

# Discrete interacting particle systems

November 7, 2016

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Preliminaries . . . . .	4
1.1.1	Graphs . . . . .	4
1.1.2	Random variables . . . . .	4
1.1.3	Notation . . . . .	5
1.2	Sweeping overview . . . . .	6
1.2.1	Random walk . . . . .	6
1.2.2	The bullet process . . . . .	6
1.2.3	The frog model . . . . .	7
1.2.4	Coalescing random walk . . . . .	7
1.2.5	Stochastic social dynamics . . . . .	7
1.2.6	Choice processes . . . . .	8
1.2.7	Spin and loop $O(n)$ models . . . . .	9
1.2.8	Random partitions . . . . .	10
<b>2</b>	<b>The bullet process</b>	<b>11</b>
2.1	Introduction . . . . .	11
2.2	Notation and statement of theorem . . . . .	12
2.3	Survival . . . . .	12
2.3.1	Warmup with random walk . . . . .	12
2.3.2	Proving survival . . . . .	13
2.4	Perishing . . . . .	18
2.4.1	Speeds in $\{1, 2, 3\}$ . . . . .	19
2.4.2	Proving a speed-1 bullet perishes . . . . .	21
2.5	Further results . . . . .	22
2.6	Excercises . . . . .	25
2.7	Possibly doable open problems . . . . .	26

<b>3</b>	<b>The frog model</b>	<b>27</b>
3.1	Related models . . . . .	27
3.2	Transience and recurrence for the frog model . . . . .	28
3.3	With drift on $\mathbb{Z}$ . . . . .	29
3.4	The frog model on trees . . . . .	31
3.4.1	Theorems to be presented . . . . .	31
3.5	Transience . . . . .	33
3.6	Poisson tree frogs . . . . .	38
3.6.1	A criterion for stochastic dominance . . . . .	40
3.6.2	Critical parameters for $d$ -ary trees . . . . .	41
3.6.3	Simplest proof of recurrence on the binary tree . . . . .	47
3.6.4	A more complicated bootstrap . . . . .	48
3.6.5	Transience regime . . . . .	54
3.7	Stochastic orders and the frog model . . . . .	55
3.7.1	Background material on stochastic orders . . . . .	55
3.7.2	Formalities . . . . .	57
3.7.3	Definition of icv and pgf statistics . . . . .	57
3.7.4	Statement of main theorem . . . . .	59
3.7.5	Applications . . . . .	60
3.7.6	Questions . . . . .	63
3.8	Recurrence with one-per-site . . . . .	64
3.8.1	Generating function recursion . . . . .	64
3.8.2	Proving recurrence . . . . .	68
3.9	Shape Theorems . . . . .	74
3.9.1	Obtaining a shape theorem . . . . .	79
3.9.2	icv and pgf monotonicity . . . . .	79
3.10	Frog wakeup time on finite $d$ -ary trees . . . . .	80
3.10.1	A Galton-Watson process . . . . .	81
3.10.2	Embedding $Z$ in the frog model . . . . .	82
3.10.3	Proof of Theorem 126 . . . . .	85
3.10.4	Speedup from multiple walks on $d$ -ary tree . . . . .	86
3.11	Frog Questions . . . . .	88
3.11.1	Open but maybe solvable with a nice idea . . . . .	88
<b>4</b>	<b>Coalescing random walk</b>	<b>88</b>
4.1	Introduction . . . . .	88
4.1.1	History . . . . .	90
4.2	Overview . . . . .	90
4.2.1	Further Questions . . . . .	91
4.2.2	Outline . . . . .	91

4.3	Site recurrence for coalescing random walk . . . . .	92
4.3.1	Site recurrence for bounded degree graphs . . . . .	95
4.3.2	Site recurrence for Galton-Watson trees . . . . .	96
4.3.3	Non-backtracking coalescing random walk on trees . . . . .	98
4.3.4	Recovering a dual . . . . .	100
4.3.5	Site recurrence . . . . .	102
<b>5</b>	<b>The power of two choices</b>	<b>103</b>
5.1	Introduction . . . . .	103
5.1.1	Discussion and further questions . . . . .	105
5.1.2	Overview . . . . .	105
5.2	Subintervals in $[0, \alpha]$ . . . . .	106
5.3	Proof of Theorem 147 . . . . .	108
5.3.1	Proposition 151 . . . . .	109
5.3.2	Proposition 152 . . . . .	112
5.3.3	Proposition 154 . . . . .	114

## 1 Introduction

Discrete interacting particle systems are toy models of real world phenomena. They are often intuitive to describe, and powerful in their applications. Moreover, there is no dearth of easy-to-state, but difficult-to-solve problems. Our goal is to present the major theorems, and open questions for some subset of the following processes: the frog model, the bullet process, coalescing random walk, stochastic social dynamics, choice interval splitting processes, and the spin and loop  $O(n)$  models.

Let's start with an informal definition.

**Definition 1.** *Discrete interacting particle system* are spatially related objects whose random behavior influences one another.

The important words here are spatial, objects, behavior and influence. Here are a couple of buzz words for what they might entail.

**spatial** graph,  $\mathbb{R}^d$ ,  $[0, 1]$ ,  $[n]$  partitions

**object** people, servers, particles, molecules

**behavior** movement, location, size, orientation

**influence** inform, impede, compete, augment

There is a huge amount of pure and applied work on these systems. We will focus on some examples that illustrate some new and old approaches to them.

## 1.1 Preliminaries

To properly explore this topic, we need a few preliminary definitions.

### 1.1.1 Graphs

**Definition 2.** A graph  $G = (V, E)$  is a collection of vertices and edges. Some common graphs for this course will be

- $K_n$  is the complete graph with  $n$  vertices.
- $C_n$  is the cycle on  $n$  vertices.
- $\mathbb{Z}_m^d$  and  $\mathbb{Z}_m^{\circ d}$  are the  $[-m, m]^d$  grid and the  $[-m, m]^d$  torus.
- $\mathbb{T}_d^n$  is the complete  $d$ -ary tree of height  $n$
- $\mathbb{T}_d$  is the infinite rooted  $d$ -ary tree
- $\mathbb{Z}^d$  is the  $d$ -ary integer lattice.

Most other commonly studied graphs are random. For example, a random graph with specified degree sequence, a Galton-Watson tree, or a percolated version of any of the above graphs ( $\mathbb{Z}^d$  and  $K_n$  are the most common).

### 1.1.2 Random variables

We will be interested in the usual random variables (r.v.'s), and will often use the notation  $X$  and  $Y$  for such. It is useful, but not essential for most of our purposes, to understand these from the measure theoretic perspective as functions on a probability measure space.

Here are some common random variables:

## DISCRETE

**Bernoulli( $p$ )** The “coinflip” random variable. It is 1 with probability  $p$  and 0 with probability  $1 - p$ .

**Binomial**( $n, p$ ) A sum of  $n$  Bernoulli( $p$ ) r.v.'s. We say that  $X \sim \text{Bin}(p)$ .

**Poisson**( $\lambda$ ) Has  $\mathbf{P}[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}$ . The nice properties of this include:

- $\text{Poi}(\lambda) + \text{Poi}(\lambda') \stackrel{d}{=} \text{Poi}(\lambda + \lambda')$ .
- Poisson thinning. That  $\text{Bin}(\text{Poi}(\lambda), p) \stackrel{d}{=} \text{Poi}(p\lambda)$  and the failures have *independent* distribution  $\text{Poi}((1-p)\lambda)$ . (Actually thinning is much stronger. We'll see this later.)

We say that  $X \sim \text{Poi}(p)$ .

**Geometric**( $p$ ) The number of trials until a success with  $\text{Ber}(p)$  trials.  $\mathbf{P}[X = k] = (1-p)^{k-1}p$ . These are nice because of the memoryless property:  $\mathbf{P}[X \geq k + j \mid X \geq k] = \mathbf{P}[X \geq j]$ . We say that  $X \sim \text{Geo}(p)$ .

## CONTINUOUS

**Uniform**(**0,1**) The quintessential continuous distribution. Any number is equally likely. The density function,  $f$ , corresponds to Lebesgue measure:  $f(x) = 1$ . We say that  $X \sim \text{Uni}(0, 1)$ .

**Exponential**( $\lambda$ ) Has density function  $f(x) = e^{-\lambda x}$ . The most useful for interacting particles and is canonically used to implement random delays between continuous time events. The two main reasons being

- Memoryless:  $\mathbf{P}[X \geq s + t \mid X \geq s] = \mathbf{P}[X \geq t]$ . This makes processes Markov.
- Minimums are also exponential:  $\min_{i \leq n} \{\exp(\lambda_i)\} \stackrel{d}{=} \exp(\sum \lambda_i)$ . Note that things are inverted, so larger  $\lambda$  means a smaller mean ( $\mathbf{E}X = 1/\lambda$ ).

We say that  $X \sim \exp(\lambda)$ .

### 1.1.3 Notation

It is basically impossible to compute closed formulas for specific cases. Most results are concerned with asymptotic behavior, or the behavior at infinity. We will use the notation  $f(n) = O(g(n))$  to denote  $f(n) \leq Cg(n)$  for some  $C$  and all large enough  $n$ . We will use  $f(n) = o(g(n))$  if  $f(n)/g(n) \rightarrow 0$ .

## 1.2 Sweeping overview

Let's take a look at all of the processes we will study. After defining each, we will describe one theorem and an open problem.

### 1.2.1 Random walk

We can define a process on a graph called *random walk*.

**Definition 3.** A particle starts at  $S(0) = v_0$ , then proceeds to uniformly select a neighbor. Let  $S(t)$  denote its position at time  $t$ .

We remark that  $t$  will often be used to indicate both discrete and continuous times. In discrete time  $S(t)$  is the position after  $t \in \mathbb{N}$  steps. In continuous time, we (to make it Markov) essentially always put an  $\exp(1)$  clock on each edge. Whichever rings first, the particle moves along. Here,  $S(t)$  does not necessarily correspond to  $t$  steps, and  $t \in [0, \infty)$ .

**Definition 4.** Let  $V = |\{t: S(t) = v_0\}|$ . A random walk is *recurrent* at  $v_0 \in G$  if  $\mathbf{P}[V = \infty] = 1$ . If  $\mathbf{P}[V = \infty] = 0$  it is called *transient*.

The most famous example of this is Polya's theorem (which we will discuss later, if need be).

**Theorem 5** (1914, Polya). *A random walk is recurrent with probability 1 on  $\mathbb{Z}^1$  and  $\mathbb{Z}^2$ , but for larger  $d$  is transient (i.e. has probability  $p_d < 1$  of returning to the origin, notice that this implies the number of visits is  $\text{Geo}(p_d)$ ).*

This theorem shows that infinite random objects can exhibit quite strange/unintuitive behavior. It is also a famous example of a *phase transition*. Namely, a dramatic change in behavior as some parameter is increased (in this case the parameter is  $d$ ).

### 1.2.2 The bullet process

**Definition 6** (Bullet process). Fire a bullet every second from the origin with i.i.d. speeds. Collisions result in mutual annihilation.

**Open Question 7** (Bullet problem). *Is there a phase transition for survival when bullet speeds are uniform(0, 1) r.v.'s?*

**Theorem 8** (2016, Dygert, J., Kinzel, Raymond, Slivken, Zhu). *When the speeds are discrete  $x_n < x_{n-1} < \dots < 3$  and equally likely, then there is a phase transition for survival. That is, sufficiently slow bullets are annihilated almost surely, and sufficiently fast bullets survive with positive probability.*

### 1.2.3 The frog model

**Definition 9** (Frog model). Place 1 awake particle at  $v_0 \in G$ , and 1 sleeping particle elsewhere. Awake particles perform random walk, and wake any sleeping particles they encounter.

We'll say that the frog model is recurrent if infinitely many frogs visit a specified vertex.

**Theorem 10** (1999, Telcs, Wormald). *The frog model is recurrent for all  $d$  on  $\mathbb{Z}^d$ .*

**Theorem 11** (2014, Hoffman, Johnson, J.). *It is recurrent on  $\mathbb{T}_2$  but transient on  $\mathbb{T}_d$  for  $d \geq 5$ .*

**Open Question 12.** *Transient or recurrent on  $\mathbb{T}_3$ ?*

### 1.2.4 Coalescing random walk

**Definition 13** (Coalescing random walk). Start an active particle at each site of a graph. They move in continuous time. When two particles collide they merge into one.

A theorem that heavily relies on past work on  $\mathbb{Z}^d$  by Bramson, Griffeath, and Kesten generalizes transience and recurrence results to a more general setting.

**Theorem 14** (2015, Benamini, Foxall, Gurel-Gurevich, J., Kesten). *Coalescing random walk is recurrent on any bounded degree graph.*

**Open Question 15.** *On  $\mathbb{Z}^d$  place  $\text{Ber}(\alpha/\|x\|^2)$  particles at each  $x$ . Is there an  $\alpha_c$  such that a phase transition for transience/recurrence occurs?*

### 1.2.5 Stochastic social dynamics

**Definition 16** (General definition). Fix a (usually finite) graph. Have neighbors meet then update states according to some rule (deterministic or random) according to exponential edge clocks.

**Definition 17** (Pandemic process). One vertex is sick. Whenever an edge adjacent to a sick particle rings, that neighbor becomes sick.

**Theorem 18** (1996, Aldous). *The expected time for every particle to become sick on  $K_n$  is  $2 \log(n) + o(\log(n))$ .*

The pandemic process is rather well-understood. [Ald13] proposes a variety of modifications. These are to be taken with a grain of salt given Aldous’s caveat [Ald13, p. 1125], “Keep in mind that it is easy to invent and simulate models, but hard to give rigorous proofs or to relate convincingly to real-world data.”

**Definition 19** (Compulsive gambler process). Each site starts with \$1. Whenever two neighbors meet, they play a fair game in which one takes all of the money. (i.e. if one has  $a$  and one has  $b$  then the first acquires all  $a + b$  with probability  $a/(a + b)$ .)

**Open Question 20.** On  $\mathbb{Z}_{n^{\circ}}^d$ , what is a typical stable distribution of the all the money?

We will talk about a few other processes like this later in the course.

### 1.2.6 Choice processes

The following processes are a very different sort of interacting particle system. These are spatially evolving process, in which the position of new points is influenced by the placements of old points.

**Definition 21** (balls and bins).  $n$  balls are placed into  $n$  bins one after the other in the following manner.

- *uniform allocation* Each is uniformly placed.
- *choice allocation* Two bins are selected uniformly. The ball is placed in the *least* full bin.

We study  $M_n$ , the bin with the most balls after the algorithm completes.

**Theorem 22** (1999, Azar, Broden, Karlin, Upfal). *For uniform allocation  $M_n \sim \log n / \log \log n$ . For choice allocation  $M_n \sim \log \log n / \log 2$ .*

Recently a continuous version of this was introduced.

**Definition 23.** Place points one after the other in  $[0, 1]$ . Consider the *uniform process* in which each point is a  $\text{uniform}(0,1)$  random variable. Compare to the *max-2 process* in which at each step two points are chosen uniformly. Both fall in some subinterval. The one in the largest subinterval is kept (break ties with a fair coin).

**Theorem 24** (1912, Weyl). *The uniform process is equidistributed.*



**Theorem 25** (2014, J.). *The max-2 process is equidistributed.*

**Open Question 26** (Benjamini). *The further-2 process is like the max-2, but instead the point that is furthest from all of the others (from two uniformly selected candidate points) is kept. Prove that the further-2 process is equidistributed.*

### 1.2.7 Spin and loop $O(n)$ models

The classical spin  $O(n)$  model is a model on a  $d$ -dimensional lattice in which a vector on the  $(n - 1)$ -dimensional sphere is assigned to every lattice site and the vectors at adjacent sites interact ferromagnetically via their inner product. Special cases include the Ising model ( $n = 1$ ), the XY model ( $n = 2$ ) and the Heisenberg model ( $n = 3$ ). We discuss questions of long-range order (spontaneous magnetization) and decay of correlations in the spin  $O(n)$  model for different combinations of the lattice dimension  $d$  and the spin dimension  $n$ .

**Definition 27** (Ising Model). Choose a configuration  $\sigma \in \{-1, 1\}^{\mathbb{Z}_n^d} := \Omega$  according to the probability measure

$$\mathbf{P}[\sigma] = \frac{1}{Z_{G,n,\beta}} \exp \left( \beta \sum_{\{u,v\} \in E} \sigma_u \sigma_v \right).$$

Notice that configurations in which neighbors have different spins are less likely. We think of  $\beta \in [0, \infty)$  as the inverse temperature of the system. So, very large  $\beta$  correspond to very low temperatures. This models atom spins at low temperatures. We are interested in the correlation

$$\rho_{x,y} = \mathbf{E} \langle \sigma_x, \sigma_y \rangle.$$

In the uniform case ( $\beta = 0$ ) we have  $\rho_{x,y} \equiv 0$ . The famous result for this is that  $\rho$  undergoes a phase transition as  $\beta$  increases.

**Theorem 28** (1944, Onsager). *Suppose that  $d \geq 2$ . There exists  $\beta_c \in (0, \infty)$  such that for  $\beta < \beta_c$  it holds that*

$$\rho_{x,y} \leq C_{d,\beta} e^{-c_{d,\beta} \|x-y\|}.$$

*While for  $\beta > \beta_c$  we have long-range order:*

$$\rho_{x,y} \geq c_{d,\beta} > 0$$

*for all  $x, y \in \mathbb{Z}_n^d$ . (Note the  $c_{d,\beta}$  are different.)*

**Open Question 29.** *Most of the major questions are generalizations of the set  $\{-1, 1\}$  to points on an  $n$ -sphere (spin models). We will get into this later.*

### 1.2.8 Random partitions

Because this involves algebra it is a bit harder to introduce. Instead we'll define a process that may seem a bit removed from DIPS. Later in the course we will tie this to spatial competition, and cycle structure of permutation.

**Definition 30** (Random sumsets). Let  $X_k \sim \text{Ber}(1/k)$  and  $M = \bigcup_{k=1}^{\infty} \mathbf{1}\{X_k = 1\}k$  be a random thinning of  $\mathbb{Z}^+$ . Define  $\text{sumset}(M) = \{\sum_{k \in A} k : A \subseteq M\}$ .

**Theorem 31** (Pemantle, Peres, Rivin, Eberhard, Ford, Green, 2015). *Let  $M_j$  be i.i.d. and define*

$$m_0 = \min\{m : \mathbf{P}[\bigcap_{j=1}^m \text{sumset}(M_j) = \emptyset] > 0\}.$$

*It holds that  $m_1 = 4$ .*

**Theorem 32** (Brita Montes de Oca, Fowler, J., Levy, 2016). *If  $X_k \sim \text{Ber}(\alpha/k)$  then*

$$m_\alpha = \begin{cases} \lceil (1 - \alpha \log 2)^{-1} \rceil, & 0 < \alpha < 1/\log 2 \\ \infty, & \alpha \geq 1/\log 2 \end{cases},$$

*so long as  $\lceil (1 - \alpha \log 2)^{-1} \rceil \notin \mathbb{Z}^+$ .*

**Open Question 33.** *What happens when  $\lceil (1 - \alpha \log 2)^{-1} \rceil \notin \mathbb{Z}^+$ .?*

**Open Question 34.** *Take  $\alpha = 2 > 1/\log 2$ . Recall that  $m_2 = \infty$ . But what if we thicken up the number of intersections. Does there exist a function  $f(n) \uparrow \infty$  such that*

$$\bigcap_{n=1}^{\infty} \bigcap_{j=1}^{f(n)} \text{sumset}(M_j \cap [1, n])$$

*is empty with positive probability?*

## 2 The bullet process

### 2.1 Introduction

We will consider a bullet process with speeds in  $\{1, 2, 3\}$ . It is trivial that an initial bullet with speed-3 survives with probability one. We will prove there is a phase transition for the slower two bullets. This uses all of the same ideas as the general discrete speeds mentioned in the introduction. See the bullets preprint on the website for a more general statement of everything. We mentioned that it is perhaps the simplest version of a ballistic annihilating system. Here are four others:

**Open Question 35** (non-backtracking annihilation). *Show there exists a transient graph  $G$  such that a 1-per-site annihilating upon contact system is recurrent when the particles perform non-backtracking random walk.*

**Open Question 36** (cannon balls). *Fire a cannon ball each second from  $0 \in \mathbb{R}^d$  with uniform  $(0, 1)$  speed and in a uniformly random direction. Collisions result in mutual annihilation. Show that the first cannon ball perishes.*

**Definition 37** (meteors, comes from Itai Benjamini). Place  $\epsilon$ -balls (meteors) in Euclidean space with centers according to a unit intensity Poisson process. At time 0 each chooses a direction uniformly randomly, and proceeds along this direction at speed-1. When two meteors collide (i.e. come within  $2\epsilon$  of one another's centers), they annihilate. Further discussion and partial progress for a related process on trees is in [BFGG<sup>+</sup>16]. Say that the origin is occupied if the center of a meteor is within  $\epsilon$  of it.

**Problem 38.** *Is the origin a.s. occupied by a meteor?*

The most important process for our purposes is the arrow-process. Although there are no published results, this was introduced by Vladas Sidoravicius and Laurent Tournier.

**Definition 39.** Fix some  $p < 1$  and assign to each arrow an i.i.d. speed. The canonical example here is speeds in  $\{-1, 0, 1\}$  with 0 being assigned with probability  $p$ , and speeds  $-1$  and  $1$  each with probability  $(1 - p)/2$ . All arrows simultaneously start moving at this speed and mutually annihilate upon colliding.

**Problem 40.** *In the  $\{-1, 0, 1\}$  arrow process, show that if the probability of a speed-0 arrow is sufficiently small, then every arrow perishes a.s.*

## 2.2 Notation and statement of theorem

Back to bullets. First we introduce some notation. Consider mutually annihilating bullets,  $\{b_1, b_2, \dots\}$ , fired from the origin along the real line at one second intervals. This delay between firings is not so important. All of the results here also hold for i.i.d. random firing times. The bullets have i.i.d. speeds  $\{s_1, s_2, \dots\}$ . We will assume that the speeds are uniformly in  $\{1, 2, 3\}$ . When two or more bullets collide, all of them are annihilated.

Let  $b_i \mapsto b_j$  denote the event  $b_i$  and  $b_j$  mutually annihilate. Define  $\tau$  to be the minimum index with  $b_\tau \mapsto b_1$ . The minimum is to account for the possibility of a simultaneous collision of several bullets. If  $b_1$  is never annihilated, set  $\tau = \infty$ . When  $\tau = \infty$ , we say that  $b_1$  *survives*. When  $\tau < \infty$  we say that  $b_1$  *perishes*.

**Theorem 41** (Dygart, J., Kinzel, Raymond, Slivken, Zhu, 2016).

$$\mathbf{P}[b_1 \text{ survives} \mid s_1 = 2] > 0, \text{ and } \mathbf{P}[b_1 \text{ survives} \mid s_1 = 1] = 0.$$

## 2.3 Survival

The idea for the proof of survival is that  $\tau$  can be recursively related back to independent copies of itself. In Proposition 45 we show that when the second bullet also has speed 2, then  $b_1$  survives “twice” as long. Also, that when the second bullet is slower than 2 it acts as a shield—increasing the survival time of  $b_1$ .

### 2.3.1 Warmup with random walk

As a warmup to our proof, we use a similar technique to show that the return time to 0 of a biased random walk started at 1 is infinite with positive probability. Call this time  $\kappa$ . Let  $S_1$  be the position of the walk after one step, and suppose  $S_1 = 2$  with probability  $p > 1/2$ . Conditioning on the first step, we can write

$$\kappa \stackrel{d}{=} \mathbf{1}_{\{S_1=0\}} + \mathbf{1}_{\{S_1=2\}}(\kappa_1 + \kappa_2 + 1).$$

Here  $\kappa_1$  and  $\kappa_2$  are i.i.d. copies of  $\kappa$ . When  $S_1 = 2$ , then the walk must return to 1, then return to 0. Each takes an independent  $\kappa$ -distributed amount of steps.

To show that  $\kappa$  is infinite with positive probability we introduce the probability generating function  $g(x) = \mathbf{E}x^\kappa$ . The distributional relationship

with  $\kappa, \kappa_1$ , and  $\kappa_2$  implies that  $g(x) = (1 - p)x + pxg(x)^2$ . After applying the quadratic formula, we have the closed form

$$g(x) = \frac{1 - \sqrt{1 - 4p(1 - p)x^2}}{2p}.$$

The series representation,  $g(x) = \sum \mathbf{P}[\kappa = j]x^j$ , implies that  $g(1) = \mathbf{P}[\kappa < \infty]$ . It is simple to check with the above formula that  $g(1) = \frac{1-p}{p} < 1$ .

### 2.3.2 Proving survival

The approach to establishing survival follows the random walk blueprint. However, we have a distributional inequality rather than equality. This requires a more subtle renewal property, and then additional work to push through the argument with generating functions.

**Lemma 42.** *If  $b_\tau \mapsto b_j$  and  $s_\tau = 3$  with  $j$  any fixed index, then the random variables  $\{\tau, s_{\tau+1}, s_{\tau+2}, \dots\}$  are independent.*

*Proof.* Necessarily  $s_\tau = 3$ , the fastest speed, so the bullets behind it do not interfere. Thus the event  $\{b_\tau \mapsto b_1\}$  depends only on the bullet speeds  $s_1, s_2, \dots, s_\tau$ .  $\square$

A longer range renewal property holds for other annihilations, where outside of a certain window the bullet speeds become independent.

**Lemma 43.** *Let  $E = E(S, i, j) = \{b_i \mapsto b_j, s_j = 2\}$  be the event that  $b_i$  annihilate  $b_j$  with  $s_j = 2$ . There exists a finite integer  $a = a(i, j)$  such that the bullet speeds  $s_{i+a}, s_{i+a+1}, \dots$  are independent of the event  $E$ .*

*Proof.* Given  $(i, j)$  let  $a$  be such that a maximal speed bullet fired at time  $i + a$  cannot reach  $b_i$  before  $b_i \mapsto b_j$ . The event  $b_i \mapsto b_j$  is thus unaffected by the bullet speeds  $s_{i+a}, s_{i+a+1}, \dots$ . The independence claim follows.  $\square$

We can now show that  $\tau$  stochastically dominates an equation involving independent copies of itself. Recall that  $X \succeq Y$  means there is a coupling such that  $X \geq Y$ , or equivalently  $\mathbf{P}[X \geq a] \geq \mathbf{P}[Y \geq a]$  for all  $a \in \mathbb{R}$ . Formally, a *coupling* is a probability space with (possibly dependent) random variables  $X'$  and  $Y'$  that have the same distribution as  $X$  and  $Y$ , respectively. A thorough reference on stochastic domination is [SS07].

**Example 44.** An elementary example of using couplings to show stochastic domination is the following. Let  $p < q$  and  $X \sim \text{Ber}(q)$  and  $Y \sim \text{Ber}(p)$ . It

holds that  $X \succeq Y$ . The coupling is obtained on a different probability space from  $U \sim \text{uniform}(0, 1)$ . Observe that

$$X' = \mathbf{1}\{U \leq q\} \sim X, \quad Y' = \mathbf{1}\{U \leq p\} \sim Y,$$

and that for all realizations of  $U$  we have  $X' \leq Y'$ .

We will deal with coupling a lot in this course. Here is our first more serious instance of this.

**Proposition 45.** *Let  $\tau_1, \dots, \tau_5$  be i.i.d. copies of  $\tau$ . Let  $X_\epsilon \sim \text{Ber}(\epsilon)$  also be independent. There exists  $\epsilon > 0$  such that*

$$\tau \succeq \mathbf{1}\{s_2 = 1\}(X_\epsilon(\tau_1 + \tau_2) + (1 - X_\epsilon)\tau_3) \tag{1}$$

$$+ \mathbf{1}\{s_2 = 2\}(\tau_4 + \tau_5) \tag{2}$$

$$+ \mathbf{1}\{s_2 = 3\}. \tag{3}$$

*Proof.* We will establish each line of the above in reverse order by conditioning on the value of  $s_2$ . When  $s_2 = 3$  as in (3), we have  $b_2 \mapsto b_1$  deterministically. Although  $\tau = 2$  on this event, it will simplify our calculations later to use the indicator function as a lower bound.

When  $s_2 = 2$  as in (2), suppose that  $b_\sigma$  destroys  $b_2$ . We have translated the original setup by one index, so  $\sigma \sim \tau + 1$ . Only the fastest bullet can destroy  $b_2$ , thus  $s_\sigma = 3$ . Lemma 42 ensures that the speeds  $s_{\sigma+1}, s_{\sigma+2}, \dots$  are independent of  $\sigma$ . Suppose that  $b_{\sigma'} \mapsto b_1$ . Once again this is the first unobstructed speed-3 bullet after  $b_\sigma$ . Thus  $\sigma' - \sigma \sim \tau$ , and this difference is independent of  $\sigma$ . This is where the term  $\tau_4 + \tau_5$  in (2) comes from (see Figure 1).

The pivotal case is (1), when  $s_2 < 2$ . The idea is that  $b_2$  acts as a shield, and causes an  $\epsilon$ -bias for the bullets close behind it to have speed 2. To see this rigorously, suppose that  $b_\gamma$  destroys  $b_2$ . First note that if  $\gamma$  is infinite with positive probability, then so is  $\tau$  and our theorem is proven. Accordingly, let us suppose that  $\gamma$  is a.s. finite. Also note that in order for  $b_\gamma \mapsto b_2$  to occur, all of the bullets  $b_3, \dots, b_{\gamma-1}$  must mutually annihilate. We can then ignore them for the remainder of the argument.

When  $s_\gamma = 3$ , it resets the model just as in the  $s_2 = 2$  case, and  $b_1$  survives until a bullet with index distributed as  $\tau + \gamma$  destroys it. There is also the possibility that  $s_2 = 2$ . When this occurs, let  $a = a(\gamma, 2)$  be the largest index for which  $b_{\gamma+a}$  could intercept  $b_\gamma$ . Bullets with indices in the set  $I = \{\gamma + 1, \dots, \gamma + a\}$  are dependent upon  $s_\gamma, s_2$ , and  $\gamma$ . In particular

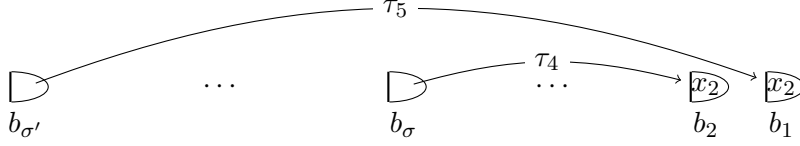


Figure 1: The picture when  $s_2 = x_2$ . The bullets  $b_2$  and  $b_{\sigma}$  are annihilated by a bullet that is fired a  $\tau$ -distributed number of seconds after it.

bullets faster than  $s_{\gamma}$  can survive to intercept  $b_{\gamma}$ . By Lemma 43 the bullets with indices larger than  $\gamma + a$  are once again independent (see Figure 2).

The model resets after  $\gamma + a$ . Let's restrict our attention to just the bullets with indices in  $I$ . That is, consider a bullet model with only  $|I|$  bullets, with speeds conditioned so that  $b_{\gamma} \mapsto b_2$  with  $s_2 = 1$ . Since  $b_{\gamma} \mapsto b_2$ , no speed-3 bullets in  $I$  can survive. As  $a$  is finite, with some positive probability,  $\epsilon > 0$ , all of the surviving bullets in  $I$  will have speed 2. For example, there is a positive probability of a sequence of alternating between bullets with speeds from  $S'$  and bullets with speed 2. When this occurs we use the argument from (2) to deduce that the index of the destroying bullet is at least the sum of two independent copies of  $\tau$ . This corresponds to  $\tau_1 + \tau_2$  in (1).

In all other realizations of the bullet speeds in  $I$ , we only have speed-1 bullets surviving. These slower bullets only prolong the survival of  $b_1$ . This is because we can repeat this argument for the largest index surviving bullet in  $I$ , and again obtain a new window of speed-2 bullets, or one of slower bullets. Thus, it is monotonically worse to remove all of the bullets in  $I$ , and restart the model with  $s_1 = 2$  and setting  $s_i = s_{\gamma+a+i}$  for  $i \geq 2$ . The survival of  $b_1$  in this settings is again distributed as  $\tau$ . This corresponds to  $\tau_3$  in (1). We then attach the Bernoulli random variable  $X_{\epsilon}$  to the event that all of the surviving bullets in  $I$  have speed 2. Otherwise, we have established that  $b_1$  survives at least a  $\tau$  distributed amount of time.  $\square$

Our goal now is to show that any random variable satisfying the recursive distributional equation at Proposition 45 will place some mass on infinity. With  $\epsilon$  as in Proposition 45, we introduce an operator,  $\mathcal{A} = \mathcal{A}(\mu)$ , that acts on nonnegative integer valued random variables. Given such a random variable  $T$ , we let  $s \in S$  be sampled according to  $\mu$ , and  $X_{\epsilon}$  a Bernoulli( $\epsilon$ ), both independent of one another. Also we take  $T_i$  to be i.i.d. copies of  $T$ .

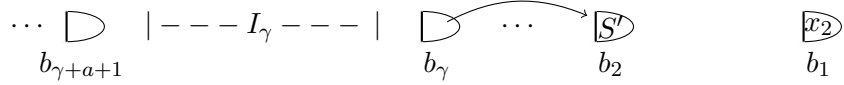


Figure 2: The picture when  $s_2 \in S'$ . If  $b_{\gamma} \mapsto b_2$ , then there is an interval of bullets behind it that contains no surviving  $x_1$ -speed bullets. With probability at least  $\epsilon$  it contains a surviving speed- $x_2$  bullet. Bullets  $b_{\gamma+a+1}$  onward are i.i.d.

We define a new random variable  $\mathcal{A}T$  to have distribution

$$\mathcal{A}T \stackrel{d}{=} \mathbf{1}\{s < x_2\}(X_{\epsilon}(T_1 + T_2) + (1 - X_{\epsilon})T_3) + \mathbf{1}\{s = x_2\}(T_4 + T_5) + \mathbf{1}\{s = x_1\}.$$

By Proposition 45 we have

$$\tau \succeq \mathcal{A}\tau. \quad (4)$$

We will show that  $\mathbf{P}[\tau = \infty] > 0$  in the following way. We first prove in Lemma 46 that  $\mathcal{A}$  is monotonic with respect to stochastic domination. Then, we show in Lemma 47 that  $\mathcal{A}^{\infty}\tau$  is equal to a unique fixed point,  $\tau^*$ . Combine this with (4), and we have  $\tau \succeq \tau^*$ . Finally, we prove in Proposition 48 that any fixed point of  $\mathcal{A}$  places positive probability on  $\infty$ . From this it follows that  $\tau$  is infinite with positive probability. We go through this rigamarole with  $\tau^*$  because, just as in the warm-up argument with the return time of a biased random walk, we seek an exact solution to the generating function relationship.

**Lemma 46.** *If  $T \succeq T'$  then  $\mathcal{A}T \succeq \mathcal{A}T'$ .*

*Proof.* Follows from the canonical coupling which sets each  $T_i \geq (T'_i)$ .  $\square$

**Lemma 47.** *Let  $\mathcal{A}^n$  denote  $n$  iterations of  $\mathcal{A}$ . It holds that  $\mathcal{A}^{\infty}\tau \stackrel{d}{=} \tau^*$  with  $\tau^*$  unique and  $\tau^* \stackrel{d}{=} \mathcal{A}\tau^*$ .*

*Proof.* Let  $F_n(k) = \mathbf{P}[\mathcal{A}^n\tau \leq k]$  be the cumulative distribution function of  $\mathcal{A}^n\tau$ . By the previous lemma we have  $\{F_n(k)\}_{n=0}^{\infty}$  is an increasing, bounded sequence. Let  $F(k)$  denote its limit. The function  $F(k)$  is non-decreasing and belongs to  $[0, 1]$ . Thus,  $F(k)$  is the density function of some random variable  $\tau^*$ . This limiting random variable must be fixed by  $\mathcal{A}$ , since an additional iteration  $\mathcal{A}(\mathcal{A}^{\infty}\tau)$  will not change the distribution.  $\square$



Note that  $\tau^* \succeq \mathbf{1}\{s = 3\}$ , and so  $\tau^*$  is not identically 0. In fact, the following proposition shows that any fixed point of  $\mathcal{A}$  places some mass on infinity.

**Proposition 48.** *With  $\epsilon$  as in Proposition 45 it holds that  $\mathbf{P}[\tau^* = \infty] > 0$ .*

*Proof.* Let  $f(x) = \mathbf{E}x^{\tau^*}$ . Denote  $f(1^-) = \lim_{x \rightarrow 1^-} f(x)$ . Since the coefficients of the power-series expansion of  $f$  are exactly the point probabilities of  $\tau^*$ , we have  $f(1^-) = 1 - \mathbf{P}[\tau^* = \infty]$ .

Using independence, we can write  $\mathbf{E}x^{A\tau^*}$  in terms of  $f$  to obtain

$$f(x) = \frac{1}{3}(\epsilon f(x)^2 + (1 - \epsilon)f(x) + f(x)^2 + x).$$

We can rewrite this as a quadratic equation in  $f(x)$

$$0 = (1 + \epsilon)f(x)^2 - (2 + \epsilon)f(x) + x \tag{5}$$

The discriminant,  $D(x) = (2 + \epsilon)^2 - 4(1 + \epsilon)x$ , is strictly decreasing, and thus minimized at  $x = 1$ . One can check that  $D(1) = \epsilon^2 > 0$ .

(5) has exactly two solutions for  $0 \leq x < 1$ :

$$f(x) = \frac{(2 + \epsilon) \pm \sqrt{D(x)}}{2(1 + \epsilon)}.$$

We use the ‘-’ root, since  $f(0) = 0$ . The quadratic formula gives the closed form

$$f(x) = \frac{(2 + \epsilon) - \sqrt{D(x)}}{2(1 + \epsilon)}.$$

Recall, we are interested in  $f(1^-)$ . It is straightforward to evaluate the above formula at  $x = 1$ . This yields

$$f(1^-) = \frac{2}{2(1 + \epsilon)},$$

which is less than one. □

*Proof of survival in Theorem 41.* By (4), Lemma 46, and Proposition 48 we have  $\tau$  is stochastically larger than a random variable that is infinite with positive probability. Hence  $\tau$  is infinite with positive probability. □

Note that this proof works the same if we replace  $\{1, 2, 3\}$  with any speeds  $x_3 < x_2 < x_1$ . We will need this to prove that a speed-1 bullet does not survive.

## 2.4 Perishing

The arrow process defined earlier is important here. We will use the notation  $a_i$  for the arrow at  $i \in \mathbb{Z}$ , and  $r_i$  for its speed. Recall that each  $r_i$  is i.i.d. in  $\mathbb{R}$ , and each arrow,  $a_i$ , begins moving simultaneously along  $\mathbb{R}$  at speed  $r_i$ . As with bullets, collisions result in mutual annihilation.

Given a realization of the bullet speeds  $\vec{s} = (s_i)_{i=1}^{\infty}$  we can partition  $\mathbb{Z}^+ := \{1, 2, \dots\}$  into the indices of bullets that mutually annihilate one another:

$$B_{\vec{s}} = \left\{ \{i \in \mathbb{Z}^+ \text{ such that } b_i \mapsto b_k\} : k \in \mathbb{Z}^+ \right\}.$$

Note this is the set of bullets that actually (as opposed to potentially) annihilate one another. We can define a similar collection for the indices of colliding arrows. Let  $\vec{r}$  be a realization of arrow speeds on  $\mathbb{Z}^+$ . Letting  $a_i \mapsto a_k$  denote annihilation of two arrows, define

$$A_{\vec{r}} = \left\{ \{i \in \mathbb{Z}^+ \text{ such that } a_i \mapsto a_k\} : k \in \mathbb{Z}^+ \right\}.$$

We say that  $\vec{s}$  and  $\vec{r}$  are *annihilation equivalent* if  $A_{\vec{r}} = B_{\vec{s}}$ . The following lemma gives a formula to translate between these two processes.

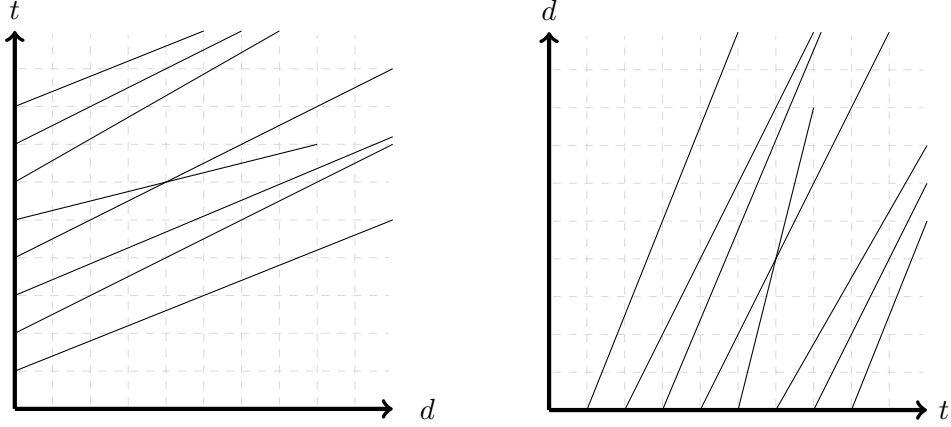
**Lemma 49.**  $\vec{s}$  and  $\vec{r} = (1/s_i)_{i=1}^{\infty}$  are annihilation equivalent.

*Proof.* Consider a graphical representation of the bullet problem with the  $x$ -axis the distance traveled, and the  $y$ -axis the time elapsed. Annihilations are then the same as the intersection of two lines. When we invert the axes, and reflect we obtain an arrow process on  $\mathbb{Z}^+$  where the bullet fired at time  $t = i$  now corresponds to an arrow with speed  $1/s_i$ . See Figure 3.  $\square$

We will also need to transform from an arrow process on  $\mathbb{Z}^-$  to a bullet problem. Given  $\vec{r}^- = (r_i)_{i=-1}^{\infty}$  we say that  $\vec{r}^-$  and  $\vec{s}$  are annihilation equivalent if  $-A_{\vec{r}^-} = B_{\vec{s}}$ .

**Lemma 50.** Suppose that arrow speeds are in the interval  $[r_n, r_1]$ . Fix a realization  $\vec{r}^-$  on  $\mathbb{Z}^-$  and let  $G(x) = (r_n + r_1 - x)^{-1}$ . It holds that  $\vec{r}^-$  and a bullet process with realization  $(G(r_i))_{i=-1}^{\infty}$  are annihilation equivalent.

*Proof.* Because arrow annihilations depend only on the relative speeds, we have the same annihilation behavior when we translate all speeds by the same constant. Shift each arrow by  $-(r_n + r_1)$ . When we reflect and rotate the axes we obtain a bullet problem with positive speeds given by  $s(b_i) = G(r_i)$ . See Figure 4.  $\square$



(a) Fire a bullet each second and plot its distance from the origin.

(b) This is equivalent to the arrow process with the inverse speeds. It is obtained by rotating  $90^\circ$ , then reflecting.

Figure 3: Transforming a bullet process into an equivalent arrow process.

Here are a few properties of  $T = G(1/s)$ .

**Lemma 51.** *Let  $T: [s_n, s_1] \rightarrow [s_n, s_1]$  be given by  $T(x) = (s_n^{-1} + s_1^{-1} - x^{-1})^{-1}$ . The function  $T$  has the following properties:*

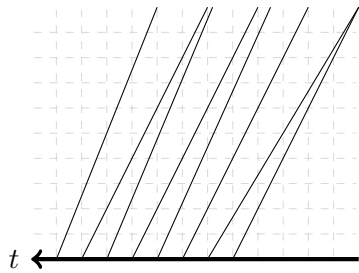
- (i)  $T$  is decreasing.
- (ii)  $T(x^*) = x^*$  for  $x^* = \frac{2s_1s_n}{s_1+s_n}$ . This fixed point is unique.
- (iii)  $T^{-1}(x) = T(x)$ .
- (iv)  $T(s_1) = s_n$  and  $T(s_n) = s_1$ .

#### 2.4.1 Speeds in $\{1, 2, 3\}$

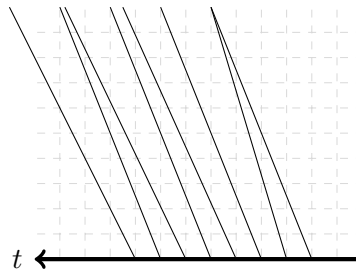
We will start by describing a transformation from the bullet problem on  $\{1, 2, 3\}$  to the arrow process from the introduction.

**Proposition 52.** *The bullet problem on  $\{1, 2, 3\}$  is equivalent to a one-sided arrow process on  $\mathbb{Z}^+$  with speeds sampled from  $\{1, \frac{1}{2}, \frac{1}{3}\}$ .*

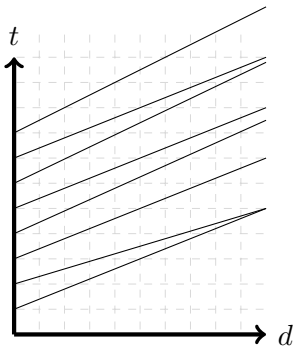
*Proof.* This comes from flipping the axes of the graphical representation for the bullet problem.  $\square$



(a) Start with an arrow process on  $\mathbb{Z}^-$ .



(b) Translate the arrow speeds by the same constant.



(c) Now when we rotate we obtain a bullet process with positive speeds.

Figure 4: Transforming an arrow process on  $\mathbb{Z}^-$  into an equivalent bullet process.

**Proposition 53.** *An arrow problem with speeds  $\{1, \frac{1}{2}, \frac{1}{3}\}$  on  $\mathbb{Z}^-$  is equivalent to a bullet problem with speeds  $\{1, \frac{6}{5}, 3\}$ .*

*Proof.* Shift the arrow speeds by  $-1 - \frac{1}{3} = -\frac{4}{3}$ , so we have  $\{-\frac{1}{3}, -\frac{5}{6}, -1\}$  then switch the axes to obtain a bullet problem with the claimed speeds.  $\square$

#### 2.4.2 Proving a speed-1 bullet perishes

The proof goes by contradiction. When we assume a slow bullet survives, we show that this implies two arrow speeds survive with positive probability in a two-sided arrow process. This contradicts the following proposition (which was shared with us by Alexandre Stauffer, Lorenzo Taggi, and Yuval Peres).

**Proposition 54.** *In an arrow process on  $\mathbb{Z}$ , there can be at most one arrow speed that occurs in the process, and survives with positive probability.*

*Proof.* Suppose there are two different arrow speeds,  $y \neq y'$ , that have a positive probability of surviving, and of occurring. Translation invariance ensures the arrow process on  $\mathbb{Z}$  is ergodic. By the Birkhoff ergodic theorem we have a positive fraction of bullets with these two different speeds that are never annihilated. More specifically, this ergodic theorem is

**Theorem 55** (Birkhoff ergodic theorem). *Let  $(\Omega, \mathcal{B}, \mu)$  be a probability space and  $T: \Omega \rightarrow \Omega$  be an ergodic measure-preserving transformation. Let  $f \in L^1(\Omega, \mathcal{B}, \mu)$ . Then*

$$\frac{1}{n} \sum_{j=1}^{n-1} f(T^j \omega) \rightarrow \int f d\mu.$$

for  $\mu$ -a.e.  $\omega \in \Omega$ .

Think of  $\Omega = R^{\mathbb{Z}}$  with  $R$  the set of possibly arrow speeds as the space of all realizations of arrow speeds. Thus an element  $\omega \in \Omega$  has  $r_i = \omega_i$ . If we let  $f(\omega) = \mathbf{1}\{r_1 = y \text{ and } r_1 \text{ survives}\}$  and  $T$  be the shift operator  $T((\omega_i)_{i \in \mathbb{Z}}) = (\omega_{i-1})_{i \in \mathbb{Z}}$  then  $T$  is ergodic (i.e. the only sets with  $\mu(T(A)) = \mu(A)$  are measure 0 or 1 sets). Notice that

$$T^j(f(\omega)) = \mathbf{1}\{r_{1-j} = y \text{ and } r_{1-j} \text{ survives}\}.$$

This means that

$$\sum_{j=1}^{n-1} f(T^j \omega) = \# \text{ of surviving speed } y \text{ bullets in } [0, n-1].$$

Also  $\int f d\mu = \mathbf{P}[r_1 = y \text{ and } r_1 \text{ survives}] > 0$ . So, if we apply the theorem we see that a dense set of speed  $y$  arrows survives. The same reasoning applies to speed  $y'$  arrows.

This is a contradiction since these infinitely surviving arrows with differing speeds must meet one another, and thus cannot survive forever.  $\square$

*Proof of Theorem 41.* Suppose that

$$\mathbf{P}[b_1 \text{ survives} \mid s_1 = 1] = p > 0.$$

Transform the bullet process with  $s_1 = 1$  to an arrow process on all of  $\mathbb{Z}$  with speeds  $\{1/3, 1/2, 1\}$  as in the previous two propositions. The probability the image of  $b_1$  is never annihilated by an arrow from the right is  $p$ , and never from the left is 1. However, we have the image of a speed 2 bullet survives with positive probability from both the left and the right (using the equivalence to a  $\{1, 6/5, 3\}$  bullet problem.) Thus, we have two arrow speeds  $1/2$  and  $1$  that survive with positive probability in the arrow process on  $\mathbb{Z}$ . This contradicts Proposition 54.  $\square$

## 2.5 Further results

The bullets have i.i.d. speeds  $\{s_1, s_2, \dots\}$ . We will assume that the speeds are obtained from a probability space  $(S, \mu)$  with  $S \subseteq (0, \infty)$  and  $\mu$  a probability measure on  $S$ .

**Theorem 56.** *Suppose that  $S = S' \cup \{x_2, x_1\}$  where  $S' \subseteq (0, x_2)$  and  $0 < x_2 < x_1 < \infty$ . If  $\mu(S') > 0$  and  $\mu(\{x_2\}) \geq \mu(\{x_1\})$ , then*

$$\mathbf{P}[b_1 \text{ survives} \mid s_1 = x_2] > 0.$$

A counterpart for slow speeds holds. When there are two slowest bullets, the slowest bullet perishes almost surely. Notice that in this setting we have a more stringent requirement on  $\mu$ .

**Theorem 57.** *Suppose that  $S = S'' \cup \{y_2, y_1\}$  where  $S'' \subseteq (y_1, \infty)$  and  $0 < y_2 < y_1$ . If  $\mu(S'') > 0$  and  $\mu(\{y_1\}) = \mu(\{y_2\})$ , then*

$$\mathbf{P}[b_1 \text{ perishes} \mid s_1 = y_2] = 1.$$

Notice the phase transition for discrete speeds follows from these.

Note that the intermediate bullet speeds do not need to be discrete. For instance, one could deduce a similar phase transition when  $S = \{1, 2, 3, 4\} \cup (2, 3)$  with the uniform measure on  $(2, 3)$ .

An extension of the idea in the proof of perishing shows that for certain speed distributions at least half of the speeds perish. This is especially powerful, because we can interpolate to absolutely continuous (with respect to Lebesgue measure) measures on bullet speeds.

**Theorem 58.** *Suppose that  $S = [x_n, x_1]$  with  $0 < x_n < x_1$ , and let  $\mu$  be any probability measure on  $S$  satisfying*

$$\mu([x_n, x]) = \mu([T(x), x_1]), \quad \forall x \in S, \quad (6)$$

with  $T(x) = (x_n^{-1} + x_1^{-1} - x^{-1})^{-1}$ . In such an  $(S, \mu)$ -bullet process

$$\mathbf{P}[b_1 \text{ survives} \mid s_1 < x^*] = 0,$$

where  $x^* = \frac{2x_1x_n}{x_1+x_n}$  is the unique fixed point of  $T$ .

*Proof.* To show a contradiction suppose that

$$\mathbf{P}[b_1 \text{ survives} \mid s(b_1) < x^*] = p > 0. \quad (7)$$

Notice that by monotonicity of the bullet process this implies that

$$\mathbf{P}[b_1 \text{ survives} \mid s(b_1) \geq x^*] = q \geq p. \quad (8)$$

Transform the bullet process conditioned on  $s(b_1) < x^*$  to an arrow process on  $\mathbb{Z}^+$  as in Lemma 49. Now, extend the arrow process to all of  $\mathbb{Z}$ . Our hypothesis and equivalence in Lemma 49 ensures that the probability  $a_1$  is never annihilated by an arrow from the right conditional on its speed being greater than  $1/x^*$  is  $p$ .

We claim that the requirement at (6) ensures that the probability that  $a_1$  is never destroyed by an arrow from the left is  $q$ . Indeed, Lemma 50 ensures that the arrow process restricted to  $a_1, a_0, a_{-1}, \dots$  is annihilation equivalent to a bullet process with speeds distributed as  $T(s(b_1))$ . The facts that  $T = T^{-1}$  and  $T$  is decreasing (see Lemma 51 (i) and (iii)) imply

$$\mathbf{P}[T(X) \leq x] = \mathbf{P}[X \geq T(x)] = \mu([T(x), s_1]).$$

By the hypothesis  $\mu([s_n, x]) = \mu([T(x), s_1])$ , we deduce that

$$\mathbf{P}[T(X) \leq x] = \mathbf{P}[X \leq x].$$

Thus,  $T(X)$  and  $X$  are identically distributed. This means that the induced arrow process from  $(S, \mu)$  on  $Z^- \cup \{0, 1\}$  is also equivalent to an  $(S, \mu)$ -bullet

process. However, the image of  $r_1$  in this new process has speed  $T(r_1)$ . By Lemma 51 (i) and (ii) we have  $T(s(b_1)) > s(b_1)$ . So, by (8) we have  $a_1$  is not annihilated by any arrow from the left with probability  $q$ .

No bullets reaching  $a_1$  from the left, and no bullets reaching  $a_1$  from the right are independent events. This independence along with the previous paragraph ensures that  $a_1$  survives with probability  $pq > 0$ . By symmetry, arrows with speed less than  $x^*$  survive with positive probability (except the probability of survival from the right is now  $q$  and from the left is now  $p$ ). We then have different arrow speeds surviving with positive probability, but this contradicts Proposition 54. □

Above we have  $\mu([x_n, x^*]) = \mu(x^*, x_1)$ , and thus  $\mu$  places half of its mass on speeds less than or equal to  $x^*$ . The statement of Theorem 58 is not so enlightening on its own. It is best explained via the following two corollaries and example. This gives a straightforward application to discrete speed distributions. Essentially we can specify half of the speeds between  $x_n$  and  $x^*$ , and the remaining (faster) speeds are determined by (6).

**Corollary 59.** *Suppose  $n = 2m$  and  $0 < x_n < x_1$  are fixed. Choose any  $x_{n-1}, x_{n-2}, x_{m+1} \in (x_n, x^*)$  with  $x_{i+1} < x_i$ . For  $0 \leq i < m$  we define the remaining speeds*

$$x_{m-i} = T(x_{m+i+1}).$$

*The measure  $\mu_n(x) = \frac{1}{n} \sum_{i=1}^n \delta_{x_i}$  satisfies (6). Thus, a bullet process with speeds  $\{x_{2m}, \dots, x_1\}$  and uniform weights has any initial bullet with speed slower than  $x_{m+1}$  perishing almost surely.*

By increasing the number of speeds in the discrete case we obtain a class of absolutely continuous (with respect to Lebesgue measure) measures for which the slower half of bullets perish.

**Corollary 60.** *Let  $(\mu_{2m})_{m=1}^\infty$  be a collection of measures as in Corollary 59. Suppose that  $\mu_{2m}([0, x]) \rightarrow \mu([0, x])$  for all  $x \in S$ , with  $\mu$  some probability measure on  $S$ . The limiting measure  $\mu$  satisfies (6), and thus the  $(S, \mu)$ -bullet process has any bullet slower than  $x^*$  perishing almost surely.*

In the following example on  $[1, 2]$  we construct  $\mu_{2m}$  so that the slow bullet speeds are uniformly distributed. This choice completely determines the distribution on faster speeds.

**Example 61.** Let  $\mu_{2m}$  be the measure on  $[1, 2]$  induced by the speeds  $x_{2m} = 1$ ,  $x_1 = 2$ , and  $x_i = 1 + i/(3m)$  for  $m < i < 2m$  as in Corollary 59.



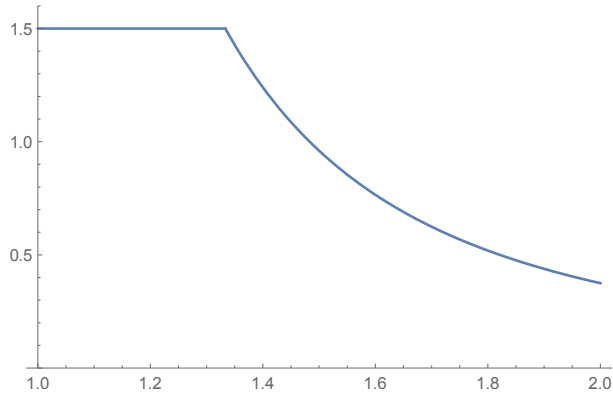


Figure 5: The density function  $f$  from Example 61. A bullet problem with speeds sampled according to  $f$  has bullets slower than  $4/3$  perishing almost surely.

We have  $\mu_{2m}([0, x]) \rightarrow \mu([0, x])$  for all  $x \in [1, 2]$  where  $\mu(A) = \int_A f(x)dx$  with

$$f(x) = \begin{cases} \frac{3}{2}, & 1 \leq x \leq 4/3 \\ \frac{6}{(2-3x)^2}, & 4/3 < x \leq 2 \end{cases}.$$

By Corollary 60 the  $([1, 2], \mu)$ -bullet process has an initial bullet slower than  $4/3$  perishing almost surely. (See Figure 5.)

Besides resembling the classical bullet problem, there is nothing particularly special about  $[1, 2]$  or the appearance of the uniform measure on  $[1, 4/3]$  in the above example. One could interpolate to a large variety of distributions on  $[x_n, x^*]$ , and (6) would dictate the distribution on  $[x^*, x_1]$ .

## 2.6 Exercises

1. Statements about survival in the  $\{1, 2, 3\}$ -speed bullet process are really about a probability measure on  $\Omega = \{1, 2, 3\}^{\mathbb{Z}^+}$ . Go through the formalities of obtaining the probability space  $(\Omega, \mathcal{F}, \mathbf{P})$  so that  $\mathbf{P}[\tau = \infty]$  is a well-defined quantity.

[I went through this rigamarole when I was trying to help undergrads come to grips with “infinite events” and found it instructive. The first couple pages of:

<http://www.math.uchicago.edu/~lawler/probnotes.pdf>

were helpful for me.]

2. Go through the formalities of applying the [Birkhoff-ergodic theorem](#) to deduce that if an arrow survives with positive probability, then there is a dense set of surviving arrows with that speed.
3. Confirm that  $\lim_{x \rightarrow 1^-} \mathbf{E}x^Y < 1$  implies that  $\mathbf{P}[Y = \infty] > 0$ .
4. Show that a bullet problem with speeds  $\{1, 2, 3\}$  is also equivalent to a bullet problem with an infinite speed bullet (such bullets instantly annihilate whatever is in front of them when fired).

## 2.7 Possibly doable open problems

1. How well can you estimate the  $\epsilon$  in Proposition 45? This would be interesting because it would let us extend to measures that put less mass on speed-2 bullets.
2. Show that a third fastest bullet survives. Or convince yourself why the proof for a second fastest does not immediately transfer to this case.

### 3 The frog model

The frog model starts with a single particle awake at the root of a graph and sleeping particles at the other vertices. The initial configuration of sleeping particles can be deterministic or random. Particles that are awake perform independent simple random walks in discrete time. When a vertex with sleeping particles is first visited, all of the particles at the site wake up and each begins its own walk. The name “frog model” was coined in 1996 by Rick Durrett; we continue the zoomorphism and refer to the particles as frogs. As with other interacting particle systems, the frog model is often motivated as a model for the spread of a rumor or infection (see [AMP02a], for instance). It and its variants have also found interest as models of combustion [RS04, CQR09, RS12], generally with particles moving in continuous time.

#### 3.1 Related models

There is *excited random walk* (also known as *cookie random walk*). The most basic version is on  $\mathbb{Z}^d$  where the first visit to each site has a  $p > 1/2$  drift in the  $\vec{e}_1$  direction. Subsequent visits are unbiased. See [KZ12] for a survey and more general versions of this.

**Theorem 62** (Benjamini, Wilson 03). *Excited random walk on  $\mathbb{Z}^d$  is transient for all  $d > 1$ .*

The closest relative to the frog model is known as *activated random walk*. In this typically  $\text{Poi}(\mu)$  particles are placed at each site of a graph. They simultaneously perform continuous time, rate-1 independent random walks, and fall asleep according to a rate- $\lambda$  Poisson process. Sleeping particles are re-activated if an active particle visits their vertex.

The system *locally fixates* if any fixed neighborhood eventually only contains sleeping particles

**Theorem 63** (Sidoravicius, Kesten, Shellef, 2009,2010). *For all  $\lambda > 0$  there is  $\mu_c \in [0, 1]$  such that the system almost surely locally fixates for  $\mu < \mu_c$ , and does not for  $\mu > \mu_c$ .*

**Conjecture 64.** *For any dimension and  $\lambda > 0$*

$$0 < \mu_c(\lambda) < 1.$$

**Theorem 65** (Sidorivius, 2015). *For  $d = 1$  we have  $\frac{\lambda}{1+\lambda} < \mu_c < 1$ .*

**Theorem 66** (Stauffer, Taggi, 2015). *For any dimension  $d$  we have  $\mu_c < 1$  for sufficiently small  $\lambda$ .*

### 3.2 Transience and recurrence for the frog model

We call a realization of the frog model *recurrent* if the root is visited infinitely often by frogs and *transient* if not. Even if each individual frog is transient, the aggregate of visits to the root can still be infinite. Thus, it is an interesting to ask about transience recurrence behavior on graphs for which a single random walk is transient.

**Theorem 67.** [KZ16, Theorem 1] *The frog model satisfies a 0-1 law for recurrence and transience on any vertex transitive graph.*

**Theorem 68.** [TW99, Theorem 1] *The frog model with one frog per site is recurrent on  $\mathbb{Z}^d$  for all  $d \geq 1$ .*

In fact a much stronger statement is known.

**Theorem 69.** [AMP02a, AMPR01, RS04] *Let  $\xi_n$  be the set of squares containing an activated site for  $FM(\mathbb{Z}^d, \eta)$ . For any dimension  $d \geq 1$  there is a nonempty convex set  $\mathcal{A} \subseteq \mathbb{R}^d$ , which depends on the distribution of  $\eta$ , such that for any  $0 < \epsilon < 1$ ,*

$$(1 - \epsilon)\mathcal{A} \subseteq \lim_{n \rightarrow \infty} \frac{\xi_n}{n} \subseteq (1 + \epsilon)\mathcal{A}.$$

The proof uses the subadditive ergodic theorem. The paper [RS04] proves something stronger with one frog per site, and frogs doing continuous time random walk.

**Theorem 70.** [RS04] *For any fixed ball,  $\mathbb{B}(0, r) \subset \mathbb{Z}^d$  the number of frogs at each site converges to a product of independent  $\text{Poi}(1)$  random variables.*

This implies a property known as *strong recurrence*, that the probability a frog is at the origin as time goes to infinity stays bounded away from 0.

Lastly, when the intensity of frogs decays a phase transition occurs.

**Theorem 71.** [Pop01, Theorem 1.1] *establishes the existence of a critical parameter  $0 < \alpha_c(d) < \infty$  for the frog model with simple random walk paths on  $\mathbb{Z}^d$  and initial configuration given by  $\eta(x) \sim \text{Ber}(p_x)$  such that*

- (i) *if  $p_x \leq \alpha/\|x\|^2$  for  $\alpha < \alpha_c(d)$  and all sufficiently large  $x$ , then the model is transient with positive probability;*
- (ii) *if  $p_x \geq \alpha/\|x\|^2$  for  $\alpha > \alpha_c(d)$  and all sufficiently large  $x$ , then the model is transient with probability zero.*

### 3.3 With drift on $\mathbb{Z}$

One interesting transient graph to study the frog model on is  $\mathbb{Z}$  with a drift,  $p > 1/2$ , to the right.

We have seen previously that a single random walk is transient. But how transient? Here is a lemma with a bit more clarity.

**Lemma 72.** *Let  $\rho = \frac{1-p}{p}$ . The probability a  $p > 1/2$  biased random walk started at  $k$  visits  $n$  before 0 is  $p_{k,n} = \frac{1-\rho^k}{1-\rho^n}$ .*

*Proof.* Let  $S_t$  be the walk's position at time  $t$ . We claim that  $\rho^{S_t}$  is a martingale. Indeed,

$$\begin{aligned} \mathbf{E}[\rho^{S_{t+1}} \mid S_t] &= (1-p)\rho^{S_t-1} + p\rho^{S_t+1} \\ &= \rho^{S_t} \left( \frac{1-p}{\rho} + p\rho \right) \\ &= \rho^{S_t} (p + 1-p) \\ &= \rho^{S_t}. \end{aligned}$$

Consider the stopping times  $\tau_0$  and  $\tau_n$ , the times that  $S_t$  reaches 0 and  $n$ . Set  $\tau_{0,n} = \min\{\tau_0, \tau_n\}$ . We can apply the optional stopping theorem to write

$$\begin{aligned} k &= \mathbf{E}\rho^{S_0} = \mathbf{E}\rho^{S_{\tau_{0,n}}} \\ &= \rho^0 \mathbf{P}[\tau_0 < \tau_n] + \rho^n \mathbf{P}[\tau_n < \tau_0] \\ &= (1 - p_{k,n}) + \rho^n p_{k,n}. \end{aligned}$$

Solving for  $p_{k,n}$  we obtain the claimed formula. □

**Theorem 73.** [GS09, Theorem 2.1] *Let  $\rho = \frac{1-p}{p}$ . The frog model with initial configuration  $\eta_j$  frogs at  $j$  is recurrent if and only if*

$$\sum_{j=1}^{\infty} \eta_j \rho^j = \infty.$$

*Proof.* Recall that  $0 < \prod a_j$  if and only if  $\sum \log a_j < \infty$  and that  $\log(1-x) \approx x$  for small  $x$ .

- (i) We show that the condition  $\sum_{j=1}^{\infty} \eta_j \rho^j = \infty$  is sufficient for recurrence. Consider the probability that the site  $-x$  with  $x \in \mathbb{Z}^+$ , is visited by

a frog starting on the right of the origin. It suffices to show that for each  $x \in \mathbb{Z}^+$  this probability is 1. Fix  $x$ . Then

$$\begin{aligned} & \mathbf{P}[\text{a frog right of the origin visits } -x] \\ &= 1 - \prod_{j=1}^{\infty} \mathbf{P}[\text{none of the frogs starting in } j \text{ visits } -x] \\ &= 1 - \prod_{j=1}^{\infty} (1 - \rho^{x+j})^{\eta_j}, \end{aligned}$$

and the infinite product equals 0, since  $\sum \eta_j \log(1 - \rho^j) \approx \sum_{j=1}^{\infty} \eta_j \rho^j = \infty$ .

- (ii) Assume that  $\sum_{j=1}^{\infty} \eta_j \rho^j < \infty$ . We show that  $\mathbf{P}$ [ the origin is visited infinitely often] is less than 1. It suffices to show that the probability  $-1$  is visited eventually is strictly less than 1.

$$\mathbf{P}[-1 \text{ is never visited}] = (1 - \rho) \prod_{j=1}^{\infty} (1 - \rho^{j+1})^{\eta_j}.$$

This is greater than 0 because

$$\sum_{j=1}^{\infty} \eta_j \log(1 - \rho^{j+1}) \approx \sum \eta_j \rho^j < \infty.$$

□

An interesting result is the main theorem from [GS09]. Note there is no dependence on  $p$ .

**Theorem 74.** [GS09, Theorem 2.2] *For all  $p > 1/2$  and  $\eta_x \sim \eta$  i.i.d. frogs per site, we have*

$$FM(\mathbb{Z}, p, \eta) = \begin{cases} \text{recurrent,} & \mathbf{E} \log \eta = \infty \\ \text{transient,} & \text{otherwise} \end{cases}.$$

*These results are for a.s. for a.e. environment. Notice there is no  $p$ -dependence.*

We give a simple proof of this from [KZ16].

*Proof.* For  $x \in \mathbb{Z}$  denote by  $R_x$  the event that waking up the frogs at  $x$  with all frogs at  $y \neq x$  still asleep, that infinitely many frogs visit  $x$ . Since  $(\eta_x)_{x \in \mathbb{Z}}$  is i.i.d. and the frogs move independently of each other, the sequence  $(\mathbf{R}_x)_{x \in \mathbb{Z}}$  is stationary (distribution doesn't change when shifting) and ergodic with respect to the shift operator. The Birkhoff ergodic theorem ensures that this sequence is either a.s. identically equal to 0, in which case  $\mathbf{P}[R_x] = 0$  for all  $x \in \mathbb{Z}$ , or a.s.  $\mathbf{1}\{R_x\} = 1$ . for infinitely many  $x \geq 1$ . In the latter case, if we wake up the frogs at a site  $v \in \mathbb{Z}$  with  $\eta_v \geq 1$ , then these frogs will be a.s. transient to the right and will therefore also visit a site  $x$  for which  $R_x$  occurs and wake up the frogs at  $x$ . This will trigger infinitely many frogs to visit  $x$ , and therefore also  $v$  so that  $R_v$  occurs as well.

It is an exercise to show that  $\mathbf{P}[\sum \eta_x \rho^x = \infty] = 1$  if and only if  $\mathbf{E} \log \eta = \infty$ .  $\square$

We also remark a result by [DP14] generalizes this a bit to  $\mathbb{Z}^d$ .

**Theorem 75.** [DP14, Theorem 2.1] For  $FM(\mathbb{Z}^d, p, \eta)$  we have  $\mathbf{E} \log^{\frac{d+1}{2}} \eta = \infty$  implies a.s. recurrent.

Note that they conjecture that  $p$  does matter here for lighter tailed distributions  $\eta$ . [KZ15] provides an example where the expectation is finite, but the frog model is transient. So there is a different regime. This would be nice to make progress on.

### 3.4 The frog model on trees

Trees are similar to a frog model with drift, because the tree-structure induces a natural drift away from the root. What is tricky (and also tricky in  $\mathbb{Z}^d$ ) is that we are not guaranteed to wake every site.

Let  $FM(\mathbb{T}_d, \eta)$  be a frog model with i.i.d.  $\eta$  frogs per vertex.

#### 3.4.1 Theorems to be presented

A variant known as the *frog model with death* has been studied on trees. In this frogs take  $\text{Geo}(1-p)$  steps after being woken up, then are removed from the system. Lebensztayn, Machado, and Popov have studied the event that a frog is always awake on  $\mathbb{T}_d^{\text{hom}}$ .

**Theorem 76.** Let  $p_c = \inf\{p: \mathbf{P}[FM(\mathbb{T}_d^{\text{hom}}, 1, p) \text{ survives}] > 0\}$  be the smallest  $p$  for which the frog model with death on  $\mathbb{T}_d^{\text{hom}}$  survives wpp. It holds that

$$\frac{d+1}{2d+1} \leq p_c \leq \frac{d+1}{2d}.$$

We'll possibly present the proof of this later, along with a covertime result on finite trees.

Our primary goal, however, is to present the following theorems:

**Theorem 77** (Hoffman, Johnson, J., 2014).  $FM(\mathbb{T}_d, 1)$  is recurrent for  $d = 2$  and transient for  $d \geq 5$ .

**Conjecture 78.**  $FM(\mathbb{T}_3, 1)$  is recurrent, while  $FM(\mathbb{T}_4, 1)$  is transient. Moreover, we believe that  $FM(\mathbb{T}_2, 1)$  is strongly recurrent, while possibly  $FM(\mathbb{T}_3, 1)$  is weakly recurrent (the probability the root is occupied goes to 0 with time).

**Theorem 79** (Hoffman, Johnson, J., 2015). There exists  $\mu_c = \mu_c(d)$  such that for  $\mu < \mu_c$  we have  $FM(\mathbb{T}_d, \text{Poi}(\mu))$  is transient, while for  $\mu > \mu_c$  it is recurrent.

**Theorem 80** (Johnson, J. 2016).  $\mu_c = \Theta(d)$ .

A thorn in our side for years was not being able to describe deterministic configurations. Ideally, we would like to load up each site with a lot of frogs and deduce recurrence. We did this in a roundabout way via an alternate stochastic ordering known as the *increasing concave order*.

Say that  $X \preceq_{\text{icv}} Y$  if  $\mathbf{E}\varphi(X) \leq \mathbf{E}\varphi(Y)$  for all increasing, concave functions  $\varphi$ . Notice that Jensen's inequality gives  $X \equiv \beta$  with  $\beta \in \mathbb{R}$  fixed is the maximal mean  $\beta$  random variable in this ordering. It is worth noting that if  $X \preceq_{\text{icv}} Y$ , then  $\mathbf{P}[X = \infty] \leq \mathbf{P}[Y = \infty]$ . The function  $t^x$  is increasing and concave. Note that as  $t \nearrow 1$ , we have  $t^x \rightarrow \mathbf{1}\{x < \infty\}$ . Thus, by the monotone convergence theorem,

$$\mathbf{E}t^X \rightarrow \mathbf{P}[X < \infty] \quad \text{and} \quad \mathbf{E}t^Y \rightarrow \mathbf{P}[Y < \infty]$$

as  $t \nearrow 1$ . Now  $\mathbf{E}t^X \geq \mathbf{E}t^Y$  for  $t \in (0, 1)$  implies that  $\mathbf{P}[X < \infty] \geq \mathbf{P}[Y < \infty]$ .

So, finally we can say something about deterministic configurations.

**Theorem 81** (Johnson, J. 2016). Suppose  $\eta \preceq_{\text{icv}} \eta'$ . Then the visits to the root in  $FM(\mathbb{T}_d, \eta)$  is icv smaller than those in  $FM(\mathbb{T}_d, \eta')$ .

**Corollary 82.**  $FM(\mathbb{T}_d, [\mu_c(d)])$  is recurrent.

In a work in progress we (believe!) we can prove that for large enough  $\mu$  the frog model is strongly recurrent.

**Theorem 83.** There is  $\lambda_c(d)$  large enough such that  $FM(\mathbb{T}_d, \text{Poi}(\lambda))$  is strongly recurrent for  $\lambda > \lambda_c(d)$ .

**Conjecture 84.**  $\lambda_c > \mu_c$  for all  $d \geq 2$



### 3.5 Transience

To prove transience we instead seek processes that dominate the frog model and can be proven transient. For example, consider a branching random walk on  $\mathbb{T}_d$  whose particles split in two at every step.

**Proposition 85.** *FM( $\mathbb{T}_d, 1$ ) is transient for  $d \geq 14$ .*

*Proof.* We will make use of independence of the BRW and a union bound.

$$\begin{aligned} \mathbf{P}[\text{frog at root at time } 2t] &= \mathbf{P}\left[\bigcup_1^{2^{2t}} \mathbf{P}[\mathbb{Z}^+ \cup \{0\} \text{ RW at } 0 \text{ at time } t]\right] \\ &\leq 2^{2t} \#\{\text{length } 2t \text{ paths to } 0\} \left(\frac{d}{d+1}\right)^t \left(\frac{1}{d+1}\right)^t. \end{aligned}$$

Such paths are called Dyck paths, and are enumerated by the Catalan numbers,  $C_{2t} = \frac{1}{t+1} \binom{2t}{t} \sim 4^t / (t^{3/2} \sqrt{\pi})$ . Thus, the above quantity is asymptotic to  $(16d/(d+1)^2)^t / t^{3/2}$ . We can solve that  $16d/(d+1)^2 = 1$  at  $d = 7 + 4\sqrt{3} \approx 13.9$ . For  $d \geq 14$  these probabilities are summable, so, the result follows from the Borel-Cantelli lemma. □

In this section, we will present a series of refinements to this argument to ultimately prove Theorem 77. In Proposition 86, we use a branching random walk on the integers and martingale techniques to prove transience for  $d \geq 6$ . We use this argument as a base for our proofs of Proposition 87, transience on the deterministic tree which alternates between five and six children, and Theorem 77, transience for  $d \geq 5$ . Both proofs use a multitype branching random walk. We included Proposition 87 because its calculations can easily be done by hand. In Theorem 77, on the other hand, we use a branching random walk with 27 types. The necessary calculations are intractable by hand, but they take only a few seconds on a computer. To get started we first address some difficulties that arise from reflection at the root.

Consider the branching random walk where each particle gives birth at each step either to one child to its left or to two children to its right. Formally, we define this as a sequence of point processes. Start with  $\xi_0$  as a single particle at 0. With probability  $1/(d+1)$ , the point process  $\xi_1$  consists of a single particle at  $-1$ ; with probability  $d/(d+1)$ , it consists of two particles at 1. After this, each particle in  $\xi_n$  produces children in  $\xi_{n+1}$  in the same way relative to its position, independently of all other particles. We will use this branching random walk to prove the frog model transient for  $d \geq 6$  and closely related processes to extend this down to  $d = 5$ .

**Proposition 86.** *For  $d \geq 6$ , the frog model on  $\mathbb{T}_d$  is almost surely transient.*

*Proof.* Consider the frog model on  $\mathbb{T}_d^{\text{hom}}$ , starting with no sleeping frogs at direct ancestors of the root. When a frog jumps backward in this process, it never spawns a new frog, and when it moves forward, it sometimes does. Thus, the projection of this frog model onto the integers can be coupled with  $(\xi_n, n \geq 0)$  so that every frog has a corresponding particle. Proving that  $\xi_n$  visits 0 finitely many times a.s. proves that the frog model on  $\mathbb{T}_d$  is transient a.s.

To determine the behavior of  $\xi_n$ , we define a weight function  $w$  on point process configurations. We refer to the position of a particle  $i$  in a point process configuration by  $P(i)$  and define

$$w(\xi) = \sum_{i \in \xi} e^{-\theta P(i)}, \quad (9)$$

with  $\theta$  to be chosen later. Letting  $\mu = \mathbf{E}w(\xi_1)$  we have

$$\mathbf{E}[w(\xi_{n+1}) \mid \xi_n] = \mu w(\xi_n),$$

and so the sequence  $w(\xi_n)/\mu^n$  is a martingale. As it is positive, it converges almost surely. When  $\mu < 1$  this means  $w(\xi_n) \rightarrow 0$ . If a particle in  $\xi_n$  occupies the origin then  $w(\xi_n) \geq 1$ , and so infinitely many visits to the origin prevents  $w(\xi_n)$  from converging. Hence,  $\mu < 1$  implies the a.s. transience of  $\xi_n$ . (In fact, this holds when  $\mu = 1$  as well, though we will not need this.) It then suffices to show that there exists  $\theta$  making  $\mu < 1$ . We compute

$$\mu = \frac{1}{d+1}e^\theta + 2\frac{d}{d+1}e^{-\theta}.$$

This is minimized by setting  $\theta = \log(2d)/2$ , which makes  $\mu = 2\sqrt{2d}/(d+1)$ . A bit of algebra shows that  $\mu < 1$  when  $d > 3 + 2\sqrt{2} \approx 5.83$ .  $\square$

By using a multitype branching process, we can extend this proof to show transience for  $\mathbb{T}_5$ . Before we do so, we will show how it works in a setting where humans can do the math without much assistance.

**Proposition 87.** *Let  $\mathbb{T}_{5,6}$  be the tree whose levels alternate between vertices with 5 children and vertices with 6 children, starting with the root having either 5 or 6 children. The frog model on this tree is transient a.s.*

*Proof.* Let  $\mathbb{T}_{5,6}^{\text{hom}}$  be the five-six children alternating homogeneous tree which contains  $\mathbb{T}_{5,6}$  and place a sleeping frog at each vertex except for direct

ancestors of the root of  $\mathbb{T}_{5,6}$ . It suffices to prove transience of this frog model on  $\mathbb{T}_{5,6}^{\text{hom}}$ .

First note that a frog at a vertex with five children has different probabilities of moving forwards or backwards than a frog at a vertex with six children. By design the tree deterministically alternates, so a frog also alternates between each state.

When a frog moves backwards there is chance it immediately jumps forward to the same vertex, which will never spawn a new frog. Similarly, when two frogs occupy the same site there is a chance both jump forward to the same vertex, spawning at most one frog, not two. The idea is to introduce additional particle types that act like frogs in these more advantageous states.

Consider a multitype branching random walk on  $\mathbb{Z}$  with six particle types,  $F_5$ ,  $D_5$ ,  $B_5$ ,  $F_6$ ,  $D_6$ , and  $B_6$ . The subscript accounts for whether a frog is at a vertex with 5 or 6 children.  $B$  particles represent frogs that have just stepped backward.  $D$  particles represent two frogs at once, the waker and wakee at a vertex where a frog has just woken up. Last,  $F$  particles represent single frogs with sleeping frogs present at all children. A visual depiction of these particle types is provided in Figure 6, and the distribution of children for each particle type is defined in Figure 7.

Let  $\zeta_n$  be the branching random walk in which particles reproduce independently with the given child distributions. These distributions are chosen to match how the projections of frogs on the integers behave. Ignoring for a moment whether a frog is at a site with five or six children, when a frog jumps back it becomes of type  $B$  and when a new frog wakes it and its waker consolidate into a type  $D$  particle. Any extra frogs become a type  $F$  particle. These particles then reproduce independently on a “fresh” tree configured so that the particles always generate at least as many frogs as the projection of the actual frog model. For this reason we can couple the integer projection of the frog model on  $\mathbb{T}_{5,6}^{\text{hom}}$  with  $\zeta_n$  so that the particles representing awake frogs are a subset of  $\zeta_n$ . It therefore suffices to prove that  $\zeta_n$  is transient.

To analyze  $\zeta_n$ , we use a generalization of the martingale from Proposition 86 to the multitype setting, introduced in [Big76]. Let  $\zeta_n = \sum_i \zeta_n^i$ , where  $i$  ranges over the six particle types and  $\zeta_n^i$  denotes the restriction of  $\zeta_n$  to particles of type  $i$ . Recalling the weight function  $w$  given by (9), we define a matrix  $\Phi(\theta)$  by

$$\Phi_{ij}(\theta) = \mathbf{E}_i \left[ w(\zeta_1^j) \right].$$

Here, we use  $\mathbf{E}_i$  to denote expectation when  $\zeta_0$  is a single particle at the

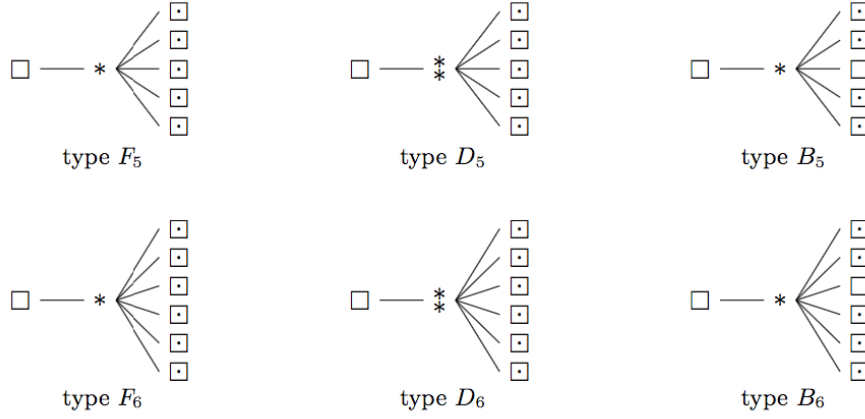


Figure 6: A depiction of the six particle types from the proof of Proposition 87. Each asterisk is a frog represented by the particle. The symbol  $\square$  signifies a vertex with a sleeping frog, and the symbol  $\square$  represents a vertex with no sleeping frog.

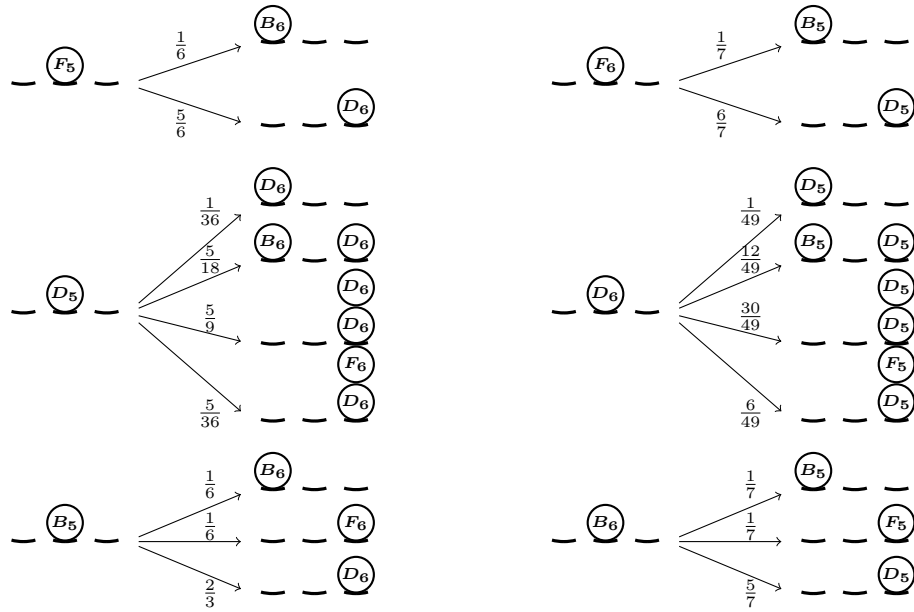


Figure 7: The distribution of children for each particle type in the proof of Proposition 87.

origin of type  $i$ . Let  $w_n$  denote a row vector whose  $i$ th entry is  $w(\zeta_n^i)$ . Then

$$\mathbf{E}[w_{n+1} \mid \zeta_n] = w_n \Phi(\theta). \quad (10)$$

Thus, for any eigenvalue  $\lambda$  and associated right eigenvector  $v$  of  $\Phi(\theta)$ ,

$$\mathbf{E}[w_{n+1}v \mid \zeta_n] = w_n \Phi(\theta)v = \lambda w_n v,$$

and so  $w_n v / \lambda^n$  is a martingale.

Since  $\Phi(\theta)$  is a nonnegative irreducible matrix, there is a positive eigenvalue  $\phi(\theta)$  equal to the spectral radius of  $\Phi(\theta)$  by the Perron–Frobenius theorem. The eigenvector  $v(\theta)$  associated with  $\phi(\theta)$  has strictly positive entries. We then have a positive martingale  $w_n v(\theta) / \phi(\theta)^n$ . If  $\phi(\theta) < 1$ , then it follows as in Proposition 86 that the branching random walk visits 0 finitely often, thus proving that the frog model is almost surely transient.

All that remains is to find some value of  $\theta$  such that  $\phi(\theta) < 1$ . Ordering the rows and columns  $F_5, D_5, B_5, F_6, D_6, B_6$  and reading off  $\mathbf{E}_i[w(\zeta_i^j)]$  from Figure 7,

$$\Phi(\theta) = \begin{bmatrix} 0 & 0 & 0 & 0 & \frac{5}{6}e^{-\theta} & \frac{1}{6}e^{\theta} \\ 0 & 0 & 0 & \frac{5}{36}e^{-\theta} & \frac{1}{36}e^{\theta} + \frac{55}{36}e^{-\theta} & \frac{5}{18}e^{\theta} \\ 0 & 0 & 0 & \frac{1}{6}e^{-\theta} & \frac{2}{3}e^{-\theta} & \frac{1}{6}e^{\theta} \\ 0 & \frac{6}{7}e^{-\theta} & \frac{1}{7}e^{\theta} & 0 & 0 & 0 \\ \frac{6}{49}e^{-\theta} & \frac{1}{49}e^{\theta} + \frac{78}{49}e^{-\theta} & \frac{12}{49}e^{\theta} & 0 & 0 & 0 \\ \frac{1}{7}e^{-\theta} & \frac{5}{7}e^{-\theta} & \frac{1}{7}e^{\theta} & 0 & 0 & 0 \end{bmatrix}.$$

Computing the eigenvalues of this matrix numerically, one can confirm that there exists  $\theta$  with  $\phi(\theta) < 1$ ; for example,  $\phi(\log 3) \approx 0.9937$ . To be completely certain that this is not an artifact of rounding, we will justify that  $\phi(\log 3) < 1$  without using floating-point arithmetic. Observe that  $\Phi(\log 3)$  has rational entries. Using the computer algebra system SAGE, we calculated  $(\Phi(\log 3))^{66}$  using exact arithmetic, and we found that its largest row sum was less than 1. (The only significance of the 66th power is that it is the lowest one for which this is true.) This implies that all eigenvalues of  $(\Phi(\log 3))^{66}$  are less than 1, which implies that all eigenvalues of  $\Phi(\log 3)$  are less than 1 as well. The source code accompanying this paper includes this matrix and has instructions so that readers can easily check these claims.  $\square$

Having proven transience for the frog model on  $\mathbb{T}_d$  with  $d \geq 6$  and on  $\mathbb{T}_{5,6}$ , we present our final refinement to prove the  $\mathbb{T}_5$  case. The proof is essentially the same as the previous one, but with more particle types and a more difficult calculation.

*Proof of transience part of Theorem 77.* We define a particle type  $P(a, b, c)$ , for  $a \geq 1$  and  $b, c \geq 0$ . A particle of type  $P(a, b, c)$  represents  $a$  frogs on one vertex. There are no sleeping frogs on at least  $b$  of the vertex’s children and on at least  $c$  of the vertex’s siblings. In this scheme, the  $F$  types from the previous proof would translate to  $P(1, 0, 0)$ , the  $D$  types would translate to  $P(2, 0, 0)$ , and the  $B$  types would translate to  $P(1, 1, 0)$ .

We use 27 of these particles,  $P(a, b, c)$  with  $1 \leq a \leq 3$  and  $0 \leq b, c \leq 2$ . For particle type  $P(a, b, c)$ , consider the frog model on the homogeneous tree, starting with  $a$  frogs at the root. As usual, remove the sleeping frogs from direct ancestors of the root. Also remove the sleeping frogs from  $b$  children of the root and from  $c$  siblings. From each of these 27 initial states, we compute all possible states to which the frog model could transition in two steps, along with the exact probabilities of doing so. We then represent each of these final states as a collection of particles of the 27 types, at levels  $-2$ ,  $0$ , and  $2$  on the tree. In this way, we determine child distributions for each particle type, as in Figure 7. There is a slight ambiguity in how to do this, as a state of frogs can be represented in more than one way by these particle types. For example, four frogs on one vertex with one sibling vertex with no sleeping frog could be represented as two particles of type  $P(2, 0, 1)$ , or as one of type  $P(3, 0, 1)$  and one of type  $P(1, 0, 1)$ . We always chose particles greedily, opting for as many 3-frog particles as possible. Whatever choice we make here, our branching random walk will still dominate the frog model, since when we assign new particles we “reset” the tree below them so the particles wake at least as many frogs as their counterpart in the frog model.

As in Proposition 87, it suffices to compute the matrix  $\Phi(\theta)$  and show that for some choice of  $\theta$ , its highest eigenvalue is less than one. Our attached source code computes  $\Phi(\theta)$  exactly. We include additional documentation there explaining how we performed this calculation and describing the steps we took to make sure it was trustworthy. To avoid rounding issues, we proceeded as with Proposition 87. We exactly computed  $(\Phi(\log 3))^{1024}$  by succesively squaring the matrix ten times, and we then checked that all of its row sums were less than 1. (There is no significance to the value  $\log 3$ ; it just happens to work.) Thus, all eigenvalues of  $(\Phi(\log 3))^{1024}$  are less than 1, implying that all eigenvalues of  $\Phi(\log 3)$  are less than 1 as well.  $\square$

### 3.6 Poisson tree frogs

We show in [HJJ16a] that the frog model on a  $d$ -ary tree with i.i.d.- $\text{Poi}(\mu)$  frogs per vertex undergoes a phase transition between transience and recurrence as  $\mu$  grows. In more detail, for each  $d \geq 2$  there exists a critical value

$\mu_c(d)$  such that the model is almost surely transient for  $\mu < \mu_c(d)$  and is almost surely recurrent for  $\mu > \mu_c(d)$ . The proof in [HJJ16a] shows that for some constants  $C, C' > 0$ , we have  $Cd \leq \mu_c(d) \leq C'd \log d$ . We sharpen this result by removing the  $\log d$  factor from the upper bound, determining the order of  $\mu_c(d)$  up to constant factors.

**Theorem 88.** *For all sufficiently large  $d$ , it holds that  $.24d \leq \mu_c(d) \leq 2.28d$ .*

Theorem 88 can be extended to initial distributions other than Poisson using the results of [JJ16b]. For example, the frog model on a  $d$ -ary tree with deterministically many frogs per site has a critical threshold of the same order; see [JJ16b, Corollary 4] for details.

An accurate description of the transition threshold on trees is especially relevant given that the frog model on  $\mathbb{Z}^d$  behaves rather differently. A transition still occurs, but it does so at a decaying density of frogs (see [Pop01, Theorem 1.1]). A natural next step would be to investigate this problem for irregular trees. Perhaps the phase transition is on the same order as the branching number of the tree? This is completely speculative, as the mere existence of a phase transition for the frog model is unknown even on a Galton–Watson tree.

Our other goal is to present as simply as possible the argument for existence of a recurrence phase on trees. While the transience phase is fairly easy to establish (see [HJJ16a, Proposition 15]), the recurrence phase is more difficult. Indeed, the question of recurrence on  $\mathbb{T}_2$  with one sleeping frog per site was first posed in [AMP02b] and was only recently answered in the affirmative [HJJ16b]. It remains open to determine if the one-per-site frog model is recurrent on  $\mathbb{T}_3$  and  $\mathbb{T}_4$ . Taking advantage of some technical improvements we have made since [HJJ16b, HJJ16a], we give a streamlined proof of recurrence on  $\mathbb{T}_2$  with i.i.d.- $\text{Poi}(\mu)$  frogs per site (see Proposition 98). Using [JJ16b, Corollary 3], this result also implies recurrence on  $\mathbb{T}_2$  with two sleeping frogs per site. This is weaker than the result in [HJJ16b], but the proof is much simpler.

## Ideas of the proof

As in [HJJ16b, HJJ16a], our proof of recurrence is based on recursion and bootstrapping. To set this up, we first show that it is enough to establish recurrence for a frog model whose paths are stopped non-backtracking walks, which we call the *self-similar frog model*. Let  $V$  be the number of visits to the root in this process. A self-similarity yields a relation between  $V$  and

a collection of independent copies of  $V$ . Such relations are called recursive distributional equations (see [AB05] for further discussion).

In the bootstrap part of the argument, we assume that  $V$  is stochastically larger than  $\text{Poi}(\lambda)$  for some  $\lambda \geq 0$ . We then analyze the recursive distributional equation to show that  $V$  is in fact stochastically larger than  $\text{Poi}(\lambda + \epsilon)$ . Iterating this argument starting at  $\lambda = 0$ , we show that  $V$  is larger than  $\text{Poi}(\epsilon)$ , then larger than  $\text{Poi}(2\epsilon)$ , and so on, with the conclusion that  $V = \infty$  a.s. Here and in [HJJ16a], this argument uses the standard stochastic order, while in [HJJ16b] it uses a more exotic stochastic order.

The bootstrap phase of our argument is more elaborate than in [HJJ16a], where we could only establish  $\mu_c(d) = \Omega(d \log d)$ . We obtain a better upper bound in this paper because the version of the self-similar frog model here better approximates the actual frog model. Put more simply, we are able to capture the contributions of more frogs. The simplification that allows us to handle the extra complexity is our use of [MSH03, Theorem 3.1(b)], a simple criterion for determining when a Poisson distribution stochastically dominates a Poisson mixture. This allows us to avoid the more difficult coupling argument used in [HJJ16a].

### 3.6.1 A criterion for stochastic dominance

Given two probability measures  $\nu$  and  $\nu'$  on the extended real numbers, we say that  $\nu$  is *stochastically smaller* than  $\nu'$  if

$$\nu((t, \infty]) \leq \nu'((t, \infty])$$

for all  $t \in \mathbb{R}$ . We denote this relationship by  $\nu \preceq_{\text{st}} \nu'$ . If  $X \sim \nu$  and  $Y \sim \nu'$ , we also write  $X \preceq_{\text{st}} Y$ ,  $X \preceq_{\text{st}} \nu'$ , and  $\nu \preceq_{\text{st}} Y$  all to mean the same thing. An alternate characterization of stochastic dominance is that  $X \preceq_{\text{st}} Y$  if and only if there exists a coupling of  $X$  and  $Y$  such that  $X \leq Y$  a.s.

Theorem 90 provides a necessary and sufficient condition for a Poisson mixture to dominate a Poisson distribution, which we will need in Section 3.6.2. We reproduce the proof given in [MSH03] for our readers' convenience. See also [Yu09] for a more general result.

**Lemma 89** ([MSH03, Lemma 3.1(b)]). *For any positive integer  $n$ , the function*

$$h_n(x) = x \sum_{k=0}^n \frac{(-\log x)^k}{k!}$$

*is increasing and concave on  $(0, 1]$ .*



*Proof.* We compute

$$h'_n(x) = \sum_{k=0}^n \frac{(-\log x)^k}{k!} - \sum_{k=1}^n \frac{(-\log x)^{k-1}}{(k-1)!} = \frac{(-\log x)^n}{n!},$$

which is positive and decreasing on  $(0, 1]$ , showing that  $h_n(x)$  is increasing and concave.  $\square$

**Theorem 90** ([MSH03, Theorem 3.1(b)]). *Let  $X \sim \text{Poi}(\lambda)$ , and let  $Y \sim \text{Poi}(U)$  for some nonnegative random variable  $U$ . Then the following are equivalent:*

- (i)  $X \preceq_{st} Y$ ,
- (ii)  $\mathbf{P}[X = 0] \geq \mathbf{P}[Y = 0]$ , and
- (iii)  $\lambda \leq -\log \mathbf{E}e^{-U}$ .

*Proof.* Conditions (ii) and (iii) are just restatements of each other, since  $\mathbf{P}[X = 0] = e^{-\lambda}$  and  $\mathbf{P}[Y = 0] = \mathbf{E}e^{-U}$ . Condition (i) implies (ii) by the definition of stochastic dominance. It remains to prove that (iii) implies (i). It suffices to show that  $\mathbf{P}[Y \leq n] \leq \mathbf{P}[X \leq n]$  for all nonnegative integers  $n$ . We compute

$$\mathbf{P}[Y \leq n] = \sum_{k=0}^n \mathbf{E} \left[ \frac{e^{-U} U^k}{k!} \right] = \mathbf{E} \left[ \zeta \sum_{k=0}^n \frac{(-\log \zeta)^k}{k!} \right] = \mathbf{E} h_n(\zeta),$$

where  $\zeta = e^{-U}$ . Our assumption is that  $\mathbf{E}\zeta = \mathbf{E}e^{-U} \leq e^{-\lambda}$ . By Lemma 89, the function  $h_n(x)$  is increasing and concave on  $(0, 1]$ , where  $\zeta$  takes values. Thus

$$\mathbf{E} h_n(\zeta) \leq h_n(\mathbf{E}\zeta) \leq h_n(e^{-\lambda}) = \sum_{k=0}^n \frac{e^{-\lambda} \lambda^k}{k!} = \mathbf{P}[X \leq n],$$

where we use that  $h_n$  is concave to apply Jensen's inequality in the first step, and we use that  $h_n$  is increasing in the second step.  $\square$

### 3.6.2 Critical parameters for $d$ -ary trees

Our argument breaks down into two parts described in the introduction. In Section 3.6.2.1, we define the self-similar frog model and define  $V$  as the number of visits to the root in this process. Then, we determine a recursive

distributional equation satisfied by the law of  $V$  (see Lemma 95). The ideas in this section can be found in [HJJ16b, HJJ16a], but they take some work to extract in the form we need. Though we do our best to avoid duplicating material, when in doubt we have opted for comprehensibility over efficiency.

Next comes the bootstrap portion of the argument. In Section 3.6.3, we use the set-up of Section 3.6.2.1 to carry this out in case of the binary tree, giving a short proof of the existence of a recurrence phase. In Section 3.6.4 we give a more complex version of this argument proving recurrence on the  $d$ -ary tree for  $\mu = \Omega(d)$ .

### 3.6.2.1 The bootstrapping set-up

**The self-similar frog model** Many basic features of the frog model depend only on the range of each frog. This yields rather nice abelian and monotonicity properties. For example, the total number of visits to the root is unaffected by the order frogs wake up in and the rate they reveal vertices in their ranges. Also, trimming the range of frogs can only reduce the number of visits to the root. Applying this observation in combination with the coupling characterization of stochastic dominance, we note the following fact. Let  $r(\eta, S)$  be the number of visits to the root in the frog model  $(\eta, S)$ .

**Fact 91.** *Consider a collection of frog paths  $S = (S_\bullet(v, i))_{v \in G, i \geq 1}$  on a graph  $G$ . Suppose that another collection of paths  $\tilde{S}$  can be coupled with  $S$  such that for all  $i$  and  $v$ , the range of  $\tilde{S}_\bullet(v, i)$  is a subset of the range of  $S_\bullet(v, i)$ . Then  $r(\eta, \tilde{S}) \preceq_{st} r(\eta, S)$ .*

From now on, let  $S = (S_\bullet(v, i), v \in \mathbb{T}_d, i \geq 1)$  denote a collection of independent simple random walks with  $S_\bullet(v, i)$  started at  $v$ , and let the components of  $\eta = (\eta(v))_{v \in \mathbb{T}_d}$  be i.i.d.- $\text{Poi}(\mu)$ , independent of  $S$ . The first step in studying the frog model  $(\eta, S)$  will be to replace  $S$  by a collection of paths  $T$  to obtain  $(\eta, T)$ , which we call the *self-similar frog model* in reference to a useful property described in Fact 93.

We define  $T$  in two steps. First, let  $S' = (S'_\bullet(v, i), v \in \mathbb{T}_d, i \geq 1)$  denote a collection of independent random non-backtracking walks stopped at  $\emptyset$ . In more detail, call a random walk a *simple random non-backtracking walk* on an arbitrary graph if it chooses from its neighbors uniformly for its first step, and then in all subsequent steps it chooses uniformly from its current neighbors except the one it just arrived from. We define  $S'_\bullet(v, i)$  to be a simple non-backtracking random walk stopped on arrival at  $\emptyset$ . The walks  $S'_\bullet(v, i)$  and  $S_\bullet(v, i)$  can be coupled so that the range of the first is a subset

of the range of the second by making  $S'_\bullet(v, i)$  a stopped, loop-erased version of  $S_\bullet(v, i)$ . This is proved in detail in [HJJ16b, Proposition 7].

Now we construct  $T$  as a modification of  $S'$ . Each path  $T_\bullet(v, i)$  will be a stopped version of  $S'_\bullet(v, i)$ . Let  $v$  be a nonroot vertex in  $\mathbb{T}_d$  with parent  $u$ . Suppose that  $v$  is visited in the frog model  $(S', \eta)$  for the first time at time  $j$ , necessarily by one or more frogs moving from  $u$  to  $v$ . Select one of these visiting frogs arbitrarily, and stop all of the other ones. (Observe that it is irrelevant which frog is allowed to continue, so long as one views frogs as indistinguishable.) If any frogs move from  $u$  to  $v$  at subsequent times, stop them at  $v$  as well. Do this for all vertices  $v \in \mathbb{T}_d$ , and let  $T$  be the resulting collection of stopped walks. As the range of each  $T_\bullet(v, i)$  is a subset of the range of  $S'_\bullet(v, i)$ , the following fact (also noted in [HJJ16b, Proposition 7]) follows:

**Fact 92.** *There is a coupling of  $S$  and  $T$  so that the range of each  $T_\bullet(v, i)$  is a subset of the range of  $S_\bullet(v, i)$ .*

By Facts 91 and 92, we have  $r(\eta, T) \preceq_{\text{st}} r(\eta, S)$ . We will now work exclusively with the self-similar frog model,  $(\eta, T)$ , and prove recurrence for it with sufficiently large  $\mu$ . Unlike all other frog models considered in this paper, the frog paths  $T$  are not independent of each other nor of  $\eta$ , because one frog's motion in  $(\eta, T)$  can cause another frog to be stopped. This is the only form of dependence, however, and frogs that have not been stopped move independently of each other. So, it is not a serious obstacle.

Let  $V = r(\eta, T)$ . Next, we discuss a self-similarity property of the model and its consequences for  $V$ . For any vertex  $v \in \mathbb{T}_d$ , let  $\mathbb{T}_d(v)$  denote the subtree made up of  $v$  and its descendants. We call  $\mathbb{T}_d(v)$  *activated* in the self-similar frog model if  $v$  is ever visited. Let  $u$  be the parent of  $v$ . By our construction of  $T$ , if  $\mathbb{T}_d(v)$  is activated, then there is a unique frog that moves from  $u$  to  $v$ , entering  $\mathbb{T}_d(v)$  and then never leaving it. The frog model viewed starting from the time of activation only at vertices  $\{u\} \cup \mathbb{T}_d(v)$  then looks identical to the original self-similar frog model viewed on  $\{\emptyset\} \cup \mathbb{T}_d(\emptyset')$ . This yields the following fact, proved in more detail in [HJJ16b, Proposition 6].

**Fact 93.** *Let  $V'$  be the number of frogs that move from  $v$  to its parent  $u$  in the self-similar frog model. The distribution of  $V'$  conditional on  $\mathbb{T}_d(v)$  being activated is identical to the distribution of  $V$ .*

The following observation shows that once a subtree  $\mathbb{T}_d(v)$  is activated, the random variable  $V'$  defined in the above fact is independent of the frog model outside of  $\mathbb{T}_d(v)$ .

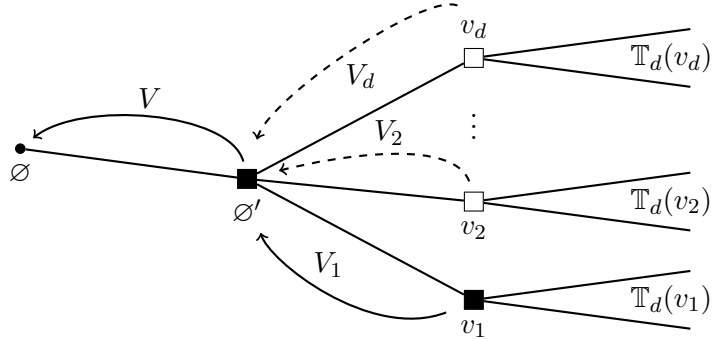


Figure 8: In the self-similar frog model on  $\mathbb{T}_d$ , the initial frog moves from  $\emptyset$  to  $\emptyset'$  to  $v_1$  and then continues down the tree. The random variables  $V$  and  $V_1$ , counting the number of frogs moving from  $\emptyset'$  to  $\emptyset$  and from  $v_1$  to  $\emptyset'$ , respectively, both have distribution  $\nu$  (see Lemma 95). For  $i \in \{2, \dots, d\}$ , the distribution of  $V_i$  conditional on a frog entering the subtree  $\mathbb{T}_d(v_i)$  is also  $\nu$ .

**Fact 94.** *Let  $V'$  be defined as in Fact 93. Conditional on  $\mathbb{T}_d(v)$  being activated,  $V'$  depends only on the path of the activator and on  $\{T_\bullet(w, i), \eta(w) : w \in \mathbb{T}_d(v), i \geq 1\}$ .*

We will use Facts 93 and 94 to express  $V$  recursively in terms of independent copies of itself, an idea expressed in Figure 8. This relation will be given in terms of an operator we define next.

**The operators  $\mathcal{B}$  and  $\mathcal{U}$**  Suppose that the initial frog in the self-similar frog model moves from  $\emptyset$  to  $\emptyset'$  to  $v_1$ . Let  $v_2, \dots, v_d$  be the remaining children of  $\emptyset'$ . Observe that since frogs are stopped at  $\emptyset$ , no children of  $\emptyset$  other than  $\emptyset'$  are ever visited. The idea of this section is to view the self-similar frog model only at the vertices mentioned above. If a vertex  $v_i$  is visited, we close our eyes to  $\mathbb{T}_d(v_i)$ , thinking of this entire subtree as a black box that eventually emits some frogs from  $v_i$  back to  $\emptyset'$ .

Enacting this view, we now define operators  $\mathcal{B}$  and  $\mathcal{U}$  on probability measures supported on the extended nonnegative integers. Informally, the operator  $\mathcal{B}$  corresponds to the number of visits to the root, and  $\mathcal{U}$  corresponds to the number of subtrees  $v_1, \dots, v_d$  that are activated. Let  $\pi$  be a probability measure on the nonnegative integers. To define  $\mathcal{B}\pi$  and  $\mathcal{U}\pi$ , we consider the following frog model. The example to keep in mind is when  $\pi$  is the law of

$V$ , in which case the following description matches up with the black box view of the self-similar frog model described above.

**Graph** a star graph with center  $\rho'$  and leaf vertices  $\rho, u_1, \dots, u_d$  (think of these as paralleling  $\emptyset'$  and  $\emptyset, v_1, \dots, v_d$ ). The root of the graph is  $\rho$ .

**Sleeping frog counts** all independent, distributed as  $\text{Poi}(\mu)$  at  $\rho'$  and as  $\pi$  at  $u_1, \dots, u_d$ . There is one frog at  $\rho$ , as is always true at the root vertex.

**Paths** All frogs have independent paths. The initial frog moves deterministically from  $\rho$  to  $\rho'$  to  $u_1$  and then remains there. All other frogs, if woken, perform simple random non-backtracking walks from their starting points, stopped on arrival at a leaf vertex.

We then define two quantities:

- $\mathcal{B}\pi$  is the distribution of the number of frogs that terminate at  $\rho$ .
- $\mathcal{U}\pi$  is the distribution of the final number of  $u_1, \dots, u_d$  that are visited by a frog.

Note that our definition of the initial frog path as deterministic is just for convenience. By symmetry, we would arrive at the same measures  $\mathcal{B}\pi$  and  $\mathcal{U}\pi$  if it were also defined as a stopped simple random non-backtracking walk.

We mention that  $\mathcal{B}$  is closely related to the operators  $\mathcal{A}$  defined in [HJJ16b] and [HJJ16a], but differs from both of them. The operator  $\mathcal{A}$  in [HJJ16b] is the same as  $\mathcal{B}$  in the  $d = 2$  case if the initial distribution at  $\rho'$  in the definition of  $\mathcal{B}$  is changed from  $\text{Poi}(\mu)$  to  $\delta_1$  (except that  $\mathcal{A}$  acts on probability generating functions rather than distributions). The operator  $\mathcal{A}$  in [HJJ16a] is the same as  $\mathcal{B}$  in the  $d = 2$  case. For  $d \geq 3$ , the two operators differ in that in the frog model defining  $\mathcal{A}$ , frogs initially at  $v_2, \dots, v_d$  do not wake other frogs.

Now we relate this system back to the frog model.

**Lemma 95.** *Let  $\nu$  be the law of  $V = r(\eta, T)$ , the number of visits to the root in the self-similar frog model on  $\mathbb{T}_d$ . It holds that  $\mathcal{B}\nu = \nu$ .*

*Proof.* Essentially, the frog model on the star graph exactly matches the black box view of the self-similar frog model described at the beginning of Section 3.6.2.1, and the result then follows from Facts 93 and 94. To make this more formal, we couple the two frog models. We take full advantage of the abelian properties of the frog model by viewing the frogs' motions in a convenient order.

Consider the frog model used to define  $\mathcal{B}\nu$  as well as the self-similar frog model. We can couple the initial number of frogs on  $\rho'$  to be the same as on  $\emptyset'$ , and we can couple the first (and only) step of each frog at  $\rho'$  with the first step of the corresponding frog at  $\emptyset'$ .

Let  $V_i$  be the number of frogs that ever move from  $v_i$  to  $\emptyset'$  in the self-similar model, and let  $U_i$  be the number of frogs initially at  $u_i$  in the star graph model. Noting that  $\mathbb{T}_d(v_1)$  is activated by the initial frog,  $V_1 \sim \nu$  by Fact 93. By Fact 94,  $V_1$  is independent of all that we have coupled so far (that is, the number and first steps of frogs initially at  $\emptyset'$ ). The random variable  $U_1$  is also independent of all we have coupled so far and is distributed identically to  $V_1$ . We can therefore couple  $U_1$  and  $V_1$  to be equal. Next, we couple the second (and final) step of each frog at  $u_1$  with the step of the corresponding frog counted by  $V_1$  after it moves from  $v_1$  to  $\emptyset'$ .

Let  $\mathcal{V}_1$  consist of the indices  $i \in \{2, \dots, d\}$  such that  $u_i$  has been visited so far. By the construction of our coupling, we can also describe  $\mathcal{V}_1$  as the set of  $i \in \{2, \dots, d\}$  such that  $\mathbb{T}_d(v_i)$  has been activated so far. Furthermore, identically many frogs have returned so far to  $\rho$  as to  $\emptyset$ . By Facts 93 and 94, conditional on the information so far, the random variables  $(V_i, i \in \mathcal{V}_1)$  are i.i.d.- $\nu$  and are independent of the information so far, as are the random variables  $(U_i, i \in \mathcal{V}_1)$ . We can therefore couple these two random vectors to be equal. We then couple the paths of the frogs at these vertices to match up as we did with the frogs at  $u_1$  and  $v_1$ .

As above, a vertex  $u_i$  is visited for the first time in this second round if and only if  $\mathbb{T}_d(v_i)$  is visited for the first time in this second round. Let  $\mathcal{V}_2$  be the set of such  $i$ . We can repeat the coupling argument of the previous paragraph, maintaining identical numbers of frogs terminating at  $\rho$  as at  $\emptyset$ , until we get an empty  $\mathcal{V}_j$  and have counted all returns to  $\rho$  and  $\emptyset$ . Thus, under this coupling, the number of frogs terminating at  $\rho$  in the star graph model is the same as the number of frogs terminating at  $\emptyset$  in the self-similar model. The first of these counts has distribution  $\mathcal{B}\nu$ , while the second has distribution  $\nu$ , showing that the two are equal.  $\square$

The next lemma is similar to [HJJ16b, Lemma 10] and [HJJ16a, Lemma 10].

**Lemma 96.** *If  $\pi_1 \preceq_{st} \pi_2$ , then  $\mathcal{B}\pi_1 \preceq_{st} \mathcal{B}\pi_2$ .*

*Proof.* This immediately follows from the coupling definition of stochastic dominance. We couple the frog models defining  $\mathcal{B}\pi_1$  and  $\mathcal{B}\pi_2$  so that the frogs in the former are a subset of the frogs of the latter model, resulting in more visits to the root.  $\square$

Just as in [HJJ16a, Lemma 11], the operator  $\mathcal{B}$  applied to a Poisson distribution yields a mixture of Poisson distributions. This is a consequence of the following property, known as *Poisson thinning*: Consider a multinomial distribution with  $\text{Poi}(\lambda)$  trials and  $n$ -types, each having probability  $p_k$ . Then the vector of outcomes is distributed as an independent collection of  $\text{Poi}(\lambda p_k)$  random variables.

**Lemma 97.** *Let  $U$  be a random variable distributed as  $\mathcal{U} \text{Poi}(\lambda)$ .*

$$\mathcal{B} \text{Poi}(\lambda) = \text{Poi}\left(\frac{\mu}{d+1} + U \frac{\lambda}{d}\right). \quad (11)$$

*Proof.* In the frog model defining  $\mathcal{B} \text{Poi}(\lambda)$ , the number of frogs at  $\rho'$  that move back to  $\rho$  is distributed as  $\text{Bin}(\text{Poi}(\mu), 1/(d+1))$ . By Poisson thinning, this is  $\text{Poi}(\mu/(d+1))$ . Each visited  $u_i$  releases  $\text{Poi}(\lambda)$  sleeping frogs. These will take a non-backtracking step back to  $\rho'$ , then with probability  $1/d$  will move to  $\rho$ . Thus, each activated  $u_i$  sends  $\text{Poi}(\lambda/d)$  frogs to  $\rho$ . It follows that

$$\mathcal{B} \text{Poi}(\lambda) \sim \text{Poi}(\mu/(d+1)) + \sum_1^d \mathbf{1}\{u_i \text{ visited}\} \text{Poi}(\lambda/d).$$

The above sum is equal to  $\sum_1^U \text{Poi}(\lambda/d)$ . By Poisson thinning, the  $\text{Poi}(\lambda/d)$  terms are independent of  $U$ . Applying additivity of Poisson random variables then brings us to the claimed formula.  $\square$

### 3.6.3 Simplest proof of recurrence on the binary tree

We now break from the main thread to give a short proof that the frog model on the binary tree with Poisson-distributed frogs has a recurrence phase. The idea of the argument is to use Lemma 97 to demonstrate that for some  $\delta > 0$ , it holds for all  $\lambda \geq 0$  that  $\mathcal{B} \text{Poi}(\lambda) \succeq_{\text{st}} \text{Poi}(\lambda + \delta)$ . Lemmas 95 and 96 then let us bootstrap our way to the conclusion that that  $V$  is stochastically larger than any Poisson distribution, and hence  $V = \infty$  a.s.

**Proposition 98.** *The frog model on  $\mathbb{T}_2$  with i.i.d.- $\text{Poi}(\mu)$  frogs per site is recurrent for  $\mu > 3 \log((1 + \sqrt{5})/2) \approx 1.4436$ .*

*Proof.* In the  $d = 2$  case,  $\mathcal{B} \text{Poi}(\lambda)$  is a particularly simple mixture:

$$\mathcal{B} \text{Poi}(\lambda) = \begin{cases} \text{Poi}\left(\frac{\mu}{3} + \frac{\lambda}{2}\right) & \text{with probability } \exp\left(-\frac{\mu}{3} - \frac{\lambda}{2}\right), \\ \text{Poi}\left(\frac{\mu}{3} + \lambda\right) & \text{with probability } 1 - \exp\left(-\frac{\mu}{3} - \frac{\lambda}{2}\right). \end{cases} \quad (12)$$

This follows from Lemma 97 once we show that

$$\mathcal{U}\text{Poi}(\lambda) = \begin{cases} 1 & \text{with probability } \exp(-\frac{\mu}{3} - \frac{\lambda}{2}), \\ 2 & \text{with probability } 1 - \exp(-\frac{\mu}{3} - \frac{\lambda}{2}). \end{cases} \quad (13)$$

Recall that  $\mathcal{U}\text{Poi}(\lambda)$  is the distribution of the number of vertices out of  $\{u_1, u_2\}$  that are visited in the frog model defining  $\mathcal{B}$ . The vertex  $u_1$  is always visited by the initial frog in this model. Each of the  $\text{Poi}(\mu)$  frogs initially at  $\rho'$  has a  $1/3$  chance of visiting  $u_2$ , and each of the  $\text{Poi}(\lambda)$  frogs initially at  $u_1$  has a  $1/2$  chance of visiting  $u_2$ , all independently of each other. By Poisson thinning, the number of these frogs that visit  $u_2$  is distributed as  $\text{Poi}(\mu/3 + \lambda/2)$ , and thus  $u_2$  is visited with probability  $1 - \exp(-\mu/3 - \lambda/2)$ . This establishes (13) and hence (12).

Theorem 90 now instructs us to compute the probability placed on 0 by  $\mathcal{B}\text{Poi}(\lambda)$ , which is

$$\begin{aligned} -\log \left[ (1 - e^{-\frac{\mu}{3} - \frac{\lambda}{2}}) e^{-\frac{\mu}{3} - \lambda} + e^{-\frac{2\mu}{3} - \lambda} \right] &= \lambda + \frac{\mu}{3} - \log \left[ 1 - e^{-\frac{\mu}{3} - \frac{\lambda}{2}} + e^{-\frac{\mu}{3}} \right] \\ &\geq \lambda + \frac{\mu}{3} - \log \left[ 1 + e^{-\frac{\mu}{3}} \right] > \lambda + \delta \end{aligned}$$

for some  $\delta > 0$  depending on  $\mu$  but not on  $\lambda$ , under our assumption that  $\mu > 3 \log((1 + \sqrt{5})/2)$ . By Theorem 90,

$$\mathcal{B}\text{Poi}(\lambda) \succeq_{\text{st}} \text{Poi}(\lambda + \delta) \quad (14)$$

for any  $\lambda \geq 0$ .

Now, we carry out the bootstrap. Recall that  $\nu$  is the distribution of  $r(\eta, T)$ , the number of visits to the root in the nonbacktracking frog model on the binary tree. As  $\nu \succeq_{\text{st}} \text{Poi}(0)$ , Lemma 96 shows that  $\mathcal{B}\nu \succeq_{\text{st}} \mathcal{B}\text{Poi}(0)$ , and so  $\mathcal{B}\nu \succeq_{\text{st}} \text{Poi}(\delta)$  by (14). But  $\nu$  is a fixed point of  $\mathcal{B}$  by Lemma 95, implying that  $\nu \succeq_{\text{st}} \text{Poi}(\delta)$ . Repeating this argument of successively applying Lemma 96, (14), and Lemma 95, we show that  $\nu \succeq_{\text{st}} \text{Poi}(2\delta)$ , and so on. Thus  $\nu$  is stochastically larger than all Poisson distributions, which implies  $\nu = \delta_\infty$ . Finally, recalling that  $r(\eta, S)$  is the number of visits to the root in the frog model and that  $r(\eta, S) \succeq_{\text{st}} \nu$  by Facts 91 and 92, we can conclude that  $r(\eta, S) = \infty$  a.s.  $\square$

### 3.6.4 A more complicated bootstrap

We now give an argument along the same lines as our proof of Proposition 98, proving recurrence for all  $d$  under assumptions on  $\mu$  that are optimal up to



constant factors, as shown by the lower bound in Theorem 88. This bound follows from [HJJ16a, Proposition 15], which is proven by coupling the frog model with a transient branching random walk. Our contribution here is the upper bound.

Recall that  $\mathcal{U} \text{Poi}(\lambda)$  is the distribution of the number of vertices  $u_1, \dots, u_d$  visited in the frog model on the star graph defined in Section 3.6.2.1. The essential difference between our proofs of recurrence for  $\mu = \Omega(d)$  here and for  $\mu = \Omega(d \log d)$  in [HJJ16a] is that here we give a better lower bound on  $\mathcal{U} \text{Poi}(\lambda)$ . For a fixed  $\lambda \geq 0$ , we define a lower bounding random variable  $U' \in \{1, \dots, d\}$  as follows. Consider the frog model used to define  $\mathcal{B} \text{Poi}(\lambda)$  and  $\mathcal{U} \text{Poi}(\lambda)$ , and observe how many of  $u_1, \dots, u_d$  are visited by the  $\text{Poi}(\mu)$  frogs starting at  $\rho'$ . If at least  $\lceil d/c \rceil$  of these vertices are visited for a yet to be determined constant  $c$ , then arbitrarily choose  $\lceil d/c \rceil$  of them and allow the frogs activated there the chance to visit the remaining  $d - \lceil d/c \rceil$  vertices. If fewer than  $\lceil d/c \rceil$  vertices are visited by the frogs at  $\rho'$ , then recall that  $u_1$  is guaranteed to be activated by the initial frog, and just use the frogs at  $u_1$  to try to activate the remaining vertices  $u_2, \dots, u_d$ . We define  $U'$  as the number of vertices out of  $u_1, \dots, u_d$  activated in the end in this scheme. This is summarized as follows:

Let  $U'_1$  be the number of vertices  $u_1, \dots, u_d$  visited by the frogs initially at  $\rho'$ .

**Case 1**  $U'_1 \geq \lceil d/c \rceil$

Arbitrarily choose  $\lceil d/c \rceil$  of the vertices counted by  $U'_1$  and denote them by  $\mathcal{V} \subseteq \{u_1, \dots, u_d\}$ . Let  $U'$  be the sum of  $\lceil d/c \rceil$  and the number of the remaining  $d - \lceil d/c \rceil$  vertices  $\{u_1, \dots, u_d\} \setminus \mathcal{V}$  visited by frogs starting in  $\mathcal{V}$ .

**Case 2**  $U'_1 < \lceil d/c \rceil$

Let  $U'$  equal one plus the number of number of vertices  $u_2, \dots, u_d$  visited by frogs returning from  $u_1$ .

As  $U'$  counts only a subset of the full collection of activated vertices, we have  $U' \preceq_{\text{st}} \mathcal{U} \text{Poi}(\lambda)$ .

We now sketch the argument in more detail. Throughout the remainder of the paper, we will assume that  $\mu = C(d+1)$  with  $C$  a yet to be determined positive constant. In Lemma 99, we prove that Case 2 occurs with exponentially small probability as  $d$  grows. Next, in Lemma 100 we give a very explicit definition of a random variable  $U''$  satisfying  $U'' \preceq_{\text{st}} U' \preceq_{\text{st}} \mathcal{U} \text{Poi}(\lambda)$ . In Lemma 101, we use this lower bound together with Lemma 97 to prove

that if  $V \succeq_{\text{st}} \text{Poi}(\lambda)$ , then  $V \succeq_{\text{st}} \text{Poi}(\lambda + \delta)$  for some  $\delta > 0$ . The same iterative argument used in Proposition 98 then implies that  $V = \infty$  a.s.

**Lemma 99.** *Recall that  $U'_1$  is the number of vertices  $u_1, \dots, u_d$  visited by the  $\text{Poi}(\mu)$  frogs initially at  $\rho'$  in the frog model defining  $\mathcal{B} \text{Poi}(\lambda)$  and  $\mathcal{U} \text{Poi}(\lambda)$ . We have*

$$\mathbf{P}[U'_1 < \lceil d/c \rceil] \leq e^{-bd} := p, \quad (15)$$

where  $b = 2(1 - e^{-C} - \frac{1}{c})^2$ .

*Proof.* It is a consequence of Poisson thinning that out of the  $\text{Poi}(\mu)$  frogs starting at  $\rho'$ , independently  $\text{Poi}(\frac{\mu}{d+1}) = \text{Poi}(C)$  move to each leaf  $u_1, \dots, u_d$ . Thus each vertex has an independent  $1 - e^{-C}$  chance of having a frog visit it from the ones starting at  $\rho'$ , showing that  $U'_1 \sim \text{Bin}(d, 1 - e^{-C})$ .

Hoeffding's inequality tailored to a binomial distribution states that  $\mathbf{P}[\text{Bin}(n, p) \leq (p - \epsilon)n] \leq \exp(-2\epsilon^2 n)$  (this follows from [Hoe63, eq. (2.3)]). If we apply the inequality to  $U'_1$  with  $\epsilon = (1 - e^{-C}) - \frac{1}{c}$ , we establish (15).  $\square$

**Lemma 100.** *Let*

$$U'' \sim \begin{cases} \lceil d/c \rceil + \text{Bin}(d - \lceil d/c \rceil, 1 - e^{-\lambda/c}) & \text{with probability } 1 - q, \\ 1 + \text{Bin}(d - 1, 1 - e^{-\lambda/d}) & \text{with probability } q, \end{cases} \quad (16)$$

where  $q = \mathbf{P}[U'_1 < \lceil d/c \rceil]$ . Then  $U'' \preceq_{\text{st}} U'$ .

*Proof.* Writing  $U' | E$  to mean  $U'$  conditioned on the event  $E$ , we claim that

$$U' | \{U'_1 \geq \lceil d/c \rceil\} \succeq_{\text{st}} \lceil d/c \rceil + \text{Bin}(d - \lceil d/c \rceil, 1 - e^{-\lambda/c}), \quad (17)$$

and

$$U' | \{U'_1 < \lceil d/c \rceil\} \succeq_{\text{st}} 1 + \text{Bin}(d - 1, 1 - e^{-\lambda/d}). \quad (18)$$

The lemma then follows because conditional stochastic dominance implies stochastic dominance [SS07, Theorem 1.A.3, (d)].

Thus it just remains to confirm (17) and (18). Suppose  $U'_1 \geq \lceil d/c \rceil$ . Then we are in Case 1, and  $U' = \lceil d/c \rceil + U'_2$ , where  $U'_2$  is the number of vertices in  $\{u_1, \dots, u_d\} \setminus \mathcal{V}$  visited by frogs returning from  $\mathcal{V}$ . Conditional on  $\mathcal{V}$ , the counts of frogs proceeding from  $\mathcal{V}$  to each of  $\{u_1, \dots, u_d\} \setminus \mathcal{V}$  form a collection of independent  $\text{Poi}(\lambda \lceil d/c \rceil / d)$  random variables. Thus each vertex in  $\{u_1, \dots, u_d\} \setminus \mathcal{V}$  has an independent probability of  $1 - e^{-\lambda \lceil d/c \rceil / d} \geq 1 - e^{-\lambda/c}$

of being visited by one of these frogs, showing that  $U_2' \succeq_{\text{st}} \text{Bin}(d - \lceil d/c \rceil, 1 - e^{-\lambda/c})$  and confirming (17).

Next, suppose that  $U_1' < \lceil d/c \rceil$ , and Case 2 is in effect. In this case,  $U' = 1 + U_2'$ , where  $U_2'$  is the number of vertices  $u_2, \dots, u_d$  visited by frogs returning from  $u_1$ . By the same reasoning as in the previous case,  $U_2' \succeq_{\text{st}} \text{Bin}(d - 1, 1 - e^{-\lambda/d})$ , confirming (18).  $\square$

**Lemma 101.** *Define*

$$h_{C,c} = h_{C,c}(\lambda, d) := \log \left[ \left( e^{-\frac{\lambda}{c} + \frac{\lambda}{d}} + 1 - e^{-\frac{\lambda}{c}} \right)^{d - \lceil d/c \rceil} + p(2 - e^{-\frac{\lambda}{d}})^{d-1} \right],$$

where  $p$  is the value defined in (15), which depends on  $C$  and  $c$ . We have

$$\mathcal{B} \text{Poi}(\lambda) \succeq_{\text{st}} \text{Poi} \left( \lambda + \frac{\mu}{d+1} - h_{C,c} \right).$$

*Proof.* Combining (11) and  $U'' \preceq_{\text{st}} \mathcal{U} \text{Poi}(\lambda)$ , it follows from [SS07, Theorem 1.A.3, (d)] that

$$\mathcal{B} \text{Poi}(\lambda) \succeq_{\text{st}} \text{Poi} \left( \frac{\mu}{d+1} + U'' \frac{\lambda}{d} \right). \quad (19)$$

In light of Theorem 90, we would like to compute  $-\log \mathbf{E} e^{-\frac{\lambda}{d} U''}$ . Recalling the definition of  $U''$  in (16), we use the fact that  $\mathbf{E} x^{\text{Bin}(n,p)} = (1 - p + px)^n$  to compute

$$\begin{aligned} \mathbf{E} e^{-\frac{\lambda}{d} U''} &= (1 - q) e^{-\frac{\lambda}{d} \lceil d/c \rceil} \left( e^{-\frac{\lambda}{c}} + (1 - e^{-\frac{\lambda}{c}}) e^{-\frac{\lambda}{d}} \right)^{d - \lceil d/c \rceil} \\ &\quad + q e^{-\frac{\lambda}{d}} \left( e^{-\frac{\lambda}{d}} + (1 - e^{-\frac{\lambda}{d}}) e^{-\frac{\lambda}{d}} \right)^{d-1}. \end{aligned}$$

Using the bound  $q \leq p$  from Lemma 99 and the trivial bound  $1 - q \leq 1$  in the first step, and factoring out  $e^{-\lambda}$  in the second step,

$$\begin{aligned} \mathbf{E} e^{-\frac{\lambda}{d} U''} &\leq e^{-\frac{\lambda}{d} \lceil d/c \rceil} \left( e^{-\frac{\lambda}{c}} + (1 - e^{-\frac{\lambda}{c}}) e^{-\frac{\lambda}{d}} \right)^{d - \lceil d/c \rceil} \\ &\quad + p e^{-\frac{\lambda}{d}} \left( e^{-\frac{\lambda}{d}} + (1 - e^{-\frac{\lambda}{d}}) e^{-\frac{\lambda}{d}} \right)^{d-1} \\ &= e^{-\lambda} \left[ \left( e^{-\frac{\lambda}{c} + \frac{\lambda}{d}} + 1 - e^{-\frac{\lambda}{c}} \right)^{d - \lceil d/c \rceil} + p(2 - e^{-\frac{\lambda}{d}})^{d-1} \right]. \end{aligned}$$

Thus,

$$-\log \mathbf{E} e^{-\frac{\lambda}{d} U''} = \lambda - h_{C,c}.$$

Using the above calculation and Theorem 90, we deduce that

$$\text{Poi}\left(\frac{\mu}{d+1} + U''\frac{\lambda}{d}\right) \succeq_{\text{st}} \text{Poi}\left(\lambda + \frac{\mu}{d+1} - h_{C,c}\right).$$

Together with (19), this completes the proof.  $\square$

*Proof of Theorem 88.* As we noted before, the lower bound is a consequence of [HJJ16a, Proposition 15], and we just need to establish the upper bound by showing that the frog model on  $\mathbb{T}_d$  is almost surely recurrent with i.i.d.- $\text{Poi}(2.28d)$  frogs per vertex for sufficiently large  $d$ . To apply our bootstrapping argument, we seek to show that for some  $\delta > 0$ , it holds for all  $\lambda \geq 0$  that  $\mathcal{B}\text{Poi}(\lambda) \succeq_{\text{st}} \text{Poi}(\lambda + \delta)$ . Considering the result of Lemma 101, we need to choose  $C$  and  $c$  such that  $\mu/(d+1) - h_{C,c}(\lambda, d) > \delta$  for all  $\lambda \geq 0$  and sufficiently large  $d$ . Recalling that  $\mu = C(d+1)$ , rearranging terms, and exponentiating both sides of the inequality, this is equivalent to showing that for some  $C, c, \delta$ , and  $d_0$  it holds that

$$\exp(h_{C,c}(\lambda, d)) < e^{C-\delta}, \quad (20)$$

on the set  $\{(\lambda, d) : \lambda \geq 0, d \geq d_0\}$ .

Towards proving this, we start with the inequality

$$\begin{aligned} \exp(h_{C,c}(\lambda, d)) &= \left(e^{-\frac{\lambda}{c} + \frac{\lambda}{d}} + 1 - e^{-\frac{\lambda}{c}}\right)^{d - \lceil d/c \rceil} + p(2 - e^{-\frac{\lambda}{d}})^{d-1} \\ &\leq \left(1 + e^{-\frac{\lambda}{c}}\left(e^{\frac{\lambda}{d}} - 1\right)\right)^{d(1 - \frac{1}{c})} + e^{-bd}2^{d-1} \end{aligned} \quad (21)$$

obtained by applying the bounds  $2 - e^{-\lambda/d} \leq 2$  and  $d - \lceil d/c \rceil \leq d(1 - 1/c)$  and substituting the value of  $p$  from (15). Note that  $b$  depends on  $C$  and  $c$ . Now we bound each of the two terms on the right hand side of (21) for the right choice of  $C, c$ , and  $d_0$ .

Some calculus shows that for any  $d$  and  $c$  satisfying  $d > c$ , the first term is maximized in  $\lambda$  when  $e^{\lambda/d} = d/(d-c)$ . This demonstrates that if  $d > c$ ,

then

$$\begin{aligned}
(1 + e^{-\frac{\lambda}{c}}(e^{\frac{\lambda}{d}} - 1))^{d(1-\frac{1}{c})} &\leq \left(1 + \left(\frac{d-c}{d}\right)^{d/c} \left(\frac{d}{d-c} - 1\right)\right)^{d(1-\frac{1}{c})} \\
&= \left(1 + \left(\frac{d-c}{d}\right)^{d/c} \frac{c}{d-c}\right)^{d(1-\frac{1}{c})} \\
&\leq \left(1 + \frac{c}{d-c}\right)^{d(1-\frac{1}{c})} \\
&\leq \exp\left(\frac{d(c-1)}{d-c}\right).
\end{aligned}$$

Thus, for any  $\epsilon > 0$ , we can choose  $d_0$  sufficiently large that for all  $d \geq d_0$  and  $\lambda > 0$ ,

$$(1 + e^{-\frac{\lambda}{c}}(e^{\frac{\lambda}{d}} - 1))^{d(1-\frac{1}{c})} \leq e^{c-1+\epsilon}. \quad (22)$$

The second term to be bounded,  $e^{-bd}2^{d-1}$ , vanishes as  $d \rightarrow \infty$  when  $b > \log 2$ . Referring back to (15) and doing some algebra, we see that  $b > \log 2$  when

$$C > -\log\left(1 - \frac{1}{c} - \sqrt{\frac{\log 2}{2}}\right). \quad (23)$$

If this inequality holds, then for any  $\epsilon > 0$ , we can choose  $d_0$  large enough that for all  $d \geq d_0$ ,

$$e^{-bd}2^{d-1} \leq \epsilon. \quad (24)$$

Finally, we set  $c = 3.26$  and  $C = 2.27$ , which satisfies (23). Applying (67) and (68), for any  $\epsilon > 0$ , there exists  $d_0$  such that for all  $d \geq d_0$  and  $\lambda \geq 0$ ,

$$\exp(h_{C,c}(\lambda, d)) \leq e^{2.26+\epsilon} + \epsilon.$$

Choosing  $\epsilon, \delta > 0$  sufficiently small, we can bound this by  $e^{2.27-\delta}$ , confirming (20).

In all, we have shown that for  $d \geq d_0$ , if  $\mu \geq 2.27(d+1)$  then for all  $\lambda \geq 0$ ,

$$\mathcal{B} \text{Poi}(\lambda) \succeq_{\text{st}} \text{Poi}(\lambda + \delta). \quad (25)$$

Increasing  $d_0$  as necessary, we can revise our assumption to  $\mu \geq 2.28d$  for  $d \geq d_0$ . The rest of the proof is to use this to bootstrap our way to the conclusion that the number of visits to the root in the frog model is almost surely infinite given these assumptions, which proceeds identically as the last paragraph of Proposition 98.  $\square$

### 3.6.5 Transience regime

The main idea of our proof of transience is to consider a *weight function* on the frog model. To analyze the weight function, we bound the frog model by a branching random walk. The weight function is the frog model analogue to a common martingale derived from branching random walk (see [Big77]).

**Proposition 102.** *If  $\mathbf{E}\eta < \frac{(d-1)^2}{4d}$ , then the frog model with an independent copy of  $\eta$  frogs per site on  $\mathbb{T}_d$  is almost surely transient.*

*Proof.* Let  $F_n$  be the set of frogs awake at time  $n$ . For  $f \in F_n$ , let  $|f|$  denote the level of  $f$  on the tree (that is, its distance from the root). We define a weight function

$$W_n = \sum_{f \in F_n} e^{-\theta|f|},$$

with  $\theta$  to be chosen shortly. Let

$$m = \frac{1}{d+1}e^\theta + \frac{d}{d+1}\mathbf{E}[\eta + 1]e^{-\theta}.$$

Before we explain the meaning of this, we minimize  $m$  by setting  $\theta = \log((\mathbf{E}\eta + 1)d)/2$ , making

$$m = \frac{2\sqrt{(\mathbf{E}\eta + 1)d}}{d+1} < 1$$

under our assumption that  $\mathbf{E}\eta < \frac{(d-1)^2}{4d}$ .

The strategy of the proof now is to show that  $W_n \rightarrow 0$ , and hence that the root eventually stops being visited. The term  $m$  gives an upper bound for the expected contribution to  $W_{n+1}$  of a frog at time  $n$  in the following way: Suppose that at time  $n$ , some frog  $f$  is at level  $i$  of the tree for any  $i \geq 1$ . With probability  $1/(d+1)$ , the next jump of  $f$  is towards the root, waking no frogs. With probability  $d/(d+1)$ , the jump is away from the root, possibly waking up an  $\eta$ -distributed number of frogs. Thus, the expected contribution to  $W_{n+1}$  from  $f$  and any frogs it wakes at time  $n+1$  is at most  $e^{-\theta i}m$ . If  $f$  is at the root at time  $n$ , then the expected contribution to  $W_{n+1}$  from  $f$  and the frogs it wakes is at most  $\mathbf{E}[\eta + 1]e^{-\theta}$ , which is bounded by  $m$  given our choice of  $\theta$ . Therefore

$$\mathbf{E}[W_{n+1} | W_n] \leq \sum_{f \in F_n} e^{-\theta|f|}m = mW_n.$$

Thus  $W_n/m^n$  is a positive supermartingale. By the martingale convergence theorem, it converges almost surely to a finite limit. Since  $m^n \rightarrow 0$ , we also have  $W_n \rightarrow 0$  a.s., which implies that eventually no frogs are present at the root.  $\square$

### 3.7 Stochastic orders and the frog model

#### 3.7.1 Background material on stochastic orders

Let  $\pi_1$  and  $\pi_2$  be probability measures on the extended nonnegative real numbers  $[0, \infty]$ , and let  $X \sim \pi_1$  and  $Y \sim \pi_2$ . The following three stochastic orders will play a role in this paper:

**Standard stochastic order**  $\pi_1 \preceq_{\text{st}} \pi_2$  if  $\mathbf{E}f(X) \leq \mathbf{E}f(Y)$  for all bounded increasing functions  $f: [0, \infty) \rightarrow \mathbb{R}$ , with  $f(\infty)$  taken as  $\lim_{x \rightarrow \infty} f(x)$ .

**Increasing concave order**  $\pi_1 \preceq_{\text{icv}} \pi_2$  if  $\mathbf{E}f(X) \leq \mathbf{E}f(Y)$  for all bounded increasing concave functions  $f: [0, \infty) \rightarrow \mathbb{R}$ , with  $f(\infty)$  taken as  $\lim_{x \rightarrow \infty} f(x)$ .

**Probability generating function order**  $\pi_1 \preceq_{\text{pgf}} \pi_2$  if  $\mathbf{E}t^X \geq \mathbf{E}t^Y$  for all  $t \in (0, 1)$ , with  $t^\infty$  interpreted as 0.

We use  $X \preceq_{\text{st}} Y$ ,  $X \preceq_{\text{st}} \pi_2$ , and  $\pi_1 \preceq_{\text{st}} Y$  all to mean that  $\pi_1 \preceq_{\text{st}} \pi_2$ , and we do the same with the other two orders.

We have listed these three stochastic orders in decreasing strength. That is,

$$\pi_1 \preceq_{\text{st}} \pi_2 \implies \pi_1 \preceq_{\text{icv}} \pi_2 \implies \pi_1 \preceq_{\text{pgf}} \pi_2. \quad (26)$$

The first implication is obvious. For the second, the map  $x \mapsto 1 - t^x$  is an increasing concave function for any  $t \in (0, 1)$ , establishing that  $\mathbf{E}t^X \geq \mathbf{E}t^Y$  for  $t \in (0, 1)$  if  $X \preceq_{\text{icv}} Y$ .

See [SS07] for a reference on stochastic dominance. We have made two slight changes from the usual definitions found there. First, in the standard and icv orders, we have required our test functions to be bounded. This apparently weaker definition is in fact equivalent to the usual one, as seen by approximating an unbounded increasing or increasing concave function by a sequence of bounded ones. Second, we have allowed our probability measures to take the value  $\infty$  with positive probability. All of the standard results on stochastic orderings are unaffected by this change. It is worth noting that if

$X \preceq_{\text{pgf}} Y$ , then  $\mathbf{P}[X = \infty] \leq \mathbf{P}[Y = \infty]$ . To see this, note that as  $t \nearrow 1$ , we have  $t^x \rightarrow \mathbf{1}\{x < \infty\}$ . Thus, by the monotone convergence theorem,

$$\mathbf{E}t^X \rightarrow \mathbf{P}[X < \infty] \quad \text{and} \quad \mathbf{E}t^Y \rightarrow \mathbf{P}[Y < \infty]$$

as  $t \nearrow 1$ . Now  $\mathbf{E}t^X \geq \mathbf{E}t^Y$  for  $t \in (0, 1)$  implies that  $\mathbf{P}[X < \infty] \geq \mathbf{P}[Y < \infty]$ . By (26), the conclusion also holds under the assumption  $X \preceq_{\text{st}} Y$  or  $X \preceq_{\text{icv}} Y$ . We also mention that a similar argument with a limit as  $t \searrow 0$  shows that if  $X \preceq_{\text{pgf}} Y$ , then  $\mathbf{P}[X = 0] \geq \mathbf{P}[Y = 0]$ .

Roughly speaking, the standard order rewards distributions for being large, while the icv order rewards them either for being large or for being concentrated. The characterizations of these two orders in terms of couplings make this more precise:  $X \preceq_{\text{st}} Y$  if and only if  $X$  and  $Y$  can be coupled so that  $X \leq Y$  a.s. [SS07, Theorem 1.A.1], and  $X \preceq_{\text{icv}} Y$  if and only if  $X$  and  $Y$  can be coupled so that  $\mathbf{E}[X | Y] \leq Y$  a.s. [SS07, Theorem 4.A.5]. Another useful equivalent condition for  $\pi_1 \preceq_{\text{st}} \pi_2$  is that  $\mathbf{P}[X > t] \leq \mathbf{P}[Y > t]$  for all  $t$ .

The following proposition shows that the maximal real- and integer-valued distributions in the icv order with a given expectation are the distributions that are as concentrated as possible.

**Proposition 103.**

- (a) If  $\mathbf{E}X \leq c$ , then  $X \preceq_{\text{icv}} c$ .
- (b) Suppose  $X$  takes nonnegative integer values and  $\mathbf{E}X \in [k, k + 1]$  for an integer  $k$ . Let  $Y$  be a random variable taking values in  $\{k, k + 1\}$  and satisfying  $\mathbf{E}X \leq \mathbf{E}Y$ . Then  $X \preceq_{\text{icv}} Y$ .

*Proof.* Part (a) follows immediately from Jensen's inequality. For part (b), let  $\varphi$  be an arbitrary increasing concave function on  $[0, \infty)$ . To simplify the algebra, let  $U = X - k$ ,  $V = Y - k$ , and  $\psi(x) = \varphi(x + k) - \varphi(k)$ . With these replacements, our goal is to show that  $\mathbf{E}\psi(U) \leq \mathbf{E}\psi(V)$ . We know that  $\mathbf{E}U \in [0, 1]$  and that  $V$  is Bernoulli, and we know that  $\psi$  is increasing and concave on  $[-k, \infty)$  and satisfies  $\psi(0) = 0$ .

Since  $V$  is Bernoulli with mean at least  $\mathbf{E}U$ ,

$$\mathbf{E}\psi(V) \geq (\mathbf{E}U)\psi(1). \tag{27}$$

Define

$$\begin{aligned} a &= \mathbf{E}[U | U \leq 0], & p &= \mathbf{P}[U \leq 0], \\ b &= \mathbf{E}[U | U \geq 1], & q &= 1 - p = \mathbf{P}[U \geq 1]. \end{aligned}$$



If  $p = 0$  or  $q = 0$ , then  $U$  is deterministic and the result is trivial because  $U$  and  $V$  have the same distribution. Thus we can assume that both conditional expectations above are well defined.

Applying Jensen's inequality,

$$\mathbf{E}\psi(U) = p\mathbf{E}[\psi(U) \mid U \leq 0] + q\mathbf{E}[\psi(U) \mid U \geq 1] \leq p\psi(a) + q\psi(b). \quad (28)$$

As  $a \leq 0$  and  $b \geq 1$ , the points  $(a, \psi(a))$  and  $(b, \psi(b))$  lie under the secant line connecting  $(0, 0)$  and  $(1, \psi(1))$  by the concavity of  $\psi$ . Thus  $\psi(a) \leq a\psi(1)$  and  $\psi(b) \leq b\psi(1)$ . Applying to this to (28) and combining with (27) gives

$$\mathbf{E}\psi(U) \leq (pa + qb)\psi(1) = (\mathbf{E}U)\psi(1) \leq \mathbf{E}\psi(V). \square$$

### 3.7.2 Formalities

Let  $G$  be a countable collection of vertices, one of which we distinguish as the root and call  $\emptyset$ . A general frog model  $(\eta, S)$  starts with one active particle at  $\emptyset$  and  $\eta(v)$  dormant particles at each  $v \neq \emptyset$ . The  $i$ th particle at  $v$  starting from its time of activation moves according to the path  $S_\bullet(v, i)$ , with  $S_0(v, i)$  assumed equal to  $v$ . When an active particle visits a site containing dormant particles, *all* of the dormant particles activate. The particles move in discrete time, though this will be unimportant since most of the properties of the frog model we consider depend only on the particles' paths and not on the time they make their moves. The particles are traditionally called frogs, and we continue the zoomorphism. Typically,  $G$  is a graph, the frog paths  $(S_\bullet(v, i))_{v \in G, i \geq 1}$  are independent random walks, the frog counts  $(\eta(v))_{v \in G}$  are either deterministic or i.i.d., and  $(S_\bullet(v, i))_{v \in G, i \geq 1}$  and  $(\eta(v))_{v \in G}$  are independent of each other. We will not labor an example like the frog model with simple random walk paths on  $\mathbb{Z}^d$  and i.i.d.- $\text{Poi}(\mu)$  frogs per vertex by stating that the frog paths are mutually independent, and that the frog counts and paths are independent. Our main result is about two classes of frog model functionals we call *icv* and *pgf statistics*.

### 3.7.3 Definition of icv and pgf statistics

We now define the classes of statistics covered by our main theorems. Roughly speaking, we call a function of the frog model an *icv statistic* if it increases when a frog is added to the model, but when two frogs are added at the same vertex it increases less than by the separate addition of each of them. The *pgf statistics* form a more restrictive class of frog model functionals that obey a higher order version of this property. Many natural counts in the frog model will fall into these classes.

Before we give the definitions, we will need to introduce some notation. Let

$$\{\eta(v), S_\bullet(v, i) : v \in G, i \geq 1\}$$

be a deterministic collection of frog counts and paths. For any path  $P_\bullet$ , let  $\sigma_{P_\bullet}(\eta, S)$  denote a new frog model with an extra frog of path  $P_\bullet$  added at  $P_0$ . That is,  $\sigma_{P_\bullet}(\eta, S) = (\eta', S')$ , where  $\eta'$  is identical to  $\eta$  except that  $\eta'(P_0) = \eta(P_0) + 1$ , and  $S'$  is identical to  $S$  except that  $S'_\bullet(P_0, \eta(P_0) + 1) = P_\bullet$ . For any frog model statistic  $f(\eta, S)$ , define

$$\Delta_{P_\bullet} f(\eta, S) = f(\sigma_{P_\bullet}(\eta, S)) - f(\eta, S),$$

the change in  $f$  when a frog with path  $P_\bullet$  is added to the model. As with the related operator  $D$  defined in (??), the operator can be applied repeatedly and expanded in the following way, which resembles (??). Let  $P_\bullet^1, \dots, P_\bullet^n$  be frog paths, and let  $U = \{u_1, \dots, u_j\} \subseteq [n]$ , where we use the notation  $[n] = \{1, \dots, n\}$ . Define

$$\sigma_U(\eta, S) = \sigma_{P_\bullet^{u_1}} \cdots \sigma_{P_\bullet^{u_j}}(\eta, S), \quad (29)$$

the frog model  $(\eta, S)$  with the addition of frogs  $P_\bullet^{u_1}, \dots, P_\bullet^{u_j}$ . If  $U$  is empty, take  $\sigma_U(\eta, S) = (\eta, S)$ . Using this notation,

$$\Delta_{P_\bullet^1} \cdots \Delta_{P_\bullet^n} f(\eta, S) = \sum_{U \subseteq [n]} (-1)^{n-|U|} f(\sigma_U(\eta, S)). \quad (30)$$

This can be proven by the same argument used in [Sta12, eq. (1.97)].

**Definition 104.** Let  $f$  be a functional of the frog model taking values in the nonnegative extended real numbers. We call  $f$  a *pgf statistic* if for any  $(\eta, S)$  and any paths  $P_\bullet^1, \dots, P_\bullet^m$  starting at the same vertex for any  $m \geq 1$ ,

$$(-1)^m \Delta_{P_\bullet^1} \cdots \Delta_{P_\bullet^m} f(\eta, S) \leq 0. \quad (31)$$

If (31) holds for any  $(\eta, S)$  and any paths  $P_\bullet^1, \dots, P_\bullet^m$  starting at the same vertex for  $m = 1, 2$ , then we call  $f$  an *icv statistic*. In either case, we call the statistic *continuous* if the condition

$$\eta_k(v) \nearrow \eta(v) \text{ as } k \rightarrow \infty \text{ for all } v \in G$$

implies that  $f(\eta_k, S) \nearrow f(\eta, S)$  as  $k \rightarrow \infty$ .

The  $m = 1$  case of (31) is the condition that  $f$  increases when a new frog is added. To make the  $m = 2$  case more transparent, we can expand condition (31) as

$$f(\sigma_{P_1} \sigma_{P_2}(\eta, S)) - f(\sigma_{P_1}(\eta, S)) - f(\sigma_{P_2}(\eta, S)) + f(\eta, S) \leq 0.$$

Shifting terms around, we have the equivalent condition

$$f(\sigma_{P_1} \sigma_{P_2}(\eta, S)) - f(\eta, S) \leq \left( f(\sigma_{P_1}(\eta, S)) - f(\eta, S) \right) + \left( f(\sigma_{P_2}(\eta, S)) - f(\eta, S) \right),$$

which states that the gain to the statistic by adding two frogs at the same vertex is less than the combined gain of adding each frog separately, as we mentioned earlier.

The prime example is the number of visits to  $\emptyset$  in the frog model  $(\eta, S)$  over all time, which we denote  $r(\eta, S)$ . A realization of the frog model is called *recurrent* if  $r(\eta, S) = \infty$  and *transient* otherwise. In [TW99], the frog model with one sleeping frog per site and simple random walk paths is shown to be recurrent on  $\mathbb{Z}^d$  for all  $d$ . This is further refined in [Pop01], which exhibits a threshold in  $\alpha$  at which a frog model with  $\text{Bernoulli}(\alpha \|x\|^{-2})$  frogs at each  $x \in \mathbb{Z}^d$  switches from transience to recurrence. A similar phenomenon occurs when the walks have a bias in one direction: [GS09] finds that on  $\mathbb{Z}$ , the model is recurrent if and only if the number of sleeping frogs per site has infinite logarithmic moment. A sufficient condition for recurrence in this setting on  $\mathbb{Z}^d$  was given in [DP14] and improved on in [KZ16]. In our papers [HJJ16b, HJJ16a, JJ16a], we prove that the frog model with simple random walk paths on a  $d$ -ary tree switches from transient to recurrent either when the density of frogs increases with  $d$  held fixed, or when  $d$  decreases with the density held fixed.

### 3.7.4 Statement of main theorem

Our main result is a comparison theorem relating icv and pgf statistics of a frog model (see Definition 104) when we vary the distribution of the initial configuration  $\eta$ . Our original motivation was that while the most convenient setting has Poisson-distributed frog counts, the most basic questions assume a deterministic number of frogs per site. As an example, in [HJJ16a] we showed the existence of a recurrence phase on the  $d$ -ary tree with Poisson frogs per site for any  $d \geq 2$ . This left open the existence of a recurrence phase for initial conditions other than i.i.d. Poisson. For instance, for large enough  $k$ , is the frog model recurrent on the  $d$ -ary tree with  $k$  frogs per site? With our previous tools, we could answer this question only for the case

$d = 2$  [HJJ16b], but our comparison theorem tidily transfers the result from Poisson to deterministic initial conditions (see Corollary 108).

If  $\eta(v)$  is dominated by  $\eta'(v)$  in the usual stochastic order, then we can couple the corresponding frog models and deduce that  $f(\eta, S)$  is dominated by  $f(\eta', S)$  for any statistic  $f$  that is increasing in  $\eta$ . This is not helpful for the problem described above, since we cannot relate a Poisson random variable to the constant  $k$  in this stochastic order. We instead turn to weaker stochastic dominance relations known as the *increasing concave order* and the *probability generating function order*. We abbreviate these as *icv order* and *pgf order*, respectively, and we denote the relations by  $\preceq_{\text{icv}}$  and  $\preceq_{\text{pgf}}$ . The icv order has come up several times in discrete probability, most notably in first passage percolation [vdBK93, Mar02]. See also [Zer98] for an application to random walk in a random environment. In these papers, the relation  $\pi_1 \preceq_{\text{icv}} \pi_2$  is referred to as  $\pi_2$  being *more variable* than  $\pi_1$ . The only use of the pgf order that we know of in discrete probability is our own in [HJJ16b], though see [TRZ11, LT14] for some applications in signal processing and wireless networks under the name *Laplace transform order*.

Our main theorem applies to functions of the frog model including the count  $r(\eta, S)$  of visits to the root and the total number of activated sites up to a given time. Our result is that such statistics are monotonic in the initial frog configuration with respect to these weaker stochastic orders.

**Theorem 105.** *Assume that the frog paths  $S_\bullet(v, i)$  and counts  $\eta(v)$  and  $\eta'(v)$  are mutually independent for all  $v$  and  $i$ , and that the paths  $S_\bullet(v, i)$  at a particular vertex  $v$  are identically distributed for all  $i$ .*

- (a) *If  $f$  is a continuous icv statistic in the sense of Definition 104 and  $\eta(v) \preceq_{\text{icv}} \eta'(v)$  for all  $v$ , then  $f(\eta, S) \preceq_{\text{icv}} f(\eta', S)$ .*
- (b) *If  $f$  is a continuous pgf statistic in the sense of Definition 104 and  $\eta(v) \preceq_{\text{pgf}} \eta'(v)$  for all  $v$ , then  $f(\eta, S) \preceq_{\text{pgf}} f(\eta', S)$ .*

The intuition behind the proof is that the extra frogs woken by the addition of two frogs at some vertex is the union of the frogs woken by the addition of each frog separately. This subadditivity property meshes neatly with concavity—for instance, the expected number of visits to the root will increase concavely as frogs are added at a vertex—and somehow this makes the frog model interact well with stochastic orders defined in terms of concave functions.

### 3.7.5 Applications

As we mentioned, our main statistic of interest fits the criteria of Theorem 105.

**Proposition 106.** *The count  $r(\eta, S)$  of visits to  $\emptyset$  in the frog model  $(\eta, S)$  is a continuous icv and pgf statistic of the frog model.*

This allows us to transfer many recurrence and transience results to different initial conditions. In the increasing concave order, the constant  $k$  dominates all mean  $k$  random variables. Theorem 105(a) and Proposition 106 therefore imply the following:

**Corollary 107.** *Consider the frog model on a graph with mutually independent frog paths and i.i.d. frogs per site with common mean  $\mu$ . If this is almost surely recurrent, then for any integer  $k \geq \mu$ , the same frog model with  $k$  frogs per site is almost surely recurrent.*

This solves our problem of showing that the frog model on a  $d$ -ary tree with deterministically  $k$  frogs per site is recurrent for large enough  $k$ . In more detail, [HJJ16a, Theorem 1] establishes that on the  $d$ -ary tree with i.i.d.- $\text{Poi}(\mu)$  frogs per site, there is a critical value  $\mu_c(d)$  such that the frog model is recurrent a.s. if  $\mu > \mu_c(d)$  and transient a.s. if  $\mu < \mu_c(d)$ . Corollary 107, together with the estimates on  $\mu_c(d)$  from [JJ16a], give us the following result:

**Corollary 108.** *For any  $d \geq 2$ , the frog model on  $\mathbb{T}_d$  with  $k$  frogs per site is almost surely recurrent for large enough  $k$ . For large enough  $d$ , the model is almost surely transient if  $k < .24d$  and almost surely recurrent if  $k > 2.28d$ .*

Another application of Theorem 105 concerns the transience regime of the  $d$ -ary tree. In [HJJ16b, Theorem 1] we show that on  $\mathbb{T}_d$  with one frog per site and simple random walk paths, the frog model is transient for  $d \geq 5$ . An immediate corollary of Theorem 105 is transience for all other mean 1 configurations.

**Corollary 109.** *For  $d \geq 5$ , the frog model on  $\mathbb{T}_d$  with  $\eta(v)$  frogs at each site and  $\mathbf{E}\eta(v) \leq 1$  for all  $v \in \mathbb{T}_d$  is almost surely transient.*

Our next application is to the frog model on  $\mathbb{Z}^d$ . As mentioned earlier, [Pop01, Theorem 1.1] establishes the existence of a critical parameter  $0 < \alpha_c(d) < \infty$  for the frog model with simple random walk paths on  $\mathbb{Z}^d$  and initial configuration given by  $\eta(x) \sim \text{Ber}(p_x)$  such that

- (i) if  $p_x \leq \alpha/\|x\|^2$  for  $\alpha < \alpha_c(d)$  and all sufficiently large  $x$ , then the model is transient with positive probability;
- (ii) if  $p_x \geq \alpha/\|x\|^2$  for  $\alpha > \alpha_c(d)$  and all sufficiently large  $x$ , then the model is transient with probability zero.

Theorem 105 allows us to extend part (i) of this result to non-Bernoulli distributions of sleeping frogs. Other results like [Pop01, Theorem 1.3] can be similarly extended.

**Corollary 110.** *For all  $\alpha < \alpha_c(d)$  and any  $(\eta(x), x \in \mathbb{Z}^d \setminus \{0\})$  satisfying  $\mathbf{E}\eta(x) \leq \alpha/\|x\|^2$  for sufficiently large  $x$ , the frog model on  $\mathbb{Z}^d$  with simple random walk paths and initial configuration  $\eta$  is transient with positive probability.*

A fundamental result for the frog model on  $\mathbb{Z}^d$  is that  $\xi_n$ , the set of squares containing activated vertices at time  $n$ , has a limiting shape. This was independently established for the discrete and continuous time one frog per site model in [AMP02a] and [RS04], respectively. When the number of frogs per site is independent and identically distributed as  $\eta$ , the result is still true. [AMPR01, Theorem 1.1] establishes that for any dimension  $d \geq 1$  there is a nonempty convex set  $\mathcal{A} \subseteq \mathbb{R}^d$ , which depends on the distribution of  $\eta$ , such that for any  $0 < \epsilon < 1$ ,

$$(1 - \epsilon)\mathcal{A} \subseteq \lim_{n \rightarrow \infty} \frac{\xi_n}{n} \subseteq (1 + \epsilon)\mathcal{A}.$$

We deduce that the limiting shape,  $\mathcal{A}$ , respects the icv and pgf orders. This mirrors the inequalities for the time constant for first passage percolation that are proven in [vdBK93].

**Corollary 111.** *Let  $\mathcal{A}$  and  $\mathcal{A}'$  be the limiting shapes for a frog model on  $\mathbb{Z}^d$  with i.i.d.  $\eta$  and  $\eta'$  particles at each site, respectively. If either  $\eta \preceq_{icv} \eta'$  or  $\eta \preceq_{pgf} \eta'$ , then  $\mathcal{A} \subseteq \mathcal{A}'$ .*

We also find applications to the frog model with death, explored in [AMP02b, FMS04, LMP05], where frogs have an independent chance  $1 - p$  of dying at each step. This is a frog model according to our general definition, taking the frog paths to be stopped random walks. In this setting, the statistic of interest has been the total number of sites visited, which undergoes a phase transition on the regular tree from being finite a.s. to being infinite with positive probability as  $p$  grows. The model is said to *die out* in the first case and to *survive* in the second. The number of sites visited is an icv and pgf statistic, as we show in Proposition 125, and we therefore obtain the following result.

**Corollary 112.** *Let  $\eta'(v) \succeq_{icv} \eta(v)$  be independent random variables indexed by the vertices  $v$  of an arbitrary graph  $G$ . If the frog model with death on  $G$  survives with  $\eta(v)$  frogs at each  $v$ , then it survives with  $\eta'(v)$  frogs at each  $v$ .*

All of the applications so far follow from either of parts (a) and (b) of Theorem 105, monotonicity in the icv and pgf orders, respectively. As (b) is the more difficult to prove, one might wonder why we bother with it. Our interest stems from the role of the pgf order in [HJJ16b] in proving recurrence for the frog model on the binary tree with one frog per site. Our argument there works by showing that the number of visits to the root is stochastically larger than any Poisson distribution in the pgf order. This hinges on [HJJ16b, Lemma 10], which shows that a certain operator is monotone with respect to the pgf order. The proof there is an unsatisfying calculation that cannot easily be extended to a general  $d$ -ary tree. But as we explain in ?THM? ??, this lemma and its analogues for  $d \geq 3$  are now immediate corollaries of Theorem 105(b). We hope this will be helpful in other problems such as establishing recurrence for the frog model on a 3-ary tree.

### 3.7.6 Questions

We will give a few open problems on the theme of comparison theorems. A wider range of problems on the frog model are listed in [HJJ16b, HJJ16a].

We are interested in how sensitive the recurrence of the frog model is to the distribution of the frog counts. We believe that recurrence depends not just on the mean number of frogs at each vertex, but on the entire distribution.

**Open Question 113.** *Give an example where  $r(\eta, S) = \infty$  a.s. and  $r(\eta', S) < \infty$  a.s. with  $\mathbf{E}\eta(v) = \mathbf{E}\eta'(v)$  for all  $v$ .*

Specifically, we would like to know that with simple random walk paths on the binary tree and i.i.d.- $\pi$  frogs per vertex with mean 1, the frog model is transient when  $\pi$  is sufficiently unconcentrated.

Another question of ours is on a stronger version of Corollary 111. In [vdBK93], van den Berg and Kesten prove that in first passage percolation, strictly decreasing the passage time distribution in the icv order yields a strictly smaller time constant (and hence a strictly smaller limiting shape). Most of their work is in establishing the strictness.

**Open Question 114.** *Let  $\mathcal{A}$  and  $\mathcal{A}'$  be the limiting shapes for a frog model on  $\mathbb{Z}^d$  with i.i.d.- $\nu$  and i.i.d.- $\nu'$  initial sleeping frogs per site, respectively. Under what conditions does it hold that  $\nu \preceq_{icv} \nu'$  implies  $\mathcal{A} \subsetneq \mathcal{A}'$ ?*

This cannot hold in full generality, because all choices of  $\nu$  with sufficiently heavy tails have the same limiting shape, the  $L^1$ -ball in  $\mathbb{R}^d$  [AMPR01, Theorem 1.2]. But it might hold under the assumption that  $\nu$  and  $\nu'$  have

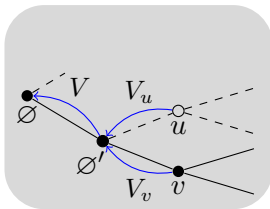


Figure 9:  $V$  is the total number of visits to  $\emptyset$  in the self-similar process,  $V_v$  and  $V_u$  are the number of visits to  $\emptyset'$  from frogs originally in  $\mathbb{T}_2(v)$  and  $\mathbb{T}_2(u)$ , respectively. In the self-similar model  $V, V_v$ , and  $V_u \mid \{u \text{ is visited}\}$  are identically distributed

finite expectations, for example. It might also hold in full generality for the continuous-time frog model, but in this setting the shape theorem has only been proven for one per site initial conditions.

Finally, we are interested in comparing frog models when the graph is altered rather than the initial configuration. As a concrete question in this vein, we ask if the  $d$ -regular tree is the most transient graph in the following sense:

**Open Question 115.** *Suppose the frog model is transient on a  $d$ -regular graph  $G$  with simple random walks. Is it necessarily transient on an infinite  $d$ -regular tree with simple random walk paths and the same initial conditions?*

### 3.8 Recurrence with one-per-site

**Theorem 116.**  $FM(\mathbb{T}_2, 1)$  is recurrent.

#### 3.8.1 Generating function recursion

**Definition 117.** Let  $V$  be the number of visits to the root in the self-similar frog model on  $\mathbb{T}_2$ . Define  $f: [0, 1] \rightarrow [0, 1]$  by  $f(x) = \mathbf{E}x^V$  with the convention that if  $V = \infty$  a.s. then  $f(1) = 0$ .

**Proposition 118.** Define  $\mathcal{A}$ , an operator on functions on  $[0, 1]$ , by

$$\mathcal{A}g(x) = \frac{x+2}{3}g\left(\frac{x+1}{2}\right)^2 + \frac{x+1}{3}g\left(\frac{x}{2}\right)\left(1 - g\left(\frac{x+1}{2}\right)\right). \quad (32)$$

The generating function  $f$  satisfies  $f = \mathcal{A}f$ .



*Proof.* If  $\mathbf{P}[V = \infty] = 1$  then  $f \equiv 0$ . This is easily checked to be a fixed point of  $\mathcal{A}$ . So, for the remainder of the argument suppose that  $\mathbf{P}[V = \infty] < 1$ .

The frog at the root in the self-similar model follows a non-backtracking path and visits one of its children and then one of this child's children; call these vertices  $\varnothing'$  and  $v$ , respectively. We label the yet to be visited child  $u$  (see Figure 9). Define  $V_v$  and  $V_u$  to be the number of frogs which visit  $\varnothing'$  that were originally sleeping in  $\mathbb{T}_2(v)$  and  $\mathbb{T}_2(u)$ , respectively.

Self-similarity guarantees that, since  $v$  has been visited, the random variable  $V_v$  is distributed identically to  $V$ . Conditionally on  $u$  being visited, the random variable  $V_u$  is also distributed identically to  $V$ . In fact, because frogs outside of  $\mathbb{T}_2(u)$  affect  $\mathbb{T}_2(u)$  only by determining whether or not  $u$  is visited, we can express  $V_u$  as  $V_u = \mathbf{1}\{u \text{ is visited}\}V'$ , where  $V'$  is distributed as  $V$  and is independent of  $V_v$ . This yields a description of  $V$  in terms of a pair of independent copies of itself:

$$V = \underbrace{\mathbf{1}\{\text{frog at } \varnothing' \text{ visits } \varnothing\}}_{\text{term 1}} + \underbrace{\mathbf{1}\{u \text{ is visited}\} \text{Bin}(V', \frac{1}{2})}_{\text{term 2}} + \underbrace{\text{Bin}(V_v, \frac{1}{2})}_{\text{term 3}}. \quad (33)$$

Term 1 accounts for a possible visit to  $\varnothing$  by the frog started at  $\varnothing'$ . The conditional binomial distributions in terms 2 and 3 arise because each frog that visits  $\varnothing'$  from  $u$  or  $v$  has a  $\frac{1}{2}$  chance of jumping back to  $\varnothing$ .

Despite the independence between  $V_v$  and  $V'$ , the three terms are dependent. For example, if term 1 is zero, then term 2 is more likely to be nonzero, since the frog at  $\varnothing'$  not visiting  $\varnothing$  makes it more likely to visit  $u$ . We unearth the pairwise independence of  $V_v$  and  $V'$  from (33) by conditioning on the following three disjoint events (see Figure 10):

- A. the frog starting at  $\varnothing'$  visits  $u$ ;
- B. the frog at  $\varnothing'$  does not visit  $u$ , and a frog returns to  $\varnothing'$  through  $v$  and visits  $u$ ;
- C. no frog ever visits  $u$ .

Event  $A$  occurs with probability  $1/3$ . Given that  $k$  frogs return to  $\varnothing'$  through  $v$ , the probability of  $C$  is  $(2/3)2^{-k}$ . Since the number of frogs returning to  $\varnothing'$  through  $v$  is distributed identically to  $V$ , the probability of  $C$  is  $\frac{2}{3}\mathbf{E}\left(\frac{1}{2}\right)^V$ , which we call  $2q/3$ . The probability of event  $B$  is  $2(1-q)/3$ . Note that under our assumption  $\mathbf{P}[V = \infty] < 1$  it follows that  $0 < q < 1$ .

Conditional on event  $A$ ,  $B$ , or  $C$ , the terms in (33) are independent. Indeed, conditioning on whether  $u$  is visited makes terms 2 and 3 independent, and conditioning further on whether the frog at  $\varnothing'$  visits  $u$  then makes term 1

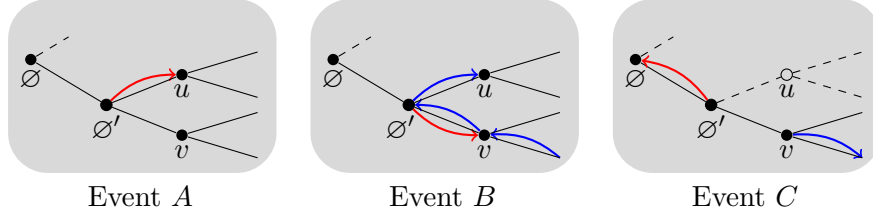


Figure 10: Outcomes that would result in events  $A$ ,  $B$  and  $C$ , respectively. The path of the frog at  $\emptyset'$  is red and the path of a frog from the subtree  $\mathbb{T}(v)$  is blue.

independent of the other two. Now, we describe the distributions of each term in (33) conditional on events  $A$ ,  $B$ , and  $C$ . For a given random variable  $X$ , we use  $\text{Bin}(X, p)$  to denote the random variable  $\sum_{i=1}^X B_i$ , where  $\{B_i\}_{i \in \mathbb{N}}$  are distributed as  $\text{Ber}(p)$ , independent of each other and of  $X$ .

- Conditional on  $A$ , term 1 is 0 and terms 2 and 3 are distributed as independent  $\text{Bin}(V, 1/2)$ .
- Conditional on  $B$ , term 1 is  $\text{Ber}(1/2)$ , term 2 is  $\text{Bin}(V, 1/2)$ , and term 3 is  $\text{Bin}(V, 1/2)$  conditional on being strictly less than  $V$  (since at least one frog will visit  $u$  and not move to  $\emptyset$ ).
- Conditional on  $C$ , term 1 is  $\text{Ber}(1/2)$ , term 2 is 0, and term 3 is  $\text{Bin}(V, 1/2)$  conditional on being equal to  $V$  (since every frog counted by  $V_v$  will return to  $\emptyset$ ).

To summarize, let  $X'$  and  $X$  be distributed as  $\text{Bin}(V, 1/2)$ . Let  $Y$  be distributed as  $\text{Bin}(V, 1/2)$  conditional on  $\text{Bin}(V, 1/2) < V$ . Let  $Z$  be distributed as  $\text{Bin}(V, 1/2)$  conditional on  $\text{Bin}(V, 1/2) = V$ . Let  $I \sim \text{Ber}(1/2)$ . Take all of these to be independent. Conditioning on events  $A$ ,  $B$ , and  $C$ , equation (33) yields

$$V \stackrel{d}{=} \begin{cases} X' + X & \text{with probability } 1/3, \\ I + X' + Y & \text{with probability } 2(1 - q)/3, \\ I + Z & \text{with probability } 2q/3. \end{cases} \quad (34)$$

From this description of the distribution of  $V$ ,

$$\begin{aligned}\mathbf{E}x^V &= \frac{1}{3}\mathbf{E}x^{X'+X} + \frac{2(1-q)}{3}\mathbf{E}x^{I+X'+Y} + \frac{2q}{3}\mathbf{E}x^{I+Z} \\ &= \frac{1}{3}\mathbf{E}x^{X'}\mathbf{E}x^X + \frac{2(1-q)}{3}\mathbf{E}x^I\mathbf{E}x^{X'}\mathbf{E}x^Y + \frac{2q}{3}\mathbf{E}x^I\mathbf{E}x^Z.\end{aligned}\quad (35)$$

Recall that a  $\text{Ber}(p)$  random variable has generating function  $px + 1 - p$  and that a random sum of i.i.d. random variables,  $\sum_1^N X_i$ , has generating function  $g_N(g_{X_1}(x))$ , where  $g_N$  and  $g_{X_1}$  are the generating functions of  $N$  and  $X_1$ . From these facts,

$$\begin{aligned}\mathbf{E}x^I &= \frac{x+1}{2}, \\ \mathbf{E}x^{X'} &= \mathbf{E}x^X = f\left(\frac{x+1}{2}\right).\end{aligned}$$

The generating functions  $\mathbf{E}x^Y$  and  $\mathbf{E}x^Z$  are a bit more complicated. The random variable  $Y$  is distributed as  $X$  conditional on  $X < V$ . Using the basic formula for conditional probability,

$$\begin{aligned}\mathbf{P}[Y = k] &= \mathbf{P}[X = k \mid X < V] = \frac{\mathbf{P}[X = k \text{ and } X < V]}{\mathbf{P}[X < V]} \\ &= \frac{\mathbf{P}[X = k] - \mathbf{P}[X = V = k]}{1 - q} \\ &= \frac{\mathbf{P}[X = k] - 2^{-k}\mathbf{P}[V = k]}{1 - q}.\end{aligned}$$

Thus, the probability generating function of  $Y$  is

$$\begin{aligned}\mathbf{E}x^Y &= \frac{1}{1-q} \sum_{k=0}^{\infty} x^k (\mathbf{P}[X = k] - 2^{-k}\mathbf{P}[V = k]) \\ &= \frac{1}{1-q} \mathbf{E}\left[x^X - \left(\frac{x}{2}\right)^V\right] \\ &= \frac{1}{1-q} \left(f\left(\frac{x+1}{2}\right) - f\left(\frac{x}{2}\right)\right).\end{aligned}\quad (36)$$

In (36) we are making use of the general fact that  $\sum(a_n - b_n) = \sum a_n - \sum b_n$  so long as each sum is finite. Similarly,

$$\mathbf{P}[Z = k] = \mathbf{P}[X = k \mid X = V] = \frac{2^{-k}\mathbf{P}[V = k]}{q},$$

and so

$$\mathbf{E}x^Z = \frac{1}{q} \sum_{k=0}^{\infty} x^k 2^{-k} \mathbf{P}[V = k] = \frac{1}{q} f\left(\frac{x}{2}\right).$$

Using all of these generating functions and (35)

$$\begin{aligned} f(x) &= \frac{1}{3} \mathbf{E}x^{X'} \mathbf{E}x^X + \frac{2(1-q)}{3} \mathbf{E}x^I \mathbf{E}x^{X'} \mathbf{E}x^Y + \frac{2q}{3} \mathbf{E}x^I \mathbf{E}x^Z \\ &= \frac{1}{3} f\left(\frac{x+1}{2}\right)^2 + \frac{2(1-q)}{3} \left(\frac{x+1}{2} f\left(\frac{x+1}{2}\right) \frac{1}{1-q} \left(f\left(\frac{x+1}{2}\right) - f\left(\frac{x}{2}\right)\right)\right) \\ &\quad + \frac{2q}{3} \left(\frac{x+1}{2q} f\left(\frac{x}{2}\right)\right) \\ &= \frac{x+2}{3} f\left(\frac{x+1}{2}\right)^2 - \frac{x+1}{3} f\left(\frac{x+1}{2}\right) f\left(\frac{x}{2}\right) + \frac{x+1}{3} f\left(\frac{x}{2}\right) = \mathcal{A}f(x), \end{aligned}$$

which establishes our claim.  $\square$

### 3.8.2 Proving recurrence

We have reduced the problem to understanding the properties of the operator  $\mathcal{A}$  defined in (32). In Lemma 119, we prove that  $\mathcal{A}$  is monotonic for functions belonging to the set  $\mathcal{S} = \{g: [0, 1] \rightarrow [0, 1], \text{ nondecreasing}\}$ . In Lemma 120, we show that  $\mathcal{A}$  maps  $\mathcal{S}$  into itself, so that we can apply Lemma 119 after applying  $\mathcal{A}$  iteratively. Finally, we show in Lemmas 121 and 123 that  $\mathcal{A}^n 1 \rightarrow 0$ . Starting at the conclusion of Proposition 118 (that the generating function  $f$  is a fixed point of  $\mathcal{A}$ ), we will then apply these results to show that  $f \equiv 0$ , thus proving that the number of visits to the root in the self-similar frog model is a.s. infinite.

**Lemma 119.** *Let  $g, h \in \mathcal{S}$ . If  $g \leq h$ , then  $\mathcal{A}g \leq \mathcal{A}h$ .*

*Proof.* This follows from Theorem 105(b). The lemma is a monotonicity result for an operator  $\mathcal{A}$  acting on probability distributions on the nonnegative integers. The operator can be defined as follows. Let  $\pi$  be a probability distribution on the nonnegative integers. Consider a binary tree truncated to four vertices as in Figure 11. Place one frog on  $\emptyset$  and one frog on  $\emptyset'$ , and independently sample from  $\pi$  to decide the number of frogs on  $u$  and  $v$ . The frog paths are random nonbacktracking walks stopped when a frog reaches a leaf. Now, run the frog model starting with the frog at  $\emptyset$  active until all frogs are stopped. Define  $\mathcal{A}\pi$  to be the distribution of frogs terminating at  $\emptyset$ . In [HJJ16b], the operator was defined in a different way, but it turns out

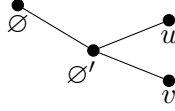


Figure 11: A graph used to define the operator  $\mathcal{A}$  in [HJJ16b].

to be equivalent. See also [HJJ16a, Section 2.2] and [JJ16a, Section 3.1.2] for similar constructions.

As visits to the root are a pgf statistic,  $\pi \preceq_{\text{pgf}} \pi'$  implies that  $\mathcal{A}\pi \preceq_{\text{pgf}} \mathcal{A}\pi'$ . As  $\mathcal{A}\pi$  is the distribution of visits to  $\emptyset$  in the frog model on the truncated graph, it is a continuous pgf statistic by Proposition 106. Thus the lemma is a consequence of Theorem 105(b).  $\square$

*Original proof.* For  $0 \leq t \leq 1$  define the interpolation between  $g$  and  $h$  by

$$i_t(x) = (1-t) \cdot g(x) + t \cdot h(x).$$

Since  $\mathcal{A}i_0 = \mathcal{A}g$  and  $\mathcal{A}i_1 = \mathcal{A}h$  it suffices to prove that  $\frac{d}{dt}\mathcal{A}i_t(x) \geq 0$ . Fix  $x$  and set  $a = i_t(\frac{x+1}{2})$  and  $b = i_t(\frac{x}{2})$  so that

$$\mathcal{A}i_t(x) = \frac{2+x}{3}a^2 + \frac{1+x}{3}b(1-a).$$

Define  $s(a, b) = \mathcal{A}i_t(x)$ . The chain rule implies

$$\frac{d}{dt}\mathcal{A}i_t(x) = \frac{\partial}{\partial a}s(a, b)\frac{da}{dt} + \frac{\partial}{\partial b}s(a, b)\frac{db}{dt}.$$

To prove  $\frac{d}{dt}\mathcal{A}i_t \geq 0$  it suffices to prove each term in the above formula is nonnegative.

- The assumption that  $g \leq h$  implies that  $\frac{d}{dt}i_t(x) = h(x) - g(x) \geq 0$  for all  $t$  and  $x$ . In particular, this implies  $\frac{da}{dt}, \frac{db}{dt} \geq 0$ .
- First we compute the partials

$$\frac{\partial}{\partial a}s(a, b) = 2a\frac{2+x}{3} - b\frac{1+x}{3} \quad \text{and} \quad \frac{\partial}{\partial b}s(a, b) = (1-a)\frac{1+x}{3}.$$

As  $g$  and  $h$  are nondecreasing,  $i_t$  is also nondecreasing in  $x$  for any fixed  $t$ . Hence  $b \leq a$ . Along with the bound  $a \leq 1$  this immediately implies both partials are positive.  $\square$

**Lemma 120.** *If  $g \in \mathcal{S}$ , then  $\mathcal{A}g \in \mathcal{S}$ .*

*Proof.* All summands in (32) are nonnegative when  $g(x) \leq 1$ , which implies that  $\mathcal{A}g \geq 0$ . By the previous lemma,  $\mathcal{A}g \leq \mathcal{A}1 \leq 1$ . We can conclude then that  $0 \leq \mathcal{A}g \leq 1$ . To see that  $\mathcal{A}g$  is nondecreasing, suppose that  $x \leq y$ , and let  $a = g\left(\frac{y+1}{2}\right) - g\left(\frac{x+1}{2}\right)$ . Then we have

$$\begin{aligned} \mathcal{A}g(y) &\geq \frac{x+2}{3}g\left(\frac{x+1}{2}\right)g\left(\frac{y+1}{2}\right) + \frac{x+1}{3}g\left(\frac{x}{2}\right)\left(1 - g\left(\frac{y+1}{2}\right)\right) \\ &= \mathcal{A}g(x) + \left(\frac{x+2}{3}g\left(\frac{x+1}{2}\right) - \frac{x+1}{3}g\left(\frac{x}{2}\right)\right)a \geq \mathcal{A}g(x). \square \end{aligned}$$

We now analyze the behavior of  $\mathcal{A}$  on the family of generating functions for Poisson random variables. Recall that the generating function of  $\text{Poi}(a)$  is  $e^{a(x-1)}$ .

**Lemma 121.** *Define  $g_a(x) = e^{a(x-1)}$  for all  $a \geq 0$ . For all  $x \in [0, 1]$ ,*

$$\mathcal{A}g_a(x) \leq g_{a+c_a}(x),$$

where

$$c_a = \begin{cases} \frac{1}{3}e^{-2} & 0 \leq a \leq 4, \\ \frac{1}{3}e^{-a/2} & a \geq 4. \end{cases} \quad (37)$$

*Proof.* Applying the operator  $\mathcal{A}$ , we have

$$\begin{aligned} \mathcal{A}g_a(x) &= \frac{x+2}{3}e^{a(x-1)} + \frac{x+1}{3}e^{ax/2-a}\left(1 - e^{a(x-1)/2}\right) \\ &= g_a(x)r_{a/2}(x), \end{aligned} \quad (38)$$

where

$$r_b(x) = \frac{2+x}{3} + \frac{1+x}{3}\left(e^{-bx} - e^{-b}\right).$$

Note that  $g_a(x)g_b(x) = g_{a+b}(x)$ . It thus suffices to establish

**Claim.** *For  $x \in [0, 1]$ , we have  $r_b(x) \leq g_{c_{2b}}(x)$ .*

*Proof of claim.* We drop subscripts and let  $r(x) = r_b(x)$  and  $c = c_{2b}$ . Calculus and a little algebra show that

$$r'(x) = \frac{1}{3}\left(1 - e^{-b} + e^{-bx}(-bx - b + 1)\right)$$

and

$$r''(x) = \frac{1}{3}e^{-bx} (b^2(x+1) - 2b).$$

We break the proof up into cases.

- If  $b \leq 1$  then  $r(x)$  is concave down on  $[0, 1]$  and the graph of  $r(x)$  lies below its tangent line at  $x = 1$ . Thus

$$\begin{aligned} r(x) &\leq 1 + r'(1)(x-1) = 1 + \frac{1}{3}[1 - 2be^{-b}](x-1) \\ &\leq \exp\left[\frac{1}{3}(1 - 2be^{-b})(x-1)\right]. \end{aligned}$$

It is easily verified that  $\frac{1}{3}(1 - 2be^{-b}) \geq \frac{1}{3}(1 - 2e^{-1}) \geq e^{-2}/3$  for  $b \leq 1$  and hence that  $r(x) \leq g_c(x)$ .

- If  $b \geq 2$  then  $r(x)$  is concave up on  $[0, 1]$  and the graph of  $r(x)$  lies below the secant line between  $(0, r(0))$  and  $(1, r(1))$ . Thus as  $r(1) = 1$  we have

$$\begin{aligned} r(x) &\leq 1 + (1 - r(0))(x-1) = 1 + \frac{1}{3}e^{-b}(x-1) \\ &\leq \exp\left[\frac{1}{3}e^{-b}(x-1)\right] = g_c(x). \end{aligned}$$

- If  $1 < b < 2$  then there is a unique inflection point at  $I = \frac{2}{b} - 1$  where  $r$  switches from concave down to concave up. Since  $r$  is concave up on  $[I, 1]$ , the graph of  $r$  lies below the line connecting  $(1, 1)$  to  $(I, r(I))$ . Since  $r$  is concave down on  $[0, I]$ , to the left of  $I$  the graph of  $r$  lies below its tangent line at  $(I, r(I))$ . Thus the line segment from  $(I, r(I))$  to  $(0, r(I) - r'(I)I)$  lies above  $r$ , as in Figure 12. Therefore  $r$  lies below the line between  $(1, 1)$  and  $(0, r(I) - r'(I)I)$ , and

$$r(x) \leq 1 + (1 - r(I) + Ir'(I))(x-1). \quad (39)$$

Next, we evaluate

$$1 - r(I) + Ir'(I) = 1 - \frac{1}{3}\left(2 + \left(\frac{4}{b} - 1\right)e^{b-2} - e^{-b}\right) \quad (40)$$

and try to bound this expression from below for  $b \in (1, 2)$ . We proceed as calculus students, looking for critical points in this interval. The derivative with respect to  $b$  is

$$-\frac{1}{3}\left(\left(\frac{4}{b} - 1 - \frac{4}{b^2}\right)e^{b-2} + e^{-b}\right),$$

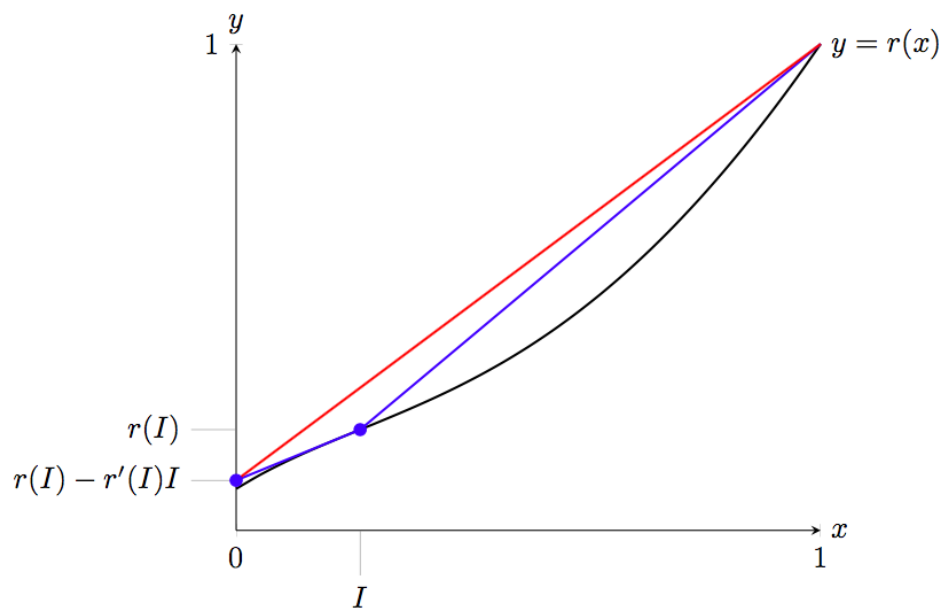


Figure 12: Above the graph of  $y = r(x)$  sits the secant line from  $(I, r(I))$  to  $(1, 1)$  and the tangent line to  $r(x)$  at  $x = I$ , both depicted in blue. Above them in red is the line  $y = 1 + (1 - r(I) + Ir'(I))(x - 1)$ .



and a bit of algebra shows that the zeros of this expression are the solutions to

$$e^{2(b-1)} \left( \frac{2-b}{b} \right)^2 = 1.$$

Taking logarithms, we are interested in solutions to

$$b - 1 + \log(2 - b) - \log b = 0.$$

on  $(1, 2)$ . On this interval we can replace the logarithms with their power series expansions around 1 to rewrite the left-hand side as

$$b - 1 + 2 \left( \frac{(b-1)^2}{2} + \frac{(b-1)^4}{4} + \frac{(b-1)^6}{6} + \dots \right),$$

which is strictly positive for  $b \in (1, 2)$ . Thus (40) has no critical values on  $(1, 2)$ , and its minimum on  $[1, 2]$  is  $e^{-2}/3$ , occurring at  $b = 2$ . Applying this to (39), we have shown that

$$r(x) \leq 1 + \frac{1}{3}e^{-2}(x-1) \leq \exp \left[ \frac{1}{3}e^{-2}(x-1) \right] = g_c(x).$$

This concludes the proof of both the claim and the lemma.  $\square$

*Remark 122.* Though the preceding lemma was an exercise in calculus, it has a probabilistic interpretation. If we think of  $\mathcal{A}$  as acting directly on distributions instead of on their generating functions, this lemma shows that the result of applying  $\mathcal{A}$  to  $\text{Poi}(a)$  is larger than  $\text{Poi}(a + c_a)$  in the probability generating function stochastic order described at the beginning of Section ???. The reason that  $\mathcal{A}g_a$  simplifies so nicely in (38) is the Poisson thinning property, and the fact that  $g_a(x)g_b(x) = g_{a+b}(x)$  is just the statement that the sum of independent Poissons is Poisson. There is a temptation to interpret  $\mathcal{A}g_a(x) = g_a(x)r_{a/2}(x)$  as saying that the distribution resulting from applying  $\mathcal{A}$  to  $\text{Poi}(a)$  is a convolution of  $\text{Poi}(a)$  and another distribution, but  $r_{a/2}(x)$  is not monotone in  $x$  and hence not the generating function of a probability distribution.

**Lemma 123.** For  $x \in [0, 1)$ ,

$$\lim_{n \rightarrow \infty} \mathcal{A}^n g_0(x) = 0.$$

*Proof.* Define the sequence  $a_n$  by  $a_0 = 0$  and  $a_{n+1} = a_n + c_{a_n}$ . By Lemmas 119, 120, and 121,

$$\mathcal{A}^n g_0(x) \leq g_{a_n}(x) = e^{a_n(x-1)}.$$

We need to show that  $a_n \rightarrow \infty$  as  $n \rightarrow \infty$ . Suppose this does not hold. Since the sequence is increasing,  $a_n \rightarrow a$  for some constant  $a$ . Looking back at (37), this implies that  $c_{a_n}$  converges to a strictly positive limit. We can then choose  $n$  sufficiently large that  $a_n + c_{a_n} > a$ , a contradiction.  $\square$

*Proof of Theorem 116.* Let  $f$  be the generating function  $f(x) = \mathbf{E}x^V$  with  $V$  the number of visits to the root in the self-similar model frog model on the binary tree. By Proposition 118 we know that  $f$  satisfies the recursion relation  $\mathcal{A}f = f$ . Since  $f$  is a probability generating function, it satisfies  $f(x) \leq 1 = g_0(x)$  for  $x \in [0, 1]$ . Proposition 118 and Lemmas 119 and 120 imply  $f(x) \leq \mathcal{A}^n g_0(x)$  for all  $n$ . By Lemma 123,  $f$  is identically zero on  $[0, 1)$ . Thus the probability of any finite number of returns to the root is 0. This implies there are a.s. infinitely many returns to the root in the self-similar model. By the coupling to the original frog model, each return in the self-similar model corresponds to a distinct return in the frog model. So, the frog model on the binary tree is a.s. recurrent.  $\square$

### 3.9 Shape Theorems

First let's talk a bit how to prove a shape theorem. Some models with shape theorems are first passage percolation, internal diffusion limited aggregation. DLA and oil and water models also ought to have limiting shapes, but this isn't known.

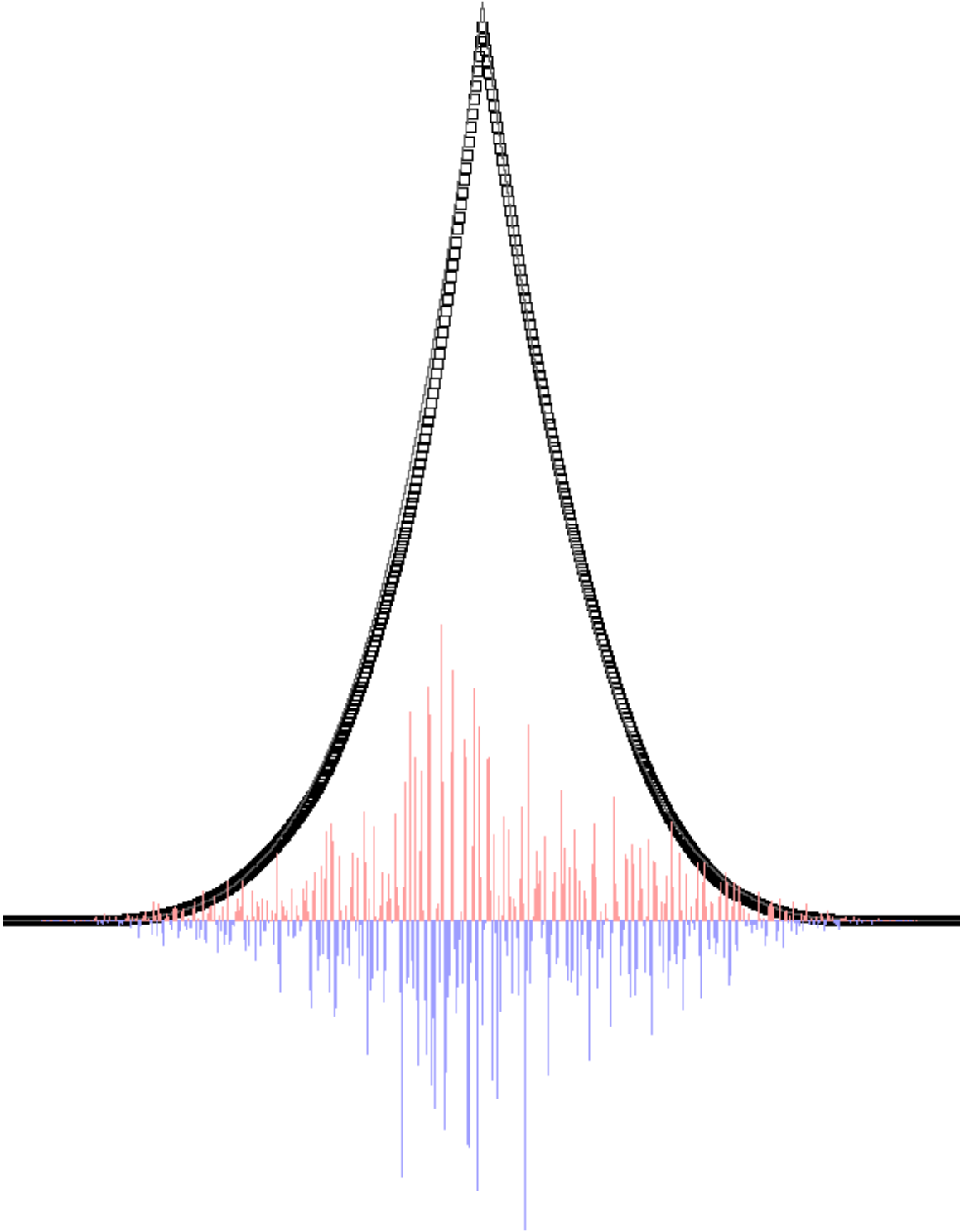


Figure 1: Graph of the odometer function of the oil and water model in  $\mathbb{Z}$  with  $n = 360000$  particles of each type started at the origin: for each  $x \in \mathbb{Z}$  a box is drawn centered at  $(x, u(x))$  where  $u(x)$  is the number of oil-water pairs fired from  $x$ . The curve  $f(x) = \frac{1}{72\pi}((18\pi n)^{1/3} - |x|)^4$  appears in gray. Red and blue vertical bars represent the final configuration of oil and water particles respectively; the height of the bar is proportional to the number of particles.

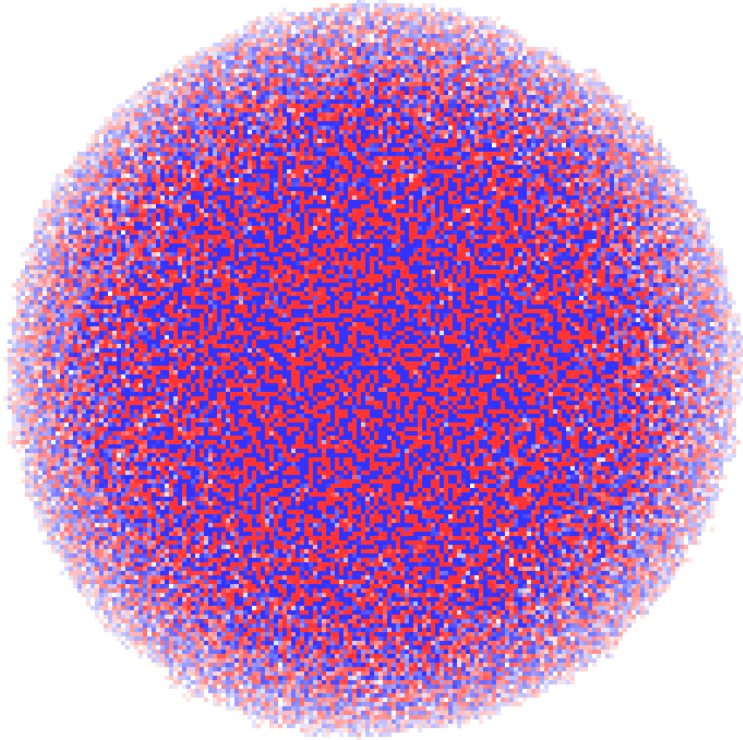


Figure 2: Oil and water in  $\mathbb{Z}^2$  with  $n = 2^{22}$  particles of each type started at the origin. Each site where particles stop is shaded red or blue according to whether oil or water particles stopped there. The intensity of the shade indicates the number of particles. We believe that the limit shape is a disk of radius of order  $n^{1/4}$ .

The oil and water model can be defined on any graph and in particular on higher-dimensional lattices. Figure 2 shows an oil and water configuration in  $\mathbb{Z}^2$ . In Section 7 we conjecture the relevant exponents in  $\mathbb{Z}^d$  for  $d \geq 2$ .

### 1.1 Related models: internal DLA and abelian networks

In *internal DLA*, each of  $n$  particles started at the origin in  $\mathbb{Z}^d$  performs a simple random walk until reaching an unoccupied site. The resulting random set of  $n$  occupied sites is close to a Euclidean ball [14]. Internal DLA is one of several models known to have an *abelian property* according to which the order of moves does not affect the final outcome.

Dhar [7] proposed certain collections of communicating finite automata as a broader class of models with this property. Until recently the only examples studied in this class have been *unary networks* (or their “block renormalizations” as proposed in [7]). Informally, a unary network is a system of local rules for moving *indistinguishable* particles on a graph, whereas a non-unary network has different types of particles. It is not as easy to construct non-unary examples with an abelian property, but they exist. Alcaraz, Pyatov and Rittenberg studied a class of non-unary examples which they termed *two-component sandpile models* [1], and asked whether there is a nontrivial example with two particle species such that the total number of particles of each type is conserved. Oil and water has this conservation property,

The main tool for proving shape theorems is the sub-additive ergodic theorem.

**Theorem 124** (Sub-additive ergodic theorem). *Suppose that  $Y(m, n)$  is a collection of positive random variables indexed by integers satisfying  $0 < m < n$  such that:*

1.  $Y(0, n) \leq Y(0, m) + Y(m, n)$  for all  $0 < m < n$  (subadditivity).
2. The joint distribution of  $\{Y(m+1, m+k+1), k \geq 1\}$  is the same as that of  $\{Y(m, m+k), k > 1\}$  for each  $m > 0$ .
3. For each  $k > 1$  the sequence of random variables  $\{Y(nk, (n+1)k), n \geq 1\}$  is a stationary ergodic process.
4.  $\mathbf{E}Y(0, 1) < \infty$ .

Then

$$\lim_{n \rightarrow \infty} \frac{Y(0, n)}{n} = \inf_{n \geq 0} \frac{\mathbf{E}Y(0, n)}{n} := \gamma.$$

Let  $T(x, y)$  be the time for a frog model with a particle awakened at  $x$  to reach  $y$ . To prove a shape theorem we show that the hypotheses of Theorem 3.1 hold for  $Y(m, n) = T(mx, nx)$ , for each fixed  $x \in \mathbb{Z}^d$ .

The first three are easy to establish. The fourth is a bit tricky. We include the page from [AMP02a] that details them. Recall that a dynamical system  $(X, \mathcal{A}, \mu, T)$  be a dynamical system, with  $T$  being the time-evolution or shift operator, is said to be strong mixing if, for any  $A, B \in \mathcal{A}$ , one has

$$\lim_{n \rightarrow \infty} \mu(A \cap T^{-n}B) = \mu(A)\mu(B).$$

Essentially this says in the time-shifted limit events are independent. It is a fact that strongly mixing systems are ergodic.

First of all observe that the set of variables  $\{T(x, y) : x, y \in \mathbb{Z}^d\}$  defined in Section 1 is *subadditive* in the sense that

$$(3.1) \quad T(x, z) \leq T(x, y) + T(y, z)$$

for all  $x, y, z \in \mathbb{Z}^d$  and *all* the realizations of the random variables  $S_n^x$ . Indeed, if site  $z$  is reached before site  $y$ , then (3.1) is evident. If that does not happen, recall that the random variables  $T(y, z), y, z \in \mathbb{Z}^d$  are constructed using the same collection of the random variables  $S_n^x$ ; that is, each particle follows the same trajectory as soon as it wakes up. So the process departing from only site  $y$  awakened [the one which gives the passage time  $T(y, z)$ ] is coupled with the original process (i.e., that started from  $x$ ), and for the latter one may have other particles awakened at time  $T(x, y)$  besides that from  $y$ . Consequently, from (1.1) it follows that  $T(x, z) - T(x, y)$ , which is the remaining time to reach site  $z$  for the original process, is less than or equal to  $T(y, z)$ , thus proving (3.1).

The second hypothesis, as well as the fact that for fixed  $x \in \mathbb{Z}^d$  and  $k \in \mathbb{N}$  the sequence  $\{T((n - 1)kx, nkx), n \geq 1\}$  is *stationary*, immediately follow from the definition. *Ergodicity* holds because the sequence of random variables  $\{T((n - 1)kx, nkx), n \geq 1\}$  is strongly mixing. That can be checked easily because the events  $\{T(n_1kx, (n_1 + 1)kx) = a\}$  and  $\{T(n_2kx, (n_2 + 1)kx) = b\}$  are independent provided that  $a + b < \|(n_1 - n_2)kx\|_1$ .

It is simple to see that the fourth hypothesis holds when  $d = 1$ . To see that remember that for  $\tau =$  *the first return to the origin of a SRW*, we can assure that  $\mathbf{P}[\tau > t] \leq Ct^{-1/2}$ . Besides that, in a random time with exponential tail we will have at least three awakened particles jumping independently in the frog model. Combining these two facts we have that  $\mathbf{E}T(0, 1) < \infty$ . So, for  $d = 1$  one gets  $T(0, n)/n \rightarrow \gamma$  a.s., and consequently we have the proof of Theorem 1.1 with  $A = [-\gamma^{-1}, \gamma^{-1}]$  in dimension 1. Thus, from now on we assume that  $d \geq 2$ .

To take care of the fourth hypothesis in dimension  $d \geq 2$ , we need the following result.

**THEOREM 3.2.** *For all  $d \geq 2$  and  $x_0 \in \mathbb{Z}^d$  there exist positive finite constants  $C = C(x_0, d)$  and  $\beta = \beta(d)$  such that*

$$\mathbf{P}[T(0, x_0) \geq n] \leq C \exp\{-n^\beta\}$$

for all  $n$ .

**PROOF.** We begin by considering the case  $d \geq 3$ . Pick  $n \geq \|x_0\|^2$ . Fix some  $0 < \varepsilon < 1$  (to be chosen later). Denote for  $1 \leq i \leq \lfloor d/2 \rfloor$ ,

$$D_{i,\varepsilon} := \{x \in \mathbb{Z}^d : \|x\| \leq in^{1/2+\varepsilon}\},$$

and define the event

$$(3.2) \quad A_1 := A_1(n, \varepsilon) := \{|\mathbf{R}_n^0 \cap D_{1,\varepsilon}| \geq r_1 n^{1-\varepsilon}\},$$

where  $r_1 = r_1(d)$  is a positive constant to be chosen later.

To prove that super-polynomial tail bound on  $T(0, x_0)$  we break up  $\mathbb{Z}^d$  into annuli and show that a positive fraction of each gets woken with high probability. Once we are to the annuli including  $x_0$  it is overwhelmingly likely that  $x_0$  be visited.

### 3.9.1 Obtaining a shape theorem

We now let  $T(x) = T(0, x)$  for  $x \in \mathbb{Z}^d$ . The subadditive ergodic theorem guarantees that  $T(x)/n \rightarrow \mu(x)$ . We can extend  $T(x, y)$  to  $\mathbb{R}^d \times \mathbb{R}^d$  by letting it be the first time the frog at the  $(-1/2, 1/2]^d$  box containing  $x$  enters the site at the box containing  $y$ . We want to then extend  $\mu$  to  $\mathbb{R}^d$ . This is done by first considering  $x \in \mathbb{Q}^d$  and observing that  $T(nx)/n \rightarrow \mu(m_x x)/m_x := \mu(x)$  with  $m_x$  the smallest integer such that  $m_x x \in \mathbb{Z}^d$ . Now we have a norm on  $\mathbb{Q}^d$  that is uniformly continuous (since all norms on fin. dim. spaces are equivalent). We can thus extend  $\mu$  to all of  $\mathbb{R}^d$ .

It remains to prove that the frog model expands at least linearly. This follows from a the super-polynomial tail bound on  $T(0, x_0)$ . By partitioning the path to  $x_0$  into iid steps of size 1, we show that each step is unlikely to take more than  $\sqrt{\|x_0\|_1}/2$  steps. Standard tail estimates imply that the sum of  $\|x_0\|_1$  of these steps is unlikely to be larger than  $\|x_0\|_1$ . From here we can use a union bound on a disk of radius  $\delta n$  to deduce that  $\bar{\xi}_n$  contains  $D(0, \delta n)$  whp for some  $\delta \in (0, 1)$ .

The limiting shape is then  $A = \{x \in \mathbb{R}^d: \mu(x) \leq 1\}$ .

### 3.9.2 icv and pgf monotonicity

We end with the proof about monotonicity of the shape when augmenting the number of frogs per site in the pgf and icv stochastic orders. We will assume the following proposition holds.

**Proposition 125.** *For  $t \in \mathbb{N} \cup \{\infty\}$  and any nonroot vertex  $u$ , let  $a_{t,u}(\eta, S)$  be an indicator on site  $u$  being visited in the frog model  $(\eta, S)$  by time  $t$ . Let  $a_t(\eta, S)$  be the total number of sites visited by time  $t$ . Then both  $a_{t,u}$  and  $a_t$  are continuous icv and pgf statistics.*

*Proof of Corollary 111.* For  $v \in \mathbb{Z}^d$ , let  $T(v)$  and  $T'(v)$  be the time that the vertex  $v$  is activated for the frog models with  $\eta$  and  $\eta'$  frogs per site, respectively. Let  $a_{t,v} = \mathbf{1}\{T(v) \leq t\}$  be an indicator that  $v$  has been activated by time  $t$ , and similarly for  $a'_{t,v}$ . By Proposition 125,  $a_{t,v}$  and  $a'_{t,v}$  are continuous icv and pgf statistics. Moreover, we can express  $T(v)$  and

$T'(v)$  in terms of these statistics:

$$T(v) = \sum_{t=0}^{\infty} (1 - a_{t,v}), \quad T'(v) = \sum_{t=0}^{\infty} (1 - a'_{t,v}).$$

By Theorem 105, we have  $a_{t,v} \preceq_{\text{icv}} a'_{t,v}$ , and hence  $\mathbf{E}a_{t,v} \leq \mathbf{E}a'_{t,v}$ . Apply this, along with Fubini's theorem, to the expressions for  $T(v)$  and  $T'(v)$  to obtain

$$\mathbf{E}T'(v) \leq \mathbf{E}T(v). \quad (41)$$

As in the proof of [AMPR01, Theorem 1.1], the limiting shapes are determined by functions  $\mu$  and  $\mu'$  with domain  $\mathbb{R}^d$ . The function  $\mu$  is obtained via Kingman's subadditive ergodic theorem by defining

$$\mu(v) = \lim_{n \rightarrow \infty} \frac{T(nv)}{n} = \inf_{n \geq 1} \frac{\mathbf{E}T(nv)}{n}, \quad (42)$$

where  $v \in \mathbb{Z}^d$ . After interpolating to all of  $\mathbb{R}^d$ , the limiting shape is given by  $\mathcal{A} = \{x \in \mathbb{R}^d : \mu(x) \leq 1\}$ . The set  $\mathcal{A}'$  is obtained in the same fashion. To deduce that  $\mathcal{A} \subseteq \mathcal{A}'$ , it then suffices to show that  $\mu'(v) \leq \mu(v)$  for  $v \in \mathbb{Z}^d$ . This follows from (41) applied to the expected value formulation of  $\mu$  at (42).  $\square$

### 3.10 Frog wakeup time on finite $d$ -ary trees

Let  $\mathcal{C} = \mathcal{C}_d^n$  be the first time all frogs are awake in the frog model on the finite tree  $\mathbb{T}_d^n$ . We call this the cover time since  $\mathcal{C}$  occurs precisely when every vertex of  $\mathbb{T}_d^n$  has been visited by a frog. A naive upper bound is  $\mathbf{E}[\mathcal{C}] = O(n^2 d^n)$ , the cover time for a single random walk [Ald91], while the diameter gives a naive lower bound of  $\mathbf{E}[\mathcal{C}] = \Omega(n)$ . The recent paper [Her16] proves stronger results than we present here for the case that there are a Poisson distributed frogs at each site. We prove an exponential speedup over the naive upper bound.

**Theorem 126.** *For each fixed  $d$  and  $\epsilon > 0$ ,*

$$\mathbf{E}[\mathcal{C}] = O(n^{4+\epsilon}(d/\sqrt{m})^n),$$

with  $m = \frac{2d}{d+1}$ .

In Theorem 126, we show that the cover time is  $O(n^{4+\epsilon}(d/\sqrt{m})^n)$  for a constant  $m > 1$ , exponentially faster than the previously available bound of  $O(n^2 d^n)$ . Our proof boils down to two steps.



**Wake-up phase** Prove that a large number of frogs are awake after a controllable amount of time.

**Clean-up phase** Bound the remaining time for these random walks to cover  $\mathbb{T}_d^n$ .

We address the first step in Sections 3.10.1 and 3.10.2. The main idea is to bound the growth of frogs from below by a Galton-Watson tree and then apply lower tail deviation estimates from [FW07]. We complete the second step and the proof in Section 3.10.3. The key ingredient here is Proposition 134, which states that  $k$  walkers on a  $d$ -ary tree cover the tree  $\sqrt{k}$  times faster than a single walker. This proposition closely follows [Sau10, Theorem 1.2], and we have held it off to the appendix.

### 3.10.1 A Galton-Watson process

We show that a large number of frogs are awakened in a controllable amount of time by embedding a Galton-Watson tree in the activated sites of the frog model. First, we define the Galton-Watson tree and collect a few facts about it. Let  $Z = (Z_n, n \geq 0)$  be a Galton-Watson process where a parent has either one child with probability  $\frac{2}{d+1}$  or two children with probability  $\frac{d-1}{d+1}$ . Let  $m = \mathbf{E}Z_1 = \frac{2d}{d+1}$ . We will use the results of [FW07] to give a lower tail bound on this process:

**Proposition 127.** *Let  $\alpha = \log_m(\frac{d+1}{2})$ . Fix any  $\epsilon > 0$ . There exists a constant  $B_\epsilon$  depending on  $\epsilon$  so that for all  $n \geq 0$  and  $k$  satisfying  $n \leq k \leq m^n/(n+1)^\epsilon$ ,*

$$\mathbf{P}[Z_n = k] \leq B_\epsilon k^{\alpha-1} m^{-n\alpha} \quad (127.1)$$

*Additionally, there is  $B > 0$  so that for all  $n \geq 0$*

$$\mathbf{P}[Z_n < n] \leq B n^\alpha m^{-n\alpha}. \quad (127.2)$$

*Proof.* Inequality (127.1) follows from Theorem 4 of [FW07] but requires some care with the asymptotics. Define  $k_n$  as the integer between  $n$  and  $m^n/(n+1)^\epsilon$  that maximizes

$$\frac{\mathbf{P}[Z_n = k_n]}{k_n^{\alpha-1} m^{-n\alpha}}.$$

Then  $k_n$  is a sequence satisfying  $k_n = o(m^n)$  and  $k_n \rightarrow \infty$ , and Theorem 4 from [FW07] shows that

$$\mathbf{P}[Z_n = k_n] = m^{-n} w(k_n/m^n)(1 + o(1)),$$

where  $w(x)$  is the density of the almost-sure limit of  $Z_n/m^n$ . This density is shown in [Dub71, Theorem 1] to satisfy

$$w(x) \leq Cx^{\alpha-1} \tag{43}$$

on  $0 < x < 1$ , for some constant  $C$ . Thus  $\mathbf{P}[Z_n = k_n] \leq Ck_n^{\alpha-1}m^{-n\alpha}(1+o(1))$ . Now choose  $B_\epsilon$  larger than the  $C(1+o(1))$  factor to obtain (127.1). Inequality (127.2) follows directly from Theorem 4 of [FW07] along with (43).  $\square$

### 3.10.2 Embedding $Z$ in the frog model

Now that we have introduced the process  $Z$  and bounded its left tail, we explain how the first  $n - 1$  generations of  $Z$  count a subset of the vertices visited by the frog model on  $\mathbb{T}_d^n$ . The results of this section are summarized in the following proposition.

**Proposition 128.** *The Galton-Watson process  $Z$  can be coupled with the frog model on  $\mathbb{T}_d^n$  so that the following holds:*

- (i) *After  $\mathcal{C}_{\text{wakeup}}$  steps (defined in section 3.10.2.1), the set of visited vertices in the frog model on  $\mathbb{T}_d^n$  is of size at least  $Z_0 + \dots + Z_{n-1}$ , and*
- (ii)  *$\mathbf{E}[\mathcal{C}_{\text{wakeup}}] = O(n^2)$ .*

*Proof of Proposition 128 (i).* Recall in the non-backtracking frog model an awake frog uniformly jumps to one of its neighbors and never backtracks, halting if it ever visits the root. Though this coupling is given on  $\mathbb{T}_d$ , it extends naturally to  $\mathbb{T}_d^n$ . The only necessary modification is we halt non-backtracking frogs upon visiting either the root or a leaf. The details are similar enough that we omit them.

Assuming frogs perform non-backtracking random walks, we introduce one more (relatively drastic) restriction. Namely, immediately halt a just woken frog if its next jump would be towards the root or to the same vertex as the frog that woke it. When a frog is woken up there are now two possibilities:

- With probability  $\frac{2}{d+1}$  a just woken frog is halted. This ensures that only one of its children will be woken.
- With probability  $\frac{d-1}{d+1}$  the woken frog jumps away from the root and to a different child than the frog that woke it. This results in two children being visited.

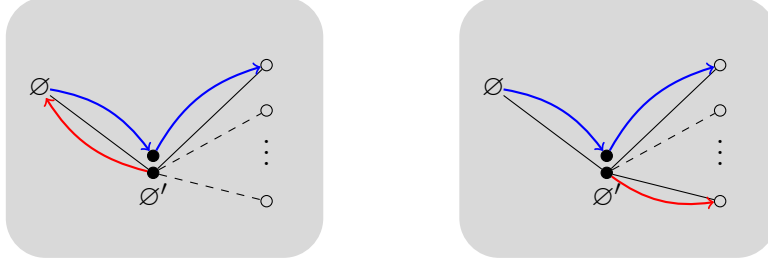


Figure 13: Outcomes that would result in  $Z_1 = 1$  and  $Z_1 = 2$ , respectively.

The subtree of starting vertices of frogs allowed to move is then equivalent to a Galton-Watson tree with generation size distributed as  $Z$ . Figure 13 illustrates how the first generation is formed. Note that since the first waking occurs at  $\emptyset'$  we have  $n - 1$  generations. □

**3.10.2.1 Time Distortion** With the usual and non-backtracking models coupled as in the proof of Proposition 128 (i) we define  $\mathcal{C}_{\text{wakeup}} = \mathcal{C}_{\text{wakeup}}^n$  to be the number of steps taken by non-halted frogs in the usual model to visit the vertices counted by  $Z_{n-1}$ .

The basic idea behind showing  $\mathbf{E}[\mathcal{C}_{\text{wakeup}}] = O(n^2)$  is to obtain a recursive upper bound. The recursion works since  $\mathcal{C}_{\text{wakeup}}^{n+1}$  is no more than  $\mathcal{C}_{\text{wakeup}}^n$  plus the time elapsed in the usual model for each awake non-backtracking frog to take one more step. To make this formal we need three lemmas. The first of which proves that a single non-backtracking step is dominated by the time for a biased random walk to last visit the origin.

**Lemma 129.** *Let  $\sigma$  be the time of last visit to 0 by a biased random walk on the integers started at 0 with probability  $\frac{d}{d+1}$  of moving right and  $\frac{1}{d+1}$  of moving left.*

*Let  $\gamma$  be the time for a frog in the usual model to visit the vertex counted by one step of its non-backtracking counterpart. It holds that  $\gamma$  is stochastically dominated by  $\sigma$ .*

*Proof.* The path of a non-backtracking frog is obtained by trimming off excursions away from the unique geodesic a simple random walk on  $\mathbb{T}_d^{\text{hom}}$  intersects infinitely many times. If we project the walk on  $\mathbb{T}_d^{\text{hom}}$  to the integers then it is exactly the random walk described in the statement of the

lemma. An excursion takes no longer than for the frog on  $\mathbb{T}^{\text{hom}}$  to last visit its current depth, which is distributed as  $\sigma$ . Thus,  $\gamma \preceq \sigma$ .  $\square$

**Lemma 130.** *Let  $\sigma_i$  be i.i.d. random variables distributed like  $\sigma$ . There is a coupling so that*

$$\mathcal{C}_{\text{wakeup}}^{n+1} \leq \mathcal{C}_{\text{wakeup}}^n + \max_{i=1, \dots, 2^n} \sigma_i.$$

*Proof.* Consider a usual frog model on  $\mathbb{T}_d^{n+1}$  coupled to a non-backtracking model where we freeze a non-backtracking frog and its coupled partner when the non-backtracking frog first visits depth  $n$ . With this freezing, after  $\mathcal{C}_{\text{wakeup}}^n$  steps in the usual model all  $Z_1 + \dots + Z_{n-1}$  frogs will be at depth  $n$  and will after each taking a non-backtracking step determine the size of  $Z_n$ . By Lemma 129 each frog in the usual model takes no more than  $\sigma$  steps to follow its non-backtracking partner. Taking a maximum and noting that  $Z_0 + \dots + Z_{n-1} \leq 2^n$  (this holds for all  $d$  since  $Z_k \leq 2^k$  when  $k \leq n-1$ ) gives the claimed bound.  $\square$

**Lemma 131.** *Let  $\sigma_i$  be as in the previous lemma. It holds that  $\mathbf{E} \max_{i=1, \dots, 2^n} \sigma_i = O(n)$ .*

*Proof.* It is a classical fact that  $\sigma$  has exponential tail. So, there is a constant  $0 < b < 1$  so that for all  $r \geq 0$  it holds that  $\mathbf{P}[\sigma > r] \leq b^r$ . Letting  $M = \max_{i=1, \dots, 2^n} \sigma_i$  a union bound implies  $\mathbf{P}[M > r] \leq 2^n \mathbf{P}[\sigma > r] \leq 2^n b^r$ . It follows that  $\mathbf{P}[M > n \log_b(1/2) + s] \leq b^s$  for all  $s \geq 0$  which implies  $\mathbf{E}M = O(n)$ .  $\square$

*Proof of Proposition 128 (??).* Taking expectation of the equation in Lemma 130 then applying Lemma 131 we obtain

$$\mathbf{E}[\mathcal{C}_{\text{wakeup}}^{n+1}] \leq \mathbf{E}[\mathcal{C}_{\text{wakeup}}^n] + O(n).$$

Since  $\mathbf{E}[\mathcal{C}_{\text{wakeup}}^1] = 1$ , we solve the recursion and obtain  $\mathbf{E}[\mathcal{C}_{\text{wakeup}}^n] = O(n^2)$ .  $\square$

*Remark 132.* The paper [LMP05] also takes the approach of embedding a Galton-Watson tree in the frog process to give a lower bound on its growth. Their Galton-Watson tree grows with mean  $\frac{2d-1}{d}$  which is faster than ours on average but like ours still converges to 2 as  $d \rightarrow \infty$ . Moreover, their process goes extinct with positive probability, rendering it unsuitable for proving a bound like Proposition 127.

We would also like to note that it is possible to improve the growth of our restricted frog model by being less stringent about halting frogs. For fixed  $d$  this process grows faster than that in [LMP05] but still approaches mean 2 growth as  $d \rightarrow \infty$ . However, being less picky introduces several complications which we decided would distract from the core idea of our proof.

### 3.10.3 Proof of Theorem 126

*Proof.* Proposition 128 guarantees that after  $\mathcal{C}_{\text{wakeup}}$  steps in the frog model on  $\mathbb{T}_d^n$ , there are at least  $Z_{n-1}$  frogs awake. Let  $T(Z_{n-1}) = T_d^n(Z_{n-1})$  be the time for these  $Z_{n-1}$  random walks (now allowed to move as simple random walks) to cover  $\mathbb{T}_d^n$ . After time  $\mathcal{C}_{\text{wakeup}} + T(Z_{n-1})$ , every vertex has been visited, and so

$$\mathcal{C} \leq \mathcal{C}_{\text{wakeup}} + T(Z_{n-1}). \quad (44)$$

Proposition 128 ensures that  $\mathbf{E}[\mathcal{C}_{\text{wakeup}}] = O(n^2)$ . So, it remains to bound  $\mathbf{E}[T(Z_{n-1})]$ .

Let  $\epsilon > 0$  and  $\kappa = m^{n-1}/n^\epsilon$ . Define the disjoint events

$$A_1 = \{Z_{n-1} < n\}, \quad A_2 = \{n \leq Z_{n-1} \leq \kappa\} \quad \text{and} \quad A_3 = \{Z_{n-1} > \kappa\}.$$

Conditioning on these outcomes we can write

$$\mathbf{E}[T(Z_{n-1})] = \sum_{i=1}^3 \mathbf{E}[T(Z_{n-1}) \mid A_i] \mathbf{P}[A_i].$$

We now estimate each of the three summands.

- I. Using the fact that  $\alpha = \log_m(\frac{d+1}{2})$ , the bound (127.2) becomes  $\mathbf{P}[A_1] \leq Bn^\alpha(\frac{2}{d+1})^n$ . At worst,  $Z_{n-1} = 1$ , and  $\mathbf{E}[T(Z_{n-1})] = O(n^2 d^n)$  by [Ald91, Theorem 7]. Multiplying these together gives the bound

$$\mathbf{E}[T(Z_{n-1}) \mid A_1] \mathbf{P}[A_1] = O(n^{2+\alpha}(\frac{2d}{d+1})^n). \quad (45)$$

- II. We further break up the event  $A_2$ :

$$\mathbf{E}[T(Z_{n-1}) \mid A_2] \mathbf{P}[A_2] = \sum_{k=n}^{\kappa} \mathbf{E}[T(Z_{n-1}) \mid Z_{n-1} = k] \mathbf{P}[Z_{n-1} = k].$$

Noting that  $m^{n-1} = O(m^n)$ , we apply the bound on  $\mathbf{P}[Z_{n-1} = k]$  from (127.1) and the speedup result of Proposition 134 to obtain

$$\begin{aligned} \sum_{k=n}^{\kappa} \mathcal{O}(n^4 d^n / k^{1/2}) \mathcal{O}(k^{\alpha-1} / m^{n\alpha}) &= O(n^4 (d/m^\alpha)^n) \sum_{k=n}^{\kappa} k^{\alpha-\frac{3}{2}} \\ &= O(n^4 (d/m^\alpha)^n \kappa^{\alpha-\frac{1}{2}}) \\ &= O(n^{4-\epsilon(\alpha-\frac{1}{2})} (d/\sqrt{m})^n). \end{aligned} \quad (46)$$

III. Last, we make the naive bound  $\mathbf{P}[A_3] \leq 1$  and again apply Proposition 134 assuming the slowest case,  $Z_{n-1} = \kappa$ . This yields

$$\mathbf{E}[T(Z_{n-1}) \mid A_3] \mathbf{P}[A_3] = O(n^{4+\frac{1}{2}\epsilon} (d/\sqrt{m})^n). \quad (47)$$

Since  $\alpha > 0$ , both (46) and (47) are equal to  $O(n^{4+\epsilon} (d/\sqrt{m})^n)$ . Therefore,

$$\mathbf{E}[T(Z_{n-1})] = O(n^{2+\alpha} (\frac{2d}{d+1})^n) + O(n^{4+\epsilon} (d/\sqrt{m})^n).$$

Recalling that  $m = \frac{2d}{d+1}$ , a simple calculation shows that for any  $d \geq 2$  it holds that  $\frac{2d}{d+1} \leq \frac{d}{\sqrt{m}}$  and so we can absorb big- $O$  terms and obtain  $\mathbf{E}[T(Z_{n-1})] = O(n^{4+\epsilon} (d/\sqrt{m})^n)$  which in light of (44) implies our theorem.  $\square$

### 3.10.4 Speedup from multiple walks on $d$ -ary tree

The following is the  $d$ -ary analogue of Lemma 4.1 from [Sau10]. The proofs are very similar to those in the original paper and we omit them.

**Lemma 133.** *Let  $N = \sum_{j=0}^n d^j$ , the number of vertices in  $\mathbb{T}_d^n$ , the  $d$ -ary tree of depth  $n$ . Consider a random walk of length  $l = N \log N$  on  $\mathbb{T}_d^n$ , starting from an arbitrary vertex. There exist positive integers  $a$ ,  $b$  and  $c$  which depend on  $d$  but are independent of  $n$  such that*

- (L.1) *With probability at least  $1 - N^{-2}$ , the random walk visits a leaf in every time interval  $[t, t + a \log N]$  with  $1 \leq t \leq l - a \log N$ .*
- (L.2) *With probability at least  $1 - N^{-2}$ , the random walk visits no vertex more often than  $b \log^2 N$  times.*
- (L.3) *With probability at least  $1/(cN \log_d N)$ , the random walk visits the root within the time interval  $[0, a \log N]$ .*

We now present our result, essentially a carbon copy of [Sau10, Lemma 4.1].

**Proposition 134.** *Let  $T(k) = T_d^n(k)$  be the time for  $k \leq N$  random walks in the least optimal initial configuration on  $\mathbb{T}_d^n$  (in terms of expected cover time) to collectively visit every vertex. For each fixed  $d$ ,*

$$\mathbf{E}[T(k)] = O(n^4 d^n / \sqrt{k}).$$

*Proof.* By (L.3), after  $t_0 = (acN \log_d N \cdot \log N) / \sqrt{k}$  steps the probability for a random walk to visit the root vertex is

$$1 - \left(1 - \frac{1}{cN \log_d^2 N}\right)^{cN \log_d N / \sqrt{k}} \geq 1 - e^{-1/\sqrt{k}} \geq \frac{2}{\sqrt{k}}.$$

For a sufficiently large constant  $k_0$ , it follows from a Chernoff bound that for  $k \geq k_0$ , at least  $\sqrt{k}/4$  walks visit the root vertex in  $t_0$  steps with probability at least  $1/2$ . (If  $k < k_0$ , the proposition follows from [Ald91, Theorem 7], which says that the cover time of  $\mathbb{T}_d^n$  by a single random walk is  $O(n^2 d^n)$ .) After a walk reaches the root, suppose that the events in (L.1) and (L.2) hold from that time on. Then, for some  $t_1$  to be chosen later, the walk will visit at least  $\frac{t_1}{b \log^2 N \cdot a \log N}$  leaves in its next  $t_1$  steps. By a union bound on the events in (L.1) and (L.2) failing to hold for any of  $\sqrt{k}/4$  walkers, this will occur for all of  $\sqrt{k}/4$  walkers with probability at least

$$1 - \frac{2\sqrt{k}}{4N^2} \geq 1 - N^{-1}$$

A random walk starting at the root is equally likely to visit each of its  $d^n$  leaves. Choosing  $t_1 = \frac{abc(d^n) \log N}{\sqrt{k}}$  yields that for any leaf  $u$ ,

$$\begin{aligned} \mathbf{P}[u \text{ is visited by time } t_0 + t_1] &\geq \frac{1}{2} (1 - N^{-1}) \left(1 - \left(1 - \frac{t_1}{d^n b \log^2 N \cdot a \log N}\right)^{\sqrt{k}/4}\right) \\ &\geq \frac{1}{2} \left(1 - \exp\left(-\frac{1}{\log^2 N}\right)\right) - N^{-1} \\ &\geq \frac{1}{4 \log^2 N} \end{aligned}$$

for sufficiently large  $N$ . Iterating this  $4 \log^2 N$  times gives,

$$\begin{aligned} \mathbf{P}[u \text{ is visited during } (t_0 + t_1) \cdot 4 \log^2 N \text{ steps}] &\geq 1 - \left(1 - \frac{1}{4 \log^2 N}\right)^{4 \log^2 N} \\ &\geq 1 - e^{-1}. \end{aligned}$$

Taking the union bound over the  $d^n$  leaves, we obtain that all leaves are visited from random walks starting at the root after  $O(N \log^4 N / \sqrt{k}) = O(n^4 d^n / \sqrt{k})$  steps with probability  $1 - e^{-1}$ . If all leaves have been visited by random walks that start from the root, then all vertices of the tree have been visited. Iterating the previous estimate shows that  $T(k)$  is stochastically dominated by  $(1 + X)O(n^4 d^n / \sqrt{k})$ , where  $X \sim \text{Geo}(1 - e^{-1})$ , completing the proof.  $\square$

### 3.11 Frog Questions

#### 3.11.1 Open but maybe solvable with a nice idea

1. What is the (order of the) cover time for the frog model on a cycle and the torus?
2. Show the frog model is transient or recurrent on a non-trivial Galton-Watson tree.
3. Consider a frog model where frogs perish when they visit a site with no sleeping frogs. Does this process survive with positive probability in  $\mathbb{Z}^d$ ?

## 4 Coalescing random walk

### 4.1 Introduction

*Coalescing random walk* (CRW) starts with one particle at each vertex of a connected, locally finite, and undirected graph. We make the graph directed by doubling the edge set, then run independent rate-1 Poisson clocks at each edge. When an edge clock rings, any particles at the tail of that edge move to the adjacent vertex. In this fashion, particles occupying the same site have “coalesced” since they proceed as one. Say that CRW is *site recurrent* if every site is almost surely visited infinitely often. If instead this occurs with probability 0, call the process *transient*. Our main tool for proving site recurrence is the following necessary and sufficient condition in terms of the expected occupation time of a vertex.

**Proposition 135.** *Site recurrence is equivalent to infinite expected occupation time at any vertex. Moreover, CRW is either site recurrent or transient (i.e. it satisfies a 0-1 law).*

Let  $p_t(v)$  be the probability a particle is at the vertex  $v$  at time  $t$ , so that site recurrence is equivalent to divergence of  $\int p_t(v) dt$ . We use duality



with the voter model to obtain non-integrable lower bounds on the following graphs:

**Theorem 136.** *CRW is site recurrent on:*

- (i) *Bounded degree graphs. If the maximum degree is  $D$ , then for all vertices  $v$*

$$p_t(v) \geq (1 + Dt)^{-1}$$

*for all  $t \geq 0$ .*

- (ii) *Galton-Watson trees with offspring distribution on  $\mathbb{Z}^+$  and the probability of  $k$  offspring bounded by  $e^{-ck}$  for some  $c > 0$  and large enough  $k$ . Here*

$$p_t(v) \geq C(t \log t)^{-1}$$

*for all vertices  $v$ , large enough  $t$ , and some  $C > 0$  that depends on  $c$ .*

Note that there are unbounded degree graphs for which CRW is not site recurrent; even the non-coalescing system of independent random walks is transient on trees with rapidly increasing degree. We are not sure how much the exponential tail hypothesis in (ii) can be weakened. See Further Questions (a) for more discussion. A corollary to Theorem 136 is a general upper bound on the probability that a vertex  $v$  is unoccupied on the interval  $(t, u)$ .

**Corollary 137.** *Let  $\sigma_t$  be the first time after  $t$  that  $v$  is occupied by a particle. It holds that*

$$\mathbf{P}(\sigma_t > u) \leq \frac{t}{t + \int_t^u p_s(v) ds}.$$

*And, for a graph with maximum degree  $D$*

$$\mathbf{P}(\sigma_t > u) \leq \frac{t}{t + \frac{1}{D}(\log(1 + Du) - \log(1 + Dt))} = O(1/\log(u)).$$

We also give a universal upper bound for  $p_t(v)$  on general graphs. It follows that the occupation probability decays to 0 for any graph. The upper bound is a small modification of an argument in [Gri78], so we also credit David Griffeath.

**Proposition 138** (Griffeath). *Let  $G$  be a connected, infinite graph. For all vertices  $v$ , any  $\epsilon > 0$ , and large enough  $t$  (depending on  $\epsilon$ ) it holds that  $p_t(v) \leq (1 + \epsilon)/(2\sqrt{\pi t})$ .*

### 4.1.1 History

The study of coalescing systems began in the 1970's with the paper of Erdős and Ney [EN74]. The duality relationship to the voter model, which we rely heavily upon, was first observed in [HL75]. Variations of coalescing random walk continue to find new applications. For example, random coalescence involving multiple types of particles, and particle interaction rules, is used to model certain chemical reactions (see [Hol83], [BL88] and [vdBK00]). Non-spatial models such as Kingman coalescence ([Kin82]) find applications in modeling ancestry in biology. A survey of coalescence models can be found in [Ber09]. Arratia [Arr83] looks at site recurrence for discrete time walks, and annihilating systems, both with possibly vacant sites in the starting configuration. CRW is applied to the voter model in [BL15]. Also, it is studied in more generality in [RV15] and [GPTZ15]. Other recent articles have focussed on different settings. Its behavior on finite graphs is of interest to computer scientists. The model on the  $d$ -dimensional torus is introduced in [Cox89]. There, they study the expected time for the process to coalesce into a single particle. In [CEOR12] the coalescence time is studied on a variety of finite graphs. Elsewhere, in a continuous spatial setting, recurrence is studied with coalescing diffusions by Cabezas, Rolla and Sidoravicius in [CRS13].

Early results for coalescing random walk focused on the lattices  $\mathbb{Z}^d$ . In [Gri78] Griffeath shows that both coalescing and annihilating random walk on  $\mathbb{Z}^d$  is a.s. *weakly recurrent*, under certain restrictions on the vacant sites in the starting configuration. Weakly recurrent means that each site is occupied infinitely often, but for a decreasing fraction of time. An important ingredient in the proof of recurrence is an estimate for the function  $p_t$ , the probability a particle occupies the origin at time  $t$ . In [BG80], Bramson and Griffeath study  $p_t$  in the coalescing case and compute its asymptotics for every  $d \geq 1$ . Rather nicely, for  $d \geq 3$  it holds that  $p_t \approx (\gamma_d t)^{-1}$  with  $\gamma_d$  the probability a random walk on  $\mathbb{Z}^d$  never returns to its starting point. The proof of this is computational; later, [vdBK02] Kesten gives a probabilistic argument that revolves around the heuristic  $p'_t \approx -\gamma_d p_t^2$ .

## 4.2 Overview

The main idea is to obtain information about  $p_t(v)$  from a dual voter model. This dual was first applied to CRW in [HL75], and subsequently utilized in [HS79, Gri78, Arr81, Arr83]. In Corollary 140 we deduce that  $p_t(v)$  is equivalent to the probability a time-changed nearest neighbor simple random

walk avoids 0 up to time  $t$ . All of our estimates come from studying this random walk.

#### 4.2.1 Further Questions

We record several questions regarding coalescing and annihilating random walk here:

- (a) Can the assumptions on the degree in Theorem 136 (ii) be weakened? We expect that our approach extends (at least) to stationary graphs with finite expected degree.
- (b) Suppose  $G$  is an infinite unimodular random graph in which each vertex is assigned an infinite trajectory in an ergodic invariant way (see [BM72]). Particles, one from each vertex, start moving along their trajectory in continuous time and annihilate when meeting. Is the resulting process recurrent? Start by showing it on Euclidean lattices.
- (c) Place  $\epsilon$ -balls (meteors) in Euclidean space with centers according to a unit intensity Poisson process. At time 0 each chooses a direction uniformly randomly and proceeds along this direction at unit speed (non-random). When two meteors collide, they annihilate. Is the origin a.s. occupied by infinitely many meteors for all  $d \geq 1$ ? This is discussed in more detail in Section 4.3.3.
- (d) Have particles perform annihilating random walk on a graph where the particle started at  $x$  steps according to an exponential clock with mean an independent uniform(0, 1) random variable. Is this model on  $\mathbb{Z}^d$  still recurrent? If so, what can be said of the limiting speed of the particles visiting the origin? Possibly slower moving particles survive longer, and the average speed of particles visiting the origin decays with time.
- (e) Have particles on  $\mathbb{Z}^d$  sample one of the  $2d$  directions then forever move according to rate-1 exponential clocks. Is this process recurrent?

#### 4.2.2 Outline

Section 4.3 starts with the proof of Theorem 148. We also establish, in Lemma 139, that infinite expected occupation time is equivalent to survival of a nearest neighbor random walk in the voter model. Corollary 140 relates this back to  $p_t$ . Sections 4.3.1 and 4.3.2 contain the proofs of Theorem 136 (i)

and (ii), respectively. Section 4.3.3 discusses some non-backtracking variants, and contains the proof of site recurrence for a non-backtracking model on bounded degree trees and Galton-Watson trees with exponential tail.

### 4.3 Site recurrence for coalescing random walk

Coalescing random walk on a graph,  $G = (V, E)$ , has a graphical representation as follows: each edge is replaced with two directed edges and an independent Poisson process with unit intensity is placed on each directed edge, indexed by time. When the bell of a Poisson process for the edge  $(u, w)$  rings we check if there's a particle at  $u$  and if so, we move it to  $w$ . If there's already a particle at  $w$ , they merge. We denote the process  $(\xi_t)_{t \geq 0}$  with  $\xi_t \in \{0, 1\}^V$  equal to the set of occupied vertices at time  $t$ , and occasionally  $\xi_t^v$  for the location at time  $t$  of the particle that began at  $v$ . In this notation we have  $p_t(v) = \mathbf{P}(\xi_t(v) = 1)$  is the probability that  $v$  is occupied by a particle at time  $t$ . Thus,  $\int_0^\infty p_t(v) dt$  is the expected occupation time of  $v$ .

*Proof of Theorem 148.* If there is positive probability of infinite occupation time at  $v$ , then the expected occupation time is infinite. For the other direction we generalize [Arr83, Lemma 2].

Suppose that  $\int_0^\infty p_t(v) dt = \infty$ . For any  $t \in [0, \infty)$  let  $\sigma_t = \inf\{s \geq t : \xi_s(v) = 1\} \in [0, \infty]$ , the first time after  $t$  that  $v$  is occupied by a particle. We wish to establish that

$$\mathbf{P}(\sigma_t < \infty) = 1, \quad \forall t \geq 0.$$

The basic coupling  $\xi_t^A = \{\xi_t^x : x \in A\}$  for  $A \subseteq V$  has the property that  $A \subseteq B \subseteq V$  implies  $\xi_t^A \subseteq \xi_t^B \subseteq \xi_t$ , so the Markov process  $(\xi_t^A : A \subseteq V)$ , with state space  $\{0, 1\}^V = \{A : A \subseteq V\}$  ordered by set inclusion, is attractive. Define  $p_t^A = \mathbf{P}(\xi_t^A(v) = 1)$ , and also  $I(t, u) = \int_t^u p_s(v) ds$ . By assumption,

$$\lim_{u \rightarrow \infty} I(t, u) = \infty, \quad \forall t \geq 0. \quad (48)$$

Let  $f_{\sigma_t}$  be the density function of  $\sigma_t$  and  $\mathbf{E}_{\sigma_t}$  denote the expectation taken over all possible realizations of  $\xi_{\sigma_t}$  given that  $\sigma_t = r$ . Let  $A = \xi_{\sigma_t}^V$ . Using the strong Markov property we have for  $t < u$ ,

$$\begin{aligned} I(t, u) &:= \mathbf{E} \int_t^u \mathbf{1}\{\xi_s(v) = 1\} ds = \int_{r=t}^u f_{\sigma_t}(r) \mathbf{E}_{\sigma_t} \left( \mathbf{E} \int_0^{u-r} \mathbf{1}\{\xi_s^A(v) = 1\} ds \right) dr \\ &\leq \int_{r=t}^u f_{\sigma_t}(r) \left( \mathbf{E} \int_0^u \mathbf{1}\{\xi_s(v) = 1\} ds \right) dr \\ &\leq \mathbf{P}(\sigma_t \leq u)(t + I(0, u)). \end{aligned}$$

Dividing both sides by  $t + I(0, u)$  we arrive at the inequality

$$\mathbf{P}(\sigma_t \leq u) \geq I(t, u)/(t + I(0, u)). \quad (49)$$

For fixed  $t$ , taking  $u \rightarrow \infty$  yields  $\mathbf{P}(\sigma_t < \infty) \geq 1$  by (48).

We conclude by describing the 0 – 1 law. The above argument establishes that if CRW occupies a site for infinite time with positive probability, then it does so with probability 1. As  $G$  is assumed to be connected, it follows that all sites are occupied infinitely often with probability 1. Therefore, the process is either site recurrent (recall this is defined as an almost sure event) or transient.  $\square$

*Proof of Corollary 137.* The lower bound on  $\mathbf{P}(\sigma_t \leq u)$  at (49) yields an upper bound on the probability  $v$  is unoccupied from time  $t$  to  $u$ :

$$\mathbf{P}(\sigma_t > u) \leq 1 - \frac{I(t, u)}{t + I(t, u)} = \frac{t}{t + \int_t^u p_s(v) ds}. \quad (50)$$

Which is the first part of the corollary. The second part follows by applying the bound on  $p_t(v)$  in Theorem 136 (i) and integrating.  $\square$

Theorem 148 allows for site recurrence to be deduced by proving  $p_t(v)$  is non-integrable. Our approach is to express  $p_t(v)$  in another way. Consider the dual process to this model, which is called the *voter model*. In the dual model we start with a partition of the space into clusters, where initially each vertex corresponds to a different cluster. When the bell at  $(u, w)$  rings the vertex  $w$  is added to the cluster containing  $u$ . We denote the process  $(\zeta_t^v)_{t \geq 0}$  where  $\zeta_t^v$  is the set of vertices belonging to the cluster that initially consists of the vertex  $v$ . If we run this model in reverse time, from time  $t$  to 0, we see the cluster that began at  $v$  at time  $t$  at time 0 consists of exactly the particles that in the coalescing model are at  $v$  at time  $t$ . In particular,

$$p_t(v) = \mathbf{P}(\xi_t(v) = 1) = \mathbf{P}(\zeta_t^v \neq \emptyset). \quad (51)$$

The advantage of working with the voter model is that the size of  $\zeta_t^v$  is a nearest-neighbor symmetric random walk with transition rate depending on the number of boundary edges. Indeed, at any moment there are some directed edges going out of the cluster, and the same number of edges coming in. More precisely, the cluster size is a skip-free process on the integers that moves with rate equal to the size of the current boundary of the cluster. We record this fact in the following lemma. Let  $|\cdot|$  denote either the counting measure of a finite set.

**Lemma 139.** Define  $\zeta_t^v \subseteq V$  to be the set of vertices in the cluster of  $v$  at time  $t$ . For each  $x \in \zeta_t^v$  let  $\partial_t(x) = \{(x, y) \in E: y \notin \zeta_t^v\}$ . Let  $\sum_{x \in \zeta_t^v} |\partial_t(x)|$  be the number of edges leading out of  $\zeta_t^v$ . The process has the following properties:

- (i)  $|\zeta_0^v| = 1$ .
- (ii) Let  $\tau = \inf\{t: |\zeta_t| = 0\}$ . For all  $t \geq \tau$  it holds that  $|\zeta_t^v| = 0$ .
- (iii) The process is a martingale that transitions to  $|\zeta_t^v| \pm 1$  at rate  $\sum_{x \in \zeta_t^v} |\partial_t(x)|$ .

*Proof.* Properties (i) and (ii) follow from the construction of  $\zeta_t^v$ . We turn our attention to property (iii). For each  $x \in \zeta_t^v$ ,  $x$  is removed from  $\zeta_t^v$  at rate  $|\partial_t(x)|$  and each of the  $|\partial_t(x)|$  sites in  $\partial_t(x)$  is added to  $\zeta_t^v$  at rate 1. Since the rates balance,

$$\mathbf{E}[|\zeta_t^v| \mid |\zeta_{t-}^v|] = |\zeta_{t-}^v|,$$

which establishes  $|\zeta_t^v|$  is a martingale. Summing the rates over  $x \in \zeta_t^v$  shows that  $|\zeta_t^v|$  transitions to  $|\zeta_t| + 1$ , and to  $|\zeta_t| - 1$ , each at rate  $\sum_{x \in \zeta_t^v} |\partial_t(x)|$ .  $\square$

This lets us describe  $p_t$  in terms of the voter model.

**Corollary 140.**  $p_t(v) = \mathbf{P}(|\zeta_t(v)| > 0)$ . This is the probability a nearest neighbor random walk with transition rate  $\sum_{x \in \zeta_t} |\partial_t(x)|$  to each of its two neighboring states, and absorbing state at 0 is yet to reach zero at time  $t$ .

It follows that  $p_t \rightarrow 0$  on any infinite, connected graph.

*Proof of Proposition 138.* The transition rate in  $\zeta_t^v$  is always at least two. By Corollary 140 and a straightforward coupling we have  $p_t$  is at most  $\tilde{p}_t = \mathbf{P}(X_s > 0, \forall s \leq t)$ , with  $X_s$  a rate-2 continuous time simple symmetric random walk started at 1. Using the reflection principle together with the local central limit theorem,  $\tilde{p}_t \sim 1/(2\sqrt{\pi t})$  as  $t \rightarrow \infty$ , and the result follows.  $\square$

*Remark 141.* For coalescing walk on  $\mathbb{Z}$  with nearest neighbour connections, since  $\zeta_t^v$  is always of the form  $\{x, x+1, \dots, x+k\}$  for some  $x \in \mathbb{Z}, k \in \mathbb{Z}^+$ , its transition rate is exactly  $2 \cdot 2 = 4$ , so the above inequality is an equality, and gives the exact asymptotics  $p_t \sim 1/(2\sqrt{\pi t})$ , as observed in [BG80]. Compared to [BG80] there is an extra factor of  $1/2$ ; our convention differs from theirs in that the transition of a particle at  $v$  is equal to  $\deg v$  and not 1, since in our case  $\deg v$  is allowed to vary.

### 4.3.1 Site recurrence for bounded degree graphs

Now we turn our attention to proving site recurrence on general graphs. Define  $\tau_v = \inf\{t: \zeta_t^v = \emptyset\}$ . Integrating over  $t$  in the duality relation (51) we find

$$\mathbf{E}\tau_v = \int_0^\infty \mathbf{P}(|\zeta_t^v| > 0)dt = \int_0^\infty \mathbf{P}(\xi_t(v) = 1)dt = \int_0^\infty p_t(v)dt.$$

So, proving site recurrence is equivalent to showing that the first hitting time of 0 for the simple random walk  $|\zeta_t^v|$  (with random and time-varying transition rate) has infinite expectation. We start with the case when  $G$  has bounded degree.

*Proof of Theorem 136 (i).* Let  $v \in V$  with the maximum degree of vertices in  $G$  bounded by  $D$ . Lemma 139 establishes that the transition rate of  $\zeta_t^v$  to either increase or decrease by 1 is less than or equal to  $D|\zeta_t^v|$ . Let  $W_t$  be a continuous time nearest-neighbour random walk on  $\mathbb{Z}^+ \cup \{0\}$  with  $W_0 = 1$ . The walk transitions from  $k \in \mathbb{Z}^+ \cup \{0\}$  to  $k \pm 1$  each at rate  $Dk$ , and is absorbed at 0. Letting  $\theta(t) = \mathbf{P}(W_t > 0)$ , it follows from Corollary 140 and a straightforward coupling of  $|\zeta_t^v|$  with  $W_t$  that  $p_t(v) \geq \theta(t)$ , so it suffices to control  $\theta(t)$ .

The process  $W_t$  can be interpreted as the number of particles in a branching process with a single initial particle in which each particle independently dies, or gives birth to a single offspring, each at rate  $D$ . Let  $\rho(t) = \mathbf{P}(W_t = 0)$ . From [AN72, Ch.III, Sec. 4, Eq.2],  $\rho(t)$  satisfies the DE

$$\frac{d}{dt}\rho(t) = u(\rho(t)) \quad \text{with} \quad u(x) = 2D \cdot (1/2 + x^2/2 - x) = D(1 - x)^2.$$

Since  $\theta(t) = 1 - \rho(t)$ , we find

$$\frac{d}{dt}\theta(t) = -\theta(t)^2,$$

and since  $\theta(0) = 1$  we obtain the unique solution  $\theta(t) = 1/(1 + Dt)$ . □

*bit more justification about  $\rho(t)$ .* By independence, for each  $k \geq 0$  we have

$$\mathbf{P}(W_{t+h} = 0 \mid W_h = k) = \rho(t)^k. \tag{52}$$

Recalling that  $W_0 = 1$  then conditioning on  $W_h$  for small  $h > 0$ ,

$$\rho(t+h) = \mathbf{P}(W_{t+h} = 0) = \sum_{k \geq 0} \mathbf{P}(W_{t+h} = 0 \mid W_h = k)\mathbf{P}(W_h = k).$$

The event that two or more transitions happens on the interval  $[0, h]$  is contained in the event that a rate  $2D$  exponential clock rings, then a rate  $4D$  exponential clock rings (i.e. we go from 1 to 2 particles then another transition happens). The probability of this is bounded by the probability that  $X + Y < h$  for  $X$  and  $Y$  rate  $4D$  exponential random variables, and has density  $f_{X+Y}(t) = \lambda^2 t e^{-\lambda t}$  with  $\lambda = 1/4D$ . Integrating on  $[0, h]$ , then taking the Taylor expansion we have  $\mathbf{P}(X + Y < h) = O(h^2)$ . Since we will be dividing by  $h$  and letting it tend to 0, we can combine all of the events that occur with two or more transitions into an  $O(h^2)$  term. Using the expression (52) this lets us write  $\rho(t + h)$  as

$$\rho(t + h) = \underbrace{1 \cdot Dh}_{\text{dies out}} + \underbrace{\rho(t)(1 - 2Dh)}_{\text{no change}} + \underbrace{\rho(t)^2 Dh}_{\text{increases by 1}} + O(h^2).$$

Subtracting  $\rho(t)$ , dividing by  $h$  and taking  $h \downarrow 0$  this converges to the equation  $\rho' = D(1 - \rho)^2$ .

So, for the survival probability  $\theta(t) = 1 - \rho(t)$  we find  $\theta' = -\theta^2$ , with  $\theta(0) = 1$ , whose unique solution is  $\theta(t) = 1/(1 + Dt)$ . □

### 4.3.2 Site recurrence for Galton-Watson trees

A more general upper bound on the transition rate is

$$|\text{maximum exposed degree}| \cdot |\zeta_t|.$$

Our hypothesis that the offspring distribution of our Galton-Watson tree has exponential tail guarantees that the maximum exposed degree is asymptotically bounded by  $\log(\text{number of steps})$ . Ultimately this lets us compare with the divergent integral  $\int_{t_0}^{\infty} (t \log t)^{-1} dt$ . This is made rigorous below.

*Proof of Theorem 136 (ii).* Again, by Corollary 140 it suffices to prove that  $\zeta_t^v$  has infinite expected survival time. For convenience we denote  $\zeta_t^v$  by  $\zeta_t$ . Let  $H_t = \cup_{s \leq t} \zeta_s$  be the vertices visited up to time  $t$ . Define the random times  $0 = t_1 < t_2 < \dots$  as when a vertex is added to  $H_t$ , and list them as  $v_1, v_2, \dots$  in the order they are discovered, with  $v_1$  being the root ( $\rho$ ). The transition rate of  $|\zeta_t| \rightarrow |\zeta_t| \pm 1$  is at most  $M_t |\zeta_t|$  where

$$M_t = \sup\{\deg v : v \in H_t\}.$$

So, our first goal is to construct the voter model in such a way that  $M_t$  can be easily controlled. A simple way to do this is to construct  $G$  “on the fly.”



That is, let  $(X_i)_{i \geq 1}$  be an i.i.d. sequence of copies of the offspring distribution, and at time  $t_i$ , sample  $X_i$  to determine the offspring distribution of  $v_i$ , which is then fixed for all  $t > t_i$ . This does not disturb the sample path distribution of  $\zeta_t$ , and has the advantage that the quantity

$$D_k = \max_{i \leq k} \deg v_i$$

is equal to  $\max_{i \leq k} X_i$  where  $X_i$  is a *fixed* (as opposed to being a randomly indexed) i.i.d. sequence. Since, by assumption,  $\mathbf{P}(X_i > x) \leq e^{-cx}$  for some  $c > 0$  and large enough  $x$ , a union bound gives  $\mathbf{P}(D_k > x) \leq ke^{-cx}$ , and setting  $x = (3/c) \log k$ ,

$$\mathbf{P}(D_k > (3/c) \log k) \leq k^{-2} \quad (53)$$

for large enough  $k$ .

Now, let  $0 = t_0 < t_1 < t_2 < \dots$  denote the jump times of  $\zeta_t$ . In what follows we will want the set of jump times to be infinite, so if  $\zeta_{t_i} = 0$  (i.e. the cluster dies out), just include jumps back to 0 at rate 1. Since  $|\zeta_{t_i}|$  is a martingale with  $\mathbf{E}|\zeta_{t_i}| = |\zeta_0| = 1$  for all  $i$ , Doob's martingale inequality implies that for all  $n > 0$

$$\mathbf{P}(\sup_i |\zeta_{t_i}| > bn) \leq (bn)^{-1} \quad (54)$$

for any  $b > 0$ . Note that  $M_{t_i}$  is nondecreasing in  $i$  and  $M_{t_i} \leq D_{1+i}$ , since vertices are exposed one at a time. Thus, the transition rate of  $|\zeta_{t_i}|$  is at most  $D_{i+1}|\zeta_{t_i}|$ . Combining these observations with (53) and (54), we find that with probability at least  $1 - (bn)^{-1} - n^{-2}$ , for  $t \leq t_{n^2}$  the transition rate in  $\zeta_t$  is at most

$$(3/c) \log(n^2 + 1)bn \leq (6b/c)n \log(n + 1). \quad (55)$$

Let  $m_n = m_n(b, c) = (6b/c)n \log(n + 1)$ . A quick summary: with high probability the first  $n^2$  transitions happen at rate no more than  $m_n$ . Equivalently, the time  $t_{n^2}$  is bounded below by the sum of  $n^2$  independent exponentials with rate  $m_n$ . This is an Erlang distribution,  $Y_{n^2}$ , with shape parameter  $n^2$  and rate  $m_n$ . Thus, we have

$$\text{mean: } \mu_n = \frac{n^2}{m_n}, \text{ and variance: } \sigma_n = \frac{\mu_n}{m_n}.$$

Chebyshev's inequality guarantees that

$$\mathbf{P}(|Y_n - \mu_n| \geq \mu_n/2) \leq \frac{\mu_n/m_n}{(\mu_n/2)^2} = \frac{4}{\mu_n m_n} = \frac{4}{n^2}.$$

One side of the above estimate is

$$\mathbf{P}(Y_{n^2} \leq \mu_n/2) \leq 4n^{-2}. \quad (56)$$

By comparison, and using (55) and (56), we have

$$\mathbf{P}(t_{n^2} \geq n^2/(2m_n)) \geq 1 - (bn)^{-1} - n^{-2} - 4n^{-2}.$$

From the well-known first passage distribution for random walk, for the random variable  $N = \inf\{n: |\zeta_{t_n}| = 0\}$  we have  $P(N > n^2) \geq c/n$  for some possibly smaller  $c > 0$ . Note that although  $m_n$  depends on  $c$ , shrinking  $c$  does not affect the estimate. Letting  $\tau = \inf\{t: |\zeta_t| = 0\}$  as before, and letting  $a_n = \mu_n/2 = cn/(12b \log(n+1))$ , for  $a_{n-1} < t < a_n$ ,

$$t \log t \geq \frac{c(n-1)}{12b \log n} (\log(n-1) - \log(12b/c) - \log \log n) \geq \frac{cn}{24b}$$

i.e.,  $n \leq (24b/c)t \log t$  for  $n$  large enough. For the same  $t$ , then,

$$\mathbf{P}(\tau > t) = \mathbf{P}(|\zeta_t| > 0) \geq \mathbf{P}(|\zeta_{t_{n^2}}| > 0, t_{n^2} > a_n).$$

Since the embedded chain of the cluster size is independent of the rate at which it transitions we have

$$\mathbf{P}(\tau > t) \geq c/n(1 - (bn)^{-1} - O(n^{-2})), \quad a_{n-1} < t < a_n.$$

The right side is at least  $c/(2n)$  for  $n$  greater than some  $n_0$ . Letting  $t_0 = a_{n_0}$  and using the upper bound on  $n$ ,

$$\mathbf{E}\tau = \int_0^\infty \mathbf{P}(\tau > t) dt \geq \frac{c^2}{48b} \int_{t_0}^\infty \frac{1}{t \log t} dt = \infty.$$

□

### 4.3.3 Non-backtracking coalescing random walk on trees

We are also interested in understanding similar, but less random processes. A lack of symmetry in these settings makes it difficult to apply known techniques. We are hopeful that progress will lead to new ideas.

The *non-backtracking coalescing random walk* is defined in the same way as the coalescing random walk with particles instead performing non-backtracking random walk. More precisely, the state of a particle is specified by a vertex-edge pair  $(u, \{u, v\})$ , and when an edge clock rings at a directed

edge  $(u, w)$ , the particle moves to  $w$  if and only if  $w \neq v$ . If the particle moves from  $u$  to  $w$ , its state is updated to  $(w, \{w, u\})$ , so that its next jump cannot be back to  $u$ . It will be convenient to assume that each particle is initialized with a uniformly chosen edge along which it cannot move, that is, the particle initially at  $v$  has state  $(v, \{v, u\})$  where  $u$  is a uniform random neighbour of  $v$ . With particles coalescing there is ambiguity about whose path to remember. There are several well-defined ways to assign priority. On a rooted tree, we analyze the special case where we always remember the path of particles moving towards the root. With the model defined in this way we do not quite have a voter model dual, but a closely related process does. Analogous to Theorem 148 we prove a necessary and sufficient condition for site recurrence.

**Proposition 142.** *Consider coalescing non-backtracking random walk on a rooted tree with priority given to particles moving towards the root. The process is site recurrent at the root if and only if the expected survival time of the cluster of the root in a certain voter model is infinite.*

We can deduce site recurrence on bounded degree trees and some trees with unbounded degree.

**Theorem 143.** *The process from Proposition 142 is site recurrent at the root of either a bounded degree tree or a Galton-Watson tree whose offspring distribution is as in Theorem 136 (ii).*

Non-backtracking removes a vital symmetry from the argument. The proof goes by, once again, constructing a dual voter process and showing the cluster of the root survives for an infinite expected amount of time. Our “priority to the root” rule is hand-picked to preserve monotonicity and the existence of a dual voter model. Neither property exists in other equally natural non-backtracking models. Further progress in these different settings will likely require a new approach. Consider the following conjecture:

**Conjecture 144.** *Non-backtracking coalescing random walk with any priority scheme is site recurrent on bounded degree trees.*

The inspiration for studying non-backtracking processes comes from the following *meteor model* on  $\mathbb{R}^d$ . Place  $\epsilon$ -balls in Euclidean space with centers according to a unit intensity Poisson process. At time 0 each chooses a direction uniformly randomly and proceeds along this direction at unit speed (non-random). When two meteors collide, they annihilate.

**Conjecture 145.** *The origin a.s. is occupied by infinitely many meteors for all  $d \geq 1$ .*

This problem appears quite difficult. It could be discretized to an annihilating system of random walks by uniformly assigning each particle a geodesic to  $\infty$  from which it never deviates and steps along according to a Poisson clock. The integer lattice is a natural graph to start with. Or, perhaps hyperbolic space—in which random walk paths stay within a logarithmic neighborhood of a geodesic—would be a more tractable place to study this problem.

#### 4.3.4 Recovering a dual

Since the graph is a tree, a particle either moves towards the root, or away, at each jump. Once it has moved away for the first time, at subsequent jumps it must always move away, since the only way back towards the root requires backtracking. To simplify matters, we suppose that at each vertex, the initial forbidden edge is chosen uniformly from the edges that lead away from the root.

Since coalescing particles may have different histories we must decide which one to remember. Therefore, upon collision we define the following rule for annihilating exactly one of the two colliding particles; in the original setting with memoryless walks, any such rule leads to the coalescing model.

- If one particle has not yet moved away from [edit] the root, and the other particle is moving away, annihilate the particle moving away.
- Otherwise, annihilate the particle currently occupying the site (i.e. keep the particle that just moved).

As currently stated, this process depends on past information. Running the process in reverse would require information about the future. Thus, it does not have a dual voter model. Still, a simple observation yields a related model that does have a dual.

Let  $X_T$  be the occupation time of the root up to time  $T$ . Notice that particles moving away from the root are inert; based on the rule above they cannot block upward moving particles, and as noted before they cannot revisit the root. Thus,  $X_T$  is unchanged if we delete particles the instant they turn away from the root. This modification gives the following model; for a vertex  $v$ , let  $d_v = \deg v$ .

- Suppose  $v$  is not the root.

- A particle at  $v$  moves towards the root at rate 1.
- A particle at  $v$  is deleted at rate  $d_v - 2$ .
- Suppose  $v = \rho$  is the root. A particle at  $\rho$  is deleted at rate  $d_\rho - 1$ .

We can think of this model as follows: each particle attempts to travel up to the root, coalescing with other particles upon collision, and particles (or coalesced collections of particles) are instantaneously zapped out of existence at some rate that depends on their present location.

We can simplify the description somewhat by introducing a single absorbing vertex  $\mathbf{a}$  and considering the process on  $V \cup \{\mathbf{a}\}$ , where every vertex has a directed edge pointing to  $v$  which rings at rate described below.

Particles at  $\mathbf{a}$  do not move. The transitions for particles at  $v \in V$  are as follows.

- Suppose  $v$  is not the root.
  - Move towards the root at rate 1.
  - Move directly to  $\mathbf{a}$  at rate  $d_v - 2$ .
- Suppose  $v$  is the root. Move to  $\mathbf{a}$  at rate  $d_v - 1$ .

A graphical representation of this model can be obtained by placing an independent Poisson process with rate 1 at each upward directed edge, with rate  $d_v - 2$  at each vertex  $v \neq \rho$ , and with rate  $d_\rho - 1$  at  $\rho$ .

The model enjoys the same monotonicity as the coalescing random walks – resetting to the initially full configuration maximizes the probability of occupying the root in the future. Then, as for the coalescing random walks, there is a dual voter model. In this case, deletion of a particle at  $v$  corresponds to the addition of  $v$  to the cluster of  $\mathbf{a}$ . Note that since the direction of motion is reversed in the voter model, clusters on the tree must expand away from the root. Altogether, the voter model has the following transitions. A *down-going* directed edge  $(w, u)$  is an edge directed away from the root. These are the rules we use in the proof of Lemma 146. We box it for emphasis:

- Along each down-going edge  $(w, u)$ , at rate 1,  $u$  is added to the cluster containing  $w$ .
- At each vertex  $v \neq \rho$  on the tree, at rate  $d_v - 2$ ,  $v$  is added to the cluster of  $\mathbf{a}$ .

- The root is added to the cluster of  $\mathbf{a}$  at rate  $d_\rho - 1$ .

The existence of a dual lets us establish Proposition 142.

*Proof of Proposition 142.* We have established that  $X_T$  in the non-backtracking coalescing random walk has the same distribution as in the simpler model. Since the simpler model has monotonicity and a voter model dual, the same argument used to prove Theorem 148 gives the desired equivalence.  $\square$

### 4.3.5 Site recurrence

Now we can turn our attention to proving infinite expected survival time of the root cluster in the voter model. The voter model from Section 4.3.4 also has the martingale property. As before, let  $\zeta_t^v$  denote the cluster that began at  $v$ .

**Lemma 146.** *Consider the voter model  $\zeta_t^v \in V \cup \{\mathbf{a}\}$  described above. For each  $v \in V$ , so long as  $|\zeta_t^v| < \infty$ , the size of the cluster  $\zeta_t^v$  is a martingale. It transitions to each of the two states  $|\zeta_t| \pm 1$  at rate  $\sum_{w \in \zeta_t} (d_w - 1) - |\{(w, u) : w, u \in \zeta_t^v\}|$ .*

*Proof.* Fix a vertex  $v \in V$  for which we will consider the cluster  $\zeta_t = \zeta_t^v$ . Let  $r_t^+ = r_t^+(v)$  and  $r_t^- = r_t^-(v)$  denote the rate at which  $|\zeta_t| \rightarrow |\zeta_t| \pm 1$ , respectively. Note that the transition rules prohibit  $\mathbf{a} \in \zeta_t$ . Moreover,  $\zeta_t$  is unchanged if we assume that initially, all vertices but  $v$  belong to the cluster of  $\mathbf{a}$ . Therefore, it is enough to check that for any finite  $W$ , if  $\zeta_t = W$  and  $\zeta_t^{\mathbf{a}} = V \cup \{\mathbf{a}\} \setminus W$  then  $r_t^+ - r_t^- = 0$ . For a vertex  $w \notin \{\mathbf{a}, \rho\}$  let  $\hat{w}$  denote its unique parent vertex, i.e., the unique vertex such that there is a down-going edge to  $w$ , and let  $\mathfrak{o}(w)$  denote the set of child vertices. From the transition rules it follows that

$$r_t^+ = \sum_{w \in W} ((d_w - \mathbf{1}(w \neq \rho)) - |\{u \in \mathfrak{o}(w) : u \in W\}|)$$

and

$$r_t^- = \sum_{w \in W} ((d_w - 1 - \mathbf{1}(w \neq \rho)) + \mathbf{1}(\hat{w} \notin W)) = \sum_{w \in W} ((d_w - \mathbf{1}(w \neq \rho)) - \mathbf{1}(\hat{w} \in W))$$

and since  $\sum_{w \in W} |\{u \in \mathfrak{o}(w) : u \in W\}|$  and  $\sum_{w \in W} \mathbf{1}(\hat{w} \in W)$  are both equal to  $|\{(w, u) : u, w \in W\}|$ ,

$$r_t^+ - r_t^- = - \sum_{w \in W} |\{u \in \mathfrak{o}(w) : u \in W\}| + \sum_{w \in W} \mathbf{1}(\hat{w} \in W) = 0.$$

□

*Proof of Theorem 143.* With Lemma 146 we can bound the transition rate of  $|\zeta_t^\rho|$  by  $\sum_{v \in \zeta_t^\rho} d_v$ . To prove the part of Theorem 143 concerning bounded degree trees we can follow the same approach as Theorem 136 (i); again we use the transition rate bound  $D \cdot |\zeta_t^\rho|$ . Similarly, we can use the same technique as the proof of Theorem 136 (ii) to deduce site recurrence for Galton-Watson trees. □

## 5 The power of two choices

### 5.1 Introduction

A sequence in  $[0, 1]$  is *equidistributed* if the limiting proportion of terms in each subinterval is equal to the subinterval's length. Over a century ago Weyl proved that  $\{\beta n \bmod 1\}_{n \geq 1}$  is equidistributed for any irrational number  $\beta$  (see [Wey10]). Since then connections have been found in ergodic theory, number theory, complex analysis and computer science ([BM72], [Vau77], [FSZ09], [CKK<sup>+</sup>07]). See [KN06] for an overview.

More recently attention has been given to equidistribution of random sequences. One way to obtain a random sequence in  $[0, 1]$  is to independently choose points uniformly. Call the resulting sequence the *uniform process*. The law of large numbers guarantees this is equidistributed almost surely.

Another random process known to equidistribute points is the *Kakutani interval splitting procedure* (introduced in [Kak76]), where at each step a point is added uniformly to the current largest subinterval. Almost sure equidistribution is proven in [Zwe78] and [Loo78] using stopping times. Because points are placed in the largest gaps they ought to spread more evenly than the uniform process. Indeed, [Pyk80] proves the size of the largest interval is asymptotic to  $2/n$ ; the same order as the average interval. Compare to  $\log n/n$  in the uniform process (see [Dar53]).

[MP14] introduces a family of interval splitting processes that exhibit a wider range of behavior. The quintessential example is the *max-2 process*. The dynamics are as follows:

- Partition  $[0, 1]$  into subintervals by placing finitely many points in any manner.
- At each step sample two points uniformly from  $[0, 1]$ . Each lies in a subinterval formed by the previous configuration.
- Keep the point contained in the larger subinterval and disregard the other point. Break a tie by flipping a fair coin.

A discrete analogue of the max-2 process appears in [ABKU99] where  $n$  balls are placed into  $n$  bins. For each ball two bins are selected uniformly and the ball is placed in the bin with fewer balls. They find that the most-filled bin has  $\approx \log_2 \log n$  balls; significantly less than  $\approx \log n / \log \log n$  if the balls were instead placed uniformly. This is studied in more detail in [MRS00] and [LM05].

In the max-2 process choosing the larger gap should spread points more evenly. Despite our intuition this is difficult to formalize, and equidistribution was a primary open problem from [MP14].

**Theorem 147.** *The max-2 process is equidistributed a.s.*

The natural counterpart is the *min-2 process* where the point contained in the smaller subinterval is kept. Unlike the previous processes, points are prone to clump together. We show that some random mixtures of the max-2, uniform and min-2 processes are equidistributed.

**Theorem 148.** *Any mixture of max-2, uniform and min-2 processes with probability  $p$  of placing a point uniformly and probability less than or equal to  $.6 - .5p$  of placing a point according to the min-2 process is equidistributed a.s.*

The formal definition of a mixture is in Section ???. As a corollary we state two examples of equidistributed processes intuitively less spread out than the uniform process.

**Example 149.** The following mixtures are equidistributed a.s.

- (i) 60%-min-2 and 40%-max-2.
- (ii) 20%-min-2 and 80%-uniform (see Figure 8).

Two generalizations are the *max- $k$  process* and *min- $k$  process*. In the max- $k$  process keep the point in the largest subinterval among  $k$  uniformly



placed points. Alternatively, in the min- $k$  process keep the point in the smallest subinterval.

Our final theorem (Theorem 1.1 in Section 1.1) informally states that given any collection of max- $k$  and min- $k$  processes there is an equidistributed mixture which places  $\epsilon$  weight on these and the other  $1 - \epsilon$  on the uniform process.

**Example 150.** The following mixtures are equidistributed a.s.

- (i)  $(1/k^2)$ -min- $k$  for fixed  $k$  and otherwise *uniform*.
- (ii)  $99.95$ -uniform and  $(5^{-k})$ -min- $k$  for  $k = 2, 3, \dots$

### 5.1.1 Discussion and further questions

Processes in our theorems satisfy the special inequality at (1.1). The reason our approach works for only certain mixtures is unclear. Numerical methods indicate the inequality fails for other processes, suggesting a different approach is needed. Hopefully, the properties we establish for general mixtures in Theorem 1.1 will aid further progress.

[MP14] conjectures that any max- $k$  or min- $k$  process is equidistributed. Based on this we suspect that any mixture of these is also equidistributed. The rate of convergence to a uniform placement of points and also the asymptotic size of the largest interval are open problems. More thorough discussion can be found in [MP14].

### 5.1.2 Overview

A more general family of interval splitting processes is introduced in [MP14]. Their main result is that, when properly scaled, the empirical distribution of subinterval lengths converges to a distribution function. The idea behind our argument is to reproduce parts of [MP14] when restricted to subintervals contained in  $[0, \alpha]$ . We find that that the empirical distribution of subinterval lengths in  $[0, \alpha]$  evolves to be essentially the same as the unrestricted version on  $[0, 1]$ . This sameness is enough to deduce equidistribution.

This article is organized to quickly arrive at the proof of Theorem 1.1. In Section 5.2 we describe the evolution of  $[0, \alpha]$  and give the major definitions. In Section 5.3 we state without proof several propositions and then prove Theorem 1.1. Section 5.4 contains the proofs for the previous section. In Section 5.5 we generalize to random mixtures. Finally, in Section 5.6 we prove that processes captured by our theorems satisfy the inequality at (1.1).

## 5.2 Subintervals in $[0, \alpha]$

Leading up to Theorem 147 we frame all of our discussion in terms of max- $k$  and min- $k$  processes. The reason being the majority of our propositions hold for any  $k$ . Moreover, we will see in Section ?? that this readily generalizes to random mixtures.

We start with a formal definition for a process to be equidistributed. Suppose  $n_0$  points are initially placed. After  $n$  iterations of an interval splitting process let  $N_n^\alpha$  be the number of the first  $n_0 + n$  terms smaller than  $\alpha$ . We say a sequence is *equidistributed* if  $n^{-1}N_n^\alpha \rightarrow \alpha$  for all  $\alpha \in [0, 1]$ . It is convenient to work in continuous time. Following [MP14] we have points arrive as a Poisson process with intensity  $e^t$ . Formal details are in Section 5.3.1. So, in continuous time equidistribution is equivalent to  $e^{-t}N_t^\alpha \rightarrow \alpha$  for all  $\alpha \in [0, 1]$ .

Fix  $k \geq 2$  and  $\alpha \in [0, 1]$ . We use the convention that a bold face letter represents a process indexed by time (i.e.  $\tilde{\mathbf{A}} = (\tilde{A}_t)_{t \geq 0}$ ). Define the joint processes  $(\tilde{\mathbf{A}}^\alpha, \tilde{\mathbf{A}}^{\alpha+}, \tilde{\mathbf{A}})$  to be the size-biased empirical distributions restricted to intervals contained in  $[0, \alpha]$ ,  $[\alpha, 1]$  and  $[0, 1]$ , respectively. Formally, letting  $I_1^{\alpha, (t)}, \dots, I_{N_t^\alpha}^{\alpha, (t)}$  be the lengths of subintervals contained in  $[0, \alpha]$  at time  $t$  we define

$$\tilde{A}_t^\alpha(x) = \sum_{j=1}^{N_t^\alpha} I_j^{\alpha, (t)} \cdot \mathbf{1}\{I_j^{\alpha, (t)} \leq x\},$$

and similarly for  $\tilde{A}_t^{\alpha+}$  and  $\tilde{A}_t$ . The spark for the refined analysis comes from the relation

$$\tilde{A}_t^\alpha(x) + \tilde{A}_t^{\alpha+}(x) = \tilde{A}_t(x), \quad \forall t, x \geq 0. \quad (57)$$

To ensure that no intervals are double counted assume the initial set of points placed in  $[0, 1]$  always contains  $\{\alpha\}$ . This assumption is only for convenience. Our proof could be adapted to omit it by running the process until two points  $\alpha_1 \leq \alpha \leq \alpha_2$  land sufficiently close to  $\alpha$ , and then using the bound  $N_t^{\alpha_1} \leq N_t^\alpha \leq N_t^{\alpha_2}$ . We further remark that the same reasoning extends our theorems to the unit circle.

For the max- $k$  process define  $\Psi(u) = u^k$  and for the min- $k$  process define  $\Psi(u) = 1 - (1 - u)^k$ . Also, let  $\psi(u) = \Psi'(u)$ . In [MP14, Section 2] the authors prove that

$$\tilde{A}_t(x) = \tilde{A}_0(x) + \int_0^t e^s x^2 \int_x^\infty \frac{\psi(\tilde{A}_s(z))}{z} d\tilde{A}_s(z) + \tilde{M}_t$$

for some martingale  $\widetilde{M}_t$ . The following proposition shows that  $\widetilde{A}_t^\alpha$  satisfies a similar equation.

**Proposition 151.** *For any max- $k$  or min- $k$  process, the joint processes  $(\widetilde{\mathbf{A}}^\alpha, \widetilde{\mathbf{A}}^{\alpha+}, \widetilde{\mathbf{A}})$  satisfy the equation*

$$\widetilde{A}_t^\alpha(x) = \widetilde{A}_0^\alpha(x) + \int_0^t e^s x^2 \int_x^\infty \frac{\psi(\widetilde{A}_s(z))}{z} d\widetilde{A}_s^\alpha(z) ds + \widetilde{M}_t^\alpha(x),$$

with  $\widetilde{M}_t^\alpha$  a martingale.

The similarity between the semimartingale decompositions of  $\widetilde{A}_t$  and  $\widetilde{A}_t^\alpha$  is paramount in obtaining our theorems. However, the details are a bit technical. To keep our momentum we delay the proof until Section ???. What follows are facts and notation essential to our main theorems.

Let non-tilde processes represent the original process scaled by  $e^{-t}$  (i.e.  $A_t(x) = \widetilde{A}_t(e^{-t}x)$ ). In light of Proposition 151, a change of variables gives the relationship

$$\mathbf{A}^\alpha = \mathcal{C}(\mathbf{A}^\alpha, \mathbf{A}) + \mathbf{M}^\alpha, \quad (58)$$

where  $\mathcal{C}: \mathcal{X} \times \mathcal{X} \rightarrow C([0, \infty), L_{\text{loc}}^1)$  is defined by

$$\mathcal{C}(\mathbf{F}, \mathbf{G})_t(x) = F_0(e^{-t}x) + \int_0^t (e^{s-t}x)^2 \int_{e^{s-t}x}^\infty \frac{\psi(G_s(z))}{z} dF_s(z) ds.$$

Here  $\mathcal{X} = \mathcal{B}([0, \infty), \mathcal{D})$  where  $\mathcal{D} = \{F: [0, \infty) \rightarrow [0, 1], \text{c\acute{a}dl\acute{a}g, increasing}\}$ . The set  $\mathcal{X}$  is a subspace of the space  $\mathcal{B}([0, \infty), L_{\text{loc}}^1)$  of measurable maps from  $[0, \infty)$  to  $L_{\text{loc}}^1$  with the topology of locally uniform convergence, which we denote by the symbol  $\xrightarrow{\mathcal{X}}$ .

We will use  $\hat{F}$  and  $F^\Psi$  interchangeably to denote the a.s. pointwise limiting distribution of  $A_t$  from [MP14, Theorem 1.1]. Also define the stationary distribution  $\hat{\mathbf{F}}^*$  so that  $\hat{F}_t^* = \hat{F}$  for all  $t \geq 0$ . With the convergence  $A_t \rightarrow \hat{F}$  in mind, we consider the operator

$$\mathcal{C}^*(\mathbf{F})_t = \mathcal{C}(\mathbf{F}, \hat{\mathbf{F}}^*)_t = F_0(e^{-t}x) + \int_0^t (e^{s-t}x)^2 \int_{e^{s-t}x}^\infty \frac{\psi(\hat{F}(z))}{z} dF_s(z) ds.$$

We will see in the proof of Theorem 147 that the limiting distribution of  $A_t^\alpha$  belongs to the set of fixed points

$$\mathfrak{F}^\alpha = \{\mathbf{F} \in \mathcal{X}_1: \mathbf{F} = \mathcal{C}^*(\mathbf{F}), F_t(+\infty) = \alpha \text{ and } (\frac{1}{\alpha}F_t)_{t \geq 0} \text{ tight}\}.$$

Here  $\mathcal{X}_1 = \mathcal{B}([0, \infty), \{F \in \mathcal{D}: \|F\|_{x^{-2}} \leq 1\})$ . We will use the following two norms on  $L_{\text{loc}}^1([0, \infty))$ :

$$\|f\|_{x^{-2}} = \int_0^\infty x^{-2}|f(x)|dx \quad \text{and} \quad \|f\|_{x^{-3/2}} = \int_0^\infty x^{-3/2}|f(x)|dx.$$

### 5.3 Proof of Theorem 147

We delay the proofs of two propositions until the next section. The first gives a sufficient condition for processes in  $\mathfrak{F}^\alpha$  to converge to  $\alpha\hat{F}$  in  $\|\cdot\|_{x^{-3/2}}$ . This is paired with a lemma proving the condition is met for the max-2 process.

**Proposition 152.** *If for all  $z \geq 0$*

$$\frac{2}{3}|z\psi'(\hat{F}(z))\hat{F}'(z) - \psi(\hat{F}(z))| - \psi(\hat{F}(z)) \leq 0, \quad (59)$$

*then  $\|F_t - \alpha\hat{F}\|_{x^{-3/2}} \leq 6e^{-t/2}$  for all  $\mathbf{F} \in \mathfrak{F}^\alpha$ .*

**Lemma 153.** *The function  $\hat{F}$  from the max-2 process satisfies (59) for all  $z \geq 0$ .*

*Proof.* This is the case  $p_2 = 1$  covered by ?THM? ??. □

We will also need several properties of max- $k$  and min- $k$  processes. Direct analogues hold for random mixtures (see ?THM? ??).

**Proposition 154.** *For the max- $k$  and min- $k$  processes:*

- (I)  $\|A_t^\alpha\|_{x^{-2}} = e^{-t}N_t^\alpha$  and  $\|\alpha\hat{F}\|_{x^{-2}} = \alpha$ .
- (II) *The collection of distribution functions  $(\frac{1}{\alpha}A_t^\alpha)_{t \geq 0}$  is tight.*
- (III) *The family  $(\mathbf{A}^{\alpha, (n)})$  defined by  $A_t^{\alpha, (n)} = A_{t+n}^\alpha$  is asymptotically equicontinuous.*
- (IV)  $\mathbf{M}^{\alpha, (n)} \xrightarrow{\mathcal{X}} 0$  as  $n \rightarrow \infty$ , where  $M_t^{\alpha, (n)}(x) = M_{t+n}^\alpha(x) - M_n^\alpha(e^{-t}x)$  for every  $t \geq 0$ .
- (V) *Define  $\mathbf{A}^{(n)}$  by  $A_t^{(n)} = A_{t+n}$ . If  $\mathbf{F}^{(n)} \xrightarrow{\mathcal{X}} \mathbf{F}$  then  $\mathcal{C}(\mathbf{F}^{(n)}, \mathbf{A}^{(n)}) \xrightarrow{\mathcal{X}} \mathcal{C}^*(\mathbf{F})$ .*

*Proof of Theorem 147.* All statements are meant to hold almost surely. Also we abbreviate items from Proposition 154 as a roman numeral. In the continuous process points are added as a Poisson process with intensity  $e^t dt$ . So, it suffices to show  $e^{-t}N_t^\alpha \rightarrow \alpha$ .

By (II), (III) and the version of the Arzelá-Ascoli theorem in [MP14, Lemma 7.3] we may choose a sequence  $(\mathbf{A}^{\alpha, (n_k)})$  which converges to a family of (scaled by  $\alpha$ ) distributions  $\mathbf{F}^{\alpha, (\infty)}$  with  $F_t^{\alpha, (\infty)}(+\infty) = \alpha$  for every  $t \geq 0$ . Taking limits in the formula at (58) we obtain

$$\mathcal{C}(\mathbf{A}^{\alpha, (n_k)}, \mathbf{A}^{(n_k)}) + \mathbf{M}^{\alpha, (n_k)} \xrightarrow{\mathcal{X}} \mathbf{F}^{\alpha, (\infty)}.$$

By (IV) and (V) we have

$$\mathcal{C}(\mathbf{A}^{\alpha, (n_k)}, \mathbf{A}^{(n_k)}) \xrightarrow{\mathcal{X}} \mathcal{C}^*(\mathbf{F}^{\alpha, (\infty)}).$$

Thus,  $\mathbf{F}^{\alpha, (\infty)}$  is in  $\mathfrak{F}^\alpha$ . Proposition 152 and Lemma 153 imply that  $\|F_t^{\alpha, (\infty)} - \alpha \hat{F}\|_{x^{-3/2}} \leq 6e^{-t/2}$ . A similar argument as the conclusion of the proof of [MP14, Theorem 7.1] gives almost sure pointwise convergence  $A_t^\alpha \rightarrow \alpha \hat{F}$ . [MP14, Theorem 1.1] states that  $A_t \rightarrow \hat{F}$  pointwise. We can then deduce from (57) that  $A_t^{\alpha+} \rightarrow (1-\alpha)\hat{F}$ . Combining pointwise convergence, (57) and Fatou's lemma we deduce that  $\|A_t^\alpha\|_{x^{-2}} \rightarrow \|\alpha \hat{F}\|_{x^{-2}}$ . Indeed,

$$\begin{aligned} \liminf \|A_t^\alpha\|_{x^{-2}} &\geq \|\alpha \hat{F}\|_{x^{-2}}, \\ \limsup \|A_t^\alpha\|_{x^{-2}} &= 1 - \liminf \|A_t^{\alpha+}\|_{x^{-2}} \leq 1 - (1-\alpha) = \|\alpha \hat{F}\|_{x^{-2}}. \end{aligned}$$

This finishes the proof since (I) states that  $\|A_t^\alpha\|_{x^{-2}} = e^{-t} N_t^\alpha$  and  $\|\alpha \hat{F}\|_{x^{-2}} = \alpha$ .  $\square$

### 5.3.1 Proposition 151

The idea is to compare subintervals selected from  $[0, \alpha]$  against subintervals from  $[\alpha, 1]$ . For example, in an iteration of the max- $k$  process we consider the event that  $j$  of the  $k$  candidate points land in  $[0, \alpha]$ . The other  $k-j$  must land in  $[\alpha, 1]$ . We obtain an interval selected from  $[0, \alpha]$  according to a max- $j$  process and another selected from  $[\alpha, 1]$  according to a max- $(k-j)$  process. If the interval from  $[0, \alpha]$  is larger, then the point is kept and accounted for by  $A_t^\alpha$ . The relationship at (57) and a combinatorial identity then yield the desired formula.

Before giving the proof we first build up some necessary definitions. Our construction is for the max- $k$  process. The definitions for the min- $k$  process are similar. Define

$$\Psi_j^\alpha(u) = (u/\alpha)^j$$

to be the distribution function for the maximum and minimum of  $j$  independent Uniform $[0, \alpha]$  random variables. We use the convention that  $\Psi_j(u) = \Psi_j^1(u)$  and  $\psi(u) = \Psi_k'(u)$ .

Let  $\xi \sim \text{Bin}(k, \alpha)$  with  $q_j = \mathbf{P}[\xi = j] = \binom{k}{j} \alpha^j (1-\alpha)^{k-j}$ . Define a Poisson random measure  $\prod$  on

$$[0, \infty) \times [0, 1] \times \{0, 1, \dots, k\} \times [0, \alpha]^k \times [0, 1-\alpha]^k$$

with intensity

$$e^t dt \otimes dv \otimes d\xi \otimes_{j=1}^k (d\Psi_j^\alpha \otimes d\Psi_{k-j}^{1-\alpha}).$$

Denote points in  $\Pi$  by  $\pi_s = (s, v, \xi, \mathbf{u}, \mathbf{w})$  where  $\mathbf{u} = u_1, \dots, u_k$ ,  $\mathbf{w} = w_0, \dots, w_{k-1}$ .

Let  $l_s^\alpha(u) = (\tilde{A}_s^\alpha)^{-1}(u)$  and  $l_s^{\alpha+}(w) = (\tilde{A}_s^{\alpha+})^{-1}(w)$  be intervals sampled from  $[0, \alpha]$  and  $[\alpha, 1]$ , respectively. Also, let  $h(v, l, x) = v\mathbf{1}\{lv \leq x\} + (1-v)\mathbf{1}\{l(1-v) \leq x\}$ . With this define

$$\tilde{B}^\alpha(\pi_s, x) = \sum_{j=1}^k \mathbf{1}\{\xi = j\} l_s^\alpha(u_j) \mathbf{1}\{l_s^\alpha(u_j) > x\} \mathbf{1}\{l_s^\alpha(u_j) > l_s^{\alpha+}(w_{k-j})\} h(v, l_s^\alpha(u_j), x),$$

so that  $\tilde{A}_t^\alpha(x) = \tilde{A}_0^\alpha(x) + \sum_{\pi_s \in \Pi, s \leq t} \tilde{B}^\alpha(\pi_s, x)$ . To help clarify we give a brief explanation for each term:

- $\mathbf{1}\{\xi = j\}$  accounts for how many of the  $k$  points land in  $[0, \alpha]$ .
- $l_s^\alpha(u_j)$  is the initial length of the interval that potentially will be added to  $\tilde{A}_t^\alpha$ .
- $\mathbf{1}\{l_s^\alpha(u_j) > x\}$  is zero if the selected interval is smaller than  $x$  and is already being counted by  $\tilde{A}_t^\alpha$ .
- $\mathbf{1}\{l_s^\alpha(u_j) > l_s^{\alpha+}(w_{k-j})\}$  indicates whether the interval from  $[0, \alpha]$  is larger than that from  $[\alpha, 1]$ . The inequality would be reversed for the min- $k$  process.
- $h(v, l_s^\alpha(u_j), x)$  “cuts” the interval  $l_s^\alpha$  and detects whether the resulting subintervals are smaller than  $x$  and so, should be added to  $\tilde{A}_t^\alpha$ .

*Proof of Proposition 151.* Our proof is for the max- $k$  process, the argument for the min- $k$  process is similar (the functions  $\Psi$  would change as would the bounds of the inside integral at (62)). We will obtain the semi-martingale decomposition of  $\tilde{A}_t^\alpha(x) = \tilde{A}_0^\alpha(x) + \sum_{\pi_s \in \Pi, s \leq t} \tilde{B}^\alpha(\pi_s, x)$  with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$  where  $\mathcal{F}_t = \sigma(\Pi|_{[0,t] \times [0,1] \times \{0,1,\dots,k\} \times [0,\alpha]^k \times [0,1-\alpha]^k})$ .

We start by computing the integral

$$\int_0^\alpha \int_0^{1-\alpha} \int_0^1 \int \tilde{B}^\alpha(\pi_s, x) d\xi dv d\Psi_{k-j}^{1-\alpha}(w_{k-j}) d\Psi_j^\alpha(u_j). \quad (60)$$

Using the fact that  $\int_0^1 h(v, l, x) dv = (x/l)^2$ , we first integrate with respect to  $\xi$  and  $v$  to write as

$$\sum_{j=1}^k q_j x^2 \int_0^\alpha \int_0^{1-\alpha} \frac{1}{l_s^\alpha(u_j)} \mathbf{1}\{l_s^\alpha(u_j) > x\} \mathbf{1}\{l_s^\alpha(u_j) > l_s^{\alpha+}(w_{k-j})\} d\Psi_{k-j}^{1-\alpha}(w_{k-j}) d\Psi_j^\alpha(u_j).$$

Integrating one step further and normalizing the  $\Psi_j^\alpha$  and  $\Psi_{k-j}^{1-\alpha}$  to  $\Psi_j$  and  $\Psi_{k-j}$  we obtain factors of  $\alpha^{-j}(1-\alpha)^{-(k-j)}$ . This lets us cancel all but the binomial coefficients from the  $q_j$  terms and obtain

$$\sum_{j=1}^k x^2 \int_0^\alpha \int_0^{1-\alpha} \frac{\mathbf{1}\{l_s^\alpha(u_j) > x\}}{l_s^\alpha(u_j)} \mathbf{1}\{l_s^\alpha(u_j) > l_s^{\alpha+}(w_{k-j})\} d\Psi_{k-j}(w_{k-j}) d\Psi_j(u_j). \quad (61)$$

Make the change of variables  $y = l_s^{\alpha+}(w_j)$  so that  $\tilde{A}_s^{\alpha+}(y) = w_{j-k}$ . Hence the above can be written as

$$\sum_{j=1}^k x^2 \int_0^\alpha \frac{\mathbf{1}\{l_s^\alpha(u_j) > x\}}{l_s^\alpha(u_j)} \int_0^{l_s^\alpha(u_j)} d\Psi_{k-j}(\tilde{A}_s^{\alpha+}(y)) d\Psi_j(u_j). \quad (62)$$

Integrate one step further and use the fact that  $\Psi_{k-j}(0) = 0$  to obtain

$$\sum_{j=1}^k x^2 \int_0^\alpha \frac{\mathbf{1}\{l_s^\alpha(u_j) > x\} \Psi_{k-j}(\tilde{A}_s^{\alpha+}(l_s^\alpha(u_j)))}{l_s^\alpha(u_j)} d\Psi_j(u_j).$$

Now apply the change of variables  $z = l_s^\alpha(u_j)$  and so  $\tilde{A}_s^\alpha(z) = u_j$  to rewrite the above as

$$x^2 \int_x^\infty \sum_{j=1}^k \binom{k}{j} \frac{\Psi_{k-j}(\tilde{A}_s^{\alpha+}(z))}{z} d\Psi_j(\tilde{A}_s^\alpha(z)).$$

Writing out  $\Psi_{k-j}(u) = u^{k-j}$  and  $\psi_j(u) = ju^{j-1}$  and using the equality  $\tilde{A}_s^{\alpha+}(z) = \tilde{A}_s(z) - \tilde{A}_s^\alpha(z)$  from (57) we can rewrite the above as

$$x^2 \int_x^\infty \sum_{j=1}^k \binom{k}{j} \frac{(\tilde{A}_s(z) - \tilde{A}_s^\alpha(z))^{k-j}}{z} j(\tilde{A}_s^\alpha)^{j-1} d\tilde{A}_s^\alpha(z).$$

The identity  $\sum_{j=1}^k \binom{k}{j} (a-b)^{k-j} j b^{j-1} = ka^{k-1} = \psi(a)$  (derived by applying the binomial theorem to  $a^k = ((a-b) + b)^k$  then differentiating both sides with respect to  $a$ ) gives (60) is equal to

$$x^2 \int_x^\infty \frac{\psi(\tilde{A}_s(z))}{z} d\tilde{A}_s^\alpha(z).$$

Finish by multiplying by  $e^s$  and integrating from 0 to  $t$ .  $\square$

### 5.3.2 Proposition 152

The proof of Proposition 152 proceeds analogously to [MP14, Lemma 4.1 and Proposition 3.4]. A significant difference is that they apply integration by parts to

$$\frac{1}{z}d\Psi(\tilde{F}_s(z)),$$

whereas our operator  $\mathcal{C}^*$  requires applying integration by parts to

$$\frac{\psi(\hat{F}(z))}{z}d\tilde{F}_s(z).$$

The requirement at (59) arises from the extra term  $\psi(\hat{F}(z))$ . Also, note that we work in the norm  $\|\cdot\|_{x^{-3/2}}$  to obtain the constant  $\frac{2}{3}$  in (59). We will need this factor to prove an inequality similar to (59) holds for processes biased towards min-2. In [MP14] they use the norm  $\|\cdot\|_{x^{-2}}$ . This change of norms does not significantly alter the argument. In fact, we could equally well work with any norm  $\|\cdot\|_{x^{-1-\delta}}$  with  $0 < \delta < 1$ .

*Proof of Proposition 152.* Let  $\mathbf{F} \in \mathfrak{F}^\alpha$ . We consider the rescaled processes  $\tilde{F}_t(x) = F(e^t x)$ ,  $\tilde{F}_t^\Psi(x) = \hat{F}(e^t x)$ . It then holds that  $\tilde{\mathbf{F}} = \mathcal{C}(\tilde{\mathbf{F}})$  where

$$\tilde{\mathcal{C}}(\tilde{\mathbf{F}})_t(x) = \tilde{F}_0(x) + \int_0^t e^s x^2 \int_x^\infty \frac{\psi(\hat{F}(z))}{z} d\tilde{F}_s(z) ds.$$

We seek to prove the distance between  $\tilde{\mathbf{F}}$  and  $\alpha\tilde{\mathbf{F}}^*$  is decreasing in  $t$ :

$$\partial_t \|\tilde{F}_t - \alpha\tilde{F}_t^\Psi\|_{x^{-3/2}} = \int_0^\infty x^{-3/2} \partial_t |\tilde{F}_t(x) - \alpha\tilde{F}_t^\Psi(x)| dx \leq 0. \quad (63)$$

Taking this for granted, we use the equality

$$\partial_t \tilde{\mathcal{C}}(\tilde{\mathbf{F}})_t(x) = e^t x^2 \int_x^\infty \frac{\psi(\hat{F}(z))}{z} d\tilde{F}_t(z)$$

to write for each  $x \geq 0$  the dynamics for the difference  $\tilde{F}_t(x) - \alpha\tilde{F}_t^\Psi(x)$  as

$$\partial_t (\tilde{F}_t(x) - \alpha\tilde{F}_t^\Psi(x)) = e^t x^2 I_t(x),$$

$$I_t(x) = \int_x^\infty \frac{\psi(\hat{F}(z))}{z} \partial_z (\tilde{F}_t(z) - \alpha\tilde{F}_t^\Psi(z)) dz.$$



Multiply both sides by  $\text{sgn}(\tilde{F}_t - \alpha\tilde{F}_t^\Psi)$  to obtain

$$e^{-t}\partial_t|\tilde{F}_t(x) - \alpha\tilde{F}_t^\Psi(x)| = x^2 \begin{cases} \text{sgn}(\tilde{F}_t(x) - \alpha\tilde{F}_t^\Psi(x))I_t(x), & \tilde{F}_t(x) \neq \alpha\tilde{F}_t^\Psi(x) \\ 0, & \tilde{F}_t(x) = \alpha\tilde{F}_t^\Psi(x) \end{cases}.$$

Let  $\hat{f}(z) = z\psi'(\hat{F}(z))\hat{F}'(z) - \psi(\hat{F}(z))$ . An application of integration by parts to the integral gives

$$I_t(x) = -\frac{\psi(\hat{F}(x))}{x}(\tilde{F}_t(x) - \alpha\tilde{F}_t^\Psi(x)) + \int_x^\infty \frac{\hat{f}(z)}{z^2}(\tilde{F}_t(z) - \alpha\tilde{F}_t^\Psi(z))dz.$$

The previous two equations therefore yield

$$e^{-t}\partial_t|\tilde{F}_t(x) - \alpha\tilde{F}_t^\Psi(x)| \leq -x\psi(\hat{F}(x))|\tilde{F}_t(x) - \alpha\tilde{F}_t^\Psi(x)| + x^2 \int_x^\infty |\hat{f}(z)| \frac{|\tilde{F}_t(z) - \alpha\tilde{F}_t^\Psi(z)|}{z^2} dz.$$

We next multiply both sides by  $x^{-3/2}$  and integrate with respect to  $x$  from 0 to infinity to obtain the bound

$$\begin{aligned} e^{-t} \int_0^\infty x^{-3/2} \partial_t |\tilde{F}_t(x) - \alpha\tilde{F}_t^\Psi(x)| dx &\leq \int_0^\infty -\psi(\hat{F}(x)) \frac{|\tilde{F}_t(x) - \alpha\tilde{F}_t^\Psi(x)|}{x^{1/2}} dx \\ &\quad + \int_0^\infty x^{1/2} \int_x^\infty |\hat{f}(z)| \frac{|\tilde{F}_t(z) - \alpha\tilde{F}_t^\Psi(z)|}{z^2} dz dx. \end{aligned}$$

An application of Fubini's theorem lets us rewrite the second integral as

$$\begin{aligned} \int_0^\infty x^{1/2} \int_x^\infty |\hat{f}(z)| \frac{|\tilde{F}_t(z) - \alpha\tilde{F}_t^\Psi(z)|}{z^2} dz dx &= \int_0^\infty |\hat{f}(z)| \frac{|\tilde{F}_t(z) - \alpha\tilde{F}_t^\Psi(z)|}{z^2} \int_0^z x^{1/2} dx dz \\ &= \int_0^\infty \frac{2}{3} |\hat{f}(z)| \frac{|\tilde{F}_t(z) - \alpha\tilde{F}_t^\Psi(z)|}{z^{1/2}} dz. \end{aligned}$$

Hence we can combine the integrals to obtain the bound

$$e^{-t} \int_0^\infty x^{-2} \partial_t |\tilde{F}_t(x) - \alpha\tilde{F}_t^\Psi(x)| dx \leq \int_0^\infty \left( \frac{2}{3} |\hat{f}(z)| - \psi(\hat{F}(z)) \right) \frac{|\tilde{F}_t(z) - \alpha\tilde{F}_t^\Psi(z)|}{z^{1/2}} dz.$$

The above is less than or equal to zero by our hypothesis (59). This establishes that

$$\|\tilde{F}_t - \alpha\tilde{F}_t^\Psi\|_{x^{-2}} \leq \|\tilde{F}_0 - \alpha\tilde{F}_0^\Psi\|_{x^{-3/2}} = \|F_0 - \alpha\hat{F}\|_{x^{-3/2}}. \quad (64)$$

A change of variables  $x = e^{-t}z$  gives

$$\begin{aligned}
\|F_t - \alpha\hat{F}\|_{x^{-3/2}} &= \int_0^\infty x^{-3/2} |F_t(x) - \alpha\hat{F}(x)| dx \\
&= e^{-t/2} \int_0^\infty z^{-3/2} |\tilde{F}_t(z) - \alpha\tilde{F}_t^\Psi(z)| dz \\
&= e^{-t/2} \|\tilde{F}_t - \alpha\tilde{F}_t^\Psi\|_{x^{-3/2}} \\
&\leq e^{-t/2} \|F_0 - \alpha\hat{F}\|_{x^{-3/2}}, \tag{65}
\end{aligned}$$

where at the last line we apply (64).

It remains to prove that  $\|F_0 - \alpha\hat{F}\|_{x^{-3/2}} \leq 6$ . By assumption,  $\mathbf{F} \in \mathcal{X}_1$  and therefore  $\|F_0\|_{x^{-2}} \leq 1$ . As  $0 \leq F_0(x) \leq 1$  we can break up the integral and use integrability of  $x^{-3/2}\mathbf{1}\{x > 1\}$ :

$$\int_0^\infty x^{-3/2} F_0(x) dx \leq \int_0^1 x^{-2} F_0(x) dx + \int_1^\infty x^{-3/2} dx \leq \|F_0\|_{x^{-2}} + 2 \leq 3.$$

Similarly,  $\|\alpha\hat{F}\|_{x^{-3/2}} \leq 3$ . Apply the triangle inequality to conclude  $\|F_0 - \alpha\hat{F}\|_{x^{-3/2}} \leq \|F_0\|_{x^{-3/2}} + \|\alpha\hat{F}\|_{x^{-3/2}} \leq 6$ .  $\square$

### 5.3.3 Proposition 154

In Proposition 154 we prove that  $A_t^\alpha$  and  $A_t$  have similar properties. Each statement requires some manipulation. Fortunately [MP14] contains much of the ‘heavy-lifting’. We make one remark concerning the proof of (V). In [MP14] they prove continuity of an operator  $\mathcal{S}^\Psi$  with domain  $\mathcal{X}$ . Our operator  $\mathcal{C}$  has domain  $\mathcal{X} \times \mathcal{X}$ . This makes the proof more involved, and also restricts us to proving continuity in sequences of the form  $(\mathbf{F}^{(n)}, \mathbf{A}^{(n)})$ .

*Proof of (I).* The equality  $\|\alpha\hat{F}\|_{x^{-2}} = \alpha$  is [MP14, Lemma 3.5]. For the other equality, take  $I_j^{\alpha,(t)}$  to be the length of an interval in  $[0, \alpha]$ . Define the measure  $\mu_t^\alpha = e^{-t} \sum_1^{N_t^\alpha} \delta_{e^t I_j^{\alpha,(t)}}$ . This gives  $\mu_t^\alpha$  is the empirical distribution of rescaled interval lengths. We can then write

$$A_t^\alpha(x) = \int_0^x y \mu_t^\alpha(dy).$$

Applying Fubini’s theorem shows that

$$\|A_t^\alpha\|_{x^{-2}} = \int_0^\infty x^{-2} \int_0^x y \mu_t^\alpha(dy) = \int_0^\infty \mu_t^\alpha(dy) = e^{-t} N_t^\alpha.$$

$\square$

*Proof of (II).* Recall that a family of distributions  $(F_t)_{t \geq 0}$  is *tight* if for all  $\epsilon > 0$  there exists  $N$  such that  $F_t(N) \geq 1 - \epsilon$  for all  $t \geq 0$ . [MP14, Proposition 6.3] implies  $(A_t)_{t \geq 0}$  is tight. Fix  $\epsilon > 0$  and let  $N$  be such that  $A_t(N) \geq 1 - \alpha\epsilon$  for all  $t \geq 0$ . The relationship at (57) ensures  $A_t^\alpha(N) + A_t^{\alpha+}(N) \geq 1 - \alpha\epsilon$ . As  $A_t^\alpha \leq \alpha$  and  $A_t^{\alpha+} \leq 1 - \alpha$ , this inequality could only hold if  $A_t^\alpha(N) \geq \alpha - \alpha\epsilon$  for all  $t \geq 0$ . Hence,  $(\frac{1}{\alpha}A_t^\alpha)_{t \geq 0}$  is tight.  $\square$

*Proof of (III).* We say that a family of functions  $(\mathbf{F}^{(n)})_{n \in \mathbb{N}}$  in  $\mathcal{X}$  is *asymptotically equicontinuous* if for every compact  $K \subset [0, \infty)$ ,

$$\lim_{\delta \rightarrow 0} \lim_{n \rightarrow \infty} \sup_{|s-t| \leq \delta} \int_K |F_s^{(n)}(x) - F_t^{(n)}(x)| dx = 0.$$

The proof is similar to [MP14, Lemma 7.5]. We omit the details and just remark that for any  $\delta > 0$  the number of points kept in  $[0, \alpha]$  from time  $t$  to  $t + \delta$  is bounded by the number of points added to  $[0, 1]$  in that same time interval. Formally, for any  $\delta > 0$  we have  $N_{t+\delta}^\alpha - N_t^\alpha \leq N_{t+\delta}^1 - N_t^1$ . This lets us use the same bounds.  $\square$

*Proof of (IV).* The proof is similar to the decay of the noise subsection in [MP14, Section 7]. The same bounds apply because points are added to  $[0, \alpha]$  no faster than to  $[0, 1]$ . This ensures that  $\tilde{B}^\alpha(\pi_s, x) \leq \tilde{B}(s, u, v, x)$ . Here  $\tilde{B}(s, u, v, x)$  is the function defined at [MP14, (3)].  $\square$

*Proof of (V).* Suppose that  $\mathbf{F}^{(n)} \xrightarrow{\mathcal{X}} \mathbf{F}$ . An equivalent notion of convergence in the topology of local uniform convergence is that  $\mathbf{F}^{(n)} \xrightarrow{\mathcal{X}} \mathbf{F}$  if and only if for all compact  $K \subset [0, \infty)$

$$\lim_{n \rightarrow \infty} \sup_{0 \leq s \leq t} \int_K |F_s^{(n)}(x) - F_s(x)| dx = 0.$$

[MP14, Theorem 7.1] implies  $\mathbf{A}^{(n)} \xrightarrow{\mathcal{X}} \mathbf{F}^*$ . Thus it suffices to prove for any fixed  $T > 0$  and  $K > 0$

$$\int_0^K |\mathcal{C}(\mathbf{F}, \mathbf{F}^*)_t(x) - \mathcal{C}(\mathbf{F}^{(n)}, \mathbf{A}^{(n)})_t(x)| dx \rightarrow 0 \quad (66)$$

uniformly for  $t \leq T$ . For fixed  $n$  we can write

$$\mathcal{C}(\mathbf{F}^{(n)}, \mathbf{A}^{(n)})_t(x) = F_0^{(n)}(x) + \int_0^t (e^{s-t}x)^2 \int_{e^{s-t}x}^\infty \frac{\psi(A_s^{(n)}(z))}{z} dF_s^{(n)}(z) ds.$$

If we write  $\psi(A_s^{(n)}(z)) = \psi(\hat{F}(z)) + \psi(A_s^{(n)}(z)) - \psi(\hat{F}(z))$  the above becomes

$$\mathcal{C}(\mathbf{F}^{(n)}, \mathbf{A}^{(n)})_t(x) = \mathcal{C}(\mathbf{F}^{(n)}, \mathbf{F}^*)_t(x) + \int_0^t (e^{s-t}x)^2 \int_{e^{s-t}x}^\infty \frac{\psi(A_s^{(n)}(z)) - \psi(\hat{F}(z))}{z} dF_s^{(n)}(z) ds.$$

We can then bound the left side of (66) by

$$\int_0^K |\mathcal{C}(\mathbf{F}, \mathbf{F}^*)_t(x) - \mathcal{C}(\mathbf{F}^{(n)}, \mathbf{F}^*)_t(x)| dx \quad (67)$$

$$+ \int_0^K \int_0^t (e^{s-t}x)^2 \int_{e^{s-t}x}^\infty \frac{|\psi(A_s^{(n)}(z)) - \psi(\hat{F}(z))|}{z} dF_s^{(n)}(z) ds dx. \quad (68)$$

It suffices to show that as  $n \rightarrow \infty$  each summand converges to zero uniformly for  $t \leq T$ .

### First summand

Start by bounding the summand at (67) by

$$\int_0^K |F_0(e^{-t}x) - F_0^{(n)}(e^{-t}x)| dx + \int_0^K \int_0^t (e^{s-t}x)^2 \left| \int_{e^{s-t}x}^\infty \frac{\psi(\hat{F}(z))}{z} d(F_s(z) - F_s^{(n)}(z)) \right| ds dx.$$

The first quantity goes to zero uniformly for  $t \leq T$  by the definition of  $\mathbf{F}^{(n)} \xrightarrow{\mathcal{X}} \mathbf{F}$  since a change of variables gives

$$\int_0^K |F_0(e^{-t}x) - F_0^{(n)}(e^{-t}x)| dx \leq e^t \int_0^K |F_0(x) - F_0^{(n)}(x)| dx.$$

Expand the interior of the second quantity with integration by parts and take the absolute value signs inside to bound it by

$$\underbrace{\frac{\psi(\hat{F}(e^{s-t}x))}{e^{s-t}x} |F_s(e^{s-t}x) dx - F_s^{(n)}(e^{s-t}x)|}_{\text{term one}} + \underbrace{\int_{e^{s-t}x}^\infty \left| \frac{d}{dz} \frac{\psi(\hat{F}(z))}{z} \right| |F_s(z) - F_s^{(n)}(z)| dz dx}_{\text{term two}}.$$

Multiply term one by  $(e^{s-t}x)^2$  and integrate so it becomes

$$\int_0^K \int_0^t (e^{s-t}x) \psi(\hat{F}(e^{s-t}x)) |F_s(e^{s-t}x) - F_s^{(n)}(e^{s-t}x)| ds dx.$$

Notice that  $\psi(u) \leq k$  for all  $u \in [0, 1]$ . Thus, the above is bounded by

$$k \int_0^K \int_0^t (e^{s-t}x) |F_s(e^{s-t}x) - F_s^{(n)}(e^{s-t}x)| dx.$$

This puts us in the case of  $I_1$  from [MP14, Lemma 3.3] and so converges to zero uniformly for  $t \leq T$ . As for term two, we differentiate to rewrite it as

$$\int_{e^{s-t}x}^\infty \frac{|z\psi'(\hat{F}(z))\hat{F}'(z) - \psi(\hat{F}(z))|}{z^2} |F_s(z) - F_s^{(n)}(z)| dz. \quad (69)$$

By ?THM? ?? we know that  $zF'(z)$  is bounded. Since  $\psi$  and  $\psi'$  are also bounded we have  $C = \sup_{0 \leq z \leq \infty} |zF'(z)\psi'(\hat{F}(z)) - \psi(\hat{F}(z))| < \infty$ . Therefore, (69) is less than

$$C \int_{e^{s-t}x}^\infty \frac{1}{z^2} |F_s(z) - F_s^{(n)}(z)| dz. \quad (70)$$

Finally we are in the position of  $I_2$  from [MP14, Lemma 3.3] and can conclude that (70) goes to zero uniformly for  $t \leq T$ .

## Second summand

Fix  $M > 0$  and for any function  $f : [0, \infty) \rightarrow [0, 1]$  define  $f^M = f|_{[0, M]}$  to be the restriction to the domain  $[0, M]$ . We have in [MP14, Theorem 7.1] that  $A^M$  converges pointwise to  $\hat{F}^M$ . Observe that each  $A_t^M$  is an increasing function with compact domain, and  $\hat{F}^M$  is continuous by [MP14, Lemma 3.5]. Together these imply (see [Rud76, exercise 7.13]) that for any  $\epsilon > 0$  there exists  $t_\epsilon$  such that for all  $z \in [0, M]$

$$\sup_{t \geq t_\epsilon} |A_t^M(z) - \hat{F}_t^M(z)| < \epsilon.$$

Because the functions  $A_t^{(n)}$  are translates of  $A_t$  it follows that for all  $n > t_\epsilon$  we have

$$\sup_{t \geq 0} |A_t^{(n), M}(z) - \hat{F}_t^M(z)| \leq \sup_{t \geq t_\epsilon} |A_t^M(z) - \hat{F}_t^M(z)| < \epsilon.$$

As  $\psi$  is uniformly continuous we conclude there exists  $n_0$  such that for all  $z \in [0, M]$

$$\sup_{t \geq 0} |\psi(A_t^{(n)}(z)) - \psi(\hat{F}(z))| < \epsilon, \quad \text{for } n \geq n_0. \quad (71)$$

We truncate the integral then apply (71) to bound the absolute value of (68) by

$$\begin{aligned} \epsilon \int_0^K \int_0^t (e^{s-t}x)^2 \int_{e^{s-t}x}^M \frac{1}{z} dF_s^{(n)}(z) ds dx & \quad (72) \\ + \int_0^K \int_0^t (e^{s-t}x)^2 \int_M^\infty \frac{|\psi(A_s^{(n)}(z)) - \psi(\hat{F}(z))|}{z} dF_s^{(n)}(z) ds dx. & \quad (73) \end{aligned}$$

Integrate the inside integral of (72) by parts to obtain,

$$\frac{F_s^{(n)}(M)}{M} - \frac{F_s^{(n)}(e^{s-t}x)}{e^{s-t}x} + \int_{e^{s-t}x}^M \frac{F_s^{(n)}(z)}{z^2} dz.$$

Using the fact that  $F_s^{(n)}(z) \leq 1$  this is bounded by  $\frac{2}{M}$ . Multiplying by  $(e^{s-t}x)^2$  and integrating gives the following bound on (72)

$$\begin{aligned} \epsilon \int_0^K \int_0^t (e^{s-t}x)^2 \int_{e^{s-t}x}^M \frac{1}{z} dF_s^{(n)}(z) ds dx & \leq \epsilon \int_0^K \int_0^t (e^{s-t}x)^2 \frac{2}{M} ds dx \\ & = \epsilon \int_0^K \frac{2x^2}{M} (1 - e^{-t}) dx \\ & \leq \epsilon \frac{K^3}{M}. \end{aligned}$$

Thus, (72) can be made arbitrarily small.

Lastly we consider (73). Since  $\psi(u) \leq k$  we start with the bound

$$\begin{aligned} \int_0^K \int_0^t (e^{s-t}x)^2 \int_M^\infty \frac{|\psi(A_s^{(n)}(z)) - \psi(\hat{F}(z))|}{z} dF_s^{(n)}(z) ds dx \\ \leq \int_0^K \int_0^t (e^{s-t}x)^2 \int_M^\infty \frac{2k}{z} dF_s^{(n)}(z) ds dx. \end{aligned}$$

Integrate by parts so the inside becomes

$$-\frac{2kF_s^{(n)}(M)}{M} + \int_M^\infty \frac{2kF_s^{(n)}(z)}{z^2} dz.$$

Once more using the bound  $F_s^{(n)}(z) \leq 1$  we conclude that for  $M$  large the above becomes arbitrarily small. Therefore, the absolute value of (68) can be bounded by any  $\epsilon > 0$  uniformly for  $t \leq T$ .  $\square$

## References

- [AB05] David J. Aldous and Antar Bandyopadhyay. A survey of max-type recursive distributional equations. *Ann. Appl. Probab.*, 15(2):1047–1110, 2005.
- [ABKU99] Yossi Azar, Andrei Z. Broder, Anna R. Karlin, and Eli Upfal. Balanced allocations. *SIAM J. Comput.*, 29(1):180–200, September 1999.
- [Ald91] David J. Aldous. Random walk covering of some special trees. *J. Math. Anal. Appl.*, 157(1):271–283, 1991.
- [Ald13] David Aldous. Interacting particle systems as stochastic social dynamics. *Bernoulli*, 19(4):1122–1149, 09 2013.
- [AMP02a] O. S. M. Alves, F. P. Machado, and S. Yu. Popov. The shape theorem for the frog model. *Ann. Appl. Probab.*, 12(2):533–546, 2002.
- [AMP02b] Oswaldo Alves, Fabio Machado, and Serguei Popov. Phase transition for the frog model. *Electron. J. Probab.*, 7:no. 16, 1–21, 2002.
- [AMPR01] O. S. M. Alves, F. P. Machado, S. Yu. Popov, and K. Ravishanker. The shape theorem for the frog model with random initial configuration. *Markov Process. Related Fields*, 7(4):525–539, 2001.
- [AN72] K.B. Athreya and P.E. Ney. *Branching Processes*. Springer, 1972.
- [Arr81] Richard Arratia. Limiting point processes for rescalings of coalescing and annihilating random walks on  $z^d$ . *Ann. Probab.*, 9(6):909–936, 12 1981.
- [Arr83] Richard Arratia. Site recurrence for annihilating random walks on  $\mathbb{Z}^d$ . *Ann. Probab.*, 11(3):706–713, 08 1983.
- [Ber09] Nathanaël Berestycki. Recent progress in coalescent theory. *Ensaïos Matemáticos*, 16(1), 2009.
- [BFGG<sup>+</sup>16] Itai Benjamini, Eric Foxall, Ori Gurel-Gurevich, Matthew Junge, and Harry Kesten. Site recurrence for coalescing random walk. *Electron. Commun. Probab.*, 21:12 pp., 2016.

- [BG80] Maury Bramson and David Griffeath. Asymptotics for interacting particle systems on  $\mathbb{Z}^d$ . *Zeitschrift fr Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 53(2):183–196, 1980.
- [Big76] J. D. Biggins. The first- and last-birth problems for a multitype age-dependent branching process. *Advances in Appl. Probability*, 8(3):446–459, 1976.
- [Big77] John D. Biggins. Martingale convergence in the branching random walk. *J. Appl. Probability*, 14(1):25–37, 1977.
- [BL88] Maury Bramson and Joel L. Lebowitz. Asymptotic behavior of densities in diffusion-dominated annihilation reactions. *Phys. Rev. Lett.*, 61:2397–2400, Nov 1988.
- [BL15] M. Balázs and A. László Nagy. Dependent augmented Branching Annihilating Random Walk. *ArXiv e-prints*, January 2015.
- [BM72] J. R. Blum and V. J. Mizel. A generalized weyl equidistribution theorem for operators, with applications. *Transactions of the American Mathematical Society*, 165(2):291–307, 06 1972.
- [CEOR12] Colin Cooper, Robert Elsässer, Hirotaka Ono, and Tomasz Radzik. Coalescing random walks and voting on graphs. *CoRR*, abs/1204.4106, 2012.
- [CKK+07] Jacek Cicho, Marek Klonowski, ukasz Krzywiecki, Bartomiej Raski, and Pawe Zieliski. Random subsets of the interval and p2p protocols. In Moses Charikar, Klaus Jansen, Omer Reingold, and JosD.P. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, volume 4627 of *Lecture Notes in Computer Science*, pages 409–421. Springer Berlin Heidelberg, 2007.
- [Cox89] J. T. Cox. Coalescing random walks and voter model consensus times on the torus in  $\mathbb{Z}^d$ . *Ann. Probab.*, 17(4):1333–1366, 10 1989.
- [CQR09] Francis Comets, Jeremy Quastel, and Alejandro F. Ramírez. Fluctuations of the front in a one dimensional model of  $X + Y \rightarrow 2X$ . *Trans. Amer. Math. Soc.*, 361(11):6165–6189, 2009.



- [CRS13] M. Cabezas, L. T. Rolla, and V. Sidoravicius. Recurrence and Density Decay for Diffusion-Limited Annihilating Systems. *ArXiv e-prints*, September 2013.
- [Dar53] D. A. Darling. On a class of problems related to the random division of an interval. *The Annals of Mathematical Statistics*, 24(2):239–253, 06 1953.
- [DP14] Christian Döbler and Lorenz Pfeifroth. Recurrence for the frog model with drift on  $\mathbb{Z}^d$ . *Electron. Commun. Probab.*, 19:no. 79, 13, 2014.
- [Dub71] M. Serge Dubuc. La densité de la loi-limite d’un processus en cascade expansif. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 19:281–290, 1971.
- [EN74] P. Erdos and P. Ney. Some problems on random intervals and annihilating particles. *Ann. Probab.*, 2(5):828–839, 10 1974.
- [FMS04] L. R. Fontes, F. P. Machado, and A. Sarkar. The critical probability for the frog model is not a monotonic function of the graph. *J. Appl. Probab.*, 41(1):292–298, 2004.
- [FSZ09] Kevin Ford, K. Soundararajan, and Alexandru Zaharescu. On the distribution of imaginary parts of zeros of the riemann zeta function, ii. *Mathematische Annalen*, 343(3):487–505, 2009.
- [FW07] Klaus Fleischmann and Vitali Wachtel. Lower deviation probabilities for supercritical Galton-Watson processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 43(2):233–255, 2007.
- [GPTZ15] B. Garrod, M. Poplavskiy, R. Tribe, and O. Zaboronski. Interacting particle systems on  $\mathbb{Z}$  as Pfaffian point processes I - annihilating and coalescing random walks. *ArXiv e-prints*, July 2015.
- [Gri78] David Griffeath. Annihilating and coalescing random walks on  $\mathbb{d}$ . *Zeitschrift fr Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 46(1):55–65, 1978.
- [GS09] Nina Gantert and Philipp Schmidt. Recurrence for the frog model with drift on  $\mathbb{Z}$ . *Markov Process. Related Fields*, 15(1):51–58, 2009.

- [Her16] J. Hermon. Frogs on trees ? *ArXiv e-prints*, September 2016.
- [HJJ16a] Christopher Hoffman, Tobias Johnson, and Matthew Junge. From transience to recurrence with Poisson tree frogs. *Ann. Appl. Probab.*, 26(3):1620–1635, 2016.
- [HJJ16b] Christopher Hoffman, Tobias Johnson, and Matthew Junge. Recurrence and transience for the frog model on trees. to appear in the *Annals of Probability*, available at arXiv:1404.6238, 2016.
- [HL75] Richard A. Holley and Thomas M. Liggett. Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probab.*, 3(4):643–663, 08 1975.
- [Hoe63] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.
- [Hol83] Richard Holley. Two types of mutually annihilating particles. *Advances in Applied Probability*, 15(1):133–148, 1983.
- [HS79] R. Holley and D. W. Stroock. Central limit phenomena of various interacting systems. *Annals of Mathematics*, 110(2):pp. 333–393, 1979.
- [JJ16a] Tobias Johnson and Matthew Junge. The critical density for the frog model is the degree of the tree. available on the arXiv, 2016.
- [JJ16b] Tobias Johnson and Matthew Junge. Stochastic orders and the frog model. available at arXiv:1602.04411, 2016.
- [Kak76] S. Kakutani. A problem of equidistribution on the unit interval  $[0, 1]$ . In Alexandra Bellow and Dietrich Klzow, editors, *Measure Theory*, volume 541 of *Lecture Notes in Mathematics*, pages 369–375. Springer Berlin Heidelberg, 1976.
- [Kin82] J.F.C. Kingman. The coalescent. *Stochastic Processes and their Applications*, 13(3):235 – 248, 1982.
- [KN06] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences*. Dover Books on Mathematics. Dover Publications, 2006.
- [KZ12] E. Kosygina and M. P. W. Zerner. Excited random walks: results, methods, open problems. *ArXiv e-prints*, April 2012.

- [KZ15] E. Kosygina and M. P. W. Zerner. A zero-one law for recurrence and transience of frog processes. *ArXiv e-prints*, August 2015.
- [KZ16] Elena Kosygina and Martin P. W. Zerner. A zero-one law for recurrence and transience of frog processes. to appear in *Probability Theory and Related Fields*, available at arXiv:1508.01953, 2016.
- [LM05] Malwina J. Luczak and Colin McDiarmid. On the power of two choices: Balls and bins in continuous time. *The Annals of Applied Probability*, 15(3):1733–1764, 08 2005.
- [LMP05] Élcio Lebensztayn, Fábio P. Machado, and Serguei Popov. An improved upper bound for the critical probability of the frog model on homogeneous trees. *J. Stat. Phys.*, 119(1-2):331–345, 2005.
- [Loo78] J.-C. Lootgieter. Sur la rpartition des suites de Kakutani (ii). *Annales de l’institut Henri Poincar (B) Probabilits et Statistiques*, 14(3):279–302, 1978.
- [LT14] Junghoon Lee and Cihan Tepedelenlioğlu. Stochastic ordering of interference in large-scale wireless networks. *IEEE Trans. Signal Process.*, 62(3):729–740, 2014.
- [Mar02] R. Marchand. Strict inequalities for the time constant in first passage percolation. *Ann. Appl. Probab.*, 12(3):1001–1038, 2002.
- [MP14] P. Maillard and E. Paquette. Choices and intervals. *ArXiv e-prints*, 2014. To appear in *Israel Journal of Mathematics*.
- [MRS00] Michael Mitzenmacher, Andra W. Richa, and Ramesh Sitaraman. The power of two random choices: A survey of techniques and results. In *in Handbook of Randomized Computing*, pages 255–312. Kluwer, 2000.
- [MSH03] Neeraj Misra, Harshinder Singh, and E. James Harner. Stochastic comparisons of Poisson and binomial random variables with their mixtures. *Statist. Probab. Lett.*, 65(4):279–290, 2003.
- [Pop01] Serguei Yu. Popov. Frogs in random environment. *J. Statist. Phys.*, 102(1-2):191–201, 2001.

- [Pyk80] Ronald Pyke. The asymptotic behavior of spacings under Kakutani's model for interval subdivision. *The Annals of Probability*, 8(1):157–163, 02 1980.
- [RS04] Alejandro F. Ramírez and Vladas Sidoravicius. Asymptotic behavior of a stochastic combustion growth process. *J. Eur. Math. Soc. (JEMS)*, 6(3):293–334, 2004.
- [RS12] Leonardo T. Rolla and Vladas Sidoravicius. Absorbing-state phase transition for driven-dissipative stochastic dynamics on  $\mathbb{Z}$ . *Invent. Math.*, 188(1):127–150, 2012.
- [Rud76] W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, New York, third edition, 1976.
- [RV15] B. Rath and D. Valesin. Percolation on the stationary distributions of the voter model. *ArXiv e-prints*, February 2015.
- [Sau10] Thomas Sauerwald. Expansion and the cover time of parallel random walks. In *Proceedings of the 29th ACM SIGACT-SIGOPS symposium on Principles of distributed computing, PODC '10*, pages 315–324, New York, NY, USA, 2010. ACM. <http://doi.acm.org/10.1145/1835698.1835776>.
- [SS07] Moshe Shaked and J. George Shanthikumar. *Stochastic orders*. Springer Series in Statistics. Springer, New York, 2007.
- [Sta12] Richard P. Stanley. *Enumerative combinatorics. Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2012.
- [TRZ11] Cihan Tepedelenlioğlu, Adithya Rajan, and Yuan Zhang. Applications of stochastic ordering to wireless communications. 10(12):4249–4257, 2011.
- [TW99] András Telcs and Nicholas C. Wormald. Branching and tree indexed random walks on fractals. *J. Appl. Probab.*, 36(4):999–1011, 1999.
- [Vau77] R. C. Vaughan. On the distribution of  $p$  modulo 1. *Mathematika*, 24:135–141, 12 1977.
- [vdBK93] J. van den Berg and H. Kesten. Inequalities for the time constant in first-passage percolation. *Ann. Appl. Probab.*, 3(1):56–80, 1993.

- [vdBK00] J. van den Berg and Harry Kesten. Asymptotic density in a coalescing random walk model. *Ann. Probab.*, 28(1):303–352, 01 2000.
- [vdBK02] J. van den Berg and Harry Kesten. Randomly coalescing random walk in dimension  $\geq 3$ . 51:1–45, 2002.
- [Wey10] Hermann Weyl. Über die gibbs’sche erscheinung und verwandte konvergenzphänomene. *Rendiconti del Circolo Matematico di Palermo*, 30(1):377–407, 1910.
- [Yu09] Yaming Yu. Stochastic ordering of exponential family distributions and their mixtures. *J. Appl. Probab.*, 46(1):244–254, 2009.
- [Zer98] Martin P. W. Zerner. Directional decay of the Green’s function for a random nonnegative potential on  $\mathbf{Z}^d$ . *Ann. Appl. Probab.*, 8(1):246–280, 1998.
- [Zwe78] W. R. Van Zwet. A proof of Kakutani’s conjecture on random subdivision of longest intervals. *The Annals of Probability*, 6(1):pp. 133–137, 1978.