

# COVER TIME FOR THE FROG MODEL ON TREES

CHRISTOPHER HOFFMAN<sup>1</sup>, TOBIAS JOHNSON<sup>1b2</sup> and  
MATTHEW JUNGE<sup>3</sup>

<sup>1</sup> Department of Mathematics, University of Washington, USA;  
email: hoffman@math.washington.edu

<sup>2</sup> Department of Mathematics, College of Staten Island, USA;  
email: tobias.johnson@csi.cuny.edu

<sup>3</sup> Department of Mathematics, Duke University, USA;  
email: jungem@math.duke.edu

Received 14 November 2018; accepted 26 July 2019

## Abstract

The frog model is a branching random walk on a graph in which particles branch only at unvisited sites. Consider an initial particle density of  $\mu$  on the full  $d$ -ary tree of height  $n$ . If  $\mu = \Omega(d^2)$ , all of the vertices are visited in time  $\Theta(n \log n)$  with high probability. Conversely, if  $\mu = O(d)$  the cover time is  $\exp(\Theta(\sqrt{n}))$  with high probability.

2010 Mathematics Subject Classification: 60K35, 60J80, 60J10

## 1. Introduction

The *frog model* is a system of interacting walks that starts with one particle awake at the root of a graph and some number, typically Poisson-distributed with mean  $\mu$ , of sleeping particles at all the other vertices. Wakened particles perform simple random walk in discrete time. They wake any sleeping particles they encounter, which then begin their own independent random walks. A long-open problem posed to us several years ago by Itai Benjamini has been to determine the time it takes to visit every vertex of the full  $d$ -ary tree of height  $n$  (that is, the cover time). One might expect a simple argument would establish fast or slow cover times when the density of particles is very high or small. In fact, for *every* density of particles, it was unknown even if this quantity was polynomial or superpolynomial in  $n$ . Here we demonstrate that both can occur.

© The Author(s) 2019. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

If we view the process as modeling the spread of an infection, finite graphs are its natural setting and cover time a fundamental measurement. Finite trees are particularly interesting because of the phase transition that occurs on infinite rooted  $d$ -ary trees with an average of  $\mu$  particles per site: As we increase  $\mu$ , the root goes from being visited finitely to infinitely many times [HJJ17b, HJJ16]. Moreover, the companion to this work [HJJ17a] proves that when  $\mu = \Omega(d^2)$ , the root is visited at a linear rate. The dramatically different regimes on infinite trees suggest that both fast and slow cover times should occur on finite trees [Her18, JJ16]. However, it is unclear how reflection at the leaves influences the spread of active particles. Indeed, dealing with the boundary is the biggest obstacle to establishing regimes for fast and slow cover times.

First we describe what was previously known. The cover time is trivially at least  $n$ , and it is bounded above by the cover time for a single random walk on a tree, which is exponential in  $n$  [Ald91]. Until recently, these were the only known results. For any fixed  $d$  and particle density, Hermon improved the lower bound to  $\Omega(n \log n)$  and the upper bound to  $\exp(O(\sqrt{n}))$  [Her18]. In this paper, we prove that if the density of particles is sufficiently large then Hermon's lower bound is sharp, and if the density is small then his upper bound is sharp. In particular, this is the first proof that there exists a  $d \geq 2$  and density of particles for which the cover time is polynomial, or a  $d \geq 2$  and density of particles for which the cover time is superpolynomial.

We mention a few other closely related topics. The *susceptibility* of the frog model on a finite graph is the minimum lifespan of frogs such that all sites are visited. This statistic has been studied on tori and expanders [BFHM18] and on trees [Her18]. In none of these cases does it exhibit a phase transition in the density of particles, making it qualitatively very different from cover time.

A process resembling the frog model was proposed by Benjamini to study the connectivity of social networks and the spread of epidemics and has been studied on finite graphs [BH16] and infinite graphs [HMQS16]. On infinite nonamenable graphs, there is a phase transition in the initial density for whether all particles are eventually socially connected. For vertex-transitive amenable graphs, there is not. This resembles the frog model, which has a phase transition between transience and recurrence on trees [HJJ16], but not on lattices [AMPR01].

**Result.** As we mentioned, [Her18] gives the first nontrivial upper and lower bounds on the cover time, which we now state in more detail. Let  $\text{Poi}(\mu)$  denote a Poisson distribution with mean  $\mu$ . We refer to the frog model with one frog awake at the root, i.i.d.- $\text{Poi}(\mu)$  frogs elsewhere, and frogs following independent simple random walk paths as having i.i.d.- $\text{Poi}(\mu)$  initial conditions. We let  $\mathbb{T}_d^n$  denote the rooted, full  $d$ -ary tree of height  $n$ . This is the tree with levels  $0, \dots, n$  in which all vertices in levels  $0, \dots, n-1$  have  $d$  children.

Let  $\mathcal{C} = \mathcal{C}(n, d, \mu)$  denote the cover time for the frog model on  $\mathbb{T}_d^n$  with i.i.d.- $\text{Poi}(\mu)$  initial conditions. In [Her18, Theorem 2], Hermon proves there exists a constant  $c > 0$  such that for any  $\mu > 0$  and  $d \geq 2$ ,

$$\lim_{n \rightarrow \infty} \mathbf{P}[\mathcal{C} \leq e^c \sqrt{n \log d}] = 1.$$

As for the lower bound, it follows from [Her18, Theorem 1] that there exists a constant  $C > 0$  such that for any  $\mu > 0$  and  $d \geq 2$ ,

$$\lim_{n \rightarrow \infty} \mathbf{P}\left[\mathcal{C} \geq \frac{Cn \log n}{\mu}\right] = 1.$$

We now give our main result, which demonstrates the existence of two distinct behaviors for the cover time depending on the initial density of frogs. With a high density, the cover time is  $O_d(n \log n / \mu)$  with high probability. By Hermon's lower bound, this determines the cover time up to constant factor for each fixed choice of  $d$ . With a low initial density of frogs, we prove that the cover time is  $\exp(\Omega(\sqrt{n \log d}))$  with high probability, which is sharp up to the constant in the exponent by Hermon's upper bound. In fact, [Her18, Theorem 2] also gives an upper bound for the cover time when  $\mu$  decays in the height of the tree; one can take  $\mu$  as small as  $\exp(-\sqrt{n \log d})$  and still obtain a bound of the same order. Thus, our lower bound shows that for small but fixed values of  $\mu$ , the cover time exhibits the same asymptotic behavior as when  $\mu$  decays rapidly as  $n$  grows.

**THEOREM 1.1.** *Let  $\mathcal{C} = \mathcal{C}(n, d, \mu)$  denote the cover time for the frog model on  $\mathbb{T}_d^n$  with one awake frog at the root and i.i.d.- $\text{Poi}(\mu)$  conditions.*

- (a) *There exist constants  $\beta_0$ ,  $C_d$ , and  $n_0(\mu, d)$  such that for all  $d \geq 2$  and  $\mu > \beta_0 d^2$ ,*

$$\mathbf{P}\left[\mathcal{C} \leq \frac{C_d n \log n}{\mu}\right] \geq 1 - d^{-n}$$

*for  $n \geq n_0(\mu, d)$ .*

- (b) *Suppose that  $\mu \leq \min(d^{1-\epsilon}, d/100)$  for  $\epsilon \in (0, 1]$ . For some absolute constant  $c > 0$ ,*

$$\mathbf{P}[\mathcal{C} \geq e^c \sqrt{\epsilon n \log d}] \geq 1 - e^{-c \sqrt{\epsilon n \log d}}$$

*for  $n \geq \log d / c^2 \epsilon$ .*

Our full versions of these bounds, Theorems 3.1 and 4.1, are slightly stronger in that we extend them to initial distributions other than Poisson. We note that our lower bound, part (b), shows that the cover time is large even when  $\mu$  is as large as  $d/100$ ; see Corollary 4.2.

Thus, our results establish a slow cover time regime when  $\mu = O(d)$  and a fast cover time regime when  $\mu = \Omega(d^2)$ . This raises the question of what happens in between. On the infinite tree, the threshold between recurrence and transience occurs when  $\mu$  is on the order of  $d$  [JJ16]. This paper's results are consistent with the possibility that the slow and fast cover time regimes on the finite tree occur at the same parameters as the transient and recurrent phases on the infinite tree. But it is not clear this is so.

**QUESTION 1.2.** Are there other phases for the cover time of the frog model on  $\mathbb{T}_d^n$  besides those described in this paper? Is there a sharp phase transition between phases? If so, how does the process behave at critical values of  $\mu$ ?

In [DMnVZ00], the *activated random walk process*, which is essentially the frog model where particles fall back asleep at random, is discussed in connection with *self-organized criticality*, a phenomenon in which some real physical systems naturally push themselves toward criticality. The idea is that while conservative systems (in which particles are neither created nor destroyed) do not exhibit self-organized criticality, their behavior at criticality can nonetheless be a good model for it (see also [RSZ17, Section 1.3], whose discussion is aimed at mathematicians). This makes the frog model's behavior at criticality on both finite and infinite trees a particularly intriguing topic.

**Description of proof.** In Section 3, we tackle the cover time upper bound. The starting point for the proof is that the infected region (that is, set of visited sites) grows linearly for the frog model on an infinite tree, which we prove in the companion paper [HJJ17a]. Naively, one might think that a polynomial cover time bound would follow as an easy corollary, but we do not believe there is a quick argument. The issue is that our strong recurrence results from [HJJ17a], that the number of visits to the root grows linearly in time, become less powerful as they are applied to a finite tree near its leaves. We describe our argument in detail here to illustrate the problem at the boundary and our resolution of it. In the rough description below, we will suppress the fact that the constants depend on  $d$  and write  $O(\cdot)$  rather than  $O_d(\cdot)$ .

Looking toward a union bound, we must show it exponentially likely that an arbitrary leaf  $v_0 \in \mathbb{T}_d^n$  is woken in time  $O(n \log n)$ . Consider the spine  $v_0, v_1, \dots, v_n = \emptyset$  leading from  $v_0$  to the root  $\emptyset$ . To  $\mathbb{T}_d^n(v_k)$ , the subtree rooted at  $v_k$ ,

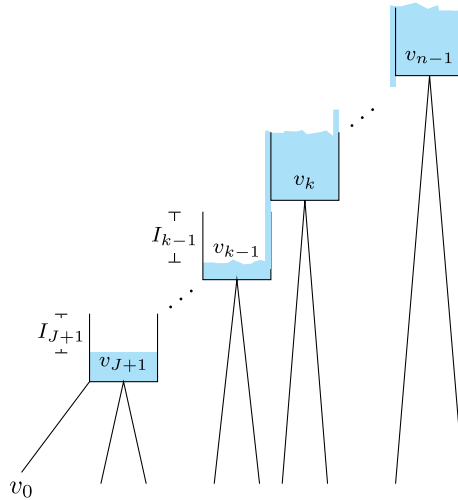


Figure 1. We show that the cover time is  $O(n \log n)$  by creating a cascade of frogs that move toward an arbitrary leaf  $v_0$  along the path  $v_0, \dots, v_n = \emptyset$ . Once  $I_k$  frogs build up at  $v_k$ , a constant stream of frogs flows to  $v_{k-1}$ . Using strong recurrence, we show that  $I_{n-1}$  frogs build up at  $v_{n-1}$  in time  $O(n)$ . This sets the cascade in motion, initiating a constant flow of frogs to  $v_{n-2}$ . After  $O(I_{n-2})$  steps, we have built up  $I_{n-2}$  frogs at  $v_{n-2}$ , setting off the next stage of the cascade, and so on. This quickly builds up  $\Omega_d(n \log n)$  frogs at distance  $J = O(\log n)$  from  $v_0$ , and it is exponentially likely that at least one will visit  $v_0$  in the next  $O(n \log n)$  steps.

we attach a random variable  $I_k$  defined as the number of frogs that must enter  $\mathbb{T}_d^n(v_k)$  to accumulate frogs that are frozen at  $v_{k-1}$  at a linear rate for  $d^k$  time steps. A possibly helpful metaphor is cascading water down a stair-step fountain (see Figure 1). Each basin needs a certain amount of water to reach a tipping point, after which it will pour water steadily into the one below it. In our proof, we wait until  $I_{n-1}$  frogs have accumulated at  $v_{n-1}$ . By definition of  $I_{n-1}$ , this sets off a linear flow of frogs which will send  $I_{n-2}$  frogs to  $v_{n-2}$  in  $O(I_{n-2})$  steps. This cascade continues until  $cn \log n$  frogs have accumulated at site  $v_J$ , at a distance  $J \approx \log_d n + \log_d \log n + C_d$  from  $v_0$ . At this point, we have built up enough frogs that we can ignore the wake-up dynamics of the frog model and instead show it is exponentially likely that at least one of  $cn \log n$  random walks started at  $v_J$  will visit  $v_0$  in the next  $O(n \log n)$  time steps. We can then apply a union bound over all leaves of  $\mathbb{T}_d^n$ .

The argument outlined above shows that the cover time is roughly  $O(I_{n-1} + \dots + I_{J+1} + n \log n)$ . Deducing a fast cover time is thus reduced to bounding

the random variables  $(I_k)$ . Since  $I_k$  is determined by the frogs within  $\mathbb{T}_d^n(v_k)$ , the random variables  $(I_k)$  are independent. We show that  $I_k$  has an exponential tail independent of  $k$ , which implies  $I_{n-1} + \dots + I_{J+1} = O(n)$  with exponentially high probability by a Chernoff bound. The proof that  $I_k$  has an exponential tail uses strong recurrence but is not an easy corollary of it. The issue is that strong recurrence only guarantees a steady flow of frogs out of  $\mathbb{T}_d(v_k)$  up to time  $k$ , while we need a flow up to time  $d^k$ , or else the argument would not work for  $v_k$  close to the leaves. Indeed, for  $k \approx J$ , strong recurrence yields a flow out of  $\mathbb{T}_d(v_k)$  for only  $O(\log n)$  steps, rather than the  $d^k \approx n \log n$  steps that we need. Thus, our challenge is to show that a steady flow of frogs out of  $\mathbb{T}_d^n(v_k)$  persists for much longer than given by strong recurrence. This argument makes up the bulk of Section 3.

In Section 4, we give our bound for the low density case. Our argument has no precursors in published work, as far as we know. We consider the number of visits to the root for the frog model on  $\mathbb{T}_d^{j^2}$  in the first  $2^j$  steps. We inductively assume that the expected number of visits to the root is  $O(1)$ , and we then try to prove that this estimate continues to hold for the frog model on  $\mathbb{T}_d^{(j+1)^2}$  after  $2^{j+1}$  time steps. To do this, we separate the tree into its first  $O(j)$  and its final  $j^2$  levels. We then push the induction forward by bounding the growth of frogs at different times in the two parts of the tree by various combinations of the inductive hypothesis, a bound given by branching random walk, and a bound of assuming all frogs are awake in a given subtree. Theorem 1.1(b) follows from considering  $n \approx j^2$ .

As we mentioned earlier, we also obtain results when the sleeping frog distributions at each vertex are not Poisson. These results are easy applications of [JJ18], in which we show that increasing the initial distributions in various stochastic orders causes certain statistics of the frog model to increase as well. We give a further introduction to these techniques in Appendix A. Some facts for random walk decompositions on trees and concentration inequalities are also contained in the appendices.

## 2. Preliminaries

Here we describe our notation, certain variants of the frog model, and also results that we will need from [HJJ17a].

**Notation.** For our purposes the frog model takes place on either the infinite rooted  $d$ -ary tree  $\mathbb{T}_d$  or on the full  $d$ -ary tree of height  $n$  denoted by  $\mathbb{T}_d^n$ . The root of whatever tree we are discussing will be denoted by  $\emptyset$ . Vertices at distance  $k$

from the root are at *level*  $k$  of the tree. For any rooted tree  $T$  and vertex  $v \in T$ , we denote the subtree of  $T$  made up of  $v$  and its descendants by  $T(v)$ .

Formally, the frog model is a pair  $(\eta, S)$  where for each vertex  $v$  other than the starting one,  $\eta(v)$  is the number of frogs initially sleeping at  $v$ , and  $S = (S_0(v, i))_{v \in G, i \geq 1}$  is a collection of walks satisfying  $S_0(v, i) = v$ . The  $i$ th particle sleeping at  $v$  on waking follows the path  $S_0(v, i)$ . When we discuss the frog model on a given graph with, say, i.i.d.- $\text{Poi}(\mu)$  initial conditions, unless we say otherwise we assume that the paths are simple random walks, and all of the random variables are independent. The root is assumed to be the starting vertex unless stated otherwise. The frog model evolves in discrete time, though it is easy to show that the results of this paper hold in continuous time as well. A realization of the frog model is called either *transient* or *recurrent* depending on whether the starting vertex is visited infinitely often by frogs. The *cover time* of a given frog model is the random variable defined as the first time all vertices in the system have been visited. Traditionally, particles are referred to as frogs, a practice we continue.

We let  $\text{Geo}(p)$  be the distribution that places probability  $(1 - p)^k p$  on  $k \geq 0$ . We also refer to the geometric distribution on  $\{1, 2, \dots\}$  with parameter  $p$ , which is the same distribution shifted by one. In a mild abuse of notation, we sometimes use  $\text{Poi}(\mu)$  and  $\text{Bin}(n, p)$  to refer to random variables with the given distributions rather than the distributions themselves, as in statements like  $\mathbf{P}[\text{Poi}(\mu) = 0] = e^{-\mu}$ .

**2.1. Modified frog models.** At times in our argument, it is helpful to consider variants of the frog model that couple to the original process. A *stopped version* of a given frog model  $(\eta, S)$  is a frog model  $(\eta, S')$  where each path  $S'_k(v, i)$  consists of  $S_0(v, i)$  stopped at some time  $T(v, i) \in \mathbb{N} \cup \{\infty\}$ . These must be stopping times for the frog model, in the sense that the decision to stop a frog at some time must be determined from the history of the stopped process up to that time. We give a quick sketch of how to formalize this. Following [KZ17, Section 2], let  $W'_j(\eta, S)$  be the set of sites visited for the first time at time  $j$  in the stopped process. Define  $\mathcal{F}_t$  as the  $\sigma$ -algebra representing all information about the stopped process revealed by time  $t$ ; formally, it is generated by the sets  $W'_j(\eta, S)$  for  $j \in \{0, \dots, t\}$ , the frog counts  $\eta(v)$  for  $v \in \bigcup_{j=0}^t W'_j(\eta, S)$ , and the frog paths  $(S'_k(v, i))_{k=0}^{t-j}$  for each  $j \in \{0, \dots, t\}$  and  $v \in W'_j(\eta, S)$ . We require the event  $\{T(v, i) \leq t\}$  to be measurable with respect to  $\mathcal{F}_t$ . As a consequence of this definition, for a stopped version of a frog model with, say, simple random walk paths, we can unstop all frogs at a given time and have them continue as independent simple random walks, since the stopping times do not impart any conditioning on the future part of the paths. By an easy coupling, in any given time stochastically fewer frogs are woken in the stopped process than in the original one.

When proving lower bounds on the growth of the frog model on the infinite tree  $\mathbb{T}_d$ , we typically work with what we call the self-similar frog model. Roughly speaking, it is the frog model with nonbacktracking frog paths, where frogs are stopped so that for each subtree of the form  $\mathbb{T}_d(v)$ , at most one frog from outside the subtree is allowed to enter it. To define it rigorously, we first define a *uniform nonbacktracking random walk* as a nearest neighbor path that samples uniformly from all adjacent edges on its first step, and then thereafter samples uniformly from all adjacent edges except the one just traversed. On  $\mathbb{T}_d$ , this is particularly simple: the path moves toward the root for some random amount of time, then takes a random nonbacktracking step away from the root, and then follows a uniformly sampled geodesic to  $\infty$ .

To define the *self-similar frog model* on  $\mathbb{T}_d$ , first let the frog paths be independent uniform nonbacktracking random walks. Now, we stop frogs as follows to enforce the rule that at most one frog enters any subtree. On a given step of the frog model, suppose that some vertex  $v \in \mathbb{T}_d \setminus \{\emptyset\}$  is visited for the first time. Let  $v'$  be the parent of  $v$ . On this step, one or more frogs move from  $v'$  to  $v$ . Stop all but one of them, and on all subsequent steps stop all frogs on moving from  $v'$  to  $v$ . Additionally, stop all frogs at  $\emptyset$  at steps 1 and beyond. We refer to [HJJ17a, Section 2.1] and [JJ16, Section 3.1.1] for more background information about the self-similar frog model. The reason for calling it self-similar is that only one external frog, that is, a frog initially at a vertex in  $\mathbb{T}_d \setminus \mathbb{T}_d(v)$ , may enter each  $\mathbb{T}_d(v)$ . Because the frog paths are nonbacktracking, the process on  $\{v'\} \cup \mathbb{T}_d(v)$  from the time a frog moves from  $v'$  to  $v$  is identical in law as on  $\{\emptyset\} \cup \mathbb{T}_d(\emptyset')$  from step 1 onward. Here  $\emptyset'$  is the child of  $\emptyset$  visited by the initial frog on its first step.

The self-similar frog model is defined on the infinite tree  $\mathbb{T}_d$ , though we will sometimes consider it on the finite tree  $\mathbb{T}_d^n$  by freezing frogs at leaves. But in proving our upper bound on cover time, we will usually consider a different process we call the nonbacktracking frog model on  $\mathbb{T}_d^n$ . To describe it, we first define a *root-biased nonbacktracking random walk from  $v_0$  on  $\mathbb{T}_d^n$*  as a walk distributed as follows. We set  $X_0 = v_0$ , and then we choose  $X_1$  uniformly from the neighbors of  $X_0$ . Conditionally on  $X_0, \dots, X_i$ , we choose  $X_{i+1}$  as follows: If  $X_i = \emptyset$ , choose  $X_{i+1}$  to be  $X_{i-1}$  with probability  $1/d^2$  and to be each of the other children of the root with probability  $(d+1)/d^2$ . If  $X_i$  is a leaf, then set  $X_{i+1}$  to be its parent (the only possibility for the next step). Otherwise, choose  $X_{i+1}$  uniformly from the neighbors of  $X_i$  other than  $X_{i-1}$ . It turns out that a simple random walk decomposes into this path plus independent excursions off of it (see [HJJ17a, Appendix A]). The odd behavior at the root results from the asymmetry of the tree there.



Finally, define the *nonbacktracking frog model* on  $\mathbb{T}_d^n$  as the frog model whose paths are independent root-biased nonbacktracking random walks on the specified tree. The following result shows that the time change for the underlying random walks speeds up the nonbacktracking model by only a constant factor compared to the usual frog model. This allows us to work with nonbacktracking frog models when we prove Theorem 1.1(a).

**PROPOSITION 2.1** [HJJ17a, Proposition 2.2]. *Let  $(\eta, S)$  and  $(\eta, S')$  be respectively the usual and the nonbacktracking frog models on  $\mathbb{T}_d^n$ , with arbitrary initial configuration  $\eta$ . There exists a coupling of the frog models  $(\eta, S)$  and  $(\eta, S')$  such that the following holds: For any  $b > \log d$ , there exists  $C = C(b)$  such that all vertices visited in  $(\eta, S')$  by time  $t$  are visited in  $(\eta, S)$  by time  $Ct$  with probability  $1 - e^{-bt}$ .*

**2.2. Adaptations of results from [HJJ17a].** We start by stating the result [HJJ17a, Theorem 3.1] which demonstrates a strong recurrence regime on infinite  $d$ -ary trees. We define the *return process* to be a point process on  $\mathbb{R}$  in which each point at  $t$  represents a frog that is occupying the root at time  $t$ . Note that this is supported on the nonnegative integers.

**THEOREM 2.2.** *Consider the self-similar frog model on  $\mathbb{T}_d$  with i.i.d.- $\text{Poi}(\mu)$  initial conditions. For any  $d \geq 2$ ,  $\alpha > 0$ , and  $\mu \geq 3d(d+1) + \alpha(d+1)$ , the return process stochastically dominates a Poisson point process with intensity measure  $\sum_{k=1}^{\infty} \alpha \delta_{2k}$ .*

This extends to  $\mathbb{T}_d^n$ , but because of the boundary, only up to time  $2n - 2$ .

**COROLLARY 2.3.** *Consider the self-similar frog model on  $\mathbb{T}_d^n$  with i.i.d.- $\text{Poi}(\mu)$  initial conditions and frogs frozen on reaching a leaf. For any  $d \geq 2$ ,  $\alpha > 0$ , and  $\mu \geq 3d(d+1) + \alpha(d+1)$ , the return process stochastically dominates a Poisson point process with intensity measure  $\sum_{k=1}^{n-1} \alpha \delta_{2k}$ .*

*Proof.* Couple the processes of Theorem 2.2 and of this corollary by having all frogs follow the same paths until reaching the boundary of the finite tree. Consider a root visit on the process on the infinite tree occurring before time  $2n$ . The combined path of frogs waking the returner together with the returning path does not reach depth  $n$  of the tree, since the return occurs before time  $2n$ . Thus, the return occurs in the process on the finite tree as well. The finite tree process therefore has all of the returns of the infinite tree process before time  $2n$ , and the result follows from Theorem 2.2.  $\square$

Last, we state [HJJ17a, Lemma 4.1], which helps us deduce a weaker version of the shape theorem [HJJ17a, Theorem 1.1] for the finite tree.

LEMMA 2.4. *Let  $\beta > 0$  and consider the self-similar frog model on  $\mathbb{T}_d$  with i.i.d.-Poi( $\mu$ ) frogs per site, where  $\mu = (3 + \beta)d(d + 1)$ . Let  $\emptyset, v_0, v_1, v_2, \dots$  be an arbitrary ray in  $\mathbb{T}_d$ , and condition the initial frog to take its first step to  $v_0$ . Let  $\tau_i$  be the number of steps after  $v_{i-1}$  is first visited that  $v_i$  is first visited. Then  $(\tau_i)_{i \geq 1}$  are i.i.d. and satisfy*

$$\mathbf{P}[\tau_i > 2t - 1] \leq e^{-\beta t} \tag{1}$$

for  $t \in \{1, 2, \dots\}$ .

An important corollary for our work here is that a self-similar frog model activates half of the leaves in the active branch of any height tree in time  $O(k)$  with probability at least  $1/2$  (recall that only sites in  $\mathbb{T}_d(\emptyset')$  are visited in the self-similar model, where  $\emptyset'$  is the child of the root first visited by the initial frog).

COROLLARY 2.5. *Consider a self-similar frog model on  $\mathbb{T}_d^{k+1}$  with i.i.d.-Poi( $\mu$ ) initial conditions where frogs are frozen at leaves, for any  $k \geq 1$ . For  $\mu \geq (3 + \beta)d(d + 1)$  and sufficiently large absolute constants  $\beta$  and  $C$ , there exists  $p = p(\beta, C)$  such that  $d^k/2$  of the leaves are visited in  $Ck$  steps with probability at least  $p$ . Moreover,  $p$  can be made arbitrarily close to 1 by choosing  $C$  and  $\beta$  sufficiently large, and in particular  $p \geq 1/2$  when  $\beta \geq 2$  and  $C \geq 8$ .*

*Proof.* Let  $v_{-1} = \emptyset$ . Let  $v_0$  be a child of  $\emptyset$ , and condition on the initial frog taking its first step to  $v_0$ . By symmetry of the tree, it suffices to prove the corollary under this assumption. Note that the children of the root other than  $v_0$  are never visited, since frogs are frozen when they visit the root in the self-similar frog model. Thus, our goal is to show that at least half the leaves descending from  $v_0$  are visited in  $Ck$  steps with probability at least  $p$ , for some  $p$  to be determined.

Let  $v_{-1}, v_0, \dots, v_k$  be the path from  $\emptyset = v_{-1}$  to an arbitrary leaf  $v_k$  descending from  $v_0$ . We will show that  $v_k$  fails to be visited in  $Ck$  steps with probability at most  $q = q(\beta, C)$ , and that  $q$  can be made arbitrarily small by choosing  $\beta$  and  $C$  large enough. Then, the expected number of leaves descending from  $v_0$  that are not visited in time  $Ck$  is at most  $qd^k$ . The lemma then follows by applying Markov’s inequality to show that the number of unvisited leaves is larger than  $d^k/2$  with probability at most  $2q$ .

For  $i \geq 1$ , let  $\tau'_i$  be the number of steps after  $v_{i-1}$  is visited that  $v_i$  is first visited. If  $v_i$  is never visited, set  $\tau'_i = \infty$ . Now, we observe that the self-similar frog model on  $\mathbb{T}_d^{k+1}$  with frogs frozen at leaves is identical to the first  $k + 1$  levels

of the self-similar frog model on  $\mathbb{T}_d$  with frogs frozen at level  $k + 1$ . This frog model can naturally be coupled with the self-similar frog model on  $\mathbb{T}_d$  with no freezing. Putting these together yields a coupling between  $(\tau'_i)_{i=1}^{k+1}$  and the random variables  $(\tau_i)_{i=1}^\infty$  defined in Lemma 2.4.

We claim that if

$$\tau_i \leq 2(k - i + 1) \quad \text{for all } 1 \leq i \leq k, \quad (2)$$

then  $\tau'_i = \tau_i$  for all  $1 \leq i \leq k$ . Indeed, suppose that  $\tau_i \leq 2(k - i + 1)$ . From the time  $v_{i-1}$  is first visited in the self-similar frog model on  $\mathbb{T}_d$ , it takes at most  $2(k - i + 1)$  steps for  $v_i$  to be visited. Since any walk from  $v_{i-1}$  to level  $k + 1$  back to  $v_i$  has length at least  $2(k - i + 1) + 1$ , this visit to  $v_i$  still occurs when frogs are frozen at level  $k + 1$ . Under the coupling, we then have  $\tau'_i = \tau_i$  for all  $1 \leq i \leq k$  as desired.

Thus, if (2) holds and  $1 + \sum_{i=1}^k \tau_i \leq Ck$ , then  $v_k$  is woken in  $Ck$  steps in the self-similar model on  $\mathbb{T}_d^{k+1}$ . Therefore,

$$q(\beta, C) \leq \mathbf{P}[\tau_i > 2(k - i + 1) \text{ for some } 1 \leq i \leq k] + \mathbf{P}\left[1 + \sum_{i=1}^k \tau_i > Ck\right].$$

We now bound the two terms on the right-hand side of this inequality. By Lemma 2.4,

$$\mathbf{P}[\tau_i > 2(k - i + 1) \text{ for some } 1 \leq i \leq k] \leq \sum_{i=1}^k e^{-\beta(k-i+1)} \leq \frac{e^{-\beta}}{1 - e^{-\beta}},$$

which can be made as small as desired by increasing  $\beta$ .

By Lemma 2.4, the random variables  $(\tau_i)$  are independent, and  $(\tau_i + 1)/2$  is stochastically dominated by the geometric distribution on  $\{1, 2, \dots\}$  with parameter  $1 - e^{-\beta}$ . By Proposition B.3,

$$\mathbf{P}\left[1 + \sum_{i=1}^k \tau_i > Ck\right] \leq \exp\left[-k\left(\frac{C(1 - e^{-\beta})}{2} - 1\right)\right] \leq \exp\left(-\frac{C(1 - e^{-\beta})}{2} + 1\right).$$

For any given  $\beta > 0$ , this can be made arbitrarily small by increasing  $C$ . This proves that  $q(\beta, C)$  is as small as desired for large enough  $\beta$  and  $C$ . In particular, plugging in numbers, we see that  $q(\beta, C) < 1/4$  if  $\beta \geq 2$  and  $C \geq 8$ . As  $2q$  is the bound on the probability of fewer than  $d^k/2$  leaves being visited, this completes the proof.  $\square$

### 3. Fast cover time for large $\mu$

We now present the most general version of the cover time upper bound.

**THEOREM 3.1.** *Let  $C$  be the cover time for the frog model on  $\mathbb{T}_d^n$  with initial frog counts given by  $(\eta(v))_{v \in \mathbb{T}_d^n \setminus \{\emptyset\}}$ . Suppose that  $\eta(v) \succeq_{\text{pgf}} \text{Poi}(\mu)$  for all  $v$ , where  $\mu > \beta_0 d^2$  for a sufficiently large absolute constant  $\beta_0$ . There exist constants  $C_d$  and  $n_0(\mu, d)$  such that*

$$\mathbf{P}[C > C_d n \log n / \mu] \leq d^{-n}$$

for all  $n \geq n_0(\mu, d)$ .

See Appendix A for the definition of the pgf stochastic order denoted by  $\succeq_{\text{pgf}}$ . Loosely speaking, the condition  $\eta(v) \succeq_{\text{pgf}} \text{Poi}(\mu)$  means that the distribution of  $\eta(v)$  is larger and more concentrated than the distribution of  $\text{Poi}(\mu)$ . In particular, if  $\eta(v) = k$  deterministically for some integer  $k \geq \mu$ , then  $\eta(v) \succeq_{\text{pgf}} \text{Poi}(\mu)$ . Thus, our result holds for the frog model with  $k$  frogs per vertex for  $k > \beta_0 d^2$ .

This theorem follows from two propositions that we explain now. Fix a leaf  $v_0 \in \mathbb{T}_d^n$ . Label the path from  $v_0$  to the root by  $v_0, \dots, v_n = \emptyset$ . In general, we will take  $\mu = (3 + \beta)d(d + 1)$  for some parameter  $\beta$ , a convenient form for applying Corollaries 2.3 and 2.5. The vertex  $v_J$ , where

$$J = J(d, n, \beta) = \lfloor \log_d n + \log_d(\log n) + 5 \log_d 10 - \log_d \beta \rfloor, \tag{3}$$

is far enough from  $v_0$  that we can show that many frogs visit  $v_J$  in time  $O(n \log n)$ . It is also close enough to  $v_0$  that one of these frogs at  $v_J$  will visit  $v_0$  in  $O(n \log n)$  steps with high probability (see Figure 2). These two statements are the content of Propositions 3.2 and 3.3, which we show under Poisson initial conditions. We complete the proof by applying Lemma A.1 to relax this assumption.

Recall the definition of a stopped version of a frog model from Section 2.1.

**PROPOSITION 3.2.** *For some constants  $\beta_0$  and  $C_d$ , the following holds. Let  $\mu = (3 + \beta)d(d + 1)$  for  $\beta > \beta_0$ . There exists a stopped version of the frog model on  $\mathbb{T}_d^n$  with i.i.d.- $\text{Poi}(\mu)$  initial conditions such that  $10n \log n$  frogs have been stopped at vertex  $v_J$  by time  $C_d n \log n / \beta$  with probability at least  $1 - d^{-3n}$  for all  $n \geq n_0(\beta, d)$ , for some constant  $n_0(\beta, d)$ .*

**PROPOSITION 3.3.** *Suppose that  $10n \log n$  simple random walks start at vertex  $v_J$  in  $\mathbb{T}_d^n$  and move independently, and that  $n \geq n_0$  for some sufficiently large absolute constant  $n_0$ . For some absolute constant  $C$ , one of the walks visits  $v_0$  within  $Cn \log n / \beta$  steps with probability  $1 - d^{-3n}$ .*

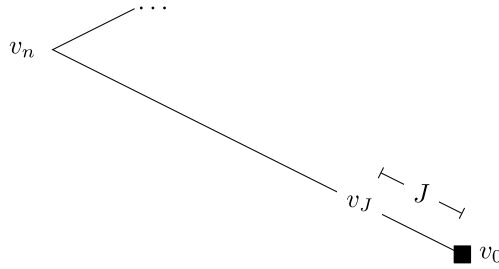


Figure 2. The basic idea of Theorem 3.1 is to show that many frogs visit  $v_J$  after  $O(n \log n)$  steps for  $J \approx \log_d n + \log_d(\log n) + C$ . Once enough frogs are built up at  $v_J$ , one of them will visit the leaf  $v_0$  with high probability within  $O(n \log n)$  steps.

The upper bound on the cover time follows from Propositions 3.2 and 3.3.

*Proof of Theorem 3.1.* First, assume that the sleeping frog counts  $(\eta(v))_v$  are i.i.d.-Poi( $\mu$ ). By Proposition 3.2, there is a stopped version of the frog model where  $10n \log n$  frogs accumulate at  $v_J$  by time  $C_d n \log n / \beta$  with probability at least  $1 - d^{-3n}$ . At time  $\lfloor C_d n \log n / \beta \rfloor$ , unfreeze all frogs and let them resume their simple random walks. By Proposition 3.3, the vertex  $v_0$  is visited in this modified process by time  $C'_d n \log n / \beta$  with probability  $1 - 2d^{-3n}$  for some constant  $C'_d$ . If this holds in this stopped and restarted frog model, then it holds in the original frog model as well, by an obvious coupling. As  $v_0$  was arbitrary, each leaf is visited with probability at least  $1 - 2d^{-3n}$ , and the expected number of leaves unvisited by time  $C'_d n \log n / \beta$  is therefore at most  $2d^{-2n}$ .

Now, we extend this to non-Poisson initial conditions. Let  $N$  be the number of leaves visited by time  $C'_d n \log n / \beta$  in the Poisson frog model, which we have shown to satisfy

$$\mathbf{E}N \geq d^n - 2d^{-2n}.$$

Let  $N'$  be the corresponding count of visited leaves for the frog model defined in the statement of this theorem. By Lemma A.1, we have  $\mathbf{E}N' \geq \mathbf{E}N$ . Thus, the expected number of unvisited leaves in this frog model is also at most  $2d^{-2n}$ , and by Markov's inequality there is an unvisited leaf with probability at most  $2d^{-2n} \leq d^{-n}$ . Once all leaves are visited, all vertices of the tree have been visited, completing the proof.  $\square$

**3.1. Establishing Proposition 3.2.** The goal of this section is to prove it overwhelmingly likely that  $\Omega(n \log n)$  frogs accumulate at  $v_J$  in time

$O(n \log n / \beta)$ , recalling the definitions of  $v_0, \dots, v_n, J$ , and  $\beta$  from the beginning of the section. Our argument is sequential: we show that many frogs flow from  $v_n$  to  $v_{n-1}$ , which spurs many frogs to flow into  $v_{n-2}$ , and so on. To make this precise, we introduce random variables  $I_k$  for  $J + 1 \leq k \leq n$ . Loosely speaking,  $I_k$  is the quantity of frogs that must start at  $v_k$  so that frogs flow steadily to  $v_{k-1}$  at a rate of  $\Omega(\beta)$  per time step. Now, imagine running the frog model until  $I_{n-1}$  frogs have built up at  $v_{n-1}$ . Once this happens, frogs will flow steadily to  $v_{n-2}$ ; allow them to build up until there are  $I_{n-2}$  there, which will take time  $O(I_{n-2}/\beta)$ . Continuing in this way, we build up  $I_{J+1}$  frogs at  $v_{J+1}$  in time

$$O\left(\sum_{k=J+1}^{n-1} I_k / \beta\right),$$

plus the time to get the first  $I_{n-1}$  frogs to  $v_{n-1}$ . This creates a steady flow of frogs to  $v_J$ , and after another  $O(n \log n / \beta)$  steps, we have produced  $\Omega(n \log n)$  visits to  $v_J$ . Thus, the main task is to show that  $\sum_{k=J+1}^{n-1} I_k / \beta$  is unlikely to be large. We do this by showing an exponential tail bound for  $I_k$ , from which it follows that it is exponentially likely that this sum is  $O(n)$ .

We mention that we use nonbacktracking frogs throughout this section. This coordinates well with our results regarding the self-similar frog model in Section 2.2. Only at the very end will we apply Proposition 2.1 to move our results back to the usual frog model.

3.1.1. *Definition of  $I_k$ .* We first define a family of processes  $\text{FM}(v_k, \ell)$ , which are frog models limited to the subtree  $\mathbb{T}_d^n(v_k)$  with an extra  $\ell$  frogs initially at  $v_k$ . Then we define  $I_k$  as the smallest  $\ell$  for which  $\text{FM}(v_k, \ell)$  produces a steady stream of frogs entering  $v_{k-1}$ :

DEFINITION 3.4 ( $\text{FM}(v_k, \bullet)$  and  $I_k$ ). Let  $\mu = (3 + \beta)d(d + 1)$ . For  $J < k < n$  and  $\ell \geq 1$ , let  $\text{FM}(v_k, \ell)$  be a frog model defined as follows. We place sleeping frogs only within  $\mathbb{T}_d^n(v_k) \setminus \mathbb{T}_d^n(v_{k-1})$ . At all of these vertices except for  $v_k$ , place  $\text{Poi}(\mu)$  frogs per site as usual. At  $v_k$  itself, we place  $\text{Poi}(\mu)$  frogs plus an extra  $\ell$  special frogs, as we will call them. The paths of the special frogs are root-biased nonbacktracking walks stopped at  $v_{k-1}$  and  $v_{k+1}$  with their first steps conditioned to move to a descendant of  $v_k$  (that is, to move away from  $v_{k+1}$ ). The paths of all other frogs are root-biased nonbacktracking walks stopped at  $v_{k-1}$  and  $v_{k+1}$ . Vertex  $v_k$  is the starting vertex for the process; all frogs there are initially awake.

For a fixed value of  $k$ , we consider  $\text{FM}(v_k, \ell)$  to be coupled for all choices of  $\ell$  in the natural way. That is, we suppose that there is an infinite pile of special frogs at  $v_k$ , and  $\text{FM}(v_k, \ell)$  uses only the first  $\ell$  of them. We denote the collection of coupled frog models  $(\text{FM}(v_k, \ell))_{\ell \geq 1}$  by  $\text{FM}(v_k, \bullet)$ .

For  $J < k < n$ , we define the random variable  $I_k$  to be the smallest integer  $\ell$  such that the number of frogs frozen at  $v_{k-1}$  by time  $t$  in  $\text{FM}(v_k, \ell)$  is at least  $\beta t/10000$  for all  $\max(3, \lceil \ell/\beta \rceil) \leq t \leq d^k$ . Observe that this test becomes vacuous when  $\ell > \beta d^k$ , and therefore  $I_k \leq \beta d^k + 1$ .

As we have remarked, one should think of  $I_k$  as the minimum number of special frogs at  $v_k$  to ensure a steady flow of frogs into  $v_{k-1}$ . This ‘steady flow’ is at rate  $\Omega(\beta)$  per time step. For technical reasons, we only require it to start at time  $\max(3, \lceil \ell/\beta \rceil)$ . We require the flow to continue only up to time  $d^k$  because it is impossible for it to continue much longer, since there are only  $O(\beta d^k)$  frogs in the entire system  $\text{FM}(v_k, \ell)$ .

**3.1.2. Exponential tail bound for  $I_k$ .** The bulk of our work in Section 3 is to prove the following exponential tail bound on  $I_k$ :

**PROPOSITION 3.5.** *For some constants  $c, C > 0$ , the following holds. Let  $\mu = (3 + \beta)d(d + 1)$ . For  $\beta \geq 10000$ , it holds for any integers  $J < k < n$  and  $1 \leq \ell \leq d^k$  that*

$$\mathbb{P}[I_k > \ell] \leq C e^{-c\ell}. \tag{4}$$

Once this is proven, a short argument shows that  $\sum_{k=J+1}^{n-1} I_k = O(n)$  is exponentially likely if the random variables  $(I_k)$  are assumed to be i.i.d. This is the most important element of the proof of Proposition 3.2. To prove Proposition 3.5, we must argue that  $\text{FM}(v_k, \ell)$  is exponentially likely in  $\ell$  to send a steady flow of frogs to  $v_{k-1}$ . There are two parts to this argument. From times  $\max(3, \lceil \ell/\beta \rceil)$  to  $k$ , we obtain the necessary quantity of frogs at  $v_{k-1}$  as a direct consequence of Corollary 2.3 (see Lemma 3.8). To show that the flow condition is maintained beyond this, we leverage Corollary 2.5 to prove it exponentially likely in  $\ell$  that we wake up a positive fraction of all frogs in  $\mathbb{T}_d^n(v_k)$  by time  $O(k)$ . We then show that enough of these frogs will move to  $v_{k-1}$  to give us our steady flow of frogs from time  $14k$  to  $d^k$  (see Lemma 3.10). To bridge the gap between times  $k$  and  $14k$ , we make  $\beta$  large enough to build up a surplus of frogs at  $v_{k-1}$  during the first  $k$  time steps. This ensures that the steady flow requirement is met until time  $14k$  even if no additional frogs visit  $v_{k-1}$  for times between  $k$  and  $14k$  (see Lemma 3.9).

We now begin working toward Proposition 3.5. We start with two technical estimates. First, we show that a frog at a leaf of  $\mathbb{T}_d^k$  hits the root in  $t$  steps with probability  $\Omega(td^{-k})$ .

**LEMMA 3.6.** *Consider a root-biased nonbacktracking random walk on  $\mathbb{T}_d^k$  starting from a leaf for  $k \geq 2$ . For any integer  $k + 2 \leq t \leq d^k$ , the walk visits the root in its first  $t$  steps with probability at least  $(t - k - 2)d^{-k}/4$ .*

*Proof.* Let  $T$  be the first time that the walk hits the root. We decompose the walk into a sequence of independent excursions from the leaves. Since each excursion reaches the root with probability  $d^{-k+1}$ , the number of unsuccessful excursions before hitting the root is  $\text{Geo}(d^{-k+1})$ . Each unsuccessful excursion has length distributed as  $2\tilde{G}$ , where  $\tilde{G}$  is a geometric random variable on  $\{1, 2, \dots\}$  with parameter  $(d-1)/d$  conditioned to be less than  $k$ . Let  $\tilde{G}^{(i)}$  be independent copies of  $\tilde{G}$ , and let  $G^{(i)}$  be independent and distributed as the unconditioned geometric distribution on  $\{1, 2, \dots\}$  with the same parameter. Thus,

$$T \stackrel{d}{=} k + 2 \sum_{i=1}^{\text{Geo}(d^{-k+1})} \tilde{G}^{(i)} \leq k + 2 \sum_{i=1}^{1+\text{Geo}(d^{-k+1})} G^{(i)} \stackrel{d}{=} k + 2(1 + \text{Geo}((d-1)d^{-k})),$$

with the last step using the fact that the sum of  $1 + \text{Geo}(p)$  many independent  $1 + \text{Geo}(q)$  random variables is a  $1 + \text{Geo}(pq)$  random variable. Therefore,

$$\begin{aligned} \mathbf{P}[T \leq t] &\geq \mathbf{P}\left[1 + \text{Geo}((d-1)d^{-k}) \leq \frac{t-k}{2}\right] \\ &= 1 - (1 - (d-1)d^{-k})^{\lfloor (t-k)/2 \rfloor} \geq 1 - \exp\left(-\frac{(t-k-2)d^{-k}}{2}\right). \end{aligned}$$

Using the bound  $1 - e^{-x} \geq x/2$  for  $x \in [0, 1]$  along with the assumption that  $t \leq d^k$ ,

$$\mathbf{P}[T \leq t] \geq \frac{(t-k-2)d^{-k}}{4}. \quad \square$$

**LEMMA 3.7.** *Suppose that  $m$  balls are placed uniformly and independently into  $n$  bins, with  $m \geq 3n$ . Let  $Z$  be the number of occupied bins. Then*

$$\mathbf{P}[Z \leq 2n/3] \leq e^{-m/54}.$$

*Proof.* Imagine that we place the balls one after another, and define  $Z_i$  as the number of bins occupied after  $i$  balls have been placed. Let  $T = \min\{i : Z_i \geq 2n/3\}$ . We need to bound the probability that  $T > m$ . We observe that  $(Z_i)_{i \geq 0}$  is a pure birth process with  $\mathbf{P}[Z_{i+1} = Z_i + 1 \mid Z_i] = 1 - Z_i/n$  and  $Z_0 = 0$ . Let  $(Y_i)_{i \geq 0}$  be a pure birth process starting at 0 and increasing at each step with probability  $1/3$ . We can couple the two processes so that  $(Y_i)$  increases only when  $(Z_i)$  does up to time  $T$ . We apply Proposition B.1 to the random variable  $Y_m$ , which is distributed as  $\text{Bin}(m, 1/3)$ , and we get

$$\mathbf{P}[T > m] \leq \mathbf{P}[Y_m \leq 2n/3] \leq \exp\left(-\frac{(1-2n/m)^2 m}{6}\right) \leq e^{-m/54},$$

using our assumption  $m \geq 3n$ . □



We are now ready to start on the proof of Proposition 3.5. Let  $X_t$  be the number of frogs frozen at  $v_{k-1}$  by time  $t$  in  $\text{FM}(v_k, \ell)$ . The basic idea is that if  $I_k > \ell$ , then  $X_t < \beta t/10000$  occurs for some  $\max(3, \lceil \ell/\beta \rceil) \leq t \leq d^k$ . Thus it suffices to show that the probability of this event decays exponentially in  $\ell$ . In the next three lemmas, we break the time interval  $\max(3, \lceil \ell/\beta \rceil) \leq t \leq d^k$  into the three segments described on page 15, and we bound the probability that  $X_t < \beta t/10000$  on any of them.

The first time segment is for length  $k$ , which is the height of the tree rooted at  $v_k$ . As we mentioned, we use the application of strong recurrence to the finite tree in Corollary 2.3 to accrue  $\Omega(k)$  frogs at  $v_j$  in time  $k$ .

LEMMA 3.8. *With the conditions of Proposition 3.5,*

$$\mathbf{P}[X_t < \beta t/10000 \text{ for some } \max(3, \lceil \ell/\beta \rceil) \leq t \leq k] \leq Ce^{-c\ell}$$

for some constants  $c, C > 0$ .

*Proof.* With probability  $1 - d^{-\ell}$ , some child of  $v_k$  other than  $v_{k-1}$  is visited by one of the  $\ell$  special initial frogs on the first step of  $\text{FM}(v_k, \ell)$ . Call this event  $E$ . Conditional on  $E$ , let  $u$  be a visited child, and couple  $\text{FM}(v_k, \ell)$  with the self-similar frog model on  $\mathbb{T}_d^k$  with frogs frozen at the leaves, as follows. Identify  $v_k$  with the root of the self-similar model. Let  $u'$  be the child of the root in  $\mathbb{T}_d^k$  first visited by the initial frog in the self-similar model. Identify  $\mathbb{T}_d^k(u)$  in  $\text{FM}(v_k, \ell)$  with  $\mathbb{T}_d^k(u')$  in the self-similar model. Make the number of initial frogs in  $\mathbb{T}_d^k(u)$  in  $\text{FM}(v_k, \ell)$  identical to the number of initial frogs in  $\mathbb{T}_d^k(u')$  in the self-similar model. Let each of these frogs in  $\text{FM}(v, \ell)$  follow the corresponding frog in the self-similar model until it is frozen. After, each frog in  $\text{FM}(v, \ell)$  continues as a root-biased nonbacktracking walk independent of the self-similar model. Similarly, let the initial frog that moved to  $u$  in  $\text{FM}(v_k, \ell)$  match the initial frog in the self-similar model until it is frozen.

By this coupling and Corollary 2.3, the count of frogs moving from  $u$  back to  $v_k$  by time  $2t$  is stochastically at least  $\text{Poi}(\beta dt)$  for any integer  $1 \leq t \leq k - 1$ , conditional on  $E$ . As each of these frogs moves next to  $v_{k-1}$  with probability  $1/d$ , we have  $X_{2t+1} \geq \text{Poi}(\beta t)$ . Restating this, conditional on  $E$ ,

$$X_t \geq \text{Poi}(\lceil (t-1)/2 \rceil \beta) \geq \text{Poi}\left(\left(\frac{\beta}{2} - 1\right)t\right) \quad (5)$$

for any  $3 \leq t \leq 2k - 1$ . Let  $t_0 = \max(3, \lceil \ell/\beta \rceil)$ . By Lemma B.4,

$$\mathbf{P}[X_t < \beta t/10000 \text{ for any } t_0 \leq t \leq k \mid E] \leq 2e^{-\Omega(\beta \lceil \ell/\beta \rceil)} \leq 2e^{-\Omega(\ell)}. \quad (6)$$

Note that for simplicity we have limited the range to  $t_0 \leq t \leq k$ , even though (6) holds for a larger interval. Combined with  $\mathbf{P}[E] = 1 - d^{-\ell}$ , this proves the claim.  $\square$

Our next time segment is from  $k$  to  $14k$ , bridging the gap between our first and third segments. The argument here is rather simple: In proving Lemma 3.8, we built up sufficiently many frogs at time  $k$  to keep  $X_t$  large enough until time  $14k$ .

LEMMA 3.9. *With the conditions of Proposition 3.5,*

$$\mathbf{P}[X_t < \beta t/10000 \text{ for some } \max(\lceil \ell/\beta \rceil, k) < t \leq 14k] \leq e^{-c\ell}$$

for some constant  $c > 0$ .

*Proof.* From (5) in the previous proof,

$$X_k \geq \text{Poi}\left(\left(\frac{\beta}{2} - 1\right)k\right).$$

Proposition B.1 then gives

$$\mathbf{P}\left[X_k < \frac{14\beta k}{10000}\right] \leq e^{-\Omega(\beta k)} \leq e^{-\Omega(\ell)}$$

if  $14k \geq \lceil \ell/\beta \rceil$ , which we can assume since otherwise the lemma is vacuous. This completes the proof, since if  $X_t < \beta t/10000$  for any  $k \leq t \leq 14k$ , then  $X_k < 14\beta k/10000$ .  $\square$

The last segment of time is the largest, from  $14k$  to  $d^k$ . The idea is to wake a large fraction of the leaves of  $\mathbb{T}_d(v_k)$  and show that this produces a steady stream of frogs to  $v_{k-1}$  up to time  $d^k$ . Corollary 2.5 ensures that with positive probability, the self-similar frog model with a single initial frog wakes a positive fraction of the leaves. Essentially, we need to show that if we start the process with  $\ell$  frogs active at the root, then the chance of waking a positive fraction of the leaves improves exponentially in  $\ell$ .

The trick to doing so is to find many independent opportunities to apply Corollary 2.5, so that we may boost the fixed probability bound to an exponential one. We start by letting the  $\ell$  initial frogs in  $\text{FM}(v_k, \ell)$  move a distance of  $L = \lfloor \log_d \ell/3 \rfloor$  down the tree. By a comparison to placing balls uniformly into bins, we show that these frogs are exponentially likely in  $\ell$  to cover at least  $2/3$  of the vertices at this level. We then apply Corollary 2.5 to the subtrees rooted at the visited vertices to show that each independently has at least probability  $1/2$  of having half its leaves wake in time  $O(k)$ . Since there are  $\Omega(\ell)$  of these

subtrees, it is exponentially likely in  $\ell$  that this occurs for a positive fraction of them. All together, this demonstrates that it is exponentially likely in  $\ell$  that a positive fraction of leaves of  $\mathbb{T}_d^n(v_k)$  are woken in time  $O(k)$ . With this many frogs awake, standard hitting estimates from a leaf to a root give us a steady flow of frogs to  $v_k$  up to time  $d^k$ . We now make this outline precise.

LEMMA 3.10. *With the assumptions of Proposition 3.5,*

$$\mathbf{P}\left[X_t < \frac{\beta t}{10000} \text{ for some } \max(\lceil \ell/\beta \rceil, 14k) \leq t \leq d^k\right] \leq Ce^{-c\ell} \quad (7)$$

for some constants  $c, C > 0$ .

*Proof.* We can assume without loss of generality that  $\ell \geq 3$ , since the  $\ell = 1, 2$  cases can be made trivial by choosing  $C$  large enough. Let  $L = \lfloor \log_d(\ell/3) \rfloor \geq 0$ . Let  $\mathcal{V}_L$  and  $\mathcal{V}_{L+1}$  be respectively the sets of distance  $L$  and  $L + 1$  descendants of  $v_k$  that are not descendants of  $v_{k-1}$ . The restriction  $\ell \leq d^k$  implies that  $v_k$  has at least  $L + 1$  generations of descendants.

For each  $v \in \mathcal{V}_{L+1}$ , couple a self-similar frog model on  $\mathbb{T}^{k-L}$  with  $\text{FM}(v_k, \ell)$  in the same way as in Lemma 3.8. This time, if  $u'$  is the child of the root in  $\mathbb{T}_d^{k-L}$  first visited by the initial frog, then  $\mathbb{T}_d^{k-L}(u')$  is identified with  $\mathbb{T}_d^n(v)$ , and the root of  $\mathbb{T}_d^{k-L}$  is identified with the parent of  $v$ . If  $v$  is ever visited in  $\text{FM}(v_k, \ell)$ , then choose one of its activators and match its path up with the initial frog in the self-similar model. All other aspects of the coupling are as in Lemma 3.8. Note that under this coupling, the self-similar models matched for each  $v \in \mathcal{V}_{L+1}$  are independent.

For  $v \in \mathcal{V}_L$ , let  $A_v$  be an indicator on  $v$  being visited by one of the  $\ell$  initial frogs in  $\text{FM}(v_k, \ell)$ , and let  $A = \sum_{v \in \mathcal{V}_L} A_v$ . The total number of vertices at level  $L$  is  $d^L \leq \ell/3$ , and each initial frog is equally likely to go to any of them. By Lemma 3.7, at least  $2/3$  of these vertices are visited with probability  $1 - e^{-\ell/54}$ . On this event,  $A \geq (2/3 - 1/d)d^L \geq d^L/6$ . Note that all of this holds even when  $L = 0$ , when  $A = A_{v_k} = 1$  deterministically.

Now, condition on  $(A_v)_{v \in \mathcal{V}_L}$ . For every child  $u$  of a vertex  $v \in \mathcal{V}_L$  satisfying  $A_v = 1$ , let  $B_u$  be an indicator on some frog woken at  $v$  moving immediately to  $u$ . By Poisson thinning, there are independently  $\text{Poi}((3 + \beta)d)$  frogs woken at  $v$  moving to  $u$ . Hence, conditional on  $(A_v)_{v \in \mathcal{V}_L}$ , the random variables  $B_u$  for such  $u$  as described above are i.i.d.- $\text{Ber}(e^{-(3+\beta)d})$ .

Call  $u \in \mathcal{V}_{L+1}$  *sustaining* if in the self-similar model coupled to it, at least  $d^{k-L-1}/2$  leaves are activated in  $12(k - L - 1)$  steps. Let  $S_u$  be an indicator on  $u$  being sustaining. The random variables  $(S_u)_{u \in \mathcal{V}_{L+1}}$  are independent of each other

and of all random variables  $A_v$  and  $B_u$  defined in the previous paragraph. Let

$$S = \sum_{\substack{u \text{ child of } v \\ v \in \mathcal{V}_L, A_v=1}} B_u S_u.$$

Conditional on  $(A_v)_{v \in \mathcal{V}_L}$ , the random variable  $S$  is the sum of independent indicators, and  $\mathbf{E}[S \mid A] \geq e^{-(3+\beta)d} Ad/2 \geq .48Ad$  by Corollary 2.5. Conditional on  $A \geq d^L/6$ , we then have  $S \geq d^{L+1}/25$  with probability at least  $1 - e^{-\Omega(d^{L+1})} \geq 1 - e^{-\Omega(\ell)}$ .

Now, we claim that if  $S > d^{L+1}/25$ , it is unlikely that  $X_t < \beta t/10000$  for any  $\lceil \ell/\beta \rceil \leq t \leq d^k$ . On the event  $\{S > d^{L+1}/25\}$ , there are stochastically at least  $\text{Poi}(\mu d^k/50)$  frogs activated by time  $L + 12(k - L - 1)$  at leaves descending from  $v_0$ . As the paths of the frogs at the leaves are independent of  $S$ , conditional on  $S > d^{L+1}/25$  their paths remain independent root-biased nonbacktracking walks. By Lemma 3.6, the number of these frogs that have visited  $v_k$  by time  $L + 12(k - L - 1) + (t + k + 2)$  is stochastically at least  $\text{Poi}(t\mu/200)$  for any  $0 \leq t \leq d^k - k - 2$ . We then thin by  $1/d$  to get the number of frogs frozen at  $v_{k-1}$  after one more step. Hence,

$$X_{13k+t} \geq X_{L+12(k-L-1)+t+k+3} \geq \text{Poi}\left(\frac{t\mu}{200d}\right) = \text{Poi}\left(\frac{t(3+\beta)(d+1)}{200}\right).$$

For  $t \geq k$ ,

$$\frac{t(3+\beta)(d+1)}{200} \geq \frac{\beta(13k+t)}{1000}.$$

By Lemma B.4,

$$\mathbf{P}\left[X_t < \frac{\beta t}{10000} \text{ for some } \max(\lceil \ell/\beta \rceil, 14k) < t \leq d^k \mid S > \frac{d^{L+1}}{25}\right] \leq 2e^{-\Omega\beta\lceil \ell/\beta \rceil} \leq 2e^{-\Omega(\ell)}.$$

Combined with the estimates on  $\mathbf{P}[S > d^{L+1}/25 \mid A \geq d^L/6]$  and on  $\mathbf{P}[A \geq d^L/6]$ , this completes the proof.  $\square$

*Proof of Proposition 3.5.* Lemmas 3.8–3.10 combine via a union bound to prove (4).  $\square$

**3.1.3. Final steps toward Proposition 3.2.** We are already done with the hard work toward proving Proposition 3.2. As we described at the beginning of Section 3.1, our argument requires us to feed  $I_{n-1}$  frogs into  $v_{n-1}$  to get a steady flow into  $v_{n-2}$ , then wait for  $I_{n-2}$  frogs to flow into  $v_{n-2}$ , and so on. What remains is to show that this happens quickly by stitching together the processes  $\text{FM}(v_k, \bullet)$

and applying Proposition 3.5. In our next lemma, we collect  $I_{n-1}$  frogs at  $v_{n-1}$  to set things in motion.

In this section, we have elected to simplify computations by frequent use of big-O notation. We will be very strict in our use of it: an expression  $O(f)$  or  $\Omega(f)$  denotes a quantity bounded respectively from above or from below by  $Cf$ , where  $0 < C < \infty$  is an absolute constant not depending on  $d, n, \mu$ , or any other parameter. For example, the expression  $O(Cn/\beta)$  in the next lemma could be replaced by  $C'n/\beta$ , where  $C'$  is an absolute constant with no dependence on  $d, n, \beta$ , or  $C$ .

LEMMA 3.11. *Consider the nonbacktracking frog model on  $\mathbb{T}_d^n$  with i.i.d.-Poi( $\mu$ ) frogs per site where  $\mu = (3 + \beta)d(d + 1)$ . Given  $C > 0$ , there exists  $\beta_0 = \beta_0(C)$  such that for  $\beta \geq \beta_0$ , there is a stopped version of the frog model with the following property: it holds with probability at least  $1 - e^{-Cdn}$  that at least  $Cdn$  frogs whose last step was from  $v_n = \emptyset$  are frozen at  $v_{n-1}$  by time  $O(Cn/\beta)$  for  $n \geq n_0(C, \beta, d)$ .*

*Proof.* Suppose that some child  $u$  of the root is visited at time  $t$  for the first time. We first mention that we can couple the frog model restricted to  $\{\emptyset\} \cup \mathbb{T}_d^n(u)$  from time  $t$  on with the self-similar frog model on  $\mathbb{T}_d^n$  with frogs frozen at leaves from time 1 on, as we have often done before: simply have all frog paths identical in both models up until time a frog is stopped in the self-similar model. By this coupling and Corollary 2.3, our original frog model has stochastically at least Poi( $cdn$ ) visits to  $\emptyset$  from  $u$  by time  $t + 2cn/\beta$ , assuming that  $\beta$  is large enough that  $2cn/\beta \leq n - 1$ .

We now apply this fact repeatedly to prove the lemma. Let  $\emptyset'$  be the child of the root first visited by the initial frog. The gist of the argument is to couple the frog model on  $\emptyset' \cup \mathbb{T}_d^n(\emptyset')$  with the self-similar model as above to obtain Poi( $cdn$ ) visits to  $\emptyset$  in time  $2cn/\beta$ . From this, we are very likely to visit, say, one third of the children of the root by time  $2cn/\beta + 1$ . For each visited child  $v$ , we couple the frog model on  $\{\emptyset\} \cup \mathbb{T}_d^n(v)$  with the self-similar frog model to get another Poi( $cdn$ ) visits to the root after another  $2cn/\beta$  steps. Summing the contributions from all  $\Omega(d)$  visited children, we have Poi( $cd^2n$ ) visits to the root, and after one more step we have  $\Omega(dn)$  frogs at  $v_{n-1}$ . We will write out this argument with all details below, but we remark that the details are less enlightening than the description we have just given.

We do the argument first in the  $d \geq 3$  case. Let  $c > 1$  be a large constant, to be specified in more detail later. In this argument, we use the phrase *with overwhelming probability* to mean with probability at least  $1 - e^{-\Omega(cdn)}$  for sufficiently large  $n$  (where the meaning of sufficiently large can depend on  $c, d$ ,

and  $\beta$ ). Each instance of the phrase might have a different constant in the  $\Omega(cdn)$  expression. Observe that by a union bound, an intersection of a bounded number of events holding with overwhelming probability also holds with overwhelming probability.

Choose  $\beta_0$  large enough that  $2cn/\beta_0 \leq n - 1$ , and assume that  $\beta \geq \beta_0$ . We then have stochastically at least  $\text{Poi}(cdn)$  visits from  $\varnothing'$  to  $\varnothing$  in time  $2cn/\beta$  by the coupling described above. Each frog that moves from  $\varnothing'$  to  $\varnothing$  moves next outside of  $\{\varnothing, v_{n-1}\}$  with probability at least  $1 - (d+2)/d^2 = \Omega(1)$ , recalling the dynamics of root-biased nonbacktracking walk from Section 2.1. Thus, by time  $2cn/\beta + 1$ , at least  $\text{Poi}(\Omega(cdn))$  frogs have done so. By Proposition B.1, this quantity of frogs is at least  $\Omega(cdn)$  with overwhelming probability. Conditional on this occurring, each of these frogs is equally likely to visit any of the children of the root other than  $\varnothing'$  and  $v_{n-1}$ . By Lemma 3.7, the number of these children visited is strictly greater than  $(d - 2)/3$  with overwhelming probability. Conditional on this, for each child of the root  $v \neq \varnothing', v_{n-1}$  visited, we couple the frog model on  $\{\varnothing\} \cup \mathbb{T}_d^n(v)$  with a self-similar model. For each  $v$ , we then obtain  $\text{Poi}(cdn)$  visits from  $v$  to  $\varnothing$  by time  $2cn/\beta + 1$ , giving us  $\text{Poi}(\Omega(cd^2n))$  such visits in all. Each frog moves next to  $v_{n-1}$  with probability  $(d + 1)/d^2$ , giving us  $\text{Poi}(\Omega(cdn))$  visits to  $v_{n-1}$  from  $\varnothing$  in time  $2cn/\beta + 2$ . Finally, by Proposition B.1, this quantity is at least  $\Omega(cdn)$  with overwhelming probability.

When  $d = 2$ , start the argument the same, obtaining  $\text{Poi}(2cn)$  visits from  $\varnothing'$  to  $\varnothing$  by time  $cn/\beta$ . Depending on whether  $\varnothing' = v_{n-1}$ , each of these frogs moves next to  $v_{n-1}$  with probability  $3/4$  or  $1/4$ . In either case, we have  $\text{Poi}(\Omega(cn))$  frogs moving from  $\varnothing$  to  $v_{n-1}$  in time  $cn/\beta + 1$ , and by Proposition B.1, there are  $\Omega(cn)$  of them with overwhelming probability.

Thus, in both cases, we have  $\Omega(cdn)$  frogs stopped at  $v_{n-1}$  after moving there from  $\varnothing$  in time  $O(cn/\beta)$  with overwhelming probability. Choosing  $c$  to equal  $C$  multiplied by a sufficiently large constant completes the proof.  $\square$

We now prove the equivalent of Proposition 3.2 for the nonbacktracking frog model on  $\mathbb{T}_d^n$ . After this, we will apply Proposition 2.1 to transfer the result to the usual frog model. Recall from (3) that  $J = \lfloor \log_d n + \log_d(\log n) + 5 \log_d 10 - \log_d \beta \rfloor$ .

**PROPOSITION 3.12.** *Consider the nonbacktracking frog model on  $\mathbb{T}_d^n$  with i.i.d.- $\text{Poi}(\mu)$  initial conditions where  $\mu = (3 + \beta)d(d + 1)$ . For any constant  $C$ , for all  $\beta \geq \beta_0(C)$  and  $n \geq n_0(\beta, d, C)$ , there is a stopped version of the model such that at least  $10n \log n$  frogs are stopped at  $v_j$  by time  $O(n \log n/\beta)$  with probability at least  $1 - e^{-Cdn}$ .*

*Proof.* This proof is somewhat long, but it just stitches together the estimates made earlier in the section. We start with an informal sketch. Start with the nonbacktracking frog model on  $\mathbb{T}_d^n$  with i.i.d.-Poi( $2\mu$ ) frogs, splitting the frogs at each site into two collections of Poi( $\mu$ ) each. With the first collection, we run the frog model to accumulate  $O(Cdn)$  frogs at  $v_{n-1}$ , which we can do in time  $O(Cn/\beta)$  with overwhelming probability by Lemma 3.11. We then abandon this first set of frogs and switch to the second, giving ourselves a fresh frog model with i.i.d.-Poi( $\mu$ ) frogs per site but with an extra  $O(Cdn)$  frogs deposited at  $v_{n-1}$ . We can now couple the process with FM( $v_{n-1}, O(Cdn)$ ). Since  $I_{n-1} \leq O(Cdn)$  with overwhelming probability by Proposition 3.5, a steady stream of frogs flows to  $v_{n-2}$ . When  $I_{n-2}$  frogs have built up there, we couple the process to FM( $v_{n-2}, I_{n-2}$ ), and we know that a steady stream of frogs will flow to  $v_{n-3}$ . Continuing in this way, we eventually feed  $I_{J+1}$  frogs in  $v_{J+1}$ , creating a steady stream of frogs into  $v_J$ . After  $O(n \log n/\mu)$  steps, enough frogs have built up at  $v_J$  and we are finished.

Now, we carry out the details. We will be proving our proposition with  $\mu$  replaced by  $2\mu$ , which is equivalent by adjusting  $\beta_0$ . We define a process based on the usual frog model with i.i.d.-Poi( $2\mu$ ) frogs per site in which frogs are repeatedly stopped and restarted. We refer to it as the *slowed process*. To define it, separate the sleeping frogs in  $\mathbb{T}_d^n$  into two independent Poi( $\mu$ )-distributed batches at each vertex. Let the initial frog at the root move as usual, as a root-biased nonbacktracking walk. For sleeping frogs in the first batch, let their paths be root-biased nonbacktracking walks stopped on moving from the root to  $v_{n-1}$ . Keep all second-batch frogs frozen for now.

Let FM( $v_k, \bullet$ ) be independent for all  $J+1 \leq k < n$ . Recall that  $I_k$  is a function of FM( $v_k, \bullet$ ), and hence  $I_{J+1}, \dots, I_{n-1}$  are independent. Once  $I_{n-1}$  frogs have been frozen at  $v_{n-1}$  in the slowed process, unfreeze all frogs accumulated there. Halt all other first-batch frogs at this time and ignore them afterward.

We now allow the second-batch frogs to work at last. When the frogs at  $v_{n-1}$  are unfrozen, couple them with the special frogs in FM( $v_{n-1}, \bullet$ ). Also couple the numbers and paths of second-batch frogs in  $\mathbb{T}_d^n(v_{n-1}) \setminus \mathbb{T}_d^n(v_{n-2})$  with the normal frogs in FM( $v_{n-1}, \bullet$ ). Thus, all frogs move (past the first step) as nonbacktracking walks frozen at  $v_{n-2}$  and  $v_n$ .

Once  $I_{n-2}$  frogs are frozen at  $v_{n-2}$ , halt all other frogs forevermore, and unfreeze these frogs. Couple them and the second-batch frogs in  $\mathbb{T}_d^n(v_{n-2}) \setminus \mathbb{T}_d^n(v_{n-3})$  with FM( $v_{n-2}, \bullet$ ) as above. Let all frogs move until  $I_{n-3}$  frogs have been frozen at  $v_{n-3}$ . We continue on in this way until  $I_{J+1}$  frogs are frozen at  $v_{J+1}$ . We then continue for one last step, unfreezing the frogs at  $v_{J+1}$ , halting all other ones permanently, and coupling the process with FM( $v_{J+1}, \bullet$ ). Finally, we run the process until  $10n \log n$  frogs are frozen at  $v_J$ .

We claim that to prove this proposition, it suffices to prove the same bound for the slowed process. This is intuitively very clear: If we remove all of the stops and restarts at vertices other than  $v_J$ , the resulting model is a stopped version of the nonbacktracking frog model. Furthermore, every frog that is stopped at  $v_J$  by a given time in the slowed process will also be stopped at  $v_J$  by this time in the stopped process. Hence, it suffices to prove that at least  $10n \log n$  frogs are stopped at  $v_J$  in the slowed process at time  $O(Cn \log n/\beta)$  with probability at least  $1 - e^{-Cdn}$ .

The rest of the proof is to show this. We claim that at least  $10n \log n$  frogs are stopped at  $v_J$  in the slowed process at time  $O(n \log n/\beta)$  if all of the following events occur:

**Event  $A_1$ :** The time to accumulate  $I_{n-1}$  frogs at  $v_{n-1}$  in the first step of the process is at most  $O(Cn/\beta)$ .

**Event  $A_2$ :** For all  $J + 1 \leq k \leq n - 1$ , it holds that  $I_k \leq d^k$ .

**Event  $A_3$ :** It holds that  $I_{n-1} + \dots + I_{J+1} \leq n \log n$ .

Indeed, suppose all these events occur. For  $J + 2 \leq k \leq n - 1$ , let

$$T_k = \max\left(3, \left\lceil \frac{10000I_{k-1}}{\beta} \right\rceil, \left\lceil \frac{I_k}{\beta} \right\rceil\right).$$

From event  $A_1$ , there will be  $I_{n-1}$  frogs at  $v_{n-1}$  by time  $O(Cn/\beta)$  or sooner, starting the stage of the slowed process in which it evolves according to  $\text{FM}(v_{n-1}, \bullet)$ . By Definition 3.4, the process  $\text{FM}(v_{n-1}, I_{n-1})$  sends at least  $\beta t/10000$  frogs to  $v_{n-2}$  in  $t$  steps for all  $\max(3, \lceil I_{n-1}/\beta \rceil) \leq t \leq d^{n-1}$ . From event  $A_2$ , we have  $I_{n-1} \leq d^{n-1}$  and  $I_{n-2} \leq d^{n-2}$ . Hence  $T_{n-1}$  lies between  $\max(3, \lceil I_{n-1}/\beta \rceil)$  and  $d^{n-1}$ , and therefore  $\text{FM}(v_{k-1}, I_{k-1})$  sends at least  $I_{n-2}$  frogs to  $v_{n-2}$  in  $T_{n-1}$  steps. By our construction of the slowed process, this kicks off the next stage of the process, which is coupled to  $\text{FM}(v_{n-2}, \bullet)$ . By identical reasoning,  $\text{FM}(v_{n-2}, I_{n-2})$  sends at least  $I_{n-3}$  frogs to  $v_{n-3}$  in  $T_{n-2}$  steps. Continuing in this way, we send at least  $I_{n-4}$  frogs to  $v_{n-4}$  in another  $T_{n-3}$  steps, and so on, culminating with the arrival of  $I_{J+1}$  frogs to  $v_{j+1}$ . Finally, let

$$T_{J+1} = \max\left(3, \left\lceil \frac{10^5 n \log n}{\beta} \right\rceil, \left\lceil \frac{I_{J+1}}{\beta} \right\rceil\right).$$

By the definition of  $J$ , we have  $10^5 n \log n/\beta \leq d^{J+1}$ . From  $A_2$ , we have  $I_{J+1} \leq d^{J+1}$ . Thus  $T_{J+1}$  lies between  $\max(3, \lceil I_{J+1}/\beta \rceil)$  and  $d^{J+1}$ , from which it follows



that  $\text{FM}(v_{J+1}, I_{J+1})$  sends at least  $10n \log n$  frogs to  $v_J$  in  $T_{J+1}$  steps. All together, we send at least  $10n \log n$  frogs to  $v_J$  in time

$$O\left(\frac{Cn}{\beta}\right) + T_{n-1} + \cdots + T_{J+1}. \quad (8)$$

Assuming event  $A_3$  holds, we have  $T_{n-1} + \cdots + T_J = O(n \log n / \beta)$ . Thus, (8) is  $O(n \log n / \beta)$  for large enough  $n$  (depending on  $C$ ), completing the proof of the claim.

All that remains is to show that  $A_1 \cap A_2 \cap A_3$  occurs with probability at least  $1 - e^{-cdn}$ . Let  $c$  be a constant to be chosen later (it will depend only on  $C$ ). To bound the probability of  $A_1$ , observe that  $cdn$  frogs are frozen at  $v_{n-1}$  in the first stage of the process by time  $O(cn/\beta)$  with probability at least  $1 - e^{-cdn}$  by Lemma 3.11. By Proposition 3.5, we have  $\mathbf{P}[I_{n-1} \leq cdn] \geq 1 - O(1)e^{-\Omega(cdn)}$ . These two facts together show that  $A_1^c$  occurs with probability  $O(1)e^{-\Omega(cdn)}$ , provided the implicit constant in big-O expression in the definition of  $A_1$  is chosen large enough.

To bound the probability of  $A_2$ , first observe that  $d^J \geq 10^5 n \log n / \beta$ . Then apply Proposition 3.5 and obtain the inequality

$$\begin{aligned} \mathbf{P}[I_k > d^k] &\leq \mathbf{P}[I_k > d^{k-J} (10^5 n \log n / \beta)] \\ &\leq O(1) \exp(-\Omega(d^{k-J} n \log n / \beta)) \end{aligned}$$

for all  $J + 1 \leq k \leq n - 1$ . Hence, by a union bound,

$$\mathbf{P}[A_2^c] \leq \sum_{k=J+1}^{n-1} O(1) \exp(-\Omega(d^{k-J} n \log n / \beta)) = O(1)e^{-\Omega(dn \log n / \beta)}.$$

For large enough  $n$  (depending only on  $c$  and  $\beta$ ), this is bounded by  $e^{-cdn}$ .

Last, we consider the event  $A_3$ . Let  $\bar{I}_k = \min(I_k, d^k)$ , so that  $\bar{I}_k$  has an exponential tail by Proposition 3.5. By Proposition B.2,

$$\mathbf{P}[\bar{I}_{n-1} + \cdots + \bar{I}_{J+1} > n \log n] \leq e^{-cdn}$$

once  $n$  is large enough relative to  $d$  and  $c$ . If  $A_2$  holds and  $\bar{I}_{n-1} + \cdots + \bar{I}_{J+1} \leq n \log n$ , then  $A_3$  holds as well, showing that  $\mathbf{P}[A_3^c] \leq 2e^{-cdn}$  for large enough  $n$ , depending on  $c$  and  $\beta$ . We now have

$$\mathbf{P}[A_1^c] + \mathbf{P}[A_2^c] + \mathbf{P}[A_3^c] = O(1)e^{-\Omega(cdn)}.$$

The proof is now completed by choosing  $c$  large enough that this is smaller than  $e^{-cdn}$ .  $\square$

*Proof of Proposition 3.2.* Set  $b = 4 \log d$  and apply Proposition 2.1 to the result of Proposition 3.12.  $\square$

**3.2. Establishing Proposition 3.3.** In this section we prove Proposition 3.3. Note that this is a result about random walks on trees, not the frog model. It will be based on the following random walk estimate. Recall from (3) that  $J = \lfloor \log_d(10^5 n \log n / \beta) \rfloor$ , where  $\mu = (3 + \beta)d(d + 1)$ .

**PROPOSITION 3.13.** *Consider a single random walk on  $\mathbb{T}_d^n$  started at  $v_J$  and assume that  $n \geq n_0$  for some sufficiently large absolute constant  $n_0$ . The walk visits  $v_0$  in less than  $4(10)^5 n \log n / \beta$  steps with probability at least  $1/3 \log_d n$ .*

Using this, the proof of Proposition 3.3 is easy:

*Proof of Proposition 3.3.* By Proposition 3.13 the probability that none of the  $10n \log n$  frogs at  $v_J$  reaches  $v_0$  in  $4(10)^5 n \log n / \beta$  steps is at most

$$\left(1 - \frac{1}{3 \log_d n}\right)^{10n \log n} \leq e^{-3n \log d}. \quad \square$$

Now we devote the rest of this section to establishing the random walk estimate. Its proof works by decomposing the random walk as a simple random walk on the spine  $\{v_0, \dots, v_n\}$  with excursions off of it. We start with a preliminary lemma to compute the expected length of the excursions.

**LEMMA 3.14.** *Let  $\tau_k$  be the number of steps to hit either  $v_{k-1}$  or  $v_{k+1}$  for a simple random walk on  $\mathbb{T}_d^n$  starting at  $v_k$ . Then  $\mathbf{E}\tau_k = d^{k-1}(d - 1)/2$ .*

*Proof.* The time to hit  $v_{k-1}$  or  $v_{k+1}$  is the same as the time by random walk starting at  $v_k$  on the weighted graph shown in Figure 3 to hit the leftmost vertex,  $\{v_{k-1}, v_{k+1}\}$ . The random walk moves at each step to a neighbor chosen with probability proportionate to the weight of the edge. The graph has been obtained from  $\mathbb{T}_d^n$  by identifying  $v_{k-1}$  and  $v_{k+1}$ , identifying all children of  $v_k$  other than  $v_{k-1}$ , and identifying all distance  $k$  descendants of  $v_k$  for each  $k \geq 2$ .

The expected hitting time is easily computed using electrical network theory. By [LP16, Proposition 2.20], which is a result first obtained in [Tet91], the hitting time has expectation  $\sum_x \pi(x)v(x)$ , where the sum is over all vertices in the graph,  $\pi(x)$  denotes the sum of the weights of edges incident to  $x$ , and  $v$  is the voltage that assigns 0 to the vertex  $\{v_{k-1}, v_{k+1}\}$  and that creates a unit current flow from  $v_k$  to  $\{v_{k-1}, v_{k+1}\}$ . This voltage assigns  $1/2$  to all vertices other than  $\{v_{k-1}, v_{k+1}\}$ . The expected hitting time is then

$$\frac{1}{2} \sum_{i=0}^{k-1} (d^i(d-1) + d^{i+1}(d-1)) + \frac{1}{2}(d+1) = \frac{1}{2}d^{k-1}(d-1). \quad \square$$

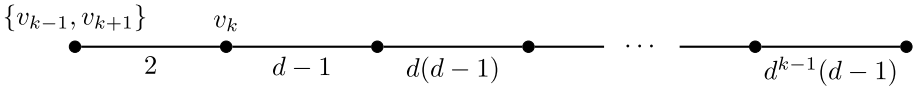


Figure 3. The following collections of vertices from  $\mathbb{T}_d^n$  have been identified in this graph:  $v_{k-1}$  and  $v_{k+1}$ ; all children of  $v_k$  other than  $v_{k-1}$ ; and the distance  $k$  descendants of  $v_k$  for each  $k \geq 2$ . Random walk moving with probability proportionate to the edge weights starting at  $v_k$  and stopping at  $\{v_{k-1}, v_{k+1}\}$  is the same as random walk on the original graph, viewing vertices as blocked together.

Next, we compute the expected number of visits to each vertex along the spine before  $v_0$  is hit. We condition on the walk hitting  $v_0$  before  $v_{J+1}$ , as this will simplify our eventual proof.

LEMMA 3.15. Consider a random walk on  $\mathbb{T}_d^n$  starting at  $v_J$ . Let  $\sigma_k$  be the first time that the walk hits  $v_k$ . Let  $V_k$  be the total number of visits to  $v_k$  up to time  $\sigma_0$ . For  $1 \leq k \leq J$ ,

$$\mathbf{E}[V_k \mid \sigma_0 < \sigma_{J+1}] = 2k \left( 1 - \frac{k}{J+1} \right).$$

*Proof.* Let  $E = \{\sigma_0 < \sigma_{J+1}\}$ . Recall that  $\mathbf{P}[\sigma_k < \sigma_{J+1}] = 1/(J+1-k)$ . Conditioned on  $E$ , the walk will visit  $v_k$  at least once for all  $k \leq J$ . The number of returns to  $v_k$  after first visiting it is a geometric random variable with parameter  $1 - \mathbf{P}[V_k = 1 \mid E]$ . We can then write

$$\begin{aligned} \mathbf{E}[V_k \mid E] &= \frac{1}{\mathbf{P}[V_k = 1 \mid \sigma_0 < \sigma_{J+1}]} = \frac{\mathbf{P}[\sigma_0 < \sigma_{J+1}]}{\mathbf{P}[\sigma_0 < \sigma_{J+1} \text{ and } V_k = 1]} \\ &= \frac{1}{(J+1)\mathbf{P}[\sigma_0 < \sigma_{J+1} \text{ and } V_k = 1]}. \end{aligned} \quad (9)$$

We claim that

$$\mathbf{P}[\sigma_0 < \sigma_{J+1} \text{ and } V_k = 1] = \frac{1}{2(J+1-k)k}. \quad (10)$$

This is because to reach  $v_1$  before  $v_{J+1}$ , the walk necessarily visits  $v_k$  before  $v_{J+1}$ , which occurs with probability  $1/(J+1-k)$ . To visit  $v_k$  only once, on arriving at  $v_k$  it must immediately move to  $v_{k-1}$ , which occurs with probability  $1/2$ . Then it must reach  $v_0$  before  $v_k$ , which occurs with probability  $1/k$ . Combining (9) and (10) gives the claimed formula.  $\square$

*Proof of Proposition 3.13.* Let  $(S_t)$  be a simple random walk on  $\mathbb{T}_d^n$  starting at  $v_k$ . Define  $\mathcal{S} = \{v_0, \dots, v_n\}$ . The Markov property of random walk shows that the restriction to  $\mathcal{S}$  of the path of  $(S_t)$  is distributed as the path of a simple random walk on  $\mathcal{S}$ . Let  $\sigma_k = \inf\{t : S_t = v_k\}$ , the hitting time of  $v_k$ , as in Lemma 3.15. Set

$$F = \{\sigma_0 \leq 4(10)^5 n \log n / \beta\},$$

$$E = \{\sigma_0 < \sigma_{J+1}\}.$$

Our goal is to bound  $\mathbf{P}[F]$  from below. A simple estimate gives

$$\mathbf{P}[F] \geq \mathbf{P}[F \cap E] = \mathbf{P}[F \mid E]\mathbf{P}[E] = \frac{\mathbf{P}[F \mid E]}{J}. \tag{11}$$

In light of (11) it suffices to prove that  $\mathbf{P}[F \mid E] \geq 1/2$ .

Let  $V_k = |\{t \leq \sigma_0 : S_t = v_k\}|$ , the total number of visits to  $v_k$  before the walk hits  $v_0$ . Let  $\tau_k(i)$  be the number of steps it takes the walk to reach  $v_{k-1}$  or  $v_{k+1}$  starting from the  $i$ th time the walk arrives at  $v_k$ . We then decompose  $\sigma_0$  as

$$\sigma_0 = \sum_{k=1}^n \sum_{j=1}^{V_k} \tau_k(j).$$

Conditional on  $E$ , the random variables  $V_k$  and  $\tau_k(j)$  are mutually independent for all  $j$  and  $k$ . By Wald’s lemma,

$$\mathbf{E}[\sigma_0 \mid E] = \sum_{k=1}^J \mathbf{E}[V_k \mid E] \mathbf{E}[\tau_k(j) \mid E]. \tag{12}$$

We need only consider  $J$  summands in (12), since conditional on  $E$  the walk does not move beyond  $v_J$  before hitting  $v_0$ . For all  $j$ , the random variable  $\tau_k(j)$  is independent of  $E$  and is distributed as  $\tau_k$  from Lemma 3.14. Therefore, by Lemmas 3.14 and 3.15,

$$\mathbf{E}[\sigma_0 \mid E] = \sum_{k=1}^J k \left(1 - \frac{k}{J+1}\right) d^{k-1} (d-1).$$

We claim that this is  $O(d^J)$ . Indeed, using the bound  $k(1 - k/(J+1)) \leq J+1 - k$  and making the substitution  $j = J+1 - k$  in the second line,

$$\begin{aligned}
 \mathbf{E}[\sigma_0 \mid E] &\leq (d-1) \sum_{k=1}^J (J+1-k)d^{k-1} \\
 &= d^{J-1}(d-1) \sum_{j=1}^J jd^{1-j} \\
 &\leq d^{J-1}(d-1) \sum_{j=1}^{\infty} jd^{1-j} = d^{J-1}(d-1)(1-d^{-1})^{-2} = \frac{d^{J+1}}{d-1}.
 \end{aligned}$$

Notice that  $d^{J+1} \leq 10^5 dn \log n / \beta$  and apply Markov's inequality to obtain

$$\mathbf{P}[\sigma_0 > 4(10)^5 n \log n / \beta \mid E] \leq \frac{d}{4(d-1)} \leq \frac{1}{2}.$$

Applying this to (11) gives

$$\mathbf{P}[\sigma_0 \leq 16(10)^5 n \log n / \beta] \geq \frac{1}{2J} \geq \frac{1}{3 \log_d n},$$

with the last inequality holding for all sufficiently large  $n$ , with no dependence on  $d$ .  $\square$

#### 4. Slow cover time for small $\mu$

We now give our lower bound on the cover time for small enough  $\mu$ .

**THEOREM 4.1.** *Let  $\mathcal{C}$  be the cover time for the frog model on  $\mathbb{T}_d^n$  with initial frog counts given by an independent collection of random variables  $(\eta(v))_{v \in \mathbb{T}_d^n \setminus \{\emptyset\}}$ , where  $\mathbf{E}\eta(v) \leq \mu$  for all vertices  $v$ . Suppose that  $\mu \leq \min(d^{1-\epsilon}, d/100)$  for any  $0 < \epsilon < 1$ . For some absolute constant  $c > 0$ ,*

$$\mathbf{P}[\mathcal{C} < e^{c\sqrt{\epsilon n \log d}}] \leq e^{-c\sqrt{\epsilon n \log d}}$$

for  $n \geq \log d / c^2 \epsilon$ .

This bound is effective even for  $\mu$  as large as  $d/100$ :

**COROLLARY 4.2.** *Let  $\mathcal{C}$  be the cover time as above, assuming only that  $\mu \leq d/100$ . Then for some absolute constant  $c > 0$ ,*

$$\mathbf{P}[\mathcal{C} < e^{c\sqrt{n}}] \leq e^{-c\sqrt{n}}$$

for  $n \geq ((\log d)/c)^2$ .

*Proof.* Apply Theorem 4.1 with  $\epsilon = \log_d 100$ . □

We extend the usual notion of the distribution  $\text{Ber}(\mu)$  to  $\mu > 1$  by setting it to be the unique distribution on  $\{\lfloor \mu \rfloor, \lceil \mu \rceil\}$  with mean  $\mu$ . For most of this section, we consider the frog model with i.i.d.- $\text{Ber}(\mu)$  initial conditions. We then apply Lemma A.1 to allow for more general initial conditions.

The proof hinges on the following result that we will prove inductively. Define  $\mathbb{T}_d^{H*}$  to be the  $d$ -ary tree of height  $H$  with an extra vertex,  $y$ , attached to the root.

**PROPOSITION 4.3.** *For some absolute constant  $C > 0$ , the following statement holds for all  $d \geq 2$  and  $\mu \leq d/100$ . Consider the frog model on  $\mathbb{T}_d^{H*}$  with one initially active frog at the root, none at  $y$ , and i.i.d.- $\text{Ber}(\mu)$  sleeping frogs at the remaining vertices, and with frogs frozen on moving to  $y$ . Let  $X^{(j,H)}$  be the number of frogs frozen at  $y$  by time  $2^j$ . Define  $H_j = H_j(d, \mu)$  by*

$$H_1 = 1, \\ H_j = \left\lceil \frac{Cj(\log(1 + \mu) + j)}{\log\left(\frac{d}{1+\mu}\right)} \right\rceil, \quad j \geq 2.$$

For any  $j \geq 1$ , if  $n \geq H_j$ , then

$$\mathbf{E}X^{(j,n)} \leq \frac{.8}{1 + \frac{2d}{d-1}\mu}. \tag{13}$$

Most of this section is devoted to proving Proposition 4.3. Before we turn to this, we prove Theorem 4.1 from it. First, it is a small task to remove the freezing of frogs from Proposition 4.3, showing that the expected number of returns to the root within time  $2^j$  in our usual frog model on  $\mathbb{T}_d^{H_j}$  is  $O(1)$ .

**COROLLARY 4.4.** *Suppose that  $\mu \leq d/100$ , and let*

$$j = j(d, n, \mu) = \max\{i : H_i \leq n - 1\},$$

where  $H_i$  is the sequence defined in Proposition 4.3. Let  $R$  be the total number of visits to the root of  $\mathbb{T}_d^n$  within time  $2^j$  in the frog model with initial frog counts given by  $\eta(v)$  for  $v \in \mathbb{T}_d^n \setminus \{\emptyset\}$ . If  $\mathbf{E}\eta(v) \leq \mu$  for all vertices  $v$ , then  $\mathbf{E}R \leq 4$ .

*Proof.* It suffices to prove this result under i.i.d.- $\text{Ber}(\mu)$  initial conditions, by Lemma A.1 and the maximality of  $\text{Ber}(\mu)$  in the pgf order mentioned in Appendix A. Now, consider the following modification of the frog model. Let the initial frog take a step. Next, run the frog model for  $2^j$  steps with frogs frozen

at the root, and kill all frogs that were woken but did not reach the root. Let  $R_1$  be the number of frogs frozen at the root. Now, let each of these frogs take one more step, and then run the frog model for another  $2^j$  steps with frogs frozen at the root, and then again kill any frogs that were woken but did not reach the root. Let  $R_2$  be the number of frogs frozen at the root after this stage. Continue in this way to define  $R_i$  for  $i \geq 3$ . As every frog is allowed to run for at least  $2^j$  steps before being killed, every visit to the root in the usual frog model in the first  $2^j$  steps occurs eventually in this modified process. Hence,  $R \leq \sum_{i=1}^{\infty} R_i$ .

Defining  $R_0 = 1$ , we claim that

$$\mathbf{E}R_{i+1} \leq (1 + \mu)\mathbf{E}X^{(j,n-1)}\mathbf{E}R_i \quad (14)$$

for all  $i \geq 0$ . We prove this statement now. After the  $i$ th step of the process, there are  $R_i$  frogs at the root. Let  $N_i$  be the number of active frogs at level 1 after they take their next steps. Fix  $1 \leq k \leq N_i$ , and suppose that the  $k$ th of these frogs follows the path  $(S_0, S_1, \dots)$  from this point on. Consider the original (unmodified) frog model with the following changes:

- (i) Add an initially active frog with path  $(S_0, S_1, \dots)$ ;
- (ii) delete all other frogs at vertex  $S_0$ , and delete the frog at the root;
- (iii) freeze frogs on moving to the root.

Let  $X_k$  be the number of frogs frozen at the root after  $2^j$  steps in this modified process. By a subadditivity property of the frog model,  $R_{i+1} \leq \sum_{k=1}^{N_i} X_k$ . Now, we think of the root vertex as  $y$ , and we think of  $X_k$  as counting the number of visits to  $y$  in a frog model on  $\mathbb{T}_d^{(n-1)*}$  with frogs frozen at  $y$ , except that because of killing frogs, some vertices of  $\mathbb{T}_d^{(n-1)*}$  have no sleeping frogs on them. Thus, conditional on  $N_i$ , we have  $X_k \leq X^{(j,n-1)}$ . Hence,

$$\mathbf{E}[R_{i+1} \mid N_i] \leq (\mathbf{E}X^{(j,n-1)})N_i.$$

Taking expectations and observing that  $\mathbf{E}[N_i \mid R_i] \leq (1 + \mu)R_i$  completes the proof of (14).

By Proposition 4.3 and our choice of  $j$ ,

$$(1 + \mu)(\mathbf{E}X^{(j,n-1)}) \leq \frac{.8(1 + \mu)}{1 + \frac{2d}{d-1}\mu} \leq .8.$$

It now follows from (14) that

$$\mathbf{E}R \leq \sum_{i=1}^{\infty} \mathbf{E}R_i \leq \sum_{i=1}^{\infty} (.8)^i = 4. \quad \square$$

Corollary 4.4 shows that in the frog model on  $\mathbb{T}_d^n$ , there are few visits to the root by time  $2^{j(d,n,\mu)}$ . To bound the cover time, we observe that once all frogs are active, many visits to the root will occur. We first give a random walk estimate.

LEMMA 4.5. *For some absolute constants  $a, b > 0$  the following statement holds. Suppose that  $n \log d/a \leq t \leq d^n$ . Then a random walk on  $\mathbb{T}_d^n$  with arbitrary starting position has probability at least  $btd^{-n}$  of hitting the root in its first  $t$  steps.*

*Proof.* One could prove more precise estimates in the same way as Lemma 3.6. Since we do not need any precise formula, we take a simpler approach. We can assume the walk starts at a leaf, as this is the worst-case scenario. Now, partition the walk into excursions away from level  $n$ . The length of each excursion has an exponential tail, since the probability that a random walk on  $\mathbb{Z}$  from 0 with a bias to the right is negative after  $k$  steps decays exponentially in  $k$ . By Proposition B.2, the probability of having  $\epsilon t$  or fewer excursions from level  $n$  in time  $t$  is at most  $e^{-ct}$  for absolute constants  $\epsilon$  and  $c$ . On each excursion, the walk has probability  $(d-1)/(d^n-1) \geq d^{-n}$  of visiting the root. Thus, in  $\lceil \epsilon t \rceil$  excursions, the probability that the root will not be visited is at most

$$(1 - d^{-n})^{\lceil \epsilon t \rceil} \leq e^{-\epsilon t d^{-n}}.$$

Combining these two estimates, the root is visited in time  $t$  with probability at least

$$1 - e^{-\epsilon t d^{-n}} - e^{-ct}.$$

Since  $t \leq d^n$ , we can apply the inequality  $1 - e^{-x} \geq x/2$ , which holds for  $x \in [0, 1]$ , to get

$$1 - e^{-\epsilon t d^{-n}} - e^{-ct} \geq \frac{\epsilon t d^{-n}}{2} - e^{-ct}.$$

Choosing  $a$  small enough, this is  $\Omega(td^{-n})$ . □

*Proof of Theorem 4.1.* Define  $j = j(d, n, \mu)$  as in Corollary 4.4. We start by estimating  $j$ . Directly calculating from the definition of  $H_i$  in Proposition 4.3, we find that if  $i \geq \log d$ , then

$$H_i \leq \frac{Ci^2}{\epsilon \log d}$$



for some absolute constant  $C$ . If we set  $i = \lceil c\sqrt{\epsilon n \log d} \rceil$  for  $c = (2C)^{-1/2}$  and assume  $n \geq c^{-2} \log d / \epsilon$  so that  $i \geq \log d$ , then we have  $H_i \leq n - 1$ . Hence  $j \geq c\sqrt{\epsilon n \log d}$ . It is also straightforward to see that  $j = O(\sqrt{n \log d})$ .

It does us no harm to assume that  $\mu > .01$ . For technical reasons, we will also assume that the expected number of sleeping frogs is exactly  $\mu$  at each site, rather than just being bounded by  $\mu$ . To see that it suffices to prove the theorem under this extra assumption, for each site with expected count strictly smaller than  $\mu$ , independently add a random number of extra frogs (distributed arbitrarily) to bring the mean up to  $\mu$ , and observe that this can only decrease the cover time.

Define the event  $A = \{C < 2^{j-1}\}$ . We will prove that  $\mathbf{P}[A] \leq C2^{-j}$  for some absolute constant  $C$ . By the lower bounds on  $j$ , this proves the theorem with an extra constant  $C$  in front of the bound, which we can eliminate by decreasing  $c_\epsilon$  or  $c$  slightly. Conditional on  $A$ , all frogs are awake at time  $2^{j-1}$ , and they move from this time on as independent simple random walks. We can apply Lemma 4.5 with  $t = 2^{j-1}$ , since  $n \log d / a \leq 2^{j-1} \leq d^n$  for large enough  $n$ , showing that each walk hits the root by time  $2^j$  with probability at least  $b2^{j-1}d^{-n}$ . Let  $R$  be the total number of visits to the root by time  $2^j$  and let  $U$  be the total number of frogs in the system. Bounding  $R$  from below by counting the visits to the root only for times in  $[2^{j-1}, 2^j]$ , we obtain

$$\mathbf{E}[R \mid A] \geq b2^{j-1}d^{-n}\mathbf{E}[U \mid A].$$

By a simple coupling, the event  $A$  is more likely the larger  $U$  is. That is, the random variables  $\mathbf{1}_A$  and  $U$  are positively associated, from which it follows that  $\mathbf{E}[U \mid A] \geq \mathbf{E}U \geq \mu d^n$ , recalling that we have assumed that each site has exactly mean  $\mu$  sleeping frogs. Thus,  $\mathbf{E}[R \mid A] \geq b2^{j-1}\mu$ . But by Corollary 4.4, we have  $\mathbf{E}R \leq 4$ . Rearranging the simple bound  $\mathbf{E}R \geq \mathbf{E}[R \mid A]\mathbf{P}[A]$  gives

$$\mathbf{P}[A] \leq \frac{4}{b2^{j-1}\mu} = O(2^{-j}),$$

under our assumption that  $\mu \geq .01$ . □

**4.1. Tagging frogs.** The remainder of Section 4 is devoted to proving Proposition 4.3. Fix integers  $H, h, j \geq 1$ , and consider the frog model on  $\mathbb{T}_d^{(H+h)*}$  with frogs frozen on moving to  $y$ , starting with one frog at the root, and with i.i.d.-Ber( $\mu$ ) frogs per site at all vertices besides the root and  $y$ . Let  $\mathcal{L}_i$  denote the set of vertices at level  $i$  of  $\mathbb{T}_d^{(H+h)*}$ , taking 0 as the level of the root and  $-1$  as the level of  $y$ .

Our plan is to advance the induction in Proposition 4.3 by supposing that  $X^{(j,H)}$  satisfies the inductive hypothesis and then showing that  $X^{(j+1,H+h)}$  does as well,

for a good choice of  $h$ . The idea of the proof is to assign each frog a tag that changes at various times in the process. When one frog wakes another, the newly woken frog starts with the same tag as its waker. If a frog is woken by two frogs with different tags arriving simultaneously, choose any procedure to decide between the frogs; this detail will prove irrelevant. In the following set of rules, when a frog changes its tag on arriving at a given vertex, the newly woken frogs inherit the new tag, not the old one.

- The initially active frog at the root has tag  $A$ .
- If an  $A$ -tagged frog reaches  $\mathcal{L}_h$ , its tag changes to  $B_0$ .
- If a  $B_i$ -tagged frog moves from  $\mathcal{L}_{h-1}$  to  $\mathcal{L}_h$ , its tag changes to  $B_{i+1}$ .
- If a  $B_i$ -tagged frog moves from  $\mathcal{L}_h$  to  $\mathcal{L}_{h-1}$  at time  $2^j$  or after, its tag changes to  $C_0$ .
- If a  $C_i$ -tagged frog moves from  $\mathcal{L}_{h-1}$  to  $\mathcal{L}_h$ , its tag changes to  $C_{i+1}$ .
- At time  $2^{j+1} + 1$ , all frogs are stripped of their tags.

Note that frogs are retagged every time they move forward in the tree to  $\mathcal{L}_h$ . The only other time a frog receives a new tag is when a  $B_i$ -tagged frog moves backward from  $\mathcal{L}_h$  to  $\mathcal{L}_{h-1}$  at time  $2^j$  or later, in which case its tag changes to  $C_0$ .

We will use three different estimates to bound the number of tagged frogs. When frogs with any tag are between the root and  $\mathcal{L}_h$ , we dominate them by branching random walks using the estimates in Lemmas 4.8 and 4.9. When a frog moves forward in the tree to a vertex  $v \in \mathcal{L}_h$  and is given tag  $B_i$ , we estimate the number of  $B_i$ -tagged frogs emerging from  $v$  back to  $\mathcal{L}_{h-1}$  using (13), the inductive hypothesis. We have very little control over the number of  $C_0$ -tagged particles emerging from  $v$  back to  $\mathcal{L}_{h-1}$ . Here, we use Lemma 4.6, which we call the *all-awake bound* since it simply assumes that all frogs in the subtree rooted at  $v$  are initially awake. The key to the argument is that we retain control over the number of  $C_{i+1}$ -tagged frogs: Whenever a  $C_i$ -tagged frog moves forward to a vertex  $v \in \mathcal{L}_h$  and is retagged as  $C_{i+1}$ , it does so after time  $2^j$ . Since we only care about the process up to time  $2^{j+1}$ , we can control the number of  $C_{i+1}$ -tagged frogs emerging from  $v$  back to  $\mathcal{L}_{h-1}$  using the inductive hypothesis rather than the all-awake bound. Thus, although there will be many  $C_0$ -tagged frogs, the number of  $C_i$ -tagged frogs for  $i \geq 1$  will not spiral out of control.

**4.2. Branching random walk and all-awake bounds.** As mentioned above, we control the frog model by dominating it by branching random walk and by

simply assuming that all frogs in a given subtree are initially awake. We start with this second bound.

LEMMA 4.6 (All-awake bound). *Consider  $\mathbb{T}_d^{H*}$  for arbitrary  $H \geq 1$  with one particle at the root, none at  $y$ , and i.i.d.- $\text{Ber}(\mu)$  particles at the remaining vertices. Let all particles perform discrete-time random walks frozen at  $y$ . Let  $W$  be the total number of particles frozen at  $y$  after  $t$  time steps. For some constant  $c_1$ ,*

$$\mathbf{E}W \leq c_1 \mu t.$$

*Proof.* For any  $0 \leq k \leq H - 1$ , a particle initially at level  $k$  of the tree visits  $y$  before the leaves with probability no more than  $d^{-k-1}$ . Initially, there is one particle at level 0 and an expected  $\mu d^k$  particles at level  $k$  for each  $1 \leq k \leq H - 1$ . Only particles starting at level  $t - 1$  or less can reach  $y$  in time  $t$ . Hence, the expected number of particles that reach the root in  $t$  steps without ever being at a leaf is at most

$$d^{-1} + \sum_{k=1}^{\min(H,t)-1} \mu d^k d^{-k-1} \leq \frac{1 + t\mu}{d}. \quad (15)$$

Now, consider a particle at a leaf. It has probability no more than  $d^{-H}$  of visiting  $y$  before revisiting level  $H$ . In time  $t$ , it makes no more than  $t$  of these excursions from the leaves. Thus, the probability that a given particle at a leaf visits  $y$  in its next  $t$  steps is at most

$$1 - (1 - d^{-H})^t \leq 1 - e^{-2d^{-H}t} \leq 2td^{-H}.$$

The first inequality above uses the bound  $1 - x \geq e^{-2x}$ , which holds for all  $x \in [0, 1/2]$ . The total expected number of particles in the tree is  $1 + \mu(d + \dots + d^H)$ . The expected number of particles that visit  $y$  before time  $t$ , starting from a leaf or after visiting a leaf, is therefore at most

$$2td^{-H}(1 + \mu(d + \dots + d^H)) \leq 2t \left( d^{-H} + \frac{\mu}{1 - d^{-1}} \right). \quad (16)$$

Combining (15) and (16),

$$\begin{aligned} \mathbf{E}W &\leq \frac{1}{d} + \left( \frac{\mu}{d} + 2d^{-H} + \frac{2\mu}{1 - d^{-1}} \right)t \\ &\leq \frac{1}{2} + \left( \frac{\mu}{2} + 1 + 4\mu \right)t = O(\mu t). \quad \square \end{aligned}$$

Next, we prove several bounds whose proofs are essentially comparisons of the frog model to branching random walk. The first step is to describe a supermartingale  $w_\theta(\xi_t)$  given as a function of the frog model.

LEMMA 4.7. Consider  $\mathbb{T}_d^{h*}$  with a single active frog at a specified vertex  $v_0$ , no frogs at the ancestors of  $v_0$  (including  $y$ ), and i.i.d.- $\text{Ber}(\mu)$  sleeping frogs at the other vertices. Run the frog model with frogs frozen on arrival to  $y$  and to  $\mathcal{L}_h$ . (When a frog arrives at  $\mathcal{L}_h$ , we consider the frogs there woken but immediately frozen.) Let  $\mathcal{F}_t$  be the  $\sigma$ -algebra representing the information revealed after  $t$  steps of this process. Let  $\xi_t$  be a point process on  $\mathbb{T}_d^{h*}$  made up of the locations of each woken frog after  $t$  steps. For any  $v \in \xi_t$ , let  $L(v)$  denote the level of  $v$  in the tree, and define

$$w_\theta(\xi_t) = \sum_{v \in \xi_t} \theta^{-L(v)}.$$

If  $\mu \leq (d - 1)^2/4d$ , then there exist positive real numbers  $\theta_0$  and  $\theta_1$  satisfying

$$\theta_0 \leq 1 + \frac{2d}{d - 1}\mu, \tag{17}$$

$$\theta_1 \geq d - \frac{2d}{d - 1}\mu \tag{18}$$

such that  $w_{\theta_0}(\xi_t)$  and  $w_{\theta_1}(\xi_t)$  are supermartingales with respect to the filtration  $\mathcal{F}_t$ .

*Proof.* Observe that

$$\mathbf{E}w_\theta(\xi_1) = \left( \frac{1}{d + 1}\theta + \frac{(1 + \mu)d}{d + 1}\theta^{-1} \right) w_\theta(\xi_0).$$

Solving a quadratic equation, we see that  $\mathbf{E}w_\theta(\xi_1) = w_\theta(\xi_0)$  if

$$\theta = \frac{d + 1 \pm \sqrt{(d + 1)^2 - 4(1 + \mu)d}}{2}.$$

Let  $\theta_0$  and  $\theta_1$  be the smaller and larger of these solutions, respectively, which are positive real numbers if  $0 \leq \mu \leq (d - 1)^2/4d$ . Let  $\mathcal{F}_t$  be the  $\sigma$ -algebra generated by the frog model up to time  $t$ . Now, suppose that  $\theta = \theta_0$  or  $\theta = \theta_1$ , and we will show that  $w_\theta(\xi_t)$  is a supermartingale. Consider a nonfrozen frog in  $\xi_t$  at level  $i$ . It jumps backward with probability  $1/(d + 1)$ , waking no frogs, and forward with probability  $d/(d + 1)$ , possibly waking a  $\text{Ber}(\mu)$ -distributed number

of frogs. Thus, its expected contribution to  $w_\theta(\xi_{t+1})$  is at most

$$\frac{1}{d+1}\theta^{-i+1} + \frac{(1+\mu)d}{d+1}\theta^{-i-1} = \theta^{-i},$$

exactly its current contribution. The contribution to  $w_\theta(\xi_{t+1})$  of each frozen frog in  $\xi_t$  is the same as its contribution to  $w_\theta(\xi_t)$ , showing that

$$\mathbf{E}[w_\theta(\xi_{t+1}) \mid \mathcal{F}_t] \leq w_\theta(\xi_t).$$

To prove (17) and (18), observe that  $\sqrt{(d+1)^2 - 4(1+\mu)d}$  is a concave function of  $\mu$ . It therefore lies above its secant line from 0 to  $(d-1)^2/4d$ , yielding

$$\sqrt{(d+1)^2 - 4(1+\mu)d} \geq d-1 - \frac{4d}{d-1}\mu.$$

Applying this to the definitions of  $\theta_0$  and  $\theta_1$  gives the desired bounds.  $\square$

**LEMMA 4.8 (BRW bound, starting at root).** *Consider  $\mathbb{T}_d^{h*}$  with one initially active frog at the root, no frogs at  $y$ , and i.i.d.- $\text{Ber}(\mu)$  sleeping frogs at the other vertices. Run the frog model with frogs frozen at  $y$  and  $\mathcal{L}_h$ . Let  $X$  and  $N$  be the number of particles eventually frozen at  $y$  and  $\mathcal{L}_h$ , respectively. (The random variable  $N$  includes in its count the frogs that are woken at  $\mathcal{L}_h$  and immediately frozen.) If  $\mu \leq (d-1)^2/4d$ , then*

$$\mathbf{E}N \leq \left(1 + \frac{2d}{d-1}\mu\right)^h, \quad \text{and} \quad \mathbf{E}X \leq \left(d\left(1 - \frac{2\mu}{d-1}\right)\right)^{-1}.$$

*Proof.* Let  $T$  be the first time when all frogs are frozen. By Lemma 4.7, the process  $w_\theta(\xi_t)$  is a supermartingale for  $\theta = \theta_0, \theta_1$ . It is bounded at all times by  $\theta \lceil \mu \rceil |\mathbb{T}_d^{h*}|$ , since the total number of frogs in the system is at most  $\lceil \mu \rceil |\mathbb{T}_d^{h*}|$  and no frog goes below level  $-1$ . Hence, the optional stopping theorem applies and shows that  $\mathbf{E}w_\theta(\xi_T) \leq 1$ . The expected contribution to  $w_{\theta_0}(\xi_T)$  by frogs frozen at  $\mathcal{L}_h$  is

$$\theta_0^{-h} \mathbf{E}N \leq \mathbf{E}w_{\theta_0}(\xi_T) \leq 1.$$

Then (17) gives the bound on  $\mathbf{E}N$ . Similarly, the expected contribution to  $w_{\theta_1}(\xi_T)$  by frogs frozen at  $y$  is

$$\theta_1 \mathbf{E}X \leq \mathbf{E}w_{\theta_1}(\xi_T) \leq 1,$$

and (18) gives us the bound on  $\mathbf{E}X$ .  $\square$

The previous lemma bounds the expected number of frogs at  $\mathcal{L}_h$  and at  $y$  when we have an initially active frog at the root. The next lemma makes similar bounds when the initially active frog is at  $\mathcal{L}_{h-1}$ .

**LEMMA 4.9 (BRW bound, starting at level  $h - 1$ ).** *Consider the frog model on  $\mathbb{T}_d^{h*}$  with one initially active frog at some vertex  $v_0 \in \mathcal{L}_{h-1}$ , no frogs at ancestors of  $v_0$  (including  $y$ ), and i.i.d.- $\text{Ber}(\mu)$  sleeping frogs elsewhere. Run the frog model with frogs frozen at  $y$  and  $\mathcal{L}_h$ . Let  $X$  and  $N$  be the number of particles eventually frozen at  $y$  and  $\mathcal{L}_h$ , respectively. (Again, the frogs woken at  $\mathcal{L}_h$  and immediately frozen are included in the count  $N$ .) If  $\mu \leq (d - 1)^2/4d$ , then*

$$\mathbf{E}N \leq 1 + \frac{2d}{d - 1}\mu, \quad \text{and} \quad \mathbf{E}X \leq \left( d \left( 1 - \frac{2\mu}{d - 1} \right) \right)^{-h}.$$

*Proof.* This has the same proof as Lemma 4.8 except that the initial value of the supermartingale  $w_\theta(\xi_t)$  is  $\theta^{-h+1}$  rather than 1. We then have

$$\begin{aligned} \theta_0^{-h} \mathbf{E}N &\leq \theta_0^{-h+1}, \\ \theta_1 \mathbf{E}X &\leq \theta_1^{-h+1}, \end{aligned}$$

and (17) and (18) from Lemma 4.7 give the bounds on  $\mathbf{E}N$  and  $\mathbf{E}X$ . □

**4.3. Estimates on tagged frogs.** Again, fix  $j, H$ , and  $h$ , and consider the frog model on  $\mathbb{T}_d^{(H+h)*}$  with the system of tags given previously. Recall that all frogs lose their tags at time  $2^{j+1} + 1$ , and so all of the following random variables count frogs only up to time  $2^{j+1}$ . See Figure 4.

- For  $\ell \in \{A, B_0, B_1, \dots, C_0, C_1, \dots\}$ , let  $X_\ell$  be the number of  $\ell$ -tagged frogs eventually frozen at  $y$ .
- For  $\ell = B_i, i \geq 0$ , or  $\ell = C_i, i \geq 1$ , let  $N_\ell$  be the number of frogs that received an  $\ell$  tag at  $\mathcal{L}_h$ . These are the frogs that move from  $\mathcal{L}_{h-1}$  to  $\mathcal{L}_h$  and change their tags to  $\ell$ , as well as the frogs sleeping at  $\mathcal{L}_h$  woken by them.
- For  $\ell = B_i, i \geq 0$ , or  $\ell = C_i, i \geq 1$ , let  $M_\ell$  be the number of  $\ell$ -tagged frogs that move from  $\mathcal{L}_h$  back to  $\mathcal{L}_{h-1}$ , not counting  $B_i$ -tagged frogs that do so at times  $2^j$  and on.
- Let  $M_{C_0}$  be the number of  $B_i$ -tagged frogs for any  $i$  that move from  $\mathcal{L}_h$  to  $\mathcal{L}_{h-1}$  at time  $2^j$  or later. These are the frogs that change tags to  $C_0$ .

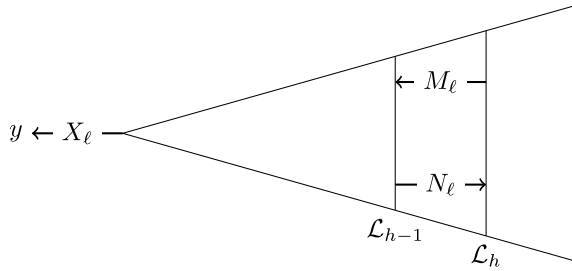


Figure 4.  $X_\ell$  counts the number of  $\ell$ -tagged frogs frozen at  $y$ ;  $N_\ell$  counts the number of frogs that move from  $\mathcal{L}_{h-1}$  to  $\mathcal{L}_h$  and are retagged as  $\ell$ , plus the  $\ell$ -tagged frogs awakened by these frogs at  $\mathcal{L}_h$ ; and  $M_\ell$  counts how many  $\ell$ -tagged frogs move from  $\mathcal{L}_h$  to  $\mathcal{L}_{h-1}$ . This decomposes the total number of visits to  $y$  as in (19).

Recall that the total number of frogs frozen at  $y$  by time  $2^{j+1}$  is  $X^{(j+1, H+h)}$ . We have just decomposed this quantity as

$$X^{(j+1, H+h)} = X_A + \sum_{i=0}^{\infty} (X_{B_i} + X_{C_i}). \quad (19)$$

Our eventual goal is bound this in expectation under the assumption that (13) holds for  $X^{(j, H)}$ , thus advancing the induction by a step.

LEMMA 4.10. *If  $\mu \leq (d-1)^2/4d$ , then*

$$\mathbf{E}N_{B_0} \leq \left(1 + \frac{2d}{d-1}\mu\right)^h, \quad \text{and} \quad \mathbf{E}X_A \leq \left(d\left(1 - \frac{2\mu}{d-1}\right)\right)^{-1}.$$

*Proof.* This follows immediately from Lemma 4.8. □

LEMMA 4.11. *For any  $i \geq 0$ , suppose either that  $\ell = B_i$  and  $\ell^+ = B_{i+1}$ , or that  $\ell = C_i$  and  $\ell^+ = C_{i+1}$ . If  $\mu \leq (d-1)^2/4d$ , then*

$$\mathbf{E}N_{\ell^+} \leq \left(1 + \frac{2d}{d-1}\mu\right)\mathbf{E}M_\ell, \quad \text{and} \quad \mathbf{E}X_\ell \leq \left(d\left(1 - \frac{2\mu}{d-1}\right)\right)^{-h}\mathbf{E}M_\ell.$$

*Proof.* Enumerate the  $\ell$ -tagged frogs that return from  $\mathcal{L}_h$  to  $\mathcal{L}_{h-1}$  as frogs  $1, \dots, M_\ell$ . Let  $v_1, \dots, v_{M_\ell} \in \mathcal{L}_h$  be the vertices that the frogs emerge from. For each  $1 \leq k \leq M_\ell$ , we define random variables  $N(k)$  and  $X(k)$  that give the portions of

$N_{\ell^+}$  and  $X_\ell$  that are attributable to frog  $k$ , in a sense that we will explain. We then estimate  $N(k)$  and  $X(k)$  using the branching random walk bounds.

To define  $N(k)$  and  $X(k)$ , consider the following modified frog model on  $\mathbb{T}_d^{(H+h)*}$ :

- at all vertices at levels 1 to  $h$  except for the ones on the path from the root to  $v_k$ , place the same sleeping frogs as in the current realization of the original frog model on  $\mathbb{T}_d^{(H+h)*}$ ;
- place an initially active frog at  $v_k$  that follows the path of frog  $k$  starting from when it moves from  $v_k$  back to  $\mathcal{L}_{h-1}$ ;
- freeze all frogs on arrival at  $\mathcal{L}_h$  and at  $y$ .

We define  $N(k)$  and  $X(k)$  as the number of frogs eventually frozen at  $\mathcal{L}_h$  and  $y$ , respectively, in this frog model. As usual, we include the frogs woken at  $\mathcal{L}_h$  and immediately frozen in the count of  $N(k)$ . We claim that  $N_{\ell^+} \leq \sum_{k=1}^{M_\ell} N(k)$ . This is because any  $\ell$ -tagged frog counted by  $N_{\ell^+}$  either is one of frogs  $1, \dots, k$  or is spawned by a sequence of frogs at levels  $1, \dots, n$  originating with one of frogs  $1, \dots, k$ . Hence, any frog counted by  $N_{\ell^+}$  must also be counted by  $X(k)$  for at least one  $k \in \{1, \dots, M_\ell\}$ . By the same argument,  $X_\ell \leq \sum_{k=1}^{M_\ell} X(k)$ .

The conditional distributions given  $M_\ell$  of  $N(k)$  and  $X(k)$ , respectively, are exactly those of  $N$  and  $X$  from Lemma 4.9. Applying this lemma,

$$\mathbf{E}[N_{\ell^+} \mid M_\ell] \leq \left(1 + \frac{2d}{d-1}\mu\right)M_\ell,$$

$$\mathbf{E}[X_\ell \mid M_\ell] \leq \left(d\left(1 - \frac{2\mu}{d-1}\right)\right)^{-h} M_\ell.$$

Now take expectations to complete the proof. □

LEMMA 4.12. *Suppose that (13) holds for the fixed  $j$  and  $H$  used in the definitions of  $X_\ell$ ,  $N_\ell$ , and  $M_\ell$ . If  $\mu \leq (d-1)^2/4d$ , then for  $\ell = B_i, i \geq 0$ , or  $\ell = C_i, i \geq 1$ ,*

$$\mathbf{E}M_\ell \leq \frac{.8\mathbf{E}N_\ell}{1 + \frac{2d}{d-1}\mu}.$$

*Proof.* We claim that

$$\mathbf{E}[M_\ell \mid N_\ell] \leq \frac{.8N_\ell}{1 + \frac{2d}{d-1}\mu}, \tag{20}$$



from which the lemma follows by taking expectations. Roughly speaking, we want to show that for each frog acquiring an  $\ell$  tag at  $v \in \mathcal{L}_h$ , the expected number of  $\ell$ -tagged frogs moving from  $v$  back to  $\mathcal{L}_{h-1}$  is at most  $.8/(1 + 2d\mu/(d - 1))$ . This follows from (13), as we will now show in detail.

Enumerate the frogs counted by  $N_\ell$  as frogs  $1, \dots, N_\ell$ . Recall that these include both the frogs that move from  $\mathcal{L}_{h-1}$  to  $\mathcal{L}_h$  and receive an  $\ell$  tag, as well as the frogs woken at  $\mathcal{L}_h$  by them. For  $1 \leq k \leq N_\ell$ , let  $v_k \in \mathcal{L}_h$  be the vertex where frog  $k$  received its  $\ell$  tag. Note that the same vertices will appear multiple times in  $v_1, \dots, v_{N_\ell}$ , though all frogs in the list are unique. For each  $k$ , we will define a random variable  $M(k)$  that gives the number of frogs counted by  $M_\ell$  attributable to frog  $k$ . As we did in the previous lemma, we then bound  $M(k)$ , this time using (13).

To define  $M(k)$ , let  $y_k$  be the parent of  $v_k$ , and consider the following frog model on  $\{y_k\} \cup \mathbb{T}_d^{H+h}(v_k)$ :

- at all descendants of  $v_k$ , place the same sleeping frogs as in the current realization of the original frog model on  $\mathbb{T}_d^{(H+h)*}$ ;
- place an initially active frog at  $v_k$  that follows that path of frog  $k$  starting from its arrival at  $v_k$ ;
- freeze all frogs on visiting  $y_k$ .

We then define  $M(k)$  as the number of frogs frozen at  $y_k$  in the first  $2^j$  steps of this frog model. We claim that  $M_\ell \leq \sum_{i=1}^{N_\ell} M(k)$ . To justify this, we first observe that any return from  $v_k$  to  $y_k$  counted by  $M_\ell$  must occur within  $2^j$  steps of when  $v_k$  is first visited by a frog that changes its label to  $\ell$ . When  $\ell = B_i$ , this is because  $M_\ell$  only counts returns up to time  $2^j$ . When  $\ell = C_i$ , it is because the first visit to  $v_k$  by a frog receiving an  $\ell$  tag occurs after time  $2^j$ , and  $M_\ell$  only counts returns up to time  $2^{j+1}$ . Now, any  $\ell$ -tagged frog counted by  $M_\ell$  is either one of frogs  $1, \dots, N_\ell$  or is spawned by a sequence of frogs at level  $h + 1$  and beyond in  $\mathbb{T}_d^{(H+h)*}$  originating with one of these frogs. It is thus counted by  $M(k)$  for some  $1 \leq k \leq N_\ell$ .

Observe that the frog model defining  $M(k)$  is just a disguised version of the frog model on  $\mathbb{T}_d^{H*}$  considered in Proposition 4.3. Hence, the distribution of  $M(k)$  conditional on  $N_\ell$  is exactly that of  $X^{(j,H)}$ . Applying (13), we have

$$\mathbf{E}[M(k) \mid N_\ell] \leq \frac{.8}{1 + \frac{2d}{d-1}\mu}.$$

Summing this over all  $k$  to bound  $\mathbf{E}[M_\ell \mid N_\ell]$  proves (20). □

Lemma 4.11 gives bounds on  $\mathbf{E}N_{B_{i+1}}$  and  $\mathbf{E}N_{C_{i+1}}$  in terms of  $\mathbf{E}M_{B_i}$  and  $\mathbf{E}M_{C_i}$ , and Lemma 4.12 gives bounds on  $\mathbf{E}M_{B_i}$  and  $\mathbf{E}M_{C_i}$  in terms of  $\mathbf{E}N_{B_i}$

and  $\mathbf{E}N_{C_i}$ . Together, these bounds show that  $\mathbf{E}N_{B_i}$ ,  $\mathbf{E}N_{C_i}$ ,  $\mathbf{E}M_{B_i}$ , and  $\mathbf{E}N_{C_i}$  decay exponentially in  $i$ .

LEMMA 4.13. For any  $i \geq 0$ , suppose either that  $\ell = B_i$  and  $\ell^+ = B_{i+1}$ , or that  $\ell = C_i$  and  $\ell^+ = C_{i+1}$ . If  $\mu \leq (d - 1)^2/4d$ , then

$$\begin{aligned} \mathbf{E}M_{\ell^+} &\leq .8\mathbf{E}M_{\ell}, \\ \mathbf{E}N_{\ell^+} &\leq .8\mathbf{E}N_{\ell}. \end{aligned}$$

Consequently,

$$\begin{aligned} \sum_{i=0}^{\infty} \mathbf{E}M_{B_i} &\leq 5\mathbf{E}M_{B_0}, & \sum_{i=0}^{\infty} \mathbf{E}M_{C_i} &\leq 5\mathbf{E}M_{C_0}, \\ \sum_{i=0}^{\infty} \mathbf{E}N_{B_i} &\leq 5\mathbf{E}N_{B_0}, & \text{and} & \sum_{i=0}^{\infty} \mathbf{E}N_{C_i} &\leq 5\mathbf{E}N_{C_0}. \end{aligned}$$

*Proof.* The bounds on  $\mathbf{E}M_{\ell^+}$  and  $\mathbf{E}N_{\ell^+}$  follow immediately from Lemmas 4.11 and 4.12. The other bounds are consequences of summing geometric series.  $\square$

LEMMA 4.14.

$$\mathbf{E}M_{C_0} \leq c_1 \mu 2^{j+1} \sum_{i=0}^{\infty} \mathbf{E}N_{B_i}.$$

*Proof.* This proof is just as for Lemmas 4.11 and 4.12, except we use the all-awake bound in place of the branching random walk bounds or the inductive hypothesis. In more detail, fix a nonnegative integer  $i$ , and number the frogs that received a  $B_i$  tag at  $\mathcal{L}_h$  as  $1, \dots, N_{B_i}$ . These are made up of the  $B_{i-1}$ -tagged frogs that moved from  $\mathcal{L}_{h-1}$  to  $\mathcal{L}_h$  (where  $B_{-1} = A$ ), as well as the frogs at  $\mathcal{L}_h$  that these frogs woke up. Let  $v_1, \dots, v_{N_{B_i}} \in \mathcal{L}_h$  be the sites where these frogs get their  $B_i$  tags. In a similar argument as we used in Lemmas 4.11 and 4.12, we define a random variable  $M(k)$  giving the number of frogs counted by  $M_{C_0}$  that are attributable to frog  $k$ . Let  $y_k$  be the parent of  $v_k$ , and define a frog model on  $\{y_k\} \cup \mathbb{T}_d^{H+h}(v_k)$  exactly as in Lemma 4.12. Define  $M(k)$  as the number of frogs frozen at  $y_k$  in the first  $2^{j+1}$  steps of this frog model. Let  $M_{C_0}^i$  be the number of  $B_i$ -tagged frogs in the original model that move from  $\mathcal{L}_h$  to  $\mathcal{L}_{h-1}$  between times  $2^j + 1$  and  $2^{j+1}$  and change tags to  $C_0$ . By similar reasoning as in the previous lemmas, we have  $M_{C_0}^i \leq \sum_{k=1}^{N_{B_i}} M(k)$ .

Conditional on  $N_{B_i}$ , the distribution of  $M(k)$  is stochastically dominated by the random variable  $W$  from Lemma 4.6. Applying the bound from Lemma 4.6 and

summing over all  $k$ , we get

$$\mathbf{E}[M_{C_0}^i \mid N_{B_i}] \leq c_1 N_{B_i} \mu 2^{j+1}.$$

Therefore,

$$\mathbf{E}M_{C_0} = \sum_{i=0}^{\infty} \mathbf{E}M_{C_0}^i \leq c_1 \mu 2^{j+1} \sum_{i=0}^{\infty} \mathbf{E}N_{B_i}. \quad \square$$

We are now ready to advance the induction in Proposition 4.3.

**PROPOSITION 4.15.** *There exists an absolute constant  $C$  so that the following statement holds. Suppose that for some specific choice of  $j$ ,  $H$ ,  $d$ , and  $\mu$  with  $j, H \geq 1$ ,  $d \geq 2$ , and  $\mu \leq d/100$ , the inductive hypothesis (13) holds. Then, for any*

$$h \geq \frac{C(j + \log(1 + \mu))}{\log\left(\frac{d}{1+4\mu}\right)}, \quad (21)$$

we have

$$\mathbf{E}X^{(j+1, H+h)} \leq \frac{.8}{1 + \frac{2d}{d-1}\mu}.$$

*Proof.* Using the decomposition of  $X^{(j+1, H+h)}$  given in (19), our goal is to show that

$$\mathbf{E}\left[X_A + \sum_{i=0}^{\infty} X_{B_i} + \sum_{i=0}^{\infty} X_{C_i}\right] \leq \frac{.8}{1 + \frac{2d}{d-1}\mu}.$$

In successive lines, we apply Lemmas 4.11, 4.13, 4.12 and 4.10 to obtain

$$\begin{aligned} \sum_{i=0}^{\infty} \mathbf{E}X_{B_i} &\leq \left(d\left(1 - \frac{2\mu}{d-1}\right)\right)^{-h} \sum_{i=0}^{\infty} \mathbf{E}M_{B_i} \\ &\leq 5\left(d\left(1 - \frac{2\mu}{d-1}\right)\right)^{-h} \mathbf{E}M_{B_0} \\ &\leq 5\left(d\left(1 - \frac{2\mu}{d-1}\right)\right)^{-h} \frac{.8\mathbf{E}N_{B_0}}{1 + \frac{2d}{d-1}\mu} \\ &\leq 5\left(\frac{1 + \frac{2d}{d-1}\mu}{d\left(1 - \frac{2\mu}{d-1}\right)}\right)^h \left(\frac{.8}{1 + \frac{2d}{d-1}\mu}\right) \leq \left(\frac{1 + \frac{2d}{d-1}\mu}{d\left(1 - \frac{2\mu}{d-1}\right)}\right)^h \left(\frac{4}{1 + 2\mu}\right). \end{aligned}$$

Applying Lemmas 4.11, 4.13 and 4.14,

$$\begin{aligned} \sum_{i=0}^{\infty} \mathbf{E}X_{C_i} &\leq \left( d \left( 1 - \frac{2\mu}{d-1} \right) \right)^{-h} \sum_{i=0}^{\infty} \mathbf{E}M_{C_i} \\ &\leq 5 \left( d \left( 1 - \frac{2\mu}{d-1} \right) \right)^{-h} \mathbf{E}M_{C_0} \\ &\leq 5 \left( d \left( 1 - \frac{2\mu}{d-1} \right) \right)^{-h} c_1 \mu 2^{j+1} \sum_{i=0}^{\infty} \mathbf{E}N_{B_i}, \end{aligned}$$

and then applying Lemmas 4.13 and 4.10,

$$\begin{aligned} \sum_{i=0}^{\infty} \mathbf{E}X_{C_i} &\leq 25 \left( d \left( 1 - \frac{2\mu}{d-1} \right) \right)^{-h} c_1 \mu 2^{j+1} \mathbf{E}N_{B_0} \\ &\leq 25 \left( \frac{1 + \frac{2d}{d-1} \mu}{d \left( 1 - \frac{2\mu}{d-1} \right)} \right)^h c_1 \mu 2^{j+1}. \end{aligned}$$

Applying these bounds together with the estimate on  $\mathbf{E}X_A$  from Lemma 4.10,

$$\begin{aligned} \mathbf{E} \left[ X_A + \sum_{i=0}^{\infty} X_{B_i} + \sum_{i=0}^{\infty} X_{C_i} \right] &\leq \frac{1}{d \left( 1 - \frac{2\mu}{d-1} \right)} + \left( \frac{1 + \frac{2d}{d-1} \mu}{d \left( 1 - \frac{2\mu}{d-1} \right)} \right)^h \left( \frac{4}{1 + 2\mu} + 25c_1 \mu 2^{j+1} \right) \\ &\leq \frac{1}{.96d} + \left( \frac{1 + 4\mu}{.96d} \right)^h (4 + 50c_1 \mu 2^j). \end{aligned}$$

From (21),

$$\begin{aligned} \left( \frac{1 + 4\mu}{.96d} \right)^h &\leq \exp \left[ \left( -\log \left( \frac{d}{1 + 4\mu} \right) + \log \left( \frac{25}{24} \right) \right) \frac{C(j + \log(1 + \mu))}{\log \left( \frac{d}{1 + 4\mu} \right)} \right] \\ &\leq \exp[-.9C(j + \log(1 + \mu))] = e^{-.9Cj} (1 + \mu)^{-.9C}. \end{aligned}$$

A bit of asymptotic analysis now shows that we can choose  $C$  large enough that for all  $\mu \geq 0$  and  $j \geq 1$ ,

$$\mathbf{E} \left[ X_A + \sum_{i=0}^{\infty} X_{B_i} + \sum_{i=0}^{\infty} X_{C_i} \right] \leq \frac{1}{.96d} + \frac{.2}{1 + 4\mu},$$

and

$$\frac{1}{.96d} + \frac{.2}{1 + 4\mu} = \frac{1 + 4\mu}{.96d} + \frac{.2}{1 + 4\mu} \leq \frac{1 + .04d}{.96d} + \frac{.2}{1 + 4\mu} \leq \frac{.8}{1 + 4\mu} \leq \frac{.8}{1 + \frac{2d}{d-1} \mu}. \quad \square$$

*Proof of Proposition 4.3.* We start by establishing (13) when  $j = 1$  and  $n \geq 1$ . Consider the frog model on  $\mathbb{T}_d^{1*}$ . Fix any  $n \geq 1$ . If the initial frog moves immediately to  $y$ , then  $X^{(1,n)} = 1$ . If instead it moves to a child of the root, then no frogs can make it to  $y$  by time 2, and  $X^{(1,n)} = 0$ . Hence,  $\mathbf{E}X^{(1,n)} = 1/(d + 1)$ , and this is easily seen to be less than the right-hand side of (13) using our assumption that  $\mu \leq d/100$ .

Applying Proposition 4.15 inductively, (13) holds for  $j \geq 2$  so long as we can show that

$$H_j \geq 1 + \sum_{i=2}^j \left\lceil \frac{C(i + \log(1 + \mu))}{\log\left(\frac{d}{1+4\mu}\right)} \right\rceil,$$

where  $C$  is the constant from Proposition 4.15. Indeed, it is straightforward to compute that

$$1 + \sum_{i=2}^j \left\lceil \frac{C(i + \log(1 + \mu))}{\log\left(\frac{d}{1+4\mu}\right)} \right\rceil = O\left(\frac{j^2 + j \log(1 + \mu)}{\log\left(\frac{d}{1+4\mu}\right)}\right). \quad \square$$

### Acknowledgements

C.H. received support from NSF grant DMS-1308645 and from a Simons Fellowship. T.J. received support from NSF grants DMS-1401479 and DMS-1811952 and PSC-CUNY Award #61540-00 49.

### Appendix A. Stochastic comparison results for the frog model

In this section, we outline the results of [JJ18], which allow us to compare two frog models on the same graph with different initial conditions. If the distribution of frog counts in the first model stochastically dominates the distribution in the other, then certain statistics of the first model will dominate the corresponding statistics in the other. This is a trivial fact with the typical definition of stochastic domination. The strength of these results is that they apply to less conventional stochastic orders, one of which is the *probability generating function order*, whose name we abbreviate to *pgf order*.

For two probability measures  $\pi_1$  and  $\pi_2$  on the nonnegative real numbers, we say that  $\pi_1$  is smaller than  $\pi_2$  in the pgf order, denoted  $\pi_1 \leq_{\text{pgf}} \pi_2$ , if for  $X \sim \pi_1$  and  $Y \sim \pi_2$  and all  $t \in (0, 1)$ , it holds that  $\mathbf{E}t^X \geq \mathbf{E}t^Y$ . We also write  $X \leq_{\text{pgf}} Y$  to mean that the law of  $X$  is stochastically smaller in the pgf order than the law of  $Y$ , and we also use mixed expressions like  $X \leq_{\text{pgf}} \pi_2$  in the obvious way.

See the introduction of [JJ18] for more on the pgf order and its relations to other stochastic orders.

We now present the result from [JJ18] as we will apply it in this paper.

**LEMMA A.1.** *Consider two frog models on  $\mathbb{T}_d^n$  with initial frog counts given by  $(\eta(v))_v$  and  $(\eta'(v))_v$  for  $v \in \mathbb{T}_d^n \setminus \{\emptyset\}$ . Assume that both counts are independent. Suppose that  $\eta(v) \preceq_{\text{pgf}} \eta'(v)$  for all  $v$ . Let  $N$  and  $N'$  be the number of leaves visited in the two models by some given time, and let  $R$  and  $R'$  be the number of visits to the root in the two models by some given time. Then  $N \preceq_{\text{pgf}} N'$  and  $R \preceq_{\text{pgf}} R'$ .*

*Proof.* By [JJ18, Theorem 3], this holds once we prove that the number of leaves visited by time  $t$  and the number of visits to the root by time  $t$  are *continuous pgf statistics*. For the first statistic, this is a very slight variation of [JJ18, Proposition 21] and has a nearly identical proof. For the second statistic, it is a consequence of [JJ18, Proposition 4], if we think of the frog models as having frog paths stopped at time  $t$ .  $\square$

We mention two basic facts about the pgf order. First, if  $X \preceq_{\text{pgf}} Y$ , then  $\mathbf{E}X \leq \mathbf{E}Y$ . Second, for any distribution  $\pi$  with expectation  $\mu$  or less, we have  $\pi \preceq_{\text{pgf}} \text{Ber}(\mu)$ , recalling our definition of  $\text{Ber}(\mu)$  for  $\mu > 1$  as the unique distribution on  $\{\lfloor \mu \rfloor, \lceil \mu \rceil\}$  with mean  $\mu$ . This fact is proven in [JJ18, Proposition 15(b)] for a different stochastic relation known as the *increasing concave order*, and it follows that it holds for the pgf order since domination in the increasing concave order implies domination in the pgf order (see [JJ18, Section 2]).

## Appendix B. Miscellaneous concentration inequalities

The following two bounds appear verbatim in [HJJ17a, Appendix C]. Both are standard results that follow from bounding the moment generating function and applying Markov's inequality.

**PROPOSITION B.1.** *Let  $\mathbf{E}Y = \lambda$ , and suppose either that  $Y$  is Poisson or that  $Y$  is a sum of independent random variables supported on  $[0, 1]$ . For any  $0 < \alpha < 1$ ,*

$$\mathbf{P}[Y \leq \alpha\lambda] \leq \exp\left(-\frac{(1-\alpha)^2\lambda}{2}\right),$$

and for any  $\alpha > 1$ ,

$$\mathbf{P}[Y \geq \alpha\lambda] \leq \exp\left(-\frac{(\alpha-1)\lambda}{\frac{2}{3} + \frac{2}{\alpha-1}}\right).$$

**PROPOSITION B.2.** *Let  $(X_i)_{i=1}^n$  be a collection of independent random variables satisfying*

$$\mathbf{P}[X_i \geq \ell] \leq Ce^{-b\ell}$$

*for some  $C$  and  $b > 0$  and all  $\ell \geq 1$ . Then for any  $b' > 0$ , there exists  $C'$  depending on  $C$ ,  $b$ , and  $b'$  such that*

$$\mathbf{P}\left[\sum_{i=1}^n X_i \geq C'n\right] \leq e^{-b'n}.$$

*We can take  $C' = 2(b' + C)/b$ .*

Next, we give a more refined version of the previous proposition that applies to random variables that are exactly geometrically distributed.

**PROPOSITION B.3.** *Let  $(G_i)_{i \geq 1}$  be a collection of independent random variables with  $G_i$  geometrically distributed on  $\{1, 2, \dots\}$  with parameter  $p$ . Let  $\mu := \mathbf{E}G_i = 1/p$ . For any  $\lambda \geq 2$ ,*

$$\mathbf{P}[G_1 + \dots + G_n \geq \lambda n \mu] \leq \exp\left[-n\left(\frac{\lambda}{2} - 1\right)\right].$$

*Proof.* If  $G_1 + \dots + G_n \geq k$ , then in the first of  $k$  independent trials with success probability  $p$ , there were at most  $n$  successes. Thus, by Proposition B.1,

$$\mathbf{P}[G_1 + \dots + G_n \geq k] = \mathbf{P}[\text{Bin}(k, p) \leq n] \leq \exp\left(-\frac{(1 - n/kp)^2 kp}{2}\right).$$

Substituting  $k = \lambda n \mu$  and  $p = 1/\mu$  gives

$$\mathbf{P}[G_1 + \dots + G_n \geq k] \leq \exp\left(-\frac{(\lambda - 2 + \lambda^{-1})n}{2}\right) \leq \exp\left(-\frac{(\lambda - 2)n}{2}\right). \quad \square$$

Last, we extend the exponential concentration bound for a single Poisson or binomial random variable to an entire sequence, via a union bound:

**LEMMA B.4.** *Let  $\gamma_1 \geq 2\gamma_2 > 0$ , and let  $\gamma_1 \geq 8$ . Suppose that  $X_i$  is Poisson or binomial with mean  $\gamma_1 i$  for all  $i \geq k$ , with no assumption on the joint distribution of  $(X_i)_{i \geq k}$ . Then*

$$\mathbf{P}[X_i < \gamma_2 i \text{ for some } i \geq k] \leq 2 \exp\left(-\frac{(1 - \gamma_2/\gamma_1)^2 \gamma_1 k}{2}\right)$$

*Proof.* By Proposition B.1,

$$\mathbf{P}[X_i < \gamma_2 i] \leq \exp\left(-\frac{(1 - \gamma_2/\gamma_1)^2 \gamma_1 i}{2}\right).$$

Applying a union bound over all  $i \geq k$  and summing the geometric series gives

$$\mathbf{P}[X_i < \gamma_2 i \text{ for some } i \geq k] \leq \frac{\exp\left(-\frac{(1 - \gamma_2/\gamma_1)^2 \gamma_1 k}{2}\right)}{1 - \exp\left(-\frac{(1 - \gamma_2/\gamma_1)^2 \gamma_1}{2}\right)},$$

and

$$1 - \exp\left(-\frac{(1 - \gamma_2/\gamma_1)^2 \gamma_1}{2}\right) \geq 1 - \exp\left(-\frac{(1 - 1/2)^2 (8)}{2}\right) \geq \frac{1}{2}. \quad \square$$

**Conflict of Interest:** None.

## References

- [Ald91] D. J. Aldous, ‘Random walk covering of some special trees’, *J. Math. Anal. Appl.* **157**(1) (1991), 271–283.
- [AMPR01] O. S. M. Alves, F. P. Machado, S. Yu. Popov and K. Ravishankar, ‘The shape theorem for the frog model with random initial configuration’, *Markov Process. Related Fields* **7**(4) (2001), 525–539.
- [BFHM18] I. Benjamini, L. R. Fontes, J. Hermon and F. P. Machado, ‘On an epidemic model on finite graphs’, Preprint, 2018, [arXiv:1610.04301](https://arxiv.org/abs/1610.04301).
- [BH16] I. Benjamini and J. Hermon, ‘Rapid social connectivity’, Preprint, 2016, [arXiv:1608.07621](https://arxiv.org/abs/1608.07621).
- [DMnVZ00] R. Dickman, M. A. Muñoz, A. Vespignani and S. Zapperi, ‘Paths to self-organized criticality’, *Braz. J. Phys.* **30** (2000), 27–41. (en).
- [Her18] J. Hermon, ‘Frogs on trees?’, *Electron. J. Probab.* **23** (2018), Paper No. 17, 40.
- [HMQS16] J. Hermon, B. Morris, C. Qin and A. Sly, ‘The social network model on infinite graphs’, Preprint, 2016, [arXiv:1610.04293](https://arxiv.org/abs/1610.04293).
- [HJJ16] C. Hoffman, T. Johnson and M. Junge, ‘From transience to recurrence with Poisson tree frogs’, *Ann. Appl. Probab.* **26**(3) (2016), 1620–1635.
- [HJJ17a] C. Hoffman, T. Johnson and M. Junge, ‘Infection spread for the frog model on trees’, *Electron. J. Probab.* **24**(112) (2019), 1–29.
- [HJJ17b] C. Hoffman, T. Johnson and M. Junge, ‘Recurrence and transience for the frog model on trees’, *Ann. Probab.* **45**(5) (2017), 2826–2854.
- [JJ16] T. Johnson and M. Junge, ‘The critical density for the frog model is the degree of the tree’, *Electron. Commun. Probab.* **21** (2016), Paper No. 82, 12.
- [JJ18] T. Johnson and M. Junge, ‘Stochastic orders and the frog model’, *Ann. Inst. Henri Poincaré Probab. Stat.* **54**(2) (2018), 1013–1030.
- [KZ17] E. Kosygina and M. P. W. Zerner, ‘A zero-one law for recurrence and transience of frog processes’, *Probab. Theory Related Fields* **168**(1–2) (2017), 317–346.



- [LP16] R. Lyons and Y. Peres, *Probability on Trees and Networks*, Cambridge Series in Statistical and Probabilistic Mathematics, 42 (Cambridge University Press, New York, 2016).
- [RSZ17] L. T. Rolla, V. Sidoravicius and O. Zeitouni, ‘Critical density for activated random walks’, Preprint, 2017, [arXiv:1707.06081](https://arxiv.org/abs/1707.06081).
- [Tet91] P. Tetali, ‘Random walks and the effective resistance of networks’, *J. Theoret. Probab.* **4**(1) (1991), 101–109.