

# Critical conditions for the coverage of complete graphs with the frog model\*

Gustavo O. Carvalho<sup>†</sup>      Fábio P. Machado<sup>‡</sup>

## Abstract

We consider a system of interacting random walks known as the frog model. Let  $\mathcal{K}_n = (\mathcal{V}_n, \mathcal{E}_n)$  be the complete graph with  $n$  vertices and  $o \in \mathcal{V}_n$  be a special vertex called the root. Initially,  $1 + \eta_o$  active particles are placed at the root and  $\eta_v$  inactive particles are placed at each other vertex  $v \in \mathcal{V}_n \setminus \{o\}$ , where  $\eta$  is a random variable taking values in  $\{0, 1, 2, \dots\}$  and  $\{\eta_v\}_{v \in \mathcal{V}_n}$  are i.i.d. random variables with the same law as  $\eta$ . At each instant of time, each active particle may die with probability  $1 - p$ . Every active particle performs a simple random walk on  $\mathcal{K}_n$  until the moment it dies, activating all inactive particles it hits along its path. Let  $V_\infty(\mathcal{K}_n, p)$  be the total number of visited vertices by some active particle up to the end of the process, after all active particles have died. In this paper, we show that  $V_\infty(\mathcal{K}_n, p_n) \geq (1 - \epsilon)n$  with high probability for any fixed  $\epsilon > 0$  whenever  $p_n \rightarrow 1$ . Furthermore, we establish the critical growth rate of  $p_n$  so that all vertices are visited. Specifically, we show that if  $p_n = 1 - \frac{\alpha}{\log n}$ , then  $V_\infty(\mathcal{K}_n, p_n) = n$  with high probability whenever  $0 < \alpha < E(\eta)$  and  $V_\infty(\mathcal{K}_n, p_n) < n$  with high probability whenever  $\alpha > E(\eta)$ .

**Keywords:** complete graph; coverage; frog model; random walks system.

**MSC2020 subject classifications:** 60K35; 05C81.

Submitted to ECP on August 26, 2024, final version accepted on March 20, 2025.

Supersedes arXiv:2407.19027.

## 1 Introduction and results

We study the frog model, which can be described by the following rules. Initially, particles are placed at the vertices of a connected graph  $\mathcal{G}$ . At time zero, all particles are inactive, except for those at a special selected vertex from  $\mathcal{G}$  called the root, which are active. Active particles perform independent, simple, nearest-neighbor, discrete time random walks on  $\mathcal{G}$ , activating any inactive particles found on the vertices visited by their walk. Inactive particles remain on their initial vertex until activated (if it happens). This model is often associated with the dynamics of the spread of a rumor or an infection in a population. A formal definition of the frog model can be found in [2].

Several variations of the frog model have been studied. Most of the results are based on cases where  $\mathcal{G}$  is infinite, usually the hypercubic lattice  $\mathbb{Z}^d$  or the regular tree  $\mathbb{T}^d$ . Part of the work has focused on finding conditions for recurrence and transience, i.e., conditions for the root to be visited infinitely often [21, 19, 13, 7, 20, 5]. Another area

---

\*Research supported by CAPES (88887.676435/2022-00), FAPESP (23/13453-5).

<sup>†</sup>University of São Paulo, Brazil. E-mail: gustavoodc@ime.usp.br

<sup>‡</sup>University of São Paulo, Brazil. E-mail: fmachado@ime.usp.br

of interest is describing the set of visited vertices and its asymptotic shape [3, 16, 10, 15]. For particles having a finite random lifetime, specifically those with a geometric distribution, some papers focus on studying conditions for the survival of the frog model [2, 9, 11]. In finite graphs, there are works that deal with cases where particles have fixed finite lifetime [12, 14, 4], as well as some other lifetime schemes [1, 18, 17, 6].

We consider the geometric lifetime version of the frog model. At each instant of time, independently of everything else, each active particle survives with probability  $p \in (0, 1)$  or dies otherwise, being permanently removed from the model. This means that the lifetime of active particles follow a geometric distribution supported on  $\{0, 1, \dots\}$  with parameter  $1 - p$ .

Let  $\mathcal{K}_n$  be the  $n$ -complete graph, i.e.,  $\mathcal{K}_n = (\mathcal{V}_n, \mathcal{E}_n)$  where  $\mathcal{V}_n$  is the set of vertices with  $|\mathcal{V}_n| = n$  and  $\mathcal{E}_n = \{\{x, y\} | x, y \in \mathcal{V}_n, x \neq y\}$  is the set of undirected edges linking every pair of distinct vertices. Let  $o \in \mathcal{V}_n$  be a special vertex called the root of  $\mathcal{K}_n$ . Initially, we consider that each vertex  $v \in \mathcal{V}_n \setminus \{o\}$  has  $\eta_v$  inactive particles and  $o$  has  $1 + \eta_o$  active particles, where  $\{\eta_v\}_{v \in \mathcal{V}_n}$  are i.i.d. with the same distribution as a variable  $\eta$ . The extra particle at the root is simply a way to ensure that the frog model does not end right from the start, but all of our results also hold without this extra particle if we consider the conditional probability on the event  $\{\eta_o \geq 1\}$ .

Excluding trivial cases where  $P(\eta = \infty) > 0$  or  $p = 1$ , there are finitely many particles which could be activated since  $\mathcal{V}_n$  is finite and each particle has a finite lifetime (given by a geometric distribution). Thus, there is a finite instant in which the process is over, i.e., there are no more active particles alive. Let  $V_\infty(\mathcal{K}_n, p)$  be the number of visited vertices when the frog model is over. When either of the trivial cases above happens, it could be the case that the process is not over at any instant of time and all vertices are visited; in this case we simply say that  $V_\infty(\mathcal{K}_n, p) := n$ .

When there is exactly one particle per vertex ( $\eta \equiv 1$ ), [6, Theorem 1.2] shows that both  $V_\infty(\mathcal{K}_n, p) \leq c \log n$  and  $V_\infty(\mathcal{K}_n, p) \geq c'n$  have positive asymptotic probabilities for  $p > 1/2$  and suitable constants  $c > 0$  and  $c' \in (0, 1)$ . The following theorem shows that when  $p = p_n$  such that  $p_n \rightarrow 1$ , w.h.p.<sup>1</sup> only the case  $V_\infty(\mathcal{K}_n, p) \geq c'n$  is possible. More than that, the same goes for every  $c' \in (0, 1)$ , therefore w.h.p. any arbitrary large proportion of vertices is visited.

**Theorem 1.1.** *Let  $\eta$  be such that  $P(\eta = 0) < 1$ . If  $(p_n)_{n \in \mathbb{N}}$  is a sequence such that  $\lim_{n \rightarrow \infty} p_n = 1$ , then for any  $\epsilon \in (0, 1)$ ,*

$$\lim_{n \rightarrow \infty} P(V_\infty(\mathcal{K}_n, p_n) \geq (1 - \epsilon)n) = 1.$$

Theorem 1.1 may appear to suggest that all vertices of  $\mathcal{K}_n$  are visited when  $p_n \rightarrow 1$ , but this is not necessarily the case. Actually, there is a significant difference between visiting a large proportion of vertices and visiting all vertices, similar to what occurs in another process that will later be used in a coupling with the frog model: the coupon collector's problem. As the collection increases, it becomes harder and harder to find new coupons in the coupon collector's problem, thus obtaining  $(1 - \epsilon)n$  different coupons demand roughly  $cn$  draws, where  $c = c(\epsilon)$  is a constant; while obtaining all  $n$  different coupons demand roughly  $n \log n$  draws.

The next theorem shows the rate at which  $p_n$  must grow in order to have w.h.p. all vertices visited. For any  $x \in \mathbb{R}$ , let  $x^+ := \max\{0, x\}$  be its positive part.

**Theorem 1.2.** *Let  $\eta$  be such that  $P(\eta = 0) < 1$ . Consider that  $p_n = (1 - \frac{\alpha}{\log n})^+$ .*

- (i) *If  $\alpha > E(\eta)$ , then  $\lim_{n \rightarrow \infty} P(V_\infty(\mathcal{K}_n, p_n) = n) = 0$ .*
- (ii) *If  $0 < \alpha < E(\eta)$ , then  $\lim_{n \rightarrow \infty} P(V_\infty(\mathcal{K}_n, p_n) = n) = 1$ .*

<sup>1</sup>We say that a sequence of events  $(E_n)_{n \in \mathbb{N}}$  happens with high probability (w.h.p.) when  $\lim_{n \rightarrow \infty} P(E_n) = 1$ .

**Remark 1.3.** When  $E(\eta) = +\infty$ , Theorem 1.2 part (ii) applies for every  $\alpha > 0$  and we have that  $\lim_{n \rightarrow \infty} P(V_\infty(\mathcal{K}_n, (1 - \frac{\alpha}{\log n})^+) = n) = 1$ . We still do not know what happens when  $\alpha = E(\eta) < +\infty$ .

Theorem 1.2 relates to Hermon’s result in [12, Proposition 1.1], where the author addresses the frog model on  $\mathcal{K}_n$ , studying cases where particles have a fixed finite lifetime and  $\eta$  is Poisson distributed with parameter  $\lambda_n$ , possibly varying with  $n$ . Among other results, Hermon studies the susceptibility  $S(\mathcal{K}_n)$ , essentially determining the shortest particles’ lifespan required for the entire graph to be visited by active particles. It is shown that  $(1 - \epsilon)\lambda_n^{-1} \log n \leq S(\mathcal{K}_n) \leq \lceil (1 + \epsilon)\lambda_n^{-1} \log n \rceil$  holds w.h.p. for any  $\epsilon \in (0, 1)$  as long as  $n\lambda_n \rightarrow \infty$ . Although Hermon’s proof approach does not apply to our case, the results obtained are comparable. In our case, when  $p_n = (1 - \frac{\alpha}{\log n})^+$ , the mean lifetime of each particle is  $\frac{p_n}{1-p_n} \approx \frac{\log n}{\alpha}$ . Since the total number of particles is roughly  $nE(\eta)$  (or  $n\lambda_n$  is Hermon’s case), the mean total lifetime of all particles together in both cases can be compared with the coupon collector time  $n \log n$ .

Simple coupling techniques show that  $P(V_\infty(\mathcal{K}_n, p) = n)$  is non-decreasing in  $p$ . So, Theorem 1.2 is also useful for several sequences  $(p_n)_{n \in \mathbb{N}}$ , depending on their asymptotic behavior. We have that  $\lim_{n \rightarrow \infty} P(V_\infty(\mathcal{K}_n, p) = n) = 0$  whenever  $p_n \leq 1 - \frac{\alpha}{\log n}$  for some  $\alpha > E(\eta)$  and sufficiently large  $n$ ; we also have that  $\lim_{n \rightarrow \infty} P(V_\infty(\mathcal{K}_n, p_n) = n) = 1$  whenever  $p_n \geq 1 - \frac{\alpha}{\log n}$  for some  $0 < \alpha < E(\eta)$  and sufficiently large  $n$ .

It is also interesting to note what happens when  $(p_n)_{n \in \mathbb{N}}$  is a sequence that slowly converges to 1. When  $\lim_{n \rightarrow \infty} p_n = 1$  and  $p_n \leq 1 - \frac{\alpha}{\log n}$  for some  $\alpha > E(\eta)$  and sufficiently large  $n$ , then  $(1 - \epsilon)n \leq V_\infty(\mathcal{K}_n, p_n) < n$  w.h.p. for any  $\epsilon \in (0, 1)$  by Theorems 1.1 and 1.2. In this case, there is a non-empty set of unvisited vertices at the end of the process, but its cardinality is smaller than any pre-established proportion.

The remainder of this article is organized as follows. In Section 2, we provide some useful lemmas, as well as definitions and results about models that can be coupled with the frog model on  $\mathcal{K}_n$ . The proofs of all theorems are presented in Section 3.

## 2 Auxiliary results

### 2.1 The coupon collector’s problem

Consider the classical coupon collector’s problem. There are  $n$  coupons and a collector who acquires a random and uniformly selected coupon at each instant of time. Let  $\tau_i$  be the number of draws needed to obtain  $i \in \{1, \dots, n\}$  different coupons, and  $\tau := \tau_n$  be the number of draws needed to complete the entire collection.

Note that  $\tau_{j+1} - \tau_j \sim Geo_1(\frac{n-j}{n})$ ,  $j \in \{0, 1, \dots, n-1\}$ , where  $\tau_0 := 0$  by definition.<sup>2</sup> Then,  $\tau_{j+1} - \tau_j$  has the same distribution as the number of steps that active particles take (if there are particles that have enough lifetime to do so) to go from  $j+1$  to  $j+2$  visited vertices in  $\mathcal{K}_{n+1}$ . Similarly,  $\tau = \sum_{j=0}^{n-1} (\tau_{j+1} - \tau_j)$  is related to the potential coverage of  $\mathcal{K}_{n+1}$  in the frog model.

Next, we show two useful lemmas regarding the coupon collector’s problem. We only show the proof for the second lemma. The first one has a simple proof involving expectation and variance of  $\tau$ , which can be seen in [8, Example 2.2.7].

**Lemma 2.1.** For any  $\epsilon \in (0, 1)$ ,

$$\lim_{n \rightarrow \infty} P[(1 - \epsilon)n \log n \leq \tau \leq (1 + \epsilon)n \log n] = 1.$$

<sup>2</sup>We let  $Geo_0(\cdot)$  and  $Geo_1(\cdot)$  denote geometric distributions with support on  $\{0, 1, \dots\}$  and on  $\{1, 2, \dots\}$ , respectively.

**Lemma 2.2.** For any  $\epsilon \in (0, 1)$ ,

$$\lim_{n \rightarrow \infty} P\left(\tau_{\lceil(1-\epsilon)n\rceil} \leq 2\left(\frac{1-\epsilon}{\epsilon}\right)n\right) = 1.$$

*Proof.* If  $Q_1 \sim Geo_1(q_1)$  and  $Q_2 \sim Geo_1(q_2)$  with  $q_1 \leq q_2$ , then  $Q_1 \succeq Q_2$  is obtained directly from their cumulative distribution functions.<sup>3</sup> Combining this with the fact that  $\frac{n-i}{n} \geq \frac{n-(1-\epsilon)n}{n} = \epsilon$  for  $i \in \{0, 1, \dots, \lceil(1-\epsilon)n\rceil - 1\}$ , we have that

$$\tau_{\lceil(1-\epsilon)n\rceil} = \sum_{i=0}^{\lceil(1-\epsilon)n\rceil-1} (\tau_{i+1} - \tau_i) \leq \sum_{i=1}^{\lceil(1-\epsilon)n\rceil} Z_i, \tag{2.1}$$

where  $Z_i \stackrel{i.i.d.}{\sim} Geo_1(\epsilon)$ .

Furthermore, as  $E(\sum_{i=1}^{\lceil(1-\epsilon)n\rceil} Z_i) = \frac{\lceil(1-\epsilon)n\rceil}{\epsilon}$  and  $\frac{\lceil(1-\epsilon)n\rceil}{\lceil(1-\epsilon)n\rceil} \xrightarrow{n \rightarrow \infty} 1$ , we have for sufficiently large  $n$  that

$$\frac{2(1-\epsilon)n}{\epsilon \lceil(1-\epsilon)n\rceil} - \frac{E\left(\sum_{i=1}^{\lceil(1-\epsilon)n\rceil} Z_i\right)}{\lceil(1-\epsilon)n\rceil} \geq \frac{1}{2\epsilon}. \tag{2.2}$$

By (2.1) and (2.2),

$$\begin{aligned} P\left(\tau_{\lceil(1-\epsilon)n\rceil} \geq \frac{2(1-\epsilon)}{\epsilon}n\right) &\leq P\left(\sum_{i=1}^{\lceil(1-\epsilon)n\rceil} Z_i \geq \frac{2(1-\epsilon)}{\epsilon}n\right) \\ &\leq P\left(\frac{\sum_{i=1}^{\lceil(1-\epsilon)n\rceil} Z_i - E(\sum_{i=1}^{\lceil(1-\epsilon)n\rceil} Z_i)}{\lceil(1-\epsilon)n\rceil} \geq \frac{1}{2\epsilon}\right). \end{aligned}$$

Therefore,  $\lim_{n \rightarrow \infty} P(\tau_{\lceil(1-\epsilon)n\rceil} \geq \frac{2(1-\epsilon)}{\epsilon}n) = 0$  by the law of large numbers. □

## 2.2 Auxiliary process

Here we establish an auxiliary process analogous to the one in [6]. The auxiliary process can be thought of as the frog model on a connected graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , but in such a way that only one particle is allowed to move per round, and additional particles are included so that the process continues infinitely.

As in the frog model, a vertex  $o \in \mathcal{V}$  is called the root of  $\mathcal{G}$ . At time zero, every vertex  $v \in \mathcal{V} \setminus \{o\}$  has  $\eta_v$  inactive particles while  $o$  has  $1 + \eta_o$  active particles, but  $\{\eta_v\}_{v \in \mathcal{V}}$  are not necessarily i.i.d. Let all  $1 + \sum_{i \in \mathcal{V}} \eta_i$  initial particles be denoted as originals. This is done because *extra* particles will be added to the process at some point. Given the initial configuration, the auxiliary process on  $\mathcal{G}$  with survival parameter  $p$  has the following rules:

- At each instant  $k$  of time, one active particle is chosen (the second and third rules will guarantee that there is at least one active particle) to participate in round  $k$ . This particle can be chosen randomly or by any selection rule. The chosen particle checks whether it survives with probability  $p$  and, if so, moves to a randomly and uniformly chosen connected vertex, activating any inactive particles at that vertex.
- Let  $R = R(\mathcal{G}, p)$  be the moment at which the last living original active particle dies, that is, the instant at which the original process ends. At this same moment, every remaining inactive particle is now called *extra inactive*. Furthermore, a new *extra active* particle is injected at the root  $o \in \mathcal{V}$  to continue the process.

<sup>3</sup>For any two random variables  $X$  and  $Y$  in the same probability space, we say that  $X$  *stochastically dominates*  $Y$  when  $P(X \geq a) \geq P(Y \geq a)$  for all  $a \in \mathbb{R}$ . We denote this relation by  $X \succeq Y$ .

- After  $R$ , each time the last living active particle dies, a new extra active particle is added at the root  $o \in \mathcal{V}$ .

Note that there are only extra particles participating from round  $R + 1$  onwards. Excluding cases in which  $\sum_{i \in \mathcal{V}} \eta_i = \infty$  or  $p = 1$ , we have that  $R < \infty$  with probability 1, since at most  $1 + \sum_{i \in \mathcal{V}} \eta_i$  original particles could be activated, each with a finite lifetime given by a geometric distribution with parameter  $1 - p$ . Despite  $R < \infty$ , the auxiliary process as a whole is infinite, as it always renews itself by injecting new extra particles on the root.

By this construction, it is possible to guarantee that there is always an active particle (original or extra) to participate in the  $k$ -th round for any  $k \in \mathbb{N}$ , regardless of whether the original process has ended or not. For  $k \in \mathbb{N}$ , define  $X_k = X_k(\mathcal{G}, p)$  depending on what happens to the particle that participates in round  $k$ :

- $X_k = 0$  if it dies.
- $X_k = 1$  if it survives and goes to a vertex which has been visited before round  $k$ ; or if it survives and goes to a previously unvisited vertex  $v$  with  $\eta_v = j - 1$ .
- $X_k = j$ ,  $j \in \{2, 3, \dots\}$ , if it survives and goes to a previously unvisited vertex  $v$  with  $\eta_v = j - 1$ .

Note that the previous random variable can be interpreted as the number of descendants (in the sense of a branching process) of the particle participating in round  $k$ . From the number of descendants in each round, we define the number  $A'_k = A'_k(\mathcal{G}, p)$  of potentially active particles at the end of the  $k$ -th round as

$$A'_0 := 1 + \eta_o; \quad A'_k := 1 + \eta_o + \sum_{j=1}^k (X_j - 1), \quad k \in \{1, 2, \dots\}. \quad (2.3)$$

The term *potential* comes from the fact that the previous expression may also be counting the descendants of extra particles. We can disregard this by denoting  $R = R(\mathcal{G}, p)$  as

$$R = \inf\{k : A'_k = 0\} \quad (2.4)$$

and defining the number  $A_k = A_k(\mathcal{G}, p)$  of active particles at the end of the  $k$ -th round as

$$A_k := A'_k \mathbb{1}_{(k < R)}, \quad k \in \{0, 1, 2, \dots\}. \quad (2.5)$$

Let  $E_k \subset \{X_k \geq 1\}$  be the event in which the particle participating in round  $k$  survives and visits a vertex that has never been visited before. Similarly to what was done previously, we define the number  $V'_k = V'_k(\mathcal{G}, p)$  of potentially visited vertices until the end of the  $k$ -th round as

$$V'_0 := 1; \quad V'_k := 1 + \sum_{j=1}^k \mathbb{1}_{E_j}, \quad k \in \{1, 2, \dots\} \quad (2.6)$$

and the number  $V_k = V_k(\mathcal{G}, p)$  of visited vertices until the end of the  $k$ -th round as

$$V_k := V'_k \mathbb{1}_{(k < R)} + V'_R \mathbb{1}_{(k \geq R)}, \quad k \in \{0, 1, 2, \dots\}. \quad (2.7)$$

We also define the total number  $V_\infty = V_\infty(\mathcal{G}, p)$  of visited vertices during the entire process by

$$V_\infty := \lim_{k \rightarrow \infty} V_k = V_R. \quad (2.8)$$

Note that  $V_\infty = V_R$  is not affected by the extra particles injected into the process, as they only play a role after  $R$ . The extra particles are only useful for the mathematical

convenience they bring, such as allowing the evaluation of the number of potentially active particles and potentially visited vertices without needing to worry about whether all the original particles have already died or not. It is also important to note that allowing only one particle to move per round maintains the same final number of visited vertices as the original dynamics of the frog model. Therefore, the total number of visited vertices in the corresponding frog model is also described by  $V_\infty$ , making it possible to study the frog model through the auxiliary process.

Specifically when we deal with complete graphs and consider that  $\{\eta_v\}_{v \in \mathcal{V}}$  is a collection of i.i.d. random variables with the same distribution as  $\eta$ , we have that  $P(X_j(\mathcal{K}_{n+1}, p) = x | V'_{j-1}(\mathcal{K}_{n+1}, p) = v)$  equals

$$\begin{cases} 1 - p, & \text{if } x = 0; \\ \frac{p(v-1)}{n} + p\left(\frac{n-v+1}{n}\right)P(\eta = 0), & \text{if } x = 1; \\ p\left(\frac{n-v+1}{n}\right)P(\eta = k - 1), & \text{if } x = k \text{ for } k = 2, 3, 4, \dots \end{cases} \quad (2.9)$$

### 2.3 Helpful lemmas

Here we provide two lemmas for later use. The first one relates the number of active particles to the total number of steps they will take before dying. This lemma will be particularly useful in combination with the results of the coupon collector's problem, as it gives us the necessary number of active particles such that the total number of remaining steps is greater than what is needed to visit a high proportion of vertices or even the entire graph.

**Lemma 2.3.** *Consider the auxiliary process on any graph and with survival parameter  $p_n$ . Consider that at some instant  $k = k(n)$  there are  $A_k$  active particles and define  $T(k)$  as the combined number of remaining steps that these  $A_k$  particles will still take (i.e.,  $T(k)$  is the sum of the remaining lifetimes of these  $A_k$  particles), with  $T(k) := 0$  if  $A_k = 0$ . Also assume that  $p_n \neq 1$  for all  $n \in \mathbb{N}$  and  $\lim_{n \rightarrow \infty} p_n = 1$  and define  $q_n := (1 - p_n)^{-1}$ . For any  $\epsilon \in (0, 1)$  and  $b > 0$ ,*

- (i)  $\lim_{n \rightarrow \infty} P(T(k) \leq (1 - \epsilon/2)bnq_n | A_k \leq (1 - \epsilon)bn) = 1;$
- (ii)  $\lim_{n \rightarrow \infty} P(T(k) \geq (1 + \epsilon/2)bnq_n | A_k \geq (1 + \epsilon)bn) = 1.$

*Proof.* Let  $T := T(k)$  be the combined total number of steps that all particles that are active at time  $k$  will still take. Let  $\{T_i\}_{i \geq 1}$  be a sequence of random variables such that  $T_i \stackrel{i.i.d.}{\sim} Geo_0(1 - p_n)$ . Note that  $T$  has the same distribution as  $\sum_{i=1}^{A_k} T_i$  because, due to lack of memory, we can ignore how many times each particle has survived before time  $k$ . Also note that

$$\begin{cases} T \leq \sum_{i=1}^{\lceil (1-\epsilon)bn \rceil} T_i, & \text{if } A_k \leq (1 - \epsilon)bn; \\ T \geq \sum_{i=1}^{\lfloor (1+\epsilon)bn \rfloor} T_i, & \text{if } A_k \geq (1 + \epsilon)bn. \end{cases} \quad (2.10)$$

As  $T_i \sim Geo_0(1 - p_n)$ , we have that  $E(T_i) = q_n - 1$  and  $Var(T_i) = q_n^2 - q_n$ . We use simple floor and ceil inequalities such as  $x - 1 \leq \lfloor x \rfloor \leq x \leq \lceil x \rceil \leq x + 1$  for any real  $x$  to obtain

$$\begin{cases} E\left(\sum_{i=1}^{\lceil (1-\epsilon)bn \rceil} T_i\right) \leq ((1 - \epsilon)bn + 1)q_n - (1 - \epsilon)bn; \\ E\left(\sum_{i=1}^{\lfloor (1+\epsilon)bn \rfloor} T_i\right) \geq ((1 + \epsilon)bn - 1)q_n - (1 + \epsilon)bn, \end{cases} \quad (2.11)$$

and

$$\begin{cases} Var\left(\sum_{i=1}^{\lceil (1-\epsilon)bn \rceil} T_i\right) \leq ((1 - \epsilon)bn + 1)(q_n^2 - q_n); \\ Var\left(\sum_{i=1}^{\lfloor (1+\epsilon)bn \rfloor} T_i\right) \leq (1 + \epsilon)bn(q_n^2 - q_n). \end{cases} \quad (2.12)$$

Since both  $n$  and  $q_n$  diverges as  $p_n \rightarrow 1$  by assumption, then  $q_n + (1 + \epsilon)bn \leq (\epsilon/4)bnq_n$  for sufficiently large  $n$ , which implies by (2.11) that

$$(1 + \epsilon/2)bnq_n - E\left(\sum_{i=1}^{\lfloor(1+\epsilon)bn\rfloor} T_i\right) \leq -(\epsilon/2)bnq_n + q_n + (1 + \epsilon)bn \tag{2.13}$$

$$\leq -(\epsilon/4)bnq_n.$$

On the other hand, for sufficiently large  $n$  it is also true that  $(\epsilon/2)bnq_n - q_n \geq (\epsilon/4)bnq_n$  and

$$(1 - \epsilon/2)bnq_n - E\left(\sum_{i=1}^{\lceil(1-\epsilon)bn\rceil} T_i\right) \geq (\epsilon/2)bnq_n + (1 - \epsilon)bn - q_n \tag{2.14}$$

$$\geq (\epsilon/4)bnq_n.$$

By (2.10), (2.12), (2.13) and Chebyshev's inequality, we have for sufficiently large  $n$  that

$$P(T \leq (1 + \epsilon/2)bnq_n | A_k \geq (1 + \epsilon)bn) \leq P\left(\sum_{i=1}^{\lfloor(1+\epsilon)bn\rfloor} T_i \leq (1 + \epsilon/2)bnq_n\right)$$

$$\leq P\left(\sum_{i=1}^{\lfloor(1+\epsilon)bn\rfloor} T_i - E\left(\sum_{i=1}^{\lfloor(1+\epsilon)bn\rfloor} T_i\right) \leq -(\epsilon/4)bnq_n\right)$$

$$\leq P\left(\left|\sum_{i=1}^{\lfloor(1+\epsilon)bn\rfloor} T_i - E\left(\sum_{i=1}^{\lfloor(1+\epsilon)bn\rfloor} T_i\right)\right| \geq (\epsilon/4)bnq_n\right)$$

$$\leq \frac{Var(\sum_{i=1}^{\lfloor(1+\epsilon)bn\rfloor} T_i)}{[(\epsilon/4)bnq_n]^2}$$

$$\leq \frac{(1 + \epsilon)(q_n^2 - q_n)}{(\epsilon/4)^2 b n q_n^2} \xrightarrow{n \rightarrow \infty} 0.$$

So, the proof of part (ii) is complete. Furthermore,  $P(T \geq (1 - \epsilon/2)bnq_n | A_k \leq (1 - \epsilon)bn) \xrightarrow{n \rightarrow \infty} 0$  can be shown in an analogous way by using (2.14) instead of (2.13), thus also completing the proof of part (i).  $\square$

The next lemma guarantees that the original auxiliary process (and so the frog model) lasts for at least some initial rounds whenever  $p_n \rightarrow 1$ . Later on, this will be important when we show that the particles collected in this early stage are already sufficient to visit a high proportion of vertices.

**Lemma 2.4.** *Consider the auxiliary process in which  $\eta_n$  are i.i.d. variables with the same distribution as  $\eta$ . Suppose  $\lim_{n \rightarrow \infty} p_n = 1$  and  $P(\eta = 0) < 1$ . For any  $c \in (0, 1)$ , we have that*

$$\lim_{n \rightarrow \infty} P(R(\mathcal{K}_{n+1}, p_n) < k_1) = 0,$$

where  $k_1 = k_1(n) = \lfloor cn \rfloor - 1$ .

*Proof.* The proof is based on the relationship between the auxiliary process in its initial rounds and a branching process. Let  $\{Z_j^{(n)}\}_{j \geq 1}$  be a sequence of i.i.d. random variables such that

$$P(Z_j^{(n)} = x) = \begin{cases} 1 - p_n, & \text{if } x = 0; \\ p_n c + p_n(1 - c)P(\eta = 0), & \text{if } x = 1; \\ p_n(1 - c)P(\eta \geq 1), & \text{if } x = 2. \end{cases} \tag{2.15}$$

When  $j \leq k_1$ , we have by (2.6) that  $V'_{j-1}(\mathcal{K}_{n+1}, p_n) \leq j \leq k_1$  and, in particular, that  $p_n(V'_{j-1}(\mathcal{K}_{n+1}, p_n) - 1) \leq \frac{p_n(k_1 - 1)}{n} \leq p_n c$ . Therefore, it is possible to create a coupling such that  $Z_j^{(n)} \leq X_j(\mathcal{K}_{n+1}, p_n)$  for  $j \leq k_1$  (compare their cumulative distribution functions by means of (2.9) and (2.15)). So,  $\sum_{j=1}^k X_j(\mathcal{K}_{n+1}, p_n) \geq \sum_{j=1}^k Z_j^{(n)}$  holds for all  $k \in \{1, \dots, k_1\}$ . Recalling the interpretation of  $X_j$  as the number of descendants of the particle participating in round  $j$ , this also implies that

$$P(R(\mathcal{K}_{n+1}, p_n) < k_1) \leq P(R_{Z^{(n)}} < k_1) \leq P(R_{Z^{(n)}} < \infty), \tag{2.16}$$

where  $R_{Z^{(n)}}$  is the number of rounds until the extinction of a branching process with the number of descendants given by  $Z^{(n)}$ . For the first inequality in (2.16) to be precise, we should consider that the rounds, as in the auxiliary process, are done by counting descendants of one particle at a time, unlike the usual rounds scheme for branching processes of counting descendants per generation. But as a whole, it does not matter as  $P(R_{Z^{(n)}} < \infty)$  is the same for both round schemes. We will prove that  $\lim_{n \rightarrow \infty} P(R_{Z^{(n)}} < \infty) = 0$ , thus completing the proof of the lemma as a consequence of (2.16).

Fix any  $\epsilon \in (0, 1)$ . We want to show that  $P(R_{Z^{(n)}} < \infty) < \epsilon$  for sufficiently large  $n$ . By classical properties of a branching process and its extinction probability, we have to show that

$$s = E(s^{Z^{(n)}}) \tag{2.17}$$

has a solution in  $(0, \epsilon)$  for sufficiently large  $n$ .

Equivalently, any root of

$$f_n(s) := E(s^{Z^{(n)}}) - s = (1 - p_n) + s[p_n c + p_n(1 - c)P(\eta = 0) - 1] + s^2 p_n(1 - c)P(\eta \geq 1)$$

is a solution of (2.17).

We can consider just the cases where  $p_n < 1$  for sufficiently large  $n$ , as  $p_n = 1$  would make  $P(Z^{(n)} = 0) = 0$  and trivially  $P(R_{Z^{(n)}} < \infty) = 0 < \epsilon$ . Then,  $f_n(0) = 1 - p_n > 0$  and  $f_n(s)$  is continuous in  $s$ . So, to show that  $f_n$  has a root in  $(0, \epsilon)$ , it is enough to show that  $f_n(\epsilon) < 0$ .

As  $P(\eta = 0) < 1$ , then  $P(\eta = 0) + \epsilon P(\eta \geq 1) < 1$ . Let

$$a := c + (1 - c)[P(\eta = 0) + \epsilon P(\eta \geq 1)] \in (0, 1).$$

Since  $\lim_{n \rightarrow \infty} p_n = 1$ , it is possible to make  $(1 - p_n) < \epsilon(1 - a)/2$  by letting  $n$  be large enough, and then

$$f_n(\epsilon) = (1 - p_n) + \epsilon(ap_n - 1) \leq (1 - p_n) - \epsilon(1 - a) < 0. \quad \square$$

### 3 Proofs

Most of the time during this section, the auxiliary process and related variables consider the graph  $\mathcal{K}_{n+1}$ , survival parameter  $p_n$  and that  $\{\eta_v\}_{v \in \mathcal{V}_{n+1}}$  are i.i.d. with the same distribution as a variable  $\eta$ . So, these will be the default to be considered when notation is omitted to make the text cleaner.

We begin by proving the following lemma, which already implies Theorem 1.1 and which will also help in the proof of the Theorem 1.2. For this lemma, we use the asymptotic notation  $o(1)$  to denote a function  $g$  such that  $\lim_{n \rightarrow \infty} g(n) = 0$ .

**Lemma 3.1.** *Consider that  $P(\eta = 0) < 1$  and  $\lim_{n \rightarrow \infty} p_n = 1$ . Fix any  $c \in (0, 1)$ ,  $\epsilon \in (0, 1)$  such that  $c < 1 - \epsilon$ . Let  $k_1 = k_1(n) := \lfloor cn \rfloor - 1$  and  $k_2 = k_2(n) := \inf\{k : V_k(\mathcal{K}_{n+1}, p_n) \geq \lceil (1 - \epsilon)n \rceil + 1\}$  with the definition that  $\inf \emptyset = \infty$ . Then*

$$\lim_{n \rightarrow \infty} P(k_2(n) < \infty) = 1$$

and for any sequence  $(x_n)_{n \in \mathbb{N}}$  of real numbers,

$$P(A_{k_2}(\mathcal{K}_{n+1}, p_n) \geq x_n) \geq P\left(\sum_{i=k_1(n)+1}^{\lceil(1-\epsilon)n\rceil} \eta_i \geq x_n\right) - o(1).$$

**Remark 3.2.** Lemma 3.1 implies Theorem 1.1 as argued below. Since the only restriction to  $(p_n)_{n \in \mathbb{N}}$  in Theorem 1.1 is that  $\lim_{n \rightarrow \infty} p_n = 1$ , we can prove in an equivalent way and without loss of generality that  $V_\infty(\mathcal{K}_{n+1}, p_n) \geq (1 - \epsilon)(n + 1)$  w.h.p. for any fixed  $\epsilon \in (0, 1)$ . With a fixed  $\epsilon \in (0, 1)$ , we can take  $0 < \epsilon' < \epsilon$  so that  $(1 - \epsilon')n \geq (1 - \epsilon)(n + 1)$  for sufficiently large  $n$ . By Lemma 3.1, we have that w.h.p.  $k_2 = \inf\{k : V_k(\mathcal{K}_{n+1}, p_n) \geq \lceil(1 - \epsilon')n\rceil + 1\} < \infty$ , and therefore w.h.p.  $V_\infty(\mathcal{K}_{n+1}, p_n) \geq V_{k_2}(\mathcal{K}_{n+1}, p_n) \geq (1 - \epsilon')n \geq (1 - \epsilon)(n + 1)$ .

*Proof.* In order to avoid some technical problems in the proof of Lemma 3.1, we consider only cases in which  $p_n$  is non-decreasing and  $p_n \neq 1$  for all  $n \in \mathbb{N}$  in addition to the original assumption of  $\lim_{n \rightarrow \infty} p_n = 1$ . These additional assumptions do not reduce the generality of  $(p_n)_{n \in \mathbb{N}}$ . In fact, if  $(p_n)_{n \in \mathbb{N}}$  is any sequence such that  $\lim_{n \rightarrow \infty} p_n = 1$ , we can construct another sequence  $(p'_n)_{n \in \mathbb{N}}$  such that  $p'_n := \min\{1 - \frac{1}{n}, \inf_{k \geq n} \{p_k\}\}$ . Since  $\lim_{n \rightarrow \infty} p'_n = 1$ ,  $p'_n$  is non-decreasing and  $p'_n \neq 1$  for all  $n \in \mathbb{N}$ , the lemma would be valid for  $p'_n$  and consequently also for  $p_n$ , because  $p_n \geq p'_n$  implies  $V_k(\mathcal{K}_{n+1}, p_n) \succeq V_k(\mathcal{K}_{n+1}, p'_n)$  and  $A_k(\mathcal{K}_{n+1}, p_n) \succeq A_k(\mathcal{K}_{n+1}, p'_n)$  for all  $k \in \mathbb{N}$ .

Initially, suppose that  $E(\eta) < \infty$ . Let  $\{Y_j^{(n)}\}_{j \in \mathbb{N}}$  be a sequence of i.i.d. random variables such that

$$P(Y_j^{(n)} = x) = \begin{cases} 1 - p_n, & \text{if } x = 0; \\ p_n c + p_n(1 - c)P(\eta = 0), & \text{if } x = 1; \\ p_n(1 - c)P(\eta = k - 1), & \text{if } x = k \text{ for } k = 2, 3, 4, \dots \end{cases} \quad (3.1)$$

Note that  $E(Y_j^{(n)}) = p_n c + p_n(1 - c)[E(\eta) + 1] \xrightarrow{n \rightarrow \infty} c + (1 - c)[E(\eta) + 1]$  as  $\lim_{n \rightarrow \infty} p_n = 1$ . Since  $P(\eta = 0) < 1$  and therefore  $E(\eta) > 0$ , we have that  $c + (1 - c)[E(\eta) + 1] = 1 + (1 - c)E(\eta) > 1$ . So, if we let  $d := (1 - c)E(\eta)/2 > 0$ , then

$$E(Y_j^{(n)}) \geq 1 + d \text{ for sufficiently large } n. \quad (3.2)$$

Let  $b > 0$  be a constant small enough so that  $\frac{(1+\epsilon)b}{c} \leq d/4$  and therefore  $\frac{(1+\epsilon)bn}{k_1} \leq \frac{(1+\epsilon)bn}{cn-2} \leq d/2$  for sufficiently large  $n$ . Using this fact and (3.2), we conclude that there is an  $n_0$  such that the following inequality, which will be used later, holds:

$$\frac{(1 + \epsilon)bn}{k_1} + 1 - E(Y_j^{(n_0)}) \leq -d/2 < 0, \text{ for } n \geq n_0. \quad (3.3)$$

Note that the comparison  $\sum_{j=1}^{k_1} X_j(\mathcal{K}_{n+1}, p_n) \succeq \sum_{j=1}^{k_1} Y_j^{(n)}$  holds (the argument is analogous to that used in Lemma 2.4 to show that  $\sum_{j=1}^{k_1} X_j(\mathcal{K}_{n+1}, p_n) \succeq \sum_{j=1}^{k_1} Z_j^{(n)}$ ). Recall that we are considering  $p_n$  to be non-decreasing, so

$$\sum_{j=1}^{k_1} X_j(\mathcal{K}_{n+1}, p_n) \succeq \sum_{j=1}^{k_1} Y_j^{(n)} \succeq \sum_{j=1}^{k_1} Y_j^{(n_0)} \text{ for } n \geq n_0. \quad (3.4)$$

When  $n \geq n_0$ , we can use the definition in (2.3) as well as (3.3) and (3.4) to conclude

that

$$\begin{aligned} P(A'_{k_1}(\mathcal{K}_{n+1}, p_n) \geq (1 + \epsilon)bn) &= P\left(\sum_{j=1}^{k_1} X_j(\mathcal{K}_{n+1}, p_n) \geq (1 + \epsilon)bn + k_1 - 1 - \eta_0\right) \\ &\geq P\left(\sum_{j=1}^{k_1} Y_j^{(n_0)} \geq (1 + \epsilon)bn + k_1\right) \\ &\geq P\left(\frac{\sum_{j=1}^{k_1} Y_j^{(n_0)} - E(\sum_{j=1}^{k_1} Y_j^{(n_0)})}{k_1} \geq -d/2\right), \end{aligned}$$

which implies that  $A'_{k_1} \geq (1 + \epsilon)bn$  w.h.p. by the law of large numbers as  $k_1(n) = \lfloor cn \rfloor - 1 \rightarrow +\infty$ .

By Lemma 2.4, we can guarantee that  $A_{k_1} = A'_{k_1}$  w.h.p. (see (2.5)), meaning that the original process does not die before round  $k_1$  and there are indeed at least  $(1 + \epsilon)bn$  active particles at the end of this round, not just potentially active particles.

Let us now recall we concluded that  $A_{k_1} \geq (1 + \epsilon)bn$  w.h.p. only for the case in which  $E(\eta) < \infty$ , in addition to the original assumption of Lemma 3.1 that  $P(\eta = 0) < 1$ . The case  $E(\eta) = \infty$  can be obtained by comparison, using  $\eta^* := \mathbb{1}(\eta \geq 1)$ . We temporarily use the notations of  $A_{k_1}$  and  $A^*_{k_1}$  to differentiate the models using  $\eta$  and  $\eta^*$ , respectively. As  $P(\eta^* = 0) < 1$  and  $E(\eta^*) < \infty$ , then w.h.p.  $A^*_{k_1} \geq (1 + \epsilon)bn$ . Combining this with the fact that  $A_{k_1} \succeq A^*_{k_1}$ , we conclude that  $A_{k_1} \geq (1 + \epsilon)bn$  w.h.p. also in the case  $E(\eta) = \infty$ .

Let us also recall that the auxiliary process on  $\mathcal{K}_n$  and all equations shown related to it do not depend on the rule for choosing which active particle is selected to participate in each round. So, now we make a non-mandatory modification of such rule whose sole purpose is to make the proof possibly easier to understand. For any  $k \geq k_1$ , it always happens that one of the  $A_{k_1}$  active particles alive at time  $k_1$  is chosen to participate in round  $k$ , unless there is no longer any living active particle at time  $k$  that was also active at time  $k_1$ . In other words, particles activated after  $k_1$  have to wait until all particles that were active at time  $k_1$  die before having the chance to participate in a round.

Let  $T(k_1)$  be the total combined number of remaining steps that all  $A_{k_1}$  active particles alive at time  $k_1$  have yet to take. Note that w.h.p.  $T(k_1) \geq (1 + \epsilon/2)bn(1 - p_n)^{-1}$  by Lemma 2.3 part (ii). Also,  $\tau_{\lceil(1-\epsilon)n\rceil} \leq n\lceil 2(\frac{1-\epsilon}{\epsilon}) \rceil$  holds w.h.p. in the coupon collector's problem by Lemma 2.2. As  $(1 - p_n)^{-1}$  diverges, we have that w.h.p.

$$T(k_1) \geq n(1 - p_n)^{-1}(1 + \epsilon/2)b \geq n\left[2\left(\frac{1 - \epsilon}{\epsilon}\right)\right] \geq \tau_{\lceil(1-\epsilon)n\rceil}.$$

Therefore, by the comparison between the coupon collector's problem with  $n$  coupons and the auxiliary process on  $\mathcal{K}_{n+1}$ , the active particles alive at time  $k_1$  w.h.p. jointly visit at least  $\lceil(1 - \epsilon)n\rceil + 1$  vertices. This completes the first part of the proof, as w.h.p.  $k_2 = \inf\{k : V_k = \lceil(1 - \epsilon)n\rceil + 1\} < \infty$  (or in other words,  $\{k : V_k = \lceil(1 - \epsilon)n\rceil + 1\} \neq \emptyset$ ).

Note that by our imposed rule for selecting which particle participates in each round, w.h.p. only particles that were activated up to instant  $k_1$  participate in rounds  $k \in (k_1, k_2] \cap \mathbb{N}$ . Therefore, the event

$$E_n := \{\text{every particle activated in rounds } k \in (k_1, k_2] \cap \mathbb{N} \text{ is alive at the end of round } k_2\}$$

holds w.h.p., since it is not possible for a particle to die without having participated in some round.

Fix any sequence of real numbers  $(x_n)_{n \in \mathbb{N}}$ . Note that the number of particles activated in rounds  $k \in (k_1, k_2] \cap \mathbb{N}$  has the same distribution as  $\sum_{i=V_{k_1}+1}^{\lceil(1-\epsilon)n\rceil+1} \eta_i$  and  $\sum_{i=V_{k_1}}^{\lceil(1-\epsilon)n\rceil} \eta_i$ . So,  $A_{k_2}$  conditioned on  $E_n$  is stochastically greater than  $\sum_{i=V_{k_1}}^{\lceil(1-\epsilon)n\rceil} \eta_i$ . Furthermore, by the definitions in (2.6) and (2.7), we have that  $V_{k_1} \leq k_1 + 1$ . Therefore, for

all  $n$ ,

$$P(A_{k_2} \geq x_n | E_n) \geq P\left(\sum_{i=V_{k_1}}^{\lceil(1-\epsilon)n\rceil} \eta_i \geq x_n\right) \geq P\left(\sum_{i=k_1+1}^{\lceil(1-\epsilon)n\rceil} \eta_i \geq x_n\right).$$

So, we conclude that

$$\begin{aligned} P(A_{k_2} \geq x_n) &\geq P(A_{k_2} \geq x_n | E_n) P(E_n) \\ &\geq P\left(\sum_{i=k_1+1}^{\lceil(1-\epsilon)n\rceil} \eta_i \geq x_n\right) [1 - P(E_n^c)] \\ &= P\left(\sum_{i=k_1+1}^{\lceil(1-\epsilon)n\rceil} \eta_i \geq x_n\right) - P\left(\sum_{i=k_1+1}^{\lceil(1-\epsilon)n\rceil} \eta_i \geq x_n\right) P(E_n^c), \end{aligned}$$

completing the proof of Lemma 3.1, as  $P(\sum_{i=k_1+1}^{\lceil(1-\epsilon)n\rceil} \eta_i \geq x_n) P(E_n^c) \leq P(E_n^c) = o(1)$ .  $\square$

*Proof of Theorem 1.2 part (i).* The proof strategy used here is similar to that used in [12, Proposition 1.1].

To be able to use the lemmas directly, we first show that  $\lim_{n \rightarrow \infty} P(V_\infty(\mathcal{K}_{n+1}, p_n) = n + 1) = 0$  and then we show that the result also holds when exchanging  $p_n$  for  $p_{n+1}$ . So, we use the graph  $\mathcal{K}_{n+1}$  and we let  $\mathcal{V} = \mathcal{V}_{n+1}$  be its corresponding set of vertices, with  $|\mathcal{V}| = n + 1$ .

The idea is to place all  $1 + \sum_{v \in \mathcal{V}} \eta_v$  original particles at the root  $o \in \mathcal{V}$ , so that every particle starts out active. In other words, we deal with another auxiliary process, but considering  $1 + \eta_o^* = 1 + \sum_{v \in \mathcal{V}} \eta_v$  and  $\eta_v^* = 0$  for all  $v \in \mathcal{V} \setminus \{o\}$ . From now on, we use the asterisk to differentiate the auxiliary process that uses  $\{\eta_v^*\}_{v \in \mathcal{V}}$  from the one that uses  $\{\eta_v\}_{v \in \mathcal{V}}$ .

We have that  $P(V_\infty^*(\mathcal{K}_{n+1}, p_n) = n + 1) \geq P(V_\infty(\mathcal{K}_{n+1}, p_n) = n + 1)$  by the structure of the complete graph, as argued below. Note that, conditioned on  $V_{j-1}' = V_{j-1}'^* = v$ , in round  $j$  we have that the probability of visiting a new vertex is  $\frac{v-1}{n} p_n$  in both processes; the probability of visiting a previously visited vertex is  $\frac{n-v+1}{n} p_n$  in both processes; the probability of the particle dying is  $1 - p_n$  in both processes. Therefore, we can couple these two models so that  $V_k' = V_k'^*$  for all  $k \in \mathbb{N}$ . Additionally, note that  $A_k'^* = 1 + \sum_{v \in \mathcal{V}} \eta_v - \sum_{i=1}^k \mathbb{1}_{(X_i^* = 0)}$  since all particles start out active and a change in the number of active particles occurs only in the case of death (compare with (2.3)); therefore  $A_k'^* \geq A_k'$  for all  $k \in \mathbb{N}$ . Then,  $R^* = \inf\{k : A_k'^* = 0\} \geq \inf\{k : A_k' = 0\} = R$ , that is, agglutinating the particles at the root at the initial instant causes the model to survive for a greater or equal number of rounds. Bringing all this information together, we conclude that  $V_\infty^* = V_{R^*}'^* \geq V_R'^* = V_R' = V_\infty$ .

Our goal is to show that, even when we accumulate all the initial particles at the root, it still holds that  $\lim_{n \rightarrow \infty} P(V_\infty^*(\mathcal{K}_{n+1}, p_n) = n + 1) = 0$ .

Let  $A_0^* = 1 + \sum_{v \in \mathcal{V}} \eta_v$  be the initial number of active particles at the root of  $\mathcal{K}_{n+1}$ . Fix  $\epsilon > 0$  such that  $(1 - \epsilon)\alpha - E(\eta) > \epsilon$ . Since  $|\mathcal{V}| = n + 1$ , we have that

$$\begin{aligned} P(A_0^* \leq (1 - \epsilon)\alpha n) &= P\left(\frac{1 + \sum_{v \in \mathcal{V}} \eta_v}{|\mathcal{V}|} \leq \frac{(1 - \epsilon)\alpha n}{n + 1}\right) \\ &= P\left(\frac{\sum_{v \in \mathcal{V}} \eta_v}{|\mathcal{V}|} - E(\eta) \leq \frac{(1 - \epsilon)\alpha n}{n + 1} - E(\eta) - \frac{1}{n + 1}\right) \end{aligned}$$

has 1 as a limit when  $n \rightarrow \infty$  by the law of large numbers, as  $\frac{(1-\epsilon)\alpha n}{n+1} - E(\eta) - \frac{1}{n+1} > \epsilon/2$  holds for sufficiently large  $n$ .

Thus, by Lemma 2.3 part (i), we know that all  $A_0^*$  particles together make w.h.p. at most  $(1 - \epsilon/2)n \log n$  steps. Now, we use the equivalence with the coupon collector's

problem and Lemma 2.1 to conclude that this number of steps w.h.p. is not enough to visit all the vertices of  $\mathcal{K}_{n+1}$ .

We proved that  $\lim_{n \rightarrow \infty} P(V_\infty(\mathcal{K}_{n+1}, p_n) = n + 1) = 0$  for all  $\alpha > E(\eta)$ . Below we show how the same would apply when using  $p_{n+1}$ , finishing the proof of Theorem 1.2 part (i).

Write  $p_n(\alpha) = 1 - \frac{\alpha}{\log(n)}$ . For any fixed  $\alpha > E(\eta)$ , select some  $\alpha^* \in (E(\eta), \alpha)$ . As  $\lim_{n \rightarrow \infty} \frac{\log(n+1)}{\log(n)} = 1$ , for a sufficiently large  $n$ , we have that

$$\frac{\alpha}{\alpha^*} \geq \frac{\log(n+1)}{\log(n)},$$

which implies that  $p_{n+1}(\alpha) = 1 - \frac{\alpha}{\log(n+1)} \leq 1 - \frac{\alpha^*}{\log(n)} = p_n(\alpha^*)$  and, due to the monotonicity of the model in  $p$ , also that  $P(V_\infty(\mathcal{K}_{n+1}, p_{n+1}(\alpha)) = n + 1) \leq P(V_\infty(\mathcal{K}_{n+1}, p_n(\alpha^*)) = n + 1)$ .

As we have already shown that part (i) of Theorem 1.2 holds for  $p_n(\alpha^*)$ , it also holds for  $p_{n+1}(\alpha)$ . □

*Proof of Theorem 1.2 part (ii).* We only focus on the case where  $E(\eta) < \infty$ . It is easy to see that proving Theorem 1.2 part (ii) in the case  $E(\eta) < \infty$  already implies proving it also for the case  $E(\eta) = \infty$ . To check this, we start with  $\eta$  such that  $E(\eta) = \infty$  and compare it with  $\eta^{(m)} := \eta \mathbb{1}_{\{\eta \leq m\}} + m \mathbb{1}_{\{\eta > m\}}$ . Since  $\lim_{m \rightarrow \infty} E(\eta^{(m)}) = E(\eta) = \infty$  by the Monotone Convergence Theorem, for any  $\alpha > 0$ , we can take a sufficiently large  $m$  such that  $\alpha < E(\eta^{(m)})$ . The comparison  $\eta \succeq \eta^{(m)}$  and this same Theorem supposedly valid for  $\eta^{(m)}$  (since  $\alpha < E(\eta^{(m)}) < \infty$ ) imply that  $P_\eta(V_\infty(\mathcal{K}_n, p_n) = n) \geq P_{\eta^{(m)}}(V_\infty(\mathcal{K}_n, p_n) = n) \xrightarrow{n \rightarrow \infty} 1$ .

Furthermore, to be able to use the lemmas directly, we focus only on the proof that  $\lim_{n \rightarrow \infty} P(V_\infty(\mathcal{K}_{n+1}, p_n) = n + 1) = 1$ . The monotonicity of the frog model in  $p$  implies that this is enough to demonstrate the same for  $p_{n+1} \geq p_n$ .

As  $\alpha < E(\eta)$ , we can fix  $\epsilon \in (0, 1)$  and  $c \in (0, 1)$  such that  $c < 1 - \epsilon$  and  $\frac{(1+\epsilon)\alpha}{1-\epsilon-c} < E(\eta)$ . Since  $(1 - \epsilon - c)n \leq \lceil (1 - \epsilon)n \rceil - \lfloor cn \rfloor + 1 \leq (1 - \epsilon - c)n + 3$ , we have that

$$\frac{(1 + \epsilon)\alpha n - E(\sum_{i=\lfloor cn \rfloor}^{\lceil (1-\epsilon)n \rceil} \eta_i)}{\lceil (1 - \epsilon)n \rceil - \lfloor cn \rfloor + 1} \leq \frac{(1 + \epsilon)\alpha}{1 - \epsilon - c} - \frac{(1 - \epsilon - c)nE(\eta)}{(1 - \epsilon - c)n + 3}. \tag{3.5}$$

By Lemma 3.1 and (3.5), if we let  $k_2 := \inf\{k : V_k(\mathcal{K}_{n+1}, p_n) = \lceil (1 - \epsilon)n \rceil + 1\}$ , then

$$\begin{aligned} P(A_{k_2} \geq (1 + \epsilon)\alpha n) &\geq P\left(\sum_{i=\lfloor cn \rfloor}^{\lceil (1-\epsilon)n \rceil} \eta_i \geq (1 + \epsilon)\alpha n\right) - o(1) \\ &\geq P\left(\frac{\sum_{i=\lfloor cn \rfloor}^{\lceil (1-\epsilon)n \rceil} \eta_i - E(\sum_{i=\lfloor cn \rfloor}^{\lceil (1-\epsilon)n \rceil} \eta_i)}{\lceil (1 - \epsilon)n \rceil - \lfloor cn \rfloor + 1} \geq \frac{(1 + \epsilon)\alpha}{1 - \epsilon - c} - \frac{(1 - \epsilon - c)nE(\eta)}{(1 - \epsilon - c)n + 3}\right) \\ &\quad - o(1). \end{aligned}$$

For the initially chosen constants  $\epsilon$  and  $c$ , we have that  $z := \frac{(1+\epsilon)\alpha}{1-\epsilon-c} - E(\eta) < 0$ . It is also true that  $\frac{(1+\epsilon)\alpha}{1-\epsilon-c} - \frac{(1-\epsilon-c)nE(\eta)}{(1-\epsilon-c)n+3} \xrightarrow{n \rightarrow \infty} z$  and therefore  $\frac{(1+\epsilon)\alpha}{1-\epsilon-c} - \frac{(1-\epsilon-c)nE(\eta)}{(1-\epsilon-c)n+3} \leq z/2 < 0$  for sufficiently large  $n$ . Then, we can apply the law of large numbers to conclude that w.h.p.  $A_{k_2}(\mathcal{K}_{n+1}, p_n) \geq (1 + \epsilon)\alpha n$ .

By Lemma 2.3 part (ii), we know that w.h.p. these  $A_{k_2}$  particles jointly take at least  $(1 + \epsilon/2)n \log n$  steps. By Lemma 2.1 and the comparison with the coupon collector's problem, this number of steps is w.h.p. enough to visit all vertices of  $\mathcal{K}_{n+1}$ . □

## References

- [1] O. S. M. Alves, E. Lebensztayn, F. P. Machado and M. Z. Martinez: Random walks systems on complete graphs. *B. Braz. Math. Soc.* **37**, (2006), 571–580. MR2284888
- [2] O. S. M. Alves, F. P. Machado and S. Popov: Phase transition for the frog model. *Electron. J. Probab.* **7**, (2002), 1–21. MR1943889
- [3] O. S. M. Alves, F. P. Machado and S. Popov: The shape theorem for the frog model. *Ann. Appl. Probab.* **12**, (2002), 533–546. MR1910638
- [4] I. Benjamini, L. R. Fontes, J. Hermon and F. P. Machado: On an epidemic model on finite graphs. *Ann. Appl. Probab.* **30**, (2020), 208–258. MR4068310
- [5] E. Beckman, N. Frank, Y. Jiang, M. Junge and S. Tang: The frog model on trees with drift. *Electron. Commun. Probab.* **24**, (2019), Paper No. 26, 10 pp. MR3962476
- [6] G. O. Carvalho and F. P. Machado: The coverage ratio of the frog model on complete graphs. *J. Stat. Phys.* **190**, (2023), no. 8, Paper No. 147, 11. MR4632181
- [7] C. Döbler, N. Gantert, T. Höfelsauer, S. Popov and F. Weidner: Recurrence and transience of frogs with drift on  $\mathbb{Z}^d$ . *Electron. J. Probab.* **23**, (2018), Paper No. 88, 23 pp. MR3858916
- [8] R. Durrett: Probability: theory and examples. Fifth edition. Cambridge Series in Statistical and Probabilistic Mathematics, 49. Cambridge University Press, Cambridge, 2019. xii+419 pp. MR3930614
- [9] L. R. Fontes, F. P. Machado and A. Sarkar: The critical probability for the frog model is not a monotonic function of the graph. *J. Appl. Probab.* **41**, (2004), no. 1, 292–298. MR2036292
- [10] A. Ghosh, S. Noren and A. Roitershtein: On the range of the transient frog model on  $\mathbb{Z}$ . *Adv. in Appl. Probab.* **49**, (2017), no. 2, 327–343. MR3668379
- [11] S. Gallo and C. Pena: Critical parameter of the frog model on homogeneous trees with geometric lifetime. *J. Stat. Phys.* **190**, (2023), no. 2, Paper No. 34, 15 pp. MR4522730
- [12] J. Hermon: Frogs on trees? *Electron. J. Probab.* **23**, (2018), Paper No. 17, 40 pp. MR3771754
- [13] C. Hoffman, T. Johnson and M. Junge: Recurrence and transience for the frog model on trees. *Ann. Probab.* **45**, (2017), no. 5, 2826–2854. MR3706732
- [14] C. Hoffman, T. Johnson and M. Junge: Cover time for the frog model on trees. *Forum Math. Sigma* **7**, (2019), Paper No. e41, 49 pp. MR4031108
- [15] C. Hoffman, T. Johnson and M. Junge: Infection spread for the frog model on trees. *Electron. J. Probab.* **24**, (2019), Paper No. 112, 29 pp. MR4017130
- [16] T. Höfelsauer and F. Weidner: The speed of frogs with drift on  $\mathbb{Z}$ . *Markov Process. Related Fields* **22**, (2016), no. 2, 379–392. MR3561142
- [17] E. Lebensztayn and M. A. Estrada: Laws of large numbers for the frog model on the complete graph. *J. Math. Phys.* **60**, (2019), no. 12, 123302, 15 pp. MR4043362
- [18] F. Machado, H. Mashurian and H. Matzinger: CLT for the proportion of infected individuals for an epidemic model on a complete graph. *Markov Process. Related Fields* **17**, (2011), no. 2, 209–224. MR2856240
- [19] S. Popov: Frogs in random environment. *J. Statist. Phys.* **102**, (2001), no. 1-2, 191–201. MR1819703
- [20] J. Rosenberg: The nonhomogeneous frog model on  $\mathbb{Z}$ . *J. Appl. Probab.* **55**, (2018), no. 4, 1093–1112. MR3899930
- [21] A. Telcs and N. C. Wormald: Branching and tree indexed random walks on fractals. *J. Appl. Probab.* **36**, (1999), no. 4, 999–1011. MR1742145

---

# Electronic Journal of Probability

## Electronic Communications in Probability

---

### Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS<sup>1</sup>)
- Easy interface (EJMS<sup>2</sup>)

### Economical model of EJP-ECP

- Non profit, sponsored by IMS<sup>3</sup>, BS<sup>4</sup>, ProjectEuclid<sup>5</sup>
- Purely electronic

### Help keep the journal free and vigorous

- Donate to the IMS open access fund<sup>6</sup> (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

---

<sup>1</sup>LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

<sup>2</sup>EJMS: Electronic Journal Management System: <https://vtex.lt/services/ejms-peer-review/>

<sup>3</sup>IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

<sup>4</sup>BS: Bernoulli Society <http://www.bernoulli-society.org/>

<sup>5</sup>Project Euclid: <https://projecteuclid.org/>

<sup>6</sup>IMS Open Access Fund: <https://imstat.org/shop/donation/>