

A note on the probability of cutting a Galton-Watson tree

Luc Devroye

School of Computer Science

McGill University

3450 University Street

Montreal H3A 2K6

Canada

luc@cs.mcgill.ca

Abstract.

The structure of Galton-Watson trees conditioned to be of a given size is well-understood. We provide yet another embedding theorem that permits us to obtain asymptotic probabilities of events that are determined by what happens near the root of these trees. As an example, we derive the probability that a Galton-Watson tree is cut when each node is independently removed with probability p , where by cutting a tree we mean that every path from root to leaf must have at least one removed node.

Key words: Random Galton-Watson tree, probabilistic analysis of algorithms, branching process.

AMS 2000 Classification: 60J80, 60J85, 60G99.

Submitted to EJP on November 14, 2010, final version accepted July 25, 2011.

1. The probabilistic model

A set of nodes of a tree is a cut if its removal hits every path from root to leaf. Given a tree t , and a fixed parameter $p \in (0, 1)$, we remove each node independently with probability p . Let $\rho(t)$ denote the probability that this removal procedure cuts the tree. The main purpose of this paper is to study $\rho(t)$ for a large class of trees, including random Galton-Watson trees conditioned on their size. A number of comparisons are carried out, including for complete d -ary trees, uniform random binary trees (Catalan trees), random Cayley trees, and random binary search trees, and we give the appropriate limits for conditional Galton-Watson trees, which are also called simply generated trees in the combinatorial literature—see, e.g., Moon (1970).

The paper is motivated by various other pieces of work, notably the destruction of terrorist cells (Farley, 2003, 2007), and the failure of a broadcast in tree-shaped networks. The phrase “cutting trees” has also been used in another model, which was studied, e.g., by Janson (2004) for complete binary trees, by Janson (2006a) for random trees including Galton-Watson trees, and by Holmgren (2010) for random split trees. Here nodes are removed sequentially, independently and uniformly from the remaining tree. After a node is removed, its entire subtree is disconnected. The parameter of interest in their model is the number of nodes that have to be removed until the tree is entirely gone.

It is clear that if t_1, \dots, t_k are the subtrees rooted at the k children of the root of t , then

$$\rho(t) = \begin{cases} p + (1-p) \prod_{i=1}^k \rho(t_i) & \text{if } k \geq 1, \\ p & \text{if } k = 0. \end{cases} \quad (1)$$

This recursive property can be used to compute $\rho(t)$ for many models of trees. Clearly,

$$p \leq p + (1-p)p^{|t|-1} \mathbf{1}_{[|t|>1]} \leq \rho(t) \leq 1 - (1-p)^{|t|} \leq 1,$$

where the lower bound is reached for a star, i.e., a root with $|t| - 1$ children, and the upper bound is reached for a chain. If nodes have at most $d > 1$ children, and $|t| - 1$ is a multiple of d , then the tree that minimizes $\rho(t)$ over all trees of the same size is the fishbone (Campos et al, 2011), where a *fishbone* is a tree in which the root has d children, of which $d - 1$ are leaves. The sole non-leaf child again has d children, of which $d - 1$ are leaves, and so forth. A simple computation shows that if we define $\rho_n = \rho(t)$ for the fishbone tree on n nodes (where $n - 1$ is a multiple of d), then

$$\rho_n = p + (1-p)p^{d-1}\rho_{n-d}, n > d; \rho(t_1) = p,$$

and the solution of this recursion is

$$\rho_n = \frac{p}{1 - (1-p)p^{d-1}} \times \left(1 - \left((1-p)p^{d-1} \right)^{\frac{n-1}{d}} \right).$$

Note that if we keep p fixed and let $n \rightarrow \infty$, then

$$\lim_{n \rightarrow \infty} \rho_n \stackrel{\text{def}}{=} \rho_\infty = \frac{p}{1 - (1-p)p^{d-1}}.$$

To compare performances of other trees with the optimal fishbones, it is instructive to compare limits as tree sizes diverge.

EXAMPLE 1: COMPLETE d -ARY TREES. If we define t_ℓ as the complete d -ary tree with ℓ full levels, then $|t_\ell| = n = 1 + d + \cdots + d^{\ell-1}$. A simple recursive argument yields

$$\rho(t_\ell) = p + (1-p)\rho(t_{\ell-1})^d, \rho_1 = p.$$

An elementary calculus shows that

$$\lim_{\ell \rightarrow \infty} \rho(t_\ell) = \rho^*$$

where for $p \in (0, 1 - 1/d)$, ρ^* is the unique solution in $(0, 1)$ of

$$\rho^* = p + (1-p)(\rho^*)^d.$$

For $p \in [1 - 1/d, 1]$, we have $\rho^* = 1$. For example, when $d = 2$, we obtain

$$\rho^* = \begin{cases} \frac{1 - \sqrt{1 - 4p(1-p)}}{2(1-p)} & \text{if } p \leq 1/2, \\ 1 & \text{if } p \geq 1/2. \quad \square \end{cases}$$

EXAMPLE 2: FAMILIES OF d -ARY TREES. For a tree t , let ℓ be the number of full levels, and let h be the the number of occupied levels (also called the height). We have

$$\rho(t_\ell) \leq \rho(t) \leq \rho(t_h).$$

But both upper and lower bound tend to the same limit as $\ell \rightarrow \infty$. Thus, for any sequence of d -ary trees with full level number $\ell \rightarrow \infty$, the limiting probability of cutting the trees tends to C_∞ . A typical example in this class is the random binary search tree T_n on n nodes. It is known that the (random) full level number $\ell = \ell(T_n)$ tends to infinity in probability (Devroye, 1987), and thus, we can conclude that

$$\rho(T_n) \rightarrow \rho^*$$

in probability and $\mathbb{E}\{\rho(T_n)\} \rightarrow \rho^*$ as $n \rightarrow \infty$, where \mathbb{E} denotes expected value. The same property holds for a large class of random trees, called split trees, introduced by Devroye (1999). \square

EXAMPLE 3: GALTON-WATSON TREES. A Galton-Watson (or Galton-Watson-Bienaymé) tree (see Athreya and Ney, 1972) is a rooted random ordered tree. Each node independently generates a random number of children drawn from a fixed offspring distribution ξ . Let T be a Galton-Watson tree determined by ξ , with generating function

$$g(s) \stackrel{\text{def}}{=} \mathbb{E} \left\{ s^\xi \right\} = \sum_{i=0}^{\infty} \mathbb{P}\{\xi = i\} s^i.$$

The recursive formula permits us to compute $\mathbb{E}\{\rho(T)\}$:

$$\mathbb{E}\{\rho(T)\} = p + (1-p)\mathbb{E} \left\{ (\mathbb{E}\{\rho(T)\})^\xi \mathbb{1}_{[\xi > 0]} \right\} = p + (1-p)(g(\mathbb{E}\{\rho(T)\}) - \mathbb{P}\{\xi = 0\}).$$

Assume that $p \in (0, 1)$ to avoid trivialities. If $\mathbb{P}\{\xi = 0\} > 0$, we see that $\mathbb{E}\{\rho(T)\}$ is the unique solution $r \in (0, 1)$ of

$$r = p + (1-p)(g(r) - \mathbb{P}\{\xi = 0\}). \quad (2)$$

If $\mathbb{P}\{\xi = 0\} = 0$ (so that the tree is necessarily infinite), then (2) has one solution at one, and if in addition $(1-p)g'(1) = (1-p)\mathbb{E}\{\xi\} > 1$, then there is a second solution in $(0, 1)$. It is easy to verify

then that $\mathbb{E}\{\rho(T)\} = 1$ if and only if $(1 - p)\mathbb{E}\{\xi\} \leq 1$. Otherwise, $\mathbb{E}\{\rho(T)\}$ is given by the unique solution $r \in (0, 1)$ of

$$r = p + (1 - p)g(r). \quad (3)$$

In what follows, we are mainly interested in critical Galton-Watson trees (i.e., those having $\mathbb{E}\{\xi = 1\}$, and $\mathbb{P}\{\xi = 1\} < 1$). For those, $\rho(T)$ is given by the unique solution of (2). \square

Moon (1970) and Meir and Moon (1978) defined the *simply generated trees* as ordered labelled trees of size n that are all equally likely given a certain pattern of labeling for each node of a given degree. The most important examples include the Catalan trees (equiprobable binary trees), random planted plane trees (equiprobable trees of unlimited degrees) and Cayley trees (equiprobable unordered rooted trees).

Let T_n be a Galton-Watson tree conditional on its size being n . It is well-known (see, e.g., Kennedy, 1975, or Kolchin, 1980, 1986) that most uniform random trees correspond to conditional Galton-Watson trees for particular choices of the offspring distribution ξ . For example, when ξ is 0 and 2 with probability 1/2 each, then we have a uniform full binary tree. When ξ is 0 or 2 with probability 1/4 and 1 with probability 1/2, we obtain the uniform binary (Catalan) tree. Uniformly random full binary trees are obtained by setting $\mathbb{P}\{\xi = 0\} = \mathbb{P}\{\xi = 2\} = 1/2$. A uniformly random d -ary tree has its offspring distributed as a binomial $(d, 1/d)$ random variable. A uniform planted plane tree is obtained for the geometric law $\mathbb{P}\{\xi = i\} = 1/2^{i+1}$, $i \geq 0$. When ξ is Poisson of parameter 1, one obtains a random rooted labeled (or Cayley) tree. For ξ uniform on $\{0, 1, 2, \dots, k\}$, T_n is like a uniform ordered tree with a maximal degree of k . All such trees can be dealt with at once.

Assume $\mathbb{E}\{\xi\} < \infty$. Find a $\theta > 0$ such that the random variable ξ^* with $\mathbb{P}\{\xi^* = i\} = c\theta^i\mathbb{P}\{\xi = i\}$ (where c is a normalization constant) has $\mathbb{E}\{\xi^*\} = 1$. Kennedy (1975) showed that the distribution of the conditional Galton-Watson tree T_n does not depend upon the value of θ . Thus, without loss of generality, we can normalize and assume that $\mathbb{E}\{\xi\} = 1$. Note however that the above construction does not work for all cases—indeed, there are heavy-tailed distributions with $\mathbb{E}\{\xi\} < 1$ such that there is no θ that yields $\mathbb{E}\{\xi^*\} = 1$. We use the notation $p_i = \mathbb{P}\{\xi = i\}$, and $g(s) = \mathbb{E}\{s^\xi\}$. We assume throughout that $p_1 < 1$. Note that $\{ip_i, i \geq 1\}$ is a probability distribution on the integers. Its generating function is $sg'(s)$. Let ξ' be a random variable with this distribution. Define the span d as the greatest common divisor of all i for which $p_i > 0$. A Galton-Watson tree with span d can only have sizes that are 1 mod d .

The main result of this paper is Theorem 1.

THEOREM 1. *Assume $p \in (0, 1)$, $\mathbb{E}\{\xi\} = 1$, $\mathbb{V}\{\xi\} = \sigma^2 \in (0, \infty)$. Let T_n be the Galton-Watson tree generated by ξ conditional on being of size n , where $n = 1 \pmod{d}$, and d is the span of ξ . Then, taking limits only for $n = 1 \pmod{d}$,*

$$\lim_{n \rightarrow \infty} \mathbb{E}\{\rho(T_n)\} = \frac{p}{1 - (1 - p)g'(r)},$$

where $r = \mathbb{E}\{\rho(T)\}$ is the unique solution on $(0, 1)$ of

$$r = p + (1 - p)(g(r) - p_0),$$

T is an unconditional Galton-Watson tree with offspring distribution ξ , and $g(s) = \mathbb{E}\{s^\xi\}$.

The proof, which is given in a later section, uses two steps. For an unconditional Galton-Watson tree T , we have a root level recursion, with ξ still denoting the number of children of the root:

$$\mathbb{E}\{\rho(T)\} = p + (1 - p)\mathbb{E}\left\{\mathbb{1}_{[\xi > 0]}(\mathbb{E}\{\rho(T)\})^\xi\right\} = p + (1 - p)(g(\mathbb{E}\{\rho(T)\}) - p_0).$$

Thus, indeed, $r = \mathbb{E}\{\rho(T)\}$. In the second step, one notes (see, e.g., Kolchin, 1986) that the root of T_n has a number of children that as $n \rightarrow \infty$ tends in distribution to ξ' . As n becomes large, all but one of these children are roots of unconditional Galton-Watson trees (for the original ξ) and one is the root of a conditional Galton-Watson tree whose size is $n - 1$ minus the sizes of the unconditional Galton-Watson trees. This structural view is asymptotically precise, and will be nailed down in a useful Lemma below. This suggests that if ρ_∞ is the limit of $\mathbb{E}\{\rho(T_n)\}$, then

$$\rho_\infty = p + (1 - p)\rho_\infty \mathbb{E}\{r^{\xi' - 1}\} = p + (1 - p)\rho g'(r),$$

so that $\rho_\infty = p/(1 - (1 - p)g'(r))$.

EXAMPLE 4: THE UNIFORM FULL BINARY TREE. For the uniform full binary tree, we have $g(s) = (1 + s^2)/2$, $g'(s) = s$,

$$r = \frac{1 - \sqrt{1 - 2p(1 - p)}}{1 - p},$$

and

$$\lim_{n \rightarrow \infty} \mathbb{E}\{\rho(T_n)\} = \frac{p}{\sqrt{1 - 2p(1 - p)}}.$$

This can be verified independently by analytic methods based on generating functions and singularity analysis (for a good account of singularity analysis, see Flajolet and Sedgewick, 2008). \square

EXAMPLE 5. THE CATALAN TREE. For the Catalan tree, we have $g(s) = (1+s)^2/4$, $g'(s) = (1+s)/2$,

$$r = p + (1-p)(r^2/4 + r/2).$$

This yields

$$r = \frac{1+p - \sqrt{1-2p+5p^2}}{1-p}.$$

We have

$$\lim_{n \rightarrow \infty} \mathbb{E}\{\rho(T_n)\} = \frac{2p}{\sqrt{1-2p+5p^2}}. \quad \square$$

EXAMPLE 6: THE RANDOM ROOTED CAYLEY TREE. A more advanced example is the random rooted Cayley tree, i.e., a uniform random rooted labeled tree. As pointed out above, this corresponds to T_n with ξ Poisson (1). We have $g(s) = g'(s) = \exp(s-1)$. The constant r is the unique solution on $(0, 1)$ of

$$r = p + (1-p)(e^{r-1} - e^{-1}).$$

Finally,

$$\lim_{n \rightarrow \infty} \mathbb{E}\{\rho(T_n)\} = \frac{p}{1 - (1-p)e^{r-1}} = \frac{p}{1 + p - r - \frac{1-p}{e}}. \quad \square$$

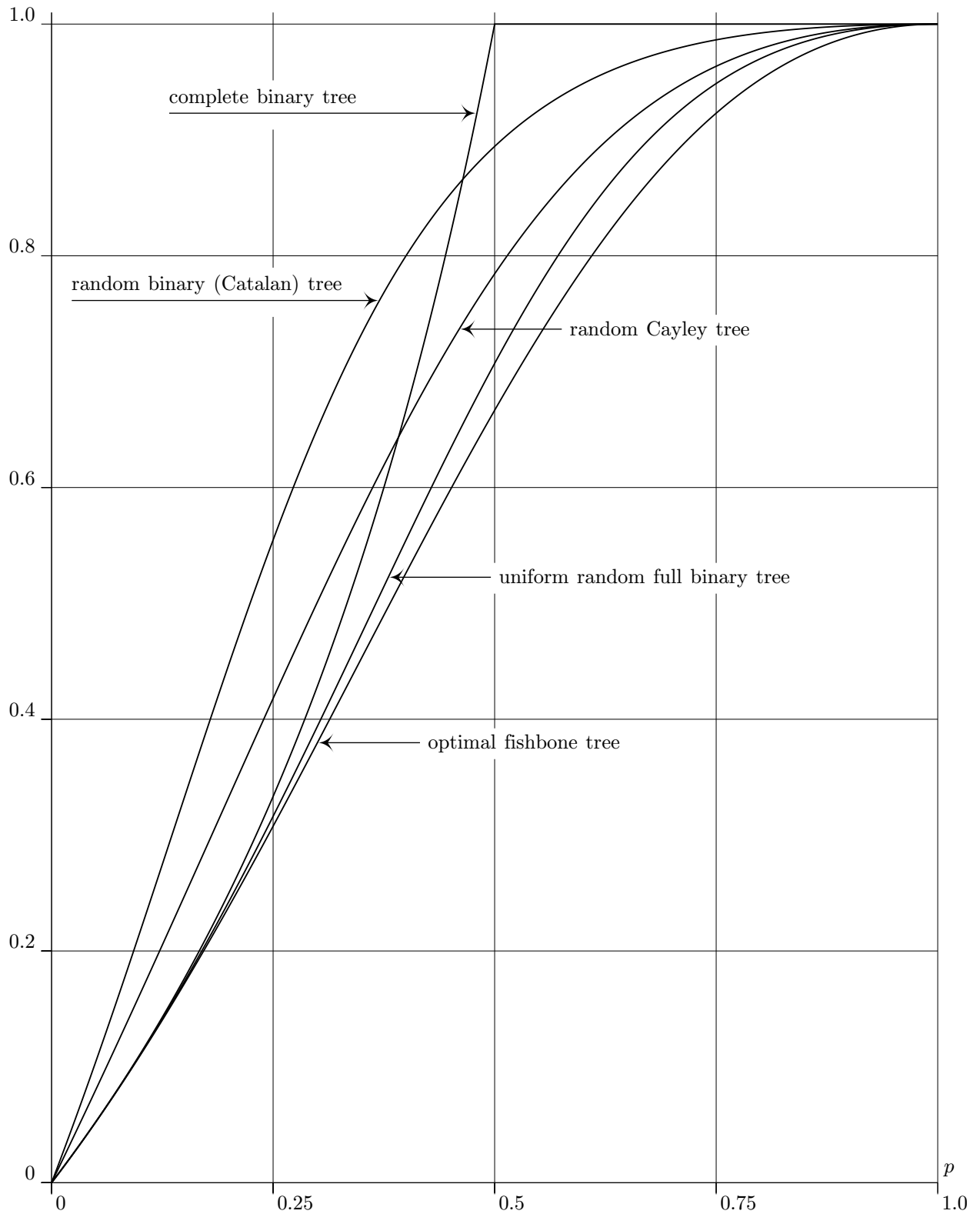


Figure 1. We show the limiting probabilities of cutting the root from the leaves for various tree models as a function of p , the probability of removing an individual node.

2. Preparing the proof: random walks

Let ξ be a random variable representing the number of children of the root in a critical Galton-Watson tree: $\mathbb{E}\{\xi\} = 1$, $\mathbb{P}\{\xi = 0\} > 0$. Set $X = \xi - 1$, and let X_1, X_2, \dots be a sequence of i.i.d. random variables distributed as X . Define the partial sums $S_0 = 1$, $S_n = 1 + \sum_{i=1}^n X_i$. There is a well-known depth first or preorder construction of a random Galton-Watson tree which lends itself well to the study of all properties of randomly selected nodes (see, e.g., Le Gall, 1989, or Aldous, 1991). Nodes in an ordered tree can be encoded with a vector of child numbers. The root corresponds to the empty vector. Its children have encodings $1, 2, \dots, \xi$. The children of the j -th child of the root are encoded by $(j, 1), (j, 2), \dots$. A preorder listing of the nodes is nothing but a lexicographic listing of the node vectors.

We can traverse a random Galton-Watson tree by visiting nodes in preorder, starting at the root. To do so, a list L of nodes to be visited is kept, which is initially of size one (having the root). When node u is visited, we consider ξ_u , the number of children of u , and remove u from L . Thus, L increases by $\xi_u - 1$. The next node in lexicographic order is taken from L , and the process continues until $L = 0$.

We denote by N the size of a random Galton-Watson tree. The size of L after n nodes have been processed is denoted by S_n . Thus, $S_0 = 1$, and

$$S_n = 1 + X_1 + \dots + X_n,$$

where $X_n = \xi_n - 1$, and indexing of the nodes is by their lexicographic rank. The root has rank one, for example. We have the identity

$$[N = n] = [S_1 > 0, S_2 > 0, \dots, S_{n-1} > 0, S_n = 0].$$

The standard circular symmetry argument for random walks (see, e.g., Dwass, 1968) shows that

$$\begin{aligned} \mathbb{P}\{N = n\} &= \mathbb{P}\{S_0 = 1, S_1 > 0, S_2 > 0, \dots, S_{n-1} > 0, S_n = 0\} \\ &= \frac{1}{n} \mathbb{P}\{S_n = 0\} \\ &= \frac{1}{n} \mathbb{P}\{X_1 + \dots + X_n = -1\}. \end{aligned}$$

The asymptotics for this probability distribution are well-known. Let $\mathbb{E}\{\xi^2\} < \infty$, $\sigma^2 = \mathbb{V}\{\xi\} > 0$ (which implies $\mathbb{P}\{\xi = 0\} > 0$), and let the span d be the greatest common divisor of all i for which $\mathbb{P}\{\xi = i\} > 0$. A Galton-Watson tree with span d can only have sizes that are $1 \pmod{d}$. From Petrov (1975, p. 197) or Kolchin (1986, p. 16), we see that when $d = 1$,

$$\lim_{n \rightarrow \infty} \sup_{k \in \mathcal{N}} \left| \sigma \sqrt{n} \mathbb{P}\{X_1 + \dots + X_n = k\} - \frac{1}{\sqrt{2\pi}} e^{-\frac{k^2}{2\sigma^2 n}} \right| = 0. \quad (0)$$

Thus, for $d = 1$,

$$\mathbb{P}\{N = n\} \sim \frac{1}{\sigma \sqrt{2\pi n^{3/2}}}.$$

In fact, if $n \rightarrow \infty$ such that $n = 1 \pmod{d}$, and $d \geq 1$, then

$$\mathbb{P}\{N = n\} \sim \frac{d}{\sigma \sqrt{2\pi n^{3/2}}}$$

(Kolchin, 1986, p. 105).

REMARK. INFINITE VARIANCE. If the variance is infinite, one can replace the estimate (0) by one that involves stable laws. This will not be pursued in the paper.

3. A structure theorem for Galton-Watson trees

Various ways have been suggested for describing the structure of a random Galton-Watson tree conditioned on its size. In particular, the idea of a spine or marked (infinite) path, or size-biasing already present in the work of Rouault (1981) has been made prominent by Lyons, Pemantle and Peres (1995), who used it to give a novel proof of the Kesten-Stigum theorem. It is useful for studying both subcritical, critical and supercritical trees. Many proofs are based on this view of marking one special node at each level to form a spine, see, e.g., Aldous and Pitman (1998), Geiger and Kaufmann (2004) and Duquesne (2009). Our intent is to give a simple lemma that is useful for proving properties that are heavily influenced by the shape of the tree near the root.

We provide a bound on the total variation distance between the top few levels of two related random trees. In this section, T is a Galton-Watson tree with typical litter size ξ , where it is assumed that $\mathbb{E}\{\xi\} = 1$, $p_1 < 1$, and $\mathbb{V}\{\xi\} = \sigma^2 < \infty$. Let T_n be T conditional on $|T| = n$.

The second tree in which the size-biasing is made explicit is denoted by T'_n . We construct T'_n by first generating ξ' children of the root according to the probability law

$$\mathbb{P}\{\xi' = i\} = ip_i, i \geq 1.$$

Note that

$$\mathbb{E}\{\xi'\} = \mathbb{E}\{\xi^2\} = 1 + \sigma^2.$$

Among the ξ' children, choose one uniformly at random, and denote its index by Z . Each child, Z excepted, is the root of an independent unconditional Galton-Watson tree drawn from the distribution of T . Denote the size of the tree of child i by N'_i . If $N \stackrel{\text{def}}{=} 1 + \sum_{i:i \neq Z} N'_i \geq n$, then remove child Z (and thus, ξ' is reduced by one). Otherwise, make child Z the root of a tree that is distributed as T_{n-N} .

One can apply the same construction recursively to child Z , and this leads to the well-known “spine” of the conditional Galton-Watson trees. However, our first result is about the total variation distance between the probability measures of the vectors of subtree sizes. In T_n , the ξ children of the root have subtree sizes $\mathcal{N} = (N_1, N_2, \dots, N_\xi)$. We pad this vector with zeroes to the right. In T'_n , we denote the subtree sizes by $\mathcal{N}' = (N'_1, \dots)$, and pad with zeroes to the right. Denote the total variation distance between the probability measures of these vectors by

$$d(\mathcal{N}, \mathcal{N}').$$

By Doeblin’s coupling lemma, we can couple the trees in such a way that with probability $1 - d(\mathcal{N}, \mathcal{N}')$, both are identical. Indeed, if the number of children matches, and the subtree sizes

match, then each subtree is by construction a Galton-Watson tree conditioned by its size. Let us use such a coupling, and let us also maximally couple the nodes marked for removal (say, by depth first order). Clearly, we have

$$\mathbb{P}\{\rho(T_n) \neq \rho(T'_n)\} \leq \mathbb{P}\{T_n \neq T'_n\} \leq d(\mathcal{N}, \mathcal{N}').$$

Thus, if $d(\mathcal{N}, \mathcal{N}') = o(1)$, then

$$\lim_{n \rightarrow \infty} (\mathbb{E}\{\rho(T_n)\} - \mathbb{E}\{\rho(T'_n)\}) = 0.$$

This shows the usefulness of the construction and Lemma 1.

LEMMA 1. Assume that $\mathbb{E}\{\xi\} = 1$, $p_1 < 1$, and $\mathbb{V}\{\xi\} = \sigma^2 < \infty$. Then

$$\lim_{n \rightarrow \infty} d(\mathcal{N}, \mathcal{N}') = 0.$$

PROOF. Let W_k be the space of all strictly positive integer vectors of length k having sum $n - 1$. Then, for all $(n_1, \dots, n_k) \in W_k$,

$$\mathbb{P}\{\mathcal{N} = (n_1, \dots, n_k, 0, 0, \dots)\} = \frac{\mathbb{P}\{\xi = k\} \prod_{j=1}^k \mathbb{P}\{|T| = n_j\}}{\mathbb{P}\{|T| = n\}}.$$

Next, we have, denoting by $T(\ell)$, $\ell \geq 1$, independent copies of T ,

$$\begin{aligned} & \mathbb{P}\{\mathcal{N}' = (n_1, \dots, n_k, 0, 0, \dots)\} \\ & \geq \mathbb{P}\{\xi' = k\} \frac{1}{k} \sum_{i=1}^k \mathbb{P}\left\{ \bigcap_{1 \leq j \leq k, j \neq i} [|T(j)| = n_j], 1 + \sum_{1 \leq j \leq k, j \neq i} |T(j)| < n - 1 \right\} \\ & = \mathbb{P}\{\xi' = k\} \frac{1}{k} \sum_{i=1}^k \prod_{1 \leq j \leq k, j \neq i} \mathbb{P}\{|T(j)| = n_j\}. \end{aligned}$$

Summing over all k and over all vectors $(n_1, n_2, \dots, n_k) \in W_k$, and using a well-known property of the total variation distance, and using the notation $u_+ = \max(u, 0)$, we have

$$\begin{aligned} & d(\mathcal{N}, \mathcal{N}') \\ & \leq \sum_{k \geq 1} \sum_{(n_1, \dots, n_k) \in W_k} \left(\frac{\mathbb{P}\{\xi = k\} \prod_{j=1}^k \mathbb{P}\{|T| = n_j\}}{\mathbb{P}\{|T| = n\}} - \mathbb{P}\{\xi = k\} \sum_{i=1}^k \prod_{1 \leq j \leq k, j \neq i} \mathbb{P}\{|T| = n_j\} \right)_+ \\ & = \sum_{k \geq 1} \sum_{(n_1, \dots, n_k) \in W_k} \mathbb{P}\{\xi = k\} \prod_{j=1}^k \mathbb{P}\{|T| = n_j\} \left(\frac{1}{\mathbb{P}\{|T| = n\}} - \sum_{i=1}^k \frac{1}{\mathbb{P}\{|T| = n_i\}} \right)_+ \\ & \leq \mathbb{P}\{\xi > K \mid |T| = n\} + \sum_{k \leq K} \sum_{(n_1, \dots, n_k) \in W_k} \mathbb{P}\{\xi = k\} \prod_{j=1}^k \mathbb{P}\{|T| = n_j\} \left(\frac{1}{\mathbb{P}\{|T| = n\}} - \sum_{i=1}^k \frac{1}{\mathbb{P}\{|T| = n_i\}} \right)_+. \end{aligned}$$

Here K is a fixed large integer, and ξ is the number of children of the root in T_n . It is understood here that there is no problem of division by zero, as zero factors will occur simultaneously in the

numerator and denominator. We will need only one term in the last sum of the upper bound. We also note that the right-hand-side is zero when $k = 1$. Incorporating this, we have

$$d(\mathcal{N}, \mathcal{N}') \leq \mathbb{P}\{\xi > K | |T| = n\} + \sum_{K \geq k \geq 2} \sum_{(n_1, \dots, n_k) \in W_k} \mathbb{P}\{\xi = k\} \prod_{j=1}^k \mathbb{P}\{|T| = n_j\} \left(\frac{1}{\mathbb{P}\{|T| = n\}} - \frac{1}{\mathbb{P}\{|T| = \max_i n_i\}} \right)_+.$$

Consider the subspace W'_k consisting of all vectors in W_k with $n_1 = \max_i n_i$. By the symmetry in our upper bound, we have

$$d(\mathcal{N}, \mathcal{N}') \leq \mathbb{P}\{\xi > K | |T| = n\} + \Delta_K,$$

where

$$\Delta_K \stackrel{\text{def}}{=} \sum_{K \geq k \geq 2} k \mathbb{P}\{\xi = k\} \sum_{(n_1, \dots, n_k) \in W'_k} \prod_{j=1}^k \mathbb{P}\{|T| = n_j\} \left(\frac{1}{\mathbb{P}\{|T| = n\}} - \frac{1}{\mathbb{P}\{|T| = n_1\}} \right)_+.$$

If n_1 is fixed, then define W''_{k-1} as the space of positive $k-1$ -vectors of sum $n-1-n_1$. Thus,

$$\Delta_K \leq \sum_{K \geq k \geq 2} k \mathbb{P}\{\xi = k\} \times \sum_{n_1: (n-1)/k \leq n_1 \leq n-k} \mathbb{P}\{|T| = n_1\} \left(\frac{1}{\mathbb{P}\{|T| = n\}} - \frac{1}{\mathbb{P}\{|T| = n_1\}} \right)_+ \times \sum_{(n_2, \dots, n_k) \in W''_{k-1}} \prod_{j=2}^k \mathbb{P}\{|T| = n_j\}.$$

The last sum is of interest. It is, in fact, nothing but the probability that the total size of a forest of $k-1$ random Galton-Watson trees is of size $n-1-n_1$. It is well-known that this is

$$\frac{k-1}{n-1-n_1} \mathbb{P}\{\xi(1) + \dots + \xi(n-1-n_1) = n-n_1-k\} = \frac{k-1}{n-1-n_1} \mathbb{P}\{S_{n-1-n_1} = n-n_1-k\},$$

where $\xi(1), \dots$ are i.i.d. and distributed as ξ (see, e.g., Kolchin, 1986, p. 104), and $S_n = \sum_{i=1}^n \xi(i)$. We now have

$$\Delta_K \leq \sum_{K \geq k \geq 2} k(k-1) \mathbb{P}\{\xi = k\} \times \sum_{n_1: (n-1)/k \leq n_1 \leq n-k} \frac{\mathbb{P}\{|T| = n_1\} \mathbb{P}\{S_{n-1-n_1} = n-n_1-k\}}{n-1-n_1} \left(\frac{1}{\mathbb{P}\{|T| = n\}} - \frac{1}{\mathbb{P}\{|T| = n_1\}} \right)_+.$$

Recalling also that

$$\mathbb{P}\{|T| = n\} = \frac{1}{n} \mathbb{P}\{S_n = n-1\},$$

we rewrite the inequality as follows:

$$\Delta_K \leq \sum_{K \geq k \geq 2} k(k-1) \mathbb{P}\{\xi = k\} \times \sum_{n_1: (n-1)/k \leq n_1 \leq n-k} \frac{\mathbb{P}\{S_{n_1} = n_1-1\} \mathbb{P}\{S_{n-1-n_1} = n-n_1-k\}}{n_1(n-1-n_1)} \left(\frac{n}{\mathbb{P}\{S_n = n-1\}} - \frac{n_1}{\mathbb{P}\{S_{n_1} = n_1-1\}} \right)_+.$$

By (0), assuming that the span of ξ is 1, we have a uniform approximation that immediately yields the following, with all the $o(\cdot)$ terms depending upon the distribution of ξ and n only, but not on k :

$$\Delta_K \leq \sum_{K \geq k \geq 2} k(k-1) \mathbb{P}\{\xi = k\} \times \sum_{n_1: (n-1)/k \leq n_1 \leq n-k} \frac{\exp\left(-\frac{(k-1)^2}{2\sigma^2(n-1-n_1)}\right) + o(1)}{n_1^{3/2}(n-1-n_1)^{3/2}\sigma\sqrt{2\pi}} \left(n^{3/2} - n_1^{3/2} + o(n^{3/2})\right)_+. \quad (4)$$

We show that the upper bound in (4) tends to zero. Note that

$$\begin{aligned} n^{3/2} - n_1^{3/2} &= n^{3/2} \left(1 - \left(1 - \frac{n-n_1}{n}\right)^{3/2}\right) \\ &\leq \frac{3}{2}\sqrt{n}(n-n_1) \quad (\text{use } (1-u)^{3/2} \geq 1-3u/2, u > 0). \end{aligned}$$

Fix k and, discarding constants, look at the sums

$$B_k \stackrel{\text{def}}{=} \sum_{n_1: (n-1)/k \leq n_1 \leq n-k} \sqrt{n}(n-n_1) \frac{\exp\left(-\frac{(k-1)^2}{2\sigma^2(n-1-n_1)}\right) + o(1)}{n_1^{3/2}(n-1-n_1)^{3/2}}$$

and

$$C_k \stackrel{\text{def}}{=} \sum_{n_1: (n-1)/k \leq n_1 \leq n-k} n^{3/2} \frac{\exp\left(-\frac{(k-1)^2}{2\sigma^2(n-1-n_1)}\right) + o(1)}{n_1^{3/2}(n-1-n_1)^{3/2}}.$$

We first show that for any fixed K ,

$$\sup_{K \geq k \geq 2} B_k = o(1), \sup_{n \geq 5} \sup_{K \geq k \geq 2} C_k < \infty. \quad (5)$$

Using (4), we then have for any fixed K ,

$$d(\mathcal{N}, \mathcal{N}') \leq \mathbb{P}\{\xi > K \mid |T| = n\} + o(1).$$

We conclude by showing that

$$\lim_{K \rightarrow \infty} \sup_{n \geq 1} \mathbb{P}\{\xi > K \mid |T| = n\} = 0. \quad (6)$$

The proof of (5) makes use of the fact that if f is a monotone decreasing function on the reals, then

$$\sum_{i=\ell}^k f(i) \leq f(\ell) + \int_{\ell}^k f(x) dx.$$

We assume throughout that n is large enough. Clearly, we have for some constants c, c' not depending upon k or n ,

$$\begin{aligned}
B_k &\leq c \sum_{n_1: (n-1)/k \leq n_1 \leq n-k} \frac{\sqrt{n}}{n_1^{3/2} \sqrt{n-1-n_1}} \\
&\leq \frac{c' k^{3/2}}{n} \sum_{n_1: (n-1)/k \leq n_1 \leq n-k} \frac{k^{3/2}}{n \sqrt{n-1-n_1}} \\
&\leq \frac{c' k^{3/2}}{n} \frac{1}{\sqrt{k-1}} + \frac{c' k^{3/2}}{n} \int_{(n-1)/k}^{n-k} \frac{1}{\sqrt{n-1-x}} dx \\
&\leq \frac{c' k^{3/2}}{n} \frac{1}{\sqrt{k-1}} + \frac{2c' k^{3/2}}{\sqrt{n}} \\
&\leq \frac{3c' K^{3/2}}{\sqrt{n}}.
\end{aligned}$$

Very similar calculations can be done for C_k :

$$\begin{aligned}
C_k &\leq cn^{3/2} \sum_{n_1: (n-1)/k \leq n_1 \leq n-k} \frac{1}{n_1^{3/2} (n-1-n_1)^{3/2}} \\
&\leq c' k^{3/2} \sum_{n_1: (n-1)/k \leq n_1 \leq n-k} \frac{1}{(n-1-n_1)^{3/2}} \\
&\leq \frac{c' k^{3/2}}{(k-1)^{3/2}} + c' k^{3/2} \int_{(n-1)/k}^{n-k} \frac{1}{(n-1-x)^{3/2}} dx \\
&\leq \frac{c' k^{3/2}}{(k-1)^{3/2}} + \frac{2c' k^{3/2}}{\sqrt{n-1-(n-1)/k}} \\
&\leq c'' K^{3/2}.
\end{aligned}$$

Thus, we only need to show (6).

Fix k . Then, using Kolchin's formula again,

$$\begin{aligned}
\mathbb{P}\{\xi = k | |T| = n\} &= \frac{p_k \mathbb{P}\{|T(1)| + \dots + |T(k)| = n-1\}}{\mathbb{P}\{|T| = n\}} \\
&= \frac{p_k \frac{k}{n-1} \mathbb{P}\{S_{n-1} = n-1-k\}}{\frac{1}{n} \mathbb{P}\{S_n = n-1\}} \\
&= p_k \frac{k \left[\exp\left(-\frac{k^2}{2\sigma^2(n-1)}\right) + o(1) \right]}{(n-1)\sqrt{n-1}} \times \frac{\sqrt{n}}{n(1+o(1))} \\
&\quad \text{(where } o(1) \text{ does not depend upon } k\text{)} \\
&= \left(\frac{n}{n-1}\right)^{3/2} k p_k \left[\exp\left(-\frac{k^2}{2\sigma^2(n-1)}\right) + o(1) \right] \\
&\leq (1+o(1)) k p_k.
\end{aligned}$$

Thus,

$$\mathbb{P}\{\xi \geq K | |T| = n\} \leq (1+o(1)) \sum_{k=K}^{\infty} k p_k,$$

from which our result follows since we assumed that $\sum kp_k = 1$. The above derivation assumes that the span of ξ is 1. It is easy to verify that the results remain valid for $d > 1$. \square

4. Proof of Theorem 1

We first deal with $\rho(T'_n)$:

$$\mathbb{E}\{\rho(T'_n)\} = p + (1-p)\mathbb{E}\{r^{\xi'-1}\rho(T'_N)\} = p + (1-p)\mathbb{E}\{r^{\xi'-1}\mathbb{E}\{\rho(T'_N)|N\}\}, \quad (7)$$

where ξ' has distribution $\{ip_i, i \geq 1\}$, and $N = \max(0, n-1 - |T(1)| - \dots - |T(\xi'-1)|)$, with the $T(i)$'s i.i.d. unconditional Galton-Watson trees. Furthermore, we agree that T'_0 is the empty tree, and that $\rho(T'_0) = 1$, to make (7) valid.

We will show that

$$\lim_{n \rightarrow \infty} \mathbb{E}\{\rho(T'_n)\} = \lambda$$

where λ is the solution of

$$\lambda = p + (1-p)\lambda\mathbb{E}\{r^{\xi'-1}\}, \quad (8)$$

i.e.,

$$\lambda = \frac{p}{1 - (1-p)\mathbb{E}\{r^{\xi'-1}\}}.$$

Before we continue, we note that g' is the generating function of $\xi' - 1$. Thus we have

$$\lambda = \frac{p}{1 - (1-p)g'(r)}.$$

To prove (8), define

$$\rho_n^+ = \sup_{\ell: \ell \geq n} \mathbb{E}\{\rho(T'_\ell)\}, \rho_n^- = \inf_{\ell: \ell \geq n} \mathbb{E}\{\rho(T'_\ell)\}.$$

From (7), we have

$$\rho_n^+ \leq p + (1-p)\mathbb{P}\{N < n/2\} + \mathbb{E}\{r^{\xi'-1}\}\rho_{\lceil n/2 \rceil}^+.$$

Take the limit on both sides. If

$$\lim_{n \rightarrow \infty} \mathbb{P}\{N < n/2\} = 0, \quad (9)$$

then we can conclude that

$$\limsup_{n \rightarrow \infty} \mathbb{E}\{\rho(T'_n)\} \leq p + (1-p)g'(r) \limsup_{n \rightarrow \infty} \mathbb{E}\{\rho(T'_n)\},$$

and thus, that

$$\limsup_{n \rightarrow \infty} \mathbb{E}\{\rho(T'_n)\} \leq \frac{p}{1 - (1-p)g'(r)}.$$

Similarly,

$$\rho_n^- \geq p - (1-p)\mathbb{P}\{N < n/2\} + \mathbb{E}\{r^{\xi'-1}\}\rho_{\lceil n/2 \rceil}^- = p - (1-p)\mathbb{P}\{N < n/2\} + g'(r)\rho_{\lceil n/2 \rceil}^-,$$

from which we conclude that

$$\liminf_{n \rightarrow \infty} \mathbb{E}\{\rho(T'_n)\} \geq \frac{p}{1 - (1-p)g'(r)}.$$

Therefore, both limits are the same, and we are done. It remains to show (9). For $\xi' > 1$, with “ $o(1)$ ” below not depending upon ξ' ,

$$\begin{aligned}
& \mathbb{P} \left\{ |T(1)| + \cdots + |T(\xi' - 1)| \geq \frac{n-2}{2} \middle| \xi' \right\} \\
&= \sum_{k: k \geq (n-2)/2} \mathbb{P}\{|T(1)| + \cdots + |T(\xi' - 1)| = k \mid \xi'\} \\
&= \sum_{k: k \geq (n-2)/2} \frac{\xi' - 1}{k} \mathbb{P}\{S_k = k - \xi' + 1 \mid \xi'\} \\
&= \sum_{k: k \geq (n-2)/2} \frac{\xi' - 1}{k} \frac{\exp\left(-\frac{(\xi' - 1)^2}{2\sigma^2 k}\right) + o(1)}{\sigma\sqrt{2\pi k}} \\
&\leq \frac{1 + o(1)}{\sigma\sqrt{2\pi}} \sum_{k: k \geq (n-2)/2} \frac{\xi' - 1}{k^{3/2}} \frac{1 + o(1)}{\sigma\sqrt{2\pi}} \\
&\leq \frac{c(\xi' - 1)}{\sqrt{n}}
\end{aligned}$$

for some constant c not depending upon ξ' or n . Thus,

$$\begin{aligned}
\mathbb{P}\{N < n/2\} &= \mathbb{P}\{n - 1 - |T(1)| - \cdots - |T(\xi' - 1)| < n/2\} \\
&= \mathbb{P}\{n/2 - 1 \leq |T(1)| - \cdots - |T(\xi' - 1)|\} \\
&= \mathbb{P}\{n/2 - 1 \leq |T(1)| + \cdots + |T(\xi' - 1)|, \xi' \geq 2\} \\
&\leq \mathbb{E} \left\{ \frac{c(\xi' - 1)}{\sqrt{n}} \right\} \\
&= \frac{c\sigma^2}{\sqrt{n}}.
\end{aligned}$$

This concludes the proof. \square

5. Other applications

Lemma 1 renders many proofs simple. For example, assuming the conditions of the Lemma, we see that if $Z_{n,1}$ is the size of the first generation in the conditional Galton-Watson tree T_n , then

$$Z_{n,1} \xrightarrow{\mathcal{L}} \xi',$$

where ξ' has the size-biased distribution $\{ip_i, i \geq 1\}$, and $\xrightarrow{\mathcal{L}}$ denotes convergence in distribution. This known fact follows immediately from the inequality on total variation distances:

$$d(Z_{n,1}, \xi') \leq d(\mathcal{N}, \mathcal{N}') = o(1).$$

All other properties of \mathcal{N} and \mathcal{N}' are shared. For example, if $M_n = \max_i N_i$ is the maximal size of a subtree of the root of T_n , then, under the conditions of Lemma 1, for every fixed $\ell \geq 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}\{M_n = n - 1 - \ell\} = \mathbb{P} \left\{ \sum_{1 \leq k < \xi'} |T(k)| = \ell \right\},$$

where $T(1), \dots$ are independent copies of T , the Galton-Watson tree for ξ , and ξ' is independent of these trees. In particular, for binary trees, we have

$$\lim_{n \rightarrow \infty} \mathbb{P}\{M_n = n - 1\} = \mathbb{P}\{\xi' = 1\} = p_1,$$

and for $\ell > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}\{M_n = n - 1 - \ell\} = (1 - p_1)\mathbb{P}\{|T| = \ell\}.$$

If in a Galton-Watson tree T , we only consider levels $0, k, 2k, 3k, \dots$, then we obtain a new Galton-Watson tree T^* . If $|T| = n$, then $|T^*|$ is random. Under reasonable conditions, we expect $|T^*| \approx n/k$ if k is fixed. In any case, $|T^*| \geq H_n/k$, where H_n is the height of T_n . Consider the population size $Z_{n,k}$ of the k -th generation in T_n , which is the size of the first generation $Z_{|T^*|,1}^*$ in the corresponding T^* . Thus,

$$d(Z_{n,k}, \xi'_k) \leq \mathbb{P}\{|T^*| \leq M\} + \sup_{m \geq M} d(Z_{m,1}^*, \xi'_k),$$

where ξ'_k is the size-biased distribution of the population in an unconditional Galton-Watson tree, i.e., if Z_k is the law of the size of the k -th generation in an ordinary critical Galton-Watson tree, then

$$\mathbb{P}\{\xi'_k = i\} = i\mathbb{P}\{Z_k = i\}, i \geq 1.$$

Since M is arbitrary and $H_n/\sqrt{n} \xrightarrow{\mathcal{L}} H$, where H is a strictly positive random variable with the theta distribution (see, e.g., Flajolet and Odlyzko, 1982), we see that

$$Z_{n,k} \xrightarrow{\mathcal{L}} \xi'_k.$$

Results of this nature have been around in various forms. For example, Janson (2006b) showed that

$$\lim_{n \rightarrow \infty} \mathbb{E}\{Z_{n,k}\} = 1 + k\sigma^2 = \mathbb{E}\{\xi'_k\},$$

where $\sigma^2 = \mathbb{V}\{\xi\}$.

Consider sums of the form

$$V_n = \sum_{k=0}^{\infty} \varphi(k)\psi(Z_{n,k}),$$

where $\varphi(k)$ is summable in k and ψ is a bounded function. Our results imply that

$$\lim_{n \rightarrow \infty} \mathbb{E}\{V_n\} = \sum_{k=0}^{\infty} \varphi(k)\mathbb{E}\{\psi(\xi'_k)\}.$$

By Fatou's lemma, it is easy to see that the right-hand side is a lower bound. Furthermore, the upper bound follows from the summability of $\varphi(k)$ and the fact that for each fixed k , $d(Z_{n,k}, \xi'_k) = o(1)$.

Lemma 1 is severely limited—it can only handle questions about the behavior of the top of T_n . In contrast, Aldous's fringe result (1991), which states that the subtree rooted at a uniform random node in T_n tends in distribution to the unconditional Galton-Watson tree T , is useful for global parameters far away from the root.

Much remains to be done. For example, if $\mathbb{V}\{\xi\} = \infty$, under which conditions is it still true that $d(\mathcal{N}, \mathcal{N}') \rightarrow 0$? Can the total variation distance be bounded in a nice manner, so that the total variation bound can be applied to a much larger chunk of T_n , with a construction of a path of marked nodes of decent length? Can one easily derive limit laws for weighted sums of functions of subtrees, and, in particular, for weighted sums for functions of subtree sizes? What is the probability of cutting a conditional Galton-Watson tree for ξ with $\mathbb{E}\{\xi\} < 1$ and not normalizable to be critical?

6. Acknowledgment

This paper was motivated by a question asked by Vasek Chvatal. Discussions are gratefully acknowledged. We also thank the referee.

6. References

- D. Aldous, “Asymptotic fringe distributions for general families of random trees,” *The Annals of Applied Probability*, vol. 1, pp. 228–266, 1991.
- D. Aldous and J. Pitman, “Tree-valued Markov chains derived from Galton-Watson processes,” *Annals of the Institute Henri Poincaré*, vol. 34, pp. 637–686, 1998.
- K. B. Athreya and P. E. Ney, *Branching Processes*, Springer Verlag, Berlin, 1972.
- V. Campos, V. Chvatal, L. Devroye, and P. Taslakian, “Transversals in trees,” *Journal of Graph Theory*, 2011. To appear.
- L. Devroye, “Branching processes in the analysis of the heights of trees,” *Acta Informatica*, vol. 24, pp. 277–298, 1987.
- L. Devroye, “Universal limit laws for depths in random trees,” *SIAM Journal on Computing*, vol. 28, pp. 409–432, 1999.
- T. Duquesne, “An elementary proof of Hawkes’ conjecture on Galton-Watson trees,” *Electronic Communications in Probability*, vol. 14, pp. 151–164, 2009.
- M. Dwass, “The total progeny in a branching process,” *Journal of Applied probability*, vol. 6, pp. 682–686, 1969.
- J. D. Farley, “Breaking al Qaeda cells: a mathematical analysis of counterterrorism operations,” *Studies in Conflict and Terrorism*, vol. 26, pp. 399–411, 2003.
- J. D. Farley, “Toward a Mathematical Theory of Counterterrorism How to Build the Perfect Terrorist Cell, I,” Technical Report, Center fo International Security and Cooperation, Stanford University, 2007.
- P. Flajolet and A. Odlyzko, “The average height of binary trees and other simple trees,” *Journal of Computer and System Sciences*, vol. 25, pp. 171–213, 1982.

- P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge, 2008.
- J. Geiger and L. Kauffmann, “The shape of large Galton-Watson trees with possibly infinite variance,” *Random Structures and Algorithms*, vol. 25, pp. 311–335, 2004.
- C. Holmgren, “Split Trees, Cuttings and Explosions,” Ph.D. dissertation, Uppsala University, 2010.
- I. A. Ibragimov, “On the accuracy of the approximation of distribution functions of sums of independent random variables by the normal distribution,” *Theory of Probability and its Applications*, vol. 11, pp. 559–580, 1966.
- S. Janson, “Random records and cuttings in complete binary trees,” in: *Mathematics and Computer Science III, Algorithms, Trees, Combinatorics and Probabilities (Vienna 2004)*, edited by M. Drmota, P. Flajolet, D. Gardy and B. Gittenberger (eds), pp. 241–253, Birkhäuser, Basel, 2004.
- S. Janson, “Random cutting and records in deterministic and random trees,” *Random Structures and Algorithms*, vol. 29, pp. 139–179, 2006a.
- S. Janson, “Conditioned Galton-Watson trees do not grow,” in: *Proceedings, Fourth Colloquium on Mathematics and Computer Science Algorithms, Trees, Combinatorics and Probabilities (Nancy, 2006)*, pp. 331–334, DMTCS Proceedings AG, 2006b.
- D. P. Kennedy, “The Galton-Watson process conditioned on the total progeny,” *Journal of Applied Probability*, vol. 12, pp. 800–806, 1975.
- V. F. Kolchin, “Branching Processes and random trees,” in: *Problems in Cybernetics, Combinatorial Analysis and Graph Theory (in Russian)*, pp. 85–97, Nauka, Moscow, 1980.
- V. F. Kolchin, *Random Mappings*, Optimization Software Inc, New York, 1986.
- J.-F. Le Gall, “Marches aléatoires, mouvement Brownien et processus de branchement,” in: *Séminaire de Probabilités XXIII*, edited by J. Azéma, P. A. Meyer and M. Yor, vol. 1372, pp. 258–274, Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1989.
- R. Lyons, R. Pemantle, and Y. Peres, “Conceptual proof of $L \log L$ criteria for mean behavior of branching processes,” *Annals of Probability*, vol. 23, pp. 1125–1138, 1995.
- A. Meir and J. W. Moon, “On the altitude of nodes in random trees,” *Canadian Journal of Mathematics*, vol. 30, pp. 997–1015, 1978.
- J. W. Moon, *Counting labelled trees*, Canadian Mathematical Congress, Montreal, 1970.
- J. Neveu and J. W. Pitman, “The branching process in a Brownian excursion,” in: *Séminaire de Probabilités XXIII*, edited by J. Azéma, P. A. Meyer and M. Yor, vol. 1372, pp. 248–257, Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1989.
- V. V. Petrov, *Sums of Independent Random Variables*, Springer-Verlag, Berlin, 1975.

A. Rouault, “Lois empiriques dans les processus de branchement spatiaux homogènes supercritiques.” *Comptes Rendus de l’Académie des Sciences de Paris Série I (Math)*, vol. 292, pp. 933–936, 1981.

F. Spitzer, *Principles of Random Walk*, 2nd ed., Springer-Verlag, New York, 1976.