

RESEARCH ARTICLE

WILEY

Random tree recursions: Which fixed points correspond to tangible sets of trees?

Tobias Johnson¹ | Moumanti Podder² | Fiona Skerman³

¹Department of Mathematics, College of Staten Island, Staten Island, New York

²Department of Mathematics, NYU-ECNU Institute of Mathematical Sciences, New York University, Shanghai, China

³Heilbronn Institute for Mathematics, Department of Mathematics, University of Bristol, Bristol, United Kingdom

Correspondence

Tobias Johnson, College of Staten Island
Department of Mathematics, 1S-215 2800
Victory Blvd. Staten Island, NY 10314.
Email: tobias.johnson@csi.cuny.edu

Funding information

This research was supported by the NSF grant DMS-1401479; DMS-1811952, NSF CAREER grant CCF:AF-1553354,

Abstract

Let $\mathcal{B}$ be the set of rooted trees containing an infinite binary subtree starting at the root. This set satisfies the metaproperty that a tree belongs to it if and only if its root has children u and v such that the subtrees rooted at u and v belong to it. Let p be the probability that a Galton-Watson tree falls in $\mathcal{B}$. The metaproperty makes p satisfy a fixed-point equation, which can have multiple solutions. One of these solutions is p , but what is the meaning of the others? In particular, are they probabilities of the Galton-Watson tree falling into other sets satisfying the same metaproperty? We create a framework for posing questions of this sort, and we classify solutions to fixed-point equations according to whether they admit probabilistic interpretations. Our proofs use spine decompositions of Galton-Watson trees and the analysis of Boolean functions.

KEYWORDS

endogeny, fixed point, Galton-Watson tree, interpretation, recursive distributional equation, tree automaton

1 | INTRODUCTION

A seminal problem in discrete probability is to determine the probability of survival of a Galton-Watson tree. For the sake of simplicity, suppose that the offspring distribution is $\text{Poi}(\lambda)$, and denote the tree by T_λ . Let $\mathcal{T}_{\text{inf}}$ denote the set of infinite rooted trees. Let p denote the survival probability, given by $\mathbf{P}[T_\lambda \in \mathcal{T}_{\text{inf}}]$. The typical solution gives p as a fixed point of a map as follows: Let Z be the number of children v of the root of T_λ such that the subtree rooted at v is infinite. Each subtree is infinite with probability p , just like the original tree. Thus $Z \sim \text{Poi}(p\lambda)$ by Poisson thinning. Since T_λ is infinite if and only if $Z \geq 1$,

$$p = 1 - e^{-\lambda p}. \quad (1)$$

As is well known [3], when $\lambda > 1$, this equation has two solutions, and the positive one is the true value of p . In arriving at (1), the only property of $\mathcal{T}_{\text{inf}}$ we used was that $t \in \mathcal{T}_{\text{inf}}$ if and only if there exists some child v of the root of t such that the subtree descending from v is in $\mathcal{T}_{\text{inf}}$. Let us call this the metaproperty of $\mathcal{T}_{\text{inf}}$ that yields (1).

Again, let T_λ be a Galton-Watson tree with child distribution $\text{Poi}(\lambda)$. It is natural to ask if there is some other set of trees $\mathcal{T}_0$ satisfying the metaproperty such that $\mathbf{P}[T_\lambda \in \mathcal{T}_0]$ is the other solution to (1), which is 0. A bit of thought reveals that $\mathcal{T}_0 = \emptyset$ fits this criteria. Vacuously, $t \in \emptyset$ if and only if the root of t has a child whose subtree is in $\emptyset$, and clearly $\mathbf{P}[T_\lambda \in \emptyset] = 0$. Thus, the metaproperty yields an equation with two solutions, and each solution gives the probability under the Galton-Watson measure of a set of trees satisfying the metaproperty. Indeed, we will later see that $\mathcal{T}_{\text{inf}}$ and $\emptyset$ are the *only* two sets of trees satisfying the metaproperty, up to measure zero changes under the Galton-Watson measure with child distribution $\text{Poi}(\lambda)$ (see Remark 1.2 for more discussion on measure zero changes).

This work was motivated by a nearly identical example that is more difficult to resolve. This time, we consider sets of trees $\mathcal{B}$ where $t \in \mathcal{B}$ if and only if the root of t has at least *two* children u and v whose subtrees are in $\mathcal{B}$. Let us call this metaproperty the *at-least-two rule*. Suppose $p = \mathbf{P}[T_\lambda \in \mathcal{B}]$ for some set of trees $\mathcal{B}$ obeying the at-least-two rule. Invoking Poisson thinning and self-similarity as in the first example, we get

$$p = 1 - e^{-\lambda p}(1 + \lambda p). \quad (2)$$

As explained in [23], which investigated the existence of a giant 3-core in a random graph, there is a critical parameter $\lambda_{\text{crit}} \approx 3.35$ where this equation changes behavior (see Figure 1). For all $\lambda > 0$, there is a trivial solution to (2) given by $p = 0$. When $\lambda < \lambda_{\text{crit}}$, this is the only solution. At $\lambda = \lambda_{\text{crit}}$, a second solution emerges, and when $\lambda > \lambda_{\text{crit}}$ there are three solutions. (We prove these statements in Example 5.5.) Let $\mathcal{B}_0$ be the set of all trees that contain an infinite binary subtree starting at the root. Note that $\mathcal{B}_0$ satisfies the at-least-two rule. It was shown by Dekking [8] (also see [22]) that $\mathbf{P}[T_\lambda \in \mathcal{B}_0]$ is the largest solution to (2) when $\lambda > \lambda_{\text{crit}}$, shown in green in Figure 1. An immediate intuition as to why the green curve is the one corresponding to $\mathbf{P}[T_\lambda \in \mathcal{B}_0]$ is that this is the only curve which is increasing in λ . Another set of trees obeying the at-least-two rule is the empty set. Obviously, $\mathbf{P}[T_\lambda \in \emptyset] = 0$, the smallest solution to (2), shown in red in Figure 1. Joel Spencer posed the question that set this work in motion: is there a set of trees to go with the middle solution (shown in blue in Figure 1)? More formally, the question asks the following:

Question 1.1 (Spencer). Let T_λ be a Galton-Watson tree with child distribution $\text{Poi}(\lambda)$. Say that a set of trees $\mathcal{B}$ follows the at-least-two rule if $t \in \mathcal{B}$ if and only if the root of t has two children u and v

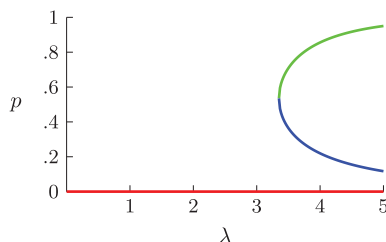


FIGURE 1 A plot showing all p satisfying (2) for given λ . For $\lambda < \lambda_{\text{crit}} \approx 3.35$, the only solution to (2) is $p = 0$. For $\lambda = \lambda_{\text{crit}}$, there are two solutions for p , and for $\lambda > \lambda_{\text{crit}}$, there are three [Color figure can be viewed at wileyonlinelibrary.com]

such that the subtrees rooted at u and v are also in $\mathcal{B}$. Suppose that $\lambda > \lambda_{\text{crit}}$. Does there exist a set of trees $\mathcal{B}$ following the at-least-two rule such that $\mathbf{P}[T_\lambda \in \mathcal{B}]$ is the middle solution of (2)?

We answer this question in the negative. More generally, our main result, Theorem 1.7, gives the answer to *any* question of this form. In the language of this paper, it is a criterion for which fixed points of tree automata admit interpretations. In this example, the tree automaton refers to the at-least-two metaproperty. For the Galton-Watson child distribution $\text{Poi}(\lambda)$, the fixed points of this automaton are the solutions to (2). An interpretation corresponds to a set of trees following the metaproperty given by the automaton. Theorem 1.7 shows that $\mathcal{B}_0$ and $\emptyset$ are the only two sets of trees following the at-least-two rule, up to measure zero changes under the Galton-Watson measure with child distribution $\text{Poi}(\lambda)$.

Remark 1.2. It is important that we consider sets of trees satisfying a metaproperty only up to measure zero changes under a Galton-Watson measure with a given child distribution. For example, let $\mathcal{T}$ be the set of trees that contain an infinite binary subtree somewhere within them (i.e., not necessarily starting from the root). This set satisfies our original metaproperty: a tree is in $\mathcal{T}$ if and only if its root has at least one child initiating a tree in $\mathcal{T}$. But on its face, $\mathcal{T}$ is neither $\mathcal{T}_{\text{inf}}$ nor $\emptyset$, which we claimed were the only sets of trees satisfying this metaproperty. The solution to this apparent paradox is that from the perspective of the Galton-Watson tree T_λ with child distribution $\text{Poi}(\lambda)$, the set $\mathcal{T}$ is in fact equivalent to either $\mathcal{T}_{\text{inf}}$ or $\emptyset$. For $\lambda < \lambda_{\text{crit}}$, there is zero probability that T_λ lies in $\mathcal{T}$, and hence $\mathcal{T}$ is a measure zero change away from $\emptyset$. For $\lambda \geq \lambda_{\text{crit}}$, the tree T_λ falls in $\mathcal{T}$ with probability 1 given that T_λ is infinite. Hence $\mathcal{T}$ is a measure zero change away from $\mathcal{T}_{\text{inf}}$ in this case.

1.1 | Summary of main result

We start by giving a nonrigorous version of our main result, since it will take some effort to state all the definitions we need for a formal statement. A *tree automaton* is a set of rules determining the color of a parent vertex in a tree from the color of its children. Let Σ be a finite set representing the possible colors. The automaton corresponding to the at-least-two rule acts on colors $\Sigma = \{0, 1\}$, assigning color 1 to the parent if and only if it has at least two children of color 1. A *fixed point* of a tree automaton is a probability distribution $\vec{v}$ on Σ such that if a Galton-Watson tree is generated and the children of the root are assigned i.i.d.- $\vec{v}$ colors, then the color of the root induced by the automaton is also distributed as $\vec{v}$. For the example presented earlier, the fixed points have the form $\text{Bernoulli}(p)$, where p satisfies (2). To define an *interpretation* of a tree automaton, suppose we have a map $\iota: \mathcal{T} \rightarrow \Sigma$, where $\mathcal{T}$ is the space of rooted trees. Now, imagine coloring each vertex v in an arbitrary tree by applying ι to the subtree rooted at v . If the resulting coloring of the tree is always consistent with the rules given by the tree automaton, then we call the map an interpretation of the automaton. We saw two interpretations in our earlier example: the first mapped a tree to 0 or 1 depending on whether it contained an infinite binary tree starting at its root, and the second mapped all trees to 0.

It is not hard to see that the color of a Galton-Watson tree assigned by an interpretation of an automaton must be distributed as a fixed point of the automaton (see Lemma 1.4). For example, if ι is the first interpretation described above and T is a Galton-Watson tree with child distribution $\text{Poi}(\lambda)$, then $\iota(T)$ is distributed as $\text{Bernoulli}(p_2)$, where p_2 is the largest solution to (2). Our main result flips this around, letting us determine for a given fixed point $\vec{v}$ whether there exists an interpretation ι such that $\iota(T) \sim \vec{v}$.

The criterion is based on an object we call the *pivot tree*. Essentially, first generate the Galton-Watson tree to level n . Then, randomly color the vertices at level n by sampling independently from the given fixed point. Apply the automaton to color the vertices at levels 0 to $n - 1$. Now, call a

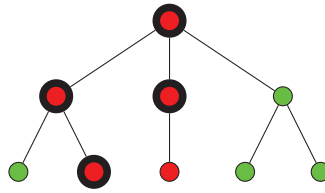


FIGURE 2 The first three levels of a tree colored consistently with the at-least-two automaton given in Example 1.3. Red denotes state 0 and green denotes state 1. Vertices in bold are pivotal, meaning that flipping their colors and recoloring above them according to the automaton causes the root to flip colors [Color figure can be viewed at wileyonlinelibrary.com]

vertex *pivotal* for this coloring if altering its color and recoloring all of its ancestors by the automaton alters the color of the root (see Figure 2). The set of all pivotal vertices to level n then forms a random subtree of the original Galton-Watson tree. There is a natural way to extend this construction beyond a fixed n to give a (possibly) infinite tree, the *pivot tree*, which turns out to be multitype Galton-Watson.

Loosely speaking, the main result of this paper is that when $|\Sigma| = 2$, a given fixed point of a tree automaton has a corresponding interpretation if and only if the associated pivot tree is subcritical or critical (or equivalently, if it is almost surely finite). If so, then it has precisely one interpretation, up to measure zero changes with respect to the Galton-Watson measure. This criterion is quite practical to check, and we do so for the at-least-two automaton and some other examples in Section 5.

When $3 \leq |\Sigma| < \infty$, we prove only that a subcritical pivot tree implies existence of an interpretation. We believe that our approach in this paper can be adapted to prove that a supercritical pivot tree implies nonexistence of an interpretation, but there are several complications (see Remark 4.16).

We now proceed to define these terms more formally. We then state our main results in Section 1.8.

1.2 | Notation

We define $\mathcal{T}$ to be the set of locally finite, ordered, rooted trees (*ordered* means that an ordering is given for the children of each vertex). This set can be viewed as a metric space (see [19, Exercise 5.2]), which we endow with its Borel σ -algebra to make a measure space. Our results will be for Galton-Watson trees with general child distributions, sometimes under mild moment conditions. We will typically denote the tree by T and the child distribution by χ . We always assume that χ puts positive probability on $\{2, 3, \dots\}$, so that T is a true tree. For any tree $t \in \mathcal{T}$, we let $V(t)$ denote its vertex set and R_t its root. Let $t(v)$ denote the subtree of t made up of v and its descendants. We let $t|_n$ denote the tree obtained by truncating t beyond its n th generation and $[t]_n \subseteq \mathcal{T}$ the set of trees that match t up to the n th generation, where the root is considered to belong to generation 0. Let $L_n(t)$ denote the set of all nodes of t in generation n , and let $\ell_n(t) = |L_n(t)|$. We abbreviate $L_n(T)$ by L_n and $\ell_n(T)$ by ℓ_n .

We will often work with *colored trees*, defined as a pair (t, τ) consisting of a tree $t \in \mathcal{T}$ together with a coloring $\tau : V(t) \rightarrow \Sigma$. We denote the space of colored trees as $\mathcal{T}_{\text{col}}$, taking the set of colors Σ as fixed in advance. For $(t, \tau) \in \mathcal{T}_{\text{col}}$, let $[t, \tau]_n \subseteq \mathcal{T}_{\text{col}}$ denote the set of colored trees that match (t, τ) up to the n th generation.

1.3 | Tree automata

Let Σ denote a finite set, to be thought of as colors or states. A *tree automaton* on the states Σ is essentially a set of rules for determining the state of a parent in the tree from the states of its children. Formally, we define an automaton as a map $A : \mathbb{N}_0^\Sigma \rightarrow \Sigma$, where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. The vector

$\vec{n} = (n_\sigma : \sigma \in \Sigma) \in \mathbb{N}_0^\Sigma$ represents the count of children in each state, and $A(\vec{n})$ represents the state assigned to the parent.

Example 1.3 (At-least-two automaton). We define an automaton A on states $\Sigma = \{0, 1\}$ that assigns state 1 to the parent if and only if at least two of its children have state 1. Formally, the automaton is the map $(n_0, n_1) \mapsto \mathbf{1}\{n_1 \geq 2\}$. As we mentioned, this automaton is implicit in Question 1.1.

Tree automata are of interest in logic and theoretical computer science. In these settings, they typically act on trees with vertex labels rather than plain trees, and there are some restrictions on them. See [7] and [17, Chapter 7] for more details on automata for finite trees, and [27, Section 6] for more on infinite trees. Tree automata can be used to determine which sets of trees can be defined by a given logic. For example, call a set of trees *regular* if there exists a tree automaton so that a tree falls into the set if and only if the automaton assigns its root one of a set of accepted states. A set of finite trees is definable in monadic second-order logic if and only if it is regular [17, Theorem 7.30, Theorem 7.34]. A similar statement holds for infinite trees as well [27, Theorem 6.19]. We will revisit logic in Section 1.9, after we state our results.

For a given tree t , we say that an assignment of colors $\tau : V(t) \rightarrow \Sigma$ is *compatible with the automaton A* if for every $v \in V(t)$, we have

$$\tau(v) = A(\vec{n}), \quad (3)$$

where $\vec{n} = (n_\sigma : \sigma \in \Sigma)$ and n_σ is the number of children of v that are colored σ under τ . If t is finite, there is only one coloring compatible with A . At each leaf, this coloring takes the value $A(0, \dots, 0)$, and then the automaton determines the colors of all other vertices. When t is infinite, however, there are typically many assignments compatible with a given automaton.

1.4 | Interpretations

An interpretation of an automaton is a deterministic classification of trees into the states of Σ such that the state of a tree can be computed from the states of the subtrees descending from the children of its root, according to the rules of the automaton. For example, assign a tree state 1 if it contains an infinite binary subtree starting at its root, and assign it state 0 otherwise. This is an interpretation of the at-least-two automaton of Example 1.3, since a tree t has state 1 if and only if its root R_t has at least two children u, v with subtrees $t(u), t(v)$ in state 1.

Formally, we define an interpretation as follows. Let χ be a probability measure on the nonnegative integers, and let $\text{GW}(\chi)$ denote the Galton-Watson measure on $\mathcal{T}$ with child distribution χ . We call a measurable map $\iota : \mathcal{T} \rightarrow \Sigma$ an *interpretation of the automaton A under $\text{GW}(\chi)$* , if for a.e.- $\text{GW}(\chi)$ tree $t \in \mathcal{T}$, the coloring $\tau : V(t) \rightarrow \Sigma$, defined as

$$\tau(v) = \iota(t(v)), \quad \text{for all } v \in V(t), \quad (4)$$

is compatible with A . Typically, we will call ι an interpretation of A without mentioning $\text{GW}(\chi)$, since the offspring distribution will be fixed throughout. For many interpretations, including our example of assigning a tree 1 if it contains an infinite binary subtree from the root, the compatibility condition holds for every tree in $\mathcal{T}$, and the measure $\text{GW}(\chi)$ is irrelevant.

1.5 | Fixed points and their connections with interpretations

Let $T \sim \text{GW}(\chi)$. If $\iota : \mathcal{T} \rightarrow \Sigma$ is an interpretation of the automaton A under $\text{GW}(\chi)$, then the distribution of $\iota(T)$ is constrained by the self-similarity of T . For example, if $\chi \sim \text{Poi}(\lambda)$ and ι is

an interpretation of the at-least-two automaton of Example 1.3, then $\iota(T) \sim \text{Bernoulli}(p)$, where p satisfies (2).

We now describe these constraints on the distribution of $\iota(T)$ when ι is an interpretation of a general automaton A and $T \sim \text{GW}(\chi)$. Let D denote the set of all probability distributions on Σ (as Σ is finite, D is a finite-dimensional simplex). We define a map $\Psi : D \rightarrow D$ that we call the *automaton distributional map corresponding to A and χ* , as follows. Fix $\vec{x} \in D$. Consider a random tree whose root has children according to χ . To each child, mutually independently, we attach a random state in Σ that follows the distribution $\vec{x}$. For every realization of this random procedure, we determine the state at the root using the rules of the automaton A . We then set $\Psi(\vec{x})$ to be the distribution of the random state thus induced at the root.

Lemma 1.4. *Let $T \sim \text{GW}(\chi)$. If ι is an interpretation for the tree automaton A , then the distribution of $\iota(T)$ is a fixed point of the automaton distribution map Ψ .*

Proof. Let the distribution of $\iota(T)$ be $\vec{y} = (y_\sigma : \sigma \in \Sigma) \in D$. Let $\tau : V(T) \rightarrow \Sigma$ be the assignment defined by $\tau(v) = \iota(T(v))$ for all $v \in V(T)$, which is almost surely compatible with A by definition of interpretation.

Under the labeling τ , the state of the root R_T is distributed as $\vec{y}$, since $\tau(R_T) = \iota(T)$. On the other hand, R_T has children according to the distribution χ ; each of these children has an independent copy of T descending from it. So, from the definition of τ , the children of R_T have i.i.d. labels distributed as $\vec{y}$. Hence the corresponding label at the root is $\Psi(\vec{y})$, by definition of Ψ . This shows that $\Psi(\vec{y}) = \vec{y}$, which is what we claimed. ■

For all tree automata, the automaton distribution map Ψ has at least one fixed point. This holds because Ψ is a continuous map from a finite-dimensional simplex to itself, and so the Brouwer fixed-point theorem guarantees the existence of a fixed point.

For a given automaton A and child distribution χ , suppose $\vec{v}$ is some fixed point of Ψ . We call ι an *interpretation of the automaton A corresponding to $\vec{v}$* if ι is indeed an interpretation of A and $\iota(T) \sim \vec{v}$. It is not hard to show that up to measure zero changes, there is at most one interpretation corresponding to a given fixed point (see Proposition 1.6). If such an interpretation exists, we call $\vec{v}$ *interpretable*; otherwise, we call it *rogue*. Our main results are a criterion for determining whether a given fixed point is rogue or interpretable when $|\Sigma| = 2$ (Theorem 1.7), as well as a sufficient condition for interpretability for $|\Sigma| \geq 3$ (Theorem 1.8). To state this criterion, we must define the two randomly colored trees explained in the next two sections.

1.6 | The random state tree

Fix a child distribution χ , automaton A , and a fixed point $\vec{v}$ of the resulting automaton distributional map Ψ . The *random state tree associated with $\vec{v}$* is a colored Galton-Watson tree. We write it as (T, ω) , where $\omega : V(T) \rightarrow \Sigma$ is a random coloring of the tree T . It is defined by the following properties:

- (i) $T \sim \text{GW}(\chi)$;
- (ii) for every n , the conditional distribution of $(\omega(v) : v \in L_n)$ given $T|_n$ is i.i.d. $\vec{v}$;
- (iii) ω is almost surely compatible with A .

Proposition 1.5. *These properties uniquely determine the distribution of (T, ω) .*

Essentially, the random state tree is defined up to height n by generating the first n levels of T , coloring the leaves i.i.d. $\vec{v}$, and then coloring the first $n-1$ levels of the tree according to the automaton.

The distributions of colored trees generated by this procedure turn out to be consistent for different values of n , which is a consequence of $\vec{v}$ being a fixed point of Ψ . Kolmogorov's extension theorem then shows the existence of the distribution of the entire colored tree. This is shown in detail in the proof of Proposition 1.5, which we give in Section 2.

The coloring of the vertices of the random state tree is reminiscent of an interpretation, which also yields a coloring of the tree via (4). But note that for a given fixed point of Ψ , the random state tree coloring always exists, and it is a random coloring (on top of the randomness of the tree). On the other hand, given a fixed point of Ψ , there may be no interpretations associated with it; if there is an interpretation, the coloring it yields is deterministic given the tree.

1.7 | Definition of the pivot tree

We now describe the pivot tree, leaving its formal definition to Section 2.3. Consider some vertex of (T, ω) , and imagine changing its color and then recoloring all the vertices above it according to the rule of the automaton. We call this recoloring operation a *switching*. If the switching changes the color at the root, then we call the vertex *pivotal* for (T, ω) . It is not hard to see that a vertex can only be pivotal if its parent is pivotal. The subgraph of T induced by the pivotal vertices is thus a subtree, which we call the *pivot tree* T_{piv} . As we will see in Proposition 2.4, the pivot tree is a multitype Galton-Watson tree.

We mention that the pivot tree is a bit more complicated when there are more than two states, because a vertex can change colors in more than one way. However, to state Theorem 1.8, we need only use the *pivot tree with maximal target set*, in which a vertex is pivotal if its color can be switched to any other color with the result of changing the color of the root in any way.

1.8 | The main result

For all of our results, fix a child distribution χ , an automaton A on a finite set of states Σ , and let $\Psi: D \rightarrow D$ be the automaton distributional map corresponding to A and χ , defined in Section 1.5.

First, as we mentioned, there is at most one interpretation for each fixed point:

Proposition 1.6. *If $\iota, \iota': \mathcal{T} \rightarrow \Sigma$ are interpretations of A under $\text{GW}(\chi)$ corresponding to the same fixed point of the automaton distribution map, then $\iota = \iota'$ a.e.- $\text{GW}(\chi)$.*

Now, we give our main results. Let $\vec{v}$ be a fixed point of Ψ . We assume that the support of the probability distribution $\vec{v}$ is all of Σ ; that is, as a vector, all entries of $\vec{v}$ are nonzero. This is in fact no restriction, since if $\vec{v}$ is supported on a subset of Σ , we can simply remove the extra elements of Σ and view A as an automaton on this smaller set. Recall that the pivot tree associated with $\vec{v}$ is a multitype Galton-Watson tree, which will be proven in Proposition 2.4. We define a multitype Galton-Watson tree to be subcritical, critical, or supercritical depending on whether its matrix of mean offspring sizes has spectral radius smaller than, equal to, or greater than 1 (see Section 2.4).

Theorem 1.7. *Suppose that $|\Sigma| = 2$ and χ has finite logarithmic moment. Then $\vec{v}$ admits an interpretation if and only if the pivot tree associated with $\vec{v}$ is subcritical or critical.*

This theorem completely classifies fixed points as interpretable or rogue when $|\Sigma| = 2$. It is practical to apply (see Section 5 for some examples), since it only takes a computation to check the criticality of a given Galton-Watson tree.

When $|\Sigma| \geq 3$, we give only a sufficient condition for existence of an interpretation.

Theorem 1.8. *If the pivot tree with maximal target set associated with $\vec{v}$ is subcritical, then $\vec{v}$ admits an interpretation.*

Our full version of this result, Proposition 3.1, is actually slightly stronger and applies in some cases when the pivot tree is critical (see Remark 3.6).

1.9 | Connections to other work

This work has some concrete connections with mathematical logic. We start by defining *first-order* and *monadic second-order* logic on trees. A sentence in the first-order language for rooted trees is a finite combination of the following:

- a constant symbol R representing the root;
- a function π where $\pi(v)$ represents the parent of vertex v ;
- a relation $=$, denoting equality of vertices;
- the Boolean connectives;
- existential and universal quantifications over vertices.

For example, a valid first-order sentence is that some vertex has exactly one child, which is expressed in the formal language by

$$\exists x \exists y \left((\pi(y) = x) \wedge (\forall z (\pi(z) = x \implies z = y)) \right).$$

The monadic second-order language adds

- existential and universal quantifications over sets of vertices;
- the relation $\in$, denoting set membership.

For example, the following sentence states that the tree is infinite:

$$\exists S \forall x \left((x \in S) \implies (\exists y (\pi(y) = x) \wedge (y \in S)) \right).$$

The *quantifier depth* of a sentence in either language is the maximal depth of nesting of existential and universal qualifiers. In the example above, the quantifier depth is 3.

Using Ehrenfeucht games, one can partition the set of rooted trees into finitely many types by the relation that two trees have the same type if they have the same truth value for all first-order sentences of quantifier depth at most k (see [17, Chapter 3]). Call this partition the *rank- k types*. One can do the same replacing first-order logic with monadic second-order logic, producing the *MSO rank- k types* [17, Section 7.2]. In both cases, one can deduce the type of a given tree t from the types of the trees rooted at the children of the root of t . This gives rise to tree automata on the set of rank- k and MSO rank- k types, both of which have interpretations given by mapping a tree to its type.

In [24, 25], this automaton is investigated for the first-order case. The most fundamental result of [25] is that its automaton distribution map is a contraction and hence has a unique fixed point. As a consequence, since the at-least-two property has multiple fixed points, the property of a tree containing an infinite binary tree starting from its root cannot be expressed in first-order logic. We discuss this further in Section 5.3. Our initial motivation for this paper was to make sense of the meaning of multiple fixed points.

Our work also has some connections to the theory of *recursive distributional equations* (RDEs) as developed by Aldous and Bandyopadhyay [1]. A prototypical example of an RDE is for the height of

a Galton-Watson tree. Given a child distribution, let N be the number of children of the root. Then the height of the tree H satisfies the distributional equation

$$H \stackrel{d}{=} 1 + \max(H_1, \dots, H_N),$$

where $H_1, H_2, \dots$ are independent copies of H .

For a given tree automaton, the automaton distribution map Ψ defines an RDE. For any choice of fixed point $\vec{v}$, the random state tree (T, ω) is an example of an object introduced by Aldous and Bandyopadhyay called a *recursive tree process* (RTP). RTPs are classified as *endogenous* or *nonendogenous*, which for (T, ω) corresponds to whether $\omega(R_T)$ is measurable with respect to T . In Proposition 2.2, we show that this is equivalent to interpretability of $\vec{v}$. Thus, Theorems 1.7 and 1.8 can be viewed as criteria for the endogeneity of an RTP, for RTPs in a certain class. This work or extensions of it might prove useful, as the endogeneity of RTPs is an actively pursued topic (see [2, 4, 13, 16], for example).

Two very recent papers have a similar flavor as ours. In [6], the authors consider critical Galton-Watson trees conditioned to have n vertices. Each vertex of the tree is given a label from a finite set. The label of a parent is a function of the labels of the children along with an independent set of randomness. (This is also the case with Aldous and Bandyopadhyay's definition of a RTP.) The main result of the paper is a limit theorem for the distribution of the label of the root as $n \rightarrow \infty$.

The paper [20] considers Galton-Watson trees labeled by elements of $[0, 1]$, cut off at level $2n$. Leaves are assigned independent labels sampled uniformly from $[0, 1]$. Then, the label at a parent at an even generation is the minimum of its children's labels; at an odd generation, it is the maximum of its children's labels. This models a game in which two players take turns, one trying to make the score big and one trying to keep it small. The paper classifies possible limit distributions for the label at the root as $n \rightarrow \infty$. It also investigates endogeneity, the question of whether the value at the root is determined by the structure of the tree.

1.10 | Outline

In Section 2, we first establish basic properties of interpretations, fixed points, the random state tree, and the pivot tree used throughout this paper. In Section 3, we prove the first direction of Theorem 1.7, existence of an interpretation when the pivot tree is almost surely finite. The main tool for this is the Kahn-Kalai-Linial inequality from the theory of Boolean functions [14]. The other direction of Theorem 1.7 is proven in Section 4 using the spine decomposition technique pioneered by Lyons, Peres, and Pemantle [18]. Finally, in Section 5, we apply these results to answer Question 1.1. We also give examples exhibiting a phase transition between interpretable and rogue for a fixed point as the child distribution of the tree is varied. In Section 6, we discuss some open questions.

2 | FOUNDATIONAL PROPERTIES OF OUR OBJECTS

In this section, we fix a child distribution χ , an automaton A on a set of states Σ , and a fixed point $\vec{v}$ of the automaton distributional map $\Psi: D \rightarrow D$ determined by A and χ . We will demonstrate some of the basic properties of fixed points, interpretations, the random state tree, and the pivot tree.

2.1 | The random state tree

We now give the proof of Proposition 1.5, establishing the existence of the random state tree (T, ω) defined in Section 1.6. We then show in Proposition 2.1 that it is a multitype Galton-Watson tree.

Proof of Proposition 1.5. To invoke the Kolmogorov extension theorem [15, Theorem 6.16], we must construct a sequence of random variables (T_n, ω_n) such that T_n is the truncation to level n of a $\text{GW}(\chi)$ -distributed tree, the distribution of $(\omega_n(v))_{v \in L_n(T_n)}$ conditional on T_n is i.i.d.- $\vec{v}$, the values of $\omega_n(v)$ for v in levels $0, \dots, n-1$ are as given by the automaton, and the truncation of (T_{n+1}, ω_{n+1}) to n levels is distributed as (T_n, ω_n) . (Formally speaking, to apply the Kolmogorov extension theorem, we view labeled trees as a sequence of their finite truncations, but we will ignore these details.)

To construct (T_n, ω_n) , we simply define T_n as the truncation of a Galton-Watson tree, then color the level n vertices i.i.d.- $\vec{v}$, and then color levels $0, \dots, n-1$ of the tree according to the automaton. The crux of the proof is showing that the truncation of (T_{n+1}, ω_{n+1}) to level n is distributed as (T_n, ω_n) . Clearly, $T_{n+1}|_n$ is distributed as T_n , and the coloring given by ω_{n+1} on levels $0, \dots, n-1$ of $T_{n+1}|_n$ is as induced by the automaton. We need only show that conditional on $T_{n+1}|_n$, the labeling ω_{n+1} assigns i.i.d. $\vec{v}$ colors to the level n vertices.

To see this, recall how we define $\Psi(\vec{x})$: We let a node have children according to distribution χ . Each of these children is assigned, mutually independently, a state according to distribution $\vec{x}$. The induced random state of the parent node, obtained via the rules of A , has distribution $\Psi(\vec{x})$. Meanwhile, each $v \in L_n(T_{n+1})$ has children according to χ , these children receive i.i.d.- $\vec{v}$ labels from ω_{n+1} , and $\omega_{n+1}(v)$ is given by applying the automaton to these labels. Hence, the distribution of $\omega_{n+1}(v)$ conditional on $T_{n+1}|_n$ is $\Psi(\vec{v})$. As $\vec{v}$ is assumed to be a fixed point, this equals $\vec{v}$. The values of $\omega_{n+1}(v)$ are independent for the different level n vertices v conditional on $T_{n+1}|_n$, showing that ω_{n+1} assigns i.i.d. $\vec{v}$ colors to the level n vertices. ■

Now that we have shown the existence of the random state tree, we prove that it is Galton-Watson with types given by ω .

Proposition 2.1. *The random state tree (T, ω) is a multitype Galton-Watson tree.*

Proof. For $\sigma_1, \dots, \sigma_k \in \Sigma$, let

$$\chi_{\text{col}}(\sigma_1, \dots, \sigma_k) = \chi(k) \vec{v}(\sigma_1) \cdots \vec{v}(\sigma_k),$$

the probability that R_T has exactly k children and that their types in order are $\sigma_1, \dots, \sigma_k$. Let $\chi_{\text{col}}^\sigma(\sigma_1, \dots, \sigma_k)$ denote the conditional probability that R_T has exactly k children and that their types in order are $\sigma_1, \dots, \sigma_k$, given that $\omega(R_T) = \sigma$. Thus, if σ is the type according to A for a vertex with children of types $\sigma_1, \dots, \sigma_k$, then

$$\chi_{\text{col}}(\sigma_1, \dots, \sigma_k) = \vec{v}(\sigma) \chi_{\text{col}}^\sigma(\sigma_1, \dots, \sigma_k). \quad (5)$$

Our goal is to prove that conditional on the first n levels of (T, ω) , each vertex v at level n independently gives birth according to the distribution given by $\chi_{\text{col}}^{\omega(v)}$. Fix any $(t, \tau) \in \mathcal{T}_{\text{col}}$. By definition of ω ,

$$\mathbf{P}[(T, \omega) \in [t, \tau]_{n+1}] = \mathbf{P}[T \in [t]_{n+1}] \prod_{u \in L_{n+1}(t)} \vec{v}(\tau(u)),$$

recalling the notation $[t]_n$ and $[t, \tau]_n$ defined in Section 1.2. For a vertex $v \in V(t)$, let $C(v)$ denote its children in t . Since T is Galton-Watson with child distribution χ ,

$$\mathbf{P}[(T, \omega) \in [t, \tau]_{n+1}] = \left(\mathbf{P}[T \in [t]_n] \prod_{v \in L_n(t)} \chi(|C(v)|) \right) \prod_{u \in L_{n+1}(t)} \vec{v}(\tau(u))$$

$$\begin{aligned}
&= \mathbf{P}\left[T \in [t]_n\right] \prod_{v \in L_n(t)} \left(\chi(|C(v)|) \prod_{u \in C(v)} \tilde{v}(\tau(u)) \right) \\
&= \mathbf{P}\left[T \in [t]_n\right] \prod_{v \in L_n(t)} \chi_{\text{col}}(\tau(u)_{u \in C(v)}).
\end{aligned}$$

By (5), this becomes

$$\begin{aligned}
\mathbf{P}\left[(T, \omega) \in [t, \tau]_{n+1}\right] &= \mathbf{P}\left[T \in [t]_n\right] \prod_{v \in L_n(t)} \tilde{v}(\tau(v)) \chi_{\text{col}}^{\tau(v)}(\tau(u)_{u \in C(v)}) \\
&= \mathbf{P}\left[(T, \omega) \in [t, \tau]_n\right] \prod_{v \in L_n(t)} \chi_{\text{col}}^{\tau(v)}(\tau(u)_{u \in C(v)}),
\end{aligned}$$

which is exactly what we set out to prove. $\blacksquare$

2.2 | Equivalent conditions for interpretability of fixed points

We start with a definition that will come up again elsewhere in the paper. Given a rooted tree t and a coloring of its level n vertices, we can repeatedly apply the automaton A to determine the state of the root. We define $A_t^n : \Sigma^{\mathcal{L}_n(t)} \rightarrow \Sigma$ to be the result of doing so, considering it as a map from the colors at level n to a color at the root.

Now, we show that a given fixed point can have at most one interpretation:

Proof of Proposition 1.6. Viewing the statement of the proposition probabilistically, our goal is to show that $\iota(T) = \iota'(T)$ a.s. Fix some $\sigma \in \Sigma$. We first show that for any n ,

$$\mathbf{P}[\iota(T) = \sigma \mid T|_n] = \mathbf{P}[\iota'(T) = \sigma \mid T|_n] \text{ a.s.} \quad (6)$$

To prove this, we start by observing that $\iota(T)$ is determined by $(\iota(T(v)))_{v \in L_n}$. Indeed, since ι is an interpretation of A and thus respects the automaton,

$$\iota(T) = A_T^n \left((\iota(T(v)))_{v \in L_n} \right).$$

Conditional on $T|_n$, each tree $T(v)$ for $v \in L_n$ is independent and distributed identically to T . Let $\vec{v}$ be the fixed point corresponding to ι and ι' . Since the distribution of $\iota(T)$ is $\vec{v}$, the distribution of $(\iota(T(v)))_{v \in L_n}$ conditional on $T|_n$ is i.i.d. $\vec{v}$. Therefore,

$$\mathbf{P}[\iota(T) = \sigma \mid T|_n] = \mathbf{P}\left[A_T^n \left((\omega(v))_{v \in L_n} \right) = \sigma \mid T|_n\right] \text{ a.s.,} \quad (7)$$

recalling that by its definition, the coloring ω of the random state tree (T, ω) also assigns colors to the level n vertices by sampling independently from $\vec{v}$, conditional on $T|_n$. The exact same reasoning shows that

$$\mathbf{P}[\iota'(T) = \sigma \mid T|_n] = \mathbf{P}\left[A_T^n \left((\omega(v))_{v \in L_n} \right) = \sigma \mid T|_n\right] \text{ a.s.,}$$

which proves (6).

Now, we take limits as $n \rightarrow \infty$ to complete the proof. The σ -fields generated by $T|_n$ form a filtration that converges to the σ -field generated by T . Hence, by Lévy's upward theorem,

$$\mathbf{P}[\iota(T) = \sigma \mid T|_n] \rightarrow \mathbf{P}[\iota(T) = \sigma \mid T] = \mathbf{1}_{\{\iota(T) = \sigma\}} \text{ a.s.} \quad (8)$$

and

$$\mathbf{P}[\iota'(T) = \sigma \mid T|_n] \rightarrow \mathbf{P}[\iota'(T) = \sigma \mid T] = \mathbf{1}_{\{\iota'(T) = \sigma\}} \text{ a.s.} \quad (9)$$

By (6), these two limits are identical. We conclude that $\mathbf{1}_{\{\iota(T) = \sigma\}} = \mathbf{1}_{\{\iota'(T) = \sigma\}}$ a.s. for all $\sigma \in \Sigma$. ■

The expression $A_T^n((\omega(v))_{v \in L_n})$ in (7) is equal to $\omega(R_T)$, since the coloring ω is compatible with A . Thus (7) can be written as

$$\mathbf{P}[\iota(T) = \sigma \mid T|_n] = \mathbf{P}[\omega(R_T) = \sigma \mid T|_n] \text{ a.s.,} \quad (10)$$

which will come up again in the next proposition. Before we state it, we mention a standard characterization of measurability [15, Lemma 1.13]: Let X and Y be random variables taking values in measurable spaces $\mathcal{X}$ and $\mathcal{Y}$, respectively, with $\mathcal{X}$ assumed to be a Polish space endowed with its Borel σ -algebra. Then the measurability of X with respect to Y is equivalent to existence of a measurable map $f: \mathcal{Y} \rightarrow \mathcal{X}$ such that $X = f(Y)$ a.s.

Proposition 2.2. *The following statements are equivalent:*

- (i) $\vec{v}$ is interpretable;
- (ii) for each $\sigma \in \Sigma$,

$$\lim_{n \rightarrow \infty} \mathbf{P}[\omega(R_T) = \sigma \mid T|_n] \in \{0, 1\} \text{ a.s.};$$

- (iii) $\omega(R_T)$ is measurable with respect to T ;
- (iv) ω is measurable with respect to T .

Proof that (i) $\implies$ (ii). Let ι be the interpretation of automaton A corresponding to $\vec{v}$ (it is unique up to GW(χ)-negligible sets by Proposition 1.6). By (10),

$$\lim_{n \rightarrow \infty} \mathbf{P}[\omega(R_T) = \sigma \mid T|_n] = \lim_{n \rightarrow \infty} \mathbf{P}[\iota(T) = \sigma \mid T|_n] = \mathbf{1}_{\{\iota(T) = \sigma\}} \in \{0, 1\} \text{ a.s.,}$$

applying Lévy's upward theorem as in (8). ■

Proof that (ii) $\implies$ (iii). Invoking Lévy's upward theorem and then (ii),

$$\mathbf{P}[\omega(R_T) = \sigma \mid T] = \lim_{n \rightarrow \infty} \mathbf{P}[\omega(R_T) = \sigma \mid T|_n] \in \{0, 1\} \text{ a.s.} \quad (11)$$

Thus, given the entire tree T , we can almost surely determine whether $\omega(R_T)$ equals σ or not. Since this is true for every $\sigma \in \Sigma$, the state $\omega(R_T)$ is almost surely equal to a deterministic function of T , showing that $\omega(R_T)$ is measurable with respect to T . ■

Proof that (iii) $\implies$ (iv). Fix any $v \in L_n$. Let $\omega|_{T(v)}$ denote the restriction of ω on the subtree $T(v)$. The conditional distribution of the colored tree $(T(v), \omega|_{T(v)})$ given $T|_n$ is the same as the unconditional distribution of (T, ω) . By (iii), we know that $\omega(v)$ is measurable with respect to $T(v)$ and is hence an almost sure function of $T(v)$. As T has countably many vertices, we can write ω as an almost sure function of T . ■

Proof that (iv) $\implies$ (i). Since ω is measurable with respect to T , so is $\omega(R_T)$. Therefore there exists a measurable map $\iota: \mathcal{T} \rightarrow \Sigma$ such that $\iota(T) = \omega(R_T)$ a.s. We claim that this will serve as the desired interpretation: Since ω is almost surely compatible with A , the assignment $v \mapsto \iota(T(v))$ is also almost surely compatible with A and is hence an interpretation. Furthermore, from the construction of ω , we know that $\omega(R_T)$ will be distributed as $\vec{v}$, and hence so is $\iota(T)$. ■

We mentioned at the end of Section 1.6 that the coloring of T given by ω and the coloring given by an interpretation via (4) are in general different. However, it is a consequence of Proposition 2.2 that when an interpretation exists for a given fixed point, the two colorings are the same:

Corollary 2.3. *The fixed point $\vec{v}$ is interpretable if and only if ω is measurable with respect to T . If this occurs, then $\omega(R_t)$ is determined by t for $\text{GW}(\chi)$ -a.e. $t \in \mathcal{T}$, and the resulting map $\mathcal{T} \rightarrow \Sigma$ given by $t \mapsto \omega(R_t)$ is the unique interpretation corresponding to the fixed point, up to a.e.- $\text{GW}(\chi)$ equivalence.*

Proof. The equivalence of interpretability and measurability of ω with respect to T is one part of Proposition 2.2. In the proof that (iv) implies (i), it is shown that $t \mapsto \omega(R_t)$ yields an interpretation corresponding to the given fixed point. The uniqueness of this interpretation is given by Proposition 1.6. ■

The equivalences proven in this section reduce the question of whether a fixed point of Ψ is rogue or interpretable to whether the coloring ω in the random state tree (T, ω) is random or deterministic given T . This question is on its face no easier than the original one. To answer it, the key will be the pivot tree, a random subtree of (T, ω) that we discuss now.

2.3 | The pivot tree

We start with some notation. Suppose we are given a colored tree (t, τ) with τ compatible with A . Suppose $v \in L_n(t)$. Now, imagine that we change the color of v to some $\gamma \in \Sigma \setminus \{\tau(v)\}$, and then recolor the vertices at levels $0, \dots, n-1$ based on this. We say that we have *switched the color at v to γ* , and we denote the new coloring by $\tau^{v \rightarrow \gamma}$. Note that $\tau^{v \rightarrow \gamma}$ is only defined on $t|_n$, and that it is consistent with the automaton at levels $0, \dots, n-1$.

Now, we give the full definition of the pivot tree. When $|\Sigma| = 2$, this definition is simple: the pivot tree of (t, τ) consists of the subgraph induced by all vertices v such that switching τ at v changes the value of the root. We denote the pivot tree of (T, ω) by T_{piv} , which we will prove shortly is indeed a tree. See Figure 2 for an example.

When $|\Sigma| \geq 3$, we sometimes demand that the color of the root change to one of a specific set of colors, known as the *target set*, complicating the definition. Given (t, τ) with τ compatible with A , let $\mathcal{A} \subseteq \Sigma \setminus \{\tau(R_t)\}$ represent this target set. Given t, τ , and $\mathcal{A}$, for any $v \in V(t)$ we define

$$B_v = \{\gamma \in \Sigma : \tau^{v \rightarrow \gamma}(R_t) \in \mathcal{A}\}.$$

In other words, B_v is the set of colors such that switching v to an element of B_v changes the color of the root to an element of $\mathcal{A}$. For any $v \in V(t)$, we say that v is *pivotal for (t, τ) with target set $\mathcal{A}$* if $B_v \neq \emptyset$.

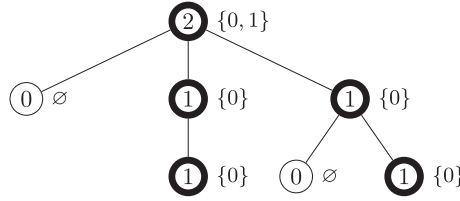


FIGURE 3 The automaton in this example is on $\{0, 1, 2\}$. The state of a parent is given by the sum of its children's states, capped at 2. Bold vertices are pivotal with target set $\{0, 1\}$. Written to the right of each vertex v is the set B_v , indicating which states v can be switched to with the effect of changing the state of the root to a value in the target set

To define the pivot tree of (T, ω) , we must specify a target set for each possible state of the root. For each $\sigma \in \Sigma$, let $\emptyset \neq \mathcal{A}_\sigma \subseteq \Sigma \setminus \{\sigma\}$ be a given (deterministic) set that we call the *target set of the root at state σ* . The most basic example is to set $\mathcal{A}_\sigma = \Sigma \setminus \{\sigma\}$ for all σ , which corresponds to requiring the color of the root to change without caring what it changes to. Let $\mathcal{A} = (\mathcal{A}_\sigma)_{\sigma \in \Sigma}$. We define the pivot tree, $T_{\text{piv}} = T_{\text{piv}}(\mathcal{A})$, as the subgraph of T induced by all vertices pivotal for (T, ω) with target set $\mathcal{A}_{\omega(R_T)}$. The pivot tree is measurable with respect to (T, ω) ; that is, T_{piv} is a measurable function of (T, ω) . Also, observe that this definition works in the $|\Sigma| = 2$ case as well. Here, there is only one possible choice of $\mathcal{A}_\sigma$, and either $B_v = \emptyset$ or B_v is a singleton set made up of the opposite color as $\omega(v)$.

Proposition 2.4. *For given target sets $(\mathcal{A}_\sigma)_{\sigma \in \Sigma}$, assign the type $(\omega(v); B_v)$ to each vertex $v \in V(T)$. With these types, both T and T_{piv} are multitype Galton-Watson trees.*

Proof. We start with proof for T . Let $\mathcal{F}_n$ denote the σ -algebra generated by $T|_n$ and by the types $(\omega(v); B_v)$ for vertices v up to level n . We will refer to these as augmented types, in contrast with the unaugmented types given by ω alone.

We must show that conditional on $\mathcal{F}_n$, the vertices at level n independently give birth according to their augmented types. First, we observe that the values of B_v for v in $T|_n$ are determined by the first n levels of (T, ω) . Hence, conditioning on $\mathcal{F}_n$ is the same as conditioning on the first n levels of (T, ω) . Thus, by Proposition 2.1, conditional on $\mathcal{F}_n$, each vertex v at level n independently gives birth to children whose number and unaugmented type are determined by the unaugmented type of v .

Now, we just need to extend this statement to the augmented types. The key fact is the following: Let $u_1, \dots, u_k$ be the children of some node v . Then for each $i = 1, \dots, k$, the set B_{u_i} is determined by $\omega(u_1), \dots, \omega(u_k)$ and B_v . Indeed, from $\omega(u_1), \dots, \omega(u_k)$, we can determine the effect on the color of v of changing u_i to have any given color. From B_v , we know whether the change will alter the color of the root to have a value in $\mathcal{A}_{\omega(R_T)}$. Thus we can determine B_{u_i} .

Let $C(v)$ denote the children of a vertex v , as in Proposition 2.1. From the fact above, conditional on $\mathcal{F}_n$, the distribution of $(B_u : u \in C(v))$ for any $v \in L_n$ is determined by $(\omega(v); B_v)$. This completes the proof that T is multitype Galton-Watson with the augmented types.

To prove the statement for T_{piv} , we first observe that T_{piv} is indeed a tree, since if a vertex u has $B_u \neq \emptyset$, then its parent v evidently satisfies $B_v \neq \emptyset$. Thus, T_{piv} is the tree formed by ignoring vertices of certain types in the Galton-Watson tree T , which always creates another Galton-Watson tree. ■

See Figure 3 for an example of a pivot tree when $|\Sigma| \geq 3$. In general, when we refer to T_{piv} as a Galton-Watson tree from now on, we mean with types given as in Proposition 2.4. When $|\Sigma| = 2$, since either $B_v = \emptyset$ or B_v is a singleton set for each v , we can think of the type $(\omega(v); B_v)$ as simply

$\omega(v)$ along with an indicator on v being pivotal. Thus T_{piv} in this case is Galton-Watson with the types given by ω alone. Naively, one might think that (T_{piv}, ω) would be Galton-Watson even when $|\Sigma| \geq 3$. We can see the problem with this in Figure 3. Let u be the 0-labeled vertex on the bottom level of the tree, and let v be its parent. Vertex u is not pivotal for the given target set of the root (or indeed, for any possible target set). However, for the subtree rooted at v , vertex u is pivotal for the target set $\{0, 2\}$. Thus, if we do not include the sets B_v in the information given by the types, the law of the progeny of a vertex would depend not just on the type of the vertex but on its ancestors.

2.4 | Regularity properties of the pivot tree

For a given multitype Galton-Watson tree, define a matrix by setting M_{ij} to the expected number of offspring of type j for a parent of type i . We classify the process as subcritical, critical, or supercritical depending on whether the spectral radius of M is smaller than, equal to, or greater than 1. If M^n has strictly positive entries for some choice of n , then the Galton-Watson process is called *positive regular*. This says that it is possible for any type to have a descendant of any other type, and that no periodic behavior occurs. The process is called *singular* if each type gives birth to exactly one child with probability one. Multitype Galton-Watson trees are nearly always considered under the assumption that they are positive regular and nonsingular. Under this assumption, the process dies out with probability one in the subcritical and critical cases, and it survives with positive probability in the supercritical case. Regardless of the starting type, the expected size of the n th generation vanishes exponentially in the subcritical case; remains of constant order in the critical case; and grows exponentially in the supercritical case.

For a Galton-Watson tree without these assumptions, the situation is messier. To illustrate, consider a process with two types A and B and matrix of means $M = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$ for $a > 0$. The expected number of vertices of each type at level n starting with a vertex of type A is given by the first row of M^n , which is $(1, an)$. Thus, even though this process is critical, the expected size of the n th generation grows to infinity, though only at a polynomial rate. On the other hand, this tree still dies out with probability one, as we can see by viewing it as a backbone of a critical single-type Galton-Watson tree of vertices of type A , each of which gives birth to critical single-type trees of vertices of type B , all of which die out with probability one.

In general, without the assumption of positive regularity and nonsingularity, it is still correct that a subcritical tree has exponentially vanishing expected n th generation and hence dies out almost surely. By [26] (see [12, Theorem 10.1]), so long as there does not exist a collection of types C such that the children of a vertex of type in C include exactly one of the types in C with probability one, a critical tree dies out almost surely; and a supercritical tree survives with positive probability from some starting state.

The Galton-Watson tree T_{piv} need not be positive regular. Nonetheless, when $|\Sigma| = 2$, many features of positive regularity still hold. We give a lemma that we will use to prove this.

Lemma 2.5. *Suppose that $\Sigma = \{0, 1\}$. Let Z_0 and Z_1 be the number of children of R_T pivotal for (T, ω) of types 0 and 1, respectively. Then*

$$\mathbf{E}[Z_0 \mid \omega(R_T) = 0] = \mathbf{E}[Z_1 \mid \omega(R_T) = 1],$$

and

$$\mathbf{E}[Z_0 \mid \omega(R_T) = 1] = \frac{\bar{v}(0)^2}{\bar{v}(1)^2} \mathbf{E}[Z_1 \mid \omega(R_T) = 0].$$

Hence, if M is the matrix of means of T_{piv} , given by $M = (m_{ij})_{i,j \in \{0,1\}}$ where $m_{ij} = \mathbf{E}[Z_j \mid \omega(R_T) = i]$, then

$$M = \begin{pmatrix} m_{00} & m_{01} \\ \frac{\vec{v}(0)^2}{\vec{v}(1)^2} m_{01} & m_{00} \end{pmatrix}. \quad (12)$$

Proof. Given a list $\sigma = (\sigma_1, \dots, \sigma_k) \in \{0, 1\}^k$ representing the states of an ordered set of children, we abuse notation slightly and write $A(\sigma)$ to mean the value that the automaton assigns to the parent given these children. For example, if $\sigma = (0, 0, 1, 0, 1)$, then we write $A(\sigma)$ to denote $A(3, 2)$, the type of the parent when there are three children of type 0 and two of type 1. We say that coordinate σ_i is pivotal if switching its value changes $A(\sigma)$. For example, if A is the at-least-two automaton of Example 1.3 and σ is as above, then σ_3 and σ_5 are pivotal.

For $a, b \in \{0, 1\}$, let

$$S_k(a, b) = \left\{ (\sigma, i) : \sigma \in \{0, 1\}^k, i \in \{1, \dots, k\}, A(\sigma) = a, \right. \\ \left. \sigma_i = b, \text{ and } \sigma_i \text{ is pivotal} \right\},$$

representing a configuration of k children making the parent have type a and a choice of a pivotal child of type b . There is a natural bijection between $S_k(a, b)$ and $S_k(1-a, 1-b)$. The map is given by sending $(\sigma, i) \in S_k(a, b)$ to $(\sigma', i) \in S_k(1-a, 1-b)$, where σ' is equal to σ except at coordinate i . Applying this bijection, keeping in mind that the states of the level 1 vertices of (T, ω) conditional on $T|_1$ are i.i.d. $\vec{v}$,

$$\begin{aligned} \mathbf{E}[Z_0 \mathbf{1}\{\omega(R_T) = a\} \mid R_T \text{ has } k \text{ children}] &= \sum_{(\sigma, i) \in S_k(a, 0)} \vec{v}^{\otimes n}(\sigma) \\ &= \sum_{(\sigma, i) \in S_k(1-a, 1)} \frac{\vec{v}(0)}{\vec{v}(1)} \vec{v}^{\otimes n}(\sigma) \\ &= \frac{\vec{v}(0)}{\vec{v}(1)} \mathbf{E}[Z_1 \mathbf{1}\{\omega(R_T) = 1-a\} \mid R_T \text{ has } k \text{ children}]. \end{aligned}$$

Here we use the notation $\vec{v}^{\otimes n}$ to denote the n -fold product measure of $\vec{v}$ with itself. Taking expectations, in the $a = 0$ case this yields

$$\frac{\mathbf{E}[Z_0 \mathbf{1}\{\omega(R_T) = 0\}]}{\vec{v}(0)} = \frac{\mathbf{E}[Z_1 \mathbf{1}\{\omega(R_T) = 1\}]}{\vec{v}(1)},$$

while in the $a = 1$ case it yields

$$\frac{\mathbf{E}[Z_0 \mathbf{1}\{\omega(R_T) = 1\}]}{\vec{v}(1)} = \left(\frac{\vec{v}(0)^2}{\vec{v}(1)^2} \right) \frac{\mathbf{E}[Z_1 \mathbf{1}\{\omega(R_T) = 0\}]}{\vec{v}(0)}.$$

■

This lets us prove that when $|\Sigma| = 2$, the pivot tree behaves nicely. In particular, at criticality T_{piv} dies out and has expected size one at every generation.

Proposition 2.6. *Suppose that $\Sigma = \{0, 1\}$ and that both entries of $\vec{v}$ are positive. Let $M = (m_{ij})_{i,j \in \{0,1\}}$ be the matrix of means of T_{piv} .*

- (a) The largest eigenvalue of M in absolute value is equal to $\mathbf{E}[\ell_1(T_{\text{piv}})]$.
- (b) For all n , it holds that $\mathbf{E}[\ell_n(T_{\text{piv}})] = \mathbf{E}[\ell_1(T_{\text{piv}})]^n$.
- (c) If T_{piv} is supercritical, then it is infinite with positive probability conditional on both $\omega(R_T) = 0$ and on $\omega(R_T) = 1$.
- (d) If T_{piv} is critical, then it is finite with probability one.

Proof of (a). By (12) from Lemma 2.5, the characteristic polynomial of M is

$$(x - m_{00})^2 - \frac{\bar{v}(0)^2}{\bar{v}(1)^2} m_{01}^2,$$

which has roots $m_{00} \pm \frac{\bar{v}(0)}{\bar{v}(1)} m_{01}$. The larger of these is $m_{00} + \frac{\bar{v}(0)}{\bar{v}(1)} m_{01}$. We then compute

$$\begin{aligned} \mathbf{E}[\ell_1(T_{\text{piv}})] &= \mathbf{E}[Z_0 + Z_1] = \bar{v}(0)(m_{00} + m_{01}) + \bar{v}(1)(m_{10} + m_{11}) \\ &= \bar{v}(0)(m_{00} + m_{01}) + \bar{v}(1) \left(\frac{\bar{v}(0)^2}{\bar{v}(1)^2} m_{01} + m_{00} \right) \\ &= (\bar{v}(0) + \bar{v}(1)) m_{00} + \frac{\bar{v}(0)(\bar{v}(1) + \bar{v}(0))}{\bar{v}(1)} m_{01} \\ &= m_{00} + \frac{\bar{v}(0)}{\bar{v}(1)} m_{01}. \end{aligned}$$

Proof of (b). The value of $\mathbf{E}[\ell_n(T_{\text{piv}})]$ is the sum of entries of the vector $\bar{v}M^n$. We can confirm by hand that $\bar{v}$ is a left eigenvector of M corresponding to the eigenvalue $\mathbf{E}[\ell_1(T_{\text{piv}})]$, from which the statement follows.

There is a more conceptual explanation for this, which we briefly sketch. Let v be a vertex at level n of T , and consider the following question: conditional on $T|_n$ and on v being pivotal, what is the distribution of $\omega(v)$? The answer is $\bar{v}$, just as if we had not conditioned on v being pivotal. This is because switching the color of v yields a bijection between colorings in which v is pivotal with color 0 and pivotal with color 1, with a ratio $\bar{v}(0)/\bar{v}(1)$ of probabilities of each corresponding state under the product measure $\bar{v}^{\otimes \ell_n}$. Thus, pivotal vertices are colored by $\bar{v}$, and so the expected number of pivotal children of a pivotal vertex is $\mathbf{E}[\ell_1(T_{\text{piv}})]$. Iterating this and applying linearity of expectation yields $\mathbf{E}[\ell_n(T_{\text{piv}})] = \mathbf{E}[\ell_1(T_{\text{piv}})]^n$.

Proof of (c). We consider two cases. First, suppose that $m_{01} = m_{10} = 0$. By Lemma 2.5, the matrix M has the form $\begin{bmatrix} m_{00} & 0 \\ 0 & m_{00} \end{bmatrix}$, and by our supercriticality assumption $m_{00} \geq 1$. Hence, T_{piv} conditional on either $\omega(R_T) = 0$ or $\omega(R_T) = 1$ is a supercritical single-type Galton-Watson tree, and it survives in both cases with positive probability.

Now, suppose it is not true that $m_{01} = m_{10} = 0$. Since the multitype Galton-Watson tree (T_{piv}, ω) is supercritical, it survives with positive probability from some starting state. Hence at least one of the two probabilities $\mathbf{P}[T_{\text{piv}} \text{ survives} \mid \omega(R_T) = 0]$ and $\mathbf{P}[T_{\text{piv}} \text{ survives} \mid \omega(R_T) = 1]$ must be positive. By Lemma 2.5, both m_{01} and m_{10} are positive. Thus, the root of T_{piv} conditioned to be type 0 has positive probability of giving birth to a pivotal vertex of type 1, and vice versa. Therefore if either of $\mathbf{P}[T_{\text{piv}} \text{ survives} \mid \omega(R_T) = 0]$ or $\mathbf{P}[T_{\text{piv}} \text{ survives} \mid \omega(R_T) = 1]$ is positive, then both of them are.

Proof of (d). As in the previous proof, we break the proof into two cases depending on whether $m_{01} = m_{10} = 0$. If so, then T_{piv} conditional on either $\omega(R_T) = 0$ or $\omega(R_T) = 1$ is a critical single-type Galton-Watson tree, which dies out with probability one unless it is singular. To rule this out suppose that a vertex of type 0 gives birth to a single pivotal vertex of type 0 with probability one. Then in

particular, a vertex of type 0 always gives birth to exactly one child of type 0, since all children of a given type have the same pivotal status. Now, we claim that a vertex of type 0 cannot give birth to any vertices of type 1. Indeed, they would be nonpivotal, and hence switching one of them would yield another configuration with multiple children of type 0 but still with a type 0 root. (Note that we have assumed that $\vec{v}$ puts positive probability on both types, meaning that the configuration after the switching still has positive probability of occurring.) Hence, a vertex of type 0 gives birth almost surely to exactly one child, which has type 0. Thus, we have deduced the automaton: it assigns a parent type 0 if and only if there is exactly one child, which has type 0. Since $\vec{v}$ is a fixed point, it satisfies $\vec{v}(0) = \chi(1)\vec{v}(0)$. But then $\vec{v}(0) \in \{0, 1\}$, contradicting our assumption that $\vec{v}$ places positive probability on both types. The same argument also shows that a vertex of type 1 does not give birth to exactly one child of type 1 in the $m_{01} = m_{10} = 0$ case.

Now, consider the case that m_{01} and m_{10} are nonzero. According to [12, Theorem 10.1], we must show that for the pivot tree, there does not exist a collection of states C such that the children of a vertex of type in C almost surely include exactly one with type in C . Suppose there exists such a set C . If $C = \{0\}$, then $m_{00} = 1$. But as the highest eigenvalue of M is $m_{00} + \frac{\vec{v}(0)}{\vec{v}(1)}m_{01}$ and m_{01} is assumed to be nonzero, T_{piv} is not critical. The same argument rules out $C = \{1\}$. If $C = \{0, 1\}$, then every vertex (of whatever type) gives birth to exactly one pivotal vertex almost surely. Since all children of the same type have the same pivotality status, this implies that every vertex must give birth almost surely to a unique child (i.e., one whose type is the opposite of all of its siblings). But this can happen only if χ is supported on $\{0, 1\}$, since otherwise choosing the number of children according to χ and then coloring them i.i.d. $\vec{v}$, there is positive probability that they all are colored the same. But this is a contradiction, since χ is assumed to assign positive weight to $\{2, 3, \dots\}$. ■

3 | SUBCRITICAL PIVOT TREES

As in Section 2, throughout this section we fix a child distribution χ , an automaton A on a finite set of states Σ , and a fixed point $\vec{v}$ of the automaton distributional map $\Psi : D \rightarrow D$ corresponding to A and χ . We let (T, ω) be the random state tree for $\vec{v}$. As usual, we let T_{piv} denote the pivot tree for (T, ω) , but in this section we fix the maximal target set $\mathcal{A}_\sigma = \Sigma \setminus \{\sigma\}$ for $\sigma \in \Sigma$. Throughout this section, when we refer to a vertex as pivotal for (T, ω) , we mean that it is pivotal with this target set (see Section 2.3). Recall from Proposition 2.4 that T_{piv} is a Galton-Watson tree with the types defined there. Our goal in this section is to prove the following:

Proposition 3.1. *Suppose that T_{piv} is almost surely finite and that $\mathbf{E}\ell_n(T_{\text{piv}}) \leq 1$ for all sufficiently large n . Then $\vec{v}$ is interpretable.*

This condition on T_{piv} holds when it is subcritical, and when $|\Sigma| = 2$ it also holds when T_{piv} is critical, as discussed in Section 2.4.

Our proof will use the theory of Boolean functions and influences (see [11, 21]). We first introduce some ideas and results from this theory, starting with pivotality in the context of Boolean functions. For a function $g : \Sigma^m \rightarrow \{0, 1\}$, we say that the i th coordinate is pivotal for g at $(s_1, \dots, s_m)$ if the map

$$s \mapsto g(s_1, \dots, s_{i-1}, s, s_{i+1}, \dots, s_m)$$

is nonconstant. To relate this to our earlier definition of a pivotal vertex in Section 2.3, recall the map $A_t^n : \Sigma^{\ell_n(t)} \rightarrow \Sigma$ defined in Section 2.2, which gives the color at the root of t according to the automaton

A as a function of the colors at level n . For some fixed $\sigma \in \Sigma$, define $g_{t,n} : \Sigma^{\ell_n(t)} \rightarrow \{0, 1\}$ by

$$(\sigma_1, \dots, \sigma_{\ell_n(t)}) \mapsto \mathbf{1}\{A_t^n(\sigma_1, \dots, \sigma_{\ell_n(t)}) = \sigma\}. \quad (13)$$

Then every pivotal coordinate for $g_{T,n}$ at $(\omega(v))_{v \in L_n(T)}$ is a pivotal vertex for (T, ω) . We mention that the converse is false: not every pivotal vertex for ω is a pivotal coordinate, because changing the label of the vertex might change the label of the root from one element of $\Sigma \setminus \{\sigma\}$ to another, leaving $g_{T,n}$ the same either way.

The *influence* of the i th coordinate of a map $g : \Sigma^m \rightarrow \{0, 1\}$, denoted by $I_i(g)$, is the probability that the i th coordinate is pivotal for $(S_1, \dots, S_m)$, where $S_1, \dots, S_m$ are independent and identically distributed as $\vec{v}$. The total influence, $I(g)$, is the sum of the influences of all the coordinates, or equivalently the expected number of pivotal coordinates for g at $(S_1, \dots, S_m)$.

The following is a variant of the BKKKL inequality [5, Theorem 1], which is itself a variant of the KKL inequality [14].

Proposition 3.2 (Theorem 3.4 from [10]). *There exists a universal constant $c > 0$ such that the following holds. Let $g : \Sigma^n \rightarrow \{0, 1\}$ be an arbitrary map, and let $p = \mathbf{P}[g(S_1, \dots, S_n) = 1]$, where $S_1, \dots, S_n$ are independent and distributed as $\vec{v}$. Then*

$$I(g) \geq c \min(p, 1 - p) \log \left(\frac{1}{\max_i I_i(g)} \right).$$

Thus, if the total and maximum influences are small, then $\min(p, 1 - p)$ is small, meaning that g is nearly constant. Our idea is to apply this to the map $g_{T,n}$ introduced in (13), which will then show that criterion (ii) of Proposition 2.2 is satisfied and hence $\vec{v}$ is interpretable.

For the rest of this section, we fix an arbitrary state $\sigma \in \Sigma$ and consider $g_{t,n}$ as defined in (13). Define

$$I_n(t) = \mathbf{E}[\ell_n(T_{\text{piv}}) \mid T|_n = t|_n].$$

When we consider the random state tree (T, ω) up to level n , there are two sources of randomness: the tree itself, which is Galton-Watson, and the colors, which are determined by coloring the level n vertices i.i.d. $\vec{v}$. We obtain $I_n(t)$ by taking an expectation only over this second source of randomness, with the structure of the tree fixed. In other words, if the level n vertices of the deterministic tree t are colored i.i.d. $\vec{v}$, then $I_n(t)$ is the expected number of these vertices that are pivotal. Thus, $I_n(T)$ is the expected number of pivotal vertices for (T, ω) conditional on $T|_n$. Since a level n vertex of T is pivotal for ω if the corresponding coordinate of $g_{T,n}$ is pivotal at $(\omega(v))_{v \in L_n(T)}$, we have $I(g_{T,n}) \leq I_n(T)$.

For a given tree t , let

$$I_n^{\max}(t) = \max_{v \in L_n(t)} \mathbf{P}[v \in T_{\text{piv}} \mid T|_n = t|_n].$$

Observe that $I_n(t)$ has the same definition except that a sum replaces the maximum. Just as $I(g_{T,n}) \leq I_n(T)$, we have $\max_i I_i(g_{T,n}) \leq I_n^{\max}(T)$.

Lemma 3.3. *If T_{piv} is almost surely finite, then $I_n^{\max}(T) \rightarrow 0$ a.s. as $n \rightarrow \infty$.*

Proof. We will show this by proving that

$$I_n^{\max}(T) \leq \mathbf{P}[T_{\text{piv}} \text{ survives to height } n \mid T|_n] \text{ a.s.} \quad (14)$$

and

$$\mathbf{P}[T_{\text{piv}} \text{ survives to height } n \mid T|_n] \rightarrow 0 \text{ a.s.} \quad (15)$$

as $n \rightarrow \infty$.

For the first claim, we start with the observation that for any $v \in L_n(T)$,

$$\mathbf{P}[v \in T_{\text{piv}} \mid T|_n] \leq \mathbf{P}[T_{\text{piv}} \text{ survives to height } n \mid T|_n] \text{ a.s.,}$$

since $v \in T_{\text{piv}}$ implies that T_{piv} survives to height n . Since

$$I_n^{\max}(T) = \max_{v \in L_n(T)} \mathbf{P}[v \in T_{\text{piv}} \mid T|_n],$$

this proves (14).

Now we turn to (15). As $n \rightarrow \infty$,

$$\mathbf{P}[T_{\text{piv}} \text{ survives to height } n] \rightarrow 0,$$

since T_{piv} is almost surely finite. Hence the convergence in (15) holds in L^1 . To get the almost sure convergence, we show that

$$\mathbf{P}[T_{\text{piv}} \text{ survives to height } n \mid T|_n] \quad (16)$$

is a supermartingale, which is more trivial than it looks at first glance. If T_{piv} survives to height $n+1$, then it survives to height n . Hence,

$$\begin{aligned} \mathbf{P}[T_{\text{piv}} \text{ survives to height } n+1 \mid T|_n, \omega|_{T|_n}] &\leq \mathbf{P}[T_{\text{piv}} \text{ survives to height } n \mid T|_n, \omega|_{T|_n}] \\ &= \mathbf{1}\{T_{\text{piv}} \text{ survives to height } n\}. \end{aligned}$$

Taking conditional expectations,

$$\mathbf{P}[T_{\text{piv}} \text{ survives to height } n+1 \mid T|_n] \leq \mathbf{P}[T_{\text{piv}} \text{ survives to height } n \mid T|_n].$$

Finally,

$$\begin{aligned} \mathbf{E}[\mathbf{P}[T_{\text{piv}} \text{ survives to height } n+1 \mid T|_{n+1}] \mid T|_n] \\ = \mathbf{P}[T_{\text{piv}} \text{ survives to height } n+1 \mid T|_n]. \end{aligned}$$

Altogether, this shows that

$$\mathbf{E}[\mathbf{P}[T_{\text{piv}} \text{ survives to height } n+1 \mid T|_{n+1}] \mid T|_n] \leq \mathbf{P}[T_{\text{piv}} \text{ survives to height } n \mid T|_n],$$

proving that (16) is a supermartingale. Thus it has an almost sure limit, which must coincide with the L^1 limit. This proves (15), which completes the proof. ■

Next, we give two easy technical lemmas to be used in the proof of Proposition 3.1.

Lemma 3.4. *Let X_n and Y_n be nonnegative random variables, and suppose that $\mathbf{E}X_n \leq 1$ for all n and $Y_n \rightarrow \infty$ a.s. Let $Z_n = \min(X_n/Y_n, 1)$. Then $\mathbf{E}Z_n \rightarrow 0$.*

Proof. Fix some large N . We then compute

$$\begin{aligned} \mathbf{E}Z_n &= \mathbf{E}[Z_n \mathbf{1}\{Y_n \geq N\}] + \mathbf{E}[Z_n \mathbf{1}\{Y_n < N\}] \\ &\leq \mathbf{E}[X_n/N] + \mathbf{P}[Y_n < N] \\ &\leq 1/N + \mathbf{P}[Y_n < N]. \end{aligned}$$

Since $\mathbf{P}[Y_n < N] \rightarrow 0$ as $n \rightarrow \infty$, we have $\limsup_{n \rightarrow \infty} \mathbf{E}Z_n \leq 1/N$. This holds for arbitrarily large values of N , confirming that $\mathbf{E}Z_n \rightarrow 0$. ■

Lemma 3.5. *Suppose that $(X_n)_{n \geq 0}$ takes values in $[0, 1]$ and is a martingale under some filtration $(\mathcal{F}_n)_{n \geq 0}$. Then $(\min(X_n, 1 - X_n))_{n \geq 0}$ is a supermartingale under the same filtration.*

Proof. Let $Y_n = \min(X_n, 1 - X_n)$. Since $Y_{n+1} \leq X_{n+1}$ and $Y_{n+1} \leq 1 - X_{n+1}$, we can take expectations to get

$$\mathbf{E}[Y_{n+1} \mid \mathcal{F}_n] \leq \mathbf{E}[X_{n+1} \mid \mathcal{F}_n] = X_n, \quad (17)$$

and

$$\mathbf{E}[Y_{n+1} \mid \mathcal{F}_n] \leq [1 - \mathbf{E}X_{n+1} \mid \mathcal{F}_n] = 1 - X_n. \quad (18)$$

On the event $X_n \leq 1/2$, which is measurable with respect to $\mathcal{F}_n$, Equation (17) gives $\mathbf{E}[Y_{n+1} \mid \mathcal{F}_n] \leq Y_n$, since $Y_n = X_n$. On the complement of this event, (18) gives $\mathbf{E}[Y_{n+1} \mid \mathcal{F}_n] \leq Y_n$, since $X_n = 1 - Y_n$. Hence $\mathbf{E}[Y_{n+1} \mid \mathcal{F}_n] \leq Y_n$ holds in both cases, proving that (Y_n) is a supermartingale. ■

Proof of Proposition 3.1. We will check that criterion (ii) of Proposition 2.2 holds. Fix a color σ in Σ . Let

$$p(T|_n) = \mathbf{P}[\omega(R_T) = \sigma \mid T|_n].$$

Our goal is to show that $p(T|_n)$ converges almost surely to 0 or 1. We can assume that σ is in the support of $\vec{v}$, since otherwise $p(T|_n) = 0$ a.s. for all n . Consider $g_{T,n}$ as defined in (13). Observe that

$$\mathbf{1}\{\omega(R_T) = \sigma\} = g_{T,n}((\omega(v))_{v \in L_n(T)}),$$

and that the conditional distribution of $(\omega(v))_{v \in L_n(T)}$ given $T|_n$ is i.i.d. $\vec{v}$. We thus apply Proposition 3.2 conditionally on $T|_n$ to obtain

$$I(g_{T,n}) \geq c \min(p(T|_n), 1 - p(T|_n)) \log \left(\frac{1}{\max_i I_i(g_{T,n})} \right).$$

Rearranging this, we obtain

$$\min(p(T|_n), 1 - p(T|_n)) \leq \frac{I(g_{T,n})}{c \log \frac{1}{\max_i I_i(g_{T,n})}} \leq \frac{I_n(T)}{c \log \frac{1}{I_n^{\max}(T)}}. \quad (19)$$

Now, we show that $\min(p(T|_n), 1 - p(T|_n))$ converges to 0 in L^1 as $n \rightarrow \infty$. Let $X_n = I_n(T)$ and $Y_n = c \log \frac{1}{T_n^{\max}(T)}$. Then $\mathbf{E}X_n$ is the expected number of pivotal vertices for (T, ω) at level n , which by assumption is bounded by 1. Since T_{piv} is assumed to be almost surely finite, Lemma 3.3 shows $Y_n \rightarrow \infty$ a.s. By Lemma 3.4, it holds that $\mathbf{E} \min(X_n/Y_n, 1) \rightarrow 0$. Since $\min(p(T|_n), 1 - p(T|_n)) \leq \min(X_n/Y_n, 1)$, this proves that $\min(p(T|_n), 1 - p(T|_n))$ converges to 0 in L^1 .

Since $p(T|_n)$ is a martingale, Lemma 3.5 shows that $\min(p(T|_n), 1 - p(T|_n))$ is a supermartingale. Hence it has an almost sure limit, which must match its L^1 limit of 0. This proves that $\lim p(T|_n) \in \{0, 1\}$ a.s. By Proposition 2.2, the fixed point $\vec{v}$ is interpretable. ■

Proof of Theorem 1.7 ($\Leftarrow$). Subcritical Galton-Watson trees have exponentially vanishing expected n th generation size and are almost surely finite (see Section 2.4). Hence the conditions of Proposition 3.1 hold in the subcritical case. In the critical case, $\mathbf{E}\ell_n(T_{\text{piv}}) = 1$ and T_{piv} is almost surely finite by parts (b) and (d) of Proposition 2.6. ■

Proof of Theorem 1.8. If T_{piv} is subcritical, then the conditions of Proposition 3.1 hold. ■

Remark 3.6. While we have stated Theorem 1.8 for subcritical pivot trees only, Proposition 3.1 in fact applies to critical pivot trees, so long as $\mathbf{E}\ell_n(T_{\text{piv}}) \leq 1$ (any constant bound would also work). As discussed in Section 2.4, Galton-Watson trees that are not positive regular can have their expected n th generation size grow to infinity even in the critical case. However, even though pivot trees are not necessarily positive regular, we are not sure if it is possible for $\mathbf{E}\ell_n(T_{\text{piv}})$ to grow to infinity when T_{piv} is critical.

4 | SUPERCRITICAL PIVOT TREES

As usual, throughout this section we fix a child distribution χ , an automaton A on a set of states Σ , and a fixed point $\vec{v}$ of the automaton distributional map $\Psi : D \rightarrow D$ corresponding to A and χ , and we let (T, ω) be the random state tree for $\vec{v}$.

Our goal is to prove that if the pivot tree is supercritical, then $\vec{v}$ is rogue in the $|\Sigma| = 2$ case. According to Proposition 2.2, rogueness of $\vec{v}$ is equivalent to nonmeasurability of ω with respect to T . Thus, we will try to show that for T in some class of trees of positive weight under the $\text{GW}(\chi)$ measure, the coloring ω is nondeterministic. The idea of the proof is that T_{piv} is supercritical, it occurs with positive probability that $\omega(R_T) = 0$ and T_{piv} survives. On this event, we randomly choose an infinite path starting from the root of T_{piv} and switch all the colors along it. This gives us a new colored tree with the same underlying tree but a different color at the root. Since the new coloring of the tree only differs at one vertex per level, it seems intuitive that it occurs with similar likelihood as the original one, meaning that ω takes different values for the same tree with positive probability.

The difficulty lies in making rigorous the idea that the switched coloring has similar probability as the original one. To do so, we use *spine decompositions* as developed by Lyons, Pemantle, and Peres [18], an elegant probabilistic method for proving two branching processes absolutely continuous or mutually singular to each other. The two processes we consider are (T, ω) , conditioned on survival of the pivot tree, and the switched version of this process described above. We prove the switched version is absolutely continuous with respect to the original. Under the assumption that $\vec{v}$ is interpretable, it is a probability one event that the color $\omega(R_T)$ is given as a deterministic function of T . By absolute continuity, the color of the root in the switched process is equal to the same function of the tree. But this is a contradiction, as we know that these colors differ while the trees are the same.

Let our set of colors be $\Sigma = \{0, 1\}$. Recall from the discussion after Proposition 2.4 that the definition of the pivot tree is simpler in the two-color case. A vertex is pivotal for (T, ω) if swapping its color results in the root swapping colors, and the pivot tree T_{piv} can be defined as the subtree of T consisting of all pivotal vertices. The pivot tree is Galton-Watson with types given by ω , with no need for the sets $\mathcal{A}$ and B_v used in the definitions when there are three or more colors.

We now formalize this and add an extra bit of information to the types, extending ω to a map $\omega_* : V(T) \rightarrow \{0d, 0s, 1d, 1s\}$ as follows. For a vertex v , the 0 or 1 in $\omega_*(v)$ is given by $\omega(v)$. To decide on d or s, consider $(T(v), \omega|_{T(v)})$, the restriction of the random state tree to v and its descendants. If this tree has an infinite pivot tree, then $\omega_*(v)$ assigns type s, for *survives*. If this tree has a finite pivot tree, then $\omega_*(v)$ assigns type d, for *dies*. For $v \in T_{\text{piv}}$, this is equivalent to assigning either s or d depending on whether T_{piv} restricted to v and its descendants is infinite or finite. We will refer to vertices as s-labeled or d-labeled according to the value assigned to them by ω_* . Define $\vec{v}_*$ as the distribution of $\omega_*(R_T)$, a probability measure on $\{0d, 0s, 1d, 1s\}$. Let $\mathcal{T}_{\text{col}*}$ denote the space of trees with vertices labeled $\{0d, 0s, 1d, 1s\}$. For $(t, \tau_*) \in \mathcal{T}_{\text{col}*}$, let $[t, \tau_*]_n$ denote the subset of $\mathcal{T}_{\text{col}*}$ made up of trees agreeing with (t, τ_*) up to the n th generation.

- Proposition 4.1.** (a) Conditional on $T|_n$, the distribution of $(\omega_*(v))_{v \in L_n(T)}$ is i.i.d. $\vec{v}_*$.
 (b) For $i \in \{0, 1\}$, let $\rho(i)$ be the probability that T_{piv} survives conditional on $\omega(R_T) = i$. Then conditional on $T|_n$ and on $\omega|_{T|_n}$, the s- and d-labels given to each vertex $v \in L_n$ by ω_* are independent, with v receiving an s-label with probability $\rho(\omega(v))$.
 (c) The labeled tree (T, ω_*) is multitype Galton-Watson.

Proof. Given $T|_n$, the distribution of $(\omega(v))_{v \in L_n(T)}$ is i.i.d. $\vec{v}$, by definition of (T, ω) . Hence, conditional on $T|_n$, the trees $(T(v), \omega|_{T(v)})$ for $v \in L_n(T)$ are independent and distributed as the (unconditional) distribution of (T, ω) . Since $\vec{v}_*$ is the distribution of $\omega_*(R_T)$, it is thus the conditional distribution given $T|_n$ of each of the independent $\omega_*(v)$ for $v \in L_n(T)$, proving (a). For (b), once we have conditioned on $T|_n$ and on $\omega|_{T|_n}$, for each $v \in L_n(T)$, the tree $(T(v), \omega|_{T(v)})$ is distributed as (T, ω) conditional on having state $\omega(v)$ at the root. Thus the pivot tree of $(T(v), \omega|_{T(v)})$ survives with probability $\rho(\omega(v))$. The s- or d-label for v depends only on $(T(v), \omega|_{T(v)})$ and hence are given independently.

The proof of (c) is nearly the same as the proof of Proposition 2.1, though we will give it now in detail. For $\sigma_1, \dots, \sigma_k \in \{0d, 0s, 1d, 1s\}$, let

$$\chi_{\text{col}*}(\sigma_1, \dots, \sigma_k) = \chi(k) \vec{v}_*(\sigma_1) \cdots \vec{v}_*(\sigma_k).$$

By the first claim, this is the probability that R_T has exactly k children whose types according to ω_* are $\sigma_1, \dots, \sigma_k$, in order. For any type $\sigma \in \{0d, 0s, 1d, 1s\}$ with $\vec{v}_*(\sigma) > 0$, let $\chi_{\text{col}*}^\sigma(\sigma_1, \dots, \sigma_k)$ denote the conditional probability that R_T gives birth to k children of types $\sigma_1, \dots, \sigma_k$ according to ω_* given that $\omega_*(R_T) = \sigma$. Observe that the value of ω_* at the root of a tree can be determined from the value of ω_* at its children: the 0 or 1 can be determined according to the automaton, and the s- or d-type can be determined according to whether there is a pivotal child of the root of s-type. Hence, if σ is the type at the root corresponding to children of types $\sigma_1, \dots, \sigma_k$, then

$$\chi_{\text{col}*}(\sigma_1, \dots, \sigma_k) = \vec{v}_*(\sigma) \chi_{\text{col}*}^\sigma(\sigma_1, \dots, \sigma_k). \quad (20)$$

Now, we seek to prove that given the first n levels of (T, ω_*) , each vertex v at level n independently gives birth according to $\chi_{\text{col}*}^{\omega_*(v)}$. Fix $(t, \tau_*) \in \mathcal{T}_{\text{col}*}$. By the first claim of this proposition,

$$\mathbf{P}\left[(T, \omega_*) \in [t, \tau_*]_{n+1}\right] = \mathbf{P}\left[T \in [t]_{n+1}\right] \prod_{u \in L_{n+1}(t)} \vec{v}_*(\tau_*(u)).$$

With $C(v)$ denoting the set of a children of a vertex v ,

$$\begin{aligned} \mathbf{P}\left[(T, \omega) \in [t, \tau_*]_{n+1}\right] &= \left(\mathbf{P}\left[T \in [t]_n\right] \prod_{v \in L_n(t)} \chi(|C(v)|)\right) \prod_{u \in L_{n+1}(t)} \vec{v}_*(\tau_*(u)) \\ &= \mathbf{P}\left[T \in [t]_n\right] \prod_{v \in L_n(t)} \chi_{\text{col}*}(\tau_*(u)_{u \in C(v)}). \end{aligned}$$

Continuing to follow the proof of Proposition 2.1, by (20),

$$\begin{aligned} \mathbf{P}\left[(T, \omega) \in [t, \tau_*]_{n+1}\right] &= \mathbf{P}\left[T \in [t]_n\right] \prod_{v \in L_n(t)} \vec{v}_*(\tau_*(v)) \chi_{\text{col}*}^{\tau_*(v)}(\tau_*(u)_{u \in C(v)}) \\ &= \mathbf{P}\left[(T, \omega) \in [t, \tau_*]_n\right] \prod_{v \in L_n(t)} \chi_{\text{col}*}^{\tau_*(v)}(\tau_*(u)_{u \in C(v)}). \end{aligned}$$

We will assume throughout the section that T_{piv} is supercritical. This implies that either $\mathbf{P}[\omega_*(R_T) = 0s] > 0$ or $\mathbf{P}[\omega_*(R_T) = 1s] > 0$, but in fact both are true by Proposition 2.6(c). Thus, it makes sense to consider the distribution of (T, ω_*) conditional on $\omega_*(R_T) = 0s$ or $\omega_*(R_T) = 1s$. With this in mind, we make a number of definitions. Most important among them are the probability measures RST^{0s} , RST^{1s} , $\text{RST}^{0s \rightarrow 1s}$, and $\text{RST}^{1s \rightarrow 0s}$ on the space $\mathcal{T}_{\text{col}*}$, with RST standing for *random state tree*. The measures RST^{0s} and RST^{1s} are the distributions of (T, ω_*) conditioned on $\omega_*(R_T) = 0s$ and $\omega_*(R_T) = 1s$, respectively. The measure $\text{RST}^{0s \rightarrow 1s}$ is the distribution of a labeled tree obtained by sampling from RST^{0s} , choosing an infinite path of pivotal vertices, and swapping every label in the path. The measure $\text{RST}^{1s \rightarrow 0s}$ is obtained in the same way, starting with RST^{1s} instead of RST^{0s} . Thus, RST^{1s} and $\text{RST}^{0s \rightarrow 1s}$ are both distributions on labeled trees with 1s at the root. Our goal, as we sketched before and will explain in more detail shortly, is to prove that $\text{RST}^{0s \rightarrow 1s}$ is absolutely continuous with respect to RST^{1s} .

Definitions 4.2 (Definitions of RST^ℓ , (T^ℓ, ω_*^ℓ) , $\mathcal{T}_{s*}$, $\mathcal{W}(t, \tau_*)$, $\mathcal{W}_n(t, \tau_*)$, $P_{\mathcal{W}(t, \tau_*)}$, $P_{\mathcal{W}_n(t, \tau_*)}$, $\tau_*^{\vec{v}}$, τ_*^x , $\text{RST}^{0s \rightarrow 1s}$, $\text{RST}^{1s \rightarrow 0s}$, $(T^{0s}, \omega_*^{0s \rightarrow 1s})$, and $(T^{1s}, \omega_*^{1s \rightarrow 0s})$). For $\ell \in \{0d, 1d, 0s, 1s\}$, let RST^ℓ be the law of (T, ω_*) conditioned on $\omega(R_T) = \ell$. Let (T^ℓ, ω_*^ℓ) be a random variable distributed as RST^ℓ . Let $\mathcal{T}_{s*} \subseteq \mathcal{T}_{\text{col}*}$ be the set of all trees (t, τ_*) labeled by $\{0d, 1d, 0s, 1s\}$ that are compatible with the automaton, have their d and s labels consistent with the tree and other labels, and have 0s or 1s at the root. This space is the union of the supports of RST^{0s} and RST^{1s} . It could also be defined as the set of all trees $(t, \tau_*) \in \mathcal{T}_{\text{col}*}$ such that $[t, \tau_*]_n$ has positive probability under RST^{0s} or RST^{1s} for all n .

Given a deterministic tree $(t, \tau_*) \in \mathcal{T}_{s*}$, let $\mathcal{W}(t, \tau_*)$ be the set of infinite paths in (t, τ_*) that start at R_t and contain only pivotal s-labeled vertices. Let $\mathcal{W}_n(t, \tau_*)$ be the set of paths from R_t of length n with the same property. Note that these sets are nonempty for any $(t, \tau_*) \in \mathcal{T}_{s*}$, since any pivotal s-labeled vertex must have a pivotal s-labeled child.

We define $P_{\mathcal{W}(t, \tau_*)}$ and $P_{\mathcal{W}_n(t, \tau_*)}$ to be distributions on $\mathcal{W}(t, \tau_*)$ and $\mathcal{W}_n(t, \tau_*)$, respectively, given as follows. Let $V_0 = R_t$. Choose V_1 uniformly from the pivotal s-labeled children of V_0 (as we said, there must be at least one). Then choose V_2 uniformly from the pivotal s-labeled children of V_1 , and so on. Let $P_{\mathcal{W}(t, \tau_*)}$ be the distribution of $(V_0, V_1, \dots)$, and let $P_{\mathcal{W}_n(t, \tau_*)}$ be the distribution of $(V_0, \dots, V_n)$.

For an assignment $\tau : R(t) \rightarrow \Sigma$ and a vertex $v \in L_n(t)$, we defined $\tau^{v \rightarrow \gamma}$ as the coloring of $t|_n$ given by switching the color of v to γ and updating the colors at levels $0, \dots, n-1$ according to the automaton. We now extend this definition to allow switching when the colors include s- and d-labels and to allow switching an infinite path. Suppose that $(t, \tau_*) \in \mathcal{T}_{\text{col}*}$. Given a path $\vec{u} = (u_0, u_1, \dots) \in \mathcal{W}(t, \tau_*)$, let $\tau_*^{\vec{u}}$ denote τ_* with labels 0s and 1s swapped along the path. It is easy to check that this new labeling is

also compatible with the automaton A , and that its s and d markings follow the same rules as before. For $\vec{u} \in \mathcal{W}_n(t, \tau_*)$, we define $\tau_*^{\vec{u}}$ in the same way, except that $\tau_*^{\vec{u}}$ is only a labeling of $t|_n$. For a vertex $x \in R(t)$, we use τ_*^x as a shorthand for $\tau_*^{\vec{u}}$, where $\vec{u}$ is the path from R_t to x .

Finally, we define $\omega_*^{0s \rightarrow 1s}$ as the switched labeling $(\omega_*^{0s})^{\vec{V}}$, where $\vec{V}$ is sampled from $P_{\mathcal{W}(T^{0s}, \omega_*^{0s})}$. To summarize, $(T^{0s}, \omega_*^{0s \rightarrow 1s})$ is formed by the following procedure: First condition (T, ω_*) on $\omega_* = 0s$ to obtain (T^{0s}, ω_*^{0s}) . Then, choose an infinite s -labeled path of pivotal vertices in (T^{0s}, ω_*^{0s}) by starting at the root and successively choosing a pivotal s -labeled child at random. Finally, swap all $0s$ and $1s$ labels along this path to obtain $(T^{0s}, \omega_*^{0s \rightarrow 1s})$. We define $\omega_*^{1s \rightarrow 0s}$ analogously, and we define $\text{RST}^{0s \rightarrow 1s}$ and $\text{RST}^{1s \rightarrow 0s}$ as the distributions of $(T^{0s}, \omega_*^{0s \rightarrow 1s})$ and $(T^{1s}, \omega_*^{1s \rightarrow 0s})$, respectively.

We now lay out our plan for the section. Our goal is to prove that $\text{RST}^{0s \rightarrow 1s} \ll \text{RST}^{1s}$. It follows quickly from this that $\vec{v}$ is rogue by an argument we sketch now. Supposing that $\vec{v}$ is interpretable, we can express $\omega(R_T)$ as $\iota(T)$ for a deterministic function $\iota: \mathcal{T} \rightarrow \Sigma$, by Proposition 2.2. By definition of RST^{0s} and RST^{1s} , we have $\iota(T^{0s}) = 0$ a.s. and $\iota(T^{1s}) = 1$ a.s. Recalling that $(T^{0s}, \omega_*^{0s \rightarrow 1s}) \sim \text{RST}^{0s \rightarrow 1s}$, absolute continuity lets us conclude from $\iota(T^{1s}) = 1$ a.s. that $\iota(T^{0s}) = 1$ a.s., a contradiction.

To prove the absolute continuity of $\text{RST}^{0s \rightarrow 1s}$ with respect to RST^{1s} , we use a technique of restricting these measures to successively larger σ -algebras and computing the Radon-Nikodym derivatives of the restricted measures. The result we use is well known:

Lemma 4.3 ([19, Lemma 12.2]). *Let μ and ν be probability measures on a σ -algebra $\mathcal{F}$. Suppose that $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots \subseteq \mathcal{F}$, and that $\cup_n \mathcal{F}_n$ generates $\mathcal{F}$. Also suppose that $\mu|_{\mathcal{F}_n}$ is absolutely continuous with respect to $\nu|_{\mathcal{F}_n}$ with Radon-Nikodym derivative X_n . Define $X = \limsup_{n \rightarrow \infty} X_n$. Then*

$$\mu \ll \nu \iff X < \infty \text{ } \mu\text{-a.e.} \iff \int X d\nu = 1,$$

and

$$\mu \perp \nu \iff X = \infty \text{ } \mu\text{-a.e.} \iff \int X d\nu = 0.$$

In our case, we will restrict $\text{RST}^{0s \rightarrow 1s}$ and RST^{1s} to the σ -algebra $\mathcal{F}_n$ generated by the first n levels of the labeled tree. That is, we define $\mathcal{F}_n$ as the σ -algebra on $\mathcal{T}_{s*}$ generated by the sets of the form $[t, \tau_*]_n$.

To investigate these Radon-Nikodym derivatives, we start by giving representations of $\text{RST}^{0s \rightarrow 1s}$ and $\text{RST}^{1s \rightarrow 0s}$ in terms of RST^{0s} and RST^{1s} :

Lemma 4.4. *For any $(t, \tau_*) \in \mathcal{T}_{s*}$,*

$$\text{RST}^{0s \rightarrow 1s}[t, \tau_*]_n = \sum_{\vec{u} \in \mathcal{W}_n(t, \tau_*)} \text{RST}^{0s}[t, \tau_*^{\vec{u}}]_n P_{\mathcal{W}_n(t, \tau_*^{\vec{u}})}(\vec{u}), \quad (21)$$

and

$$\text{RST}^{1s \rightarrow 0s}[t, \tau_*]_n = \sum_{\vec{u} \in \mathcal{W}_n(t, \tau_*)} \text{RST}^{1s}[t, \tau_*^{\vec{u}}]_n P_{\mathcal{W}_n(t, \tau_*^{\vec{u}})}(\vec{u}). \quad (22)$$

Proof. These statements follow very directly from the definitions. Recall that $(T^{0s}, \omega_*^{0s \rightarrow 1s})$ differs from (T^{0s}, ω_*^{0s}) along a random path $(V_i)_{i \geq 0}$ sampled from $P_{\mathcal{W}(T^{0s}, \omega_*^{0s})}$, in which all 0s and 1s labels have been swapped. Hence,

$$(T^{0s}, \omega_*^{0s \rightarrow 1s}) \in [t, \tau_*]_n$$

holds if and only if

$$(T^{0s}, \omega_*^{0s}) \in [t, \tau_*^{\vec{u}}] \text{ for some } \vec{u} \in \mathcal{W}_n(t, \tau_*^{\vec{u}}), \text{ and } (V_0, \dots, V_n) = \vec{u}. \quad (23)$$

Since $\vec{u} \in \mathcal{W}_n(t, \tau_*^{\vec{u}})$ if and only if $\vec{u} \in \mathcal{W}_n(t, \tau_*)$, we can refine (23) to

$$(T^{0s}, \omega_*^{0s}) \in [t, \tau_*^{\vec{u}}] \text{ for some } \vec{u} \in \mathcal{W}_n(t, \tau_*), \text{ and } (V_0, \dots, V_n) = \vec{u}. \quad (24)$$

As the events in (24) are disjoint for different choices of $\vec{u}$,

$$\begin{aligned} \mathbf{P}[(T^{0s}, \omega_*^{0s \rightarrow 1s}) \in [t, \tau_*]_n] \\ &= \sum_{\vec{u} \in \mathcal{W}_n(t, \tau_*)} \mathbf{P}[(T^{0s}, \omega_*^{0s}) \in [t, \tau_*^{\vec{u}}]_n \text{ and } (V_0, \dots, V_n) = \vec{u}] \\ &= \sum_{\vec{u} \in \mathcal{W}_n(t, \tau_*)} \mathbf{P}[(T^{0s}, \omega_*^{0s}) \in [t, \tau_*^{\vec{u}}]_n] P_{\mathcal{W}_n(t, \tau_*)}(\vec{u}), \end{aligned}$$

which is a restatement of (21). The proof of (22) is identical. ■

Define a map r_n on $\mathcal{T}_{s*}$ as follows. For $(t, \tau_*) \in \mathcal{T}_{s*}$ with 0s at the root, let

$$r_n(t, \tau_*) = \frac{\text{RST}^{1s \rightarrow 0s}[t, \tau_*]_n}{\text{RST}^{0s}[t, \tau_*]_n}, \quad (25)$$

and for $(t, \tau_*) \in \mathcal{T}_{s*}$ with 1s at the root, let

$$r_n(t, \tau_*) = \frac{\text{RST}^{0s \rightarrow 1s}[t, \tau_*]_n}{\text{RST}^{1s}[t, \tau_*]_n}. \quad (26)$$

Thus, $r_n(t, \tau_*)$ matches the Radon-Nikodym derivative either of $\text{RST}^{1s \rightarrow 0s}|_{\mathcal{F}_n}$ with respect to $\text{RST}^{0s}|_{\mathcal{F}_n}$ or of $\text{RST}^{0s \rightarrow 1s}|_{\mathcal{F}_n}$ with respect to $\text{RST}^{1s}|_{\mathcal{F}_n}$, depending on $\tau_*(R_t)$. To prove that $\text{RST}^{0s \rightarrow 1s} \ll \text{RST}^{1s}$, it therefore suffices by Lemma 4.3 to show that

$$\limsup_{n \rightarrow \infty} r_n(T^{0s}, \omega_*^{0s \rightarrow 1s}) < \infty \text{ a.s.} \quad (27)$$

We define

$$f_n(t, \tau_*) = \sum_{\vec{u} \in \mathcal{W}_n(t, \tau_*)} P_{\mathcal{W}_n(t, \tau_*)}(\vec{u}),$$

with $f_0(t, \tau_*)$ taken to be 1. According to the next lemma, we can use this simpler function f_n as a stand-in for r_n .

Lemma 4.5. For some constant $1 \leq C < \infty$ depending on $\vec{v}_*$, it holds for all $n \geq 0$ and all $(t, \tau_*) \in \mathcal{T}_{s*}$ that

$$\frac{1}{C}f_n(t, \tau_*) \leq r_n(t, \tau_*) \leq Cf_n(t, \tau_*).$$

Proof. We aim to show that for some $1 \leq C < \infty$, it holds for all $n \geq 0$, $\vec{u} \in \mathcal{W}_n(t, \tau_*)$, and $(t, \tau_*) \in \mathcal{T}_{s*}$ with $\tau_*(R_t) = 1s$ that

$$\frac{1}{C}\text{RST}^{1s}[t, \tau_*]_n \leq \text{RST}^{0s}[t, \tau_*^{\vec{u}}]_n \leq C\text{RST}^{1s}[t, \tau_*]_n,$$

and it holds for all $n \geq 0$, $\vec{u} \in \mathcal{W}_n(t, \tau_*)$, and $(t, \tau_*) \in \mathcal{T}_{s*}$ with $\tau_*(R_t) = 0s$ that

$$\frac{1}{C}\text{RST}^{0s}[t, \tau_*]_n \leq \text{RST}^{1s}[t, \tau_*^{\vec{u}}]_n \leq C\text{RST}^{0s}[t, \tau_*]_n.$$

Once we prove this, the result follows immediately from Lemma 4.4 and the definition of r_n .

To prove these statements, we go back to the unconditioned tree (T, ω_*) . Let

$$C_1 = \frac{\max_{\sigma \in \{0s, 1s\}} \vec{v}_*(\sigma)}{\min_{\sigma \in \{0s, 1s\}} \vec{v}_*(\sigma)}.$$

By Proposition 4.1(a),

$$\begin{aligned} \mathbf{P}\left[(T, \omega_*) \in [t, \tau_*^{\vec{u}}]_n\right] &= \mathbf{P}[T|_n = t|_n] \prod_{x \in L_n(t)} \vec{v}_*(\tau_*^{\vec{u}}(x)) \\ &\leq C_1 \mathbf{P}[T|_n = t|_n] \prod_{x \in L_n(t)} \vec{v}_*(\tau_*(x)) = C_1 \mathbf{P}\left[(T, \omega_*) \in [t, \tau_*]_n\right] \end{aligned}$$

since $\tau_*^{\vec{u}}$ and τ_* match each other on $L_n(t)$ except at a single vertex, where one assigns 0s and the other assigns 1s.

Suppose that $\tau_*(R_t) = 1s$. Then

$$\begin{aligned} \text{RST}^{0s}[t, \tau_*^{\vec{u}}]_n &= \frac{\mathbf{P}\left[(T, \omega_*) \in [t, \tau_*^{\vec{u}}]_n\right]}{\vec{v}_*(0s)} \\ &\leq \frac{C_1 \mathbf{P}\left[(T, \omega_*) \in [t, \tau_*]_n\right]}{\vec{v}_*(0s)} \\ &\leq \frac{C_1^2 \mathbf{P}\left[(T, \omega_*) \in [t, \tau_*]_n\right]}{\vec{v}_*(1s)} = C_1^2 \text{RST}^{1s}[t, \tau_*]_n. \end{aligned}$$

The lower bound on $\text{RST}^{0s}[t, \tau_*^{\vec{u}}]_n$ and the bounds on $\text{RST}^{1s}[t, \tau_*^{\vec{u}}]_n$ follow by nearly identical proofs. ■

Next, we recast $f_n(t, \tau_*)$ as a weighted sum over paths.

Definitions 4.6 ($N(x, t, \tau_*)$ and $w_{t, \tau_*}(x, y)$). For a vertex $x \in V(t)$, let $N(x, t, \tau_*)$ be the number of pivotal s -labeled children of x in (t, τ_*) . Suppose that x and y are respectively a vertex and its child in

some path in $\mathcal{W}(t, \tau_*)$. Define

$$w_{t, \tau_*}(x, y) = N(x, t, \tau_*^y)^{-1},$$

which we will view as a weight on the edge from x to y . We will shorten this to $w(x, y)$ when the tree (t, τ_*) is clear from context.

In words, $w(x, y)$ is the reciprocal of the number of pivotal s -labeled children of x after swapping all labels on the path from the root to x to y . Note that this count of pivotal s -labeled children is never zero for such an x and y , since y is always s -labeled and pivotal for (t, τ_*) as a consequence of belonging to a path in $\mathcal{W}(t, \tau_*)$.

Lemma 4.7. *For any $(t, \tau_*) \in \mathcal{T}_{s*}$,*

$$f_n(t, \tau_*) = \sum_{\vec{u} \in \mathcal{W}_n(t, \tau_*)} \prod_{i=0}^{n-1} w(u_i, u_{i+1}), \quad (28)$$

which we can express recursively as

$$f_n(t, \tau_*) = \sum_x w(R_t, x) f_{n-1}(t(x), \tau_*|_{t(x)}), \quad (29)$$

where x ranges over the pivotal s -labeled children of R_t .

Proof. To prove (28), we need to show that for any $\vec{u} = (u_0, \dots, u_n) \in \mathcal{W}_n(t, \tau_*)$,

$$P_{\mathcal{W}_n(t, \tau_*)}(\vec{u}) = \prod_{i=0}^{n-1} N(u_i, t, \tau_*^{u_{i+1}})^{-1}.$$

This is evident, as $P_{\mathcal{W}_n(t, \tau_*)}(\vec{u})$ is the probability that $\vec{u}$ is selected by the procedure of starting at the root in (t, τ_*) and uniformly picking a pivotal s -labeled child, then another pivotal s -labeled child, and so on. Equation (29) follows from (28) by partitioning $\mathcal{W}_n(t, \tau_*)$ into paths going through each of the s -labeled children of the root. ■

Recall that $\omega_*^{0s \rightarrow 1s}$ is formed by swapping labels in ω_*^{0s} along a random path $\vec{V} = (V_i)_{i \geq 0}$, where $V_0 = R_{T^{0s}}$. Call this path the *spine* of $(T^{0s}, \omega_*^{0s \rightarrow 1s})$. We now give some terminology for describing the weights (in the sense of Definitions 4.6) of edges along and hanging off the spine.

Definitions 4.8 (Definitions of $V_{i,j}$, W_i , $W_{i,j}$, T_n^{spine} , T_n^{spine} , $\mathcal{G}$, and $\mathcal{G}_n$). Let $V_{i,1}, \dots, V_{i,k_i}$ be the pivotal s -labeled children of V_i in $(T^{0s}, \omega_*^{0s \rightarrow 1s})$ other than V_{i+1} . Let $W_i = w_{T^{0s}, \omega_*^{0s \rightarrow 1s}}(V_i, V_{i+1})$, and let $W_{i,j} = w_{T^{0s}, \omega_*^{0s \rightarrow 1s}}(V_i, V_{i,j})$.

Let $T_n^{\text{spine}} \subseteq T^{0s}$ be the subtree consisting of $\vec{V}$ and all vertices $V_{i,j}$. Let T_n^{spine} be the restriction of T^{spine} to height n . Let $\mathcal{G}$ be the σ -algebra generated by $\vec{V}$, T^{spine} and $\omega_*^{0s \rightarrow 1s}|_{T^{\text{spine}}}$. Let $\mathcal{G}_n$ be the σ -algebra generated by $(V_0, \dots, V_n)$, T_n^{spine} and by $\omega_*^{0s \rightarrow 1s}|_{T_n^{\text{spine}}}$. See Figure 4 for a depiction of the information captured by these σ -algebras.

The key idea in analyzing $f_n(T^{0s}, \omega_*^{0s \rightarrow 1s})$ is that while $(T^{0s}, \omega_*^{0s \rightarrow 1s})$ behaves unusually along the spine, starting from any vertex $V_{i,j}$ it is a multitype Galton-Watson tree with the same child distributions as (T, ω_*) . This is formally expressed in Proposition 4.9. Thus, understanding the weights along the

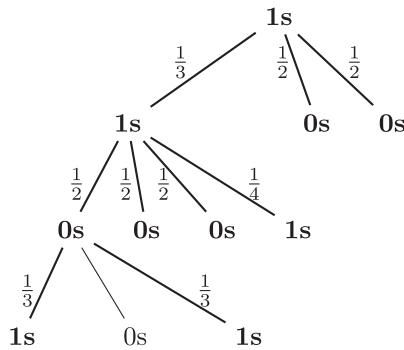


FIGURE 4 A subset of $(T^{0s}, \omega^{0s \rightarrow 1s})$. The tree automaton in this example assigns 1 to a vertex if and only if it has exactly one 1-labeled child. The spine, $(V_i)_{i \geq 0}$, is on the left side, and vertices $V_{i,1}, \dots, V_{i,k_i}$ hang off each V_i . Pivotal vertices and the edges between them are in bold. Alongside each edge is its weight. For example, the weight from V_0 to V_1 is $\frac{1}{3}$, because if V_0 and V_1 have their labels swapped from 1s to 0s, then V_0 has three pivotal s-labeled children

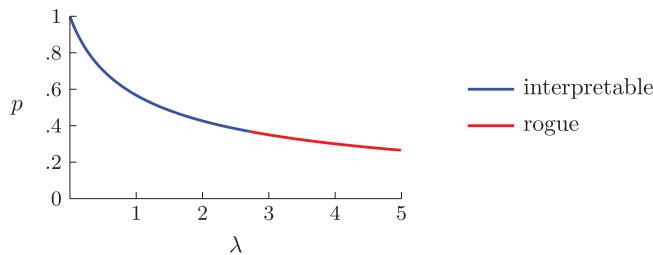


FIGURE 5 Plot for Example 5.7, showing the fixed point for the automaton assigning the parent state “1” if and only if there are zero “1” children with a $\text{Poi}(\lambda)$ child distribution [Color figure can be viewed at wileyonlinelibrary.com]

spine of $(T^{0s}, \omega_*^{0s \rightarrow 1s})$ as well as the weights on (T, ω_*) is enough to understand the weights on all of $(T^{0s}, \omega_*^{0s \rightarrow 1s})$. This is the same idea used by Pemantle, Peres, and Lyons to prove the Kesten-Stigum theorem (see [18, Section 3] or [19, Chapter 12]).

Proposition 4.9.

- (a) The random variables $\{W_i, W_{i,j}\}_{i < n, 1 \leq j \leq k_i}$ are measurable with respect to $\mathcal{G}_n$.
 (b) Conditional on $\mathcal{G}$, the subtrees

$$\left\{ (T^{0s}(x), \omega_*^{0s \rightarrow 1s}|_{T^{0s}(x)}) : x = V_{i,j} \text{ for some } i \geq 0, 0 \leq j \leq k_i \right\}$$

are independent.

- (c) For any $x = V_{i,j}$, the subtree $(T^{0s}(x), \omega_*^{0s \rightarrow 1s}|_{T^{0s}(x)})$ is distributed conditional on $\mathcal{G}$ as RST^{0s} if $\omega_*^{0s \rightarrow 1s}(x) = 0s$ and as RST^{1s} if $\omega_*^{0s \rightarrow 1s}(x) = 1s$.
 (d) The subtree $(T^{0s}(V_n), \omega_*^{0s \rightarrow 1s}|_{T^{0s}(V_n)})$ is distributed conditional on $\mathcal{G}_n$ as $\text{RST}^{0s \rightarrow 1s}$ if $\omega_*^{0s \rightarrow 1s}(V_n) = 1s$ and as $\text{RST}^{1s \rightarrow 0s}$ if $\omega_*^{0s \rightarrow 1s}(V_n) = 0s$.

Proof. Part (a) follows directly from the definition. For parts (b) and (c), note that for any $x = V_{i,j}$, the subtree $(T^{0s}(x), \omega_*^{0s \rightarrow 1s}|_{T^{0s}(x)})$ is identical to $(T^{0s}(x), \omega_*^{0s}|_{T^{0s}(x)})$. We can also recharacterize $\mathcal{G}$ as the σ -algebra generated by $\vec{V}$, T^{spine} , and $\omega_*^{0s}|_{T^{\text{spine}}}$. Conditioning on $\mathcal{G}$ is then just revealing part of

(T^{0s}, ω_*^{0s}) , which is Galton-Watson by Proposition 4.1(c). Under this conditioning, the subtrees

$$\{ (T^{0s}(x), \omega_*^{0s} |_{T^{0s}(x)}) : x = V_{i,j} \text{ for some } i \geq 0, 0 \leq j \leq k_i \}$$

are unrevealed except for the labels of their roots, and hence they evolve independently according to RST^{0s} or RST^{1s} .

To prove part (d), we observe that conditioning on $\mathcal{G}_n$ reveals $(V_0, \dots, V_n)$, and it reveals a portion of (T^{0s}, ω_*^{0s}) with V_n as a leaf. Thus $(T^{0s}(V_n), \omega_*^{0s} |_{T^{0s}(V_n)})$ evolves either as RST^{0s} or as RST^{1s} conditional on $\mathcal{G}_n$, depending on $\omega_*^{0s}(V_n)$. Also, by its definition, $(V_i)_{i \geq n}$ conditional on $\mathcal{G}_n$ is distributed as $P_{\mathcal{W}(T^{0s}(V_n), \omega_*^{0s} |_{T^{0s}(V_n)})}$. Thus, $(T^{0s}(V_n), \omega_*^{0s \rightarrow 1s} |_{T^{0s}(V_n)})$ is distributed conditionally on $\mathcal{G}_n$ as stated. ■

Now, we can start evaluating $\limsup_n f_n(T^{0s}, \omega_*^{0s \rightarrow 1s})$. First, we expand $f_n(T^{0s}, \omega_*^{0s \rightarrow 1s})$ in terms of the weights along and off the spine.

Lemma 4.10. *We can express $f_n(T^{0s}, \omega_*^{0s \rightarrow 1s})$ as*

$$\begin{aligned} f_n(T^{0s}, \omega_*^{0s \rightarrow 1s}) &= \sum_{i=0}^{n-1} (W_0 \cdots W_{i-1}) \sum_{j=1}^{k_i} W_{i,j} f_{n-i-1}(T^{0s}(V_{i,j}), \omega_*^{0s \rightarrow 1s} |_{T^{0s}(V_{i,j})}) \\ &\quad + W_0 \cdots W_{n-1} \\ &\leq \sum_{i=0}^{n-1} (W_0 \cdots W_{i-1}) \sum_{j=1}^{k_i} W_{i,j} C r_{n-i-1}(T^{0s}(V_{i,j}), \omega_*^{0s \rightarrow 1s} |_{T^{0s}(V_{i,j})}) \\ &\quad + W_0 \cdots W_{n-1}, \end{aligned} \quad (30)$$

where C is the constant from Lemma 4.5.

Proof. The equality holds by successively applying (29) from Lemma 4.7, and the inequality is an application of Lemma 4.5. ■

It is odd that we have bounded f_n by r_n when f_n is a simpler quantity that we typically prefer to work with. But in the proof of Lemma 4.15, it will be easier to work with the Radon-Nikodym derivative itself rather than an approximation.

Now, in Lemmas 4.11-4.14, we prove some technical facts that help us bound (30).

Lemma 4.11. *For any i and $1 \leq j \leq k_i$, the process*

$$\left(r_{n-i-1}(T^{0s}(V_{i,j}), \omega_*^{0s \rightarrow 1s} |_{T^{0s}(V_{i,j})}) \right)_{n \geq i+1}$$

conditional on $\mathcal{G}$ is a nonnegative martingale in n with mean one.

Proof. By definition of r_n , given in (25) and (26), the process is nonnegative. By Proposition 4.9(c), the conditional distribution of

$$(T^{0s}(V_{i,j}), \omega_*^{0s \rightarrow 1s} |_{T^{0s}(V_{i,j})})$$

given $\mathcal{G}$ is either RST^{0s} or RST^{1s} , depending on the value of $\omega_*^{0s \rightarrow 1s}(V_{i,j})$. For the sake of concreteness, suppose that $\omega_*^{0s \rightarrow 1s}(V_{i,j}) = 0s$ so that its conditional distribution is RST^{0s} . Then

$r_{n-i-1}(T^{0s}(V_{ij}), \omega_*^{0s \rightarrow 1s} |_{T^{0s}(V_{ij})})$ conditional on $\mathcal{G}$ is the Radon-Nikodym derivative of $\text{RST}^{1s \rightarrow 0s} |_{\mathcal{G}_n}$ with respect to $\text{RST}^{0s} |_{\mathcal{G}_n}$, applied to an RST^{0s} -distributed random variable. Hence, conditional on $\mathcal{G}$, it is a martingale in n [9, Lemma 5.3.4]. The same logic shows that $r_{n-i-1}(T^{0s}(V_{ij}), \omega_*^{0s \rightarrow 1s} |_{T^{0s}(V_{ij})})$ is a martingale conditional on $\mathcal{G}$ when $\omega_*^{0s \rightarrow 1s}(V_{ij}) = 1s$. The initial value of either martingale, when $n = i + 1$, is 1. ■

Lemma 4.12. *For some $c < 1$, it holds for all $n \geq 0$ that*

$$\mathbf{E}[W_n W_{n+1} \mid \mathcal{G}_n] \leq c \text{ a.s.}$$

Proof. First, we claim that W_n is the reciprocal of the number of pivotal s-labeled children of V_n in the original unswitched tree (T^{0s}, ω_*^{0s}) ; that is,

$$W_n = N(V_n, T^{0s}, \omega_*^{0s})^{-1}. \quad (31)$$

Indeed, by the definition of W_n in Definitions 4.8 and then the definition of $w_{t, \tau_*}(x, y)$ in Definitions 4.6,

$$W_n = w_{T^{0s}, \omega_*^{0s \rightarrow 1s}}(V_i, V_{i+1}) = N(V_n, T^{0s}, (\omega_*^{0s \rightarrow 1s})^{V_{n+1}})^{-1}.$$

The labels $(\omega_*^{0s \rightarrow 1s})^{V_{n+1}}$ consist of the original labels ω_*^{0s} switched along the spine and then switched back again, yielding (31).

We now seek to analyze this expression conditional on $\mathcal{G}_n$. By Proposition 4.9(d), the distribution of $(T^{0s}(V_n), \omega_*^{0s \rightarrow 1s} |_{T(V_n)})$ conditional on $\mathcal{G}_n$ is either $\text{RST}^{1s \rightarrow 0s}$ or $\text{RST}^{0s \rightarrow 1s}$, depending on $\omega_*^{0s}(V_n)$. Equivalently, the distribution of $(T^{0s}(V_n), \omega_*^{0s} |_{T(V_n)})$ conditional on $\mathcal{G}_n$ is RST^ℓ where $\ell = \omega_*^{0s}(V_n)$, which is measurable with respect to $\mathcal{G}_n$. Hence, $N(V_n, T^{0s}, \omega_*^{0s})$ conditional on $\mathcal{G}_n$ is distributed as $N(R_{T^\ell}, T^\ell, \omega_*^\ell)$, where $\ell = \omega_*^{0s}(V_n)$.

Let $c_\ell = \mathbf{E}N(R_{T^\ell}, T^\ell, \omega_*^\ell)^{-1}$ for $\ell = 0s, 1s$. Recall that $N(R_{T^\ell}, T^\ell, \omega_*^\ell) \geq 1$, since a pivotal s-labeled vertex must give birth to another pivotal s-labeled vertex. Hence, we have $c_\ell = 1$ if and only if $N(R_{T^\ell}, T^\ell, \omega_*^\ell) = 1$ a.s. If $c_{0s} < 1$ and $c_{1s} < 1$, then set $c = \max(c_{0s}, c_{1s})$ and use the easy bound $W_{n+1} \leq 1$ to get

$$\mathbf{E}[W_n W_{n+1} \mid \mathcal{G}_n] \leq c \text{ a.s.}$$

The troublesome case is when $c_{0s} = 1$ or $c_{1s} = 1$. Suppose $c_{0s} = 1$. If $c_{1s} = 1$, the proof is identical with the roles of 0s and 1s switched. The argument has two steps:

- (a) $c_{1s} < 1$;
- (b) a vertex of type 0s gives birth to a pivotal 1s-labeled vertex with positive probability.

Suppose that (a) is false. Then $c_{0s} = c_{1s} = 1$, and consequently no vertex type in (T, ω_*) ever gives birth to more than one s-labeled pivotal child. This implies that no vertex gives birth to more than one pivotal child. Indeed, according to Proposition 4.1(b), given the colors of the children of a vertex, their s- and d-labels are assigned independently. Thus, if a vertex of color 0 or 1 had positive probability of having multiple pivotal children, it would also have positive probability of having multiple s-labeled pivotal children. Since vertices of either color give birth to at most one pivotal child, the highest eigenvalue of the matrix of means of (T_{piv}, ω) is at most one. But this contradicts our assumption throughout this section that the pivot tree is supercritical, establishing (a).

For (b), from $c_{0s} = 1$ we know that a vertex of type 0s always gives birth to exactly one pivotal s-labeled child. As above, this implies that it always gives birth to exactly one pivotal child. In fact, a vertex of type 0 must always give birth to zero or one pivotal children, since if it had positive probability of giving birth to two or more, then it would have positive probability of giving birth to two or more pivotal s-labeled vertices, and so a vertex of type 0s would have positive probability of giving birth to two or more pivotal children. Hence $m_{00} + m_{01} \leq 1$, in the language of Lemma 2.5.

Suppose that (b) is false and a vertex of type 0s never gives birth to a pivotal 1s-labeled vertex. Then, a vertex of type 0 never gives birth to a pivotal vertex of type 1 (if it had positive probability of doing so, then a vertex of type 0s would have positive probability of giving birth to a pivotal vertex of type 1s). Hence, $m_{01} = 0$. By Lemma 2.5, the matrix of means for T_{piv} has the form $\begin{bmatrix} m_{00} & 0 \\ 0 & m_{00} \end{bmatrix}$ where $m_{00} \leq 1$. But this contradicts the supercriticality of T_{piv} , proving (b).

Now, we are ready to evaluate $\mathbf{E}[W_n W_{n+1} \mid \mathcal{G}_n]$ when $c_{0s} = 1$. Let p be the probability that a vertex of type 0s has a pivotal child of type 1s, which we know to be positive by (b). If $\omega_*^{0s}(V_n) = 0s$, then

$$\mathbf{E}[W_n W_{n+1} \mid \mathcal{G}_n] = 1 - p + p c_{1s} < 1.$$

If $\omega_*^{0s}(V_n) = 1s$, then using the bound $W_{n+1} \leq 1$, we have

$$\mathbf{E}[W_n W_{n+1} \mid \mathcal{G}_n] \leq c_{1s} < 1.$$

We then take c as the maximum of these two values to complete the proof. ■

We mention that the difficult case in this proof, where one of c_{0s} and c_{1s} is zero, can truly occur. For example, let A be the two-state automaton assigning color 1 to a parent if and only if the parent has at least three children, all of which have the same color. If a vertex has color 0, then it has a pivotal child only when it has three or more children and all but one of them are the same color, in which case it has exactly one pivotal child (the odd-colored one). Thus, conditional on being type 0s, a vertex has exactly one pivotal s-labeled child, which makes $c_{0s} = 1$.

The following lemma is well known, though it is most commonly stated with its converse under the additional assumption that $X_1, X_2, \dots$ are independent (see [19, Exercise 12.2]). We sketch the proof here.

Lemma 4.13. *Let $X_1, X_2, \dots$ be nonnegative random variables with a common distribution. If this distribution has finite log-moment, then*

$$\sum_{n=1}^{\infty} c^n X_n < \infty \text{ a.s.}$$

for all $c \in (0, 1)$.

Proof. Apply the Borel-Cantelli lemma to show that $\limsup_{n \rightarrow \infty} \frac{1}{n} \log X_n = 0$ a.s. Then it follows for some finite random N that $X_n \leq (2c)^{-n}$ for $n \geq N$. ■

Lemma 4.14. *If the child distribution χ has finite log-moment, then*

$$\sum_{i=0}^{\infty} W_0 \cdots W_{i-1} N(V_i, T^{0s}, \omega_*^{0s \rightarrow 1s}) < \infty \text{ a.s.}$$

Proof. First, we claim that

$$\mathbf{E}[W_0 \cdots W_{i-1}] \leq c^{\lfloor i/2 \rfloor}$$

for some $c < 1$. This is proven by applying Lemma 4.12 to take conditional expectations given $\mathcal{G}_{i-2}$, then given $\mathcal{G}_{i-4}$, and so on. Choosing any $b \in (\sqrt{c}, 1)$ and applying Markov's inequality,

$$\sum_{i=1}^{\infty} \mathbf{P}[W_0 \cdots W_{i-1} > b^i] \leq \sum_{i=1}^{\infty} \frac{c^{\lfloor i/2 \rfloor}}{b^i} < \infty.$$

By the Borel-Cantelli lemma, it holds almost surely that $W_0 \cdots W_{i-1} \leq b^i$ for all but finitely many values of i . Hence, it suffices to show that

$$\sum_{i=0}^{\infty} b^i N(V_i, T^{0s}, \omega_*^{0s \rightarrow 1s}) < \infty \text{ a.s.} \quad (32)$$

Now, it remains to apply Lemma 4.13. Since $N(V_i, T^{0s}, \omega_*^{0s \rightarrow 1s})$ is the number of s -pivotal offspring of V_i in $(T^{0s}, \omega_*^{0s \rightarrow 1s})$, it is bounded by the total number of offspring of V_i in T^{0s} . By Proposition 4.9(c), the distribution of $(T(V_i), \omega_*^{0s \rightarrow 1s} |_{T(V_i)})$ conditional on $\mathcal{G}$ is either RST^{0s} or RST^{1s} . Thus, conditional on $\mathcal{G}$, the random variable $N(V_i, T^{0s}, \omega_*^{0s \rightarrow 1s})$ is stochastically dominated by the number of vertices at level 1 of either T^{0s} or T^{1s} . Let A_{0s} and A_{1s} be random variables with these distributions, respectively. Since χ is assumed to have finite log-moment, both A_{0s} and A_{1s} have finite log-moment as well. Now, let X_i have any distribution that stochastically dominates A_{0s} and A_{1s} and has finite log-moment. For example, one could take $X_i = A_{0s} + A_{1s}$ where A_{0s} and A_{1s} are independent. Thus, $N(V_i, T^{0s}, \omega_*^{0s \rightarrow 1s})$ conditional on $\mathcal{G}$ is stochastically dominated by X_i , and so there exists a coupling in which $N(V_i, T^{0s}, \omega_*^{0s \rightarrow 1s}) \leq X_i$ for all i . (Note that we do not care about the joint distribution of $X_1, X_2, \dots$) Under this coupling,

$$\sum_{i=0}^{\infty} b^i N(V_i, T^{0s}, \omega_*^{0s \rightarrow 1s}) \leq \sum_{i=0}^{\infty} b^i X_i,$$

which is almost surely finite by Lemma 4.13. This proves (32), from which the lemma follows. $\blacksquare$

Finally, we are ready to achieve what we have been building towards by bounding the right-hand side of (30).

Lemma 4.15. *Assume that the child distribution χ has finite log-moment. Then*

$$\limsup_{n \rightarrow \infty} f_n(T^{0s}, \omega_*^{0s \rightarrow 1s}) < \infty \text{ a.s.}$$

Proof. Let

$$Y_n = \sum_{i=0}^{n-1} (W_0 \cdots W_{i-1}) \sum_{j=1}^{k_i} W_{ij} Cr_{n-i-1}(T^{0s}(V_{ij}), \omega_*^{0s \rightarrow 1s} |_{T^{0s}(V_{ij})}),$$

one of the terms on the right-hand side of (30). We will show that Y_n converges almost surely to a finite value. The idea is to use Lemmas 4.11 and 4.14 to show that the conditional distribution of $(Y_n)_{n \geq 0}$ given $\mathcal{G}$ is that of a submartingale bounded in L^1 .

First, consider Y_n conditionally on $\mathcal{G}$. By Proposition 4.9(a), the random variables $W_0, W_1, \dots$ are constants. By Proposition 4.9(c), the processes

$$\left(r_{n-i-1} \left(T^{0s}(V_{ij}), \omega_*^{0s \rightarrow 1s} |_{T^{0s}(V_{ij})} \right) \right)_{n \geq i+1}$$

are independent for different values of i . By Lemma 4.11, these processes are martingales. Hence, Y_n conditional on $\mathcal{G}$ is a sum of independent martingales with an additional martingale added at each step, which makes it a submartingale conditional on $\mathcal{G}$.

To prove that $\sup_n \mathbf{E}[Y_n | \mathcal{G}] < \infty$ a.s., by Lemma 4.11,

$$\mathbf{E} \left[r_{n-i-1} \left(T^{0s}(V_{ij}), \omega_*^{0s \rightarrow 1s} |_{T^{0s}(V_{ij})} \right) \mid \mathcal{G} \right] = 1 \text{ a.s.}$$

Hence,

$$\begin{aligned} \sup_n \mathbf{E}[Y_n | \mathcal{G}] &= \sup_n \sum_{i=0}^{n-1} (W_0 \cdots W_{i-1}) \sum_{j=1}^{k_i} W_{i,j} \\ &\leq \sup_n \sum_{i=0}^{n-1} W_0 \cdots W_{i-1} N(V_i, T^{0s}, \omega_*^{0s \rightarrow 1s}) \\ &\leq \sum_{i=0}^{\infty} W_0 \cdots W_{i-1} N(V_i, T^{0s}, \omega_*^{0s \rightarrow 1s}) < \infty \text{ a.s.} \end{aligned} \quad (33)$$

The first inequality uses the bound $W_{ij} \leq 1$ along with the fact that $k_i = N(V_i, T^{0s}, \omega_*^{0s \rightarrow 1s}) - 1$. The last inequality is the statement of Lemma 4.14.

We have now shown that $(Y_n)_{n \geq 0}$ conditional on $\mathcal{G}$ is a submartingale bounded in L^1 . It hence converges almost surely to a finite limit. Since $W_0 \cdots W_{n-1}$ is a decreasing positive sequence in n , it also converges as $n \rightarrow \infty$. By Lemma 4.10, we have shown that $f_n(T^{0s}, \omega_*^{0s \rightarrow 1s})$ is bounded by a process converging almost surely to a finite limit as $n \rightarrow \infty$. ■

Proof of Theorem 1.7 ($\Rightarrow$). Suppose that T_{piv} is supercritical. By Lemmas 4.5 and 4.15,

$$\limsup_{n \rightarrow \infty} r_n(T^{0s}, \omega_*^{0s \rightarrow 1s}) < \infty \text{ a.s.}$$

By Lemma 4.3, we have $\text{RST}^{0s \rightarrow 1s} \ll \text{RST}^{1s}$. Therefore any almost sure event under RST^{1s} is almost sure under $\text{RST}^{0s \rightarrow 1s}$ as well.

Suppose that $\vec{v}$ is interpretable. By Proposition 2.2, the random variable $\omega(R_T)$ is measurable with respect to T . Hence there exists a measurable map $\varphi: \mathcal{T} \rightarrow \{0, 1\}$ such that $\omega(R_T) = \varphi(T)$ a.s. Since T^{0s} and T^{1s} are distributed as T conditioned on subevents of $\omega(R_T) = 0$ and $\omega(R_T) = 1$, respectively, we have $\varphi(T^{0s}) = 0$ a.s. and $\varphi(T^{1s}) = 1$ a.s. Stating the second of these facts in terms of measure theory, the event $\{(t, \tau_*) \in \mathcal{T}_{s*} : \varphi(t) = 1\}$ has probability one under RST^{1s} . Hence it has probability one under $\text{RST}^{0s \rightarrow 1s}$ as well. Since $(T^{0s}, \omega_*^{0s \rightarrow 1s}) \sim \text{RST}^{0s \rightarrow 1s}$, this shows that $\varphi(T^{0s}) = 1$ a.s., a contradiction. ■

Remark 4.16. The main difficulty in extending this proof to the case $|\Sigma| \geq 3$ is that the regularity properties proven in Section 2.4 for $|\Sigma| = 2$ do not necessarily hold when $|\Sigma| = 3$. For example, when $|\Sigma| = 2$, if the pivot tree is supercritical, then it survives with positive probability conditional on either $\omega(R_T) = 0$ or $\omega(R_T) = 1$ by Proposition 2.6(c), which let us define measures RST^{0s} and RST^{1s} . When

$|\Sigma| \geq 3$, if the pivot tree is supercritical, it must survive with positive probability from some starting state, but it is not obvious that it must do so from multiple starting states. Nonetheless, we expect that it can be done and plan to address it in future work.

5 | APPLICATIONS OF THE MAIN RESULTS

5.1 | Monotone tree automata

We introduce here a special class of tree automata, which we term *monotone tree automata*. This class encompasses tree automata that arise out of many naturally occurring EMSO properties of rooted trees.

Consider an automaton A with set of colors Σ . Suppose that Σ has a total ordering on it, so that without loss of generality, we set $\Sigma = \{0, \dots, s\}$. Let $\vec{n} = (n_i : 0 \leq i \leq s)$ and $\vec{m} = (m_i : 0 \leq i \leq s)$, where n_i and m_i represent counts of children of type i . If $\sum_{i=0}^s n_i = \sum_{i=0}^s m_i$, then we write $\vec{n} \leq \vec{m}$ if one can modify the configuration of children from $\vec{n}$ to become $\vec{m}$ by only increasing the colors of children. (For example, $(1, 2, 1) \leq (1, 1, 2)$, since one moves from children 0,1,1,2 to 0,1,2,2 by increasing the color of a child from 1 to 2.) The automaton A is called *monotone* if $\vec{n} \leq \vec{m}$ implies that $A(\vec{n}) \leq A(\vec{m})$.

In the following lemma we state a notable characteristic of the pivot tree when we consider a monotone tree automaton A on two states.

Lemma 5.1. *Consider a tree automaton A on two colors. Then A is monotone if and only if pivotal children always have the same color as their parents.*

Proof. Let $\Sigma = \{0, 1\}$. Assume pivotal children always have the same state as their parents, and consider two configurations of children $\vec{n} \leq \vec{m}$. Suppose that $A(\vec{n}) = 1$. We can move from $\vec{n}$ to $\vec{m}$ only by changing vertices from state 0 to 1. These vertices are never pivotal, so $A(\vec{m}) = 1$. Since $A(\vec{n}) = 1$ implies $A(\vec{m}) = 1$, the automaton A is monotone.

Conversely, suppose that a node can have a pivotal child of the opposite state of itself. Then swapping this child's state changes the parent's state in the opposite direction, showing that A is not monotone. ■

Lemma 5.2. *Let A be a monotone tree automaton with $\Sigma = \{0, 1\}$, and let T_{piv} be the pivot tree associated with some fixed point. Then*

$$\mathbf{E}[\ell_1(T_{\text{piv}}) \mid \omega(R_T) = 0] = \mathbf{E}[\ell_1(T_{\text{piv}}) \mid \omega(R_T) = 1].$$

That is, the expected number of pivotal children that a vertex has is the same regardless of whether the vertex is labeled 0 or 1.

Proof. By Lemma 5.1,

$$\mathbf{E}[\ell_1(T_{\text{piv}}) \mid \omega(R_T) = 0] = \mathbf{E}[Z_0 \mid \omega(R_T) = 0],$$

and

$$\mathbf{E}[\ell_1(T_{\text{piv}}) \mid \omega(R_T) = 1] = \mathbf{E}[Z_1 \mid \omega(R_T) = 1],$$

using the notation of Lemma 2.5. By this lemma, these quantities are equal. ■

When $|\Sigma| = 2$, the automaton distribution map Ψ maps a distribution $\text{Bernoulli}(x)$ to $\text{Bernoulli}(y)$. We thus abuse notation and treat Ψ as a map from $[0, 1]$ to itself, writing $\Psi(x) = y$ instead of $\Psi(\text{Bernoulli}(x)) = \text{Bernoulli}(y)$. We also say that $p \in [0, 1]$ is a fixed point of Ψ rather than saying that $\text{Bernoulli}(p)$ is.

In the next lemma, we give a convenient way of determining whether a fixed point corresponding to a given monotone automaton with two colors is rogue or not.

Lemma 5.3. *Suppose A is a monotone automaton with set of colors $\Sigma = \{0, 1\}$. For a fixed point $\vec{v} = \text{Bernoulli}(p)$ with $0 < p < 1$, the growth rate of the pivot tree is equal to $\Psi'(p)$.*

Proof. In Lemma 2.6, we show that if M is the matrix of means for T_{piv} , then the spectral radius is equal to the expected number of pivotal children of the root, which is the growth rate of T_{piv} . Thus all we have to establish is that $\Psi'(p)$ is equal to the spectral radius of M .

The value of $\Psi(x)$ is given by the following procedure: Sample a number of children from χ and assign them i.i.d. $\text{Bernoulli}(x)$ states. Then, apply the automaton to determine the state of the parent. Then $\Psi(x)$ is the expected value of the parent. Abusing notation slightly by letting A act on a sequence of states as we did in the proof of Lemma 2.5, we have

$$\Psi(x) = \mathbf{E}^x[A(X_1, \dots, X_K)],$$

where $K \sim \chi$ and $(X_i)_{i \geq 1}$ are i.i.d. $\text{Bernoulli}(x)$ under $\mathbf{E}^x$. Let P denote the number of coordinates of $(X_1, \dots, X_k)$ that are pivotal for A at $(X_1, \dots, X_k)$. By the Margulis-Russo formula [11, Theorem 3.2],

$$\frac{d}{dx} \mathbf{E}^x[A(X_1, \dots, X_k) \mid K] = \mathbf{E}^x[P \mid K].$$

Taking expectations,

$$\Psi'(x) = \mathbf{E}^x[P].$$

Under $\mathbf{E}^p$, the random variable P has the distribution of the number of pivotal children of the root of (T, ω) . Hence,

$$\Psi'(p) = \mathbf{E}[\ell_1(T_{\text{piv}})]. \quad \blacksquare$$

Remark 5.4. Since a fixed point p of Ψ is attractive if $|\Psi'(p)| < 1$, this lemma together with Theorem 1.7 shows that for a monotone two-state automaton, an attractive fixed point is always interpretable. We mention that this is not true for nonmonotone automata. For example, the fixed point in Example 5.8 can be computed to be rogue for $\lambda = 4$ despite being attractive.

We are finally ready to answer Question 1.1. Recall that the at-least-two automaton assigns the parent state 1 if and only if at least two children have state 1. We mentioned in the introduction that with Poisson child distribution, this automaton has either one, two, or three fixed points depending on λ . We will prove this in detail now, and we will classify the fixed points as rogue or interpretable.

Example 5.5. Let A be the at-least-two automaton, and let $\chi \sim \text{Poi}(\lambda)$. As we saw in (2), the automaton distribution map is

$$\Psi(x) = 1 - e^{-\lambda x}(1 + \lambda x).$$

Define

$$\lambda_{\text{crit}} = \min_{x>0} \frac{x}{1 - e^{-x}(1+x)} \approx 3.35.$$

The function $x/(1 - e^{-x}(1+x))$ is convex on $(0, \infty)$ and hence has a unique minimizer, which we denote by x^* . Now, substituting λx for x in the function to be minimized, consider the inequality

$$\frac{\lambda x}{\Psi(x)} \leq \lambda \quad (34)$$

on $(0, \infty)$. If $\lambda < \lambda_{\text{crit}}$, it has no solutions, since $\lambda x/\Psi(x) \geq \lambda_{\text{crit}}$. Hence $\Psi(x) < x$ for $x > 0$, demonstrating that Ψ has no fixed points other than the trivial $x = 0$. If $\lambda = \lambda_{\text{crit}}$, then (34) has exactly one solution. The solution is $x = x^*/\lambda_{\text{crit}}$, and equality occurs in (34), making it a fixed point of Ψ . Since $x^* < \lambda_{\text{crit}}$, the solution lies in $(0, 1)$. Hence Ψ has one nontrivial fixed point in this case. If $\lambda > \lambda_{\text{crit}}$, then (34) has an interval of solutions $[a, b]$, which contains $x^*/\lambda \in (0, 1)$. It is easy to check that $a > 0$ and $b < 1$. Thus $\Psi(x)$ lies below the line $y = x$ on $(0, a)$, then sits above it on (a, b) , and then lies below it on $(b, 1]$, giving Ψ fixed points a, b in addition to 0.

Now, assume that $\lambda > \lambda_{\text{crit}}$, so that Ψ has fixed points 0, a , and b . Question 1.1 asks whether there exists a classification of trees into states $\{0, 1\}$ such that a tree has state 1 if and only if it has at least two children of state 1, and a Galton-Watson tree with $\text{Poi}(\lambda)$ child distribution has state 1 with probability a . In other words, the question is whether a is rogue or interpretable. By Lemma 5.3, we can answer this question by finding $\Psi'(a)$. Since $\Psi(x)$ lies under the curve $y = x$ up until $x = a$ and then rises above it, its derivative at $x = a$ exceeds 1. Therefore a is a rogue solution by Lemma 5.3 and Theorem 1.7.

The following result uses a similar approach:

Proposition 5.6. *The highest and lowest fixed points of a monotone two-state automaton are always interpretable.*

Proof. If $\Psi(1) = 1$, then the highest fixed point is 1, which has the trivial interpretation $t \mapsto 1$. Otherwise $\Psi(1) < 1$, and at the highest fixed point the graph of Ψ is either crossing from above the line $y = x$ to below, or it has $y = x$ as a tangent line (note that Ψ is continuously differentiable). In either case $\Psi'(x) \leq 1$ at the fixed point, making it interpretable by Lemma 5.3 and Theorem 1.7. The same argument shows that the smallest fixed point is also interpretable. ■

5.2 | More examples of two-state automata

We give some examples of tree automata for which a fixed point undergoes a phase transition from interpretable to rogue as a parameter of the child distribution is varied. We also give a numerical example to show that a two-state automaton can have many fixed points. As in the previous section, we write $\Psi(x) = y$ rather than $\Psi(\text{Bernoulli}(x)) = \text{Bernoulli}(y)$ for two-state automata.

Example 5.7. Consider the automaton where a node is in state 1 if and only if it has zero children in state 1, given formally by the map $(n_0, n_1) \mapsto \mathbf{1}\{n_1 = 0\}$. The distributional map corresponding to this automaton with $\text{Poi}(\lambda)$ offspring distribution is

$$\Psi(x) = e^{-\lambda x}.$$

Notice that if we consider the function $f_\lambda(x) = e^{-\lambda x} - x$, then $f''_\lambda(x) = \lambda^2 e^{-\lambda x} > 0$, showing that it is a convex function. Moreover, $f_\lambda(0) = 1$ and $f_\lambda(1) = e^{-\lambda} - 1 < 0$ for all $\lambda > 0$. Hence f_λ has a unique root in $(0, 1)$, which tells us that there is a unique fixed point $x(\lambda)$ of Ψ .

For this automaton, a node in state 0 has pivotal children if and only if it has exactly one child in state 1 (whatever the number of 0-state children may be), and this child will be pivotal. A node in state 1 has pivotal children if and only if it has at least one child, in which case every child will be pivotal. If X_0 denotes the total number of 0-state children and X_1 the total number of 1-state children of the root, then

$$\begin{aligned} \mathbf{E}[\ell_1(T_{\text{piv}})] &= \mathbf{E}[\mathbf{1}\{X_1 = 1\} + X_0 \mathbf{1}\{X_1 = 0\}] \\ &= \lambda x e^{-\lambda x} + \mathbf{E}[X_0] \mathbf{P}[X_1 = 0] \\ &= \lambda x e^{-\lambda x} + \lambda(1 - x) e^{-\lambda x} = \lambda e^{-\lambda x}. \end{aligned}$$

Thus, to determine if the fixed point is rogue or interpretable with Theorem 1.7, we have to determine if $\lambda e^{-\lambda x(\lambda)} \leq 1$.

Rewriting the equation $x(\lambda) = e^{-\lambda x(\lambda)}$, we find that

$$\log x(\lambda) = -\lambda x(\lambda) \iff y(\lambda) \log y(\lambda) = \lambda, \quad (35)$$

where $y(\lambda) = (x(\lambda))^{-1}$. Noting that the function $u \log u$ is strictly increasing over all $u > 1$ (which is the range we care about), we conclude that $\lambda > e$ if and only if $y(\lambda) > e$. In that case, from the first equation of (35), we have

$$\lambda e^{-\lambda x(\lambda)} = \lambda x(\lambda) = \log y(\lambda) > 1.$$

This shows that $E[\ell_1(T_{\text{piv}})] > 1$ for $\lambda > e$, and $E[\ell_1(T_{\text{piv}})] \leq 1$ for $\lambda \leq e$. By Theorem 1.7 and Proposition 2.6(a), the fixed point is interpretable for $\lambda \leq e$ and rogue for $\lambda > e$. This is illustrated in Figure 5.

Example 5.8. Consider the automaton A on color set $\Sigma = \{0, 1\}$, where a node is in state 1 if and only if it has at least one child in state 0 and at least one child in state 1. That is, the automaton is the map $(n_0, n_1) \mapsto \mathbf{1}\{(n_0 \geq 1) \wedge (n_1 \geq 1)\}$. The distributional map for this automaton with child distribution $\text{Poi}(\lambda)$ is given by

$$\Psi(x) = 1 - e^{-\lambda(1-x)} - e^{-\lambda x} + e^{-\lambda}. \quad (36)$$

Note that

$$\Psi'(x) = -\lambda e^{-\lambda(1-x)} + \lambda e^{-\lambda x}, \quad (37)$$

and

$$\Psi''(x) = -\lambda^2 e^{-\lambda(1-x)} - \lambda^2 e^{-\lambda x} < 0, \quad (38)$$

showing that Ψ is strictly concave. We observe that $\Psi(0) = 0$ and $\Psi'(0) = \lambda(1 - e^{-\lambda})$. Let $\lambda_0 \approx 1.35$ be the unique solution to $\lambda(1 - e^{-\lambda}) = 1$. If $\lambda \leq \lambda_0$, we have $\Psi'(0) \leq 1$, and the graph of $\Psi(x)$ stays below the line $y = x$ for all $x > 0$. Thus 0 is the only fixed point of $\Psi(x)$ in this case. If $\lambda > \lambda_0$, then

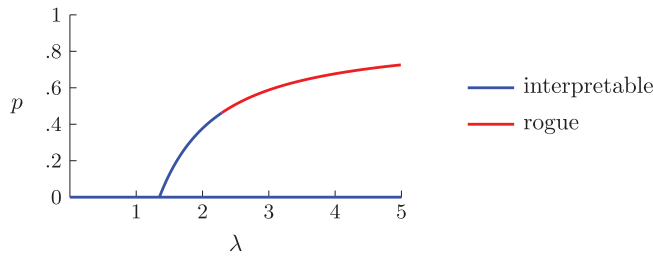


FIGURE 6 Plot for Example 5.8, showing the fixed point for the automaton assigning the parent state “1” if and only if it has at least one “0” child and at least one “1” child with a $\text{Poi}(\lambda)$ child distribution [Color figure can be viewed at wileyonlinelibrary.com]

$\Psi'(0) > 1$. Since $\Psi(1) = 0$, this implies that the graph of $\Psi(x)$ rises above the line $y = x$ and then dips back below it, giving rise to a nontrivial fixed point we denote by $x(\lambda)$.

Let X_i be the number of children of the root in state i . We summarize all configurations in which any of these children are pivotal:

Root in state 0

- $X_0 \geq 2, X_1 = 0$: X_0 pivotal children
- $X_0 = 0, X_1 \geq 2$: X_1 pivotal children

Root in state 1

- $X_0 = X_1 = 1$: two pivotal children
- $X_0 \geq 2, X_1 = 1$: one pivotal child
- $X_0 = 0, X_1 \geq 2$: one pivotal child

Thus,

$$\begin{aligned}
 \mathbf{E}[\ell_1(T_{\text{piv}})] &= \mathbf{E}\left[X_0 \mathbf{1}\{X_0 \geq 2, X_1 = 0\} + X_1 \mathbf{1}\{X_0 = 1, X_1 \geq 2\} \right. \\
 &\quad \left. + \mathbf{1}\{X_0 = X_1 = 1\} + \mathbf{1}\{X_0 \geq 2, X_1 = 1\} + \mathbf{1}\{X_0 = 1, X_1 \geq 2\} \right] \\
 &= e^{-\lambda x} \lambda (1-x) (1 - e^{-\lambda(1-x)}) + e^{-\lambda(1-x)} \lambda x (1 - e^{-\lambda x}) \\
 &\quad + \mathbf{P}[X_0 \geq 1, X_1 = 1] + \mathbf{P}[X_0 = 1, X_1 \geq 1] \\
 &= \lambda (e^{-\lambda x} + e^{-\lambda(1-x)} - 2e^{-\lambda}).
 \end{aligned}$$

Substituting from (36), we get

$$\mathbf{E}[\ell_1(T_{\text{piv}})] = \lambda (1 - x(\lambda) - e^{-\lambda}).$$

For $\lambda = \lambda_c$, we have $x(\lambda) = 0$ and $\mathbf{E}[\ell_1(T_{\text{piv}})] = \lambda_0(1 - e^{-\lambda_0}) = 1$. With some laborious calculus, one can establish that as λ increases, the quantity $\mathbf{E}[\ell_1(T_{\text{piv}})]$ decreases and then increases, reaching 1 at $\lambda_1 \approx 2.30$. Thus, by Theorem 1.7 and Proposition 2.6(a), this fixed point $x(\lambda)$ is interpretable for $\lambda \in [\lambda_0, \lambda_1]$ and rogue for $\lambda > \lambda_1$. A plot showing the behavior of the fixed points is given in Figure 6.

Example 5.9. Finally we present an example to demonstrate that the automaton may have many fixed points. Consider the automaton A on color set $\Sigma = \{0, 1\}$, where a node is in state 1 unless it has x children in state 1 for $x \in \{1, 2, 3, 4, 5\} \cup \{8, 9, 10, 11\}$ in which case it is state 0. That is, the

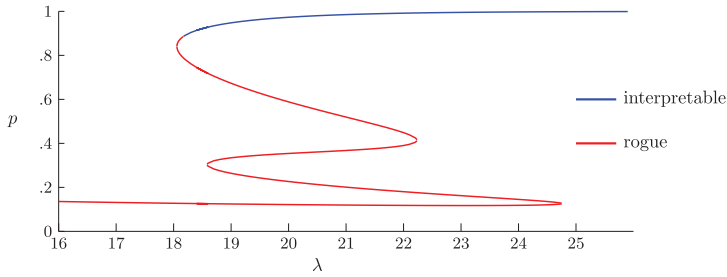


FIGURE 7 A plot of the fixed points for the automaton in Example 5.9 with $\text{Poi}(\lambda)$ child distribution [Color figure can be viewed at wileyonlinelibrary.com]

automaton is the map $(n_0, n_1) \mapsto \mathbf{1}\{(n_1 \in \{0, 6, 7\}) \vee (n_1 \geq 12)\}$. The plot of the fixed points for this automaton with child distribution $\text{Poi}(\lambda)$ is shown in Figure 7.

5.3 | First-order interpretations

As we mentioned in the introduction, the papers [24, 25] investigated tree automata and fixed points corresponding to statements of first-order logic. The goal of [25] is to study the probability that $T \sim \text{GW}(\text{Poi}(\lambda))$ satisfies some given first-order sentence of quantifier depth k . Recall from Section 1.9 that there is an automaton on the set of rank k types and an interpretation given by sending a tree to its type. Assuming that the child distribution is $\text{Poi}(\lambda)$, the automaton distribution map for the tree automaton is then shown to be a contraction [25, Theorem 3.2], which implies that it has a unique fixed point. This fixed point is also shown to be a smooth function of λ [25, Theorem 2.4]. Finally, this is applied to the original problem: since the set of trees satisfying a given first-order sentence φ of quantifier depth k is the union of a collection of rank k types, the probability that T satisfies φ is also a smooth function of λ .

All of this work was done with no explicit mention of the concept of interpretations. Our goal here is to put it more comfortably into this paper's framework. We call $\iota: \mathcal{T} \rightarrow \Sigma$ a *first-order interpretation* if each set of trees $\{t \in \mathcal{T} : \iota(t) = \sigma\}$ for $\sigma \in \Sigma$ can be defined in the first-order language described in Section 1.9.

To avoid reproving results in [24, 25], we continue to assume that $\chi \sim \text{Poi}(\lambda)$, but we expect that the results should hold for general child distributions. As usual, the assumption that a fixed point has no zero entries causes no loss of generality, since the set Σ can be shrunk and the automaton considered as one on a smaller set of states.

Theorem 5.10. *Assume that $\chi \sim \text{Poi}(\lambda)$, and let Σ be any finite set of states. Let $\iota: \mathcal{T} \rightarrow \Sigma$ be an interpretation of an automaton A corresponding to a fixed point $\vec{v}$, which we assume has strictly positive entries. If ι is a first-order interpretation, then $\vec{v}$ is the only fixed point of the automaton distribution map.*

Proof. Let $\mathcal{T}_n^{\text{taut}} \subseteq \mathcal{T}$ consist of all trees t on which $\iota(t)$ is tautologically determined by $t|_n$. That is, $\mathcal{T}_n^{\text{taut}}$ consists of all trees t such that $\iota(t) = \iota(t')$ for all $t' \in [t]_n$. It follows from [25, Lemma 5.6] that

$$\lim_{n \rightarrow \infty} \mathbf{P}[T \in \mathcal{T}_n^{\text{taut}}] = 1. \quad (39)$$

Let (t, τ) be an arbitrary tree whose coloring is compatible with the automaton A . We claim that if $t \in \mathcal{T}_n^{\text{taut}}$, then $\tau(R_t) = \iota(t)$. Indeed, condition on $T \in [t]_n$. Under this conditioning, the vector $(\iota(T(v)))_{v \in L_n(t)}$ is i.i.d. $\vec{v}$. Since $\vec{v}$ has strictly positive entries, this vector takes on each value in $\Sigma^{\ell_n(t)}$

with positive probability. Since $t \in \mathcal{T}_n^{\text{taut}}$, we have $\iota(T) = \iota(t)$ a.s. But $\iota(T)$ is given by iteratively applying the automaton to $(\iota(T(v)))_{v \in L_n(t)}$, from which we can conclude that applying the automaton in this way to *any* vector in $\Sigma^{\ell_n(t)}$ yields $\iota(t)$. Thus, since $\tau(R_t)$ is given by applying the automaton to $(\iota(t(v)))_{v \in L_n(t)}$, it too is equal to $\iota(t)$.

Now, suppose that $\vec{v}'$ is another fixed point of the automaton map, and let (T, ω) be the random state tree associated with $\vec{v}'$. If $T \in \mathcal{T}_n^{\text{taut}}$ for any n , then $\omega(R_T) = \iota(T)$ by the claim we have just proved. By (39), it holds with probability one that $T \in \mathcal{T}_n^{\text{taut}}$ for some value of n (observe that $T \in \mathcal{T}_n^{\text{taut}}$ forms an increasing sequence of events). Hence $\omega(R_T) = \iota(T)$ a.s. Thus $\omega(R_T) \sim \vec{v}$, since $\iota(T) \sim \vec{v}$. But by the definition of the random state tree, $\omega(R_T) \sim \vec{v}'$, demonstrating that $\vec{v} = \vec{v}'$. ■

6 | FURTHER QUESTIONS

The most straightforward open problem is to extend Theorem 1.7 to $3 \leq |\Sigma| < \infty$. Theorem 1.8 already provides one direction of the theorem, leaving the critical and supercritical cases. As we discussed in Remarks 3.6 and 4.16, it might be possible to adapt the two-state proofs. In both cases, the difficulty is that the pivot tree need not be positive regular. In fact, it seems to us that when the pivot tree is positive regular, all proofs go through as is, and Theorem 1.7 holds in general for $|\Sigma| < \infty$ (though we have not checked every detail).

Beyond this, two generalizations interest us. First, extending the results to infinite state spaces would allow the theory to address situations like those considered in [20]. Second, one could consider randomized automata: give each vertex v an independent source of randomness X_v and then allow the automaton to determine the state of a vertex from the states of its children together with X_v . This situation often arises in practice and is the model considered in [1]. Extending the theory to this case might yield answers to questions about endogeneity, as discussed in Section 1.9.

In a different direction, we wonder what configurations of fixed points are possible. For example, when $3 \leq |\Sigma| < \infty$, can an automaton have infinitely many fixed points? (This can be ruled out when $|\Sigma| = 2$ by arguing that the automaton distributional map is analytic.) In the case $|\Sigma| = 2$, for any specified finite set of rogue and interpretable fixed points, is there an automaton and a child distribution to match them? Even restricting ourselves to two-state monotone automata, it is not clear which sets of rogue and interpretable fixed points can occur.

Section 5.3 also raises some questions. For example, Theorem 5.10 provides a condition on an interpretation that makes the corresponding tree automaton have a unique fixed point. This suggests the problem of giving conditions on the tree automaton itself that force its automaton distribution map to have a unique fixed point.

ACKNOWLEDGMENTS

We thank Joel Spencer, who set this work in motion and generously advised us. We also thank Leonid Libkin for guiding us through the literature on tree automata in logic.

REFERENCES

1. D.J. Aldous and A. Bandyopadhyay, *A survey of max-type recursive distributional equations*, Ann. Appl. Probab. **15** (2005), 1047–1110. MR 2134098 (2007e:60010).
2. G. Alsmeyer and M. Meinert, *Fixed points of the smoothing transform: Two-sided solutions*, Probab. Theory Related Fields. **155** (2013), 165–199. MR 3010396.
3. K.B. Athreya and P.E. Ney, *Branching processes*. Die Grundlehren der mathematischen Wissenschaften, Band 196, Springer-Verlag, New York-Heidelberg, 1972. MR 0373040.

4. A. Bandyopadhyay, *Endogeny for the logistic recursive distributional equation*, *Z. Anal. Anwend.* **30** (2011), 237–251. MR 2793003.
5. J. Bourgain, J. Kahn, G. Kalai, Y. Katznelson, and N. Linial, *The influence of variables in product spaces*, *Israel J. Math.* **77** (1992), 55–64. MR 1194785.
6. N. Broutin, L. Devroye, and N. Fraiman, *Recursive functions on conditional Galton-Watson trees*, 2018, arXiv:1805.09425.
7. H. Comon, M. Dauchet, R. Gilleron, C. Löding, F. Jacquemard, D. Lugiez, S. Tison, and M. Tommasi, *Tree automata techniques and applications*, 2007, available at <http://www.grappa.univ-lille3.fr/tata>.
8. F.M. Dekking, *Branching processes that grow faster than binary splitting*, *Amer. Math. Monthly.* **98** (1991), 728–731. MR 1130682.
9. R. Durrett, *Probability: Theory and examples*, 4.1th ed., April 21, 2013, available at <http://services.math.duke.edu/rtd/PTE/pte.html>. Fourth edition published by Cambridge University Press in 2010.
10. E. Friedgut and G. Kalai, *Every monotone graph property has a sharp threshold*, *Proc. Amer. Math. Soc.* **124** (1996), 2993–3002. MR 1371123.
11. C. Garban and J.E. Steif, *Noise sensitivity of Boolean functions and percolation*. Institute of Mathematical Statistics Textbooks, Cambridge University Press, New York, 2015. MR 3468568.
12. T.E. Harris, *The theory of branching processes*. Die Grundlehren der Mathematischen Wissenschaften, Bd. 119, Springer-Verlag, Berlin; Prentice-Hall, Inc., Englewood Cliffs, NJ, 1963. Also available as RAND report R-381-PR, <https://www.rand.org/pubs/reports/R381.html>, MR 0163361.
13. S. Jacka and M. Sheehan, *The noisy veto-voter model: A recursive distributional equation on $[0, 1]$* , *J. Appl. Probab.* **45** (2008), 670–688. MR 2455177.
14. J. Kahn, G. Kalai, and N. Linial, *The influences of variables on Boolean functions*. 29th Annual Symposium on Foundations of Computer Science (White Plains, 1988), IEEE Computer Society Press, Washington, D.C., pp. 68–80, 1988.
15. O. Kallenberg, *Foundations of modern probability*. Probability and its Applications (New York), Springer-Verlag, New York, 2002, 2nd ed. MR 1876169.
16. V. Kleptsyn and M. Triestino, *Cut-off method for endogeny of recursive tree processes*, 2016, arXiv:1610.06946.
17. L. Libkin, *Elements of finite model theory*. Texts in Theoretical Computer Science. An EATCS Series, Springer-Verlag, Berlin, 2004. MR 2102513.
18. R. Lyons, R. Pemantle, and Y. Peres, *Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes*, *Ann. Probab.* **23** (1995), 1125–1138. MR 1349164.
19. R. Lyons and Y. Peres, *Probability on trees and networks*. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, New York, Vol. 42, 2016. MR 3616205.
20. J.B. Martin and R. Stasiński, *Minimax functions on Galton-Watson trees*, 2018, arXiv:1806.07838.
21. R. O'Donnell, *Analysis of Boolean functions*, Cambridge University Press, New York, 2014. MR 3443800.
22. A.G. Pakes and F.M. Dekking, *On family trees and subtrees of simple branching processes*, *J. Theoret. Probab.* **4** (1991), 353–369. MR 1100239.
23. B. Pittel, J. Spencer, and N. Wormald, *Sudden emergence of a giant k -core in a random graph*, *J. Combin. Theory Ser. B.* **67** (1996), 111–151. MR 1385386.
24. M. Podder and J. Spencer, *First order probabilities for Galton-Watson trees*. A Journey Through Discrete Mathematics, Springer, Cham, pp. 711–734, 2017. MR 3726620.
25. M. Podder and J. Spencer, *Galton-Watson probability contraction*, *Electron. Commun. Probab.* **22** (2017). Paper No. 20, 16. MR 3627009.
26. B. A. Sevast'yanov, *On the theory of branching random processes*, *Doklady Akad. Nauk SSSR (N.S.)*. **59** (1948), 1407–1410. MR 0024090.
27. W. Thomas, *Languages, automata, and logic*. Handbook of Formal Languages, Springer, Berlin, Vol. 3, 1997, pp. 389–455. MR 1470024.

How to cite this article: Johnson T, Podder M, Skerman F. Random tree recursions: Which fixed points correspond to tangle sets of trees?. *Random Struct Alg.* 2020;56:796–837. <https://doi.org/10.1002/rsa.20895>