



# Forests on wired regular trees

Gourab Ray and Ben Xiao

DTB A440, University of Victoria, V8W 2Y2, Canada.

*E-mail address:* [gourabray@uvic.ca](mailto:gourabray@uvic.ca)

*E-mail address:* [benssiao@gmail.com](mailto:benssiao@gmail.com)

**Abstract.** The Arboreal gas model on a finite graph  $G$  is the Bernoulli bond percolation on  $G$  conditioned on the event that the sampled subgraph is a forest. In this short note we study the arboreal gas on a regular tree wired at the leaves and obtain a comprehensive description of the weak limit of this model.

## 1. Introduction

Let  $G = (V, E)$  be a finite graph. A forest is a possibly disconnected graph with no cycles. A connected forest is a tree. The **Arboreal gas** model with parameter  $p \in [0, 1)$  is a probability measure on  $\{0, 1\}^E$  defined as

$$\mathbb{P}_{G,p}(f) = \frac{p^{|o(f)|}(1-p)^{|c(f)|} \mathbb{1}_{f \text{ is a forest}}}{Z_{G,p}}, \quad f \in \{0, 1\}^E, \quad (1.1)$$

where  $o(f)$  is  $\{e \in E : f(e) = 1\}$  (called **open** edges) and  $c(f) = E \setminus o(f)$  (called **closed edges**) and  $Z_{G,p}$  is the appropriate partition function. Note that if  $G$  has no cycles then this model is simply the Bernoulli bond percolation model with parameter  $p$ . As  $p \rightarrow 1$  this model converges to the uniform spanning tree on  $G$ . Although the uniform spanning tree on finite and infinite graphs has been studied in great depth, (we refer to [Lyons and Peres \(2016\)](#) and references therein for the interested reader) the Arboreal gas model is much less studied even when  $p$  is close to 1. The reason for this stems from the fact that the tools used to study uniform spanning tree, viz Wilson's algorithm and Electrical network theory is no longer applicable for any  $p < 1$ .

The goal of this note is to study the Arboreal gas model on a regular tree which we define now. Fix an integer  $d \geq 2$  and let  $T_{n,d}$  denote the ordered  $d$ -ary tree of **depth**  $n$  described as follows:  $T_{n,d}$  is a tree where all degrees are  $d$  except one vertex which has degree  $d - 1$  (called the **root vertex**, denoted  $\rho$ ) and the leaves (degree 1 vertices) and furthermore all leaves are at graph distance exactly  $n$  from  $\rho$ . Let  $T_{n,d}^w$  denote the tree  $T_{n,d}$  where all the leaves are identified into a single vertex denoted by  $\partial$  (this enforces a **wired boundary**). We emphasise here that the tree we consider is not a

---

*Received by the editors August 26th, 2021; accepted April 19th, 2022.*

*2010 Mathematics Subject Classification.* 60G99, 60K35, 60G70.

*Key words and phrases.* Arboreal gas, regular tree, Hidden Markov model.

First Author has been supported by NSERC 50311-57400.

$d$ -regular tree of depth  $n$  but a  $d$ -ary tree. We denote the vertex set of  $T_{n,d}^w$  by  $V_{n,d}$ , the edge set by  $E_{n,d}$ .

Let  $T_d$  denote the infinite  $d$ -ary tree, i.e., it is the unique infinite tree with where every vertex has degree  $d$  except one vertex, also called the root (denoted  $\rho$ ). Note that if we identify all vertices which are at distance strictly more than  $n - 1$  from the root into a single vertex and remove all the self loops then we obtain the graph  $T_{n,d}^w$ . We now state the main result of this note. We write  $\mathbb{P}_{n,d,p}$  in short for  $\mathbb{P}_{T_{n,d}^w,p}$ .

**Theorem 1.1.** *Fix  $p \in [0, 1)$ . As  $n \rightarrow \infty$ ,  $\mathbb{P}_{n,d,p}$  converges weakly to a limiting law  $\mathbb{P}_{d,p}$ . Furthermore the law  $\mathbb{P}_{d,p}$  can be described explicitly as a hidden tree indexed Markov chain model.*

We point out in above that the convergence is in the ‘local’ sense, meaning for any finite collection of edges in  $T_d$ , the joint distribution of the status of those edges converge. The explicit description of the tree indexed Markov chain model appears later in Theorem 3.2. For now let us mention that by storing the extra information of whether an open edge is connected to the leaf or not turns the model into a process with nice Markovian properties. This idea is inspired by a paper of Häggström [Häggström \(1996\)](#). This also allows an explicit description of the limiting law which is the content of the next theorem.

**Theorem 1.2.** *The limiting law  $\mathbb{P}_{d,p}$  undergoes a phase transition as follows.*

- For  $p \leq 1/d$ ,  $\mathbb{P}_{d,p}$  is an i.i.d. Bernoulli percolation on  $T_d$  with parameter  $p$ .
- If  $p \in (1/d, 1)$  then  $\mathbb{P}_{d,p}$  has infinitely many infinite components. Furthermore each infinite component is one ended and can be described as follows.
  - Suppose  $e = (u, v)$  be an edge with  $u$  being closer to the root. Conditioned on  $e$  being closed,  $v$  has a probability  $q = \frac{dp-1}{p(d-1)}$  of being in an infinite component which can be sampled as follows. Declare a child edge  $(v, v_1)$  of  $v$  to be open uniformly at random, then inductively having defined  $v_i$ , declare a child edge  $(v_i, v_{i+1})$  to be open uniformly at random. This defines an infinite path  $P$ .
  - Each vertex in such a path  $P$  has an independent  $\text{Bin}(d-1, 1/d)$  many edges attached to it which do not belong to  $P$  obtained by doing Bernoulli percolation on the  $d-1$  edges with parameter  $1/d$ . The other endpoints of these edges are attached to almost surely finite trees which are distributed as independent critical branching processes with offspring distribution given by  $\text{Bin}(d, 1/d)$ , and obtained by sampling a Bernoulli bond percolation with parameter  $1/d$ .
- Furthermore, all the finite components are also distributed as independent critical branching process with offspring distribution  $\text{Bin}(d, 1/d)$ .

We believe our results should have straightforward counterparts for a  $d$ -regular tree which we do not pursue in this article (see Remark 3.4).

The random forest model on complete graphs was studied in [Łuczak and Pittel \(1992\)](#); [Martin and Yeo \(2018\)](#) and more recently in the regular lattice  $\mathbb{Z}^d$  in [Bauerschmidt et al. \(2021b\)](#) (for  $d = 2$ ) and [Bauerschmidt et al. \(2021a\)](#) (for  $d \geq 3$ ). We refer to these papers for more references and history of the model. Let us mention that both the papers [Bauerschmidt et al. \(2021a,b\)](#) use a connection between the two point function in the Arboreal gas model and a non linear sigma model with hyperbolic target space. In contrast, our techniques are quite elementary and relies heavily on the regularity of the tree which leads to an explicit recursion. What is perhaps surprising is that this recursion can be explicitly solved. There are of course natural questions regarding the global behaviour of these forests. Some results regarding the emergence of giant components can be found, for example, in [Goel et al. \(2015\)](#). We do not pursue such avenues in this article.

The results we have are in line with the curious phenomenon also present in [Bauerschmidt et al. \(2021b,a\)](#): in the supercritical phase, the finite components behave like a critical component, in

particular their diameter and volume have polynomial tails, which is in sharp contrast with more standard models like Bernoulli percolation. However there is one observation which might only be seen in the tree or in general in a nonamenable graph. The Bernoulli percolation model with parameter  $p$  can be seen as the ‘free’ Arboreal gas model on the tree. It is well known that in the supercritical phase of a branching process, the finite components behave like a subcritical branching process and their volume has exponential tail [Athreya and Ney \(1972\)](#). This is in contrast to the behaviour in the wired tree where all finite clusters behave like a critical branching process. Comparing the behaviour of the wired and the free Arboreal gas models in a general graph remains a natural open question.

While preparing this manuscript it has come to our attention that P. Easo has been independently and simultaneously working on a similar problem [Easo \(2022\)](#), although his proof takes a different approach than ours. Our articles appeared in the arXiv on the same day.

**Acknowledgement.** We thank the anonymous referee for several valuable comments.

## 2. Recursion

Given a forest  $f = (f_e)_{e \in E_{n,d}} \in \{0, 1\}^{E_{n,d}}$ , We say  $x$  is connected to  $y$  in  $f$  there is an open path in  $f$  connecting  $x$  and  $y$  and denote this by  $x \leftrightarrow y$ . Also given an edge  $e$ , let  $e_-$  denote the vertex of  $e$  closer to the root  $\rho$  and let  $e_+$  be the vertex further to the root. The edges attached to  $e_+$  which are not  $e$  are called the **children edges** of  $e$ . The edge attached to  $e_-$  whose other endpoint strictly closer to the root is called the **parent edge** of  $e$ . The children edges of the parent edge of  $e$  except  $e$  are called the **siblings** of  $e$ . Note that every edge has  $d - 1$  siblings.

We start with a few definitions.

- If  $f$  is a forest in  $T_{n,d}^w$ , then define  $w(f) = p^{o(f)}(1 - p)^{c(f)}$  to be the weight of  $f$ .
- We define  $Z_{n,d,p}$  to be the partition function associated with the Arboreal gas measure  $\mathbb{P}_{n,d,p}$ . We write  $Z_{n,d,p} = Z_{n,d,p}^S + Z_{n,d,p}^\times$  where  $Z_{n,d,p}^S$  is the sum over the weights of all forests where  $\rho \leftrightarrow \partial$  and  $Z_{n,d,p}^\times$  is the sum over the rest.
- We define  $q_{n,d,p} = \mathbb{P}_{n,d,p}(0 \leftrightarrow \partial)$ . Note that in this notation

$$q_{n,d,p} = \frac{Z_{n,d,p}^S}{Z_{n,d,p}}; \quad 1 - q_{n,d,p} = \frac{Z_{n,d,p}^\times}{Z_{n,d,p}}.$$

We think of this quantity as the probability of ‘survival’ if we interpret the Arboreal gas as a branching process.

**Lemma 2.1.** *Fix  $d, p \in [0, 1)$ . We have*

$$Z_{n,d,p}^S = dpZ_{n-1,d,p}^S((1 - p)Z_{n-1,d,p}^S + Z_{n-1,d,p}^\times)^{d-1} \text{ and } Z_{n,d,p}^\times = ((1 - p)Z_{n-1,d,p}^S + Z_{n-1,d,p}^\times)^d.$$

*Proof:* We drop  $d, p$  from the subscript in the notation for simplicity. The only way to survive is if  $\rho$  is connected to  $\partial$  through a path which goes through one of the edges attached to the root and none of the other edges attached to the root reach  $\partial$ . The child which survives contributes a weight  $pZ_{n-1}^S$ . To take care of the other children, we need to select  $i$  many to be open but must lead to extinction and  $d - 1 - i$  many to be closed but the paths may or may not survive. This has weight

$$\sum_{i=0}^{d-1} \binom{d-1}{i} (pZ_{n-1}^\times)^i ((1 - p)Z_{n-1})^{d-1-i}.$$

Overall, we see that

$$Z_n^S = dpZ_{n-1}^S \sum_{i=0}^{d-1} \binom{d-1}{i} (pZ_{n-1}^\times)^i ((1 - p)Z_{n-1})^{d-1-i}.$$

Applying the binomial theorem yields

$$Z_n^S = dpZ_{n-1}^S((1-p)Z_{n-1} + pZ_{n-1}^\times)^{d-1}.$$

Substituting  $Z_{n-1} = Z_{n-1}^S + Z_{n-1}^\times$  and simplifying we get the desired expression as claimed.

For  $Z_n^\times$ , we need to pick  $i$  many open edges but they must lead to extinction and  $d - i$  many closed edges and they may or may not be extinct and so we get

$$\sum_{i=0}^d \binom{d}{i} (pZ_{n-1}^\times)^i ((1-p)Z_{n-1})^{d-i},$$

and so we arrive again at our desired expression after applying the Binomial theorem and simplifying. □

Our goal is to now solve the family of recursions  $Z_n^S$  and  $Z_n^\times$  with the initial conditions  $Z_0^S = 1$  and  $Z_0^\times = 0$  by convention (for  $n = 0$ , the root is connected to itself). Set

$$K_{n,d,p} = \frac{Z_{n,d,p}^\times}{Z_{n,d,p}^S} = \frac{(1-p)Z_{n-1,d,p}^S + Z_{n-1,d,p}^\times}{dpZ_{n-1,d,p}^S}.$$

where the last equality is a consequence of Lemma 2.1. Note that  $Z_{n,d,p}^S$  is 0 only when  $p = 0$  and since that case is not so interesting, we will from now on assume  $p > 0$ .

**Lemma 2.2.** *Set  $p \in (0, 1), n \geq 0$ . Then*

$$K_{n,d,p} = \frac{(1-p)}{(dp)^n} \sum_{i=0}^{n-1} (dp)^i.$$

Consequently, if  $p \neq 1/d$ , then

$$K_{n,d,p} = \frac{1-p}{dp-1} (1 - (dp)^{-n}).$$

And if  $p = \frac{1}{d}$ ,

$$K_{n,d,p} = \left(1 - \frac{1}{d}\right) n.$$

*Proof:* We drop the subscripts  $d, p$  from the notation. Using Lemma 2.1, we can write

$$K_n = \frac{(1-p)Z_{n-1}^S + Z_{n-1}^\times}{dpZ_{n-1}^S} = \frac{1-p}{dp} + \frac{Z_{n-1}^\times}{dpZ_{n-1}^S} = \frac{1-p}{dp} + \frac{1}{dp} K_{n-1}.$$

We claim now the first equality holds and we prove it by induction. Since  $K_0 = 0$  by convention and the right hand side is 0 as well (empty sum is 0), the base case holds. Assume the formula is true for  $K_n$ . Then

$$K_{n+1} = \frac{1-p}{dp} + \frac{1}{dp} K_n = \frac{1-p}{dp} + \frac{1}{dp} \frac{(1-p)}{(dp)^n} \sum_{i=0}^{n-1} (dp)^i = \frac{1}{(dp)^{n+1}} [(1-p)(dp)^n + (1-p) \sum_{i=0}^{n-1} (dp)^i],$$

and factoring out the  $1 - p$  terms gives us the desired expression. By induction, the formula holds for all  $n$ . The second equality comes from applying the geometric series formula to the sum. □

An easy consequence of Lemma 2.2 is that

**Corollary 2.3.** *If  $p \leq 1/d, q_{n,d,p} \rightarrow 0$  as  $n \rightarrow \infty$ . On the other hand if  $p \in (1/d, 1)$ ,*

$$\lim_{n \rightarrow \infty} q_{n,d,p} = \frac{dp-1}{p(d-1)} =: q_{d,p}. \tag{2.1}$$

We call  $q_{d,p}$  the **survival probability** drawing an analogue from the classical branching process. It also follows from the expression in Lemma 2.2 that if  $p < 1/d$  then  $q_{n,d,p}$  converges to 0 exponentially fast, while if  $p = 1/d$ ,  $q_{n,d,p} = d/(d-1)n^{-1}$ . This is reminiscent of similar results for branching processes and indeed resonates with Theorem 1.2 that if  $p \leq 1/d$  then the Arboreal gas converges to an i.i.d. Bernoulli( $p$ ) process.

### 3. A hidden Markov model

This section is inspired by the ideas in Häggström Häggström (1996). Given a forest  $f = (f_e)_{e \in E_{n,d}} \in \{0, 1\}^{E_{n,d}}$  in we introduce a bijective mapping  $\phi : \{0, 1\}^{E_{n,d}} \rightarrow \{0', 1', 2'\}^{E_{n,d}}$  as follows.

- If  $f_e = 0$ ,  $\phi$  always maps it to  $0'$ .
- If  $f_e = 1$ ,  $\phi$  maps it to  $2'$  if  $e_+ \leftrightarrow \partial$  in  $f$  and to  $1'$  otherwise.

Note that while  $\phi$  depends on global properties of  $f$ ,  $\phi^{-1}$  is a local map which maps  $\{1', 2'\}$  to 1 and  $0'$  to 0. Denote the pushforward of  $\mathbb{P}_{n,d,p}$  under  $\phi$  by  $\mathbb{P}'_{n,d,p}$ . We now show that this mapping converts the Arboreal gas model into a Markovian model with states  $\{0', 1', 2'\}$ . We record this observation in the following lemma. Fix an edge  $e \in E_{n,d}$  and note that removing  $e$  breaks up  $T_{n,d}$  into two components, call the component containing the root  $C_\rho = (V(C_\rho), E(C_\rho))$ . Suppose  $\xi \in \{0', 1', 2'\}^{E(C_\rho)}$  which is valid in the sense that  $\mathbb{P}'_{n,d}$  probability of  $C_\rho$  having configuration  $\xi$  is strictly positive. Number the children edges  $(e_1, e_2, \dots, e_d)$  ordered from left to right. For any element  $f' \in \{0', 1', 2'\}^{E_{n,d}}$  and  $S \subset E_{n,d}$ , let  $f'|_S$  denote the restriction of  $f'$  to  $S$ . Let  $\mathcal{S} = \{e_1, e_2, \dots, e_d\}$ . Suppose  $e_+$  is at graph distance  $k < n$  from  $\rho$ . In the following lemma, we drop  $d, p$  from the subscripts of the partition functions for the sake of brevity.

**Lemma 3.1.** *Fix  $p \in [0, 1)$  and  $e, \mathcal{S}, \xi, C_\rho$  be as above and let  $m = n - k$ . Suppose  $\mathbb{P}'_{n,d,p}(f'|_{C_\rho} = \xi) > 0$ . Then the following holds for all  $e_j$  and for all  $\mathcal{S}_1 \subseteq \mathcal{S} \setminus \{e_j\}$ .*

$$\begin{aligned} & \mathbb{P}'_{n,d,p} \left( f'_{e_j} = 2', f'_g = 1' \forall g \in \mathcal{S}_1, f'_h = 0' \forall h \in \mathcal{S} \setminus (\mathcal{S}_1 \cup \{e_j\}) \mid f'_e = 0', f'|_{C_\rho} = \xi \right) \\ &= \frac{pZ_{m-1}^S (pZ_{m-1}^\times)^{|\mathcal{S}_1|} ((1-p)Z_{m-1})^{d-|\mathcal{S}_1|-1}}{dpZ_{m-1}^S (pZ_{m-1}^\times + (1-p)Z_{m-1})^{d-1} + (pZ_{m-1}^\times + (1-p)Z_{m-1})^d}. \end{aligned} \tag{3.1}$$

Also for any  $\mathcal{S}_1 \subseteq \mathcal{S}$ ,

$$\begin{aligned} & \mathbb{P}'_{n,d,p} \left( f'_g = 1' \forall g \in \mathcal{S}_1, f'_h = 0' \forall h \in \mathcal{S} \setminus \mathcal{S}_1 \mid f'_e = 0', f'|_{C_\rho} = \xi \right) \\ &= \frac{(pZ_{m-1}^\times)^{|\mathcal{S}_1|} ((1-p)Z_{m-1})^{d-|\mathcal{S}_1|}}{dpZ_{m-1}^S (pZ_{m-1}^\times + (1-p)Z_{m-1})^{d-1} + (pZ_{m-1}^\times + (1-p)Z_{m-1})^d}. \end{aligned} \tag{3.2}$$

For any  $\mathcal{S}_1 \subseteq \mathcal{S}$ ,

$$\begin{aligned} & \mathbb{P}'_{n,d,p} \left( f'_g = 1' \forall g \in \mathcal{S}_1, f'_h = 0' \forall h \in \mathcal{S} \setminus \mathcal{S}_1 \mid f'_e = 1', f'|_{C_\rho} = \xi \right) \\ &= \frac{(pZ_{m-1}^\times)^{|\mathcal{S}_1|} ((1-p)Z_{m-1})^{d-|\mathcal{S}_1|}}{(pZ_{m-1}^\times + (1-p)Z_{m-1})^d}. \end{aligned} \tag{3.3}$$

Finally for all  $1 \leq j \leq d$  and for all  $\mathcal{S}_1 \subseteq \mathcal{S} \setminus \{e_j\}$ .

$$\begin{aligned} & \mathbb{P}'_{n,d,p} \left( f'_{e_j} = 2', f'_g = 1' \forall g \in \mathcal{S}_1, f'_h = 0' \forall h \in \mathcal{S} \setminus (\mathcal{S}_1 \cup \{e_j\}) \mid f'_e = 2', f'|_{C_\rho} = \xi \right) \\ &= \frac{pZ_{m-1}^S (pZ_{m-1}^\times)^{|\mathcal{S}_1|} ((1-p)Z_{m-1})^{d-|\mathcal{S}_1|-1}}{dpZ_{m-1}^S (pZ_{m-1}^\times + (1-p)Z_{m-1})^{d-1}}. \end{aligned} \tag{3.4}$$

In particular, all the expressions in the right hand side above are independent of  $\xi$ .

*Proof:* We only calculate the first expression since the rest follow from similar arguments. Write

$$\mathbb{P}'_{n,d,p} \left( f'_{e_j} = 2', f'_g = 1' \forall g \in \mathcal{S}_1, f'_h = 0' \forall h \in \mathcal{S} \setminus (\mathcal{S}_1 \cup \{e_j\}) \mid f'_e = 0', f'|_{C_\rho} = \xi \right) = \frac{N}{D}$$

where

$$N = \mathbb{P}'_{n,d,p} \left( f'_{e_j} = 2', f'_g = 1' \forall g \in \mathcal{S}_1, f'_h = 0' \forall h \in \mathcal{S} \setminus (\mathcal{S}_1 \cup \{e_j\}), f'_e = 0', f'|_{C_\rho} = \xi \right),$$

$$D = \mathbb{P}'_{n,d,p} \left( f'_e = 0', f'|_{C_\rho} = \xi \right).$$

Let us first determine  $D$ . Note that in any configuration with  $f'_e = 0$  and  $f'|_{C_\rho} = \xi$  must have term  $w(\xi)(1 - p)$  as a multiplicative factor. Now, for the children of  $e$  and their descendants, we have two cases based on whether or not  $e_+ \leftrightarrow \partial$ . This gives us the term

$$dpZ_{m-1}^S \sum_{i=0}^{d-1} \binom{d-1}{i} (pZ_{m-1}^\times)^i ((1-p)Z_{m-1})^{d-1-i} + \sum_{i=0}^d \binom{d}{i} (pZ_{m-1}^\times)^i ((1-p)Z_{m-1})^{d-i}.$$

Applying the binomial theorem, we get

$$D = \frac{1}{Z_n} w(\xi)(1 - p) \left[ dpZ_{m-1}^S (pZ_{m-1}^\times + (1-p)Z_{m-1})^{d-1} + (pZ_{m-1}^\times + (1-p)Z_{m-1})^d \right].$$

Now we turn to  $N$ . Note that we also must include the multiplicative factor  $(1 - p)w(\xi)$  and the only difference is that the value of  $f'$  on the children of  $e$  is fixed. The weight of the configuration on the children of  $e$  is

$$pZ_{m-1}^S (pZ_{m-1}^\times)^{|\mathcal{S}_1|} ((1-p)Z_{m-1})^{d-|\mathcal{S}_1|-1}$$

since we must have one surviving path and  $\mathcal{S}_1$  determines the edges with value  $1'$ . Therefore, the numerator is

$$\frac{1}{Z_n} (1 - p)w(\xi) \left[ pZ_{m-1}^S (pZ_{m-1}^\times)^{|\mathcal{S}_1|} ((1-p)Z_{m-1})^{d-|\mathcal{S}_1|-1} \right].$$

After simplifications, we see that  $\frac{N}{D}$  matches the expression as claimed and furthermore the expression does not depend on  $\xi$ . The remaining expressions follow from similar calculations.  $\square$

*Remark 3.2.* We point out that by definition, none of the children of  $1'$  can be a  $2'$  and exactly one child of a  $2'$  must be a  $2'$ . Also note how the conditional probabilities in Lemma 3.1 only depend on the state of  $e$  and so a Markov property holds. Furthermore, one can see that if we sum over all  $\mathcal{S}_1 \subseteq \mathcal{S}$ , the conditional probabilities sum to 1. For example, consider the numerator of eq. (3.2) when it is summed over all  $\mathcal{S}_1 \subseteq \mathcal{S}$ ,

$$\begin{aligned} & \sum_{\mathcal{S}_1 \subseteq \mathcal{S}} (pZ_{m-1}^\times)^{|\mathcal{S}_1|} ((1-p)Z_{m-1})^{d-|\mathcal{S}_1|} \\ &= \sum_{k=0}^d \binom{d}{k} (pZ_{m-1}^\times)^k ((1-p)Z_{m-1})^{d-k}, \end{aligned}$$

and by the binomial theorem, we arrive at

$$\frac{(pZ_{m-1}^\times + (1-p)Z_{m-1})^d}{dpZ_{m-1}^S (pZ_{m-1}^\times + (1-p)Z_{m-1})^{d-1} + (pZ_{m-1}^\times + (1-p)Z_{m-1})^d}.$$

Similarly, we can show eq. (3.1) sums to

$$\frac{dpZ_{m-1}^S (pZ_{m-1}^\times + (1-p)Z_{m-1})^{d-1}}{dpZ_{m-1}^S (pZ_{m-1}^\times + (1-p)Z_{m-1})^{d-1} + (pZ_{m-1}^\times + (1-p)Z_{m-1})^d}.$$

Therefore, summing over all possibilities conditioned on  $\{f_e = 0', f'|_{C_\rho}\}$  gives us 1. Similarly, one can show the same result when conditioned on  $\{f_e = 1', f'|_{C_\rho}\}$  and  $\{f_e = 2', f'|_{C_\rho}\}$ . Thus the

equations in Lemma 3.1 describe the transition probabilities of a Markov chain with state space  $\{0', 1', 2'\}$ .

We now wish to take a limit of the above expressions essentially using our computation in Corollary 2.3. Fix  $e$  so that  $e_+$  is at graph distance  $k$  from  $\rho$  for a fixed  $k$ .

**Lemma 3.3.** *Fix  $p \in [0, 1)$  and  $e, \mathcal{S}, C_\rho$  be as above and let  $q = q_{d,p}$  as in Corollary 2.3. Suppose  $\xi_n$  is a sequence of configurations with  $\mathbb{P}'_{n,d,p}(f'|_{C_\rho} = \xi_n) > 0$  for all  $n$ . Then the following holds for all  $e_j$  and for all  $\mathcal{S}_1 \subset \mathcal{S} \setminus \{e_j\}$ .*

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}'_{n,d,p} \left( f'_{e_j} = 2', f'_g = 1' \forall g \in \mathcal{S}_1, f'_h = 0' \forall h \in \mathcal{S} \setminus (\mathcal{S}_1 \cup \{e_j\}) \mid f'_e = 0', f'|_{C_\rho} = \xi_n \right) \\ = \frac{pq(p(1-q))^{|\mathcal{S}_1|} (1-p)^{d-|\mathcal{S}_1|-1}}{dpq(p(1-q) + (1-p))^{d-1} + (p(1-q) + (1-p))^d} \end{aligned} \quad (3.5)$$

Also for any  $\mathcal{S}_1 \subset \mathcal{S}$ ,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}'_{n,d,p} \left( f'_g = 1' \forall g \in \mathcal{S}_1, f'_h = 0' \forall h \in \mathcal{S} \setminus \mathcal{S}_1 \mid f'_e = 0', f'|_{C_\rho} = \xi_n \right) \\ = \frac{(p(1-q))^{|\mathcal{S}_1|} (1-p)^{d-|\mathcal{S}_1|}}{dpq(p(1-q) + (1-p))^{d-1} + (p(1-q) + (1-p))^d}. \end{aligned} \quad (3.6)$$

For any  $\mathcal{S}_1 \subset \mathcal{S}$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P}'_{n,d,p} \left( f'_g = 1' \forall g \in \mathcal{S}_1, f'_h = 0' \forall h \in \mathcal{S} \setminus \mathcal{S}_1 \mid f'_e = 1', f'|_{C_\rho} = \xi_n \right) = \frac{(p(1-q))^{|\mathcal{S}_1|} (1-p)^{d-|\mathcal{S}_1|}}{(p(1-q) + (1-p))^d}. \quad (3.7)$$

Finally for all  $e_j$  and for all  $\mathcal{S}_1 \subseteq \mathcal{S} \setminus \{j\}$  and  $p \in (1/d, 1)$ ,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}'_{n,d,p} \left( f'_{e_j} = 2', f'_g = 1' \forall g \in \mathcal{S}_1, f'_h = 0' \forall h \in \mathcal{S} \setminus (\mathcal{S}_1 \cup \{e_j\}) \mid f'_e = 2', f'|_{C_\rho} = \xi_n \right) \\ = \frac{pq(p(1-q))^{|\mathcal{S}_1|} (1-p)^{d-|\mathcal{S}_1|-1}}{dpq(p(1-q) + (1-p))^{d-1}}. \end{aligned} \quad (3.8)$$

*Proof:* For each expression in the right hand side of Lemma 3.1, we divide the numerator and the denominator by  $Z_{m-1}^d$  and use Corollary 2.3.  $\square$

This motivates us to define the following tree indexed Markov process on  $\{0', 1', 2'\}$  with parameters  $\theta, \alpha$  as follows. Assume the parent of the edges incident to  $\rho$  has value  $0'$  (to initiate the process).

- If an edge  $e$  has value  $0'$  then toss a coin with success probability  $\theta$ .
  - If success occurs we choose uniformly at random one of the children edges of  $e$  to be  $2'$ . Then for each of its siblings, decide to put value  $1'$  or  $0'$  independently with probability  $\alpha$  or  $1 - \alpha$ .
  - If failure occurs, then for each of its children, decide to put value  $1'$  or  $0'$  independently with probability  $\alpha$  or  $1 - \alpha$ .
- If an edge  $e$  has value  $2'$  choose one of its children uniformly at random and assign value  $2'$ . For its siblings, assign value  $1'$  or  $0'$  with probability  $\alpha, 1 - \alpha$ .
- If an edge has value  $1'$ , assign independently each of its children value  $1'$  or  $0'$  with probability  $\alpha$  or  $1 - \alpha$ .

Note that if  $\theta = 0$ , the above process describes an i.i.d. Bernoulli ( $\alpha$ ) percolation process since we never get a  $2'$  in the configuration. Call the probability measure induced by this Markov process by  $\mathbb{P}'_{d,p}$ .

**Theorem 3.1.** Fix  $p \in (1/d, 1)$ . The probability measures  $\mathbb{P}'_{n,d,p}$  converges to the Markov process described above with the following parameters

$$\theta = \frac{dpq}{dpq + (1 - pq)} = q, \quad \alpha = \frac{p(1 - q)}{p(1 - q) + (1 - p)} = \frac{1}{d}.$$

If  $p \leq 1/d$ ,  $\mathbb{P}'_{n,d,p}$  converges to an i.i.d. Bernoulli edge percolation process with parameter  $p$ .

*Proof:* Let  $f'_n$  denote a sample from  $\mathbb{P}'_{n,d,p}$ . We first show that  $f'_n$  converges in law to  $f'$  where  $f'$  is a tree indexed Markov process with state space  $\{0', 1', 2'\}$  whose transition probabilities are described by Lemma 3.3 (the chain starts from  $0'$  being the parent of  $\rho$ ). In particular, the joint distribution of the children given the parent is  $0'$  is described by (3.5), (3.6), that if the parent is  $1'$  is described by (3.7) and that if the parent is  $2'$  is described by (3.8). We see this by induction. The edges attached to the root has the right joint distribution from the limits given by (3.1), (3.5) and (3.2), (3.6) (applied without any conditioning on  $\xi_n$ ). Suppose we sample  $f'_n$  step by step, always sampling all the children of an edge in a particular step. Suppose we have sampled  $f'_n|_{T'} = \xi_k$  for some subtree  $T'$  containing  $k$  edges. We pick an edge  $e$  in  $T'$  whose children are yet to be sampled. Note that even conditioned on all possible extensions of  $\xi_k$  to a valid configuration  $\xi_n$  on the component of the root in  $T_{n,d} \setminus \{e\}$ , the joint distribution of all the children of  $e$  converge to that described by the Markov chain described in Lemmas 3.1 and 3.3. Therefore taking the average over all possible such  $\xi_n$  which are extensions of  $\xi_k$ , the convergence also holds since the conditional distribution does not depend on the extension. Since  $T'$  was an arbitrary subtree the limit of  $f'_n$  has the above described law.

It remains to show that  $f' \sim \mathbb{P}'_{d,p}$ . This is a straightforward verification. Looking at (3.7), we have the probability of getting  $1'$  in  $\mathcal{S}_1$  and  $0'$  in the rest is given by

$$(\alpha)^{|\mathcal{S}_1|}(1 - \alpha)^{d-|\mathcal{S}_1|} = \frac{(p(1 - q))^{|\mathcal{S}_1|}(1 - p)^{d-|\mathcal{S}_1|}}{(p(1 - q) + (1 - p))^{|\mathcal{S}_1|}},$$

which matches the description. For (3.5), the probability of producing the event is given by

$$\frac{\theta}{d} \alpha^{|\mathcal{S}_1|} (1 - \alpha)^{d-1-|\mathcal{S}_1|},$$

which matches the expression on the right hand side of (3.5) after plugging in the value. For (3.6), the probability of the event is

$$(1 - \theta) \alpha^{|\mathcal{S}_1|} (1 - \alpha)^{d-|\mathcal{S}_1|},$$

which also matches with the right hand side of (3.6). Finally the probability of the event in (3.8) in the description is given by

$$\frac{1}{d} \alpha^{|\mathcal{S}_1|} (1 - \alpha)^{d-|\mathcal{S}_1|},$$

which also matches with the right hand side of (3.8). □

As an immediate consequence, we obtain an improved version of Theorem 1.1 with an explicit description of the limiting process.

**Theorem 3.2.** The law  $\mathbb{P}_{n,d,p}$  converges in law to pushforward under  $\phi^{-1}$  of  $\mathbb{P}'_{d,p}$ .

*Proof:* Clearly  $\phi^{-1}$  is continuous in the topology of pointwise convergence, from which the result is immediate. □

*Proof of Theorem 1.2:* This is immediate from the description of Markov process  $\mathbb{P}'_{d,p}$ . Indeed if  $p \leq 1/d$ ,  $\theta = 0$ , and the process reduces to a Bernoulli ( $p$ ) percolation. On the other hand if  $p \in (1/d, 1)$ , since  $\alpha > 0$ , it is clear (for example applying an appropriate Borel–Cantelli) that there are infinitely many infinite components. Each component is one ended since  $2'$  produces exactly one

$2'$ . Furthermore the offsprings of  $1'$  which is a sibling of a  $2'$  are distributed as a branching process with parameter  $\text{Bin}(d, 1/d)$  (critical branching process).  $\square$

*Remark 3.4.* We finish with a brief comment about working with a  $d$ -regular tree rather than  $d$ -ary. Suppose an oriented edge  $(e_-, e_+)$  is fixed as the root edge and  $e_-$  is fixed as the root vertex. The  $d$ -regular tree of height  $n$  is the subgraph of the infinite  $d$ -regular tree induced by all vertices at graph distance at most  $n$  from  $e_-$ .

It is easy to see that one can decompose the partition function of the Arboreal gas model on a  $d$ -regular tree of height  $n$ , denoted  $Z_{n,d,p}^{\text{reg}}$ , as

$$Z_{n,d,p}^{\text{reg}} = p(Z_{n-1,d,p} Z_{n,d,p}^{\times} + Z_{n-1,d,p}^{\times} Z_{n,d,p}) + (1-p)Z_{n-1,d,p} Z_{n,d,p}$$

From this observation it is straightforward to obtain the correct variation of Theorem 1.1 in this context.

## References

- Athreya, K. B. and Ney, P. E. *Branching processes*. Die Grundlehren der mathematischen Wissenschaften, Band 196. Springer-Verlag, New York-Heidelberg (1972). [MR0373040](#).
- Bauerschmidt, R., Crawford, N., and Helmuth, T. Percolation transition for random forests in  $d \geq 3$ . *ArXiv Mathematics e-prints* (2021a). [arXiv: 2107.01878](#).
- Bauerschmidt, R., Crawford, N., Helmuth, T., and Swan, A. Random spanning forests and hyperbolic symmetry. *Comm. Math. Phys.*, **381** (3), 1223–1261 (2021b). [MR4218682](#).
- Easo, P. The wired arboreal gas on regular trees. *Electron. Commun. Probab.*, **27**, Paper No. 22, 10 (2022). [MR4402795](#).
- Goel, A., Khanna, S., Raghvendra, S., and Zhang, H. Connectivity in random forests and credit networks. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 2037–2048. SIAM, Philadelphia, PA (2015). [MR3451159](#).
- Häggström, O. The random-cluster model on a homogeneous tree. *Probab. Theory Related Fields*, **104** (2), 231–253 (1996). [MR1373377](#).
- Łuczak, T. and Pittel, B. Components of random forests. *Combin. Probab. Comput.*, **1** (1), 35–52 (1992). [MR1167294](#).
- Lyons, R. and Peres, Y. *Probability on trees and networks*, volume 42 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, New York (2016). ISBN 978-1-107-16015-6. [MR3616205](#).
- Martin, J. B. and Yeo, D. Critical random forests. *ALEA Lat. Am. J. Probab. Math. Stat.*, **15** (2), 913–960 (2018). [MR3845513](#).