# Probability Theory on Galton-Watson trees

by

Alex Perlin

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Mathematics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2001

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
AUG 0 7 2001
LIBRARIES

Copyright. Alex Perlin. 2001. All rights reserved.

The author hereby grants to MIT permission to reproduce and distribute publicly paper an electronic copies of this thesis document in whole or part and to grant others the right to do so

Author ..............................................................
Department of Mathematics
May 18, 2001

Certified by ..............................................................
Daniel W. Stroock
Professor
Thesis Supervisor

Accepted by ..............................................................
Tomasz Mrowka
Chairman, Department Committee on Graduate Students

ARCHIVES

## Probability Theory on Galton-Watson trees

by

Alex Perlin

Submitted to the Department of Mathematics
on May 18, 2001, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Mathematics

### Abstract

By a Galton-Watson tree $T$ we mean an infinite rooted tree that starts with one node and where each node has a random number of children independently of the rest of the tree. In the first chapter of this thesis, we prove a conjecture made in [7] for Galton-Watson trees where vertices have bounded number of children not equal to 1. The conjecture states that the electric conductance of such a tree has a continuous distribution. In the second chapter, we study rays in Galton-Watson trees. We establish what concentration of vertices with is given number of children is possible along a ray in a typical tree. We also gauge the size of the collection of all rays with given concentrations of vertices of given degrees.

Thesis Supervisor: Daniel W. Stroock
Title: Professor

# Acknowledgments

It is a great pleasure to acknowledge that research presented in this dissertation has been conducted under the guidance of Professor Daniel W. Stroock. This thesis just would not be the same without his insights on Probability Theory (especially on Large Deviations) expressed in many conversations over the course of the last five years and comments on the style of presentation. This writer is and the reader should be very much indebted to Dr. Balint Virag whose remarks significantly simplified the most convoluted parts of the original proofs. In particular, Dr. Virag suggested the line of reasoning behind the results of section 2.7.

The author has also benefited from discussing the material included in the dissertation with Professors Sara Billey, Russell Lyons and Gilbert Strang. Dr. Lydia Gladkova has read the draft of the thesis and pointed out quite a few typos. All of the aforementioned mathematicians, as well as Victor Perlin, Misha Kogan and Ed Goldstein have been extremely encouraging throughout the course of graduate studies.

It seems fit to close the acknowledgment section by thanking those who stimulated author's interest in Mathematics long before the graduate school. They are Professor Ildar Ibragimov, Dr. Sergey Rukshin, and, of course, the parents Emilia Smorgonskaya and Eugene Perlin, who is also a professor!

# Chapter 1

# Conductance of Galton-Watson trees

## 1.1 Introduction

Given a random variable $\xi$ that takes on positive integer values, a branching process starts with one particle that has $\xi$ children. Each of the children in turn has a random number of children with the law of $\xi$. The children of children also have children and the process continues forever. We can draw a diagram of the process by associating a node with each particle and then connecting each node to the nodes representing its children. This diagram, a random infinite graph, is called a Galton-Watson tree. Let $Q_T$ be the probability that the simple random walk started at the root of the tree (node representing the original particle) $T$ never returns to the root.

**Theorem 1** *If $2 \leq \xi \leq k_0$, then the distribution of the random variable $Q_T$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}$ and the probability density of $Q_T$ is bounded.*

**Remark 1.** This writer believes he has a proof that the constraints $2 \leq \xi \leq k_0$ can be lifted, but it would make the exposition even more cumbersome.

**Remark 2.** If each edge of the tree is a wire of unit conductance, then the effective conductance of the whole tree (from the root to infinity) is given by $\frac{Q_T}{Q_T+1}$.

Therefore, the distribution of the effective conductance is also absolutely continuous.

**Remark 3.** Theorem 1 confirms a conjecture made in [7].

This chapter is devoted to the proof of this theorem. Before going into technicalities, we present a very vague idea of the proof. The theorem will follow from the statement of the type

$$P(Q_T \in I) \leq \frac{P(Q_T \in J)}{2} + \text{small error}, \tag{1.1}$$

where $I$ and $J$ are intervals located roughly at the same place, $J$ being twice as long as $I$. An inequality like (1.1) would have been relatively easy to verify for intervals of a certain size if $Q_T$ were a sum of sufficiently large number of independent random variables. It is not, but it is possible to condition $Q_T$ on a (random) subtree $T_0 \subset T$ in such a way that the conditional distribution of $Q_T$ is very closely approximated by the sum of independent random variables for almost all likely structures of $T_0$. Thus,

$$\begin{aligned} P(Q_T \in I) = E(P(Q_T \in I \,|\, T_0)) &\leq \frac{1}{2}E(P(Q_T \in J \,|\, T_0)) + \text{small error} \\ &\leq P(Q_T \in J) + \text{small error}. \end{aligned}$$

The reason why such an approximation for $Q_T$ given $T_0$ exists is because $T \setminus T_0$ will turn out to have a large number of components that are far apart. Their contributions to the effective conductance of the network are nearly independent and they nearly add up.

The organization of this chapter is as follows. Section 2 introduces the notation used throughout this chapter. It takes advantage of the symmetry of Galton-Watson trees to demonstrate that sampling according to the uniform flow measure is equivalent to choosing a specific node of the tree. Section 3 describes why changes in the tree that are made far from a specific node have a very little effect on the probability of visiting that node. Section 4 establishes that for the overwhelming majority of realizations of a Galton-Watson tree, one can choose a lot of vertices that are visited by the random walk with more or less the same probabilities. Section 5 studies the effect of removal of one or more subtrees from a tree on $Q_T$. In particular, it tells how closely $Q_T$ conditioned on some subtree of the graph can be approximated by a sum of independent random variables. Section 6 provides a regularity result for a slightly perturbed sum of independent random variables. Section 7 gives a sufficient scaling condition that guarantees that a measure is absolutely continuous with respect to Lebesgue measure. Section 8 draws upon sections 2 through 7 to prove the theorem.

## 1.2 Basic notation and the uniform flow measure.

The vertices of the trees we shall consider will be elements $(n_1, n_2, \dots) \in \mathbb{N}^\infty$, with all but finitely many $n_i$'s equal to zero. Occasionally we will abridge such a sequence to a finite one by removing all elements following the last non-zero term. The sequence of all zeros is called the root. It may be abridged to an empty sequence. We turn $\mathbb{N}^\infty$ into a tree by connecting all nodes of the form $(n_1, n_2, \dots, n_k)$ to $(n_1, n_2, \dots, n_k, n_{k+1})$ $(n_k, n_{k+1} > 0)$. For a vertex $v = (n_1, n_2, \dots, n_k)$, we will write $|v| = k$. More generally, the length of the shortest path in some subgraph $T$ of $\mathbb{N}^\infty$ between two vertices $v$ and $w$ will be expressed by $dist(v, w)$. Also the set of vertices $v \in \mathbb{N}^\infty$ with $|v| = k$ will be denoted $L_k$. It is called the $k$-th *generation* or the $k$-th *level* of the tree.

Define a collection of i.i.d random variables $\xi_v$ indexed by $v \in \mathbb{N}^\infty$ that take on positive integer values. If one imposes the constraint that $(n_1, n_2, \dots, n_k)$ is connected to $(n_1, n_2, \dots, n_k, n_{k+1})$ if and only if

$$0 < n_{k+1} \leq \xi_{(n_1, n_2, \dots, n_k)},$$

then the resulting subgraph of $\mathbb{N}^\infty$ is a forest. The component $T$ of the forest containing the root (which we will denote $rt$) is called the Galton-Watson tree. We will say a subtree of $\mathbb{N}^\infty$ is admissible if it contains the root and all nodes of the form $(n_1, \dots, n_{k-1}, \bar{n})$ with $\bar{n} \leq n_k$ once it contains $(n_1, \dots, n_{k-1}, n_k)$. The Galton-Watson tree is a random admissible tree.

If $v = (n_1, \dots, n_k)$ and $w = (n_1, \dots, n_l)$ are two vertices in an admissible tree $T$ and $k < l$, $v$ is said to be a *predecessor* or an *ancestor* of $w$ and $w$ is a *descendant* of $v$. In the particular case, when $l = k+1$, $v$ is said to be the *parent* of $w$, and $w$ is a *child* of $v$. If $n_{k+1} = \dots = n_l = 1$, we call $w$ the principal descendant of $v$ of order $l-k$, and write $w = pd(v, l-k)$. The principal descendants are well-defined because $\xi_v \geq 1$ for all nodes $v$.

For a vertex $v$ in a tree $T = (V, E)$, we will mean by $T^v$ the subtree of all descendants of $v$.

We will use $|I|$ to denote the length of $I$, when $I$ is an interval. Expression $\#A$ will always stand for the cardinality of a finite set $A$.

A functional $F(T, v_1, v_2)$ defined on trees and pairs of their vertices will be called *equivariant* if for any isomorphism of graphs $\pi : T_1 \to T_2$, and any two vertices $v_1$ and $v_2$ in $T_1$

$$F(T_2, \pi(v_1), \pi(v_2)) = F(T_1, v_1, v_2).$$

Given an admissible tree $T$, let $\mu_n$ be the measure on $L_n = \{v : |v| = n\}$, such that for any $v \in L_n$

$$\mu_n(\{v\}) = \frac{1}{deg(rt)} \Pi_w \frac{1}{deg(w) - 1},$$

where the product is taken over all vertices $w$ on the path between $v$ and the root. The measure $\mu_n$ is very well known. It corresponds to the uniform flow on $T$.

**Proposition 1** *Let $F$ be an equivariant functional taking on positive values. Then for any Galton-Watson tree $T$ and a positive integer $k$*

$$E(F(T, rt, pd(rt, k))) = E\left(\int F(T, rt, v)\, d\mu_k(v)\right) \tag{1.2}$$

**Proof of proposition 1.** Fix $k$ and $F$. For $m > 0$ introduce a functional

$$F_m(T, v, w) = \begin{cases} F(T, v, w), & \text{if } deg(v') \leq m \text{ for all } v' \text{ such that } dist(v', w) \leq k \\ 0, & \text{otherwise} \end{cases}$$

Clearly, the $F_m$'s are equivariant and increasing as $m$ grows. If one knew that (1.2) were true for all $F_m$, one would conclude that

$$\begin{aligned} E(F(T,\, rt,\, pd(rt,\, k))) &= E(\lim_{m\to\infty} F_m(T,\, rt,\, pd(rt,\, k))) \\ &= \lim_{m\to\infty} E(F_m(T,\, rt,\, pd(rt,\, k))) \\ &= \lim_{m\to\infty} E\left(\int F_m(T,\, rt,\, v)\, d\mu_k(v)\right) \\ &= E\left(\lim_{m\to\infty}\left(\int F_m(T,\, rt,\, v)\, d\mu_k(v)\right)\right) \\ &= E\left(\int F(T,\, rt,\, v)\, d\mu_k(v)\right). \end{aligned}$$

Therefore, it is sufficient to prove (1.2) for $F = F_m$ and in doing it we drop the subscript $m$.

Let group $\Gamma$ be the product of the first $m$ symmetric groups

$$\Gamma = \Pi_{i=1}^{m} \mathbb{S}_i.$$

(We view the elements of $\mathbb{S}_i$ as a bijective function from $\{1, \ldots, i\}$ to itself.) The canonical projection on the $i$-th component will be denoted $\pi_i : \Gamma \to \mathbb{S}_i$. We need many copies of $\Gamma$. In fact, we will have a separate $\Gamma_v$ for each $v \in V_m$ where $V_m$ stands for the set of sequences $\{(n_1, \ldots, n_l)\}$ with $0 \leq l \leq k$, subject to $1 \leq n_j \leq m$ for all $j$ between 1 and $l$. Let

$$\Gamma_0 = \Pi_{v \in V_m} \Gamma_v.$$

The canonical projections in $\Gamma_v$ will be denoted

$$\pi_i^v : \Gamma_v \to \mathbb{S}_i^v.$$

Canonical projections from $\Gamma_0$ onto $\Gamma_v$ will be called $\pi^v$.

For each $\gamma \in \Gamma_0$ we will define a map

$$\phi_\gamma : \mathbb{T}_{k,\, m} \to \mathbb{T}_{k,\, m}.$$

Here $\mathbb{T}_{k,m}$ is the set of all admissible trees such that $deg(v') \leq m$ for any $v'$ at the distance $k$ or less from the root.

Let $v = (n_1, \ldots, n_l)$ be a vertex in a tree $T \in \mathbb{T}_{k,m}$ and $\gamma \in \Gamma_0$. Suppose $v_0$ the parent of $v$. (If $v = rt$, then $\gamma$ preserves $v$.) Then we define $\phi_\gamma(v)$ to be the sequence

$$\phi_\gamma v = (\bar{n}_1, \ldots, \bar{n}_l),$$

where

$$\bar{n}_j = \begin{cases} ((\pi^{v_0}_{deg(v_0)-1}(\pi^{v_0}))(\gamma))(n_j), & \text{if } 1 \leq j \leq l \text{ and } v_0 \neq rt \\ ((\pi^{v_0}_{deg(v_0)}(\pi^{v_0}))(\gamma))(n_j), & \text{if } 1 \leq j \leq l \text{ and } v_0 = rt \\ n_j, & \text{if } l < j \end{cases}$$

Once it is known where $\phi_\gamma$ sends all the nodes of $T \subset \mathbb{N}^\infty$ we can properly connect the nodes in the image to get a tree $\gamma T \subset \mathbb{N}^\infty$. It is clear that $\phi_\gamma(T) \in \mathbb{T}_{k,m}$. Furthermore, $\phi_\gamma(T)$ is isomorphic to $T$. We extend the domain of $\phi_\gamma$ to include the set of all admissible trees requiring it to be the identity on the trees not in $\mathbb{T}_{k,m}$. It is routine to check that every $\phi_\gamma$ preserves the distribution of the Galton-Watson measure on the set of all admissible trees.

Consider a tree $T$ and a node $v \in T$ with $|v| = k$. Let $\Gamma(v) \subset \Gamma_0$, be the set of all $\gamma \in \Gamma_0$ such that $\phi_\gamma(v) = pd(rt, k)$. If $v_0 = rt$, $v_1, \ldots, v_{k-1}$ are the vertices on the path connecting $rt$ and $v = (n_1, \ldots, n_k)$, then $\phi_\gamma(v) = pd(rt, k)$ is equivalent to the requirement that

$$((\pi^{v_l}_{deg(v_l)[-1]}(\pi^{v_l}))(\gamma))(n_l) = 1 \quad (0 \leq l < k), \tag{1.3}$$

where $-1$ in the square brackets is used unless $l = 0$. Note that for $l > 0$ relation $\sigma(n_l) = 1$ is satisfied by $(deg(v_l) - 2)!$ elements $\sigma \in \mathbb{S}^v_{deg(v_l)-1}$, whereas for $l = 0$ it is true for $(deg(rt) - 1)!$ elements $\sigma$. It follows that for $l > 0$

$$\#\{s \in \Gamma_{v_l} \mid \pi^v_{deg(v_l-1)}(s)(n_l) = 1\} = \frac{(deg(v_l) - 2)!}{(deg(v_l) - 1)!}\#\Gamma_{v_l} = \frac{1}{deg(v_l) - 1}\#\Gamma_{v_l}. \tag{1.4}$$

Similarly,

$$\#\{s \in \Gamma_{rt} \mid \pi^{v}_{deg(rt)}(s)(n_l) = 1\} = \frac{(deg(rt)-1)!}{(deg(v_l))!}\#\Gamma_{rt} = \frac{1}{deg(rt)}\#\Gamma_{rt} \tag{1.5}$$

Equations (1.4) and (1.5) together imply that

$$\#\Gamma(v) = \#\Gamma\frac{1}{deg(rt)}\Pi_{i=1}^{k-1}\frac{1}{deg(v_i)-1} = \#\Gamma\mu_k(v) \tag{1.6}$$

On the other hand, since $\phi_\gamma$ is an isomorphism, given $\gamma \in \Gamma_0$, in any admissible tree $T$ there exists exactly one $v \in T$ such that $\phi_\gamma(v) = pd(rt, k)$.

With these preliminaries in place we are ready to attack proposition 1 (for the case $F = F_m$). According to a longstanding tradition $E(X; A)$ expresses the integral of the random variable $X$ over the event $A$.

$$\begin{aligned}
&E\left(\int F(T, rt, v)\, d\mu_k(v)\right) = E\left(\sum_{v:|v|=k} F(T, rt, v)\frac{\#\Gamma(v)}{\#\Gamma_0}\right)\\
&= \frac{1}{\#\Gamma_0,\ |v|=k}\sum_{\gamma\in\Gamma_0} E(F(T, rt, v); \phi_\gamma(v) = pd(rt,k)) \text{ use that } F \text{ is equivariant}\\
&= \frac{1}{\#\Gamma_0}\sum_{\gamma\in\Gamma_0,\,|v|=k} EF(\phi_\gamma(T), \phi_\gamma(rt), \phi_\gamma(v)); \phi_\gamma(v) = pd(rt,k))\\
&= \frac{1}{\#\Gamma_0}\sum_{\gamma\in\Gamma_0} E(F(\phi_\gamma(T), rt, pd(rt,k)))\\
&= \frac{1}{\#\Gamma_0}\sum_{\gamma\in\Gamma_0} E(F(T, rt, pd(rt,k)))\\
&= E(F(T, rt, pd(rt,k))).
\end{aligned}$$

The next to last line is a consequence of the fact that Galton-Watson measure is preserved under $\phi_\gamma$. Proposition 1 is now proved.

## 1.3 Small changes in trees.

The main idea of the proof of Theorem 1 is to choose a set of vertices, say $V_0$, in such a way that the contributions of the subtrees of their descendants $T^v$ ($v \in V_0$) to the probability of return to the origin are nearly independent. The technical problem that arises is that we may not look at these subtrees while making the choice. Therefore, we will have a set of candidate vertices $V_1$ which may or may not be chosen to serve in $V_0$. Proposition 2 below ensures that the subtrees $T^v$ with $v \in V_1 \setminus V_0$ do not spoil the picture. This means the contributions of the subtrees $T^v$ of $v \in V_0$ will be almost independent regardless of what the subtrees $T^v$ for $v \in V_1 \setminus V_0$ are. What is being said here is strictly informal and the notation of this paragraph will not be used in future.

The following simple lemma will be helpful throughout.

**Lemma 1.** Suppose in an infinite tree $T = (V, E)$ there is only one node of degree less than 3. Then for any two nodes $v$ and $w$ with $dist(v, w) = d$,

$$P_T(v \to w) \leq \left(\frac{5}{6}\right)^{d-1},$$

where $P_T(v \to w)$ denotes the probability that the random walk on $T$ started at $v$ visits $w$.

**Proof of lemma**. Let $\bar{v} \in V$ be the node of degree less than 3. Label the nodes of the path connecting $v$ to $w$, $v = v_0, v_1, \ldots, v_d = w$. Consider the random walk $(X_i)_{i \geq 0}$ on $T$ started at $v$, and set

$$\tau_i = \inf\{j \,:\, X_j = v_i\} \quad (0 \leq i \leq d)$$

By the strong Markov property,

$$P_T(v \to w) = P(\tau_d < \infty) = \Pi_{i=0}^{d-1} P(\tau_{i+1} < \infty \,|\, \tau_i < \infty). \tag{1.7}$$

We claim that if $v_i \neq \bar{v}$ and $v_i$ cannot be connected to $\bar{v}$ by a path entirely avoiding all other $v_j$'s $(0 \leq j \leq d)$, then

$$P(\tau_{i+1} < \infty \,|\, \tau_i < \infty) \leq \frac{5}{6}$$

Since there are at least $d-1$ vertices $v_i$ that meet the assumptions of the claim, lemma 1 will have been proved once the claim is verified. Let $A_i$ be the event that $X_{\tau_i+1}$ does not belong to the path from $v$ to $w$. Since $v_i$ has at least 3 neighbors, two (or just one) of them being part of the path from $v$ to $w$, $P(A_i \,|\, \tau_i < \infty) \geq \frac{1}{3}$. Let $B_i$ be the event that the random walk never returns to $v_i$ after visiting it once. Then using lemma 1 from section 5, one gets

$$\begin{aligned}
P(\tau_{i+1} < \infty \,|\, \tau_i < \infty) &\leq 1 - P(A_i \cap B_i \,|\, \tau_i < \infty) \\
&\leq 1 - P(B_i \,|\, A_i \cap \{\tau_i < \infty\}) P(A_i \,|\, \tau_i < \infty) \\
&\leq 1 - \left(\frac{1}{2}\right)\left(\frac{1}{3}\right) = \frac{5}{6}
\end{aligned}$$

as desired.

**Remark 1.** The conclusion of the lemma remains valid if $T$ has no nodes of degree smaller than 3. The same proof with obvious simplifications applies.

**Proposition 2** *Suppose $m < n$ are two positive integers. Let $T_1 = (V_1, E_1)$ and $T_2 = (V_2, E_2)$ be two admissible subtrees of $\mathbb{N}^\infty$ such that any vertex in the symmetric difference $T_1 \Delta T_2$ is a descendant of some vertex of the form $pd(v, n-m)$, where $v \in L_m \cap (V_1 \cup V_2)$. If $v_0 \in L_m \cap V_1$*

$$\begin{aligned}
&|P_{T_1}(rt \to pd(v_0, n-m)) - P_{T_2}(rt \to pd(v_0, n-m))| \\
&\qquad \leq \left(\frac{5}{6}\right)^{n-m-1} P_{T_1^{v_0}}(v_0 \to pd(v_0, n-m)).
\end{aligned}$$

**Proof of proposition 2.** We will only prove

$$\begin{aligned}
&P_{T_1}(rt \to pd(v_0, n-m)) - P_{T_2}(rt \to pd(v_0, n-m)) \\
&\qquad \leq \left(\frac{5}{6}\right)^{n-m-1} P_{T_1^{v_0}}(v_0 \to pd(v_0, n-m)). \qquad (1.8)
\end{aligned}$$

The other inequality required to get the conclusion of the lemma is derived in the same way, with $T_1$ and $T_2$ interchanged.

We use a coupling procedure to construct random walks on $T_1$ and $T_2$ on the same probability space. To wit, associate with each node in $v \in V_1$ a sequence of vertices $Y_j^v$ $(j = 1, 2, \ldots)$ selected independently among the nodes adjacent to $v$ in $T_1$. For any node $w$ adjacent to $v$ and any $j$, $P(Y_j^v = w) = \frac{1}{deg(v)}$. Let $V_0$ be the set of nodes $v$ that belong to and have the same sets of adjacent vertices in $T_1$ and $T_2$. For $v \in V_2 \setminus V_0$ produce a similar sequence $(Z_j^v)_{j \geq 1}$. It should be uniformly distributed on the neighbors of $v$ in $T_2$.

The following statement defines a version of the random walk $X^1 = (X_i^1)_{i \geq 0}$ on $T_1$ (started at the root). If $X_i^1$ is the $j$-th visit of $X^1$ to some vertex $w$, then $X_{i+1}^1 = Y_j^w$. For the random walk $X^2 = (X_i^2)_{i \geq 0}$ on $T_2$, if $(X_i^2)$ is the $j$-th visit of $X^2$ to some vertex $w$, then

$$X_{i+1}^2 = \begin{cases} Y_j^w, & \text{if } w \in V_0 \\ Z_j^w, & \text{if } w \in V_2 \setminus V_0 \end{cases}$$

Define

$$\tau = \inf\{i : X_i^2 = pd(v, n-m) \text{ for some } v \in L_m \cap T_1\}.$$

It is evident from the assumptions about $T_1$ and $T_2$, that $X_i^1 = X_i^2$ for all $i \leq \tau$. Consequently,

$$\{X^1 \text{ visits } pd(v_0, n-m)\} \subset \{X^2 \text{ visits } pd(v_0, n-m)\} \cup \{\tau < \infty\}.$$

It follows that

$$P_{T_1}(rt \to pd(v_0, n-m)) - P_{T_2}(rt \to pd(v_0, n-m)) \leq$$
$$\leq P(\{\tau < \infty\} \cap (X^1 \text{ visits } pd(v_0, n-m)) \cap (X_\tau^1 \neq pd(v_0, n-m)))$$

By the strong Markov property this expression is bounded by

$$\max_{v \in L_m \cap (V_1 \setminus \{v_0\})} P_{T_1}(pd(v, n-m) \to pd(v_0, n-m)).$$

Using lemma 1, the Markov property, and the fact that $v_0$ disconnects $v$ from $T_1^{v_0}$

and, in particular, from node $pd(v_0,\, n-m)$, write

$$\begin{aligned}
&P_{T_1}(pd(v,\, n-m) \to pd(v_0,\, n-m)) \\
&\qquad = P_{T_1}(pd(v,\, n-m) \to v) P_{T_1}(v \to pd(v_0,\, n-m)) \\
&\qquad \le \left(\frac{5}{6}\right)^{n-m-1} P_{T_1^{v_0}}(v_0 \to pd(v_0,\, n-m)).
\end{aligned}$$

The inequality (1.8) is now proved.

## 1.4 Nodes visited with approximately equal probability.

The objective of this section is to choose many nodes in the $n$-th generation of the tree that are far apart and are visited by a random walk started at the root of the tree with more or less equal probability. It is also important to know how these probabilities depend on $n$. The following proposition introduces a possible answer.

**Proposition 3** *On a Galton-Watson tree $T$ where each node has between 2 and $k_0$ children, define $Q_n = P_T(rt \to pd(rt,\, n))$. Then as $n \to \infty$ the limit $\lim \frac{\log Q_n}{n}$ exists almost surely. It is strictly negative and not random.*

**Proof of proposition 3.** Once the existence of $\lim \frac{\log Q_n}{n}$ is verified, the fact that it is not random will follow from Kolmogorov's 0-1 Law.

It is convenient to view a Galton-Watson tree $T$ as a subset of a larger infinite random tree $\bar{T}$. Intuitively one obtains $\bar{T}$ from $T$ by introducing the root's predecessors and their children other than the root. The formal construction goes as follows.

Let $BT_0$ be a set of infinite sequences of non-negative integers $(n_i)_{i \ge 0}$ subject to the properties

(BT1) If $n_i = 0$ for some $i > 0$, then $n_j = 0$ for all $j > i$.

(BT2) Only finitely many among $n_i$'s can be different from 0.

Let $\xi_v$ be a collection of i.i.d random variables having the same distribution as $\xi_{rt}$ indexed by $v \in BT_0$. An element of $BT_0$, say $v = (n_0,\, n_1, \dots, n_k,\, 0,\, \dots)$ with

$n_k \neq 0$, will be a vertex in $\bar{T}$ if and only if for all $2 \leq l \leq k$

$$n_l \leq \xi_{(n_0, n_1, \dots, n_{l-1}, 0, 0, \dots)}$$

and

$$n_1 \leq \begin{cases} \xi_{(n_0, 0, 0, \dots)} - 1 & \text{if } n_0 > 0 \\ \xi_{(n_0, 0, 0 \dots)} & \text{if } n_0 = 0 \end{cases}$$

Suppose $v = (n_i)_{i \geq 0}$, $w = (m_i)_{i \geq 0} \in \bar{T}$. We connect $v$ to $w$ and say that $v$ is the parent of $w$ in $\bar{T}$ if $n_i = m_i$ for all $i$ except for one index $i = i_0$, $m_j = 0$ for all $j > i_0$ and either (A0), (A1) or (A2) below is true.

(A0) $i_0 = 0$ and $n_0 - m_0 = 1$

(A1) $i_0 > 1$, $n_{i_0} = 0$, $m_{i_0} \leq \xi_{(n_0, n_1, \dots, n_{i_0 - 1}, 0, 0, \dots)}$

(A2) $i_0 = 1$, $n_1 = 0$ and

$$m_1 \leq \begin{cases} \xi_{(n_0, 0, 0, \dots)} - 1 & \text{if } n_0 > 0 \\ \xi_{(n_0, 0, 0 \dots)} & \text{if } n_0 = 0 \end{cases}$$

One can easily identify the Galton-Watson tree $T$ with the set of $v = (n_0, n_1, \dots)$ for which $n_0 = 0$.

For integer $i$ introduce random variables $Q_{i,i+1} = P_{\bar{T}}(w_i \to w_{i+1})$, $R_i = -\log Q(i)$. (As usual, the subscript $\bar{T}$ means that we are considering the random walk on $\bar{T}$ rather than on $T$.)

The very construction of the tree $\bar{T}$ makes the family $(R_i)_{i \geq 0}$ stationary in the sense that the joint distributions of $(R_i)_{i \in \mathbb{Z}}$ and $(R_{i+1})_{i \in \mathbb{Z}}$ coincide.

By lemma 1 from section 3

$$P_{\bar{T}}(rt \to w_n) \leq \left(\frac{5}{6}\right)^{n-1}.$$

On the other hand,

$$P_{\bar{T}}(rt \to w_n) \geq \left(\frac{1}{k_0 + 1}\right)^{n}.$$

Therefore, the $R_i$'s have finite expectations and

$$-\infty < \lim_{n\to\infty} \frac{\log Q_n}{n} < 0$$

if the limit exists.

Note that $P_{\bar{T}}(rt \to w_n) = \Pi_{i=0}^{n-1} Q_{i,\,i+1}$. Therefore,

$$\lim_{n\to\infty} \frac{\log P_{\bar{T}}(rt \to w_n)}{n} = -\lim_{n\to\infty} \frac{\sum_{i=1}^{n} R_i}{n},$$

so the limit does exist by the ergodic theorem.

Since $\frac{1}{2}Q_n \leq P_{\bar{T}}(rt \to w_n) \leq Q_n$ (cf lemma 1 in section 5), the preceding paragraph implies the statement of the proposition.

From now on we use $\alpha$ to denote the limit whose existence was proved in proposition 3.

Let $T = (V, E)$ be an admissible tree, $c_1$, $C_1$ and $\sigma$ positive numbers, $m < n$ positive integers. In a tree $T$ we define $GV(c_1, C_1, m, n, \sigma)$ to be the set of all vertices $v$ with $|v| = m$, such that

$$c_1 e^{-(\alpha+\sigma)n} \leq P_T(rt \to pd(v, n-m)) \leq C_1 e^{-(\alpha-\sigma)n}, \tag{1.9}$$

$$c_1 e^{-(\alpha+\sigma)(n-m)} \leq P_{T^v}(v \to pd(v, n-m)) \leq C_1 e^{-(\alpha-\sigma)(n-m)}. \tag{1.10}$$

**Proposition 4** *Let $\rho \in (0, 1)$. Then there exists $\theta_1 > 0$ such that for any choice of $\sigma > 0$ and $\theta_5 > 0$, there exist positive numbers $c_1$, $C_1$, $D$ such that for any integers $n > m > 2$, satisfying $|m - \rho n| < 1$*

$$P(\#GV(c_1, C_1, m, n, \sigma) \geq \exp(\theta_1 n)) \geq 1 - D\exp(-\theta_5 n),$$

*where $\#$... denotes the cardinality of the set.*

**Proof of proposition 4.** We will take advantage of the following elementary fact.

**Lemma 1.** If $Z$ is a random variable taking values in $[0,\ 1]$ and $E(Z) \geq 0.9$, then $P(Z \geq 2/3) \geq 0.7$.

**Proof of lemma 1.** By Markov's inequality

$$\begin{aligned} P(Z \geq 2/3) &= P(1 - Z \leq 1/3) \\ &= 1 - P(1 - Z \geq 1/3) \geq 1 - 3E(1 - Z) = 3E(Z) - 2 \geq 0.7. \end{aligned}$$

Before stating our lemma 2, we remark that by proposition 3 in this section there exists $K > 0$ such that for $k \geq K$

$$P\left(\frac{\log(P_T(rt \to pd(rt, k)))}{k} \in H\right) \geq 0.9, \text{ where } H = (-\alpha - \sigma/2, -\alpha + \sigma/2). \tag{1.11}$$

**Lemma 2.** Suppose integers $m_1$ and $m_2$ are greater than $K$. In a Galton-Watson tree $T$ define two following subsets of $L_{m_1} = \{v\ :\ |v| = m_1\}$

$$A_{m_1} = \left\{ w \in L_{m_1}\ :\ \frac{\log(P_T(rt \to w))}{m_1} \in H \right\},$$

$$B_{m_1} = \left\{ w \in L_{m_1}\ :\ \frac{\log(P_{T^w}(w \to pd(w,\ m_2)))}{m_2} \in H \right\}.$$

Then $P(\mu_{m_1}(A_{m_1} \cap B_{m_1}) \geq 1/3) \geq 0.4$. (As usual, $\mu_{m_1}$ is the uniform flow measure on $L_{m_1}$.)

**Proof.** Observe that $\mu_{m_1}$ depends on the structure of the tree before level $m_1$, while $P_{T^w}(w \to pd(w,\ m_2))$ is defined in terms of the structure of the tree at the level $m_1$ and higher. If $T_{m_1}$ stands for the (finite) tree cut off at the level $m_1$, this type of independence implies

$$
\begin{aligned}
E(\mu_{m_1}(B_{m_1})) &= E(E(\mu_{m_1}(B_{m_1}) \,|\, T_{m_1})) \\
&= E\left( \sum_{w \in L_{m_1}} \mu_{m_1}(\{w\}) P\left( \frac{\log(P_{T^w}(w \to pd(w,\, m_2)))}{m_2} \in H \,|\, T_{m_1} \right) \right) \\
&= P\left( \frac{\log(P_T(rt \to pd(rt,\, m_2)))}{m_2} \in H \right) E\left( \sum_{w \in L_{m_1}} \mu_{m_1}(\{w\}) \right) \\
&\geq (0.9)\mu_{m_1}(L_{m_1}) = 0.9.
\end{aligned}
$$

where the last inequality follows from the choice of $m_2 > K$. By lemma 1,

$$P(\mu_{m_1}(B_{m_1}) \geq 2/3) \geq 0.7. \tag{1.12}$$

Furthermore, by proposition 1 applied to

$$F(T, v, w) = \begin{cases} 1, & \text{if } \frac{\log P_T(v \to w)}{dist(v,w)} \in H \\ 0, & \text{otherwise,} \end{cases}$$

$$E(\mu_{m_1}(A_{m_1})) = E\left( \int F(T,\, rt,\, w) d\mu_{m_1}(w) \right) = E(F(T,\, rt,\, pd(rt,\, m_1))).$$

The latter quantity exceeds 0.9, since $m_1 > K$. Lemma 1 now implies

$$P(\mu_{m_1}(A_{m_1}) \geq 2/3) \geq 0.7. \tag{1.13}$$

Note that if both $\mu_{m_1}(A_{m_1}) \geq 2/3$ and $\mu_{m_1}(A_{m_1}) \geq 2/3$, then

$$\mu_{m_1}(A_{m_1} \cap B_{m_1}) = \mu_{m_1}(A_{m_1}) + \mu_{m_1}(B_{m_1}) - \mu_{m_1}(A_{m_1} \cup B_{m_1}) \geq 2/3 + 2/3 - 1 = 1/3.$$

Consequently,

$$
\begin{aligned}
P(\mu_{m_1}(A_{m_1} \cap B_{m_1}) \geq 1/3) &\geq P((\mu_{m_1}(A_{m_1}) \geq 2/3) \cap (\mu_{m_1}(B_{m_1}) \geq 2/3)) \\
&\geq P(\mu_{m_1}(A_{m_1}) \geq 2/3) + P(\mu_{m_1}(B_{m_1}) \geq 2/3) \\
&- P((\mu_{m_1}(A_{m_1}) \geq 2/3) \cup (\mu_{m_1}(B_{m_1}) \geq 2/3)) \\
&\geq 0.7 + 0.7 - 1 = 0.4.
\end{aligned}
$$

Lemma 2 is now proved.

**Proof of proposition 4.** Let $l = [\sqrt{m}]$, $m_1 = m - l$, $m_2 = n - m$ Assume temporarily that $m_1$ and $m_2$ are greater than $K$. For any $v \in L_l$ define $\mu_v$ to be the measure on $L_m$ that has the same distribution as $\mu_{m_1}$ corresponding to $T^v$, the tree of the descendants of $v$. (In other words, $\mu_v$ corresponds to the uniform flow started at $v$ and stopped at the level $m$ of the original tree.) Continuing with the same $v$, set

$$A_v = \left\{ w \in L_m \, : \, \frac{\log(P_{T^v}(v \to w))}{m_1} \in H \right\},$$

$$B_v = \left\{ w \in L_m \, : \, \frac{\log(P_{T^w}(w \to pd(w, m_2)))}{m_2} \in H \right\}.$$

Lemma 2 allows us to write

$$P(\mu_v(A_v \cap B_v) \geq 1/3) \geq 0.4. \tag{1.14}$$

Note that the events $\mu_v(A_v \cap B_v) \geq 1/3$ are mutually independent as $v$ runs through $L_l$, a set of at least $2^l$ elements. Indeed, these events only have to do with trees $T^v$.

Let $BAD$ be the event that

$$\#\{v \in L_l \, : \, \mu_v(A_v \cap B_v) < 1/3\} < 2^{l-2}.$$

With (1.14) at hand, it is a simple large deviation estimate that

$$P(BAD) \leq 2\exp(-\nu 2^l) \tag{1.15}$$

for some absolute constant $\nu$.

Moreover, $\mu_v(\{w\}) \leq 2^{-m_1}$ for any $v \in L_l$ and $w \in L_m$. Hence, if the inequality (1.14) is a true for some $v$, it has at least $\frac{2^{m_1}}{3}$ descendants at the level $m$ that are both in $A_v$ and $B_v$. Therefore, if $BAD^c$ takes place,

$$\# \cup_{v \in L_l} (A_v \cap B_v) \geq 2^{l-2}\frac{2^{m_1}}{3} = \frac{1}{12}2^m \tag{1.16}$$

**Remark 1.** Since $|m - \rho n| < 1$ the right-hand side of (1.16) is greater or equal to $\exp(\theta_1 n)$ where $n \geq n_0$ for a suitable choice of $\theta_1$ that may depend on $\rho$ and where quantity $n_0$ also depends only on $\rho$.

Fix for a moment $w \in A_v \cap B_v$. Since $w \in B_v$, it is immediately clear that $w$ satisfies (1.10), with any choice of $c_1 < 1$, $C_1 > 1$. In addition,

$$\begin{aligned} P_T(rt \to pd(w, n-m)) &\leq P_{T^v}(v \to pd(w, n-m)) \\ &\leq P_{T^v}(v \to w) P_{T^w}(w \to pd(w, n-m)) \\ &\leq \exp(-(\alpha - \sigma/2)(n-l)). \end{aligned}$$

Note that $l \leq \sqrt{\rho n + 1}$. Therefore the second inequality in (1.9) will hold for a suitable choice of $C_1 > 1$. (Here $C_1$ is allowed to depend on $\rho$, $\alpha$ and $\sigma$ but, of course, not on $n$.) Assuming each vertex in $T$ has no more than $k_0$ children

$$\begin{aligned} P_T(rt \to pd(w, n-m)) &\geq \left(\frac{1}{2}\right) P_T(rt \to v) P_{T^v}(v \to pd(w, n-m)) \\ &\geq \left(\frac{1}{2}\right) \frac{1}{(k_0+1)^l} \left(\frac{1}{2}\right) P_{T^v}(v \to w) P_{T^w}(w \to pd(w, n-m)) \\ &\geq \frac{1}{4(k_0+1)^l} \exp(-(n-l)(\alpha + \sigma/2)) \end{aligned}$$

Since $l \leq \sqrt{\rho n + 1}$, the quantity in the previous line is greater than $c_1 \exp(-(\alpha+\sigma)n)$ for a suitable choice of $c_1$.

We conclude that the event $BAD^c$ guarantees

$$\#GV(c_1, C_1, m, n, \sigma) \geq \exp(\theta_1 n)$$

where the choice of constants $c_1$, $C_1$ and $\theta_1$ only depended on $\rho$ and $\sigma$ and $\alpha$. It is evident from (1.15) that

$$P(BAD^c) \geq 1 - D\exp(-\theta_5 n)$$

where the choice of $D$ may depend on the value of $\rho$ as $l \approx \sqrt{\rho n}$. Finally the constraints $m_1 \approx \rho n - \sqrt{\rho n} > K$, $m_2 \approx (1 - \rho n) > K$, $n > n_0$ can be lifted, by increasing $D$ in such a way that $1 - D\exp(\theta_5) < 0$ for all (small) $n$ that do not meet those constraints. Proposition 4 is proved.

Fix an admissible tree $T$ in which any node has $k_0$ or fewer children. Let $A_{2v}$ be the event that the random walk on $T$ started at the root visits at least two vertices of the form $pd(v, n-m)$ with $v \in GV(c_1, C_1, m, n, \sigma)$.

**Proposition 5** *In the above notation,*

$$P(A_{2v}) \leq 3C_1^3 k_0^{2m+1} \exp(-3(n-m)(\alpha - \sigma)).$$

**Proof of proposition 5.** For any two nodes $v_1$ and $v_2$ in $GV(c_1, C_1, m, n, \sigma)$, let $A_{v_1, v_2}$ be the event that the random walk visits $pd(v_1, n-m)$ and then it visits $pd(v_2, n-m)$. Clearly,

$$A_{2v} \subset \cup_{v_1, v_2} A_{v_1, v_2},$$

whence

$$P(A_{2v}) \leq (\#GV(c_1, C_1, m, n, \sigma))^2 \max P(A_{v_1, v_2}), \tag{1.17}$$

where the union and the maximum are taken over all pairs of distinct vertices $v_1$ and $v_2$ in $GV(c_1, C_1, m, n, \sigma)$. For any such pair, using Markov property and (1.21) of the next section, we have

$$
\begin{aligned}
P(A_{v_1,v_2}) &\le P_{T^{v_1}}(v_1 \to pd(v_1,\, n-m)) \times \\
&\times P_{T^{v_1}}(pd(v_1,\, n-m) \to v_1) P_{T^{v_2}}(v_2 \to pd(v_2,\, n-m)) \\
&\le (C_1 \exp(-2(n-m)(\alpha-\sigma)) \times \\
&\times (3k_0(C_1 \exp(-(n-m)(\alpha-\sigma))) \\
&\le 3k_0 C_1^3 \exp(-3(\alpha-\sigma/2)(n-m)).
\end{aligned}
$$

To prove the proposition combine the inequality just obtained with (1.17) and the fact that

$$\#GV(c_1,\, C_1,\, m,\, n,\, \sigma) \le \#L_m \le k_0^m.$$

## 1.5 Approximating the probability of no return to the root by a sum.

The purpose of this section is to investigate how a local change in a tree influences the probability that the random walk started at its root never visits the root again.

Let $T = (V, E)$ be an infinite tree with root $rt$. Suppose $V_1 \subset V \setminus \{rt\}$ is a finite collection of vertices such that the shortest path between any element of $V_1$ and $r$ contains no other elements of $V_1$. For $v \in V_1$, let $T_v$ be a subtree of $T$ obtained by cutting off the descendants of all vertices $w \in V_1 \setminus \{v\}$. (In other words, to get $T_v$ we eliminate from $T$ the vertices that would become disconnected from the root if all the vertices in $V_1$ but $v$ were removed.) Let $T_0$ be the subtree obtained by cutting off all the descendants of all $w \in V_1$. Clearly, $T_0 \subset T_v \subset T$. For any subtree $S \subset T$, let $Q_S$ be the probability that the simple random walk on the tree $S$ started at $rt$ never returns to $rt$. Also let $A$ stand for the event that the random walk on $T$ started at the root visits at least two points in $V_1$.

**Proposition 6** *In the preceding setup,*

$$|Q_T - Q_{T_0} - \sum_{v \in V_1} (Q_{T_v} - Q_{T_0})| \le P(A) \# V_1,$$

*where $\# V_1$ denotes the cardinality of $V_1$.*

**Proof of proposition 6.** We begin by associating with each node $v \in V$ a sequence of vertices $Y_n^v$ (here $n = 1, 2, \dots$) that are selected independently and uniformly from the set of nodes adjacent to $v$ in $T$. We may and will assume that the random walk $(X_k)_{k \ge 0}$ is being run in such a way, that if $X_l$ is the $n$-th visit to some vertex $w$, then $X_{l+1} = Y_n^w$. The random walk $(\bar{X}_k)_{k \ge 0}$ on $T_0$ is being run according to the same rule when $X_l \in V \setminus V_1$, and if $X_l \in V_1$ then $X_{l+1}$ is the parent of $X_l$. For $v \in V_1$, we consider the random walk $(X^v)_{k \ge 0}$ on $T_v$ started at the root. It is run much in the same way as $(\bar{X}_k)_{k \ge 0}$, but with the provision that $X_{l+1}$ must be the parent of $X_l$ applied only when $X_l \in V_1 \setminus \{v\}$. To summarize, we have defined random walks on $T$, $T_0$ and $T_v$'s on the same probability space. Call this probability space $\Omega$.

Note that $\Omega$ can be represented as a disjoint union

$$\Omega = A \cup \left( \bigcup_{v \in V_1} A_v \right) \cup A_0,$$

where $A_v$ is the event that $v$ is the only vertex of $V_1$ visited by the random walk $(X_k)_{k \ge 0}$, and $A_0$ is the event that it visits no nodes in $V_1$. Furthermore, let $B_T$ ($B_0$, $B_v$) be the event that the random walk on $T$ ($T_0$, $T_v$ respectively) doesn't return to the root. Finally, set $\tau_v = \inf\{k : X_k = v\}$, $\tau = \inf\{\tau_v : v \in V_1\}$. Clearly, $\tau_v = +\infty$ on $\Omega \setminus (A \cup A_v)$, and $\tau = \infty$ on $A_0$.

We claim that for any $v \in V_1$,

$$B_0 \subset B_v \subset B_T. \tag{1.18}$$

Indeed, suppose $\omega \in B_v^c$. Then for some $k > 0$, $X_k^v(\omega) = r$. Let $i_1 < i_2 \dots < i_l = k$ be all the indexes $i$ such that neither $X_{i-1}^v(\omega)$ nor $X_i^v(\omega)$ is a descendant of $v$. Then it is not hard to show by induction on $s$ that $\bar{X}_s(\omega) = X^v_{i_s}(\omega)$. In particular, $\bar{X}_l(\omega) = r$, whence, $\omega \in B_0^c$. Thus, $B_v^c \subset B_0^c$. The first inclusion in (1.18) follows.

The second inclusion can be proved in a similar fashion. Assume $\omega \in B_T^c$. Then for some $k > 0$, $X_k(\omega) = r$. Let $i_1 < i_2 \ldots < i_l = k$ be all the indexes $i$ such that neither $X_{i-1}(\omega)$ nor $X_i(\omega)$ is a descendant of some $w \in V \setminus \{v\}$. Then again by induction on $s$, we get $X^v_s(\omega) = X_{i_s}(\omega)$. Therefore, $\bar{X}_l(\omega) = r$. Hence, $\omega \in B_v^c$. Thus, $B_T^c \subset B_v^c$, which concludes the proof of (1.18).

With (1.18) at our disposal, to prove the proposition we need to establish

$$|P(B_T \setminus B_0) - \sum_{v \in V_1} P(B_v \setminus B_0)| \leq P(A). \tag{1.19}$$

Given our construction of the random walk, a simple induction shows $X^v_k = \bar{X}_k$ for all $k \leq \tau_v$. It follows that $B_v \setminus B_0 \subset A \cup A_v$. Also $X_k = \bar{X}_k$ for $k \leq \tau$. Thus, $(B_T \setminus B_0) \cap A_0 = \emptyset$. Moreover,

$$(B_T \setminus B_0) \cap A_v = (B_v \setminus B_0) \cap A_v.$$

In fact, on $A_v$ $(X_k)_{k\geq 0} = (X^v)_{k \geq 0}$, because $A_v \subset \cap_{w \neq v}\{\tau_w = \infty\}$.

Consequently,

$$\begin{aligned}\left|P(B_T \setminus B_0) - \sum_{v\in V_1} P(B_v \setminus B_0)\right| &= |\sum_{w \in V_1} P((B_T \setminus B_0) \cap A_w) + P((B_T \setminus B_0) \cap A) \\ &\quad - \sum_{w\in V_1} P((B_v \setminus B_0) \cap A_w) - \sum_{v \in V_1} P((B_v \setminus B_0) \cap A)| \\ &= \left|P((B_T \setminus B_0) \cap A) - \sum_{v\in V_1} P((B_v \setminus B_0)\cap A)\right| \\ &\leq P(A)\#V_1,\end{aligned}$$

which shows (1.19). Proposition 6 is now proved.

Suppose $T = (V, E)$ is a tree with root $rt$ and $v \in V$ is a leaf (node of degree 1) in $T$. Let $G$ be an arbitrary *random* rooted tree. Form a tree $T_G$ by attaching $G$ to $T$ at $v$. Thus, the tree of descendants of $v$ in $T_G$ is isomorphic to $G$. Assume that all vertices in $T_G$ have between 2 and $k_0$ children.

Let $Q_{T_G}$ be the conditional probability given $G$ that the random walks on $T_G$ started at $rt$ does not return to $rt$. Since $G$ is random, $Q_{T_G}$ is a random variables.

Recall that $P_T(v \to w)$ stands for the probability that random walk on $T$ started at $v$ visits $w$. Set $q_v = P_T(rt \to v)$.

**Proposition 7** *Referring to the preceding setup,*

$$0 \leq Q_{T_G} - Q_T \leq 3k_0 q_v^2$$

**Lemma 1.** Let $S$ be a tree with root $rt$, $v$ be a child of $rt$. If $v$ and all of its descendants have at least two children, then $P_S(v \to rt) \leq 1/2$.

**Proof of lemma 1.** If $X_0, X_1, \ldots$ is a random walk started at $v$, $Y_i$ denotes the distance between $X_i$ and $r$, and $F_i$ is the $\sigma$-field generated by the first $i$ steps of the random walk, then

$$\frac{2}{3} \leq P(Y_i - Y_{i-1} = 1 \mid F_i) = 1 - P(Y_i - Y_{i-1} = -1 \mid F_i).$$

The statement of the lemma is now the result of comparison of $Y_i$ to one-dimensional random walk (also known as gambler's ruin problem) with bias equal to 2/3. Such a random walk visits 0 with probability 1/2 if it was started at 1.

**Lemma 2.** Let $v$ be a vertex in a rooted tree $S$ where any vertex except $v$ has at least two children. Define $G_S(v, v)$ to be the expected number of visits to $v$ (counting time 0) for a random walk on $S$ started at $v$. Then $G_S(v, v) \leq 6$.

**Proof of lemma 2.** Let $(Y_i)_{i \geq 0}$ be a random walk on $S$ started at $v$. Let $D$ be the event that $Y_2$ is a child of $Y_1$ and $Y_2 \neq v$. Clearly $P(D) \geq \frac{1}{3}$. Let $R_k$ be the event that $(Y_i)_{i \geq 0}$ makes fewer than $k$ visits to $v$ after time 0. By lemma 1, $P(R_1 \mid D) \geq 1/2$. Therefore,

$$P(R_1) = P(R_1 \mid D)P(D) \geq \frac{1}{6}.$$

Hence,

$$P(R_k) \geq \frac{1}{6}\left(\frac{5}{6}\right)^{k-1} \tag{1.20}$$

To conclude the proof of lemma 2 write,

$$G_S(v, v) = 1 + \sum_{k=1}^{\infty}(1 - P(R_k)) \leq \sum_{k\geq 0}\left(\frac{5}{6}\right)^k = 6.$$

**Lemma 3.** Let $v$ be a vertex in a rooted tree $T$ where each vertex has between 2 and $k_0$ children. Then

$$\frac{1}{3k_0}P_T(rt \to v) \leq P_T(v \to rt) \leq 3k_0 P_T(rt \to v). \tag{1.21}$$

**Proof of lemma 3.** We will only prove the second inequality. The proof of the first one goes along the same lines. Let $q_1$ be the probability that a particle starting its random walk at $rt$ visits $v$ before returning to $rt$. Let $q_2$ be the probability that a particle starting its random walk at $v$ would visit $rt$ before returning to $v$. Since random walk is a reversible Markov chain,

$$q_2 = \frac{deg\, rt}{deg\, v} q_1 \leq \frac{k_0}{2} P_T(rt \to v). \tag{1.22}$$

Moreover, by a standard Markov chain argument and lemma 2

$$P_T(v \to rt) \leq G_T(v, v) q_2 \leq 6 q_2. \tag{1.23}$$

The statement of lemma 3 is now a simple combination of (1.22) and (1.23).

**Proof of Proposition 7.** In this proof we will deal with $Q_{T_G}$ evaluated given that $G = G_0$, where $G_0$ is some non-random graph. Therefore, the randomness and the probabilities involved will have to do with the stochastic nature of the random walk, rather than with the fact that a random tree was attached to $T$. Furthermore, we assume that the random walks $X_k$ and $X_k^G$ on $T$ and $T_G$ respectively started at $rt$

are defined on the same probability space (as in the proof of proposition 6.) Observe that

$$Q_{T_G} - Q_T = P(B_{T_G} \setminus B_T) = P(B_T^c \setminus B_{T_G}^c),$$

where $B_{T_G}$ ($B_T$ )is the event that the random walk on $T_G$ (respectively $T$) never returns to $rt$. (By our coupling $B_T \subset B_{T_G}$.) Consequently, $Q_{T_G} - Q_T \geq 0$. Let

$$\tau_v = \inf\{k : X_k = v\}, \ \tau_{rt} = \inf\{k > 0 : X_k = rt\}.$$

For $k \leq \tau_v$, $X_k = X_k^G$, whence

$$B_T^c \setminus B_{T_G}^c \subset \{\tau_v < \infty\} \cap \{\tau_v \leq \tau_{rt}\} \cap \{\tau_{rt} < \infty\}.$$

Now by the strong Markov property of the random walk and lemma 3,

$$P(B_T^c \setminus B_{T_G}^c) \leq P_T(rt \to v) P_T(v \to rt) \leq 3k_0 q_v^2,$$

which proves the right-hand inequality of proposition 7. Proposition 7 is proved.

Our last goal in this section is to give a lower bound for $Var(Q_{T_G})$. We continue with the same $T$, $v$ and $v_0$ as before proposition 7.

**Proposition 8** *Assume that the probability that a random walk on G returns to the root of G is less than 1 and not constant a.s. Then*

$$Var(Q_{T_G}) \geq \frac{C_G}{k_0^2}(q_v)^4$$

*for some $C_G > 0$ may depend on $G$ but not on $T$.*

**Proof of proposition 8.** In the proof of this proposition we will distinguish between two probability measures. Namely, $RW$ will stand for the probabilities of the events related solely to the random walk on $T$. At the same time $RG$ will denote the probabilities of the events due to the randomness of $G$.

**Lemma 4.** Let $N_v$ be the number of times the random walk on $T$ started at $rt$

visits $v$ before its first return to $rt$. (We set $N_v = 0$ if the root is never visited again.) Then

$$RW(N_v = 1) \geq \frac{1}{108k_0} q_v^2, \text{ and} \tag{1.24}$$

$$RW(N_v = n) = P(N_v = 1)\gamma^{n-1},$$

for some $\gamma \leq \frac{5}{6}$ and all $n \geq 1$.

**Proof of lemma 4.** Using lemmas 2 and 3, we obtain

$$RW(N_v \geq 1) \geq \frac{RW_T(rt \to v)}{G_T(rt, rt)} RW_T(v \to rt) \geq \frac{1}{(6)(3k_0)} (RW_T(r \to v))^2 = \frac{1}{18k_0} q_v^2 \tag{1.25}$$

Therefore, to establish lemma it will suffice to prove that

$$RW(N_v = l \,|\, N_v \geq l) \geq \frac{1}{6}, \tag{1.26}$$

and that the quantity in the left-hand side of (1.26) does not depend on $l > 0$. Since $v$ is a leaf in $T$,

$$RW(N_v = l \,|\, N_v \geq l) = 1 - \delta, \tag{1.27}$$

where $\delta$ is the conditional probability that the random walk on $T$ started at $v_0$ visits $v$ before the root, given that it eventually visits the root. Formula (1.27) makes it clear that the left-hand side of (1.26) does not depend on $l$.

Let $\tau_v$ be the time of the first visit to $v$ by a random walk on $T$ started at the root. Modify $\tau_v$ by setting $\tau_v = \infty$ if the random walk returns to $rt$ before visiting $v$. Then,

$$\begin{aligned} RW(N_v = 1 \,|\, N_v \geq 1) &= \frac{RW(N_v = 1)}{RW(N_v \geq 1)} = \frac{RW(N_v = 1,\, \tau_v < \infty)}{RW(N_v \geq 1,\, \tau_v < \infty)} \\ &= \frac{RW(N_v = 1 \,|\, \tau_v < \infty)}{RW(N_v \geq 1 \,|\, \tau_v < \infty)} \end{aligned} \tag{1.28}$$

By the strong Markov property, $RW(N_V \geq 1 \,|\, \tau_v < \infty) = RW(v \to rt)$. On the other hand, $RW(N_v = 1 \,|\, \tau_v < \infty)$ equals to the probability that a random walk started at $v$ will visit the root before returning to $v$. Consequently,

$$\frac{RW(N_v = 1 \,|\, \tau_v < \infty)}{RW(N_V \geq 1 \,|\, \tau_v < \infty)} \geq \frac{1}{G_T(v,\, v)} \geq \frac{1}{6}, \tag{1.29}$$

where the second inequality comes from lemma 2. Estimate (1.26) can now be obtained by combining (1.28) and (1.29). Lemma 4 is proved.

Getting back to proposition 8, let $Q_G$ be the RG-probability that random walk on $T_G$ started at $v$ eventually visits $v_0$. Clearly $Q_G$ depends only on structure of the graph $G$. Observe that if we fix $G$, then for any positive $n$ $RW(Q_{T_G} - Q_T \,|\, N_v = n) = 1 - Q_G^n$. By the formula of complete probability

$$Q_{T_G} - Q_T = \sum_{n=1}^{\infty} (1 - Q_G^N) RW(N_v = n).$$

Hence,

$$\begin{aligned} Var_{RG}(Q_{T_G}) &= Var_{RG}\left( \sum_{n=1}^{\infty} Q_G^N RW(N_v = n) \right) \\ &= Var_{RG}\left( RW(N_v = 1) \sum_{n=1}^{\infty} Q_G^N \gamma^{n-1} \right) \\ &= [RW(N_v = 1)]^2 Var_{RG}\left( \frac{Q_G}{1 - \gamma Q_G} \right). \end{aligned} \tag{1.30}$$

Note that

$$Var\left( \frac{Q_G}{1 - \gamma Q_G} \right)$$

is a continuous function of $\gamma$ that attains its minimum $C_G$ on the interval $[0,\, 5/6]$. Since $Q_G$ is not a constant random variable, $C_G > 0$. By (1.30) and lemma 4,

$$Var_{RG}(Q_{T_G}) \geq C_G [RW(N_v = 1)]^2$$

In view of (1.24), the last line allows us to write

$$Var_{RG}(Q_{T_G}) \geq \frac{C_G}{108^2 k_0^2} q_v^4$$

which proves proposition 8.

## 1.6 A regularity result for sums of independent random variables

Let $RV(n, \theta_1, A_2, \theta_2, A_3, \theta_3)$ be the collection of random variables $\bar{S}_N$ that can be represented in the form

$$\bar{S}_N = X_1 + X_2 + \ldots + X_N + Y,$$

where

(i) $N$ is some integer satisfying $\exp(\theta_1 n) \leq N \leq 2\exp(\theta_1 n)$.

(ii) The random variables $X_1, \ldots, X_N$ are mutually independent. (But they may depend on $Y$.)

(iii) For any integer $i$ $(1 \leq i \leq N)$ $Var(X_i) \geq A_2 \exp(-2\theta_2 n)$.

(iv) For the same $i$ as above $P\left(|X_i| \leq A_3 \exp(-\theta_3 n)\right) = 1$.

(v) Perturbation $Y$ satisfies $P(|Y| \leq A_3 \exp(-\theta_3 n)) = 1$.

We will always be assuming that $\theta_2 \geq \theta_3$, since otherwise assumptions (iii) and (iv) would be inconsistent for large $n$.

Use $|I|$ to denote the length of an interval $I$. Let $IL(n, \theta_4)$ be the collection of all pairs of intervals $I$ and $J$ on the real line such that

(vi) $|J| = 2|I|$

(vii) $e^{-\theta_4(n+1)} \leq |I| \leq e^{-\theta_4 n}$

(viii) $I \subset J$

**Proposition 9** *Suppose positive numbers* $\theta_1,\ A_2,\ \theta_2,\ A_3,\ \theta_3,\ \theta_4$ *satisfy*

$$\theta_3 - 3(\theta_2 - \theta_3) > \theta_4 \textit{ and} \tag{1.31}$$

$$\theta_4 > \theta_2 - \frac{\theta_1}{2} \tag{1.32}$$

*Then there exist positive numbers* $A$, $B$ *and* $\theta$ *such that for any* $n$, *any random variable* $S_N \in RV(n, \theta_1,\ A_2, \theta_2, A_3, \theta_3)$ *and* $(I_0, J_0) \in IL(n, \theta_4)$

$$P(\bar{S}_N \in I_0) \le \frac{(1 + A\exp(-\theta n))}{2} P(\bar{S}_N \in J_0) + B|I|^2 \tag{1.33}$$

The constants $\theta_1,\ A_2,\ \theta_2,\ A_3,\ \theta_3,\ \theta_4$ will be referred to as parameters of the setting. In the course of the proof we will choose a lot of other positive constants. Such a choice will be called legitimate if it only depends on the parameters of the setting but not on $n$, $I_0$, $J_0$ or random variables involved.

**Proof of proposition.** Changing $A_3$ if necessary, we may assume that $E(X_i) = 0$ for all $i$. Choose $\bar{S}_N$, $I_0$, $J_0$ as in the proposition and let $S_N = \bar{S}_N - Y$. Moreover, let $F_i$ be the cumulative distribution function (c. d. f.) for $X_i$, and $F$ be the c.d.f. for $S_N$.

We begin by deriving an inequality that will take care of $Y$, the small perturbation of $S_N$. Attach two segments of length $A_3\exp(-\theta_3 n)$ to the ends of $I_0$, thereby obtaining a larger interval $I$. Reduce $J_0$ to a smaller interval $J$ by removing a subinterval of length $A_3\exp(-\theta_3 n)$ at each end. It is clear from the definition of $IL(n, \theta_4)$ that

$$\sup J - \inf I \le 2\exp(-\theta_4 n). \tag{1.34}$$

Furthermore,

$$P(\bar{S}_N \in I_0) \le P(S_N \in I), \tag{1.35}$$

$$\frac{P(\bar{S}_N \in I_0)}{P(\bar{S}_N \in J_0)} \leq \frac{P(S_N \in I)}{P(S_N \in J)}. \tag{1.36}$$

Our assumptions dictate that $\theta_4 < \theta_3$ (see (1.31)) and $|J_0| = 2|I_0| \geq \exp(-\theta_4(n+1))$. Hence,

$$\frac{|I|}{|J|} \leq \frac{|I| + 2A_3 \exp(-\theta_3 n)}{2|I| - 2A_3 \exp(-\theta_3 n)} \leq \frac{1 + a_0 \exp(-\delta_0 n)}{2} \tag{1.37}$$

for some legitimate choice of $a_0$ and $\delta_0$.

Without loss in generality we may assume that $\sup I > 0$. Set $x = \max(0, \inf I)$. Then

$$J \subset [x - |I_0|,\, x + 2|I_0|]. \tag{1.38}$$

In view of (1.32) it is possible to choose $\beta$ in such a way, that

$$\theta_2 - \theta_1/2 < \beta < \theta_4 \tag{1.39}$$

Set $\lambda_0 = \exp(n\beta)$,

$$y = \frac{\int_{-\infty}^{\infty} t \exp(\lambda_0 t)\, dF(t)}{\int_{-\infty}^{\infty} \exp(\lambda_0 t)\, dF(t)}. \tag{1.40}$$

Quantity $y$ is well defined, because $S_N$ is bounded. Since

$$\int_{-\infty}^{\infty} t \exp(\lambda t)\, dF(t) = \frac{d}{d\lambda}\left[\int_{-\infty}^{\infty} \exp(\lambda t)\, dF(t)\right],$$

$$\begin{aligned} y &= \frac{d}{d\lambda}\left[\log\left(E(\exp(\lambda S_N))\right)\right]|_{z=\lambda_0} = \sum_{i=1}^{N} \frac{d}{d\lambda}\left[\log\left(E(\exp(\lambda X_i))\right)\right]|_{z=\lambda_0} \\ &= \sum_{i=1}^{N} \frac{\int_{-\infty}^{\infty} t \exp(\lambda_0 t)\, dF_i(t)}{\int_{-\infty}^{\infty} \exp(\lambda_0 t)\, dF_i(t)}. \end{aligned} \tag{1.41}$$

We will treat the following two cases separately.

**Case 1.** (Large Deviation zone.) $x \geq y$.

**Case 2.** (Central Limit Theorem zone.) $0 \leq y \leq x$.

We claim that in case 1 interval $I$ is so far from the center of the distribution of $S_N$ that

$$P(S_N \in I) \leq B|I|^2 \tag{1.42}$$

for some $B > 0$. In view of (1.35) this is sufficient to establish proposition.

To start the proof of (1.42) observe that

$$P(S_N \in I) \leq P(S_n \geq y) \leq \exp(-\lambda_0 y)E(\exp(\lambda_0 S_N)) = \exp(-\lambda_0 y)\Pi_{i=1}^N E(\exp(\lambda_0 X_i)). \tag{1.43}$$

Combining (1.41) and (1.43), we have

$$P(S_N \in I) \leq \Pi_{i=1}^N \left[ E(\exp(\lambda_0 X_i)) \exp\left( -\lambda_0 \frac{\int_{-\infty}^{\infty} t \exp(\lambda_0 t)\, dF_i(t)}{\int_{-\infty}^{\infty} \exp(\lambda_0 t)\, dF_i(t)} \right) \right]. \tag{1.44}$$

Since $\beta < \theta_4 < \theta_3$ (cf. (1.31) and (1.39)),

$$|\lambda_0 X_i| = \exp(\beta n) X_i \leq A_3 \exp((\beta - \theta_3)n), \tag{1.45}$$

so for all sufficiently large $n$, and each $i \leq n$

$$|\exp(\lambda_0 X_i) - 1 - \lambda_0 X_i| \leq \min(0.1\lambda_0 |X_i|,\, 0.6\lambda_0^2 X_i^2). \tag{1.46}$$

**Remark 1.** The clause requiring that $n$ be large enough means $n$ should be greater than some legitimately chosen $n_0$. Of course, we only need to establish (1.33) for $n$ which are large enough. Once this is done, just increase $A$ and $B$ to make (1.33) true for any $n$.

Recall that $E(X_i) = 0$, whence (1.46) gives

$$E(\exp \lambda_0 X_i)) \leq 1 + .6\lambda_0^2 E(X_i^2) \leq \exp(0.6(\lambda_0 X_i)^2). \tag{1.47}$$

Inequality (1.46) also yields

$$\int_{-\infty}^{\infty} t\exp(\lambda_0 t)\,dF_i(t) = E(X_i\exp(\lambda X_i)) \geq .9\lambda_0 E(X_i^2). \tag{1.48}$$

for sufficiently large $n$. Putting together (1.47) and (1.48), we get

$$\begin{aligned} E(\exp(\lambda_0 X_i))\exp&\left(-\lambda_0\frac{\int_{-\infty}^{\infty} t\exp(\lambda_0 t)\,dF_i(t)}{\int_{-\infty}^{\infty}\exp(\lambda_0 t)\,dF_i(t)}\right) \\ &\leq \exp\left(0.6\lambda_0^2 Var(X_i) - \frac{0.9\lambda_0^2 Var(X_i)}{1+0.6\lambda_0^2 Var(X_i)}\right) \\ &\leq \exp(-0.2\lambda_0^2 Var(X_i)) \end{aligned} \tag{1.49}$$

where the last inequality is true for sufficiently large $n$ because of (1.45). It follows from (1.43) and (1.49) that

$$\begin{aligned} P(S_N \in I) &\leq \Pi_{i=1}^N \exp(-0.2\lambda_0^2 Var(X_i)) \leq \exp(-0.2N\exp(2\beta n)A_2\exp(-2\theta_2 n)) \\ &\leq \exp(-0.2A_2\exp((\theta_1+2\beta-2\theta_2)n)). \end{aligned} \tag{1.50}$$

By the choice of $\beta$ (cf (1.39)), $\theta_1 + 2\beta - 2\theta_2 > 0$, whence

$$\exp(-0.2A_2\exp((\theta_1+2\beta-2\theta_2)n)) \leq B\exp(-2\theta_4(n+1)) = B|I|^2 \tag{1.51}$$

for sufficiently large $B$. Now (1.42) follows from (1.50) and (1.51). This concludes the analysis of case 1.

To begin the analysis of the second case, recall that the function

$$H_F(\lambda) = \frac{\int_{-\infty}^{\infty} t\exp(\lambda t)\,dF(t)}{\int_{-\infty}^{\infty}\exp(\lambda t)\,dF(t)}$$

is a continuous monotonically increasing function of $\lambda$, $H_F(0) = 0$, and $H_F(\lambda_0) = y$.

Therefore, in case 2, there exists $\lambda_1 \in [0, \lambda_0]$ such that

$$H_F(\lambda_1) = x. \tag{1.52}$$

For this $\lambda_1$ introduce independent random variables $Y_i$ having c.d.f. $G_i(t)$, so that

$$G_i(t) = \frac{\int_{-\infty}^{t} \exp(\lambda_1 t)\, dF_i(t)}{\int_{-\infty}^{\infty} \exp(\lambda_1 t)\, dF_i(t)}.$$

Set $R_N = \sum_{i=1}^{N} Y_i$, $G(t) = P(R_N < t)$. Our general strategy is to connect the distribution of $R_N$ to that of $S_N$, and then use Berry-Esseen inequality to compare $P(R_N \in I)$ to $P(R_N \in J)$.

Since random variables $(Y_i)_{1\le i\le N}$ are independent,

$$F(t) = \left(\int_{-\infty}^{\infty} \exp(\lambda_1 t)\, dF(t)\right)\left(\int_{-\infty}^{t} \exp(-\lambda_1 t)\, dG(t)\right), \tag{1.53}$$

(cf formula (2.11) in chapter 8 of [8]).

The connection between the distributions of $R_N$ and $S_N$ is clear from the following lemma.

**Lemma 1.** For any interval $D \subset \mathbb{R}$,

$$P(R_N \in D) \exp(-\lambda_1 \sup D) \left(\int_{-\infty}^{\infty} \exp(\lambda_1 s) dF(s)\right) \le P(S_N \in D)$$
$$\le P(R_N \in D) \exp(-\lambda_1 \inf D) \left(\int_{-\infty}^{\infty} \exp(\lambda_1 s) dF(s)\right).$$

**Proof of lemma 1.** Lemma 1 follows from the fact that

$$P(S_N \in D) = \left(\int_{-\infty}^{\infty} \exp(\lambda_1 t)\, dF(t)\right)\left(\int_{D} \exp(-\lambda_1 t) dG(t)\right)$$

which is itself a consequence of (1.53).

**Lemma 2.** For some legitimate $a_1$ and $\delta_1$

$$\frac{P(S_N \in I)}{P(S_N \in J)} \leq \frac{P(R_N \in I)}{P(R_N \in J)}(1 + a_1 \exp(-\delta_1 n)). \tag{1.54}$$

**Proof of lemma 2.** It is evident from lemma 1 that

$$\frac{P(S_N \in I)}{P(S_N \in J)} \leq \frac{P(R_N \in I)}{P(R_N \in J)} \exp(-\lambda_1(\inf I - \sup J)). \tag{1.55}$$

Because of (1.34) and the fact that $\lambda_1 < \lambda_0 = \exp(\beta n)$

$$\exp(-\lambda_1(\inf I - \sup J)) \leq \exp(2\lambda_1|I|) \leq \exp(2\exp(\beta - \theta_4)n) \tag{1.56}$$

Assumption $\beta < \theta_4$ (made in (1.39)) implies that

$$\exp(2\exp(\beta - \theta_4)n) \leq 1 + a_1 \exp(-\delta_1 n) \tag{1.57}$$

for some legitimate $a_1$ and $\delta_1$. Inequality (1.54) can be obtained by substituting (1.57) into (1.56) and then (1.56) into (1.55). Lemma 2 is proved.

Since we are going to use Berry-Esseen inequality, it is vital to have some control over the the first three moments of the $Y_i$'s. It is not hard to deduce from (1.53) that

$$G(t) = \frac{\int_{-\infty}^{t} \exp(\lambda_1 t)\, dF(t)}{\int_{-\infty}^{\infty} \exp(\lambda_1 t)\, dF(t)}.$$

Consequently,

$$E(Y_1 + \ldots + Y_n) = H_F(\lambda_1) = x. \tag{1.58}$$

Clearly, for any $i \leq N$

$$E(|Y_i|^3) \leq A_3^3 \exp(-3\theta_3 n). \tag{1.59}$$

**Lemma 3.** For any $i \geq n$, $Var(Y_i) \geq c_2 \exp(-2\theta_2 n)$, with $c_2 = A_2 \exp(-4A_3)$.

**Proof of lemma 3.** Let $Y_i'$ be a random variable independent of $Y_i$ whose c.d.f. is also $G_i(t)$. Then

$$\begin{aligned}
Var(Y_i) &= \frac{1}{2}E((Y_i - Y_i')^2) \\
&= \frac{1}{2}\frac{\int_{|u|,|v|\leq A_3\exp(-\theta_3 n)}(u-v)^2\exp(\lambda_1 u)\exp(\lambda_1 v)\,dF_i(u)dF_i(v)}{\left(\int_{-A_3\exp(-\theta_3 n)}^{A_3\exp(-\theta_3 n)}\exp(\lambda_1 t)\,dF_i(t)\right)^2} \\
&= \frac{1}{2}\frac{\int_{|u_1|,|v_1|\leq A_3\exp(-\theta_3 n)}(u_1-v_1)^2\exp(\lambda_1 u_1)\exp(\lambda_1 v_1)\,dF_i(u_1)dF_i(v_1)}{\int_{|u_2|,|v_2|\leq A_3\exp(-\theta_3 n)}\exp(\lambda_1 u_2)\exp(\lambda_1 v_2)\,dF_i(u_2)dF_i(v_2)}.
\end{aligned} \tag{1.60}$$

Since in the last of line of (1.60)

$$\max(|u_1 - u_2|, |v_1 - v_2|) \leq 2A_3\exp(-\theta_3 n),$$

and $\lambda_1 \leq \lambda_0 = \exp(\beta n)$, we get from (1.60)

$$\begin{aligned}
Var(Y_i) &\geq \exp(-4A_3\lambda_1\exp(-\theta_3 n))\frac{1}{2}\int_{|u_1|,|v_1|\leq A_3\exp(-\theta_3 n)}(u_1-v_1)^2\,dF_i(u_1)dF_i(v_1) \\
&\geq \exp(-4A_3\exp(\beta-\theta_3)n))Var(X_i) \\
&\geq (A_2\exp(-4A_3))\exp(-2\theta_2 n),
\end{aligned} \tag{1.61}$$

where the last inequality uses assumption (iii), and the fact that $\beta < \theta_4 < \theta_3$ (see (1.31) and (1.39)). Lemma 3 is proved.

With (1.59) and lemma 3 in our arsenal it is easy to find a bound on the Berry-Esseen ratio. Set $v_N = Var(R_n)$. Generate a standard normal variable $Z$.

**Lemma 4.** There exists a legitimate $C_1 > 0$ such that for any interval $[t_1, t_2]$

$$\left|P\left(\frac{R_N - x}{\sqrt{v_N}} \in [t_1,\, t_2]\right) - P\left(Z \in [t_1,\, t_2]\right)\right| \leq C_1\exp\left(-\left(\frac{\theta_1}{2} - 3(\theta_2 - \theta_3)\right)n\right). \tag{1.62}$$

**Proof of lemma 4.** It follows from lemma 3 that

$$Nc_2 \exp(-2\theta_2 n) \le v_N. \tag{1.63}$$

By Berry-Esseen inequality, (1.59), (1.63) and the definition of $RV(n, \theta_1, \dots)$, for some absolute constant $C_2$

$$\begin{aligned}
|P\left((R_N - x)v_N^{-1/2} \in [t_1, t_2]\right) - P(Z \in [t_1, t_2])| \\
&\le \frac{C_2 \sum_{i=1}^{N} E(Y_i^3)}{v_N^{3/2}} \\
&\le \frac{NA_3^3 \exp\left(-3\theta_3 n\right)}{(Nc_2 \exp(-2\theta_2 n))^{3/2}} \\
&\le \frac{C_2 A_3^3 \exp(3(\theta_2 - \theta_3))n}{\sqrt{c_2}^3 \sqrt{N}} \\
&\le C_1 \exp\left(-\left(\frac{\theta_1}{2} - 3(\theta_2 - \theta_3)\right) n\right).
\end{aligned}$$

(Here $C_1 = C_2 \frac{A_3^3}{\sqrt{c_2^3}}$.) Lemma 4 is proved.

Next we estimate $P(R_N \in I)$ and $P(R_N \in J)$. To this end translate and dilate intervals $I$ and $J$: $I_1 = \frac{I-x}{\sqrt{v_N}}$, $J_1 = \frac{J-x}{\sqrt{v_N}}$.

**Lemma 5.** For some legitimate $a_2$ and $\delta_2$

$$P(R_N \in I) \le \frac{|I|}{\sqrt{2\pi v_N}}(1 + a_2 \exp(-\delta_2 n)). \tag{1.64}$$

**Proof of lemma 5.** Since the density of the standard normal distribution is bounded by $\sqrt{2\pi}^{-1}$, lemma 4 gives

$$\begin{aligned}
P(R_N \in I) &\le P(Z \in I_1) + C_1 \exp\left(-\left(\frac{\theta_1}{2} - 3(\theta_2 - \theta_3)\right) n\right) \\
&\le \frac{|I|}{\sqrt{2\pi v_N}} + \exp\left(-\left(\frac{\theta_1}{2} - 3(\theta_2 - \theta_3)\right) n\right).
\end{aligned} \tag{1.65}$$

Note that the first term in the last line can be bounded below as follows

$$\frac{|I|}{\sqrt{2\pi v_N}} \geq b_1 \exp((\theta_3 - \theta_4 - \theta_1/2)n) \text{ with } b_1 = \frac{\exp(-\theta_4 - \theta_1)}{\sqrt{2\pi} A_3}. \tag{1.66}$$

Assumption (1.31) ensures that

$$\theta_3 - \theta_4 - \theta_1/2 > -(\theta_1/2 - 3(\theta_2 - \theta_3)). \tag{1.67}$$

Therefore, in view of (1.66) estimate (1.65) can be simplified to the form

$$P(R_N \in I) \leq \frac{|I|}{\sqrt{2\pi v_N}}(1 + a_2 \exp(-\delta_2 n)) \tag{1.68}$$

for some legitimate $a_2$ and $\delta_2$. Lemma 5 is proved.

**Lemma 6.** For some legitimate choice of $a_3$ and $\delta_3$

$$P(R_N \in J) \geq \frac{|J|}{\sqrt{2\pi v_N}}(1 - a_3 \exp(-\delta_3 n)). \tag{1.69}$$

**Proof of lemma 6.** Due to (1.38), $J_1$ is contained within $2|I_0|{v_N}^{-1/2}$ neighborhood of zero. Inequality (1.63) implies

$$2|I_0|{v_N}^{-1/2} \leq \frac{2}{\sqrt{c_2}} \exp((-\theta_4 - \theta_1/2 + \theta_2)n)$$

Since $-\theta_4 - \theta_1/2 + \theta_2 < 0$ (see (1.32)) and the density of the standard normal distribution, call it $\phi(z)$, is smooth,

$$\phi(z) \geq \frac{1}{\sqrt{2\pi}}(1 - a_4 \exp(-\delta_4 n)) \text{ for } z \in J_1, \tag{1.70}$$

where $a_4$ and $\delta_4$ are chosen legitimately. It follows from lemma 4 and (1.70) that

$$\begin{aligned} P(R_N \in J) &\geq P(Z \in J_1) - C_1 \exp\left(-\left(\frac{\theta_1}{2} - 3(\theta_2 - \theta_3)\right) n\right) \\ &\geq \frac{|J|}{\sqrt{2\pi v_N}}(1 - a_4 \exp(-\delta_4 n)) - C_1 \exp\left(-\left(\frac{\theta_1}{2} - 3(\theta_2 - \theta_3)\right) n\right) \\ &\geq \frac{|J|}{\sqrt{2\pi v_N}}(1 - a_3 \exp(-\delta_3 n)) \end{aligned} \tag{1.71}$$

for some legitimate $c_3$ and $\delta_3$. The last inequality in (1.71) was obtained using (1.66), (1.67) and the fact that $|J| \approx 2|I|$. Lemma 6 is proved.

It can be inferred from lemmas 5 and 6 and inequality (1.37) that

$$\frac{P(R_N \in I)}{P(R_N \in J)} \leq \frac{\frac{|I|}{\sqrt{2\pi v_N}}(1 + a_2 \exp(-\delta_2 n))}{\frac{|J|}{\sqrt{2\pi v_N}}(1 - a_3 \exp(-\delta_3 n))} \leq \frac{1 + a_5 \exp(-\delta_5 n)}{2} \tag{1.72}$$

for some legitimate $a_5$ and $\delta_5$.

It remains to combine (1.36) with lemma 2 and (1.72) to establish the proposition in case 2.

## 1.7 Absolute continuity of a measure under a scaling condition.

**Proposition 10** *Suppose $B$, $\eta$, $\theta$, $n_0$ and $\theta_4$ are positive numbers. Assume that a probability measure $\mu$ is supported on the interval $[0, 1]$ and for any $n \geq n_0$ and any pair of intervals $(I, J) \in IL(n, \theta_4)$*

$$\mu(I) \leq \frac{1 + \exp(-\theta n)}{2} \mu(J) + B|I|^{1+\eta}. \tag{1.73}$$

*Then $\mu$ is absolutely continuous with respect to Lebesgue measure. Its Radon-Nikodim derivative with respect to Lebesgue measure is bounded almost everywhere.*

**Remark 1.** The proof below is straightforward but a little tedious. Appendix A contains a much more elegant proof due to Daniel Stroock.

**Proof of proposition 10.** Given any interval $I_0 \subset [0, 1]$ of length less than $1/2\exp(-\theta_4 n_0)$, one can construct an interval $I_1$ such that $|I_1| = 2|I_0|$ and $I_0 \subset I_1 \subset [0,1]$. If the length of $I_1$ is also less than $1/2\exp(-\theta_4 n_0)$, construct $I_2$ such that $|I_2| = 2|I_1|$ and $I_0 \subset I_1 \subset I_2$. Iterate this procedure until we get some interval $|I_k|$ of length between $1/2\exp(-\theta_4 n_0)$ and $\exp(-\theta_4 n_0)$. We now point out for the future reference that

$$|I_l| = \frac{1}{2^{k-l}}|I_k| \leq 2^{l-k}\exp(-\theta_4 n_0). \tag{1.74}$$

We claim for every integer $l$ such that $0 \leq l \leq k-1$, there exists $m_l \geq n_0$ such that

$$(I_l,\, I_{l+1}) \in IL(m_l,\, \theta_4)$$

In fact, conditions (vi) and (vii) from the definition of the classes $IL(n,\, \theta_4)$ are obvious. To address (viii), set $J_n = [e^{-\theta_4(n+1)}\, e^{-\theta_4 n})$ $(n \geq n_0)$. Clearly,

$$\cup_{n=n_0}^{\infty} J_n = (0,\, \exp(-\theta_4 n_0))$$

Since $|I_l| < \exp(-\theta_4 n_0)$, $|I_l| \in J_{m_l}$ for some $m_l$. Condition (vii) is satisfied for this choice of $m_l$.

The assumption (1.73) yields for any $l$

$$\mu(I_l) \leq \frac{(1+\exp(-\theta n_l))\mu(I_{l+1})}{2} + B|I_l|^{1+\eta}. \tag{1.75}$$

Set $d_l = \frac{\mu(I_l)}{|I_l|}$. Then (1.75) is easily seen to imply

$$\frac{d_l}{d_{l+1}} \leq 1 + \exp(-\theta m_l) + B\exp(-\eta m_l). \tag{1.76}$$

Multiplying (1.76) for all $l$ from 0 through $k-1$ and noting that $d_k \leq 2\exp(\theta_4 n_0)$, because $|I_k| \geq \frac{1}{2}\exp(-\theta_4 n_0)$, we obtain

$$d_0 \leq 2\exp(\theta_4 n_0)\Pi_{l=0}^{k-1}(1+\exp(-\theta m_l)+B\exp(-\eta m_l)). \tag{1.77}$$

Recalling (1.74) and (viii) from the definition of $IL(n, \theta_4)$, write

$$\exp(-\theta_4(m_l+1)) \leq |I_l| \leq 2^{l-k}\exp(-\theta_4 n_0). \tag{1.78}$$

It follows from (1.78) that

$$\exp(-\theta m_l) \leq (2^{(l-k)\frac{\theta}{\theta_4}}). \tag{1.79}$$

Similarly,

$$\exp(-\eta m_l) \leq (2^{(l-k)\frac{\eta}{\theta_4}}). \tag{1.80}$$

Then the change of variables $m = k - l$ in inequality (1.77) leads to

$$d_0 \leq 2\Pi_{m\geq 0}\left(1+2^{-m\frac{\theta}{\theta_4}}+B2^{-m\frac{\eta}{\theta_4}}\right). \tag{1.81}$$

The product in the right-hand side of (1.79) converges to some number, say $C$, which is independent of the choice of $|I_0|$. Thus, for any interval $I_0 \subset [0,1]$,

$$\mu(I_0) \leq C|I_0|.$$

This inequality implies the proposition.

## 1.8 Proof of the theorem.

Referring to $\alpha$ in section 3 choose $\rho > 0$ so small that

$$3k_0^{3\rho}\exp(-3(1-\rho)\alpha) < \exp(-2.5\alpha) \text{ and} \tag{1.82}$$

$$\left(\frac{5}{6}\right)^{1-\rho} < \exp(-2\alpha\rho). \tag{1.83}$$

Then choose $\theta_1 = \theta_1(\rho)$ as in proposition 4. Next pick $\sigma > 0$ in such a way that

$$3k_0^{3\rho} \exp(-3(1-\rho)(\alpha-\sigma)) < \exp(-2.5\alpha), \tag{1.84}$$

$$10\sigma < \alpha, \tag{1.85}$$

$$32\sigma < \theta_1 \quad \text{and} \tag{1.86}$$

$$3\sigma < \alpha\rho. \tag{1.87}$$

Set $\theta_2 = 2(\alpha + \sigma)$, $\theta_3 = 2(\alpha - \sigma)$ and pick any $\theta_4 > 0$ in the open interval

$$\left(\theta_2 - \frac{\theta_1}{2},\, \theta_3 - 3(\theta_2 - \theta_3)\right).$$

The interval under consideration is non degenerate due to relation (1.86). It contains positive numbers because of (1.85). Pick any $\theta_5 > 2\theta_4$. Let $c_1$, $C_1$, and $D$ be as in proposition 4. Find $n_0 \geq 2$ and $c_2 > 0$ such that for any $n \geq n_0$, $\rho n > 1$ and

$$c_2 \exp(-(\alpha+\sigma)n) \leq c_1 \exp(-(\alpha+\sigma)n) - \frac{6C_1}{5} \exp(-(\alpha+2\sigma)n). \tag{1.88}$$

Pick an arbitrary $n \geq n_0$ and set $m = [\rho n]$. Then take any pair of intervals

$$(I,\, J) \in IL(n,\, \theta_4).$$

with some $n \geq n_0$. Let $\bar{V} \subset \mathbb{N}^\infty$ be the set of all sequences $(n_i)_{i\geq 1} \in \mathbb{N}^\infty$ such that $n_l = 1$ for all integer $l \in [m, n-1]$.

Without loss in generality we may assume that the basic probability space $\Omega$ on which random variables $\xi_v$ are defined is a product of two (independent) probability spaces $\Omega_1$ and $\Omega_2$ with measures $P_1$ and $P_2$ respectively. It can be further assumed that $\xi_v$ with $v \in \bar{V}$ are defined on $\Omega_1$, whereas all other $\xi_v$'s are defined on $\Omega_2$. In this framework, Galton-Watson tree could be viewed as a map $T$ from $\Omega_1 \times \Omega_2$ into the ensemble of admissible trees. Then $Q_T$, the probability of no return to the origin on $T$, is a measurable function on $\Omega_1 \times \Omega_2$.

By Fubini's theorem,

$$P(Q_T \in I) = \int_{\Omega_2} P_1(Q_T(\omega_1, \omega_2) \in I)\, dP_2(\omega_2).$$

Let us break up $\Omega_2$ into a disjoint union of two sets,

$$\Omega_2 = CMMN \cup RARE.$$

A sample point $\omega_2 \in CMMN$ if for some $\omega_1$, for the realization of Galton-Watson tree $T(\omega_1, \omega_2)$

$$\#GV(c_1, C_1, m, n, \sigma) \geq \exp(\theta_1 n). \tag{1.89}$$

Otherwise we say $\omega_2 \in RARE$.

By proposition 4 and the choice of $\theta_5$,

$$\begin{aligned} \int_{RARE} P_1(Q_T(\omega_1, \omega_2) \in I)\, dP_2(\omega_2) &\leq D\exp(-\theta_5 n) \\ &\leq D(\exp(-\theta_4 n))^2 \\ &\leq D_1 |I|^2, \end{aligned} \tag{1.90}$$

where $D_1 = D\exp(2\theta_4)$. Take an arbitrary $\omega_2 \in CMMN$. Pick any $\omega_1 \in \Omega_1$ so that (1.89) holds. Consider a tree $T_1 = T(\omega_1, \omega_2)$. For the sake of convenience denote $V_0 = GV(c_1, C_1, m, n, \sigma)$ in $T_1$. Choose any $V_1 \subset V_0$, such that

$$\exp(\theta_1 n) \leq \#V_1 \leq 2\exp(\theta_1 n), \tag{1.91}$$

Without loss in generality we may assume that $(\Omega_1, P_1)$ is a direct product of probability spaces $(\Omega_3, P_3)$ and $(\Omega_4, P_4)$ so that $\xi_w$ is defined on $\Omega_3$ if $w$ is a descendant of a vertex of the form $pd(v, n-m)$ with $v \in V_1$. However, if $w$ is a descendant of a vertex of the form $pd(v, n-m)$ with $v \in L_m \setminus V_1$, then $\xi_w$ is defined on $\Omega_4$. Then a generic element $\omega_1 \in \Omega_1$ can be written as a pair $\omega_1 = (\omega_3, \omega_4)$.

Let $T_2$ be some other admissible tree subject that coincides with $T_1$ such that,

$$T_2 \cap (\mathbb{N}^\infty \setminus \bar{V}) = T_1 \cap (\mathbb{N}^\infty \setminus \bar{V})$$

and $v_0 \in V_0$. Then by proposition 2 and (1.10)

$$\begin{aligned}
|P_{T_1}(rt \to pd(v_0, n-m)) - P_{T_2}(root \to pd(v_0, n-m))| \\
&\leq \left(\frac{5}{6}\right)^{n-m-1} P_{T_1^{v_0}}(v_0 \to pd(v_0, n-m)) \\
&\leq C_1 \left(\frac{5}{6}\right)^{n(1-\rho)-1} e^{-(\alpha-\sigma)(n-m)} \quad \text{use (1.83)} \\
&\leq \frac{6C_1}{5} \exp(-2\alpha\rho n) e^{-(\alpha-\sigma)(n(1-\rho))} \\
&\leq \frac{6C_1}{5} \exp((-(\alpha-\sigma)-\rho\alpha)n) \text{ use (1.87)} \\
&\leq \frac{6C_1}{5} \exp(-(\alpha+2\sigma)n)
\end{aligned}$$

The last line together with (1.88) and (1.9) in section 4 yields

$$c_2 \exp(-(\alpha+\sigma)n) \leq P_{T_2}(root \to pd(v_0, n-m)) \leq 2C_1 \exp(-(\alpha-\sigma)n) \qquad (1.92)$$

Obviously, $P_{T^{v_0}}(v \to pd(v, n-m))$ is the same in $T_1$ and $T_2$. Therefore, inequalities (1.92) imply that

$$v_0 \in GV(c_2, 2C_1, m, n, \sigma)$$

in $T_2$.

Fix $\omega_4 \in \Omega_4$. Our next goal is to apply proposition 6 to $T = T((\cdot, \omega_4), \omega_2)$, a random tree on $\Omega_3$ and $V = \{pd(v, n-m) \,|\, v \in V_1\}$. Compare the following properties with the stipulations made in section 6.

(i) It can be seen from (1.91) that

$$\#V = \#V_1 \in [\exp(\theta_1 n),\, 2\exp(\theta_1)n]$$

by the choice of $V_1$.

(ii) Random variables $Q_{T_v} - Q_T$ are mutually independent, because they are determined by disjoint collections of $\xi_v$'s.

(iii) Proposition 8, formula (1.9) and self-similarity of Galton-Watson trees give

$$Var(Q_{T_v} - Q_{T_0}) \geq \frac{c_2^4 \exp(-4(\alpha+\sigma)n) C_{GW}}{k_0^2}$$

where $C_{GW}$ is $C_G$ from Proposition 8 corresponding to $G$ being a Galton-Watson tree.

(iv) Proposition 7 and (1.9) for $GV(c_2,\, 2C_1,\, m,\, n,\, \sigma)$ imply that for $v \in V$

$$0 \leq Q_{T_v} - Q_{T_0} \leq 3k_0 C_1 \exp(-2(\alpha - \sigma)).$$

(v) If we set

$$Y = (Q_T - Q_{T_0}) - \sum_{v \in V_1} (Q_{T_v} - Q_{T_0}),$$

then propositions 5 and 6 would give

$$\begin{aligned} |Y| \leq P(A_{2,v})\#V &\leq (3C_1^3 k_0^{2m+1} \exp(-3(n-m)(\alpha-\sigma))) k_0^m \\ &\leq (k_0 C_1^3) 3 k_0^{3\rho n} \exp(-3(1-\rho)(\alpha-\sigma)n) \\ &\leq k_0 C_1^3 \exp(-2.5\alpha n) \leq k_0 C_1^3 \exp(-2(\alpha+\sigma)n), \end{aligned}$$

the last two inequalities being consequences of (1.84) and (1.85).

Analysis of the items (i) through (v) shows that

$$Q_T((\cdot, \omega_4),\, \omega_2) - \text{constant} \in RV\left(n,\, \theta_1,\, \frac{c_2^4 C_{GW}}{k_0^2},\, \theta_2,\, \max(k_0 C_1^3, 3k_0 C_1),\, \theta_3\right).$$

Proposition 9 now implies that for a suitable choice of positive numbers $A$, $B$ and $\theta$

$$P_3(Q_T((\cdot,\omega_4),\omega_2)\in I)\leq\frac{(1+A\exp(-\theta n))}{2}P_3(Q_T((\cdot,\omega_4),\omega_2)\in J)+B|I|^2. \quad (1.93)$$

Integrating (1.93) over $\Omega_4$ with respect to $P_4$ while keeping $\omega_2$ fixed, we obtain

$$P_1(Q_T(\cdot,\omega_2)\in I)\leq\frac{(1+A\exp(-\theta n))}{2}P_1(Q_T(\cdot,\,\omega_2)\in J)+B|I|^2$$

for any $\omega_2\in CMMN$. Next integrate over CMMN

$$\int_{CMMN}P_1(Q_T(\omega_1,\,\omega_2)\in I)\,dP_2(\omega_2)\leq\frac{(1+A\exp(-\theta n))}{2}\times$$
$$\times\int_{CMMN}P_1(Q_T(\omega_1,\,\omega_2)\in J)dP_2(\omega_2)+B|I|^2\;\Rightarrow$$

$$\int_{CMMN}P_1(Q_T(\omega_1,\,\omega_2)\in I)\,dP_2(\omega_2)\leq\frac{(1+A\exp(-\theta n))}{2}P(Q_T\in J)+B|I|^2. \quad (1.94)$$

Finally, add (1.90) and (1.94) to get

$$\begin{aligned}P(Q_T\in I)&=\int_{RARE}P_1(Q_T(\omega_1,\,\omega_2)\in I)\,dP_2(\omega_2)\\&\quad+\int_{CMMN}P_1(Q_T(\omega_1,\omega_2)\in I)\,dP_2(\omega_2)\\&\leq D|I|^2+\frac{(1+A\exp(-\theta n))}{2}P(Q_T\in J)+B|I|^{1+\eta}\\&\leq\frac{(1+A\exp(-\theta n))}{2}P(Q_T\in J)+B'|I|^2,\end{aligned}$$

with $B'=B+D_1$.

The assumptions of proposition 10 are satisfied for

$$\mu(A)=(P_1\times P_2)(Q_T(\omega_1,\,\omega_2)\in A).$$

Proposition 10 applies to show that $Q_T$ is absolutely continuous with bounded density.

# Chapter 2

# The boundary of a Galton-Watson tree.

## 2.1 Introduction

Consider an infinite rooted tree $T$. A *ray* in $T$ is a sequence of vertices $v_0 = rt$, $v_1$, $v_2, \ldots$ such that for any $i \geq 0$ $v_i$ is the parent of $v_{i+1}$. The set of all rays in $T$ is called the *boundary* of the tree $T$ and is denoted $\delta T$. It can be turned into a metric space by defining the distance between the rays $(v_i)_{i\geq 0}$ and $(w_i)_{i\geq 0}$ to equal

$$dist((v_i)_{i\geq 0}, (w_i)_{i\geq 0}) = \exp(-j), \text{ where } j = \max\{i : v_i = w_i\}. \tag{2.1}$$

Considering the path connecting a vertex $v \neq rt$ in an infinite tree $T = (V, E)$ to the root of the tree, define $\Delta_k(v)$ to be the fraction of the nodes with $k$ children in the path. More formally,

$$\Delta_k(v) = \frac{\#|\{w \in V,\ deg(w) = k+1,\ w \text{ disconnects } v \text{ from } rt\}|}{|v|}.$$

Given a Galton-Watson tree $T$, and a sequence $r = (r_k)_{k\geq 1}$ of non-negative numbers that add up to 1, let $A_r$ be the set of rays $(v_i)_{i\geq 0}$ in $\delta T$ such that

$$\lim_{i\to\infty} \Delta_k(v_i) = r_k.$$

In this chapter we are computing the Hausdorff dimension of $A_r$.

Roughly speaking we would like to know whether it is possible to find a ray in $T$, in which the concentration of vertices with $k$ children is $r_k$ for any $k$. If such rays exist, we would like to be able to tell how many of them are present in the tree.

Of course, the quantity in question depends on the parameters of the Galton-Watson tree. Set $p_k = P(\xi_{rt} = k)$, $m_0 = E(\xi_{rt})$, $q_k = \frac{kp_k}{m_0}$. We will be assuming that $p_0 = 0$ and $m_0 < \infty$.

It was proved in [3] that the Hausdorff dimension of $\delta T$ is equal to $\log m_0$ with probability 1. The following theorem describes the typical behavior of concentration of vertices with given numbers of children along the rays of $\delta T$.

**Theorem 2** *For almost every ray (in the sense of Hausdorff measure $H_{\log m_0}$) $(v_i)_{i\geq 0} \in \delta T$ in almost every Galton-Watson tree $T$*

$$\lim_{i\to\infty} \Delta_k(v_i) = q_k \quad \textit{for all } k.$$

Thus, $A_{(q_k)}$ is the thickest of all the sets $A_r$. It is natural to expect that the closer is the sequence $(r_k)_{k\geq 1}$ to $(q_k)_{k\geq 1}$, the higher the dimension of $A_r$ is. The exact statements will be made in terms of the *relative entropy* also known as the *Kullback-Leibler distance* between distributions. We will only need this metric in the case of integer-valued random variables. Therefore, for our purposes it suffices to define it as a distance between sequences of non-negative numbers. Let $x = (x_k)_{k\geq 1}$ and $y = (y_k)_{k\geq 1}$ be two sequences of non-negative real numbers that add up to 1 such that $y_k = 0$ whenever $x_k = 0$. Then define the *relative entropy* of $y$ with respect to $x$ to be

$$H(y|x) = \sum_{k\geq 1\,:\,y_k\neq 0}^{\infty} y_k \log\left(\frac{y_k}{x_k}\right).$$

The theorem below is the main result of this chapter.

**Theorem 3** *Referring to the preceding setup, assume that $r_k = 0$ whenever $p_k = 0$. Then for almost every tree $T$ in the sense of Galton-Watson measure*

$$dim(A_r) = \max(\log m_0 - H(r|q), 0).$$

One possible application for the theorem would be to obtain inequalities relating the dimension and the speed of the biased random walks on Galton-Watson trees. (See [5] and [6] for the relevant definitions and interesting results.) Some natural estimates of this sort were conjectured in [7] and proved in [9], but the method based on Theorem 3 should produce different bounds.

The following paragraph describes the breakdown of the chapter into sections. Section 3 contains the proof of theorem 2. The proof uses a simple lemma on Hausdorff presented at the end of section 2. Section 2 also explains how to convert the information about deterministic trees into bounds on Hausdorff dimension of its boundary. Unfortunately, it imposes stringent requirement on the trees it treats, namely, the trees have to grow fast everywhere. This is why we need section 4 explaining how to select a subtree of a random tree that grows almost as fast as the original tree while doing it absolutely everywhere. Sections 6 and 7 give some upper and lower bounds for the number of small balls in $\delta T$ necessary to cover the finite analogs of $A_r$. The strategy there is to use Sanov's theorem to estimate the expected number of the balls in the coverings. By Markov' inequality, the bound for the expectation will imply similar bounds for individual trees. It is possible to use the homogeneity property of Galton-Watson tree to prove the lower bound as well. Section 5 lays the groundwork for sections 6 and 7. Namely, it proves a certain independence statement that makes Sanov's theorem relevant in subsequent sections. Propositions 19 and 20 proved in section 8 are equivalent to the theorem.

**Remark.** It was brought to author's attention that a similar result in slightly different setting was proved in [4].

## 2.2 Estimates for Hausdorff dimension.

The main purpose of this section is to prove a lower bound for the Hausdorff dimension of the boundary of an infinite deterministic tree $T$ that in some sense grows uniformly. This means $T$ does not have pockets of slow growth.

At the end of the section we establish a lower bound for the dimension of a subset of a general metric set in terms of the size of its $\epsilon$-net. It will follow from a sufficient condition that ensures that the Hausdorff measure of certain dimension is zero for some sets. Statements of this type are as plentiful as mushrooms, because it is easier to prove one than to look it up.

Consider an infinite rooted tree $T$ and its boundary $\delta T$. (The relevant definition are included into the introduction to this chapter.) If $x = (v_i)_{i \geq 0} \in \delta T$, then $\bar{B}_r(x)$, the closed ball of radius $r$ centered at $x$, consists of all rays $(w_i)_{i \geq 0}$ that agree with $x$ up to the vertex $v_{[-\log r]}$. Thus, $\bar{B}_r(x)$ is precisely the set of rays going through $v_{[-\log r]}$. We will say that $v$ is the *pivotal* vertex for $B_r(x)$.

It is not hard to see that $\delta T$ is compact.

Recall the notions of Hausdorff measure and dimension. Given a subset $A$ of a metric space $X$ and a positive number $\alpha$, the Hausdorff measure of dimension $\alpha$ is given by

$$H_\alpha(A) = \lim_{\epsilon \to 0} \left( \inf \{ \sum_{i=1}^{\infty} r_i^\alpha \ : \ A \subset \cup B_{r_i}(x_i) \text{ for some } x_i \in X \text{ and } r_i < \epsilon \text{ for all } i \} \right)$$

Then the Hausdorff dimension of $A$ is defined by the relation

$$dim(A) = \inf\{\alpha \ : \ H_\alpha(A) = 0\}.$$

**Remark 1.** For any $\beta > dim(A)$ and any $\epsilon > 0$, there exists a covering of $A$ by balls $B_{r_i}(x_i)$ with radii $r_i < \epsilon$ centered at $x_i$'s, such that

$$\sum_i r_i^\beta < 1.$$

The following proposition derives a lower bound on Hausdorff dimension of $\delta T$ for a tree $T$ whose behavior can be controlled only at certain levels $L_{S_i} = \{v : |v| = S_i\}$ .

**Proposition 11** *Suppose $(S_i)_{i\geq 0}$ is a strictly increasing sequence of integers with $S_0 = 0$ such that*

$$\lim_{i\to\infty}(S_{i+1} - S_i) = \infty$$

*and*

$$S_{i+1} - S_i = o(S_i) \qquad as\ 1 \to \infty. \tag{2.2}$$

*Assume that in a rooted tree $T = (V, E)$ for any $i$ and any vertex $v \in L_{S_i}$*

$$\#(T^v \cap L_{S_{i+1}}) \geq c\exp(\alpha(S_{i+1} - S_i)) \tag{2.3}$$

*for some positive numbers $c$ and $\alpha$ independent of $i$ and $v$. Then $dim(\delta T) \geq \alpha$.*

**Proof of proposition 11.** Pick any positive number $\beta < \alpha$. We will show that

$$dim(\delta T) \geq \beta, \tag{2.4}$$

and the proposition will follow, because $\beta$ is arbitrary.

We will argue by contradiction. Assume that $\beta > dim(\delta T)$. Set $\gamma = \frac{\alpha+\beta}{2}$. By (2.3), the generation $L_{S_1}$ has at least $c\exp(\alpha_{S_1})$ vertices. Each of these vertices has at least $c\exp(\alpha(S_2 - S_1))$ descendants in the generation $L_{S_2}$, hence $\#L_{S_2} \geq c^2\exp(\alpha S_2)$. Continuing this argument, we see that

$$\#L_{S_i} \geq c^i\exp(\alpha S_i) = \exp(\alpha S_i + i\log c).$$

It is clear from our assumptions on $S_i$, that $\frac{S_i}{i} \to \infty$ as $i \to \infty$. Hence, there exists $i_0$ such that $|L_{S_{i_0}}| \geq \exp(\gamma i_0)$. Without loss in generality, $i_0$ can be chosen so large that for any $i \geq i_0$

$$c\exp((\alpha-\gamma)(S_{i+1}-S_i)) \geq 2, \tag{2.5}$$

$$S_i\beta < S_{i-1}\gamma - 1. \tag{2.6}$$

By remark 1 and our assumption $\beta > dim(\delta T)$, there exists a collection of open balls $B_{r_i}(x_i)$ with $r_i < \exp(-i_0)$ whose union covers the entire space $\delta T$, and

$$\sum_i r_i^\beta < 1. \tag{2.7}$$

By compactness of $\delta T$, we may assume this collection of balls is finite. Let $v_i$ be the pivotal vertex for $B_{r_i}(x_i)$. Since $r_i < \exp(-i_0)$, $|v_i| > i_0$. Define $V_1 = \cup_i\{v_i\}$.

Let $F_i$ be the set of all $v \in L_{S_i} \setminus V_1$ that don't have predecessors in $V_1$. We will prove by induction on $j$ that for $j \geq i_0$

$$\#F_j \geq \exp(\gamma S_j). \tag{2.8}$$

If $j = i_0$, then $F_i = L_{S_i}$, because for any $i$ $|v_i| > i_0$. Therefore, (2.8) is true by the choice of $i_0$.

Next assume that (2.8) has been established for all $j < h$ where $h$ is some integer greater than $i_0$. Set

$$V_2 = V_1 \cap (\cup_{k=S_{h-1}+1}^{S_h} L_k),$$

$$W = \{w \in F_{h-1} : w \text{ has no descendants in } V_2\}.$$

For any $v_i \in V_2$, the corresponding $r_i$ is not less than $\exp(-S_h)$. It follows from (2.7) that

$$(\#V_2)\exp(-\beta S_h) \leq \sum_{v_i \in V_2} r_i^\beta < 1 \;\Rightarrow\; \#V_2 \leq \exp(\beta S_h)$$

Since each element of $V_2$ has one predecessor in $F_{h-1}$, the assumption of induction for

$j = h - 1$ gives

$$\#W \geq \#F_{h-1} - \exp(\beta S_h) \geq \exp(\gamma S_{h-1}) - \exp(\beta S_h). \tag{2.9}$$

Inequality (2.6) gives a lower bound for the right-hand side of (2.9), and we write

$$\#W \geq (1 - e^{-1}) \exp(\gamma S_{h-1}). \tag{2.10}$$

Observe that for any $w \in W$,

$$T^w \cap L_{S_h} \subset F_h. \tag{2.11}$$

Indeed, if $w_1 \in T^w \cap L_{S_h}$, then any $w_2$, a predecessor of $w_1$, is either a predecessor of $w$ or a descendant of $w$. By the definition of $F_{h-1}$ the predecessors of $w$ are not in $V_1$. If $w_2$ is a descendant of $w$, then

$$S_{h-1} = |w| < |w_2| < |w_1| = S_h.$$

Therefore, if $w_2$ were in $V_1$, it would also be in $V_2$ which would contradict $w \in W$. Hence $w_1$ has no predecessors in $V_1$. The inclusion (2.11) is verified.

As a consequence of (2.11), we get

$$(\cup_{w \in W} T^w) \cap L_{S_h} \subset F_h$$

The union in the left-hand side of the last line is disjoint. Using (2.3) and (2.10), conclude

$$\begin{aligned}
\#F_h &\geq c(1 - e^{-1}) \exp(\alpha(S_h - S_{h-1})) \exp(\gamma S_{h-1}) \\
&\geq c \exp(\gamma S_h) e^{(\alpha-\gamma)(S_h - S_{h-1})} (1 - e^{-1}) \\
\text{use (2.5)} \quad &\geq \exp(\gamma S_h),
\end{aligned}$$

which proves inequality (2.8) for $j = h$.

On the other hand, the fact that

$$\delta T \subset \cup B_{r_i}(x_i)$$

means any ray in $\delta T$ goes through one of the $v_i$'s. Since $V_1$ is finite, there exists $S_k > \max\{|v_i| \mid v_i \in V_1\}$, and we may also assume $k > i_0$. Then each vertex $w \in L_{S_k}$ has to have a predecessor in $V_1$, because any ray through such a vertex has to contain an element $v \in V_1$, and this $v$ cannot be a descendant of $w$, since $|v| < S_k = |w|$. However, nodes in $F_{S_k}$ have no predecessors in $V_1$ and $|F_{S_k}| \geq \exp(\gamma S_k) > 0$. The contradiction just obtained proves (2.4).

**Proposition 12** *Let $A_N$ with $N \geq 1$ be subsets of a metric space $X$. Assume that each $A_N$ is a union of $h_N$ balls of radii* $\exp(-N)$. *If for some positive number $\alpha$,*

$$\sum_{N=1}^{\infty} h_N \exp(-\alpha N) < \infty,$$

*then $H_\alpha(\limsup_{N\to\infty} A_N) = 0$, where as usual*

$$\limsup_{N\to\infty} A_N = \cap_{i=1}^{\infty} \cup_{j=i}^{\infty} A_j.$$

**Proof of proposition 12.** Take an arbitrary positive integer $n$. Note that

$$\limsup_{N\to\infty} A_N \subset \cup_{N\geq n} A_N$$

is a covering of $\limsup_{N\to\infty} A_N$ by a countable number of balls $B_{r_i}(x_i)$. Among them, exactly $h_N$ balls have radii $\exp(-N) \leq \exp(-n)$. Therefore,

$$H_\alpha(\limsup_{n\to\infty} A_N) \leq \lim_{n\to\infty} \left( \sum_{r_i \leq \exp(-n)} r_i^\alpha \right) \leq \lim_{n\to\infty} \sum_{N\geq n} h_N \exp(-\alpha N) = 0.$$

Proposition 12 is proved.

**Proposition 13** *Let $A_N \subset X$ and $h_N$ $(N \geq 1)$ be as in proposition 12. Then*

$$dim(\limsup_{N\to\infty} A_N) \leq \limsup_{N\to\infty} \frac{\log h_N}{N}. \tag{2.12}$$

**Proof of proposition 13.** Choose an arbitrary $\beta > \limsup \frac{\log h_N}{N}$. Then the series

$$\sum_N h_N \exp(-\beta N)$$

converges. Therefore, by proposition 2 $H_\beta(\limsup_{N\to\infty} A_N) = 0$.

Since the choice of $\beta > \limsup \frac{\log h_N}{N}$ was arbitrary, inequality (2.12) is true.

## 2.3 Analysis of typical rays.

This section is devoted to the proof of theorem 2. The idea here is to find exactly the expected value of some random variable closely related to the "raywise" concentration of vertices with given numbers of children. This value will provide us with all the information we need to establish the theorem.

Fix $J$, an arbitrary collection of positive integers. Define

$$U_J(v) = \#\{w \in V,\ deg(w) - 1 \in J,\ 0 \leq |w| < |v|,\ w \text{ disconnects } v \text{ from } rt\}.$$

For our purposes it will be essential to consider random variable

$$X_n = \sum_{v\,:\,|v|=n} \exp(\alpha U_J(v)),$$

where $\alpha$ is some positive number to be chosen later.

**Lemma 1.** $E(X_n) = \left(m_0 + \left(\sum_{k\in J} kp_k\right)(\exp(\alpha) - 1)\right)^n$.

**Proof of lemma 1.** Lemma 1 is based on the following simple observation. If $v$

is a child of $w$, then

$$U_J(v) - U_J(w) = \begin{cases} 1, & \text{the number of children of } w \text{ is in } J \\ 0, & \text{otherwise} \end{cases}$$

Consequently, for any vertex $w$ in a Galton-Watson tree

$$E\left(\sum_{v \text{ is a child of } w} \exp(\alpha(U_k(v) - U_k(w)))\right) = m_0 + \left(\sum_{k\in J} kp_k\right)(\exp(\alpha) - 1), \quad (2.13)$$

and (2.13) holds even if we condition the expectation involved on the development of the tree before vertex $w$ appears.

For a Galton-Watson tree $T$, let $T_{n-1}$ be its finite subtree obtained by cutting off the vertices of the $n$-th and higher generations. Then

$$\begin{aligned} E(X_n) &= E(E(X_n \mid T_{n-1})) \\ &= E\left(\sum_{w\in L_{n-1}} \sum_{v \text{ is a child of } w} E(\exp(\alpha U_J(v)) \mid T_{n-1})\right) \\ &= E\left(\sum_{w\in L_{n-1}} \exp(\alpha U_J(w)) \sum_{v \text{ is a child of } w} E(\exp(\alpha(U_J(v) - U_J(w))) \mid T_{n-1})\right) \\ \text{by}(2.13) \quad &= E\left(\sum_{w\in L_{n-1}} \exp(\alpha U_J(w))\right) \times \\ &\times \left(m_0 + \left(\sum_{k\in J} kp_k\right)(\exp(\alpha) - 1)\right) \\ &= \left(m_0 + \left(\sum_{k\in J} kp_k\right)(\exp(\alpha) - 1)\right) E(X_{n-1}). \end{aligned} \quad (2.14)$$

Since $E(X_0) = 1$, the statement of the lemma can be obtained by repeated application of (2.14).

Set $q_J = \sum_{k\in J} q_k$.

**Lemma 2.** Given $\epsilon > 0$ there exists $\delta > 0$, such that in the Galton-Watson tree $T$

$$P\left(\#\{v \in L_n : U_J(v) > (q_J + \epsilon)\, n\} > \frac{m_0^n}{n^2}\right) \leq 2\exp(-\delta n).$$

**Proof of lemma 2.** Note that if $T$ were such that

$$\#\{v \in L_n : U_J(v) > (q_J + \epsilon)\, n\} > \frac{m_0^n}{n^2},$$

then

$$X_n \geq \frac{m_0^n}{n^2} \exp(\alpha(q_J + \epsilon)n).$$

Therefore, by Markov's inequality

$$\begin{aligned} & P\left(\#\{v \in L_n : U_J(v) > (q_J + \epsilon)\, n\} > \frac{m_0^n}{n^2}\right) \\ & \leq \exp\left(-\alpha\,(q_J + \epsilon)\, n\right) \frac{E(X_n) n^2}{m_0^n} \\ & \leq \exp\left(-\alpha\,(q_J + \epsilon)\, n\right) \frac{n^2 \left[m_0 + \left(\sum_{k \in J} k p_k\right)(\exp(\alpha) - 1)\right]^n}{m_0^n} \\ & \leq n^2 \left[\exp\left(-\alpha\,(q_J + \epsilon)\right)\left(1 + q_J(\exp(\alpha) - 1)\right)\right]^n \end{aligned} \tag{2.15}$$

The statement of lemma 2 follows from (2.15) and the fact that

$$\exp\left(-\alpha\,(q_J + \epsilon)\right)\left(1 + q_J(\exp(\alpha) - 1)\right) = 1 - \epsilon\alpha + O(\alpha^2) < 1 \quad (\alpha \to 0).$$

for sufficiently small $\alpha$.

We return to the proof of theorem 2. In a Galton-Watson tree $T$ define $A_N \subset \delta T$ to be the set of all rays $(v_i)_{i \geq 0}$ such that $U_J(v_N) \geq (q_J + \epsilon)N$. Then $A_N$ is a union of $h_N$ balls of radii $\exp(-N)$ in $\delta T$, where

$$h_N = \#\{v \in L_N \,:\, U_J(v) \geq (q_J + \epsilon)N\}.$$

Moreover, since $U_J(v) = |v| \sum_{k\in J} \Delta_k(v)$,

$$\limsup_{N\to\infty} A_N = \{(v_i)_{i\geq 0} \,:\, \limsup_{i\to\infty} \sum_{k\in J} \Delta_k(v_i) \geq q_J + \epsilon\}$$

Let $Z_n$ be the size of the $n$-th generation in $T$. Apply Borel-Cantelli lemma and lemma 2 to conclude that for almost every Galton-Watson tree $T$ for all sufficiently large $N$

$$h_N \leq \frac{m_0^N}{N^2}. \tag{2.16}$$

Since $\lim_{n\to\infty} Z_n m_0^{-n}$ exists and is positive almost surely the series

$$\sum_N h_N m_0^{-N}$$

converges with probability 1. Apply proposition 12 with $\alpha = \log m_0$ to conclude that

$$H_{\log m_0}\left(\{(v_i)_{i\geq 0} \,:\, \limsup_{i\to\infty} \sum_{k\in J} \Delta_k(v_i) \geq q_J + \epsilon\}\right) = 0.$$

Since $\epsilon > 0$ was arbitrary, for $H_{\log m_0}$ almost all rays in $(v_i)_{i\geq 0} \in \delta T$

$$\limsup_{i\to\infty} \left(\sum_{k\in J} \Delta_k(v_i)\right) \leq q_J \tag{2.17}$$

almost surely.

Using (2.17) with $J = \{n\}$, we get

$$\limsup_{i\to\infty} \Delta_n(v_i) \leq q_n \quad \text{for a.e. } (v_i)_{i\geq 0} \text{ in a. e. } T. \tag{2.18}$$

On the other hand employing 2.17 with $J = \mathbb{Z}_+ \setminus \{n\}$, we get

$$\limsup_{i\to\infty} \left(\sum_{k\neq n} \Delta_k(v_i)\right) \leq 1 - q_n, \quad \Rightarrow$$

$$\liminf_{i\to\infty} \Delta_n(v_i) \geq q_n \quad \text{for a.e. } (v_i)_{i\geq 0} \text{ in a. e. } T. \tag{2.19}$$

Since (2.18) and (2.19) have been established for all $n$, theorem 2 is proved.

## 2.4 Subtrees of uniform growth.

Suppose most vertices in a Galton-Watson tree $T$ have at least $m$ children. In other words, $P(\xi_{rt} \geq m)$ is close to 1. There will still be pockets in $T$, where most vertices have small degrees. These pockets would violate the assumptions important to our approach to the lower bound for the Hausdorff dimension of certain subtrees. We would like to remove these pockets without slowing down the growth of $T$. That is, we would like to select a subtree $T' \subset T$, that branches quickly everywhere, while growing almost as fast as $T$ does.

An additional twist is that we will have to prove a result of this type for a branching tree somewhat more general than a Galton-Watson tree. By an *inhomogeneous* Galton-Watson tree we will mean a measure on the set of all (admissible) trees constructed in the same way as Galton-Watson measure, except that the assumption that $\xi_v$ are identically distributed is dropped. The random variables $\xi_v$ are still independent. Contrary to our usual practice, in this section we do not exclude Galton-Watson trees with leaves.

Getting down to a rigorous presentation, fix $\epsilon > 0$ and assume that for some sequence $(m_n)_{n\geq 0}$ and an inhomogeneous Galton-Watson tree $T$

$$P(\xi_v \geq m_n) \geq 1 - \epsilon \tag{2.20}$$

for any vertex $v \in L_n$. Define $A$ to be the event that the (inhomogeneous) Galton-Watson tree $T = (V, E)$ has a subtree $T' = (V', E')$ containing the root of $T$ such that any vertex $v \in V' \cap L_n$ has at least $m_n(1 - 3\epsilon)$ children in $V'$.

**Proposition 14** *In the preceding setup, there exist positive numbers $C_\epsilon$ and $\rho_\epsilon$*

$$P(A \,|\, \xi_{rt} \geq m_0) \geq 1 - C_\epsilon \exp(-\rho_\epsilon \min_i m_i),$$

*uniformly over all sequences $(m_n)_{n\geq 1}$ and the distributions of $\xi_v$ subject to (2.20).*

**Proof of proposition 14.** Let $(m_n)_{n\geq 0}$ be as in the statement of proposition. Without loss in generality, we may and will assume that $P(\xi_v \neq m_n) \geq 1-\epsilon$, when $|v| = n$. If some vertex $v \in L_n$ has more than $m_n$ children, just cut off the children in excess of $m_n$ and their descendants.

In the inhomogeneous Galton-Watson tree $T = (V, E)$, define subsets $(V_k) \subset V$ $(k \geq 0)$ inductively. Let $V_0$ be the set of all vertices $v$ in $V$ with fewer than $m_{|v|}$ children. Once $V_k$ is constructed, we put a vertex $v$ into $V_{k+1}$, if it either has at least $3\epsilon m_{|v|}$ children in $V_k$ or is in $V_k$ itself.

**Lemma 1.** There exists some positive number $\bar{m}$ such that if $\min_i m_i \geq \bar{m}$, then for any $k \geq 0$

$$P(rt \in V_k \,|\, \xi_{rt} = m_0) < \epsilon. \tag{2.21}$$

**Proof of lemma 1.** Inequality (2.21) is vacuously true for $k = 0$, regardless of the value of $\bar{m}$. Suppose it has been proved for $k = k_0$ and some $\bar{m}$. Let us try to verify it for $k = k_0 + 1$. Assume $\xi_{rt} = m_0$ and label the children of the root $v_1, \ldots, v_{m_0}$. Note that for any $i$

$$P(v_i \in V_{k_0}) \leq P(v_i \in V_0) + P(v_i \in V_{k_0} \,|\, \xi_{v_i} = m_1). \tag{2.22}$$

The first quantity in the right-hand side of (2.22) is smaller than $\epsilon$ by (2.20). The second one is bounded by $\epsilon$ as well, because (2.21) was proved for $k = k_0$. (We are applying the result with $k = k_0$ to $T^{v_i}$ and the sequence of the $m$'s started at $m_1$.) Thus, $P(v_i \in V_{k_0}) \leq 2\epsilon$. Since the events $v_i \in V_{k_0}$ are independent $(1 \leq i \leq m_0)$,

$$P(rt \in V_{k_0+1} \,|\, \xi_{rt} = m_0) \leq P(B(m_0, 2\epsilon) \geq 3\epsilon m_0), \tag{2.23}$$

where $B(m_0, 2\epsilon)$ is a binomial random variables with parameters $m_0$ and $2\epsilon$. By the law of large numbers,

$$\lim_{m\to\infty} P(B(m, 2\epsilon) \geq 3\epsilon m) = 0.$$

Hence, $\bar{m}$ can be chosen so that $P(B(m, 2\epsilon) \geq 3\epsilon m) < \epsilon$ for $m \geq \bar{m}$. Then (2.21) holds for $k = k_0 + 1$ by virtue of (2.23). Lemma 1 is now proved.

**Lemma 2.** If $v \in L_n \cap (V \setminus (\cup_k V_k))$, then $v$ has at least $m_n(1 - 3\epsilon)$ children in $V \setminus (\cup_k V_k)$.

**Proof of lemma 2.** Since $v \in L_n \cap V_0^c$, $v$ has $m_n$ children. Arguing by contradiction, assume that $v_1, \dots, v_j$ are distinct children of $v$, $j > 3\epsilon m_n$ and all the $v_i$'s just introduced are in $\cup_k V_k$. Then for some $k_i$'s, $v_i \in V_{k_i}$. Since the $V_i$'s are increasing, all the $v_i$'s with $1 \leq i \leq j$ belong to $V_{\max\{k_i:1\leq i\leq j\}}$. Then $v \in V_{\max\{k_i:1\leq i\leq j\}+1}$, which contradicts the assumptions of lemma 2.

Returning to the proof of proposition 14, observe that if the root does not belong to any of the $V_k$s, then $A$ happens. Indeed, then the root has $m_0$ children. By lemma 2, among them at least $m_0(1 - 3\epsilon)$ do not belong to any of the $V_k$'s either. These $m_0(1 - 3\epsilon)$ children must have degree $m_1$, since they are not in $V_0$. Each of them of them will have at least $m_1(1 - 3\epsilon)$ children of his own that are not in $\cup_k V_k$ either. Looking at their children, the children of their children and so on and repeatedly using lemma 2, we construct an infinite tree. In this tree no vertex is in $\cup_k V_k$ and each vertex of the $n$-th generation has at least $m_n(1 - 3\epsilon)$ children.

To complete the proof of the proposition we need to give an upper bound for $P(A^c \mid \xi_{rt} = m_0)$. The reasoning here is almost the same as in lemma 1. Assume that $\min_i m_i \geq \bar{m}$. Again label the children of the root $v_1, \dots, v_{m_0}$. Note that for any $i \leq m_0$

$$\begin{aligned} P(v_i \in \cup_k V_k) &\leq P(v_i \in V_0) + P(v_i \in \cup_k V_k \mid \xi_{v_i} = m_1) \\ &\leq \epsilon + \lim_{k\to\infty} P(v_i \in V_k \mid \xi_{v_i} = m_1) \leq 2\epsilon, \end{aligned}$$

where the last inequality is due to lemma 1. Since the events $v_i \in \cup_k V_k$ are independent,

$$P(A^c \,|\, \xi_{rt} = m_0) = P(rt \in \cup_k V_k \,|\, \xi_{rt} = m_0) \leq P(B(m_0,\, 2\epsilon) \geq 3\epsilon m_0)$$
$$\leq C_\epsilon \exp(-\rho_\epsilon m_0) \leq C_\epsilon \exp(-\rho_\epsilon \min_i m_i), \tag{2.24}$$

where next to the last inequality is just a simple large deviation statement for binomial random variables. Increasing $C_\epsilon$ if necessary, we can make sure that (2.24) also holds when $\min_i m_i \leq \bar{m}$. Proposition 14 is proved.

## 2.5 Degrees of vertices in a random path

Observe the first $N$ generations of a Galton-Watson tree $T$ and choose a vertex $v$ in the $N$-th generation at random. That is, all vertices in $T$ with $|v| = N$ have the same chance to be chosen. Let $v_0 = rt,\, v_1, \dots, v_N = v$ be the shortest path connecting $v$ to the root. Set $D_i = \xi_{v_i}$, the number of children $v_i$. In general, the $D_i$'s are neither independent nor identically distributed. However, as $N$ grows large, the joint distribution of the sequence of the $D$'s with large indexes becomes very close to that of a sequence of i.i.d. random variables. The purpose of this section is to make a precise statement about the $D's$ that would be almost as convenient to work with as the true independence. In fact, we will show that the $D's$ do become independent if one knows about each level $L_i$ $(i \leq N)$ how many vertices of each degree it contains. (In this context "knows" means "conditions upon")

The information of this sort will be stored in what we call *plausible degree* functions. Fix a positive integer $N$. A function of two non-negative integer arguments $\kappa(m,\, n)$ $(n \leq N)$ will be called a *plausible degree* function if it takes on non-negative integer values and for any $0 \leq n < N$

$$\sum_{m=1}^{\infty} m\kappa(m,\, n) = \sum_{m=1}^{\infty} \kappa(m,\, n+1).$$

A plausible degree function $\kappa$, is said to *summarize* a rooted tree $T$ (up to the level $N$) if for all suitable pairs $(m,\, n)$, the number of nodes with $m$ children located at the distance $n$ from the root of $T$ equals $\kappa(m, n)$. Any rooted tree $T$ is summarized by one plausible degree function which we will denote $\kappa(T)$ or $\kappa_N(T)$. (The latter piece of notation will only be used in subsequent sections where we do not assume $N$ to be fixed.)

Suppose $\kappa_0$ be a plausible degree function.

**Proposition 15** *For a Galton-Watson tree $T$ and $n < N$*

$$P(D_n = m \mid \kappa(T) = \kappa_0) = \frac{m\,\kappa_0(m, n)}{\sum_m m\,\kappa_0(m, n)}. \tag{2.25}$$

*Moreover, given that $\kappa(T) = \kappa_0$, the random variables $D_i$ $(0 \leq i < N)$ are independent.*

**Proof of the proposition 15.** Let $t_n = \sum_m \kappa_0(m, n)$ be the total number of vertices in the $n$-th generation of the Galton-Watson tree $T$ under the condition $\kappa(T) = \kappa_0$.

Assume that the Galton-Watson tree $T$ (conditioned on $\kappa(T) = \kappa_0$) is originally defined on some probability space $(\Omega_0, \mathbb{F}_0, P_0)$. That is, for each $\omega \in \Omega_0$ $T(\omega)$ is an admissible tree, and the distribution of $T(\omega)$ under $P_0$ is the (conditional) Galton-Watson measure. Let $\Omega_1$ be the product of symmetric groups

$$\Omega_1 = \Pi_{i=1}^N S_{t_i}$$

with the uniform measure $P_1$. We will define the Galton-Watson tree on the direct product

$$\Omega = \Omega_0 \times \Omega_1$$

with measure $P = P_0 \times P_1$. Pick $\omega \in \Omega_0$ and permutations $\pi_i \in S_{t_i}$ $(1 \leq i \leq N)$. Note that $L_1 \subset \mathbb{N}^\infty$ (and any other generation $L_k$ for that matter) admits a deterministic ordering. We may assume that this ordering is chosen once and for all and label the vertices of $T(\omega) \cap L_1$ $w_1, \ldots, w_{t_1}$ in such a way that $w_1 < \ldots < w_{t_1}$. Now we cut off the $T^{w_j}$s, $(1 \leq j \leq t_1)$ the trees of descendants of the $w_j$'s, and permute them using $\pi_1$. Then reattach them to the $w$'s. As a result of this procedure, what used to be $T^{w_j}$

is now attached to $w_{\pi_1(j)}$, and it thereby becomes $T^{w_{\pi_1(j)}}$. In the tree just obtained label all the vertices of the second generation $z_1, \ldots, z_{t_2}$ in the order of increasing, and then permute the $T^{z_j}$'s in accordance with $\pi_2$. Iterate this procedure $N$ times to get a tree $\bar{T}$. Even though $T$ and $\bar{T}$ are not necessarily isomorphic, it is clear that $\kappa(\bar{T}) = \kappa(T) = \kappa_0$.

Due to the homogeneity property of Galton-Watson trees, for any fixed set of permutations $(\pi_i)_{1\le i\le N}$ the distribution of $\bar{T}(\omega)$ under $P_0$ is the same as that of $T(\omega)$. Consequently, $\bar{T}$ is the Galton-Watson tree conditioned upon the event $\kappa(T) = \kappa_0$ defined on $\Omega$.

Let $\bar{D}_1, \ldots, \bar{D}_{N-1}$ be the numbers of children of the nodes on the path connecting the root to a randomly chosen vertex of $v_0$ in the $N$-th generation of $\bar{T}$. It follows from the preceding paragraph that the joint distribution of the $D_i$'s (conditional upon $\kappa(T) = \kappa_0$) is the same as that of the $\bar{D}_i$'s.

Given any sequence of integers $(m_i)_{1\le i<N}$ and a fixed tree $T(\omega)$ with $\kappa(T) = \kappa_0$, we will compute

$$P_1\left(\cap_{i=1}^{N-1}(\bar{D}_i = m_i)\right) = \Pi_{i=1}^{N-2} P_1\left(\bar{D}_i = m_i \mid \cap_{j=i+1}^{N-1}\left(\bar{D}_j = m_j\right)\right) P_1(D_{N-1} = m_{N-1}), \tag{2.26}$$

where $P_1$ stands for the uniform probability measure on $\Omega_1$.

To do this, note that whether $\cap_{j=i+1}^{N-1}(\bar{D}_j = m_j) \subset \Omega_1$ is in the $\sigma$-field generated by the the permutations $\pi_j$ with $j > i+1$. Moreover given the permutations $\pi_j$ with $j > i+1$, the event $\bar{D}_i = m_i$ is determined solely by what $\pi_{i+1}$ is. To wit, let $T'$ be the tree $T$ transformed by the first $i$ iterations described above. Backtracking the last $N-(i-1)$ iterations, we can determine which vertex $v_1 \in L_N \cap T'$ that will become $v_0$ after these $N-(i-1)$ iterations. Then looking at the same iterations we can find the vertex $u$ in the $(i+1)-st$ generation of $T'$ which is the predecessor of $v_1$. The event $\bar{D}_i = m_i$ happens if and only if $\pi_{i+1}$ acts in such way that $u$ becomes a child of a vertex with $m_i$ children. Among the total of $\sum_m m\kappa_0(m,\, i)$ vertices in the $(i+1)-st$ generation of $T'$, precisely $m_i\kappa_0(m_i,\, i)$ are children of parents with

$m_i$ children. Thus,

$$P_1\left(\bar{D}_i = m_i \mid \cap_{j=i+1}^{N-1} (\bar{D}_j = m_j)\right) = \frac{m_i \kappa_0(m_i,\, i)}{\sum_m m\kappa_0(m,\, i)}.$$

By the same token,

$$P_1\left(\bar{D}_{N-1} = m_i\right) = \frac{m_{N-1}\kappa_0(m_{N-1},\, N-1)}{\sum_m m\kappa_0(m,\, N-1)}.$$

Then formula (2.26) yields

$$P_1\left(\cap_{i=1}^{N-1}(\bar{D}_i = m_i)\right) = \Pi_{i=1}^{N-1}\left(\frac{m_i\kappa_0(m_i,\, i)}{\sum_m m\kappa_0(m,\, i)}\right). \tag{2.27}$$

Note that the right-hand side of (2.27) is independent of the choice of $T(\omega)$ provided $\kappa(T) = \kappa_0$. Consequently, the subscript 1 in $P_1$ can be dropped:

$$P\left(\cap_{i=1}^{N-1}(\bar{D}_i = m_i)\right) = \Pi_{i=1}^{N-1}\frac{m_i\kappa_0(m_i,\, i)}{\sum_m m\kappa_0(m,\, i)}. \tag{2.28}$$

Proposition 15 is an immediate consequence of (2.28) and the fact the joint conditional distribution of the $D_i's$ is that of $\bar{D}_i$'s.

## 2.6 Upper bound for the number of covering balls.

Recall the definitions of $\Delta_k(v)$ and the relative entropy $H(r|q)$ given in the introduction to Chapter 2. The purpose of this section is to give the upper bound for the number of vertices $v \in L_N$ in a Galton-Watson tree such that

$$\Delta_k(v) \approx r_k,$$

simultaneously for all $k$. Here $(r_k)_{k\geq 1}$ is a given sequence of non-negative numbers that add up to 1, $N$ is a large integer.

We remind our reader that the parameters of the Galton-Watson tree we are studying are $p_k = P(\xi_{rt} = k)$, $m_0 = E(\xi_{rt})$, and $q_k = \frac{kp_k}{m_0}$. Throughout our presentation the sequence $r = (r_k)_{k\geq 1}$, the desired frequencies of the vertices of degree $k$ along the path, will be fixed. We will be assuming that $r_k = 0$ whenever $p_k = 0$.

**Proposition 16** *Given any $\epsilon > 0$, there exist positive $\sigma$ and $K$ such that for almost every Galton-Walton tree*

$$\limsup_{N\to\infty} \frac{\log\left(\#\{v \in L_N \,|\, |\Delta_k(v) - r_k| \leq \sigma \text{ for all } k \leq K\}\right)}{N} \leq \log m_0 - H(r|q) + \epsilon,$$

*where $q = (q_k)_{k\geq 1}$.*

Before attacking the proposition we prove a minor generalization of the famous Borel-Cantelli lemma and do a simple large deviation computation.

**Lemma 1.** Suppose the events $(A_k)_{k\geq 1}$, $(B_k)_{k\geq 1}$ are such that $A_k \subset A_{k-1}$ for all $k > 1$, $P(\cap_k A_k) = 0$, and $\sum_k P(B_k \setminus A_k) < \infty$. Then almost surely only finitely many of the $B_k$'s occur.

**Proof of lemma 1.** Our assumptions on $(A_k)_{k\geq 1}$ ensure that with probability 1 only finitely many of the $A_k$'s occur. By the standard version of Borel-Cantelli, with probability 1 only finitely many of the events $(B_k \setminus A_k)_{k\geq 1}$ occur. Conclude that with probability 1 only finitely many of the events

$$B_k \subset A_k \cup (B_k \setminus A_k) \quad (k \geq 1)$$

occur.

**Lemma 2.** Suppose independent random variables $X_1, \ldots, X_N$ take on values 0 and 1. Assume that $P(X_i = 1) \leq \frac{1}{N^2}$ for each $i$. Then for any $\sigma$ and $L > 0$, there exists $C > 0$ such that

$$P(X_1 + \ldots + X_N \geq \sigma n) \leq C \exp(-LN).$$

**Proof of lemma 2.** By Markov's inequality

$$\begin{aligned}
P(X_1+\ldots+X_N \geq \sigma N) &\leq \exp(-\sigma N\log N)E(\exp(\sigma(\log N)(X_1+\ldots+X_N))) \\
&\leq \exp(-\sigma N\log N)(\max_i(E(\exp((\log N)X_i))))^N \\
&\leq \exp(-\sigma N\log N)\left(1+\frac{1}{N^2}(e^{\log N}-1)\right)^N \\
&\leq \exp(1-\sigma N\log N) \leq C\exp(-LN),
\end{aligned}$$

which proves the lemma.

**Proof of proposition 16.** Fix an $\epsilon > 0$. Our first task is to select $K$ appropriately. It should be so large that the relative entropy $H(r|q)$ changes very little if we cut off both $r$ and $q$ after the first $K$ terms. In fact, define the sequence $q' = (q'_k)_{k\geq 1}$ in the following way.

$$q'_k = \begin{cases} q_k, & \text{if } k \leq K \\ \sum_{l>K} q_l, & \text{if } k = K+1 \\ 0, & \text{if } k > K+1 \end{cases}.$$

(The definition of $r'$ is similar.) It is a trivial exercise in calculus to see that one can choose $K$ to be so large that

$$H(r'|q') \geq H(r|q) - \frac{\epsilon}{3}. \tag{2.29}$$

It is possible to choose a $\sigma \in (0,\, 1/2)$ to be so small that

$$H(r|q) - \frac{5\epsilon}{6} < \left(H(r|q) - \frac{2\epsilon}{3}\right)\left(1-\frac{\sigma}{4}\right), \quad \text{and} \tag{2.30}$$

$$H(\bar{r}|q') \geq H(r'|q') - \frac{\epsilon}{3} \tag{2.31}$$

for any sequence of non-negative numbers $\bar{r} = (\bar{r}_k)_{1\leq k\leq K+1}$ such that

$$|r_k - \bar{r}_k| \leq (2K+1)\sigma \text{ for } k \leq K+1 \text{ and } \sum_{k=1}^{K+1} \bar{r}_k = 1.$$

For this choice of $\sigma$, let $B_N$ be the event that

$$\frac{\log\left(\#\{v \in L_N \,:\, |\Delta_k(v) - r_k| \leq \sigma \text{ for all } k \leq K\}\right)}{N} > \log m_0 - H(r|q) + \epsilon.$$

Because of the deviations that might be present in the first few generations of the tree, it is not very convenient to deal with $\Delta_k(v)$. We will introduce its modification $\bar{\Delta}_k(v)$. Set

$$M_N = [\sigma N/4]. \tag{2.32}$$

Abusing notation slightly, we will be omitting index $N$. For a vertex $v \in L_N$, let $m(v)$ be the vertex in $L_M$ located on the shortest path from $v$ to the root. Then define a modified quantity $\bar{\Delta}_k(v)$

$$\bar{\Delta}_k(v) = \frac{\#\{w \in V,\ deg(w) = k+1,\ w \text{ disconnects } v \text{ from } m(v)\}}{N-M}.$$

(For the sake of definiteness, we adopt the following agreement. If $m(v)$ has $k$ children it contributes to the cardinality in the numerator. On the other hand $v$ is not counted even when it has $k$ children. Whenever we define a $\Delta$ of some sort, only the predecessor may be counted.)

The values of $\bar{\Delta}_k$ and $\Delta_k$ are close to one another, but $\bar{\Delta}_k$ is easier to handle.

**Lemma 3.** Referring to the preceding, for any $v \in L_N$ and any $k$

$$|\bar{\Delta}_k(v) - \Delta_k(v)| \leq \frac{\sigma}{2}.$$

**Proof of lemma 3.** Since there are only $M = [\sigma N/4]$ vertices on the path between the root and $m(v)$,

$$\#\{w \in V,\ deg(w) = k+1,\ w \text{ disconnects } v \text{ from } m(v)\} \leq$$
$$\#\{w \in V,\ deg(w) = k+1,\ w \text{ disconnects } v \text{ from } rt\} \leq$$
$$\#\{w \in V,\ deg(w) = k+1,\ w \text{ disconnects } v \text{ from } m(v)\} + \frac{\sigma N}{4}.$$

Dividing the first inequality by $N - M$, in view of (2.32) we get

$$\bar{\Delta}_k(v) \leq \Delta_k(v)\frac{N}{N-M} \leq \Delta_k(v)(1+\sigma/2) \leq \Delta_k(v) + \frac{\sigma}{2}.$$

Divide the second inequality of the proof by $N$, to get

$$\Delta_k(v) \leq \frac{N-M}{N}\bar{\Delta}_k(v) + \frac{\sigma}{4} \leq \bar{\Delta}_k(v) + \frac{\sigma}{4},$$

which establishes of lemma 3.

Consider $\bar{B}_N$, a modified version of $B_N$, defined to be the event that

$$\frac{\log\left(\#\{v \in L_N\ :\ |\bar{\Delta}_k(v) - r_k| \leq 2\sigma \text{ for all } k \leq K\}\right)}{N} > \log m_0 - H(r|q) + \epsilon.$$

It follows from lemma 3, that

$$B_N \subset \bar{B_N} \tag{2.33}$$

Therefore, the proof of proposition 16 boils down to establishing that for almost every Galton-Watson tree only finitely many of the $\bar{B}_N$'s (and whence of the $B_N$'s) occur.

Our plan is to divide the plausible degree functions into two classes, called $RGLR_N$ and $EXCP_N$. (Since $N$ will vary, the plausible degree functions and the classes we are about to introduce will be supplied with a subscript.) Informally, trees in $RGLR_N$ branch as one would expect from the Law of Large Numbers, while trees in $EXCP_N$ exhibit deviant behavior at one or more levels. Moreover, at each level we will impose on the trees in $RGLR_N$ only finitely many conditions. This means we will have no direct control over the values of $\kappa_N(m, n) \in RGLR_N$ for large $m$.

We will say $\kappa_N \in RGLR_N$ if the following holds.

(i) $\sum_m \kappa(m, M) \geq m_0^{M/2}$.

(ii) For all integer $n \in [M, N]$ and all $m \leq K$

$$\left|\kappa(m, n) - p_m \sum_i \kappa(i, n)\right| \leq \frac{1}{(K+1)^2 N^2} \sum_i \kappa(i, n).$$

(iii) $N^{-1} m_0^N \leq \sum_m \kappa(m, N) \leq N m_0^N$.

Otherwise $\kappa_N \in EXCP_N$.

Let $A_N$ be the event that for some $n \geq N$ $\kappa_n(T) \in EXCP_n$. (The function $\kappa_n(T)$ summarizes the tree $T$ up to the level $n$.)

Of course, it is our goal to establish that the sequences of events $(A_N)_{N\geq 1}$ and $\bar{B}_{N\geq 1}$ satisfy the assumptions the lemma 1. The inclusion property of the sequence $(A_N)_{N\geq 1}$ is evident. Therefore, to prove proposition 16 we need prove the following lemma 4 and lemma 5.

**Lemma 4.** In the above notation, $\sum_N P(\bar{B}_N \setminus A_N) < \infty$.

**Lemma 5.** For the events $(A_N)_{N\geq 1}$ introduced above, $P(\cap_N A_N) = 0$.

**Proof of lemma 4.** By the telescoping property of conditional probabilities

$$P(\bar{B}_N \setminus A_N) = E\left(P(\bar{B}_N \setminus A_N \,|\, \kappa(T))\right) = E(P(\bar{B}_N \,|\, \kappa(T));\ \kappa(T) \in RGLR_N), \tag{2.34}$$

because if the tree $T$ is such that $\bar{B}_N \setminus A_N$ occurs, then $\kappa_N(T) \in RGLR_N$.

Pick any $\kappa_0 \in RGLR_N$ and estimate $P(\bar{B}_N \,|\, \kappa(T) = \kappa_0)$. By Markov's inequality and the definition of $\bar{B}_N$,

$$\begin{aligned} P(\bar{B}_N \,|\, \kappa_N(T) = \kappa_0) &\leq \exp(-N(\log m_0 - H(r|q) + \epsilon)) \times \\ &\times E(\#\{v \in L_N \,:\, |\bar{\Delta}_k(v) - r_k| \leq 2\sigma \text{ for all } k \leq K\} \,|\, \kappa_N(T) = \kappa_0). \end{aligned} \tag{2.35}$$

Set $t_N = \sum_m \kappa_0(m, N)$. Since $\kappa_0$ is assumed to belong to $RGLR_N$, condition (iii) yields

$$N^{-1}m_0^N \le t_N \le N m_0^N. \tag{2.36}$$

Let $v$ be a randomly chosen vertex among the $t_N$ vertices of the $N$-th generation in $T$, the $D_i$'s be as in section 2.4. As a matter of notation when we use the $D$'s, all the probabilities involved are meant to be conditional upon $\kappa(T) = \kappa_0$. Define $\mu_D$ to be the empirical measure on the integers:

$$\mu_D = \frac{1}{N-M}\sum_{i=M}^{N-1} \delta_{D_i}.$$

Then

$$\begin{aligned} E(\#\{v \in L_N \,:\, |\bar{\Delta}_k(v) - r_k| \le 2\sigma \text{ for all } k \le K\} \mid \kappa_N(T) = \kappa_0) \\ = t_N P(|\mu_D(\{k\}) - r_k| \le 2\sigma \text{ for all } k \le K). \end{aligned} \tag{2.37}$$

We would like to apply Sanov's Theorem to estimate the latter quantity, but there are two problems. First, the random variables $(D_i)_{M \le i \le N-1}$ are not identically distributed. We need a coupling argument to rectify the situation. Second, neither we have control over the tails of $D_i$, nor does the right-hand side of (2.37) call for the information on these tails. Therefore, we will have to trim $D_i$ at $K$. To this end, we need $f(x) = \min(x,\, K+1)$ and

$$\mu_{f(D)} = \frac{1}{N-M}\sum_{i=M}^{N-1} \delta_{f(D_i)}.$$

Observe that if $X$ is a positive integer random variable such that $|P(X=k) - r_k| \le 2\sigma$ for all $k \le K$, then

$$\begin{aligned} P(f(X) = k+1) - r'_{K+1} &= \left(1 - \sum_{k \le K} P(f(X) = k)\right) - \left(1 - \sum_{k \le K} r_k\right) \\ &\le \sum_{k \le K} |P(X=k) - r_k| \le 2K\sigma. \end{aligned}$$

Using this observation for an $X$ whose distribution is $\mu_D$, write

$$
\begin{aligned}
P(|\mu_D(\{k\}) - r_k| \leq & 2\sigma \text{ for all } k \leq K) \\
& \leq P(|\mu_{f(D)}(\{k\}) - r'_k| \leq 2K\sigma \text{ for all } k \leq K+1).
\end{aligned} \tag{2.38}
$$

Apply proposition 15 and the assumption (ii) on $\kappa_0$ to obtain

$$
\begin{aligned}
\sum_{j=1}^{K} |P(D_i = j \,|\, \kappa_N(T) = \kappa_0) - q_j| &= \sum_{j=1}^{K} \left| \frac{j\,\kappa_0(j,\,i)}{\sum_m m\,\kappa_0(m,\,i)} - q_j \right| \\
&\leq \sum_{j=1}^{K} \frac{2j}{(K+1)^2 N^2} \leq \frac{1}{N^2}.
\end{aligned}
$$

Since the $D_i$'s are independent (given $\kappa_N(T) = \kappa_0$), it may be assumed without loss in generality, that there are independent random variables $\eta_i$'s defined on the same probability space as $D_i$, such that for any $i \geq M$

$$
P(f(D_i) \neq \eta_i) \leq \frac{1}{N^2}, \quad P(\eta_i = m) = q'_m. \tag{2.39}
$$

Define the empirical measure corresponding to $\eta$'s:

$$
\mu_\eta = \frac{1}{N-M} \sum_{i=M}^{N-1} \delta_{\eta_i}.
$$

We are now in a position to use our coupling argument. Write

$$
\begin{aligned}
P(|\mu_{f(D)}(\{k\}) - r'_k| \leq 2K\sigma \text{ for all } k \leq K+1) &\leq P(\#\{i : f(D_i) \neq \eta_i\} \geq \sigma(N-M)) \\
&+ P(|\mu_\eta(\{k\}) - r'_k| \leq (2K+1)\sigma \text{ for } k \leq K+1).
\end{aligned} \tag{2.40}
$$

By Sanov's Theorem (cf [2, p. 70]), (2.29) and (2.31),

$$
\begin{aligned}
\limsup_{N\to\infty} \frac{P(|\mu_\eta(\{k\}) - r'_k| \leq (2K+1)\sigma \text{ for all } k \leq K+1)}{N-M} &\leq -H(r'|q') + \frac{\epsilon}{3} \\
&\leq -H(r|q) + \frac{2\epsilon}{3}.
\end{aligned} \tag{2.41}
$$

Furthermore, in view of the first inequality in (2.39), lemma 2 gives

$$\limsup_{N\to\infty} \frac{P(\#\{i : f(D_i) \neq \eta_i\} \geq \sigma(N-M))}{N} = -\infty. \tag{2.42}$$

Recall that $M = [\sigma N/4]$. Putting together (2.40), (2.41) and (2.42), we have

$$\limsup_{N\to\infty} \frac{P(|\mu_{f(D)}(\{k\}) - r'_k| \leq 2K\sigma \text{ for } k \leq K+1)}{N} \leq \left(1 - \frac{\sigma}{4}\right)\left(-H(r|q) + \frac{2\epsilon}{3}\right). \tag{2.43}$$

Then use (2.37) and (2.38) in the left-hand side of the last line, and (2.30) in the right-hand-side to conclude that

$$\limsup_{N\to\infty} \frac{E(\#\{v \in L_N : |\bar{\Delta}_k(v) - r_k| \leq 2\sigma \text{ for } k \leq K\} \,|\, \kappa_N(T) = \kappa_0)}{N} \leq$$
$$\leq \limsup_{N\to\infty} \frac{\log t_N}{N} - H(r|q) + \frac{5\epsilon}{6} \tag{2.44}$$

Since $t_N$ is constrained by (2.36),

$$\lim_{N\to\infty} \frac{\log t_N}{N} = \log m_0. \tag{2.45}$$

Infer from (2.35), (2.44) and (2.45) that

$$\limsup_{N\to\infty} \frac{P(\bar{B}_N \,|\, \kappa_N(T) = \kappa_0)}{N} \leq -\frac{\epsilon}{6} \tag{2.46}$$

for $\kappa_0 \in RGLR_N$. Lemma 4 follows from (2.34) and (2.46).

**Proof of lemma 5.** Let $Z_k = \#L_k$ stand for the size of the $k$-th generation. It is very well known from the theory of branching processes (e.g. see [1, p. 9]) that $\lim_{k\to\infty} \frac{Z_k}{m_0^k}$ exists and is positive almost surely. It follows that for all $k > k_0$

$$Z_k = \sum_m \kappa_k(T)(m,k) \geq m_0^{k/2} \text{ and}$$

$$\frac{m_0^k}{k} \leq Z_k \leq k m_0^k.$$

Therefore, condition (iii) in the definition of $RGLR_N$ is satisfied for $\kappa_N(T)$ whenever $N > k_0$. If in addition $N$ is so large that $M = [\sigma N/4] > k_0$, then (i) is also satisfied.

By the very construction of the Galton-Watson tree, given $Z_k$, $\kappa_n(T)(m, k)$ (here $n > k$) is distributed binomially with parameters $Z_k$ and $p_m$. Therefore, by Chebishev's inequality

$$P\left(\left|\frac{\kappa_n(T)(m,k)}{Z_k} - p_m\right| \geq \frac{1}{(K+1)^2\left(\frac{4k}{\sigma}\right)^2} \mid Z_k > m_0^{k/2}\right)$$
$$\leq \frac{(256(K+1)^4 k^4) p_m (1-p_m)}{\sigma^4} m_0^{-k/2}.$$

Clearly,

$$\sum_k \frac{256(K+1)^4 k^4 p_m k^2 (1-p_m)}{\sigma^4} m_0^{-k/2} < \infty.$$

It follows that for all but finitely many $M \geq k_0$ and all $m \leq K$

$$|\kappa(m, M) - p_m \sum_i \kappa(i, M)| \leq \frac{1}{(K+1)^2 N^3} \sum_i \kappa(i, n). \quad (M = [\sigma N/4])$$

almost surely. This is precisely condition (ii). Lemma 5 and proposition 16 are now proved.

## 2.7 Lower bounds for the number of covering balls.

The purpose of this section is to give a lower bound for the number of vertices $v \in L_N$ in a Galton-Watson tree $T$ such that $\Delta_k(v) \approx r_k$ simultaneously for several $k$'s. The simplest result that can be established here is the bound for the expected number of such vertices. However, it turns out that using some clever trick (which was shown to this writer by Balint Virag), this result can be upgraded to a bound that holds with probability close to 1. The notation in this section is the same as in the preceding one.

**Proposition 17** *Let $\beta$, $\sigma$, and $K$ be positive numbers. Assume that*

$$\beta < \log m_0 - H(r|q). \tag{2.47}$$

*Then there exists $N_0$ such that for any $N > N_0$*

$$E\left(\#\{v \in L_N : |\Delta_k(v) - r_k| \leq \sigma \text{ for all } k \leq K\}\right) \geq \exp(\beta N). \tag{2.48}$$

**Proof of proposition 17.** We begin by reducing the unconditional expectation in (2.48) to a conditional one. The condition will be that the tree $T$ (up to the level $N$) is summarized by a plausible degree function $\kappa_N(T) \in RGLR_N$. According to lemma 5 of the preceding section,

$$\lim_{N\to\infty} P(\kappa_N(T) \in RGLR_N) \geq 1 - \lim_{N\to\infty} P(\cap_N A_N) = 1.$$

Consequently, if $N_0$ is chosen to be large enough, then for $N \geq N_0$

$$P(\kappa_N(T) \in RGLR_N) \geq 1/2.$$

Hence, for such $N$

$$\begin{aligned} &E\left(\#\{v \in L_N : |\Delta_k(v) - r_k| \leq \sigma \text{ for all } k \leq K\}\right) \geq \\ &\geq \frac{1}{2} E\left(\#\{v \in L_N : |\Delta_k(v) - r_k| \leq \sigma \text{ for all } k \leq K\} \mid \kappa_N(T) \in RGLR_N\right). \end{aligned} \tag{2.49}$$

The next step we need to take is to replace $\Delta_k(v)$ by $\bar{\Delta}_k(v)$ described in the preceding section. Without loss in generality we may assume (cf (2.47)) that

$$\beta < \log m_0 - \frac{H(r|q)}{1-\sigma}. \tag{2.50}$$

Indeed, decreasing the value of $\sigma$, we will make inequality (2.48) even sharper.

Set $M = [\sigma N/4]$ and define $\bar{\Delta}_k(v)$ as in the previous section. It follows from lemma 3 of that section that if $|\bar{\Delta}_k(v) - r_k| \leq \sigma/2$ for some node $v \in L_N$, then $|\Delta_k(v) - r_k| \leq \sigma$. Conclude from (2.49) that

$$\begin{aligned}&E\left(\#\{v \in L_N : |\Delta_k(v) - r_k| \leq \sigma \text{ for all } k \leq K\}\right) \geq \\ &\geq \frac{1}{2} E\left(\#\{v \in L_N : |\bar{\Delta}_k(v) - r_k| \leq \sigma/2 \text{ for all } k \leq K\} \mid \kappa_N(T) \in RGLR_N\right).\end{aligned} \tag{2.51}$$

Fix $\kappa_0 \in RGLR_N$ and let the $D_i$'s, $\mu_D$, $\mu_{f(D)}$, the $\eta_i$'s, $\mu_\eta$ and $t_N$ stand for what they did in the preceding section. Write

$$\begin{aligned}E(\#\{v \in L_N : |\bar{\Delta}_k(v) - r_k| \leq \sigma/2 \text{ for all } k \leq K\} | \kappa_N(T) = \kappa_0) \\ = t_N P(|\mu_D(\{k\}) - r_k| \leq \sigma/2 \text{ for all } k \leq K).\end{aligned} \tag{2.52}$$

This time use the coupling argument in the following way:

$$\begin{aligned}&P(|\mu_D(\{k\}) - r_k| \leq \sigma/2 \text{ for all } k \leq K) \\ &\geq P\left(|\mu_\eta(\{k\}) - r_k| \leq \sigma/4 \text{ for } k \leq K\right) \\ &- P\left(\#\{i \,|\, f(D_i) \neq \eta_i\} \geq (\sigma(N-M))/4\right).\end{aligned} \tag{2.53}$$

By lemma 2 of the previous section

$$\limsup_{N\to\infty} \frac{\log\left(P\left(\#\{i : f(D_i) \neq \eta_i\} \geq \frac{\sigma(N-M)}{4}\right)\right)}{N} = -\infty. \tag{2.54}$$

By Sanov's theorem (cf. [2, p. 70]),

$$\liminf_{N\to\infty} \frac{\log\left(P(|\mu_\eta(\{k\}) - r_k| \leq \sigma/4 \text{ for } k \leq K)\right)}{N-M} \geq -H(r|q). \tag{2.55}$$

Combining (2.53), (2.54) and (2.55), we arrive at

$$\liminf_{N\to\infty} \frac{\log\left(P(|\mu_D(\{k\}) - r_k| \leq \sigma/2 \text{ for all } k \leq K)\right)}{N-M} \geq -H(r|q). \tag{2.56}$$

Due to the fact $M = [\sigma N/4]$ and (2.52), estimate (2.56) implies the following inequality

$$\liminf_{N\to\infty} \log \left(E\left(\#\{v \in L_N \ : \ |\bar{\Delta}_k(v) - r_k| \le \sigma/2 \text{ for all } k \le K\} \mid \kappa_N(T) = \kappa_0\right)\right)$$
$$\ge \liminf_{N\to\infty} \frac{t_N}{N} - \frac{H(r|q)}{1-\sigma} > \beta, \tag{2.57}$$

where the last inequality is just a restatement of (2.50). (That limit involving $t_N$ was handled via (2.45).) Since the choice of $\kappa_0 \in RGLR_N$ was arbitrary, inequalities (2.51) and (2.57) together give that

$$\liminf_{N\to\infty} \frac{\log\left(E\left(\#\{v \in L_N \ : \ |\Delta_k(v) - r_k| \le \sigma \text{ for all } k \le K\}\right)\right)}{N} > \beta$$

Proposition 17 is now proved.

**Proposition 18** *Under the assumptions of proposition 17, for any $\epsilon > 0$ there exists a positive $N$ such that*

$$P\left(\#\{v \in L_N \ : \ |\Delta_k(v) - r_k| \le \sigma \textit{ for all } k \le K\} > \exp \beta N\right) \ge 1 - \epsilon. \tag{2.58}$$

**Proof of proposition 18.** We will actually prove a stronger statement. Namely, we will establish, that for almost every tree $T$ in the sense of Galton-Watson measure

$$\liminf_{N\to\infty} \frac{\log\left(\#\{v \in L_N \ : \ |\Delta_k(v) - r_k| \le \sigma \text{ for all } k \le K\}\right)}{N} > \beta. \tag{2.59}$$

Our strategic goal is to reduce the inequality (2.59) to a survival statement for a certain Galton-Watson tree. (Of course, the underlying distribution of the number of children in this new tree will be different.)

Take an arbitrary $\beta_1 \in (\beta, \log m_0 - H(r|q))$. By proposition 17 applied with $\sigma/2$ and $\beta_1$ in lieu of $\sigma$ and $\beta$, we are able to find an $N_1$ such that

$$E\left(\#\{v \in L_{N_1} : |\Delta_k(v) - r_k| \leq \sigma/2 \text{ for all } k \leq K\}\right) \geq \exp(\beta_1 N). \tag{2.60}$$

Fix for a while a vertex $v$ in the original tree $T$, and focus our attention on $T^v$, the tree of descendants of $v$. For any vertex, $w \in T^v$, such that $|w| \geq |v| + N_1$, let $u(w)$ be its predecessor such that $dist(u(w), w) = N_1$. Define

$$\Delta'_k(w) = \frac{\#\{u \in T^v : deg(u) = k+1 \text{ and } u \text{ disconnects } w \text{ from } u(w)\}}{N_1}.$$

Next we construct a Galton-Watson tree $S_v$, which we will call a *supertree*. The root of $S_v$ is $v$. The children of $v$ in $S_v$ are all the vertices $w$ in $T^v \cap L_{|v|+N_1}$ such that

$$|\Delta'_k(w) - r_k| \leq \frac{\sigma}{2} \text{ for all } k \leq K.$$

More generally, if $u$ is a vertex in the supertree $S_v$, then the children of $u$ in $S_v$ are the vertices $w$ in $T^v \cap L_{|u|+N_1}$ such that

$$|\Delta'_k(w) - r_k| \leq \frac{\sigma}{2} \text{ for } k \leq K. \tag{2.61}$$

It is clear that $S_v$ is a Galton-Watson tree. The expected number of children of a vertex in $S_v$ equals to

$$E\left(\#\{v \in L_{N_1} : |\Delta_k(v) - r_k| \leq \sigma/2 \text{ for all } k \leq K\}\right). \tag{2.62}$$

Call this quantity $m_1$. It is clear from proposition 17 that if $N_1$ is large enough $m_1 > 1$. Therefore, the survival probability of $S_v$ (i.e. the probability that $S_v$ is infinite), is positive.

In the original Galton-Watson tree one can choose an infinite sequence of vertices $v_1, v_2, \ldots$ in such a way that no $v_i$ is a descendant of some other $v_j$ and all the $|v_i|$'s are divisible by $N_1$. For this sequence, the survivals of the $(S_{v_i})_{i \geq 1}$ are mutually independent events. (Indeed each survival only depends on the structure of $T^{v_i}$ and these trees do not overlap.) These events have the same positive probability. Hence, a Galton-Watson tree will almost surely have a vertex $v_0$ such that $S_{v_0}$ survives.

Let $Y_l$ be the number of children in the $l$-th generation of $S_{v_0}$. By a standard theorem,

$$\lim_{l\to\infty} \frac{Y_l}{m_1^l}$$

exists and is positive almost surely. Using estimate (2.60) for $m_1$ one can write

$$Y_l \geq C \exp(\beta_1 N_1 l), \tag{2.63}$$

where the constant $C$ may depend on the original tree $T$ and the choice of $v_0$.

**Lemma 1.** There exists $l_0$ (which may depend on $v_0$ and $T$), such that for any $l \geq l_0$, and any vertex $w$ in the $l$-th generation of $S_{v_0}$,

$$|\Delta_k(w) - r_k| \leq \sigma \text{ for all } k \leq K.$$

**Proof of lemma 1.** Let $w_0 = v_0$, $w_1, \dots, w_l = w$, be the path connecting $w$ to $v_0$ in $S_{v_0}$. Then the path between $w$ and the root of $T$ can be subdivided into pieces by vertices $w_0$, $w_1 \dots, w_{l-1}$. In any piece between some $w_i$ and $w_{i+1}$, the fraction of the vertices with $k$ children doesn't deviate from $r_k$ by more than $\sigma/2$ by the definition of $S_{v_0}$ (cf. (2.61)). The concentration of vertices with $k$ children in the piece between the root and $v_0$ is not known, but as $l$ grows its contribution to $\Delta_k(w)$ becomes negligible.

More formally,

$$\Delta_k(w) = \frac{\Delta_k(v_0)|v_0| + N_1 \sum_{i=1}^{l} \Delta_k'(w_i)}{lN_1 + |v_0|}.$$

In view of (2.61), the last line implies that

$$
\begin{aligned}
|\Delta_k(w) - r_k| &\le \left| \frac{\Delta_k(v_0)|v_0|}{lN_1 + |v_0|} - r_k \frac{|v_0|}{lN_1 + |v_0|} \right| + \frac{N_1}{lN_1 + |v_0|} \sum_{i=1}^{l} |\Delta'_k(w_i) - r_k| \\
&\le \left| \frac{\Delta_k(v_0)|v_0| - |v_0| r_k}{lN_1 + |v_0|} \right| + \left( \frac{lN_1}{lN_1 + |v_0|} \right) \left( \frac{\sigma}{2} \right) \\
&\le \frac{|v_0|}{lN_1 + |v_0|} + \sigma/2 < \sigma,
\end{aligned}
$$

if $l$ is sufficiently large. Lemma 1 is proved.

Observe that the $l$-th generation of $S_{v_0}$ is a subset of $L_{|v_0|+N_1 l}$. Thus, we can infer from lemma 1, that

$$
\begin{aligned}
&\liminf_{l\to\infty} \frac{\log \left( \#\{v \in L_{|v_0|+N_1 l} \; : \; |\Delta_k(v) - r_k| \le \sigma \text{ for all } k \le K\} \right)}{|v_0| + N_1 l} \\
&\qquad \ge \liminf_{l\to\infty} \frac{\log Y_l}{|v_0| + N_1 l} \\
\text{use (2.63)} \qquad &\qquad \ge \liminf_{l\to\infty} \frac{\log C + \beta_1 N_1 l}{|v_0| + N_1 l} > \beta.
\end{aligned}
$$

This proves statement (2.59) along the sequence of $N = |v_0| + N_1 l$ which includes all sufficiently large multiples of $N_1$. Obviously, in the same way we can establish (2.59) for the sequences of $N$'s that have other residues modulo $N_1$. This proves proposition 18.

## 2.8 Proof of the main theorem.

Given a Galton-Watson tree $T$, let $A_r$ be the set of rays $(v_i)_{i\ge 0}$ in $\delta T$ such that $\lim_{i\to\infty} \Delta_k(v_i) = r_k$. for all positive integers $k$.

**Proposition 19** *For almost every Galton-Watson tree $T$*

$$dim(A_r) \ge \log m_0 - H(r|q).$$

**Proof of Proposition 19**. Assume $H(r|q) < \log m_0$, since otherwise there is nothing to prove. Pick any $\beta < \log m_0 - H(r|q)$. According to proposition 18, for any integer $K$ there exists $N_K$ such that

$$P\left(\#\{v \in L_{N_K} : |\Delta_k(v) - r_k| \leq K^{-1} \text{ for all } k \leq K\} > \exp(\beta N_K)\right) \geq \frac{3}{4}. \quad (2.64)$$

We may further assume that the sequence $(N_K)_{K \geq 1}$ is strictly increasing and $N_1$ is so large that

$$1 - C_{1/4} \exp(-\rho_{1/4} \exp(\beta N_1)) > 0, \quad (2.65)$$

where $C_{1/4}$ and $\rho_{1/4}$ are as in proposition 14. It is easy to choose positive integers $(i_K)_{K \geq 1}$ in such a way that

$$N_K = o\left(\sum_{j=1}^{K-1} i_j N_j\right) \text{ as } K \to \infty \quad (2.66)$$

Put together an infinite sequence $N_1, \ldots, N_1, N_2, \ldots, N_2, N_3 \ldots, N_3, \ldots$ where each $N_K$ is taken $i_K$ times. Let $S_i$ stand for the sum of the first $i$ terms of this sequence. (Of course, $S_0 = 0$.) Then, it is easy to see from (2.66) that

$$S_i - S_{i-1} = o(S_i) \text{ as } i \to \infty. \quad (2.67)$$

It is our intention to define an inhomogeneous Galton-Watson tree $R$ whose $i$-th generation consists of some vertices in the $S_i - th$ generation $L_{S_i}$ of the original Galton-Watson tree $T = (V, E)$. Suppose $i$ is a positive integer, and $K$ is the index for which $S_i - S_{i-1} = N_K$. Then vertices $v \in L_{S_{i-1}}$ and $w \in L_{S_i}$ will be connected to one another in $R$ if $v$ is a predecessor of $w$, and for any $k \leq K$

$$\left|\frac{\#\{u \in V : deg(u) = k+1, \ u \text{ disconnects } w \text{ from } v\}}{N_K} - r_k\right| \leq \frac{1}{K}. \quad (2.68)$$

Set

$$m_{i-1} = \exp(\beta N_K). \tag{2.69}$$

Connecting vertices of $L_{S_{i-1}}$ and $L_{S_i}$ as above for all integer $i$ we get a graph. We define $R$ to be the component of the graph containing the root. It is clear that $R$ is an inhomogeneous Galton-Watson tree. Moreover, comparing (2.64) to (2.68), one notes that $N_k$ were chosen in such a way, that the assumptions of proposition 14 are satisfied with $\epsilon = 1/4$ and the $(m_i)$'s introduced by (2.69).

Apply the conclusion of that proposition. Due to (2.65), with positive probability $R$ will have an infinite subtree $R'$ such that, each vertex in the $i$-th generation of $R'$ has at least $\frac{m_i}{4}$ children in $R'$. Consider a subtree $T' = (V', E')$ of $T = (V, E)$ defined in the following way. A vertex $v \in V$ will belong to $V'$ if it has descendants in $R'$. Clearly all vertices of $R'$ itself are also in $T'$. Apply proposition 11 to $T'$ with $c = 1/4$, $\alpha = \beta$ and the sequence $(S_i)_{i\geq 1}$ defined above. Then $dim(\delta T') \geq \beta$.

Summarizing the conclusion of the two preceding paragraphs, write

$$P(\dim(\delta T') \geq \beta) > 0. \tag{2.70}$$

**Lemma 1.** Referring to the preceding setup, $\delta T' \subset A_r$.

**Proof of lemma 1.** Let $(v_i)_{i\geq 0}$ be a ray in $\delta T'$. Clearly, $v_{S_i} \in R'$ for all integer $i$. Fix integers $k$ and $K \geq k$, and set $h = \sum_{j<K} i_j N_j$. By the construction of $R'$ (cf. (2.68)), for any $g \geq h$

$$\left| \frac{\#\{u \in V : deg(u) = k+1, \ u \text{ disconnects } v_{S_g} \text{ from } v_{S_{g+1}}\}}{N_{K_1}} - r_k \right| \leq \frac{1}{K}, \tag{2.71}$$

where $K_1$ is the index for which $S_{g+1} - S_g = N_{K_1}$. (In fact, we ought to write $1/K_1$ instead of $1/K$ in the right-hand side of (2.71), but it is clear that $K_1 \geq K$.)

It follows from (2.71) that for any $g > h$,

$$\left| \# \frac{\{u \in V : deg(u) = k+1, \ u \text{ disconnects } v_{S_h} \text{ from } v_{S_g}\}}{S_g - S_h} - r_k \right| \leq \frac{1}{K}. \tag{2.72}$$

Therefore,

$$\liminf_{g\to\infty} \Delta_k(v_{S_g}) - \frac{1}{K} \le r_k \le \liminf_{g\to\infty} \Delta_k(v_{S_g}) + \frac{1}{K}. \tag{2.73}$$

Since $K \ge k$ was arbitrary

$$\lim_{g\to\infty} \Delta_k(v_{S_g}) = r_k.$$

In view of relation (2.67), the existence of the limit of the $\Delta_l$s along the subsequence of $(S_k)_{k\ge 1}$ implies the existence of the limit along the whole sequence:

$$\lim_{i\to\infty} \Delta_k(v_i) = r_k.$$

Lemma 1 is proved.

As a corollary of lemma 1 and the remark immediately preceding it, we get

$$P(dim(A_r) \ge \beta) > 0.$$

However, a little though reveals that Kolmogorov zero-one law applies to the event $dim(A_r) \ge \beta$. Hence, $P(dim(A_r) \ge \beta) = 1$. Since $\beta < \log m_0 - H(r|q)$ was arbitrary,

$$dim(A_r) \ge \log m_0 - H(r|q)$$

almost surely, as desired.

**Proposition 20** *If $H(r|q) < \log m_0$ then for almost every Galton-Watson tree $T$*

$$dim(A_r) \le \log m_0 - H(r|q).$$

**Proof of proposition 20.** Pick an arbitrary $\epsilon > 0$. Find $\sigma$ and $K$ as in proposition 16. For a tree $T$ define $A_N \subset \delta T$ to be the set of rays $(v_i)_{i\ge 0}$, such that

$$|\Delta_k(v_N) - r_k| \le \sigma \text{ for all } k \le K.$$

It is evident that $A_N$ is a union of $h_N$ balls of radii $\exp(-N)$, where

$$h_N = \#\{v \in L_N \,:\, |\Delta_k(v) - r_k| \le \sigma \text{ for all } k \le K\}.$$

Moreover, $A_r \subset \limsup_{N\to\infty} A_N$. On the other hand, by proposition 16

$$\lim_{N\to\infty} \frac{\log h_N}{N} \le \log m_0 - H(r|q) + \epsilon.$$

Applying proposition 13 with $X = \delta T$, we obtain

$$dim(A_r) \le \log m_0 - H(r|q) + \epsilon.$$

Proposition 20 now follows from the fact the $\epsilon > 0$ was chosen arbitrarily.

# Appendix A

Here we give a cute proof of Proposition 10 due to Daniel Stroock.

Let $\mathbb{F}_n$ be the $\sigma$-field generated by the intervals $[k2^{-n}, (k+1)2^{-n}]$ with integer $k$. Define $\mu_n$ to be the restriction of $\mu$ to $F_n$. Then $\mu_n$ is absolutely continuous with respect to the restriction of Lebesgue measure to $\mathbb{F}_n$. We denote the corresponding Radon-Nikodim derivative $f_n(x)$. It is clear from (1.73) that

$$f_{n+1}(x) \leq (1 + \exp(-\theta n))f_n(x) + 2^{-\eta n} \tag{A.1}$$

for sufficiently large $n$. It is routine to infer from (A.1) that the sequence of functions $(f_n)_{n\geq 1}(x)$ is uniformly bounded, whence also uniformly integrable. By a standard result from the basic measure theory, $\lim_{n\to\infty} f_n(x)$ exists almost surely, is bounded and equal to the Radon-Nikodim derivative of $\mu$ with respect to Lebesgue measure. Proposition 10 follows.

# Bibliography


[1] Krishna Athreya, Peter Ney. (1972) *Branching Processes.* Springer, New York

[2] Jean-Dominique Deuschel, Daniel Stroock. (1989) *Large Deviations.* Academic Press, San Diego

[3] T. Hawkes. (1981) Trees generated by a simple branching process. *J. London Math. Soc.* **24** 373-384.

[4] S. Lalley, T. Sellke (2000) An extension of Hawkes' theorem on the Hausdorff dimension of a Galton-Watson tree. *Probab. Theory Relat. Fields* **116** 41-56.

[5] R. Lyons, R. Pemantle, and Y. Peres. (1995) Ergodic Theory on Galton-Watson trees: speed of random walk and dimension of harmonic measure. *Egrodic Theory Dyn. Systems* **15** 593-619.

[6] R. Lyons, R. Pemantle, and Y. Peres. (1996) Biased random walks on Galton-Watson trees. *Probab. Theory Relat. Fields* **106** 249-264.

[7] R. Lyons, R. Pemantle, and Y. Peres. (1997) Unsolved problems concerning random walks on trees. *Classical and Modern Branching Processes,* K. Athreya and P. Jagers (editors) 223-238. Springer, New York

[8] Valentin Petrov. (1975) *Sums of independent random variables.* Springer, New York

[9] B. Virag. (2000) On the speed of random walks on graphs. *Ann. Probab.* **28**, 379-394