Mixing time for a random walk on rooted trees

Jason Fulman

Department of Mathematics
University of Southern California, Los Angeles, CA, USA

fulman@usc.edu

Submitted: Aug 8, 2009; Accepted: Nov 3, 2009; Published: Nov 13, 2009
Mathematics Subject Classification: 60J10, 05E99

Abstract

We define an analog of Plancherel measure for the set of rooted unlabeled trees on n vertices, and a Markov chain which has this measure as its stationary distribution. Using the combinatorics of commutation relations, we show that order n^2 steps are necessary and suffice for convergence to the stationary distribution.

1 Introduction

The Plancherel measure of the symmetric group is a probability measure on the irreducible representations of the symmetric group which chooses a representation with probability proportional to the square of its dimension. Equivalently, the irreducible representations of the symmetric group are parameterized by partitions λ of n, and the Plancherel measure chooses a partition λ with probability

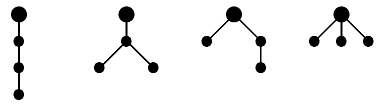
$$\frac{n!}{\prod_{x \in \lambda} h(x)^2} \tag{1}$$

where the product is over boxes in the partition and $h(x)$ is the hooklength of a box. The hooklength of a box x is defined as 1 + number of boxes in same row as x and to right of x + number of boxes in same column of x and below x. For example we have filled in each box in the partition of 7 below with its hooklength

$$\begin{array}{cccc} \boxed{6} & \boxed{4} & \boxed{2} & \boxed{1} \\ \boxed{3} & \boxed{1} & & \\ \boxed{1} & & & \end{array},$$

and the Plancherel measure would choose this partition with probability $\frac{7!}{(6\times4\times3\times2)^2}$. There has been significant interest in statistical properties of partitions chosen from Plancherel measure of the symmetric group; for this the reader can consult [4], [5], [11] and the many references therein.

In this paper we define a similar measure on the set of rooted, unlabeled trees on n vertices. We place the root vertex on top, and the four rooted trees on 4 vertices are depicted below:



This measure chooses a rooted tree with probability

$$\pi(t) = \frac{n \cdot 2^{n-1}}{|SG(t)| \prod_{v \in t} h(v)^2}, \tag{2}$$

where $h(v)$ is the size of the subtree with root v, and $|SG(t)|$ is a certain symmetry factor associated to the tree t (precise definitions are given in Section 3). We do not know that this measure has applications similar to the Plancherel measure of the symmetric group, but the resemblance is striking. Moreover, there are Hopf algebras in the physics literature whose generators are rooted trees (Kreimer's Hopf algebra [9],[21] a Hopf algebra of Connes and Moscovici [10], and a Hopf algebra of Grossman and Larson [18]), and as a paper of Hoffman [19] makes clear, the combinatorics of these Hopf algebras is very close to the combinatorics we use in this paper.

In fact the main object we study is a Markov chain K which has π as its stationary distribution; this Markov chain is defined in Section 3 and involves removing a single terminal vertex and reattaching it. There are several ways of quantifying the convergence rate of a Markov chain on a state space X to its stationary distribution; we use the maximal separation distance after r steps, defined as

$$s^*(r) := \max_{x,y \in X} \left[1 - \frac{K^r(x,y)}{\pi(y)} \right],$$

where $K^r(x,y)$ is the chance of transitioning from x to y after r steps. In general it can be quite tricky even to determine which x, y attain the maximum in the definition of $s^*(r)$. We do this, and prove that for $c > 0$ fixed,

$$\lim_{n \to \infty} s^*(cn^2) = \sum_{i=3}^{\infty} \frac{(-1)^{i-1}}{2} (2i-1)(i+1)(i-2) e^{-ci(i-1)}.$$

There are very few Markov chains for which such precise asymptotics are known. Our proof method uses a commutation relation of a growth and pruning operator on rooted trees (due to Hoffman [19]), a formula for the eigenvalues of K, and ideas from [15]. Details appear in Section 4.

We mention that the Markov chain K is very much in the spirit of the down-up chains (on the state space of partitions) studied in [6], [7], [15], [17], [22]. There are also similarities to certain random walks on phylogenetic trees (cladograms) studied in [1],

[14], [23]. Our methods only partly apply to these walks (the geometry of the two spaces of trees is different), so this will be studied in another work.

To close the introduction, we mention two reasons why it can be useful to understand a Markov chain K whose stationary distribution π is of interest. First, in analogy with Plancherel measure of the symmetric group, one can hope to use Stein's method ([17]) or other techniques ([6]) to study statistical properties of π. Second, convergence rates of K can lead to concentration inequalities for statistics of π [8].

2 Background on Markov chains

We will be concerned with the theory of finite Markov chains. Thus X will be a finite set (in our case the set of rooted unlabeled trees on n vertices) and K a matrix indexed by $X \times X$ whose rows sum to 1. Let π be a probability distribution on X such that K is reversible with respect to π; this means that $\pi(x)K(x,y) = \pi(y)K(y,x)$ for all x, y and implies that π is a stationary distribution for the Markov chain corresponding to K (i.e. that $\pi(x) = \sum_y \pi(y)K(y,x)$ for all x).

A common way to quantify convergence rates of Markov chains is to use separation distance, introduced by Aldous and Diaconis [2],[3]. They define the separation distance of a Markov chain K started at x as

$$s(r) = \max_y \left[1 - \frac{K^r(x,y)}{\pi(y)}\right]$$

and the maximal separation distance of the Markov chain K as

$$s^*(r) = \max_{x,y} \left[1 - \frac{K^r(x,y)}{\pi(y)}\right].$$

They show that the maximal separation distance has the nice properties:

- $\displaystyle \frac{1}{2} \max_x \sum_y |K^r(x,y) - \pi(y)| \leqslant s^*(r)$
- (monotonicity) $s^*(r_1) \leqslant s^*(r_2)$, $r_1 \geqslant r_2$
- (submultiplicativity) $s^*(r_1 + r_2) \leqslant s^*(r_1)s^*(r_2)$

3 Combinatorics of rooted trees

For a finite rooted tree t, we let $|t|$ denote the number of vertices of t; $\mathcal{T}_n$ will be the set of rooted unlabeled trees on n vertices. For example $\mathcal{T}_1 = \{\bullet\}$ consists of only the root vertex, and the four elements of $\mathcal{T}_4$ were depicted in the introduction. Letting $T_n = |\mathcal{T}_n|$ and $T_0 = 0$, there is a recursion

$$\sum_{n \geqslant 1} T_n \cdot x^n = x \prod_{n \geqslant 1} (1 - x^n)^{-T_n}$$

from which one obtains $T_1 = 1, T_2 = 1, T_3 = 2, T_4 = 4, T_5 = 9, T_6 = 20$, etc. (see [24] for more information on this sequence).

A rooted tree can be viewed as a directed graph by directing all edges away from the root, and a vertex is called terminal if it has no outgoing edge. There is a partial order $\preceq$ on the set $\mathcal{T}$ of all finite rooted trees defined by letting t be covered by t' exactly when t can be obtained from t' by removing a single terminal vertex and the edge into it; we denote this by $t \nearrow t'$ or $t' \searrow t$.

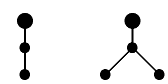
When $t \nearrow t'$, one can define two quantities

$$n(t,t') = |\text{vertices of t to which a new edge can be added to get t}'|$$

and

$$m(t,t') = |\text{edges of t}' \text{ which when removed give t}|.$$

These need not be equal, as can be seen by taking t, t' to be:



Then $n(t,t') = 1$ and $m(t,t') = 2$.

Let $\mathbb{C}\mathcal{T}_n$ denote the complex vector space with basis the elements of $\mathcal{T}_n$. For $n \geqslant 1$, Hoffman [19] defines a growth operator $G : \mathbb{C}\mathcal{T}_n \mapsto \mathbb{C}\mathcal{T}_{n+1}$ by

$$G(t) = \sum_{t' \searrow t} n(t,t')t',$$

and for $n \geqslant 2$ a pruning operator $P : \mathbb{C}\mathcal{T}_n \mapsto \mathbb{C}\mathcal{T}_{n-1}$ by

$$P(t) = \sum_{t' \nearrow t} m(t',t)t'.$$

One sets $P(\bullet) = 0$.

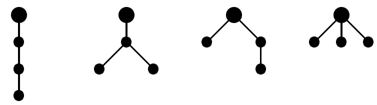
One can extend the definitions of $m(t,t')$ and $n(t,t')$ to any pair of rooted trees t, t' with $|t'| - |t| = k \geqslant 0$ by setting

$$G^k(t) = \sum_{|t'|=|t|+k} n(t,t')t'$$

and

$$P^k(t') = \sum_{|t|=|t'|-k} m(t,t')t.$$

Since $\bullet \preceq t$ for all t, one can think of $n(\bullet,t)$ as the number of ways to build up t, and of $m(\bullet,t)$ as the number of ways to take t apart by sequentially removing terminal edges. To simplify notation, we let $n(t) = n(\bullet,t)$ and $m(t) = m(\bullet,t)$. For example, the reader can check that the four trees t_1, t_2, t_3, t_4



satisfy $n(t_1) = 1, m(t_1) = 1; n(t_2) = 1, m(t_2) = 2; n(t_3) = 3, m(t_3) = 3; n(t_4) = 1, m(t_4) = 6$ respectively.

There is a "hook-length" type formula for $m(t)$ in the literature. Namely if t has n vertices,

$$m(t) = \frac{n!}{\prod_{v \in t} h(v)} \tag{3}$$

where $h(v)$ is the number of vertices in the subtree with root v; see Section 22 of [25] or Exercise 5.1.4-20 of [20] for a proof.

As for $n(t)$, it is also known as the "Connes-Moscovici weight" [21]. To give a formula for it, we use the concept of the symmetry group $SG(t)$ of a tree. For v a vertex of T with children $\{v_1, \cdots, v_k\}$, $SG(t, v)$ is the group generated by the permutations that exchange the trees with roots v_i and v_j when they are isomorphic rooted trees; then $SG(t)$ is defined as the direct product

$$SG(t) = \prod_{v \in T} SG(t, v).$$

It is proved in [21] that

$$n(t) = \frac{m(t)}{|SG(t)|}. \tag{4}$$

More generally, Proposition 2.5 of [19] shows that

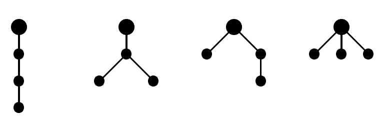
$$n(s,t)|SG(t)| = m(s,t)|SG(s)| \tag{5}$$

when $|s| \leqslant |t|$.

Definition 1 We define a probability measure π_n on the set of rooted (unlabeled) trees of size n by

$$\pi_n(t) = \frac{m(t)n(t)}{\prod_{i=2}^{n} \binom{i}{2}} = \frac{n \cdot 2^{n-1}}{|SG(t)| \prod_{v \in t} h(v)^2}. \tag{6}$$

It follows from Proposition 2.8 of [19] that π is in fact a probability measure (i.e. that the probabilities sum to 1). The second equality in (6) follows from equations (3) and (4). The reader can check that the four trees t_1, t_2, t_3, t_4



are assigned probabilities $1/18, 1/9, 1/2, 1/3$ respectively.

Definition 2 We define upward transition probabilities from $t \in \mathcal{T}_{n-1}$ to $t' \in \mathcal{T}_n$ by

$$P_u(t, t') = \frac{m(t,t')n(t')}{\binom{n}{2} n(t)} = \frac{n(t,t')m(t')}{\binom{n}{2} m(t)}$$

and downward transition probabilities from $t \in \mathcal{T}_n$ to $t' \in \mathcal{T}_{n-1}$ by

$$P_d(t,t') = \frac{m(t',t)m(t')}{m(t)}.$$

It is clear from the definitions that the downward transition probabilities sum to 1. The second equality in the definition of P_u is from (4) and (5), and it follows from Proposition 2.8 of [19] that the upward transition probabilities sum to 1. We define a "down-up" Markov chain with state space $\mathcal{T}_n$ by composing the down chain with the up chain, i.e.

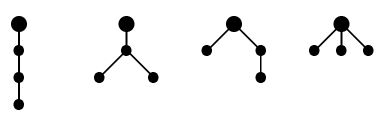
$$\begin{aligned} K(t,t') &= \sum_{s \nearrow t,t'} P_d(t,s)P_u(s,t') \\ &= \sum_{s \nearrow t,t'} \frac{m(s,t)m(s)}{m(t)} \frac{n(s,t')m(t')}{\binom{n}{2}m(s)} \\ &= \frac{m(t')}{\binom{n}{2}m(t)} \sum_{s \nearrow t,t'} m(s,t)n(s,t'). \end{aligned}$$

Thus we deduce the crucial relation

$$K_n = \frac{1}{\binom{n}{2}} AGPA_n^{-1} \tag{7}$$

where A is the diagonal matrix which multiplies a tree t by $m(t)$, and P, G are the pruning and growth operators. The subscript n indicates that the chain is on trees of size n.

For example, ordering the four elements of $\mathcal{T}_4$ as



one calculates the transition matrix

$$(K(t,t')) = \begin{pmatrix} 1/6 & 1/3 & 1/2 & 0 \\ 1/6 & 1/3 & 1/2 & 0 \\ 1/18 & 1/9 & 1/2 & 1/3 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix}.$$

The following lemma will be useful.

Lemma 3.1. *1. If s is chosen from the measure π_{n-1} and one moves from s to t with probability $P_u(s,t)$, then t is distributed according to the measure π_n.*

2. If t is chosen from the measure π_{n+1} and one moves from t to s with probability $P_d(t,s)$, then s is distributed according to the measure π_n.

3. The "down-up" Markov chain K_n on rooted trees of size n is reversible with respect to π_n.

Proof. For part 1, one calculates that

$$\begin{aligned}\sum_{s \nearrow t} \pi_{n-1}(s) P_u(s,t) &= \sum_{s \nearrow t} \frac{m(s)n(s)}{\prod_{i=2}^{n-1} \binom{i}{2}} \frac{n(s,t)m(t)}{m(s)\binom{n}{2}} \\ &= \frac{m(t)}{\prod_{i=2}^{n} \binom{i}{2}} \sum_{s \nearrow t} n(s)n(s,t) \\ &= \frac{m(t)n(t)}{\prod_{i=2}^{n} \binom{i}{2}} = \pi_n(t).\end{aligned}$$

For part 2, one computes that

$$\begin{aligned}\sum_{t \searrow s} \pi_{n+1}(t) P_d(t,s) &= \sum_{t \searrow s} \frac{m(t)n(t)}{\prod_{i=2}^{n+1} \binom{i}{2}} \frac{m(s,t)m(s)}{m(t)} \\ &= \frac{m(s)}{\prod_{i=2}^{n+1} \binom{i}{2}} \sum_{t \searrow s} n(t)m(s,t) \\ &= \frac{m(s)n(s)}{\prod_{i=2}^{n} \binom{i}{2}} = \pi_n(s),\end{aligned}$$

where the last line follows since the upward transition probabilities from s sum to 1.

For part 3, one calculates that

$$\begin{aligned}\pi_n(t)K(t,t') &= \frac{n(t)m(t')}{\binom{n}{2}\prod_{i=2}^{n} \binom{i}{2}} \sum_{s \nearrow t,t'} m(s,t)n(s,t') \\ &= \frac{n(t)m(t')}{\binom{n}{2}\prod_{i=2}^{n} \binom{i}{2}} |SG(t)| \sum_{s \nearrow t,t'} \frac{n(s,t)n(s,t')}{|SG(s)|} \\ &= \frac{n(t)m(t')}{\binom{n}{2}\prod_{i=2}^{n} \binom{i}{2}} |SG(t)| \sum_{s \nearrow t,t'} \frac{n(s,t)m(s,t')}{|SG(t')|} \\ &= \frac{n(t')m(t)}{\binom{n}{2}\prod_{i=2}^{n} \binom{i}{2}} \sum_{s \nearrow t,t'} n(s,t)m(s,t') \\ &= \pi_n(t')K(t',t).\end{aligned}$$

Note that equation (5) was used in equalities 2 and 3 and that equation (4) was used in the fourth equality. □

The final combinatorial fact we need about rooted trees is the following commutation relation between the growth and pruning operators (Proposition 2.2 of [19]) :

$$PG_n - GP_n = nI, \tag{8}$$

for all $n \geqslant 1$. Here I is the identity operator, so the right hand side multiplies a tree by its size.

4 Proof of main results

The purpose of this section is to obtain precise asymptotics for the maximal separation distance $s^*(r)$ of the Markov chain K after r iterations. To do this we use equation (7), the commutation relation (8), and the methodology of [15]. To begin we determine the eigenvalues of the Markov chain K. The multiplicities involve the numbers T_i of rooted unlabeled trees of size i, discussed in Section 3.

Proposition 4.1. *The eigenvalues of the Markov chain K are:*

$$\begin{array}{ll} 1 & \textit{multiplicity } 1 \\ 1-\frac{\binom{i}{2}}{\binom{n}{2}} & \textit{multiplicity } T_i - T_{i-1} \ (3 \leqslant i \leqslant n) \end{array}$$

Proof. Since $K_n = \frac{1}{\binom{n}{2}} AGPA_n^{-1}$, it suffices to determine the eigenvalues of GP_n; these follow from the commutation relation (8) and Theorem 2.6 of [26]. ☐

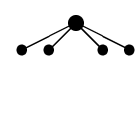
Recall that our interest is in studying the behavior of

$$s^*(r) = \max_{t,t'} \left[1 - \frac{K^r(t,t')}{\pi(t')}\right].$$

Proposition 4.2 determines the pairs (t,t') where this maximum is obtained.

Proposition 4.2. *For all values of r, the quantity $1 - \frac{K^r(t,t')}{\pi(t')}$ is maximized by letting t be the unique rooted tree with one terminal vertex and t' be the unique tree with $n-1$ terminal vertices, or by letting t' be the unique rooted tree with one terminal vertex and t be the unique tree with $n-1$ terminal vertices.*

For instance when $n = 5$ the two relevant trees are



Proof. By relation (7), we seek the t, t' minimizing

$$\frac{K^r(t,t')}{\pi(t')} = \frac{m(t')(GP)_n^r[t,t']}{\binom{n}{2}^r m(t)\pi(t')}.$$

By the commutation relation (8) and Proposition 4.5 of [15],

$$(GP)_n^r = \sum_{k=0}^{n} A_n(r,k) G^k P_n^k$$

where the $A_n(r,k)$ solve the recurrence

$$A_n(r,k) = A_n(r-1,k-1) + A_n(r-1,k)\left[\binom{n}{2} - \binom{n-k}{2}\right]$$

with initial conditions $A_n(0,0)=1$, $A_n(0,m)=0$ for $m\neq 0$. Thus

$$\frac{K^r(t,t')}{\pi(t')} = \frac{m(t')\sum_{k=0}^{n} A_n(r,k)\,G^k P_n^k[t,t']}{\binom{n}{2}^r m(t)\pi(t')}. \tag{9}$$

The proposition now follows from three observations:

- All terms in (9) are non-negative. Indeed, this is clear from the recurrence for $A_n(r,k)$.
- If t is the unique rooted tree with one terminal vertex and t' is the unique rooted tree with $n-1$ terminal vertices (or the same holds with t,t' swapped), then the summands in (9) for $0\leqslant k\leqslant n-3$ all vanish. Indeed, in order to move from t to t' by pruning k vertices and then reattaching them, one must prune at least $n-2$ vertices.
- The $k=n-2$ and $k=n-1$ summands in (9) are independent of t,t'. Indeed, for the $k=n-1$ summand, one has that

$$\begin{aligned}
& \frac{m(t')A_n(r,n-1)\,G^{n-1}P_n^{n-1}[t,t']}{\binom{n}{2}^r m(t)\pi(t')} \\
= \quad & \frac{m(t')A_n(r,n-1)\,G^{n-1}[\bullet,t']}{\binom{n}{2}^r \pi(t')} \\
= \quad & \frac{m(t')A_n(r,n-1)n(t')}{\binom{n}{2}^r \pi(t')} \\
= \quad & A_n(r,n-1)\frac{\prod_{i=2}^{n}\binom{i}{2}}{\binom{n}{2}^r}.
\end{aligned}$$

  A similar argument shows that the $k=n-2$ summand is equal to $A_n(r,n-2)\frac{\prod_{i=2}^{n}\binom{i}{2}}{\binom{n}{2}^r}$.

□

Remark: The proof of Proposition 4.2 shows that

$$s^*(r) = 1-\frac{\prod_{i=2}^{n}\binom{i}{2}}{\binom{n}{2}^r}\left[A_n(r,n-2)+A_n(r,n-1)\right],$$

where $A_n(r,k)$ is the solution to the recurrence in the proof of Proposition 4.2.

In Theorem 4.3, we give an explicit formula for $s^*(r)$ and determine its asymptotic behavior.

Theorem 4.3. *Let $s^*(r)$ be the maximal separation distance after r iterations of the down-up Markov chain K on the space of rooted trees on n vertices.*

1. *For $r \geqslant 1$, $s^*(r)$ is equal to*

$$\sum_{i=3}^{n-1} (-1)^{i-1} \frac{(2i-1)(i+1)(i-2)(n!)^2}{2n(n-i)!(n+i-1)!} \left(1 - \frac{\binom{i}{2}}{\binom{n}{2}}\right)^r.$$

2. *For $c > 0$ fixed,*

$$\lim_{n\to\infty} s^*(cn^2) = \sum_{i=3}^{\infty} \frac{(-1)^{i-1}}{2}(2i-1)(i+1)(i-2)e^{-ci(i-1)}.$$

Proof. By Proposition 4.2, the maximal separation distance is attained when t is the unique rooted tree with one terminal vertex and t' is the unique rooted tree with $n-1$ terminal vertices. Note that it takes $n-2$ iterations of the Markov chain K to move from t to t'. By Proposition 4.1, K has $n-1$ distinct eigenvalues (one more than the Markov chain distance between t and t'), so it follows from Proposition 5.1 of [16] that

$$s^*(r) = \sum_{i=3}^{n} \lambda_i^r \left[\prod_{j\neq i} \frac{1-\lambda_j}{\lambda_i - \lambda_j}\right], \tag{10}$$

where 1, $\lambda_i = 1 - \frac{\binom{i}{2}}{\binom{n}{2}}$, $i = 3, \cdots, n$ are the distinct eigenvalues of K. For $r \geqslant 1$, this is equal to

$$\sum_{i=3}^{n-1} \left(1 - \frac{\binom{i}{2}}{\binom{n}{2}}\right)^r \prod_{\substack{j\neq i\\ 3\leqslant j\leqslant n}} \frac{\binom{j}{2}}{\binom{j}{2} - \binom{i}{2}} = \sum_{i=3}^{n-1} \left(1 - \frac{\binom{i}{2}}{\binom{n}{2}}\right)^r \prod_{\substack{j\neq i\\ 3\leqslant j\leqslant n}} \frac{j(j-1)}{(j+i-1)(j-i)} \tag{11}$$

and the first assertion follows by elementary simplifications.

For part 2 of the theorem, it is enough to show that for $c > 0$ fixed, there is a constant i_c (depending on c but not on n) such that for $i \geqslant i_c$, the summands in part 1 of the theorem are decreasing in magnitude (and alternating in sign). Part 2 follows from this claim, since then one can take limits for each fixed i. For $i \geqslant 2\sqrt{n}$ one checks that

$$\frac{(2i-1)(i+1)(i-2)(n!)^2}{2n(n-i)!(n+i-1)!}$$

is a decreasing function of i. To handle the case of $i \leqslant 2\sqrt{n}$, one need only show that

$$\frac{(n-i)(2i+1)(i+2)(i-1)}{(n+i)(2i-1)(i+1)(i-2)} \frac{\exp(cn^2\log(1-\binom{i+1}{2}/\binom{n}{2}))}{\exp(cn^2\log(1-\binom{i}{2}/\binom{n}{2}))} < 1 \tag{12}$$

for $i \geqslant i_c$, a constant depending on c but not on n. This is easily established, since using the inequalities $\log(1-x) \leqslant -x$ for $x > 0$ in the numerator and $\log(1-x) \geqslant -x - x^2$ for $0 < x < \frac{1}{2}$ in the denominator gives that

$$\frac{\exp(cn^2\log(1-\binom{i+1}{2}/\binom{n}{2}))}{\exp(cn^2\log(1-\binom{i}{2}/\binom{n}{2}))} \leqslant \exp\left[\frac{-cn^2}{\binom{n}{2}}\left(i - \frac{\binom{i}{2}^2}{\binom{n}{2}}\right)\right],$$

and (12) follows as $i \leqslant 2\sqrt{n}$. □

Some authors who work on Markov chains similar to that studied here but on different state spaces (e.g. [7], [22]) prefer to work with up-down chains instead of down-up chains. Proposition 4.4 shows the study of maximal separation for these two chains to be equivalent.

Proposition 4.4. *Let $s^*_{UD_n}(r)$ denote the maximal separation distance after r iterations of the down-up chain on $\mathcal{T}_n$, and let $s^*_{DU_n}(r)$ be the corresponding quantity for the up-down chain. Then*

$$s^*_{DU_n}(r) = s^*_{UD_{n+1}}(r+1)$$

for all $n, r \geqslant 1$.

Proof. An argument similar to that used to prove equation (7) gives that

$$DU_n = \frac{1}{\binom{n+1}{2}} APGA_n^{-1} \tag{13}$$

where A is the diagonal matrix which multiplies a tree by $m(t)$, and P, G are the pruning and growth operators. Combining this with the commutation relation (8), it follows that

$$\begin{aligned} (DU)_n^r &= \frac{1}{\binom{n+1}{2}^r} [A(nI + GP)A_n^{-1}]^r \\ &= \frac{1}{\binom{n+1}{2}^r} \sum_{l=0}^{r} \binom{r}{l} n^{r-l} A(GP)^l A_n^{-1}. \end{aligned}$$

Arguing as in Proposition 4.2, one concludes that the same t, t' maximize the separation distance. Moreover, one sees from (13), commutation relation (8), and Proposition 4.1 that the distinct eigenvalues of the up-down chain on trees of size n are 1 and $\mu_i = 1 - \frac{\binom{i}{2}}{\binom{n+1}{2}}$, $i = 3, \cdots, n$. Thus the argument of Theorem 4.3 gives that

$$s^*_{DU_n}(r) = \sum_{i=3}^{n} \left(1 - \frac{\binom{i}{2}}{\binom{n+1}{2}}\right)^r \prod_{\substack{j \neq i \\ 3 \leqslant j \leqslant n}} \frac{\binom{j}{2}}{\binom{j}{2} - \binom{i}{2}}.$$

The proposition now follows by making the replacements $r \to r+1$ and $n \to n+1$ in the left hand side of equation (11); indeed, the $j = n+1$ term is then equal to $\left(1 - \binom{i}{2} / \binom{n+1}{2}\right)^{-1}$. ∎

To close, we note the following probabilistic interpretation of $s^*(r)$. We use the convention that a random variable X is called geometric with parameter (probability of success) p if $\mathbb{P}(X = n) = p(1-p)^{n-1}$ for all $n \geqslant 1$.

Proposition 4.5. *Letting $s^*(r)$ be as in Theorem 4.3, one has that $s^*(r) = \mathbb{P}(T > r)$, where $T = \sum_{i=3}^{n} X_i$, and the X_i's are independent geometrics with parameters $\frac{\binom{i}{2}}{\binom{n}{2}}$.*

Proof. This is immediate from equation (10) and Proposition 2.4 of [15]. □

We remark that representations of separation distance similar to that in Proposition 4.5 are in the literature for stochastically monotone birth-death chains with non-negative eigenvalues ([12], [13]) and for some random walks on partitions [15]. Of course the Markov chain K studied in this paper is not one-dimensional.

Acknowledgments

The author was supported by NSA grant H98230-08-1-0133 and NSF grant DMS 0802082. We thank Persi Diaconis for pointers to the literature.

References

[1] Aldous, D., Mixing time for a Markov chain on cladograms, *Combin. Probab. Comput.* **9** (2000), 191-204.

[2] Aldous, D. and Diaconis, P., Shuffling cards and stopping times, *Amer. Math. Monthly* **93** (1986), 333-348.

[3] Aldous, D. and Diaconis, P., Strong uniform times and finite random walks, *Adv. in Appl. Math.* **8** (1987), 69-97.

[4] Aldous, D. and Diaconis, P., Longest increasing subsequences: from patience sorting to the Baik-Deift-Johansson theorem, *Bull. Amer. Math. Soc. (N.S.)* **36** (1999), 413-432.

[5] Borodin, A., Okounkov, A., and Olshanski, G., Asymptotics of Plancherel measures for symmetric groups, *J. Amer. Math. Soc.* **13** (2000), 481-515.

[6] Borodin, A. and Olshanski, G., Markov processes on partitions, *Probab. Theory Related Fields* **135** (2006), 84-152.

[7] Borodin, A. and Olshanski, G., Infinite-dimensional diffusions as limits of random walks on partitions, *Probab. Theory Related Fields* **144** (2009), 281-318.

[8] Chatterjee, S., Concentration of Haar measures, with an application to random matrices, *J. Funct. Anal.* **245** (2007), 379-389.

[9] Connes, A. and Kreimer, D., Hopf algebras, renormalization and noncommutative geometry, *Comm. Math. Phys.* **199** (1998), 203-242.

[10] Connes, A. and Moscovici, H., Hopf algebras, cyclic cohomology and the transverse index theorem, *Comm. Math. Phys.* **198** (1998), 199-246.

[11] Deift, P., Integrable systems and combinatorial theory, *Notices Amer. Math. Soc.* **47** (2000), 631-640.

[12] Diaconis, P. and Fill, J., Strong stationary times via a new form of duality, *Ann. Probab.* **18** (1990), 1483-1522.

[13] Diaconis, P. and Saloff-Coste, L., Separation cutoffs for birth death chains, *Ann. Appl. Probab.* **16** (2006), 2098-2122.

[14] Ford, D., *Probabilities on cladograms: introduction to the alpha model*, Ph.D. thesis, Stanford University, 2006.

[15] Fulman, J., Commutation relations and Markov chains, *Probab. Theory Related Fields* **144** (2009), 99-136.

[16] Fulman, J., Separation cutoffs for random walk on irreducible representations, to appear in *Ann. Comb.*, arXiv: math.PR/0703921 (2007).

[17] Fulman, J., Stein's method and random character ratios, *Trans. Amer. Math. Soc.* **360** (2008), 3687-3730.

[18] Grossman, R. and Larson, R., Hopf-algebraic structure of families of trees, *J. Algebra* **126** (1989), 184-210.

[19] Hoffman, M., Combinatorics of rooted trees and Hopf algebras, *Trans. Amer. Math. Soc.* **355** (2003), 3795-3811 (electronic).

[20] Knuth, D., *The art of computer programming. Volume 3. Sorting and searching.* Addison-Wesley Series in Computer Science and Information Processing. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1973.

[21] Kreimer, D., Chen's iterated integral represents the operator product expansion, *Adv. Theor. Math. Phys.* **3** (1999), 627-670.

[22] Petrov, L., A two-parameter family of infinite-dimensional diffusions in the Kingman simplex, arXiv:0708.1930.

[23] Schweinsberg, J., An $O(n^2)$ bound for the relaxation time of a Markov chain on cladograms, *Random Structures Algorithms* **20** (2002), 59-70.

[24] Sloane, N., Online encyclopedia of integer sequences, Sequence A000081, www.research.att.com/∼njas/sequences.

[25] Stanley, R., *Ordered structures and partitions*, Memoirs Amer. Math. Soc. 119. American Mathematical Society, Providence, R.I., 1972.

[26] Stanley, R., Variations on differential posets, in *Invariant theory and tableaux*, 145-165, IMA Vol. Math. Appl., 19, Springer, New York, 1990.