Crossing bridges between percolation models and Bienaymé-Galton-Watson trees

Airam Blancas*, María Clara Fittipaldi†, Saraí Hernández-Torres‡

November 15, 2024

Abstract

In this survey, we explore the connections between two areas of probability: percolation theory and population genetic models. Our first goal is to highlight a construction on Galton-Watson trees, which has been described in two different ways: Bernoulli bond percolation and neutral mutations. Next, we introduce a novel connection between the Divide-and-Color percolation model and a particular multi-type Galton-Watson tree. We provide a gentle introduction to these topics while presenting an overview of the results that connect them.

Keywords: Bienaymé-Galton-Watson trees; Bernoulli percolation; multi-allelic neutral mutations.
MSC2020 subject classifications: 60J80; 92D25; 60K35; 82B43.

1 Introduction

Percolation and branching processes are two flourishing areas within probability theory that have received increased attention in recent years. Both fields present rich mathematical problems that have fostered dynamic research communities. This survey explores connections between these areas. We begin with a brief introduction to the models involved.

A percolation model is associated with dispersion in porous media, where the physical phenomenon of percolation can occur. This phenomenon refers to the passage of a fluid from the surface to the center of a material. One of the simplest yet most intriguing percolation models is Bernoulli bond percolation, defined by Broadbent and Hammersley in the 1950s [13].

In Bernoulli bond percolation, a graph G represents the underlying space. A random subgraph of G is generated by randomly deleting edges of G according to a percolation parameter $p \in [0, 1]$. The remaining edges are considered to be open (a precise construction is given in Subsection 4.1). Referring to the initial description of percolation phenomena, a liquid can advance through an open edge.

Roughly speaking, the percolation event occurs when there is an open connected component of size comparable to the underlying graph G. We refer to an open connected component as a *cluster*.

*Department of Statistics, ITAM, Mexico. Email: airam.blancas@itam.mx
†Facultad de Ciencias, UNAM, Mexico. Email: mcfittipaldi@ciencias.unam.mx
‡Instituto de Matemáticas, UNAM, Mexico. Email: saraiht@im.unam.mx

When G is an infinite graph, the percolation event corresponds to the existence of an infinite open cluster. In particular, we may look at the open cluster connected to origin 0. We denote this cluster by $\mathcal{C}_0$ and corresponds to the subset of vertices connected to 0 by open edges. In the terms described above, if a liquid advances through open edges, these are the vertices reachable from the origin.

Bernoulli percolation exhibits a phase transition between a subcritical phase and a supercritical phase, depending on the density of the porosity. In the subcritical phase, the density is low enough that the percolation event will not occur with probability one. However, as the density increases, there is a sudden shift, and the percolation event occurs with probability one.

Since the introduction of the model, some of the most important questions for the percolation research community have been closely related to conjectures about the system's behavior at the critical parameter, often originating from physics. For instance, questions about universality and conformal invariance at criticality were initially proposed in the physics literature. We discuss these questions and achievements in this area in Subsection 4.1.3.

This survey, however, takes a different direction. Its motivation is to explore and interpret percolation models in ways that relate to branching processes and population dynamics.

The study of branching processes began in the 1845 with Bienaymé [11] and was advanced in the 1873 by Galton and Watson [52]. We refer to [32] for an accout of the origins of this model. Branching processes model the random behavior of a population's genealogy as successive generations reproduce over time. The main model in this field is the Bienaymé-Galton-Watson process, valued for its simplicity, wide-ranging applications, and flexibility in incorporating variations to account for more realistic scenarios. For example, to represent populations with multiple types of individuals, one can employ the multi-type Bienaymé-Galton-Watson process (detailed in Subsection 3.2).

Another approach to incorporating evolutionary dynamics into a branching process is through the Bienaymé-Galton-Watson process with neutral mutations (discussed in Subsection 3.3). Bertoin introduced and analyzed this model in [8, 7]. In this model, mutations change the genotype of individuals without affecting their reproductive laws, which follow a standard Bienaymé-Galton-Watson process. Since mutations occur in ancestral lineages, offspring may not inherit the genetic type of their mothers (referred to as an *allele* in biological terms). Additionally, we assume the population has an infinite number of alleles, meaning each mutation event produces a different allele for a mutant, which is an individual with a new allele. This dynamic is known as neutral evolution, where reproductive fitness remains unaffected by evolutionary changes, as proposed in Kimura's model of infinite alleles [34].

The assumption of infinite alleles is prevalent in mutation models, though it may not always be the most suitable assumption. In particular, some works in the biological literature emphasize the importance of working with models that consider a finite number of alleles [37, 47, 54].

Finite allele models are particularly suitable for DNA evolution due to the inherent chemistry of DNA, which allows for a finite number of combinations of the nitrogenous bases. Some examples of this were given by Jukes and Cantor [31] and Kimura [35, 36], among others. A well-known model with these assumptions is the *parent-independent mutation* (P.I.M.) model, an extension of these models in which the type of a mutant does not depend on the type of its mother. In Subsection 3.4, we introduce a version of the P.I.M. model along a Bienaymé-Galton-Watson tree.

Initially, it may appear that the percolation theory and population genetics are unrelated. However, previous work has revealed deep relations through branching processes. These connec-

tions go both ways: they unveil large-scale properties of percolation on graphs and properties of mutations along genealogies.

The work of Bertoin is seminal in the applications of Bernoulli percolation to population genetics. In [7], he considers a (sub-)critical Bienaymé-Galton-Watson process with neutral mutations (infinite alleles model), and decompose the entire population into clusters of individuals carrying the same allele. The author studied the law of this allelic partition in terms of the distribution of the number of clone-children, and the number of mutant-children of a typical individual. In a follow-up work [8], Bertoin focused in the situation where the initial population is large and the mutation rate small. He proves that, for an appropriate regime, the process of the sizes of the allelic sub-families converges in distribution to a certain continuous-state branching process in discrete time. This last work was extended in two directions by Blancas and Rivero [12]. In the critical case, they construct the version of Bertoin's model conditioned not to be extinct. Additionally, they establish a version of the limit theorems presented in [8] when the reproduction law has an infinite variance and it is in the domain of attraction of an α-stable distribution, both for the unconditioned process and for the process conditioned to non-extinction.

On the other hand, interpreting Bernoulli percolation on a random tree as neutral mutations along the genealogical tree has proven to be a successful strategy. For instance, Berzunza [10] studied Bernoulli percolation on trees associated with the genealogy of Crump-Mode-Jagers processes, which model populations where individuals may give birth at different points in time during their lives. The main result in [10] is the existence of a giant percolation cluster for large Crump-Mode-Jagers trees. In [9], Bertoin and Uribe Bravo considered supercritical Bernoulli bond percolation on large scale-free trees, closely related to Yule processes. With the neutral mutations interpretation, the authors in [9] obtained a weak limit theorem for the sizes of the clusters.

In this survey, we focus on the connections between percolation theory and the two extensions of BGW processes presented above: the multi-type Bienaymé-Galton-Watson process and the Bienaymé-Galton-Watson process with neutral mutations.

Our starting point is the key observation in [7]: a Bienaymé-Galton-Watson process with neutral mutations and an *infinite number of alleles* is equivalent to Bernoulli percolation on a Galton-Watson tree. This type of representation serves as the main bridge between the two topics in this survey, and this note aims to explain it carefully.

As we explained above, in some situations a finite number of alleles is a more appropriate model. The corresponding bridge between a Bienaymé-Galton-Watson process with neutral mutations and a *finite number of alleles* requires a variation of Bernoulli percolation, known as Divide-and-Color percolation, introduced by Häggström in [28].

Divide-and-Color percolation is a simple and natural model for coloring the vertices of a graph. Its construction is detailed in Subsection 4.3. In this survey, we propose an novel interpretation of Divide-and-Color percolation as a model for neutral mutations along a genealogy when the number of alleles is finite.

The remainder of the survey is structured as follows. In Section 2, we provide essential definitions on trees both as a graph and as a genealogical structure. Section 3 introduces the Bienaymé-Galton-Watson (BGW) process and its variations, including the multi-type BGW process, BGW process with neutral mutations and infinite alleles, and the BGW process with neutral mutations and finite alleles. For additional background on branching processes, we recommend classical works such as Athreya and Ney [5], as well as more recent notes by Lambert [38]. In Section 4 we introduce Bernoulli percolation on finite and infinite graphs, as well as Divide-and-Color

percolation. The classical reference on Bernoulli percolation is Grimmett [26]. For the theory of percolation on trees, we also refer to [46]. Finally, in Section 5, we discuss the connections between the models introduced in the previous two sections.

2 Trees

In this section, we look at trees as mathematical objects from two perspectives. From a combinatorial viewpoint, we first define a tree as a graph, following the standard notation presented in Lyons and Peres [46]. Afterwards, we consider a tree as a genealogical structure. For further details on this latter perspective, we refer to Abraham and Delmas [3] and Duquesne and Le Gall [20].

2.1 Trees as graphs

A *graph* is a pair $G = (V, E)$, where V is the set of vertices and E is a subset of $V \times V$ called the *edge set*. We consider the edges of E to be *unoriented*, and accordingly, we use the notation $\{x, y\}$ for an element of E. If $\{x, y\} \in E$, then we call x and y *adjacent* or *neighbors*, and we write $x \sim y$. The *degree* of a vertex is the number of its neighbors. Whenever we have more than one graph under consideration, we shall distinguish the vertex and edge sets of G by writing $V(G)$ and $E(G)$. A graph G is *finite* if the cardinality $|V| < \infty$; otherwise, we say that G is an infinite graph. An infinite graph is *locally finite* is every vertex has finite degree. From now on, we assume that all infinite graphs are locally finite as well.

A *subgraph* G' of G is a graph whose vertex set is a subset of $V(G)$ and whose edge set is a subset of $E(G)$. If G' satisfies that its vertex subset is V, we refer to G' as a *spanning subgraph*.

A graph *isomorphism* between $G = (V, E)$ and $G' = (V', E')$ is a bijection $\varphi : V \to V'$ for which $\varphi(x) \sim \varphi(v)$ if and only if $x \sim v$. An isomorphism between a graph G and itself is called an automorphism. A graph is transitive if for any $u, v \in V$ there exists an automorphism $\varphi_{u,v}$ such that $\varphi_{u,v}(u) = v$.

A *path* in a graph is a sequence of vertices $[v_0, v_1, \dots,]$ where each successive pair of vertices forms an edge in the graph. A finite path connects its first and last vertices. When a path does not pass through any vertex (resp., edge) more than once, we will call it *simple*. The length of a finite path $[v_0, v_1, \dots, v_n]$ is equal the number of edges connecting each pair of vertices, so we write $\text{length}([v_0, v_1, \dots, v_n]) = n$.

A finite path with at least one edge and whose first and last vertices are the same is called a *cycle*. A cycle is called *simple* if no pair of vertices are the same except for its first and last ones. A graph is *connected* if, for each pair $x \neq y$ of its vertices, there is a path connecting x to y.

Let $G = (V, E)$ be a finite, or locally finite, connected graph. The *graph metric* $d_G : V \times V \to [0, \infty)$ is defined by

$$d_G(x, y) \coloneqq \inf\{\text{length}(\mathbf{p}) \ : \ \mathbf{p} \text{ is a simple path between } x \text{ and } y\}. \tag{1}$$

The closed ball of radius $r > 0$ with center at x is the set of vertices

$$B_G(x, r) \coloneqq \{y \in V \ : \ \ d_G(x, y) \leq r\}.$$

If G is a transitive graph, any two closed balls or radius $r > 0$ are isomorphic as graphs.

A graph with no cycles is called a *forest*, and a connected forest is a *tree*. We will work with *rooted trees*, meaning that some vertex is designated as the root, denoted by $\emptyset$. Once we set a root, we can label the tree using different partial order for its vertices. We are interested in a labeling that convey genealogical data, which we introduce in the next subsection.

2.2 Trees as genealogical structures

In this note, we will work in the following setting. We imagine that the tree is growing away from its root, and its growth specifies a population genealogy: the root has edges leading to its children, then these vertices branch out to their own children, and we advance in this way throughout the generations. The ancestry of an individual in the tree determines its label in the Ulam-Harris notation, which we introduce below.

Let us define the set of finite sequences of positive integers as

$$\mathcal{U} = \bigcup_{n \geq 0} \mathbb{N}^n,$$

following the convention $\mathbb{N}^0 = \{\emptyset\}$. For $n \geq 1$ and $v = (v_1 \dots v_n) \in \mathcal{U}$, we denote by $\|v\| = n$ the *length* of v, where the root satisfies $\|\emptyset\| = 0$. Let $u, v \in \mathcal{U}$. We say that v is an *ancestor of* the vertex u, and write $v \prec u$, if there exists $w \in \mathcal{U}$, $w \neq \emptyset$, such that we can write $u = wv$, the concatenation of the two sequences. In this case, the element w is called *the most recent common ancestor* of $\{v, u\}$. Observe that $u \preceq u$ for any $u \in \mathcal{U}$. This ancestry relation defines a lexicographic order on $\mathcal{U}$, which we will use to set the index of each vertex on a given tree, starting from the root. Within this framework, a *genealogical tree* $\mathbf{t}$ is a subset of $\mathcal{U}$ satisfying the following conditions:

1. $\emptyset \in \mathbf{t}$.
2. If $u \in \mathbf{t}$, then $A_u \coloneqq \{v \in \mathcal{U} : v \prec u\} \subset \mathbf{t}$.
3. For every $u \in \mathbf{t}$, there exists $k_u(\mathbf{t}) \in \mathbb{N}_0 \cup \{+\infty\}$ such that for every $i \in \mathbb{N}$, $ui \in \mathbf{t}$ if and only if $1 \leq i \leq k_u(\mathbf{t})$.

The set A_u corresponds to the *set of ancestors* of u, and the integer $k_u(\mathbf{t})$ indicates the number of offspring of the vertex $u \in \mathbf{t}$. See Figure 1 for an example.

For any $u \in \mathbf{t}$, we define the sub-tree $\mathbf{t}_u$ of $\mathbf{t}$ with root u as:

$$\mathbf{t}_u \coloneqq \{v \in \mathcal{U} : uv \in \mathbf{t}\}.$$

Then, the tree $\mathbf{t}_u$ corresponds to the genealogical tree of the individual u and its descendants.

We denote by $\mathbb{T}$ the set of trees without vertices of infinite degree (also known as locally finite trees):

$$\mathbb{T} = \{\mathbf{t} : k_u(\mathbf{t}) < +\infty \quad \forall u \in \mathbf{t}\}.$$

For $\mathbf{t} \in \mathbb{T}$, we denote by $|\mathbf{t}| = \operatorname{Card} \mathbf{t}$ the cardinality of its vertex set. Let $\mathbb{T}_0$ be the set of finite trees:

$$\mathbb{T}_0 = \{\mathbf{t} \in \mathbb{T} : |\mathbf{t}| < +\infty\}.$$

3 Genealogical structures for branching processes

3.1 Bienaymé-Galton-Watson process

Consider a population evolving in discrete time, where each time step denotes a generation. At each given time, every individual dies and gives birth to a random number ξ of children in the next generation, according to an *offspring distribution* $\mu = (\mu(k), k \in \mathbb{N}_0)$, so that $\mathbb{P}(\xi = k) = \mu(k)$.

For each $j \in \mathbb{N}$ and $n \in \mathbb{N}_0$, we write $\xi_{n+1,j}$ for the number of children alive in the $(n+1)$-th generation, of the j-th individual in generation n. The random variables $((\xi_{n+1,j}; j \in \mathbb{N}), n \in \mathbb{N})$ are independent and identically distributed with common distribution μ. In order to avoid trivial cases, we assume throughout that $\mu(0) + \mu(1) < 1$ and $\mu(k) \neq 1$ for any k. A Bienaymé-Galton-Watson (BGW) process with offspring distribution μ is a discrete-time Markov chain $(X_n, n \in \mathbb{N}_0)$ defined as $X_0 = 1$ and

$$X_{n+1} = \sum_{j=1}^{X_n} \xi_{n+1,j}, \qquad \forall n \in \mathbb{N}_0,$$

so that X_n is the size of the population at generation n. The law of a BGW process starts from x individuals (i.e $X_0 = x$) is denoted by $\mathbb{P}_x$.

Let $(X_n(x), n \in \mathbb{N}_0)$ be a BGW process starting from x individuals. One of the key properties of a BGW process is the *branching property*. It establishes that

$$X(x+y) \stackrel{d}{=} \bar{X}(x) + \widetilde{X}(y),$$

where $\bar{X}$ and $\widetilde{X}$ are independent copies of X starting from x and y individuals, respectively. In other terms, for each $x, y \in \mathbb{N}_0$, the probability measure $\mathbb{P}_{x+y}$ is equal to the convolution of $\mathbb{P}_x$ and $\mathbb{P}_y$.

A *BGW tree*, denoted as $\mathcal{T}$, is the genealogical tree resulting from the BGW process originating from a single individual. Therefore, $\mathcal{T}$ is a $\mathbb{T}$-valued random variable, with a root denoted by $\emptyset$ to signify the initial individual, and edges establishing connections between parents and their children. For each vertex v within $\mathcal{T}$, $k_v(\mathcal{T})$ denotes the number of offspring it has, following a distribution μ.

We follow the Ulam-Harris notation to label the genealogical tree $\mathcal{T}$. According to this notation, an individual indexed by $u = (u_1 \dots u_n u_{n+1})$ in the $(n+1)$-th generation is identified as the u_{n+1}-th child of the individual $(u_1 \dots u_n)$ in the n-th generation. This labeling allows us to trace backward in time along the genealogy until reaching the u_1-th individual in the first generation, who is a child of the root denoted by $\emptyset$.

As a consequence of the definition of the BGW process, the subtree $\mathcal{T}u$, rooted in $u \in \mathcal{T}$, is a BGW tree with the same distribution as $\mathcal{T}$. It is worth noting that $\mathcal{T}u$ represents the genealogical tree associated with a BGW process $(\widetilde{X}_n, n \geq 0)$ starting from the individual u.

By the branching property, conditioned on the event $\{k_\emptyset(\mathcal{T}) = n\}$, the sub-trees $(\mathcal{T}^{(1)}, \dots, \mathcal{T}^{(n)})$ are independent and follow the distribution of the original tree $\mathcal{T}$.

3.2 Multi-type BGW processes

We now aim to describe a branching population with d types of individuals; these types are indexed by the set $[d] := \{1, 2 \dots, d\}$. For every $i \in [d]$, individuals of type i will produce offspring of

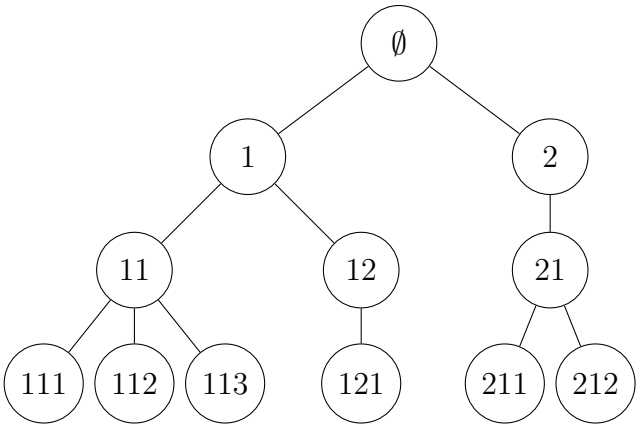

Figure 1: This figure shows an example of a BGW tree with Ulam-Harris labeling. In this instance, the offspring distribution is defined as $\mu(k) = \frac{1}{3}\mathbb{1}_{\{1,2,3\}}(k)$.

possibly each type, independently of each other. In order to avoid trivial cases, we assume that the event that all the children of an individual inherit its type occurs with probability less than one. Formally, let $\boldsymbol{\xi}^{(i)} = (\xi^{(i,j)})_{j\in[d]}$ denote the random vector where the j-th entry is the number of children that an individual of type i gives birth. Its distribution is μ_i, that is

$$\mu_i(\mathbf{v}) = \mathbb{P}(\xi^{(i,1)} = v_1, \xi^{(i,2)} = v_2, \dots, \xi^{(i,d)} = v_d), \qquad \mathbf{v} \in \mathbb{N}_0^d. \tag{2}$$

We set $\boldsymbol{\xi}^{(i)}_{(n+1,k)} = (\xi^{(i,1)}_{n+1,k}, \xi^{(i,2)}_{n+1,k}, \dots, \xi^{(i,d)}_{n+1,k})$ for the number of children in the $(n+1)$-th generation beget by the k-th individual of type i in generation n. Let us assume that, for each $i \in [d]$, $((\boldsymbol{\xi}^{(i)}_{n+1,k}, k \in \mathbb{N}), n \in \mathbb{N}_{\mathbb{0}})$ are independent random vectors with common distribution μ_i. In other words, $\mu_i(\mathbf{v})$ is the probability that such individual gives birth $|\mathbf{v}| = \sum_{j=1}^d v_j$ children, of which v_j are of type j.

Let $Y_n^{(i)}$ denote the number of individuals of type i in the n-th generation. Without loss of generality, we assume that the population starts with one individual of type i. Thus, $(\mathbf{Y}_n = (Y_n^{(1)}, \dots, Y_n^{(d)}), n \in \mathbb{N}_0)$, with law $\mathbb{P}$, is a discrete-time Markov chain recursively defined as $\mathbf{Y}_0 = \mathbf{e}_i$, where $\mathbf{e}_i$ denotes the canonical vector of $\mathbb{N}_0^d$, and

$$Y_{n+1}^{(j)} = \sum_{i=1}^{d} \sum_{k=1}^{Y_n^{(i)}} \xi^{(i,j)}_{n+1,k}, \qquad i \in [d], \quad n \in \mathbb{N}_0. \tag{3}$$

The process $(\mathbf{Y}_n, n \in \mathbb{N}_0)$ is known as a *multi-type BGW* with offspring distribution $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d)$. It is clear that when $d = 1$ the process coincides with the BGW process.

As in the one-dimensional case, the multitype BGW process $\mathbf{Y}$ satisfies the **branching property**, i.e.

$$\mathbf{Y}(\mathbf{u} + \mathbf{v}) \stackrel{d}{=} \bar{\mathbf{Y}}(\mathbf{u}) + \widetilde{\mathbf{Y}}(\mathbf{v}),$$

where $\bar{\mathbf{Y}}$ and $\widetilde{\mathbf{Y}}$ are independent copies of $\mathbf{Y}$, starting from $\mathbf{u} = (u_1, \dots, u_d)$ and $\mathbf{v} = (v_1, \dots, v_d)$ individuals respectively. As before, this implies that the probability measure $\mathbb{P}_{\mathbf{u}+\mathbf{v}}$ is equal to the convolution of $\mathbb{P}_{\mathbf{u}}$ and $\mathbb{P}_{\mathbf{u}}$ for any $\mathbf{u}, \mathbf{v} \in \mathbb{N}_0^d$.

A *multitype BGW tree* $\boldsymbol{\mathcal{T}}^i$ is the genealogical tree associated with the multitype BGW process $(\mathbf{Y}_n, n \in \mathbb{N})$ starting from an individual of type i. To include the individual type information, we label $\boldsymbol{\mathcal{T}}^i$ according to a generalization of the Ulam-Harris notation. We identify every vertex with a pair (i, u), so that the first coordinate $i \in [d]$ indicates the type and $u \in \mathbb{U}$ as before. We point out that the multitype BGW tree hold the branching property in a similar manner as in the BGW with one type.

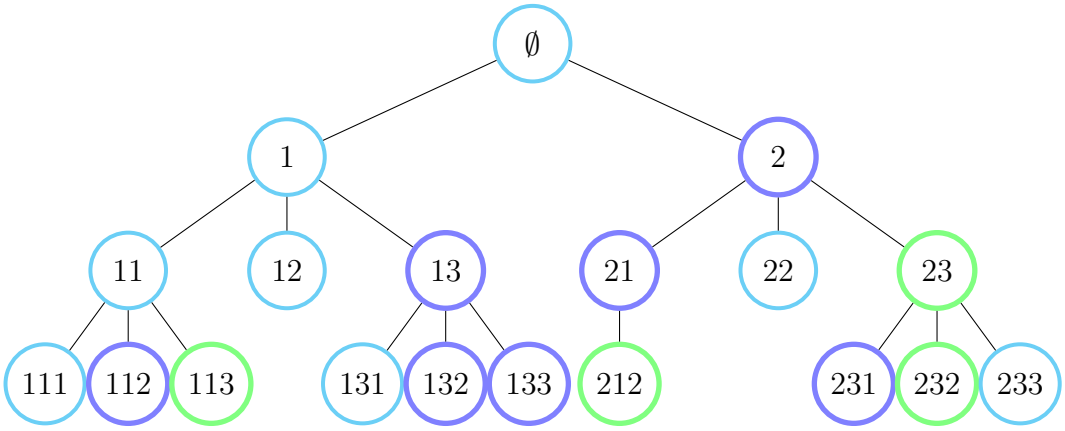

Figure 2: Example of a genealogical tree of our multi-type process. In this case we label the tree using the usual Ulam-Harris notation, but we include a node shape that represent the *type* of each individual, let's say cyan for type 1, blue for type 2 and green for type 3.

3.3 BGW processes with neutral mutations and infinite alleles

As a further extension of the BGW model, Bertoin [7, 8], studied the Galton-Watson process with neutral mutations, that we will call *BGW process with infinite neutral mutations*. This model emerges assuming that the mutations modify the genotype of individuals but not the dynamic of the population, which evolves as the BGW described in Section 3.1. Since mutations appear in the ancestral lines of the population, each individual begets children that do not necessarily inherit its genetic type. From the biological point of view, a genetic type corresponds to an allele. In addition, we suppose that the population has infinite alleles, that is, each mutation event originates a different type. Following the terminology of [8], the children of individuals sharing the same type as its mother will be called *clones*, whereas the children who are of a different type than their mother will be called *mutants*. Additionally, $\xi^{(c)}$ and $\xi^{(m)}$ are non-negative integer-valued random variables, not necessarily independent of each other, which describe respectively the number of clones and the number of mutants of a given individual with distribution $\mu^{(c)}$ and $\mu^{(m)}$, respectively. Hence, any individual (no matter if is a mutant or not) has a random number $\xi^{(+)} := \xi^{(c)} + \xi^{(m)}$ of children, independently of all the others individuals, with distribution μ. In particular, we are interested in the case where mutations affect each child according to a fixed probability r and independently of its other siblings. More precisely, the conditional distribution of $\xi^{(m)}$ given $\{\xi^{(+)} = v_1 + v_2\}$ is Binomial with parameter $(v_1 + v_2, r)$, i.e.

$$\mathbb{P}(\xi^{(c)} = v_1, \xi^{(m)} = v_2) = \binom{v_1 + v_2}{v_1} (1-r)^{v_1} r^{v_2}. \tag{4}$$

It is possible to view the BGW with infinite neutral mutations as a multitype BGW process with countably many types. To be precise, consider a population with types given by $\mathbb{N}_0$, so that every time that individuals of type i reproduces gives birth to individuals either of type i (clones) or type $i+j$, for some $j \in \mathbb{N}$ in a way that the type $i+j$ has never been seen in the population (mutants). Let $[d_n]$ be the set of types observed up to generation n. Hence, if $Y_n^{(i)}$ denotes the number of type i individual in the n-generation, we have that

$$\begin{cases} Y_{n+1}^{(i)} = \sum_{h=1}^{Y_n^{(i)}} \xi_{n+1,h}^{(c)} & \text{for any } i \in [d_n], \\ \\ Y_{n+1}^{(d_n+i(h))} = \xi_{n+1,h}^{(m)} & h = 1, \dots, \sum_{i=0}^{d_n} Y_n^{(i)} \text{ if } \xi_{n+1,h}^{(m)} \neq 0, \end{cases}$$

where $\{(\xi_{n+1,h}^{(c)}, h \in \mathbb{N}), n \in \mathbb{N}\}$ are independent copies of $\xi^{(c)}$ and $i(h) = \sum_{k=1}^{h} \mathbb{1}_{\{\xi_{n+1,k}^{(m)} \neq 0\}}$.

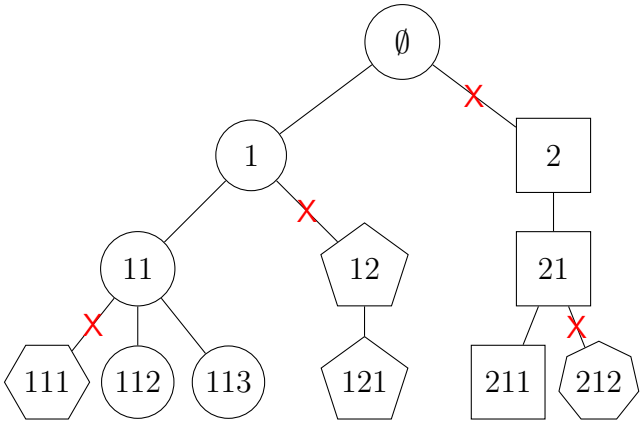

Figure 3: An example of a BGW tree with infinite allele type mutation. This tree corresponds to the tree in Figure 1 with $r = 0.5$. Note that after a new mutation, appears a new type of individual (never seen before).

3.4 BGW processes with neutral mutations and finite alleles

We are also interested in studying BGW processes with neutral mutation in the case of finitely-many alleles. In this framework, we propose two models of multi-allelic neutral mutation for a BGW process. First, we propose a mother-dependent model, where a mutant cannot be of the same type of its mother. We drop this assumption in the second model.

3.4.1 Mother-dependent mutation model

In the same framework of [8], we consider neutral mutations in a BGW process with probability r but assuming that the population has finitely-many alleles. Formally, we can view this model as a multitype BGW process, introduced in Section 3.2. In particular, we have the discrete-time Markov chain $(\mathbf{Y}_t, t \geq 0)$ as in (3) where for $i \in [d]$, the offspring distribution μ_i defined in (2) is

multinomial, that is

$$\begin{aligned}\mu_i(\mathbf{v}) &= \mu(|\mathbf{v}|)\binom{|\mathbf{v}|}{v_1,\ldots,v_d}(1-r)^{v_i}\prod_{j\neq i}\left(\frac{r}{d-1}\right)^{v_j} \\ &= \mu(|\mathbf{v}|)\binom{|\mathbf{v}|}{v_i}(1-r)^{v_i}r^{|\mathbf{v}|-v_i}\binom{|\mathbf{v}|}{v_1,\ldots,v_{i-1},v_{i+1},\ldots,v_d}\left(\frac{1}{d-1}\right)^{|\mathbf{v}|-v_i},\end{aligned} \tag{5}$$

where $r \in [0,1]$, $|\mathbf{v}| = \sum_{j=1}^{d} v_j$. Here μ is the probability measure in $\mathbb{N}_0$ associated with

$$\xi^{(+)} \overset{d}{=} \xi^{(+,i)} = \sum_{j=1}^{d} \xi^{(i,j)}, \tag{6}$$

the total number of children of any individual, independently of its type.

Hence, the above described evolution correspond to the following mechanism: at each generation, every individual reproduces with probability $(\mu(k), k \in \mathbb{N}_0)$, independently of others. Each child will be a copy of its mother with probability $1-r$, independently of their siblings. If it is not, its new type will be chosen uniformly from among the $d-1$ remaining types.

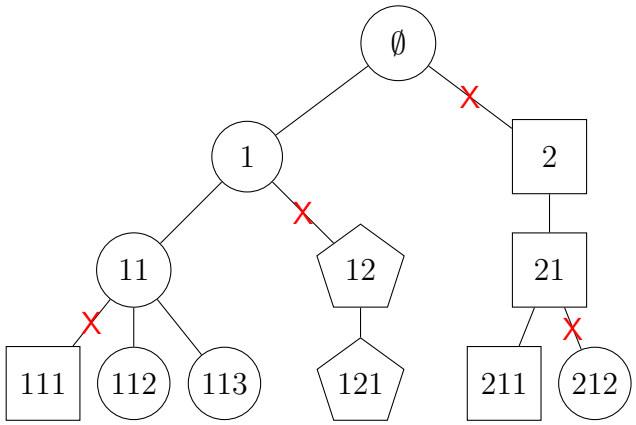

Figure 4: An example of a BGW tree with mother-dependent finite allele mutation. This tree correspond to the tree in Figure 1 with $r = 0.5$ and three types (circle, square and pentagon). Note that in this case, each mutation event produces an individual that chooses its type uniformly among all types, but different from its mother's.

3.4.2 Mother-independent mutation model

In this case, we allow mutant children to have the same type of its mother. Similarly to the mother-dependent mutation (M.D.M.) model, we have the discrete-time Markov chain $(\mathbf{Z}_t, t \geq 0)$ as in (3) where for $i \in [d]$, the offspring distribution μ_i defined in (2) is multinomial, that is

$$\mu_i(\mathbf{v}) = \mu(|\mathbf{v}|)\sum_{k=1}^{v_i}\binom{|\mathbf{v}|}{k}(1-r)^{k}r^{|\mathbf{v}|-k}\binom{|\mathbf{v}|-k}{v_1,\ldots,v_i-k,\ldots,v_d}\left(\frac{1}{d}\right)^{|\mathbf{v}|-k}. \tag{7}$$

Here, the first terms corresponds to offspring distribution in (6). The index k corresponds to the number of clones of type i. The last two factors correspond to our coloring, where $v_i - k$ individuals are mutant children that randomly chose the same type of its mother.

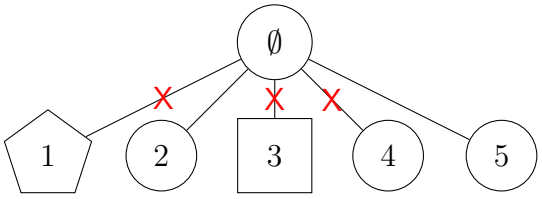

Figure 5: An illustration of a branching event for the mother-independent mutation (M.I.M.) model in the case where we have three types (circle, square and pentagon). For this model, each mutant child chooses its type uniformly from the set of all possible types.

3.5 Phase Transition for branching processes

3.5.1 Extinction for BGW processes

Given a BGW process $(X_n, n \in \mathbb{N}_0)$, *extinction* occurs if $X_n = 0$ for some generation n. Note that if $X_n = 0$, then $X_{n+k} = 0$ for every $k \in \mathbb{N}$. We denote this event by $\{\text{Ext}\}$, and thus,

$$\{\text{Ext}\} = \{ \text{ there exist } n \in \mathbb{N} : X_n = 0\} = \bigcup_{n\in\mathbb{N}} \{X_n = 0\} = \lim_{n\to+\infty} \{X_n = 0\},$$

as $(\{X_n = 0\})_{n\in\mathbb{N}}$ are increasing events. We point out that the extinction event in the BGW process corresponds to the event that $\mathcal{T}$ is a finite tree, i.e. $\{\text{Ext}\} = \{\mathcal{T} \in \mathbb{T}_0\}$.

The generating function of ξ, is given by

$$f(s) = \sum_{k\in\mathbb{N}_0} s^k \mu(k), \qquad s \in [0, 1]$$

and its mean is $\mathbb{E}(\xi) = m$. It is well-known that the fixed point equation $f(s) = s$ has at most two solutions in $[0, 1]$, and the extinction probability $q := \mathbb{P}(\text{Ext})$ is the smallest root of this equation. In fact, BGW processes display a phase transition, from almost surely extinction ($q = 1$) if $m < 1$ to positive probability of survival ($q < 1$) if $m > 1$. We say then that a BGW process $(X_n, n \in \mathbb{N}_0)$ (resp. a BGW tree $\mathcal{T}$) is *sub-critical*, *critical* or *super-critical* depending on $m < 1$, $m = 1$ or $m > 1$, respectively.

The extinction conditions for a multi-type BGW are written in terms of the matrix recording the expected number of type j offspring of a single type i individual in one generation. The precise condition is in [5, Theorem 2, Chapter V].

Analogous to the one dimensional case, the extinction conditions for a multi-type BGW are given in terms of a parameter ρ, which is the maximum eigenvalue of the mean matrix $\mathbf{M} = (m_{ij})_{i,j\in[d]}$, given by $m_{ij} = \mathbb{E}(\xi^{(i,j)})$. Given $\mathbf{M}$, a strictly positive matrix, if $\rho \leq 1$ then $\mathbb{P}(Y_t^{(i)} = 0 \text{ for some } t) = 1$ for all $i \in [d]$. If $\rho > 1$ then $\mathbb{P}(Y_t^{(i)} = 0 \text{ for some } t) < 1$ for all $i \in [d]$. The precise results (for the discrete-time case) can be found in [5, Chapter V.3].

3.5.2 Phase transition for BGW process with neutral mutation

For the BGW process with neutral mutations, the process of the total number of individuals per generation depends on the law of the underlying branching process, independently if the model has finitely-many or infinite alleles. In fact, the underlying BGW tree $\mathcal{T}$ will be finite almost surely if $m \leq 1$, where

$$m = \mathbb{E}\left[\xi^+\right], \tag{8}$$

where ξ^+ is given by (6). If $m > 1$, $\mathcal{T}$ has positive probability to be infinite. In this case, the relevant question is whether there exist an infinite subtree of individuals of the same type, and if such a subtree includes the root. We will address these matters later, in Subsection 5.2.

4 A glance at Bernoulli percolation

In this section, we will present a quick overview of Bernoulli percolation.

Let us recall our graph notations introduced in Section 2.1 and let $G = (V, E)$ be a finite graph. In the sequel, it will be helpful to consider an spanning subgraph G' of G as an element of $\mathcal{S} \coloneqq \{0,1\}^E = \{s : E \to \{0,1\}\}$. Observe that each $s \in \mathcal{S}$ is uniquely identified by its subset of edges $E_s \coloneqq \{e \in E \,:\, s(e) = 1\}$. We will refer to the edges on E_s as *open* on the configuration defined by s. Hence s is in one-to-one correspondence with the subgraph (V, E_s). With this correspondence in mind, it is convenient to refer to subgraphs as elements of $\mathcal{S}$. We will describe a random spanning subgraph as a probability distribution over the measureable space $(\mathcal{S}, \mathcal{P}(\mathcal{S}))$, where $\mathcal{P}(\mathcal{S})$ is the power set of $\mathcal{S}$.

4.1 Bernoulli percolation

Bernoulli bond percolation on $G = (V, E)$ with parameter $p \in [0,1]$ is defined by the probability measure $\mathbf{P}_p^G$ on the measurable space $(\mathcal{S}, \mathcal{P}(\mathcal{S}))$ given by

$$\mathbf{P}_p^G(s) = \prod_{e \in E} p^{s(e)}(1-p)^{1-s(e)}, \qquad s \in \mathcal{S}. \tag{9}$$

By the correspondence established above between $\mathcal{S}$ and spanning subgraphs of G, we can define a random spanning subgraph $\mathcal{O}$ so that its edge set is E_s with probability $\mathbf{P}_p^G(s)$. In words, the subgraph $\mathcal{O}$ is given by the following rule for each edge $e \in E$:

the edge e is open with probability p and
the edge e is closed with probability $1 - p$,

independently from other edges. We thus have that $E_s = \{e \in E \,:\, e \text{ is open}\}$. See Figure 6. One may think that, for each edge $e \in E$ the marginal of the measure $\mathbf{P}_p^G$ follows a Bernoulli distribution with parameter p.

From the perspective of percolation theory, we are interested in the geometry of the random spanning subgraph $\mathcal{O}$. On a discrete setting, questions about the geometry of a subgraph first concern connectivity properties, including the number of connected components and their size (i.e. their number of vertices). A subgraph $G' = (V', E')$ is *connected* if for any $x, y \in V'$ there exists a path of vertices in G' between x and y. Given a percolation configuration $\mathcal{O}$, a (random) connected subgraph $\mathcal{C}$ of open edges is called a *cluster*. See Figure 6.

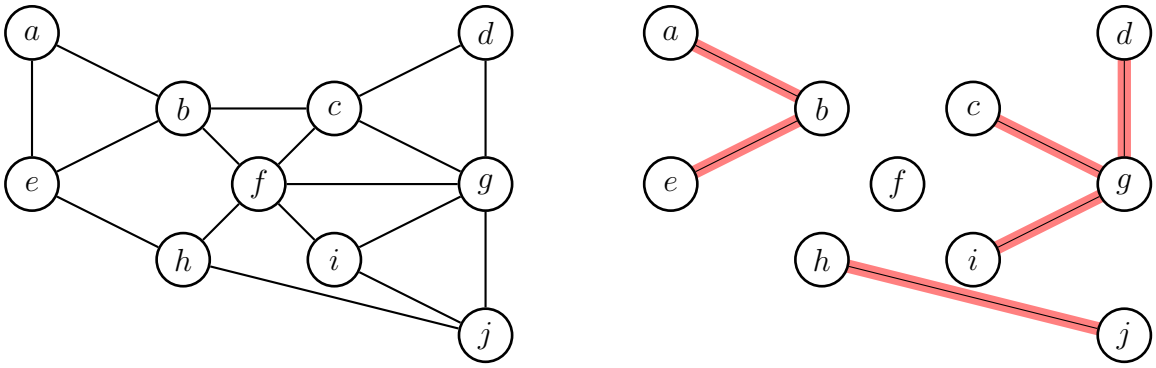

Figure 6: On the left-hand side we have a finite connected graph G. The right-hand side figure illustrates a realization of a Bernoulli percolation with parameter p. The red edges corresponds to the set of open edges E_s. The clusters of E_s are $\{a, b, e\}$, $\{c, d, g, i\}$, $\{f\}$ and $\{h, j\}$. The closed edges of G were deleted.

4.1.1 Phase transition for Bernoulli percolation on finite graphs

A remarkable property of Bernoulli percolation on a graph G is that it undergoes a phase transition of the size of the largest cluster, as the parameter p varies. A cluster is called *giant* if it contains a positive proportion of the vertices in G. The *supercritical* phase of Bernoulli percolation corresponds to the existence of a giant cluster with probability greater than a given constant. We give a quantitative definition below.

Fix $r \in (0, 1)$ and $\varepsilon > 0$. A cluster K of $\mathcal{O}$ is r-giant if

$$|K| \geq r|V|,$$

where $|\cdot|$ denotes the cardinality of the vertex set of K.

Let $(K_n)_{n\geq 1}$ be the clusters of $\mathcal{O}$ ranked in decreasing size order, so that K_1 be the largest cluster. With this in mind, for each $\varepsilon > 0$, the percolation parameter p is ε-supercritical for G if $|V| \geq 2\varepsilon^{-3}$ and

$$\mathbf{P}^G_{(1-\varepsilon)p}\left(|K_1(G)| \geq n\varepsilon\right) \geq \varepsilon.$$

For an asymptotic definition of the phase transition in percolation on finite graphs, let us consider a sequence of finite graphs $(G_n)_{n\geq 1} = ((V_n, E_n))_{n\geq 1}$ such that $|V_n| \to \infty$. The sequence of parameters $(p_n)_{n\geq 1}$ is *supercritical* if there exists $\varepsilon > 0$ and $N < \infty$ such that

$$\mathbf{P}^{G_n}_{(1-\varepsilon)p_n}(K(G_n) \geq \varepsilon|V_n|) \geq \varepsilon \qquad \text{for every } n \geq N.$$

This means that, on the supercritical regime we have a giant component asymptotically almost surely.

4.1.2 Bernoulli percolation on infinite graphs

The theory of phase transitions in Bernoulli percolation was originally formulated and developed for infinite graphs. Here, we briefly review this setting and refer to [26] for a detailed account of the subject.

Let $G = (V, E)$ be a infinite, connected and locally finite graph, and consider its space of subgraphs $\mathcal{S} = \{0, 1\}^E$. In this space, a *cylinder event* B is defined by a finite set $F = \{e_1, \dots, e_j\} \subset E$ and values $a_1, \dots, a_j \in \{0, 1\}$ such that

$$B = \{(s_e)_{e \in E} \in \mathcal{S} \,:\, (s_{e_1}, \dots, s_{e_j}) = (a_1, \dots, a_j)\}.$$

We endow $\mathcal{S}$ with the product σ-algebra $\mathcal{F}$, generated by cylinder events. The Bernoulli bond percolation measure with parameter $p \in [0, 1]$ on G is defined by

$$\mathbf{P}_p^G(B) = \prod_{i=1}^{j} p^{a_i}(1-p)^{1-a_i}, \tag{10}$$

for any cylinder event B as above.

For a transitive, connected, infinite graph G, we define the phase transition by the existence of an infinite open cluster. The *percolation density* $\theta : [0, 1] \to [0, 1]$ is the function defined by the probability that $\mathcal{C}_\rho$, the cluster containing a fixed vertex ρ, is infinite:

$$\theta(p) = \mathbf{P}_p^G\left(|\mathcal{C}_\rho| = \infty\right), \qquad p \in [0, 1],$$

where $\mathbf{P}_p^G$ is the Bernoulli percolation measure on G with parameter p. The *percolation critical parameter* on G is defined by

$$p_c^G = \inf\{p \in [0, 1] \,:\, \theta(p) > 0\}.$$

4.1.3 Critical phenomena, universality, and locality

In this subsection, we overview two fundamental concepts driving research on Bernoulli percolation: *universality* and *locality*. This part of the survey is independent of the rest of the work and provides perspective on the research landscape in this area.

Universality refers to the shared behavior of various physical systems at and near critical. When two systems share this behaviour, we say that they belong to the same *universality class*. One instance where this behavior appears is in the decay of certain quantities associated with the system, where systems in the same universality class exhibit common exponents for the power-law decay of these quantities as they approach a critical point. For example, in the case of Bernoulli bond percolation on an transitive, connected, infinite graph G with critical parameter p_c, it is expected that

$$\theta(p) = (p - p_c)^{\beta + o(1)}, \qquad \text{as } p \downarrow p_c. \tag{11}$$

The exponent β, common to all systems in the same universality class, is referred to as a *universal critical exponent*. It is believed that the critical exponent β for Bernoulli percolation depends only on the spatial dimension of the underlying graph.

On transitive graphs, the concept of dimension is defined through the *volume growth rate*, determined by the asymptotic behavior of the volume of closed balls in the graph metric (defined above in (1)) as their radius $r \to \infty$. For instance, in the lattice graph $\mathbb{Z}^d$, there exist constants $c, C > 0$ such that

$$cn^d \le |B_{\mathbb{Z}^d}(0, r)| \le Cn^d, \qquad \text{for all } n \ge 1.$$

Thus, $\mathbb{Z}^d$ has a polynomial growth rate, with volume growth dimension d.

In the 1970s, *critical phenomena* was enlightened in physics through Wilson's work on the renormalization group, for which he received the 1982 Nobel Prize in Physics [53]. One consequence of this formalism is the conformal invariance in the scaling limit of critical two-dimensional statistical models such as Bernoulli percolation [1, 2].

A deeper understanding of conformally invariant processes in the complex plane emerged with Schramm's introduction of the Schramm-Loewner evolution (SLE), a family of stochastic processes on the plane indexed by a parameter $\kappa > 0$ [49]. In particular, SLE with parameter $\kappa = 6$ describes the scaling limit of two-dimensional critical Bernoulli percolation. With this tool, Lawler, Werner and Schramm had a fruitful collaboration on the properties of SLE and its application as a tool to understand critical phenomena, as well as the fractal geometry of two-dimensional Brownian motion [39, 40, 41, 42, 43, 44]. For these results, Werner received the Fields Medal in 2006 [48].

A rigorous proof of the conformal invariance of the scaling limit of two-dimensional critical Bernoulli percolation was achieved by Smirnov, who proved the conjectured conformal invariance for the triangular lattice [50]. This landmark work also led to Smirnov's receipt of a Fields Medal in 2010 [33]. A summary of the results by Lawler, Werner, Schramm, and Smirnov, along with Kesten's scaling relations, appears in [51]. In this work, Smirnov and Werner establish the existence and values of universal critical exponents for Bernoulli percolation on the triangular lattice, confirming $\beta = 5/36$, as predicted in the physics literature.

Recently, percolation theory has extended the techniques of Bernoulli percolation to a larger class of percolation models, allowing for dependencies. One significant model in this class is the *random-cluster model* (also called FK percolation), introduced by Fortuin and Kasteleyn in the 1970s [24, 22, 23]. It generalizes Bernoulli percolation and provides a unifying framework for models such as the Ising and Potts models, as well as Bernoulli percolation itself. Taking limits on the parameters, the random-cluster model includes the uniform spanning tree and the β-arboreal gas models [30]. For a comprehensive treatment, see [27].

Duminil-Copin and collaborators have achieved significant progress in understanding the phase transitions and the critical and near-critical behavior of the random-cluster model [4, 6, 14, 16, 17, 18]. This work has also led to new insights into Bernoulli percolation, simplifying the proofs of key results, such as a proof for the sharpness of the phase transition [19]. Duminil-Copin's work earned him a Fields Medal in 2022 [29].

In Bernoulli percolation, several questions remain open. One of the biggest challenges is to understand the phase transition and critical behavior in three dimensions, such as on $\mathbb{Z}^3$. Another rich area of research explores Bernoulli percolation on graphs beyond $\mathbb{Z}^d$, where this model has generated deep connections with other mathematical fields, including geometric group theory.

Studying Bernoulli percolation on diverse graphs has highlighted different properties of the model. Since the early days of percolation theory, numerical analysis showed that p_c^G is not universal; rather, it depends on the graph G. In fact, for $p_c < 1$, the value depends on the graph's *local* geometry. It is a remarkable phenomenon, since locality is a opposite concept to universality.

The local geometry of a graph is measured by the *local topology* (also known as the *Benjamini-Schramm topology*). Let $(G_n)_{n\geq 1}$ be a sequence of connected infinite transitive graphs converging to a connected infinite transitive graph G in the local topology. This means that, for each $r > 0$, the closed balls of radius r in G_n are isomorphic to those in G for all sufficiently large n.

In the space of connected infinite transitive graphs, let $(G_n)_{n\geq 1}$ be a sequence converging to a graph G in the local topology as $n \to \infty$. If $\sup_{n\geq 1} p_c(G_n) < 1$, Schramm conjectured that $p_c(G_n) \to p_c(G)$ as $n \to \infty$.

A key step towards understanding this conjecture was achieved in [15], where the authors showed that the condition $p_c(G_n) = 1$ is equivalent to G_n being one-dimensional, meaning that its volume growth satisfies $|B_{G_n}(v, r)| = O(r)$ as $r \to \infty$.

Given [15], it is sufficient to assume that each element in the sequence $(G_n)_{n\geq 1}$ has a superlinear growth rate. Schramm's locality conjecture was recently settled in a breakthrough work by Hutchcroft and Easo. They proved that, given a sequence of connected infinite transitive graphs $(G_n)_{n\geq 1}$ with superlinear volume growth, if $G_n \to G$ as $n \to \infty$ in the local topology, then $p_c(G_n) \to p_c(G)$ as $n \to \infty$ [21].

4.2 Bernoulli percolation on trees

We now turn to a specific class of graphs and begin with a brief account of Bernoulli percolation on infinite trees.

Proposition 4.1. Let T_d be a d-ary tree with $d \geq 2$ and consider Bernoulli percolation of parameter $p \in [0, 1]$ on T_d. If $\mathcal{T}_\emptyset$ is the open cluster of the root then $\mathcal{T}_\emptyset$ follows the distribution of a BGW tree with progeny distribution $\mathrm{Bin}(d, p)$.

A direct consequence of Proposition 4.1 and the phase transition on the survival of a BGW tree (as explained in Subsection 3.5.1) we have the following.

Proposition 4.2. Let T_d be a d-ary tree with $d \geq 2$. Then the critical percolation parameter is $p_c(T_d) = 1/d$.

A generalization of the previous theorem to general trees involves de definition of the *branching number*. On a infinite tree, the branching number corresponds to the average number of branches per vertex. The definition of branching number arises from the work of Furstenberg on the Hausdorff dimension of a tree [25]. Its precise definition goes beyond the scope of this survey, and we refer the reader to [45] or [46, Section 1.2]. However, we want to point out that the branching number of a d-regular tree is d. In this sense, the following theorem extends Proposition 4.2.

Theorem 4.3 ([45])**.** Let T be a infinite tree with branching number b. Then the critical percolation parameter is $p_c(T) = 1/b$.

4.3 Coloring percolation clusters at random

In [28], Häggström introduced a percolation model on a graph G called the *divide and color model*. This model introduces a random coloring of the vertices of G. Hereafter we will write *DaC*, for short. This model arises by first performing Bernoulli bond percolation with parameter p on the graph G (see Figure 6), and then assigning a random color, chosen according to a prescribed probability distribution on the finite set $[d]$, to each of the connected components, independently for different components, see Figure 7.

In the following we introduce the model more precisely for the case of our interest, that is, when G is a tree.

4.3.1 Coloring percolation on trees

Let $\mathbf{t} = (V, E)$ be an infinite tree and $d \geq 2$ different colors indexed by the set $[d]$. Additionally to the percolation parameters p used to defined the Bernoulli percolation, we consider a probability

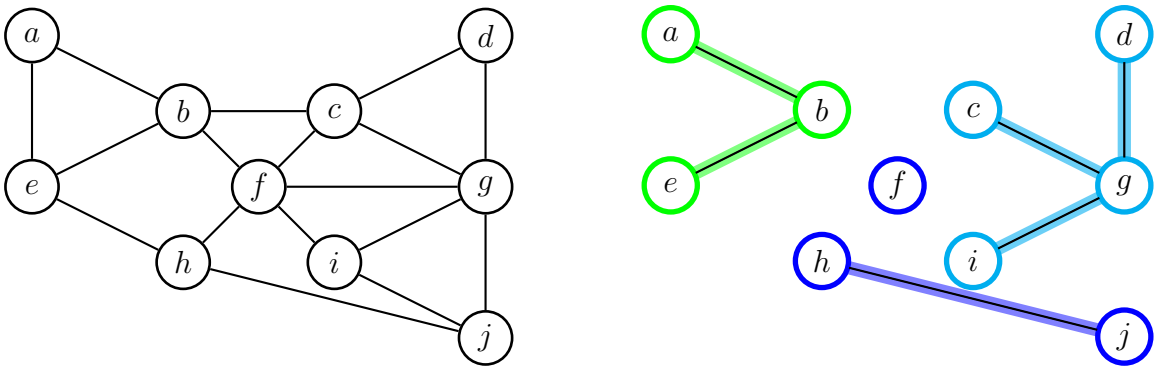

Figure 7: On the left-hand side, we have a finite connected graph, while on the right-hand side, we see a realization of a DaC percolation. This was obtained by applying a random coloring to the Bernoulli percolation illustrated in Figure 6.

distribution $\mathbf{a} = (a_1, \ldots, a_d)$ over the set of colors $[d]$. The coloring is done according to the following two-step procedure.

> **Step 1**: We perform a Bernoulli percolation on $\mathbf{t}$ (as described in Subsection 4.2). As a result, we obtain a configuration on $\mathcal{S}$ denoted by s. The set of clusters obtained is denoted by $\mathcal{C}$.
>
> **Step 2**: For each vertex on the same cluster of $\mathcal{C}$ we assign the same color. This color is chosen according to the distribution $\mathbf{a}$ independently for different connected components. The resulting coloring Π is defined by a probability measure on $[d]^V$ and it will be denoted by $\mathbb{P}^{\mathbf{t}}_{p,d,\mathbf{a}}$.

Throughout this paper, we shall work with the DaC model with just two colors. Hence, we use the notation $\mathbb{P}_{p,2,a_1}$, where a_1 is the probability of coloring a vertex on the graph with the color 1. Note that $1 - a_1$ is the probability of not to color the vertex with color 1. The development in [28] include the study of the DaC for infinite trees and phase transition percolation results for $d = 2$. The reason for the author to consider only two colors is due to the fact that the question of whether the DaC measure $\mathbb{P}^{\mathbf{t}}_{p,d,\mathbf{a}}$ produces an infinite connected component of color $i \in [d]$ with $d \geq 2$ can be answered immediately using the measure $\mathbb{P}^{\mathbf{t}}_{p,2,a_i}$, since $1 - a_i$ is the probability of not to color the vertex with color i. We have summarized the phase transition results of [28] in the following proposition.

Proposition 4.4 ([28, Propositions 2.7 and 2.8])**.** Let Π be the resulting coloring of applying the DaC model on an infinite tree $\mathbf{t}$ with parameters $p \in [0, 1]$, $d = 2$ and $\mathbf{a} = (a_1, a_2)$, we have

$$\mathbb{P}^{\mathbf{t}}_{p,2,a_1}(X \text{ contains an infinite connected component of the color of } 1) > 0.$$

if and only if

$$\mathbf{P}^{\mathbf{t}}_{\widetilde{p}}(\mathcal{O} \text{ contains an infinite connected component }) > 0$$

where $\widetilde{p} = 1 - (1 - p)(1 - a_1)$ and $\mathbf{P}^{\mathbf{t}}_{\widetilde{p}}$ is the probability measure defined in (10).

According to [28, Proof of Proposition 2.7], the idea behind the proof of the propositions above is that the connected component of a given color containing a given vertex has the same

distribution as a Bernoulli percolation. However, one needs to be careful because the probability law of DaC on $\mathbf{t}$ is not equivalent to Bernoulli site percolation on the vertices of $\mathbf{t}$ (see [28, Remark p.227]). This can be clarified, without loss of generality, by analyzing a genealogical tree with root $\emptyset$ and its children. Note that the individuals who are children of the root have two possibilities for inheriting the same color of its mother. Namely,

1) The children of the root belongs to the same cluster as their mother; this occurs with probability p.

2) The children of the root belong to a different cluster from their mother, but this connected component chooses the color of the mother. This event has probability $(1-p)a_\emptyset$, where $a_\emptyset$ is the probability of coloring a cluster with the same color of the root.

Therefore, a child of the root has the same color as its mother with probability

$$p + (1-p)a_\emptyset = 1 - (1-p)(1-a_\emptyset).$$

Let us denote by $\widetilde{\mathbf{t}}_\emptyset$ the subtree where every individual has the same color as the root and have no ancestors of other color. Thanks to the previous observation we can directly conclude the following result.

Proposition 4.5. Consider the DaC model on an infinite tree $\mathbf{t}$ with parameters $p \in [0,1]$, $d = 2$ and $a_\emptyset$ denoting the probability of coloring a cluster with the same color of the root. We have that

$$\mathbb{P}^{\mathbf{t}}_{p,d,a_\emptyset}\left(\widetilde{\mathbf{t}}_\emptyset \text{ has infinite size }\right) > 0$$

if and only if

$$\mathbf{P}^{\mathbf{t}}_{\widetilde{p}}(\mathcal{C}_\emptyset \text{ has infinite size }) > 0,$$

where $\widetilde{p} = 1-(1-p)(1-a_\emptyset)$ and $\mathcal{C}_\emptyset$ is the cluster of the root obtained to apply a Bernoulli percolation with parameter $\widetilde{p}$ on $\mathbf{t}$.

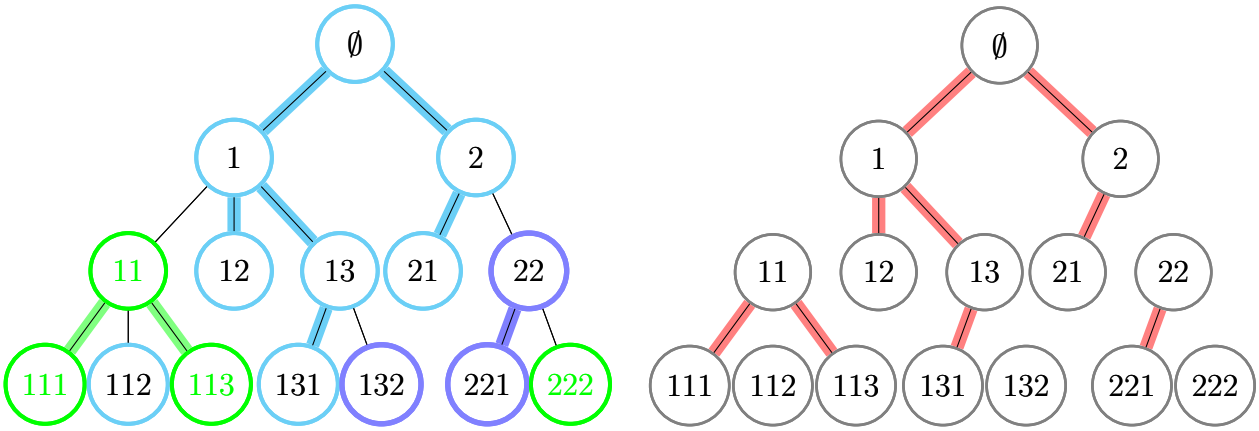

Figure 8: The figure on the left shows a realization of the *DaC* in a finite tree, with parameter p and $d = 3$. The figure on the right illustrates a realization of the correspondent Bernoulli percolation, where the connected components are given by color.

5 Connections between percolation models and branching structures

This section is devoted to provide a representation of the BGW branching process with neutral mutations, as presented in Section 3.3, in terms of Bernoulli percolation and the DaC model.

In [7], Bertoin introduced the model of neutral mutations on infinite alleles and implicitly used a percolation in BGW trees. This connection is made explicit in [9]. In Subsection 5.1, we generalize this relationship in two ways: from the point of view of population genetics (BGW process with multi-allelic neutral mutations) and from the point of view of percolation theory (DaC percolation model).

First we will discuss the relation between BGW processes with infinite allele neutral mutations and Bernoulli percolation introduced by Bertoin [7] and studied also in [8]. Next, we will introduce the relation between the DaC percolation and the MIM model. Finally, at the end of Subsection 5.1 we will present the *restricted DaC percolation model*, related to MDM model.

5.1 From Bernoulli percolation to neutral mutations and back

5.1.1 Bernoulli percolation on BGW trees

In Subsection 4.1, we define Bernoulli percolation with retention parameter p on a deterministic graph. We can apply the same procedure to a random tree, in particular, on a BGW tree. Note that by taking $p = 1 - r$, where r is the mutation parameter on BGW process with infinite alleles (neutral mutation), we obtain a unique correspondence between both models. Roughly speaking, a mutation corresponds to closing an edge. In terms of population genetics, opening an edge can be interpreted as the mother and child sharing the same type. Indeed, by the branching property it is enough to analyze the root $\emptyset$ and its children. Conditioned on $\{\xi^{(+)} = v\}$ the distribution of the number of clones is given by

$$\mathbb{P}(\xi^{(c)} = u) = \binom{v}{u}(1-r)^u r^{v-u},$$

i.e. conditioned on $\xi^{(c)}\,|_{\{\xi^{(+)}=v\}} \sim \text{Binomial}(v, 1-r)$. Therefore, for each child we have a Bernoulli distribution with parameter $1-r$, this corresponds with the marginal of the measure $\mathbf{P}^G_{1-r}$ defined in (10).

5.1.2 Divide and color percolation on BGW trees

In Subsection 4.3, we define the DaC model on an infinite tree $\mathbf{t}$ with parameters $p \in [0,1]$, $d \geq 2$ and $\mathbf{a} = (a_1, \ldots, a_d)$. We can apply the same model to a BGW tree. As before, we can find a correspondence with the mother-independent mutation model defined in 3.4.1. In order to do that, we take the parameter $p = 1 - r$, where r is the mutation probability. We also consider d the number of types and set $a_i = 1/d$ for all $i \in [d]$, since each mutant child chooses its color uniformly.

Again, by the branching property it is enough to analyze the root $\emptyset$ and its children. Assume that the type of the root is i. Following equations (6) and (7), conditioned on $\{\xi^{(+,i)} = |\mathbf{u}|\}$,

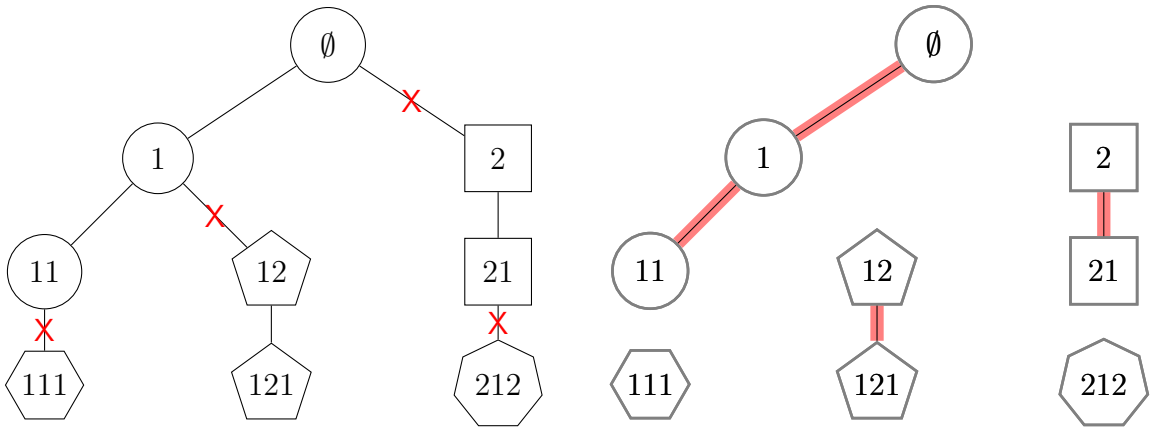

Figure 9: Comparison between infinite alleles model and Bernoulli percolation

the probability that the root have $\boldsymbol{\xi}^i := (\xi^{(i,1)}, \xi^{(i,2)}, \ldots, \xi^{(i,d)}) = (u_1, u_2, .., u_d)$ children, with $\sum_{i=1}^{d} u_i = |\mathbf{u}|$ is given by

$$\mathbb{P}(\boldsymbol{\xi}^i = (u_1, u_2, .., u_d)) = \sum_{k=1}^{u_i} \binom{|\mathbf{u}|}{k} (1-r)^k r^{|\mathbf{u}|-k} \binom{|\mathbf{u}|-k}{u_1, \ldots, u_i - k, \ldots, u_d} \left(\frac{1}{d}\right)^{|\mathbf{u}|-k}. \tag{12}$$

This measure coincides with the DaC measure $\mathbb{P}_{p,d,\mathbf{a}}$ applied to the root and its children at the first generation (see Figure 5).

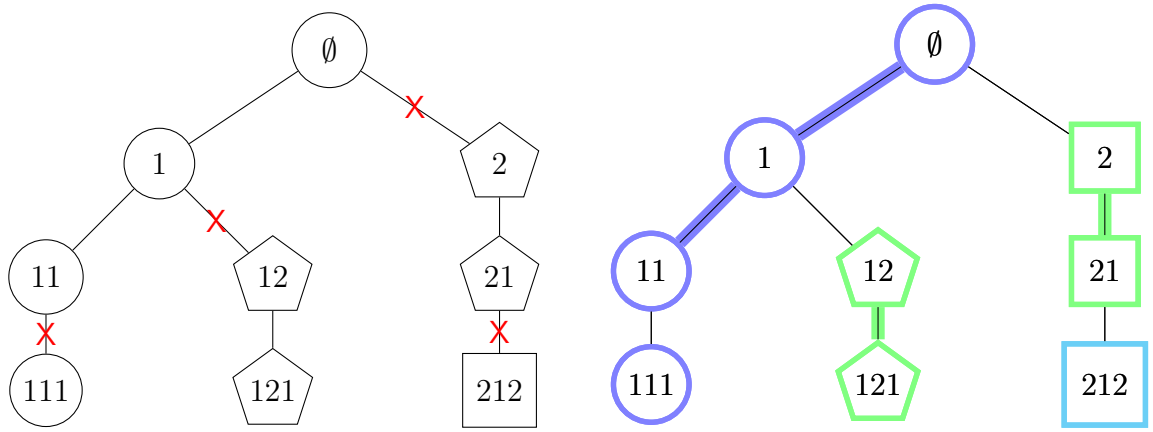

Figure 10: Comparison between MIM model and DaC percolation.

5.1.3 Restricted divide and color percolation on BGW trees

Let us now consider the (uniform) DaC percolation model in a BGW tree but restricted to the case when mother and child belonging to different clusters cannot have the same color. As in 4.3.1, we can describe the coloring step procedure as follows:

> **Step 1**: We perform a Bernoulli percolation on $\mathbf{t}$ (as described in Subsection 4.2). As a result, we obtain a configuration on $\mathcal{S}$ denoted by s. The set of clusters obtained is denoted by $\mathcal{C}$. Note that each cluster K is in a particular a subtree with a root $\emptyset_K$.

Step 2: Given an assigned color for the cluster of the initial root, we will color the other clusters according of the lexicographical order of their roots. The color of $\emptyset_K$ determines the color of all the vertex of K. Then, the color of $\emptyset_K$ is chosen uniformly over the set $[d] \setminus \{i\}$ if the color of its mother is i. The resulting coloring $\widetilde{\Pi}$ is defined by a probability measure on $[d]^V$ and it will be denoted by $\widetilde{\mathbb{P}}^{\mathbf{t}}_{p,d,unif}$.

Similarly to the previous subsection, we can compare this percolation model with the mother-dependent mutation model. In this case, following equation (5), conditioned on $\{\xi^{(+,i)} = |\mathbf{u}|\}$, the probability that the root have $\boldsymbol{\xi}^i := (\xi^{(i,1)}, \xi^{(i,2)}, \dots, \xi^{(i,d)}) = (u_1, u_2, .., u_d)$ children, with $\sum_{i=1}^d u_i = |\mathbf{u}|$ is given by

$$\mathbb{P}(\boldsymbol{\xi}^i = (u_1, u_2, .., u_d)) = \binom{|\mathbf{u}|}{u_i}(1-r)^{u_i} r^{|\mathbf{u}|-u_i} \binom{|\mathbf{u}| - u_i}{u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_d} \left(\frac{1}{d}\right)^{|\mathbf{u}|-u_i}. \tag{13}$$

This coincide with the restricted uniform DaC measure $\widetilde{\mathbb{P}}^{\mathbf{t}}_{p,d,unif}$ defined in the step-procedure above, applied to the root and its children at the first generation.

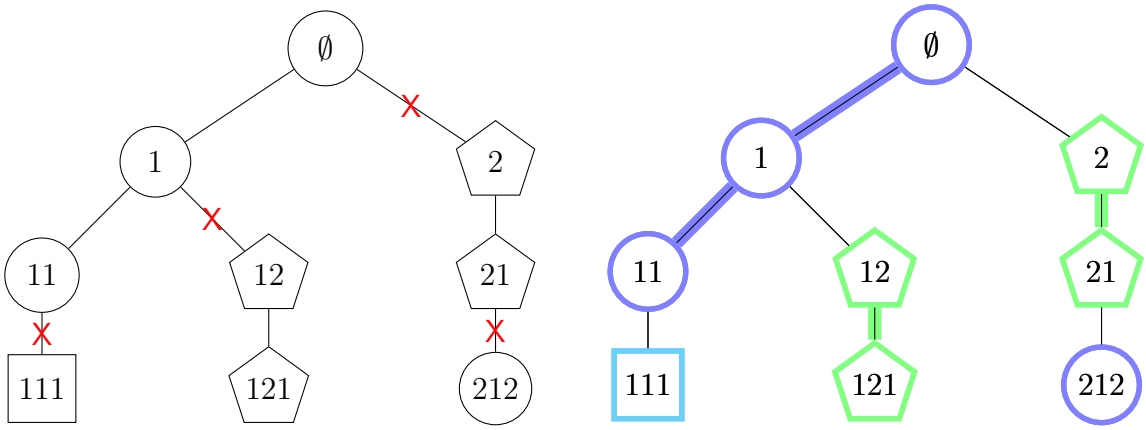

Figure 11: Comparison between MDM model and restricted DaC percolation.

5.2 The percolation phase transition from the point of view of neutral mutation models

For the Bernoulli phase transition on a tree $\mathcal{T}$, the classical interpretation of the percolation supercritical regime corresponds to the existence of an infinite cluster with probability 1. From the point of view of population genetics, for a BGW tree with mutation we can interpret the cluster of the root as a subtree $\mathcal{T}_\emptyset$ of individuals of the same type of the root $\emptyset$. Note that a subtree is connected, so $\mathcal{T}_\emptyset$ does not include individuals of the same type of $\emptyset$ but with ancestors of a different type.

On the percolation supercritical regime, the size of $\mathcal{T}_\emptyset$ has a positive probability of being of the same order to the size of the full genealogy. If the full genealogical tree is finite, it means that the size of $\mathcal{T}_\emptyset$ is proportional to the size of $\mathcal{T}$. On the other hand, if $\mathcal{T}$ is infinite, this means that $\mathcal{T}_\emptyset$ is infinite as well.

In the following we will present some percolation results on BGW trees conditioned to be infinite and we describe it in terms of the mutation models presented above. First, we recall some results of Lyons for percolation in BGW trees and its interpretation in terms of the infinite alleles mutation model and the MDM model. Finally, we will extend Proposition 4.4 for the BGW case, and we describe its meaning for the MIM model.

5.2.1 Percolation phase transition for BGW trees

We now recall the generalization of Proposition 4.2 to branching trees.

Theorem 5.1 (Lyons, 1990)**.** Let $\mathcal{T}$ be a BGW tree whose offspring distribution has mean $m > 1$. That is, if ξ is a random variable following the offspring distribution, then $m = \mathbb{E}(\xi) > 1$. Conditioned on non-extinction, we have that $p_c^{\mathcal{T}} = 1/m$ almost surely.

In the case of infinite alleles model, we consider the probabilities

$$\begin{aligned} q_\infty &= \mathbb{P}\left(\text{There exists an infinite sub-tree with the same type}\right) \\ q_\emptyset &= \mathbb{P}\left(\mathcal{T}_\emptyset \text{ is infinite}\right), \end{aligned} \tag{14}$$

where $\mathcal{T}_\emptyset$ correspond to the sub-tree of clones. Therefore, we interpret Theorem 5.1 as a phase transition for the probabilities in (14). Specifically,

$$q_\infty = \begin{cases} 1 & \text{if } 1-r \leq p_c^{\mathcal{T}} \\ 0 & \text{if } 1-r > p_c^{\mathcal{T}}. \end{cases} \tag{15}$$

and

$$q_\emptyset \begin{cases} > 0 & \text{if } 1-r \leq p_c^{\mathcal{T}} \\ = 0 & \text{if } 1-r > p_c^{\mathcal{T}}. \end{cases} \tag{16}$$

Similarly, for the mother-dependent mutation model, we can make a similar interpretation for the probabilities $q_\infty^\star$ and $q_\emptyset^\star$. We define $q_\infty^\star$ just as q_∞. The variation for $q_\emptyset^\star$ is that instead of $\mathcal{T}_\emptyset$ we look at the subtree $\mathcal{T}_\emptyset^\star$ of individuals of the same type of the root $\emptyset$. Then we define

$$q_\emptyset^\star = \mathbb{P}\left(\text{The sub-tree } \mathcal{T}_\emptyset^\star \text{ is infinite}\right). \tag{17}$$

A phase transition holds as in (15) and (16).

5.2.2 Divide and color percolation phase transition for BGW trees

Finally, we consider DaC percolation on a BGW tree conditioned to be infinite. We will compare this model with mother-independent mutation model introduced in 3.4.2. For this propose, we extend below Proposition 4.4 from deterministic trees to BGW trees.

Theorem 5.2. Let $\mathcal{T}$ be a BGW tree whose offspring distribution has mean $m > 1$, as in Theorem 5.1. We have that

$$\widetilde{p}_c^{\mathcal{T}} = \frac{1 - m a_\emptyset}{m(1 - a_\emptyset)} \qquad \text{a.s. given non-extinction.}$$

Proof. As in the deterministic case, the connected component of a given color containing a given vertex is distributed as in Bernoulli percolation in a BGW tree with retention parameter $\widetilde{p} = 1-(1-p)(1-a_\emptyset)$. As a consequence of this fact, we have that the Proposition 5.9 in [46, Section 5] is also valid in this case, so we have that that the cluster of the root (by color) will be finite a.s. if $m\widetilde{p} \leq 1$. From this, we deduce the critical probability for this case. □

We now provide an interpretation of Theorem 5.2 for the mother-independent mutation model, as in (15) and (16) in the previous subsection. In this case, we define the probabilities

$$\begin{aligned} \widetilde{q}_\infty &= \mathbb{P}^{\mathcal{T}}_{p,2,1/d}\left(\text{There exists an infinite sub-tree of the same type of the root }\right) \\ \widetilde{q}_\emptyset &= \mathbb{P}^{\mathcal{T}}_{p,2,1/d}\left(\widetilde{\mathcal{T}}_\emptyset \text{ is infinite}\right), \end{aligned} \tag{18}$$

where here $\widetilde{\mathcal{T}}_\emptyset$ correspond to the subtree of individuals that are of the same type as the root and have no ancestors of other types.

We then interpret Theorem 5.2 as a phase transition for the probabilities in (18). Specifically,

$$\widetilde{q}_\infty = \begin{cases} 1 & \text{if } 1-r \leq \widetilde{p}_c^{\mathcal{T}} \\ 0 & \text{if } 1-r > \widetilde{p}_c^{\mathcal{T}}, \end{cases} \tag{19}$$

and

$$\widetilde{q}_\emptyset \begin{cases} > 0 & \text{if } 1-r \leq \widetilde{p}_c^{\mathcal{T}} \\ = 0 & \text{if } 1-r > \widetilde{p}_c^{\mathcal{T}}. \end{cases} \tag{20}$$

Acknowledgements

MCF's research is supported by UNAM-PAPIIT grant IN109924. SHT would like to acknowledge the support of UNAM-PAPIIT grant IA103724.

The authors would like to thank the research network Aleatorias & Normales connecting female Latin American researchers in probability and statistics. The discussions facilitated by the *Aleatorias & Normales Seminar* were the starting point of this project.

References

[1] A. M. Polyakov A. A. Belavin and A. B. Zamolodchikov. Infinite conformal symmetry in two-dimensional quantum field theory. *Nuclear Phys. B.*, 241(2):333–380, 1984.

[2] A. M. Polyakov A. A. Belavin and A. B. Zamolodchikov. Infinite conformal symmetry of critical fluctuations in two dimensions. *J. Statist. Phys.*, 34(5-6):763–774, 1984.

[3] R. Abraham and J.-F. Delmas. An introduction to Galton-Watson trees and their local limits. Available at hal-01164661v2, 2015.

[4] M. Aizenman, H. Duminil-Copin, and V. Sidoravicius. Random currents and continuity of Ising model's spontaneous magnetization. *Comm. Math. Phys.*, 334(2):719–742, 2015.

[5] K. B. Athreya and P. E. Ney. *Branching Processes*. Springer-Verlag, 1972.

[6] V. Beffara and H. Duminil-Copin. The self-dual point of the two-dimensional random-cluster model is critical for $q \geq 1$. *Probab. Theory Related Fields*, 153(3-4):511–542, 2012.

[7] J. Bertoin. The structure of the allelic partition of the total population for galton–watson processes with neutral mutations. *Ann. Probab.*, 37(4):1502–1523, 2009.

[8] J. Bertoin. A limit theorem for trees of alleles in branching processes with rare neutral mutations. *Stoch. Process. Their Appl.*, 120(5):678–697, 2010.

[9] J. Bertoin and G. Uribe Bravo. Supercritical percolation on large scale-free random trees. *Ann. Appl. Probab.*, 25(1):81–103, 2015.

[10] G. Berzunza. The existence of a giant cluster for percolation on large crump–mode–jagers trees. *Adv. Appl. Probab.*, 52(1):266–290, 2020.

[11] I. J. Bienaymé. De la loi de multiplication et de la duree des families. *Soc. Philomat. Paris Extraits, Sér 5*, pages 37–39, 1845.

[12] A. Blancas and V. Rivero. On branching process with rare neutral mutation. *Bernoulli*, 24(2):1576 – 1612, 2018.

[13] S. Broadbent and J. Hammersley. Percolation processes: I. Crystals and mazes. *Proceedings of the Cambridge Philosophical Society*, 53.

[14] H. Duminil-Copin, M. Gagnebin, M. Harel, I. Manolescu, and V. Tassion. Discontinuity of the phase transition for the planar random-cluster and Potts models with $q > 4$. *Ann. Sci. Éc. Norm. Supér*, 54(6):1363–1413, 2021.

[15] H. Duminil-Copin, S. Goswami, A. Raoufi, F. Severo, and A. Yadin. Existence of phase transition for percolation using the gaussian free field. *Duke Math. J.*, 169(18):3539–3563, 2020.

[16] H. Duminil-Copin, K. K. Kozlowski, D. Krachun, and I. Manolescu. Rotational invariance in critical planar lattice models. Preprint, available at arXiv:2012.11672, 2020.

[17] H. Duminil-Copin, A. Raoufi, and V. Tassion. Sharp phase transition for the random-cluster and Potts models via decision trees. *Ann. Math.*, 189(1):75–99, 2019.

[18] H. Duminil-Copin, V. Sidoravicius, and V. Tassion. Continuity of the phase transition for planar random-cluster and Potts models with $1 \leq q \leq 4$. *Comm. Math. Phys.*, 349(1):47–107, 2017.

[19] H. Duminil-Copin and V. Tassion. A new proof of the sharpness of the phase transition for Bernoulli percolation and the Ising model. *Comm. Math. Phys.*, 343:725–745, 2016.

[20] T. Duquesne and J.-F. Le Gall. *Random trees, Lévy processes and spatial branching processes*, volume 281. Société mathématique de France Paris, France, 2002.

[21] P. Easo and T. Hutchcroft. The critical percolation probability is local. Preprint, available at arXiv:2310.10983, 2023.

[22] C. Fortuin. On the random-cluster model: II. The percolation model. *Phyisica*, 58, 1972.

[23] C. Fortuin. On the random-cluster model: III. The simple random-cluster model. *Phyisica*, 59, 1972.

[24] C. Fortuin and P. Kasteleyn. On the random-cluster model: I. Introduction and relation to other models. *Phyisica*, 57, 1972.

[25] H. Furstenberg. Intersections of Cantor sets and transversality of semigroups. In *Problems in Analysis. Symp in Honor of Salomon Bochner, Princeton Univ.*, pages 41–59. Princeton University Press, 1971.

[26] G. Grimmett. *Percolation*. Springer, 1999.

[27] G. Grimmett. *The Random-Cluster Model*. Springer, 2006.

[28] O. Haggström. Coloring percolation clusters at random. *Stoch. Process. Their Appl.*, 96:213––242, 2001.

[29] M. Hairer. The work of Hugo Duminil-Copin. In *Proceedings of the International Congress of Mathematicians 2022 (ICM 2022) Vol. I: Plenary Lectures and Ceremonies*, pages 1–23, 2022.

[30] J. L. Jacobsen, J. Salas, and A. D. Sokal. Spanning forests and the q-state Potts model in the limit $q \to 0$. *J. Stat. Phys*, 119:1153–1281, 2005.

[31] T.H. Jukes and C.R. Cantor. Evolution of protein molecules. *Mammalian Protein Metabolism*, 3(24):21–132, 1969.

[32] D. G. Kendall. The genealogy of genealogy branching processes before (and after) 1873. *B. Lond. Math. Soc.*, 7(3):225–253, 1975.

[33] H. Kesten. The work of stanislav smirnov. In *Proceedings of the International Congress of Mathematicians 2010 (ICM 2010) Vol. I: Plenary Lectures and Ceremonies*, pages 72–84, 2010.

[34] M. Kimura. Theoretical foundation of population genetics at the molecular level. *Theor. Popul. Biol*, 2(2):174–208, 1971.

[35] M. Kimura. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.*, 16(2):111–120, 1980.

[36] M. Kimura. Estimation of evolutionary distances between homologous nucleotide sequences. *Proc. Natl. Acad. Sci. U.S.A.*, 78(1):454–458, 1981.

[37] J. Kuipers, K. Jahn, B. J. Raphael, and N. Beerenwinkel. Single-cell sequencing data reveal widespread recurrence and loss of mutational hits in the life histories of tumors. *Genome Res.*, 27(11):1885–1894, 2017.

[38] A. Lambert. Population dynamics and random genealogies. *Stoch. Models*, 24(sup1):45–163, 2008.

[39] G. F. Lawler, O. Schramm, and W. Werner. Values of Brownian intersection exponents, I: Half-plane exponents. *Acta Math.*, 187:237–273, 2001.

[40] G. F. Lawler, O. Schramm, and W. Werner. Values of Brownian intersection exponents, II: Plane exponents. *Acta Math.*, 187:275–308, 2001.

[41] G. F. Lawler, O. Schramm, and W. Werner. Analyticity of intersection exponents for planar Brownian motion. *Acta Math.*, 189(2):179–201, 2002.

[42] G. F. Lawler, O. Schramm, and W. Werner. One-arm exponent for critical 2D percolation. *Electron. J. Probab.*, 7:1–13, 2002.

[43] G. F. Lawler, O. Schramm, and W. Werner. Values of Brownian intersection exponents III: Two-sided exponents. *Ann. Inst. Henri Poincaré B*, 38(1):109–123, 2002.

[44] G. F. Lawler, O. Schramm, and W. Werner. Conformal invariance of planar loop-erased random walks and uniform spanning trees. *Ann. Probab.*, 32(1B):939–995, 2004.

[45] R. Lyons. Random walks and percolation on trees. *Ann. Probab.*, 18(3):931–958, 1990.

[46] R. Lyons and Y. Peres. *Probability on trees and networks*, volume 42. Cambridge University Press, 2017.

[47] L. A. Mathew, P. R. Staab, L. E. Rose, and D. Metzler. Why to account for finite sites in population genetic studies and how to do this with Jaatha 2.0. *Ecol. Evol.*, 3(11):3647–3662, 2013.

[48] C. M. Newman. The work of Wendelin Werner. In *International Congress of Mathematicians 2006 (ICM 2006). Vol. I: Plenary Lectures and Ceremonies*, pages 88–95, 2007.

[49] O. Schramm. Scaling limits of loop-erased random walks and uniform spanning trees. *Isr. J. Math.*, 118:221–288, 2000.

[50] S. Smirnov. Critical percolation in the plane: conformal invariance, Cardy's formula, scaling limits. *C. R. Acad. Sci.*, 333:239–244, 2001.

[51] S. Smirnov and Wendelin W. Critical exponents for two-dimensional percolation. *C. R. Acad. Sci.*, 333:239–244, 2001.

[52] H. W. Watson and F. Galton. On the probability of the extinction of families. *J. R. Anthropol. Inst.*, 4:138–144, 1875.

[53] K. G. Wilson. The renormalization group and critical phenomena. *Rev. Mod. Phys.*, 55(5-6):763–774, 1983.

[54] H. Zafar, A. Tzen, N. Navin, K. Chen, and L. Nakhleh. SiFit: inferring tumor trees from single-cell sequencing data under finite-sites models. *Genome Biol.*, 18:1–20, 2017.