These lecture notes are for a short course given in spring 2016 at the Vietnam Institute for Advanced Study in Mathematics. The material in the first two lectures is discussed (in much greater detail) in the upcoming book by Lyons and Peres [3]. The material in the third lecture is taken from papers by Evans et al. [2] and by Borgs et al. [1].

1 Infinite trees

Definition 1.1. *A* labelled, rooted tree T *is a non-empty set of finite sequences of non-negative integers such that for every* $n \ge 1$ *and* $(i_1, \dots, i_n) \in T$,

1. *the sequence* $(i_1, \dots, i_{n-1}) \in T$*; and*
2. *for every* $1 \le j \le i_n$*, the sequence* $(i_1, \dots, i_{n-1}, j) \in T$.

Applying the first property in Defintion 1.1 repeatedly, it follows that every labelled, rooted tree contains the empty sequence $\emptyset$; we call this element the *root*. The *children* of $u = (i_1, \dots, i_n) \in T$ (denoted $C(u)$) is the set of $v \in T$ of the form $(i_1, \dots, i_n, i_{n+1})$. The *parent* of a non-empty $u = (i_1, \dots, i_n) \in T$ is $(i_1, \dots, i_{n-1})$. A *ray* of T is an infinite sequence $(i_1, i_2, \dots)$ with the property that $(i_1, \dots, i_n) \in T$ for every n. We write $|\cdot|$ for the length of a sequence (so that $|(i_1, \dots, i_n)| = n$), and T_n for $\{x \in T : |x| = n\}$. If $u = (i_1, \dots, i_n) \in T$, then we write $T^{(u)}$ for the *subtree at* u:

$$T^{(u)} = \{(j_1, \dots, j_m) : (i_1, \dots, i_n, j_1, \dots, j_m) \in T\}.$$

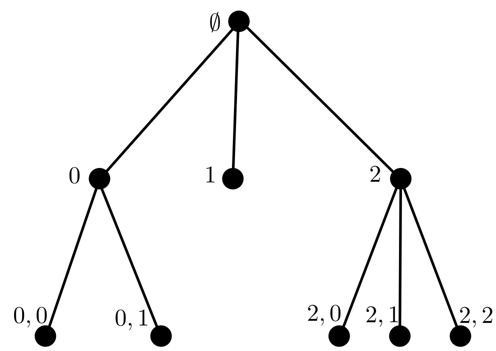
We draw a tree in rows, starting with a dot for $\emptyset$ at the top. In row n, we draw a dot for each $u \in T_n$ (from left to right in lexicographic order) and a line connecting it to its parent in T_{n-1}. Here is an example:

Definition 1.2. *The labelled, rooted tree* T *is* locally finite *if* $|T_n| < \infty$ *for all* n.

In these notes, we will only be talking about labelled, rooted, locally finite trees; from now on, we will just call them *trees*.

Definition 1.3 (Growth rate)**.** *For a tree* T*, define the lower and upper growth rates of* T *by*

$$\underline{\mathrm{gr}}(T) = \liminf_{n\to\infty} |T_n|^{1/n}$$
$$\overline{\mathrm{gr}}(T) = \limsup_{n\to\infty} |T_n|^{1/n}.$$



If $\underline{\mathrm{gr}}(T) = \overline{\mathrm{gr}}(T)$*, define the* growth rate *of* T *by*

$$\mathrm{gr}(T) = \underline{\mathrm{gr}}(T) = \overline{\mathrm{gr}}(T) = \lim_{n\to\infty} |T_n|^{1/n}.$$

The growth rate of a tree is a reasonable way of measuring how "large" it is, but it has at least two clear flaws: first, it is not defined for every tree; even if it is defined, it depends on the structure of the tree in a very coarse way. It turns out that there is a different way of measuring the "size" of a tree that is in many senses better than the growth rate.

1.1 Constrained flows and the branching number

Definition 1.4. *A* flow θ *on the tree* T *is a function* $\theta : T \to [0, \infty)$ *such that for every* $u \in T$*,*

$$\theta(u) = \sum_{v\in C(u)} \theta(v). \tag{1}$$

The strength *of* θ *is* $|\theta| = \theta(\emptyset)$*.*

If we think of the tree T as a collection of pipes, then we can think of $\theta(u)$ as the amount of water flowing into u from its parent (unless $u = \emptyset$, in which case $\theta(u)$ is the amount of water that we are pouring into u). The condition (1) says that the amount of water flowing into u is the same as the amount flowing out. Note that $\theta \equiv 0$ is a flow; we call it the *trivial flow.*

Exercise 1.1.

1. *For any flow* θ *and any* $n \geq 0$*,* $\sum_{u\in T_n} \theta(u) = |\theta|$*.*
2. *If* $|T| < \infty$ *then the only flow on* T *is* $\theta \equiv 0$*.*

Definition 1.5. *For $\lambda > 0$, let $\Theta(T, \lambda)$ be the set of flows θ on T satisfying $|\theta(u)| \le \lambda^{-|u|}$ for all $u \in T$; we say $\Theta(T, \lambda)$ is trivial if it contains only the trivial flow. The* branching number *of T is*

$$\begin{aligned} \operatorname{br}(T) &= \inf\{\lambda > 0 : \Theta(T, \lambda) \text{ is trivial}\} \\ &= \sup\{\lambda > 0 : \Theta(T, \lambda) \text{ is non-trivial}\}. \end{aligned}$$

To see that the two definition of $\operatorname{br}(T)$ are equivalent, note that $\Theta(T, \lambda) \subseteq \Theta(T, \lambda')$ for all $\lambda \ge \lambda'$. It follows that $\Theta(T, \lambda)$ is trivial for every $\lambda > \operatorname{br}(T)$, while $\Theta(T, \lambda)$ is non-trivial for every $\lambda < \operatorname{br}(T)$.

By part 1 of Exercise 1.1, there is a one-sided relationship between growth rate and branching number: for any $\lambda > 0$, any $n \ge 0$, and any $\theta \in \Theta(T, \lambda)$, Exercise 1.1 implies that

$$|\theta| = \sum_{u \in T_n} \theta(u) \le |T_n| \lambda^{-n}.$$

Taking $n \to \infty$, we have $|\theta| \le \liminf_n (|T_n|^{1/n}/\lambda)^n$; if $\lambda > \underline{\operatorname{gr}}(T)$ then the right hand side is zero and we conclude that $\Theta(T, \lambda)$ is trivial. Hence,

$$\operatorname{br}(T) \le \underline{\operatorname{gr}}(T). \tag{2}$$

Example 1.6. *Let T be the k-ary tree – that is, T consists of all finite strings over $\{0, \dots, k-1\}$. Since $|T_n| = k^n$ for all n, $\operatorname{gr}(T) = k$. By (2), $\operatorname{br}(T) \le k$. On the other hand, $\theta(u, v) = k^{-|v|}$ (for $v \in C(u)$) is a non-zero flow that belongs to $\Theta(T, k)$; hence, $\operatorname{br}(T) = \operatorname{gr}(T) = k$.*

Example 1.7. *We can construct a tree T with $|T_n| = 2^n$ for every n (and hence $\operatorname{gr}(T) = 2$), but with branching number 1. We will describe it in words, and leave some details as an exercise: first, let $\emptyset$ have two children. Inductively for $n \ge 1$, write down T_n in lexicographic order. The first 2^{n-1} elements get three children each; the second 2^{n-1} elements get one child each.*

Since this describes a tree of infinite height, $\operatorname{br}(T) \ge 1$. To see that $\operatorname{br}(T) = 1$, take any $\lambda > 1$ and suppose that $\theta \in \Theta(T, \lambda)$. Now, for any $u \ne 0^n \in T$, $\operatorname{gr}(T^{(u)}) = 1$ (by Exercise 1.2). Hence, (2) applied to $T^{(u)}$ implies that $\theta(u) = 0$. Since $\theta(\emptyset) = \sum_{u \in T_n} \theta(u)$ for all n, we must have $\theta(\emptyset) = \theta(0^n) \le \lambda^{-n}$ for all n, and so $\theta \equiv 0$.

Exercise 1.2.

1. *Write down T from Example 1.7 explicitly as a set of finite strings of non-negative integers.*

2. *Show that* $\mathrm{gr}(T^u) = 1$ *for all* $u \neq 0^n \in T$.

3. *For every* $\alpha > 1$, *give an example of a tree* T *with* $\mathrm{gr}(T) = \alpha$ *but* $\mathrm{br}(T) = 1$.

1.2 Electricity and the branching number

We defined the branching number in terms of contrained flows (or, intuitively, water flowing through pipes of limited capacity). It turns out that the branching number can also be characterized in terms of the "energy" of a "resistor network." Given $\lambda > 0$ and a flow θ, we define the energy of θ with respect to λ by

$$\mathscr{E}_\lambda(\theta) = \sum_{u \in T} \lambda^{|u|} \theta^2(u).$$

The intuition here is that we construct the tree T out of electrical resistors, where the resistor on the edge leading into u has a resistance of $\lambda^{|u|}$ ohms. Now we send an electrical current through the network, where $\theta(u)$ amps flow into u. If this were a real circuit, $\mathscr{E}_\lambda(\theta)$ would be the amount of energy dissipated by the resistors. It turns out that this electrical analogy may be carried much further; see [3] for more details.

Theorem 1.8.

$$\mathrm{br}(T) = \sup\{\lambda : \exists \text{ non-trivial flow } \theta \text{ with } \mathscr{E}_\lambda(\theta) < \infty\}.$$

Proof of one direction. We will only prove one direction of Theorem 1.8 (the direction that we will need later); for the other direction, see [3, Chapter 3]. Specifically, we will show that

$$\mathrm{br}(T) \leq \sup\{\lambda > 0 : \exists \text{ non-trivial flow } \theta \text{ with } \mathscr{E}_\lambda(\theta) < \infty\}.$$

To show this, take $\lambda < \mathrm{br}(T)$ and choose some $\lambda' \in (\lambda, \mathrm{br}(T))$. Then there is some non-trivial flow $\theta \in \Theta(T, \lambda')$. For this θ,

$$\begin{aligned}
\mathscr{E}_\lambda(\theta) &= \sum_{n \geq 0} \sum_{u \in T_n} \theta^2(u) \lambda^{|u|} \\
&\leq \sum_{n \geq 0} \sum_{u \in T_n} \theta(u) (\lambda/\lambda')^{|u|} \\
&= \sum_{n \geq 0} (\lambda/\lambda')^n \sum_{u \in T_n} \theta(u).
\end{aligned}$$

Recalling from Exercise 1.1 that $\sum_{u \in T_n} \theta(u) = |\theta|$, it follows that

$$\mathscr{E}_\lambda(\theta) \le |\theta| \sum_{n \ge 0} (\lambda/\lambda')^n < \infty. \qquad \square$$

Example 1.9. *Let T be the binary tree (that is, the k-ary tree for $k = 2$). In Example 1.6, we saw that* $\mathrm{br}(T) = 2$ *and that "water flows" at the critical value (i.e., $\Theta(T, 2)$ is non-trivial). By Theorem 1.8, 2 is also the critical value for "electricity to flow" on T. However, electricity does not flow at the critical value: indeed, if θ is a non-trivial flow on T then*

$$\mathscr{E}_2(\theta) = \sum_{n \ge 0} 2^n \sum_{u \in T_n} \theta^2(u) \ge \sum_{n \ge 0} \left(\sum_{u \in T_n} \theta(u) \right)^2 = \infty,$$

where the inequality follows by the Cauchy-Schwarz inequality and the last equality follows from Exercise 1.1.

Exercise 1.3. *The purpose of this exercise is to construct a tree T such that electricity does flow at the critical value. Let a_n be some non-decreasing sequence such that $a_1 = 1$ and $a_n \lambda^{-n} \to 0$ for all $\lambda > 1$. Given such a sequence, let T be the tree constructed recursively as follows: if $3|T_{n-1}| \le a_n 2^n$ then let every $u \in T_{n-1}$ have 3 children; otherwise, let $u \in T_{n-1}$ have 2 children.*

1. *Show that* $\mathrm{br}(T) = 2$.
2. *Construct a sequence a_n and a flow θ on the resulting tree T such that $\mathscr{E}_2(\theta) < \infty$.*

1.3 Cutsets and the branching number

A third characterization of the branching number comes from the notion of a cutset.

Definition 1.10. *A set $\Pi \subset T$ is a* cutset *if for every ray $(i_1, i_2, \dots)$ there exists some n such that $(i_1, \dots, i_n) \in \Pi$. The cutset Π is* minimal *if for every $(i_1, \dots, i_n) \in T$, there is at most one $m \le n$ such that $(i_1, \dots, i_m) \in \Pi$.*

It is easy to check that a cutset Π is minimal if and only if it is *inclusion-minimal*, in the sense that every cutset $\Pi' \subseteq \Pi$ satisfies $\Pi' = \Pi$.

Exercise 1.4. *Show that every minimal cutset is finite. Hint: suppose we have an infinite, minimal cutset Π. Take a sequence $u_n \in T$, $|u_n| \to \infty$ such that u_n has a child in Π. Use the local finiteness property and a diagonalization argument to construct a ray that does not pass through Π.*

Theorem 1.11. *For any tree T,*

$$\mathrm{br}(T) = \inf\{\lambda > 0 : \inf_{\Pi} \sum_{u\in\Pi} \lambda^{-|u|} = 0\},$$

where the inner infimum runs over all cutsets Π*.*

Theorem 1.11 is essentially a special case of the well-known "max flow min cut" theorem in graph theory. In order to be self-contained, we will give a proof.

Proof. For $\lambda < \mathrm{br}(T)$, there exists some non-trivial $\theta \in \Theta(T, \lambda)$. Then for every minimal cutset Π, $\sum_{u\in\Pi} \lambda^{-|u|} \geq \sum_{u\in\Pi} \theta(u) = |\theta|$. (This is an extension of Exercise 1.1, and we leave it as an exercise.) Hence,

$$\mathrm{br}(T) \leq \inf\{\lambda > 0 : \inf_{\Pi} \sum_{u\in\Pi} \lambda^{-|u|} = 0\}.$$

The other direction is a little more difficult. We fix λ and suppose that $\inf_\Pi \sum_{u\in\Pi} \lambda^{-|u|} = \epsilon > 0$; we will prove Theorem 1.11 by constructing a non-trivial flow in $\Theta(T, \lambda)$.

First, we restrict ourselves to finite levels of T: say that $\theta : T \to [0, \infty)$ is a *flow up to n* if $\theta(u) = 0$ for all $|u| > n$ and

$$\theta(u) = \sum_{v\in C(u)} \theta(v)$$

for all $|u| < n$. Let $\Theta_n(T, \lambda)$ be the set of flows θ up to n satisfying $\theta(u) \leq \lambda^{-|u|}$ for every u. Note that we may define the energy of flows up to n in the same way that we defined it for flows. We say that a cutset Π is a *cutset up to n* if every $u \in \Pi$ satisfies $|u| \leq n$.

Now fix $n \in \mathbb{N}$. Note that (by the local finiteness property), $\Theta_n(T, \lambda)$ is compact; we claim that

$$\max\{|\theta| : \theta \in \Theta_n(T, \lambda)\} \geq \epsilon. \qquad (3)$$

By compactness, we may choose some θ such that $|\theta|$ is maximal among all $\theta \in \Theta_n(T, \lambda)$. Say that $u \in T, |u| \leq n$ is *augmentable* if $\theta(v) < \lambda^{-|v|}$ for every v that is an ancestor of, or equal to, u. Since θ has maximal strength, every augmentable u must satisfy $|u| < n$ (otherwise, we could increase θ's strength by increasing the flow slightly for every ancestor of u). Let $A \subset T$ be the set of augmentable vertices; note that if $u \in A$ then u's parent is also in A. Moreover, we may assume that $\emptyset \in A$; otherwise we are done, because $|\theta| = \theta(\emptyset) = 1$, which is at least ϵ because $\Pi = \{\emptyset\}$ is a cutset.

Let Π be the set of $u \in T$ such that $u \notin A$ but u's parent is in A. We claim that

(i) Π is a cutset up to n,

(ii) Π is minimal, and

(iii) $\theta(u) = \lambda^{-|u|}$ for every $u \in \Pi$.

To prove (i), note first that $u \in \Pi$ implies that the parent (v, say) of u is in A; hence, $|v| < n$ and so $|u| \leq n$. Moreover, Π is a cutset because every $u \in T_n$ has some ancestor in A; hence, it also has some ancestor (or possibly u itself) in Π. For (ii) note that if $u \in \Pi$ then u's parent is augmentable and so every ancestor of u is augmentable (and hence not in Π). To prove (iii), if u's parent were augmentable and $\theta(u)$ were strictly smaller than $\lambda^{-|u|}$ then u would also be augmentable (and hence not in Π). Hence,

$$|\theta| = \sum_{u \in \Pi} \theta(u) = \sum_{u \in \Pi} \lambda^{-|u|} \geq \epsilon,$$

where the first equality follows from (i) and (ii) and the second follows from (iii). This proves (3).

To complete the proof, for every n take $\theta_n \in \Theta_n(T, \lambda)$ with $|\theta_n| \geq \epsilon$. By a diagonalizaton argument, there is a subsequence n_k such that $\theta_{n_k}(u)$ converges for every u; define θ by $\theta(u) = \lim_{k \to \infty} \theta_{n_k}(u)$. Then

(i) $\theta(u) \leq \lambda^{-|u|}$ because every $\theta_{n_k}(u) \leq \lambda^{-|u|}$;

(ii) $\theta(\emptyset) \geq \epsilon$ because every $\theta_{n_k}(\emptyset) \geq \epsilon$; and

(iii) θ is a flow: for every $u \in T$ and all k large enough so that $n_k > |u|$,

$$\theta_{n_k}(u) = \sum_{v \in C(u)} \theta_{n_k}(v),$$

and so the same holds for θ.

In particular, $\theta \in \Theta(T, \lambda)$ and $\theta \not\equiv 0$. $\square$

2 Galton-Watson processes

Choose a sequence of numbers $\{p_k : k \geq 0\}$ satisfying $p_k \geq 0$ and $\sum_{k\geq 0} p_k = 1$. We will define a random tree recursively: set $T_0 = \{\emptyset\}$. Then for every $n \geq 0$, do the following: take independent random variables $\{X_u : u \in T_n\}$ with distribution $P(X_u = k) = p_k$. Define T_{n+1} by taking $u \in T_n$ to have X_u children in T_{n+1}. The resulting tree T is called a *Galton-Watson* tree.

It is sometimes instructive to consider just the generation sizes $Z_n := |T_n|$. From the construction above, it is clear that the distribution of Z_n is characterized by the recursion

$$Z_{n+1} = \sum_{i=1}^{Z_n} X_{n,i},$$

where $X_{n,i}$ are independent random variables with $P(X_{n,i} = k) = p_k$.

Clearly, if $Z_n = 0$ then $Z_m = 0$ for all $m \geq n$. We call the event $\{\exists n : Z_n = 0\}$ *extinction*; let q be its probability. If $p_0 = 0$ then $q = 0$; if $p_0 > 0$ then $q > 0$ (because the probability of going extinct in the first generation is p_0). The next interesting question is whether $q = 1$ or $q < 1$.

Theorem 2.1. *$q = 1$ if and only if $p_1 \neq 1$ and*

$$\sum_{k\geq 0} kp_k \leq 1.$$

In order to prove Theorem 2.1, we introduce the *probability generating function* of p_k:

$$f(s) = \sum_{k\geq 0} p_k s^k,$$

which converges at least for $s \in [0, 1]$ (taking the convention that $0^0 = 1$). Let $f^{(n)}(s)$ denote f composed with itself n times (i.e., $f^{(0)}(s) = s$ and $f^{(n+1)}(s) = f(f^{(n)}(s))$).

Lemma 2.2. *For all $s \in [0, 1]$ and $n \geq 0$, $\mathbb{E}[s^{Z_{n+1}} \mid Z_1, \ldots, Z_n] = f(s)^{Z_n}$. Moreover, $\mathbb{E}[s^{Z_n}] = f^{(n)}(s)$.*

Proof. To prove the first claim,

$$\begin{aligned}\mathbb{E}[s^{Z_{n+1}} \mid Z_1, \dots, Z_n] &= \mathbb{E}\Big[\prod_{i=1}^{Z_n} s^{X_{n,i}} \mid Z_n\Big] \\ &= \prod_{i=1}^{Z_n} \mathbb{E}[s^{X_{n,i}}] \\ &= f(s)^{Z_n}.\end{aligned}$$

The second claim follows by induction: it is obvious for $n = 0$. For the inductive step,

$$\mathbb{E}[s^{Z_{n+1}}] = \mathbb{E}\mathbb{E}[s^{Z_{n+1}} \mid Z_1, \dots, Z_n] = \mathbb{E}[f(s)^{Z_n}] = f(f^{(n)}(s)) = f^{(n+1)}(s).$$

□

Proof of Theorem 2.1. Assume that $p_0 > 0$ and $p_1 < 1$ (otherwise, the proof is easy). Note that f is an increasing, convex function with $f(0) = p_0 > 0$. Moreover,

$$f'(1) = \sum_{k \geq 0} k p_k$$

(which may be infinite). Some elementary calculus shows that the equation $f(s) = s$ has a solution in $(0, 1)$ if and only if $f'(1) > 1$. If this solution exists, then it is unique and $f^{(n)}(0)$ converges to it; otherwise, $f^{(n)}(0)$ converges to one. Since

$$q = \Pr(Z_n \to 0) = \lim_{n \to \infty} \Pr(Z_n = 0) = \lim_{n \to \infty} \mathbb{E}[0^{Z_n}] = \lim_{n \to \infty} f^{(n)}(0),$$

Theorem 2.1 follows. □

2.1 The growth rate of a Galton-Watson tree

From now on, we will be interested in Galton-Watson trees that don't go extinct. Therefore, we will assume that $m := \sum_{k \geq 0} k p_k > 1$ – these are called *supercritical* Galton-Watson trees. One might expect that on the event that if the tree doesn't go extinct then it will grow exponentially with rate m; indeed, this is more-or-less true:

Theorem 2.3 (Seneta-Hyde)**.** *If $1 < m < \infty$ then there exists a sequence $c_n > 0$ such that*

(i) $\lim_{n \to \infty} Z_n / c_n$ converges a.s. to a finite limit;

(ii) $\lim Z_n/c_n > 0$ *a.s. on the event that* Z_n *does not go extinct; and*

(iii) $c_{n+1}/c_n \to m$*.*

Exercise 2.1. *Prove that on the event of non-extinction,* $\mathrm{gr}(T) = m$ *a.s.*

Definition 2.4. *A set* $\mathscr{P}$ *of trees is called an* inherited property *if*

(i) $\mathscr{P}$ *contains all finite trees; and*

(ii) if $T \in \mathscr{P}$ *then* $T^{(u)} \in \mathscr{P}$ *for all* $u \in C(\emptyset)$*.*

Lemma 2.5. *If* $\mathscr{P}$ *is an inherited property then* $\Pr(T \in \mathscr{P}) \in \{q, 1\}$*.*

Proof.

$$\Pr(T \in \mathscr{P}) = \mathbb{E}\Pr(T \in \mathscr{P} \mid Z_1) \le \mathbb{E}\Pr(T^{(1)}, \dots, T^{(Z_1)} \in \mathscr{P} \mid Z_1).$$

Since $T^{(1)}, \dots, T^{(Z_1)}$ are independent and have the same distribution as T, the right hand side above is equal to $f(\Pr(T \in \mathscr{P}))$. That is, $a \le f(a)$ where $a = \Pr(T \in \mathscr{P})$. On the other hand, $a \ge q$ because $T \in \mathscr{P}$ on the event of extinction. By the properties of f that we saw before, $a \in \{q, 1\}$. □

Proof of Theorem 2.3. Choose some $s_0 \in (q, 1)$ and define s_n by $s_{n+1} = f^{-1}(s_n)$; then $s_n \to 1$. By Lemma 2.2, $s_n^{Z_n}$ is a martingale; since it is also bounded, it has an a.s. limit $Y \in [0, 1]$. Moreover, $\mathbb{E}[Y] = s_0 \in (q, 1)$.

Now define $c_n = -1/\log(s_n)$. By l'Hôpital's rule,

$$\frac{c_n}{c_{n-1}} = \frac{\log(s_{n-1})}{\log(s_n)} = \frac{\log(f(s_n))}{\log(s_n)} \to f'(1) = m$$

as $n \to \infty$ and $s_n \to 1$. This proves (iii).

To prove (i), note that $Y = \lim e^{-Z_n/c_n}$, and so Z_n/c_n a.s. has a (possibly infinite) limit. Since $Y \le 1$ and $\mathbb{E}[Y] = s_0 > q$, it follows that $\Pr(Y > 0) = \Pr(\lim Z_n/c_n < \infty) > q$. Since the property $\{\lim |T_n|/c_n < \infty\}$ is inherited, it follows that Z_n/c_n a.s. has a finite limit; this proves (i).

Similarly, the property $\{|T_n|/c_n \to 0\}$ is also inherited and so

$$\Pr(Y = 1) = \Pr(Z_n/c_n \to 0) \in \{q, 1\}.$$

Since $\mathbb{E}[Y] < 1$, we must have $\Pr(Y = 1) = q$; since $Y = 1$ on the event of extinction, we must have $Y < 1$ a.s. on the event of non-extinction. □

2.2 The branching number of a Galton-Watson tree

Theorem 2.6. *If T is a Galton-Watson tree with mean number of offspring $m \in (1, \infty)$ then* $\mathrm{br}(T) = m$ *a.s. on the event of non-extinction.*

Proof. By Exercise 2.1, $\mathrm{br}(T) \le \mathrm{gr}(T) = m$. For the other direction, first note that the property $\{\mathrm{br}(T) \le \lambda\}$ is inherited. Hence, $\mathrm{br}(T)$ is almost surely equal to a constant on the event of non-extinction; let λ^* be this constant. Now take $\lambda > \lambda^*$ and consider the tree T' that we obtain by the following procedure: for every $u \in T$ independently, delete u and all its descendents with probability $1 - 1/\lambda$. Then T' is a Galton-Watson tree with mean number of offspring m/λ (this is left as an exercise).

Since $\lambda > \mathrm{br}(T)$ a.s., Theorem 1.11 implies that, for every $\epsilon > 0$ there exists a cutset Π_ϵ with

$$\sum_{u \in \Pi_\epsilon} \lambda^{-|u|} \le \lambda \epsilon.$$

For $u \in \Pi_\epsilon$, the event that u belongs to T' is the event that neither u nor any of its ancestors were deleted from T; conditioned on T, this has probability $\lambda^{-|u|-1}$. Hence

$$\Pr(\Pi_\epsilon \cap T' \neq \emptyset \mid T) \le \sum_{u \in \Pi_\epsilon} \Pr(u \in T' \mid T) = \sum_{u \in \Pi_\epsilon} \lambda^{-|u|-1} \le \epsilon. \qquad (4)$$

Now consider the event that $\Pi_\epsilon \cap T'$ is empty: letting $n = \max_{u \in \Pi_\epsilon} |u|$, we see that every $u \in T$ with $|u| > n$ has some ancestor in Π_ϵ; hence u also has some ancestor that was deleted in the construction of T'. In particular, T' is finite on the event that $\Pi_\epsilon \cap T' = \emptyset$. By taking $\epsilon \to 0$ in (4), we have $\Pr(|T'| < \infty) = 1$. By Theorem 2.1 applied to T', we must have $m/\lambda \le 1$ and hence $\lambda \ge m$. Since $\lambda > \lambda^*$ was arbitrary, it follows that $\lambda^* \ge m$. □

Exercise 2.2. *Prove that the random tree T' given in the proof of Theorem 2.6 is a Galton-Watson tree. Write down its offspring distribution.*

Exercise 2.3. *Let $p_k = (1-p)^k p$ for $k \ge 0$ and let T be a Galton-Watson tree with this offspring distribution. Show that conditioned on $|T| = k$, T is uniformly distributed on the set of all trees of size k.*

3 The broadcast process

Fix a tree T; a *configuration* on T is a function $\sigma : T \to \{-1, 1\}$. Given a parameter $\lambda \in [-1, 1]$, consider the following way of producing a random configuration: first, choose $\sigma(\emptyset)$ uniformly at random. Then, recursively and independently for every $u \in T$ and every $v \in C(u)$, let $\sigma(v) = \sigma(u)$ with probability $\frac{1+\lambda}{2}$ and $\sigma(v) = -\sigma(u)$ otherwise. Equivalently, let $\{\xi(u) : u \in T\}$ be independent random variables in $\{-1, 1\}$ where $\xi(\emptyset)$ is uniformly random and $\Pr(\xi(u) = 1) = \frac{1+\lambda}{2}$; define $\sigma(u) = \prod_v \xi(v)$ where the product ranges over u and all of its ancestors.

One motivation for this process is as a model of evolution with mutation: suppose you have a population with two types of individuals that reproduce asexually. Usually, an individual has offspring of its own type, but sometimes a random mutation occurs and the child will be of the opposite type. Given such a model, it is natural to ask the following "reconstruction" problem: if we observe a family tree descended from a single individual, but we only observe the types of the descendents that are currently alive, then can we say anything about the type of the original ancestor?

To make this question precise, we recall the definition of total variation distance: for probability measures $\mathbb{P}$ and $\mathbb{Q}$,

$$d_{\mathrm{TV}}(\mathbb{P}, \mathbb{Q}) = \sup_A |\mathbb{P}(A) - \mathbb{Q}(A)|,$$

where the supremum ranges over all measurable sets. Now, let $\mathbb{P}_n^+$ denote the distribution of $\sigma(T_n)$ given $\sigma(\emptyset) = 1$ and let $\mathbb{P}_n^-$ denote the distribution of $\sigma(T_n)$ given $\sigma(\emptyset) = -1$. Suppose that we have a procedure for guessing $\sigma(\emptyset)$ after seeing $\sigma(T_n)$; let $A_n \subset \{-1, 1\}^{T_n}$ be the set of inputs for which this procedure will guess that $\sigma(T_n) = 1$. The success probability of this procedure is then

$$\begin{aligned}
&\Pr(\sigma(T_n) \in A_n \text{ and } \sigma(\emptyset) = 1) + \Pr(\sigma(T_n) \notin A_n \text{ and } \sigma(\emptyset) = -1) \\
&= \frac{1}{2}\left(\mathbb{P}_n^+(\sigma(T_n) \in A_n) + 1 - \mathbb{P}_n^-(\sigma(T_n) \in A_n)\right) \\
&\le \frac{1}{2} + \frac{1}{2} d_{\mathrm{TV}}(\mathbb{P}_n^+, \mathbb{P}_n^-).
\end{aligned}$$

On the other hand, if we choose $A_n = \{\tau \in \{-1, 1\}^{T_n} : \mathbb{P}_n^+(\tau) > \mathbb{P}_n^-(\tau)\}$ then a similar calculation shows that the success probability of this procedure is exactly $\frac{1}{2} + \frac{1}{2} d_{\mathrm{TV}}(\mathbb{P}_n^+, \mathbb{P}_n^-)$. In other words, $d_{\mathrm{TV}}(\mathbb{P}_n^+, \mathbb{P}_n^-)$ quantifies how accurately one can guess $\sigma(\emptyset)$ after observing $\sigma(T_n)$. In particular, the following theorem tells us when we can learn something non-negligible about $\sigma(\emptyset)$ given $\sigma(T_n)$ for large n.

Theorem 3.1. *If $\lambda^2 \mathrm{br}(T) > 1$ then*

$$\liminf_{n \to \infty} d_{\mathrm{TV}}(\mathbb{P}_n^+, \mathbb{P}_n^-) > 0.$$

If $\lambda^2 \mathrm{br}(T) < 1$ *then* $d_{\mathrm{TV}}(\mathbb{P}_n^+, \mathbb{P}_n^-) \to 0$.

3.1 A reconstruction algorithm

Take a tree T, some $\lambda \in [-1, 1]$ and a flow θ on T with $|\theta| = 1$. For $n \geq 0$ define

$$S_n = \sum_{u \in T_n} \lambda^{-n} \theta(u) \sigma(u).$$

Lemma 3.2. *For the* S_n *defined above,* $\mathbb{E}[S_n \mid \sigma(\emptyset)] = \sigma(\emptyset)$ *and*

$$\mathbb{E}[S_n^2] = \mathbb{E}[S_n^2 \mid \sigma(\emptyset)] = \lambda^2 + (1 - \lambda^2) \sum_{|u| \leq n} \lambda^{-2|u|} \theta^2(u).$$

Before proving Lemma 3.2, let us see how it can be used to prove the first claim of Theorem 3.1. that $\mathrm{br}(T) > \lambda^{-2}$, there exists some flow θ with $|\theta| = 1$ and

$$\lambda^2 + (1 - \lambda^2) \sum_{u \in T} \lambda^{-2|u|} \theta^2(u) = \lambda^2 + (1 - \lambda^2) \mathscr{E}_{\lambda^{-2}}(\theta) \leq K < \infty.$$

Applying Lemma 3.2 with this θ, we have $\mathbb{E}[S_n^2] \leq K$ for any n.

Now, let ν^+ be the distribution of S_n conditioned on $\sigma(\emptyset) = 1$ and let ν^- be the distribution of S_n conditioned on $\sigma(\emptyset) = -1$. We recall another characterization of total variation distance:

$$d_{\mathrm{TV}}(\nu^+, \nu^-) = \inf_{S^+, S^-} \Pr(S^+ \neq S^-),$$

where the infimum runs over all couplings (S^+, S^-) of ν^+ and ν^-. Now, Lemma 3.2 implies that for any such coupling, $\mathbb{E}[S^+ - S^-] = 2$ and

$$\mathbb{E}[(S^+ - S^-)^2] \leq 2\mathbb{E}[(S^+)^2] + 2\mathbb{E}[(S^-)^2] \leq 4K.$$

Now, the Cauchy-Schwarz inequality gives

$$2 = \mathbb{E}[(S^+ - S^-) 1_{\{S^+ \neq S^-\}}] \leq \sqrt{\mathbb{E}[(S^+ - S^-)^2] \Pr(S^+ \neq S^-)},$$

and we conclude that $\Pr(S^+ \neq S^-) \geq K^{-1}$. Hence $d_{\mathrm{TV}}(\nu^+, \nu^-) \geq K^{-1}$; since ν^+ and ν^- are push-forwards of $\mathbb{P}_n^+$ and $\mathbb{P}_n^-$ respectively, we conclude that

$$d_{\mathrm{TV}}(\mathbb{P}_n^+, \mathbb{P}_n^-) \geq \frac{1}{K},$$

which is bounded away from zero as $n \to \infty$. This proves the first part of Theorem 3.1.

Proof of Lemma 3.2. If u is the parent of v then $\mathbb{E}[\sigma(v) \mid \sigma(u)] = \lambda\sigma(u)$. By induction,

$$\mathbb{E}[\sigma(u) \mid \sigma(\emptyset)] = \lambda^{|u|}\sigma(\emptyset).$$

Summing over $u \in T_n$, $\mathbb{E}[S_n \mid \sigma(\emptyset)] = \sigma(\emptyset)\sum_{u\in T_n}\theta(u) = \sigma(\emptyset)$, as claimed.

In order to compute the second moment, we introduce the notation $u \wedge v$ to be the most recent common ancestor of u and v. By the representation of the broadcast process as a product of independent variables, we have $\mathbb{E}[\sigma(u)\sigma(v)] = \lambda^{|u|+|v|-2|u\wedge v|}$. Expanding the square,

$$\mathbb{E}[S_n^2] = \sum_{u,v\in T_n} \lambda^{-2|u\wedge v|}\theta(u)\theta(v)$$

Let $A(u)$ denote the set consisting of u and all of its ancestors; note that $\lambda^{-2|u|} = \lambda^2 + (1-\lambda^2)\sum_{v\in A(u)}\lambda^{-2|v|}$. Since $\sum_{u,v\in T_n}\theta(u)\theta(v) = |\theta|^2 = 1$,

$$\begin{aligned}\mathbb{E}[S_n^2] &= \sum_{u,v\in T_n} \lambda^{-2|u\wedge w|}\theta(u)\theta(v)\\ &= \lambda^2 + (1-\lambda^2)\sum_{u,v\in T_n}\sum_{w\in A(u\wedge v)}\lambda^{-2|w|}\theta(u)\theta(v).\end{aligned}$$

Changing the order of summation in the inner sum,

$$\begin{aligned}\sum_{u,v\in T_n}\sum_{w\in A(u\wedge v)}\lambda^{-2|w|}\theta(u)\theta(v) &= \sum_{|w|\le n}\lambda^{-2|w|}\sum_{u,v\in T_n}1_{\{w\in A(u\wedge v)\}}\theta(u)\theta(v)\\ &= \sum_{|w|\le n}\lambda^{-2|w|}\sum_{u,v\in T_n}1_{\{w\in A(u)\}}1_{\{w\in A(v)\}}\theta(u)\theta(v)\\ &= \sum_{|w|\le n}\lambda^{-2|w|}\left(\sum_{u\in T_n}1_{\{w\in A(u)\}}\theta(u)\right)^2\\ &= \sum_{|w|\le n}\lambda^{-2|w|}\theta^2(w)\end{aligned}$$

□

3.2 Non-reconstruction

For now, suppose that T is a finite tree, and fix some set $L \subset T$ of "leaves." For a labelling $\sigma : T \to \{-1, 1\}$, write $\sigma(L)$ for $\{\sigma(u) : u \in L\}$. Let $\mathbb{P}_L^+$ denote the distribution of $\sigma(L)$ given that $\sigma(\emptyset) = 1$ and let $\mathbb{P}_L^-$ denote the distribution of $\sigma(L)$ given that $\sigma(\emptyset) = -1$. Let $\mathbb{P}_L = \frac{1}{2}(\mathbb{P}_L^+ + \mathbb{P}_L^-)$ be the marginal distribution of $\sigma(L)$. Note that if $L = T_n$ then $\mathbb{P}_{T_n}^+$ is the same as the distribution $\mathbb{P}_n^+$ that we defined before. Define the *magnetization* by

$$X_{T,L} = \mathbb{E}[\sigma(\emptyset) \mid \sigma(L)].$$

Lemma 3.3.

1. $\frac{\mathbb{P}_L^+}{\mathbb{P}_L} = 1 + X_{T,L}$

2. $\frac{\mathbb{P}_L^-}{\mathbb{P}_L} = 1 - X_{T,L}$

3. $\mathbb{E}^+[X_{T,L}] = -\mathbb{E}^-[X_{T,L}] = \mathbb{E}^+[X_{T,L}^2] = \mathbb{E}^-[X_{T,L}^2]$.

Lemma 3.3 allow us to see the link between magnetization and total variation distance: by the first two parts of the lemma,

$$d_{\mathrm{TV}}(\mathbb{P}_n^+, \mathbb{P}_n^-) = \frac{1}{2}\mathbb{E}\left|\frac{\mathbb{P}_n^+}{\mathbb{P}_n} - \frac{\mathbb{P}_n^-}{\mathbb{P}_n}\right| = \mathbb{E}|X_{T,T_n}|.$$

In particular, $d_{\mathrm{TV}}(\mathbb{P}_n^+, \mathbb{P}_n^-) \to 0$ if and only if $X_{T,T_n} \to 0$ in probability. By the third part of the lemma, this is equivalent to $\mathbb{E}^+[X_{T,T_n}] \to 0$.

Proof of Lemma 3.3. By Bayes' formula, for any $\tau : L \to \{-1, 1\}$

$$\frac{\mathbb{P}_L^+(\tau)}{\mathbb{P}_L(\tau)} = \frac{\mathbb{P}(\sigma(\emptyset) = 1 \mid \sigma(L) = \tau)}{\mathbb{P}(\sigma(\emptyset) = 1)} = 2\mathbb{P}(\sigma(\emptyset) = 1 \mid \sigma(L) = \tau).$$

On the other hand,

$$X_{T,L} = \mathbb{P}(\sigma(\emptyset) = 1 \mid \sigma(L)) - \mathbb{P}(\sigma(\emptyset) = -1 \mid \sigma(L)) = 2\mathbb{P}(\sigma(\emptyset) = 1 \mid \sigma(L)) - 1.$$

This proves the first claim, and the proof of the second is analogous.

For the third claim, note that the distribution of $X_{T,L}$ under $\mathbb{P}^+$ is the same as the distribution of $-X_{T,L}$ under $\mathbb{P}^-$; the first and third equalities follow, and it also follows that $\mathbb{E}[X_{T,L}^2] = \mathbb{E}^+[X_{T,L}^2]$. Therefore, it suffices to show that $\mathbb{E}^+[X_{T,L}] = \mathbb{E}[X_{T,L}^2]$. By the first claim,

$$\mathbb{E}^+[X_{T,L}] = \mathbb{E}\Big[X_{T,L}\frac{\mathbb{P}_L^+}{\mathbb{P}_L}\Big] = \mathbb{E}[X_{T,L}(1 + X_{T,L})] = \mathbb{E}[X_{T,L}] + \mathbb{E}_{T,L}[X_{T,L}^2].$$

Finally, note that $\mathbb{E}_{T,L}[X_{T,L}] = \frac{1}{2}(\mathbb{E}_{T,L}^+[X_{T,L}] + \mathbb{E}_{T,L}^-[X_{T,L}]) = 0$. □

The main step to complete the proof of Theorem 3.1 is to study how $X_{T,L}$ changes as we change T. Therefore, let T_1 and T_2 be two trees, and let L_1 and L_2 be subsets of T_1 and T_2 respectively. Let T_2' be T_2 with a new root added as a parent of the original root, and let T be the tree where we "merge" T_1 and T_2', identifying their roots (see Figure 1). Take $L = L_1 \cup L_2$. Our goal is to study the relationship between X_{T_1,L_1}, X_{T_2,L_2}, and $X_{T,L}$. To

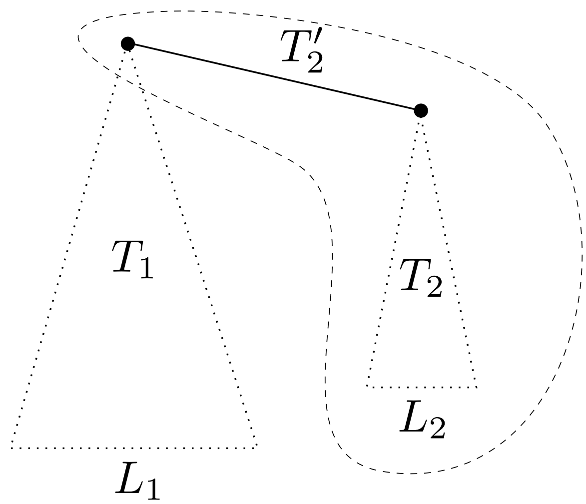


Figure 1:
The tree merging operation.

simplify the notation, we will call these X_1, X_2, and X respectively. We write $\emptyset$ for the root of T (which is also the root of T_1 and the root of T_2' and $\emptyset_2$ for the root of T_2. Let $\mathbb{P}_2^+$ be the distribution of $\sigma(T_2)$ conditioned on $\emptyset_2 = +$, and define $X_2' = X_{T_2', L_2'}$.

Lemma 3.4. *$X_2' = \lambda X_2$. Moreover, $\mathbb{E}^+[X_2'] = \lambda^2 \mathbb{E}_2^+[X_2]$ and $\mathbb{E}^+[X_2'^2] = \lambda^2 \mathbb{E}_2^+[X_2]$.*

Proof. Recall from Lemma 3.3 that

$$1 + X_2' = \frac{\mathbb{P}^+(\sigma(L_2))}{\mathbb{P}(\sigma(L_2))}$$

Next, condition on $\sigma(\emptyset_2)$ to obtain

$$1 + X_2' = \frac{\mathbb{P}^+(\sigma(\emptyset_2) = 1)\mathbb{P}_2^+(\sigma(L_2)) + \mathbb{P}^+(\sigma(\emptyset_2) = -1)\mathbb{P}_2^-(\sigma(L_2))}{\mathbb{P}(\sigma(L_2))}$$

Since $\mathbb{P}^+(\sigma(\emptyset_2) = \tau) = (1 + \tau\lambda)/2$, this can be rearranged to give

$$
\begin{aligned}
1 + X_2' &= \frac{(1+\lambda)\mathbb{P}_2^+(\sigma(L_2)) + (1-\lambda)\mathbb{P}_2^-(\sigma(L_2))}{2\mathbb{P}(\sigma(L_2))} \\
&= 1 + \lambda \frac{\mathbb{P}_2^+(\sigma(L_2)) - \mathbb{P}_2^-(\sigma(L_2))}{2\mathbb{P}(\sigma(L_2))} \\
&= 1 + \lambda X_2.
\end{aligned}
$$

This proves the first claim. For the second, condition on $\sigma(\emptyset_2)$ to obtain

$$
\mathbb{E}^+[X_2] = \frac{1+\lambda}{2}\mathbb{E}_2^+[X_2] + \frac{1-\lambda}{2}\mathbb{E}_2^-[X_2] = \lambda \mathbb{E}_2^+[X_2],
$$

where the second equality comes from the fact (Lemma 3.3, part 3) that $\mathbb{E}_2^+[X_2] = -\mathbb{E}_2^-[X_2]$. Applying our first claim gives $\mathbb{E}^+[X_2'] = \lambda\mathbb{E}^+[X_2] = \lambda^2\mathbb{E}_2^+[X_2]$. The final claim is very similar:

$$
\mathbb{E}^+[X_2^2] = \frac{1+\lambda}{2}\mathbb{E}_2^+[X_2^2] + \frac{1-\lambda}{2}\mathbb{E}_2^-[X_2^2] = \mathbb{E}_2^+[X_2^2] = \mathbb{E}_2^+[X_2],
$$

where both the second and third inequalities follow from Lemma 3.3. Applying the first claim gives $\mathbb{E}^+[X_2'^2] = \lambda^2\mathbb{E}^+[X_2] = \lambda^2\mathbb{E}_2^+[X_2]$. □

Lemma 3.5. $X = \dfrac{X_1 + X_2'}{1 + X_2 X_2'}$.

Proof. By Lemma 3.3 and the fact that $\sigma(L_1)$ and $\sigma(L_2)$ are independent given $\sigma(\emptyset)$,

$$
1+X = \frac{\mathbb{P}^+(\sigma(L))}{\mathbb{P}(\sigma(L))} = \frac{\mathbb{P}^+(\sigma(L_1))\mathbb{P}^+(\sigma(L_2))}{\mathbb{P}(\sigma(L))} = \frac{\mathbb{P}(\sigma(L_1))\mathbb{P}(\sigma(L_2))}{\mathbb{P}(\sigma(L))}(1+X_1)(1+X_2').
$$

On the other hand, since $\sigma(L_1)$ and $\sigma(L_2)$ are independent once we condition on $\sigma(\emptyset)$,

$$
\begin{aligned}
\frac{\mathbb{P}(\sigma(L))}{\mathbb{P}(\sigma(L_1))\mathbb{P}(\sigma(L_2))} &= \frac{\mathbb{P}^+(\sigma(L)) + \mathbb{P}^-(\sigma(L))}{2\mathbb{P}(\sigma(L_1))\mathbb{P}(\sigma(L_2))} \\
&= \frac{\mathbb{P}^+(\sigma(L_1))\mathbb{P}^+(\sigma(L_2)) + \mathbb{P}^-(\sigma(L_1))\mathbb{P}^-(\sigma(L_2))}{2\mathbb{P}(\sigma(L_1))\mathbb{P}(\sigma(L_2))} \\
&= \frac{(1+X_1)(1+X_2') + (1-X_1)(1-X_2')}{2} \\
&= 1 + X_1 X_2'.
\end{aligned}
$$

Combining these two calculations and rearranging proves the claim. □

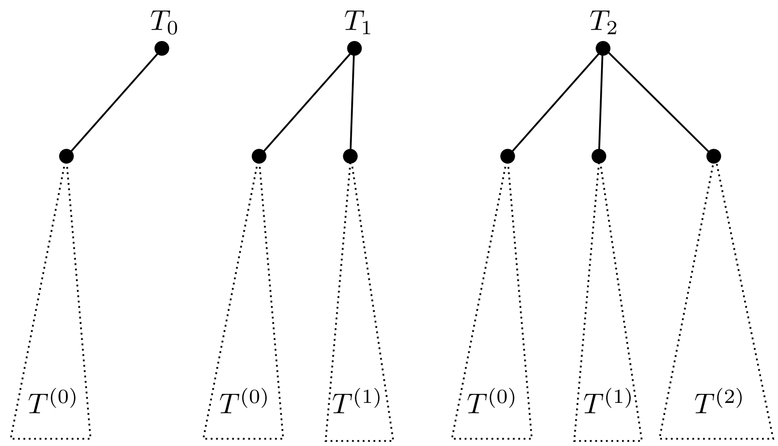


Figure 2:
Building a tree by adding its children one-by-one.

By combining Lemma 3.5 with the formula $\frac{r}{1+r} = 1 - r + \frac{r^2}{1+r}$ and the fact that $|X| \leq 1$,

$$\begin{aligned} X &= X_1 + X_2' - (X_1 + X_2')X_1X_2' + (X_1X_2')^2 X \\ &\leq X_1 + X_2' - (X_1 + X_2')X_1X_2' + (X_1X_2')^2. \end{aligned} \tag{5}$$

Define $x = \mathbb{E}^+[X]$, $x_1 = \mathbb{E}^+[X_1]$, and $x_2 = \mathbb{E}_2^+[X_2]$. Recall from Lemma 3.3 that $\mathbb{E}^+[X_1^2] = \mathbb{E}^+[X_1^2] = x_1$ and recall from Lemma 3.4 that $\mathbb{E}^+[X_2'] = \mathbb{E}^+[X_2'^2] = \lambda^2 x_2$. Taking the expectation of (5) with respect to $\mathbb{P}^+$ (and recalling that X_1 and X_2' are independent under $\mathbb{P}^+$),

$$x \leq x_1 + \lambda^2 x_2 - \lambda^2 x_1 x_2 \leq x_1 + \lambda^2 x_2. \tag{6}$$

From this, we can obtain a recursion for the expected magnetization of the root in terms of the magnetizations of its children: given a tree T and some $L \subset T$, suppose that $0, \ldots, k-1$ are the children of the root. Let $T_{\leq i}$ be the tree obtained from T by deleting $T^{(i+1)}, \ldots, T^{(k-1)}$ (see Figure 2). Let $L_{\leq i} = T_{\leq i} \cap L$ and $L^{(i)} = T^{(i)} \cap L$, and let $\mathbb{P}^{u,+}$ be the distribution of σ given that $\sigma(u) = 1$.

Now apply Lemma 3.4 with $T_2 = T^{(0)}$ and $T_2' = T_{\leq 0}$: this gives

$$\mathbb{E}^+[X_{T_{\leq 0}, L_{\leq 0}}] = \lambda^2 \mathbb{E}^{0,+}[X_{T^{(0)}, L^{(0)}}].$$

Then apply (6) with $T_1 = T_{\leq 0}$ and $T_2 = T^{(1)}$: this gives

$$\mathbb{E}^+[X_{T_{\leq 1}, L_{\leq 1}}] \leq \lambda^2 \mathbb{E}^{0,+}[X_{T^{(0)}, L^{(0)}}] + \lambda^2 \mathbb{E}^{1,+}[X_{T^{(1)}, L^{(1)}}].$$

Repeating the last step, we obtain

$$\mathbb{E}^{+}[X_{T,L}] = \mathbb{E}^{+}[X_{T_{\leq k-1},L_{\leq k-1}}] \leq \lambda^2 \sum_{i=0}^{k-1} \mathbb{E}^{i,+}[X_{T^{(i)},L^{(i)}}]. \tag{7}$$

Now we are ready to complete the proof of Theorem 3.1. Assuming that $\mathrm{br}(T) < \lambda^{-2}$, for every $\epsilon > 0$ there exists a minimal cutset Π such that $\sum_{u\in\Pi} \lambda^{2|u|} \leq \epsilon$. Then take n large enough so that $|u| < n$ for all $u \in \Pi$. Taking $L = T_n$, $L^{(u)} = T_n \cap T^{(u)}$, and applying (7) recursively on T, we have

$$\mathbb{E}^{+}[X_{T,T_n}] \leq \sum_{u\in\Pi} \lambda^{2|u|}\mathbb{E}^{u,+}[X_{T^{(u)},L^{(u)}}] \leq \sum_{u\in\Pi} \lambda^{2|u|} \leq \epsilon.$$

Since $\epsilon > 0$ was arbitrary, we conclude that $\mathbb{E}^{+}[X_{T,T_n} \to 0]$ as $n \to \infty$. This completes the proof of Theorem 3.1.

Exercise 3.1. *Theorem 3.1 doesn't make any claims about what happens when $\lambda^2\mathrm{br}(T) = 1$. It turns out that both behaviors are possible.*

(a) Show that for the k-ary tree, if $\lambda^2\mathrm{br}(T) = 1$ then $d_{\mathrm{TV}}(\mathbb{P}_n^+, \mathbb{P}_n^-) \to 0$. (Hint: imitate the last part of the proof of Theorem 3.1, but use the left-hand inequality of (6)*.)*

(b) Construct a tree T such that $\liminf d_{\mathrm{TV}}(\mathbb{P}_n^+, \mathbb{P}_n^-) > 0$ *when* $\lambda^2\mathrm{br}(T) = 1$. *(Hint: Exercise 1.3.)*

References

[1] C. Borgs, J. Chayes, E. Mossel, and S. Roch. The Kesten-Stigum reconstruction bound is tight for roughly symmetric binary channels. In *47th Annual IEEE Symposium on Foundations of Computer Science*, pages 518–530, 2006.

[2] W. Evans, C. Kenyon, Y. Peres, and L. J. Schulman. Broadcasting on trees and the Ising model. *The Annals of Applied Probability*, 10(2):410–433, 05 2000.

[3] R. Lyons and Y. Peres. *Probability on Trees and Networks*. Cambridge University Press, 2016. Available at `http://pages.iu.edu/~rdlyons/`.