Sapienza
Università di Roma

Oil and Water and Internal Diffusion Limited Aggregation

Facoltà di Scienze Matematiche, Fisiche e Naturali
Corso di Laurea Magistrale in Matematica Applicata

Candidate
Emanuele Pasqui
ID number 1746611

Thesis Advisor
Prof. Lorenzo Taggi

Academic Year 2021/2022

Oil and Water and Internal Diffusion Limited Aggregation
Master's thesis. Sapienza – University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Author's email: emanuelepasqui97@gmail.com

Contents

Introduzione

I sistemi di particelle hanno avuto molti utilizzi in diversi campi e continuano a rivelarsi uno strumento utile in applicazioni sempre nuove. Tali impieghi si coniugano in campi come la fisica matematica, la biologia delle popolazioni, la sociologia e la modellizzazione di fenomeni come il processo decisionale collettivo o il contagio del rischio di credito nella finanza e di numerosi altri problemi.
In questo lavoro studiamo diversi aspetti di alcuni processi di particelle dove, partendo da una configurazione iniziale, il sistema evolve seguendo delle regole ben precise. Questi modelli fanno parte di sistemi dinamici che mostrano criticità auto-organizzate e in particolare appartengono all'insieme delle cosiddette reti abeliane, reti di automi la cui configurazione finale è indipendente dall'ordine in cui essi agiscono. Da qualche decennio tali modelli sono largamente studiati al fine di comprendere meglio l'affioramento delle criticità ordinarie e di quelle auto-organizzate, che appaiono in un grande numero di situazioni nella vita reale come nel comportamento collettivo di gruppi di animali, l'azione coordinata di cellule cerebrali, la diffusione degli incendi boschivi e la distribuzione della magnitudo dei terremoti. In tali sistemi, il comportamento di un singolo individuo è il risultato degli stimoli provenienti da tutti gli altri, d'altra parte il numero di soggetti è così elevato che sarebbe troppo complicato far dipendere ogni azione dal comportamento di poche variabili libere. Tuttavia, si nota che questi sistemi si comportano in maniera più cooperativa e che in essi "i singoli gradi di libertà si mantengono l'un l'altro in un equilibrio più o meno stabile, che non può essere descritto come una perturbazione di qualche stato disaccoppiato, né in termini di pochi gradi di libertà collettivi", come scrivono Bak, Tang, e Wiesenfeld in [4], uno degli articoli più importanti in questo campo; ad esempio, nei sistemi ecologici le diverse specie "si sostengono" a vicenda,

ma il modo in cui lo fanno non può essere compreso analizzando isolatamente i singoli costituenti del sistema, inoltre la stessa interdipendenza delle specie rende l'ecosistema molto suscettibile a piccoli cambiamenti, detti fluttuazioni o rumore.

In questo elaborato, dopo aver esposto alcuni risultati preliminari nel Capitolo 1, in primo luogo tratteremo, nel Capitolo 2, il problema dell'internal aggregation, cioè della forma dell'aggregato (l'insieme dei siti toccati da almeno una particella) che si crea partendo da una configurazione in cui tutte le particelle sono nell'origine di $\mathbb{Z}^d$; ci porremo tale quesito per il modello Internal Diffusion Limited Aggregation (IDLA), un sistema unario, nel senso che in esso è presente un solo tipo di particella, e conservativo, vale a dire il numero delle particelle rimane invariato dall'inizio alla fine del processo. Questa è una dinamica sul reticolo intero d-dimensionale in cui si costruisce un aggregato di siti occupati cominciando al tempo 0 dall'origine e generando, ogni volta che non ci sono particelle che si muovono, una particella sull'origine stessa, la quale eseguirà una passeggiata aleatoria semplice e simmetrica sul grafo fino ad incontrare un sito non occupato, dove si fermerà: in questo modo, ogni volta che una particella si arresta il cluster conterrà esattamente un nuovo sito rispetto all'istante precedente. La questione fu risolta da Lawler, Bramson e Griffeath in [15], dove dimostrarono che l'aggregato che si forma con n particelle, per n abbastanza grande, è arbitrariamente vicino una palla euclidea di raggio $O\left(n^{1/d}\right)$.

Figure 0.1. Aggregato per IDLA con 100, 1600 e 25600 particelle.
Figura pubblicata in [18].

Nel Capitolo 3 ci interesseremo invece al modello Oil and Water, il quale, a differenza del precedente, non è unario, ma vi sono due tipi di particelle, detti olio e acqua. In tale processo, considerando un grafo sottostante infinito, non orientato e vertice transitivo con grado finito, si parte da una configurazione iniziale di particelle distribuita come un prodotto di variabili aleatorie che seguono una misura di probabilità ν su $\mathbb{N}_0$. A ciascun sito è associato un processo di Poisson di tasso 1 a ogni occorrenza del quale il sito corrispondente, se occupato da almeno due particelle di tipo diverso, fa scegliere a entrambe indipendentemente con probabilità uniforme un sito vicino, su cui saltano (da qui il nome dei tipi di particella). Tuttavia, come IDLA, questa dinamica è conservativa. Potremmo pensare, a priori, che il sistema Oil and Water possa avere due comportamenti, detti, in inglese, "*fixation*", dove ogni vertice invia sui vicini una coppia olio-acqua un numero finito di volte, e "*activity*", in cui questo succede un numero infinito di volte. Noi ci porremo il problema della transizione di fase del comportamento del modello tra fixation e activity, rispetto alla densità di particelle μ della configurazione iniziale (cioè il valore atteso della distribuzione derivante da ν). Infatti, indipendentemente dal grafo sottostante (a patto che abbia le caratteristiche sopra elencate), se $\mu < \infty$, come già provato da Candellero, Stauffer e Taggi in [9], il processo derivante dalla dinamica è transitorio nello spazio delle configurazioni, ossia esiste per ogni vertice un tempo finito dopo il quale nessuna coppia è più inviata da quel vertice sui vicini, dunque il sistema converge verso una configurazione che non permette più alle particelle di muoversi; dunque tali configurazioni costituiscono una famiglia di stati assorbenti per il processo, che perciò non attraversa una transizione di fase, indipendentemente dalla densità di particelle (purché finita), andando sempre incontro a un regime di fixation. Questo non è un comportamento tipico nelle reti abeliane, infatti i due modelli ad oggi più studiati nel campo, Activated Random Walks e Stochastic Sandpiles, esibiscono una transizione di fase, ovvero esiste una soglia critica non banale di un parametro (detto parametro d'ordine) superata la quale il sistema cambia il suo comportamento, come succede ad esempio all'acqua una volta attraversata la temperatura d'ebollizione a pressione fissata. Nel caso di questi modelli il parametro d'ordine è la densità di particelle per sito e la soglia critica è stata dapprima congetturata tramite simulazioni e poi è

stato analiticamente individuato l'intervallo a cui essa appartiene, cioè $(0, 1)$. Quasi certamente, partendo da una configurazione iniziale in cui il numero di particelle su ogni sito è distribuito secondo una variabile aleatoria di Poisson indipendente di parametro μ, questi modelli cadono sotto il regime di fixation quando μ è sotto questa soglia e di activity altrimenti.
Activated Random Walks è un modello di reazione-diffusione, in cui le particelle possono essere in uno di due stati, attivo o dormiente, indicati rispettivamente con A e S: le particelle nello stato A si muovono sul grafo seguendo la legge di una passeggiata aleatoria semplice e simmetrica a tempo continuo con tasso di salto 1 e quando saltano su un sito con almeno una particella nello stato S convertono istantaneamente tutte le particelle di quel sito allo stato A. Tuttavia, le particelle nello stato A passano allo stato S a un tasso positivo λ. Abbiamo dunque due reazioni tra loro contrapposte: una catalizzata, $A + S \to 2A$, che rappresenta il diffondersi dell'attività, e una spontanea, $A \to B$, la quale simboleggia invece la tendenza di questa attività a morire. La relazione tra il tasso di conversione allo stato dormiente e della densità di particelle determinerà quindi la continuazione o la cessazione dell'attività per tempi arbitrariamente lunghi.
Notiamo che, per $\lambda = +\infty$, Activated Random Walks segue le stesse regole di evoluzione di IDLA ma con tempo continuo e iniziando da una configurazione distribuita sull'interno grafo come prodotto di distribuzioni di Poisson con parametro μ indipendenti, dunque i risultati sulla transizione di fase sono estesi anche a quest'ultimo modello. Si può allora rispondere anche per una dinamica che segue le regole di evoluzione di IDLA al problema che avevamo originariamente sollevato per il sistema Oil and Water.

Allo stesso modo, cercheremo nel Capitolo 4 una risposta al problema della forma dell'aggregato per il sistema Oil and Water quando nella configurazione iniziale ci sono n coppie di particelle di acqua e olio nell'origine. La domanda è ad oggi ancora aperta, ma vi sono congetture (ad opera di Candellero, Ganguly, Hoffman e Levine), dimostrate solo in dimensione 1, che ipotizzano che l'aggregato che si ottiene, per n abbastanza grande, sia arbitrariamente vicino a una palla euclidea di raggio $O\left(n^{1/(d+2)}\right)$.

Per studiare più a fondo il problema, introdurremo un altro modello chiamato Internal Diffusion Limited Aggregation con Assorbimento (IDLAA), un sistema simile a IDLA ma con una condizione in più che rende la crescita dell'aggregato "più lenta"; infatti anche questo modello è a tempo discreto e tutte le particelle partono dall'origine, ma ad ogni passo temporale una particella viene selezionata tra le altre con distribuzione uniforme. Tale particella può o saltare su un sito vicino, secondo la distribuzione di una passeggiata aleatoria semplice e simmetrica sul grafo, o venire eliminata dal sistema, a seconda dello stato di un processo stocastico associato a ogni vertice. In particolare, quest'ultimo processo sarà, per ogni sito, una passeggiata aleatoria semplice e simmetrica su $\mathbb{Z}$ che parte da 0 al tempo iniziale. Questo modello non è presente nella letteratura scientifica e noi crediamo che l'ordine di grandezza del raggio del suo aggregato e di quello della dinamica Oil and Water siano lo stesso. Tali congetture saranno supportate da argomentazioni e simulazioni (prodotte con il software Matlab), mettendo in luce le analogie tra i due modelli. Nell'ultima parte del lavoro, infatti, ci occuperemo di uno studio numerico di questi processi, in particolare definiremo le fluttuazioni del raggio dell'aggregato intorno al suo valore atteso e ne caratterizzeremo l'ampiezza. Studieremo le fluttuazioni dividendole in due componenti, errori interni ed esterni: gli errori interni sono la differenza fra il valore atteso per il raggio dell'aggregato e il raggio effettivo della palla inscritta ad esso al termine del processo, mentre gli errori esterni sono la differenza fra il raggio effettivo della palla circoscritta all'aggregato al termine del processo e il valore atteso per il raggio. Detta $\mathfrak{B}_n$ la "palla reticolata" di raggio n, ossia l'intersezione della palla euclidea di raggio n con $\mathbb{Z}^d$, come dimostrato in [3] da Asselah e Gaudillière, in IDLA questi errori, per un sistema di $|\mathfrak{B}_n|$ particelle, sono entrambi $O(\log n)$ per $d = 2$ e $O(\sqrt{\log n})$ per $d \geq 3$. Le nostre simulazioni per $d = 2$ per Oil and Water e IDLAA, invece, sono coerenti con lo stesso risultato per gli errori interni, mentre per quelli esterni, per le limitate capacità di archiviazione delle nostre macchine, indicano che probabilmente con un numero di particelle maggiore il risultato sarebbe riscontrabile. Tuttavia, per gli errori esterni, le simulazioni sono coerenti con il fatto che siano $o(\log^2 n)$. D'altronde, fu inizialmente provato anche per IDLA, da Asselah e Gaudillière in [2], che, per $d \geq 2$, gli errori interni sono $O(\log n)$ e quelli esterni $O(\log^2 n)$. Dunque, un futuro studio analitico potrebbe analizzare la congettura per la quale per Oil and Water e IDLAA per $d = 2$ gli errori interni sono $O(\log n)$ e quelli esterni $o(\log^2 n)$.

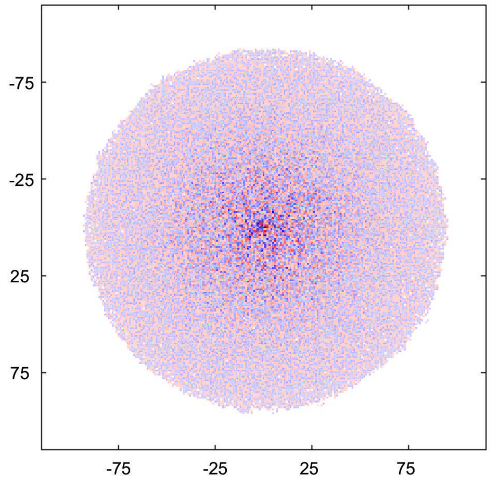


Figure 0.2. Configurazione finale per Oil and Water partendo con 2^{22} coppie olio-acqua nell'origine nella configurazione iniziale. I siti colorati di rosso e di blu sono quelli su cui si trovano particelle rispettivamente di olio e di acqua. L'intensità del colore indica il numero di particelle presenti sul sito: più il colore è vivido, più particelle sono presenti.

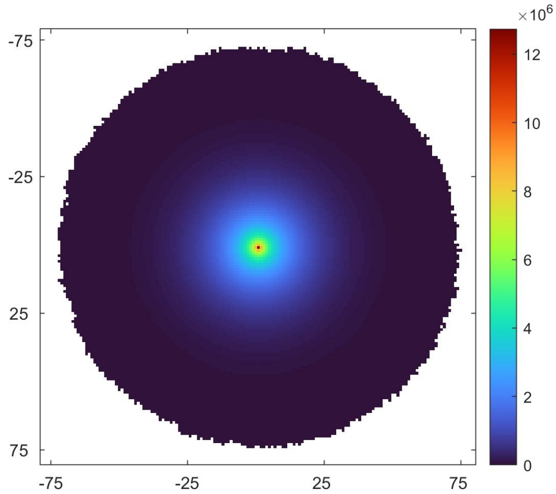


Figure 0.3. Aggregato per IDLAA con 2^{22} particelle nella configurazione iniziale. I colori rappresentano il numero di particelle che hanno toccato ciascun sito.

Introduction

Particle systems have had many uses in different fields and they keep proving to be a useful tool in ever-new applications. These uses are conjugated in fields such as mathematical physics, population biology, sociology and modeling of phenomena like collective decision-making or contagion of credit risk in finance and several other problems.
In this work we study different aspects of some particles processes where, starting from an initial configuration, the system evolves following defined rules. These models belong to the dynamic systems that show self-organized criticalities and in particular they belong to the set of the abelian networks, networks of automata whose final configuration is independent from the order in which they act. These models have been extensively studied for some decades in order to better understand the emergence of the ordinary criticalities and the self-organized ones, which appear in a large number of real life situations as in the collective behavior of animal groups, the coordinated action of brain cells, the spread of forest fires and the distribution of the magnitude of earthquakes. In such systems, the behavior of a single individual is the result of the stimuli from all the others; on the other hand, the number of entities is so high that it would be too complicated to make every action depend on the behavior of a few free variables. However, it is noticed that these systems behave in a more cooperative way and that in them “the individual degrees of freedom keep each other in a more or less stable balance, which cannot be described as a perturbation of some decoupled state, nor in terms of a few collective degrees of freedom”, as Bak, Tang and Wiesenfeld write in [4], one of the most important articles in this field; for example, in ecological systems the different species “support” each other, but the way they do it cannot be understood by analysing in isolation the individual

constituent of the system, moreover the same interdependence of species makes the ecosystem very susceptible to small changes, called fluctuations or noise.

In this thesis, after presenting some preliminary results in Chapter 1, in the first place we will discuss the topic of the internal aggregation, namely of the shape of the aggregate (the set of sites that are touched by at least one particle) that is created starting from a configuration in which all the particles are on the origin of $\mathbb{Z}^d$; we will ask this question for the model Internal Diffusion Limited Aggregation (IDLA), an unary system, in the sense that there is only one type of particle, and conservative, i.e. the number of particles stays unchanged during the whole process. This is a discrete time dynamics on the d-dimensional integer lattice in which an aggregate of occupied sites is built, starting from the origin at time 0 and generating, each time there are no moving particles, a particle on the origin itself, that will perform a simple symmetric random walk on the graph until it reaches an unoccupied site, where it will stop: in this way, each time a particle stops the cluster will contain exactly a new site with respect to the previous instant. The question was solved by Lawler, Bramson and Griffeath in [15], where they proved that the aggregate which is formed with n particles, for n large enough, is arbitrarily close to an Euclidean ball with radius $O\left(n^{1/d}\right)$.

Figure 0.4. Aggregates for IDLA with 100, 1600 and 25600 particles. Figure published in [18].

In Chapter 3 we will look at the Oil and Water model, that is not unary, differently from the previous one, but there are two types of particles, oil and water. In this process, considering an infinite, undirected and vertex-transitive graph with finite degree, we start from an initial configuration of particles distributed on the whole graph as a product of random variables distributed as a probability measure ν on $\mathbb{N}_0$. Each site is associated with a Poisson process of rate 1 to each occurrence of which the corresponding site, if occupied by at least two particles of different types, makes both independently choose a neighboring site with uniform probability, on which they jump (this is why the particle types have those names). However, like IDLA, this dynamics is conservative. We could think, *a priori*, that the Oil and Water system could have two behaviors, called "*fixation*", where every vertex send to its neighbors an oil-water pair finitely many times, and "*activity*", where this happens an infinite number of times. We will tackle the phase transition problem of the behavior of the model between fixation and activity, with respect to the particles density μ in the initial configuration (namely the expected value of the distribution deriving from ν). Indeed, independently from the underlying graph (as long as it has the properties listed above), if $\mu < \infty$, as already proven by Candellero, Stauffer and Taggi in [9], the process deriving from the model is transitory in the space of configurations, namely there exists for every vertex a finite after which no pair is sent from that vertex to the neighbors, therefore the system converges to a configuration that no longer allows particles to move; therefore these configurations form a family of absorbing states for the process, which therefore does not undergo a phase transition, regardless of the density of particles (supposing that it is finite), always moving towards a regime of fixation. This is not a typical behavior in the abelian networks, indeed the two most studied models to date, Activated Random Walks and Stochastic Sandpiles, exhibit a phase transition, that is there exists a nontrivial critical threshold of a parameter (called order parameter) beyond which the system changes its behavior, like what happens for example to water once it has crossed the boiling temperature for fixed pressure. In these models, the order parameter is the particles density per site and the critical threshold has been firstly conjectured via simulations and then the interval to which it belongs, namely $(0, 1)$, has been analytically located. Almost surely, starting from an initial configuration

where the number of particles on each site is distributed like an independent Poisson random variable with parameter μ, these models fall into the regime of fixation when μ is under this threshold and of activity otherwise.
Activated Random Walks is a reaction-diffusion model, where particles can be in one of two states, active or sleeping, indicated respectively with A and S: A-particles move on the graph according to a simple symmetric continuous-time random walk with jump rate 1 and when they jump on a site with at least one S-particle, any such particle at this site is immediately converted to the A state. However, particles in state A, when isolated in a vertex, switch to state S at a positive rate λ. We have therefore two contrasting reactions: the catalyzed one, $A + S \to 2A$, which represents the spread of activity, and the spontaneous one, $A \to B$, that symbolizes the tendency of this activity to die. The relation between λ and the particles density will therefore determine the continuation or termination of the activity for arbitrarily long times.
Notice that, for $\lambda = +\infty$, Activated Random Walks follows the same evolution rules as IDLA but with continuous time and starting from a configuration distributed as a product of independent Poisson distributions with parameter μ, therefore the results on the phase transition are extended to this model, too. Hence, we can answer also for a dynamics which follows the evolution rules of IDLA to the problem we originally arose for the Oil and Water system.

In the same way, in Chapter 4 we will look for a solution to the question of the shape of the aggregate for the Oil and Water system when in the initial configuration there are n pairs of oil and water particles in the origin. The problem is still open, but there are conjectures (by Candellero, Ganguly, Hoffman and Levine), proven just for dimension 1, which hypothesize that the aggregate, for n large enough, is arbitrarily close to an Euclidean ball with radius $O\left(n^{1/(d+2)}\right)$.
In order to study the problem even further, we will introduce a new model, Internal Diffusion Limited Aggregation with Absorption (IDLAA), a system similar to IDLA but with one more condition which "slows down" the growth of the aggregate; indeed also in this model all the particles start on the origin, but at each time step a particle

is selected uniformly at random among the other ones. Such a particle can either jump on a neighboring site, according to the distribution of a simple symmetric random walk on the graph, or be deleted from the system, depending on the state of a stochastic process associated to each vertex. In particular, this last process will be, for each site, a simple symmetric random walk on $\mathbb{Z}$ starting from 0 at the initial time. This model is not present in the scientific literature and we think that the order of magnitude of the radius of its aggregate and the one for the Oil and Water dynamics are the same.
These conjectures will be supported by reasonings and simulations (created with the software Matlab), highlighting the analogies between the two models. In the last part of the work, indeed, we will numerically study these processes, in particular we will define the fluctuations of the radius of the aggregate around its expected value and we will characterize their width. We will study the fluctuations dividing them in two components, inner and outer errors: the inner errors are the difference between the expected value for the radius of the aggregate and the actual radius of the ball inscribed in it at the end of the process, while the outer errors are the difference between the actual radius of the ball circumscribed to the aggregate at the end of the process and the expected value for the radius. Calling $\mathfrak{B}_n$ the "lattice ball" with radius n, i.e. the intersection of the Euclidean ball with radius n with $\mathbb{Z}^d$, as proven in [3] by Asselah and Gaudillière, in IDLA these errors, for a system with $|\mathfrak{B}_n|$ particles, are both $O(\log n)$ for $d = 2$ and $O(\sqrt{\log n})$ for $d \geq 3$. Our simulations for $d = 2$ for Oil and Water and IDLAA, instead, are consistent with the same results for the inner errors, while for the outer ones, because of the limited capabilities of our computers, indicate that probably with a larger number of particles the result would be noticeable. However, for outer errors, simulations are consistent with the fact that they are $o(\log^2 n)$. Indeed it was initially proven for IDLA, too, by Asselah and Gaudillière in [2], that, for $d \geq 2$, the inner errors are $O(\log n)$, and the outer ones are $O(\log^2 n)$. Therefore, a future analytical study could analyze the conjecture for which for Oil and Water and IDLAA for $d = 2$ the inner errors are $O(\log n)$ and the outer ones are $o(\log^2 n)$.

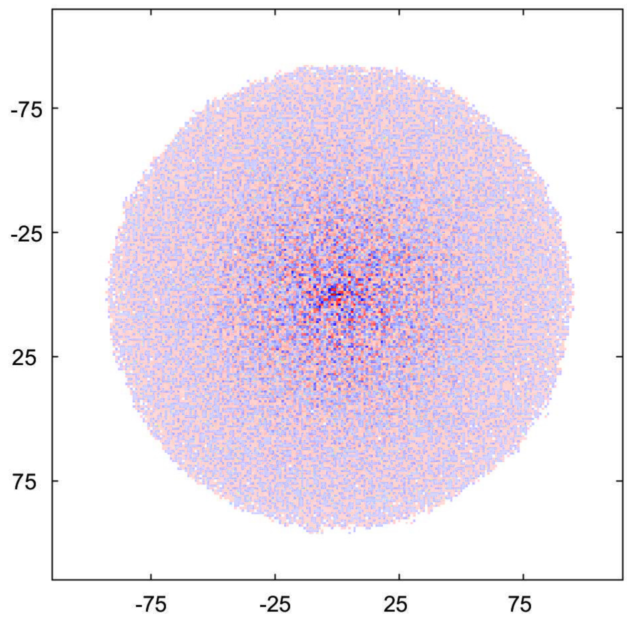


Figure 0.5. Final configuration for Oil and Water starting with 2^{22} oil-water pairs on the origin in the initial configuration. The blue sites are the ones occupied by water particles, while the red sites are occupied by oils. The intensity of the shade indicates the number of particles: the more particles there are, the more vivid the color.

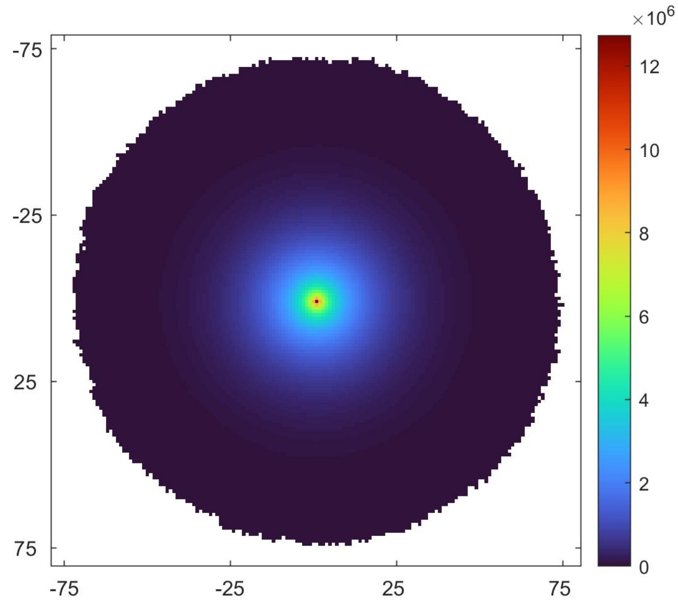


Figure 0.6. Aggregate for IDLAA with 2^{22} particles in the initial configuration. The colors represent the number of particles that touched each site

Chapter 1

Random walks on graphs and harmonic functions

In this chapter we explore harmonic functions on graphs and exploit them to derive several results on Green's function for simple symmetric random walks, which we will use in the next chapters.

Let $G = (V, E)$ be an undirected, vertex-transitive graph with finite degree. Recall that G is vertex-transitive if for every pair of vertices $v, u \in V$ there is some group automorphism $f : V \to V$ such that $f(u) = v$.

Let $d : V \times V \to \mathbb{N}_0$ be a distance on V where $d(x, y)$ is the length of the shortest path on G from x to y, for $x, y \in V$. We will write $x \sim y$ to indicate that $d(x, y) = 1$, i.e. that $\{x, y\} \in E$.

Let $\{X_t\}_{t\in\mathbb{N}_0}$ be a symmetric simple random walk on G, namely a Markov chain that begins at some vertex and at each time step moves to another vertex, choosing it uniformly at random among its neighbors, that is

$$P(X_{t+1} = y | X_t = x) = \begin{cases} \dfrac{1}{d_x} & \text{if } y \sim x, \\ 0 & \text{otherwise} \end{cases} \qquad \forall\ t \in \mathbb{N}_0,\ x, y \in V$$

where d_x is the degree of x. We will use the notation $P_x(\cdot) := P(\cdot | X_0 = x)$ and indicate with $E_x(\cdot)$ the expected value associated to this probability measure.

1.1 Harmonic functions on a graph and their relation with random walks

Let $\mathbf{0} \in V$ be a vertex that we will call the *origin*.
Consider a symmetric simple random walk $\{X_t\}_{t\in\mathbb{N}_0}$ on G.

Definition (Discrete Laplacian and harmonic function) *The discrete Laplacian of $g : V \to \mathbb{R}$ is the function $\Delta g : V \to \mathbb{R}$ defined as*

$$\Delta g(x) := \frac{1}{d_x} \sum_{y\sim x} (g(y) - g(x)) \tag{1.1}$$

and we say that g is harmonic in $K \subseteq V$ if $\Delta g(x) = 0$ for every $x \in K$.

Remark *If $f : V \to \mathbb{R}$, for $x \in V$*

$$\Delta f(x) = E_x\left[f(X_1) - f(X_0)\right], \tag{1.2}$$

indeed $E_x\left[f(X_1) - f(X_0)\right] = E_x\left[f(X_1)\right] - f(x) = \sum_{y\sim x} \frac{1}{d_x} f(y) - f(x) = \Delta f(x)$.

For $A \subseteq V$ we will indicate with

$$\tau_A := \inf\{t \geq 0 : X_t \notin A\} \qquad \text{and} \qquad \tau_A^+ := \inf\{t \geq 1 : X_t \notin A\}$$

respectively the exit time and the positive exit time for A.
Notice that if $X_0 \in A$, $\tau_A^+ = \tau_A$.
Let now $G = (V, E)$ be the d-dimensional hypercubic lattice, namely the graph with

$$V = \mathbb{Z}^d \qquad \text{and} \qquad E = \bigcup_{x\in\mathbb{Z}^d} \{\{x, y\} : \|x - y\| = 1\}.$$

Proposition 1 *If $f : \mathbb{Z}^d \to \mathbb{R}$ is bounded and harmonic in $A \subseteq \mathbb{Z}^d$, $M_t = f(X_{t\wedge\tau_A})$ is a martingale with respect to the natural filtration $\{\mathcal{F}_t\}_{t\in\mathbb{N}_0}$.*

Proof. Let $X_0 = x$.

$$E_x\left[f(X_{t+1})|\mathcal{F}_t\right] \underset{\substack{\uparrow \\ \text{Strong Markov property}}}{=} E_{X_t}\left[f(X_1)\right] \overset{\substack{(1.2) \\ \downarrow}}{=} f(X_t) + \Delta f(X_t)$$

hence

$$\begin{aligned}
E\left[M_{t+1}|\mathcal{F}_t\right] &= E\left[f(X_{(t+1)\wedge\tau_A})\mathbb{1}_{\{\tau_A>t\}}\middle|\mathcal{F}_t\right] + E\left[f(X_{(t+1)\wedge\tau_A})\mathbb{1}_{\{\tau_A\le t\}}\middle|\mathcal{F}_t\right]\\
&= E\left[f(X_{t+1})\mathbb{1}_{\{\tau_A>t\}}\middle|\mathcal{F}_t\right] + E\left[f(X_{t\wedge\tau_A})\mathbb{1}_{\{\tau_A\le t\}}\middle|\mathcal{F}_t\right]\\
&= \mathbb{1}_{\{\tau_A>t\}}E\left[f(X_{t+1})\middle|\mathcal{F}_t\right] + M_t\mathbb{1}_{\{\tau_A\le t\}}\\
&= \mathbb{1}_{\{\tau_A>t\}}\left(f(X_t)+\Delta f(X_t)\right) + M_t\mathbb{1}_{\{\tau_A\le t\}}\\
&\underset{\substack{\uparrow\\ \Delta f(X_t)=0 \text{ on } \{\tau_A>t\}}}{=} f(X_t)\mathbb{1}_{\{\tau_A>t\}} + M_t\mathbb{1}_{\{\tau_A\le t\}} = M_t.
\end{aligned}$$

□

Lemma 1 *For $A\subset\mathbb{Z}^d$ finite there exist $c=c_A<\infty, \rho=\rho_A<1$ such that for every $x\in A$*

$$P_x(\tau_A^+\ge n)\le c\rho^n.$$

Proof. Let $R=\sup\{\|x\|: x\in A\}$.

For every $x\in A$ there exists a path with length $R+1$ that starts in x and ends outside of A, therefore

$$P_x(\tau_A^+\le R+1)\ge\left(\frac{1}{2d}\right)^{R+1}.$$

For $k\in\mathbb{N}$,

$$\begin{aligned}
P_x(\tau_A^+>k(R+1)) &= P_x(\tau_A^+>(k-1)(R+1))P_x\left(\tau_A^+>k(R+1)\middle|\tau_A^+>(k-1)(R+1)\right)\\
&\underset{\substack{\uparrow\\ \text{Markov}}}{\le} P_x\left(\tau_A^+>(k-1)(R+1)\right)\left(1-\left(\frac{1}{2d}\right)^{R+1}\right)
\end{aligned}$$

thus $P_x(\tau_A^+>k(R+1))\le\rho^{k(R+1)}$ with $\rho=\left(1-\left(\frac{1}{2d}\right)^{R+1}\right)^{\frac{1}{R+1}}$.

Since if $n\in\mathbb{N}$ then $n=k(R+1)+j$ for some $k\in\mathbb{N}$ with $j\in\{1,\dots,R+1\}$,

$$P_x(\tau_A^+\ge n)\le P_x(\tau_A^+>k(R+1))\le\rho^{k(R+1)}\underset{\substack{\uparrow\\ n-(R+1)\ge k(R+1)}}{\le}\rho^{-(R+1)}\rho^n.$$

□

Remark 1 *If $A\subset V$ is a finite set of vertices, it follows from Lemma 1 that $\tau_A,\tau_A^+<\infty$ a.s. (regardless of X_0).*

Now, for $G = \left(\mathbb{Z}^d, \bigcup_{x\in\mathbb{Z}^d} \{\{x,y\} : \|x-y\| = 1\}\right)$, let, for $n \in \mathbb{N}$,

$$\mathfrak{B}_n := \{x \in \mathbb{Z}^d : \|x\| < n\}$$

be the "lattice ball" of radius n and define

$$\tau_n := \tau_{\mathfrak{B}_n} \qquad \text{and} \qquad \tau_n^+ := \tau_{\mathfrak{B}_n}^+.$$

The following result will be very useful in Chapter 2.

Lemma 2 *For $z \in \mathfrak{B}_n$*

$$n^2 - \|z\|^2 \le E_z(\tau_n) \le (n+1)^2 - \|z\|^2.$$

Proof. $M_t = \|X_t\|^2 - t$ is a martingale because for every $t \ge 1$

$$E\left(\|X_t\|^2 \middle| \mathcal{F}_{t-1}\right) = E\left(\|(X_t - X_{t-1}) + X_{t-1}\|^2 \middle| \mathcal{F}_{t-1}\right)$$
$$= E\left(\|X_t - X_{t-1}\|^2 \middle| \mathcal{F}_{t-1}\right) + 2E\left(X_{t-1}\cdot(X_t - X_{t-1})|\mathcal{F}_{t-1}\right) + E\left(\|X_{t-1}\|^2 \middle| \mathcal{F}_{t-1}\right);$$

now, since

$$X_m = \mu + \sum_{i=1}^m W_i$$

with $\mu = E(X_m)$ for all $m \ge 0$ and $W_i = (W_{i(1)}, \dots, W_{i(d)})$ for every $i \ge 1$, where $W_{i(j)}$ are i.i.d. random variables such that

$$P\left(W_{i(j)} = +1\right) = P\left(W_{i(j)} = -1\right) = \frac{1}{2d} \qquad \forall\ i \ge 1, j \in \{1, \dots, d\},$$

we have that $X_t - X_{t-1} = W_t$ is independent from $\mathcal{F}_{t-1}$, and using that X_{t-1} is $\mathcal{F}_{t-1}$-measurable, we obtain

$$E\left(\|X_t - X_{t-1}\|^2 \middle| \mathcal{F}_{t-1}\right) = E\left(\|X_t - X_{t-1}\|^2\right) = E\left(\|W_t\|^2\right) = 1 = t - (t-1),$$
$$2E(X_{t-1}\cdot(X_t - X_{t-1})|\mathcal{F}_{t-1}) = 2X_{t-1}E(W_t|\mathcal{F}_{t-1}) = 2X_{t-1}E(W_t) = 0,$$
$$E\left(\|X_{t-1}\|^2 \middle| \mathcal{F}_{t-1}\right) = \|X_{t-1}\|^2,$$

so that $E\left(\|X_t\|^2 \middle| \mathcal{F}_{t-1}\right) = \|X_{t-1}\|^2 + t - (t-1)$, i.e. M_t is a martingale.
For Lemma 1, if $\|z\| < n$, hence $z \in \mathfrak{B}_n$, there exists $c < \infty$ such that

$$E_z\left(|M_t|\mathbb{1}_{\{\tau_n^+ \ge t\}}\right) = |M_t|P_z(\tau_n^+ \ge t) \le |M_t|c\rho^t = |\,\|X_t\| - t|c\rho^t$$

where $\|X_t\|^2 \leq (n+1)^2$ since $\tau_n^+ \geq t$, so that

$$|\,\|X_t\| - t|c\rho^t \leq |(n+1)^2 - t|c\rho^t \xrightarrow[t\to\infty]{} 0$$

$\Longrightarrow$ we can use the optional sampling theorem (Theorem 12, see Appendix) to say that

$$E_z(M_{\tau_n^+}) = E_z(M_0) = \|z\|^2$$

and, since $n^2 - \tau_n^+ \leq \underbrace{\left\|X_{\tau_n^+}\right\|^2 - \tau_n^+}_{= M_{\tau_n^+}} \leq (n+1)^2 - \tau_n^+$,

$$\underbrace{E_z(n^2 - \tau_n^+)}_{= n^2 - E_z(\tau_n^+)} \leq \underbrace{E_z(M_{\tau_n^+})}_{= \|z\|^2} < \underbrace{E_z((n+1)^2 - \tau_n^+)}_{= (n+1)^2 - E_z(\tau_n^+)},$$

namely

$$n^2 - \|z\|^2 \leq E_z(\tau_n^+) < (n+1)^2 - \|z\|^2$$

with $\tau_n^+ = \tau_n$ because $z \in \mathfrak{B}_n$

$$\Longrightarrow n^2 - \|z\|^2 \leq E_z(\tau_n) < (n+1)^2 - \|z\|^2 .$$

$\square$

The results we have proven so far allow us to enunciate the following theorem which assures the uniqueness of the solution for Poisson-Dirichlet problems on a finite subset of $\mathbb{Z}^d$. In the following chapters we will also use this result for finite subsets of graphs that are undirected and vertex transitive with finite degree, which have the same proof.

Theorem 1 *Let $A \subset \mathbb{Z}^d$ finite, $F : \partial A \to \mathbb{R}$, $g : A \to \mathbb{R}$. The unique function $f : \overline{A} \to \mathbb{R}$ satisfying*

$$(a)\ \Delta f(x) = -g(x),\ x \in A$$

$$(b)\ f(x) = F(x),\ x \in \partial A$$

is

$$f(x) = E_x\left[F(X_{\tau_A}) + \sum_{t=0}^{\tau_A - 1} g(X_t)\right].$$

Proof. Notice that, for Lemma 2, $E_x\left[\sum_{t=0}^{\tau_A-1}|g(X_t)|\right] \le \|g\|_\infty E_x[\tau_A] < \infty$, hence f is well defined.

It is easy to prove that f satisfies (a) and (b), indeed

- for $x \in \partial A$ $\tau_A = 0$ P_x-a.s., so $f(x) = E_x\left[F(X_{\tau_A})\right] + E_x\left[\sum_{t=0}^{\tau_A-1} g(X_t)\right] = F(x)$ verifies (b);

- for $x \in A$ $\tau_A > 0$ P_x-a.s., so

$$\begin{aligned}
f(x) &= E_x\left[F(X_{\tau_A})\right] + E_x\left[\sum_{t=0}^{\tau_A-1} g(X_t)\right] \\
&= \frac{1}{2d}\sum_{y\sim x}\left(E_x\left[F(X_{\tau_A})\middle|X_1=y\right] - E_x\left[\sum_{t=1}^{\tau_A-1} g(X_t)\middle|X_1=y\right]\right) - g(x) \\
&\underset{\substack{\uparrow\\ \text{Markov}}}{=} \frac{1}{2d}\sum_{y\sim x}\left(E_y\left[F(X_{\tau_A})\right] - E_y\left[\sum_{t=0}^{\tau_A-1} g(X_t)\right]\right) - g(x) \\
&= \frac{1}{2d}\sum_{y\sim x} f(y) - g(x) \qquad \text{satifies } (a);
\end{aligned}$$

To show that f is unique, let f be a function verifying (a) and (b); then

$$M_t = f(X_{t\wedge\tau_A}) - \sum_{j=0}^{(t-1)\wedge(\tau_A-1)} \Delta f(X_j) = f(X_{t\wedge\tau_A}) - \sum_{j=0}^{(t-1)\wedge(\tau_A-1)} g(X_j)$$

is a martingale, indeed

$$\begin{aligned}
E\left[M_{t+1}|\mathcal{F}_t\right] = E&\left[\left(f(X_{(t+1)\wedge\tau_A}) - \sum_{j=0}^{t\wedge(\tau_A-1)} \Delta f(X_j)\right)\mathbb{1}_{\{\tau_A>t\}}\middle|\mathcal{F}_t\right] \\
&+ E\left[\left(f(X_{(t+1)\wedge\tau_A}) - \sum_{j=0}^{t\wedge(\tau_A-1)} \Delta f(X_j)\right)\mathbb{1}_{\{\tau_A\le t\}}\middle|\mathcal{F}_t\right] \\
= \underbrace{\mathbb{1}_{\{\tau_A>t\}}E\left[f(X_{t+1})|\mathcal{F}_t\right]}_{\mathbb{1}_{\{\tau_A>t\}}(f(X_t)+\Delta f(X_t))}& + f(X_{t\wedge\tau_A})\mathbb{1}_{\{\tau_A\le t\}} \\
&- \mathbb{1}_{\{\tau_A>t\}}\sum_{j=0}^{(t-1)\wedge(\tau_A-1)} \Delta f(X_j) - \mathbb{1}_{\{\tau_A>t\}}\Delta f(X_t) \\
&- \mathbb{1}_{\{\tau_A\le t\}}\sum_{j=0}^{(t-1)\wedge(\tau_A-1)} \Delta f(X_j) \\
= M_t.&
\end{aligned}$$

Furthermore

$$E_x\left[|M_t|\mathbb{1}_{\{\tau_A\geq t\}}\right]\leq(\|f\|_\infty+t\,\|g\|_\infty)\,P_x(\tau_A\geq t)\underset{\substack{\uparrow\\ \text{Lemma 1}}}{\leq} c\rho^t\xrightarrow[t\to\infty]{}0$$

for a $c<\infty$;

$$\xRightarrow[\text{optional sampling theorem}]{} f(x)=E_x(M_0)=E_x(M_{\tau_A})=E_x\left[F(X_{\tau_A})+\sum_{t=0}^{\tau_A-1}g(X_t)\right].$$

□

1.2 Green's function

The Green's function will be our primary tool to get our main results, since the stochastic processes we will study have behaviors which are referable to the ones of random walks.

1.2.1 Green's function for a set

Definition *Given $A\subseteq V$, the Green's function for A is $G_A:V\times V\to\mathbb{R}_{\geq 0}$ with*

$$G_A(x,y):=E_x\left(\sum_{t=0}^{\tau_A-1}\mathbb{1}_{\{X_t=y\}}\right).$$

Hence, $G_A(x,y)$ is the expected number of visits to y performed by a simple random walk started at x and killed upon exiting the set A.
We put $G_n:=G_{\mathfrak{B}_n}$.

Lemma 3 (Properties of G_A) *For $A\subseteq V$*

1) $G_A(x,y)=0$ *unless $x,y\in A$;*

2) $G_A(x,y)=G_A(y,x)$ *for all x,y;*

3) *for $x\in A$, $\Delta_x G_A(x,y)=-\delta_x(y)$ with* $\delta_x(y):=\begin{cases}1 & \text{if } y=x\\ 0 & \text{if } y\neq x\end{cases}.$

In particular, if $h(y)=G_A(x,y)$, then h vanishes on $V\setminus A$, is harmonic in $A\setminus\{x\}$ and $\Delta h(x)=-1$;

4) *for $y\in A$,* $G_A(y,y)=\dfrac{1}{P_y(\tau_A^+<\tau_{V\setminus\{y\}}^+)}<\infty;$

5) *for* $x, y \in A$, $G_A(x,y) = P_x(\tau_{V\setminus\{y\}} < \tau_A^+) G_A(y,y)$.

6) *If the graph is* $G = \left(\mathbb{Z}^d, \bigcup_{x\in\mathbb{Z}^d} \{\{x,y\} : \|x-y\| = 1\}\right)$, $G_A(x,y) = G_{A-x}(\mathbf{0}, y-x)$ *where* $A - x := \{z - x : z \in A\}$.

Proof. The first point and the last one are obvious.

For the second one, since, if $z_0 = x, z_1, \dots, z_{n-1}, z_n = y \in A$ then $P_x(X_1 = z_1, X_2 = z_2, \dots, X_n = y) = P_y(X_1 = z_{n-1}, X_2 = z_{n-2}, \dots, X_n = x)$ for the reversibility of the random walk, the thesis follows from

$$P_x(X_t = y, t < \tau_A^+) = P_y(X_t = x, t < \tau_A^+).$$

For the third property, let $x \in A$, hence $\tau_A \geq 1$ P_x-a.s.
Let $y \in A$. $\Delta_x G(x,y) = \frac{1}{d_y} \sum_{z\sim y} G_A(x,z) - G_A(x,y)$.

- if $y \neq x$

$$G_A(x,y) = G_A(y,x) = E_y\left(\sum_{t=0}^{\tau_A-1} \mathbb{1}_{\{X_t=x\}}\right) = \sum_{z\sim y} E_y\left(\sum_{t=0}^{\tau_A-1} \mathbb{1}_{\{X_t=x\}} \middle| X_1 = z\right) \underbrace{P_y(X_1 = z)}_{1/d_y}$$

$$= \frac{1}{d_y} \sum_{z\sim y} E_z\left(\sum_{t=0}^{\tau_A-1} \mathbb{1}_{\{X_t=x\}}\right) = \frac{1}{d_y} \sum_{z\sim y} G_A(z,x)$$

$$\implies G_A(x,y) = \frac{1}{d_y} \sum_{z\sim y} G_A(x,z);$$

- if $y = x$

$$G_A(x,x) = E_x\left(\sum_{t=0}^{\tau_A-1} \mathbb{1}_{\{X_t=x\}}\right) = \underbrace{E_x\left(\mathbb{1}_{\{X_0=x\}}\right)}_{1} + E_x\left(\sum_{t=1}^{\tau_A-1} \mathbb{1}_{\{X_t=x\}}\right)$$

$$= 1 + \sum_{z\sim x} \frac{1}{d_x} E_x\left(\sum_{t=1}^{\tau_A-1} \mathbb{1}_{\{X_t=x\}} \middle| X_1 = z\right)$$

$$= 1 + \sum_{z\sim x} \frac{1}{d_x} E_z\left(\sum_{t=1}^{\tau_A-1} \mathbb{1}_{\{X_t=x\}}\right)$$

$$\underset{\substack{\uparrow \\ z \neq x}}{=} 1 + \frac{1}{d_x} \sum_{z\sim x} E_z\left(\sum_{t=0}^{\tau_A-1} \mathbb{1}_{\{X_t=x\}}\right)$$

$$= 1 + \frac{1}{d_x} \sum_{z\sim x} G_A(z,x)$$

$$\implies \frac{1}{d_x} \sum_{z\sim x} G_A(z,x) - G_A(x,x) = -1.$$

To show (4),

$$G_A(y,y) = \underbrace{E_y\left(\sum_{t=0}^{\tau_A-1} \mathbb{1}_{\{X_t=y\}} \middle| \tau_A^+ < \tau_{V\setminus\{y\}}^+\right)}_{1} P_y\left(\tau_A^+ < \tau_{V\setminus\{y\}}^+\right)$$
$$+ \underbrace{E_y\left(\sum_{t=0}^{\tau_A-1} \mathbb{1}_{\{X_t=y\}} \middle| \tau_{V\setminus\{y\}}^+ < \tau_A^+\right)}_{1+G_A(y,y)} \left(1 - P_y\left(\tau_A^+ < \tau_{V\setminus\{y\}}^+\right)\right)$$
$$= 1 + G_A(y,y) - G_A(y,y)P_y\left(\tau_A^+ < \tau_{V\setminus\{y\}}^+\right)$$

$$\implies 0 = 1 - G_A(y,y)P_y\left(\tau_A^+ < \tau_{V\setminus\{y\}}^+\right).$$

To prove the last point, let $x, y \in A$.

$$G_A(x,y) = E_x\left(\sum_{t=0}^{\tau_A-1} \mathbb{1}_{\{X_t=y\}}\right) = P_x(\tau_{V\setminus\{y\}} < \tau_A^+)E_x\left(\sum_{t=0}^{\tau_A-1} \mathbb{1}_{\{X_t=y\}} \middle| \tau_{V\setminus\{y\}} < \tau_A^+\right)$$
$$= P_x(\tau_{V\setminus\{y\}} < \tau_A^+)E_y\left(\sum_{t=0}^{\tau_A-1} \mathbb{1}_{\{X_t=y\}}\right).$$

□

Let now $G = (V,E) = \left(\mathbb{Z}^d, \bigcup_{x\in\mathbb{Z}^d} \{\{x,y\} : \|x-y\| = 1\}\right)$.

Lemma 4 *If $y \in B \subseteq A$ with $A \subset \mathbb{Z}^d$ finite and $z \in \partial A$*

$$P_y\left(X_{\tau_A} = z\right) = \sum_{x\in B} G_A(y,x)P_x\left(X_{\tau_{A\setminus B}^+} = z\right).$$

Proof. Let $\eta^+ = \inf\{t \geq 1 : X_t \in \{y\} \cup \partial A\}$. For $z \in \partial A$

$$P_y(X_{\eta^+} = z) = P_y(X_{\eta^+} \neq y, X_{\eta^+} = z) = \underbrace{P_y(X_{\eta^+} = z | X_{\eta^+} \neq y)}_{P_y(X_{\tau_A}=z)} P_y(X_{\eta^+} \neq y)$$

$$\Rightarrow\ P_y(X_{\tau_A} = z) = P_y(X_{\eta^+} = z)\frac{1}{P_y(X_{\eta^+} \neq y)} = P_y(X_{\eta^+} = z)\frac{1}{P_y(\tau_A^+ < \tau_{V\setminus\{y\}}^+)}$$
$$= P_y(X_{\eta^+} = z)G_A(y,y) \underset{\substack{\uparrow \\ \text{reversibility}}}{=} P_z(X_{\eta^+} = y)G_A(y,y)$$
$$= G_A(y,y)\sum_{x\in B} \underbrace{P_z(X_{\eta^+} = y | X_{\tau_{A\setminus B}^+} = x)}_{P_x(X_{\eta^+}=y)} P_z\left(X_{\tau_{A\setminus B}^+} = x\right)$$
$$= \sum_{x\in B} P_z\left(X_{\tau_{A\setminus B}^+} = z\right) G_A(x,y).$$

where in the last equality we used that

$$G_A(x,y) = P_x(\tau_{V\setminus\{y\}} < \tau_A^+) G_A(y,y) = P_x(X_{\eta^+} = y) G_A(y,y).$$

$\square$

1.2.2 Green's function, transient case

For $V = \mathbb{Z}^d$, let now $d \geq 3$.

Definition 1 *For $d \geq 3$ we define the Green's function as $G : \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}_{\geq 0}$,*

$$G(x,y) := E_x \left(\sum_{t=0}^{\infty} \mathbb{1}_{\{X_t = y\}} \right).$$

Notice that the sum is never infinite because for $d \geq 3$ the walk is transient.

Lemma 5 (Properties of G) *If $d \geq 3$*

1) *for $x \in \mathbb{Z}^d$, $\Delta_x G(x,y) = -\delta_x(y)$, namely $h(y) = G(x,y)$ is harmonic in $\mathbb{Z}^d \setminus \{x\}$ and $\Delta h(x) = -1$;*

2) *for $x, y \in \mathbb{Z}^d$, $G(x,y) = G(y,x)$;*

3) *for $x, y \in \mathbb{Z}^d$, $G(x,y) = G(\mathbf{0}, x-y)$.*

Proof. $\Delta_x G(x,y) = \dfrac{1}{2d} \displaystyle\sum_{z \sim y} G(x,z) - G(x,y)$;

- if $y \neq x$

$$G(x,y) = G(y,x) = E_y \left(\sum_{t=0}^{\infty} \mathbb{1}_{\{X_t = x\}} \right) = \sum_{z \sim y} E_y \left(\sum_{t=0}^{\infty} \mathbb{1}_{\{X_t = x\}} \middle| X_1 = z \right) \underbrace{P_y(X_1 = z)}_{1/(2d)}$$

$$= \frac{1}{2d} \sum_{z \sim y} E_z \left(\sum_{t=0}^{\infty} \mathbb{1}_{\{X_t = x\}} \right) = \frac{1}{2d} \sum_{z \sim y} G(z,x)$$

$$\implies \Delta_x G(x,y) = 0;$$

- if $y = x$

$$G(x,x) = \underbrace{E_x \left(\mathbb{1}_{\{X_0 = x\}} \right)}_{1} + E_x \left(\sum_{t=1}^{\infty} \mathbb{1}_{\{X_t = x\}} \right) = 1 + \frac{1}{2d} \sum_{z \sim x} E_x \left(\sum_{t=1}^{\infty} \mathbb{1}_{\{X_t = x\}} \middle| X_1 = z \right)$$

$$= 1 + \frac{1}{2d} \sum_{z \sim x} E_z \left(\sum_{t=1}^{\infty} \mathbb{1}_{\{X_t = x\}} \right) \underset{\substack{\uparrow \\ z \neq x}}{=} 1 + \frac{1}{2d} \sum_{z \sim x} G(x,z)$$

$$\implies \Delta_x G(x,x) = -1,$$

therefore we have point (1). Point (2), similarly to the Green's function for a set, follows from the reversibility of the walk, while point (3) is obvious. □

Proposition 2 *For $A \subset \mathbb{Z}^d$ finite ($d \geq 3$) and $x, z \in A$*

$$G_A(x,z) = G(\mathbf{0}, z-x) - \sum_{y \in \partial A} P_x(X_{\tau_A} = y) G(\mathbf{0}, z-y).$$

Proof.

$$G_A(x,z) = E_x\left(\sum_{t=0}^{\tau_A - 1} \mathbb{1}_{\{X_t = z\}}\right) = E_x\left(\sum_{t=0}^{\infty} \mathbb{1}_{\{X_t = z\}} - \sum_{t=\tau_A}^{\infty} \mathbb{1}_{\{X_t = z\}}\right)$$
$$= \underbrace{G(x,z)}_{G(\mathbf{0},x-z)} - \sum_{y \in \partial A} P_x(X_{\tau_A} = y) \underbrace{G(y,z)}_{G(\mathbf{0},z-y)}.$$

□

From now on, we will indicate with ω_d the volume of the d-dimensional Euclidean ball with radius 1.

We will now present two results on the asymptotic behavior of $G_n(x, \mathbf{0})$ when x is far from the origin.

Proposition 3 *For $d \geq 3$, let $\eta = \inf\{t \geq 0 : X_t \in \{\mathbf{0}\} \cup \partial\mathfrak{B}_n\}$ and $x \in \mathfrak{B}_n$. Then, for $\|x\| \to \infty$,*

1) $P_x(X_\eta = \mathbf{0}) = \dfrac{a_d}{G(\mathbf{0},\mathbf{0})}\left(\|x\|^{2-d} - n^{2-d}\right) + O\left(\|x\|^{1-d}\right)$ *and*

2) $G_n(x, \mathbf{0}) = a_d\left(\|x\|^{2-d} - n^{2-d}\right) + O\left(\|x\|^{1-d}\right)$

with $a_d = \dfrac{d}{2}\Gamma\left(\dfrac{d}{2} - 1\right)\pi^{-d/2} = \dfrac{2}{(d-2)\omega_d}$.

Proof. $G(\mathbf{0}, x)$ is harmonic in $\mathbb{Z}^d \setminus \{\mathbf{0}\} \supseteq \mathbb{Z}^d \setminus (\{\mathbf{0}\} \cup \partial\mathfrak{B}_n)$ and bounded because of Theorem 7 (see Appendix), $\eta = \inf\left\{t \geq 0 : X_t \notin \mathbb{Z}^d \setminus (\{\mathbf{0}\} \cup \partial\mathfrak{B}_n)\right\}$
$\Rightarrow$ for Proposition 1, if $X_0 = x \in \mathbb{Z}^d \setminus (\{\mathbf{0}\} \cup \partial\mathfrak{B}_n)$, $M_t = G(\mathbf{0}, X_{t\wedge\eta})$ is a bounded martingale.

Then, for the optional sampling theorem,

$$G(\mathbf{0}, x) = E_x(M_0) = E_x(M_\eta) = E_x\left(G(\mathbf{0}, X_\eta)\right)$$
$$= E_x\left(G(\mathbf{0}, X_\eta) | X_\eta = \mathbf{0}\right) P_x(X_\eta = \mathbf{0}) + E_x\left(G(\mathbf{0}, X_\eta) | X_\eta \in \partial\mathfrak{B}_n\right) P_x(X_\eta \in \mathfrak{B}_n)$$

and, since $G(\mathbf{0},x) = a_d \|x\|^{2-d} + o(\|x\|^{1-d})$ for Theorem 7 and $E_x\left(G(\mathbf{0},X_\eta)\middle| X_\eta \in \partial\mathfrak{B}_n\right) \underset{\substack{\uparrow \\ \text{for } y\in\partial\mathfrak{B}_n,\ n\le \|y\|<n+1}}{=} a_d n^{2-d} + O(n^{1-d})$, we have the first result.
For the second one, notice that

$$G_n(x,\mathbf{0}) = P_x(X_\eta = \mathbf{0})G_n(\mathbf{0},\mathbf{0}) = \frac{a_d}{G(\mathbf{0},\mathbf{0})}G_n(\mathbf{0},\mathbf{0})\left(\|x\|^{2-d} - n^{2-d}\right) + O\left(\|x\|^{1-d}\right)$$

and for Proposition 2

$$\begin{aligned} G_n(\mathbf{0},\mathbf{0}) &= G(\mathbf{0},\mathbf{0}) - \sum_{y\in\partial\mathfrak{B}_n} P_{\mathbf{0}}(X_{\tau_n} = y)G(\mathbf{0},-y) \\ &= G(\mathbf{0},\mathbf{0}) - \sum_{y\in\partial\mathfrak{B}_n} P_{\mathbf{0}}(X_{\tau_n} = y)\underbrace{a_d\|y\|^{2-d}}_{O(n^{2-d})} + \sum_{y\in\partial\mathfrak{B}_n}\underbrace{o(\|y\|^{1-d})}_{O(n^{1-d})} \\ &= G(\mathbf{0},\mathbf{0}) + O(n^{2-d}). \end{aligned}$$

□

Proposition 4 *Suppose $n < m$ and $A = \{z \in \mathbb{Z}^d : n < \|z\| < m\}$ with $d \geq 3$. For $x \in A$ and $n \to \infty$*

$$P_x(\|X_{\tau_A}\| \le n) = \frac{\|x\|^{2-d} - m^{2-d} + O(n^{1-d})}{n^{2-d} - m^{2-d}}.$$

Proof. Consider the bounded martingale $M_t = G(\mathbf{0}, X_{t\wedge\tau_A})$. For the optional sampling theorem,

$$\begin{aligned} G(\mathbf{0},x) &= E_x(M_0) = E_x(M_{\tau_A}) \\ &= E_x\left(M_{\tau_A}\middle|\, \|X_{\tau_A}\| \le n\right)P_x\left(\|X_{\tau_A}\| \le n\right) + E_x\left(M_{\tau_A}\middle|\, \|X_{\tau_A}\| \ge m\right)\left(1 - P_x(\|X_{\tau_A}\| \le n)\right). \end{aligned}$$

For Theorem 7

$$\begin{aligned} G(\mathbf{0},x) &= a_d\|x\|^{2-d} + o\left(\|x\|^{1-d}\right), \\ E_x\left(M_{\tau_A}\middle|\, \|X_{\tau_A}\| \le n\right) &= a_d n^{2-d} + O(n^{1-d}), \\ E_x\left(M_{\tau_A}\middle|\, \|X_{\tau_A}\| \ge m\right) &= a_d m^{2-d} + O(m^{1-d}) \end{aligned}$$

and solving for $P_x\left(\|X_{\tau_A}\| \le n\right)$ we get the result. □

1.2.3 Recurrent case

Green's function $G(x,y)$ is infinite if $d \leq 2$ because on $\mathbb{Z}^2$ or $\mathbb{Z}$ the simple symmetric random walk is recurrent. However, there is another useful quantity called the *potential kernel.*

Definition (Potential kernel) *The potential kernel is the function $a : V \to \mathbb{R}_{\geq 0}$ with*

$$a(x) := \lim_{n\to\infty} E_{\mathbf{0}}\left[\sum_{t=0}^{n}\left(\mathbb{1}_{\{X_t=\mathbf{0}\}} - \mathbb{1}_{\{X_t=x\}}\right)\right]. \tag{1.3}$$

We will explore its relation with the Green's function for a set and its asymptotic behavior for $\|x\| \to \infty$.

Proposition 5 *For $d \leq 2$, $A \subset \mathbb{Z}^d$ finite and $x, z \in A$*

$$G_A(x,z) = \left[\sum_{y\in\partial A} P_x(X_{\tau_A} = y)a(y-z)\right] - a(z-x).$$

Proof. $h(x) = a(x-z)$ satisfies $\Delta h(x) = -\delta_{\mathbf{0}}(z-x)$ therefore for Theorem 1

$$\underset{a(x-z)}{\overset{\displaystyle h(x)}{\|}} = E_x\left[h(X_{\tau_A})\right] - E_x\left[\sum_{t=0}^{\tau_A-1}\mathbb{1}_{\{X_t=z\}}\right] = \sum_{y\in\partial A}\underbrace{h(x)}_{a(y-z)} P_x(X_{\tau_A}=y) - G_A(x,z).$$

□

Theorem 2 *For $d \leq 2$, $n \to \infty$*

$$G_n(\mathbf{0},\mathbf{0}) = \frac{2}{\pi}\log n + k + O\left(\frac{1}{n}\right).$$

Proof. Proposition 5 gives us

$$G_n(\mathbf{0},\mathbf{0}) = \sum_{y\in\partial\mathfrak{B}_n} P_{\mathbf{0}}(X_{\tau_A} = y)a(y) = \frac{2}{\pi}\log n + k + O\left(\frac{1}{n}\right)$$

where for the last equality we used Theorem 8 (see Appendix) with $\alpha = 1$ to get

$$a(y) = \frac{2}{\pi}\log\|y\| - k + O(\|y\|^{-1}) \underset{\substack{\uparrow\\ y\,\in\,\partial\mathfrak{B}_n}}{=} \frac{2}{\pi}\log n + k + O(n^{-1}).$$

□

Proposition 6 *For $d \leq 2$, let $x \in \mathfrak{B}_n$, $\eta = \inf\{t \geq 1 : X_t \in \{\mathbf{0}\} \cup \partial\mathfrak{B}_n\}$ and $\alpha < 2$. Then, for $n \to \infty$*

1) $P_x(X_\eta = \mathbf{0}) = \dfrac{1}{\log n}\left[\log n - \log\|x\| + o\left(\|x\|^{-\alpha}\right) + O\left(\dfrac{1}{\log n}\right)\right]$;

2) $G_n(x,\mathbf{0}) = \dfrac{2}{\pi}\left[\log n - \log\|x\|\right] + o\left(\|x\|^{-\alpha}\right) + O\left(\dfrac{1}{n}\right)$.

Proof. Let $X_0 = x$.
$\Delta a(x) = \delta(x)$ namely a is harmonic in $\mathbb{Z}^2 \setminus \{\mathbf{0}\} \supseteq (\{\mathbf{0}\} \cup \partial\mathfrak{B}_n)^c$; moreover, a is bounded, therefore, for Proposition 1, $M_t = a(X_{t\wedge\eta})$ is a bounded martingale, thus for the optional sampling theorem

$$\begin{aligned} a(x) &= M_0 = E_x(M_\eta) = E_x\left(a(X_\eta)\right) \\ &= \underbrace{E_x\left(a(X_\eta)\middle|X_\eta = \mathbf{0}\right)}_{0} P_x(X_\eta = \mathbf{0}) + E_x\left(a(X_\eta)\middle|X_\eta \in \partial\mathfrak{B}_n\right)\left(1 - P_x(X_\eta = \mathbf{0})\right) \\ &= E_x\left(a(X_\eta)\middle|\, \|X_\eta\| \geq n\right)\left(1 - P_x(X_\eta = \mathbf{0})\right) \\ &= \left(\frac{2}{\pi}\log n + k + O\left(\frac{1}{n}\right)\right)\left(1 - P_x(X_\eta = \mathbf{0})\right) \end{aligned}$$

where for the last equality we used Theorem 8 with $\alpha = 1$.
For the same theorem, $a(x) = \dfrac{2}{\pi}\log\|x\| + k + o\left(\|x\|^{-\alpha}\right)$. Plugging it in the equation above and solving for $P_x(X_\eta = \mathbf{0})$ we get

$$P_x(X_\eta = \mathbf{0}) = \frac{\frac{2}{\pi}\left(\log n - \log\|x\|\right) + o\left(\|x\|^{-\alpha}\right)}{\frac{2}{\pi}\log n + k + O\left(\frac{1}{n}\right)}$$

that is the first result. For the second one,

$$\begin{aligned} G_n(x,\mathbf{0}) &= P_x(X_\eta = \mathbf{0})G_n(\mathbf{0},\mathbf{0}) \underset{\substack{\uparrow \\ \text{Theorem 2}}}{=} P_x(X_\eta = \mathbf{0})\left[\frac{2}{\pi}\log n + k + O\left(\frac{1}{n}\right)\right] \\ &= \frac{1}{\log n}\left[\log n - \log\|x\| + o\left(\|x\|^{-\alpha}\right) + O\left(\frac{1}{\log n}\right)\right]\left[\frac{2}{\pi}\log n + k + O\left(\frac{1}{n}\right)\right] \\ &= \frac{2}{\pi}\left(\log n - \log\|x\|\right) + O\left(\frac{1}{n}\right). \end{aligned}$$

□

Proposition 7 *Suppose $n < m$, $A = \{z \in \mathbb{Z}^2 : n < \|z\| < m\}$. For $x \in A$ and $n \to \infty$*

$$P_x\left(\|X_{\tau_A}\| \leq n\right) = \frac{\log m - \log\|x\| + O\left(\frac{1}{n}\right)}{\log m - \log n}.$$

Proof. Consider the bounded martingale $M_t = a(X_{t\wedge\tau_A})$. For the optional sampling theorem

$$
\begin{aligned}
a(x) &= E_x(M_0) = E_x(M_{\tau_A}) \\
&= E_x\left(M_{\tau_A}\middle|\,\|X_{\tau_A}\| \le n\right) P_x\left(\|X_{\tau_A}\| \le n\right) + E_x\left(M_{\tau_A}\middle|\,\|X_{\tau_A}\| \ge m\right)\left(1 - P_x\left(\|X_{\tau_A}\| \le n\right)\right).
\end{aligned}
$$

For Theorem 8 with $\alpha = 1$

$$
\begin{aligned}
a(x) &= \frac{2}{\pi}\log\|x\| + k + o\left(\|x\|^{-1}\right), \\
E_x\left(M_{\tau_A}\middle|\,\|X_{\tau_A}\| \le n\right) &= \frac{2}{\pi}\log n + k + O\left(\frac{1}{n}\right), \\
E_x\left(M_{\tau_A}\middle|\,\|X_{\tau_A}\| \ge m\right) &= \frac{2}{\pi}\log m + k + O\left(\frac{1}{m}\right)
\end{aligned}
$$

and solving for $P_x\left(\|X_{\tau_A}\| \le n\right)$ we get the result. $\square$

The following theorems are useful to control the variation of harmonic functions and will be the last preliminary results.

Theorem 3 (Harnack inequality) *For every $r < 1$ there exists $c = c_r < \infty$ such that if $f : \overline{\mathfrak{B}}_n \to [0, \infty)$ is harmonic in $\mathfrak{B}_n$*

$$f(x) \le cf(y) \qquad \forall\ x, y \in \overline{\mathfrak{B}}_{rn}.$$

Proof. Firstly we will prove the result for $r = \dfrac{1}{16}$.
Let $z \in \partial\mathfrak{B}_{n/4}$ and $x \in \overline{\mathfrak{B}}_{n/16} = \overline{\mathfrak{B}}_{rn}$.
Since $\mathfrak{B}_{2n/3} + x \subseteq \mathfrak{B}_n$ and $\mathfrak{B}_n - x \subseteq \mathfrak{B}_{3n/2}$, we have

$$G_{2n/3}(\mathbf{0}, z - x) \le G_n(z, x) \le G_{3n/2}(\mathbf{0}, z - x).$$

Moreover, it is easy to prove that

$$\frac{3}{16}n - 1 \le \|z - x\| \le \frac{5}{16}n + 1.$$

Therefore, for Proposition 6 or Proposition 4, there exists c such that

$$G_n(z_1, x_1) \le cG_n(z_2, x_2) \qquad \forall\ z_1, z_2 \in \partial\mathfrak{B}_{n/4},\ x_1, x_2 \in \overline{\mathfrak{B}}_{n/16}.$$

If $y \in \partial\mathfrak{B}_n$ for lemma 4

$$P_x(X_{\tau_n}) = \sum_{z\in\partial\mathfrak{B}_{n/4}} P_y(X_{\tau^+_{\mathfrak{B}_n\setminus\mathfrak{B}_{n/4}}} = z)G_n(z, x)$$

hence if $x_1, x_2 \in \overline{\mathfrak{B}}_{n/16}$

$$P_{x_1}(X_{\tau_n} = y) \le cP_{x_2}(X_{\tau_n} = y)$$

and, for Theorem 1,

$$f(x_1) = E_{x_1}[f(X_{\tau_n})] = \sum_{y \in \partial \mathfrak{B}_n} \underbrace{E_{x_1}[f(X_{\tau_n})|X_{\tau_n} = y]}_{\substack{f(y) \\ \| \\ E_{x_2}[f(X_{\tau_n})|X_{\tau_n}=y]}} P_{x_1}(X_{\tau_n} = y)$$

$$\le c \sum_{y \in \partial \mathfrak{B}_n} E_{x_2}[f(X_{\tau_n})|X_{\tau_n} = y] P_{x_2}(X_{\tau_n} = y) = cf(x_2).$$

Now, let $r < 1$ and $x_1, x_2 \in \overline{\mathfrak{B}}_{rn}$, then $\|x_1\|, \|x_2\| \le rn$.

- if $\|x_1 - x_2\| \le \frac{1}{16}(1-r)n$ we can apply the above result to $\mathfrak{B}_{(1-r)n}(x_1) = \{z \in \mathbb{Z}^d : z - x \in \mathfrak{B}_{(1-r)n}\}$ to get $f(x_1) \le cf(x_2)$. By induction, if $\|x_1\|, \dots, \|x_{k+1}\| \le rn$ and $\|x_{i+1} - x_i\| \le \frac{1}{16}(1-r)n$ for $i = 1, \dots, k$,

$$f(x_1) \le c^k f(x_2);$$

- if $\|x_1 - x_2\| > \frac{1}{16}(1-r)n$ there exist $z_1 = x_1, z_2, \dots, z_{k_r+1} = x_2$ with $k_r \in \mathbb{N}$ large enough, $\|z_j\| \le rn$ for $j = 1, \dots, k_r + 1$ and $\|z_{i+1} - z_i\| \le \frac{1}{16}(1-r)n$ for $i = 1, \dots, k_r$, hence

$$f(x_1) \le c^{k_r} f(x_2).$$

□

The second part of this last proof can be applied to more general sets than $\mathfrak{B}_{(1-r)n}$. If $U \subset \mathbb{R}^d$ is a compact subset contained in an open set V, then U can be covered by a finite number of open balls with centers in U and radius at most $\frac{1}{2} dist(U, \partial V)$. Using this idea, we can prove the following result, that will be useful in the next chapter.

Theorem 4 (Harnack principle) *Let $U \subset \mathbb{R}^d$ be a compact set contained in an open set V, $U_n = nU \cap \mathbb{Z}^d := \{nu : u \in U\} \cap \mathbb{Z}^d$, $V_n = nV \cap \mathbb{Z}^d := \{nv : v \in V\} \cap \mathbb{Z}^d$ and $f : \overline{V}_n \to [0, \infty)$ harmonic in V_n. Then there exists $c = c_{U,V} < \infty$ such that*

$$f(x) \le cf(y) \qquad \forall\ x, y \in U_n.$$

Chapter 2

Internal Diffusion Limited Aggregation

In this chapter, we will study a growth process on the lattice $\mathbb{Z}^d$ that starts with only the origin occupied and grows a random cluster over time.

One by one, particles perform independent simple symmetric d-dimensional discrete-time random walks

$$X_t^i, \qquad i = 1, 2, \dots \ .$$

Each particle starts from the origin $\mathbf{0}$ and moves until it reaches a site that has not been visited previously, at which point it stops.
This model is called *internal diffusion limited aggregation* or *internal DLA* or *IDLA*.
We will focus on the shape of the cluster of the particles when there are no more jumps and each particle does not move anymore.
Let $A(n)$ be the set of the occupied sites after the n-th particle stops and set $A(0) := \{\mathbf{0}\}$ by convention. $A(n)$ is called the *aggregate* of the system with n particles.
Then, $\{A(n)\}_{n\in\mathbb{N}_0}$ is a Markov chain with transition probability

$$A(n+1) = A(n) \cup \{x\} \qquad \text{with probability } h(A(n)^c, x)$$

where $h_z(B, C)$ is the probability that a random walk that starts in z hits the set of sites B at some site in set C and the abbreviation $h(B, x) := h_{\mathbf{0}}(B, \{x\})$ will be used.

Extend the process $A(\cdot)$ to real $t \geq 0$ by setting $A(t) = A(\lfloor t \rfloor)$.

The main result of this chapter is the following theorem.

Theorem 5 *At time $\omega_d n^d$, internal DLA occupies a set of sites close to a d-dimensional ball of radius n. More precisely, for any $\varepsilon > 0$,*

$$\mathfrak{B}_{n(1-\varepsilon)} \subseteq A\left(\omega_d n^d\right) \subseteq \mathfrak{B}_{n(1+\varepsilon)} \qquad \textit{a.s. for all sufficiently large } n. \tag{2.1}$$

In order to prove it, firstly we will enunciate some preliminary probabilistic potential theory results, exploiting the facts we proved in chapter 1. Secondly, we will prove the lower bound for the shape of the aggregate in (2.1) and lastly its upper bound.

2.1 Preliminary probabilistic potential theory results

Recall that $\tau_n = \min\{t \geq 0 : X_t \notin \mathfrak{B}_n\}$, $G_n(y,z) = E_y\left(\sum_{t=0}^{\tau_n - 1} \mathbb{1}_{\{X_t = z\}}\right)$ and $G(y,z) = E_y\left(\sum_{t=0}^{\infty} \mathbb{1}_{\{X_t = z\}}\right)$, where G can be defined just for $d \geq 3$, otherwise the sum would diverge.

For $d \geq 3$, since the number of visits in $\mathbf{0}$ of a random walk that starts from $z \in \mathfrak{B}_n$ is the sum of those visits that happen before and after τ_n,

$$\begin{aligned} G(\mathbf{0},z) = G(z,\mathbf{0}) &= E_z\left(\sum_{t=0}^{\infty} \mathbb{1}_{\{X_t=\mathbf{0}\}}\right) = E_z\left(\sum_{t=0}^{\tau_n-1} \mathbb{1}_{\{X_t=\mathbf{0}\}}\right) + E_z\left(\sum_{t=\tau_n}^{\infty} \mathbb{1}_{\{X_t=\mathbf{0}\}}\right) \\ &= G_n(z,\mathbf{0}) + E_z\left(\sum_{t=\tau_n}^{\infty} \mathbb{1}_{\{X_t=\mathbf{0}\}}\right) = G_n(z,\mathbf{0}) + E_z\left[E_{X_{\tau_n}}\left(\sum_{t=0}^{\infty} \mathbb{1}_{\{X_t=\mathbf{0}\}}\right)\right] \end{aligned}$$

therefore we have the relation

$$G(\mathbf{0},z) = G(z,\mathbf{0}) = G_n(z,\mathbf{0}) + E_z[G(X_{\tau_n},\mathbf{0})]. \tag{2.2}$$

In the following results, to avoid exceptional cases when $z = \mathbf{0}$, the notation $[\![z]\!] := \max\{\|z\|, 1\}$ will be used.

Lemma 6 *For $z \in \mathfrak{B}_n$, $n \to \infty$*

$$G_n(\mathbf{0},z) = \frac{2}{\pi}\log\frac{n}{[\![z]\!]} + o\left(\frac{1}{[\![z]\!]}\right) + O\left(\frac{1}{n}\right) \qquad \textit{for } d = 2, \tag{2.3}$$

$$G_n(\mathbf{0}, z) = G(\mathbf{0}, z) + O(n^{2-d}) \qquad \textit{for } d \geq 3, \tag{2.4}$$

$$G_n(\mathbf{0}, z) = \frac{2}{d-2}\frac{1}{\omega_d}([\![z]\!]^{2-d} - n^{2-d}) + O\left([\![z]\!]^{1-d}\right) \qquad \textit{for } d \geq 3 \tag{2.5}$$

and if $z \in \mathfrak{B}_{n(1-\varepsilon)}$

$$G_{\varepsilon n}(\mathbf{0}, \mathbf{0}) \leq G_n(z, z) \leq G_{2n}(\mathbf{0}, \mathbf{0}). \tag{2.6}$$

Proof. For the proof of (2.3)

- if $z \neq \mathbf{0}$, for Proposition 6 with $\alpha = 1$, $G_n(z, \mathbf{0}) = \dfrac{2}{\pi} \log \dfrac{n}{\|z\|} + o\left(\dfrac{1}{\|z\|}\right) + O\left(\dfrac{1}{n}\right)$;

- if $z = \mathbf{0}$, for Theorem 2, $G_n(\mathbf{0}, \mathbf{0}) = \dfrac{2}{\pi} \log n + k + O\left(\dfrac{1}{n}\right)$

therefore we have the desired result

$$G_n(\mathbf{0}, z) = \frac{2}{\pi} \log \frac{n}{[\![z]\!]} + o\left(\frac{1}{[\![z]\!]}\right) + O\left(\frac{1}{n}\right) \qquad \forall\ z \in \mathfrak{B}_n.$$

In order to prove (2.5), we use Proposition 3, hence

- if $z \neq \mathbf{0}$, $G_n(z, \mathbf{0}) = \dfrac{2}{(d-2)\omega_d}(\|z\|^{2-d} - n^{2-d}) + O(\|z\|^{1-d})$;

- if $z = \mathbf{0}$, since $\Delta_x G_n(x, y) = -\delta_x(y)$,

$$\begin{aligned}
G_n(\mathbf{0}, \mathbf{0}) &= 1 + \frac{1}{2d} \sum_{y \sim \mathbf{0}} G_n(y, \mathbf{0}) \\
&= 1 + \frac{1}{2d} \sum_{y \sim \mathbf{0}} \left[\frac{2}{(d-2)\omega_d}(\|y\|^{2-d} - n^{2-d}) + O(\|y\|^{1-d})\right] \\
&= 1 + \frac{2}{(d-2)\omega_d}(1^{2-d} - n^{2-d}) + O(1^{1-d}) \\
&= \frac{2}{(d-2)\omega_d}(1^{2-d} - n^{2-d}) + O(1).
\end{aligned}$$

Therefore we have the relation (2.5) for each $z \in \mathfrak{B}_n$.

Now, for (2.2), $G_n(\mathbf{0}, z) = G(\mathbf{0}, z) - E_z[G(X_{\tau_n}, \mathbf{0})]$ with $G(X_{\tau_n}, \mathbf{0}) = O(n^{2-d})$ for $n \to \infty$, because of Theorem 7, hence

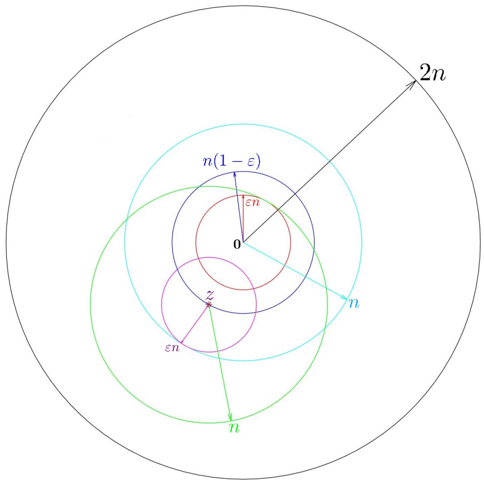
$$G_n(\mathbf{0}, z) = G(\mathbf{0}, z) + O(n^{2-d})$$

that is relation (2.4).

Now we will prove (2.6): for $z \in \mathfrak{B}_{n(1-\varepsilon)}$, calling, for $r > 0$,

$\mathfrak{B}_r(z) = \{x : \|x - z\| < r\} \cap \mathbb{Z}^d$ and $\tilde{\tau}_r = \inf\{t \geq 0 : X_t \notin \mathfrak{B}_r(z)\}$ respectively the lattice ball centered in z with radius r and the exit time for it,

$$G_{\varepsilon n}(\mathbf{0},\mathbf{0}) = G_{\mathfrak{B}_{\varepsilon n}(z)}(z,z) \underset{\substack{\uparrow \\ \mathfrak{B}_{\varepsilon n}(z) \subseteq \mathfrak{B}_n}}{\leq} G_n(z,z) = G_{\mathfrak{B}_n(z)}(\mathbf{0},\mathbf{0}) \underset{\substack{\uparrow \\ \mathfrak{B}_n(z) \subseteq \mathfrak{B}_{2n}}}{\leq} G_{2n}(\mathbf{0},\mathbf{0}).$$



□

The next lemma tells us that for $z \in \mathfrak{B}_{n(1-\varepsilon)}$ the mean of $G_n(y,z)$ on all the $y \in \mathfrak{B}_n$ is bounded above by $G_n(\mathbf{0},z)$.

Lemma 7 *Let $\varepsilon > 0$. Then, for sufficiently large n and $z \in \mathfrak{B}_{n(1-\varepsilon)}$,*

$$\sum_{y \in \mathfrak{B}_n} G_n(y,z) \leq \omega_d n^d G_n(\mathbf{0},z).$$

Proof.

$$\begin{aligned} E_z(\tau_n) &= E_z\left(\sum_{t=0}^{\tau_n - 1} 1\right) = E_z\left(\sum_{t=0}^{\tau_n - 1} \mathbb{1}_{\{X_t \in \mathfrak{B}_n\}}\right) = E_z\left(\sum_{t=0}^{\tau_n - 1} \sum_{y \in \mathfrak{B}_n} \mathbb{1}_{\{X_t = y\}}\right) \\ &= \sum_{y \in \mathfrak{B}_n} E_z\left(\sum_{t=0}^{\tau_n - 1} \mathbb{1}_{\{X_t = y\}}\right) = \sum_{y \in \mathfrak{B}_n} G_n(y,z) \leq (n+1)^2 - \|z\|^2, \end{aligned} \tag{2.7}$$

using Lemma 2 for the inequality. Now

- if $d = 2$, for Lemma 6, $\omega_d n^d G_n(\mathbf{0},z) = 2n^2 \log \frac{n}{[\![z]\!]} + o\left(\frac{1}{[\![z]\!]}\right) n^2 + O(n)$, so it sufficies to prove that for $d = 2$, for sufficiently large n

$$(n+1)^2 - \|z\|^2 \leq 2n^2 \log \frac{n}{[\![z]\!]} + o\left(\frac{1}{[\![z]\!]}\right) n^2 + O(n),$$

but this happens if and only if

$$\begin{aligned}
\frac{(n+1)^2}{[\![z]\!]^2} - 1 \leq 2\left(\frac{n}{[\![z]\!]}\right)^2 \log\frac{n}{[\![z]\!]} + o\left(\frac{1}{[\![z]\!]}\right)\left(\frac{n}{[\![z]\!]}\right)^2 + O(n)\frac{1}{[\![z]\!]} \quad &\Longleftrightarrow \\
O(n)\frac{1}{[\![z]\!]} + \frac{n^2}{[\![z]\!]^2} + \frac{1+2n}{[\![z]\!]^2} \leq 1 + 2\left(\frac{n}{[\![z]\!]}\right)^2 \log\frac{n}{[\![z]\!]} + o\left(\frac{1}{[\![z]\!]}\right)\left(\frac{n}{[\![z]\!]}\right)^2 \quad &\Longleftrightarrow \\
O(n)\frac{1}{[\![z]\!]} + \frac{n^2}{[\![z]\!]^2} \leq 1 + 2\left(\frac{n}{[\![z]\!]}\right)^2 \log\frac{n}{[\![z]\!]} + o\left(\frac{1}{[\![z]\!]}\right)\left(\frac{n}{[\![z]\!]}\right)^2 \quad &\Longleftrightarrow \\
\left(\frac{n}{[\![z]\!]}\right)^2\left(1 + O\left(\frac{1}{n}\right)\right) \leq 1 + 2\left(\frac{n}{[\![z]\!]}\right)^2 \log\frac{n}{[\![z]\!]} + o\left(\frac{1}{[\![z]\!]}\right)\left(\frac{n}{[\![z]\!]}\right)^2 &
\end{aligned}$$

hence we have to prove that this last inequality holds for $d = 2$ and sufficiently large n.

Considering the following inequality in $\delta = \delta(\varepsilon)$

$$(1+\delta)x^2 < 1 + 2(1-\delta)x^2 \log x \qquad \text{for } x \in \left[\frac{1}{1-\varepsilon}, \infty\right),$$

we have that it is true for

$$\delta < \frac{1 - x^2 + 2x^2\log x}{x^2 + 2x^2 \log x} \qquad \text{which is positive for } x \in \left[\frac{1}{1-\varepsilon}, \infty\right),$$

therefore, since $\frac{n}{[\![z]\!]} \in \left[\frac{1}{1-\varepsilon}, \infty\right)$ for every $z \in \mathfrak{B}_n$, we can take n sufficiently large such that

$$O\left(\frac{1}{n}\right) < \frac{1 - \left(\frac{n}{[\![z]\!]}\right)^2 + 2\left(\frac{n}{[\![z]\!]}\right)^2 \log\frac{n}{[\![z]\!]}}{\left(\frac{n}{[\![z]\!]}\right)^2 + 2\left(\frac{n}{[\![z]\!]}\right)^2 \log\frac{n}{[\![z]\!]}}$$

so we have proven the lemma for $d = 2$.

- if $d \geq 3$, again for Lemma 6

$$\omega_d n^d G_n(\mathbf{0}, z) = \frac{2}{d-2}([\![z]\!]^{2-d} - n^{2-d})n^d + O([\![z]\!]^{1-d})n^d$$

and it sufficies to show that for $d \geq 3$, for sufficiently large n

$$(n+1)^2 - [\![z]\!]^2 \leq \frac{2n^d}{d-2}([\![z]\!]^{2-d} - n^{2-d}) + O([\![z]\!]^{1-d})n^d,$$

but this happens if and only if

$$\frac{(n+1)^2}{[\![z]\!]^d} - [\![z]\!]^{2-d} \le \frac{2}{d-2}\left(\frac{n}{[\![z]\!]}\right)^d ([\![z]\!]^{2-d} - n^{2-d}) + O\left(\frac{1}{[\![z]\!]^{d-1}}\right)\left(\frac{n}{[\![z]\!]}\right)^d \iff$$
$$\left(\frac{n}{[\![z]\!]}\right)^2 + \frac{1+2n}{[\![z]\!]^2} \le 1 + \left\{\frac{2}{d-2} + O\left(\frac{1}{[\![z]\!]}\right)\right\}\left(\frac{n}{[\![z]\!]}\right)^d - \frac{2}{d-2}\left(\frac{n}{[\![z]\!]}\right)^2 \iff$$
$$\frac{2}{d-2}\left(\frac{n}{[\![z]\!]}\right)^2 + \left(\frac{n}{[\![z]\!]}\right)^2 + \frac{1+2n}{[\![z]\!]^2} \le 1 + \left\{\frac{2}{d-2} + O\left(\frac{1}{[\![z]\!]}\right)\right\}\left(\frac{n}{[\![z]\!]}\right)^d \iff$$
$$\left(\frac{n}{[\![z]\!]}\right)^2 \left(\underbrace{\frac{2}{d-2}+1}_{d/(d-2)} + O\left(\frac{1}{n}\right)\right) \le 1 + \left\{\frac{2}{d-2} + O\left(\frac{1}{[\![z]\!]}\right)\right\}\left(\frac{n}{[\![z]\!]}\right)^d \iff$$
$$\frac{d}{d-2}\left(1 + O\left(\frac{1}{n}\right)\right)\left(\frac{n}{[\![z]\!]}\right)^2 \le 1 + \left\{\frac{2}{d-2} + O\left(\frac{1}{[\![z]\!]}\right)\right\}\left(\frac{n}{[\![z]\!]}\right)^d,$$

that will be our new thesis.

Consider the inequality in $\delta = \delta(\varepsilon)$

$$\frac{d}{d-2}(1+\delta)x^2 < 1 + \frac{2}{d-2}(1-\delta)x^d \qquad \text{for } x \in \left[\frac{1}{1-\varepsilon}, \infty\right)$$

which is true for

$$\delta < \frac{d-2+2x^d-dx^2}{2x^d+2x^2} \qquad \text{that is positive for } x \in \left[\frac{1}{1-\varepsilon}, \infty\right),$$

therefore, since $\frac{n}{[\![z]\!]} \in \left[\frac{1}{1-\varepsilon}, \infty\right)$ for every $z \in \mathfrak{B}_n$, we can take n sufficiently large such that

$$O\left(\frac{1}{n}\right) < \frac{d-2+2\left(\frac{n}{[\![z]\!]}\right)^d - d\left(\frac{n}{[\![z]\!]}\right)^2}{2\left(\frac{n}{[\![z]\!]}\right)^d + 2\left(\frac{n}{[\![z]\!]}\right)^2}$$

and we have

$$\frac{d}{d-2}\left(1 + O\left(\frac{1}{n}\right)\right)\left(\frac{n}{[\![z]\!]}\right)^2 < 1 + \left\{\frac{2}{d-2} + O\left(\frac{1}{n}\right)\right\}\left(\frac{n}{[\![z]\!]}\right)^d;$$

now

$$1 + \left\{\frac{2}{d-2} + O\left(\frac{1}{n}\right)\right\}\left(\frac{n}{[\![z]\!]}\right)^d \le 1 + \left\{\frac{2}{d-2} + O\left(\frac{1}{[\![z]\!]}\right)\right\}\left(\frac{n}{[\![z]\!]}\right)^d \iff$$
$$O\left(\frac{1}{n}\right) \le O\left(\frac{1}{[\![z]\!]}\right)$$

and there exists $r > 0$ such that the thesis is true for every z with $[\![z]\!] \ge r$ (for fixed n).

For $[\![z]\!] < r$, we still have to show that for $d \geq 3$ for sufficiently large n we have $(n+1)^2 - [\![z]\!] \leq \omega_d n^d G_n(\mathbf{0}, z)$. For Lemma 6 this is equivalent to say that

$$(n+1)^2 - [\![z]\!] \leq \omega_d n^d \left(G(\mathbf{0}, z) + O(n^{2-d}) \right) \qquad \forall\ z \text{ with } [\![z]\!] < r$$

and we just have to take n big enough such that this is true simultaneously for all $[\![z]\!] < r$.

$\square$

Next, we formulate a very general large deviation estimate for sums of independent $\{0,1\}$-valued random variables.

Lemma 8 *Let* $S = \sum_{k=1}^{n} \mathbb{1}_{A_k}$ *with* $\mathbb{1}_{A_1}, \ldots, \mathbb{1}_{A_n}$ *independent random variables, and* $\mu = E(S)$. *Then, for all* $0 < \gamma < \frac{1}{2}$ *and sufficiently large* μ

$$P(|S - \mu| \geq \mu^{\frac{1}{2}+\gamma}) \leq 2e^{-\frac{1}{4}\mu^{2\gamma}}.$$

Proof. $S = \sum_{k=1}^{n} \mathbb{1}_{A_k}$ with $P(A_k) = p_k$, so $\mu = \sum_{k=1}^{n} p_k$.

$$\begin{aligned} E\left(e^{\lambda(S-\mu)}\right) &= E\left(e^{\sum \lambda \mathbb{1}_{A_k} - \lambda\mu}\right) = \prod_{k=1}^{n} E\left(e^{\lambda \mathbb{1}_{A_k}}\right) e^{-\lambda\mu} \\ &= \left[\prod_{k=1}^{n} \left(p_k e^{\lambda} + (1-p_k)\right)\right] \underset{\displaystyle\prod_{k=1}^{n} e^{-\lambda p_k}}{\underset{\|}{e^{-\lambda\mu}}} = \prod_{k=1}^{n} \left[p_k e^{\lambda(1-p_k)} + (1-p_k) e^{-\lambda p_k}\right]; \end{aligned}$$

for all $\lambda > 0$

$$\underset{\displaystyle P\left(S-\mu \geq \mu^{\frac{1}{2}\gamma}\right)}{\underset{\|}{P\left(e^{\lambda(S-\mu)} \geq e^{\lambda\mu^{\frac{1}{2}+\gamma}}\right)}} \underset{\substack{\uparrow \\ \text{Markov inequality}}}{\leq} \frac{E\left(e^{\lambda(S-\mu)}\right)}{e^{\lambda\mu^{\frac{1}{2}}\gamma}} = e^{-\lambda\mu^{\frac{1}{2}\gamma}} \prod_{k=1}^{n} \left[(1-p_k)e^{-\lambda p_k} + p_k e^{\lambda(1-p_k)}\right] =$$

$$= e^{-\lambda\mu^{\frac{1}{2}+\gamma}} e^{\log\left(\prod\left[(1-p_k)e^{-\lambda p_k} + p_k e^{\lambda(1-p_k)}\right]\right)} = e^{-\lambda\mu^{\frac{1}{2}+\gamma}} e^{\sum \log\left[(1-p_k)e^{-\lambda p_k} + p_k e^{\lambda(1-p_k)}\right]}$$

and we can prove that, for λ small enough,

$$e^{\sum \log\left[(1-p_k)e^{-\lambda p_k} + p_k e^{\lambda(1-p_k)}\right]} \leq e^{\lambda^2\mu}$$

so that we have the inequality

$$P\left(S - \mu \geq \mu^{\frac{1}{2}\gamma}\right) \leq e^{-\lambda\mu^{\frac{1}{2}+\gamma}} e^{\lambda^2\mu} = e^{-\lambda\mu^{\frac{1}{2}+\gamma}(1-\lambda\mu^{\frac{1}{2}-\gamma})}$$

and for $\lambda = \frac{1}{2}\mu^{\gamma-\frac{1}{2}}$ we get the first half of the thesis

$$P\left(S-\mu \geq \mu^{\frac{1}{2}\gamma}\right) \leq e^{-\frac{1}{2}\mu^{\gamma-\frac{1}{2}+\frac{1}{2}+\gamma}(1-\frac{1}{2}\mu^{\gamma-\frac{1}{2}+\frac{1}{2}-\gamma})} = e^{-\frac{1}{2}\mu^{2\gamma}(1-\frac{1}{2})} = e^{-\frac{1}{4}\mu^{2\gamma}}.$$

Indeed,

$$\sum_{k=1}^{n} \log\left[(1-p_k)e^{-\lambda p_k} + p_k e^{\lambda(1-p_k)}\right] \leq \sum_{k=1}^{n} \log\left(1+\lambda^2 p_k\right)$$

holds if for every $k \in \{1,\dots,n\}$

$$(1-p_k)e^{-\lambda p_k} + p_k e^{\lambda(1-p_k)} \leq (1+\lambda^2 p_k),$$

but this is true because, calling $p = p_k$, for Taylor's expansion around zero,

$$\begin{gathered}(1-p)e^{-\lambda p} + pe^{\lambda(1-p)} \\ = (1-p)\left(1-\lambda p+\frac{\lambda^2p^2}{2}+O(\lambda^3)\right)+p\left(1+\lambda(1-p)+\frac{\lambda^2(1-p)^2}{2}+O(\lambda^3)\right) \\ = 1-\lambda p+\frac{\lambda^2p^2}{2}+O(\lambda^3)-p+\lambda p^2-\frac{\lambda^2p^3}{2}+p+\lambda p-\lambda p^2+\frac{\lambda^2p}{2}+\frac{\lambda^2p^3}{2}-2\frac{\lambda^2p^2}{2} \\ = 1-\frac{\lambda^2p^2}{2}+\frac{\lambda^2p}{2}+O(\lambda^3),\end{gathered}$$

and

$$1-\frac{\lambda^2p^2}{2}+\frac{\lambda^2p}{2}+O(\lambda^3) \leq 1+\lambda^2 p \iff O(\lambda^3) \leq \lambda^2 p - \lambda^2 p^2.$$

By definition of $O(\lambda^3)$, there is a neighborhood of 0, I_0, such that $O(\lambda^3) \leq \lambda^3$ for every $\lambda \in I_0$, hence, in this neighborhood the inequality we want to show is implied by the following one,

$$\lambda^3 \leq \lambda^2 p - \lambda^2 p^2,$$

which holds if and only if

$$\lambda \leq p(1-p),$$

so it sufficies to take

$$\underset{\frac{1}{2}\mu^{\gamma-\frac{1}{2}}}{\overset{\lambda}{\shortparallel}} \in I_0 \cap \left(\bigcap_{k=0}^{n}\left(0, \frac{1-p_k}{p_k}\right]\right),$$

therefore (since $\gamma - \frac{1}{2} < 0$) μ sufficiently large.
Proceeding similarly for $P\left(S-\mu \leq -\mu^{\frac{1}{2}+\gamma}\right)$, we get

$$P\left(S-\mu \leq -\mu^{\frac{1}{2}+\gamma}\right) \leq e^{-\frac{1}{4}\mu^{2\gamma}}.$$

$\square$

Let us now define the shells

$$\mathscr{S}_k := \{x : k \le \|x\| < k+1\}$$

and their (positive) hitting times

$$T_k = \min\{t \ge 1 : X_t \in \mathscr{S}_k\}.$$

In the following proofs, we will often use the fact that a random walk cannot exit the ball $\mathfrak{B}_k$ without hitting $\mathscr{S}_k$.

Lemma 9 *There exists $J = J_d < \infty$ such that for $j < k$ and $\Delta = k - j$*
1) *if $z \in \mathscr{S}_k$*

$$P_z(T_j < T_k) \le \frac{J}{\Delta}; \tag{2.8}$$

2) *if $y \in \mathscr{S}_j$ and $B \subseteq \mathscr{S}_k$,*

$$h_y(\mathscr{S}_k, B) \le \frac{J|B|}{\Delta^{d-1}} \tag{2.9}$$

(recall that $h_y(\mathscr{S}_k, B)$ is the probability that a random walk starting at y first hits the set $\mathscr{S}_k$ at some site in set B).

Proof. For part (1), $\mathscr{S}_k = \{z : k \le \|z\| < k+1\} = A \setminus \{z : \|z\| = k\}$ with $A = \{z : k < \|z\| < k+1\}$;

- for $d \ge 3$, for Proposition 4 we know that for every $z \in A$

$$P_z(\|X_{\tau_A}\| \le k) = \frac{\|z\|^{2-d} - (k+1)^{2-d} + O(k^{1-d})}{k^{2-d} - (k+1)^{2-d}}.$$

Let $z \in \mathscr{S}_k,\ j < k$ and $\hat{\tau}_k = \inf\{t \ge 0 : X_t \notin \mathscr{S}_k\}$; then

$$\begin{aligned} P_z(T_j < T_k) &= P_z(\|X_{\hat{\tau}_k}\| < k) P_z\left(T_j < T_k \middle| \|X_{\hat{\tau}_k}\| < k\right) \\ &= \frac{\|z\|^{2-d} - (k+1)^{2-d} + O(k^{1-d})}{k^{2-d} - (k+1)^{2-d}} P_z\left(T_j < T_k \middle| \|X_{\hat{\tau}_k}\| < k\right). \end{aligned}$$

Now, if $j + 1 = k$, i.e. $\Delta = k - j = 1$, our aim is to prove that

$$\frac{\|z\|^{2-d} - (k+1)^{2-d} + O(k^{1-d})}{k^{2-d} - (k+1)^{2-d}} < \frac{c}{1} = c \qquad \text{with } c < \infty,$$

but this is true because, since $\dfrac{1}{\|z\|^{d-2}} \leq \dfrac{1}{k^{d-2}}$,

$$\frac{\|z\|^{2-d}-(k+1)^{2-d}+O(k^{1-d})}{k^{2-d}-(k+1)^{2-d}} = \frac{\dfrac{1}{\|z\|^{d-2}}-\dfrac{1}{(k+1)^{d-2}}+O\left(\dfrac{1}{k^{d-1}}\right)}{\dfrac{1}{k^{d-2}}-\dfrac{1}{(k+1)^{d-2}}}$$

$$\leq \frac{\dfrac{1}{k^{d-2}}+O\left(\dfrac{1}{(k+1)^{d-2}}\right)}{O\left(\dfrac{1}{k^{d-2}}\right)} = \frac{O\left(\dfrac{1}{k^{d-2}}\right)}{O\left(\dfrac{1}{k^{d-2}}\right)} = O(1) < c \qquad \text{with } c < \infty;$$

if $j+1<k$

$$P_z(T_j < T_k) = \underbrace{\frac{\|z\|^{2-d}-(k+1)^{2-d}+O(k^{1-d})}{k^{2-d}-(k+1)^{2-d}}}_{\leq c} P_z\left(T_j < T_k \middle| \|X_{\hat{\tau}_k}\| < k\right).$$

Let $X_{\hat{\tau}_k} = y$. Then, on the event $\{\|X_{\hat{\tau}_k}\| < k\}$, $j+1 \leq \|y\| < k$ and $P_z\left(T_j < T_k \middle| \|X_{\hat{\tau}_k}\| < k\right) = P_y(T_j < T_k)$, therefore

$$P_z(T_j < T_k) \leq cP_y(T_j < T_k)$$

and, defining $\tilde{A} = \{z : j+1 \leq \|z\| < k\}$,

$$cP_y(T_j < T_k) = c\underbrace{P_y\left(T_j < T_k \middle| \left\|X_{\tau_{\tilde{A}}}\right\| < j\right)}_{=1} P_y\left(\left\|X_{\tau_{\tilde{A}}}\right\| < j\right)$$

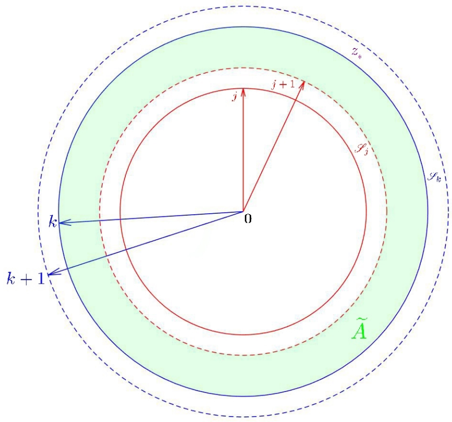
therefore, we have to prove that

$$P_y\left(\left\|X_{\tau_{\tilde{A}}}\right\| < j\right) \leq \frac{1}{\Delta} \qquad \text{with } c' < \infty.$$

For Proposition 4 applied to $\tilde{A}$ we get

$$\begin{aligned} P_y\left(\left\|X_{\tau_{\tilde{A}}}\right\| < j\right) &= \frac{\|y\|^{2-d}-k^{2-d}+O((j+1)^{1-d})}{O((k-j)^{2-d})} \\ &\leq \frac{O((k-j)^{2-d})}{O((k-j)^{3-d})} = O\left(\frac{1}{k-j}\right) \leq \frac{c'}{\Delta}. \end{aligned}$$

- For $d \leq 2$ the proof is similar, using the potential kernel $a(x)$ and Proposition 7.



So we have proven part (1).

Now, to show that if $y \in \mathscr{S}_j$ then

$$h_y(\mathscr{S}_k, B) \leq \frac{J|B|}{\Delta^{d-1}},$$

since for $\overline{\Delta} \geq \Delta$ we have $\dfrac{J|B|}{\overline{\Delta}^{d-1}} \leq \dfrac{J|B|}{\Delta^{d-1}}$, the thesis is implied by

$$h_y(\mathscr{S}_k, B) \leq \frac{J|B|}{\Delta^{d-1}} \qquad \text{for } \Delta \text{ sufficiently large,}$$

and since $h_y(\mathscr{S}_k, B) = \sum_{z \in B} h_y(\mathscr{S}_k, z)$, it sufficies to prove that

$$h_y(\mathscr{S}_k, z) \leq \frac{J}{\Delta} \qquad \text{for sufficiently large } \Delta,\ \forall\ z \in B.$$

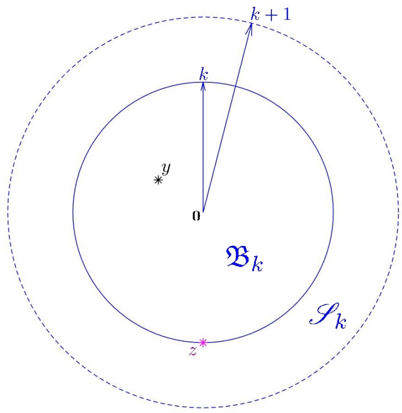
Let $\rho = \rho_{y,k} = \inf\{t \geq 1 : X_t \in \mathscr{S}_k \cup \{y\}\}$;

$$h_y(\mathscr{S}_k, z) = P_y(X_{T_k} = z) = \begin{cases} 0 & \text{if } z \notin \partial\mathfrak{B}_k \\ P_y(X_{\tau_k} = z) & \text{if } z \in \partial\mathfrak{B}_k \end{cases}.$$

$y \in \{y\} \subseteq \mathfrak{B}_k$, therefore, for Lemma 4, for $z \in \partial\mathfrak{B}_k$,

$$h_y(\mathscr{S}_k, z) = G_k(y, y) P_y\left(X_{\tau^+_{\mathfrak{B}_k \setminus \{y\}}} = z\right) = G_k(y, y) P_y(X_\rho = z)$$

since $\tau^+_{\mathfrak{B}_k \setminus \{y\}} = \inf\{t \geq 1 : X_t \notin \mathfrak{B}_k \setminus \{y\}\} = \inf\{t \geq 1 : X_t \in \mathscr{S}_k \cup \{y\}\} = \rho$.



By the reversibility of random walk, $P_y(X_\rho = z) = P_z(X_\rho = y)$, thus

$$h_y(\mathscr{S}_k, z) = G_k(y, y)P_z(X_\rho = y).$$

Now, let $m = m_{k,\Delta} = \left\lfloor k - \frac{\Delta}{2} \right\rfloor$.

For sufficiently large Δ we have $y \notin \mathscr{S}_m$, indeed $\|y\| < j + 1$ so we just have to take Δ such that $j + 1 < \left\lfloor k - \frac{k-j}{2} \right\rfloor = \left\lfloor \frac{k+j}{2} \right\rfloor$ $\Big($an example, since $j + 1 < \frac{k+j}{2} \Leftrightarrow j + 2 < k$, is to take $k > j + 2\Big)$.

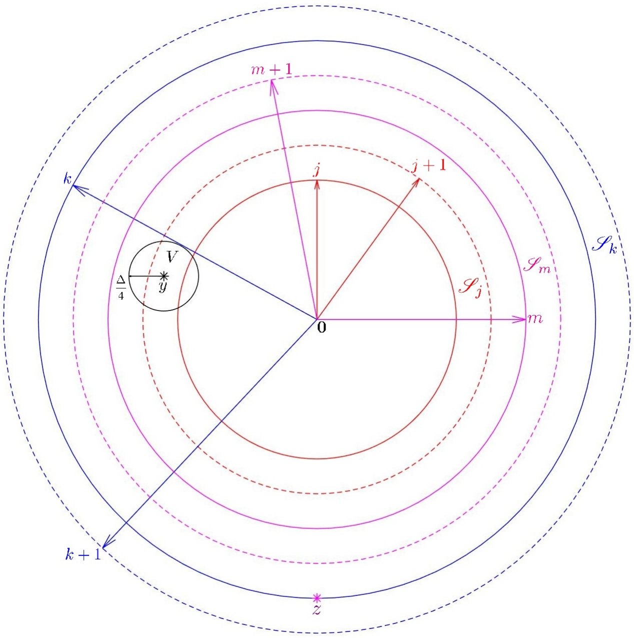
$y \notin \mathscr{S}_m$ so

$$P_z(X_\rho = y) = P_z(X_\rho = y|T_m < T_k)P_z(T_m < T_k) + \underbrace{P_z(X_\rho = y|T_k < T_m)}_{=0} P_z(T_k < T_m)$$

$$\leq P_z(T_m < T_k) \sup_{x \in \mathscr{S}_m} P_x(X_\rho = y) \underset{\substack{\uparrow \\ (1)}}{\leq} \sup_{x \in \mathscr{S}_m} P_x(X_\rho = y)$$

$$\implies h_y(\mathscr{S}_k, z) \leq \frac{J}{\Delta} G_k(y, y) \sup_{x \in \mathscr{S}_m} P_x(X_\rho = y) = \frac{J}{\Delta} \sup_{x \in \mathscr{S}_m} G_k(x, y)$$

where in the last equality we have exploited Green's function properties applied to the set $\mathfrak{B}_k$.



Here and throughout the remainder of this proof we use J for an absolute constant that may change from line to line.

Therefore, we want to show that

$$\sup_{x \in \mathscr{S}_m} G_k(x, y) \leq \frac{J}{\Delta^{d-2}}.$$

Let $V = V_{y,\Delta} = \left\{ w \in \mathbb{Z}^d : \|w - y\| \leq \frac{\Delta}{4} \right\}$.

If $x \in \mathscr{S}_m$ and Δ is large enough, the function $h(w) = G_k(x, w)$ is harmonic in $\left\{ w \in \mathbb{Z}^d : \|w - y\| < \frac{3\Delta}{8} \right\} \supseteq V \ni y$. For the Harnack principle (Theorem 4) then there exists $J < \infty$ such that

$$G_k(x, y) \leq J G_k(x, w) \qquad \forall\ w \in V, x \in \mathscr{S}_m$$

$$\implies \sup_{x \in \mathscr{S}_m} G_k(x, y) \leq J \sup_{x \in \mathscr{S}_m} G_k(x, w) \qquad \forall\ w \in V,$$

thus

$$\sup_{x \in \mathscr{S}_m} G_k(x, y) \leq J \frac{1}{\Delta^d} \sup_{x \in \mathscr{S}_m} \sum_{w \in V} G_k(x, w) = \frac{J}{\Delta^d} \sup_{x \in \mathscr{S}_m} E_x \left[\sum_{w \in V} \sum_{t=0}^{\tau_k - 1} \mathbb{1}_{\{X_t = w\}} \right]$$

and, if we call $Y = \sum_{t=0}^{\tau_k - 1} \mathbb{1}_{\{X_t \in V\}}$ the number of visits in V before leaving $\mathfrak{B}_k$, this is equal to

$$J \frac{1}{\Delta^d} \sup_{x \in \mathscr{S}_m} E_x[Y] \leq J \frac{1}{\Delta^d} \sup_{w \in V} E_w[Y].$$

Therefore, it sufficies to show the following inequality

$$\sup_{w \in V} E_w(Y) \leq J\Delta^2.$$

If $X_0 = w \in V$, $X_t = w + W_1 + \cdots + W_t$ with W_i random vectors on $\mathbb{Z}^d$ i.i.d. with mean $\boldsymbol{\mu} = \mathbf{0}$ and covariance matrix I_d ($d \times d$ identity matrix), we can use the central limit theorem to say that for time Δ^2 sufficiently large, there exists $0 < \varepsilon < P\left(\|Z\| \geq \frac{9}{4}\right)$ such that

$$|P\left(\|R\| \geq k\right) - P_w\left(\|X_{\Delta^2}\| \geq k\right)| < \varepsilon \qquad \text{with } R \sim \mathcal{N}(w, \Delta^2 I_d),$$

but $|P\left(\|R\| \geq k\right) - P_w\left(\|X_{\Delta^2}\| \geq k\right)| \geq P\left(\|R\| \geq k\right) - P_w\left(\|X_{\Delta^2}\| \geq k\right)$

$$\begin{aligned} \implies P_w\left(\|X_{\Delta^2}\| \geq k\right) &\geq P\left(\|R\| \geq k\right) - \varepsilon \\ &= P\left(\|w + \Delta Z\| \geq k\right) - \varepsilon \qquad \text{with } Z \sim \mathcal{N}(\mathbf{0}, I_d). \end{aligned} \tag{2.10}$$

Since $\|w\| - \|y\| \leq \|w - y\| \leq \frac{k-j}{4}$, we obtain $\|w\| \leq \frac{k-j}{4} + \|y\| < \frac{k-j}{4} + j + 1$ and $\|w + \Delta Z\| \geq \|\Delta Z\| - \|w\| > \Delta^2 \|Z\| - \frac{\Delta}{4} + j + 1$, therefore

$$\begin{aligned} P\left(\|w + \Delta Z\| \geq k\right) &\geq P\left(\Delta^2 \|Z\| - \frac{\Delta}{4} + j + 1 \geq k\right) = P\left(\|Z\| \geq \frac{5}{4\Delta} + \frac{1}{\Delta^2}\right) \\ &\geq P\left(\|Z\| \geq \frac{5}{4} + 1\right) = P\left(\|Z\| \geq \frac{9}{4}\right) \end{aligned}$$

Thus, from (2.10) we have

$$P_w(\|X_{\Delta^2}\| \geq k) \geq P\left(\|Z\| \geq \frac{9}{4}\right) - \varepsilon > 0.$$

Defining $\theta = P\left(\|Z\| \geq \frac{9}{4}\right) - \varepsilon$,

$$\begin{aligned} P_w(\tau_k \leq \Delta^2) &= P_w(\exists\, t \in \{1, \dots \Delta^2\} : \|X_t\| \geq k) \\ &= P_w\left(\bigcup_{t=1}^{\Delta^2} \{\|X_t\| \geq k\}\right) \geq P(\|X_{\Delta^2}\| \geq k) \\ &\implies P_w(\tau_k \leq \Delta^2) \geq \theta \end{aligned}$$

with $\theta > 0$ which does not depend on k.

On a time interval of length Δ^2, at most Δ^2 sites in V can be visited; each time that we restart the process at some $w \in V$, the above bound on the exit probability holds. Formally,

$$P_w(\tau_k \leq \Delta^2) \geq \theta \;\Rightarrow\; P_w(\tau_k > \Delta^2) \leq 1 - \theta,$$

$$\forall\, w \in V \quad P_w(Y > l\Delta^2) = P_w\left(\sum_{t=0}^{\tau_k - 1} \mathbb{1}_{\{X_t \in V\}} > l\Delta^2\right)$$

$$\underset{\mathfrak{B}_k \supseteq V}{\underset{\uparrow}{\leq}} P_w\left(\sum_{t=0}^{\tau_k - 1} \mathbb{1}_{\{X_t \in \mathfrak{B}_k\}} > l\Delta^2\right) = P_w(\tau_k > l\Delta^2)$$

$$\leq \sum_{x_1,\ldots,x_{l-1} \in \mathfrak{B}_k} P_w(X_{\Delta^2} = x_1, \ldots, X_{(l-1)\Delta^2} = x_{l-1}) \underbrace{P_w(\tau_k > l\Delta^2 | X_{\Delta^2} = x_1, \ldots, X_{(l-1)\Delta^2} = x_{l-1})}_{\underbrace{\underset{(1-\theta)}{\overset{|\wedge}{P_{X_{l-1}}(\tau_k > \Delta^2)}} \cdots \underset{(1-\theta)}{\overset{|\wedge}{P_{X_1}(\tau_k > \Delta^2)}} \underset{(1-\theta)}{\overset{|\wedge}{P_w(\tau_k > \Delta^2)}}}_{l \text{ times}}},$$

therefore

$$\sup_{w \in V} P_w(Y > l\Delta^2) \leq (1-\theta)^l \qquad \forall\, l \in \mathbb{Z}^+,$$

namely $\dfrac{Y}{\Delta^2}$ has a geometrical tail.

This, since $E(Y) = \sum_{l=0}^{\infty} P(Y \geq l) \leq \Delta^2 \sum_{l=0}^{\infty} P(Y \geq l\Delta^2)$ because $P(Y \geq l)$ does not increase in l, implies

$$\sup_{w \in V} E_w(Y) \leq J\Delta^2$$

and completes the proof. □

2.2 The lower bound

The aim of this paragraph is to prove that

$$\forall\, \varepsilon > 0 \quad \mathfrak{B}_{n(1-\varepsilon)} \subseteq A\left(\omega_d n^d\right) \qquad \text{a.s. for all sufficiently large } n. \tag{2.11}$$

It sufficies to show that

$$\forall\, \varepsilon > 0 \quad \mathfrak{B}_{n(1-\varepsilon)} \subseteq A\left(\omega_d n^d (1+\varepsilon)\right) \qquad \text{a.s. for all sufficiently large } n;$$

indeed, in that case we would have, for $\varepsilon \to 0$, $\mathfrak{B}_n \subseteq A\left(\omega_d n^d\right)$, hence, for every $\tilde{\varepsilon} > 0$, $\mathfrak{B}_{n(1-\tilde{\varepsilon})} \subseteq \mathfrak{B}_n \subseteq A\left(\omega_d n^d\right)$.

Let each independent random walk X_t^i, constituent of the internal DLA cluster, evolve forever, even after it has left the cluster of occupied sites.

Consider the random times

- ⋄ $\sigma^i = \min\{t : X_t^i \notin A(i-1)\}$ the time it takes the i-th particle to leave the occupied cluster,

◇ $\tau^i_{\mathbb{Z}^d\setminus\{z\}} = \min\{t : X^i_t = z\}$ the time it takes the i-th walk to hit the site z,

◇ $\tau^i_n = \min\{t : X^i_t \notin \mathfrak{B}_n\}$ the time it takes the i-th walk to leave $\mathfrak{B}_n$.

Let $F_z(n) = \{\sigma^i < \tau^i_{\mathbb{Z}^d\setminus\{z\}}\ \forall\ i \leq n\}$ be the event for which the site z does not belong to the cluster $A(n)$

$\Longrightarrow$ for each set Λ, $\{\Lambda \nsubseteq A(n)\} = \bigcup\limits_{z\in\Lambda} F_z(n)$

$\xRightarrow[\text{Borel-Cantelli}]{}$ a sufficient condition for the thesis is that

$$\sum_n \sum_{z\in\mathfrak{B}_{n(1-\varepsilon)}} P\left(F_z\left(\omega_d n^d(1+\varepsilon)\right)\right) < \infty, \tag{2.12}$$

indeed

$$\sum_n \sum_{z\in\mathfrak{B}_{n(1-\varepsilon)}} P\left(F_z\left(\omega_d n^d(1+\varepsilon)\right)\right) = \sum_n P\left(\mathfrak{B}_{n(1-\varepsilon)} \nsubseteq A\left(\omega_d n^d(1+\varepsilon)\right)\right),$$

hence Borel-Cantelli lemma would imply that

$$P\left(\limsup_{n\to\infty}\left\{\mathfrak{B}_{n(1-\varepsilon)} \nsubseteq A\left(\omega_d n^d(1+\varepsilon)\right)\right\}\right) = 0$$

but

$$\begin{aligned}
\limsup_{n\to\infty}\left\{\mathfrak{B}_{n(1-\varepsilon)} \nsubseteq A\left(\omega_d n^d(1+\varepsilon)\right)\right\} &= \bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}\left\{\mathfrak{B}_{m(1-\varepsilon)} \nsubseteq A\left(\omega_d m^d(1+\varepsilon)\right)\right\}\\
&= \bigcap_{n=1}^{\infty}\left\{\exists\ m > n : \mathfrak{B}_{m(1-\varepsilon)} \nsubseteq A\left(\omega_d m^d(1+\varepsilon)\right)\right\}\\
&= \left\{\forall\ n\ \exists\ n_0 > n : \mathfrak{B}_{n_0(1-\varepsilon)} \nsubseteq A\left(\omega_d n_0^d(1+\varepsilon)\right)\right\}
\end{aligned}$$

$$\Longrightarrow \quad \underbrace{P\left(\left\{\forall\ n\ \exists\ n_0 > n : \mathfrak{B}_{n_0(1-\varepsilon)} \nsubseteq A\left(\omega_d n_0^d(1+\varepsilon)\right)\right\}^c\right)}_{P\left(\left\{\exists\ n\ \forall\ n_0>n:\mathfrak{B}_{n_0(1-\varepsilon)}\subseteq A\left(\omega_d n_0^d(1+\varepsilon)\right)\right\}\right)} = 1,$$

which is our thesis.

Therefore, let us prove that

$$\sum_{n=0}^{\infty} \sum_{z\in\mathfrak{B}_{n(1-\varepsilon)}} P\left(F_z\left(\omega_d n^d(1+\varepsilon)\right)\right) < \infty. \tag{2.13}$$

Fix n and $z \in \mathfrak{B}_{n(1-\varepsilon)}$.

Consider the random variables

◇ $N = \sum\limits_{i\leq\omega_d n^d(1+\varepsilon)} \mathbb{1}_{\{\tau^i_{\mathbb{Z}^d\setminus\{z\}} < \sigma^i\}}$ the number of particles that visit z before stopping (i.e. before leaving the occupied cluster),

◇ $M = \sum_{i \leq \omega_d n^d (1+\varepsilon)} \mathbb{1}_{\{\tau^i_{\mathbb{Z}^d \setminus \{z\}} < \tau^i_n\}}$ the number of walks that visit z before leaving $\mathfrak{B}_n$,

◇ $L = \sum_{i \leq \omega_d n^d (1+\varepsilon)} \mathbb{1}_{\{\sigma^i \leq \tau^i_{\mathbb{Z}^d \setminus \{z\}} < \tau^i_n\}}$ the number of walks that visit z before leaving $\mathfrak{B}_n$ but after the related particle stops.

Clearly, $M \geq L$ and $M - L \leq N$. Then for any a

$$\underbrace{P\left(F_z\left(\omega_d n^d (1+\varepsilon)\right)\right)}_{\left\{z \notin A\left(\omega_d n^d (1+\varepsilon)\right)\right\}} = P(N=0) \underset{\substack{\uparrow \\ N=0 \Rightarrow M-L=0 \Leftrightarrow M=L \Rightarrow M \leq a \text{ or } L \geq a}}{\leq} P(M \leq a \text{ or } L \geq a) \leq P(M \leq a) + P(L \geq a)$$

Now,

$$\begin{aligned} E(M) &= E\left(\sum_{i \leq \omega_d n^d (1+\varepsilon)} \mathbb{1}_{\{\tau^i_{\mathbb{Z}^d \setminus \{z\}} < \tau^i_n\}}\right) = \sum_{i \leq \omega_d n^d (1+\varepsilon)} E\left(\mathbb{1}_{\{\tau^i_{\mathbb{Z}^d \setminus \{z\}} < \tau^i_n\}}\right) \\ &= \sum_{i \leq \omega_d n^d (1+\varepsilon)} P\left(\tau^i_{\mathbb{Z}^d \setminus \{z\}} < \tau^i_n\right) \underset{\substack{\uparrow \\ \text{the random walks are i.i.d.}}}{=} \lfloor \omega_d n^d (1+\varepsilon) \rfloor P\left(\tau_{\mathbb{Z}^d \setminus \{z\}} < \tau_n\right) \end{aligned}$$

with $\tau_{\mathbb{Z}^d \setminus \{z\}} = \min\{t : X^1_t = z\}, \ \tau_n = \min\{t : X^1_t \notin \mathfrak{B}_n\}$.

Notice that the fact that the random walks are i.i.d. implies that also the addends in M are i.i.d., but does not imply that the addends of L are the same: indeed, they are neither identically distributed nor independent.

In L, just the addends which correspond to indexes i such that $X^i_{\sigma^i} \in \mathfrak{B}_n$ contribute to the sum and for any $y \in \mathfrak{B}_n$ there is at most one index i with $X^i_{\sigma^i} = y$ and the corresponding random walk $\left(X^i_t\right)_{t > \tau^i_{\mathbb{Z}^d \setminus \{y\}}}$, with $\tau^i_{\mathbb{Z}^d \setminus \{y\}} = \min\left\{t : X^i_t = y\right\}$, is then independent.

Therefore, to avoid dependence in L, we enlarge the set of indexes to all $\mathfrak{B}_n$ and we call

$$\tilde{L} = \sum_{y \in \mathfrak{B}_n} \mathbb{1}_{\{\tau^{(y)}_{\mathbb{Z}^d \setminus \{z\}} < \tau^{(y)}_n\}}$$

with

$$\begin{aligned} \tau^{(y)}_{\mathbb{Z}^d \setminus \{z\}} &= \min\left\{t : X^{(y)}_t = z\right\}, \\ \tau^{(y)}_n &= \min\left\{t : X^{(y)}_t \notin \mathfrak{B}_n\right\} \end{aligned}$$

where $\{X^{(y)}_t\}_{t \in \mathbb{N}_0}$ is a simple symmetric random walk on $\mathbb{Z}^d$ starting at y which eventually coincides with $\{X^i_t\}_{t \in \mathbb{N}_0}$ if there exists $i \leq \omega_d n^d (1+\varepsilon)$ such that $X^i_{\sigma^i} = y$.

$$E(\tilde{L}) = \sum_{y \in \mathfrak{B}_n} E\left(\mathbb{1}_{\{\tau^{(y)}_{\mathbb{Z}^d \setminus \{z\}} < \tau^{(y)}_n\}}\right) = \sum_{y \in \mathfrak{B}_n} P\left(\tau^{(y)}_{\mathbb{Z}^d \setminus \{z\}} < \tau^{(y)}_n\right) = \sum_{y \in \mathfrak{B}_n} P_y\left(\tau_{\mathbb{Z}^d \setminus \{z\}} < \tau_n\right);$$

$L \leq \tilde{L}$ (a.s.), hence $P(L \geq a) \leq P(\tilde{L} \geq a)$

$$\implies \quad P\left(F_z\left(\omega_d n^d(1+\varepsilon)\right)\right) \leq P(\tilde{L} > a) + P\left(M \leq a\right). \tag{2.14}$$

We will now prove the following fact.

For fixed $\varepsilon > 0$ and n sufficiently large,

$$\begin{aligned} P\left(\tilde{L} \geq \left(1+\frac{\varepsilon}{4}\right) E[\tilde{L}]\right) &\leq e^{-c_d n}, \\ P\left(M \geq \left(1+\frac{\varepsilon}{4}\right) E[\tilde{L}]\right) &\leq e^{-c_d n} \end{aligned} \tag{2.15}$$

$\forall\ z \in \mathfrak{B}_{n(1-\varepsilon)}$ and appropriate costants $c_d = c_d(\varepsilon) > 0$.

Indeed, for lemma 3 we have

$$P_y(\tau_{\mathbb{Z}^d \setminus \{z\}} < \tau_n) = \frac{G_n(y,z)}{G_n(z,z)}, \qquad P(\tau_{\mathbb{Z}^d \setminus \{z\}} < \tau_n) = \frac{G_n(\mathbf{0},z)}{G_n(z,z)},$$

where we are using P to indicate the measure $P_{\mathbf{0}}$, indeed we know that every random walk starts from the origin.

$$\begin{aligned} \implies E(M) = \lfloor \omega_d n^d(1+\varepsilon) \rfloor \frac{G_n(\mathbf{0},z)}{G_n(z,z)} &\underset{\substack{\uparrow \\ \text{Lemma 7}}}{\geq} \left(1+\frac{\varepsilon}{2}\right) \sum_{y \in \mathfrak{B}_n} \frac{G_n(y,z)}{G_n(z,z)} \\ = \left(1+\frac{\varepsilon}{2}\right) \sum_{y \in \mathfrak{B}_n} P_y(\tau_{\mathbb{Z} \setminus \{z\}} < \tau_n) &= \left(1+\frac{\varepsilon}{2}\right) E(\tilde{L}) \end{aligned} \tag{2.16}$$

$$E(\tilde{L}) = \sum_{y \in \mathfrak{B}_n} P_y(\tau_{\mathbb{Z}^d \setminus \{z\}} < \tau_n) = \sum_{y \in \mathfrak{B}_n} \frac{G_n(y,z)}{G_n(z,z)} \underset{\substack{\uparrow \\ (2.7)}}{=} \frac{E_z(\tau_n)}{G_n(z,z)} \underset{\substack{\uparrow \\ (2.6)}}{\geq} \frac{E_z(\tau_n)}{G_{2n}(\mathbf{0},\mathbf{0})} \underset{\substack{\uparrow \\ \text{Lemma 2}}}{\geq} \frac{n^2 - \|z\|^2}{G_{2n}(\mathbf{0},\mathbf{0})}$$

and we can prove that

$$E(\tilde{L}) \geq \begin{cases} \beta_2 \dfrac{n^2}{\log n} & \text{if } d = 2 \\ \beta_d n^2 & \text{if } d \geq 3 \end{cases} \tag{2.17}$$

for every $z \in \mathfrak{B}_{n(1-\varepsilon)}$, eventually in n, for a suitable constant $\beta_d = \beta_d(\varepsilon) > 0$.

Indeed,

- if $d = 2$

$$\frac{n^2 - \|z^2\|}{G_{2n}(\mathbf{0},\mathbf{0})} \overset{\substack{(2.3) \\ \downarrow}}{=} \frac{n^2 - \|z\|^2}{\frac{2}{\pi}\log 2n + o(1) + O\left(\frac{1}{n}\right)} \geq \frac{n^2 - \|z\|^2}{\log n + c} \geq \frac{n^2}{\tilde{c}\log n}$$

for n large enough with $c, \tilde{c} < \infty$ constants, since $\log n + c < \tilde{c}\log n \iff \tilde{c} > \frac{\log n + c}{\log n}$ which is continuous and decreasing for n sufficiently large.

- if $d \geq 3$

$$\frac{n^2 - \|z^2\|}{G_{2n}(\mathbf{0},\mathbf{0})} \overset{(2.5)}{\underset{\downarrow}{=}} \frac{n^2 - \|z\|^2}{\frac{2}{(d-2)\omega_d}(1 - n^{2-d}) + O(1)} \geq \frac{n^2}{c'_d},$$

hence we have (2.17).

For (2.16), the same lower bound holds for $E(M)$.

M and $\tilde{L}$ are sums of independent indicator random variables, so for Lemma 8 with $\gamma = \frac{1}{3}$

$$\begin{cases} P\left(\tilde{L} \geq E(\tilde{L}) + E(\tilde{L})^{5/6}\right) \leq 2e^{\frac{1}{4}E(\tilde{L})^{2/3}} \leq e^{-c_d n}, \\ P\left(M \leq E(M) - E(M)^{5/6}\right) \leq 2e^{\frac{1}{4}E(M)^{2/3}} \leq e^{-c_d n} \end{cases} \quad \forall\ z \in \mathfrak{B}_{n(1-\varepsilon)} \qquad (2.18)$$

for n sufficiently large, where $c_d > 0$ is constant such that the last inequalities hold, indeed $2e^{\frac{1}{4}E(M)^{2/3}} \leq e^{-c_d n} \ \Leftrightarrow\ \log 2 - \dfrac{1}{4}E(M)^{2/3} \leq -c_d n \ \Leftrightarrow$
$c_d \leq \dfrac{1}{n}\left(\dfrac{1}{4}E(M)^{2/3} - \log 2\right)$ and the same is true for $2e^{\frac{1}{4}E(\tilde{L})^{2/3}}$,

$$\Longrightarrow\ c_d \leq \min\left\{\frac{1}{n}\left(\frac{1}{4}E(M)^{2/3} - \log 2\right), \frac{1}{n}\left(\frac{1}{4}E(\tilde{L})^{2/3} - \log 2\right)\right\}$$

and $c_d > 0$ because $\dfrac{1}{4}E(\tilde{L})^{2/3} - \log 2 > 0 \Leftrightarrow E(\tilde{L}) > (4\log 2)^{3/2}$ and, for (2.17), we can take n large enough such that this is satisfied, and the same for $E(M)$.

Now, we can take n large enough such that

$$P\left(\tilde{L} \geq \left(1 + \frac{\varepsilon}{4}\right)E(\tilde{L})\right) \leq P\left(\tilde{L} \geq E(\tilde{L}) + E(\tilde{L})^{5/6}\right),$$

so that we get the first part of (2.15); indeed, this happens when
$E(\tilde{L}) + \dfrac{\varepsilon}{4}E(\tilde{L}) \geq E(\tilde{L}) + E(\tilde{L})^{5/6}$, that is true if and only if $E(\tilde{L}) \geq \left(\dfrac{4}{\varepsilon}\right)^6$, hence we just have to take n sufficiently large so that this is satisfied.

We can also take n big enough in order to have

$$P\left(M \leq \left(1 + \frac{\varepsilon}{4}\right)E(\tilde{L})\right) \leq P\left(M \leq E(M) - E(M)^{5/6}\right) \leq 2e^{\frac{1}{4}E(M)^{2/3}},$$

so we can fully prove (2.15); indeed, that is true when
$E(M) - E(M)^{5/6} \geq \left(1 + \dfrac{\varepsilon}{4}\right)E(\tilde{L})$, but

$$\begin{aligned} E(M) - E(M)^{5/6} &= E(M)^{5/6}\left(E(M)^{1/6} - 1\right) \\ &\underset{\substack{\uparrow \\ (2.16)}}{\geq} \left(1 + \frac{1}{\varepsilon}\right)^{5/6} E(\tilde{L})^{5/6}\left[\left(1 + \frac{\varepsilon}{2}\right)^{1/6} E(\tilde{L})^{1/6} - 1\right] \\ &= \left(1 + \frac{1}{\varepsilon}\right)E(\tilde{L}) - \left(1 + \frac{1}{\varepsilon}\right)^{5/6} E(\tilde{L})^{5/6} \end{aligned}$$

and we can take n large enough to have $\left(1+\frac{1}{\varepsilon}\right)^{5/6} E(\tilde{L})^{5/6} \leq \frac{\varepsilon}{4} E(\tilde{L})$
$\Leftrightarrow 0 \leq E(\tilde{L})^{5/6} \left[E(\tilde{L})^{1/6}\frac{\varepsilon}{4} - \left(1+\frac{\varepsilon}{2}\right)^{5/6}\right] \Leftrightarrow E(\tilde{L}) \geq \frac{4}{\varepsilon}\left(1+\frac{\varepsilon}{2}\right)^{5/6}$, hence to have
$E(M) - E(M)^{5/6} \geq \left(1+\frac{1}{\varepsilon}\right) E(\tilde{L}) - \left(1+\frac{1}{\varepsilon}\right)^{5/6} E(\tilde{L})^{5/6} \geq \left(1+\frac{1}{\varepsilon}\right) E(\tilde{L}) - \frac{\varepsilon}{4} E(\tilde{L})$.
Now that we have shown (2.15), it is easy to conclude the proof of (2.13); indeed, for $a = \left(1+\frac{\varepsilon}{4}\right) E(\tilde{L})$ in (2.14),

$$P\left(F_z\left(\omega_d n^d(1+\varepsilon)\right)\right) \leq P\left(\tilde{L} \geq \left(1+\frac{\varepsilon}{4}\right) E(\tilde{L})\right) + P\left(M \leq \left(1+\frac{\varepsilon}{4}\right) E(\tilde{L})\right) \underset{\underset{(2.15)}{\uparrow}}{\leq} 2e^{-c_d n}$$

for $n > n_0$ for a suitable $n_0 \in \mathbb{N}$. Thus,

$$\sum_{n \geq n_0} \sum_{z \in \mathfrak{B}_{n(1-\varepsilon)}} P\left(F_z\left(\omega_d n^d(1+\varepsilon)\right)\right) \leq \sum_{n \geq n_0} \sum_{z \in \mathfrak{B}_{n(1-\varepsilon)}} 2e^{-c_d n} \leq \sum_{n \geq n_0} 2\omega_d n^d e^{-c_d n} < \infty$$

which is (2.12).

2.3 The upper bound

The aim of this paragraph is to prove that

$$\forall\, \varepsilon > 0 \quad A\left(\omega_d n^d\right) \subseteq \mathfrak{B}_{n(1+\varepsilon)} \qquad \text{a.s. for all sufficiently large } n. \tag{2.19}$$

The general strategy is to use the lower bound (2.11) and Lemma 9 to prove that the rate of growth of the number of occupied sites in each shell $\mathscr{S}_k$ outside of $\mathfrak{B}_n$ can be controlled.
It sufficies to prove that

$$\forall\, \varepsilon > 0 \quad P\left(A\left(\omega_d n^d\right) \subseteq \mathfrak{B}_{n(1+K\varepsilon^{1/d})}\right) \geq 1-\varepsilon \qquad \text{for all sufficiently large } n, \tag{2.20}$$

for a suitable $K = K_d < \infty$; indeed, in that case we would have, for $\varepsilon \to 0$, a.s. $A\left(\omega_d n^d\right) \subseteq \mathfrak{B}_n$, therefore, for every $\tilde{\varepsilon} > 0$, $A\left(\omega_d n^d\right) \subseteq \mathfrak{B}_n \subseteq \mathfrak{B}_{n(1+\tilde{\varepsilon})}$.
For the lower bound,

$$P\left(\mathfrak{B}_{n(1-\varepsilon)} \subseteq A\left(\omega_d n^d\right) \text{ for } n \geq n_0\right) \geq 1-\varepsilon \qquad \text{for } n_0 \text{ big enough}, 0 < \varepsilon < 1.$$

Since most of the first $\lfloor \omega_d n^d \rfloor$ particles stop to fill out $\mathfrak{B}_{n(1-\varepsilon)} \subseteq \mathfrak{B}_n$ with only a small portion left over,

$$P\left(\left|A\left(\omega_d n^d\right) \cap \mathfrak{B}_n^c\right| < K_0 \varepsilon n^d \quad \text{for } n > n_0\right) \geq 1-\varepsilon \tag{2.21}$$

for a suitable constant $K_0 = K_{0,d} < \infty$, indeed

$$\left|A\left(\omega_d n^d\right) \cap \mathfrak{B}_n^c\right| = \left|A\left(\omega_d n^d\right) \setminus \mathfrak{B}_n\right| \underset{\substack{\uparrow \\ A\left(\omega_d n^d\right) \setminus \mathfrak{B}_n \subseteq A\left(\omega_d n^d\right) \setminus \mathfrak{B}_{n(1-\varepsilon)}}}{\leq} \left|A\left(\omega_d n^d\right) \setminus \mathfrak{B}_{n(1-\varepsilon)}\right|$$

and with probability $1-\varepsilon$ for $n \geq n_0$ this is equal to

$$\left|A\left(\omega_d n^d\right)\right| - \left|\mathfrak{B}_{n(1-\varepsilon)}\right| \text{ and } \omega_d n^d - \omega_d n^d(1-\varepsilon)^d = \omega_d n^d \varepsilon \left[\frac{1-(1-\varepsilon)^d}{\varepsilon}\right] \leq \omega_d d\varepsilon n^d,$$

therefore we have (2.21).

Let us rename as Y^j the particles X^{i_j} that exit $\mathfrak{B}_n$ during the interval of time steps from 0 to $\omega_d n^d$ and consider the process $\tilde{A}(j) = A(i_j)$ (i.e. the time is now rescaled in terms of the particles leaving $\mathfrak{B}_n$).

Choose $k_0 = \lfloor n(1+\varepsilon^{1/d}) \rfloor + 1$ and introduce

$$Z_k(j) = |\tilde{A}(j) \cap \mathscr{S}_{k_0+k}|,$$
$$\mu_k(j) = E(Z_k(j))$$

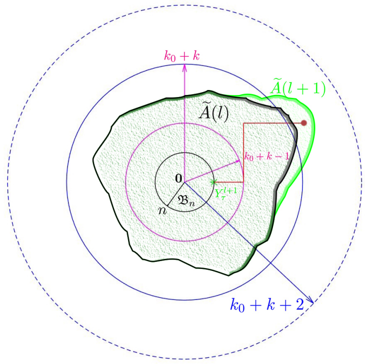
for $k \geq 1$.

$\mu_k(j)$ is the expected number of occupied vertices in $\mathscr{S}_{k_0+k}$ after the particles $Y^1, \ldots, Y^j$ have stopped.

We will control the growth of $\tilde{A}(j)$ by estimating $\mu_k(j)$. Clearly, $\mu_1(j) \leq j$ and $\mu_k(0) = 0$, while for the general case we will prove the following result:

$$\forall\, j,k \quad \mu_k(j) \leq n^{d-1}\left[J_1 \frac{j}{k} \varepsilon^{(1-d)/d} n^{1-d}\right]^k \qquad \text{with } J_1 = J_{1,d} < \infty. \tag{2.22}$$

The fact the proof exploits is that the rate at which the particles fill a shell $\mathscr{S}_{k_0+k}$ is restricted by the rate at which particles exit $\mathfrak{B}_n$ and penetrate the preceding shell $\mathscr{S}_{k_0+k-1}$. For point (2) of Lemma 9, the probability of this penetration is bounded by the portion of $\mathscr{S}_{k_0+k-1}$ which is occupied. In this way, we will get a recursion formula for μ_k in terms of μ_{k-1}, whose iteration will give (2.22).



First, we condition on the time $\tau = \tau_n^{i_{l+1}}$ at which Y^{l+1} exits $\mathfrak{B}_n$, getting

$$\mu_k(l+1) - \mu_k(l) = E\left(Z_k(l+1)\right) - E\left(Z_k(l)\right) = E\underbrace{\left(\left|\tilde A(l+1) \cap \mathscr{S}_{k_0+k}\right| - \left|\tilde A(l) \cap \mathscr{S}_{k_0+k}\right|\right)}_{\substack{\left|\left(\tilde A(l+1)\cap\mathscr{S}_{k_0+k}\right)\setminus\tilde A(l)\right| \\ \| \\ \mathbb{1}_{\left\{Y^{l+1}\text{ exits }\tilde A(l)\text{ through a vertex in }\mathscr{S}_{k_0+k}\right\}}}} =$$

$$= E\left[h_{Y_\tau^{l+1}}\left(\tilde A(l)^c, \mathscr{S}_{k_0+k}\right)\right].$$

Since each walk stopping in $\mathscr{S}_{k_0+k}$ has to stay in the cluster while it hits the preceding shell $\mathscr{S}_{k_0+k-1}$, we have that the event for which the walk starting from Y_τ^{l+1} exits $\tilde A(l)$ through a vertex in $\mathscr{S}_{k_0+k}$ is contained in the event according to which the walk starting from Y_τ^{l+1} hits for the first time $\mathscr{S}_{k_0+k-1}$ in a site in $\tilde A(l)$, therefore

$$E\left[h_{Y_\tau^{l+1}}\left(\tilde A(l)^c, \mathscr{S}_{k_0+k}\right)\right] \leq E\left[h_{Y_\tau^{l+1}}\left(\mathscr{S}_{k_0+k-1}, \tilde A(l)\right)\right] \leq \max_{y\in\mathscr{S}_n} E\left[h_y\left(\mathscr{S}_{k_0+k-1}, \tilde A(l)\right)\right]$$

and for Lemma 9 with $j \to n$, $k \to k_0+k-1$, $\Delta \to k_0+k-n-1 = \lfloor n(1+\varepsilon^{1/d})\rfloor +$ $+1+k-n-1 = \lfloor n\varepsilon^{1/d}\rfloor + k \geq n\varepsilon^{1/d}$, we have that for each $y \in \mathscr{S}_n$

$$\mu_k(l+1) - \mu_k(l) \leq \max_{y\in\mathscr{S}_n} E\left[h_y\left(\mathscr{S}_{k_0+k-1}, \tilde A(l)\right)\right] \leq J\mu_{k-1}(l)\left(\frac{1}{n\varepsilon^{1/d}}\right)^{d-1}$$

and summing over $l = 0, \ldots, j-1$

$$\mu_k(j) - \underbrace{\mu_k(0)}_{0} \leq J\left(\frac{1}{n\varepsilon^{1/d}}\right)^{d-1}\sum_{l=1}^{j-1}\mu_{k-1}(l). \tag{2.23}$$

It is easy to prove by induction that for fixed j

$$\mu_k(j) \leq \left[\left(J\frac{1}{n\varepsilon^{1/d}}\right)^{d-1}\right]^{k-1}\frac{j^k}{k!},$$

indeed, for $k = 1$ we have $\mu_1(j) \leq j$ and assuming it true till $k-1$, we get that

$$\mu_k(j) \leq J\left(\frac{1}{n\varepsilon^{1/d}}\right)^{d-1}\sum_{l=1}^{j-1}\mu_{k-1}(l) \leq J\left(\frac{1}{n\varepsilon^{1/d}}\right)^{d-1}\sum_{l=1}^{j-1}\left[J\left(\frac{1}{n\varepsilon^{1/d}}\right)^{d-1}\right]^{k-2}\frac{l^{k-1}}{(k-1)!}$$

$$= \left[J\left(\frac{1}{n\varepsilon^{1/d}}\right)^{d-1}\right]^{k-1}\frac{1}{(k-1)!}\underbrace{\sum_{l=1}^{j-1}l^{k-1}}_{\leq\frac{j^k}{k}}.$$

Using this in (2.23), we obtain

$$\begin{aligned}\mu_k(j) &\leq \left[J\left(\frac{1}{n\varepsilon^{1/d}}\right)^{d-1}\right]^{k-1}\frac{j^k}{k!}\underset{\substack{\uparrow\\ k!\geq k^k e^{-k}}}{\leq} n^{d-1+k-dk}\frac{j^k}{k^k}\varepsilon^{(k-dk)/d}e^k J^{k-1}\underbrace{\varepsilon^{(d-1)/d}}_{\leq 1 \text{ for } 0<\varepsilon<1}\\ &\leq n^{d-1}\left[J_1\frac{j}{k}\varepsilon^{(1-d)/d}n^{1-d}\right]^k\end{aligned}$$

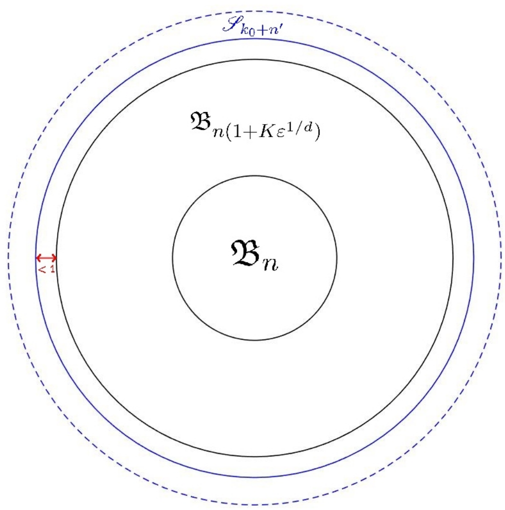
which is (2.22) and where we have set $J_1 = e\tilde{J}$ with $\tilde{J} = \begin{cases}1 & \text{if } J<1\\ J & \text{if } J>1\end{cases}$.

Now we can continue with the proof of (2.20).

Let F be the event in (2.21),

$$F = \left\{\left|A\left(\omega_d n^d\right)\cap\mathfrak{B}_n^c\right| < K_0\varepsilon n^d \quad \text{for } n>n_0\right\}.$$

For $n \geq n_0$, choosing $n' = \left\lceil (K-1)\varepsilon^{1/d}n - 1\right\rceil$ and noticing that
$k_0 + \lceil (K-1)\varepsilon^{1/d}n - 1\rceil = \lfloor n(1+\varepsilon^{1/d})\rfloor + \lceil (K-1)\varepsilon^{1/d}n\rceil$ and
$n(1+\varepsilon^{1/d}) + (K-1)\varepsilon^{1/d}n = n + nK\varepsilon^{1/d}$,



$$
\begin{aligned}
&P\left(A\left(\omega_d n^d\right) \nsubseteq \mathfrak{B}_{n(1+K\varepsilon^{1/d})}, F\right)\\
&= P\underbrace{\left(\left\{A(\omega_d n^d) \nsubseteq \mathfrak{B}_{n(1+K\varepsilon^{1/d})}\right\}, \left\{\left|A\left(\omega_d m^d\right) \cap \mathfrak{B}^c_m\right| < K_0\varepsilon m^d \ \ \forall\, m > n_0\right\}\right)}_{\subseteq\left\{\left|\tilde{A}\left(\left\lfloor K_0\varepsilon n^d\right\rfloor\right)\cap \mathscr{S}_{k_0+\left\lceil (K-1)\varepsilon^{1/d}n-1\right\rceil}\right|\geq 1\right\}}\\
&\leq P\left(Z_{n'}\left(\left\lfloor K_0\varepsilon n^d\right\rfloor\right) \geq 1\right) \leq \sum_{l=0}^{\infty} P\left(Z_{n'}\left(\left\lfloor K_0\varepsilon n^d\right\rfloor\right) \geq l\right)\\
&= E\left(Z_{n'}\left(\left\lfloor K_0\varepsilon n^d\right\rfloor\right)\right) = \mu_{n'}\left(\left\lfloor K_0\varepsilon n^d\right\rfloor\right)
\end{aligned}
$$

and, for (2.22) with $j \to \lfloor K_0\varepsilon n^d \rfloor$, $k \to n'$,

$$
\mu_{n'}\left(\left\lfloor K_0\varepsilon n^d\right\rfloor\right) \leq n^{d-1}\left[J_1 \frac{\left\lfloor K_0\varepsilon n^d\right\rfloor}{\left\lceil (K-1)\varepsilon^{1/d}n-1\right\rceil}\varepsilon^{(1-d)/d}n^{1-d}\right]^{n'} \leq n^{d-1}\left(\frac{J_2}{K}\right)^{n'}
$$

for large n and K and a suitable $J_2 = J_{2,d} < \infty$. Indeed, we have that the last inequality holds if and only if $K\dfrac{\left\lfloor K_0\varepsilon n^d\right\rfloor}{\left\lceil (K-1)\varepsilon^{1/d}n-1\right\rceil}\varepsilon^{(1-d)/d}n^{1-d} \leq \tilde{J}_2$ for a suitable $\tilde{J}_2$ and this is true because

$$
\begin{aligned}
K\frac{\left\lfloor K_0\varepsilon n^d\right\rfloor}{\left\lceil (K-1)\varepsilon^{1/d}n-1\right\rceil}\varepsilon^{(1-d)/d}n^{1-d} &\leq K\frac{K_0\varepsilon n^d}{(K-1)\varepsilon^{1/d}n-1}\varepsilon^{(1-d)/d}n^{1-d}\\
&= K\frac{K_0\varepsilon^{1/d}n}{(K-1)\varepsilon^{1/d}n-1} \xrightarrow{n\to\infty} \frac{KK_0}{K-1}\\
&\leq 2K_0 \qquad \text{for } K \geq 2.
\end{aligned}
$$

For $K > J_2$, $n^{d-1}\left(\dfrac{J_2}{K}\right)^{n'} \overset{n\to\infty}{\longrightarrow} 0$ faster than $e^{-\alpha n}$ for some $\alpha > 0$; indeed,

$$
\begin{aligned}
&n^{d-1}\left(\frac{J_2}{K}\right)^{\left\lceil (K-1)\varepsilon^{1/d}n-1\right\rceil} \leq n^{d-1}\left(\frac{J_2}{K}\right)^{(K-1)\varepsilon^{1/d}n} \qquad \text{for } n \text{ large enough and}\\
&n^{d-1}\left(\frac{J_2}{K}\right)^{(K-1)\varepsilon^{1/d}n} \leq e^{-\alpha n} \iff \underbrace{-\frac{\log\left(n^{d-1}\right)}{n}}_{\xrightarrow[n\to\infty]{}0}\underbrace{-\log\left(\frac{J_2}{K}\right)^{(K-1)\varepsilon^{1/d}}}_{>0 \text{ because } \frac{J_2}{K}<1} \geq \alpha
\end{aligned}
$$

hence there exists $n_1 \in \mathbb{N}$ such that for each $n > n_1$,
$-\dfrac{\log n^{d-1}}{n} - \log\left(\dfrac{J_2}{K}\right)^{(K-1)\varepsilon^{1/d}} > \alpha$ for a fixed $\alpha > 0$.
Thus, for n_1 large enough,

$$
\sum_{n\geq n_1} P\left(A\left(\omega_d n^d\right) \nsubseteq \mathfrak{B}_{n(1+K\varepsilon^{1/d})}, F\right) \leq \sum_{n\geq n_1} e^{-\alpha n} < \infty
$$

$$\xRightarrow[\text{Borel-Cantelli}]{} P \underbrace{\left(\limsup_n \left\{A\left(\omega_d n^d\right) \not\subseteq \mathfrak{B}_{n(1+K\varepsilon^{1/d})}\right\} \cap F\right)}_{\underbrace{\left\{\forall\, n \;\; \exists\, m \geq n : A\left(\omega_d m^d\right) \not\subseteq \mathfrak{B}_{m(1+K\varepsilon^{1/d})}\right\}}_{=M} \cap F} = 0$$

$$\implies \quad P(M \cap F) = 0$$

$$\implies \quad \underbrace{P(M \cup F)}_{\in[0,1]} = P(M) + \underbrace{P(F)}_{\geq 1-\varepsilon} - \underbrace{P(M \cap F)}_{=0}$$

$$\implies \quad P(M) \leq \varepsilon$$

which is (2.20).

2.4 Further observations on internal DLA

We have explicitly covered the cases in which the dimension was at least 2, but the same procedure could be applied when it is 1; however, in this scenario a more detailed analysis is present in literature, for example in [11], where Freedman exploits Friedman's "safety campaign" urn model ([12]) to prove that in this dimension both edges of $A(n)$ (which is, of course, an interval $[a, b]$ with $a, b \in \mathbb{Z}$) have Gaussian fluctuations around $n/2$ with standard deviation $c\sqrt{n}$ for a computable constant c. After the work of Lawler, Bramson and Griffeath in [15] the internal DLA was studied with more perspectives, considering for example systems with multiple sources of particles, where there would be multiple growing balls with center on each source and if the number of particles is large enough they would collide, creating non trivial new interactions and shapes which would be quite complicate. However, once the cluster reaches a size which is much bigger than the maximal intersource distance, these collisions would become more and more insignificant and the same holds for the source the individual particles were generated by, bringing this system back to the case with just one source of particles that we studied in this thesis, hence Theorem 5 is still valid in this scenario.

One more generalization of IDLA is its continuous-time variant, defined in [15], where particles are generated at the origin of $\mathbb{Z}^d$ according to a Poisson clock of rate 1 and each of them performs an independent rate 1 continuous time simple symmetric random walk killed upon visiting an unoccupied site, where it stops. In this new process, several "active" particles can occupy the same vertex. Indicating with $B_t(n)$ the cluster of sites which have been visited by the first n particles at time t, we can show that

$$\lim_{t\to\infty} B_t(n) = A(n)$$

and through this it is possible to prove Theorem 5 also for B_t in dimension $d \geq 3$, namely that for any $\varepsilon > 0$

$$B_{t(1-\varepsilon)} \subseteq B_{\omega_d t^d} \subseteq B_{t(1+\varepsilon)} \qquad \text{a.s. for all sufficiently large } t.$$

Chapter 3

Oil and water model

In this chapter we will study the *Oil and Water* model, an interacting particle system on a graph $G = (V, E)$.
In the initial configuration, for each site, the numbers of oil and water particles are independent and both sampled from the distribution given by the probability measure ν on $\mathbb{N}_0$, and we denote by $\mu = \mu(\nu)$ the expected number of particles at a given vertex. We will call the model with this initial configuration *Oil and Water at density* μ.

Each vertex has an independent Poisson clock of rate 1. Starting from the above configuration, every time the clock of a vertex $x \in V$ rings,

· if there are at least one oil and one water particle on x then the vertex *fires* an oil-water pair, namely each of the two particles chooses randomly and independently a neighboring vertex of x according to the uniform distribution and jumps there (in other words, the oil and water particles both perform an independent step of simple random walk on G);

· if there are no oil-water pairs on x (i.e. x hosts no particles or particles of just one type), then the vertex does not fire; in this case we say that x is *stable.*

Therefore, if the system reaches a configuration where every vertex of G is stable, which we will call *stable configuration*, no vertex fires from that time on and the dynamics stops, thus stable configurations are absorbing states for the dynamics.

There are two possible behaviors for the system: either it is *active*, when each vertex fires infinitely many times, or it *fixates*, when each vertex fires finitely many times (it is easy to check that one vertex fires infinitely many times if and only if all the vertices fire infinitely many times).

The main result of this chapter is the following theorem, which states that, for *any* vertex-transitive graph G, the Oil and Water model does not undergo a phase transition, namely there does not exist a nontrivial critical density μ_c at which the behavior of the system changes between *activity* and *fixation*.

Let $\mathbb{P}_\nu$ be the probability law of the Oil and Water dynamics starting from a configuration of density $\mu = \mu(\nu)$ as above.

Theorem 6 *Let G be an infinite, vertex-transitive graph of finite degree. Then, for any ν with $\mu = \mu(\nu) < \infty$,*

$$\mathbb{P}_\nu(\textit{Oil and Water fixates}) = 1.$$

This is not a common behavior of abelian networks, for instance Activated Random Walks exhibits a phase transition. It has been proven that, when the random walks performed by the moving particles are simple and symmetric, for every underlying vertex-transitive and amenable graph G (G is amenable when $\inf\{|\partial A|/|A| : A \text{ is a non-empty finite subgraph of } G\} = 0$) where the random walk is transient, the critical threshold μ_c belongs to the interval $(0,1)$ and that the same result holds for vertex-transitive and non-amenable graphs (Stauffer & Taggi in [21] and Taggi in [22]). It has also been shown that, on the graph $\mathbb{Z}$, $\mu_c \in \left[\frac{\lambda}{\lambda+1}, 1\right]$, where $\lambda > 0$ is the sleeping rate (Trivellato Rolla & Sidoravicius in [25]) and that, when the graph is $\mathbb{Z}^d$, $\mu_c \leq 1$ and it is non decreasing in λ for $d \geq 2$, too (Shellef in [20]). Other remarkable results are that $\mu_c \xrightarrow{\lambda \to 0} 0$ for vertex-transitive graphs on which the random walk is transient and $\mu_c \xrightarrow{\lambda \to +\infty} 1$ for all vertex-transitive graphs.

Notice that in the extreme case where $\lambda = +\infty$ on a vertex-transitive graph, hence $\mu_c = 1$, when an active particle visits an empty site, it becomes a sleeping particle instantaneously. This case is therefore a system which follows the same evolution rules as the continuous-time version of internal DLA but starts from a configuration distributed on the whole graph as a product of independent Poisson distributions with parameter μ. Therefore, the critical threshold for this dynamics with the same evolution rules as internal DLA is $\mu_c = 1$.

Let us now mathematically formalize the Oil and Water model.

3.1 Mathematical formalization

Let $G = (V, E)$ be an infinite, undirected, vertex-transitive graph of finite degree $\mathbf{d}$ (vertex-transitivity implies that all vertices have the same degree) and let $\mathbf{0} \in V$ be the origin.

We consider the space of configurations

$$\Omega = \mathbb{N}_0^V \times \mathbb{N}_0^V;$$

$\eta = (\eta^o(x), \eta^w(x))_{x \in V} \in \Omega$ is a configuration, where, for every $x \in V$, $\eta^o(x), \eta^w(x) : \Sigma \to \mathbb{N}$ are independent random variables on a probability space $(\Sigma, \mathcal{F}, Q)$, and are also independent with respect to $x \in V$. For each vertex $x \in V$, they represent respectively the number of oil and water particles in x.

Let $\mu(\nu) = \mu = E\left(\eta^o(\mathbf{0}) + \eta^w(\mathbf{0})\right)$ be the expected number of particles in a site at time 0. Since $\eta^o(\mathbf{0}) \overset{d}{\sim} \eta^w(\mathbf{0})$, $E\left(\eta^o(x)\right) = E\left(\eta^w(x)\right) = \dfrac{\mu}{2}$ for every $x \in V$.

We will say that a vertex $x \in V$ is *stable* if $\eta^o(x) \wedge \eta^w(x) = 0$, that is, when the number of pairs composed by an oil and a water particle in x is 0, so there are no pairs. Otherwise, we will say that x is *unstable*.

Notice that the measure $\mathbb{P}_\nu$ can be defined on the infinite, vertex-transitive graph with finite degree by defining the truncated configurations: given the Oil and Water dynamics starting from a configuration η_0 of density $\mu(\nu) < \infty$, let, for $M \in \mathbb{N}$, η_0^M be the truncated configuration given by $\eta_0^M(x) = \begin{cases} \eta_0(x) & \text{if } d(x, \mathbf{0}) < M \\ 0 & \text{otherwise} \end{cases}$, and let ν_M be its distribution. Then, $\mathbb{P}_{\nu_M}$ is well defined and is the law of an Oil and Water model where a finite number of vertices fires. It follows from a construction due to Andjel ([1]) that $\mathbb{P}_\nu$ is well defined and, moreover,

$$\mathbb{P}_\nu(E) = \lim_{M \to \infty} \mathbb{P}_{\nu_M}(E)$$

for every event E which depends on a finite space-time window, i.e.
$E \in \sigma\left(\eta_s(x) : d(x, \mathbf{0}) < t, s \in [0,t] \text{ for some } t < \infty\right)$, where η_s is the configuration at time s.
However, the discussion on the existence of $\mathbb{P}_\nu$ on infinite graphs is not the goal of this thesis. Indeed, we will work with a different formulation which does not imply the time variable thanks to an abelian property. The probability measure of such an alternative formulation is denoted by $\mathcal{P}_\nu$ and is introduced below. The equivalence between these two formulations will be stated in Lemma 12.

Let, for every $t \geq 0$, $x \in V$, $u_t(x)$ denote the number of fires occured at x by time t. u_t is called *odometer*.
We will also say that the process *fixates* when for every finite $A \subseteq V$ there exists a random time $\tau_A < \infty$ a.s. such that for all $x \in A$

$$u_t(x) = u_{\tau_A}(x) \qquad \forall\, t \geq \tau_A,$$

namely no vertex in A fires after time τ_A.
Otherwise, we will say that the process is *active*.
In order to formalize the dynamics of the system, for every $x \in V$ and $(y_o, y_w) \in V^2$ pair of vertices such that $y_o, y_w \sim x$, we define a pair of operators

$$(\tau^o_{x,y_o}, \tau^w_{x,y_w}) : \Omega \to \Omega$$

that brings the configuration $\eta = (\eta^o, \eta^w)$ in the configuration $\eta_1 = (\eta^o_1, \eta^w_1)$ with

$$\eta^i_1(z) = \begin{cases} \eta^i(z) - 1 & \text{if } z = x \\ \eta^i(z) + 1 & \text{if } z = y_i \\ \eta^i(z) & \text{otherwise} \end{cases} \qquad \text{for } i \in \{o, w\}.$$

The operator $\left(\tau^o_{x,y_o}, \tau^w_{x,y_w}\right)$, which is a pair of instructions, makes one oil jump from x to y_o and one water from x to y_w.
Now, we fixate an array

$$\tau = \left(\tau^{x,0}, \tau^{x,1}, \tau^{x,2}, \dots\right)_{x \in V} = \left(\tau^{x,j}\right)_{x \in V, j \in \mathbb{N}_0}$$

where each $\tau^{x,j} = \left(\tau^{x,j,o}, \tau^{x,j,w}\right)$ is a pair of instructions in the form $(\tau^o_{x,y_o}, \tau^w_{x,y_w})$, namely it is an element of the set $\left\{(\tau^o_{x,y_o}, \tau^w_{x,y_w}) : y_o, y_w \sim x\right\}$.
We also define a function $h : V \to \mathbb{N}_0$ which counts the number of pairs of instructions used on each vertex.
Given the counter h, we say that $x \in V$ *fires* or that we *topple* (borrowing the notation from the abelian sandpile context) when we use on the pair (η, h) the operator $\Phi_x : \Omega \times \mathbb{N}_0^V \to \Omega \times \mathbb{N}_0^V$,

$$\Phi_x(\eta, h) := \left(\tau^{x,h(x)+1}\eta, h + \delta_x\right).$$

Taken the configuration η and the counter h, Φ_x makes one oil and one water particle jump from x simultaneously and then updates h.
We will say that Φ_x is a *legal* operation for (η, h) if x is unstable for η, otherwise we will call it *illegal*.
We will now mention some properties of this formalization that will be useful for our purposes.
Given $\alpha = (x_1, x_2, \dots, x_k) \in V^k$ $(k \in \mathbb{N})$ sequence of vertices, we define

$$\Phi_\alpha := \Phi_{x_k}\Phi_{x_{k-1}} \dots \Phi_{x_1}.$$

Φ_α (or α itself) is *legal* for η if Φ_{x_l} is legal for $\Phi_{(x_1,\dots,x_{l-1})}(\eta, \hat{0})$ $\forall\ l \in \{2, \dots, k\}$ where $\hat{0}$ is the counter that is 0 on every vertex; otherwise, Φ_α is *illegal*. Given a configuration $\eta \in \Omega$, a legal sequence of vertices α and an array of instructions τ, we put $\Phi_\alpha\eta := \Phi_\alpha(\eta, \hat{0})$, namely Φ_α is the configuration (with counter) obtained from η toppling the vertices according to the sequence α.
Let, for all $x \in V$,

$$m_\alpha(x) := \sum_{l=1}^{k} \mathbb{1}_{\{x=x_l\}} \tag{3.1}$$

be the times x appears in the sequence α.
We will write

$m_\alpha \geq m_\beta$ if $m_\alpha(x) \geq m_\beta(x)$ for every $x \in V$;

$\eta_1 \geq \eta_2$ if $\eta_1^i(x) \geq \eta_2^i(x)$ for every $x \in V$, $i \in \{o, w\}$;

$(\eta', h') \geq (\eta, h)$ if $\eta' \geq \eta$ and $h' \geq h$.

Let $K \subset V$ finite. We will say that

· A configuration η is *stable* in K if every $x \in K$ is stable for η;

· A sequence of vertices α is *contained* in K if all its components are in K;

· α *stabilizes* η in K if $\Phi_\alpha \eta$ is stable in K, indicating, with an abuse of notation, the first component of $\Phi_\alpha \eta$ with $\Phi_\alpha \eta$ itself.

We will now enunciate the abelian property for this model, which states that the final configuration is independent from the order with which the vertices fire.

Lemma 10 (Abelian property) *Let $K \subset V$ finite, $\eta \in \Omega$ and α, β two legal sequences of vertices for η contained in K that stabilize η in K. Then*

1) $m_\alpha = m_\beta$;

2) $\Phi_\alpha \eta = \Phi_\beta \eta$.

All the necessary abelian networks theory for the proof of the abelian property for the Oil and Water model can be found in 6.2 in Appendix. We can see the Oil and Water model as an abelian network $\mathcal{N}$ with underlying graph $G = (V, E)$ and a processor P_v for each $v \in V$, that has

· input alphabet $A_v = \{o_v, w_v\}$ ($A_v \cap A_u = \emptyset$ for $u \neq v$), where o_v and w_v represent respectively an oil and a water particle falling on the vertex v;

· state space $Q_v = \mathbb{N}_0 \times \mathbb{N}_0$;

· transition function $T_v : A_v \times Q_v \to Q_v$,

$$T_v(a_v, (q_o, q_w)) = \begin{cases} \begin{cases} (q_o + 1, q_w) & \text{if } (q_o + 1) \wedge q_w = 0 \\ (q_o, q_w - 1) & \text{if } (q_o + 1) \wedge q_w > 0 \end{cases} & \text{if } a_v = o_v, \\ \begin{cases} (q_o, q_w + 1) & \text{if } q_o \wedge (q_w + 1) = 0 \\ (q_o - 1, q_w) & \text{if } q_o \wedge (q_w + 1) > 0 \end{cases} & \text{if } a_v = w_v \end{cases}$$

· message passing function $T_{(v)} : A_v \times Q_v \times \overline{\Omega} \to \bigsqcup_{u \sim v} A_u^*$ (where $(\overline{\Omega}, \mathcal{H}, \overline{P})$ is a space of probability) with components, for each $\{v, u\} \in E$, $T_{\{v,u\}} : A_v \times Q_v \times \overline{\Omega} \to A_u^*$

with

$$T_{\{v,u\}}(a_v,(q_o,q_w),\omega)=\begin{cases}\begin{cases}\varepsilon & \text{if } (q_o+1)\wedge q_w=0\\ X_{\{v,u\}}(\omega) & \text{if } (q_o+1)\wedge q_w>0\end{cases} & \text{if } a_v=o_v,\\ \begin{cases}\varepsilon & \text{if } q_o\wedge (q_w+1)=0\\ X_{\{v,u\}}(\omega) & \text{if } q_o\wedge (q_w+1)>0\end{cases} & \text{if } a_v=w_v\end{cases}$$

where, using $\varepsilon \in A_u^*$ to indicate the empty word,

$$X_{\{v,u\}}(\omega)=\begin{cases}o_u & \text{if } \omega\in B_1\\ w_u & \text{if } \omega\in B_2\\ o_uw_u & \text{if } \omega\in B_3\\ w_uo_u & \text{if } \omega\in B_4\\ \varepsilon & \text{if } \omega\in B_5\end{cases}.$$

with $B_1,B_2,B_3,B_4,B_5\in\mathcal{H}$ such that $\overline{P}(B_1)=\overline{P}(B_2)=\frac{1}{\mathbf{d}}\frac{\mathbf{d}-1}{\mathbf{d}}$,
$\overline{P}(B_3)=\overline{P}(B_4)=\frac{1}{\mathbf{d}}\frac{1}{\mathbf{d}}$, $\overline{P}(B_5)=1-\frac{1}{\mathbf{d}}-\frac{1}{\mathbf{d}}$.

We can then extend the domain of T_v and $T_{\{v,u\}}$ respectively to $A_v^*\times Q_v$ and $\left(\bigcup_{i=1}^{+\infty}\left(A_v^*\times Q_v\times\overline{\Omega}^i\right)\right)\setminus B$, where
$B=\bigcup_{n=1}^{+\infty}\{(W,q,(\omega_1,...,\omega_n))\in A_v^*\times Q_v\times\overline{\Omega}^n : n \text{ is smaller than the length of } W\}$,
in this way: if $W=aW'$ is a word in the alphabet A_v starting with letter a, then

$$T_v(W,q)=T_v(W',T_v(a,q)),$$

and, for $u\sim v$, indicating $T_{\{v,u\}}(W,q,(\omega_1,...,\omega_n))$ with $T_{\{v,u\}}(W,q,\omega_1,...,\omega_n)$ in order to simplify notation,

$$T_{\{v,u\}}(W,q,\omega_1,...,\omega_n)=$$

$$=\begin{cases}T_{\{v,u\}}(a,q,\omega_1)T_{\{v,u\}}(W',T_v(a,q),\omega_2,...,\omega_n) & \begin{array}{l}\text{if } a=o_v \text{ and } (q_o+1)\wedge q_w>0\\ \text{or if } a=w_v \text{ and } q_o\wedge(q_w+1)>0\end{array}\\ \varepsilon T_{\{v,u\}}(W',T_v(a,q),\omega_1,...,\omega_n) & \text{otherwise}\end{cases},$$

while for the empty word ε we put $T_v(\varepsilon,q)=q$ and $T_{\{v,u\}}(\varepsilon,q,\omega_1,...,\omega_d)=\varepsilon$ for every $d\geq 1$. Given a vertex $v\in V$ and word $W\in A_v^*$, let $|W|\in\mathbb{N}_0^{A_v}$ be the sequence indexed by the elements of A_v with $|W|_a$ number of times a appears in W. We will now prove that, if $W, W' \in A_v^*$ are such that $|W| = |W'|$, then

$$T_v(W,q) = T_v(W',q) \qquad \text{and} \qquad |T_{\{v,u\}}(W,q,\omega)| = |T_{\{v,u\}}(W',q,\omega)|$$

for all $q \in Q_v$, $\{v,u\} \in E$, $\omega \in \Omega^d$ such that $(W,q,\omega),(W',q,\omega)$ are in the domain of $T_{\{v,u\}}$. We are going to show it for two words $W, W' \in A_v^*$ of length 2 that are of course not equal (otherwise it is obvious), namely $W = o_v w_v$, $W' = w_v o_v$, then we can easily get by induction the result for words of any length.

Proof that $T_v(o_v w_v, q) = T_v(w_v o_v, q)$. For $(q_o, q_w) = q \in Q_v$

$$T_v(a_1 a_2, q) = T_v(a_2, T_v(a_1,(q_o,q_w))), \qquad T_v(a_2 a_1, q) = T_v(a_1, T_v(a_2,(q_o,q_w)))$$

so we have to prove that $T_v(a_2, T_v(a_1,(q_o,q_w))) = T_v(a_1, T_v(a_2,(q_o,q_w)))$.
Let us compute, case by case, $T_v(a_1, T_v(a_2,(q_o,q_w)))$ for $a_1 \neq a_2$.

1) If $a_2 = o_v$ and $a_1 = w_v$

 1.1) if $(q_o+1) \wedge q_w = 0$ we have $T_v(a_1,(q_o+1,q_w))$

  1.1.1) if $(q_o+1) \wedge (q_w+1) = 0$ we have (q_o+1, q_w+1), this case is impossible

  1.1.2) if $(q_o+1) \wedge (q_w+1) > 0$ we have (q_o,q_w)

 1.2) if $(q_o+1) \wedge q_w > 0$ we have $T_v(a_1,(q_o,q_w-1))$

  1.2.1) if $q_o \wedge q_w = 0$ we have (q_o,q_w)

  1.2.2) if $q_o \wedge q_w > 0$ we have (q_o-1, q_w-1);

2) if $a_2 = w_v$ and $a_1 = o_v$

 2.1) if $q_o \wedge (q_w+1) = 0$ we have $T_v(a_1,(q_o,q_w+1))$

  2.1.1) if $(q_o+1) \wedge (q_w+1) = 0$ we have (q_o+1, q_w+1), this case is impossible

  2.1.2) if $(q_o+1) \wedge (q_w+1) = 0$ we have (q_o,q_w)

 2.2) if $q_o \wedge (q_w+1) > 0$ we have $T_v(a_1,(q_o-1,q_w))$

  2.2.1) if $q_o \wedge q_w = 0$ we have (q_o,q_w)

  2.2.2) if $q_o \wedge q_w > 0$ we have (q_o-1, q_w-1)

Let us now compute, case by case, $T_v(a_2, T_v(a_1,(q_o,q_w)))$ for $a_1 \neq a_2$.

1) If $a_1 = o_v$ and $a_2 = w_v$

 1.1) if $(q_o+1) \wedge q_w = 0$ we have $T_v(a_2,(q_o+1,q_w))$

1.1.1) if $(q_o+1) \wedge (q_w+1) = 0$ we have (q_o+1, q_w+1), this case is impossible

1.1.2) if $(q_o+1) \wedge (q_w+1) > 0$ we have (q_o, q_w)

1.2) if $(q_o+1) \wedge q_w > 0$ we have $T_v(a_2, (q_o, q_w - 1))$

1.2.1) if $q_o \wedge q_w = 0$ we have (q_o, q_w)

1.2.2) if $q_o \wedge q_w > 0$ we have $(q_o - 1, q_w - 1)$;

2) if $a_1 = w_v$ and $a_2 = o_v$

2.1) if $q_o \wedge (q_w+1) = 0$ we have $T_v(a_2, (q_o, q_w + 1))$

2.1.1) if $(q_o+1) \wedge (q_w+1) = 0$ we have (q_o+1, q_w+1), this case is impossible

2.1.2) if $(q_o+1) \wedge (q_w+1) = 0$ we have (q_o, q_w)

2.2) if $q_o \wedge (q_w+1) > 0$ we have $T_v(a_2, (q_o - 1, q_w))$

2.2.1) if $q_o \wedge q_w = 0$ we have (q_o, q_w)

2.2.2) if $q_o \wedge q_w > 0$ we have $(q_o - 1, q_w - 1)$

Thus, the possible cases are the following ones.

1) $a_1 = o_v$, $a_2 = w_v$ and $(q_o+1) \wedge q_w = 0$: in this case we have

$T_v(a_1, T_v(a_2, (q_o, q_w))) = (q_o, q_w) = T_v(a_2, T_v(a_1, (q_o, q_w)))$;

2) $a_1 = o_v$, $a_2 = w_v$, $(q_o+1) \wedge q_w > 0$ and $q_o \wedge q_w = 0$: in this case we have

$T_v(a_1, T_v(a_2, (q_o, q_w))) = (q_o, q_w) = T_v(a_2, T_v(a_1, (q_o, q_w)))$;

3) $a_1 = o_v$, $a_2 = w_v$, $(q_o+1) \wedge q_w > 0$ and $q_o \wedge q_w > 0$: in this case we have

$T_v(a_1, T_v(a_2, (q_o, q_w))) = (q_o - 1, q_w - 1) = T_v(a_2, T_v(a_1, (q_o, q_w)))$;

4) $a_2 = o_v$, $a_1 = w_v$ and $(q_o+1) \wedge q_w = 0$: in this case we have

$T_v(a_1, T_v(a_2, (q_o, q_w))) = (q_o, q_w) = T_v(a_2, T_v(a_1, (q_o, q_w)))$;

5) $a_2 = o_v$, $a_1 = w_v$, $(q_o+1) \wedge q_w > 0$ and $q_o \wedge q_w = 0$: in this case we have

$T_v(a_1, T_v(a_2, (q_o, q_w))) = (q_o, q_w) = T_v(a_2, T_v(a_1, (q_o, q_w)))$;

6) $a_2 = o_v$, $a_1 = w_v$, $(q_o+1) \wedge q_w > 0$ and $q_o \wedge q_w > 0$: in this case we have

$T_v(a_1, T_v(a_2, (q_o, q_w))) = (q_o - 1, q_w - 1) = T_v(a_2, T_v(a_1, (q_o, q_w)))$.

Hence we have the desired result. $\square$

Proof that $|T_{\{v,u\}}(a_1a_2, q, \omega_1, \omega_2)| = |T_{\{v,u\}}(a_2a_1, q, \omega_1, \omega_2)|$. For $(q_o, q_w) = q \in Q_v$ let us compute, case by case, $T_{\{v,u\}}(a_1a_2, (q_o, q_w), \omega_1, \omega_2)$ for $a_1 \neq a_2$.

1) If $a_1 = o_v$ and $a_2 = w_v$

1.1) if $(q_o+1) \wedge q_w = 0$ we have $\varepsilon T_{\{v,u\}}(w_v, (q_o+1, q_w), \omega_1, \omega_2)$

1.1.1) if $(q_o + 1) \wedge (q_w + 1) = 0$ we have $\varepsilon\varepsilon = \varepsilon$, this case is impossible

1.1.2) if $(q_o + 1) \wedge (q_w + 1) > 0$ we have $\varepsilon X_{\{v,u\}}(\omega_1)$

1.2) if $(q_o + 1) \wedge q_w > 0$ we have $X_{\{v,u\}}(\omega_1) T_{\{v,u\}}(w_v, (q_o, q_w - 1), \omega_2)$

1.2.1) if $q_o \wedge q_w = 0$ we have $X_{\{v,u\}}(\omega_1)\varepsilon$

1.2.2) if $q_o \wedge q_w > 0$ we have $X_{\{v,u\}}(\omega_1) X_{\{v,u\}}(\omega_2)$;

2) if $a_1 = w_v$ and $a_2 = o_v$

2.1) if $q_o \wedge (q_w + 1) = 0$ we have $\varepsilon T_{\{v,u\}}(o_v, (q_o, q_w + 1), \omega_1, \omega_2)$

2.1.1) if $(q_o + 1) \wedge (q_w + 1) = 0$ we have $\varepsilon\varepsilon = \varepsilon$, this case is impossible

2.1.2) if $(q_o + 1) \wedge (q_w + 1) > 0$ we have $\varepsilon X_{\{v,u\}}(\omega_1)$

2.2) if $q_o \wedge (q_w + 1) > 0$ we have $X_{\{v,u\}}(\omega_1) T_{\{v,u\}}(o_v, (q_o - 1, q_w), \omega_2)$

2.2.1) if $q_o \wedge q_w = 0$ we have $X_{\{v,u\}}(\omega_1)\varepsilon$

2.2.2) if $q_o \wedge q_w > 0$ we have $X_{\{v,u\}}(\omega_1) X_{\{v,u\}}(\omega_2)$.

Let us now compute, case by case, $T_{\{v,u\}}(a_2 a_1, (q_o, q_w), \omega_1, \omega_2)$.

1) If $a_2 = o_v$ and $a_1 = w_v$

1.1) if $(q_o + 1) \wedge q_w = 0$ we have $\varepsilon T_{\{v,u\}}(w_v, (q_o + 1, q_w), \omega_1, \omega_2)$

1.1.1) if $(q_o + 1) \wedge (q_w + 1) = 0$ we have $\varepsilon\varepsilon = \varepsilon$, this case is impossible

1.1.2) if $(q_o + 1) \wedge (q_w + 1) > 0$ we have $\varepsilon X_{\{v,u\}}(\omega_1)$

1.2) if $(q_o + 1) \wedge q_w > 0$ we have $X_{\{v,u\}}(\omega_1) T_{\{v,u\}}(w_v, (q_o, q_w - 1), \omega_2)$

1.2.1) if $q_o \wedge q_w = 0$ we have $X_{\{v,u\}}(\omega_1)\varepsilon$

1.2.2) if $q_o \wedge q_w > 0$ we have $X_{\{v,u\}}(\omega_1) X_{\{v,u\}}(\omega_2)$;

2) if $a_2 = w_v$ and $a_1 = o_v$

2.1) if $q_o \wedge (q_w + 1) = 0$ we have $\varepsilon T_{\{v,u\}}(o_v, (q_o, q_w + 1), \omega_1, \omega_2)$

2.1.1) if $(q_o + 1) \wedge (q_w + 1) = 0$ we have $\varepsilon\varepsilon = \varepsilon$, this case is impossible

2.1.2) if $(q_o + 1) \wedge (q_w + 1) > 0$ we have $\varepsilon X_{\{v,u\}}(\omega_1)$

2.2) if $q_o \wedge (q_w + 1) > 0$ we have $X_{\{v,u\}}(\omega_1) T_{\{v,u\}}(o_v, (q_o - 1, q_w), \omega_2)$

2.2.1) if $q_o \wedge q_w = 0$ we have $X_{\{v,u\}}(\omega_1)\varepsilon$

2.2.2) if $q_o \wedge q_w > 0$ we have $X_{\{v,u\}}(\omega_1) X_{\{v,u\}}(\omega_2)$.

Thus, the possible cases are the following ones.

1) $a_1 = o_v$, $a_2 = w_v$ and $(q_o + 1) \wedge q_w = 0$: in this case we have

$$T_{\{v,u\}}(a_1a_2,(q_o,q_w),\omega_1,\omega_2) = \varepsilon X_{\{v,u\}}(\omega_1),$$

$$T_{\{v,u\}}(a_2a_1,q,\omega_1,\omega_2) = \begin{cases} \varepsilon X_{\{v,u\}}(\omega_1) & \text{if } q_o = 0 \\ X_{\{v,u\}}(\omega_1)\varepsilon & \text{if } q_o > 0 \end{cases};$$

2) $a_1 = o_v$, $a_2 = w_v$, $(q_o+1)\wedge q_w > 0$ and $q_o \wedge q_w = 0$: in this case we have

$$T_{\{v,u\}}(a_1a_2,(q_o,q_w),\omega_1,\omega_2) = X_{\{v,u\}}(\omega_1)\varepsilon,\ \ T_{\{v,u\}}(a_2a_1,q,\omega_1,\omega_2) = \varepsilon X_{\{v,u\}}(\omega_1);$$

3) $a_1 = o_v$, $a_2 = w_v$, $(q_o+1)\wedge q_w > 0$ and $q_o \wedge q_w > 0$: in this case we have

$$T_{\{v,u\}}(a_1a_2,(q_o,q_w),\omega_1,\omega_2) = X_{\{v,u\}}(\omega_1)X_{\{v,u\}}(\omega_2),$$

$$T_{\{v,u\}}(a_2a_1,q,\omega_1,\omega_2) = X_{\{v,u\}}(\omega_1)X_{\{v,u\}}(\omega_2);$$

4) $a_2 = o_v$, $a_1 = w_v$ and $(q_o+1)\wedge q_w = 0$: in this case we have

$$T_{\{v,u\}}(a_1a_2,(q_o,q_w),\omega_1,\omega_2) = \begin{cases} \varepsilon X_{\{v,u\}}(\omega_1) & \text{if } q_o = 0 \\ X_{\{v,u\}}(\omega_1)\varepsilon & \text{if } q_o > 0 \end{cases},$$

$$T_{\{v,u\}}(a_2a_1,q,\omega_1,\omega_2) = \varepsilon X_{\{v,u\}}(\omega_1);$$

5) $a_2 = o_v$, $a_1 = w_v$, $(q_o+1)\wedge q_w > 0$ and $q_o \wedge q_w = 0$: in this case we have

$$T_{\{v,u\}}(a_1a_2,(q_o,q_w),\omega_1,\omega_2) = \varepsilon X_{\{v,u\}}(\omega_1),\ \ T_{\{v,u\}}(a_2a_1,q,\omega_1,\omega_2) = X_{\{v,u\}}(\omega_1)\varepsilon;$$

6) $a_2 = o_v$, $a_1 = w_v$, $(q_o+1)\wedge q_w > 0$ and $q_o \wedge q_w > 0$: in this case we have

$$T_{\{v,u\}}(a_1a_2,(q_o,q_w),\omega_1,\omega_2) = X_{\{v,u\}}(\omega_1)X_{\{v,u\}}(\omega_2),$$

$$T_{\{v,u\}}(a_2a_1,q,\omega_1,\omega_2) = X_{\{v,u\}}(\omega_1)X_{\{v,u\}}(\omega_2).$$

Each case satisfies $|T_{\{v,u\}}(a_1a_2,(q_o,q_w),\omega_1,\omega_2)| = |T_{\{v,u\}}(a_2a_1,q,\omega_1,\omega_2)|$. □

Hence, this is an abelian network. Notice that the time at which the processor $\mathcal{P}_v$ processes its input word is given by the time at which the Poisson clock of rate 1 associated to the vertex v rings. Like all the abelian networks, it satisfies the abelian property.

For every $K \subset V$ finite, $\eta \in \Omega$, τ array of instructions, let, for $x \in V$, $m_{K,\eta,\tau}(x)$ be the number of times x fires during the stabilization of K starting from the configuration η following the instructions in τ.
For the abelian property, $m_{K,\eta,\tau}$ is well defined.

Lemma 11 (Monotonicity for Oil and Water) *Let $K \subseteq K' \subset V$ with K' finite set and $\eta, \eta' \in \Omega$ such that $\eta \leq \eta'$. Then*

$$m_{K,\eta,\tau} \leq m_{K',\eta',\tau}.$$

Proof. For a fixed array of instructions τ, let α and β be two legal sequences that stabilize respectively η' in K' and η in K. Then α also stabilizes η' in $K \subseteq K'$. Since if x is stable in η' and $\eta'(x) \geq \eta(x)$ then x is stable in η, and since if moreover $\eta' \geq \eta$ then $\Phi_x \eta' \geq \Phi_x \eta$, we have that if α stabilizes η' in K then it also stabilizes η in K.

Therefore, for all $x \in V$, the number of times x fires according to the sequence α in the stabilization of K is equal to the one according to the sequence β.

Since by definition this sequence could not have stabilized any vertex in $K' \setminus K$ (K' cannot be stable until K is not stable), we get the desired result from (4.6). □

Monotonicity gives that if Oil and Water at density μ fixates, it also fixates for all $\mu' \in [0, \mu]$.

Therefore, given $V_1 \subseteq V_2 \subseteq \cdots \subseteq V$ such that $\lim_{t\to\infty} V_t = V$,

$$\exists \lim_{t\to\infty} m_{V_t,\eta,\tau} =: m_{\eta,\tau} \qquad \text{and it does not depend on the sequence } \{V_t\}_{t\in\mathbb{N}}.$$

We now allow the array of instructions τ, therefore the instructions $\tau^{x,j}$, to be stochastic, in the following way.

$$\tau^{x,j} := (\tau^{x,j,o}, \tau^{x,j,w}) : \Gamma \times \Gamma \to \Big\{ (\tau^o_{z,y_o}, \tau^w_{z,y_w}) \Big| z \in V, y_o \sim z, y_w \in z \Big\}$$

will be a random variable on the probability space $(\Gamma \times \Gamma, \mathcal{G}, \mathcal{P})$ with $\mathcal{P}$ such that the pairs $(\tau^{x,j,o}, \tau^{x,j,w})$ are

· independent for different values of x:

$$\begin{aligned} &\mathcal{P}\Big((\tau^{x,j,o}, \tau^{x,j,w}) = (\tau^o_{x,y_o}, \tau^w_{x,y_w}), (\tau^{z,j,o}, \tau^{z,j,w}) = (\tau^o_{z,y_o}, \tau^w_{z,y_w}) \Big) \\ = &\mathcal{P}\Big((\tau^{x,j,o}, \tau^{x,j,w}) = (\tau^o_{x,y_o}, \tau^w_{x,y_w}) \Big) \mathcal{P}\Big((\tau^{z,j,o}, \tau^{z,j,w}) = (\tau^o_{z,y_o}, \tau^w_{z,y_w}) \Big); \end{aligned}$$

· independent for different values of j:

$$\begin{aligned} &\mathcal{P}\Big((\tau^{x,j,o}, \tau^{x,j,w}) = (\tau^o_{x,y_o}, \tau^w_{x,y_w}), (\tau^{x,k,o}, \tau^{x,k,w}) = (\tau^o_{x,y_o}, \tau^w_{x,y_w}) \Big) \\ = &\mathcal{P}\Big((\tau^{x,j,o}, \tau^{x,j,w}) = (\tau^o_{x,y_o}, \tau^w_{x,y_w}) \Big) \mathcal{P}\Big((\tau^{x,k,o}, \tau^{x,k,w}) = (\tau^o_{x,y_o}, \tau^w_{x,y_w}) \Big) \end{aligned}$$

and such that $\tau^{x,j,o}, \tau^{x,j,w}$ are independent for different values of $i \in \{o, w\}$:

$$\mathcal{P}(\tau^{x,j,o} = \tau^o_{x,y_o}, \tau^{x,j,w} = \tau^w_{x,y_w}) = \mathcal{P}(\tau^{x,j,o} = \tau^o_{x,y_o}) \mathcal{P}(\tau^{x,j,w} = \tau^w_{x,y_w}),$$

therefore

$$\mathcal{P}\Big(\underset{\tau^{x,j}}{\underset{\|}{(\tau^{x,j,o},\tau^{x,j,w})}} = (\tau^{o}_{x,y_o},\tau^{w}_{x,y_w})\Big) = \mathcal{P}(\tau^{x,j,o}=\tau^{o}_{x,y_o})\mathcal{P}(\tau^{x,j,w}=\tau^{w}_{x,y_w})$$

and we define $\mathcal{P}$ through the degree $\mathbf{d}$ of x:

$$\mathcal{P}(\tau^{x,j,o}=\tau^{o}_{x,y_o}) = \mathcal{P}(\tau^{x,j,w}=\tau^{w}_{x,y_w}) = \frac{1}{\mathbf{d}} \qquad \forall\ y_o, y_w \sim x, \tag{3.2}$$

namely under $\mathcal{P}$ the instructions induce every particle they are used on to perform a step of independent simple random walk.

Let now $\mathcal{P}_\nu$ be the joint probability measure of the configurations η and the instructions arrays τ.

With an abuse of notation, we will sometimes indicate $\mathcal{P}_\nu$ with $\mathcal{P}$ itself.

Lemma 12 (0-1 law)

$$\mathbb{P}_\nu(\textit{Oil and Water fixates}) = \mathcal{P}_\nu(m_{\eta,\tau}(\mathbf{0})<\infty) \in \{0,1\}.$$

The proof of this result is analogous to the one for Activated Random Walks and Stochastic Sandpiles, which can be found in [25].

For simplicity, we will now use the notation $m_K := m_{K,\eta,\tau}$ and $m := m_{\eta,\tau}$.

Given a configuration $\eta = (\eta^o,\eta^w)$,

· $\eta^o(x)\wedge\eta^w(x)$ is the *number of oil-water pairs* in the vertex x;

· $\eta^o(x)\vee\eta^w(x) - \eta^o(x)\wedge\eta^w(x)$ is the *number of unpaired particles* in the vertex x;

· we will call the vertex x a *hole* for η if the number of oil and water particles in x is the same, i.e. $\eta^o(x)=\eta^w(x)$, or, equivalently, if the number of unpaired particles is zero, namely $\eta^o(x)\vee\eta^w(x)-\eta^o(x)\wedge\eta^w(x)=0$.

3.2 Number of waters falling into holes and ghost-pair stabilization

Notice that when a vertex x fires, if no neighbor of x is a hole, calling d_w and d_o the number of neighbors of x which have respectively unpaired waters and unpaired oils, we have $d_w + d_o = \mathbf{d}$ and

$$\text{the number of oil-water pairs changes by} \begin{cases} +1 & \text{with probability } \dfrac{d_w d_o}{\mathbf{d}^2} \\ -1 & \text{with probability } \dfrac{d_w d_o}{\mathbf{d}^2} \\ 0 & \text{with probability } 1 - 2\dfrac{d_w d_o}{\mathbf{d}^2} \end{cases},$$

namely in this case the total number of pairs in the system behaves like a martingale. When at least one neighbor of x is a hole, instead, an additional pair is lost each time a hole is chosen by exactly one particle as the site to jump on, and the total number of pairs behaves like a supermartingale.

Now we want to stabilize any finite set $K \subset V$ following a legal sequence and we will formalize this operation through the definition of a *strategy*.

Definition 2 (Strategy for stabilizing a finite set of vertices) *Given a finite set $K \subset V$, a strategy for stabilizing K is $F_K : \Omega \to V \cup \{\emptyset\}$,*

$$F_K(\eta) = \begin{cases} x \textit{ with } x \in K \textit{ stable} & \textit{if } \eta \textit{ is not stable in } K \\ \emptyset & \textit{if } \eta \textit{ is stable in } K \end{cases}.$$

We say that we stabilize η in K following the strategy F_K if:

- beginning from $\eta_0 := \eta \in \Omega$,
- we apply F_K to η_0: if $F_K(\eta_0) = \emptyset$, η_0 is stable in K and we stop; otherwise, we topple the vertex $F_K(\eta_0)$ and we denote $\eta_1 \in \Omega$ the resulting configuration;
- we apply F_K to η_1: if $F_K(\eta_1) = \emptyset$, η_1 is stable in K and we stop; otherwise, we topple the vertex $F_K(\eta_1)$ and we denote $\eta_2 \in \Omega$ the resulting configuration;
- we iterate until we reach the random time

$$T_{F_K} := \inf \{i \in \mathbb{N}_0 : F_K(\eta_i) = \emptyset\}$$

  at which K is stabilized.

Remark 2 $T_{F_K} < \infty$ *a.s., indeed K is stabilized in a finite time if and only if for all $x \in K$ the odometer $u(x) < \infty$. If we consider $\widetilde{F}_K : \Omega \to V \cup \{\emptyset\}$,*

$$\widetilde{F}_K(\eta) = \begin{cases} x \textit{ with } x \in K \textit{ such that } \eta(x) \neq (0,0) & \textit{if } \exists\, y \in K \textit{ such that } \eta(y) \neq (0,0) \\ \emptyset & \textit{otherwise} \end{cases}$$

and the same stabilization procedure we defined for F_K applied to $\widetilde{F}_K$, we get that in this new procedure for all $x \in K$ the oil and water particles which are on x jump whether there is an oil-water pair on x or not. Therefore, calling $\tilde{u}(x)$ the odometer of this new process, we have $u(x) \leq \tilde{u}(x)$ for each $x \in K$. But in the new process, for all $x \in K$, for every particle p of oil or water, p performs a simple random walk until it exits K, and for Remark 1 there exists a finite time T at which there are no oil nor water particles on x, and we have $u(x) \leq \tilde{u}(x) \leq T < \infty$, therefore $T_{F_K} < \infty$ a.s.

Thus, we have a sequence of configurations $(\eta_i)_{i \in [0, T_{F_K}]}$ with $\eta_{T_{F_K}}$ stable in K.
We define now the configurations $(\eta_i)_{i > T_{F_K}}$ as follows:
for $i \geq T_{F_K}$, choosing the origin $\mathbf{0}$ in K,

- if $F_K(\eta_i) = \emptyset$ (namely η_i is stable in K), we add an oil-water pair in the origin $\mathbf{0}$, getting the new configuration η_{i+1}, unstable in K (notice that no vertex fires in step i in this case)

- if $F_K(\eta_i) \neq \emptyset$ (namely η_i is unstable in K), the vertex $F_K(\eta_i)$ fires and we get the new configuration η_{i+1} (which can be stable or unstable in K).

Thus, we have the sequence $(\eta_i)_{i \in \mathbb{N}_0}$. For every $x \in V$ let

$$H_{K,F_K,\eta_0,\tau}(x) := \left| \{ 0 \leq i \leq T_{F_K} - 1 : \eta_i^w(x) = \eta_i^o(x) \text{ and } \eta_{i+1}^w(x) = \eta_{i+1}^o(x) + 1 \} \right|$$

be the *number of times a water particle falls in a hole in x following the strategy F_K starting from the configuration η_0*, and for simplicity we will write $H_{K,F_K}(x)$ in place of $H_{K,F_K,\eta_0,\tau}(x)$.

Lemma 13 *If the system starting from a configuration η distributed as a product of the measure ν is active a.s., for every $\varepsilon > 0$, $M \in \mathbb{N}$ there exists $D = D_{\nu,\varepsilon,M} < \infty$ large enough such that*

$$\inf_{\substack{K \subset V \textit{ finite:} \\ d(\mathbf{0},K^c) > D}} \ \inf_{\substack{F_K \\ \textit{strategy}}} \mathcal{P}_\nu \left(H_{K,F_K}(\mathbf{0}) > M \right) \geq 1 - \varepsilon.$$

Proof. Let $K \subset V$ be a finite set such that $B_D := \{x : \|x\| < D\} \subseteq K$. We will now stabilize K following a strategy F_K.

For each $j \in \mathbb{N}_0$, let t_j be the j-th time a neighbor of $\mathbf{0}$ fires, i.e. let

$\mathcal{N}_{\mathbf{0}} = \{x \in V : x \sim \mathbf{0}\}$ be the set of neighbors of the origin and $t_0 = \mathbf{0}$,

$t_j = \inf\{i > t_{j-1} : F_K(\eta_{i-1}) \in \mathcal{N}_{\mathbf{0}}\}$ for $j \in \mathbb{N}$ be the first time after t_{j-1} a firing occurs in a vertex in $\mathcal{N}_{\mathbf{0}}$.

Let $N_A = \sum\limits_{x \in \mathcal{N}_{\mathbf{0}}} m_A(x)$ be the number of times a firing occurs in a neighbor of $\mathbf{0}$ during the stabilization of $A \subseteq V$ and let

$$R_j = \eta^w_{t_j}(\mathbf{0}) - \eta^o_{t_j}(\mathbf{0}) \qquad \forall\, j \in \mathbb{N}_0$$

be a sequence of random variables which are the difference between the number of water and oil particles in the origin every time a firing occurs in a vertex in $\mathcal{N}_{\mathbf{0}}$.

$\{R_j\}_{j \in \mathbb{N}_0}$ is distributed as a symmetric lazy random walk on $\mathbb{Z}$ with given initial value R_0, indeed

$$\begin{aligned}
&\mathcal{P}_\nu(R_{j+1} = R_j + 1 | R_j) = \mathcal{P}_\nu(\eta^w_{t_{j+1}}(\mathbf{0}) = \eta^w_{t_j}(\mathbf{0}) + 1, \eta^o_{t_{j+1}}(\mathbf{0}) = \eta^o_{t_j}(\mathbf{0})) = \frac{1}{\mathbf{d}}\frac{\mathbf{d}-1}{\mathbf{d}},\\
&\mathcal{P}_\nu(R_{j+1} = R_j - 1 | R_j) = \mathcal{P}_\nu(\eta^w_{t_{j+1}}(\mathbf{0}) = \eta^w_{t_j}(\mathbf{0}), \eta^o_{t_{j+1}}(\mathbf{0}) = \eta^o_{t_j}(\mathbf{0}) + 1) = \frac{\mathbf{d}-1}{\mathbf{d}}\frac{1}{\mathbf{d}},\\
&\mathcal{P}_\nu(R_{j+1} = R_j | R_j) = 1 - 2\frac{1}{\mathbf{d}}\frac{\mathbf{d}-1}{\mathbf{d}}.
\end{aligned}$$

For all $j \in \mathbb{N}$ let

$$\mathcal{J}(j) = \Big|\{k \in [0, j) : R_k = 0 \text{ and } R_{k+1} = 1\}\Big|$$

be the number of times the random walk jumps from 0 to 1 in the first j steps. Then, by definition, $H_{K,F_K}(\mathbf{0}) = \mathcal{J}(N_K)$, therefore, for all $M, \varphi \in \mathbb{N}_0$

$$\begin{aligned}
\mathcal{P}_\nu(H_{K,F_K}(\mathbf{0}) > M) &\geq \mathcal{P}_\nu(\underset{\substack{\| \\ \mathcal{J}(N_K)}}{H_{K,F_K}(\mathbf{0})} > M, N_K > \varphi) \geq \mathcal{P}_\nu(\mathcal{J}(\varphi) > M, N_K > \varphi)\\
&\geq \mathcal{P}_\nu(\mathcal{J}(\varphi) > M) - \mathcal{P}_\nu(N_K \leq \varphi)\\
&\underset{\substack{\uparrow \\ N_{B_D} \underset{\substack{\uparrow \\ B_D \subseteq K}}{\leq} N_K \text{ for monotonicity, therefore } \{N_K \leq \varphi\} \subseteq \{N_{B_D} \leq \varphi\}}}{\geq} \mathcal{P}_\nu(\mathcal{J}(\varphi) > M) - \mathcal{P}_\nu(N_{B_D} \leq \varphi).
\end{aligned}$$

Since $R_0 = \eta^w_0(\mathbf{0}) - \eta^o_0(\mathbf{0}) < \infty$ a.s. and the symmetric lazy random walk on $\mathbb{Z}$ is recurrent, we have thtat for all $\varepsilon \in (0, 1), M \in \mathbb{N}_0, \nu$ with $\mu(\nu) < \infty$ there exists

$\varphi = \varphi_{\nu,\varepsilon,M}$ large enough such that

$$\mathcal{P}_\nu(\mathcal{J}(\varphi) > M) \geq 1 - \frac{\varepsilon}{2}.$$

By assumption the system is active a.s., therefore there exists $D = D_{\varepsilon,\varphi}$ large enough such that

$$\mathcal{P}_\nu(N_{B_D} \leq \varphi) \leq \frac{\varepsilon}{2}$$

uniformly in F_K by the abelian property
$\Rightarrow$ for every ε, M there exists $D = D_{\nu,\varepsilon,M}$ large enough such that, uniformly in $K \supseteq B_D$ and in F_K,

$$\mathcal{P}_\nu(H_{K,F_K}(\mathbf{0}) > M) \geq 1 - \frac{\varepsilon}{2} - \frac{\varepsilon}{2} = 1 - \varepsilon.$$

$\square$

This lemma gives us that for an active system the number of times a water particle falls into a hole in a given vertex x is arbitrarily large. Noticing that when a particle falls into a hole the total number of oil-water pairs in the process decreases, we will exploit this lemma to present a *reductio to absurdum* argument which assumes the system to be active a.s. and then arrives to a contradiction.
In order to show this, since it is extremely hard to control the evolution of the system consisting exclusively of oil-water pairs, because this requires controlling the evolution of the configuration of holes and pairs at the same time, which are strongly correlated, we introduce a new type of auxiliary particles to compensate the pairs that are lost when we fire a vertex neighboring a hole, which we will call *ghosts*. These new particles will not interact with oils nor waters and perform independent simple random walks.
In order to introduce this new procedure, we define a new set of configurations

$$\widetilde{\Omega} := \mathbb{N}_0^V \times \mathbb{N}_0^V \times \mathbb{N}_0^V,$$

where in a configuration $\tilde{\eta} = (\tilde{\eta}^o(x), \tilde{\eta}^w(x), \tilde{\eta}^g(x))_{x\in V} \in \widetilde{\Omega}$, for every $x \in V$, $\tilde{\eta}^o(x), \tilde{\eta}^w(x)$ and $\tilde{\eta}^g(x)$ represent respectively the number of oil, water and ghost particles in x.
Given $K \subset V$ finite, $\sigma \in \Omega$ unstable configuration (consisting of just oil and water

particles), we will call *ghost-pair stabilization* the following procedure:
at time 0, we start from the configuration $\tilde{\eta}_0 \in \widetilde{\Omega}$ such that $\sigma = (\tilde{\eta}_0^o, \tilde{\eta}_0^w)$ and $\tilde{\eta}_0^g(x) = 0$ for all $x \in V$ (i.e. at time 0 there are no ghost particles); inductively for every $t \in \mathbb{N}_0$, we perform the following steps:

(i) either a ghost or an oil-water pair in $\tilde{\eta}_t$ which are on a vertex in K perform a step of simple random walk (where for the oil-water pair it means that an oil particle and a water one which are on the same vertex take each one an independent step according to the distribution of a simple random walk). This brings to a new configuration $\theta_t \in \tilde{\Omega}$.

(ii) - if during (i) a water particle falls into a hole $x \in K$ (namely $\tilde{\eta}_t^o(x) = \tilde{\eta}_t^w(x)$ and $\theta_t^w(x) = \theta_t^o(x) + 1$), then a ghost particle is created on that vertex, that is,

$$\tilde{\eta}_{t+1}^g := \theta_t^g + \delta_x, \qquad \tilde{\eta}_{t+1}^q := \theta_t^q \qquad \text{for } q \in \{o, w\}.$$

- otherwise nothing happens, i.e. $\tilde{\eta}_{t+1} := \theta_t$.

This defines $\tilde{\eta}_{t+1}$.
Each step of the procedure corresponds either to an oil-water pair performing a simple random walk step from an unstable vertex, or a ghost performing a simple random walk step and, at any given step of the process, at most one ghost is created. Since K is finite, for the same reasoning as in Remark 2, after an almost surely finite time no oil-water pair and no ghost particle are present in K and the procedure stops. Therefore, let

$$T = T_K = \inf\{t \geq 0 : (\tilde{\eta}_t^o, \tilde{\eta}_t^w) \text{ is stable in } K \text{ and } \tilde{\eta}_t^g(x) = 0 \ \ \forall\, x \in K\}$$

and, for all $t \geq T$, we set

$$\tilde{\eta}_t := \tilde{\eta}_T.$$

Let, for all $y \in K$, $\tilde{m}(y)$ be the number of times that either a ghost or an oil-water pair jumps from y, and we denote with $\tilde{P}_{K,\sigma}$ the law of the ghost-pair stabilization. The next lemma expresses $\tilde{m}(y)$ in terms of the Green's function for K and of the number of pairs in the starting configuration.

Lemma 14 *For $K \subset V$ finite, $y \in K$ and $\sigma = (\tilde{\eta}_0^o, \tilde{\eta}_0^w) \in \Omega$ unstable,*

$$\tilde{E}_{K,\sigma}\left[\tilde{m}(y)\right] = \sum_{x \in K} \left(\tilde{\eta}_0^o(x) \wedge \tilde{\eta}_0^w(x)\right) G_K(x,y)$$

being $\tilde{E}_{K,\sigma}$ the expected value associated to the probability measure $\tilde{P}_{K,\sigma}$.

Proof. Let $K \subset V$ finite and $y \in K$.

For Theorem 1, there exists a unique function $g : V \to \mathbb{R}$ which satisfies

$$(a)\ \Delta g(x) = 0, \text{ for } x \in K \setminus \{y\}$$
$$(b)\ g(x) = F(x), \text{ for } x \in \partial K \cup \{y\} \qquad \text{with } F(x) = \mathbb{1}_{\{x=y\}}$$
$$(c)\ g(x) = 0, \text{ for } x \in K^c,$$

i.e. g is harmonic in $K \setminus \{y\}$ and such that $g(y) = 1$, $g(z) = 0$ for $z \in K^c$; that function is $g(x) = \begin{cases} P_x\left(\tau_{V\setminus\{y\}} < \tau_K\right) & \text{if } x \in K \\ 0 & \text{if } x \notin K \end{cases}$ and is such that

$$-\Delta g(y) = 1 - P_y(\tau^+_{V\setminus\{y\}} < \tau_K), \tag{3.3}$$

indeed $\Delta g(y) = \frac{1}{\mathbf{d}} \sum_{z \sim y} g_z - 1$ and

$$1 - P_y(\tau^+_{V\setminus\{y\}} < \tau_K) = 1 - E_y\left(\mathbb{1}_{\{\tau^+_{V\setminus\{y\}} < \tau_K\}}\right) = 1 - \sum_{z\sim y} \underbrace{E_y\left(\mathbb{1}_{\{\tau^+_{V\setminus\{y\}} < \tau_K\}} \middle| X_1 = z\right)}_{E_z\left(\mathbb{1}_{\{\tau_{V\setminus\{y\}} < \tau_K\}}\right)} \frac{1}{\mathbf{d}}$$
$$= 1 - \frac{1}{\mathbf{d}} \sum_{z\sim y} P_z(\tau_{V\setminus\{y\}} < \tau_K) = 1 - \frac{1}{\mathbf{d}} \sum_{z \sim y} g_z = -\Delta g(y).$$

Let $\tilde{\eta}_t$ be the configuration at time t.

By convention, we will refer to the transition between $\tilde{\eta}_{t-1}$ and $\tilde{\eta}_t$ as "step t".

Let x_t be the vertex from which an oil-water pair or a ghost particle jumps at step t.

For all $t \in \mathbb{N}_0$ let

$$M_t = \sum_{x \in K} \left(\tilde{\eta}_t^o(x) \wedge \tilde{\eta}_t^w(x) + \tilde{\eta}_t^g(x)\right) g(x) - \Delta g(y) \sum_{i=1}^{t} \mathbb{1}_{\{x_i = y\}}.$$

Let $\left(\tilde{\Sigma}, \tilde{\mathcal{F}}, \tilde{P}_{K,\sigma}\right)$ be the probability space on which the process $\{\tilde{\eta}_t\}_{t\in\mathbb{N}_0}$ is defined (with natural filtration $\{\tilde{\mathcal{F}}_t\}_{t\in\mathbb{N}_0}$ associated to $\{\tilde{\eta}_t^o, \tilde{\eta}_t^w, \tilde{\eta}_t^g\}_{t\in\mathbb{N}_0}$).

We will now show that $\{M_t\}_{t\in\mathbb{N}_0}$ is a martingale, i.e. $\tilde{E}_{K,\sigma}(M_t | \tilde{\mathcal{F}}_{t-1}) = M_{t-1}$.

- if at step t a ghost particle jumps from $x_t = b \in K$

$$\begin{aligned}
\tilde{E}_{K,\sigma}(M_t|\tilde{\mathcal{F}}_{t-1}) = &\sum_{\substack{x\in K\\ x\nsim b\\ x\neq b}} \left(\tilde{\eta}^o_{t-1}(x) \wedge \tilde{\eta}^w_{t-1}(x) + \tilde{\eta}^g_{t-1}(x)\right) g(x)+\\
&+ \left(\tilde{\eta}^o_{t-1}(b) \wedge \tilde{\eta}^w_{t-1}(b) + \tilde{\eta}^g_{t-1}(b) - 1\right) g(b)+\\
&+ \sum_{\substack{x\in K\\ x\sim b}} \left(\tilde{\eta}^o_{t-1}(x) \wedge \tilde{\eta}^w_{t-1}(x) + \tilde{\eta}^g_{t-1}(x)\right) g(x)+\\
&+ \sum_{z\sim b} \underbrace{\tilde{E}_{K,\sigma}(\tilde{\eta}^g_t(z)|\tilde{\mathcal{F}}_{t-1})}_{\substack{\left(\tilde{\eta}^g_{t-1}(z)+1\right)\frac{1}{\mathbf{d}}+\tilde{\eta}^g_{t-1}(z)\frac{\mathbf{d}-1}{\mathbf{d}}\\ \parallel\\ \tilde{\eta}^g_{t-1}(z)+\frac{1}{\mathbf{d}}}} g(z) - \Delta g(y)\sum_{i=1}^{t-1}\mathbb{1}_{\{x_i=y\}} - \Delta g(y)\mathbb{1}_{\{b=y\}}\\
&= \sum_{x\in K} \left(\tilde{\eta}^o_{t-1}(x) \wedge \tilde{\eta}^w_{t-1}(x) + \tilde{\eta}^g_{t-1}(x)\right) g(x) - g(b)+\\
&+ \frac{1}{\mathbf{d}}\sum_{z\sim b} g(z) - \Delta g(y)\sum_{i=1}^{t-1}\mathbb{1}_{\{x_i=y\}} - \Delta g(y)\mathbb{1}_{\{b=y\}}\\
&= M_{t-1} \underbrace{-g(b) + \frac{1}{\mathbf{d}}\sum_{z\sim b} g(z) - \mathbb{1}_{\{b=y\}}\Delta g(y)}_{=0 \text{ because } g \text{ is harmonic in } K\setminus\{y\}} = M_{t-1};
\end{aligned}$$

- if at step t an oil-water pair jumps from $x_t = b \in K$, let

$\mathcal{N}^o_{b,t} = \{z \in V : z \sim b, \tilde{\eta}^o_{t-1}(z) - \tilde{\eta}^w_{t-1}(z) \geq 0\}$ and

$\mathcal{N}^w_{b,t} = \{z \in V : z \sim b, \tilde{\eta}^o_{t-1}(z) - \tilde{\eta}^w_{t-1}(z) < 0\}$, both $\tilde{\mathcal{F}}_{t-1}$-measurable.

$$\begin{aligned}
\tilde{E}_{K,\sigma}(M_t|\tilde{\mathcal{F}}_{t-1}) = &\sum_{\substack{x\in K\\ x\nsim b\\ x\neq b}} \left(\tilde{\eta}^o_{t-1}(x) \wedge \tilde{\eta}^w_{t-1}(x) + \tilde{\eta}^g_{t-1}(x)\right) g(x)+\\
&+ \left(\tilde{\eta}^o_{t-1}(b) \wedge \tilde{\eta}^w_{t-1}(b) + \tilde{\eta}^g_{t-1}(b) - 1\right) g(b)+\\
&+ \sum_{\substack{x\in K\\ x\sim b}} \tilde{E}_{K,\sigma}\left(\tilde{\eta}^o_{t-1}(x) \wedge \tilde{\eta}^w_{t-1}(x) + \tilde{\eta}^g_{t-1}(x)\middle|\tilde{\mathcal{F}}_{t-1}\right) g(x)+\\
&- \Delta g(y)\sum_{i=1}^{t-1}\mathbb{1}_{\{x_i=y\}} - \Delta g(y)\mathbb{1}_{\{b=y\}}
\end{aligned}$$

and since, denoting in the following sum the destination of the oil and the water particle respectively with z_o and z_w, we have

$$\sum_{\substack{x\in K\\ x\sim b}} \tilde{E}_{K,\sigma}\left(\tilde{\eta}^o_{t-1}(x)\wedge\tilde{\eta}^w_{t-1}(x)+\tilde{\eta}^g_{t-1}(x)\Big|\tilde{\mathcal{F}}_{t-1}\right)g(x)$$

$$=\sum_{\substack{x\in K\\ x\sim b}}\left(\tilde{\eta}^o_{t-1}(x)\wedge\tilde{\eta}^w_{t-1}(x)+\tilde{\eta}^g_{t-1}(x)\right)g(x)+$$

$$+\frac{1}{\mathbf{d}^2}\sum_{\substack{z_w\in\mathcal{N}^w_{b,t}\\ z_o\in\mathcal{N}^w_{b,t}}} g(z_o)+\frac{1}{\mathbf{d}^2}\sum_{\substack{z_w\in\mathcal{N}^o_{b,t}\\ z_o\in\mathcal{N}^o_{b,t}}} g(z_w)+\frac{1}{\mathbf{d}^2}\sum_{\substack{z_w\in\mathcal{N}^o_{b,t}\\ z_o\in\mathcal{N}^w_{b,t}}}(g(z_o)+g(z_w))$$

(where the last two sums come from the fact that when $z_w\in\mathcal{N}^o_{b,t}$, both if $\tilde{\eta}^o_{t-1}(z_w)>\tilde{\eta}^w_{t-1}(z_w)$ or if $\tilde{\eta}^w_{t-1}(z_w)$, the quantity $\tilde{\eta}^o_{t-1}(z_w)\wedge\tilde{\eta}^w_{t-1}(z_w)+\tilde{\eta}^g_{t-1}(z_w)$ is incremented by 1 at time t, respectively because a new oil-water pair is formed or because a new ghost particle is created), we get

$$\tilde{E}_{K,\sigma}(M_t|\tilde{\mathcal{F}}_{t-1})=M_{t-1}-g(b)+\frac{1}{\mathbf{d}^2}\sum_{\substack{z_w\in\mathcal{N}^w_{b,t}\\ z_o\in\mathcal{N}^w_{b,t}}} g(z_o)+\frac{1}{\mathbf{d}^2}\sum_{\substack{z_w\in\mathcal{N}^o_{b,t}\\ z_o\in\mathcal{N}^o_{b,t}}} g(z_w)+$$

$$+\frac{1}{\mathbf{d}^2}\sum_{\substack{z_w\in\mathcal{N}^o_{b,t}\\ z_o\in\mathcal{N}^w_{b,t}}}(g(z_o)+g(z_w))-\Delta g(y)\mathbb{1}_{\{b=y\}}$$

$$=M_{t-1}-g(b)+\frac{|\mathcal{N}^w_{b,t}|+|\mathcal{N}^o_{b,t}|}{\mathbf{d}^2}\sum_{z\sim b}g(z)-\Delta g(y)\mathbb{1}_{\{b=y\}}$$

$$=M_{t-1}\underbrace{-g(b)+\frac{1}{\mathbf{d}}\sum_{z\sim b}g(z)-\Delta g(y)\mathbb{1}_{\{b=y\}}}_{=0\text{ because } g \text{ is harmonic in } K\setminus\{y\}}=M_{t-1}.$$

$\implies \{M_t\}_{t\in\mathbb{N}_0}$ is a martingale.

Recalling that T is the first time in which K is stable and there are no ghost particles in K and the fact that it is finite a.s., therefore $E(T)<\infty$ a.s., since $\{M_t\}_{t\in\mathbb{N}_0}$ has bounded increments, for the optional sampling theorem

$$\tilde{E}_{K,\sigma}(M_T)=\tilde{E}_{K,\sigma}(M_0)\overset{\substack{\text{the initial configuration is } \sigma \text{ with no ghost particles}\\ \downarrow}}{=}\sum_{x\in K}\left(\tilde{\eta}^o_0(x)\wedge\tilde{\eta}^w_0(x)\right)g(x)$$

$$\implies -\Delta g(y)\tilde{E}_{K,\sigma}[\tilde{m}(y)]=-\Delta g(y)\tilde{E}_{K,\sigma}\left(\sum_{t=1}^{\infty}\mathbb{1}_{\{x_t=y\}}\right)$$

$$=\tilde{E}_{K,\sigma}(M_T)=\sum_{x\in K}\left(\tilde{\eta}^o_0(x)\wedge\tilde{\eta}^w_0(x)\right)g(x)$$

$$
\begin{aligned}
\implies \tilde{E}_{K,\sigma}[\tilde{m}(y)] &= \sum_{x\in K} \left(\tilde{\eta}_0^o(x) \wedge \tilde{\eta}_0^w(x)\right) g(x) \frac{1}{-\Delta g(y)} \\
&\overset{(3.3)}{\overset{\downarrow}{=}} \sum_{x\in K} \left(\tilde{\eta}_0^o(x) \wedge \tilde{\eta}_0^w(x)\right) g(x) \frac{1}{1 - P_y(\tau^+_{V\setminus\{y\}} < \tau_K)} \\
&= \sum_{x\in K} \left(\tilde{\eta}_0^o(x) \wedge \tilde{\eta}_0^w(x)\right) g(x) G_K(y,y) \\
&= \sum_{x\in K} \left(\tilde{\eta}_0^o(x) \wedge \tilde{\eta}_0^w(x)\right) P_x(\tau_{V\setminus\{y\}} < \tau_K) G_K(y,y) \\
&= \sum_{x\in K} \left(\tilde{\eta}_0^o(x) \wedge \tilde{\eta}_0^w(x)\right) G_K(x,y).
\end{aligned}
$$

□

3.3 Proof of Theorem 6

In this section we prove Theorem 6 using, as we said before, a contradiction, which will involve the ghost-pair stabilization, where ghosts were introduced to compensate the loss of pairs we have when an oil or a water particle falls into a hole, in such a way that the sum of the number of oil-water pairs and ghost particles which are present in the system at any given time is a martingale. The proof of Theorem 6 supposes, per impossible, that the system is active a.s., but for Lemma 13 this implies that a large number of pairs is lost in most vertices, then a large number of ghost particles is produced in those vertices and the proof will show that it is not possible to create such a large number of ghosts if we start with a finite density of pairs, leading to the desired contradiction.
In the proof we will need two preliminary results, which we will state in the following two lemmas.

Lemma 15 *Let $B \subseteq Q \subseteq V$ such that $\mathbf{0} \in Q$. Then*

$$
\sum_{y\in B} G_Q(y,\mathbf{0}) = G_Q(\mathbf{0},\mathbf{0}) \left[\delta_{\mathbf{0}\in B} + E_{\mathbf{0}}\left(\left|\{t \le \tau^+_{Q\setminus\{\mathbf{0}\}} : X_t \in B \setminus \{\mathbf{0}\}\}\right|\right)\right]
$$

where $\delta_{\mathbf{0}\in B} = \begin{cases} 1 & \textit{if } \mathbf{0} \in B \\ 0 & \textit{if } \mathbf{0} \notin B \end{cases}.$

Proof. For every $y \in B$, $G_Q(y, \mathbf{0}) \underset{\substack{\uparrow \\ \text{Lemma 3}}}{=} P_y(\tau_{V\setminus\{\mathbf{0}\}} < \tau_Q) G_Q(\mathbf{0}, \mathbf{0})$, therefore

$$\begin{aligned}\sum_{y\in B} G_Q(y,\mathbf{0}) &= G_Q(\mathbf{0},\mathbf{0}) \sum_{y\in B} P_y(\tau_{V\setminus\{\mathbf{0}\}} < \tau_Q) \\ &= G_Q(\mathbf{0},\mathbf{0}) \left[\delta_{\mathbf{0}\in B} + \sum_{y\in B\setminus\{\mathbf{0}\}} P_y(\tau_{V\setminus\{\mathbf{0}\}} < \tau_Q)\right].\end{aligned}$$

Let T be the number of times the random walk goes back to y before arriving in $\mathbf{0}$ or exiting Q and let $\tilde{\tau}_i$ be the i-th time the walk returns in y, $i \in \{1, \dots, T\}$. Then we have

$$\begin{aligned}P_y(\tau_{V\setminus\{\mathbf{0}\}} < \tau_Q) &= \sum_{n=0}^{\infty} P_y(\tau_{V\setminus\{\mathbf{0}\}} < \tau_Q, T = n) \\ &= \sum_{n=0}^{\infty} P_y(\tau_{V\setminus\{\mathbf{0}\}} < \tau_Q, T = n, X_t \neq y \ \ \forall\, t \in [\tilde{\tau}_T + 1, \tau_{V\setminus\{\mathbf{0}\}}]) \\ &= \sum_{n=0}^{\infty} P_y(T = n) P_y(\tau_{V\setminus\{\mathbf{0}\}} < \tau_Q, X_t \neq y \ \ \forall\, t \in [\tilde{\tau}_T + 1, \tau_{V\setminus\{\mathbf{0}\}}] | T = n) \\ &\underset{\substack{\uparrow \\ \text{Strong Markov property}}}{=} P_y(\tau^+_{V\setminus\{y\}} < \tau_{Q\setminus\{\mathbf{0}\}})^n P_y(\tau_{V\setminus\{\mathbf{0}\}} < \tau^+_{Q\setminus\{y\}}) \\ &= \frac{1}{1 - P_y(\tau^+_{V\setminus\{y\}} < \tau_{Q\setminus\{\mathbf{0}\}})} P_y(\tau_{V\setminus\{\mathbf{0}\}} < \tau^+_{Q\setminus\{\mathbf{0}\}}) \\ &= \frac{1}{1 - P_y(\tau^+_{V\setminus\{y\}} < \tau_{Q\setminus\{\mathbf{0}\}})} P_{\mathbf{0}}(\tau_{V\setminus\{y\}} < \tau^+_{Q\setminus\{\mathbf{0}\}})\end{aligned}$$

where the last identity follows from the reversibility of the simple random walk. Thus we have

$$\sum_{y\in B} G_Q(y,\mathbf{0}) = G_Q(\mathbf{0},\mathbf{0}) \left[\delta_{\mathbf{0}\in B} + \sum_{y\in B\setminus\{\mathbf{0}\}} P_{\mathbf{0}}(\tau_{V\setminus\{y\}} < \tau^+_{Q\setminus\{\mathbf{0}\}}) \frac{1}{1 - P_y(\tau^+_{V\setminus\{y\}} < \tau_{Q\setminus\{\mathbf{0}\}})}\right].$$

Since for every $y \neq \mathbf{0}$ we have $P_{\mathbf{0}}(\tau_{V\setminus\{y\}} < \tau^+_{Q\setminus\{\mathbf{0}\}}) = E_{\mathbf{0}}\left(\mathbb{1}_{\{\tau_{V\setminus\{y\}} < \tau^+_{Q\setminus\{\mathbf{0}\}}\}}\right)$, we get

$$\sum_{y\in B} G_Q(y,\mathbf{0}) = G_Q(\mathbf{0},\mathbf{0}) \left[\delta_{\mathbf{0}\in B} + \sum_{y\in B\setminus\{\mathbf{0}\}} E_{\mathbf{0}}\left(\mathbb{1}_{\{\tau_{V\setminus\{y\}} < \tau^+_{Q\setminus\{\mathbf{0}\}}\}}\right) \frac{1}{1 - P_y(\tau^+_{V\setminus\{y\}} < \tau_{Q\setminus\{\mathbf{0}\}})}\right].$$

Now,

$$\underset{\displaystyle E_{\mathbf{0}}\left(\left|\{t\le\tau^+_{Q\setminus\{\mathbf{0}\}}:X_t=y\}\right|\right)}{\underset{\shortparallel}{E_{\mathbf{0}}\left(\sum_{t=0}^{\tau^+_{Q\setminus\{\mathbf{0}\}}}\mathbb{1}_{\{X_t=y\}}\right)}} = \underbrace{E_{\mathbf{0}}\left(\sum_{t=0}^{\tau^+_{Q\setminus\{\mathbf{0}\}}}\mathbb{1}_{\{X_t=y\}}\middle|\tau_{V\setminus\{y\}}<\tau^+_{Q\setminus\{\mathbf{0}\}}\right)}_{E_y\left(\sum_{t=0}^{\tau^+_{Q\setminus\{\mathbf{0}\}}}\mathbb{1}_{\{X_t=y\}}\right)} P_{\mathbf{0}}(\tau_{V\setminus\{y\}}<\tau^+_{Q\setminus\{\mathbf{0}\}})$$

$$= E_{\mathbf{0}}\left(\mathbb{1}_{\{\tau_{V\setminus\{y\}}<\tau^+_{Q\setminus\{\mathbf{0}\}}\}}\right)\underbrace{E_y\left(\sum_{t=0}^{\tau_{Q\setminus\{\mathbf{0}\}}}\mathbb{1}_{\{X_t=y\}}\right)}_{G_{Q\setminus\{\mathbf{0}\}}(y,y)}$$

$$= E_{\mathbf{0}}\left(\mathbb{1}_{\{\tau_{V\setminus\{y\}}<\tau^+_{Q\setminus\{\mathbf{0}\}}\}}\right)\frac{1}{1-P_y(\tau^+_{V\setminus\{y\}}<\tau_{Q\setminus\{\mathbf{0}\}})},$$

therefore we have

$$\begin{aligned}\sum_{y\in B}G_Q(y,\mathbf{0}) &= G_Q(\mathbf{0},\mathbf{0})\left[\delta_{\mathbf{0}\in B}+\sum_{y\in B\setminus\{\mathbf{0}\}}E_{\mathbf{0}}\left(\left|\{t\le\tau^+_{Q\setminus\{\mathbf{0}\}}:X_t=y\}\right|\right)\right]\\ &= G_Q(\mathbf{0},\mathbf{0})\left[\delta_{\mathbf{0}\in B}+E_{\mathbf{0}}\left(\left|\{t\le\tau^+_{Q\setminus\{\mathbf{0}\}}:X_t\in B\setminus\{\mathbf{0}\}\}\right|\right)\right].\end{aligned}$$

□

Lemma 16 *For each $D\in\mathbb{N}$ there exists $L_0=L_0(D)$ large enough such that, for every $L>L_0$,*

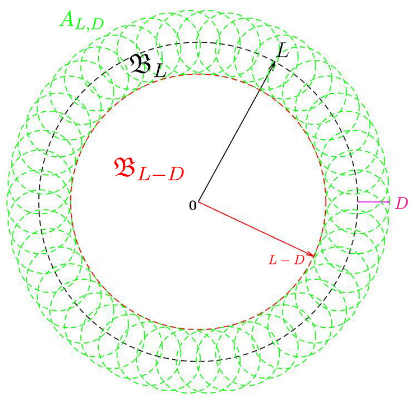
$$\sum_{x\in B_L}G_{B_L}(x,\mathbf{0})<10\sum_{\substack{x\in B_L:\\ B_D(x)\subseteq B_L}}G_{B_L}(x,\mathbf{0})$$

where, for $r>0$, $B_r(x)=\{z\in V:\|z-x\|<r\}$ and $B_L=B_L(\mathbf{0})$.

Proof. Let $L_0\ge D(1+d^D)$ and, for $L\ge L_0$, let

$$A_{L,D}=\bigcup_{x\in\partial B_L}B_D(x)$$

be the union of the balls with radius D and center in the points that are in ∂B_L. By construction, $B_{L-D}\subseteq B_L\setminus A_{L,D}$.



Notice that it sufficies to show

$$-\sum_{y\in B_L\setminus A_{L,D}} 9G_{B_L}(y,\mathbf{0}) + \sum_{y\in A_{L,D}\cap B_L} G_{B_L}(y,\mathbf{0}) < 0, \tag{3.4}$$

so that we have

$$\begin{aligned}\sum_{x\in B_L} G_{B_L}(x,\mathbf{0}) &= \sum_{x\in B_L\cap A_{L,D}} G_{B_L}(x,\mathbf{0}) + \sum_{x\in B_L\setminus A_{L,D}} G_{B_L}(x,\mathbf{0})\\ &< \sum_{x\in B_L\setminus A_{L,D}} 9G_{B_L}(x,\mathbf{0}) + \sum_{x\in B_L\setminus A_{L,D}} G_{B_L}(x,\mathbf{0})\\ &= 10\sum_{x\in B_L\setminus A_{L,D}} G_{B_L}(x,\mathbf{0}) \le 10\sum_{x\in \overline{B}_{L-D}} G_{B_L}(x,\mathbf{0})\\ &= 10\sum_{\substack{x\in B_L:\\ B_D(x)\subseteq B_L}} G_{B_L}(x,\mathbf{0}).\end{aligned}$$

For every $B\subseteq Q\subseteq V$ such that $\mathbf{0}\in Q$, let

$$E_{\mathbf{0}}\mathcal{R}(B,Q^c) := E_{\mathbf{0}}\left(\sum_{t=0}^{\tau_{Q\setminus\{\mathbf{0}\}}-1} \mathbb{1}_{\{X_t\in B\setminus\{\mathbf{0}\}\}}\right)$$

be the expected "range" (from which the symbol $\mathcal{R}$) performed by a random walk started at $\mathbf{0}$ inside the set $B\setminus\{\mathbf{0}\}$ before leaving Q or coming back to $\{\mathbf{0}\}$.

Applying Lemma 15 to (3.4) first with $B\to B_L\setminus A_{L,D}$, $Q\to B_L$ and then with $B\to A_{L,D}\cap B_L$, $Q\to B_L$, our new objective becomes to show that

$$-9\left[\delta_{\mathbf{0}\in B_L\setminus A_{L,D}} + E_{\mathbf{0}}\mathcal{R}(B_L\setminus A_{L,D}, B_L^c)\right] + \delta_{\mathbf{0}\in A_{L,D}\cap B_L} + E_{\mathbf{0}}\mathcal{R}(A_{L,D}\cap B_L, B_L^c) < 0$$

which is actually, since $B_{L-D} \subseteq B_L \setminus A_{L,D}$ and therefore $\delta_{\mathbf{0} \in B_L \setminus A_{L,D}} = 1$ and $\delta_{\mathbf{0} \in A_{L,D} \cap B_L} = 0$,

$$-9\left[1 + E_{\mathbf{0}}\mathcal{R}(B_L \setminus A_{L,D}, B_L^c)\right] + E_{\mathbf{0}}\mathcal{R}(A_{L,D} \cap B_L, B_L^c) < 0,$$

namely

$$9 + 9E_{\mathbf{0}}\mathcal{R}(B_L \setminus A_{L,D}, B_L^c) > E_{\mathbf{0}}\mathcal{R}(A_{L,D} \cap B_L, B_L^c)$$

that is implied by the inequality

$$E_{\mathbf{0}}\mathcal{R}(B_L \setminus A_{L,D}, B_L^c) \geq E_{\mathbf{0}}\mathcal{R}(A_{L,D} \cap B_L, B_L^c),$$

but $B_{L-D} \subseteq B_L \setminus A_{L,D}$ hence

$$\begin{aligned}
E_{\mathbf{0}}\mathcal{R}(B_L \setminus A_{L,D}, B_L^c) &= E_{\mathbf{0}}\left(\sum_{t=0}^{\tau_{B_L \setminus \{\mathbf{0}\}}-1} \mathbb{1}_{\{X_t \in B_L \setminus (A_{L,D} \cup \{\mathbf{0}\})\}}\right) \\
&= E_{\mathbf{0}}\left(\sum_{t=0}^{\tau_{B_L \setminus \{\mathbf{0}\}}-1} \mathbb{1}_{\{X_t \in B_L \setminus (A_{L,D} \cup \{\mathbf{0}\})\}} \middle| \tau_{B_L \setminus A_{L,D}} < \tau^+_{V \setminus \{\mathbf{0}\}}\right) P_{\mathbf{0}}(\tau_{B_L \setminus A_{L,D}} < \tau^+_{V \setminus \{\mathbf{0}\}}) \\
&\underset{\substack{\uparrow \\ B_{L-D} \subseteq B_L \setminus A_{L,D}}}{\geq} (L-D)P_{\mathbf{0}}(\tau_{B_L \setminus A_{L,D}} < \tau^+_{V \setminus \{\mathbf{0}\}}) \underset{\substack{\uparrow \\ L \geq L_0 \geq D(1+d^D) \Rightarrow L-D \geq Dd^D}}{\geq} Dd^D P_{\mathbf{0}}(\tau_{B_L \setminus A_{L,D}} < \tau^+_{V \setminus \{\mathbf{0}\}})
\end{aligned}$$

therefore our new goal is to prove that

$$Dd^D P_{\mathbf{0}}(\tau_{B_L \setminus A_{L,D}} < \tau^+_{V \setminus \{\mathbf{0}\}}) \geq E_{\mathbf{0}}\mathcal{R}(A_{L,D} \cap B_L, B_L^c). \tag{3.5}$$

Let $T_{L,D} = \left|\{t \geq 0 : X_t \in A_{L,D} \cap B_L\}\right|$ be the number of times the walk steps on $A_{L,D} \cap B_L$. By construction, every vertex $x \in A_{L,D} \cap B_L$ is such that $d(x, B_L^c) \leq D$. Then we can prove by induction that, for every $k \in \mathbb{N}$, $x \in A_{L,D} \cap B_L$,

$$P_x(T_{L,D} > kD) \leq P(S > k)$$

with S random variable geometrically distributed with success parameter $\dfrac{1}{d^D}$. Indeed, for $k = 1$,

$$P_x(T_{L,D} \leq D) \underset{\substack{\uparrow \\ \text{taking each step in the same direction to exit } B_L}}{\geq} \frac{1}{d^{d(x,B_L^c)}} \geq \frac{1}{d^D} = P(S = 1),$$

hence

$$P_x(T_{L,D} > D) \leq 1 - \frac{1}{d^D}.$$

For $k \in \mathbb{N} \setminus \{1\}$, supposing that $P_x(T_{L,D} > (k-1)D) \le \left(1 - \frac{1}{d^D}\right)^{k-1}$ and calling

$$T_k^{(L,D)} = \inf\left\{ t \ge 0 : \sum_{t=0}^{T_k^{(L,D)}-1} \mathbb{1}_{\{X_t \in A_{L,D} \cap B_L\}} = kD,\ X_t \in A_{L,D} \cap B_L \right\}$$

the first time the walk reaches $A_{L,D} \cap B_L$ after having already reached it kD times,

$$\begin{aligned}
P_x(T_{L,D} > kD) &= P_x(T_{L,D} > kD | T_{L,D} > (k-1)D) \underbrace{P_x(T_{L,D} > (k-1)D)}_{\le \left(1-\frac{1}{d^D}\right)^{k-1}} \\
&\le \sum_{y \in A_{L,D} \cap B_L} \underbrace{P_y(T_{L,D} > D)}_{\le 1 - \frac{1}{d^D}} P_x(X_{T_{k-1}^{(L,D)}} = y) \left(1 - \frac{1}{d^D}\right)^{k-1} \\
&\le \left(1 - \frac{1}{d^D}\right)^k .
\end{aligned}$$

Therefore, the random variable $\dfrac{T_{L,D}}{D}$ is stochastically dominated by S, thus

$$E_x\left(\frac{T_{L,D}}{D}\right) \le E_x(S) = d^D$$

which is equivalent to say that

$$E_x(T_{L,D}) \le D d^D \qquad \forall\, x \in A_{L,D} \cap B_L. \tag{3.6}$$

Now,

$$\begin{aligned}
E_{\mathbf{0}} \mathcal{R}(A_{L,D} \cap B_L, B_L^c) &= E_{\mathbf{0}} \left(\sum_{t=0}^{\tau_{B_L \setminus \{\mathbf{0}\}}-1} \mathbb{1}_{\{X_t \in A_{L,D} \cap B_L\}} \right) \\
&= E_{\mathbf{0}} \left(\sum_{t=0}^{\tau_{B_L \setminus \{\mathbf{0}\}}-1} \mathbb{1}_{\{X_t \in A_{L,D} \cap B_L\}} \middle| \tau_{B_L \setminus A_{L,D}} < \tau^+_{V \setminus \{\mathbf{0}\}} \right) P_{\mathbf{0}}(\tau_{B_L \setminus A_{L,D}} < \tau^+_{V \setminus \{\mathbf{0}\}}) \\
&= E_{\mathbf{0}}(T_{L,D} | \tau_{B_L \setminus A_{L,D}} < \tau^+_{V \setminus \{\mathbf{0}\}}) P_{\mathbf{0}}(\tau_{B_L \setminus A_{L,D}} < \tau^+_{V \setminus \{\mathbf{0}\}}) \\
&= \sum_{x \in A_{L,D} \cap B_L} E_x(T_{L,D}) P_{\mathbf{0}}(X_{\tau_{B_{L-D}}} = x | \tau_{B_L \setminus A_{L,D}} < \tau^+_{V \setminus \{\mathbf{0}\}}) P_{\mathbf{0}}(\tau_{B_L \setminus A_{L,D}} < \tau^+_{V \setminus \{\mathbf{0}\}}) \\
&\underset{\substack{\uparrow \\ (3.6)}}{\le} P_{\mathbf{0}}(\tau_{B_L \setminus A_{L,D}} < \tau^+_{V \setminus \{\mathbf{0}\}}) D d^D
\end{aligned}$$

that is (3.5). $\square$

Now we are ready to begin the proof of Theorem 6.

Proof of Theorem 6. For fixed L, we stabilize B_L following the ghost-pair stabilization, therefore, during the process each time a water particle falls into a hole a new ghost particle is created in that vertex; the ghost particles perform simple independent random walks which are killed upon leaving B_L.
For every $x \in B_L$ let

- ◇ $\tilde{m}_L(x)$ the number of oil-water pairs or ghost particles which jump from x during the stabilization of B_L;
- ◇ $m_L(x)$ the times x fires during the stabilization of B_L;
- ◇ $w_L(x)$ the number of ghost particles which jump from x during the stabilization of B_L;
- ◇ $H_L(x)$ the number of ghost particles created in x during the stabilization of B_L.

For $L \in \mathbb{N}$ for Lemma 14, if $\sigma = (\tilde{\eta}_0^o, \tilde{\eta}_0^w)$ is the initial configuration, $\tilde{E}_{B_L,\sigma}[\tilde{m}(y)] = \sum_{x \in B_L} (\tilde{\eta}_0^o(x) \wedge \tilde{\eta}_0^w(x))\, G_{B_L}(x,y)$. Averaging over the initial configuration σ and noticing that for every $x \in V$ the expected value of $\tilde{\eta}_0^o(x) \wedge \tilde{\eta}_0^w(x)$ cannot be greater than the expected value of $\tilde{\eta}_0^o(x) + \tilde{\eta}_0^w(x)$ which is μ, we have, indicating with $\tilde{E}_\nu$ the expectation of the measure which is defined in the enlarged probability space of oil, water and ghost particles,

$$\tilde{E}_\nu[\tilde{m}_L(x)] \leq \sum_{y \in B_L} \mu G_{B_L}(y,x). \tag{3.7}$$

Moreover, from linearity of the expected value and from the fact that every ghost particle performs an independent simple random walk until it leaves B_L, it follows that

$$\tilde{E}_\nu[w_L(x)] = \sum_{y \in B_L} \tilde{E}_\nu[H_L(y)] G_{B_L}(y,x). \tag{3.8}$$

Assume now, *per impossible*, that the system is active a.s., then for Lemma 13 for every $\varepsilon, M > 0$ there exists D large enough such that

$$\inf_{\substack{K \subset V \text{ finite:} \\ d(\mathbf{0},K^c) > D}} \; \inf_{\substack{F_K \\ \text{strategy}}} \mathcal{P}_\nu \left(H_{K,F_K}(\mathbf{0}) > M \right) \geq 1 - \varepsilon$$

and since G is vertex-transitive, this not only holds for $\mathbf{0}$ but for every $x \in V$, namely for each $\varepsilon, M > 0$ there exists D large enough such that

$$\inf_{\substack{K\subset V \text{ finite:} \\ d(x,K^c)>D}} \inf_{\substack{F_K \\ \text{strategy}}} \mathcal{P}_\nu \left(H_{K,F_K}(x) > M\right) \geq 1-\varepsilon.$$

Taking $L > D$ we have that for every $x \in B_L$ such that $d(x, B_L^c) > D$ (i.e. such that $B_D(x) \subseteq B_L$)

$$\inf_{\substack{F_{B_L} \\ \text{strategy}}} \mathcal{P}_\nu \left(H_{B_L,F_{B_L}}(x) > M\right) \geq 1-\varepsilon.$$

For $M = 10\mu$ we get $\inf_{\substack{F_{B_L} \\ \text{strategy}}} \mathcal{P}_\nu \left(H_{B_L,F_{B_L}}(x) > 10\mu\right) \geq 1-\varepsilon$, which implies, since a ghost particle is created every time a water falls into a hole, that for every $\varepsilon > 0$

$$\inf_{\substack{F_{B_L} \\ \text{strategy}}} \mathcal{P}_\nu \left(H_L(x) > 10\mu\right) \geq 1-\varepsilon \xrightarrow[\varepsilon\to 0]{} 1.$$

$\Rightarrow$ for every strategy F_{B_L} we get $\mathcal{P}_\nu\left(H_L(x) > 10\mu\right) = 1$, thus, since $H_L(x)$ is a $\mathbb{N}_0$-valued random variable, $\mathcal{P}_\nu\left(H_L(x) \geq \lceil 10\mu\rceil\right) = 1$ and $\mathcal{P}_\nu\left(H_L(x) < \lceil 10\mu\rceil\right) = 0$

$$\begin{aligned}
\Rightarrow\ \tilde{E}_\nu[H_L(x)] &= \sum_{l=1}^{\infty} l\mathcal{P}_\nu\left(H_L(x) = l\right) = \sum_{l=\lceil 10\mu\rceil}^{\infty} l\mathcal{P}_\nu\left(H_L(x) = l\right) \\
&\geq \lceil 10\mu\rceil \sum_{l=\lceil 10\mu\rceil}^{\infty} \mathcal{P}_\nu\left(H_L(x) = l\right) = \lceil 10\mu\rceil \underbrace{\mathcal{P}_\nu\left(H_L(x) \geq \lceil 10\mu\rceil\right)}_{1}
\end{aligned}$$

$$\Rightarrow\ \tilde{E}_\nu[H_L(x)] \geq 10\mu \qquad \forall\, x \in B_L \text{ such that } B_D(x) \subseteq B_L. \tag{3.9}$$

Obviously,

$$\underbrace{m_L(x)}_{\substack{\text{number of times} \\ \text{a pair} \\ \text{jumps from } x}} = \underbrace{\tilde{m}_L(x)}_{\substack{\text{number of times} \\ \text{a pair} \\ \text{or a ghost particle} \\ \text{jump from } x}} - \underbrace{w_L(x)}_{\substack{\text{number of times} \\ \text{a ghost particle} \\ \text{jumps from } x}}$$

and we have

$$\begin{aligned}
\tilde{E}_\nu[m_L(x)] &\overset{\substack{(3.7)\text{ and }(3.8) \\ \downarrow}}{\leq} \sum_{y\in B_L} \mu G_{B_L}(y,x) - \sum_{y\in B_L} \tilde{E}_\nu[H_L(y)] G_{B_L}(y,x) \\
&\leq \sum_{y\in B_L} \mu G_{B_L}(y,x) - \sum_{\substack{y\in B_L: \\ B_D(y)\subseteq B_L}} \tilde{E}_\nu[H_L(y)] G_{B_L}(y,x) \\
&\underset{\substack{\uparrow \\ (3.9)}}{\leq} \mu \left[\sum_{y\in B_L} G_{B_L}(y,x) - 10 \sum_{\substack{y\in B_L: \\ B_D(y)\subseteq B_L}} G_{B_L}(y,x)\right].
\end{aligned}$$

For Lemma 16, we conclude that $\tilde{E}_\nu[m_L(\mathbf{0})] < 0$ for L large enough.
This is the desired contradiction because the number of times an oil-water pair jumps from $\mathbf{0}$ cannot be negative

$$\implies \mathbb{P}_\nu(\text{the system is active}) < 1 \xRightarrow[\text{0-1 law}]{} \mathbb{P}_\nu(\text{the system is active}) = 0.$$

□

3.4 Further observations on the Oil and Water model

It is easy to prove that the Oil and Water model where the oil and water particles jump on the neighboring vertices with different (but still independent) probability distributions is an abelian network. Indeed, the mathematical formalization of this model is the same as the one in the previous section for the classic Oil and Water dynamics, except for the fact that in the definition of $\mathcal{P}$ we set, in place of (3.2), the following definition

$$\mathcal{P}(\tau^{x,j,o} = \tau^o_{x,y_o}) = p^o_{y_o}, \qquad \mathcal{P}(\tau^{x,j,w} = \tau^w_{x,y_w}) = p^w_{y_w} \qquad \forall\ y_o, y_w \sim x,$$

with $\sum\limits_{v\sim x} p^i_v = 1$ and $p^i_u > 0$ for every $u \in V$, $i \in \{o, w\}$, where p^o_v and p^w_v are the probabilities that respectively the oil and the water particle chooses v to jump on when a firing happens in one of its neighbors.
The proof of this being an abelian system is also the same for the classic Oil and Water model, except for the fact that the random variable $X_{\{v,u\}}$ is defined with the probabilities p^o_u and p^w_u:

$$X_{\{v,u\}}(\omega) = \begin{cases} o_u & \text{if } \omega \in B_1 \\ w_u & \text{if } \omega \in B_2 \\ o_u w_u & \text{if } \omega \in B_3 \\ w_u o_u & \text{if } \omega \in B_4 \\ \varepsilon & \text{if } \omega \in B_5 \end{cases}$$

where $B_1, B_2, B_3, B_4, B_5 \in \mathcal{H}$ such that $\overline{P}(B_1) = p^o_u(1 - p^w_u)$, $\overline{P}(B_2) = p^w_u(1 - p^o_u)$, $\overline{P}(B_3) = \overline{P}(B_4) = p^o_u p^w_u$, $\overline{P}(B_5) = 1 - p^o_u - p^w_u$.

Hence, this is another abelian network and satisfies the abelian property.

In the same way, we can expand this model even further and define a new one that, like Oil and Water, is a non-unary network, in the sense that there is more than one type of particle (while internal DLA is an example of unary system). In particular, we can add a third type, letting the vertices fire if and only if a triplet composed by particles of three different types is located on it and in the firing each one of these particles will independently choose a neighboring vertex and jump on it; in the same way we did for Oil and Water, it is possible to prove that, independently from the distributions according to which the three types of particle choose the neighbor to jump on, this is an abelian network.

Since in Chapter 3 we have found, through Activated Random Walks, an answer to the phase transition problem between fixation and activity for a model which follows the same evolutions rules as internal DLA, we would like to find a result on the shape of the aggregate (intended as the set of vertices touched by at least one particle) for the Oil and Water model when we start with all the particles on the origin. This dynamics has not a complete theory about this question, however there are studies which propose conjectures, proven just in dimension 1, about the shape of the final configuration, like [8] by Candellero, Ganguly, Hoffman and Levine; we will exploit and support through simulations these conjectures in Chapter 4. In the same paper, they also provide a scaling limit for the odometer in dimension 1.

Chapter 4

Conjectures on the aggregate for Oil and Water and IDLAA

On $G = \left(\mathbb{Z}^d, \bigcup_{x\in\mathbb{Z}^d} \{\{x,y\} : \|x-y\| = 1\}\right)$, starting from a configuration $\eta_0 \in \mathbb{N}_0^{(\mathbb{Z}^d)} \times \mathbb{N}_0^{(\mathbb{Z}^d)}$ with n oil-water pairs on the origin, namely

$$\eta_0^o(x) = \eta_0^w(x) = n\delta_{\mathbf{0}}(x)$$

for every vertex $x \in \mathbb{Z}^d$, we define the first time in which all the sites are stable

$$\tau = \inf\{t \in \mathbb{N}_0 : \eta_t^o(x) \wedge \eta_t^w(x) = 0 \;\; \forall\, x \in \mathbb{Z}^d\} < \infty$$

and the aggregate we get when the Oil and Water system stabilizes, i.e. the set of vertices visited by at least one particle,

$$A(n) := \{x \in \mathbb{Z}^d : \exists\, t \in [0,\tau] \text{ such that } \eta_t^o(x) \vee \eta_t^w(x) > 0\}.$$

In order to study the shape of the aggregate in the Oil and Water model, we introduce now a simpler system that we will call *internal diffusion limited aggregation with absorption* (IDLAA). From what we know, this model is not present in the scientific literature and here we will raise a conjecture which is supported by some facts. Indeed, we will build an analogy between Oil and Water and IDLAA.
In this new dynamics a simple symmetric random walk on $\mathbb{Z}$ is associated to each site of the graph. We will refer to these random walks as the local random walks of the vertices. Every time a particle is selected, if the local random walk of the vertex

on which it is located is on 0, the particle is removed from the system, otherwise jumps on a neighbor chosen uniformly at random. In both cases the local random walk of the site is updated.

Typically, the vertices visited by many particles are the ones responsible for the absorption of many particles. The mechanism of multiple particles absorption "slows down" the growth of the aggregate of IDLAA with respect to the one of IDLA, where it is near the Euclidean ball with radius $O\left(n^{1/d}\right)$.

4.1 Shape of the aggregate

We will now define IDLAA in a more formal way.

Consider the lattice $\mathbb{Z}^d$ and indicate with $\mathbf{0}$ the origin. Associate to each vertex $v \in \mathbb{Z}^d$

· $\{X_t^v\}_{t\in\mathbb{N}_0}$ a simple symmetric random walk on $\mathbb{Z}$ with $X_0^v = 0$;

· $\{h_t(v)\}_{t\in\mathbb{N}_0}$ a $\mathbb{N}_0$-valued function called *counter*, with $h_t(v) = 0$.

The system evolves according to the following dynamics.

At time $t = 0$ there are $n \in \mathbb{N}$ particles on the origin $\mathbf{0}$.

For each time step $t \in \mathbb{N}_0$ a particle is randomly chosen among all the present particles with uniform distribution; let v be the vertex on which such particle is located. Firstly, sample $X^v_{h_t(v)+1} = \begin{cases} X^v_{h_t(v)} + 1 & \text{with probability } \frac{1}{2} \\ X^v_{h_t(v)} - 1 & \text{with probability } \frac{1}{2} \end{cases}$; then,

- if $X^v_{h_t(v)+1} = 0$ the particle is *absorbed*, namely it is removed from the system;
- if $X^v_{h_t(v)+1} \neq 0$ the particle randomly chooses with uniform distribution a vertex among the neighbors of v and jumps on it;

lastly, set, in both cases, $h_{t+1} = h_t + \delta_v$.

Let T_n be the first time in which all the particles are absorbed in IDLAA, i.e., calling, for $t \in \mathbb{N}$,

$$R_v(t) = \left|\, \{i \in \{1, \dots, t\} : X_i^v = 0\} \,\right|$$

the number of times the random walk associated to the vertex $v \in \mathbb{Z}^d$ returns on zero by time t, which is equal to the number of particles absorbed at v by time t,

$$T_n = \inf \left\{ t \in \mathbb{N} : \sum_{v \in \mathbb{Z}^d} R_v(t) = n \right\} + 1.$$

This random time is finite almost surely because the symmetric simple random walk on $\mathbb{Z}$ is recurrent, therefore it visits the value zero infinitely often, namely any arbitrarily large number of particles can be absorbed in finite time. We define the *aggregate* for IDLAA

$$A(n) := \left\{ v \in \mathbb{Z}^d : h_{T_n}(v) > 0 \right\}$$

which is the set of vertices visited by at least one particle.

We focus on the asymptotic shape of $A(n)$ when n is large.

Since the vertices closer to the origin are typically visited by many particles, this yields that they are responsible for the majority of the absorptions with respect to the vertices which are closer to the boundary of the aggregate.

As already said, we may expect this process to have an aggregate which grows *slower* than the one of IDLA. Indeed, while in IDLA each vertex can "absorb" at most one particle, in IDLAA the number of particles absorbed at each vertex $v \in \mathbb{Z}^d$ increases by one each time the random walk associated to v is on zero, and the expected number of returns to zero by time t of a symmetric simple random walk on $\mathbb{Z}$, $\{X_t\}_{t \in \mathbb{N}_0}$, grows as $\sqrt{t}$, more precisely $E_0\left(\sum_{n=0}^{t} \mathbb{1}_{\{X_t=0\}}\right) \sim \sqrt{\frac{2}{\pi} t} + o(1)$ for $t \in 2\mathbb{N}_0$, $t \to \infty$, indeed for such a t it holds that

$$E_0\left(\sum_{n=0}^{t} \mathbb{1}_{\{X_t=0\}}\right) = (t+1)\binom{t}{t/2}\frac{1}{2^t} \tag{4.1}$$

(for the proof of (4.1) see Appendix), hence

$$E_0\left(\sum_{n=0}^{t} \mathbb{1}_{\{X_t=0\}}\right) = (t+1)\binom{t}{t/2}\frac{1}{2^t} \underset{\substack{\uparrow \\ \text{Stirling's formula (see Appendix)}}}{\sim} \sqrt{\frac{2}{\pi} t} + o(1).$$

Noticing that the counter of each site is updated every time that site is selected leads us to the following remark.

Remark 3 *In IDLAA, the expected number of particles absorbed by each vertex $v \in \mathbb{Z}^d$ grows as the square root of the number of times v is selected.*

This is the crux of our analogy between IDLAA and Oil and Water: we will show that the Oil and Water model has with the oil-water pairs the same behavior IDLAA has with its single particles.
Indeed, in Oil and Water, the number of lost oil-water pairs is the number of times a single particle falls into a hole while the system is active, but the expected value of this number is equal to the expected number of holes created. Then, for $v \in \mathbb{Z}^d$, let $R_j^v = \eta^w_{\tilde{t}_j}(v) - \eta^o_{\tilde{t}_j}(v)$, $j \in \mathbb{N}_0$ with $\tilde{t}_j$ the j-th time in which a neighboring vertex of v fires. In this way, v is a hole at time $\tilde{t}_j$ if and only if $R_j^v = 0$. In the proof of Lemma 13 we have shown that $\{R_j^v\}_{j\in\mathbb{N}_0}$ is a symmetric lazy random walk, namely

$$\mathcal{P}_\nu(R^v_{j+1} = R^v_j + 1 | R^v_j) = \mathcal{P}_\nu(\eta^w_{\tilde{t}_{j+1}}(v) = \eta^w_{\tilde{t}_j}(v) + 1, \eta^o_{\tilde{t}_{j+1}}(v) = \eta^o_{\tilde{t}_j}(v)) = \frac{\mathbf{d}-1}{\mathbf{d}^2},$$
$$\mathcal{P}_\nu(R^v_{j+1} = R^v_j - 1 | R^v_j) = \mathcal{P}_\nu(\eta^w_{\tilde{t}_{j+1}}(v) = \eta^w_{\tilde{t}_j}(v), \eta^o_{\tilde{t}_{j+1}}(v) = \eta^o_{\tilde{t}_j}(v) + 1) = \frac{\mathbf{d}-1}{\mathbf{d}^2},$$
$$\mathcal{P}_\nu(R^v_{j+1} = R^v_j | R^v_j) = 1 - 2\frac{\mathbf{d}-1}{\mathbf{d}^2}.$$

Define the process $\{U_l^v\}_{l\in[0,T_v]}$ which keeps track just of the times R_j^v changes its state, i.e.

$$U_l^v = R^v_{s_l} \qquad \text{with } s_0 = 0,\ s_{l+1} = \inf\{t > s_l : R^v_t \neq R^v_{s_l}\}.$$

$T_v < \infty$ a.s. because the system fixates. Then, $\{U_l^v\}_{l\in[0,T_v]}$ is a symmetric simple random walk on $\mathbb{Z}$ and the number of times v becomes a hole increases by one each time $\{U_l^v\}_{l\in[0,T_v]}$ visits 0. The expected number of returns to zero by time t of a simple symmetric random walk on $\mathbb{Z}$, again for (4.1), grows as $\sqrt{t}$.
Notice that $\{U_l^v\}_{l\in[0,T_v]}$ is updated every time a neighboring vertex of v sends exactly one particle on v, and the number of times this happens is proportional to the number of times v is selected (i.e. its Poisson clock rings). This implies the following remark.

Remark 4 *In Oil and Water, the expected number of holes created at v, therefore the expected number of lost oil-water pairs, grows as the square root of the number of times v is selected.*

Therefore, from Remarks 3 and 4, in IDLAA we can interpret the fact that the local random walk of v is on 0 as the presence of a hole at v in Oil and Water and the number of times there is a new hole in v in Oil and Water as the number of absorbed

particles at v in IDLAA, and this number grows as the square root of the number of times v is selected.
Hence, we can see IDLAA as a simplification of the Oil and Water model.

This analogy leads us to think that the shape of the aggregate in IDLAA behaves like the one in the Oil and Water model when the numbers of oil and water particles are both n, which is conjectured to have as a limit shape an Euclidean ball with radius $O\left(n^{1/(d+2)}\right)$. Hence, the conjecture for IDLAA can be formalized as follows, recalling that $\mathfrak{B}_r$ is the "lattice ball" of radius $r > 0$ centered in the origin.

Conjecture 1 *For every $d \in \mathbb{N}$ there exists a positive constant K_d such that for every $\varepsilon > 0$*

$$\mathfrak{B}_{K_d n^{1/(d+2)}(1-\varepsilon)} \subseteq A(n) \subseteq \mathfrak{B}_{K_d n^{1/(d+2)}(1+\varepsilon)} \qquad \textit{a.s. for sufficiently large } n. \tag{4.2}$$

This means that the radius of the ball which is the limit shape for the aggregate in IDLAA, for n sufficiently large, has the same order of the one in the Oil and Water model.

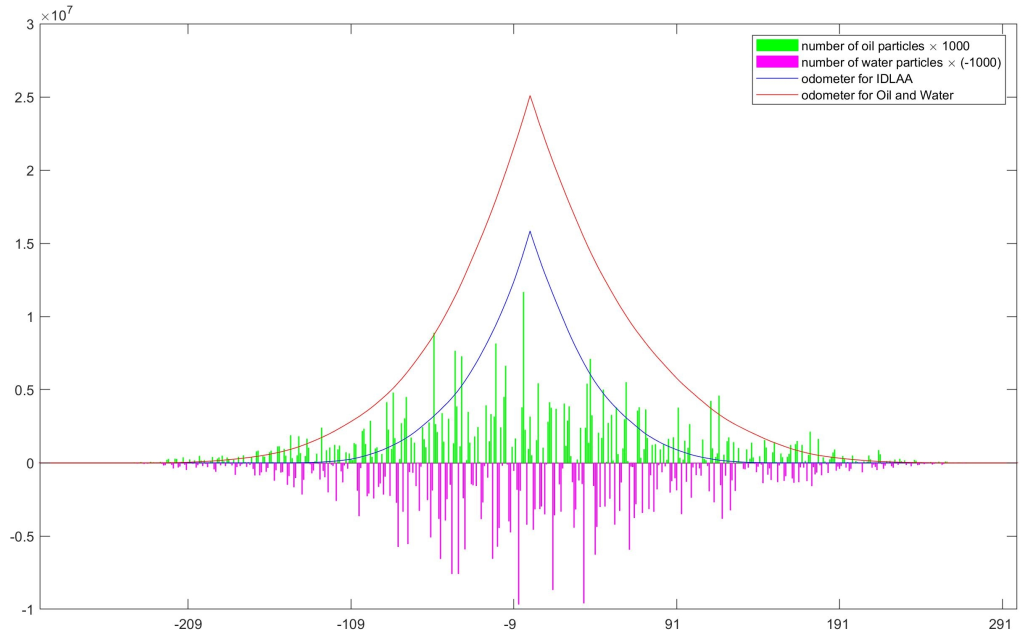


Figure 4.1. In blue and in red, the odometer respectively in the Oil and Water system starting with $n = 360000$ pairs on the origin and in IDLAA starting with n particles on the origin. Notice that the red line is always under the blue one. For each site, the length of the bars is proportional to the number of oils or waters in that site in the final configuration, with green bars for oils and magenta ones for waters: in particular, the length of the green bars is the number of oil particles multiplied by 1000, while the length of the magenta ones is the number of water particles multiplied by -1000.

In Figure 4.1 we simulated the two models in dimension $d = 1$. We can see that the number of firings of each vertex in IDLAA is smaller than the one in Oil and Water for the same vertex, since the odometer (the function which counts the number of times each vertex fires) of the former is always under the one of the latter.
This suggests us also that, if the conjectures about the shapes of the aggregates in IDLAA and Oil and Water were true, namely that the radii of their limit shapes are both $O\left(n^{1/(d+2)}\right)$, then smaller or equal respectively to $K_d^{(IDLAA)} n^{1/(d+2)}$ and $K_d^{(OW)} n^{1/(d+2)}$, we can take $K_1^{(IDLAA)} < K_1^{(OW)}$. A confirmation for this is given in Figure 4.2.

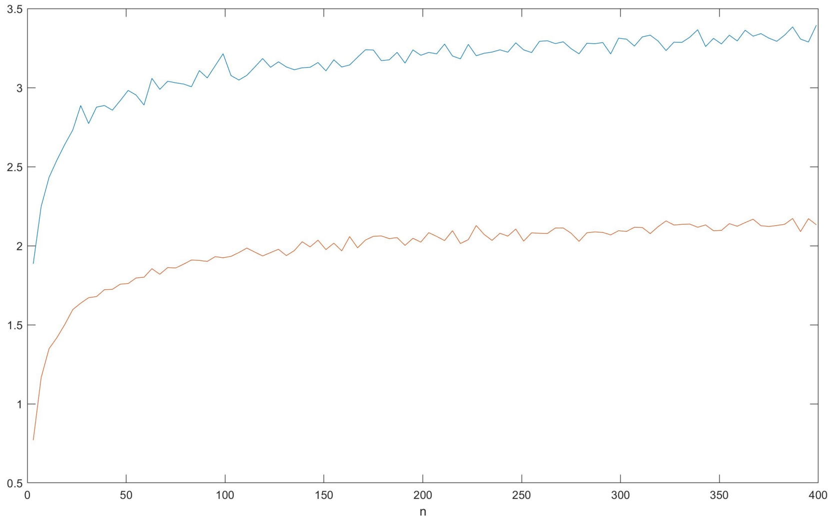


Figure 4.2. These plots are the results of 100 simulations for each $n \in \{|\mathfrak{B}_r| : r \in 2\{1, \ldots, 100\}\}$. In blue $\frac{r^{(OW)}(n)}{n^{1/(1+2)}}$ and in red $\frac{r^{(IDLAA)}(n)}{n^{1/(1+2)}}$, with $r^{(j)}(n) = \frac{1}{100}\sum_{i=1}^{100} r_i^{(j)}(n)$ for $j \in \{OW, IDLAA\}$, where $r_i^{(j)}(n)$ is the smallest radius of the lattice ball which contains the aggregate of the model j in the i-th simulation. The values of these functions are estimates for $K_1^{(OW)}$ and $K_1^{(IDLAA)}$. Notice that the red curve is always under the blue one.

Then, we simulated the two models in dimension $d = 2$. In Figure 4.3 we can see the Oil and Water odometer over the IDLAA one. The fact that the odometer of IDLAA is never above the odometer of Oil and Water suggests that also in dimension 2 the constant for the Oil and Water model is bigger than the IDLAA one.

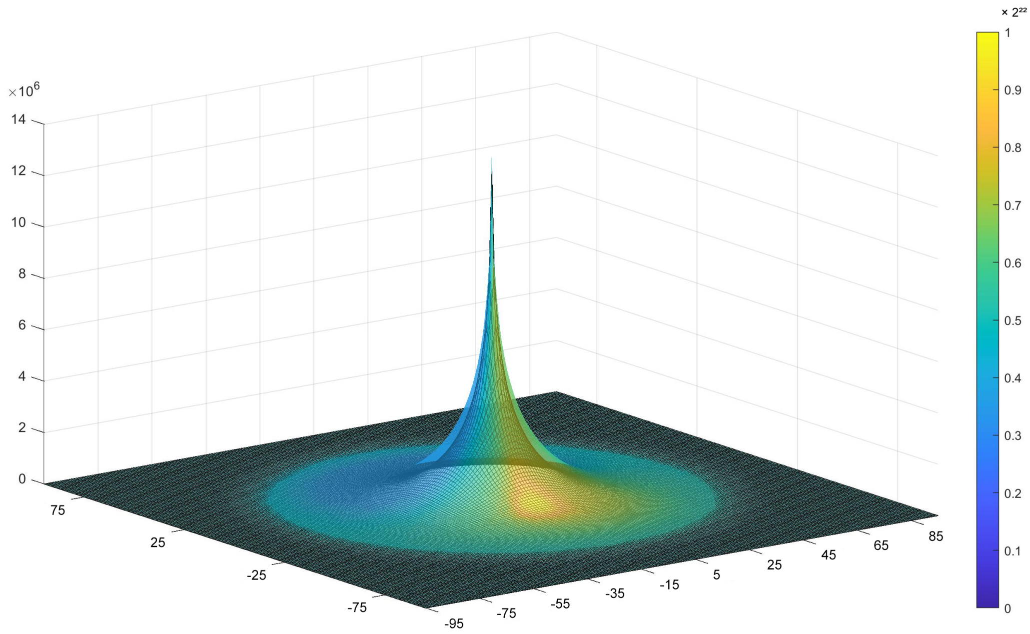


Figure 4.3. In transparency, according to the colorbar on the side, the odometer in the Oil and Water system starting with 2^{22} pairs on the origin. Under this odometer, in black, the odometer in IDLAA starting with 2^{22} particles in the origin.

To get a confirmation of this, in the same way we did for $d = 1$ in Figure 4.2, we produced Figure 4.4.

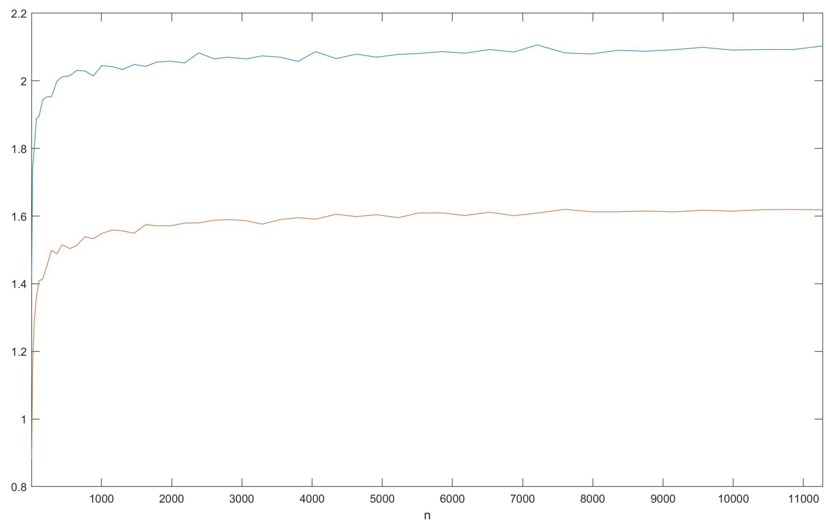


Figure 4.4. Estimates for $K_2^{(OW)}$, in blue, and for $K_2^{(IDLAA)}$, in red, for $n \in \{|\mathfrak{B}_r| : r \in (1.2)\{1, \ldots, 50\}\}$. Again, the red curve is always under the blue one.

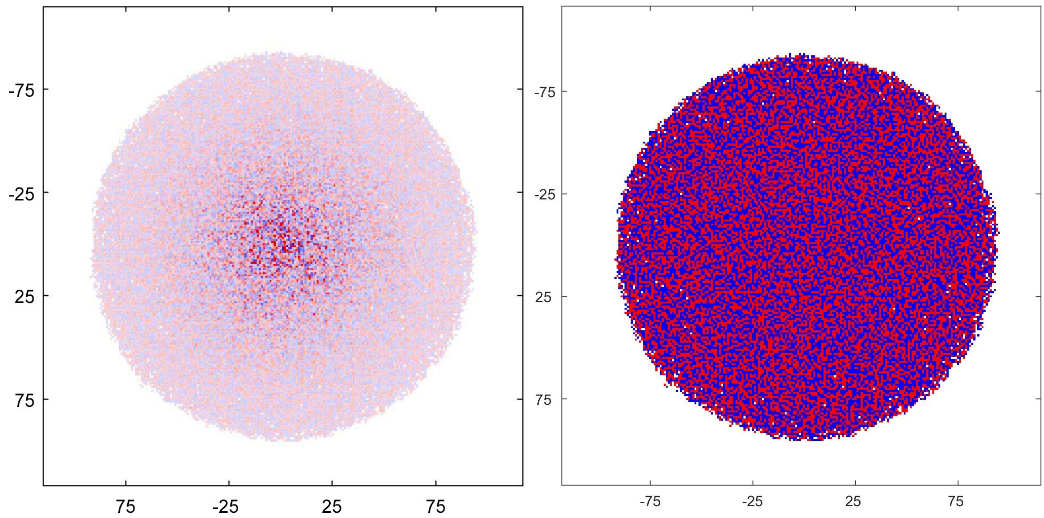


Figure 4.5. The final configuration of an Oil and Water system in $\mathbb{Z}^2$ which started with $n = 2^{22}$ oil-water pairs in the origin. On the left, the blue sites are the ones occupied by water particles, while the red sites are occupied by oils. The intensity of the shade indicates the number of particles: the more particles there are, the more vivid the color. On the right, we have the same configuration without shading intensity: the red sites are occupied by oil particles while the blue sites are occupied by water particles. It is conjectured that the limit shape is an Euclidean ball of radius which is of order $n^{1/4}$.

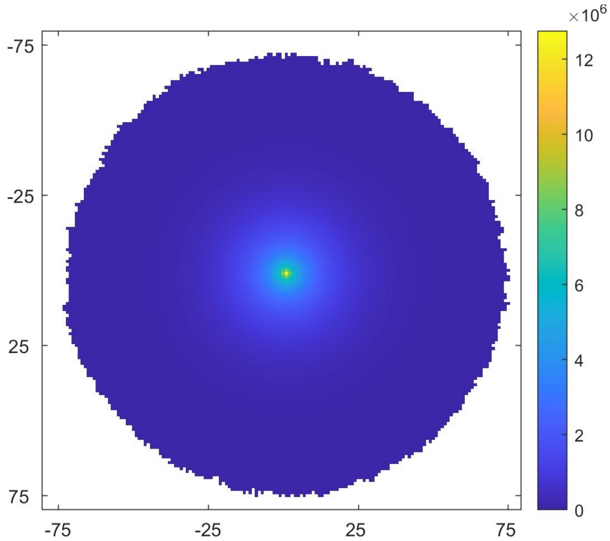


Figure 4.6. Aggregate of IDLAA in $\mathbb{Z}^2$ starting with $n = 2^{22}$ particles in the origin. According to conjecture 1, the limit shape is an Euclidean ball of radius which is of order $n^{1/4}$. The colors represent the number of particles that touched each site.

The difference in size is also noticeable in Figures 4.5 and 4.6, where we represented respectively the final configuration for Oil and Water and the aggregate for IDLAA.

We can also see that in the final configuration of Oil and Water in Figure 4.5 the sites which hold no particles (the white vertices in the figure) increase their number near the boundary of the cluster and decrease it near the origin. It is not a coincidence and has a simple explanation.

Since the difference between the number of oil particles and water ones at a given site v in the final configuration is distributed like the position of a simple symmetric lazy random walk on $\mathbb{Z}$ at time $T^{(v)}$, where $T^{(v)}$ is the number of firings occurred at v during the stabilization of the system, the probability that v does not host any particle in the final configuration is the probability that the simple symmetric lazy random walk is on 0 at time $u(v)$, where u is the odometer for Oil and Water.

We will consider, for simplicity, a simple symmetric random walk $\{Z_t\}_{t\in\mathbb{N}_0}$ on $\mathbb{Z}$, instead of the simple symmetric lazy random walk on $\mathbb{Z}$, since the asymptotic behavior of their probability to be on 0 at time t is the same for $t \to \infty$. For every even $t \in \mathbb{N}_0$

$$P_0(Z_t = 0) = \frac{1}{2^t}\binom{t}{t/2} = \frac{1}{2^t}\frac{t!}{(t/2)!(t/2)!} \underset{\substack{\uparrow \\ \text{Stirling's formula}}}{\sim} \sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{t}} \qquad \text{for } t \to \infty.$$

Therefore we get that the probability that v does not host any particle in the final configuration is asymptotically equivalent to

$$\sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{T^{(v)}}}.$$

Let us consider a vertex near the boundary of the aggregate and another one close to $\mathbf{0}$, which we will call respectively v_{ext} and v_{int}, with $\|v_{\text{ext}}\| \geq \|v_{\text{int}}\|$.

As we can see in Figure 4.3 from the shape of the odometer u, due to the initial configuration, the farther v is from the origin, the more $u(v)$ decreases. Thence (as we already noticed for IDLAA), the nearest sites to $\mathbf{0}$ will be responsible for the majority of the total firings and $T^{(v_{\text{int}})} \geq T^{(v_{\text{ext}})}$, thus $\sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{T^{(v_{\text{ext}})}}} \geq \sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{T^{(v_{\text{int}})}}}$.

Therefore we have obtained that the probability that a site hosts no particles increases with the distance of the site from the origin.

This result holds for every vertex in the lattice $\mathbb{Z}^d$ for every $d \in \mathbb{N}$ because all the vertices, even the origin, are holes in the initial configuration.

In [3], where Asselah and Gaudillière analyzed the internal DLA, the inner and outer errors were defined, respectively as

$$\delta_I(n) = n - \sup\{r \geq 0 : \mathfrak{B}_r \subseteq A(|\mathfrak{B}_n|)\},$$
$$\delta_O(n) = \inf\{r \geq 0 : A(|\mathfrak{B}_n|) \subseteq \mathfrak{B}_r\} - n,$$

according to Theorem 5. These quantities compose the fluctuations of the shape of the aggregate. In the same paper, they proved that in dimension $d = 2$ the following relations hold almost surely

$$\limsup_{n\to\infty} \frac{\delta_I(n)}{\log n} \leq \alpha_2, \qquad \limsup_{n\to\infty} \frac{\delta_O(n)}{\log n} \leq \beta_2 \tag{4.3}$$

and that for $d \geq 3$ we have almost surely

$$\limsup_{n\to\infty} \frac{\delta_I(n)}{\sqrt{\log n}} \leq \alpha_d, \qquad \limsup_{n\to\infty} \frac{\delta_O(n)}{\sqrt{\log n}} \leq \beta_d$$

where α_d and β_d are constants for every $d \geq 2$.
This means that, in internal DLA, $\delta_I(n)$ and $\delta_O(n)$ are both $O(\log n)$ for $d = 2$ and $O(\sqrt{\log n})$ for $d \geq 3$.
In an analogous way, we defined the inner and the outer errors for Oil and Water and IDLAA, according to the conjectures on the shapes of their aggregate, in the following way:

$$\delta_I(n) = K_d \left(\omega_d n^d\right)^{1/(d+2)} - \sup\{r \geq 0 : \mathfrak{B}_r \subseteq A(|\mathfrak{B}_n|)\}, \tag{4.4}$$

$$\delta_O(n) = \inf\{r \geq 0 : A(|\mathfrak{B}_n|) \subseteq \mathfrak{B}_r\} - K_d \left(\omega_d n^d\right)^{1/(d+2)} \tag{4.5}$$

where K_d is the positive constant which appears in the conjectures and $A(l)$ is the aggregate (the Oil and Water ones or the IDLAA ones, according to which model we are considering) that we get starting with l particles in IDLAA or with l pairs in Oil and Water.
In order to replicate for Oil and Water and IDLAA the results Asselah and Gaudillière obtained in [3] for internal DLA, we got an estimate of K_d, for $d = 2$, in the same way

we did for Figures 4.2 and 4.4, using a system which started with $N = |\mathfrak{B}_{120}| = 45213$ particles or pairs (depending on the model), getting

$$K_2^{(IDLAA)} \backsimeq 1.6431 \qquad \text{and} \qquad K_2^{(OW)} \backsimeq 2.1083.$$

Plugging these estimates in (4.4) and (4.5), we simulated the two models for different values of n (with 100 simulations for each value to obtain more accurate estimates) and plotted the outcomes in Figure 4.7, which are consistent with (4.3) for $\delta_I(n)$, while for $\delta_O(n)$, due to the limited storage capacities of our computers, the simulations we performed indicate that it could be possible to notice the result for larger values of n.

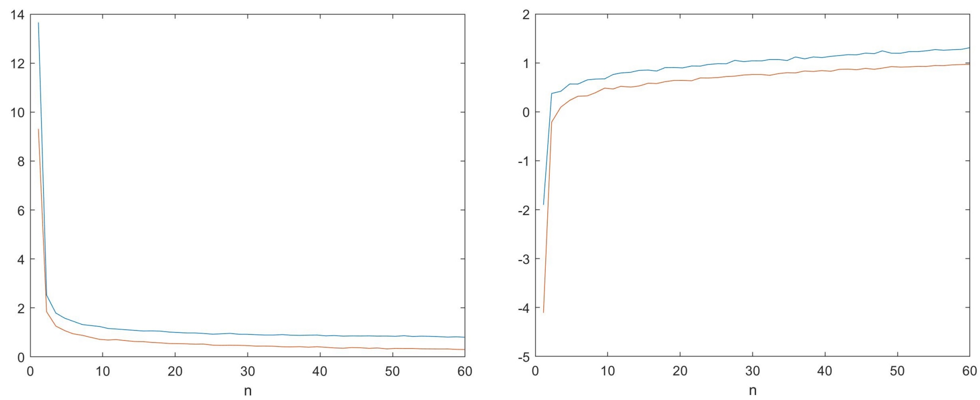


Figure 4.7. On the left, $\frac{\tilde{K}_2\left(\omega_2 n^2\right)^{1/(2+2)} - \sup\{r \geq 0 : \mathfrak{B}_r \subseteq A(|\mathfrak{B}_n|)\}}{\log n}$ for Oil and Water (in blue) and IDLAA (in red), with $\tilde{K}_2 > K_2$ for both models, therefore this function is bigger than $\frac{\delta_I(n)}{\log n}$ in both cases and it is the same for its limit superior, which is bounded above by a constant α_2. On the right, $\frac{\delta_O(n)}{\log n}$ for Oil and Water (in blue) and IDLAA (in red). We can also notice that in each case the red line is under the blue one.

However, in Figure 4.8 the simulations are consistent with the result

$$\limsup_{n \to \infty} \frac{\delta_O(n)}{(\log n)^2} \leq 0,$$

i.e. that $\delta_O(n) = o(\log^2 n)$. Indeed it was initially proven for IDLA, too, by Asselah and Gaudillière in [2], that, for $d \geq 2$, $\delta_I(n) = O(\log n)$ and $\delta_O(n) = O(\log^2 n)$. Future analysis could therefore prove the conjecture according to which for Oil and Water and IDLAA, for $d = 2$, $\delta_I(n) = O(\log n)$ and $\delta_O(n) = o(\log^2 n)$.

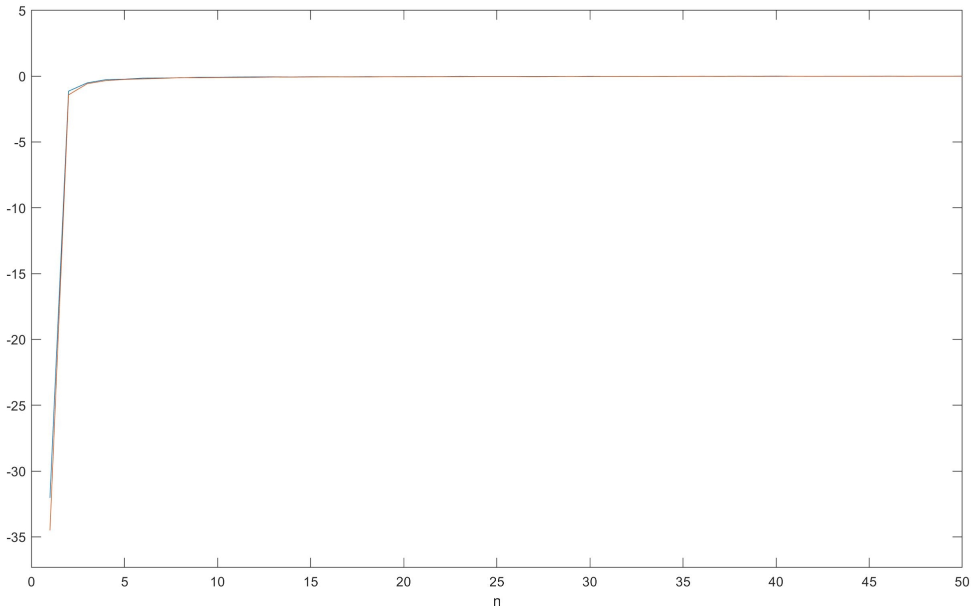


Figure 4.8. $\frac{\delta_O(n)}{(\log n)^2}$ for Oil and Water (in blue) and IDLAA (in red).

4.2 Mathematical formalization for IDLAA

We can give a mathematical formalization for IDLAA in the same way we did for Oil and Water.

Consider the graph $G = \left(\mathbb{Z}^d, \bigcup_{x \in \mathbb{Z}^d} \{\{x, y\} : \|x - y\| = 1\}\right)$ and the space of configurations $\Omega = \mathbb{N}_0^{(\mathbb{Z}^d)}$, where if

$(\eta(v))_{v \in \mathbb{Z}^d} \in \Omega$ is a configuration then, for $v \in \mathbb{Z}^d$, $\eta(v)$ represents the number of particles that are on v.

Let us define, for $v \in \mathbb{Z}^d$, the operator $\tau_v^A : \Omega \to \Omega$,

$$\tau_v^A(\eta_t) := \eta_{t+1} \qquad \text{with} \qquad \eta_{t+1}(z) = \begin{cases} \eta_t(z) - 1 & \text{if } z = v \\ \eta_t(z) & \text{if } z \neq v \end{cases}$$

and, for every vertex u such that $u \sim v$, the operator $\tau_{v,u}^J \Omega \to \Omega$,

$$\tau_{v,u}^J(\eta_t) := \eta_{t+1} \qquad \text{with} \qquad \eta_{t+1}(z) = \begin{cases} \eta_t(z) - 1 & \text{if } z = v \\ \eta_t(z) + 1 & \text{if } z = u \\ \eta_t(z) & \text{otherwise} \end{cases} .$$

In words, the operator τ_v^A absorbs a particle in v, while $\tau_{v,u}^J$ makes a particle in v jump on the neighboring vertex u.

Consider the array

$$\tau = \left(\tau^{v,0}, \tau^{v,1}, \tau^{v,2}, \dots\right)_{v\in\mathbb{Z}^d} = \left(\tau^{v,j}\right)_{v\in\mathbb{Z}^d, j\in\mathbb{N}_0}$$

where every $\tau^{v,j}$ is an instruction in the form τ_v^A or $\tau_{v,u}^J$, namely it is an element of the set $\{\tau_v^A, \tau_{v,u}^J : u \sim v\}$.

Let us also define a function $h \in \mathbb{N}_0^{(\mathbb{Z}^d)}$ which counts the number of instructions used on each vertex.

In this model, given a probability space $(\Gamma, \mathcal{L}, \overline{P})$, if $(\omega_1, \dots \omega_k) \in \Gamma^k$ with $k \in \mathbb{N}$ and $X \in \mathbb{Z}$, we will say that we topple the vertex v when we use on $(\eta, h, X, (\omega_1, \dots, \omega_k))$ the operator $\Phi_v : \Omega \times \mathbb{N}_0^{(\mathbb{Z}^d)} \times \mathbb{Z}^{(\mathbb{Z}^d)} \times \left(\bigcup_{k\geq 1} \Gamma^k\right) \to \Omega \times \mathbb{N}_0^{(\mathbb{Z}^d)} \times \mathbb{Z}^{(\mathbb{Z}^d)} \times \left(\bigcup_{k\geq 1} \Gamma^{k-1}\right)$,

$$\Phi_v\left(\eta, h, X, (\omega_1, \dots, \omega_k)\right) := \left(\eta_1, h_1, X_1(\omega_1), (\omega_2, \dots, \omega_k)\right)$$

where

$$\eta_1 = \tau^{v,h(v)+1}, \qquad h_1 = h + \delta_v,$$

$$X_1(z, \omega_1) = \begin{cases} X(z) & \text{if } z \neq v \\ \begin{cases} X(z)+1 & \text{if } \omega_1 \in B_1 \\ X(z)-1 & \text{if } \omega_1 \in B_2 \end{cases} & \text{if } z = v \end{cases}$$

with $B_1, B_2 \in \mathcal{L}$ such that $\overline{P}(B_1) = \overline{P}(B_2) = 1/2$.

We call the vertex v *unstable* in η if $\eta(v) > 0$ and *stable* if $\eta(v) = 0$.

In this interpretation, a vertex that topples does not necessarily send a particle to a neighbor, but it could also absorb it.

We say that Φ_v is *legal* for $(\eta, h, X, (\omega_1, \dots, \omega_k))$ if $\eta(v) > 0$ and one of the following conditions is satidfied:

- $\tau^{v,h(v)+1} = \tau_v^A$ and $X(v) = 0$,
- $\tau^{v,h(v)+1} = \tau_{v,u}^J$ for some $u \sim v$ and $X(v) \neq 0$.

Given a sequence of vertices $\alpha = (v_1, v_2, \dots, v_k)$, $k \in \mathbb{N}$, we define

$$\Phi_\alpha := \Phi_{v_k}\Phi_{v_{k-1}} \dots \Phi_{v_1}$$

and we say that it (or α itself) is *legal* for $(\eta,(\omega_1,\dots,\omega_{k_1}))$ where $k_1 \geq k$ if Φ_{v_l} is legal for $\Phi_{(v_1,\dots,v_{l-1})}\left(\eta,\hat{0},\widetilde{0},(\omega_1,\dots,\omega_{k_1})\right)$ for every $l \in \{2,\dots,k\}$, where $\hat{0} \in \mathbb{N}_0^{(\mathbb{Z}^d)}$ and $\widetilde{0} \in \mathbb{Z}^{(\mathbb{Z}^d)}$ are equal to zero on every vertex.

In order to simplify notation, we put $\Phi_\alpha^{(\omega_1,\dots,\omega_k)}\eta := \Phi_\alpha\left(\eta,\hat{0},\widetilde{0},(\omega_1,\dots,\omega_k)\right)$, therefore the first component of $\Phi_\alpha^{(\omega_1,\dots,\omega_k)}\eta$ is the configuration obtained from η when the vertices topple according to the sequence α.

Let, for all $v \in \mathbb{Z}^d$,

$$m_\alpha(v) := \sum_{l=1}^{k} \mathbb{1}_{\{v=v_l\}} \tag{4.6}$$

be the times v appears in the sequence α.

We will write $m_\alpha \geq m_\beta$ if $m_\alpha(v) \geq m_\beta(v)$ for every $v \in \mathbb{Z}^d$.

Let $K \subset \mathbb{Z}^d$ finite. We will say that

- A configuration η is *stable* in K if every $v \in K$ is stable for η;
- A sequence of vertices α is *contained* in K if all its components are in K;
- $\alpha = (\alpha_1,\dots,\alpha_k)$ *stabilizes* $(\eta,(\omega_1,\dots,\omega_{k_1}))$ in K, with $k_1 \geq k$, if $\Phi_\alpha^{(\omega_1,\dots,\omega_{k_1})}\eta$ is stable in K, indicating, with an abuse of notation, the first component of $\Phi_\alpha^{(\omega_1,\dots,\omega_{k_1})}\eta$ with $\Phi_\alpha^{(\omega_1,\dots,\omega_{k_1})}\eta$ itself.

It is possible to prove that IDLAA satisfies the abelian property, too. We will show it by enunciating a preliminary lemma, which is the least action principle for this model.

Lemma 17 (Least action principle for IDLAA) *Let $\alpha = (\alpha_1,\dots,\alpha_k)$ and $\beta = (\beta_1,\dots,\beta_l)$ be two legal sequences of vertices for $(\eta,(\omega_1,\dots,\omega_n))$, with $n \geq \max\{k,l\}$. Suppose that β is contained in a finite subset K of $\mathbb{Z}^d$ and that α stabilizes $(\eta,(\omega_1,\dots,\omega_n))$ in K. Then, $m_\beta \leq m_\alpha$.*

Proof. Let, per impossibile, $m_\alpha \ngeq m_\beta$. Then, there exists $j \leq l$ such that $m_{(\beta_1,\dots,\beta_j)}(x) > m_\alpha(x)$ for some $x \in \mathbb{Z}^d$. Let $\beta^{(j)} = (\beta_1,\dots,\beta_j)$ for $j \leq l$ and $r = \max\{j \leq l : m_{\beta^{(j)}} \leq m_\alpha\} < l$. Let $y = \beta_{r+1} \in K$. Since β is legal, y is not stable in the configuration $\Phi_{\beta^{(r)}}^{(\omega_1,\dots,\omega_n)}\eta$, namely $\Phi_{\beta^{(r)}}^{(\omega_1,\dots,\omega_n)}\eta(y) > 0$. Moreover, by definition of r, we have that $m_{\beta^{(r)}} \leq m_\alpha$ and $m_{\beta^{(r)}}(y) = m_\alpha(y)$. Since toppling sites other than y cannot decrease the number of particles at y, we get

$$\Phi_{\alpha}^{(\omega_1,\dots,\omega_n)}\eta(y) \geq \Phi_{\beta(r)}^{(\omega_1,\dots,\omega_n)}\eta(y) > 0,$$

namely y is not stable in the configuration $\Phi_{\alpha}^{(\omega_1,\dots,\omega_n)}\eta$, too, and we have the desired contradiction. □

Lemma 18 (Abelian property) *Let $\alpha = (\alpha_1, \dots, \alpha_k)$ and $\beta = (\beta_1, \dots, \beta_l)$ be two legal sequences of vertices for $(\eta, (\omega_1, \dots, \omega_n))$, with $n \geq \max\{k, l\}$, contained in a finite subset K of $\mathbb{Z}^d$ and assume that they both stabilize $(\eta, (\omega_1, \dots, \omega_n))$ in K. Then $m_\alpha = m_\beta$ and in particular $\Phi_{\alpha}^{(\omega_1,\dots,\omega_n)}\eta = \Phi_{\beta}^{(\omega_1,\dots,\omega_n)}\eta$.*

Proof. For the first result we just apply Lemma 17 in two directions to conclude that $m_\alpha \leq m_\beta \leq m_\alpha$. The second result follows from the fact, simple to prove, that for a legal sequence of vertices $g = (g_1, \dots, g_s)$ for $(\eta, (\omega_1, \dots, \omega_a))$, with $a \geq s$, $\Phi_{g}^{(\omega_1,\dots,\omega_a)}\eta$ depends on g only through m_g. □

Now we allow the array τ, therefore its components τ_j, to be stochastic in the following way.

$$\tau_j(\eta_t, X_t) : \Lambda_1 \to \{\tau^{v,j}(\eta_t, X_t) : v \in \mathbb{Z}^d\}$$

will be a random variable on the probability space $(\Lambda_1, \mathcal{G}_1, P_1)$ such that

$$P_1\left(\tau_j(\eta_t, X_t) = \tau^{v,j}(\eta_t, X_t)\right) = \frac{\eta_t(v)}{\sum\limits_{x \in \mathbb{Z}^d} \eta_t(x)}$$

and under P_1 the variables τ_j are independent for different values of $j \in \mathbb{N}_0$. This reflects the fact that each particle is chosen with uniform distribution, therefore the sites which contain more particles have more chances to be selected. Now, the instruction $\tau^{v,j}$ can either be a jump or an absorption instruction, namely

$$\tau^{v,j}(\eta_t, X_t) = \begin{cases} \tau_v^A(\eta_t) & \text{if } X_t(v) = 0 \\ \tau^{v,j,J}(\eta_t) & \text{if } X_t(v) \neq 0 \end{cases} \qquad \text{with}$$

$$\tau^{v,j,J}(\eta_t) : \Lambda_2 \to \{\tau_{v,u}^J(\eta_t) : u \sim v\}$$

random variable on the probability space $(\Lambda_2, \mathcal{G}_2, P_2)$ such that

$$P_2\left(\tau^{v,j,J}(\eta_t) = \tau_{v,u}^J(\eta_t)\right) = \frac{1}{2d} \qquad \forall\ u \sim v$$

and under P_2 the variables $\tau^{v,j,J}$ are independent for different values of $v \in \mathbb{Z}^d$, $j \in \mathbb{N}_0$.

In words, the random variable $\tau^{v,j,J}$ makes the particle choose the neighbor to jump on with uniform probability.

In general, indicating with $\mathcal{P}$ the joint probability law of τ_j and $\tau^{v,j,J}$, we have that under $\mathcal{P}$, for every $j \in \mathbb{N}_0$

$$\tau_j(\eta_t, X_t) = \begin{cases} \tau_v^A(\eta_t) & \text{with probability } \dfrac{\eta_t(v)}{\sum\limits_{x \in \mathbb{Z}^d} \eta_t(x)} \quad \forall\, v \in \mathbb{Z}^d & \text{if } X_t = 0 \\ \tau_{v,u}^J(\eta_t) & \text{with probability } \dfrac{\eta_t(v)}{\sum\limits_{x \in \mathbb{Z}^d} \eta_t(x)} \dfrac{1}{2d} \quad \forall\, u \sim v \quad \forall\, v \in \mathbb{Z}^d & \text{if } X_t \neq 0 \end{cases}$$

and under $\mathcal{P}$ the random variables $\tau_j(\eta_t, X_t)$ are independent for different values of $j \in \mathbb{N}_0$.

Since we start with n particles on the origin and nothing elsewhere, the initial configuration of the model will be $\eta_0 = n\delta_{\mathbf{0}} \in \Omega$.

Chapter 5

Conclusions

In this work we analyzed different models, exploiting their random walk based behavior to prove different claims. In order to do that, we proved some preliminary theorems on harmonic functions and their relation with simple random walks, focusing on the Green's function. Using these results, firstly, we showed that the limit shape of the aggregate in the internal DLA on $\mathbb{Z}^d$ is an Euclidean ball centered in the source of the particles with radius $O\left(n^{1/d}\right)$. Subsequently, we considered a new non-unary model, Oil and Water, and we proved that it does not undergo a phase transition between fixation (the case in which every vertex of the graph fires finitely many times) and activity (when every vertex fires infinitely many times) independently from the initial configuration (provided that the particles density per vertex is finite and the underlying graph is vertex-transitive), in particular the system always fixates. Then we compared the behavior of the Oil and Water dynamics with another abelian network, Activated Random Walks, which exhibits a phase transition at a positive nontrivial critical density. We recognized that the evolution rules of internal DLA are a particular case of Activated Random Walks, more precisely when the sleeping rate is infinity, namely every isolated particle instantly becomes a sleeping particle. In this way, we were able to answer to the question about the transition phase also for a model which follows the same evolution rules as IDLA but with continuous time and starting from a configuration distributed as a product of independent Poisson distributions with parameter μ, asserting that its critical threshold is $\mu_c = 1$. Therefore, we thought about a solution for the problem

originally exposed for internal DLA, namely the internal aggregation, for the Oil and Water model. This is still an open problem, hence we considered conjectures according to which the limit shape of the aggregate for the Oil and Water dynamics that starts with n oil-water pairs on the origin of $\mathbb{Z}^d$ is an Euclidean ball with radius $O\left(n^{1/(d+2)}\right)$, which is currently proven just for $d = 1$. To support this conjecture, we introduced a system simpler than the Oil and Water one, internal DLA with Absorption, whose aggregate has, we think, a similar behavior to that of the Oil and Water system. We believe that the radii of the aggregates of these models have the same order of magnitude. We supported this comparison with analogies and numerical simulations, whose results have stimulated other observations and also highlighted the differences between the two dynamics. In particular, we tried to characterize the behavior of the fluctuations of their aggregate in order to compare them with the ones of the internal DLA, exposing the limits and the conjectures our simulations arose, namely that, dividing the fluctuations in inner and outer errors, in dimension $d = 2$ the inner ones are $O(\log n)$ and the outer ones $o(\log^2 n)$, when the systems starts with as many particles as the vertices contained in a ball with radius n, which could be proven in the future.
This comparison could give rise to analogous approaches for the two problems, applying the techniques created for one case to the other one and vice versa. Nevertheless, the problem of the proof for these conjectures is still open, like the analysis of other correlated aspects, for example the time it takes for each model to fixate.

The more general study of abelian networks has still many open questions, however important and rather general results have already been shown and even today the number of applications is very large and the research is being prolific and stimulating.

Chapter 6

Appendix

6.1 Further properties of the Green's function and potential kernel

The following two results are proven in [14], where they are respectively Theorem 1.5.4 and Theorem 1.6.2.
Let $a_d = \frac{d}{2}\Gamma\left(\frac{d}{2}-1\right)\pi^{-d/2} = \frac{2}{(d-2)\omega_d}$.

Theorem 7 *For $d \geq 3$ and $\|x\| \to \infty$, $x \in \mathbb{Z}^d$,*

$$G(\mathbf{0}, x) \sim a_d \|x\|^{2-d}$$

and more precisely

$$G(\mathbf{0}, x) = a_d \|x\|^{2-d} + o\left(\|x\|^{1-d}\right).$$

Theorem 8 *For $\|x\| \to \infty$, $x \in \mathbb{Z}^2$, there exists a constant k such that if $\alpha < 2$*

$$\lim_{\|x\|\to\infty} \|x\|^{\alpha}\left(a(x) - \frac{2}{\pi}\log\|x\| - k\right) = 0.$$

6.2 Abelian networks

Abelian networks are automata networks that are subject to axioms which ensure that the final output does not depend on the order in which the automata process their inputs.

Definition (Automaton) *An automaton is a 5-tuple $M = (\Sigma, \Gamma, Q, \delta, \lambda)$, where*

· Σ *is a finite set of symbols, called the input alphabet;*

· Γ *is a finite set of symbols, called the output alphabet;*

· Q *is a set called the state space;*

· $\delta : \Sigma \times Q \to Q$ *is called next-state function or transition function and brings the pair state-input in the next state;*

· $\lambda : \Sigma \times Q \to \Gamma$ *is called next-output function or message passing function and brings the pair state-input in the output.*

Definition (Word, or string) *Given a group G and $S \subseteq G$, a word (or string) in S is an expression in the form*

$$s_1^{\varepsilon_1} s_2^{\varepsilon_2} \dots s_n^{\varepsilon_n}$$

where $s_1, \dots, s_n \in S$, $\varepsilon_1, \dots, \varepsilon_n \in \{+1, -1\}$. The number n is the length of the word. Each word in S represent an element of G, namely the product of the expression above.
By convention, the (unique) identity element is represented by the empty word, which is the only word of length zero.

Consider the automaton $M = (\Sigma, \Gamma, Q, \delta, \lambda)$.

· M reads a finite string of symbols $a_1 a_2 \dots a_n$ with $a_i \in \Sigma$, called *input word*. The set of all input words for M is denoted with Σ^*.

· A sequence of states $q_0, q_1, \dots, q_n$, where $q_i = \delta(a_i, q_{i-1})$ for all $i \in \{1, \dots, n\}$, is a *run* of the automaton on the input $a_1 a_2 \dots a_n \in \Sigma^*$ that starts from state q_0.

In other words, first of all at the beginning the automaton is in the state q_0 and receives input a_1; for a_1 and for every successive a_i in the input string, M assumes the next state q_i according to the transition function $\delta(a_i, q_{i-1})$ until the last symbol a_n is read, leaving the automaton in the final state q_n of the run.
Similarly, at every step M emits an output according to the output function $\lambda(a_i, q_{i-1})$. Such output will be a k-uple of symbols, where k is the number of components of $\lambda(a_i, q_{i-1})$.

We will now define an automata network and then an abelian network.

Let $G = (V, E)$ be an oriented graph (which can have self-loops and multiple edges). Associated to each vertex $v \in V$ there is a *processor* $\mathcal{P}_v$, which is an automaton with a single input port (from which it reads the symbols of its input alphabet in first-in-first-out order), multiple output ports, one for each edge $(v, u) \in E$, input alphabet A_v (whose elements are called *letters*), state space Q_v, transition function $T_v : A_v \times Q_v \to Q_v$ (new internal state) and message passing function $T_{(v)} : A_v \times Q_v \to \bigcup_{u \sim v} A_u^*$ with components, for each $(v, u) \in E$, $T_{(v,u)} : A_v \times Q_v \to A_u^*$ (letters sent from v to u).

We will suppose, with no loss of generality, $A_v \cap A_u = \emptyset$ for all $u \neq v$ in V, so that a given letter can only belong to the input alphabet of a single processor (hence $\bigcup_{u \sim v} A_u^* = \bigsqcup_{u \sim v} A_u^*$).

Therefore, if $\mathcal{P}_v$ is in the state $q \in Q_v$ and it elaborates input $a \in A_v$:

1) $\mathcal{P}_v$ changes its internal state in $T_v(a, q)$ and

2) for all $(v, u) \in E$, $\mathcal{P}_u$ receives input $T_{(v,u)}(a, q)$.

If more than one processor $\mathcal{P}_v$ has inputs to process, changing the order in which the processors act may change the order of the letters that arrive at other processors. However, there exists a family of processors networks, called abelian, for which the behavior of the network does not depend on the order in which the letters are processed.

In particular, we look for networks for which the following characteristics do not depend on the order in which the single processors act:

a.1) the *halting status* (namely, wheather or not processing eventually stops);

a.2) the *final state* of the processors;

a.3) the *run time* (total number of letters processed by all $\mathcal{P}_v$);

a.4) the *local run times* (number of letters processed by a given $\mathcal{P}_v$);

a.5) the *detailed local run times* (number of times a given $\mathcal{P}_v$ processes a given $a \in A_v$).

To better formalize processing whole words, we extend the domain of T_v and $T_{(v,u)}$ to $A_v^* \times Q_v$ in this way: if $w = aw'$ is a word in alphabet A_v beginning with a, then $T_v(w,q) := T_v(w', T_v(a,q))$ and $T_{(v,u)}(w,q) := T_{(v,u)}(a,q)T_{(v,u)}(w', T_v(a,q))$. For the empty word ε, we set $T_v(\varepsilon, q) := q$ and $T_{(v,u)}(\varepsilon, q) := \varepsilon$.
Let us define the function $|\cdot| : \bigsqcup_{v\in V} A_v^* \to \bigsqcup_{v\in V} \mathbb{N}^{A_v}$, $|w| = (|w|_i)_{i\in A_v}$ where $|w|_i$ is the number of times i appears in the word w.
In particular, two words w and w' satisfy $|w| = |w'|$ if and only if w' is a permutation of w.

Definition (Abelian processor) *A processor $\mathcal{P}_v$ is abelian if for every $w, w' \in A_v^*$ such that $|w| = |w'|$*

$$T_v(w,q) = T_v(w',q) \qquad \textit{and} \qquad |T_{(v,u)}(w,q)| = |T_{(v,u)}(w',q)|.$$

For an abelian processor $\mathcal{P}_v$ therefore changing the order of the input letters does not change the resulting state and may change the output word sent to $u \sim v$, but only by permuting its letters.
It is easy to prove that if the condition in the definition of abelian processor is satisfied for words with length 2, then it is true for words with arbitrary length.

Definition (Abelian network) *An abelian network on a discrete graph $G = (V, E)$ is a collection of automata $\mathcal{N} = (\mathcal{P}_v)_{v\in V}$ with $\mathcal{P}_v$ abelian for all $v \in V$.*

Given an abelian network $\mathcal{N}$ with underlying graph $G = (V, E)$, we may view $\mathcal{N}$ as a single automaton with (input) alphabet $A = \bigsqcup_{v\in V} A_v$ and state space $\mathbb{Z}^A \times Q$ where $Q = \prod_{v\in V} Q_v$: if its state is $(\mathbf{x}, \mathbf{q})$ with $\mathbf{x} \in \mathbb{Z}^A$ and $\mathbf{q} \in Q$, it means that the configuration of the network is such that

· for all $a \in A$, there are $\mathbf{x}_a$ letters a waiting to be processed and

· for all $v \in V$, $\mathcal{P}_v$ is in state q_v.

Since the alphabets A_v are disjoint, we can think to $\mathbf{x}$ as a set of piles of letters, one pile for each vertex, where the $\mathbf{x}_a$ letters a are in the pile of the vertex $v \in V$ such that $a \in A_v$.

The order of execution is encoded by a word $w = a_1 \dots a_n$ with $a_i \in A$, instructing the network first to process a_1, then a_2, etc.

The fact that $\mathbf{x}$ belongs to $\mathbb{Z}^A$ and not to $\mathbb{N}_0^A$ means that we are allowing the possibility of "illegal" moves, i.e. processing a letter a even if the number $\mathbf{x}_a$ of a's waiting to be processed is not positive, namely even if a is not in the pile.

Now we will define the transition function of $\mathcal{N}$.

For $v \in V$, $a \in A_v$, let

$$t_a : Q \to Q, \qquad t_a(\mathbf{q})v := \begin{cases} T_v(a, q_v) & \text{if } a \in A_v \\ q_v & \text{if } a \notin A_v \end{cases}.$$

The transition function of $\mathcal{N}$ is $\pi : A \times \left(\mathbb{Z}^A \times Q\right) \to \mathbb{Z}^A \times Q$,

$$\pi(a, (\mathbf{x}, \mathbf{q})) = \pi_a(\mathbf{x}, \mathbf{q}) := (\mathbf{x} - \mathbf{1}_a + N(a, \mathbf{q}), t_a(\mathbf{q})) \tag{AN1}$$

where $(\mathbf{1}_a)_b = \begin{cases} 1 & \text{if } b = a \\ 0 & \text{if } b \neq a \end{cases}$ and $N(a, \mathbf{q}) = \sum_{u:(v,u)\in E} |T_e(a, q_v)|$ with $v \in V$ such that $a \in A_v$, i.e. $N(a, \mathbf{q})_b$ is the number of b's produced when $\mathcal{P}_v$ in state q_v processes letter a.

In other words, for all $a \in A$, the function $\pi_a : \mathbb{Z}^A \times Q \to \mathbb{Z}^A \times Q$ describes the new state of the network when it processes letter a.

We now enlarge the domain of π to $A^* \times \left(\mathbb{Z}^A \times Q\right)$ by putting, $\pi\left(\varepsilon, (\mathbf{x}, \mathbf{q})\right) = (\mathbf{x}, \mathbf{q})$ and, for a word $w = aw' \in A^*$ which starts with letter a, $\pi(w, \mathbf{q}) := \pi\left(w', \pi(a, \mathbf{q})\right)$; in terms of the functions π_a for the single letters, this is equivalent to define, for $w = a_1 \dots a_n \in A^*$ with each $a_i \in A$, $\pi_w := \pi_{a_n} \circ \dots \circ \pi_{a_1}$, and $\pi\left(w, (\mathbf{x}, \mathbf{q})\right) = \pi_w(\mathbf{x}, \mathbf{q})$.

To generalize equation (AN1), we also extend the domain of N to $A^* \times Q$ as follows; let $\mathbf{q}^{i-1} = \left(t_{a_{i-1}} \circ \dots \circ t_{a_1}\right) \mathbf{q}$ and

$$N(w, \mathbf{q}) := \sum_{i=1}^{n} N\left(a_i, \mathbf{q}^{i-1}_{v(i)}\right)$$

where $v(i)$ is the unique vertex such that $a_i \in A_{v(i)}$. Notice that if $a \in A_v$ and $b \in A_u$ for $u \neq v$, then

$$N(ab, \mathbf{q}) \underset{\substack{\uparrow \\ t_a \text{ acts as identity on } Q_u \text{ and } t_b \text{ on } Q_v}}{=} N(a, \mathbf{q}) + N(b, \mathbf{q}). \tag{AN2}$$

Therefore we have

$$\pi\left(w,(\mathbf{x},\mathbf{q})\right)=\pi_w(\mathbf{x},\mathbf{q})=(\mathbf{x}-|w|+N(w,\mathbf{q}),t_w(\mathbf{q})) \tag{AN3}$$

with $t_w:=t_{a_n}\circ\cdots\circ t_{a_1}$.

The following inequalities and equalities on vectors are coordinatewise.

Lemma 19 (Monotonicity) *Let $w,w'\in A^*$ such that $|w|\le|w'|$. Then for every $\mathbf{q}\in Q$*

$$N(w,\mathbf{q})\le N(w',\mathbf{q}).$$

Proof. For each $v\in V$ let $p_v:A^*\to A_v^*$, $p_v(a)=\begin{cases}a & \text{if } a\in A_v^*\\ \varepsilon & \text{if } a\notin A_v^*\end{cases}$.

$N(w,\mathbf{q})\overset{\overset{(\mathrm{AN2})}{\downarrow}}{=}\sum_{v\in V}N\left(p_v(w),\mathbf{q}\right)$ with $p_v(w)=p_v(a_1)\dots p_v(a_n)$ if $w=a_1\dots a_n$, hence it sufficies to prove the lemma for $w,w'\in A_v^*$, for $v\in V$.

Fixed $v\in V$ and $w,w'\in A_v^*$ with $|w|\le|w'|$, there exists $w''\in A_v^*$ such that $|ww''|=|w'|$.

For $a\in A_u$

if $(v,u)\notin E$,

$$N(w,\mathbf{q})_a=N(w'\mathbf{q})_a=0;$$

if $(v,u)\in E$,

$$\begin{aligned}N(w',\mathbf{q})_a=|T_{(v,u)}(w',q_v)|_a&\underset{\underset{\mathcal{P}_v\text{ abelian}}{\uparrow}}{=}|T_{(v,u)}(ww'',q_v)|_a\\&=\underbrace{|T_{(v,u)}(w,q_v)|_a}_{N(w,\mathbf{q})_a}+\underbrace{|T_{(v,u)}\left(w'',T_v(w,q_v)\right)|_a}_{\ge 0}\ge N(w,\mathbf{q})_a.\end{aligned}$$

□

Lemma 20 *$w,w'\in A^*$ such that $|w|=|w'|\implies\pi_w=\pi_{w'}$.*

Proof. $|w|=|w'|$, then for lemma 19, for every $q\in Q$, $N(w\mathbf{q})=N(w',\mathbf{q})$.

t_a and t_b commute for every $a,b\in A$ (because every processor is abelian), hence

$$t_w(\mathbf{q})=t_{w'}(\mathbf{q})$$

$$\Longrightarrow \underset{\pi(w,(\mathbf{x},\mathbf{q}))}{\underset{\|}{\pi_w(\mathbf{x},\mathbf{q})}} \overset{\text{(AN3)}}{\overset{\downarrow}{=}} (\underbrace{\mathbf{x}-|w|+N(w,\mathbf{q})}_{\mathbf{x}-|w'|+N(w',\mathbf{q})}, \underbrace{t_w(\mathbf{q})}_{t_{w'}(\mathbf{q})}) \overset{\text{(AN3)}}{\overset{\downarrow}{=}} \pi_{w'}(\mathbf{x},\mathbf{q}) \qquad \forall\ (\mathbf{x},\mathbf{q}) \in \mathbb{Z}^A \times Q.$$

□

We will call *execution* a word in A^*.

Fix an initial state $(\mathbf{x},\mathbf{q}) \in \mathbb{Z}^A \times Q$;

- a letter $a \in A$ is called a *legal move* from $(\mathbf{x},\mathbf{q})$ if $\mathbf{x}_a \geq 1$;
- an execution $w = a_1 \dots a_n \in A^*$ is called *legal* for $(\mathbf{x},\mathbf{q})$ if a_i is a legal move from $\pi_{a_1 \dots a_{i_1}}(\mathbf{x},\mathbf{q})$ for all $i = 1, \dots, n$;
- an execution $w \in A^*$ is called *complete* for $(\mathbf{x},\mathbf{q})$ if $\pi_w(\mathbf{x},\mathbf{q}) = (\mathbf{y},\mathbf{q}')$ for some $\mathbf{q}' \in Q$ and $\mathbf{y} \in \mathbb{Z}^A$ with $\mathbf{y}_a \leq 0$ for all $a \in A$.

Notice that a complete execution is not necessarily legal.

In chapter 3 we called these three concepts respectively legal operation, legal sequence of vertices and sequence of vertices that stabilizes a configuration; indeed, we have shown that the Oil and Water model is an abelian network.

The next lemma states that every processor in an abelian network performs the minimum possible number of operations to remove all the letters from the network.

Lemma 21 (Least action principle) *If $w \in A^*$ is legal for $(\mathbf{x},\mathbf{q})$ and $w' \in A^*$ is complete for $(\mathbf{x},\mathbf{q})$, then*

$$|w| \leq |w'|.$$

Proof. Let $w = a_1 \dots a_n$ and let, per impossible, $|w| \not\leq |w'|$.

Let i be the smallest index such that $|a_1 \dots a_i| \not\leq |w'|$.

Denote $u = a_1 \dots a_{i-1}$ and $a = a_i$. Then $|u|_a \geq |w'|_a$ and $|u|_b \leq |w'|_b$ for all $b \in \{a_1, \dots, a_{i-1}\}$.

w is legal for $(\mathbf{x},\mathbf{q})$, therefore at least one letter a is present in $\pi_u(\mathbf{x},\mathbf{q})$, namely

$$1 \leq \pi_u(\mathbf{x},\mathbf{q})_i = \mathbf{x}_a - |u|_a + N(u,\mathbf{q})_a \leq \mathbf{x}_a - |w'|_a + N(w',\mathbf{q}_a) = \pi_{w'}(\mathbf{x},\mathbf{q})_i$$
$$\underset{w' \text{ complete for } (\mathbf{x},\mathbf{q})}{\underset{\uparrow}{=}} (\mathbf{y},\mathbf{q}')_i = \mathbf{y}_i \leq 0 \qquad \text{for some } (\mathbf{y},\mathbf{q}') \in \mathbb{Z}^A \times Q$$

and we have the desired contradiction. □

Lemma 22 (Halting dichotomy) *Given an initial state* $(\mathbf{x}, \mathbf{q}) \in \mathbb{Z}^A \times Q$ *for an abelian network* $\mathcal{N}$*, either*

1) *there does not exist a finite complete execution for* $(\mathbf{x}, \mathbf{q})$*; or*

2) *every legal execution for* $(\mathbf{x}, \mathbf{q})$ *is finite, there exists a legal complete execution for* $(\mathbf{x}, \mathbf{q})$ *and every pair* (w, w') *of legal executions for* $(\mathbf{x}, \mathbf{q})$ *satisfies* $|w| = |w'|$*.*

Proof. If there exists a finite complete execution, let s be its length;
$\xRightarrow[\text{Lemma 21}]{}$ every legal execution has length smaller or equal to s.
There exists a legal execution (the empty word ε) and every legal execution of maximal length is complete (otherwise it could be extended by a legal move).
If w and w' are legal complete executions, for Lemma 21 $|w| \leq |w'| \leq |w|$. □

In case (1) every legal finite execution w can be extended by a legal move: since w is not complete, there exists $a \in A$ such that wa is legal, therefore in this case there is an infinite word $a_1 a_2 \dots$ such that $a_1 \dots a_n$ is a legal execution for every $n \geq 1$.
The halting status (a.1) is therefore the distinction that the halting dichotomy enunciates: in case (2), starting from the initial state $(\mathbf{x}, \mathbf{q})$, the process eventually stops, and in this scenario we say that $\mathcal{N}$ *halts* on input $(\mathbf{x}, \mathbf{q})$; in case (1), the opposite happens.

Definition 3 (Odometer) *If* $\mathcal{N}$ *halts on input* $(\mathbf{x}, \mathbf{q})$*,* $[\mathbf{x}, \mathbf{q}] \in \mathbb{N}^A$*, defined as* $[\mathbf{x}, \mathbf{q}]_a := |w|_a$ *(the total number of letters* a *processed during a complete legal execution* w *of* $(\mathbf{x}, \mathbf{q})$*) is called odometer of* $(\mathbf{x}, \mathbf{q})$*.*
Notice that by the halting dichotomy it does not depend on the chosen complete legal execution w*.*

In chapter 3, in the Oil and Water model, we had called odometer the function $(u(x))_{x \in V}$.

Theorem 9 *Abelian networks have properties (a.1), (a.2), (a.3), (a.4) and (a.5).*

Proof. For the halting dichotomy the halting status does not depend on the execution (item (a.1)); moreover, given $\mathcal{N}, \mathbf{x}, \mathbf{q}$, the odometer does not depend on the chosen legal execution (items (a.3), (a.4) and (a.5)); since the odometer and the initial state $\mathbf{q}$ are the ones determining the final state $t_w(\mathbf{q})$, we have item (a.2). □

6.3 Remarks and used theorems

Theorem 10 (Central Limit Theorem) *Let* $\boldsymbol{X}_1 = \left(X_{1(1)} \ldots, X_{1(k)}\right)$, $\boldsymbol{X}_2 = \left(X_{2(1)}, \ldots, X_{2(k)}\right), \ldots$ *be independent and identically distributed real random k-vectors with expected value* $\boldsymbol{\mu} = \boldsymbol{E}(\boldsymbol{X}_1) = \left(E(X_{1(1)}), \ldots, E(X_{1(k)})\right)$ *and finite covariance matrix* Σ. *For* $n \in \mathbb{N}$, *let* $\overline{\boldsymbol{X}}_n = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{X}_i$. *Then*

$$\sqrt{n}\left(\overline{\boldsymbol{X}}_n - \boldsymbol{\mu}\right) \xrightarrow{d} \boldsymbol{Z}, \qquad \text{with } \boldsymbol{Z} \sim \mathcal{N}(\boldsymbol{0}, \Sigma).$$

Remark 5 *The expected number of returns by time* t *of a symmetric simple random walk on* $\mathbb{Z}$, $\{X_t\}_{t\in\mathbb{N}_0}$, *is, for* $t \in 2\mathbb{N}_0$,

$$E_0\left(\sum_{n=0}^{t} \mathbb{1}_{\{X_t=0\}}\right) = (t+1)\binom{t}{t/2}\frac{1}{2^t} \tag{6.1}$$

and we can prove it by induction: for $t = 0$ *we have* $E_0\left(\mathbb{1}_{\{X_0=0\}}\right) = 1$ *and assuming the relation to be true for* t *we get*

$$\begin{aligned} E_0\left(\sum_{n=0}^{t+2} \mathbb{1}_{\{X_t=0\}}\right) &= E_0\left(\sum_{n=0}^{t} \mathbb{1}_{\{X_t=0\}}\right) + P_0(X_{t+2}=0) \\ &= (t+1)\binom{t}{t/2}\frac{1}{2^t} + \binom{t+2}{t/2+1}\frac{1}{2^{t+2}} \\ &= ((t+2)+1)\binom{t+2}{t/2+1}\frac{1}{2^{t+2}}, \end{aligned}$$

therefore we have (6.1).

Theorem 11 (Stirling's formula) *As* $n \to \infty$

$$n! \sim \sqrt{2\pi} n^{n+1/2} e^{-n}.$$

We present now two recalls and then enunciate Doob's optional sampling theorem.

Recall (Markov inequality) *If* X *is a real random variable and* $f : [0, +\infty) \to [0, +\infty)$ *is monotone increasing, for any* $\varepsilon > 0$ *with* $f(\varepsilon) > 0$,

$$P(|X| \geq \varepsilon) \leq \frac{E[f(|X|)]}{f(\varepsilon)}.$$

Recall (Measurable, adapted, progressively measurable, predictable process, martingale, submartingale, supermartingale) *Let* $X = \{X_t\}_{t\in I}$ *be a stochastic process on a probability space* $(\Omega, \mathcal{F}, P)$ *to* $(E, \mathcal{E})$ *and* $\{\mathcal{F}_t\}_{t\in I}$ *a filtration.*

◇ *X is **measurable** if the map $\Omega \times I \to E$ defined by $(\omega, t) \longmapsto X_t(\omega)$ is $(\mathcal{F} \otimes \mathcal{B}(I))$-measurable.*

◇ *X is **adapted** with respect to $\{\mathcal{F}_t\}_{t\in I}$ if X_t is $\mathcal{F}_t$-measurable for every $t \in I$.*

◇ *X is **progressively measurable** with respect to $\{\mathcal{F}_t\}_{t\in I}$ if, for every $T \in I$, the map $\Omega \times (I \cap [0,T]) \to E$ defined by $(\omega, t) \longmapsto X_t(\omega)$ is $(\mathcal{B}(I \cap [0,T]) \otimes \mathcal{F}_T)$-measurable.*

◇ *if $X = \{X_t\}_{t\in I}$ is real-valued, X is a **martingale** with respect to $\{\mathcal{F}_t\}_{t\in I}$ if*

a) $X_t \in L^1(\Omega, \mathcal{F}_t, P)$ for each $t \in I$;

b) $E(X_t|\mathcal{F}_s) = X_s$ a.s. for every $s, t \in I$ such that $s \le t$.

*If the condition (b) is $E(X_t|\mathcal{F}_s) \ge X_s$, then X is a **submartingale**.*
*If the condition (b) is $E(X_t|\mathcal{F}_s) \le X_s$, then X is a **supermartingale**.*

◇ *if $I = \{0, \ldots, T\}$ $(T \in \mathbb{N}_0 \cup \{\infty\})$, $X = \{X_t\}_{t\in I}$ is **predictable** with respect to $\{\mathcal{F}_t\}_{t\in I}$ if X_t is $\mathcal{F}_{t-1}$-measurable for every $t \in \{0, \ldots, T\} \setminus \{0\}$.*

Theorem 12 (Optional sampling theorem) *Let $\{M_t\}_{t\in I}$ be a discrete-time martingale and τ a stopping time with values in $\mathbb{N}_0 \cup \{\infty\}$, both with respect to a filtration $\{\mathcal{F}_t\}_{t\in\mathbb{N}_0}$. Assume that one of the following three conditions holds:*

a) *τ is almost surely bounded, i.e. there exists a constant $c > 0$ such that $\tau \le c$ a.s.;*

b) *$E(\tau) < \infty$ and there exists a constant c such that $E\left(|M_{t+1} - M_t||\mathcal{F}_t\right) \le c$ a.s. on the event $\{\tau > t\}$ for all $t \in \mathbb{N}_0$;*

c) *there exists a constant c such that $|M_{t\wedge\tau}| \le c$ for all $t \in \mathbb{N}_0$.*

Then M_τ is an almost surely well defined random variable and

$$E(M_\tau) = E(M_0).$$

Similarly, if M_t is a submartingale or a supermartingale instead and one of the above three conditions holds, then

$$E(M_\tau) \ge E(M_0)$$

for a submartingale, and

$$E(M_\tau) \leq E(M_0)$$

for a supermartingale.

Bibliography

[1] Andjel, E. D., (1982), *Invariant Measures for the Zero Range Process*, The Annals of Probability, Vol. 10, No. 3, pp. 525-547, Institute of Mathematical Statistics.
http://www.jstor.org/stable/2243365

[2] Asselah, A., & Gaudillière, A., (2013), *From logarithmic to subdiffusive polynomial fluctuations for internal DLA and related growth models*, The Annals of Probability, Vol. 41, No. 3A, pp. 1115-1159, Institute of Mathematical Statistics.
https://doi.org/10.1214/12-AOP762

[3] Asselah, A., & Gaudillière, A., (2013), *Sublogarithmic fluctuations for internal DLA*, The Annals of Probability, Vol. 41, No. 3A, Institute of Mathematical Statistics.
https://arxiv.org/abs/1011.4592

[4] Bak, P., Tang, C., & Wiesenfeld, K., (1988), *Self-organized criticality*, Physical Review A, Vol. 38, No. 1, pp. 364-374, American Physical Society,
https://link.aps.org/doi/10.1103/PhysRevA.38.364

[5] Bond, B. & Levine, L., (2015), *Abelian networks II. Halting on all inputs.*
https://arxiv.org/abs/1409.0169

[6] Bond, B. & Levine, L., (2016), *Abelian networks I. Foundations and examples.*
https://arxiv.org/abs/1309.3445

[7] Breiman, L., (1992), *Probability*, Classics in applied mathematics, Philadelphia (PA), Society for Industrial and Applied Mathematics.

[8] Candellero, E., Ganguly S., Hoffman, C., & Levine, L., (2014), *Oil and water: a two-type internal aggregation model.*
https://arxiv.org/abs/1408.0776

[9] Candellero, E., Stauffer, A. & Taggi, L., (2019), *Abelian oil and water dynamics does not have an absorbing-state phase transition*,
https://arxiv.org/abs/1901.08425

[10] Dickman, R., Trivellato Rolla, L., & Sidoravicius, V., (2010), *Activated Random Walkers: Facts, Conjectures and Challenges*, Journal of Statistical Physics, Vol. 138, No. 1-3, pp. 126–142, Springer Science and Business Media LLC.
http://dx.doi.org/10.1007/s10955-009-9918-7

[11] Freedman, D. A., (1965), *Bernard Friedman's Urn*, The Annals of Mathematical Statistics, Vol. 36, No. 3, pp. 956–970.
http://www.jstor.org/stable/2238205

[12] Friedman, B., (1949), *A simple urn model*, Communications on Pure & Applied Mathematics, Vol. 2, No. 1, pp. 59-70.

[13] Kelnke, A., (2014), *Probability Theory, A Comprehensive Course*, (Second Edition), Universitext, Heidelberg, Springer London.

[14] Lawler, G. F., (1991), *Intersections of Random Walks*, Probability and Its Applications, Boston (MA), Birkhäusen.

[15] Lawler, G. F., Bramson, M., & Griffeath, D., (1992), *Internal Diffusion Limited Aggregation*, The Annals of Probability, Vol. 20, No. 4, pp. 2117–2140.
http://www.jstor.org/stable/2244742

[16] Lawler, G. F., & Limic, V., (2010), *Random Walk: A Modern Introduction*, Cambridge Studies in Advanced Mathematics, New York, Cambridge University Press.

[17] Lyons, R., & Peres, Y., (2016), *Probability on trees and networks*, Cambridge Series in Statistical and Probabilistic Mathematics, Vol. 42, New York, Cambridge University Press.

[18] Moore, C., & Machta, J., (2000), *Internal Diffusion-Limited Aggregation: Parallel Algorithms and Complexity* Journal of Statistical Physics No. 99, pp. 661–690, Springer Science and Business Media LLC. https://doi.org/10.1023/A:1018627008925

[19] Roch, S., (2020), *Modern Discrete Probability: An Essential Toolkit*, Chap. 4, pp. 187-267.
https://people.math.wisc.edu/~roch/mdp/roch-mdp-chap4.pdf

[20] Shellef, E., (2010), *Nonfixation for Activated Random Walks.*
https://arxiv.org/abs/0910.3338

[21] Stauffer, A., & Taggi, L., (2017), *Critical density of activated random walks on transitive graphs.*
https://arxiv.org/abs/1512.02397

[22] Taggi, L., (2021), *Active phase for activated random walks on* $\mathbb{Z}^d$*,* $d \geq 3$*, with density less than one and arbitrary sleeping rate.*
https://arxiv.org/abs/1712.05292

[23] Taggi, L., (2021), *Essential enhancements in Abelian networks: continuity and uniform strict monotonicity.*
https://arxiv.org/abs/2003.00932

[24] Trivellato Rolla, L., (2008), *Generalized Hammersley process and phase transition for Acrivated Random Walk models*, Doctoral Thesis, PhD in Sciences, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, advisor: Vladas Sidoravicius.
http://leorolla.scienceontheweb.net/papers/rolla_thesis.pdf

[25] Trivellato Rolla, L., & Sidoravicius, V., (2011), *Absorbing-state phase transition for driven-dissipative stochastic dynamics on* $\mathbb{Z}$, Inventiones mathematicae, Vol. 188, No. 1, pp. 127-150, Springer Science and Business Media LLC.
https://arxiv.org/abs/0908.1152