e.* $S_{k'/2} A_n[\infty] \overset{\text{law}}{=} A_n[\infty]$.

2.2 Abelian property

We give here the Abelian property, which states that altering the order in which particles are sent does not change the *law* of the aggregate. We begin by defining the Diaconis-Fulton *smash sum*: (see [19]). For $A \subset \mathbb{Z}^d$ (possibly random) and $z \in \mathbb{Z}^d$:

- if $z \notin A$, then $A \oplus \{z\} = A \cup \{z\}$;
- if $z \in A$, then $A \oplus \{z\}$ is the random set obtained by adding to A the vertex on which a simple random walk started in z, independent of A, exits A.

Proposition 2.5 (Abelian property)**.** *Let* A *and* $\{z_1, \dots, z_k\}$ *be subsets of* $\mathbb{Z}^d$*. The distribution of*

$$((A \oplus \{z_1\}) \oplus \{z_2\}) \oplus \cdots \oplus \{z_k\}$$

does not depend on the order of the z_i*'s. That is, if we take* $\sigma \in \mathfrak{S}_k$ *a permutation of* $\{1, \dots, k\}$*, then:*

$$((A \oplus \{z_1\}) \oplus \{z_2\}) \oplus \cdots \oplus \{z_k\} \overset{\text{law}}{=} ((A \oplus \{z_{\sigma(1)}\}) \oplus \{z_{\sigma(2)}\}) \oplus \cdots \oplus \{z_{\sigma(k)}\}.$$

2.3 Proofs of Propositions 2.3 and 2.4

We only show the proof of Proposition 2.3 since Proposition 2.4 can be dealt with in a similar manner. Our main idea is to build an auxiliary aggregate $A'_n[\infty]$ with the same law as $A_n[\infty]$, but for which it is simpler to show translation invariance. To do so, we construct $A'_n[\infty]$ in the same spirit as $A_n[\infty]$. Let $M \geq 0$. We define $A'_n[M]$ similarly to $A_n[M]$, by sending n particles per source of $\mathcal{H}_M := \{0\} \times [\![-M, M]\!]^{d-1}$, but this time the order is given by random clocks. More precisely, let $(\mathcal{U}_{z,i})_{z\in\mathcal{H},\ 1\leq i\leq n}$ be a family of i.i.d. uniform random variables on $[0,1]$. For each $z \in \mathcal{H}$ we can order these n random variables in order to get an increasing family of clocks $(\tau_{z,i})_{1\leq i\leq n}$ in $[0,1]$. Now, with the collection of random clocks $\{\tau_{z,i} :\ z \in \mathcal{H},\ 1 \leq i \leq n\}$ we can associate a family of independent symmetric random walks $\{S_{z,i} :\ z \in \mathcal{H},\ 1 \leq i \leq n\}$ on $\mathbb{Z}^d$, issued in z and independent of the family of clocks. Just like above, at time $\tau_{z,i}$, the i-th particle is sent from source $z \in \mathcal{H}$ and follows a trajectory given by $S_{z,i}$, adding a new site to the current aggregate. Let us specify that each particle's trajectory is instantly realized and that it settles immediately. The aggregate $A'_n[M]$ is obtained following the same protocol as above by sending particles up to level M according to the random clocks given by our family $(\mathcal{U}_{z,i})_{\|z\|\leq M,\ 1\leq i\leq n}$. Using the Abelian property, we have

$$A'_n[M] \overset{\text{law}}{=} A_n[M].$$

By adapting Lemma 2.1 of [16], we can easily show that a.s. for all $n, M \geq 0$, $A'_n[M] \subset A'_n[M+1]$. Then we define $A'_n[\infty]$ as the increasing union:

$$A'_n[\infty] := \bigcup_{M\geq 0} A'_n[M] \qquad \text{a.s.}$$

Since both sequences $(A'_n[M])_{M\geq 0}$ and $(A_n[M])_{M\geq 0}$ are almost surely increasing and that $A'_n[M] \overset{\text{law}}{=} A_n[M]$ for all $M \geq 0$, we have $A'_n[\infty] \overset{\text{law}}{=} A_n[\infty]$.

We are now prepared to prove Proposition 2.3. Indeed, since $A'_n[\infty] \overset{\text{law}}{=} A_n[\infty]$, it is sufficient to prove the same type of result for $A'_n[\infty]$. For $x, y \in \mathbb{Z}^{d-1}$, we let $Q_{x\to y}$ denote the number of particles sent from $(0, x)$ that settle on the line $\mathbb{Z} \times \{y\}$ in the construction of $A'_n[\infty]$. Now, for $A, B \subset \mathbb{Z}^{d-1}$, we define:

$$Q(A, B) := \mathbb{E}\left[\sum_{x\in A,\ y\in B} Q_{x\to y}\right].$$

In particular, for all $y \in \mathbb{Z}^{d-1}$, we have

$$\mathbb{E}\left[\#\left(A'_n[\infty] \cap (\mathbb{Z} \times \{y\})\right)\right] = Q(\mathbb{Z}^{d-1}, \{y\}).$$

Since $Q(\{y\}, \mathbb{Z}^{d-1}) = n$, it is sufficient to prove that $Q(\mathbb{Z}^{d-1}, \{y\}) = Q(\{y\}, \mathbb{Z}^{d-1})$. We show this using a mass transport argument (see Theorem 5.2 of [8]). It is sufficient to show that Q is diagonally invariant, that is: $Q(A + w, B + w) = Q(A, B)$ for all $w \in \mathbb{Z}^{d-1}$. This holds, since

$$\begin{aligned} Q(A + w, B + w) &= \sum_{x\in A,\ y\in B} \mathbb{E}\left[Q_{x+w\to y+w}\right] \\ &= \sum_{x\in A,\ y\in B} \mathbb{E}\left[Q_{x\to y}\right] \\ &= Q(A, B), \end{aligned}$$

where the second line comes from the fact that

$$\begin{aligned} Q_{x+w\to y+w} &= Q_{x+w\to y+w}\left((\tau_{z,i})_{\substack{z\in\mathcal{H} \\ 1\leq i\leq n}},\ (S_{z,i})_{\substack{z\in\mathcal{H} \\ 1\leq i\leq n}}\right) \\ &\overset{\text{a.s}}{=} Q_{x\to y}\left((\tau_{z-w,i})_{\substack{z\in\mathcal{H} \\ 1\leq i\leq n}},\ (S_{z-w,i})_{\substack{z\in\mathcal{H} \\ 1\leq i\leq n}}\right) \\ &\overset{\text{law}}{=} Q_{x\to y}. \end{aligned} \tag{2}$$

Note that the computations in (2) are specific to $A'_n[\infty]$, and are *not* true for $A_n[\infty]$. The key argument here is that the particles in $A'_n[\infty]$ are not sent according to a specific order but according to a family of independent uniform clocks, implying that all particles play the same role for the aggregate.

3 The donut method

In this section, we introduce what will be commonly referred to throughout this paper as the *donut method*. This argument will be particularly useful when coupled with the global upper bound given in Section 4 to control the trajectory of a given particle. The method consists in building donuts, starting from the origin and up to any given level, and showing that a particle is unlikely to cross multiple donuts without settling beforehand. Let us begin by detailing the construction of our donuts, for which it is necessary to first define *cones*. For $\varepsilon > 0$, we define the cone of angle ε as:

$$\mathscr{C}_\varepsilon := \bigcup_{l\geq 0}\Big\{z\in\mathbb{Z}^d,\ \|p_{\mathcal{H}}(z)\|=l,\ |z_1|\leq \varepsilon l\Big\}, \tag{3}$$

where $p_{\mathcal{H}}$ is the operator realizing the orthogonal projection on $\mathcal{H}$. Let $\mathbb{B}_{d-1}(r)$ denote the $(d-1)$-dimensional lattice ball of radius r, that is

$$\forall r>0,\ \mathbb{B}_{d-1}(r) := \{x\in\mathbb{Z}^{d-1} : \|x\|\leq r\}.$$

Given a decreasing family of real numbers $(l_i)_{i\geq 0}$, we define the donut $\mathbf{D}^i$ as:

$$\forall i\geq 0,\ \mathbf{D}^i := [\![-\varepsilon l_i, \varepsilon l_i]\!]\times(\mathbb{B}_{d-1}(l_i)\setminus\mathbb{B}_{d-1}(l_{i+1}))\,.$$

We build each donut $\mathbf{D}^i$ so that its length, which equals $2\varepsilon l_i$, is equal to its width $l_i - l_{i+1}$ (see Figure 2.3: one may see this figure as the view along a vertical cut of our donuts in dimension 3). This gives the following condition on $(l_i)_{i\geq 0}$:

$$\forall i\geq 0,\ l_{i+1} = (1-2\varepsilon)l_i,$$

with $\varepsilon < 1/2$. By induction, we get the general expression:

$$\forall i\geq 0,\ l_i = (1-2\varepsilon)^i l,$$

where $l = l_0$. We consider the number of donuts between levels l and M, with $M < l$, and define $k = k(l, M, \varepsilon)$ as the greatest integer such that:

$$\sum_{i=0}^{k} 2\varepsilon l_i \leq l - M.$$

Since $l_i = (1-2\varepsilon)^i l$, for ε taken small enough, we have:

$$k \geq \underbrace{\frac{-1}{2\log(1-2\varepsilon)}}_{K(\varepsilon)}\times\log\left(\frac{l}{M}\right). \tag{4}$$

Notice here that $K(\varepsilon)$ can be taken arbitrarily large by taking ε arbitrarily small.

Let us now briefly explain the reasoning behind the construction of our donuts. Our method will be particularly useful to show that a particle sent far away from the origin is highly unlikely to travel close to the origin while staying *within* the cone. For a particle to do so it necessarily has to travel through many donuts without ever exiting the cone, since the donuts are built in such a way that they wrap around the cone $\mathscr{C}_\varepsilon$. Such an event is handled by the following Proposition.

Proposition 2.6. *Let $M\geq 1$ and $\varepsilon > 0$. Fix $(S_t)_{t\geq 0}$ a simple symmetric random walk starting from some source of $\mathcal{H}\setminus\mathcal{H}_M$ and consider the cone $\mathscr{C}_\varepsilon$ defined as in (3). For $i\geq 1$, let*

$$A_i = \left\{\begin{array}{c}\textit{The walk crosses the } i \textit{ donuts } \boldsymbol{D}^0,\ldots,\ \boldsymbol{D}^{i-1}\\ \textit{without exiting the cone } \mathscr{C}_\varepsilon\end{array}\right\},$$

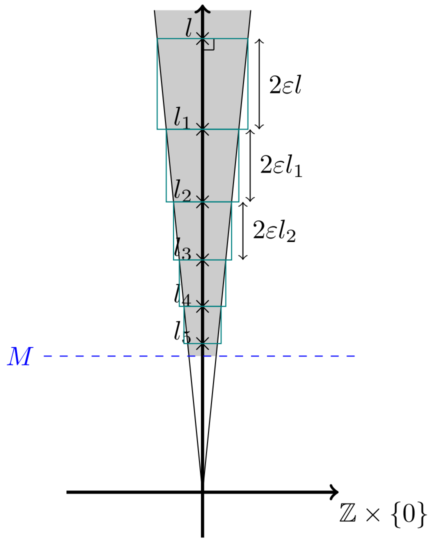


Figure 2.3: Partition into donuts

and let $A_0 = \Omega$. *Then, for any* $i \geq 0$,

$$\mathbb{P}(A_i) \leq (1-c)^i,$$

where $c = (2d)^{-2}$.

Note that what we mean by a walk or particle crossing donut $\mathbf{D}^i$ is for it to reach the inner ring of $\mathbf{D}^i$ without ever exiting $\mathbf{D}^i$.

Notice that for a walk to cross a donut (from the outer ring to the inner ring), it already needs to get through the middle of that donut. To deal with this property, let us introduce the notion of 'middling slice' of a donut. Let $i \geq 0$ and consider the i-th donut $\mathbf{D}^i = [\![-\varepsilon l_i, \varepsilon l_i]\!] \times (\mathbb{B}_{d-1}(l_i) \setminus \mathbb{B}_{d-1}(l_{i+1}))$. Define the lattice sphere of radius s as:

$$\forall s \geq 0,\ \mathbb{S}_{d-1}(s) := \{x \in \mathbb{Z}^{d-1},\ \|x\| = s\}.$$

Now, notice that $\frac{l_i + l_{i+1}}{2} = (1-\varepsilon) l_i$. Define the middling slice of $\mathbf{D}^i$ as

$$\mathbf{m}_i := [\![-\varepsilon l_i, \varepsilon l_i]\!] \times \mathbb{S}_{d-1}\left((1-\varepsilon) l_i\right).$$

Additionally, define the exterior border of $\mathbf{D}^i$ as

$$\mathbf{D}^i_{\text{ext}} = \Big(]\!] -\infty, -\varepsilon l_i]\!] \cup [\![\varepsilon l_i, +\infty[\![\Big) \times \Big(\mathbb{B}_{d-1}(l_i) \setminus \mathbb{B}_{d-1}(l_{i+1})\Big).$$

The following result shows that a walk started from the middling slice of a donut has a positive probability of exiting the donut through $\mathbf{D}^i_{\text{ext}}$ and will be used to derive Proposition 2.6.

Lemma 2.7. *Let* $y \in \mathbf{m}_i$ *and let* $(S_t)_{t\geq 0}$ *be a simple symmetric random walk on* $\mathbb{Z}^d$ *started at*

y. For all $i \geq 0$, we introduce the stopping time $\tau_y = \inf\{t \geq 0,\ S_t \notin \mathbb{B}_d(y, \varepsilon l_i)\}$. We have:

$$\mathbb{P}_y\left(S_{\tau_y} \in \mathbf{D}^i_{\text{ext}}\right) \geq \frac{1}{2d}.$$

Proof of Lemma 2.7: Let $y \in \mathbf{m}_i$. Notice that $\mathbb{B}_d(y, \varepsilon l_i) \subseteq \mathbb{Z} \times \Big(\mathbb{B}_{d-1}(l_i) \setminus \mathbb{B}_{d-1}(l_{i+1})\Big)$, and that $\mathbb{B}_d(y, \varepsilon l_i)$ has at least one of its $2d$ faces, say $\mathbf{F}$, included in $\mathbf{D}^i_{\text{ext}}$. By an argument of symmetry, we have

$$\mathbb{P}_y(S_{\tau_y} \in \mathbf{F}) = \frac{1}{2d}.$$

Now, since $\mathbb{P}_y(S_{\tau_y} \in \mathbf{F}) \leq \mathbb{P}_y(S_{\tau_y} \in \mathbf{D}^i_{\text{ext}})$, we have the desired result. □

Proof of Proposition 2.6. The case where $i = 0$ is trivial. Let $i \geq 1$. Notice that the sequence of events $(A_i)_{i\geq 0}$ is decreasing, so $\mathbb{P}(A_i) = \mathbb{P}(A_i|A_{i-1})\mathbb{P}(A_{i-1})$. Thus, it is sufficient to prove that $\mathbb{P}(A_i|A_{i-1}) \leq (1-c)$. Since we are considering events where the walk crosses donuts from outer ring to inner ring, we will refer to *good sides* as sides orthogonal to the 'x' axis, whereas *bad sides* will refer to sides that are *not good*.

Let us define the following events:

$$M_i = \Big\{\text{The random walk reaches } \mathbf{m}_i\Big\},$$

$$D_i = \left\{\begin{array}{c}\text{The random walk exits the } i\text{-th donut } \mathbf{D}^{i-1} \\ \text{through one of the bad sides}\end{array}\right\}.$$

Additionally, define the sequence of stopping times $T_i := \inf\{t \geq 0,\ S_t \in \mathbf{m}_i\}$. As mentioned earlier, for the walk to cross a donut (from outer ring to inner ring) it necessarily has to cross the middling slice of the donut, and since $\mathscr{C}_\varepsilon \cap \mathbb{Z}^c_M \subset \bigcup_{j\geq 0} \mathbf{D}^j$, this implies that on the event A_i, the walk crossed the i donuts $\mathbf{D}^0, \ldots,\ \mathbf{D}^{i-1}$ without ever exiting through a good side. Therefore:

$$\begin{aligned}
\mathbb{P}(A_i|A_{i-1}) &\leq \mathbb{P}(M_i \cap D_i|A_{i-1}) \\
&\leq \sum_{m\in\mathbf{m}_i} \mathbb{E}\left[\mathbb{1}_{D_i\cap M_i}\mathbb{1}_{S_{T_i}=m}|\ A_{i-1}\right] \\
&\leq \sum_{m\in\mathbf{m}_i} \mathbb{E}\left[\mathbb{1}_{D_i}\mathbb{1}_{S_{T_i}=m}|\ A_{i-1}\right] \\
&\leq \sum_{m\in\mathbf{m}_i} \mathbb{P}\left(D_i|\ S_{T_i}=m,\ A_{i-1}\right)\mathbb{P}(S_{T_i}=m|\ A_{i-1}).
\end{aligned}$$

Now, by the Markov property, for all $m \in \mathbf{m}_i,\ \mathbb{P}\left(D_i|\ S_{T_i}=m,\ A_{i-1}\right) \leq \mathbb{P}_m(D_i)$. It remains to bound $\mathbb{P}_m(D_i)$, which is an immediate consequence of Lemma 2.7. Let us first define $\partial\mathbf{D}^i_{\text{ext}} := \{-\varepsilon l_i - 1, \varepsilon l_i + 1\} \times \Big(\mathbb{B}_{d-1}(l_i) \setminus \mathbb{B}_{d-1}(l_{i+1})\Big)$. Notice that if the walk hits a site of $\partial\mathbf{D}^i_{\text{ext}}$, then it has necessarily exited the donut through a good side. Hence, using the result of Lemma 2.7:

$$\begin{aligned}
\mathbb{P}_m(D_i^c) &\geq \mathbb{P}_m\left(\{S_{\tau_m} \in \mathbf{D}^i_{\text{ext}}\} \cap \{S_{\tau_m+1} \in \partial\mathbf{D}^i_{\text{ext}}\}\right) \\
&\geq \mathbb{P}_m(S_{\tau_m+1} \in \partial\mathbf{D}^i_{\text{ext}}|S_{\tau_m} \in \mathbf{D}^i_{\text{ext}})\mathbb{P}_m(S_{\tau_m} \in \mathbf{D}^i_{\text{ext}}) \\
&\geq \mathbb{P}_{\mathbf{D}^i_{\text{ext}}}(S_1 \in \partial\mathbf{D}^i_{\text{ext}})\mathbb{P}_m(S_{\tau_m} \in \mathbf{D}^i_{\text{ext}}) \\
&\geq \frac{1}{2d} \times \frac{1}{2d}.
\end{aligned}$$

This concludes the proof, since

$$\begin{aligned}\mathbb{P}(A_i|A_{i-1}) &\leq \sum_{m\in\mathbf{m}_i} \mathbb{P}_m(D_i)\mathbb{P}(S_{T_i}=m|\ A_{i-1})\\ &\leq \left(1-\frac{1}{4d^2}\right)\sum_{m\in\mathbf{m}_i}\mathbb{P}(S_{T_i}=m|\ A_{i-1}) \leq 1-\frac{1}{4d^2}.\end{aligned}$$

□

4 A rough global upper bound

As seen in dimension 2 (see [16], Section 6), when restricted to a certain level, the aggregate $A_n[\infty]$ is contained within a rectangle of length roughly n, with high probability. In this section, we prove that above a certain level the aggregate is entirely contained within a cone with high probability. To state the result, we first give some notation. For any source $z \in \mathcal{H}$ and given a realization of $A_n[\infty]$, we define:

$$X_z(n) := \max\Big\{|z_1'|,\ z' \in A_n[\infty],\ z_i' = z_i \quad \forall i = 2,\ldots,d\Big\}.$$

The random variable $X_z(n)$ is the absolute value of the first coordinate of the furthest occupied site on the line of level z. Moreover, for any $0 < \varepsilon$ and $M \geq 1$ we let:

$$\mathbf{Over}(M,\varepsilon,n) = \bigcup_{l\geq M}\{\exists z \in \mathcal{H} : \|z\| = l,\ X_z(n) > \varepsilon l\}.$$

The event $\mathbf{Over}(M,\varepsilon,n)$ describes the situation where one or more particles have settled at a distance greater than εl on some line of distance $l \geq M$ from the origin. The following proposition shows that such an event occurs with small probability.

Theorem 2.8. *Let $n \geq 1$. For all $L > 1$, for all $\varepsilon > 0$, there exists a positive constant $C_{\varepsilon,n}$ such that for all $M \geq 2$:*

$$\mathbb{P}\left(\mathbf{Over}(M,\varepsilon,n)\right) \leq \frac{C_{\varepsilon,n}}{M^L}\ .$$

As a consequence of the above result, a.s. there exists a random integer M_0 such that for any $M \geq M_0$, the aggregate $A_n[\infty] \cap \mathbb{Z}^c_M$ is included in $\mathscr{C}_\varepsilon$. Theorem 2.8 can be understood as the fact that $A_n[\infty] \cap \mathbb{Z}^c_M$ is included within $\mathscr{C}_\varepsilon$ with high probability since

$$\mathbf{Over}(M,\varepsilon,n)^c = \{A_n[\infty] \cap \mathbb{Z}^c_M \subset \mathscr{C}_\varepsilon\}.$$

The property $A_n[\infty] \cap \mathbb{Z}^c_M \subset \mathscr{C}_\varepsilon$ is referred to as the *rough global upper bound.* This upper bound will be very useful when coupled with the donut argument of Section 3, as it will allow us to show that particles are unlikely to travel long distances while staying within the cone.

The proof of Theorem 2.8 will be shown by induction over n. Our idea is that if for some fixed n, the aggregate $A_n[\infty]$ is contained within a cone of angle ε for some $\varepsilon > 0$, then we show that after launching an additional particle from each source, the resulting aggregate is very likely contained in a slightly larger cone of angle $\varepsilon' > \varepsilon$. To prove Theorem 2.8, we first show a stronger version in Proposition 2.9. Before we give this result, we need to build two increasing sequences $(M_n)_{n\geq 0}$ and $(\varepsilon_n)_{n\geq 0}$, corresponding to levels of particles and successive angles of cones. Let

$\varepsilon \in]0,\,1[,\ M \geq 1$. We define the sequences $(M_n)_{n\geq 0}$ and $(\varepsilon_n)_{n\geq 0}$ by induction:

$$\begin{cases} M_0 = M \\ \forall n \geq 0,\ M_{n+1} = M_n \left(1 - \dfrac{\varepsilon}{2^{n+1}}\right)^{-1} \end{cases} \qquad \begin{cases} \varepsilon_0 = \varepsilon \\ \forall n \geq 0,\ \varepsilon_{n+1} = \varepsilon_n + \dfrac{\varepsilon}{2^n} \end{cases}$$

Note that for all $n \geq 0$, we have $\varepsilon \leq \varepsilon_n \leq 2\varepsilon$ and $M \leq M_n < 2M$, for ε small enough.
We now give a stronger version of Theorem 2.8. To avoid any heavy notation, in what follows, we will write $\mathbf{Over}(M_n, \varepsilon_n)$ rather than $\mathbf{Over}(M_n, \varepsilon_n, n)$. Additionally, we continue to omit writing the floor function $\lfloor \cdot \rfloor$.

Proposition 2.9. *For all $L > 1$, for all $\varepsilon > 0$, for all $n \geq 0$, there exists a constant $C_{\varepsilon,n} > 0$ such that for all $M \geq 1$,*

$$\mathbb{P}\left(\mathbf{Over}(M_n, \varepsilon_n)\right) \leq \frac{C_{\varepsilon,n}}{M^L}.$$

Proof of Theorem 2.8: Let n be fixed. Using the fact that for all $M \geq 1$ and for all ε small enough, $\varepsilon \leq \varepsilon_n \leq 2\varepsilon$ and $M \leq M_n < 2M$, we have $\mathbf{Over}(2M, 2\varepsilon, n) \subset \mathbf{Over}(M_n, \varepsilon_n)$, hence

$$\mathbb{P}\left(\mathbf{Over}(2M, 2\varepsilon, n)\right) \leq \mathbb{P}\left(\mathbf{Over}(M_n, \varepsilon_n)\right) \leq \frac{C_{\varepsilon,n}}{M^L}.$$

□

Proof of Proposition 2.9: We prove our result by induction over n. Take $L > 1$. Our induction statement is the following:

$$\forall n \geq 0,\ \mathcal{P}(n) : \forall \varepsilon \in]0,\,1[,\ \exists C_{\varepsilon,n} > 0, \forall M \geq 1,\ \mathbb{P}\left(\mathbf{Over}(M_n, \varepsilon_n)\right) \leq \frac{C_{\varepsilon,n}}{M^L}.$$

When $n = 0$, we have that $A_n[\infty] \overset{\text{a.s}}{=} \emptyset$ and hence $A_n[\infty] \cap \mathbb{Z}^c_M \overset{\text{a.s}}{\subset} \mathscr{C}_\varepsilon$, therefore $\mathbb{P}\left(\mathbf{Over}(M_n, \varepsilon_n)\right) = 0$. Let $n \geq 0$ and suppose $\mathcal{P}(n)$ holds. We write:

$$\mathbb{P}\left(\mathbf{Over}(M_{n+1}, \varepsilon_{n+1})\right) \leq \mathbb{P}\left(\mathbf{Over}(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}(M_n, \varepsilon_n)^c\right) + \mathbb{P}\left(\mathbf{Over}(M_n, \varepsilon_n)\right).$$

The right-hand term is handled by our induction hypothesis. We now focus on the left-hand term. On the event $\mathbf{Over}(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}(M_n, \varepsilon_n)^c$, we have $A_n[\infty] \cap \mathbb{Z}^c_{M_n} \subset \mathscr{C}_{\varepsilon_n}$, but when launching one additional particle from each source of $\mathcal{H}$, the new aggregate obtained spills over $\mathscr{C}_{\varepsilon_{n+1}}$ on $\mathbb{Z}^c_{M_{n+1}}$. This implies the existence of three random sites $(Z, Z^*, Z_{n+1}) \in \mathbb{Z}^d$ such that:

- Z^* is the source from which the first overflowing particle is emitted
- Z_{n+1} is the site on which this particle settles
- Z is the orthogonal projection of Z_{n+1} on $\mathcal{H}$.

Note that the coordinates of Z_{n+1} are given by

$$Z_{n+1} = Z \pm (\varepsilon_{n+1} \|Z\|) \cdot e_1,$$

where $e_1 = (1, 0, \ldots, 0)$, and that these coordinates only depend on Z (up to the sign), meaning it suffices to know the location of Z to know precisely where the overflowing particle settled.
Additionally, we call A_{Z^*} the aggregate (restricted to $\mathbb{Z}^c_{M_{n+1}}$) made up of $A_n[\infty]$ and each additional particle sent from $\mathcal{H}$ in the usual order up to site Z^*, and $A^-_{Z^*}$ the aggregate (restricted

to $\mathbb{Z}^c_{M_{n+1}}$) made up of $A_n[\infty]$ and each additional particle sent from $\mathcal{H}$ in the usual order up to site Z^* *excluded.* We know that this aggregate is contained inside of $\mathscr{C}_{\varepsilon_{n+1}} \cap \mathbb{Z}^c_{M_{n+1}}$. Notice that $A_{Z^*} = A^-_{Z^*} \cup \{Z_{n+1}\}$.
We write:

$$\begin{aligned}\mathbb{P}\left(\mathbf{Over}(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}(M_n, \varepsilon_n)^c\right) \\ \leq \sum_{l \geq M_{n+1}} \sum_{\|z\|=l} \mathbb{P}\left(Z = z,\ \mathbf{Over}(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}(M_n, \varepsilon_n)^c\right). \quad (5)\end{aligned}$$

Now, fix $l \geq M_{n+1}$ and $z \in \mathcal{H}$ such that $\|z\| = l$, and let $z_{n+1} = z \pm (\varepsilon_{n+1}\|z\|) \cdot e_1$. To deal with the probability in (5), we consider two cases, which we show are both unlikely. The first case is the case where many particles have settled in a ball around z_{n+1}, and the second is the case where a thin 'tentacle' has branched out towards z_{n+1}. The following lemma is an adaptation of Lemma 2 of [31], which deals with the case of tentacles:

Lemma 2.10. *There exist positive universal constants b, K_0, c such that for all real numbers $r > 0$ and all $z \in \mathcal{H}$ with $0 \notin \mathring{\mathbb{B}}(z_{n+1}, r)$,*

$$\begin{aligned}\mathbb{P}\left(Z = z,\ \mathbf{Over}(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}(M_n, \varepsilon_n)^c,\ \#\left(A_{Z^*} \cap \mathbb{B}(z_{n+1}, r)\right) \leq br^d\right) \\ \leq K_0 e^{-cr^2}.\end{aligned}$$

Let us first explain the choice of the radius for the ball centered around z_{n+1}. This is where the construction of (M_n) and (ε_n) comes into play. When building the ball around z_{n+1}, we need to take a radius small enough to ensure that our ball does not intersect $\mathscr{C}_{\varepsilon_n}$ as well as the strip $\mathbb{Z}_{M_n} := \mathbb{Z} \times [\![-M_n, M_n]\!]^{d-1}$, since we need to consider only new particles (particles contributing to $A_{n+1}[\infty] \setminus A_n[\infty]$). To do it, we define the radius $r_{n+1} = r_{n+1}(l, \varepsilon)$ as

$$r_{n+1} = \frac{\varepsilon_{n+1} - \varepsilon_n}{2}\|z\| = \frac{\varepsilon}{2^{n+1}} l. \quad (6)$$

This choice of r_{n+1} ensures that $\mathbb{B}\left(z_{n+1}, r_{n+1}\right) \cap \mathscr{C}_{\varepsilon_n} = \emptyset$, since z_{n+1} is necessarily at a distance $(\varepsilon_{n+1} - \varepsilon_n)\, l$ of $\mathscr{C}_{\varepsilon_n}$. Moreover, $\mathbb{B}(z_{n+1}, r_{n+1}) \cap \mathbb{Z}_{M_n} = \emptyset$, since $M_{n+1} \leq l$ thus $r_{n+1} \leq \|z\| - M_n$. To deal with (5), we write:

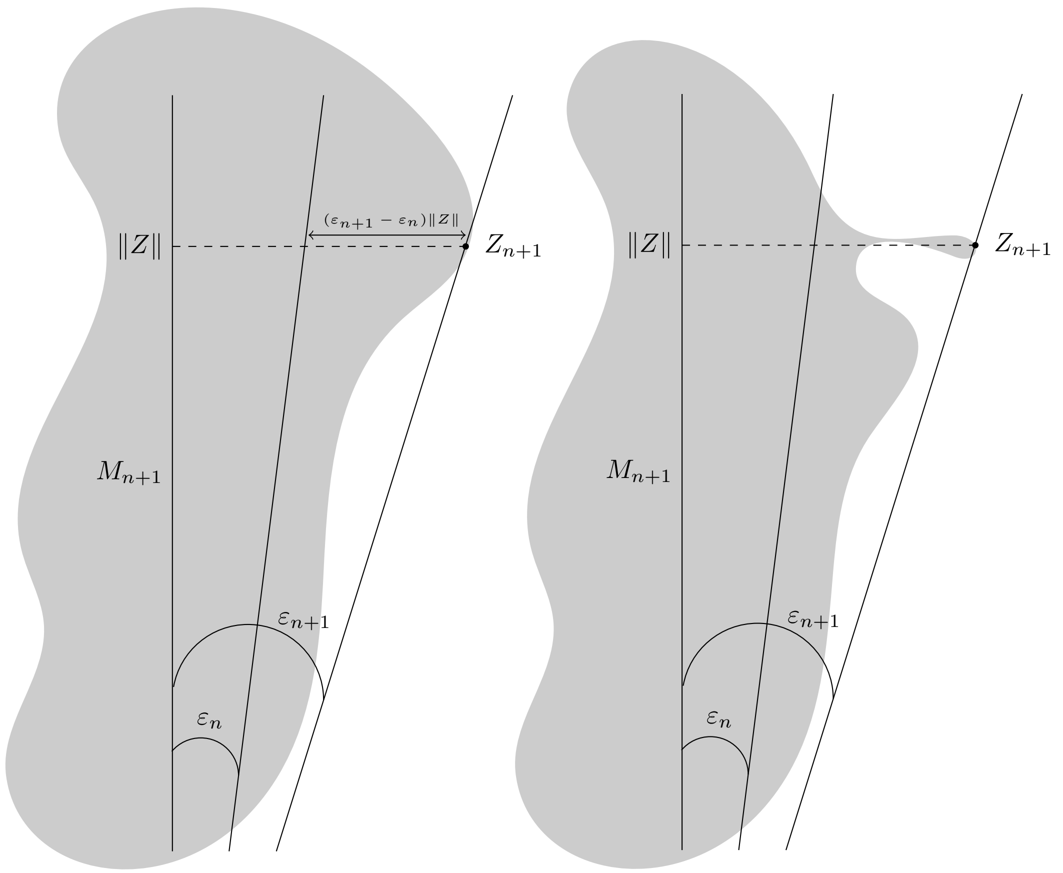
$$\begin{aligned}&\mathbb{P}\left(Z = z,\ \mathbf{Over}(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}(M_n, \varepsilon_n)^c\right) \\ \leq &\mathbb{P}\left(Z = z,\ \mathbf{Over}(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}(M_n, \varepsilon_n)^c,\ \#\left(A_{Z^*} \cap \mathbb{B}(z_{n+1}, r_{n+1})\right) \leq br^d_{n+1}\right) \quad (7) \\ + &\mathbb{P}\left(Z = z,\ \mathbf{Over}(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}(M_n, \varepsilon_n)^c,\ \#\left(A_{Z^*} \cap \mathbb{B}(z_{n+1}, r_{n+1})\right) > br^d_{n+1}\right) \quad (8)\end{aligned}$$

The term (7) is handled by Lemma 2.10 with $r = r_{n+1}$ as in (6). This gives:

$$\begin{aligned}\mathbb{P}\left(Z = z,\ \mathbf{Over}(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}(M_n, \varepsilon_n)^c,\ \#\left(A_{Z^*} \cap \mathbb{B}(z_{n+1}, r_{n+1})\right) \leq br^d_{n+1}\right) \\ \leq K_0 e^{-c_1 l^2}, \quad (9)\end{aligned}$$

where $c_1 = c_1(n, \varepsilon) = \dfrac{c\varepsilon^2}{4^{n+1}}$.

To deal with (8), we use the following argument: in order to have more than $br^d_{n+1} = b\varepsilon^d 2^{-d(n+1)}\|z\|^d$ *new* particles gathered in a ball around z_{n+1}, and knowing only *one*



(a) Illustration of $\mathbf{Over}(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}(M_n, \varepsilon_n)^c$

(b) Case of a tentacle reaching out to Z_{n+1}

additional particle is thrown from each site, this implies that $\|Z - Z^*\| \geq K\|z\|^{\frac{d}{d-1}}$ (where $K = K(\varepsilon, n)$ is a positive constant). Indeed, since $\mathcal{H}$ is a hyperplane of dimension $d-1$, for any $\delta > 0$, the number of sources of $\mathbb{B}(Z, \|z\|^{\delta}) \cap \mathcal{H}$ is of order $\|z\|^{\delta(d-1)}$. Each of these sources emits only *one* additional particle, so the total number of particles emitted within $\mathbb{B}(Z, \|z\|^{\delta}) \cap \mathcal{H}$ is also of order $\|z\|^{\delta(d-1)}$. In order for this quantity to be of the same order as br_{n+1}^d, which is of order $\|z\|^d$, it is hence necessary that $\delta = \frac{d}{d-1}$. This gives:

$$
\begin{aligned}
&\mathbb{P}\left(Z = z,\ \mathbf{Over}(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}(M_n, \varepsilon_n)^c,\ \#\left(A_{Z^*} \cap \mathbb{B}(z_{n+1}, r_{n+1})\right) > b r_{n+1}^d\right) \\
\leq &\mathbb{P}\left(Z = z,\ \|Z^* - z\| \geq K l^{\frac{d}{d-1}},\ \mathbf{Over}(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}(M_n, \varepsilon_n)^c\right) \\
\leq &\mathbb{P}\left(\bigcup_{h \geq K l^{\frac{d}{d-1}}} \bigcup_{\|z'-z\|=h} \left\{\text{the particle sent from } z' \text{ reaches level } l \text{ while staying within } \mathscr{C}_{\varepsilon_{n+1}}\right\}\right) \\
\leq &\sum_{h \geq K l^{\frac{d}{d-1}}} \sum_{\|z'-z\|=h} \mathbb{P}\left(\text{the particle sent from } z' \text{ reaches level } l \text{ while staying within } \mathscr{C}_{\varepsilon_{n+1}}\right).
\end{aligned}
$$

Proceeding in the same way as Section 3, we can build donuts between levels h and l within the cone $\mathscr{C}_{\varepsilon_{n+1}}$. The previous probability is then handled using a donut argument as in Proposition 2.6, and is smaller than $(1-c)^k$, with $c = (2d)^{-2}$ and with $k = k(h, l, \varepsilon_{n+1})$ such that

$$
k \geq \frac{-1}{2\log(1 - 2\varepsilon_{n+1})} \times \log\left(\frac{h}{l}\right), \tag{10}
$$

given that ε is small enough and M is large enough (here, we used the fact that z_{n+1} is at the same level as z). Therefore,

$$
\begin{aligned}
\mathbb{P}\left(Z = z,\ \mathbf{Over}(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}(M_n, \varepsilon_n)^c,\ \#\left(A_{Z^*} \cap \mathbb{B}(z_{n+1}, r_{n+1})\right) > b r_{n+1}^d\right) \\
\leq \sum_{h \geq K l^{\frac{d}{d-1}}} \sum_{\|z'-z\|=h} (1-c)^k.
\end{aligned}
$$

Now, using (10), standard computations yield

$$
\sum_{h \geq K l^{\frac{d}{d-1}}} \sum_{\|z'-z\|=h} (1-c)^k \leq K_{d,\varepsilon,n} l^{d - \frac{C}{d-1}},
$$

where $C := C(\varepsilon_{n+1}) = \frac{\log(1-c)}{2\log(1-2\varepsilon_{n+1})}$ can be taken as large as we want (given once again that ε is small enough) and $K_{d,\varepsilon,n}$ denotes a positive constant depending only on d, ε and n. Combining this with (9), we get:

$$
\begin{aligned}
&\mathbb{P}\left(\mathbf{Over}(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}(M_n, \varepsilon_n)^c\right) \\
&\qquad \leq \sum_{l \geq M_{n+1}} \sum_{\|z\|=l} K_0 e^{-c_1 l^2} + \sum_{l \geq M_{n+1}} \sum_{\|z\|=l} K_{d,\varepsilon,n} l^{d - \frac{C}{d-1}} \\
&\qquad \leq K_d \sum_{l \geq M_{n+1}} l^{d-2} e^{-c_1 l^2} + K_{d,\varepsilon,n} \sum_{l \geq M_{n+1}} l^{d-2} l^{d - \frac{C}{d-1}}.
\end{aligned}
$$

Notice that the first term of the previous sum can be bounded by $K e^{-\frac{c_1 M_{n+1}^2}{2}}$ for some constant K. Since $M \leq M_{n+1}$, it is clear that $K e^{-\frac{c_1 M_{n+1}^2}{2}} \leq \frac{C'_{\varepsilon,n}}{M^L}$ for some constant $C'_{\varepsilon,n} > 0$. To deal with the second term, recall that $C := C(\varepsilon_{n+1})$ can be taken as large as necessary (by taking ε

sufficiently small). We can therefore choose ε small enough such that:

$$\sum_{l\geq M_{n+1}} l^{2d-2-\frac{C}{d-1}} \leq \frac{C''_{\varepsilon,n}}{M^L}.$$

Recall that from our induction hypothesis, $\mathbb{P}\left(\mathbf{Over}(M_n,\varepsilon_n)\right) \leq \frac{C_{\varepsilon,n}}{M^L}$. This implies

$$\begin{aligned}\mathbb{P}\left(\mathbf{Over}(M_{n+1},\varepsilon_{n+1})\right) &\leq \mathbb{P}\left(\mathbf{Over}(M_{n+1},\varepsilon_{n+1}) \cap \mathbf{Over}(M_n,\varepsilon_n)^c\right) + \mathbb{P}\left(\mathbf{Over}(M_n,\varepsilon_n)\right)\\ &\leq \frac{C'_{\varepsilon,n}}{M^L} + \frac{C''_{\varepsilon,n}}{M^L} + \frac{C_{\varepsilon,n}}{M^L}\end{aligned}$$

and concludes the proof of Proposition 2.9. □

5 Strong stabilization

This section is devoted to the proof of Theorem 2.2 which heavily lies on the donut method and the global upper bound (see Sections 3 and 4).

Fix an integer $\alpha \geq 2$. For $M \geq 1,\ j \geq 0$, recall the definition of :

$$\mathrm{Ann}(M,j) := \mathcal{H} \cap \big(\mathbb{B}((j+2)M^\alpha) \setminus \mathbb{B}((j+1)M^\alpha)\big),\ j \geq 0,$$

and let $E_{M,j}$ be the following event:

$$E_{M,j} = \left\{ \begin{array}{c} \text{At least one of the particles starting from } \mathrm{Ann}(M,j) \\ \text{visits the strip } \mathbb{Z}_M \text{ before exiting the current aggregate} \end{array} \right\}.$$

According to the Borel-Cantelli lemma, it is sufficient to show that

$$\sum_{M\geq 1} \mathbb{P}\left(\bigcup_{j\geq 0} E_{M,j} \right) < +\infty.$$

To do it, we write:

$$\mathbb{P}\left(\bigcup_{j\geq 0} E_{M,j} \right) \leq \sum_{j\geq 0} \mathbb{P}\left(E_{M,j} \cap \mathbf{Over}(M,\varepsilon,n)^c\right) + \mathbb{P}\left(\mathbf{Over}(M,\varepsilon,n)\right).$$

We focus on the left-hand term. To deal with it, let $j \geq 0$, take $l = M^\alpha(j+1)$ and take $k = k(l,M,\varepsilon)$ as in (4). Define $N_{tot} = N_{tot}(n,M,j)$ as the total number of particles sent from

$\mathrm{Ann}(M,j)$. We have:

$$\begin{aligned}
&\mathbb{P}\left(E_{M,j}\cap \mathbf{Over}(M,\varepsilon,n)^c\right)\\
&\quad= \mathbb{P}\left(\bigcup_{i=1}^{N_{tot}}\Big\{\text{particle } i \text{ visits } \mathbb{Z}_M \text{ before exiting the aggregate}\Big\}\cap \mathbf{Over}(M,\varepsilon,n)^c\right)\\
&\quad\leq \sum_{i=1}^{N_{tot}}\mathbb{P}\left(\Big\{\text{particle } i \text{ visits } \mathbb{Z}_M \text{ before exiting the aggregate}\Big\}\cap \mathbf{Over}(M,\varepsilon,n)^c\right)\\
&\quad\leq \sum_{i=1}^{N_{tot}}\mathbb{P}\left(\Big\{\text{particle } i \text{ crosses } k \text{ donuts before exiting the aggregate}\Big\}\cap \mathbf{Over}(M,\varepsilon,n)^c\right)\\
&\quad\leq \sum_{i=1}^{N_{tot}}\mathbb{P}\left(\Big\{\text{the } \textit{walk} \text{ associated with particle } i \text{ crosses } k \text{ donuts before exiting } \mathscr{C}_\varepsilon\cap\mathbb{Z}_M^c\Big\}\right)\\
&\quad\leq N_{tot}(1-c)^k,
\end{aligned}$$

where the last line comes from Proposition 2.6.
Notice here that the global upper bound $\mathbf{Over}(M,\varepsilon,n)^c$ is used to deduce two different arguments. The first time, it allows us to say that if a particle reaches $\mathbb{Z}_M$ before exiting the aggregate, then it necessarily crosses the k donuts $\mathbf{D}^0,\ldots,\ \mathbf{D}^{k-1}$ before exiting the aggregate, since the aggregate is contained within the cone, which itself is contained within the union of the donuts. The second time, it allows us to say that if a particle crosses k donuts without exiting the aggregate, then in particular, it does so without exiting the cone, and the same is true for its associated walk. Now, using (4) with $l = M^\alpha(j+1)$, we get:

$$\begin{aligned}
N_{tot}(1-c)^k &\leq N_{tot}\exp\left(K(\varepsilon)\log\left(\frac{M^\alpha(j+1)}{M}\right)\log(1-c)\right)\\
&\leq N_{tot}M^{C(1-\alpha)}(j+1)^{-C},
\end{aligned}$$

where $C := -K(\varepsilon)\log(1-c) = \dfrac{\log(1-c)}{2\log(1-2\varepsilon)}$ can be taken arbitrarily large, by taking ε arbitrarily small.
Since $N_{tot}\leq K_d M^{\alpha(d-1)}j^{d-2}n$, we have:

$$\begin{aligned}
&\sum_{M\geq 1}\mathbb{P}\left(\bigcup_{j\geq 0}E_{M,j}\right)\\
&\qquad\leq \sum_{M\geq 1}\sum_{j\geq 0}\mathbb{P}\left(E_{M,j}\cap\mathbf{Over}(M,\varepsilon,n)^c\right)+\sum_{M\geq 1}\mathbb{P}\left(\mathbf{Over}(M,\varepsilon,n)\right)\\
&\qquad\leq K_d n\sum_{M\geq 1}M^{C(1-\alpha)+\alpha(d-1)}\sum_{j\geq 0}j^{d-2}(j+1)^{-C}+\sum_{M\geq 1}\mathbb{P}\left(\mathbf{Over}(M,\varepsilon,n)\right).
\end{aligned}$$

The left hand-term of the sum is finite since $\alpha > 1$ and since we can pick $C = C(\varepsilon)$ sufficiently large, while the second term is handled using Theorem 2.8. This concludes the proof of Theorem 2.2.

6 Shape theorem

In this section, we prove Theorem 2.1 following the same strategy as [4, 3] and [16], by splitting the proof into two parts: the lower bound and the upper bound. We begin by showing the lower bound, as we will be using it later for the proof of the upper bound.

Let us recall that, for any real number $x > 0$, the slab $\mathcal{R}_x$ and the strip $\mathbb{Z}_x$ are defined as

$$\mathcal{R}_x = [\![-\lfloor x \rfloor, \lfloor x \rfloor]\!] \times \mathbb{Z}^{d-1}$$

and

$$\mathbb{Z}_x = \mathbb{Z} \times [\![-\lfloor x \rfloor, \lfloor x \rfloor]\!]^{d-1}.$$

6.1 Proof for the lower bound

In this section, we show the lower bound of Theorem 2.1, which is the following result: for any integers $d \geq 3$ and any $\alpha \geq 1$, there exists a constant $C = C(d, \alpha) > 0$ such that, almost surely, there exists $N \geq 1$ such that for any $n \geq N$,

$$\mathcal{R}_{n/2 - C\sqrt{\log n}} \cap \mathbb{Z}_{n^\alpha} \subset A_n[\infty] \cap \mathbb{Z}_{n^\alpha}.$$

We adopt in this section the notation of [1, 4, 16] and denote by $A(\eta)$ the aggregate generated by an initial configuration η. Since we are launching n particles from each site of $\mathcal{H}$, we will mostly be using the notation $A(n\mathbb{1}_{\mathcal{H}})$ to refer to $A_n[\infty]$.

For $k \in \mathbb{N}$, we define the shell S_k by

$$S_k = \left(\mathcal{R}_{(k+1)\sqrt{\log n}} \setminus \mathcal{R}_{k\sqrt{\log n}}\right) \cap \mathbb{Z}_{n^\alpha}.$$

We let

$$\partial\mathcal{R}_{k\sqrt{\log n}} = \{-\lfloor k\sqrt{\log n} \rfloor, \lfloor k\sqrt{\log n} \rfloor\} \times \mathbb{Z}^{d-1}.$$

and write

$$\partial_{k,n} = \partial\mathcal{R}_{k\sqrt{\log n}} \cap \mathbb{Z}_{n^\alpha}.$$

Now, for $z \in \partial\mathcal{R}_{k\sqrt{\log n}}$ we define the *tile* and *cell* centered in z as

$$\tau(z) = \mathbb{B}\left(z, \frac{\sqrt{\log n}}{2}\right) \cap \partial\mathcal{R}_{k\sqrt{\log n}} \quad \text{and} \quad \mathcal{C}(z) = \mathbb{B}\left(z, \sqrt{\log n}\right) \cap \mathcal{R}^c_{k\sqrt{\log n}}.$$

The strategy to prove the lower bound is to show that each tile of $\mathcal{R}_{k\sqrt{\log n}}$ is likely visited by many particles, and then show that if many particles reached a tile, they are likely to fill up the corresponding cell. The idea here is similar to that of a floodgate. Each tile τ can be seen as a floodgate, and the particles as water. We stop the particles once they reach τ, and let them accumulate on the tile, just like a floodgate would store water. Then, when there is a sufficient number of particles accumulated on τ, we release the particles and show that they are likely to fill up the corresponding cell. This is the same as opening the floodgates and letting the water run free again. For some configuration η and $B \subset \mathcal{R}_{k\sqrt{\log n}}$, we will denote by:

- $W_{k\sqrt{\log n}}(\eta, B)$ the number of particles with initial configuration η that hit set B before or when exiting $\mathcal{R}_{k\sqrt{\log n}}$ (and before leaving the aggregate).

- $M_{k\sqrt{\log n}}(\eta, B)$ the number of random walks with initial configuration η that hit set B before or when exiting $\mathcal{R}_{k\sqrt{\log n}}$.

We say that set B is not covered if $B \not\subset A(n\mathbb{1}_{\mathcal{H}})$. It is sufficient to show that there exists a constant C such that for all $L > 1,\ n \geq 1$ and $k \leq \frac{n}{2\sqrt{\log n}} - C$, we have

$$\mathbb{P}\left(S_k \text{ is not covered }\right) \leq \frac{c}{n^L}. \tag{11}$$

This implies that

$$\sum_{n\geq 1} \mathbb{P}\left(S_k \text{ is not covered }\right) \leq \sum_{n\geq 1} \frac{c}{n^L} < +\infty,$$

which, combined with the Borel-Cantelli lemma, suffices to prove the lower bound. Now, let

$$\mathcal{T}_{k\sqrt{\log n}} = \{\tau(z),\ z \in \partial_{k,n}\},$$

and for a tile $\tau = \tau(z) \in \mathcal{T}_{k\sqrt{\log n}}$,

$$\mu(\tau) = \mathbb{E}\left[M_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau)\right] - \mathbb{E}\left[M_{k\sqrt{\log n}}\left(\mathcal{R}_{k\sqrt{\log n}} \setminus \mathcal{Z}, \tau\right)\right],$$

where $\mathcal{Z}$ is the following set:

$$\mathcal{Z} = \{z' \in \mathcal{R}_{k\sqrt{\log n}} : d(z', \tau(z)) \leq b\sqrt{\log n}\}$$

for some $b > 0$ which will be chosen later, and with $d(x, A) = \inf_{y\in A} \|x - y\|$. Now, fix $C > 0$. For $k \leq \frac{n}{2\sqrt{\log n}} - C$, we write:

$$\mathbb{P}\left(S_k \text{ is not covered }\right) \leq \mathbb{P}\left(\exists \tau \in \mathcal{T}_{k\sqrt{\log n}},\ W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) < \frac{1}{3}\mu(\tau)\right) \tag{12}$$

$$+\mathbb{P}\left(\forall \tau \in \mathcal{T}_{k\sqrt{\log n}},\ W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) \geq \frac{1}{3}\mu(\tau),\ S_k \text{ is not covered }\right). \tag{13}$$

The second term of expression (13) will be handled later in Section 6.1. To handle (12), essentially, we must control the probability that few particles hit a tile. This is achieved using Lemma 2.11 below. Notice that working on strip $\mathbb{Z}_{n^\alpha}$ allows us to use a union bound on $\mathbb{P}\left(\exists \tau \in \mathcal{T}_{k\sqrt{\log n}},\ W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) < \frac{1}{3}\mu(\tau)\right)$, since

$$\#\mathcal{T}_{k\sqrt{\log n}} \leq 2(2n^\alpha + 1)^{d-1} \leq Kn^{\alpha(d-1)}.$$

It is sufficient to prove that for any tile $\tau \in \mathcal{T}_{k\sqrt{\log n}}$,

$$\mathbb{P}\left(W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) < \frac{1}{3}\mu(\tau)\right) \leq \exp\left(-\kappa C^2 \log(n)\right). \tag{14}$$

The above inequality is a consequence of the following lemma (Lemma 2.5 of [4]):

Lemma 2.11. *Suppose that a sequence of random variables* $\{W_n, M_n, L_n, \widetilde{M}_n; n \geq 0\}$ *and a sequence of real numbers* $(c_n)_{n\geq 0}$ *satisfy for any* $n \geq 0$*:*

$$W_n + L_n + c_n \geq \widetilde{M}_n \qquad \text{and} \qquad \widetilde{M}_n \overset{\text{law}}{=} M_n.$$

Assume that W_n *and* L_n *are independent and that* L_n *and* M_n *are both sums of independent Bernoulli random variables with finite first moment. Assume also that*

(H1) *the Bernoulli variables* $\{Y_1^{(n)},\ Y_2^{(n)}, \dots\}$ *whose series is* L_n *satisfy for some* $\eta > 1$*:*

$$\sup_n \sup_i \mathbb{E}\left[Y_i^{(n)}\right] < \frac{\eta - 1}{\eta};$$

(H2) $\mu_n = \mathbb{E}\left[M_n - L_n\right] \geq 0.$

Then, for any $n \geq 0$ *and* $\xi_n \in \mathbb{R}$*, we have for any* $\lambda \geq 0$*,*

$$\mathbb{P}\left(W_n < \xi_n\right) \leq \exp\left(-\lambda\left(\mu_n - \xi_n - c_n\right) + \frac{\lambda^2}{2}\left(\mu_n + \eta \sum_{i=1}^{\infty} \mathbb{E}\left[Y_i^{(n)}\right]^2\right)\right).$$

We first need to check that both hypotheses of Lemma 2.11 are satisfied. To do so, we use a similar strategy of [34] and [4] to stochastically dominate the variable $M_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau)$. We have:

$$W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) + M_{k\sqrt{\log n}}\left(A_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}), \tau\right) \overset{\text{law}}{=} M_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau), \tag{15}$$

where $A_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}) = A(n\mathbb{1}_{\mathcal{H}}) \cap \mathcal{R}_{k\sqrt{\log n}}$.

We explain the general idea behind (15). Consider a random walk starting on some site of $\mathcal{H}$ and hitting τ when exiting $\mathcal{R}_{k\sqrt{\log n}}$. Such a walk is accounted for in $M_{k\sqrt{\log n}}$. Now, it is possible for the particle associated to that random walk to also hit τ when exiting $\mathcal{R}_{k\sqrt{\log n}}$. In that case, it is accounted for in $W_{k\sqrt{\log n}}$. If, however, it settles beforehand, say on some site $y \in \mathcal{R}_{k\sqrt{\log n}}$, then we can find a coupling such that the trajectory of the walk starting after the particle has settled is the same as the trajectory of a random walk started on y hitting τ when exiting $\mathcal{R}_{k\sqrt{\log n}}$. Such a term is accounted for in $M_{k\sqrt{\log n}}\left(A_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}), \tau\right)$. This stochastic equality will be of use when applying Lemma 2.11, using it with $M_n = M_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau)$ and

$$\widetilde{M}_n = \widetilde{M}_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) := W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) + M_{k\sqrt{\log n}}\left(A_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}), \tau\right).$$

Now, simply using the fact that $A_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}) \subset \mathcal{R}_{k\sqrt{\log n}}$, we have:

$$\widetilde{M}_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) \leq W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) + M_{k\sqrt{\log n}}\left(\mathcal{R}_{k\sqrt{\log n}}, \tau\right) \quad \text{a.s.} \tag{16}$$

Now, we let $c_n = \#\mathcal{Z} \leq c\left(\sqrt{\log n}\right)^d$, where $c = c(d, b) > 0$. Using (16) gives:

$$\widetilde{M}_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) \leq W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) + M_{k\sqrt{\log n}}\left(\mathcal{R}_{k\sqrt{\log n}} \setminus \mathcal{Z}, \tau\right) + c_n.$$

Note that this inequality is similar to the one in Lemma 2.11. The idea will be to apply Lemma

2.11 with

$$\begin{cases} W_n = W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) \\ M_n = M_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) \\ \tilde{M}_n = \widetilde{M}_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) \\ L_n = M_{k\sqrt{\log n}}\left(\mathcal{R}_{k\sqrt{\log n}} \setminus \mathcal{Z}, \tau\right) \end{cases}$$

Note that both L_n and M_n can be written as sums of independent Bernoulli random variables, as:

$$L_n = \sum_{y \in \mathcal{R}_{k\sqrt{\log n}} \setminus \mathcal{Z}} \mathbb{1}^y_{S(H_{k\sqrt{\log n}}) \in \tau} \quad \text{and} \quad M_n = \sum_{y \in n\mathbb{1}_{\mathcal{H}}} \mathbb{1}^y_{S(H_{k\sqrt{\log n}}) \in \tau},$$

where the indicators $\mathbb{1}^y$ correspond to independent simple symmetric random walks beginning at y, and $H_{k\sqrt{\log n}}$ denotes the hitting time of $\mathcal{R}_{k\sqrt{\log n}}$. Before applying Lemma 2.11, we must ensure that hypotheses **(H1)** and **(H2)** hold. This is done in the next subsection.

Checking hypotheses (H1) and (H2)

Let us begin by giving the following lemma, which ensures hypothesis **(H1)** of Lemma 2.11 for some $b > 0$ in the definition of $\mathcal{Z}$. This lemma is analogous to Lemma 5.1 of [1], and is adapted to the case of *slabs*. We omit its proof.

Lemma 2.12. *There exists a positive constant $\kappa_0 > 0$ such that for all $r > 0$, for any $y \in \mathcal{R}_r$ and $x \in \partial\mathcal{R}_r \setminus \{y\}$, we have*

$$\mathbb{P}_y\left(S(H_r) = x\right) \leq \frac{\kappa_0}{\|x - y\|^{d-1}},$$

where H_r denotes the hitting time of $\partial\mathcal{R}_r$ for the simple random walk $(S(t))_{t\geq 0}$.

This ensures **(H1)**, since for all $y \in \mathcal{R}_{k\sqrt{\log n}} \setminus \mathcal{Z}$,

$$\begin{aligned} \mathbb{E}\left[\mathbb{1}^y_{S(H_{k\sqrt{\log n}}) \in \tau}\right] &= \sum_{x \in \tau} \mathbb{P}_y\left(S(H_{k\sqrt{\log n}}) = x\right) \\ &\leq \sum_{x \in \tau} \frac{\kappa_0}{\|x - y\|^{d-1}} \\ &\leq \#\tau \frac{\kappa_0}{(b\sqrt{\log n})^{d-1}} \\ &\leq \frac{c}{b^{d-1}}, \end{aligned}$$

where the last line comes from the fact that $\#\tau$ is of order $(\sqrt{\log n})^{d-1}$. Thus taking b large enough in the definition of $\mathcal{Z}$ ensures **(H1)**. We now show that hypothesis **(H2)** holds as well. We show that if τ is a tile at a distance $C\sqrt{\log n}$ of $\partial\mathcal{R}_{n/2}$, then for some positive constant $\kappa > 0$, we have

$$\mu(\tau) \geq \kappa C \log(n)^{d/2}. \tag{17}$$

We also show the following:

$$\sum_{y \in \mathcal{R}_{k\sqrt{\log n}} \setminus \mathcal{Z}} \mathbb{P}_y\left(S(H_{k\sqrt{\log n}}) \in \tau\right)^2 \leq c\log(n)^{d-1}. \tag{18}$$

The proofs of (17) and (18) are given in Section 6.3.

Application of Lemma 2.11

We now have all the tools in hand to prove (14). To do so, we use the previous results and Lemma 2.11. Notice that $C = C(b)$ can be taken large enough such that (17) gives $\mu(\tau) \geq 3c_n$. Therefore, applying Lemma 2.11 gives for all $\lambda \geq 0$:

$$\mathbb{P}\left(W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) < \frac{1}{3}\mu(\tau)\right) \leq \exp\left(-\frac{\lambda}{3}\mu(\tau) + \frac{\lambda^2}{2}\left(\mu(\tau) + \eta c \log(n)^{d-1}\right)\right). \tag{19}$$

After optimizing in λ, and using that $\mu(\tau) \geq \kappa C \log(n)^{d/2}$ we get that

$$\mathbb{P}\left(W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) < \frac{1}{3}\mu(\tau)\right) \leq \exp\left(-\kappa' C^2 \log(n)\right),$$

for some constant $\kappa' > 0$. Therefore,

$$\begin{aligned}\mathbb{P}\left(\exists \tau \in \mathcal{T}_{k\sqrt{\log n}},\ W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) < \frac{1}{3}\mu(\tau)\right) &\leq \#\mathcal{T}_{k\sqrt{\log n}} \exp\left(-\kappa' C^2 \log(n)\right) \\ &\leq K n^{\alpha(d-1)} \exp\left(-\kappa' C^2 \log(n)\right),\end{aligned}$$

which for C large enough decreases faster than any power of n^{-1}. The proof of the optimization is given in Section 6.3.

Handling (13)

We now focus on giving an upper bound of (13). The idea here is that if many particles hit a tile, they are very likely to fill the corresponding cell. Note that we have the following inclusion

$$S_k \subset \bigcup_{z \in \partial_{k,n}} \mathcal{C}(z),$$

where $\partial_{k,n} = \partial \mathcal{R}_{k\sqrt{\log n}} \cap \mathbb{Z}_{n^\alpha}$. Now, using the previous inclusion and the fact that $\mu(\tau) \geq \kappa C \log(n)^{d/2}$, we get

$$\begin{aligned}&\mathbb{P}\left(\forall \tau \in \mathcal{T}_{k\sqrt{\log n}},\ W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) \geq \frac{1}{3}\mu(\tau),\ S_k \text{ is not covered}\right) \\ &\quad\leq \mathbb{P}\left(\bigcup_{z \in \partial_{k,n}} \mathcal{C}(z) \text{ is not covered },\ \forall \tau \in \mathcal{T}_{k\sqrt{\log n}},\ W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) \geq \frac{\kappa C}{3} \log(n)^{d/2}\right) \\ &\quad\leq \sum_{z \in \partial_{k,n}} \mathbb{P}\left(\mathcal{C}(z) \text{ is not covered } \middle|\ \forall \tau \in \mathcal{T}_{k\sqrt{\log n}},\ W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau) \geq \frac{\kappa C}{3} \log(n)^{d/2}\right).\end{aligned}$$

The following lemma shows that if many particles are initially contained inside a ball of radius $R/2$, they have a high probability of filling the ball of radius R. (See Lemma 1.3 of [4], which describes the idea mentioned above of floodgates being open and filling up a given area).

Lemma 2.13. *Choose R and A large enough. Assume that $\lfloor AR^d \rfloor$ particles lie initially on $\mathbb{B}(0, R/2)$. We call η the initial configuration of these particles and $A(\eta)$ the aggregate they* *produce. There are positive constants* $\{\kappa_d,\ d \geq 3\}$ *independent of* R *and* A *such that*

$$\mathbb{P}\left(\mathbb{B}(0,R) \not\subset A(\eta)\right) \leq \exp\left(-\kappa_d A R^2\right).$$

In our case, we know that each tile has been hit with at least $\frac{\kappa C}{3}\log(n)^{d/2}$ particles. Now, recall that a cell's radius is twice the radius of a tile, meaning we can apply Lemma 2.13 directly, and get

$$\mathbb{P}\left(\mathcal{C}(z) \text{ is not covered } \middle| \forall \tau \in \mathcal{T}_{k\sqrt{\log n}},\ W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}},\tau) \geq \frac{\kappa C}{3}\log(n)^{d/2}\right)$$
$$\leq \exp\left(-\kappa_d C \log(n)\right).$$

Now, using that

$$\#\partial_{k,n} = \#\left(\partial\mathcal{R}_{k\sqrt{\log n}} \cap \mathbb{Z}_{n^\alpha}\right) = 2 \times \left(2n^\alpha + 1\right)^{d-1} \leq K n^{\alpha(d-1)}$$

we can conclude on the upper bound of (13). We have

$$\mathbb{P}\left(\forall \tau \in \mathcal{T}_{k\sqrt{\log n}},\ W_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}},\tau) \geq \frac{1}{3}\mu(\tau),\ S_k \text{ is not covered }\right)$$
$$\leq K n^{\alpha(d-1)} \exp\left(-\kappa_d C \log(n)\right)$$

which decreases, for C large enough, faster than any given power of n^{-1}. This concludes the proof of the lower bound.

We understand thanks to Lemma 2.13 the source of the sublogarithmic fluctuations for the aggregate. We use Lemma 2.13 in hopes of applying the Borel-Cantelli Lemma and getting a bound decreasing faster than any power of n. The bound we obtain is of order $\exp\left(-\kappa_d A R^2\right)$, hence choosing R of order $\sqrt{\log n}$ means $\exp\left(-\kappa_d A R^2\right)$ is of order $n^{-\kappa A}$, with $\kappa > 0$ and $A = \Theta(1)$ (*i.e.* if there exist constants such that $c_1 \leq A \leq c_2$). We therefore pick shells of size $R \approx \sqrt{\log n}$, leading to sublogarithmic fluctuations. Note that the bound in Lemma 2.13 for $d = 2$ is equal to $\exp\left(-\kappa_2 \frac{AR^2}{\log R}\right)$, so choosing $R \approx \sqrt{\log n}$ no longer grants a functional bound for the Borel-Cantelli Lemma.

6.2 Proof for the upper bound

In this section, we prove the upper bound of Theorem 2.1, that is: for any integers $d \geq 3$ and $\alpha \geq 1$, there exists $C > 0$ such that, almost surely, there exists $N \geq 1$ such that for any $n \geq N$,

$$A_n[\infty] \cap \mathbb{Z}_{n^\alpha} \subset \mathcal{R}_{n/2 + C\sqrt{\log n}} \cap \mathbb{Z}_{n^\alpha}.$$

The proof follows the same lines as in [16]. Fix $\alpha \geq 1$ and take $C > 0$.

From the Borel-Cantelli lemma, it is sufficient to prove that

$$\mathbb{P}\left(A_n[\infty] \cap \mathbb{Z}_{n^\alpha} \not\subset \mathcal{R}_{n/2 + C\sqrt{\log n}} \cap \mathbb{Z}_{n^\alpha}\right) \tag{20}$$

is smaller than any power of n^{-1}. To do so, we define the random variable

$$X(n) = \max\left\{|z_1|,\ z \in A_n[\infty] \cap \mathbb{Z}_{n^\alpha}\right\}.$$

Now, notice that we can bound (20) by $\mathbb{P}\left(X(n) > \frac{n}{2} + C\sqrt{\log n}\right)$. Hence, we write :

$$\mathbb{P}\left(X(n) > \frac{n}{2} + C\sqrt{\log n}\right) \leq \mathbb{P}\left(\exists z \in A_n[\infty] \cap \mathbb{Z}_{n^\alpha},\ X(n) = |z_1| > \frac{n}{2} + C\sqrt{\log n}\right)$$
$$\leq \sum_{z \in \mathbb{Z}_{n^\alpha}} \mathbb{P}\left(z \in A_n[\infty],\ X(n) = |z_1| > \frac{n}{2} + C\sqrt{\log n}\right).$$

Now, since $A_n[\infty]$ is translation-invariant (with respect to $T_k,\ k \in \mathcal{H}$), this quantity only depends on the first coordinate of z. For symmetry reasons, we can focus only on the case where $z_1 \geq 0$. So for $l \geq 0$, we define $z(l)$ as $z(l) := (l + \frac{n}{2}, 0, \dots, 0)$, and write:

$$\mathbb{P}\left(X(n) > \frac{n}{2} + C\sqrt{\log n}\right)$$
$$\leq 2(2n^\alpha + 1)^{d-1}\mathbb{P}\left(\exists l > C\sqrt{\log n},\ z(l) \in A_n[\infty],\ X(n) = l + \frac{n}{2}\right)$$
$$\leq 2(2n^\alpha + 1)^{d-1} \sum_{l \geq C\sqrt{\log n}} \mathbb{P}\left(z(l) \in A_n[\infty],\ X(n) = l + \frac{n}{2}\right).$$

Now, fix $l \geq C\sqrt{\log(n)}$. When bounding $\mathbb{P}\left(z(l) \in A_n[\infty],\ X(n) = l + \frac{n}{2}\right)$, let us split this probability into two parts, claiming that if $z(l)$ belongs to the aggregate, either a thin tentacle of settled particles branches out to that point, or there are many particles settled in a ball around $z(l)$. We will once again be using Lemma 2 of [31], which was previously used in Section 4. We write

$$\mathbb{P}\left(z(l) \in A_n[\infty],\ X(n) = z_1(l) = l + \frac{n}{2}\right)$$
$$\leq \mathbb{P}\left(\#\left(A_n[\infty] \cap \mathbb{B}(z(l), l)\right) > \beta l^d,\ X(n) = z_1(l) = l + \frac{n}{2}\right)$$
$$+ \mathbb{P}\left(z(l) \in A_n[\infty],\ \#\left(A_n[\infty] \cap \mathbb{B}(z(l), l)\right) \leq \beta l^d\right), \quad (21)$$

where β is chosen as in Lemma 2 of [31], allowing us to handle the last term of (21). This gives

$$\mathbb{P}\left(z(l) \in A_n[\infty],\ \#\left(A_n[\infty] \cap \mathbb{B}(z(l), l)\right) \leq \beta l^d\right) \leq K_0 e^{-cl^2}. \quad (22)$$

We write

$$\sum_{l \geq C\sqrt{\log n}} K_0 e^{-cl^2} \leq e^{-\frac{cC^2 \log n}{2}} \sum_{l \geq C\sqrt{\log n}} e^{-\frac{cl^2}{2}} \leq K e^{-\frac{cC^2 \log n}{2}},$$

which, by taking C large enough, is smaller than any power of n^{-1}. Let us now define a new random aggregate, namely $\tilde{A}$, built using the following protocol. For an initial configuration η, we throw particles according to the usual IDLA protocol, but we freeze all particles once they reach $\partial\mathcal{R}_{n/2}$. Once all particles have been thrown from η, we unfreeze the particles stationed on $\partial\mathcal{R}_{n/2}$ and let them continue their trajectory (until they exit the aggregate).

From the Abelian property, we know that $\tilde{A}(\eta) \overset{\text{law}}{=} A(\eta)$. In particular, we have that

$$\#\left(A_n[\infty] \cap \mathbb{B}(z(l), l)\right) \overset{\text{law}}{=} \#\left(\tilde{A}_n[\infty] \cap \mathbb{B}(z(l), l)\right), \quad (23)$$

where $\tilde{A}_n[\infty]$ is obtained by taking the increasing union over $M \geq 0$ of $\tilde{A}_n[M]$. To handle the first term of the right-hand side of (21), we follow the same method as in [16]. For an initial configuration η and a set $B \subset \mathbb{Z}^d$, with respect to the aggregate $\tilde{A}(\eta)$, we denote by $M^*_{n/2+l}(\eta, B)$ the number of random walks with initial configuration η and which satisfy the following conditions:

- The random walks intersect B before they exit $\mathcal{R}_{n/2+l}$
- The particles associated with the random walks hit $\partial\mathcal{R}_{n/2}$

The path of such a random walk is illustrated by Figure 2.5 below: the aggregate is contained inside the dashed blue lines, the particle's path is represented by the red line, and the associated random walk's path by the dashed red line.
Just as we did for $A_n[\infty]$, we can define the random variable

$$\tilde{X}(n) = \max\left\{|z_1|,\ z \in \tilde{A}_n[\infty] \cap \mathbb{Z}_{n^\alpha}\right\}.$$

Since $\tilde{A}_n[\infty] \overset{\text{law}}{=} A_n[\infty]$, we have that $\tilde{X}(n) \overset{\text{law}}{=} X(n)$.
Notice that on the event $\{\tilde{X}(n) = z_1(l)\}$, the aggregate's furthermost point (on the x axis) is therefore $z_1(l)$ and has a coordinate equal to $z_1(l) = l + \frac{n}{2}$. This implies the following inequality:

$$\#\left(\tilde{A}_n[\infty] \cap \mathbb{B}(z(l), l)\right) \overset{\text{a.s.}}{\leq} M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}_{n^\gamma}}, \mathbb{B}^-(z(l), l)\right), \tag{24}$$

where $\mathbb{B}^-(z(l), l) := \mathbb{B}(z(l), l) \cap \mathcal{R}_{n/2+l}$. Indeed, on the event $\tilde{X}(n) = z_1(l)$, the aggregate cannot go any further than $\mathcal{R}_{n/2+l}$ and therefore any particle of $\tilde{A}_n[\infty] \cap \mathbb{B}(z(l), l)$ necessarily hits $\partial\mathcal{R}_{n/2}$ before exiting the aggregate, and the random walk associated to that particle also necessarily intersected $\mathbb{B}^-(z(l), l)$ before exiting $\mathcal{R}_{n/2+l}$. The crucial point is that no particle can escape from $\mathcal{R}_{n/2+l}$ on this event. This implies inequality (24). Combining this with (23), we have that

$$\#\left(A_n[\infty] \cap \mathbb{B}(z(l), l)\right) \overset{\text{sto.}}{\leq} M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}_{n^\gamma}}, \mathbb{B}^-(z(l), l)\right). \tag{25}$$

Therefore,

$$\mathbb{P}\left(\#\left(A_n[\infty] \cap \mathbb{B}(z(l), l)\right) > \beta l^d,\ z_1(l) = X(n)\right) \leq \mathbb{P}\left(M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}_{n^\gamma}}, \mathbb{B}^-(z(l), l)\right) > \beta l^d\right).$$

The trajectories of the walks counted by M^* are not independent. This comes from the fact that the *particles* with which they are associated are killed upon exiting the current aggregate. Since the trajectories of the particles are highly dependent, this remains true for M^*. However, we only count walks for which the associated particle has hit $\partial\mathcal{R}_{n/2}$ before exiting the aggregate, meaning that the random walks evolve *independently* after hitting $\partial\mathcal{R}_{n/2}$. In particular, they evolve independently once they reach $\mathbb{B}^-(z(l), l)$. We know that the walks counted my M^* are independent after they reach $\partial\mathcal{R}_{n/2}$, so in particular after reaching $\partial\mathbb{B}^-(z(l), l)$. Recall that a random walk started in $x \in \partial\mathbb{B}^-(z(l), l)$ has probability at least $\rho > 0$ to hit $\mathbb{B}(z(l), 2l)^c \cap \mathcal{R}^c_{n/2+l}$ when it exits $\mathbb{B}(z(l), 2l)$. It is important to note that ρ does not depend on n or l. In particular, random walks counted in $M^*_{n/2+l}$ have probability at least ρ of hitting tile $\tilde{\tau}(z(l))$, where

$$\tilde{\tau}(z(l)) = \mathbb{B}(z(l), 2l) \cap \partial\mathcal{R}_{n/2+l}.$$

We now split our probability into two, conditioning on the event where a sufficient amount of

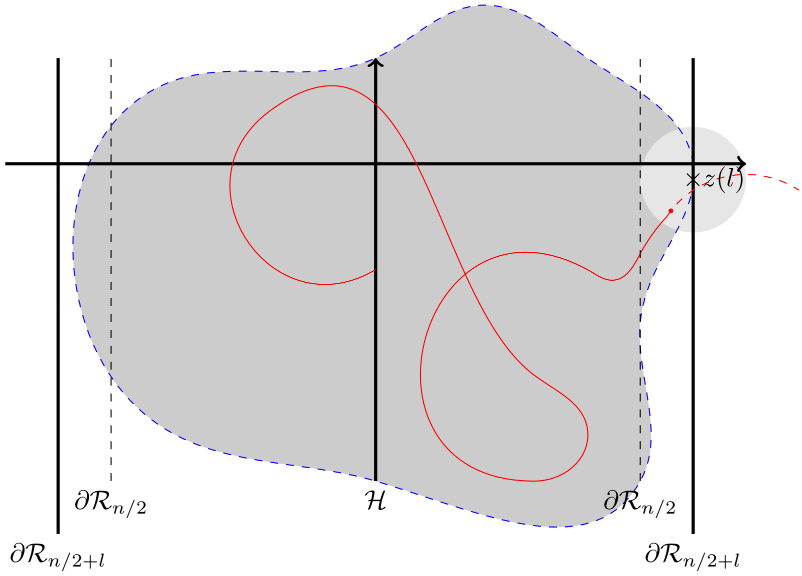


Figure 2.5: An example of a walk in $M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}_{n^\gamma}}, \mathbb{B}(z(l),l)\right)$

random walks of $M^*_{n/2+h(n)}$ have hit the tile $\tilde{\tau}(z(l))$ defined above. This gives

$$\begin{aligned}
&\mathbb{P}\left(M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}_{n^\gamma}}, \mathbb{B}^-(z(l),l)\right) > \beta l^d\right) \\
&\qquad \le \mathbb{P}\left(M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}_{n^\gamma}}, \tilde{\tau}(z(l))\right) \le \frac{\beta l^d \rho}{2} \;\middle|\; M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}_{n^\gamma}}, \mathbb{B}^-(z(l),l)\right) > \beta l^d\right) \\
&\qquad\qquad + \mathbb{P}\left(M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}_{n^\gamma}}, \tilde{\tau}(z(l))\right) > \frac{\beta l^d \rho}{2}\right). \qquad (26)
\end{aligned}$$

From what we said above, we know that the random walks are independent after they hit $\partial\mathbb{B}^-(z(l),h(n))$, and that any walk started from a point $x \in \partial\mathbb{B}^-(z(l),l)$ has probability at least ρ to hit $\tilde{\tau}(z(l))$.

Therefore, conditional on the event $\left\{M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}_{n^\gamma}}, \mathbb{B}^-(z(l),l)\right) > \beta l^d\right\}$, the random variable $M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}_{n^\gamma}}, \tilde{\tau}(z(l))\right)$ stochastically dominates a binomial distribution with parameters βl^d and ρ, denoted by $\mathcal{B}\left(\beta l^d, \rho\right)$. This allows us to handle the first term in (26). Indeed, using the fact that $\mathbb{E}\left[\mathcal{B}\left(\beta l^d, \rho\right)\right] = \beta l^d \rho$, standard concentration inequality theory (e.g [12]) gives:

$$\mathbb{P}\left(\mathcal{B}\left(\beta l^d, \rho\right) \le \frac{\beta l^d \rho}{2}\right) \le \exp\left(-\frac{\beta l^d \rho}{8}\right).$$

Therefore, using the stochastic domination we mentioned above,

$$\mathbb{P}\left(M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}_{n^\gamma}}, z(l), \tilde{\tau}(z(l))\right) \leq \frac{\beta l^d \rho}{2} \,\middle|\, M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}_{n^\gamma}}, \mathbb{B}^-(z(l),l)\right) > \beta l^d\right) \leq \exp\left(-\frac{\beta l^d \rho}{8}\right).$$

Proceeding as with (22), we can show that

$$\sum_{l \geq C\sqrt{\log n}} \exp\left(-\frac{\beta l^d \rho}{8}\right) \leq K \exp\left(-\frac{\beta C^d \log(n)^{d/2}}{16}\right),$$

which, for C taken large enough, decreases faster than any power of n^{-1}.

Switching our focus to the second term (26), it remains to prove that for all $L > 0$ and n large,

$$\sum_{l \geq C\sqrt{\log n}} \mathbb{P}\left(M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}_{n^\gamma}}, \tilde{\tau}(z)\right) > \frac{\beta l^d \rho}{2}\right) \leq n^{-L}.$$

We will know be using a lemma similar to Lemma 2.11, which we give below.

Lemma 2.14. *(Lemma 2.5 of [4]) Suppose that a sequence of random variables $W_n,\ L_n,\ M_n,\ \widetilde{M}_n$ and an event $\mathcal{A}_n$ satisfy for each $n \in \mathbb{N}$:*

$$(W_n + L_n)\mathbb{1}_{\mathcal{A}_n} \underset{sto}{\leq} \widetilde{M}_n \quad \text{and} \quad M_n \overset{\text{law}}{=} \widetilde{M}_n.$$

Assume that W_n and L_n are independent, and that L_n and M_n are series of independent Bernoulli random variables with finite expectations, with $L_n = \sum_{i\geq 0} Y_i^{(n)}$. Finally, assume that $\mu_n = \mathbb{E}[M_n] - \mathbb{E}[L_n] \geq 0$.
Then, for all $n \geq 0$, $\xi_n \in \mathbb{R}$ and $\lambda \in [0, \log 2]$,

$$\mathbb{P}(W_n \geq \xi_n, \mathcal{A}_n) \leq \exp\left(-\lambda(\xi_n - \mu_n) + \lambda^2\left(\mu_n + 4\sum_{i\geq 0} \mathbb{E}{Y_i^{(n)}}^2\right)\right).$$

Using the same arguments as for (15), we can establish the following equality:

$$M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \tilde{\tau}(z(l))\right) + M_{n/2+l}\left(A_{n/2}(n\mathbb{1}_{\mathcal{H}}), \tilde{\tau}(z(l))\right) \overset{\text{law}}{=} \tilde{M}_n, \tag{27}$$

where $A_{n/2}(n\mathbb{1}_{\mathcal{H}}) = A(n\mathbb{1}_{\mathcal{H}}) \cap \mathcal{R}_{n/2}$, and $\tilde{M}_n$ is an independent copy of $M_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \tilde{\tau}(z(l))\right)$.

The idea here is to once again consider a walk counted by $M_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \tilde{\tau}(z(l))\right)$, and consider the trajectory of its associated particle. Either the particle has hit $\partial\mathcal{R}_{n/2}$ before exiting the aggregate, and is therefore counted by $M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \tilde{\tau}(z(l))\right)$, or the particle has settled before on some site x. In that case, we can launch a new random walk from $x \in A_{n/2}(n\mathbb{1}_{\mathcal{H}})$, which is accounted for by $M_{n/2+l}\left(A_{n/2}(n\mathbb{1}_{\mathcal{H}}), \tilde{\tau}(z(l))\right)$. Now, take $\alpha' > \alpha$. Let us denote by $\delta_I(n)$ the inner error of $A_{n/2}(n\mathbb{1}_{\mathcal{H}})$ on $\mathbb{Z}_{n^{\alpha'}}$, that is

$$\delta_I(n) = \max\left\{\frac{n}{2} - |z_1|,\ z \in \left(\mathcal{R}_{n/2} \setminus A_{n/2}(n\mathbb{1}_{\mathcal{H}})\right) \cap \mathbb{Z}_{n^{\alpha'}}\right\}.$$

This quantity is illustrated in Figure 2.6.

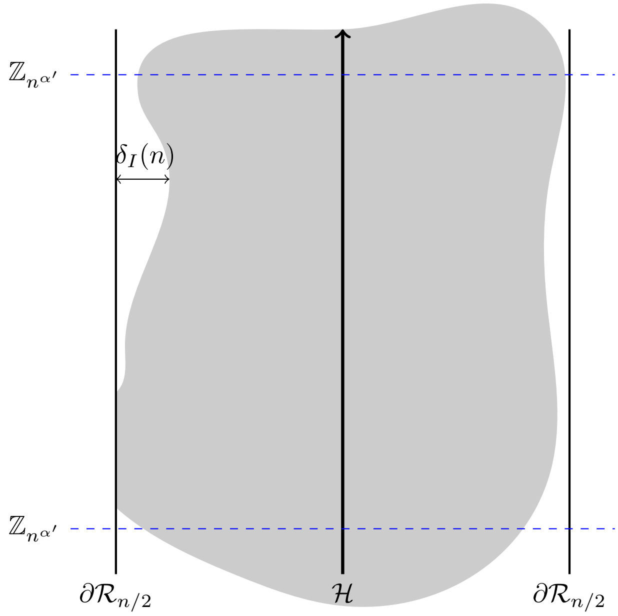


Figure 2.6: Illustration of the inner error

We will be using the work we did in the previous section to bound this inner error. Using the definition of $\delta_I(n)$, we have $\mathcal{R}_{n/2-\delta_I(n)} \cap \mathbb{Z}_{n^{\alpha'}} \subset A_{n/2}(n\mathbb{1}_{\mathcal{H}})$, so combining this with (27), we get

$$M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \tilde{\tau}(z(l))\right) + M_{n/2+l}\left(\mathcal{R}_{n/2-\delta_I(n)} \cap \mathbb{Z}_{n^{\alpha'}}, \tilde{\tau}(z(l))\right) \overset{\text{sto.}}{\leq} \tilde{M}_n.$$

Now, for some $\alpha_d > 0$ which will be chosen later, on the event $\left\{\delta_I(n) \leq \frac{\alpha_d l}{2C}\right\}$, we have $\mathcal{R}_{n/2-\frac{\alpha_d l}{2C}} \cap \mathbb{Z}_{n^{\alpha'}} \subset \mathcal{R}_{n/2-\delta_I(n)} \cap \mathbb{Z}_{n^{\alpha'}}$, with the convention that $\mathcal{R}_m = \mathbb{1}_{\mathcal{H}}$ when $m \leq 0$. This gives:

$$\left(M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \tilde{\tau}(z(l))\right) + M_{n/2+l}\left(\mathcal{R}_{n/2-\frac{\alpha_d l}{2C}} \cap \mathbb{Z}_{n^{\alpha'}}, \tilde{\tau}(z)\right)\right) \mathbb{1}_{\delta_I(n) \leq \frac{\alpha_d l}{2C}} \overset{\text{sto.}}{\leq} \tilde{M}_n.$$

This stochastic inequality is similar to the one required by Lemma 2.14. We can directly apply this lemma with

$$\begin{cases} W_n = M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \tilde{\tau}(z(l))\right) \\ L_n = M_{n/2+l}\left(\mathcal{R}_{n/2-\frac{\alpha_d l}{2C}} \cap \mathbb{Z}_{n^{\alpha'}}, \tilde{\tau}(z(l))\right) \\ \widetilde{M}_n = M_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \tilde{\tau}(z(l))\right) \\ \mathcal{A}_n = \left\{\delta_I(n) \leq \frac{\alpha_d l}{2C}\right\} \\ \xi_n = \frac{\beta l^d \rho}{2} \end{cases}$$

We show in Section 6.3 that the hypotheses in Lemma 2.14 hold. Now, we write

$$\mathbb{P}\left(M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \tilde{\tau}(z(l))\right) > \frac{\beta l^d \rho}{2}\right) \\ \leq \mathbb{P}\left(M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \tilde{\tau}(z(l))\right) > \frac{\beta l^d \rho}{2},\ \delta_I(n) \leq \frac{\alpha_d l}{2C}\right) + \mathbb{P}\left(\delta_I(n) > \frac{\alpha_d l}{2C}\right).$$

We first focus our attention on the second term. We will be using our computations from the proof of the lower bound to control this term. We have:

$$\begin{aligned}\sum_{l \geq C\sqrt{\log n}} \mathbb{P}\left(\delta_I(n) > \frac{\alpha_d l}{2C}\right) &\leq \sum_{l \geq \sqrt{\log n}} \mathbb{P}\left(\delta_I(n) > \frac{\alpha_d l}{2}\right) \\ &\leq \sum_{l \geq \sqrt{\log n}} \exp(-\kappa \alpha_d l^2) \\ &\leq K \exp\left(-\frac{\kappa \alpha_d \log n}{2}\right).\end{aligned}$$

Taking α_d sufficiently large, this term becomes smaller than any given power of n^{-1}.

We now shift our focus to the first term of the sum. Fix α_d large enough so that the previous term is smaller than any power of n^{-1}.

After an application of Lemma 2.14 and an optimization detailed in Section 6.3, we get for some constant $\kappa > 0$, if C is chosen large enough:

$$\mathbb{P}\left(M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \mathbb{B}(z(l), \tilde{\tau}(z(l)))\right) > \frac{\beta l^d \rho}{2},\ \delta_I(n) \leq \frac{\alpha_d l}{2C}\right) \leq \exp\left(-\kappa l^2\right). \tag{28}$$

Hence,

$$\begin{aligned}\sum_{l \geq C\sqrt{\log n}} \mathbb{P}\left(M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \mathbb{B}(z(l), \tilde{\tau}(z(l)))\right) > \frac{\beta l^d \rho}{2},\ \delta_I(n) \leq \frac{\alpha_d l}{2C}\right) &\leq \sum_{l \geq C\sqrt{\log n}} \exp\left(-\kappa l^2\right) \\ &\leq K \exp\left(-\frac{\kappa C^2 \log n}{2}\right),\end{aligned}$$

Which goes to zero faster than any power of n^{-1}, given C is large enough, and concludes the proof of our theorem.

6.3 Auxiliary proofs

Proofs from the lower bound

We start by giving the proof of (17). To do it, we apply the following lemma, which is a simple extension of Lemma 6.4 of [16].

Lemma 2.15. *Let $r \leq r'$ and let $\tau \subset \partial \mathcal{R}_{r'}$ be finite. Then*

$$\mathbb{E}\left[M_{r'}(\mathcal{R}_r, \tau)\right] = \frac{2r+1}{2} \# \tau.$$

In particular, for any $r' \geq 1$,

$$\mathbb{E}\left[M_{r'}(\mathbb{1}_{\mathcal{H}}, \tau)\right] = \frac{\#\tau}{2}. \tag{29}$$

Now, to get (17), we write

$$\begin{aligned}
\mu(\tau) &= \mathbb{E}\left[M_{k\sqrt{\log n}}(n\mathbb{1}_{\mathcal{H}}, \tau)\right] - \mathbb{E}\left[M_{k\sqrt{\log n}}\left(\mathcal{R}_{k\sqrt{\log n}} \setminus \mathcal{Z}, \tau\right)\right] \\
&\geq n\mathbb{E}\left[M_{k\sqrt{\log n}}(\mathbb{1}_{\mathcal{H}}, \tau)\right] - \mathbb{E}\left[M_{k\sqrt{\log n}}\left(\mathcal{R}_{k\sqrt{\log n}}, \tau\right)\right] \\
&\geq \frac{\#\tau}{2}\left(n - 2k\sqrt{\log n} - 1\right) \\
&\geq cC\left(\sqrt{\log n}\right)^d.
\end{aligned}$$

We now show the bound given by (18). Recall that k is taken such that $k \leq \frac{n}{2\sqrt{\log n}} - C$. Using Lemma 2.12, we have

$$\begin{aligned}
\mathbb{P}_y\left(S(H_{k\sqrt{\log n}}) \in \tau\right) &\leq \#\tau \max_{x\in\tau} \mathbb{P}_y\left(S(H_{k\sqrt{\log n}}) = x\right) \\
&\leq \#\tau \max_{x\in\tau} \frac{\kappa}{\|y-x\|^{d-1}} \\
&\leq \#\tau \frac{c}{\|y-z\|^{d-1}}.
\end{aligned}$$

Hence

$$\sum_{y\in\mathcal{R}_{k\sqrt{\log n}}\setminus\mathcal{Z}} \mathbb{P}_y\left(S(H_{k\sqrt{\log n}}) \in \tau\right)^2 \leq (c\#\tau)^2 \sum_{y\in\mathcal{R}_{k\sqrt{\log n}}\setminus\mathcal{Z}} \frac{1}{\left(\|y-z\|^2\right)^{d-1}}.$$

Since $\#\tau$ is of order $(\sqrt{\log n})^{d-1}$, it suffices to show that $\sum_{y\in\mathcal{R}_{k\sqrt{\log n}}\setminus\mathcal{Z}} \frac{1}{(\|y-z\|^2)^{d-1}}$ is bounded by a finite constant, and therefore that

$$\sum_{j=1}^{2k\sqrt{\log n}} \int_{[1,\infty[^{d-1}} \frac{\mathrm{d}x_1 \ldots \mathrm{d}x_{d-1}}{\left(j^2 + x_1^2 + \ldots x_{d-1}^2\right)^{d-1}}$$

is as well. This is true since

$$\begin{aligned}
\sum_{j=1}^{2k\sqrt{\log n}} \int_{[1,\infty[^{d-1}} \frac{\mathrm{d}x_1 \ldots \mathrm{d}x_{d-1}}{\left(j^2 + x_1^2 + \ldots x_{d-1}^2\right)^{d-1}} &\leq c_d \sum_{j=1}^{2k\sqrt{\log n}} \int_1^\infty \frac{r^{d-2}}{\left(j^2+r^2\right)^{d-1}}\,\mathrm{d}r \\
&\leq c_d \sum_{j=1}^{\infty} \int_{1/j}^\infty \frac{j^{d-2}r^{d-2}}{j^{2(d-1)}\left(1+r^2\right)^{d-1}} j\mathrm{d}r \\
&\leq c_d \sum_{j=1}^{\infty} \frac{1}{j^{d-1}} \int_0^\infty \frac{r^{d-2}}{\left(1+r^2\right)^{d-1}}\mathrm{d}r,
\end{aligned} \tag{30}$$

where c_d denotes the volume of the $(d-2)$-dimensional sphere. Now, since $d \geq 3$, we have that

(30) is finite. Thus, for some $c > 0$,

$$\sum_{y \in \mathcal{R}_{k\sqrt{\log n}} \setminus \mathcal{Z}} \mathbb{P}_y \left(S(H_{k\sqrt{\log n}}) \in \tau \right)^2 \leq c \log(n)^{d-1}.$$

Optimization in (19)**:** In this section, we detail the computations of the optimization in (19). In all that follows, κ' denotes a generic constant.

We wish to minimize $\lambda \mapsto \exp\left(-\frac{\lambda}{3}\mu(\tau) + \frac{\lambda^2}{2}\left(\mu(\tau) + \eta \log(n)^{d-1}\right)\right)$ on $\mathbb{R}_+$. Recall from (17) that $\mu(\tau) \geq \kappa C \log(n)^{d/2}$. Note that it suffices to minimize the function within the exponential, which happens to be a second degree polynomial. Pick λ^* minimizing the polynomial, that is

$$\lambda^* = \frac{\mu(\tau)}{3\left(\mu(\tau) + \eta \log(n)^{d-1}\right)}.$$

This gives

$$\begin{aligned} -\frac{\lambda^*}{3}\mu(\tau) + \frac{(\lambda^*)^2}{2}\left(\mu(\tau) + \eta \log(n)^{d-1}\right) &\leq -\frac{\mu(\tau)^2}{18\left(\mu(\tau) + \eta \log(n)^{d-1}\right)} \\ &\leq -\frac{\mu(\tau)}{18\left(1 + \frac{\eta \log(n)^{d-1}}{\mu(\tau)}\right)}. \end{aligned}$$

Now,

$$1 + \frac{\eta \log(n)^{d-1}}{\mu(\tau)} \leq 1 + \frac{\eta \log(n)^{d-1}}{\kappa C \log(n)^{d/2}} \leq \frac{\kappa'}{C} \log(n)^{d/2-1}.$$

Therefore, combining this with the fact that $\mu(\tau) \geq \kappa C \log(n)^{d/2}$, we get

$$\begin{aligned} -\frac{\mu(\tau)}{18\left(1 + \frac{\eta \log(n)^{d-1}}{\mu(\tau)}\right)} &\leq -\frac{C\mu(\tau)}{\kappa' \log(n)^{d/2-1}} \\ &\leq -\frac{\kappa C^2 \log(n)^{d/2}}{\kappa' \log(n)^{d/2-1}} \\ &\leq -\kappa' C^2 \log n. \end{aligned}$$

Proofs from the upper bound

We now show proofs concerning the results for the upper bound of Theorem 2.1. Let us begin by showing that the hypotheses in Lemma 2.14 hold. In all that follows, we fix $l \geq C\sqrt{\log n}$. Recall that $z(l)$ is such that $z(l) = (\frac{n}{2} + l, 0, \ldots, 0)$.

Hypotheses of Lemma 2.14: We first need to check that $\mu_n = \mathbb{E}[M_n] - \mathbb{E}[L_n] \geq 0$. This once again uses Lemma 2.15. To lighten notation, we define $\mathcal{R}(l, n, \alpha_d, C) := \mathcal{R}_{n/2 - \frac{\alpha_d l}{2C}}$. We have:

$$\begin{aligned}\mu_n &= \mathbb{E}\left[M_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \tilde{\tau}(z(l))\right)\right] - \mathbb{E}\left[M_{n/2+l}\left(\mathcal{R}(l,n,\alpha_d,C)\cap \mathbb{Z}_{n^{\alpha'}}, \tilde{\tau}(z(l))\right)\right] \\ &\geq \mathbb{E}\left[M_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \tilde{\tau}(z(l))\right)\right] - \mathbb{E}\left[M_{n/2+l}\left(\mathcal{R}(l,n,\alpha_d,C), \tilde{\tau}(z(l))\right)\right] \\ &= \frac{\#\tilde{\tau}(z(l))}{2}\left(n - 2\left(n/2 - \min\left(\frac{\alpha_d l}{2C},\ n/2\right)\right) - 1\right) \\ &= \frac{\#\tilde{\tau}(z(l))}{2}\left(\min\left(\frac{\alpha_d l}{C},\ n\right) - 1\right).\end{aligned}$$

In the case where $\frac{\alpha_d l}{C} \geq n$, then the minimum is n, and so $\mu_n \geq cn\#\tilde{\tau}(z(l))$ for some positive constant c. Since $\#\tilde{\tau}(z(l))$ is of order l^{d-1} we have that $\mu_n \geq cl^{d-1}$.

In the other case, the minimum is $\frac{\alpha_d l}{C}$, and so in this case $\mu_n \geq c\#\tilde{\tau}(z(l))\frac{\alpha_d l}{C} = \frac{c\alpha_d}{C}l^d$, for some positive constant c. In both cases, the condition $\mu_n \geq 0$ is verified.

Now, following the same computations as the ones used to get (18), we get

$$\begin{aligned}\sum_{y\in\mathcal{R}(l,n,\alpha_d,C)\cap\mathbb{Z}_{n^{\alpha'}}} \mathbb{P}_y\left(S(H_{n/2+l}) \in \tilde{\tau}(z(l))\right)^2 &\leq \sum_{y\in\mathcal{R}(l,n,\alpha_d,C)} \mathbb{P}_y\left(S(H_{n/2+l}) \in \tilde{\tau}(z(l))\right)^2 \\ &\leq cl^{2d-2}.\end{aligned}$$

Control of μ_n: We showed just above that the hypotheses of Lemma 2.14 were satisfied. However, when applying this lemma and optimizing, we need an upper bound on μ_n, given by the following proposition.

Lemma 2.16. *There exist positive constants c, C_0 such that for n large enough, for all $l > C\sqrt{\log n}$,*

$$\mu_n \leq \frac{c\alpha_d l^d}{C} + C_0.$$

Proof of Lemma 2.16: Here, we must once again consider two cases. When $n \leq \frac{\alpha_d l}{C}$, then we write

$$\begin{aligned}\mu_n &= \mathbb{E}\left[M_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \tilde{\tau}(z(l))\right)\right] - \mathbb{E}\left[M_{n/2+l}\left(\mathcal{R}(l,n,\alpha_d,C)\cap \mathbb{Z}_{n^{\alpha'}}, \tilde{\tau}(z(l))\right)\right] \\ &\leq \mathbb{E}\left[M_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \tilde{\tau}(z(l))\right)\right] \\ &= n\frac{\#\tilde{\tau}(z(l))}{2} \\ &\leq cl^{d-1}n \\ &\leq cl^{d-1}\cdot\frac{\alpha_d l}{C} \\ &= \frac{c\alpha_d l^d}{C}\end{aligned}$$

In this case, $C_0 = 0$.

Now, consider the other case where $n > \frac{\alpha_d l}{C}$. We write

$$\begin{aligned}\mu_n &= \mathbb{E}\left[M_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \tilde{\tau}(z(l))\right)\right] - \mathbb{E}\left[M_{n/2+l}\left(\mathcal{R}(l,n,\alpha_d,C), \tilde{\tau}(z(l))\right)\right] \\ &\quad + \mathbb{E}\left[M_{n/2+l}\left(\mathcal{R}(l,n,\alpha_d,C), \tilde{\tau}(z(l))\right)\right] - \mathbb{E}\left[M_{n/2+l}\left(\mathcal{R}(l,n,\alpha_d,C)\cap \mathbb{Z}_{n^{\alpha'}}, \tilde{\tau}(z(l))\right)\right].\end{aligned}$$

The first line has already been computed above, and can be bounded from above by $\frac{c\alpha_d}{C}l^d$, for some $c > 0$. We now shift our focus on the second line. Notice that this quantity is equal to:

$$\mathbb{E}\left[M_{n/2+l}\left(\mathcal{R}(l,n,\alpha_d,C)\cap\mathbb{Z}^c_{n^{\alpha'}},\tilde{\tau}(z(l))\right)\right], \tag{31}$$

and it remains to show that this is bounded uniformly on n.
Now, recall that $z(l) \in \mathbb{Z}_{n^\alpha}$. Walks counted by $M_{n/2+l}\left(\mathcal{R}(l,n,\alpha_d,C)\cap\mathbb{Z}^c_{n^{\alpha'}},\tilde{\tau}(z(l))\right)$ necessarily stay within $\mathcal{R}_{n/2+l}$ between levels $n^{\alpha'}$ and n^α before exiting through $\tilde{\tau}(z(l))$. We can therefore use a donut argument (just like in the proof of Theorem 2.2) between levels n^α and levels $n^{\alpha'}$ and beyond, by building hypercubes of length $n+2l$ between these levels. The length of our hypercubes is imposed to us by the width of $\mathcal{R}_{n/2+l}$, which equals $n+2l$. Since we want the walk to stay inside the slab, we choose our boxes to have the same width. We illustrate this argument with Figure 2.7. Now, for a walk started on some site of $\mathcal{R}(l,n,\alpha_d,C)\cap\mathbb{Z}^c_{n^{\alpha'}}$ to exit $\mathcal{R}_{n/2+l}$

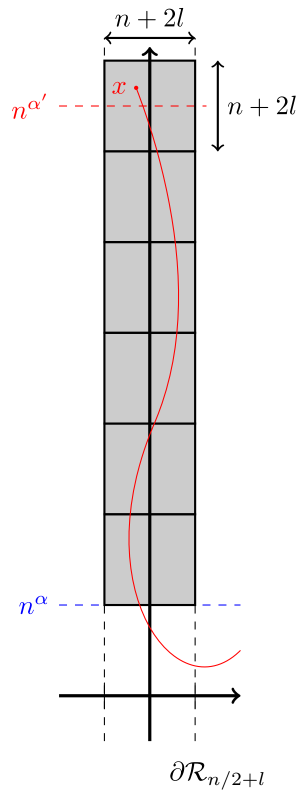


Figure 2.7: Illustration of the box method

through $\tilde{\tau}(z(l))$, it necessarily has to cross a certain amount of boxes. Just like in Proposition 2.6, we use the same reasoning to say that:

$$\forall k \in \mathbb{N},\ \mathbb{P}\left(\text{the walk goes through at least } k \text{ boxes}\right) \leq (1-c)^k,$$

with $c = \frac{1}{4d^2}$. We need to determine the minimum amount of cubes of length $n+2l$ that can fit between n^α and $n^{\alpha'}$. This is equal to $\frac{n^{\alpha'}-n^\alpha}{n+2l}$, which is greater than n given n is large enough and $\alpha' > \max(\alpha, 2)$. This uses the fact that $n+2l$ is of order n. Let us number these boxes by

$B_0,\ B_1, \dots$ starting from level n^α. We have:

$$\mathcal{R}(l,n,\alpha_d,C)\cap \mathbb{Z}^c_{n^{\alpha'}} \subset \bigcup_{k\geq n} B_k.$$

Hence:

$$\begin{aligned}
&\mathbb{E}\left[M_{n/2+l}\left(\mathcal{R}(l,n,\alpha_d,C)\cap \mathbb{Z}^c_{n^{\alpha'}}, \tilde{\tau}(z(l))\right)\right] \\
&\qquad = \mathbb{E}\left[M_{n/2+l}\left(\bigcup_{k\geq n} B_k, \tilde{\tau}(z(l))\right)\right] \\
&\qquad \leq \sum_{k\geq n}\sum_{x\in B_k} \mathbb{P}_x\left(S(H_{n/2+l})\in \tilde{\tau}(z(l))\right) \\
&\qquad \leq \sum_{k\geq n}\sum_{x\in B_k} \mathbb{P}_x\left(\text{the walk goes through at least } k-1 \text{ boxes}\right) \\
&\qquad \leq \sum_{k\geq n} (n+2l)^d(1-c)^{k-1} \\
&\qquad \leq Kn^d(1-c)^{n-1},
\end{aligned}$$

which tends to 0 when n tends to infinity, and is consequently uniformly bounded on n.

In this case, we have that

$$\mu_n \leq \frac{c\alpha_d l^d}{C} + C_0,$$

for some constant C_0 that does not depend on n. □

Optimization in (28) **:** We end by detailing the optimization in (28). This computation follows the same spirit as the optimization in the lower bound (see (19)). We will be using the previous bound on μ_n given by Lemma 2.16 to conclude.

In all that follows, recall that α_d is fixed. Let κ denote a generic positive constant. After applying Lemma 2.14, we get that for all $\lambda > 0$:

$$\begin{aligned}
&\mathbb{P}\left(M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \mathbb{B}(z(l),\tilde{\tau}(z(l)))\right) > \frac{\beta l^d\rho}{2},\ \delta_I(n)\leq \frac{\alpha_d l}{2C}\right) \\
&\quad \leq \exp\left(-\lambda\left(\frac{\beta l^d \rho}{2} - \mu_n\right) + \lambda^2\left(\mu_n + 4\sum_{y\in\mathcal{R}(l,n,\alpha_d,C)\cap\mathbb{Z}_{n^{\alpha'}}} \mathbb{P}_y\left(S(H_{n/2+l})\in\tilde{\tau}(z(l))\right)^2\right)\right).
\end{aligned}$$

Once again, minimizing in $\lambda > 0$ yields:

$$\mathbb{P}\left(M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \mathbb{B}(z(l), \tilde{\tau}(z(l)))\right) > \frac{\beta l^d \rho}{2},\ \delta_I(n) \leq \frac{\alpha_d l}{2C}\right)$$
$$\leq \exp\left(-\frac{\left(\mu_n - \frac{\beta l^d \rho}{2}\right)^2}{4\left(\mu_n + 4\sum_{y\in\mathcal{R}(l,n,\alpha_d,C)\cap\mathbb{Z}_{n^{\alpha'}}} \mathbb{P}_y\left(S(H_{n/2+l}) \in \tilde{\tau}(z(l))\right)^2\right)}\right).$$

Using the bound of Lemma 2.16, we can take C sufficiently large in order for $\left(\mu_n - \frac{\beta l^d \rho}{2}\right)^2$ to be of order l^{2d}. Therefore, we have for C large enough, $(\frac{\beta l^d \rho}{2} - \mu_n)^2 \geq \kappa l^{2d}$. Now, recall that

$$\sum_{y\in\mathcal{R}(l,n,\alpha_d,C)\cap\mathbb{Z}_{n^{\alpha'}}} \mathbb{P}_y\left(S(H_{n/2+h(n)}) \in \tilde{\tau}(z)\right)^2 \leq \kappa l^{2d-2}.$$

Using the bound of Lemma 2.16 and the fact that $d > 2$, we have:

$$4\left(\mu_n + 4\sum_{y\in\mathcal{R}(l,n,\alpha_d,C)\cap\mathbb{Z}_{n^{\alpha'}}} \mathbb{P}_y\left(S(H_{n/2+l}) \in \tilde{\tau}(z(l))\right)^2\right) \leq \kappa l^{2d-2}.$$

Combining both results gives:

$$\mathbb{P}\left(M^*_{n/2+l}\left(n\mathbb{1}_{\mathcal{H}}, \mathbb{B}(z(l), \tilde{\tau}(z(l)))\right) > \frac{\beta l^d \rho}{2},\ \delta_I(n) \leq \frac{\alpha_d l}{2C}\right) \leq \exp\left(-\frac{\kappa l^{2d}}{\kappa l^{2d-2}}\right)$$
$$= \exp\left(-\kappa l^2\right).$$

Chapter 3

Construction of ergodic IDLA forests in $\mathbb{Z}^d$

In this chapter, we prove the existence of ergodic (with respect to translations of $\mathcal{H}$) IDLA forests on $\mathbb{Z}^d$, with $d \geq 2$, based on the protocol used to build $A_n^\dagger[\infty]$ (see Section 2.1 for details about this aggregate). We extend the results of the previous chapter shown for $A_n[\infty]$ to $A_n^\dagger[\infty]$. Notably, we improve the aggregate stabilization result as well as the global upper bound.

Let $n,\ M \geq 0$. We can naturally build a family of *finite-volume* IDLA forests from $A_n^\dagger[M]$, denoted by $\mathcal{F}_n[M]$, simply by considering the edges of $\mathbb{Z}^d$ from which particles exit the current aggregate. We know from the coupling between $A_n^\dagger[M]$ and $A_n^\dagger[M+1]$ that the vertex set of $\mathcal{F}_n[M]$, namely $V(\mathcal{F}_n[M])$ is such that $V(\mathcal{F}_n[M]) \subset V(\mathcal{F}_n[M+1])$. However, this is *not* true for the set of edges. This is because vertices of $\mathcal{F}_n[M]$ and $\mathcal{F}_n[M+1]$ can be reached using different particles, hence through different edges. This can occur through a complex phenomenon called *chains of changes*. To define the infinite-volume forest, we are tempted to take the limit of $\mathcal{F}_n[M]$ as $M \to \infty$, but such a strategy can not work due to the fact that the finite-volume forests are not consistent.

To bypass this problem, we provide a stabilization result for the family of finite-volume forests. This result states that the finite-volume forests are consistent, when restricted to strips near the origin. This result then allows us to take a limit *in space*, i.e. when $M \to \infty$, allowing us to define the *infinite-volume* IDLA *forest up to time n*, which we denote by $\mathcal{F}_n$. Then, we define the infinite-volume directed IDLA forest by taking the limit *in time*, i.e. as $n \to \infty$:

$$\mathcal{F}_\infty := \bigcup \uparrow \mathcal{F}_n.$$

One of the main difficulties concerning the stabilization of the forests is the fact that the Abelian property can not be exploited. The reason behind this is that the *edges* of the forest crucially rely on the trajectory of each particle. When looking at IDLA aggregates, the Abelian property allows us to completely ignore the different trajectories of particles: this allows us to pause certain particles before launching them again, and so on. When looking at the edges of the forest, we can no longer use these types of arguments, making the study of IDLA forests a much more complex problem than the study of aggregates. We come up with an innovative tool to prove stabilization of the forests, relying on arguments of boolean percolation. We essentially transform the issue of chains of changes into a percolation problem, and show that our model does not percolate. This allows us to obtain our stabilization result for the forests.

The study of our percolation model uses many results on $A_n^\dagger[\infty]$, as both models are heavily linked. The adaptation (and improvement) of results from the previous chapter is crucial to show absence of percolation. In particular, we greatly improve the stabilization result for the aggregate and the global upper-bound, proved for $A_n[\infty]$ in the previous chapter, to the aggregate $A_n^\dagger[\infty]$. We get a much sharper bound, which allows us to sharpen our stabilization result for the aggregate. In the previous chapter, we saw that we could compare the aggregates $A_n[\infty]$ and $A_n[M^\alpha]$ inside the strip $\mathbb{Z}_M$, granted M is sufficiently large and $\alpha > 1$. Our improved result claims that we can now compare $A_n[\infty]$ and $A_n[2M]$, given that M is sufficiently large. This 'linear' stabilization makes it much easier to adapt a multi-scale argument from [22] in order to show absence of percolation.

The previous results are featured in a submitted article [14] in collaboration with Nicolas Chenavier, David Coupier and Arnaud Rousselle.

Outline of the current chapter

1 Introduction

The Internal Diffusion Limited Aggregation (IDLA) model gives a protocol to build random aggregates $(A_n)_{n\geq 0}$ recursively in $\mathbb{Z}^d$. Initially, we assume that $A_0 = \emptyset$. Then, at some step $n \geq 1$, given A_{n-1}, the first site visited outside of A_{n-1} by a random walk started at the origin is added to A_{n-1} in order to obtain A_n. In this context, such random walks are called *particles*. IDLA was initially introduced by Meakin and Deutch in [39] to model an industrial chemical technique known as electropolishing. The goal of such a process is to eliminate a small coat of material off of a metallic surface in order to make it smoother. It became pertinent to quantify how smooth the surface of a polished metal could get through such a process.

A first shape theorem was established by Lawler, Bramson and Griffeath in [34] to describe the asymptotic shape of A_n, as n tends to infinity, as a Euclidean ball. This result was later made sharper with the works of Jerison, Levine and Sheffield [31, 30, 28] and Asselah and Gaudillière [1, 4, 2]. In their works, the bounds for fluctuations around the limit shape are improved from linear to logarithmic in dimension $d = 2$ and sublogarithmic in dimensions $d \geq 3$. See also [23] for a continuous time version. From then on, variants of the IDLA model have been studied on many other graphs, such as on cylinder graphs in [29, 37, 48], on supercritical percolation clusters in [20, 47], on comb lattices in [5, 27] or on non-amenable graphs in [11, 26].

While the previously cited IDLA models are single-source, multi-source IDLA models have also been considered in [18, 17, 35] and by the authors in [15]. This question was originally investigated by Diaconis and Fulton [19] in the context of the *smash sum* of two domains in which they discover the famous *Abelian Property* of IDLA aggregates, meaning that modifying the order in which particles are launched does not change the distribution of the final aggregate. This beautiful property will be a powerful tool for the study of IDLA models.

In the present paper, we introduce new random graphs on $\mathbb{Z}^d$ with $d \geq 2$, called *IDLA forests*, whose construction is based on a multi-source IDLA protocol. We consider an infinite set of sources, namely the hyperplane $\mathcal{H} := \{0\} \times \mathbb{Z}^{d-1}$. Basically, in addition to the site at which the current particle exits the aggregate and stops, we also retain the edge by which the particle reaches that site. This procedure leads to a random forest on the lattice $\mathbb{Z}^d$, i.e. a collection of disjoint random trees rooted at sources of $\mathcal{H}$. Unlike IDLA aggregates, trajectories of particles really matter for the IDLA forests, meaning the Abelian property is no longer true for this new model, making it more difficult to show the existence of these forests.

However, we prove in Theorem 3.1 (our main result) the existence of the infinite-volume IDLA forests, generated by the infinite set of sources $\mathcal{H}$. See Figure 3.1 for a simulation in dimension $d = 2$. Moreover, our construction does not favor any source in the sense that (roughly speaking), at any time, the next source to emit a particle is selected 'uniformly' among all the sources of $\mathcal{H}$. This remarkable feature is stated in Theorem 3.2 in which we prove that the IDLA forests are ergodic w.r.t. the translations of $\mathcal{H}$.

Let us notice that the existence of bi-dimensional IDLA forests has been already explored by the authors in [16] but their proof strongly used the one-dimensional aspect of the set of sources–which is $\{0\} \times \mathbb{Z}$ when $d = 2$–and then completely collapses in higher dimensions. More than a generalization of [16], we think that the method developed here and based on percolation tools is original in the context of IDLA and is certainly promising to deal with graphs built from IDLA protocols with infinitely many sources.

Construction of the finite-volume IDLA forests

Let us start by describing our random inputs. Let us first consider a family of i.i.d. Poisson Point Processes (PPP) on $\mathbb{R}_+$, with intensity 1 and denoted by $\{\mathcal{N}_z\}_{z \in \mathcal{H}}$. Each PPP $\mathcal{N}_z$ provides a sequence $(\tau_{z,j})_{j \geq 1}$ of successive tops which act as random clocks for the emission of particles from the source z. Thus, to the sequence $(\tau_{z,j})_{j \geq 1}$, is associated a sequence of i.i.d. simple random walks $(S_{z,j})_{j \geq 1}$ on $\mathbb{Z}^d$ starting at z. Note that the $S_{z,j}$'s, for $z \in \mathcal{H}$ and $j \geq 1$, are independent from each other and also independent from the PPP $\mathcal{N}_z$'s. Hence, at time $\tau_{z,j}$, a particle is emitted from the source z (precisely, the j-th one coming from z), and follows the trajectory given by the random walk $S_{z,j}$ until exiting the current aggregate. To avoid having multiple particles alive at the same time, we assume that particles realize their trajectories instantaneously (w.r.t. the Poisson clocks).

Let $M, n \geq 0$ be integers. Set $\mathcal{H}_M := \{0\} \times [\![-M, M]\!]^{d-1}$. Let us consider the random set $A_n^\dagger[M] \subset \mathbb{Z}^d$ defined as the IDLA aggregate generated by particles emitted from the sources of

$\mathcal{H}_M$ and during the time interval $[0, n]$. Since the number of particles involved in $A_n^\dagger[M]$ is a.s. finite ($n(2M+1)^{d-1}$ in mean), this aggregate is a.s. well defined. See the left hand side of Figure 3.1 for a simulation of $A_{20}^\dagger[50]$ in dimension $d = 3$.

Now, we can build quite naturally from $A_n^\dagger[M]$ a finite-volume IDLA forest $\mathcal{F}_n[M]$, simply by considering the edges of $\mathbb{Z}^d$ from which the particles involved in $A_n^\dagger[M]$ exit the current aggregate. Let $\kappa := \sum_{z \in \mathcal{H}_M} \#\mathcal{N}_z([0, n])$ be the total number of such particles. Let us enumerate them according to their starting times, say $0 < \tau_1 < \tau_2 < ... < \tau_\kappa < n$ (they are a.s. different). For $j \in [\![1, \kappa]\!]$, we denote by $A[j]$ the aggregate obtained until time τ_j, including the site added by the particle sent at time τ_j. We then have $A[0] = \emptyset$ (with $\tau_0 = 0$) and $A[\kappa] = A_n^\dagger[M]$. We proceed by induction to build the associated forest $\mathcal{F}_n[M]$. Let us first set $\mathcal{F}_n[M, 0] = (\emptyset, \emptyset)$. Now, for $j \in [\![1, \kappa]\!]$, given the random graph $\mathcal{F}_n[M, j-1] = (V_{j-1}, E_{j-1})$, we define $\mathcal{F}_n[M, j] = (V_j,\ E_j)$ as follows. Let x be the new site added to $A[j-1]$ by the j-th particle.

- If x is the source from which the j-th particle is emitted, then this particle actually is the first one emitted from x, and x will be the root of a new tree in the graph. We set $V_j = V_{j-1} \cup \{x\}$ and $E_j = E_{j-1}$.
- Otherwise, let x' be the last site of $A[j-1]$ visited by the j-th particle before reaching x. Then we set $V_j = V_{j-1} \cup \{x\}$ and $E_j = E_{j-1} \cup \{(x', x)\}$.

Finally, we define $\mathcal{F}_n[M] := \mathcal{F}_n[M, \kappa]$. See the right hand side of Figure 3.1 for a simulation of $\mathcal{F}_{30}[20]$ in dimension $d = 2$. This construction ensures that $\mathcal{F}_n[M]$ is a finite union of trees with roots in $\mathcal{H}_M$ and whose vertex set is equal to $A_n^\dagger[M]$.

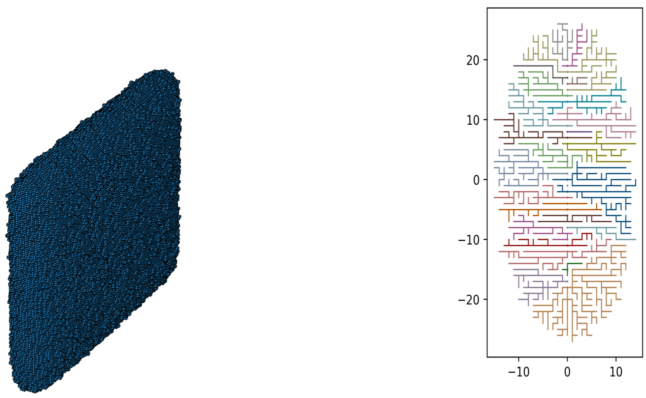


Figure 3.1: To the left: A realization in dimension $d = 3$ of the multi-source aggregate $A_{20}^\dagger[50]$ with particles emitted from $\mathcal{H}_{50}$ and on the time interval $[0, 20]$, looking like a 'bar of soap'. Each settled particle is represented by a blue cube. To the right: A realization in dimension $d = 2$ of the finite-volume IDLA forest $\mathcal{F}_{30}[20]$ associated to the aggregate $A_{30}^\dagger[20]$. Each tree is represented in a different color. Unfortunately, for visual reasons, we will only represent IDLA forests in dimension $d = 2$.

No monotonicity because of chains of changes

Thanks to the *natural coupling* defined in Section 2.2, one can construct on the same probability space the aggregates $A_n^\dagger[M]$, for all $M \geq 0$, in such a way that a.s. $A_n^\dagger[M] \subset A_n^\dagger[M+1]$ for any

M. This monotonicity property allows us to a.s. define a limiting aggregate as

$$A_n^\dagger[\infty] := \bigcup_{M\geq 0} \uparrow A_n^\dagger[M] \ . \tag{1}$$

However the same monotonicity property does not hold for the sequence $(\mathcal{F}_n[M])_{M\geq 0}$ of associated IDLA forests, as depicted in Figure 3.2. Consequently, we cannot define an infinite-volume forest as the increasing union of finite-volume forests, in the same spirit as (1). Let $M' \geq M \geq 0$. Although their vertex sets satisfy the inclusion $V(\mathcal{F}_n[M]) = A_n^\dagger[M] \subset V(\mathcal{F}_n[M']) = A_n^\dagger[M']$ (thanks to the natural coupling), this is no longer true for their edge sets. Indeed, some vertices present in both forests $\mathcal{F}_n[M]$ and $\mathcal{F}_n[M']$ are reached using different particles and possibly through different edges. This contributes to an edge in $\mathcal{F}_n[M]$ which is not present in $\mathcal{F}_n[M']$ (and conversely). These discrepancies between $\mathcal{F}_n[M]$ and $\mathcal{F}_n[M']$ can occur through a tricky phenomenon called *chains of changes* that we detail now. At first, the reader may skip this part and go directly to the result section below.

Let $M' \geq M \geq 0$ and $n \geq 0$. To explain what a chain of changes is, we need to slightly describe the natural coupling we use. The basic idea of that coupling consists in using the same random clocks $(\mathcal{N}_z)$ and the same random walks $(S_{z,j})$ for both aggregates $A_n^\dagger[M]$ and $A_n^\dagger[M']$. Hence a particle starting from a source in $\mathcal{H}_M$ will work for both aggregates, i.e. it will add a new site to both aggregates (but not necessarily the same site). However, a particle starting from a source in $\mathcal{H}_{M'} \setminus \mathcal{H}_M$ will only work for the larger aggregate $A_n^\dagger[M']$. Now, consider such a particle starting at time $t_1 \in (0, n)$ from a source in $\mathcal{H}_{M'} \setminus \mathcal{H}_M$, it adds a site x_1 to the larger aggregate. Precisely, if we write $A_{t_1^-}^\dagger[M']$ the current aggregate produced right before sending that particle, then we get $A_{t_1}^\dagger[M'] = A_{t_1^-}^\dagger[M'] \cup \{x_1\}$, while $A_{t_1}^\dagger[M] = A_{t_1^-}^\dagger[M]$ remains unchanged. At this time, x_1 belongs to the larger aggregate $A_{t_1}^\dagger[M']$ but not to the smaller one $A_{t_1}^\dagger[M]$. If no other future particles starting from $\mathcal{H}_M$ visit the site x_1, then x_1 will remain a discrepancy until time n between both aggregates. Conversely, assume x_1 is visited at time $t_2 \in (t_1, n)$ (and for the first time) by a particle coming from $\mathcal{H}_M$, then both aggregates are updated as follows:

- The site x_1 is added to the smaller aggregate: $A_{t_2}^\dagger[M] = A_{t_2^-}^\dagger[M] \cup \{x_1\}$.
- Since x_1 already belongs to $A_{t_2^-}^\dagger[M']$, the particle continues its trajectory until exiting the larger aggregate, say through a new site x_2, that it is then added: $A_{t_2}^\dagger[M'] = A_{t_2^-}^\dagger[M'] \cup \{x_2\}$.

At time t_2, the site x_1 is no longer a discrepancy between both aggregates but it has been reached by two different particles: x_1 actually is a blue point using the color code of Figure 3.2, i.e. both edges leading to x_1 in $\mathcal{F}_n[M]$ and $\mathcal{F}_n[M']$ could be different. Moreover, the site x_2 has become a discrepancy between both aggregates $A_{t_2}^\dagger[M]$ and $A_{t_2}^\dagger[M']$. In other words, the discrepancy has been *relayed* from x_1 to x_2 by the particle emitted at time t_2.

From then on, one can imagine a scenario where, at a random time $t_3 \in (t_2, n)$, the discrepancy at x_2 is relayed (by a third particle from $\mathcal{H}_M$) to a new site x_3 and possibly becomes another difference between both forests (i.e. a blue point), and so on. Such a phenomenon is referred to as a *chain of changes*. We point out that, even if it is initiated by a particle from $\mathcal{H}_{M'} \setminus \mathcal{H}_M$–i.e. quite far away from the e_1-axis when M is large–a chain of changes can spread up to the e_1-axis thanks to several relays. See the blue points in the top left picture of Figure 3.2.

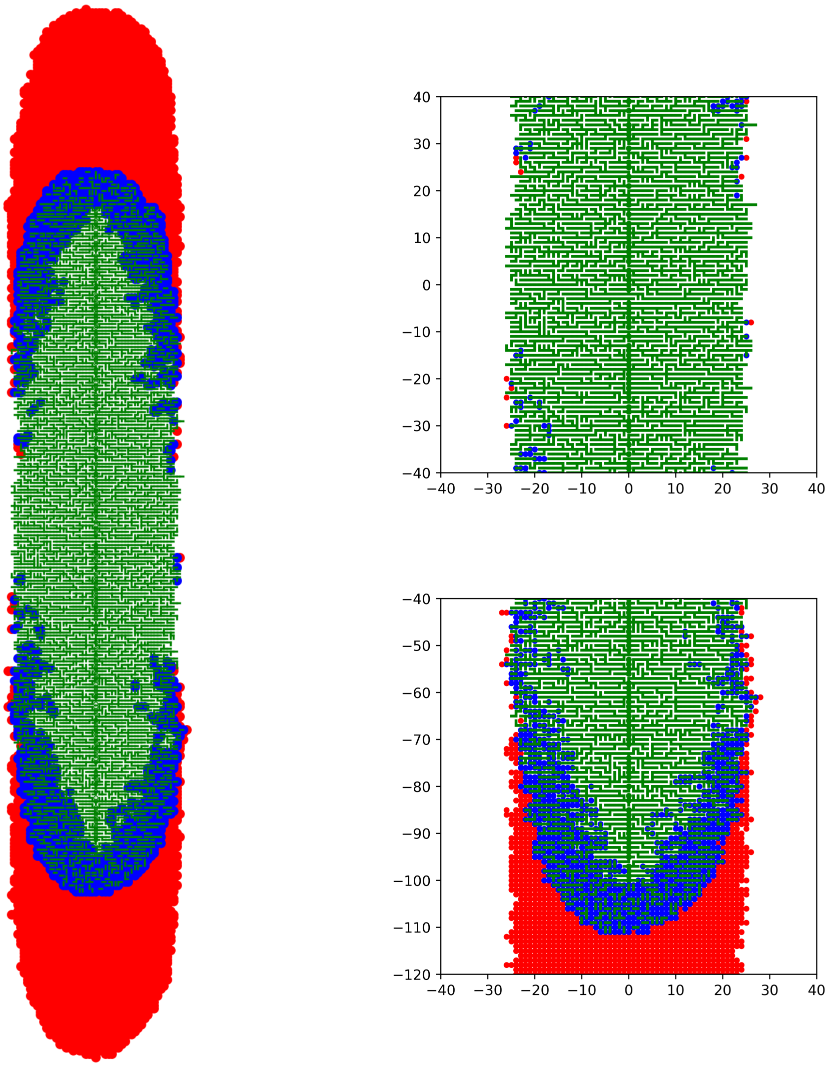


Figure 3.2: A realization of $\mathcal{F}_{50}[100]$ and $\mathcal{F}_{50}[150]$ using the natural coupling, i.e. their vertex sets are such that $A^{\dagger}_{50}[100] \subset A^{\dagger}_{50}[150]$. The green edges are common to both forests. The blue points are vertices common to both aggregates, but reached by different particles, and may lead to discrepancies between $\mathcal{F}_{50}[100]$ and $\mathcal{F}_{50}[150]$. The red points are vertices of $A^{\dagger}_{50}[150] \setminus A^{\dagger}_{50}[100]$. Both pictures on the right are zooms of the one on the left. One can see the presence of many blue points, especially on both extremities of $A^{\dagger}_{50}[100]$. Remark also that some of them appear in the vicinity of the e_1-axis (e_1 denoting the first vector of the canonical basis): see the top left picture. These possible discrepancies are produced by chains of changes.

Results

Our main task consists in controlling the chains of changes described in the previous section and proving that they cannot spread up too much inside the aggregate. This leads to the following stabilization result for the sequence of IDLA forests $(\mathcal{F}_n[M])_{M\geq 0}$. For that purpose, we need to define the strip $\mathbb{Z}_K := \mathbb{Z} \times [\![-K, K]\!]^{d-1}$ for any integer $K \geq 0$.

Theorem 3.1 (Forest stabilization result)**.** *Let $d \geq 2$. For all $n \geq 1$ and $K \geq 1$, the following holds with probability one:*

$$\exists N_0 = N_0(n, K) \geq 0, \forall N \geq N_0, \mathcal{F}_n[N] \cap \mathbb{Z}_K = \mathcal{F}_n[N_0] \cap \mathbb{Z}_K \ , \tag{2}$$

where the above identity means that all vertices and edges of $\mathcal{F}_n[N]$ and $\mathcal{F}_n[N_0]$ inside the strip $\mathbb{Z}_K$ coincide.

From then on, Theorem 3.1 allows us to take the limit $M \to \infty$ in the sequence $(\mathcal{F}_n[M])_{M\geq 0}$ in order to obtain an infinite-volume forest $\mathcal{F}_n$. First remark that the sequence $(N_0(n, K))_{n,K}$ in (3.1) can be chosen increasing in K, which implies for any $K' \geq K$,

$$\mathcal{F}_n[N_0(n, K)] \cap \mathbb{Z}_K = \mathcal{F}_n[N_0(n, K')] \cap \mathbb{Z}_K \subset \mathcal{F}_n[N_0(n, K')] \cap \mathbb{Z}_{K'} \ . \tag{3}$$

Inclusion (3) compensates for the lack of monotonicity of the sequence $(\mathcal{F}_n[M])_{M\geq 0}$ and allows us to define the *infinite-volume* IDLA *forest up to time n* denoted by $\mathcal{F}_n$ as

$$\mathcal{F}_n := \bigcup_{K\geq 1} \uparrow \mathcal{F}_n[N_0(n, K)] \cap \mathbb{Z}_K \ , \tag{4}$$

a.s. and for any $n \geq 1$. A realization of $\mathcal{F}_{100}[200]$ is given in Figure 3.3 seen through the strip $\mathbb{Z}_{40}$.

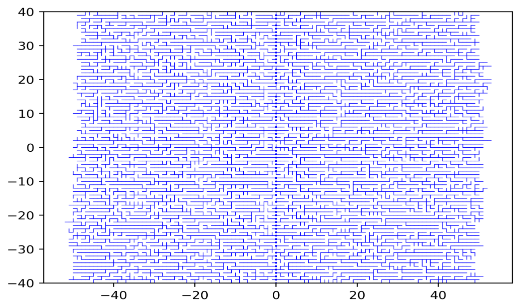


Figure 3.3: Here is a realization of $\mathcal{F}_{100}[200]$ intersected with the strip $\mathbb{Z}_{40}$. Taking $M = 200$ large enough, one can expect the infinite-volume forest $\mathcal{F}_{100}$ and the finite volume forest $\mathcal{F}_{100}[200]$ to coincide on $\mathbb{Z}_{40}$ thanks to Theorem 3.1.

After taking the limit $M \to \infty$ in space, let us take the limit $n \to \infty$ in time. The sequence $(N_0(n, K))_{n,K}$ can also be chosen increasing in n. So using once again the stabilization result of Theorem 3.1, we can write:

$$\begin{aligned} \mathcal{F}_n \cap \mathbb{Z}_K = \mathcal{F}_n[N_0(n, K)] \cap \mathbb{Z}_K &= \mathcal{F}_n[N_0(n+1, K)] \cap \mathbb{Z}_K \\ &\subset \mathcal{F}_{n+1}[N_0(n+1, K)] \cap \mathbb{Z}_K = \mathcal{F}_{n+1} \cap \mathbb{Z}_K, \end{aligned}$$

where the inclusion $\mathcal{F}_n[M] \subset \mathcal{F}_{n+1}[M]$ is merely due to extra particles emitted during the time interval $(n, n+1]$ (from $\mathcal{H}_M$). Hence, the sequence of random graphs $(\mathcal{F}_n)_{n\geq 1}$ is increasing in the sense that a.s. for any $n \geq 1$, $V(\mathcal{F}_n) \subset V(\mathcal{F}_{n+1})$ and $E(\mathcal{F}_n) \subset E(\mathcal{F}_{n+1})$. We then define the *infinite-volume* IDLA *forest* $\mathcal{F}_\infty$ by

$$\mathcal{F}_\infty := \bigcup_{n\geq 1} \uparrow \mathcal{F}_n \quad \text{a.s.} \tag{5}$$

The infinite-volume IDLA forests $\mathcal{F}_\infty$ and $(\mathcal{F}_n)_{n\geq 0}$ previously defined are built from an infinite set of sources, namely the hyperplane $\mathcal{H} = \{0\} \times \mathbb{Z}^{d-1}$. Theorem 3.2 below asserts that they are invariant in distribution w.r.t. translations of $\mathcal{H}$, meaning that all sources in $\mathcal{H}$ play the same role in their constructions. This gives a mathematical sense to what was announced at the beginning about $\mathcal{F}_\infty$ and $(\mathcal{F}_n)_{n\geq 0}$; at each time, the next source to emit a particle is chosen 'uniformly' among $\mathcal{H}$.

For $k \in \mathcal{H}$, let us denote by T_k the translation operator on $\mathbb{Z}^d$ defined by $\forall x \in \mathbb{Z}^d$, $T_k(x) = x + k$.

Theorem 3.2. *Let $d \geq 2$. The infinite-volume* IDLA *forests $\mathcal{F}_\infty$ and $(\mathcal{F}_n)_{n\geq 0}$ satisfy the following properties:*

1. *Almost surely, the set of vertices of $\mathcal{F}_\infty$ satisfies $V(\mathcal{F}_\infty) = \mathbb{Z}^d$.*
2. *The distributions of $(\mathcal{F}_n)_{n\geq 0}$ and $\mathcal{F}_\infty$ are invariant w.r.t. translations $T_k,\ k \in \mathcal{H}$.*
3. *The distributions of $(\mathcal{F}_n)_{n\geq 0}$ and $\mathcal{F}_\infty$ are α-mixing, and then ergodic, w.r.t. translations $T_k,\ k \in \mathcal{H}$.*

Strategy for proving Theorem 3.1

Let us call *level M* the set of sources in $\mathcal{H}$ at distance M from the origin (w.r.t. the infinite norm $\|(z_1, \ldots, z_d)\| := \max_i |z_i|$). Our main task consists in proving the stabilization result Theorem 3.1, i.e. given $n \geq 1$ and $K \geq 1$, that (2) recalled below occurs with probability 1:

$$\exists N_0 = N_0(n, K) \geq 0,\ \forall N \geq N_0,\ \mathcal{F}_n[N] \cap \mathbb{Z}_K = \mathcal{F}_n[N_0] \cap \mathbb{Z}_K\ .$$

Since the event $\{\mathcal{F}_n[N] \cap \mathbb{Z}_K \neq \mathcal{F}_n[N_0] \cap \mathbb{Z}_K\}$ implies the existence of a chain of changes between $A_n^\dagger[N]$ and $A_n^\dagger[N_0]$, a natural approach would be, using the Borel-Cantelli Lemma combined with a union bound, to show that

$$\sum_{N_0} \sum_{N \geq N_0} \mathbb{P}\big(\mathcal{F}_n[N] \cap \mathbb{Z}_K \neq \mathcal{F}_n[N_0] \cap \mathbb{Z}_K\big) < \infty\ .$$

However, it is difficult to obtain any upper bounds which decrease with respect to N *and* make the corresponding series summable. Indeed, this event provides no control on the level from which that chain of changes is initiated: it could be initiated anywhere from level $N_0 + 1$, independently of N. A different strategy is therefore required.

Our original approach is to interpret the chain of changes phenomenon with a percolation point of view. Consider $N \geq N_0 \geq K$ and a chain of changes between the forests $\mathcal{F}_n[N_0]$ and $\mathcal{F}_n[N]$ creating (at least) a discrepancy inside the strip $\mathbb{Z}_K$. Then, the sequence of successive relays will be interpreted as a sequence of successive overlapping balls, centered at the sources emitting the relay particles, whose cluster goes from level N_0 to the strip $\mathbb{Z}_K$. In particular, when $N_0 \gg K$, a very large cluster will correspond to this chain of changes.

Given $n, K \geq 1$, we proceed by contradiction and assume that, with positive probability,

$$\forall N_0,\ \exists N \geq N_0,\ \mathcal{F}_n[N] \cap \mathbb{Z}_K \neq \mathcal{F}_n[N_0] \cap \mathbb{Z}_K\ . \tag{6}$$

Throughout the whole paper, we will refer to (6) as the *Absurd hypothesis*. According to our percolation viewpoint, (6) leads to the existence of a Boolean model, say Σ, which percolates (i.e. admits an unbounded cluster) with positive probability. To get a contradiction, we will also state that Σ is actually subcritical with probability 1, concluding the proof of Theorem 3.1. To do it, we will act on three characteristics of the Boolean model Σ: its intensity (the density of its centers), its radii distribution and its long-range correlations.

First, the relay particles involved in a chain of changes are emitted by space-time points (z, t), i.e. from a source z and at time t. Denoting by (z_i, t_i)'s the sequence of emitting space-time points of a given chain of changes, the time sequence (t_i) is increasing by construction. Taking advantage of this monotonicity property, we prove that the intensity of the Boolean model Σ can be chosen as small as we want.

Thus, recall that the (infinite) aggregate $A_n^\dagger[\infty]$ defined in (1) is also the vertex set of the IDLA forest $\mathcal{F}_n$. So any information about it will help us to analyze the chain of changes phenomenon. In particular, in the Boolean model Σ, the radius of a ball associated to a relay particle is given by the maximal fluctuations performed by that particle from its source to exiting $A_n^\dagger[\infty]$. So stating a global upper bound (see Proposition 3.5) for this infinite aggregate–within a kind of cone–will allow us to control these fluctuations and prove that the radii distribution of Σ satisfies good integrability conditions.

Finally, in order to show that Σ is subcritical, we will apply a multiscale argument in the manner of [22]. A crucial ingredient making this strategy successful is to quantify how much Σ, when restricted to a finite window, depends on what happens far away. An important step towards such a local property satisfied by the Boolean model Σ lies in the stabilization result below. Theorem 3.3, interesting in itself, asserts that with high probability, the infinite aggregate $A_n^\dagger[\infty]$, when restricted to the strip $\mathbb{Z}_M$, does not depend on particles launched from levels larger than $2M$.

Theorem 3.3 (Aggregate stabilization result)**.** *Let $d \geq 2$ and $n \geq 1$. There exists a positive constant $C = C(n, d)$ such that for any $M, L \geq 1$,*

$$\mathbb{P}\big(A_n^\dagger[\infty] \cap \mathbb{Z}_M = A_n^\dagger[2M] \cap \mathbb{Z}_M\big) \geq 1 - \frac{C}{M^L} \ .$$

Proof of Theorem 3.3 is based on the global upper bound for $A_n^\dagger[\infty]$ mentioned above, on a donut argument already used in [16] and on a variant of the natural coupling between two aggregates, called the *special coupling*.

Why does the proof of [16] collapse in higher dimensions?

In [16], the authors proved the existence of IDLA forests $(\mathcal{F}_n)_{n\geq 1}$ and $\mathcal{F}_\infty$–see (4) and (5)–in the case of dimension $d = 2$, i.e. with the set of sources $\mathcal{H} = \{0\} \times \mathbb{Z}$. Their proof only works for the dimension $d = 2$ and cannot be generalized to higher dimensions: let us explain why.

For any integer n, let us define the *vacant set* $\mathcal{V}_n \subset \mathcal{H}$ by

$$\mathcal{V}_n := \{z \in \mathcal{H} : L(z) \cap A_n^\dagger[\infty] = \emptyset\}$$

where $L(z) := z + \{(k, 0) : k \in \mathbb{Z}\}$ is the horizontal line passing by z. It is proved in Corollary 5.2 of [16] that the random set $\mathcal{V}_n$ contains a.s. infinitely many sources. Due to the dimension $d = 2$, this implies that the aggregate $A_n^\dagger[\infty]$ is made up of (infinitely many) disjoint, finite connected components. The aggregate $A_n^\dagger[\infty]$ being the vertex set of the IDLA forest $\mathcal{F}_n$, it is then impossible for a chain of changes initiated from a very far away source to spread and create a discrepancy inside a given strip $\mathbb{Z}_K$ (the relay particles cannot jump from a connected component of $A_n^\dagger[\infty]$ to another one). Corollary 5.2 of [16] is certainly still true in dimension $d \geq 3$ (in the sense that $A_n^\dagger[\infty]$ contains an infinite number of empty lines $\mathbb{Z} \times \{y\}$, for $y \in \mathbb{Z}^{d-1}$) but its consequence about $A_n^\dagger[\infty]$, due to planarity, definitely collapses in higher dimensions.

However, one could ask whether the aggregate $A_n^\dagger[\infty]$ would still be an union of disjoint, finite connected components in dimension $d \geq 3$. Actually that is wrong whenever n is large enough. Indeed, let us call *rooted* a source z having emitted at least one particle during the time interval $[0, n]$. Hence, z is rooted if and only if the corresponding PPP $\mathcal{N}_z$ has generated at least one top in $[0, n]$, which occurs with probability $1 - e^{-n}$. The events $\{z \text{ is rooted}\}$, $z \in \mathcal{H}$, being independent from each other and their (common) probability tending to 1 as $n \to \infty$, we get that the set of rooted sources percolates in $\mathcal{H}$ for n large enough. Since a rooted source belongs to $A_n^\dagger[\infty]$, we conclude that this aggregate contains a (unique) infinite connected component.

Organization of the paper

Our paper is organized as follows. In Section 2, are gathered several properties about the aggregate $A_n^\dagger[\infty]$ that will be useful in the sequel. The natural and special couplings are detailed and the *Aggregate stabilization* result Theorem 3.3 is proved. In Section 3, we explain how chains of changes can be interpreted in terms of percolation. In particular, we define in Section 3.3 a discrete Boolean model $\hat{\Sigma}_\varepsilon$, with intensity $p_\varepsilon \to 0$ as $\varepsilon \to 0$, which is proved to be supercritical (i.e. percolating) for any $\varepsilon > 0$ under the *Absurd hypothesis* (6). Conversely, in Section 4, we adapt to our context a multiscale argument due to [22] which allows us to prove that $\hat{\Sigma}_\varepsilon$ does not percolate provided the intensity p_ε is small enough, leading to a contradiction with the conclusion of Section 3. Finally, Section 5 is devoted to the proof of Theorem 3.2. A detailed proof of Proposition 3.5 is given in the Appendix 6.

2 The infinite aggregate $A_n^\dagger[\infty]$

Let $n \geq 1$. Recall that the infinite aggregate $A_n^\dagger[\infty]$ has been defined in (1) as the limit of the increasing sequence of (finite) aggregates $(A_n^\dagger[M])_{M\geq 1}$. In this section, are gathered several results about $A_n^\dagger[\infty]$ which will be helpful for the proof of our main result Theorem 3.1. Indeed, the infinite aggregate $A_n^\dagger[\infty]$ is intended to be the vertex set of the infinite-volume IDLA forest $\mathcal{F}_n$. Results about $A_n^\dagger[\infty]$ are stated in Section 2.1. Both natural and special couplings are given in Section 2.2. Finally, we prove the stabilization result for $A_n^\dagger[\infty]$ (Theorem 3.3) in Section 2.3.

2.1 Results

Let us start with an invariance property in distribution for the infinite aggregate $A_n^\dagger[\infty]$. This result is based on the fact that the random ingredients generating $A_n^\dagger[\infty]$, namely the collection of PPP $\{\mathcal{N}_z : z \in \mathcal{H}\}$ and the random walks $\{S_{z,j} : z \in \mathcal{H}, j \geq 1\}$, are iid. See Proposition 2.2 of [16] for the same result but in dimension $d = 2$. The same proof actually works for any $d \geq 2$.

Proposition 3.4. *The distribution of $A_n^\dagger[\infty]$ is invariant w.r.t. translations of the source set $\mathcal{H}$:*

$$T_k A_n^\dagger[\infty] \overset{\text{law}}{=} A_n^\dagger[\infty]$$

where T_k, for $k \in \mathcal{H}$, is defined on $\mathbb{Z}^d$ by $T_k(x) = x + k$.

The following result, given by Proposition 3.5, provides a global control of the shape of $A_n^\dagger[\infty]$. It is referred to as a *Global Upper Bound* and is necessary in the proof of Theorem 3.3 in Section

2.3. Let us begin by introducing some notations. For $0 < \alpha < 1$ and $\varepsilon > 0$, let us consider the cone $\mathscr{C}_\varepsilon^\alpha$ as

$$\mathscr{C}_\varepsilon^\alpha = \bigcup_{\ell \geq 0} \left\{ z \in \mathbb{Z}^d, \ \|p_{\mathcal{H}}(z)\| = \ell, \ |z_1| \leq \varepsilon \ell^\alpha \right\},$$

where $p_{\mathcal{H}}$ denotes the orthogonal projection onto $\mathcal{H}$. Then, for any integer $M \geq 0$, we define the event

$$\mathbf{Over}_\alpha^\dagger(M, n, \varepsilon) := \{ A_n^\dagger[\infty] \cap \mathbb{Z}_M^c \not\subset \mathscr{C}_\varepsilon^\alpha \} , \tag{7}$$

meaning that the aggregate $A_n^\dagger[\infty]$ exceeds the cone $\mathscr{C}_\varepsilon^\alpha$ outside the strip $\mathbb{Z}_M$. Proposition 3.5 states an upper bound for the probability of $\mathbf{Over}_\alpha^\dagger(M, n, \varepsilon)$ which implies that above a certain level, the aggregate is almost surely contained inside the cone $\mathscr{C}_\varepsilon^\alpha$.

Proposition 3.5. *(Global upper bound) For any $\varepsilon > 0$ and $\alpha \in (1 - 1/d, 1)$, there exists a positive constant $C = C(d, n, \varepsilon, \alpha)$ such that for all integers $L, M > 1$,*

$$\mathbb{P}\left(\mathbf{Over}_\alpha^\dagger(M, n, \varepsilon) \right) \leq \frac{C}{M^L} .$$

In particular, with probability 1, there exists a (random) level from which $A_n^\dagger[\infty] \cap \mathbb{Z}_M^c$ is included in the cone $\mathscr{C}_\varepsilon^\alpha$.

Proposition 3.5 actually is a refined version of Theorem 4.1 of [15] which states the same result but for $\alpha = 1$ (i.e. for a wider cone). So only a short proof of Proposition 3.5 is given in the Appendix, focusing on the differences due to the use of the thinner cone $\mathscr{C}_\varepsilon^\alpha$, with $\alpha \in (1-1/d, 1)$. However this refinement (using $\mathscr{C}_\varepsilon^\alpha$ instead of $\mathscr{C}_\varepsilon^1$) is required here to get Theorem 3.3–and then our main result Theorem 3.1–as explained at the end of Section 2.3.

2.2 Two couplings

Let $n \geq 1$ and $M' > M$. In this section, we detail two different couplings allowing to construct both aggregates $A_n^\dagger[M]$ and $A_n^\dagger[M']$ on the same probability space in such a way that

$$A_n^\dagger[M] \subset A_n^\dagger[M'] \quad \text{a.s.} \tag{8}$$

The first one, called the *natural coupling*, will be used intensively in Section 3 to describe the chains of changes. It has been introduced in [16]. In this paper, we will require a variant of the natural coupling, called the *special coupling*, ensuring a special property ($\star$) (see below) in addition to 8. The special coupling will be used in Section 2.3 to get Theorem 3.3. Hence, we first recall the natural coupling in details and thus its variant.

Let us begin by describing the natural coupling. Let $\kappa := \sum_{z \in \mathcal{H}_{M'}} \#\mathcal{N}_z([0, n])$ be the total number of particles sent from $\mathcal{H}_{M'}$ during the time interval $[0, n]$. Let us build two sequences of aggregates $(A_i)_{0 \leq i \leq \kappa}$ and $(B_i)_{0 \leq i \leq \kappa}$ such that

$$\text{for all } 0 \leq i \leq \kappa, \ A_i \subset B_i \ \text{ and } \ A_\kappa = A_n^\dagger[M], \ B_\kappa = A_n^\dagger[M'] .$$

We proceed by induction on $i \in [\![0, \kappa]\!]$ by sorting the κ particles according to their starting times (from time 0 to n). When $i = 0$ (no particles have been emitted), we have that $A_0 = B_0 = \emptyset$. Now, suppose $i \geq 0$ and $A_i \subset B_i$ and let us say that the $(i+1)$-th particle is sent from a source $z \in \mathcal{H}_{M'}$.

- If $z \in \mathcal{H}_{M'} \backslash \mathcal{H}_M$ then the $(i+1)$-th particle only contributes to B_i. It adds a random site

x to B_i while A_i remains unchanged:

$$A_{i+1} := A_i \subset B_i \subset B_i \cup \{x\} =: B_{i+1}.$$

- If $z \in \mathcal{H}_M$, the $(i+1)$-th particle contributes to both aggregates. Since $A_i \subset B_i$, it exits A_i before B_i, and adds a random site x to A_i. Now, we must consider two cases.
  - If $x \notin B_i$, then x is added to B_i. Hence,

$$A_{i+1} := A_i \cup \{x\} \subset B_i \cup \{x\} =: B_{i+1} \ .$$

  - If $x \in B_i$, then the $(i+1)$-th particle does not exit B_i in x, and continues its trajectory until exiting B_i on some site $x' \neq x$. In this case,

$$A_{i+1} := A_i \cup \{x\} \subset B_i \subset B_i \cup \{x'\} =: B_{i+1} \ .$$

Let us now detail the special coupling. The total number of particles sent from $\mathcal{H}_{M'}$ during $[0, n]$ is still denoted by κ and, as before, we build two sequences of aggregates $(\tilde{A}_i)_{0 \le i \le \kappa}$ and $(\tilde{B}_i)_{0 \le i \le \kappa}$ by induction on $i \in [\![0, \kappa]\!]$. Our construction will ensure that

$$\text{for all } 0 \le i \le \kappa,\ \ \tilde{A}_i \subset \tilde{B}_i \ \text{ and } \ \tilde{A}_\kappa = A_n^\dagger[M],\ \tilde{B}_\kappa \overset{\text{law}}{=} A_n^\dagger[M'] \ ,$$

while also ensuring that the following condition holds, for any $0 \le i \le \kappa$:

($\star$) Any element $x \in \tilde{B}_i \backslash \tilde{A}_i$ is produced by a particle emitted from a source in $\mathcal{H}_{M'} \backslash \mathcal{H}_M$.

Let us build the $\tilde{A}_i$'s and $\tilde{B}_i$'s. When $i = 0$, we have that $\tilde{A}_0 = \tilde{B}_0 = \emptyset$. Let us assume for some $i \ge 0$ that $\tilde{A}_i \subset \tilde{B}_i$, and that they both satisfy condition ($\star$). Let us say that the $(i+1)$-th particle is sent from a source $z \in \mathcal{H}_{M'}$ (and moves according to a random walk S_z).

- If $z \in \mathcal{H}_{M'} \backslash \mathcal{H}_M$ then the $(i+1)$-th particle only contributes to B_i. As for the natural coupling, it adds a random site x to B_i while A_i remains unchanged:

$$\tilde{A}_{i+1} := \tilde{A}_i \subset \tilde{B}_i \subset \tilde{B}_i \cup \{x\} =: \tilde{B}_{i+1}$$

  and the couple $(\tilde{A}_{i+1}, \tilde{B}_{i+1})$ still satisfies ($\star$).

- If $z \in \mathcal{H}_M$, the $(i+1)$-th particle contributes to both aggregates. Since $\tilde{A}_i \subset \tilde{B}_i$, it exits $\tilde{A}_i$ before $\tilde{B}_i$, and adds a random site x to $\tilde{A}_i$. Once again, we consider two cases.
  - If $x \notin \tilde{B}_i$ then we proceed as for the natural coupling: x is also added to $\tilde{B}_i$ which again implies

$$\tilde{A}_{i+1} := \tilde{A}_i \cup \{x\} \subset \tilde{B}_i \cup \{x\} =: \tilde{B}_{i+1} \ .$$

    The couple $(\tilde{A}_{i+1}, \tilde{B}_{i+1})$ still satisfies ($\star$) since the just added site x belongs to $\tilde{A}_{i+1}$ and $\tilde{B}_{i+1}$.
  - If $x \in \tilde{B}_i \backslash \tilde{A}_i$, then thanks to condition ($\star$), the site x was reached by a particle (say with index $i' < i$) originating from some source $z' \in \mathcal{H}_{M'} \backslash \mathcal{H}_M$ (and moving according to a random walk $S_{z'}$). Then, the $(i+1)$-th particle settles at x, wakes up the i'-th particle which continues its trajectory (according to $S_{z'}$) until exiting $\tilde{B}_i$ on some site y. In this case,

$$\tilde{A}_{i+1} := \tilde{A}_i \cup \{x\} \subset \tilde{B}_i \subset \tilde{B}_i \cup \{y\} =: \tilde{B}_{i+1} \ .$$

Note that the couple $(\tilde{A}_{i+1}, \tilde{B}_{i+1})$ satisfies $(\star)$ since y, which is a discrepancy between $\tilde{A}_{i+1}$ and $\tilde{B}_{i+1}$, has been produced by a particle sent from $\mathcal{H}_{M'} \setminus \mathcal{H}_M$.

To conclude, let us remark that in both couplings, the aggregates A_i and $\tilde{A}_i$ are built in the same way. They are a.s. equal and then $\tilde{A}_\kappa = A_\kappa = A_n^\dagger[M]$. This is not the case for the $\tilde{B}_i$'s: even if $\tilde{B}_i = B_i$, the site y added following the random walk $S_{z'}$ could be different from the site x' added following S_z. However, in distribution, they are equal:

$$\tilde{B}_\kappa \stackrel{\text{law}}{=} B_\kappa = A_n^\dagger[M']$$

since the trajectory used to add the site y to $\tilde{B}_i$ is the concatenation of S_z from the source z to x, thus $S_{z'}$ after x. This trajectory is a random walk.

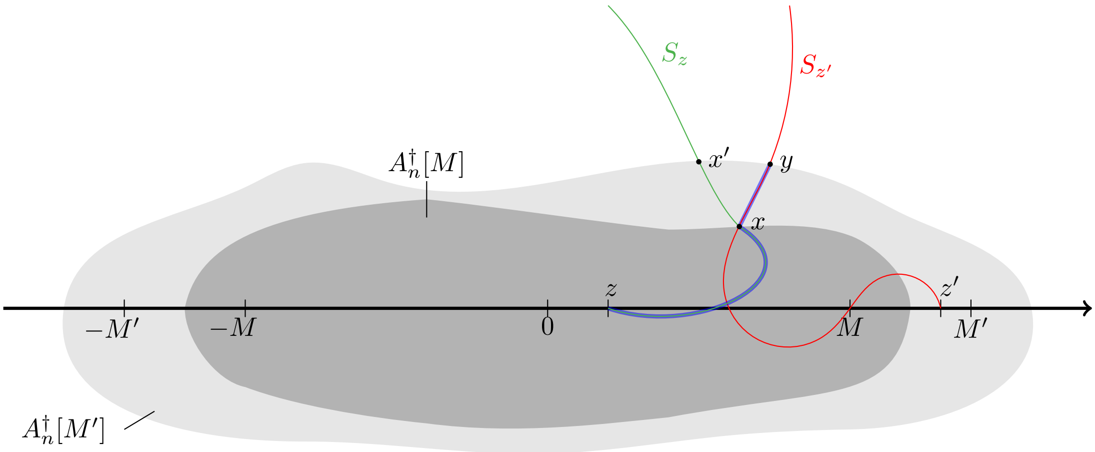


Figure 3.4: The aggregates $A_n^\dagger[M]$ and $A_n^\dagger[M']$ are represented in dark and light gray. The trajectories of random walks S_z and $S_{z'}$ are respectively depicted in green and red. In the natural coupling, the site x' is added to $\tilde{B}_i$ using solely S_z whereas in the special coupling, the site y is added to $\tilde{B}_i$ using first S_z until exiting $\tilde{A}_i = A_i$ and then $S_{z'}$. We highlight in blue the actual path that is realized when doing so.

2.3 Proof of Theorem 3.3

Let $n \geq 1$. Our goal is to prove that there exists a positive constant $C = C(n, d)$ such that for any $M, L \geq 1$,

$$\mathbb{P}\big(A_n^\dagger[\infty] \cap \mathbb{Z}_M = A_n^\dagger[2M] \cap \mathbb{Z}_M\big) \geq 1 - \frac{C}{M^L} .$$

To do it, let us introduce the event

$$D_M := \left\{ \begin{array}{c} \text{The trajectory of any random walk associated with a particle of } A_n^\dagger[\infty] \text{ and starting} \\ \text{from a level greater than } 2M \text{ does not visit the strip } \mathbb{Z}_M \text{ before exiting } \mathscr{C}_\varepsilon^\alpha \end{array} \right\} .$$

The event D_M has a probability tending to 1 as $M \to \infty$:

Lemma 3.6. *For $\varepsilon > 0$ and $\alpha \in (0,1)$, there exists a positive constant $C = C(d, n, \varepsilon, \alpha)$ such that for all $M, L \geq 1$,*

$$\mathbb{P}\left(D_M\right) \geq 1 - \frac{C}{M^L}.$$

The proof of Theorem 3.3 works in two steps. We first explain how to conclude from Lemma 3.6 and then we prove this auxiliary result.

Let us pick $\varepsilon > 0$ and $\alpha \in (1-1/d, 1)$. Let $M, L \geq 1$. Let us consider the event $\mathcal{G}_M$ defined by

$$\mathcal{G}_M := \left\{ A_n^\dagger[\infty] \cap \mathbb{Z}_M^c \subset \mathscr{C}_\varepsilon^\alpha \right\} \cap D_M \ .$$

Thanks to Proposition 3.5 and Lemma 3.6, there exists a positive constant $C = C(d, n, \varepsilon, \alpha)$ such that $\mathbb{P}(\mathcal{G}_M) \geq 1 - CM^{-L}$. Given $M' > 2M$, we consider the aggregates $A_n^\dagger[2M]$ and $A_n^\dagger[M']$ under the special coupling. Hence, a.s. $A_n^\dagger[2M]$ is included in $A_n^\dagger[M']$ and any element x in $A_n^\dagger[M'] \setminus A_n^\dagger[2M]$ is produced by a particle emitted from a source in $\mathcal{H}_{M'} \setminus \mathcal{H}_{2M}$ thanks to the condition $(\star)$. On the event $\mathcal{G}_M$, the random walk associated to this particle necessarily exited $\mathscr{C}_\varepsilon^\alpha$ before reaching $\mathbb{Z}_M$, and hence necessarily exited $A_n^\dagger[M']$ before reaching $\mathbb{Z}_M$. This means that, on the event $\mathcal{G}_M$, both aggregates $A_n^\dagger[2M]$ and $A_n^\dagger[M']$ coincide on $\mathbb{Z}_M$ and leads to

$$\mathbb{P}\left(A_n^\dagger[M'] \cap \mathbb{Z}_M = A_n^\dagger[2M] \cap \mathbb{Z}_M\right) \geq 1 - \frac{C}{M^L}, \ \forall M' > 2M \ . \tag{9}$$

The key argument here is the special coupling which ensures that any discrepancy between both aggregates is due to a single particle and not to a chain of changes.

We can now conclude. On the one hand, thanks to (9),

$$\liminf_{M' \to \infty} \mathbb{P}\left(A_n^\dagger[M'] \cap \mathbb{Z}_M \subset A_n^\dagger[2M] \cap \mathbb{Z}_M\right) \geq 1 - \frac{C}{M^L} \ .$$

On the other hand, the infinite aggregate $A_n^\dagger[\infty]$ being the limit of the increasing sequence $(A_n^\dagger[M'])_{M'}$ (thanks to the natural coupling), we have

$$\lim_{M' \to \infty} \mathbb{P}\left(A_n^\dagger[M'] \cap \mathbb{Z}_M \subset A_n^\dagger[2M] \cap \mathbb{Z}_M\right) = \mathbb{P}\left(A_n^\dagger[\infty] \cap \mathbb{Z}_M \subset A_n^\dagger[2M] \cap \mathbb{Z}_M\right) \ .$$

As a consequence, we can write

$$\mathbb{P}\left(A_n^\dagger[\infty] \cap \mathbb{Z}_M \subset A_n^\dagger[2M] \cap \mathbb{Z}_M\right) \geq 1 - \frac{C}{M^L}$$

and the same holds for $\mathbb{P}(A_n^\dagger[\infty] \cap \mathbb{Z}_M = A_n^\dagger[2M] \cap \mathbb{Z}_M)$ since $A_n^\dagger[2M] \subset A_n^\dagger[\infty]$ with probability 1. This concludes the proof of Theorem 3.3.

Proof of Lemma 3.6. Let $\varepsilon > 0$, $\alpha \in (0,1)$ and $M, j \geq 1$. Let us set

$$E_{M,j} := \left\{ \begin{array}{c} \text{At least one random walk starting from } \mathrm{Ann}(M,j) \\ \text{visits the strip } \mathbb{Z}_M \text{ before exiting } \mathscr{C}_\varepsilon^\alpha \end{array} \right\}$$

where $\mathrm{Ann}(M,j) := \mathcal{H}_{(j+2)M} \setminus \mathcal{H}_{(j+1)M}$. Hence, the event D_M^c is equal to $\cup_{j \geq 1} E_{M,j}$ and we focus on bounding each term $\mathbb{P}(E_{M,j})$.

As in the proof of Theorem 1.2 of [15], our strategy consists in building donuts from level $(j+1)M$ down to level M, symmetric w.r.t. the hyperplane $\mathcal{H}$ and containing the cone $\mathscr{C}_\varepsilon^\alpha$ (restricted to $\mathbb{Z}_{(j+1)M} \setminus \mathbb{Z}_M$). The largest donut is the one built at level $(j+1)M$. Its width is equal to $2\varepsilon((j+1)M)^\alpha$ and all the following donuts have smaller sizes. Therefore, the number of donuts $k^o = k^o(j, M, \varepsilon, \alpha)$ we can build from level $(j+1)M$ to level M verifies

$$k^o \geq \frac{(j+1)M - M}{2\varepsilon(j+1)^\alpha M^\alpha} \geq \frac{(jM)^{1-\alpha}}{4\varepsilon}, \tag{10}$$

where the last inequality is due to $(j+1)^\alpha \leq 2j$.

We know that if a random walk sent from a level greater than $(j+1)M$ reaches the strip $\mathbb{Z}_M$ while staying inside the cone $\mathscr{C}_\varepsilon^\alpha$, it necessarily crossed over the k^o donuts previously built. The probabilistic cost to cross any given donut while staying inside the cone is at most $1-c$ with $c := (2d)^{-2}$. So the probability for such random walk to cross the k^o donuts before exiting $\mathscr{C}_\varepsilon^\alpha$ is at most $(1-c)^{k^o}$. See Proposition 3.1 of [15] for details.

Besides, let us denote by $N_{tot}^{(j)} = N_{tot}^{(j)}(d,n,M,j)$ the total number of particles sent from $\mathrm{Ann}(M,j)$ during the time interval $[0,n]$:

$$N_{tot}^{(j)} := \sum_{z\in\mathrm{Ann}(M,j)} \mathcal{N}_z([0,n]) \ .$$

The next result allows us to bound from above $N_{tot}^{(j)}$ with high probability.

Lemma 3.7. *Let $M, j \geq 1$ and set $C_{M,j} := \#\mathrm{Ann}(M,j)$. Then,*

$$\mathbb{P}\big(N_{tot}^{(j)} > 2nC_{M,j}\big) \leq \exp\left(-c_0 j^{d-2} M^{d-1}\right),$$

where $c_0 = c_0(d,n)$ denotes a positive constant.

Proof of Lemma 3.7. The searched inequality is a direct consequence of the concentration inequality for Poisson variables (11) stated below, applied to $N_{tot}^{(j)}$ which is distributed as a Poisson random variable with parameter $nC_{M,j}$ and to the fact that $C_{M,j}$ is of order $j^{d-2}M^{d-1}$.

If X is a Poisson random variable of parameter $\lambda > 0$, then for any $t \geq 0$, the following holds:

$$\mathbb{P}\left(X - \mathbb{E}[X] \geq t\right) \leq \exp\left(-\frac{t^2}{2\left(\lambda + \frac{t}{3}\right)}\right). \tag{11}$$

□

All the ingredients are gathered, we can compute:

$$\begin{aligned}
\mathbb{P}(E_{M,j}) &\leq \mathbb{P}\big(E_{M,j} \cap \{N_{tot}^{(j)} \leq 2nC_{M,j}\}\big) + \mathbb{P}\big(N_{tot}^{(j)} > 2nC_{M,j}\big) \\
&\leq \sum_{i=1}^{2nC_{M,j}} \mathbb{P}\big(\{\text{walk } i \text{ visits } \mathbb{Z}_M \text{ before exiting } \mathscr{C}_\varepsilon^\alpha\}\big) + \exp\big(-c_0 j^{d-2} M^{d-1}\big) \\
&\leq \sum_{i=1}^{2nC_{M,j}} \mathbb{P}\big(\{\text{walk } i \text{ crosses } k^o \text{ donuts before exiting } \mathscr{C}_\varepsilon^\alpha\}\big) + \exp\big(-c_0 j^{d-2} M^{d-1}\big) \\
&\leq 2nC_{M,j}(1-c)^{k^o} + \exp\big(-c_0 j^{d-2} M^{d-1}\big) \ .
\end{aligned}$$

Thus, (10) leads to

$$2nC_{M,j}(1-c)^{k^o} \leq C_1 M^{d-1}(j+1)^{d-2} \exp\big(-C_2(jM)^{1-\alpha}\big)$$

where C_1, C_2 are positive constants depending only on d, n, ε. Summing over $j \geq 1$, we get

$$\mathbb{P}\Big(\bigcup_{j\geq 1} E_{M,j}\Big) \leq \sum_{j\geq 1} \mathbb{P}(E_{M,j}) \leq \sum_{j\geq 1} C_1 M^{d-1}(j+1)^{d-2} \exp\big(-C_2(jM)^{1-\alpha}\big) + \sum_{j\geq 1} \exp\big(-c_0 j^{d-2} M^{d-1}\big) \; .$$

Since $1-\alpha > 0$, both terms of the upper bound above are summable and decrease faster than any power of M^{-1}, which concludes the proof.

Remark that the previous conclusion holds only if $d \geq 3$. When $d = 2$, we have to proceed slightly differently. Each $\mathrm{Ann}(M, j)$ has the same cardinality, equal to M. So, in the previous computation, it suffices to intersect the event $E_{M,j}$ with $\{N_{tot}^{(j)} \leq 2nMj^\beta\}$ (for some $\beta > 0$) since (11) allows to bound the probability of $\{N_{tot}^{(j)} > 2nMj^\beta\}$ by $\exp(-c'Mj^\beta)$–which is summable w.r.t. j and M. □

Theorem 4.1 of [15] is a weaker version of Proposition 3.5: it gives the same result but in the particular case of a linear cone $\mathscr{C}_\varepsilon^1$, i.e. with $\alpha = 1$. Taking $\alpha = 1$ in the computation (10) leads to a constant number of donuts k^o (no longer depending on j, M) which prevents us to conclude as above. Hence, in [15], in order to get sufficiently many donuts to make unlikely the crossing to $\mathbb{Z}_M$ for particles coming far away, we required more space. This argument led to a stabilization result (Theorem 1.2 of [15]) weaker than our Theorem 3.3 since it claimed that we need to go above levels M^γ, $\gamma > 1$, to stabilize $A_n^\dagger[\infty] \cap \mathbb{Z}_M$.

However, to apply with success the multiscale argument of Section 4, we need a stabilization result for $A_n^\dagger[\infty] \cap \mathbb{Z}_M$ requiring a linear number of levels of M (rather than M^γ). This is why we had to improve the stabilization result Theorem 1.2 of [15] into Theorem 3.3. This justifies the use of the thinner cone $\mathscr{C}_\varepsilon^\alpha$ leading to the refined global upper bound Proposition 3.5.

3 From chains of changes to percolation models

To get the stabilization result Theorem 3.1, we proceed by contradiction by assuming the *Absurd hypothesis* (6) that we recall now: there exist positive integers n_0, K_0 (fixed for the whole section) such that

$$\forall N_0,\, \exists N \geq N_0,\, \mathcal{F}_{n_0}[N] \cap \mathbb{Z}_{K_0} \neq \mathcal{F}_{n_0}[N_0] \cap \mathbb{Z}_{K_0}$$

occurs with positive probability.

In section 3.1, we use a space-time representation of a chain of changes between the forests $\mathcal{F}_{n_0}[N_0]$ and $\mathcal{F}_{n_0}[N]$ to describe the propagation of discrepancies between the corresponding aggregates $A_{n_0}^\dagger[N_0]$ and $A_{n_0}^\dagger[N]$. In Section 3.2, a percolation model Σ with good properties (Lemmas 3.8 and 3.9) is introduced. Under the *Absurd hypothesis*, we prove that Σ percolates in the sense that it contains an *infinite descending chain* with positive probability. Finally, in Section 3.3, we take advantage of the monotonicity in time of such infinite descending chain in order to state that it actually appears instantaneously. The final result of Section 3 is summarized in Proposition 3.11 saying that a discrete Boolean model $\hat{\Sigma}_\varepsilon$ percolates even if its intensity tends to 0 with $\varepsilon \to 0$.

For $z \in \mathcal{H}$ and $r \in \mathbb{N}$, let us set $\mathbb{B}(z, r) := z + \mathcal{H}_r$ the $(d-1)$-dimensional ball, included in $\mathcal{H}$, centered at z and with radius r (for the infinity norm). We also denote by $p_\mathcal{H} : \mathbb{Z}^d \to \mathcal{H}$ the orthogonal projection onto the hyperplane $\mathcal{H}$.

3.1 A space-time representation of chains of changes

Let (N_0, N), with $N \geq N_0$, a couple of integers such that the forests $\mathcal{F}_{n_0}[N]$ and $\mathcal{F}_{n_0}[N_0]$ do not coincide on the strip $\mathbb{Z}_{K_0}$. From now on, we consider their vertex sets, namely the aggregates $A^{\dagger}_{n_0}[N]$ and $A^{\dagger}_{n_0}[N_0]$, under the natural coupling defined in Section 2.2, i.e. satisfying a.s. the inclusion $A^{\dagger}_{n_0}[N_0] \subset A^{\dagger}_{n_0}[N]$. A (random) space-time couple (z, t) where z is a source with $\|z\| \leq N$ and $t \in [0, n_0]$ is a 'top' of the PPP $\mathcal{N}_z$, is called a *starting point*. It means that, at time t, a particle using the random walk $S_{z,t} = (S_{z,t}(k))_{k\geq 0}$ and launched from the source z, contributes to the construction of $A^{\dagger}_{n_0}[N]$–and also of $A^{\dagger}_{n_0}[N_0]$ if $\|z\| \leq N_0$. As explained in the introduction, the difference between both forests $\mathcal{F}_{n_0}[N]$ and $\mathcal{F}_{n_0}[N_0]$ inside $\mathbb{Z}_{K_0}$ is due to a chain of changes, i.e. a sequence of $\kappa \geq 1$ particles coming from starting points $(z_i, t_i)_{1\leq i\leq \kappa}$ satisfying the following conditions:

$$\begin{cases} N_0 < \|z_1\| \leq N \text{ and } \|z_i\| \leq N_0, \text{ for } 2 \leq i \leq \kappa \\ 0 < t_1 < t_2 < \cdots < t_\kappa < n_0 \\ \text{for } 1 \leq i \leq \kappa - 1, \text{ the } i\text{-th particle is relayed by the } (i+1)\text{-th one} \\ \text{the } \kappa\text{-th particle exists } A^{\dagger}_{t_\kappa^-}[N] \text{ through } \mathbb{Z}_{K_0} \end{cases} \tag{12}$$

Recall that 'the i-th particle is relayed by the $(i + 1)$-th one' means that the discrepancy $x_i \in A^{\dagger}_{t_i}[N] \setminus A^{\dagger}_{t_i}[N_0]$ created by the i-th particle at time t_i, is visited by the $(i + 1)$-th particle at time t_{i+1}, which contributes to both aggregates. So, at time t_{i+1}, x_i is now longer a discrepancy and is replaced with a new one $x_{i+1} \in A^{\dagger}_{t_{i+1}}[N] \setminus A^{\dagger}_{t_{i+1}}[N_0]$ which actually is the site at which the $(i + 1)$-th particle settles when it exits the current aggregate $A^{\dagger}_{t_{i+1}^-}[N]$.

Associated with a given starting point (z, t) and with the corresponding particle, we define the radius $R_N(z, t)$ as follows:

$$R_N(z,t) := \min \left\{ r \in \mathbb{N} : \mathbb{B}(z,r) \text{ contains } p_{\mathcal{H}}(S_{z,t}(0)), p_{\mathcal{H}}(S_{z,t}(1)), \ldots, p_{\mathcal{H}}(S_{z,t}(\tau)) \right\}$$

where

$$\tau = \tau(z, t, N) := \min\{k : S_{z,t}(k) \notin A^{\dagger}_{t^-}[N]\}$$

denotes the time at which the particle moving according to $S_{z,t}$ exits the current aggregate $A^{\dagger}_{t^-}[N]$. In other words, the ball $\mathbb{B}(z, R_N(z, t))$ contains the part of the projected trajectory $p_{\mathcal{H}}(S_{z,t})$ until $S_{z,t}$ exits $A^{\dagger}_{t^-}[N]$. It is worth pointing out here that $R_N(z, t)$ only depends on the random walk $S_{z,t}$ and the current aggregate $A^{\dagger}_{t^-}[N]$.

Now, let us come back to the sequence of $\kappa \geq 1$ particles satisfying (12). The fact that the i-th particle is relayed by the $(i + 1)$-th one means that

$$\mathbb{B}(z_i, R_N(z_i, t_i)) \cap \mathbb{B}(z_{i+1}, R_N(z_{i+1}, t_{i+1})) \neq \emptyset \ .$$

Figure 3.5 properly illustrates our argument. On the left hand side, aggregates $A^{\dagger}_n[N_0]$ and $A^{\dagger}_n[N]$ are respectively in dark and light gray–the first one being included in the second one according to the natural coupling. The trajectory of the first particle, starting at (z_1, t_1), is depicted in red. Since $\|z_1\| > N_0$, it works only for the larger aggregate and creates a first discrepancy x_1. The trajectory of the second particle, starting at (z_2, t_2), is depicted in black and green. This particle works for both aggregates until it visits x_1 for the first time (the black path): x_1 is then added to the smaller aggregate and is no longer a discrepancy. Thus, the second particle continues its trajectory (the green path) but now working only for the larger aggregate: it creates a new discrepancy x_2 between both aggregates. Note that the radii $R_N(z_1, t_1)$ and

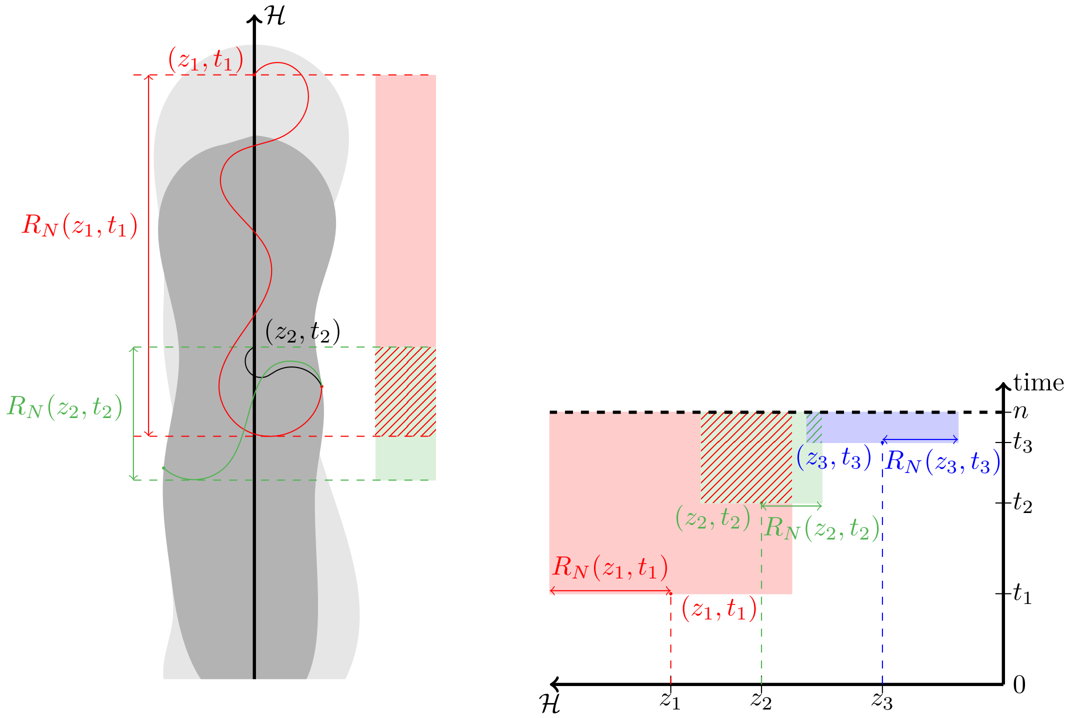


Figure 3.5: Representations of a chain of changes

$R_N(z_2, t_2)$ are represented on the left hand side of Figure 3.5. The hatched area emphasizes the fact that the balls $\mathbb{B}(z_1, R_N(z_1, t_1))$ and $\mathbb{B}(z_2, R_N(z_2, t_2))$ overlap.

This phenomenon is perhaps better visualized in the right hand side of Figure 3.5 where the extra time parameter is taken into account. Remark that not only the colored rectangles have to intersect, but they need to do so with respect to increasing times. This additional constraint will turn out to be crucial in our proof.

Conclusion. Given (N_0, N) such that the forests $\mathcal{F}_{n_0}[N]$ and $\mathcal{F}_{n_0}[N_0]$ do not coincide on $\mathbb{Z}_{K_0}$, we can assert that there exists a sequence of $\kappa \geq 1$ particles coming from starting points $(z_i, t_i)_{1 \leq i \leq \kappa}$ and satisfying the following conditions:

$$\begin{cases} N_0 < \|z_1\| \leq N \text{ and } \|z_i\| \leq N_0, \text{ for } 2 \leq i \leq \kappa \\ 0 < t_1 < t_2 < \cdots < t_\kappa < n_0 \\ \text{for } 1 \leq i \leq \kappa - 1,\, \mathbb{B}(z_i, R_N(z_i, t_i)) \cap \mathbb{B}(z_{i+1}, R_N(z_{i+1}, t_{i+1})) \neq \emptyset \\ \mathbb{B}(z_\kappa, R_N(z_\kappa, t_\kappa)) \cap \mathbb{Z}_{K_0} \neq \emptyset \end{cases} \tag{13}$$

At this stage, the radii $R_N(z_i, t_i)$'s we consider are far from easy to handle, as the law of $R_N(z_i, t_i)$ strongly depends on the shape of $A^{\dagger}_{t_i}[N]$ as well as the starting point (z_i, t_i), which is random. Consequently, the radii $R_N(z_i, t_i)$'s are neither independent, nor identically distributed. To overcome this latter obstacle, we will replace in the next section the aggregate $A^{\dagger}_{t_i^-}[N]$ involved in the definition of $R_N(z_i, t_i)$ with the larger aggregate $A^{\dagger}_T[\infty]$ (with $T \geq n_0$) whose distribution is translation invariant.

3.2 Existence of infinite descending chains under the *Absurd hypothesis*

Let $T \geq n_0$. Associated to the starting point (z,t), with $z \in \mathcal{H}$ and $t \in [0,n_0]$, let us introduce the random radius $R((z,t),T)$ defined by

$$R((z,t),T) := \min\left\{r \in \mathbb{N} : \mathbb{B}(z,r) \text{ contains } p_{\mathcal{H}}(S_{z,t}(0)), p_{\mathcal{H}}(S_{z,t}(1)), \ldots, p_{\mathcal{H}}(S_{z,t}(\tau'))\right\}$$

where

$$\tau' = \tau'(z,t,N) := \min\{k : S_{z,t}(k) \notin A_T^{\dagger}[\infty]\}$$

denotes the time at which the particle moving according to $S_{z,t}$ exits the infinite aggregate $A_T^{\dagger}[\infty]$. The new radius $R((z,t),T)$ is defined similarly as $R_N(z,t)$, but we are now considering the trajectory of $S_{z,t}$ until it exits $A_T^{\dagger}[\infty]$ rather than $A_{t-}^{\dagger}[N]$. Growing the aggregate both in time and space, from $A_{t-}^{\dagger}[N]$ to $A_T^{\dagger}[\infty]$, we then a.s have

$$R_N(z,t) \leq R((z,t),T) \ .$$

Hence, given (N_0,N) such that the forests $\mathcal{F}_{n_0}[N]$ and $\mathcal{F}_{n_0}[N_0]$ do not coincide on $\mathbb{Z}_{K_0}$, we get the existence of a sequence of $\kappa \geq 1$ particles coming from starting points $(z_i,t_i)_{1\leq i\leq\kappa}$ and satisfying the following conditions:

$$\begin{cases} N_0 < \|z_1\| \\ 0 < t_1 < t_2 < \cdots < t_\kappa < n_0 \\ \text{for } 1 \leq i \leq \kappa-1, \ \mathbb{B}(z_i, R((z_i,t_i),T)) \cap \mathbb{B}(z_{i+1}, R((z_{i+1},t_{i+1}),T)) \neq \emptyset \\ \mathbb{B}(z_\kappa, R((z_\kappa,t_\kappa),T)) \cap \mathbb{Z}_{K_0} \neq \emptyset \end{cases} \tag{14}$$

So, we now consider the space-time percolation model $\Sigma = \Sigma(n_0,T)$ defined as the collection of balls $\mathbb{B}(z, R((z,t),T))$, for any starting point (z,t) with $z \in \mathcal{H}$ and $t \in [0,n_0]$. Note also that considering larger radii, i.e. replacing $R_N(z,t)$ with $R((z,t),T)$, all dependency on the parameter N disappears in (14) and this allows us to take advantage of the stationarity property of $A_T^{\dagger}[\infty]$, which will greatly facilitate the study of Σ.

In the rest of this section, we state two properties about the percolation model Σ; a finite moment property for its radii (Lemma 3.8) and a finite degree property (Lemma 3.9).

Let us start by introducing the random radius $R_T(z)$, for $z \in \mathcal{H}$ and $T \geq n_0$, defined as

$$R_T(z) := \max_{t \in \mathcal{N}_z([0,T])} R((z,t),T) \tag{15}$$

where $\mathcal{N}_z$ denotes the PPP with intensity 1 associated to the source z. Also, $t \in \mathcal{N}_z([0,T])$ means that t is a 'top' of the PPP $\mathcal{N}_z$ occurring in $[0,T]$. Since the aggregate $A_T^{\dagger}[\infty]$ is translation invariant in distribution (w.r.t. translations of $\mathcal{H}$, see Proposition 3.4) then all the radii $R_T(z)$, $z \in \mathcal{H}$, have the same distribution. Setting $R_T := R_T(0)$, the next result states that the $R_T(z)$'s admit finite moments.

Lemma 3.8. *Let $T \geq n_0$ and $L > 0$ be real numbers. There exists a constant $C > 0$ such that for any integer M,*

$$\mathbb{P}\big(R_T \geq M\big) \leq CM^{-L} \ .$$

In particular $\mathbb{E}[(R_T)^k]$ is finite for any integer k.

Proof. Let $M \geq T$ be an a positive integer, fix $\alpha \in (1-1/d,\ 1)$. Let us first write:

$$\mathbb{P}\big(R_T \geq 2M\big) \leq \mathbb{P}\left(R_T \geq 2M,\ \#\mathcal{N}_0([0,T]) \leq M,\ \mathbf{Over}^\dagger_\alpha(M,T,\varepsilon)^c\right) + \mathbb{P}\left(\#\mathcal{N}_0([0,T]) > M\right) + \mathbb{P}\left(\mathbf{Over}^\dagger_\alpha(M,T,\varepsilon)\right)\ , \tag{16}$$

where $\mathbf{Over}^\dagger_\alpha(M,T,\varepsilon) := \{A^\dagger_T[\infty] \cap \mathbb{Z}^c_M \not\subset \mathscr{C}^\alpha_\varepsilon\}$. We know from Proposition 3.5 that the probability of $\mathbf{Over}^\dagger_\alpha(M,T,\varepsilon)$ is smaller than any power of M^{-1} provided M is sufficiently large. Additionally, we also know that $\mathbb{P}\left(\#\mathcal{N}_0([0,T]) > M\right)$ decreases faster than any power of M^{-1}, given M is large enough. It then remains to focus on the first term of (16). Hence,

$$\begin{aligned}
&\mathbb{P}\left(R_T \geq 2M,\ \#\mathcal{N}_0([0,T]) \leq M,\ \mathbf{Over}^\dagger_\alpha(M,T,\varepsilon)^c\right)\\
&\quad = \mathbb{P}\left(\exists t \in \mathcal{N}_0([0,T]),\ R((0,t),T) \geq 2M,\ \#\mathcal{N}_0([0,T]) \leq M,\ \mathbf{Over}^\dagger_\alpha(M,T,\varepsilon)^c\right)\\
&\quad \leq \mathbb{E}\left[\sum_{t\in\mathcal{N}_0([0,T])} \mathbb{1}_{R((0,t),T)>2M}\,\mathbb{1}_{\#\mathcal{N}_0([0,T])\leq M}\,\mathbb{1}_{\mathbf{Over}^\dagger_\alpha(M,T,\varepsilon)^c}\right]\\
&\quad = \mathbb{E}\left[\mathbb{1}_{\#\mathcal{N}_0([0,T])\leq M} \sum_{t\in\mathcal{N}_0([0,T])} \mathbb{P}\left(R((0,t),T) > 2M,\ \mathbf{Over}^\dagger_\alpha(M,T,\varepsilon)^c \mid \mathcal{N}_0([0,T])\right)\right].
\end{aligned}$$

Now, the event $\{R((0,t),T) > 2M\}$ implies that the random walk traveled a distance at least $2M$, and in particular, that it traveled from levels M to $2M$, while staying contained inside $A^\dagger_T[\infty]$. Now, working on $\mathbf{Over}^\dagger_\alpha(M,T,\varepsilon)^c$ implies that $A_T[\infty]$ is contained inside $\mathscr{C}^\alpha_\varepsilon$ above levels M. Hence working on both events implies that the random walk traveled from level M to level $2M$, all while staying contained inside $\mathscr{C}^\alpha_\varepsilon$. By taking $j=1$ in (10), we know that the number of boxes between levels M and $2M$, denoted by k^o, is such that

$$k^o \geq \frac{(2M)^{1-\alpha}}{4\varepsilon} = \frac{M^{1-\alpha}}{2^{1+\alpha}\varepsilon}.$$

Hence, using a box argument similar to the one used in the proof of Theorem 3.3, we can show that:

$$\begin{aligned}
\sum_{t\in\mathcal{N}_0([0,T])} \mathbb{P}\left(R((0,t),T) > 2M,\ \mathbf{Over}^\dagger_\alpha(M,T,\varepsilon)^c \mid \mathcal{N}_0([0,T])\right) &\leq \sum_{t\in\mathcal{N}_0([0,T])} \left(1-\frac{1}{4d^2}\right)^{k^o}\\
&= \#\mathcal{N}_0([0,T]) \exp\left(-c_0 M^{1-\alpha}\right),
\end{aligned}$$

where $c_0 = -\frac{1}{2^{1+\alpha}\varepsilon}\ln\left(1-\frac{1}{4d^2}\right) > 0$. Hence,

$$\begin{aligned}
\mathbb{P}\left(R_T \geq 2M,\ \#\mathcal{N}_0([0,T]) \leq M,\ \mathbf{Over}^\dagger_\alpha(M,T,\varepsilon)^c\right) &\leq \mathbb{E}\left[\#\mathcal{N}_0([0,T]) e^{-c_0 M^{1-\alpha}}\mathbb{1}_{\#\mathcal{N}_0([0,T])\leq M}\right]\\
&\leq M \exp(-c_0 M^{1-\alpha})\ ,
\end{aligned}$$

which decreases faster than any power of M^{-1}. Finally, we have proved that $\mathbb{P}(R_T \geq 2M)$ decreases faster than any power of M^{-1} from which one easily concludes. ☐

Lemma 3.8 paves the way to a finite degree property for the percolation model $\Sigma = \Sigma(n_0, T)$. Precisely, any given ball $\mathbb{B}(z, R((z,t),T))$ of Σ overlaps only a finite number of other balls

$\mathbb{B}(z', R((z',t'),T))$ of Σ, with probability 1.

Lemma 3.9. *Let $T \geq n_0$ and $z \in \mathcal{H}$. Then, with probability one, for any starting point (z,t) with $t \leq n_0$, the number of starting points (z',t'), with $z' \in \mathcal{H}$ and $t' \leq n_0$, such that $\mathbb{B}(z, R((z,t),T)) \cap \mathbb{B}(z', R((z',t'),T)) \neq \emptyset$ is finite.*

Proof. Since the radius $R_T(z)$ is by definition larger than any $R((z,t),T)$–see (15)–it is enough to prove that a.s.

$$\#\big\{z' \in \mathcal{H} : \mathbb{B}(z, R_T(z)) \cap \mathbb{B}(z', R_T(z')) \neq \emptyset\big\} < \infty \ .$$

Thus using the translation invariance property of the $R_T(z)$'s, it is also enough to state that the expectation

$$\mathbb{E}\Big[\#\big\{z \in \mathcal{H} : \mathbb{B}(0, R_T(0)) \cap \mathbb{B}(z, R_T(z)) \neq \emptyset\big\}\Big]$$

is finite. This follows from the next computation:

$$\begin{aligned}
\mathbb{E}\Big[\#\big\{z \in \mathcal{H} : \mathbb{B}(0, R_T(0)) \cap \mathbb{B}(z, R_T(z)) \neq \emptyset\big\}\Big] &= \mathbb{E}\Big[\sum_{z\in\mathcal{H}} \mathbb{1}_{R_T(0)+R_T(z)\geq\|z\|}\Big] \\
&= \sum_{z\in\mathcal{H}} \mathbb{P}\big(R_T(0)+R_T(z) \geq \|z\|\big) \\
&\leq \sum_{z\in\mathcal{H}} \mathbb{P}\big(\{R_T(0) \geq \|z\|/2\} \cup \{R_T(z) \geq \|z\|/2\}\big) \\
&\leq 2\sum_{\ell=0}^{\infty} \sum_{\|z\|=\ell} \mathbb{P}\big(R_T(0) \geq \|z\|/2\big) \\
&\leq C_d \sum_{\ell=0}^{\infty} \ell^{d-2}\, \mathbb{P}\big(R_T(0) \geq \ell\big) \ .
\end{aligned}$$

This latter sum is finite thanks to Lemma 3.8. □

Conclusion. Let us interpret the percolation model Σ as the (undirected) graph $\mathcal{G}$ whose vertex set is given by the starting points (z,t), with $z \in \mathcal{H}$ and $t \leq n_0$ and whose edge set is made of pairs $\{(z,t),(z',t')\}$ such that the corresponding balls $\mathbb{B}(z, R((z,t),T))$ and $\mathbb{B}(z', R((z',t'),T))$ overlap. Lemma 3.9 asserts that each vertex of the graph $\mathcal{G}$ almost surely has a finite degree.

Now, combining this finite degree property with the *Absurd hypothesis*, we get the existence of an *infinite descending chain*. Let us explain why. Recall that $T \geq n_0$ and K_0 are fixed parameters. The *Absurd hypothesis* says that, with positive probability, for any integer N_0, there exists a sequence of $\kappa = \kappa(N_0) \geq 1$ particles satisfying (14). With each of these sequences can be associated a cluster in the percolation model $\Sigma = \Sigma(n_0, T)$ joining the outside of $\mathbb{Z}_{N_0}$ to $\mathbb{Z}_{K_0}$ whose centers have increasing times. Roughly speaking, these sequences connect the strip $\mathbb{Z}_{K_0}$ to regions as far as desired through the percolation model Σ. Then, using the finite degree property and proceeding step by step, we can extract an infinite sequence of starting points $((z_i,t_i))_{i\geq1}$ such that

$$\begin{cases} \|z_i\| \to \infty \\ \text{for any index } i,\ 0 < t_{i+1} < t_i < n_0 \\ \text{for any index } i,\ \mathbb{B}(z_i, R((z_i,t_i),T)) \cap \mathbb{B}(z_{i+1}, R((z_{i+1},t_{i+1}),T)) \neq \emptyset \\ \mathbb{B}(z_1, R((z_1,t_1),T)) \cap \mathbb{Z}_{K_0} \neq \emptyset \end{cases} \tag{17}$$

Note that, w.r.t. (14), we have reversed the indices of the time sequence $(t_i)_{i\geq1}$. On the one hand, this sequence of starting points $((z_i,t_i))_{i\geq1}$ is *infinite* and means that Σ percolates since it contains an unbounded cluster. On the other hand, it is also *descending* since the sequence of starting times $(t_i)_{i\geq 1}$ is decreasing.

Let $\mathcal{K}_0 := \mathcal{H}_{K_0}$. The previous analysis allows us to say that, under the *Absurd hypothesis*, the following holds

$$\mathbb{P}\big(\mathbf{Perco}(n_0, \mathcal{K}_0, n_0)\big) > 0 \ , \tag{18}$$

where, for any $0 \leq t \leq T$ and any compact set $\mathcal{K} \subset \mathcal{H}$,

$$\mathbf{Perco}(t, \mathcal{K}, T) := \left\{ \begin{array}{c} \exists \text{ a sequence } ((z_i, t_i))_{i\geq 1} \text{ of starting points s.t. } \|z_i\| \to \infty, \\ \mathbb{B}(z_1, R((z_1, t_1), T)) \cap \mathcal{K} \neq \emptyset, \text{ and for any } i \geq 1, \ t_{i+1} < t_i < t \\ \text{and } \mathbb{B}(z_i, R((z_i, t_i), T)) \cap \mathbb{B}(z_{i+1}, R((z_{i+1}, t_{i+1}), T)) \neq \emptyset \end{array} \right\} .$$

In the next section, we will take advantage of the monotonicity of the time sequence $(t_i)_{i\geq 1}$ to establish that the infinite descending chain mentioned above appears instantaneously.

3.3 Instantaneous percolation

For $T \geq t \geq 0$, let us set:

$$\mathbf{Perco}(t, T) := \bigcup_{\mathcal{K}} \mathbf{Perco}(t, \mathcal{K}, T) \ , \tag{19}$$

where the union above is taken over all compact subsets of $\mathcal{H}$. The event $\mathbf{Perco}(t, T)$ ensures the existence of an infinite descending chain made up with balls coming from the percolation model Σ, anchored at some (random) $\mathcal{K} \subset \mathcal{H}$. The event $\mathbf{Perco}(t, T)$ is monotone w.r.t. parameters t and T:

Lemma 3.10. *Let $0 \leq t \leq t' \leq T \leq T'$. Then,*

$$\mathbf{Perco}(t, T) \subset \mathbf{Perco}(t, T') \ \textit{ and } \ \mathbf{Perco}(t, T) \subset \mathbf{Perco}(t', T) \ .$$

Proof. Let $0 \leq t \leq T \leq T'$. Recall that the infinite aggregate $A^{\dagger}_T[\infty]$ corresponds to particles launched from the whole set of sources $\mathcal{H}$ and during the time interval $[0, T]$. Hence, $A^{\dagger}_T[\infty]$ and $A^{\dagger}_{T'}[\infty]$ can be naturally coupled so that a.s. $A^{\dagger}_T[\infty] \subset A^{\dagger}_{T'}[\infty]$. This leads to $R((z_i, t_i), T) \leq R((z_i, t_i), T')$ whatever the starting point (z_i, t_i) with $t_i \leq t$. So, $\mathbf{Perco}(t, \mathcal{K}, T)$ is included in $\mathbf{Perco}(t, \mathcal{K}, T')$, for any compact set $\mathcal{K}$, and the same holds for $\mathbf{Perco}(t, T)$ and $\mathbf{Perco}(t, T')$.

The second inclusion is easy to prove. Indeed, replacing t with $t' \geq t$ amounts to relax the upper bound on the decreasing sequence $(t_i)_{i\geq 1}$ appearing in the event $\mathbf{Perco}(t, \mathcal{K}, T)$. So $\mathbf{Perco}(t, \mathcal{K}, T)$ is included in $\mathbf{Perco}(t', \mathcal{K}, T)$, for any $\mathcal{K}$, and the same holds for $\mathbf{Perco}(t, T)$ and $\mathbf{Perco}(t', T)$. □

From now on, we fix $T := n_0 + 1$. Let us introduce the critical percolation time as

$$t_c = t_c(T) := \inf \big\{ 0 \leq t \leq T : \mathbb{P}\big(\mathbf{Perco}(t, T) \big) > 0 \big\} \in [0, T] \ .$$

Combining to the monotone property given by Lemma 3.10, we can then write:

$$\left\{ \begin{array}{l} \mathbb{P}\left(\mathbf{Perco}(t, T)\right) = 0 \text{ for any } 0 \leq t < t_c, \\ \mathbb{P}\left(\mathbf{Perco}(t, T)\right) > 0 \text{ for any } t_c < t \leq T. \end{array} \right. \tag{20}$$

Both statements of (20) become meaningful provided the critical percolation time t_c is non-trivial, i.e. different from 0 and T. This is where the *Absurd hypothesis* steps in since (18) and

Lemma 3.10 imply that

$$\mathbb{P}\big(\mathbf{Perco}(n_0,T)\big) \geq \mathbb{P}\big(\mathbf{Perco}(n_0,n_0)\big) \geq \mathbb{P}\big(\mathbf{Perco}(n_0,\mathcal{K}_0,n_0)\big) > 0$$

that is to say

$$t_c \leq n_0 < T \ . \tag{21}$$

Actually, the condition $t_c < T$–ensured by the *Absurd hypothesis*–leads to a phenomenon of *instantaneous percolation* for the percolation model Σ that we describe below.

Let us first assume that the critical percolation time t_c is positive. The case $t_c = 0$ is similar and will be treated after. Hence for any $\varepsilon > 0$ small enough (i.e. such that $t_c - \varepsilon \geq 0$ and $t_c + \varepsilon \leq T$), we have $\mathbb{P}(\mathbf{Perco}(t_c+\varepsilon,T)) > 0$ while $\mathbb{P}(\mathbf{Perco}(t_c-\varepsilon,T)) = 0$ by (20), meaning that

$$\mathbb{P}\big(\mathbf{Perco}(t_c+\varepsilon,T)\backslash\mathbf{Perco}(t_c-\varepsilon,T)\big) > 0 \ .$$

Let us analyze what happens on the event $\mathbf{Perco}(t_c+\varepsilon,T)\backslash\mathbf{Perco}(t_c-\varepsilon,T)$. First of all, the event $\mathbf{Perco}(t_c+\varepsilon,T)$ asserts the existence of an infinite descending chain in Σ associated to a sequence of starting points $((z_i,t_i))_{i\geq 1}$ with, for any $i \geq 1$, $t_{i+1} < t_i < t_c + \varepsilon$. Besides, the event $\mathbf{Perco}(t_c-\varepsilon,T)^c$ forces all the t_i's to be larger than $t_c - \varepsilon$. Otherwise there would exist some index i_0 such that $t_{i_0} < t_c - \varepsilon$. In that case, one would have an infinite descending chain associated to the sequence of starting points $((z_i,t_i))_{i\geq i_0}$ anchored at $\mathbb{B}(z_{i_0}, R((z_{i_0},t_{i_0}),T))$ and such that for any i, $t_{i+1} < t_i \leq t_{i_0} < t_c - \varepsilon$, meaning that $\mathbf{Perco}(t_c-\varepsilon,T)$ actually occurs. In conclusion, the sequence of starting points $((z_i,t_i))_{i\geq 1}$ satisfies $t_c - \varepsilon < t_{i+1} < t_i < t_c + \varepsilon$ for any index i. This means that this infinite descending chain appears entirely during the time interval $[t_c-\varepsilon, t_c+\varepsilon]$, with length 2ε, for $\varepsilon > 0$ arbitrarily small. This is why we talk about instantaneous percolation. Let us emphasize that the previous argument works because the radii $R((z_i,t_i),T)$'s remain the same in both events $\mathbf{Perco}(t_c-\varepsilon,T)$ and $\mathbf{Perco}(t_c+\varepsilon,T)$ since they have the same second parameter T.

In the particular case $t_c = 0$, we know that there is no percolation at the critical time t_c since no particles have been launched yet. So we apply the previous analysis with $t_c + \varepsilon = \varepsilon$ and $t_c - \varepsilon$ replaced with 0. As before, an infinite descending chain appears entirely during the time interval $[0,\varepsilon]$, for any small $\varepsilon > 0$.

In order to simplify the argumentation and lighten notations, let us reduce the first case $t_c > 0$ to the second one $t_c = 0$. Indeed, because the radii $R((z_i,t_i),T)$'s remain unchanged in the events $\mathbf{Perco}(\cdot,T)$, the time interval during which the infinite descending chain entirely occurs matters in distribution only through its length (the PPP $\mathcal{N}_z$'s have stationary increments). Henceforth,

$$\Big(\forall\varepsilon > 0,\ \mathbb{P}\big(\mathbf{Perco}(t_c+\varepsilon,T)\backslash\mathbf{Perco}(t_c-\varepsilon,T)\big) > 0\Big) \implies \Big(\forall\varepsilon > 0,\ \mathbb{P}\big(\mathbf{Perco}(\varepsilon,T)\big) > 0\Big) \ .$$

Conclusion. We have taken advantage of the monotonicity in time of the sequence of starting points $((z_i,t_i))_{i\geq 1}$ to state that the corresponding infinite descending chain in the percolation model Σ appears instantaneously. Precisely, we have proven under the *Absurd hypothesis* that

$$\forall\varepsilon > 0,\ \mathbb{P}\big(\mathbf{Perco}(\varepsilon,T)\big) > 0 \ . \tag{22}$$

We can now forget the time dimension of our model: (22) allows us to exhibit a *supercritical* discrete Boolean model, denoted by $\hat{\Sigma}_\varepsilon$, whose intensity tends to 0 as $\varepsilon \to 0$.

Let us set, for any source $z \in \mathcal{H}$,

$$Y_z := \mathbb{1}_{\#\mathcal{N}_z([0,\varepsilon])>0} \ .$$

So $Y_z = 1$ means that at least one particle has been emitted during the time interval $[0, \varepsilon]$. The Y_z's are Bernoulli random variables with common parameter

$$p_\varepsilon := \mathbb{P}\left(Y_z = 1\right) = 1 - e^{-\varepsilon} \tag{23}$$

which tends to 0 as $\varepsilon \to 0$. By hypothesis on the PPP $\mathcal{N}_z$'s, the random variables Y_z, $z \in \mathcal{H}$, are also independent from each other. Let us denote by

$$\chi_\varepsilon := \left\{z \in \mathcal{H} : Y_z = 1\right\}$$

the random set of emitting sources during the time interval $[0, \varepsilon]$. The set χ_ε can be interpreted as a discrete PPP on $\mathcal{H}$ with intensity p_ε: it will play the role of the center set for the Boolean model $\hat{\Sigma}_\varepsilon$. In all that follows, let us keep in mind that when $\varepsilon \to 0$ there are very few centers, and so very few balls, in the Boolean model $\hat{\Sigma}_\varepsilon$.

Thus, for $z \in \mathcal{H}$, let us define in the same spirit of (15) the radius $R_T(z; \varepsilon)$ as

$$R_T(z; \varepsilon) := \max_{t \in \mathcal{N}_z([0,\varepsilon])} R((z, t), T) \ .$$

The only difference between radii $R_T(z; \varepsilon)$ and $R_T(z)$ defined in (15) is that $R_T(z; \varepsilon)$ involves particles launched during $[0, \varepsilon]$ while $R_T(z)$ involves particles of $[0, T]$, so that $R_T(z; \varepsilon) \leq R_T(z)$ (and then $R_T(z; \varepsilon)$ also satisfies the finite moment property of Lemma 3.8). They both refer to radii $R(\cdot, T)$ defined from the aggregate $A_T^\dagger[\infty]$ (with $T = n_0 + 1$).

We are now ready to define the (discrete) Boolean model $\hat{\Sigma}_\varepsilon$ by

$$\hat{\Sigma}_\varepsilon := \bigcup_{z \in \chi_\varepsilon} \mathbb{B}\big(z, R_T(z; \varepsilon)\big) \ .$$

Statement (22) says that, for any $\varepsilon > 0$ and with positive probability, there exists a sequence of starting points $((z_i, t_i))_{i \geq 1}$ such that $\|z_i\| \to \infty$ and, for any $i \geq 1$, $t_{i+1} < t_i < \varepsilon$ and the balls $\mathbb{B}(z_i, R((z_i, t_i), T))$ and $\mathbb{B}(z_{i+1}, R((z_{i+1}, t_{i+1}), T))$ overlap. So, the z_i's all belong to χ_ε and the larger balls $\mathbb{B}(z_i, R_T(z_i; \varepsilon))$ and $\mathbb{B}(z_{i+1}, R_T(z_{i+1}; \varepsilon))$ overlap too. In other words, the Boolean model $\hat{\Sigma}_\varepsilon$ contains an unbounded cluster.

Finally, the argumentation of the whole Section 3 provides the following result:

Proposition 3.11. *Under the Absurd hypothesis (6), the following holds:*

$$\forall \varepsilon > 0, \ \mathbb{P}\Big(\hat{\Sigma}_\varepsilon \text{ percolates}\Big) > 0 \ . \tag{24}$$

4 A multiscale argument

In this section, we study the $(d-1)$-dimensional, discrete Boolean model

$$\hat{\Sigma}_\varepsilon = \bigcup_{z \in \chi_\varepsilon} \mathbb{B}\big(z, R_T(z; \varepsilon)\big) \ ,$$

defined in the previous section, and whose intensity p_ε tends to 0 as $\varepsilon \to 0$. We adapt the strategy of [22] to our context to state :

Proposition 3.12. *For any $\varepsilon > 0$ small enough, the Boolean model $\hat{\Sigma}_\varepsilon$ does not percolate with probability 1.*

To implement the strategy of [22], we need to make our model more local. This is the role of $\hat{\Sigma}^{\text{loc}}_\varepsilon$ defined below. For that purpose, the stabilization result for the infinite aggregate $A^\dagger_n[\infty]$ (Theorem 3.3) is an important ingredient to ensure that the localized model $\hat{\Sigma}^{\text{loc}}_\varepsilon$ is a good approximation of $\hat{\Sigma}_\varepsilon$.

Proof of Theorem 3.1. Propositions 3.11 and 3.12 together say that the *Absurd hypothesis* (6) leads to a contradiction. We then conclude that, with probability 1, there exists N_0 such that, for any $N \geq N_0$, the forests $\mathcal{F}_{n_0}[N]$ and $\mathcal{F}_{n_0}[N_0]$ coincide on the strip $\mathbb{Z}_{K_0}$. This concludes the proof of Theorem 3.1. □

4.1 The localized Boolean models $\hat{\Sigma}^{\text{loc}}_\varepsilon$

Recall that $T = n_0 + 1$ and $\varepsilon > 0$ is thought to be small. Given a source $x \in \mathcal{H}$, we denote by $A^\dagger_T[\mathbb{B}(x, 20M)]$ the aggregate $A^\dagger_T[\cdot]$ using only sources of $\mathbb{B}(x, 20M)$. Associated to the starting point (z, t), with $z \in \mathcal{H}$ and $t \in [0, \varepsilon]$, let us introduce the local radius $R^{\text{loc}}_{x,M}((z, t), T)$ defined by

$$R^{\text{loc}}_{x,M}((z,t),T) := \min\left\{r \in \mathbb{N} : \mathbb{B}(z,r) \text{ contains } p_{\mathcal{H}}(S_{z,t}(0)), p_{\mathcal{H}}(S_{z,t}(1)), \ldots, p_{\mathcal{H}}(S_{z,t}(\tau''))\right\} \tag{25}$$

where

$$\tau'' := \min\{k : S_{z,t}(k) \notin A^\dagger_T[\mathbb{B}(x, 20M)]\}$$

denotes the time at which the particle moving according to $S_{z,t}$ exits $A^\dagger_T[\mathbb{B}(x, 20M)]$. Taking the maximum over starting points (z, t) with $t \in \mathcal{N}_z([0, \varepsilon])$, we get

$$R^{\text{loc}}_T(z;\varepsilon) = R^{\text{loc}}_{T,x,M}(z;\varepsilon) := \max_{t\in\mathcal{N}_z([0,\varepsilon])} R^{\text{loc}}_{x,M}((z,t),T) \ .$$

Let us now define the *localized Boolean model* $\hat{\Sigma}^{\text{loc}}_\varepsilon(x, M)$ which is a version of $\hat{\Sigma}_\varepsilon$ localized to the neighborhood of x:

$$\hat{\Sigma}^{\text{loc}}_\varepsilon = \hat{\Sigma}^{\text{loc}}_\varepsilon(x,M) := \bigcup_{z\in\chi_\varepsilon\cap\mathbb{B}(x,10M)} \mathbb{B}\Big(z, R^{\text{loc}}_{T,x,M}(z;\varepsilon)\Big) \ .$$

Unlike $\hat{\Sigma}_\varepsilon$, the localized Boolean model $\hat{\Sigma}^{\text{loc}}_\varepsilon$ is local in the following sense. It only uses sources of χ_ε contained inside $\mathbb{B}(x, 10M)$. The radii that it considers depend only on $A^\dagger_T[\mathbb{B}(x, 20M)]$.

As in [22], we consider the event

$$G_\varepsilon(x,M) := \left\{ \begin{array}{c} \text{The connected component of } x \text{ in } \hat{\Sigma}^{\text{loc}}_\varepsilon(x,M) \cup \mathbb{B}(x,M) \\ \text{is not included in } \mathbb{B}(x,8M) \end{array} \right\} .$$

Since our model is translation invariant, we have for any x:

$$\pi_\varepsilon(M) := \mathbb{P}\big(G_\varepsilon(0,M)\big) = \mathbb{P}\big(G_\varepsilon(x,M)\big) \ .$$

Let us denote by $\hat{C}_\varepsilon(0)$ the cluster of the source 0 in the discrete Boolean model $\hat{\Sigma}_\varepsilon$ whose diameter $\operatorname{diam} \hat{C}_\varepsilon(0)$ is defined as the minimal integer r such that $\hat{C}_\varepsilon(0) \subset \mathbb{B}(0,r)$. The next two results allow us to conclude.

Proposition 3.13. *There exists a constant $C = C_{T,d} > 0$ such that for all $M, L \geq 1$,*

$$\mathbb{P}\big(\operatorname{diam} \hat{C}_\varepsilon(0) \geq 8M\big) \leq \pi_\varepsilon(M) + \frac{C}{M^L} \ .$$

Proposition 3.14. *For any $\varepsilon > 0$ small enough we have*

$$\liminf_{M\to\infty} \pi_\varepsilon(M) = 0 \ .$$

Proof of Proposition 3.12. Taking $\varepsilon > 0$ small enough according to Proposition 3.14, the two previous results imply that

$$\begin{aligned} \lim_{M\to\infty} \mathbb{P}\big(\operatorname{diam} \hat{C}_\varepsilon(0) \geq 8M\big) &= \liminf_{M\to\infty} \mathbb{P}\big(\operatorname{diam} \hat{C}_\varepsilon(0) \geq 8M\big) \\ &\leq \liminf_{M\to\infty} \pi_\varepsilon(M) \ = \ 0 \ , \end{aligned}$$

meaning that the source 0 almost surely belongs to a finite connected component in $\hat{\Sigma}_\varepsilon$. This discrete Boolean model being translation invariant in distribution, we conclude that it a.s. admits only finite connected components. I.e. $\hat{\Sigma}_\varepsilon$ does not percolate with probability 1. □

The next two sections are devoted to the proofs of Propositions 3.13 and 3.14.

4.2 Proof of Proposition 3.13

Let us introduce for any $M \geq 1$ the event $H(M)$ defined by

$$H(M) := \Big\{\text{each ball of } \hat{\Sigma}_\varepsilon \text{ intersecting } \mathbb{B}(0,10M) \text{ has a radius smaller than } M\Big\} \ .$$

The following lemma gives a control for the probability of $H(M)$.

Lemma 3.15. *There exists a positive constant $C = C(d,\varepsilon)$ such that for any $M, L \geq 1$,*

$$\mathbb{P}\big(H(M)^c\big) \leq \frac{C}{M^L} \ .$$

Let us first prove Proposition 3.13 from Lemma 3.15 and in a second step, Lemma 3.15 will be proven.

Proof of Proposition 3.13. We define the event

$$\mathcal{G}(M) := H(M) \cap \Big\{A_T^\dagger[\infty] \cap \mathbb{Z}_{10M} = A_T^\dagger[\mathbb{B}(0,20M)] \cap \mathbb{Z}_{10M}\Big\} .$$

Now, we use the stabilization result Theorem 3.3 which, combined to Lemma 3.15, provides the following control on the probability of $\mathcal{G}(M)$:

$$\mathbb{P}\left(\mathcal{G}(M)\right) \geq 1 - \frac{C}{M^L},$$

for some positive constant C and for any $M, L \geq 1$. Recalling that $\pi_\varepsilon(M)$ denotes the probability of $G_\varepsilon(0, M)$, it is then sufficient to prove the inclusion:

$$\left\{\operatorname{diam} \hat{C}_\varepsilon(0) \geq 8M\right\} \cap \mathcal{G}(M) \subset G_\varepsilon(0, M) \ .$$

Let us assume that the event $\{\operatorname{diam} \hat{C}_\varepsilon(0) \geq 8M\} \cap \mathcal{G}(M)$ holds. This implies the existence of a cluster $\mathcal{C}$ of $\hat{\Sigma}_\varepsilon$ containing 0 and going beyond $\mathbb{B}(0, 8M)$. Each ball of $\mathcal{C}$ overlaps $\mathbb{B}(0, 8M)$ and then has a radius smaller than M thanks to $H(M)$. So they are included in $\mathbb{B}(0, 10M)$ and their centers belong to $\chi_\varepsilon \cap \mathbb{B}(0, 10M)$.

Given a vertex z, the radii $R_T^{\text{loc}}(z; \varepsilon)$ and $R_T(z; \varepsilon)$ used resp. for $\hat{\Sigma}_\varepsilon^{\text{loc}}$ and $\hat{\Sigma}_\varepsilon$ are possibly different since they resp. refer to $A_T^\dagger[\mathbb{B}(0, 20M)]$ and $A_T^\dagger[\infty]$. However, we prove that on the event $\{\operatorname{diam} \hat{C}_\varepsilon(0) \geq 8M\} \cap \mathcal{G}(M)$, these radii are equal for the balls involved in the cluster $\mathcal{C}$. Indeed, these balls are included in $\mathbb{B}(0, 10M)$ and, on $\mathcal{G}(M)$, the aggregates $A_T^\dagger[\infty]$ and $A_T^\dagger[\mathbb{B}(0, 20M)]$ coincide on $\mathbb{Z}_{10M}$. So the radii $R_T^{\text{loc}}(z; \varepsilon)$ and $R_T(z; \varepsilon)$ coincide for each center z of balls involved in $\mathcal{C}$. Therefore, on $\{\operatorname{diam} \hat{C}_\varepsilon(0) \geq 8M\} \cap \mathcal{G}(M)$, there exists a cluster of balls of $\hat{\Sigma}_\varepsilon$ such that:

- all these balls have their centers in $\chi_\varepsilon \cap \mathbb{B}(0, 10M)$;
- all these balls have their radii given by $R_T^{\text{loc}}(z; \varepsilon)$;
- the cluster contains 0 and goes beyond $\mathbb{B}(0, 8M)$.

Then, this cluster of balls is also a cluster of $\hat{\Sigma}_\varepsilon^{\text{loc}}$ and the event $G_\varepsilon(0, M)$ occurs. □

Proof of Lemma 3.15. Fix $M, L \geq 1$. Let us introduce the following events:

$$\textbf{Inside}(M) := \left\{ \begin{array}{c} \text{There exists a ball of } \hat{\Sigma}_\varepsilon \text{ centered } \textbf{inside } \mathbb{B}(0, 20M) \\ \text{that intersects } \mathbb{B}(0, 10M) \text{ with a radius greater than } M \end{array} \right\},$$

$$\textbf{Outside}(M) := \left\{ \begin{array}{c} \text{There exists a ball of } \hat{\Sigma}_\varepsilon \text{ centered } \textbf{outside } \mathbb{B}(0, 20M) \\ \text{that intersects } \mathbb{B}(0, 10M) \end{array} \right\}.$$

Hence, the event $H(M)^c$ can be written as the union $\textbf{Inside}(M) \cup \textbf{Outside}(M)$ and we have to work with the probability of these two events. We begin by handling the event $\textbf{Inside}(M)$. Let us write:

$$\begin{aligned} \mathbb{P}(\textbf{Inside}(M)) &= \mathbb{P}\left(\bigcup_{z \in \chi_\varepsilon \cap \mathbb{B}(0, 20M)} \left\{ R_T(z; \varepsilon) \geq M \right\} \right) \\ &\leq \sum_{z \in \mathbb{B}(0, 20M)} \mathbb{P}\left(R_T(z; \varepsilon) \geq M\right) \\ &\leq C_d M^{d-1} \mathbb{P}\left(R_T(0; \varepsilon) \geq M\right) \ . \end{aligned}$$

Using Lemma 3.8, $\mathbb{P}(R_T(0; \varepsilon) \geq M)$ decreases faster than any power of M^{-1}, which handles this case.

We switch our focus to the event $\textbf{Outside}(M)$ which is trickier to handle since it deals with an infinite number of particles outside of $\mathbb{B}(0, 20M)$. Let us recall the event

$$\textbf{Over}_\alpha^\dagger(10M, T, \delta) = \left\{ A_T^\dagger[\infty] \cap \mathbb{Z}_{10M}^c \not\subset \mathscr{C}_\delta^\alpha \right\}$$

introduced in Section 2.1, where

$$\mathscr{C}_\delta^\alpha = \bigcup_{\ell \geq 0} \left\{ z \in \mathbb{Z}^d,\ \|p_{\mathcal{H}}(z)\| = \ell,\ |z_1| \leq \delta \ell^\alpha \right\} .$$

For $\delta > 0$ and $\alpha \in (1 - 1/d, 1)$, we know thanks to the global upper bound Proposition 3.5 that the probability of $\mathbf{Over}_\alpha^\dagger(10M, T, \delta)$ decreases faster than any power of M^{-1}. So we can restrict our attention to the event $\mathbf{Outside}(M) \cap \mathbf{Over}_\alpha^\dagger(10M, T, \delta)^c$. The event $\mathbf{Outside}(M)$ provides the existence of a ball of the Boolean model $\hat{\Sigma}_\varepsilon$ that intersects $\mathbb{B}(0, 10M)$ and whose center is beyond level $20M$. This ball is due to a particle starting (during the time interval $[0, \varepsilon]$) from a source beyond level $20M$ and visiting the strip $\mathbb{Z}_{10M}$ before exiting the aggregate $A_T^\dagger[\infty]$. Thanks to $\mathbf{Over}_\alpha^\dagger(10M, T, \delta)^c$, we can assert that the random walk associated to that particle starts from a level greater than $20M$ and visits $\mathbb{Z}_{10M}$ before exiting the cone $\mathscr{C}_\delta^\alpha$. This implies the event D_{10M}^c introduced in Section 2.3 whose probability is smaller than any power of M^{-1} thanks to Lemma 3.6. This concludes the proof. □

4.3 Proof of Proposition 3.14

This section is an adaptation of [22]. Lemmas 3.16 and 3.17 together imply, by induction, that $\pi_\varepsilon(10^n M) \to 0$ as $n \to \infty$ for some M and ε small enough. Lemma 3.16 is the induction step, allowing to go from scale $10^n M$ to $10^{n+1} M$, while Lemma 3.17 is the base step.

Lemma 3.16. *There exist positive constants $c = c(d)$ and $C = C(T, d)$ such that for any $M, L \geq 1$,*

$$\pi_\varepsilon(10M) \leq c\pi_\varepsilon(M)^2 + \frac{C}{M^L} .$$

Lemma 3.17. *There exists a positive constant $C' = C'(d)$ such that for all $M \geq 1$ and all $\varepsilon > 0$:*

$$\pi_\varepsilon(M) \leq \varepsilon C' M^{d-1} .$$

Let us prove Proposition 3.14 from Lemmas 3.16 and 3.17.

Proof of Proposition 3.14. This is an adaptation of Lemma 3.7 of [22]. Setting $f(M) := c\pi_\varepsilon(M)$ and $g(M) := 10cC/M$, Lemma 3.16 provides

$$f(10M) \leq f(M)^2 + g(10M) . \tag{26}$$

Since $g(M) \to 0$, we can pick M_0 such that, for any $M \geq M_0$, $g(M) \leq 1/4$. Thus, using Lemma 3.17, there exists $\varepsilon_0 = \varepsilon_0(M_0) > 0$ small enough such that, for $\varepsilon \leq \varepsilon_0$ and $M \leq M_0$,

$$f(M) = c\pi_\varepsilon(M) \leq c\varepsilon_0 C' M_0^{d-1} \leq 1/2 .$$

So, the function f is bounded by $1/2$ on the interval $(0, M_0]$. Let us first extend this bound on $(M_0, 10M_0]$. To do it, let us fix $\varepsilon \in (0, \varepsilon_0)$. Thanks to (26), we can write for any $M \in (M_0, 10M_0]$:

$$f(M) \leq f(M/10)^2 + g(M) \leq \left(\frac{1}{2}\right)^2 + \frac{1}{4} = \frac{1}{2} .$$

Iterating this argument, we prove by induction that $f(M) \leq 1/2$ for any $M > 0$.

As a consequence, we deduce from (26) that, for any $M > 0$,

$$f(10M) \leq \frac{1}{2} f(M) + g(10M) \leq \frac{1}{2} + g(10M) ,$$

from which is not difficult to get, once again by induction, that the following holds for any integer n,

$$f(10^n M) \leq \frac{1}{2^n} + g(10^n M) + \frac{g(10^{n-1}M)}{2} + \ldots + \frac{g(10M)}{2^{n-1}} \ . \tag{27}$$

Henceforth, using (27) and $g(10^n M) \to 0$ as $n \to \infty$, we prove that $f(10^n M)$ tends to 0 as $n \to \infty$, which is the searched result. ☐

Proof of Lemma 3.15. Let us first consider the event $F(M)$ defined by

$$F(M) := \{\forall z \in \chi_\varepsilon \cap \mathbb{B}(0, 100M),\ R_T(z;\varepsilon) \leq M\}$$

for which, any ball of $\hat{\Sigma}_\varepsilon$ centered at some $z \in \chi_\varepsilon \cap \mathbb{B}(0, 100M)$ has a radius $R_T(z;\varepsilon)$ smaller than M. Since these radii are identically distributed and satisfy Lemma 3.8), the event $F(M)$ is very likely. There exists a constant $C > 0$ such that for any $M, L \geq 1$,

$$\mathbb{P}\big(F(M)^c\big) \leq \frac{C}{M^L} \ . \tag{28}$$

Besides, for $M \geq 1$, let us consider the event

$$\mathbf{Stab}(0, 100M) := \bigcap_{z \in \mathbb{B}(0,100M)} \mathbf{Stab}^M(z)$$

where, for any $z \in \mathcal{H}$,

$$\mathbf{Stab}^M(z) := \left\{ \begin{array}{c} A_T^\dagger[\mathbb{B}(z, 20M)] \cap (\mathbb{Z} \times \mathbb{B}(z, 10M)) = A_T^\dagger[S] \cap (\mathbb{Z} \times \mathbb{B}(z, 10M))\,, \\ \text{for any } S \subset \mathcal{H} \text{ such that } \mathbb{B}(z, 20M) \subset S. \end{array} \right\} .$$

In particular, $\mathbf{Stab}^M(0)$ means that $A_T^\dagger[\mathbb{B}(0, 20M)]$ (also denoted by $A_T^\dagger[20M]$ for short) and $A_T^\dagger[S]$ coincide on $\mathbb{Z}_{10M}$, for any $\mathcal{H}_{20M} \subset S \subset \mathcal{H}$. Let us prove that there exists a constant $C > 0$ such that for any $M, L \geq 1$,

$$\mathbb{P}\big(\,\mathbf{Stab}(0, 100M)^c\big) \leq \frac{C}{M^L} \ . \tag{29}$$

Let us start by writing, using translation invariance of the aggregates $A_T^\dagger[\cdot]$,

$$\mathbb{P}\big(\,\mathbf{Stab}(0, 100M)^c\big) \leq \sum_{z \in \mathbb{B}(0,100M)} \mathbb{P}\big(\,\mathbf{Stab}^M(z)^c\big) \leq c(100M)^{d-1}\mathbb{P}\big(\,\mathbf{Stab}^M(0)^c\big) \ .$$

So it is sufficient to show that $\mathbf{Stab}^M(0)^c$ has a probability decreasing faster than any power of M^{-1}. The *Aggregate stabilization* result Theorem 3.3 asserts that, with probability larger than $1 - CM^{-L}$, the aggregates $A_T^\dagger[20M]$ and $A_T^\dagger[\infty]$ coincide on the strip $\mathbb{Z}_{10M}$. Then the same holds for $A_T^\dagger[20M]$ and $A_T^\dagger[S]$, for any $\mathcal{H}_{20M} \subset S \subset \mathcal{H}$ since $A_T^\dagger[20M] \subset A_T^\dagger[S] \subset A_T^\dagger[\infty]$ by the natural coupling. So $\mathbf{Stab}^M(0)$ occurs with probability larger than $1 - CM^{-L}$, and then (29) is proven.

Let $\mathbb{S}_r$ be the $(d-2)$-dimensional sphere centered at the origin, with radius r and included in the source set $\mathcal{H}$: $\mathbb{S}_r := \{z \in \mathcal{H} : \|z\| = r\}$. We claim that the following key inclusion holds for any M,

$$G_\varepsilon(0,10M)\cap F(M)\cap \mathbf{Stab}(0,100M)\subset\Big(\bigcup_{c\in\mathbb{S}_{10}} G_\varepsilon(Mc,M)\Big)\cap\Big(\bigcup_{c'\in\mathbb{S}_{80}} G_\varepsilon(Mc',M)\Big)\ . \qquad (30)$$

Lemma 3.15 actually appears as a straight consequence of the inclusion (30) combined with (28) and (29). Let us first explain why and, in a second step, we will establish (30).

Inequalities (28) and (29) allow to write

$$\begin{aligned}\pi_\varepsilon(10M)=\mathbb{P}\big(G_\varepsilon(0,10M)\big) &\leq \mathbb{P}\big(G_\varepsilon(0,10M)\cap F(M)\cap\mathbf{Stab}(0,100M)\big)+\frac{C}{M^L}\\ &\leq \mathbb{P}\Big(\Big(\bigcup_{c\in\mathbb{S}_{10}} G_\varepsilon(Mc,M)\Big)\cap\Big(\bigcup_{c'\in\mathbb{S}_{80}} G_\varepsilon(Mc',M)\Big)\Big)+\frac{C}{M^L}\ .\end{aligned}$$

Recall that the event $G_\varepsilon(Mc,M)$ involves balls of $\hat{\Sigma}^{\text{loc}}_\varepsilon$ whose centers are included in $\mathbb{B}(Mc,10M)$ and radii are defined w.r.t. the aggregate $A^\dagger_T[\mathbb{B}(Mc,20M)]$. So the event $\bigcup_{c\in\mathbb{S}_{10}} G_\varepsilon(Mc,M)$ only concerns random inputs (i.e. Poisson clocks and random walks) associated to sources of $\mathcal{H}_{30M}$. In the same way, $\bigcup_{c'\in\mathbb{S}_{80}} G_\varepsilon(Mc',M)$ only concerns random inputs associated to sources outside of $\mathcal{H}_{59M}$. So they are independent from each other. It is then easy to conclude:

$$\begin{aligned}\pi_\varepsilon(10M) &\leq \mathbb{P}\Big(\bigcup_{c\in\mathbb{S}_{10}} G_\varepsilon(Mc,M)\Big)\,\mathbb{P}\Big(\bigcup_{c'\in\mathbb{S}_{80}} G_\varepsilon(Mc',M)\Big)+\frac{C}{M^L}\\ &\leq |\mathbb{S}_{10}||\mathbb{S}_{80}|\pi_\varepsilon(M)^2+\frac{C}{M^L}\ .\end{aligned}$$

As a consequence, it only remains to establish the inclusion (30) and, to do it, let us assume that $G_\varepsilon(0,10M)\cap F(M)\cap\mathbf{Stab}(0,100M)$ occurs. On the event $G_\varepsilon(0,10M)$, the localized Boolean model

$$\hat{\Sigma}^{\text{loc}}_\varepsilon(0,10M)=\bigcup_{z\in\chi_\varepsilon\cap\mathbb{B}(0,100M)}\mathbb{B}\Big(z,R^{\text{loc}}_{T,0,10M}(z;\varepsilon)\Big)$$

contains a cluster $\mathcal{C}$ joining $\mathbb{B}(0,10M)$ to $\mathbb{B}(0,80M)^c$. The balls of $\hat{\Sigma}^{\text{loc}}_\varepsilon(0,10M)$ are centered at vertices z in $\mathbb{B}(0,100M)$ and their radii $R^{\text{loc}}_{T,0,10M}(z;\varepsilon)$ are by definition relative to the aggregate $A^\dagger_T[\mathbb{B}(0,200M)]$. Since $A^\dagger_T[\mathbb{B}(0,200M)]\subset A^\dagger_T[\infty]$, we have that, for any $z\in\chi_\varepsilon\cap\mathbb{B}(0,100M)$, $R^{\text{loc}}_{T,0,10M}(z;\varepsilon)\leq R_T(z;\varepsilon)\leq M$ on the event $F(M)$. This means that all the balls of $\mathcal{C}$ have radii smaller than M. We can then extract from $\mathcal{C}$ a sub-cluster, say $\mathcal{C}'$, joining $\mathbb{B}(Mc,M)$ for some (random) $c\in\mathbb{S}_{10}$ to $\mathbb{B}(Mc,8M)^c$. Indeed, the sphere $\mathbb{S}_{10M}$ is covered by the union of balls $\mathbb{B}(Mc,M)$, with $c\in\mathbb{S}_{10}$. Now, we have to prove that $\mathcal{C}'$ is also a cluster of $\hat{\Sigma}^{\text{loc}}_\varepsilon(Mc,M)$, ensuring the occurrence of $G_\varepsilon(Mc,M)$. Hence, we have to prove that each ball $\mathbb{B}(z,R^{\text{loc}}_{T,0,10M}(z;\varepsilon))$ involved in $\mathcal{C}'$ satisfies the two following properties:

- Its center z belongs to $\chi_\varepsilon\cap\mathbb{B}(Mc,10M)$. This is clear since, by construction, each ball of $\mathcal{C}'$ overlaps $\mathbb{B}(Mc,8M)$ and has a radius smaller than M.
- Its radius $R^{\text{loc}}_{T,0,10M}(z;\varepsilon)$ is actually equal to $R^{\text{loc}}_{T,Mc,M}(z;\varepsilon)$. This is where the event $\mathbf{Stab}(0,100M)$ comes into play. The previous item implies that the ball $\mathbb{B}(z,R^{\text{loc}}_{T,0,10M}(z;\varepsilon))$ is completely included in $\mathbb{B}(Mc,10M)$. Its radius is defined w.r.t. the aggregate $A^\dagger_T[\mathbb{B}(0,200M)]$ (see (25)), but only through

$$A^\dagger_T[\mathbb{B}(0,200M)]\cap\Big(\mathbb{Z}\times\mathbb{B}(Mc,10M)\Big)$$

since $\mathbb{B}(z, R^{\text{loc}}_{T,0,10M}(z;\varepsilon)) \subset \mathbb{B}(Mc, 10M)$. Thanks to $\mathbf{Stab}(0, 100M)$, in particular $\mathbf{Stab}^M(cM)$ applied with $S := \mathbb{B}(0, 200M) \supset \mathbb{B}(Mc, 20M)$, we have

$$A^\dagger_T[\mathbb{B}(0,200M)] \cap \Big(\mathbb{Z} \times \mathbb{B}(Mc,10M)\Big) = A^\dagger_T[\mathbb{B}(Mc,20M)] \cap \Big(\mathbb{Z} \times \mathbb{B}(Mc,10M)\Big) . \quad (31)$$

Since the radius $R^{\text{loc}}_{T,Mc,M}(z;\varepsilon)$ is defined w.r.t. the aggregate $A^\dagger_T[\mathbb{B}(Mc, 20M)]$, the identity (31) implies that $R^{\text{loc}}_{T,Mc,M}(z;\varepsilon)$ and $R^{\text{loc}}_{T,0,10M}(z;\varepsilon)$ are equal.

Thus, we have proven that the event $G_\varepsilon(Mc, M)$ holds for some $c \in \mathbb{S}_{10}$. We can show in a similar fashion that $G_\varepsilon(Mc', M)$ also occurs for some $c' \in \mathbb{S}_{80}$. Inclusion (30) is established. □

Proof of Lemma 3.17. Let $M \geq 1$. Note that the occurrence of the event $G_\varepsilon(0, M)$ forces the random set $\chi_\varepsilon \cap \mathbb{B}(0, 10M)$ to be non-empty. Therefore,

$$\begin{aligned} \mathbb{P}(G_\varepsilon(0,M)) &\leq \mathbb{P}(\#(\mathbb{B}(0,10M) \cap \chi_\varepsilon) \geq 1) \\ &\leq \mathbb{E}[\#(\mathbb{B}(0,10M) \cap \chi_\varepsilon)] \\ &= \sum_{z \in \mathbb{B}(0,10M)} \mathbb{E}[\mathbb{1}_{z \in \chi_\varepsilon}] \\ &= \#\mathbb{B}(0,10M) p_\varepsilon . \end{aligned}$$

Using $p_\varepsilon = 1 - e^{-\varepsilon} \leq \varepsilon$ and $\#\mathbb{B}(0, 10M) \leq C_d M^{d-1}$, we get the desired result. □

5 Proof of Theorem 3.2

Before giving the proof of Theorem 3.2, we give the following lemma:

Lemma 3.18. *For any finite subset $S \subset \mathbb{Z}^d$,*

$$\mathbb{P}\left(S \subset A^\dagger_n[\infty]\right) \underset{n\to\infty}{\longrightarrow} 1.$$

Proof of Lemma 3.18. To do so, we will be using the Shape Theorem for standard IDLA, in the case where exactly n particles are sent from the origin. Let us denote this aggregate by $A(n)$. We know from Theorem 1 of [34] that

$$\mathbb{P}(S \subset A(n)) \underset{n\to\infty}{\longrightarrow} 1. \quad (32)$$

Now, recall that particles of $A^\dagger_n[\infty]$ are given according to a family of PPP's in $\mathbb{R}_+$, denoted by $(\mathcal{N}_z)_{z\in\mathcal{H}}$. Let $N_0 = \mathcal{N}_0([0, n])$ denote the number of particles sent from the origin. Since N_0 is a Poisson random variable of parameter n, we know from concentration inequality theory that for all $0 < \varepsilon < 1/2$,

$$\mathbb{P}\left(N_0 - n \leq n^{1/2+\varepsilon}\right) \leq \exp\left(-\frac{n^{2\varepsilon}}{2}\right).$$

Hence, define $\mathscr{E}_n := \left\{N_0 \le n - n^{1/2+\varepsilon}\right\}$. Let S denote a finite subset of $\mathbb{Z}^d$. We write:

$$\begin{aligned}\mathbb{P}\left(S \subset A(n - n^{1/2+\varepsilon})\right) &\le \mathbb{P}\left(\left\{S \subset A(n - n^{1/2+\varepsilon})\right\} \cap \mathscr{E}_n{}^c\right) + \mathbb{P}\left(\mathscr{E}_n\right) \\ &\le \mathbb{P}\left(S \subset A_n^\dagger[0]\right) + \exp\left(-\frac{n^{2\varepsilon}}{2}\right) \\ &\le \mathbb{P}\left(S \subset A_n^\dagger[\infty]\right) + \exp\left(-\frac{n^{2\varepsilon}}{2}\right).\end{aligned}$$

Hence, for any finite subset $S \subset \mathbb{Z}^d$, we have that

$$\mathbb{P}\left(S \subset A_n^\dagger[\infty]\right) \ge \mathbb{P}\left(S \subset A(n - n^{1/2+\varepsilon})\right) - \exp\left(-\frac{n^{2\varepsilon}}{2}\right).$$

Using (32), we have the desired result. □

Proof of Theorem 3.2. We begin by showing 1. For any $n \ge 0$, $\mathcal{F}_n \subset \mathcal{F}_\infty$, so $V\left(\mathcal{F}_n\right) \subset V\left(\mathcal{F}_\infty\right)$. Now, consider S a finite subset of $\mathbb{Z}^d$. Since the vertex set of $\mathcal{F}_n$ is $A_n^\dagger[\infty]$, we have that for any $n \ge 0$,

$$\mathbb{P}\left(S \subset V\left(\mathcal{F}_\infty\right)\right) \ge \mathbb{P}\left(S \subset V\left(\mathcal{F}_n\right)\right) = \mathbb{P}\left(S \subset A_n^\dagger[\infty]\right).$$

This result is immediate using Lemma 3.18.
We move to the proof of 2. Consider a compact K of $\mathbb{R}^d$. Fix $k \in \mathcal{H}$ and $n \ge 0$. It is sufficient to show that $\mathcal{F}_n$ and $\mathcal{F}_\infty$ have the same probability of intersecting K. Assume that this holds for $\mathcal{F}_n$, that is:

$$\mathbb{P}\left(\mathcal{F}_n \cap K \neq \emptyset\right) = \mathbb{P}\left(T_k\mathcal{F}_n \cap K \neq \emptyset\right). \tag{33}$$

Note that for any subset C such that $(C + \mathbb{B}(0,1)) \cap \mathbb{Z}^d \subset A_n^\dagger[\infty]$, then $\mathcal{F}_n \cap C = \mathcal{F}_\infty \cap C$. Again using Lemma 3.18, we have that

$$\mathbb{P}\left((K + \mathbb{B}(0,1)) \cap \mathbb{Z}^d \subset A_n^\dagger[\infty]\right) \underset{n\to\infty}{\longrightarrow} 1.$$

Thus, there exists an integer n_0 such that $\mathbb{P}\left((K + \mathbb{B}(0,1)) \cap \mathbb{Z}^d \subset A_{n_0}^\dagger[\infty]\right) \ge 1 - \varepsilon$. We have:

$$\begin{aligned}\left|\mathbb{P}\left(\mathcal{F}_\infty \cap K \neq \emptyset\right) - \mathbb{P}\left(\mathcal{F}_{n_0} \cap K \neq \emptyset\right)\right| &\le \mathbb{P}\left(\mathcal{F}_\infty \cap K \neq \mathcal{F}_{n_0} \cap K\right) \\ &\le \mathbb{P}\left((K + \mathbb{B}(0,1)) \cap \mathbb{Z}^d \not\subset A_{n_0}^\dagger[\infty]\right) \le \varepsilon.\end{aligned} \tag{34}$$

Similarly, we can show that

$$\left|\mathbb{P}\left(T_k\mathcal{F}_\infty \cap K \neq \emptyset\right) - \mathbb{P}\left(T_k\mathcal{F}_{n_0} \cap K \neq \emptyset\right)\right| \le \varepsilon. \tag{35}$$

We can now conclude for $\mathcal{F}_\infty$, since

$$\begin{aligned}&\left|\mathbb{P}\left(\mathcal{F}_\infty \cap K \neq \emptyset\right) - \mathbb{P}\left(T_k\mathcal{F}_\infty \cap K \neq \emptyset\right)\right| \le \\ &\qquad \left|\mathbb{P}\left(\mathcal{F}_\infty \cap K \neq \emptyset\right) - \mathbb{P}\left(\mathcal{F}_{n_0} \cap K \neq \emptyset\right)\right| + \left|\mathbb{P}\left(\mathcal{F}_{n_0} \cap K \neq \emptyset\right) - \mathbb{P}\left(T_k\mathcal{F}_{n_0} \cap K \neq \emptyset\right)\right| \\ &\qquad + \left|\mathbb{P}\left(T_k\mathcal{F}_{n_0} \cap K \neq \emptyset\right) - \mathbb{P}\left(T_k\mathcal{F}_\infty \cap K \neq \emptyset\right)\right|.\end{aligned}$$

Now, from (33), we know that the middle term is equal to 0, and from (34) and (35), we get that the first and third term are bounded by ε. Thus,

$$\left|\mathbb{P}\left(\mathcal{F}_\infty \cap K \neq \emptyset\right) - \mathbb{P}\left(T_k\mathcal{F}_\infty \cap K \neq \emptyset\right)\right| \le 2\varepsilon.$$

It now remains to show (33). Fix $n \geq 0$. Take $M \geq 1$ sufficiently large such that $K \cap \mathbb{Z}^d \subset \mathbb{Z}_M$. From Theorem 3.1, there exists a random integer N_0 such that for any $N' \geq N_0$, with probability greater than $1-\varepsilon$, we have

$$\mathcal{F}_n[N'] \cap \mathbb{Z}_M = \mathcal{F}_n \cap \mathbb{Z}_M, \quad T_k\mathcal{F}_n[N'] \cap \mathbb{Z}_M = T_k\mathcal{F}_n \cap \mathbb{Z}_M. \tag{36}$$

Then, we have:

$$\begin{aligned}
&|\mathbb{P}(\mathcal{F}_n \cap K \neq \emptyset) - \mathbb{P}(T_k\mathcal{F}_n \cap K \neq \emptyset)| \leq \\
&\quad |\mathbb{P}(\mathcal{F}_n \cap K \neq \emptyset) - \mathbb{P}(\mathcal{F}_n[N_0] \cap K \neq \emptyset)| + |\mathbb{P}(\mathcal{F}_n[N_0] \cap K \neq \emptyset) - \mathbb{P}(T_k\mathcal{F}_n[N_0] \cap K \neq \emptyset)| + \\
&\quad |\mathbb{P}(T_k\mathcal{F}_n[N_0] \cap K \neq \emptyset) - \mathbb{P}(T_k\mathcal{F}_n \cap K \neq \emptyset)| \\
&\leq |\mathbb{P}(\mathcal{F}_n[N_0] \cap K \neq \emptyset) - \mathbb{P}(T_k\mathcal{F}_n[N_0] \cap K \neq \emptyset)| + 2\varepsilon.
\end{aligned}$$

Now, we grow the forests $\mathcal{F}_n[N_0]$ and $T_k\mathcal{F}_n[N_0]$ to obtain two forests $\mathfrak{F}_1$ and $\mathfrak{F}_2$. We obtain $\mathfrak{F}_1$ by sending the particles used to build $\mathcal{F}_n[N_0]$, and add the additional particles from $(T_k\mathcal{H}_{N_0}) \cap {\mathcal{H}_{N_0}}^c$. To build $\mathfrak{F}_2$, consider the forest induced by particles used to build $\mathcal{F}_n[N_0]$ and additional particles from $(T_{-k}\mathcal{H}_{N_0}) \cap {\mathcal{H}_{N_0}}^c$. Let $\mathfrak{F}$ denote this forest. We define $\mathfrak{F}_2$ as the translation of vector k of this forest, that is $\mathfrak{F}_2 := T_k\mathfrak{F}$. Now, from (36), we know that $\mathfrak{F}_1$ and $\mathcal{F}_n[N_0]$ coincide on the strip $\mathbb{Z}_M$ (and hence on K) with probability greater than $1-\varepsilon$. The same is true for $\mathfrak{F}_2$ and $T_k\mathcal{F}_n[N_0]$. Therefore, we have

$$|\mathbb{P}(\mathcal{F}_n \cap K \neq \emptyset) - \mathbb{P}(T_k\mathcal{F}_n \cap K \neq \emptyset)| \leq |\mathbb{P}(\mathfrak{F}_1 \cap K \neq \emptyset) - \mathbb{P}(\mathfrak{F}_2 \cap K \neq \emptyset)| + 4\varepsilon.$$

Lemma 3.19. *The set of sources used to build $\mathfrak{F}_1$ and $\mathfrak{F}_2$ are identical.*

Now, the forest $\mathfrak{F}_1$ and $\mathfrak{F}_2$ are built using the same IDLA protocol, using the same set of sources, with i.i.d Poisson clocks over the time interval $[0,n]$. They therefore have same distribution, which implies that

$$|\mathbb{P}(\mathcal{F}_n \cap K \neq \emptyset) - \mathbb{P}(T_k\mathcal{F}_n \cap K \neq \emptyset)| \leq 4\varepsilon,$$

which concludes the proof.

We now prove 3. We show that for any compacts $C_1,\ C_2$ of $\mathbb{R}^2$

$$\lim_{k\in\mathcal{H},\ \|k\|\to\infty} \mathbb{P}(\mathcal{F}_\infty \cap (C_1 \cup T_kC_2) = \emptyset) = \mathbb{P}(\mathcal{F}_\infty \cap C_1 = \emptyset)\,\mathbb{P}(\mathcal{F}_\infty \cap C_2 = \emptyset). \tag{37}$$

Fix $\varepsilon > 0$ and let $C_1,\ C_2$ be two compact sets of $\mathbb{R}^2$. Let $r > 0$ be such that $C_1 \cup C_2$ is included in $\mathbb{B}(0, r-1)$. From Lemma 3.18, we can pick n large enough such that

$$\mathbb{P}\left(\mathbb{B}(0,r) \cap \mathbb{Z}^2 \subset A_n^\dagger[\infty]\right) \geq 1-\varepsilon.$$

On the event $\left\{\mathbb{B}(0,r) \cap \mathbb{Z}^2 \subset A_n^\dagger[\infty]\right\}$, $\mathcal{F}_\infty \cap C_i$ and $\mathcal{F}_n \cap C_i$ are equal for any $i \in \{1,2\}$. Since the distribution of $A_n^\dagger[\infty]$ is invariant with respect to T_k, we have that for any $k \in \mathcal{H}$ (independent of ε),

$$\mathbb{P}\left(T_k(C_1 \cup C_2) \subset A_n^\dagger[\infty]\right) \geq 1-\varepsilon.$$

This implies:

$$\mathbb{P}(\mathcal{F}_\infty \cap (C_1 \cup T_kC_2) = \mathcal{F}_n \cap (C_1 \cup T_kC_2)) \geq 1-2\varepsilon.$$

Therefore,

$$\begin{aligned}&\left|\mathbb{P}\left(\mathcal{F}_\infty \cap (C_1 \cup T_k C_2) = \emptyset\right) - \mathbb{P}\left(\mathcal{F}_\infty \cap C_1 = \emptyset\right) \mathbb{P}\left(\mathcal{F}_\infty \cap C_2 = \emptyset\right)\right| \\ &\qquad \leq \left|\mathbb{P}\left(\mathcal{F}_n \cap (C_1 \cup T_k C_2) = \emptyset\right) - \mathbb{P}\left(\mathcal{F}_n \cap C_1 = \emptyset\right) \mathbb{P}\left(\mathcal{F}_n \cap C_2 = \emptyset\right)\right| + 4\varepsilon.\end{aligned}$$

It remains to show that $\mathcal{F}_n$ is mixing with respect to T_k. From Theorem 3.1, there exists N_0 such that

$$\mathbb{P}\left(\mathcal{F}_n \cap C_1 \neq \mathcal{F}_n[N_0] \cap C_1\right) \leq \varepsilon \text{ and } \mathbb{P}\left(\mathcal{F}_n \cap C_2 \neq \mathcal{F}_n[N_0] \cap C_2\right) \leq \varepsilon. \tag{38}$$

We have, by shifting all the random clocks by a vector $-k \in \mathcal{H}$:

$$\begin{aligned}\mathbb{P}\left(\mathcal{F}_n \cap C_2 \neq \mathcal{F}_n[N_0] \cap C_2\right) &= \mathbb{P}\left(\mathcal{F}_n(\omega) \cap C_2 \neq \mathcal{F}_n[N_0](\omega) \cap C_2\right) \\ &= \mathbb{P}\left(\mathcal{F}_n(\omega - k) \cap C_2 \neq \mathcal{F}_n[N_0](\omega - k) \cap C_2\right) \\ &= \mathbb{P}\left(\mathcal{F}_n(\omega) \cap T_k C_2 \neq \mathcal{F}_n[\mathbb{B}(k, N_0)](\omega) \cap T_k C_2\right).\end{aligned}$$

Hence, from (38), for any $k \in \mathcal{H}$ (independent of ε), there exists N_0 such that

$$\mathbb{P}\left(\mathcal{F}_n(\omega) \cap T_k C_2 \neq \mathcal{F}_n[\mathbb{B}(k, N_0)](\omega) \cap T_k C_2\right) \leq \varepsilon. \tag{39}$$

Now, we write:

$$\begin{aligned}&\left|\mathbb{P}\left(\mathcal{F}_n \cap (C_1 \cup T_k C_2) = \emptyset\right) - \mathbb{P}\left(\mathcal{F}_n \cap C_1 = \emptyset\right) \mathbb{P}\left(\mathcal{F}_n \cap C_2 = \emptyset\right)\right| \\ &= \left|\mathbb{P}\left(\{\mathcal{F}_n \cap C_1 = \emptyset\} \cap \{\mathcal{F}_n \cap T_k C_2 = \emptyset\}\right) - \mathbb{P}\left(\mathcal{F}_n \cap C_1 = \emptyset\right) \mathbb{P}\left(\mathcal{F}_n \cap C_2 = \emptyset\right)\right| \\ &\leq \left|\mathbb{P}\left(\{\mathcal{F}_n \cap C_1 = \emptyset\} \cap \{\mathcal{F}_n \cap T_k C_2 = \emptyset\}\right) - \mathbb{P}\left(\{\mathcal{F}_n[N_0] \cap C_1 = \emptyset\} \cap \{\mathcal{F}_n[\mathbb{B}(k, N_0)] \cap T_k C_2 = \emptyset\}\right)\right| \\ &+ \left|\mathbb{P}\left(\{\mathcal{F}_n[N_0] \cap C_1 = \emptyset\} \cap \{\mathcal{F}_n[\mathbb{B}(k, N_0)] \cap T_k C_2 = \emptyset\}\right) - \mathbb{P}\left(\mathcal{F}_n \cap C_1 = \emptyset\right) \mathbb{P}\left(\mathcal{F}_n \cap C_2 = \emptyset\right)\right|.\end{aligned}$$

Lemma 3.20. *Let $A, A', B,$ and B' denote 4 events. Suppose $\mathbb{P}(A \neq A') \leq \varepsilon$ and $\mathbb{P}(B \neq B') \leq \varepsilon$. Then*

$$\left|\mathbb{P}\left(\{A = \emptyset\} \cap \{B = \emptyset\}\right) - \mathbb{P}\left(\{A' = \emptyset\} \cap \{B' = \emptyset\}\right)\right| \leq 2\varepsilon.$$

To alleviate notation, we define the following sets:

$$\begin{cases} A = \mathcal{F}_n \cap C_1, \\ B = \mathcal{F}_n \cap T_k C_2, \end{cases} \qquad \text{and} \qquad \begin{cases} A' = \mathcal{F}_n[N_0] \cap C_1, \\ B' = \mathcal{F}_n[\mathbb{B}(k, N_0)] \cap T_k C_2. \end{cases}$$

We can rewrite the previous inequality as:

$$\begin{aligned}&\left|\mathbb{P}\left(\mathcal{F}_n \cap (C_1 \cup T_k C_2) = \emptyset\right) - \mathbb{P}\left(\mathcal{F}_n \cap C_1 = \emptyset\right) \mathbb{P}\left(\mathcal{F}_n \cap C_2 = \emptyset\right)\right| \\ &\qquad \leq \left|\mathbb{P}\left(\{A = \emptyset\} \cap \{B = \emptyset\}\right) - \mathbb{P}\left(\{A' = \emptyset\} \cap \{B' = \emptyset\}\right)\right| \\ &\qquad + \left|\mathbb{P}\left(\{A' = \emptyset\} \cap \{B' = \emptyset\}\right) - \mathbb{P}(A = \emptyset)\, \mathbb{P}(B = \emptyset)\right|.\end{aligned}$$

Note that for the last term, we replaced $\mathbb{P}(\mathcal{F}_n \cap C_2 = \emptyset)$ by $\mathbb{P}(B = \emptyset)$. We can do so since the distribution of $\mathcal{F}_n$ is invariant with respect to T_k, so $\mathcal{F}_n$ has same probability of intersecting C_2 or $T_k C_2$.

Notice that for any k such that $\|k\| > 2N_0$, the events A' and B' are independent, since the two forests considered are built using disjoint sets of sources. Additionally, from (34) and (35), we have that $\mathbb{P}(A \neq A') \leq \varepsilon$ and $\mathbb{P}(B \neq B') \leq \varepsilon$. Therefore, using the result of Lemma 3.20, we get, for any $\|k\| > 2N_0$:

$$|\mathbb{P}(\mathcal{F}_n \cap (C_1 \cup T_k C_2) = \emptyset) - \mathbb{P}(\mathcal{F}_n \cap C_1 = \emptyset)\mathbb{P}(\mathcal{F}_n \cap C_2 = \emptyset)| \\ \leq |\mathbb{P}(A' = \emptyset)\mathbb{P}(B' = \emptyset) - \mathbb{P}(A = \emptyset)\mathbb{P}(B = \emptyset)| + 2\varepsilon.$$

Now, since $\mathbb{P}(A \neq A') \leq \varepsilon$ and $\mathbb{P}(B \neq B') \leq \varepsilon$, one can show (in the same spirit of Lemma 3.20) that

$$|\mathbb{P}(A' = \emptyset)\mathbb{P}(B' = \emptyset) - \mathbb{P}(A = \emptyset)\mathbb{P}(B = \emptyset)| \leq 2\varepsilon.$$

Therefore,

$$|\mathbb{P}(\mathcal{F}_\infty \cap (C_1 \cup T_k C_2) = \emptyset) - \mathbb{P}(\mathcal{F}_\infty \cap C_1 = \emptyset)\mathbb{P}(\mathcal{F}_\infty \cap C_2 = \emptyset)| \leq 8\varepsilon,$$

which concludes the proof. □

6 Appendix: Proof of Proposition 3.5

Before proving Proposition 3.5, we must first introduce a new family of random aggregates. Let $n \geq 1$. Just like for $A_n^\dagger[\infty]$, we begin by building a family of *finite* random aggregates $(A_n^*[M])_{M\geq 0}$. Like $A_n^\dagger[\infty]$, the number of particles sent from each source z is random, given this time by a Poisson random variable N_z of parameter n, but unlike $A_n^\dagger[\infty]$, these are sent in a predetermined order. Let $(N_z)_{z\in\mathcal{H}}$ denote a family of i.i.d Poisson random variables of parameter n. When $M = 0$, $A_n^*[0]$ is the aggregate obtained after launching N_0 particles from the origin. Given a realization of $A_n[M-1]$, we throw N_z particles from each source z of level M according to the lexicographical order. So $A_n^*[M]$ is defined as the aggregate produced by $A_n^*[M-1]$ and the new sites added by particles launched at level M. By construction, $(A_n^*[M])_{M\geq 0}$ is increasing with respect to inclusion, so we can define $A_n^*[\infty]$ as

$$A_n^*[\infty] :=\uparrow \bigcup_{M\geq 0} A_n^*[M] \quad \text{a.s.}$$

A consequence of the Abelian Property (see [19], p. 97) is that for all $M \geq 0$,

$$A_n^\dagger[M] \overset{\text{law}}{=} A_n^*[M]. \tag{40}$$

Since both families of aggregates $(A_n^*[M])_{M\geq 0}$ and $(A_n^\dagger[M])_{M\geq 0}$ are increasing, we deduce from (40) that $A_n^\dagger[\infty] \overset{\text{law}}{=} A_n^*[\infty]$.

The global upper bound provided by Proposition 3.5 is a refined version of Theorem 4.1 of [15]. Although our result is finer and deals with a random number of emitted particles, the strategy of the proof is essentially the same.

We will be proving Proposition 3.5 by induction over n. Since $A_n^\dagger[\infty] \overset{\text{law}}{=} A_n^*[\infty]$, we will show the result for $A_n^*[\infty]$ instead, since this aggregate is built by sending particles in the *usual order.* We define the event $\mathbf{Over}_\alpha^*$ in the same way as $\mathbf{Over}_\alpha^\dagger$ in (7) but with respect to the aggregate $A_n^*[\infty]$. We show that if for some fixed n, the aggregate $A_n^*[\infty]$ is contained within $\mathscr{C}_\varepsilon^\alpha$ for some $\varepsilon > 0$, and if we launch N_z' additional particles from each source z of $\mathcal{H}$, where $(N_z')_{z\in\mathcal{H}}$ is an independent family of Poisson variables of parameter 1, then the resulting aggregate is contained within a slightly larger cone $\mathscr{C}_{\varepsilon'}^\alpha$ $(\varepsilon' > \varepsilon)$ with high probability. We keep the same notations as in the proof of Theorem 4.1 of [15] and define the sequences $(M_n)_{n\geq 0}$ and $(\varepsilon_n)_{n\geq 0}$ in the same way. Just like for Theorem 4.1 of [15], it is sufficient to prove the following proposition:

Proposition 3.21. *For all $\alpha \in (1-1/d, 1)$, for all $\varepsilon > 0$, for all $n \geq 1$, there exists a constant $C = C(\varepsilon, n, \alpha, d) > 0$ such that for all $M \geq 1$ and all $L > 1$,*

$$\mathbb{P}\left(\mathbf{Over}^*_\alpha(M_n, \varepsilon_n)\right) \leq \frac{C}{M^L}.$$

Proof of Proposition 3.21: We show our result by induction over n. Take $L > 1$. Our induction statement is the following:
$\forall n \geq 0,\ \mathcal{P}(n) : \forall \alpha \in (1-1/d, 1),\ \forall \varepsilon \in (0,1),\ \forall M \geq 1,\ \exists C = C(\varepsilon, n, \alpha, d) > 0,$

$$\mathbb{P}\left(\mathbf{Over}^*_\alpha(M_n, \varepsilon_n)\right) \leq \frac{C}{M^L}.$$

When $n = 0$, this is clear since $A^*_n[\infty] \overset{\text{a.s}}{=} \emptyset$, hence $A^*_n[\infty] \cap \mathbb{Z}^c_M \overset{\text{a.s}}{\subset} \mathscr{C}^\alpha_\varepsilon$.
Let $n \geq 1$ and suppose $\mathcal{P}(n)$ holds. Fix $\alpha \in (1-1/d, 1)$. We write:

$$\mathbb{P}\left(\mathbf{Over}^*_\alpha(M_{n+1}, \varepsilon_{n+1})\right) \leq \mathbb{P}\left(\mathbf{Over}^*_\alpha(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}^*_\alpha(M_n, \varepsilon_n)^c\right) + \mathbb{P}\left(\mathbf{Over}^*_\alpha(M_n, \varepsilon_n)\right).$$

The last term is handled by our induction hypothesis. We switch our focus to the first term. On the event $\mathbf{Over}^*_\alpha(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}^*_\alpha(M_n, \varepsilon_n)^c$, we have $A^*_n[\infty] \cap \mathbb{Z}^c_{M_n} \subset \mathscr{C}^\alpha_{\varepsilon_n}$, but when launching N'_z new particles from each source $z \in \mathcal{H}$, the new aggregate obtained spills over $\mathscr{C}^\alpha_{\varepsilon_{n+1}}$ on $\mathbb{Z}^c_{M_{n+1}}$. This implies the existence of three random sites $(Z, Z^*, Z_{n+1}) \in \mathbb{Z}^d$, and aggregates A_{Z^*} and $A^-_{Z^*}$, defined just like in [15], but defined with respect to $\mathscr{C}^\alpha_\varepsilon$ rather than $\mathscr{C}_\varepsilon$. Note now that we have : $Z_{n+1} = Z \pm (\varepsilon_{n+1}\|Z\|^\alpha) \cdot e_1$, where $e_1 = (1, 0, \ldots, 0)$.

We must control that no unreasonable amount of particles is emitted from $\mathcal{H}$. To do so, we introduce the following event:

$$\mathscr{E}_M(\gamma) := \bigcap_{l \geq 0} \left\{\forall z \in \mathcal{H},\ \|z\| = l,\ N'_z \leq 1 + \max\{l, M\}^\gamma\right\},$$

We can show using (11) that $\mathbb{P}\left(\mathscr{E}^c_M(\gamma)\right)$ decreases faster than any power of M^{-1}. Fix $\gamma \in (0,1)$ such that $\gamma < (\alpha - 1)d + 1$ (such a value exists since $\alpha \in (1-1/d, 1)$). We explain this choice later. We write

$$\begin{aligned}
&\mathbb{P}\left(\mathbf{Over}^*_\alpha(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}^*_\alpha(M_n, \varepsilon_n)^c\right) \\
&\quad \leq \mathbb{P}\left(\mathbf{Over}^*_\alpha(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}^*_\alpha(M_n, \varepsilon_n)^c \cap \mathscr{E}_M(\gamma)\right) + \mathbb{P}\left(\mathscr{E}^c_M(\gamma)\right) \\
&\quad \leq \sum_{l \geq M_{n+1}} \sum_{\|z\| = l} \mathbb{P}\left(Z = z,\ \mathbf{Over}^*_\alpha(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}^*_\alpha(M_n, \varepsilon_n)^c \cap \mathscr{E}_M(\gamma)\right) + \mathcal{O}(1).
\end{aligned}$$

Fix $l \geq M_{n+1}$ and $z \in \mathcal{H}$ such that $\|z\| = l$, and let $z_{n+1} = z \pm (\varepsilon_{n+1}\|z\|^\alpha) \cdot e_1$. We consider the case where a ball of particles has settled around z_{n+1}, and the case where a thin tentacle reaches out to z_{n+1}. We use an adaptation of Lemma 2 of [31] to deal with the event of tentacles.

Lemma 3.22. *There exist positive universal constants $b,\ K_0,\ c$ such that for all real numbers $r > 0$ and all $z \in \mathcal{H}$ with $0 \notin \mathbb{B}(z_{n+1}, r)$,*

$$\mathbb{P}\left(Z = z,\ \mathbf{Over}^*_\alpha(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}^*_\alpha(M_n, \varepsilon_n)^c,\ \#\left(A_{Z^*} \cap \mathbb{B}(z_{n+1}, r)\right) \leq br^d\right) \leq K_0 e^{-cr^2}.$$

We apply this Lemma with $r = r_{n+1} = \frac{\varepsilon l^\alpha}{2^{n+1}}$, in the same spirit as in [15]. This gives:

$$\begin{aligned}
&\mathbb{P}\left(Z = z,\ \mathbf{Over}^*_\alpha(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}^*_\alpha(M_n, \varepsilon_n)^c \cap \mathscr{E}_M(\gamma)\right) \\
\leq &\mathbb{P}\left(Z = z,\ \mathbf{Over}^*_\alpha(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}^*_\alpha(M_n, \varepsilon_n)^c,\right. \\
&\qquad \left.\mathscr{E}_M(\gamma),\ \#\left(A_{Z^*} \cap \mathbb{B}(z_{n+1}, r_{n+1})\right) > b r^d_{n+1}\right) + K_0 e^{-c_1 l^{2\alpha}}, \qquad (41)
\end{aligned}$$

where $c_1 = c_1(n, \varepsilon) = \frac{c\varepsilon^2}{4^{n+1}}$. The term (41) requires a little more work than in the deterministic global upper bound. Note that r^d_{n+1} is of order $l^{\alpha d}$. We are working on the event where (roughly) more than $l^{\alpha d}$ particles have settled around z_{n+1}, knowing that each source $\tilde{z} \in \mathcal{H}$ has emitted at most $1 + \|\tilde{z}\|^\gamma$ particles. We show that this implies that $\|Z - Z^*\| \geq K l^\eta$, where $\eta > 1$ and K denotes some positive constant. Let us now explain why $\eta > 1$. Suppose, contrary to our claim, that $\eta \leq 1$. The number of sources inside $\mathbb{B}\left(z, K l^\eta\right) \cap \mathcal{H}$ is of the order $l^{\eta(d-1)}$ since $\mathcal{H}$ is a hyperplane of dimension $d-1$. Working on the event $\mathscr{E}_M$, the largest amount of particles emitted by a single source within this $(d-1)$-dimensional ball is $1 + (l + K l^\eta)^\gamma$, which is of the order l^γ when $\eta \leq 1$. In the worst case scenario, if all of the sources within $\mathbb{B}\left(z, K l^\eta\right) \cap \mathcal{H}$ emit of the order of l^γ particles, the total number of particles emitted will be of the order of $l^{\eta(d-1)} \times l^\gamma = l^{\eta(d-1)+\gamma}$. Now, for this to be of the order (or greater) than $l^{\alpha d}$, it is thus necessary that

$$\eta(d-1) + \gamma \geq \alpha d \Longleftrightarrow \eta \geq \frac{\alpha d - \gamma}{d-1}.$$

However, due to our choice of γ, this necessarily implies that $\eta > 1$, which contradicts our assumption that $\eta \leq 1$. Therefore, in order for *more* than $l^{\alpha d}$ particles to settle inside $\mathbb{B}\left(z_{n+1}, r_{n+1}\right)$, it is necessary that one of these particles has been emitted from a source outside of $\mathbb{B}\left(z, K l^\eta\right) \cap \mathcal{H}$, with $\eta > 1$. It is thus necessary that $\|Z - Z^*\| \geq K l^\eta$. Since Z^* is defined as the source from which the *first* overflowing particle is emitted, this implies that the aggregate before sending from Z^* is strictly contained within $\mathscr{C}^\alpha_{n+1}$, allowing us to use a donut argument (similar to the one used in the proof of Lemma 3.15) to control the trajectory of the overflowing particle. This implies that one of the particles sent from z', with $\|z' - z\| \geq K l^\eta$, has crossed multiple donuts. We detail below how we compute a lower bound on the total number of donuts a particle needs to cross from z' to z. Fix $h \geq K l^\eta$ and z' such that $\|z' - z\| = h$. We build donuts from level $\|z'\|$ to $\|z\| = l$. The first donut, which is the *largest* donut, has dimensions at most $2\varepsilon_{n+1}(l+h)^\alpha \leq 2\varepsilon_{n+1}(2h)^\alpha \leq 4\varepsilon_{n+1} h^\alpha$. All other donuts will have smaller dimensions, which implies that the number of donuts $k = k(h, l, \varepsilon_{n+1}, \alpha)$ between z' and z is such that

$$k \geq \frac{h}{4\varepsilon_{n+1} h^\alpha} = \frac{h^{1-\alpha}}{4\varepsilon_{n+1}}.$$

Now, the number of particles sent from z' at distance h of z is (up to a multiplicative factor) at most $h^{d-2+\gamma}$. Therefore, the donut argument gives:

$$\begin{aligned}
&\sum_{h \geq K l^\eta} \sum_{\|z'-z\|=h} \mathbb{P}\left(\left\{\begin{array}{c} \text{a particle sent from } z' \text{ reaches level } l \\ \text{while staying within } \mathscr{C}^\alpha_{\varepsilon_{n+1}} \end{array}\right\} \cap \mathscr{E}_M(\gamma)\right) \\
\leq &K_d \sum_{h \geq K l^\eta} h^{d-2+\gamma} (1-c)^k = K_d \sum_{h \geq K_d l^\eta} h^{d-2+\gamma} \exp\left(-c_0\left(\varepsilon_{n+1}\right) h^{1-\alpha}\right),
\end{aligned}$$

where $c_0\left(\varepsilon_{n+1}\right) = -\frac{\log(1-c)}{4\varepsilon_{n+1}}$. Throughout the rest of the proof, K_d will denote a generic constant depending only on d, whose value may vary from line to line. Now, using the fact that $1-\alpha > 0$, standard computations yield:

$$K_d \sum_{h \geq K l^\eta} h^{d-2+\gamma} \exp\left(-c_0\left(\varepsilon_{n+1}\right) h^{1-\alpha}\right) \leq K_d \exp\left(-\frac{1}{2} c_0(\varepsilon_{n+1}) l^{\eta(1-\alpha)}\right).$$

Combining this result with (41), we get

$$\begin{aligned}
&\mathbb{P}\left(\mathbf{Over}^*_\alpha(M_{n+1}, \varepsilon_{n+1}) \cap \mathbf{Over}^*_\alpha(M_n, \varepsilon_n)^c \cap \mathscr{E}_M(\gamma)\right) \\
&\qquad \leq \sum_{l \geq M_{n+1}} \sum_{\|z\|=l} K_0 e^{-c_1 l^{2\alpha}} + \sum_{l \geq M_{n+1}} \sum_{\|z\|=l} K_d \exp\left(-\frac{1}{2} c_0(\varepsilon_{n+1}) l^{\eta(1-\alpha)}\right) \\
&\qquad \leq K_d \sum_{l \geq M_{n+1}} l^{d-2} e^{-c_1 l^{2\alpha}} + K_d \sum_{l \geq M_{n+1}} l^{d-2} \exp\left(-\frac{1}{2} c_0(\varepsilon_{n+1}) l^{\eta(1-\alpha)}\right) \\
&\qquad \leq K_d \exp\left(-\frac{c_1 M_{n+1}^{2\alpha}}{2}\right) + K_d \exp\left(-\frac{1}{4} c_0(\varepsilon_{n+1}) M_{n+1}^{\eta(1-\alpha)}\right).
\end{aligned}$$

Since $M \leq M_{n+1}$, it is clear that both of these terms can be bounded by $\frac{C'}{M^L}$ for some constant $C' = C(\varepsilon, n, \alpha, d) > 0$. □

Chapter 4

Perspectives

Outline of the current chapter

In the final chapter of this thesis, we present a few thinking points regarding our multi-source IDLA model.

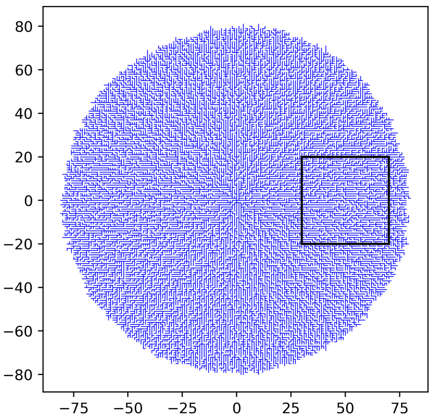
1 The IDLA Tree

It seems quite natural to mention the IDLA tree as a first application of our results. After all, all the work done so far on the aggregates $A_n[\infty]$ and $A_n^\dagger[\infty]$ was motivated by the study of the IDLA tree. As a reminder, this tree is constructed by considering the edge through which each particle exits the classical IDLA model (see Section 1.5 for more details). Among the questions we have concerning the tree, we can cite the existence of multiple infinite branches, and in the case where they exist, the asymptotic direction of these branches, as well as the existence of infinite branches in all directions. Another question would be to determine if the branches of this tree are straight. As mentioned earlier, the radial aspect of the IDLA tree makes it particularly difficult to study. A preferred strategy would be to compare the IDLA tree, restricted to a ball (along an axis) far from the origin , to the IDLA forest constructed in Chapter 3. The idea behind this argument is to say that far from the origin, the radial aspect of the IDLA tree seems to fade away. One is tempted to say that, when restricted to a ball far from the origin, both the tree and the forest have the same distribution. The method of approximating the distribution of a tree by the distribution of a forest is classical, used for example by Baccelli and Bordenave in [6] to compare the distribution of the *Radial Spanning Tree* (RST) to that of the *Directed Spanning Forest* (DSF). We would then need to show that the IDLA forest $\mathcal{F}_\infty$ is a good candidate to approximate the IDLA tree $\mathcal{T}_\infty$.

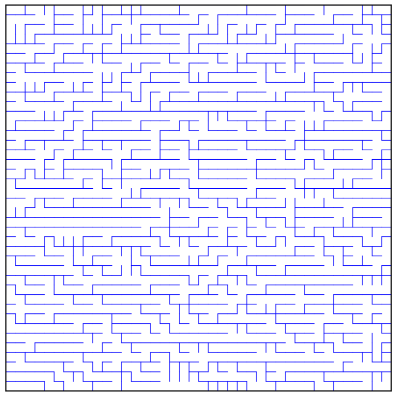
Let us detail what we mean when we say that we approach the tree with the forest. Fix $r > 0$, let $B_n := \mathbb{B}(n \cdot e_1, r)$, where $e_1 = (1, 0, 0, \dots)$. We would like to prove a result of the form:

$$\lim_{n \to +\infty} \mathbb{P}\left(\mathcal{T}_\infty \cap B_n \neq \mathcal{F}_\infty \cap B_n\right) = 0.$$

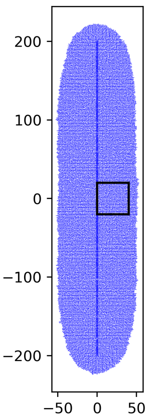
Note that in the approximation of the tree by the forest, it is crucial that the parameter r in the definition of B_n is taken independently of n.



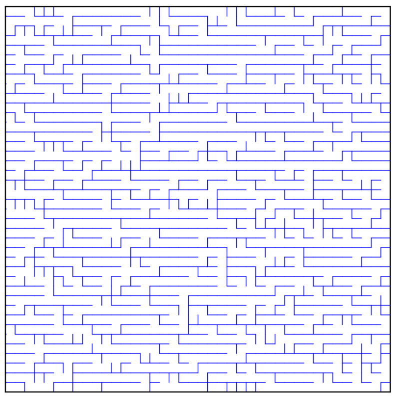
(a) A realization of $\mathcal{T}_{20000}$



(b) A zoom inside the black box



(c) A realization of $\mathcal{F}_{100}[200]$



(d) A zoom inside the black box

Figure 4.1: A comparison of IDLA trees and forests

We can compare with Figure 4.1 the branches of the IDLA tree with those of the directed IDLA forest. We observe a realization of the IDLA tree in a window far from the origin: we chose the window $[30, 70] \times [-20, 20]$. The same is done with the directed IDLA forest, which is observed in the window $[0,40] \times [-20,-20]$. One can notice that the IDLA tree no longer presents such a pronounced radial aspect as it does near the origin. The branches observed in this window are in some way similar to those obtained for the directed forest. Note however that we still see a bit of the radial aspect on the left side of this simulation, where the branches seem to be directed toward the origin. On the right side of the simulation, this feature is much less pronounced.

2 The IDLA forest

Another perspective would be the study of the infinite-volume directed IDLA forest, whose construction and existence is shown in Chapter 3. The initial goal behind the construction of this forest was to use it in order to approach the IDLA tree, but some interesting questions also arise around this object. We showed that this forest was ergodic with respect to translations of $\mathcal{H}$. A natural question concerns the trees that make up the forest, specifically their length. Are they all almost surely finite? In that case, what can we say about the law of the length of a 'typical tree' (in a sense that needs to be clarified) of the forest?

Another interesting result would be to prove stationarity 'far away' from the sources of $\mathcal{H}$. For the obvious reason that the forest has all of its roots in $\mathcal{H}$, $\mathcal{F}_\infty$ is not invariant with respect to horizontal translations T_k, with $k \in \mathbb{Z} \times \{0\}^{d-1}$. However, it would be interesting to study the long-term behavior of the forest, hence far from $\mathcal{H}$, and see if this forest exhibits stationary behavior. This would imply that after a long time, the branches 'forget' where they come from, and would be distributed following a stationary distribution. These questions are challenging, and perhaps require a finer study of the properties of the random aggregate $A_n^\dagger[\infty]$.

3 Links with the GFF

The *Gaussian Free Field* (GFF) is a random field that has been very popular in mathematical physics since the early 1970s. It can be interpreted as the generalization of a Brownian bridge, where the time parameter is replaced by a multidimensional parameter. Due to its very singular behavior, it is defined as a random distribution. Jerison, Levine and Sheffield show in [28] that the fluctuations of the classical IDLA aggregate, when properly renormalized in time and space, converge to a variant of the GFF. Therefore, if we define

$$E_t = \frac{1}{r} \sum_{x \in \mathbb{Z}^2} \left(\mathbb{1}_{x \in A(T(t))} - \mathbb{1}_{x \in \mathbb{B}(r)} \right) \delta_{x/r},$$

where $T(t)$ denotes a Poisson random variable with parameter t, $A(\cdot)$ the classical IDLA aggregate, and where $r = \sqrt{t/\pi}$, then E_t converges in the sense of distributions to a variant of the GFF. That is, for any family of test functions $\phi_1, \dots, \phi_k$ (from a suitably chosen space), the joint distribution of $(\langle \phi_1, E_t \rangle, \dots, \langle \phi_1, E_t \rangle)$ converges to $(\langle \phi_1, h \rangle, \dots, \langle \phi_1, h \rangle)$, where h denotes a variant of the GFF, called the *augmented GFF*.

We briefly comment on this result. When launching $T(t)$ particles, we expect the shape of $A(T(t))$, according to Theorem 1.2, to be close to that of $\mathbb{B}(r)$. The random measure E_t should be seen as the measure of the average error between the IDLA aggregate $A(T(t))$ and its expected shape $\mathbb{B}(r)$. Some of these deviations from the mean will cancel out, and it can then be shown that these fluctuations, suitably renormalized, converge in the sense of distributions to a variant of the GFF.

It would be interesting to see if such a result can be obtained for the the fluctuations of the aggregate $A_n[\infty]$. However, this seems like a more complex task, given that the limiting shape of the aggregate $A_n[\infty]$ is harder to handle than that of the classical aggregate $A(\cdot)$. Notably, all the isotropic properties of the Euclidean ball can not be applied in the case of the aggregate $A_n[\infty]$. One would need to manipulate the block $\mathcal{R}_{n/2} = [\![-n/2, n/2]\!] \times \mathbb{Z}^{d-1}$, which has worse properties, adding an additional difficulty to the problem.

4 Competition models

Up to now, we have always presented variants of IDLA in which particles are sent one after the other, without ever having several particles simultaneously alive. Here, we present an IDLA model where this is the case, and where particles compete with each other. To define this model, we use two families of Poisson Point Processes (PPP), $(\mathcal{N}_z^b)_{z\in\mathbb{Z}^d}$ and $(\mathcal{N}_z^m)_{z\in\mathbb{Z}^d}$, with respective intensities λ and μ. The first family of clocks corresponds to the birth times of particles, while the second family corresponds to the movement times of particles.

Thus, particles in this model are born at rate λ and move at rate μ. Given a realization of the aggregate $A_{\lambda,\mu}(t) \subset \mathbb{Z}^d$, the birth and movement rules are as follows:

- Particles are born at each site z according to the clocks given by $\mathcal{N}_z^b$, provided that z is in the current aggregate. If z is not in the aggregate, then nothing happens. That is to say, a particle can only be born at a site z if this site is already in the aggregate and if no other live particle occupies it.

- A particle located at site z moves according to the clocks given by $\mathcal{N}_z^m$. In this case, the particle chooses a neighboring site z' uniformly at random among its $2d$ neighbors. If a particle is alive at z', then the particle at z does not move (at least until the next clock of $\mathcal{N}_z^m$). However, if z' is unoccupied, the particle moves to z'. Its next move will then be determined by the next clock of $\mathcal{N}_{z'}^m$.

- If a particle moves to a site z that is not part of the aggregate, then the particle dies, and site z is added to the current aggregate. In particular, particles can now be born at site z.

Thus, starting from $A_{\lambda,\mu}(0) := \{0\}$, for $t > 0$, we define $A_{\lambda,\mu}(t)$ as the aggregate obtained up to time t using the protocol described above, and using clocks of $\mathcal{N}_z^b([0,t])$ and $\mathcal{N}_z^m([0,t])$.

We believe this model serves as a transit between a model called *uniform IDLA* or uIDLA, and the *Eden* model, also known as Richardson's model. The protocol of uIDLA is almost identical to the standard IDLA protocol, with the only difference being that at step $n+1$, the emitted particle starts from a site chosen uniformly at random in the current aggregate, instead of starting from the origin. This model has been particularly studied in [7]. The authors demonstrate a *shape theorem* with the limiting shape being the Euclidean ball of radius n, when a total of $|\mathbb{B}(n)|$ particles are emitted. In the Eden model, we start with an aggregate consisting of $\{0\}$ at time 0, and a new site x is added to the aggregate after an exponential time, with a rate proportional to the number of occupied neighbors of x. A shape theorem also exists for this model: Richardson shows in [45] the existence of an implicit limiting shape. This shape is convex, but it is conjectured that it is not the Euclidean ball.

We are interested in the behavior of our model depending on the value of the ratio λ/μ. In particular, we believe that when $\lambda/\mu \ll 1$, the model should behave in a similar manner as uIDLA. Our intuition is that if $\lambda/\mu \ll 1$, then particles are born at a very low rate compared to the movement rate, and the model should behave as though there is only one active particle at a time. The aspect where the particles compete with each other should be nonexistent, and we should end up in the case of uIDLA where a particle is born uniformly at random in the current aggregate and follows its trajectory until it exits the aggregate.

In the case where $\lambda/\mu \gg 1$, we believe the opposite occurs. There should be many active particles simultaneously, and the current aggregate should then have all of its sites occupied by active particles. In this case, all the particles would block each other out, except for particles located at the boundary of the aggregate, who could move (outwards) to add a new site to the aggregate. Recall that particles in our model can only move to unoccupied sites. We then presume that this model should have a behavior similar to that of the *Eden model*.

We give a simulation of both regimes ($\lambda/\mu \gg 1$ and $\lambda/\mu \ll 1$) in Figure 4.2. Both aggregates are made up of 5000 particles. The simulation on the left would correspond to the Eden regime, while the one on the right would correspond to the uIDLA regime. In the case of uIDLA, the limiting shape resembles a euclidean ball, whereas the limiting shape in the Eden model is a lot less clear. We provide another simulation of the Eden model in Figure 4.3.

(a) A simulation with $\lambda = 10^4, \ \mu = 1$ (b) A simulation with $\lambda = 1, \ \mu = 10^4$

Figure 4.2: A comparison of both regimes

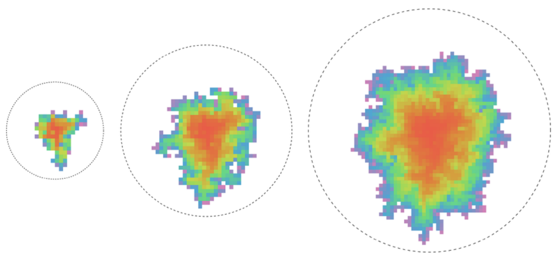


Figure 4.3: A simulation of the Eden model with 100, 500 and 1500 particles (source: [25])

Bibliography

[1] A. Asselah and A. Gaudillière. "From logarithmic to subdiffusive polynomial fluctuations for internal DLA and related growth models." English. In: *The Annals of Probability.* 41.3A (2013), pp. 1115–1159. ISSN: 0091-1798; 2168-894X/e.

[2] A. Asselah and A. Gaudillière. "Lower bounds on fluctuations for internal DLA." English. In: *Probab. Theory Relat. Fields* 158.1-2 (2014), pp. 39–53. ISSN: 0178-8051; 1432-2064/e.

[3] A. Asselah and A. Gaudillière. "On outer fluctuations for internal DLA". In: *arXiv preprint arXiv:1904.10168* (2019).

[4] A. Asselah and A. Gaudillière. "Sublogarithmic fluctuations for internal DLA." English. In: *The Annals of Probability.* 41.3A (2013), pp. 1160–1179. ISSN: 0091-1798; 2168-894X/e.

[5] A. Asselah and H. Rahmani. "Fluctuations for internal DLA on the comb". In: *Ann. Inst. Henri Poincaré Probab. Stat.* 52.1 (2016), pp. 58–83. ISSN: 0246-0203,1778-7017. DOI: 10.1214/14-AIHP629. URL: https://doi.org/10.1214/14-AIHP629.

[6] F. Baccelli and C. Bordenave. "The radial spanning tree of a Poisson point process". In: *The Annals of Applied Probability* 17.1 (2007), pp. 305–359.

[7] I. Benjamini, H. Duminil-Copin, C. Kozma, and C. Lucas. "Internal diffusion-limited aggregation with uniform starting points". In: *Ann. Inst. Henri Poincaré Probab. Stat.* 56.1 (2020), pp. 391–404. ISSN: 0246-0203. DOI: 10.1214/19-AIHP965. URL: https://doi.org/10.1214/19-AIHP965.

[8] I. Benjamini and O. Schramm. "Percolation in the hyperbolic plane". In: *J. Amer. Math. Soc.* 14.2 (2001), pp. 487–507. ISSN: 0894-0347,1088-6834. DOI: 10.1090/S0894-0347-00-00362-3. URL: https://doi.org/10.1090/S0894-0347-00-00362-3.

[9] I. Benjamini and A. Yadin. "Upper bounds on the growth rate of diffusion limited aggregation". In: *arXiv preprint arXiv:1705.06095* (2017).

[10] S. Blachère. "Internal diffusion limited aggregation on discrete groups of polynomial growth". In: *Random walks and geometry.* Walter de Gruyter, Berlin, 2004, pp. 377–391.

[11] S. Blachère and S. Brofferio. "Internal diffusion limited aggregation on discrete groups having exponential growth". In: *Probab. Theory Related Fields* 137.3-4 (2007), pp. 323–343. ISSN: 0178-8051,1432-2064. DOI: 10.1007/s00440-006-0009-2. URL: https://doi.org/10.1007/s00440-006-0009-2.

[12] S. Boucheron, C. Lugosi, and P. Massart. *Concentration inequalities. A nonasymptotic theory of independence.* eng. Oxford University Press, 2013.

[13] J. P. Chen, W. Huss, E. Sava-Huss, and A. Teplyaev. "Internal DLA on Sierpinski gasket graphs". In: *Analysis and geometry on graphs and manifolds* 461 (2020), pp. 126–155.

[14] N. Chenavier, D. Coupier, K. Penner, and A. Rousselle. “Construction of ergodic IDLA forests in $\mathbb{Z}^d$”. In: *arXiv preprint arXiv:2506.10476* (2025).

[15] N. Chenavier, D. Coupier, K. Penner, and A. Rousselle. “IDLA with sources in a hyperplane of $\mathbb{Z}^d$”. In: *arXiv preprint arXiv:2403.12590* (2024).

[16] N. Chenavier, D. Coupier, and A. Rousselle. “The bi-dimensional Directed IDLA forest”. In: *The Annals of Applied Probability* 33.3 (2023), pp. 2247–2290.

[17] D. Darrow. “A convergence rate for extended-source internal DLA in the plane”. In: *Potential Anal.* 61.1 (2024), pp. 35–64. ISSN: 0926-2601.

[18] D. Darrow. “Scaling limits of fluctuations of extended-source internal DLA”. In: *J. Anal. Math.* 150.2 (2023), pp. 449–484. ISSN: 0021-7670.

[19] P. Diaconis and W. Fulton. “A growth model, a game, an algebra, Lagrange inversion, and characteristic classes”. In: *Rend. Sem. Mat. Univ. Pol. Torino* 49.1 (1991), pp. 95–119.

[20] H. Duminil-Copin, C. Lucas, A. Yadin, and A. Yehudayoff. “Containing internal diffusion limited aggregation.” English. In: *Electron. Commun. Probab.* 18 (2013). Id/No 50, p. 8. ISSN: 1083-589X/e.

[21] U. Freiberg, N. Heizmann, R. Kaiser, and E. Sava-Huss. “Internal aggregation models with multiple sources and obstacle problems on Sierpiński gaskets”. In: *Journal of Fractal Geometry* 11.1 (2023), pp. 111–160.

[22] J.-B. Gouéré. “Subcritical regimes in the Poisson Boolean model of continuum percolation”. In: *Ann. Probab.* 36.4 (2008), pp. 1209–1220. ISSN: 0091-1798.

[23] J. Gravner and J. Quastel. “Internal DLA and the Stefan problem.” In: *Ann. Probab.* 28.4 (2000), pp. 1528–1562. ISSN: 0091-1798.

[24] N. Heizmann. “On the fluctuations of Internal DLA on the Sierpinski gasket graph”. In: *Mathematical and Computational Applications* 28.3 (2023), p. 73.

[25] D. M. Hua, F. Manin, T. Queer, and T. Wang. “Local behavior of the Eden model on graphs and tessellations of manifolds”. In: *Journal of Applied and Computational Topology* 8.6 (2024), pp. 1607–1647.

[26] W. Huss. “Internal diffusion-limited aggregation on non-amenable graphs”. In: *Electron. Commun. Probab.* 13 (2008), pp. 272–279. DOI: 10.1214/ECP.v13-1374. URL: https://doi.org/10.1214/ECP.v13-1374.

[27] W. Huss and E. Sava. “Internal aggregation models on comb lattices”. In: *Electron. J. Probab.* 17 (2012), no. 30, 21. DOI: 10.1214/EJP.v17-1940. URL: https://doi.org/10.1214/EJP.v17-1940.

[28] D. Jerison, L. Levine, and S. Sheffield. “Internal DLA and the Gaussian free field”. In: *Duke Math. J.* 163.2 (2014), pp. 267–308. ISSN: 0012-7094. DOI: 10.1215/00127094-2430259. URL: https://doi.org/10.1215/00127094-2430259.

[29] D. Jerison, L. Levine, and S. Sheffield. “Internal DLA for cylinders”. In: *Advances in analysis: the legacy of Elias M. Stein.* Vol. 50. Princeton Math. Ser. Princeton Univ. Press, Princeton, NJ, 2014, pp. 189–214.

[30] D. Jerison, L. Levine, and S. Sheffield. “Internal DLA in higher dimensions”. In: *Electron. J. Probab.* 18 (2013), No. 98, 14. DOI: 10.1214/EJP.v18-3137. URL: https://doi.org/10.1214/EJP.v18-3137.

[31] D. Jerison, L. Levine, and S. Sheffield. "Logarithmic fluctuations for internal DLA". In: *J. Amer. Math. Soc.* 25.1 (2012), pp. 271–301. ISSN: 0894-0347. DOI: 10.1090/S0894-0347-2011-00716-9. URL: https://doi.org/10.1090/S0894-0347-2011-00716-9.

[32] H. Kesten. "How long are the arms in DLA?" In: *Journal of Physics A: Mathematical and General* 20.1 (1987), p. L29.

[33] G. F. Lawler. "Subdiffusive fluctuations for internal diffusion limited aggregation." English. In: *The Annals of Probability* 23.1 (1995), pp. 71–86. ISSN: 0091-1798; 2168-894X/e.

[34] G. F. Lawler, M. Bramson, and D. Griffeath. "Internal diffusion limited aggregation". In: *The Annals of Probability* (1992), pp. 2117–2140.

[35] L. Levine and Y. Peres. "Scaling limits for internal aggregation models with multiple sources". In: *J. Anal. Math.* 111 (2010), pp. 151–219. ISSN: 0021-7670. DOI: 10.1007/s11854-010-0015-2. URL: https://doi.org/10.1007/s11854-010-0015-2.

[36] L. Levine and Y. Peres. "Strong spherical asymptotics for rotor-router aggregation and the divisible sandpile". In: *Potential Analysis* 30.1 (2009), pp. 1–27.

[37] L. Levine and V. Silvestri. "How long does it take for internal DLA to forget its initial profile?" In: *Probab. Theory Related Fields* 174.3-4 (2019), pp. 1219–1271. ISSN: 0178-8051. DOI: 10.1007/s00440-018-0880-7. URL: https://doi.org/10.1007/s00440-018-0880-7.

[38] C. Lucas. "The limiting shape for drifted internal diffusion limited aggregation is a true heat ball". In: *Probab. Theory Related Fields* 159.1-2 (2014), pp. 197–235. ISSN: 0178-8051. DOI: 10.1007/s00440-013-0505-0. URL: https://doi.org/10.1007/s00440-013-0505-0.

[39] P. Meakin and J. M. Deutch. "The formation of surfaces by diffusion limited annihilation". In: *The Journal of chemical physics* 85.4 (1986), pp. 2320–2325.

[40] Y. Mu, E. B. Procaccia, and Y. Zhang. "Scaling limit of DLA on a long line segment". In: *Transactions of the American Mathematical Society* 375.12 (2022), pp. 8769–8806.

[41] V. B. Priezzhev, D. Dhar, A. Dhar, and S. Krishnamurthy. "Eulerian walkers as a model of self-organized criticality". In: *Physical Review Letters* 77.25 (1996), p. 5079.

[42] E. B. Procaccia, J. Ye, and Y Zhang. "Stationary harmonic measure as the scaling limit of truncated harmonic measure". In: *arXiv preprint arXiv:1811.04793* (2018).

[43] E. B. Procaccia, J. Ye, and Y. Zhang. "Stationary DLA is well defined". In: *Journal of Statistical Physics* 181 (2020), pp. 1089–1111.

[44] E. B. Procaccia and Y. Zhang. "Stationary harmonic measure and DLA in the upper half plane". In: *Journal of Statistical Physics* 176.4 (2019), pp. 946–980.

[45] D. Richardson. "Random growth in a tessellation". In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 74. 3. Cambridge University Press, 1973, pp. 515–528.

[46] E. Sava-Huss. "From fractals in external DLA to internal DLA on fractals". In: *Fractal geometry and stochastics VI*. Springer. 2021, pp. 273–298.

[47] E. Shellef. "IDLA on the supercritical percolation cluster." English. In: *Electron. J. Probab.* 15 (2010). Id/No 24, pp. 723–740. ISSN: 1083-6489/e.

[48] V. Silvestri. "Internal DLA on cylinder graphs: fluctuations and mixing". In: *Electron. Commun. Probab.* 25 (2020), Paper No. 61, 14. ISSN: 1083-589X. DOI: 10.3390/mca25010013. URL: https://doi.org/10.3390/mca25010013.

[49] T. A. Witten and L. M. Sander. "Diffusion-limited aggregation". In: *Physical review B* 27.9 (1983), p. 5686.

IDLA aggregates and forests with sources in a hyperplane of $\mathbb{Z}^d$

Abstract

Internal Diffusion Limited Aggregation (IDLA) is a process used to construct random aggregates in $\mathbb{Z}^d$. At the start, it was introduced to model certain industrial chemical techniques such as electropolishing. Since then, many variants of this model have been studied, and have led to results describing the overall shape of the aggregate. Such results are referred to as *shape theorems*. This thesis handles several IDLA models in $\mathbb{Z}^d$, with $d \geq 2$. We study various families of random aggregates, whose construction is based on an IDLA protocol with an infinite number of sources. We establish stabilization results for these aggregates as well as shape theorems.

We begin by studying a multi-source IDLA model built by sending a deterministic number of particles from infinitely many sources. We establish properties of stationarity for this model, along with a global upper bound, which roughly describes the shape of the aggregate far away from the origin. This result allows us to then show a stabilization result, which turns out to be crucial in the second part of our study.

The second part is devoted to the study of a multi-source model similar to the previous one, built this time by sending a random number of particles from infinitely many sources, in a random order. We study this model with the goal of building a translation invariant forest, called the directed infinite-volume IDLA forest. Its existence is non-trivial due to issues of consistancy, and forces us to adapt results obtained for the previous model to this one, requiring even to sharpen some of these results. In this second part, we prove a stabilization result for forests, whose proof is based on arguments of percolation, allowing us to prove the existence of the directed infinite-volume IDLA forest.

Keywords: IDLA, random walks, random forests, random graphs, growth models, shape theorems, stabilization, percolation

---

Agrégats et forêts IDLA avec sources dans un hyperplan de $\mathbb{Z}^d$

Résumé

L'Agrégation Limitée par Diffusion Interne (IDLA) est un processus permettant de construire des agrégats aléatoires dans $\mathbb{Z}^d$. Il a initialement été introduit pour modéliser des problèmes de chimie industrielle tels que l'électro-polissage. Depuis, de nombreuses variantes de ce modèle ont été étudiées, et ont donné lieu à des résultats permettant de décrire la forme des agrégats obtenus. De tels résultats sont qualifiés dans la littérature de *shape theorems*. Cette thèse traite de plusieurs modèles d'agrégation limitée par diffusion interne sur $\mathbb{Z}^d$, avec $d \geq 2$. Nous étudions différentes familles d'agrégats aléatoires, dont la construction se base sur un protocole IDLA comportant une infinité de sources. Nous établissons des résultats de stabilisation pour ces agrégats, ainsi que des *shape theorems*.

Dans un premier temps, nous étudions un modèle d'IDLA multi-sources construit en envoyant un nombre déterministe de particules depuis un nombre infini de sources. Nous établissons des propriétés de stationnarité pour ce modèle, ainsi qu'une borne globale permettant de contrôler de façon grossière l'agrégat loin de l'origine. Ce résultat nous permet ensuite de montrer un résultat de stabilisation, qui se révèle être crucial pour la seconde partie de notre étude.

La seconde partie est devouée à l'étude d'un modèle multi-sources similaire au premier, mais construit cette fois-ci en envoyant un nombre aléatoire de particules depuis un nombre infini de sources, selon un ordre aléatoire. Nous étudions ce modèle dans le but de construire une forêt invariante par translation, appelée forêt IDLA dirigée de volume infini. L'existence de cette forêt est non-triviale à cause de problèmes de consistance, et nous amène à adapter et à affiner les résultats du précédent modèle pour le second. Nous démontrons dans cette seconde partie un résultat de stabilisation pour des forêts, dont la preuve est basée sur un argument de percolation, nous permettant de démontrer l'existence de la forêt IDLA dirigée de volume infini.

Mots clés : IDLA, marches aléatoires, forêts aléatoires, graphes aléatoires, modèles de croissance, *shape theorems*, stabilisation, percolation

---

LMPA

Maison de la Recherche Blaise Pascal – 50, rue Ferdinand Buisson – CS 80699 – 62228 Calais Cedex – France