# Tree/Endofunction Bijections and Concentration Inequalities

Steven Heilman[*]

Department of Mathematics
University of Southern California
Los Angeles, CA 90089-2532, U.S.A.

stevenmheilman@gmail.com

Submitted: Jul 9, 2021; Accepted: Mar 15, 2022; Published: May 20, 2022

© The author. Released under the CC BY-ND license (International 4.0).

**Abstract**

We demonstrate a method for proving precise concentration inequalities in uniformly random trees on $n$ vertices, where $n \geqslant 1$ is a fixed positive integer. The method uses a bijection between mappings $f\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$ and doubly rooted trees on $n$ vertices. The main application is a concentration inequality for the number of vertices connected to an independent set in a uniformly random tree, which is then used to prove partial unimodality of its independent set sequence. While inequalities for random trees often use combinatorial arguments, our argument is perhaps more probabilistic.

**Mathematics Subject Classifications:** 60C05,05C80,60E15,05C69

## 1 Introduction

Let $f\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$ be a mapping with associated directed graph $G(f)$ with vertices $\{1,\ldots,n\}$ and directed edges $\{(x, f(x))\colon 1 \leqslant x \leqslant n\}$. It is well known that $G(f)$ can be written as a union of trees connected to the cycles of $G(f)$. Deleting or rearranging some edges within the cycles of $G(f)$ can then produce a tree. For example, we could remove one edge from each cycle in $G(f)$ and then string each of these cycles together, while maintaining the structure of all non-cyclic vertices (see Figure 2 below). If the cycles are connected in a reversed order according to their smallest elements (as in Figure 3), then we get a bijection $R$ between mappings $f\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$ and doubly rooted trees on $n$ vertices. The two roots in a tree correspond to the beginning and the end of the path of cycles, respectively. This bijection requires deleting about $\log n$ edges from

[*]Supported by NSF Grants DMS 1839406 and CCF 1911216.

the mapping directed graph $G(f)$, with high probability (see Lemma 21). Consequently, random quantities depending on edges in $G(f)$ are essentially the same after the bijection $R$ is applied to $f$. So, e.g. a concentration inequality for $G(f)$ depending on edges applies essentially unchanged to the image of $f$ under $R$. Thus, the existence of the bijection $R$ (Theorem 1) can lead to concentration inequalities for random trees (Lemma 2).

The main point of this paper is a demonstration of a method for proving concentration inequalities on random trees by first proving an inequality for random mappings and then transferring that inequality to random trees via the bijection $R$.

### 1.1 History of the Rényi-Joyal Bijection

Joyal's bijection [15] between mappings from $\{1, \ldots, n\}$ to itself and doubly rooted trees on $n$ vertices used any bijection between linear orders and permutations, when specifying the action of the mapping bijection on the core of the mapping. That is, if $S_n$ denotes the set of permutations on $n$ elements, then any bijection $\pi\colon S_n \to S_n$ yields a corresponding lifted bijection $R_\pi$ between mappings from $\{1, \ldots, n\}$ to itself and doubly rooted trees on $n$ vertices. Choosing $\pi$ to be the Rényi bijection [18, page 11] between linear orderings and permutations, described in the previous paragraph by arranging the cycles of a permutation in reverse order of their smallest elements, seems most natural. Rényi's bijection $\pi$ [18] is often attributed to Foata [9], e.g. it is referred to as Foata's transition lemma in https://en.wikipedia.org/wiki/Permutation, though Rényi's preceding publication [18] was pointed out by Stanley [21, page 106]. Given a doubly rooted tree $(T, r_1, r_2)$, with $r_1, r_2 \in \{1, \ldots, n\}$, let $\rho(T, r_1, r_2) := T$ be the un-rooted tree $T$. We then study the properties of un-rooted tree $\rho(R_\pi(f))$.

By adding an additional randomization to the Joyal bijection, the authors of [2, 4] (and also [3]) define a coupling between random walks on mapping directed graphs and random walks on trees. Their coupling also works for non-uniform distributions on the set of functions $f\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$.

### 1.2 Concentration Inequalities on Random Trees

Our main goal is a concentration inequality on random trees with the best possible constants. Some of these concentration inequalities appear in the literature, albeit with sub-optimal constants for our particular application. Alternative approaches to proving concentration inequalities on random trees include the following non-exhaustive list:

- Give direct bounds on a moment generating function [23, 19], to e.g. deduce a central limit theorem,

- Using extensions of bounded difference inequalities to dependent random variables, such as [17, Theorem 3.7], or

- Using martingales [12] (e.g. the Aldous-Broder algorithm [6, 1]) and the Azuma-Hoeffding inequality, Lemma 18.

In [23, 19], a moment generating function is manipulated to obtain bounds on all moments of a so-called additive tree parameter, and a central limit theorem follows from these moment bounds. The bounds on the moment generating function itself might imply a concentration inequality we could use, though it is unclear to the author if the method of [23, 19] extends to the function on trees we consider in this work.

The constants that appear in bounded difference inequalities such as McDiardmid's inequality are far from optimal in our application, so that approach seems unnatural for proving our desired result. The latter approach proves concentration, but it is often sub-optimal, since e.g. the Azuma-Hoeffding inequality is not sharp for random quantities with small expected value. One might hope to somehow use Talagrand's convex distance inequality [14, Theorem 2.29] in place of the Azuma-Hoeffding inequality in order to improve the constants in the inequality, but Talagrand's inequality requires independence, so it might not be clear how to apply this inequality to random trees.

In [12], martingales are used to prove a Central Limit Theorem for Lipschitz tree parameters, and an inverse polynomial type error bound for this central limit theorem is given. Since we require concentration inequalities with exponential bounds rather than inverse polynomial bounds, it is unclear whether or not the argument of [12] could be adapted to our setting.

For more on random mappings, see [22, 16] and the references therein.

### 1.3 Our Contribution

Below, we let $E(\cdot)$ denote the set of undirected edges of a graph, we let $\Delta$ denote the symmetric difference of sets, and we let $c(f)$ denote the number of cycles in the mapping directed graph of $f\colon \{1,\dots,n\} \to \{1,\dots,n\}$.

**Theorem 1.** *There exists a bijection $R$ from the set*

$$\{f\colon \{1,\dots,n\} \to \{1,\dots,n\}\}$$

*to the set of doubly rooted trees on $n$ vertices such that for all maps $f$*

$$|E(R(f)) \,\Delta\, E(G(f))| \leqslant 2c(f).$$

### 1.4 The Main Application

The following Chernoff-type bounds are used in [11] to prove partial unimodality of the independent set sequence of uniformly random labelled trees, with high probability as the number of vertices $n$ goes to infinity. An independent set in a graph is a subset of vertices no two of which are connected by an edge.

**Lemma 2** (**Main Application to Independent Sets**)**.** *Let $S \subseteq \{1,\dots,n\}$ with $|S| = o((n-|S|)^2)$. Let $T$ be a uniformly random tree on $n$ vertices, conditioned on $S$ being an independent set. Let $N$ be the number of vertices in $S^c$ not connected to $S$. Let $\alpha := |S|/n$. Then*

$$\mathbb{P}(|N - \mathbb{E}N| > s\mathbb{E}N + 1) \leqslant e^{-\min(s,s^2)n(1-\alpha)^2 e^{-\alpha/(1-\alpha)}/3}, \qquad \forall\ s > 0.$$

*More generally,*

$$\mathbb{P}(N < (1-s)\mathbb{E}N - 1) \leqslant e^{-s^2 n(1-\alpha)^2 e^{-\alpha/(1-\alpha)}/2}, \qquad \forall\ 0 < s < 1,$$
$$\mathbb{P}(N > (1+s)\mathbb{E}N + 1) \leqslant e^{-s^2 n(1-\alpha)^2 e^{-\alpha/(1-\alpha)}/(2+s)}, \qquad \forall\ s \geqslant 0.$$

When $\alpha \in (0,1)$ is fixed, we have $|S| = \alpha n$ as $n \to \infty$ in Lemma 2. In the asymptotic regime where $\alpha \to 1^-$, Lemma 2 is vacuous.

### 1.5 Sketch of the Proof of Lemma 2

Lemma 2 is proven using the following general strategy.

- Begin with a random variable $N(G(f))$ defined on random mappings

$$f\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$$

  that is a function of the edges in the mapping directed graph $G(f)$ of $f$.

- Apply the (variant of the) Joyal bijection $R$ to $f$, to obtain a doubly rooted tree $R(f)$ on $n$ vertices and a corresponding random variable $N(R(f))$ defined now on $R(f)$.

- Theorem 1 says that the value of the random variables $N(G(f))$ and $N(R(f))$ do not differ too much. (In our particular example, these two quantities differ by at most 1, and this accounts for adding and subtracting some 1's in Lemma 2.)

- By the previous item, if $N(G(f))$ satisfies a concentration inequality, then so does $N(R(f))$. Lemma 2 is the resulting concentration inequality for $N(G(f))$. (In Lemma 2, $N(G(f))$ satisfies Chernoff bounds as in Lemma 17 by some properties of negatively associated random variables.)

So, one could generalize Lemma 2 in the following way. Suppose $N(G(f))$ is a function of the edges of a graph satisfying a concentration inequality, where $f\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$ is a random mapping. Then the doubly rooted random tree on $n$ vertices satisfies the same concentration inequality, up to an additive difference between $N(G(f))$ and $N(R(f))$ (which is bounded in the worst case by the number of cycles in $f$, by Theorem 1.) To this end, we also present a classical concentration inequality for the number of cycles of a random mapping, in Lemma 21. So, in this generality, we could replace the 1's appearing in Lemma 2 with $\log n$ terms, if we are willing to have a different exponential term in Lemma 2 as well, coming from Lemma 21.

### 1.6 Organization

- Theorem 1 is stated and proved as Lemma 9 below.

- Joyal's original bijection is presented in Lemma 6.

- A restriction of Theorem 1 is required to prove the main application, Lemma 2. That is, we need restrict $R^{-1}$ to doubly rooted trees such that a fixed set of vertices $S$ is an independent set. This restriction is demonstrated in Lemma 11.
- Section 4 lists some concentration inequalities cited elsewhere in this paper.
- Section 5 gives an algorithmic interpretation of Theorem 1. That is, we specify an algorithm for sampling uniformly random labelled trees on $n$ vertices.
- The Appendix, Section 6, proves some concentration inequalities for the number of cycles in a random map. Theorem 1 and Lemma 2 are proven without Section 6. Both Lemmas 21 and 24 explain the comment from the introduction that the tree/endofunction bijection deletes about $\log n$ edges from the mapping directed graph, with high probability. This property is rather crucial, since deleting too many edges during the bijection can ruin the transfer of the concentration inequality between the trees and endofunctions.

# 2 Tree/Endofunction Bijections

As a warmup to the more technical Section 2.3, we present the Joyal bijection and the Renyi-Joyal bijection. In fact, we will use the details of the present section during the more technical Section 2.3.

Let $n$ be a positive integer. We refer to a function $f\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$ as a **mapping** or **endofunction**. A **tree** on $n$ labelled vertices $\{1,\ldots,n\}$ is a connected, undirected graph with no cycles and no self-loops. A **doubly rooted tree** is a tree together with an ordered pair of roots $(r_1,r_2) \in \{1,\ldots,n\}^2$.

## 2.1 Joyal Bijection

We now describe the concepts from the introduction more precisely.

**Definition 3** (**Core**)**.** Let $f\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$. For any integer $j \geqslant 1$ let $f^j$ denote the composition of $f$ with itself $j$ times. Define the **core** of the mapping $f$ to be

$$\mathcal{M} = \mathcal{M}(f) := \{x \in \{1,\ldots,n\} \colon \exists\, j \geqslant 1 \text{ such that } f^j(x) = x\}.$$

**Definition 4** (**Mapping Directed Graph**)**.** Let $f\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$. Define the directed edge set

$$E(f) := \{(x, f(x)) \colon x \in \{1,\ldots,n\}\}.$$

The directed graph $(\{1,\ldots,n\}, E(f))$ is called the **mapping directed graph** of $f$. This graph has $c(f)$ cycles, where $c(f)$ is the number of cycles of the permutation $f|_{\mathcal{M}}$ (using Lemma 5).

When $T$ is a tree, we let $E(T)$ denote the set of (undirected) edges of the tree.

**Lemma 5.** *Let $f\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$. Then $f|_{\mathcal{M}}$ is a permutation on $\mathcal{M}$.*

*Proof.* Denote $\ell := f|_{\mathcal{M}}$. We denote gcd as the greatest common divisor of a set of positive integers. If $\ell(x) = \ell(y)$ for some $x, y \in \mathcal{M}$, then let $j := \gcd\{q \geqslant 1\colon \ell^q(x) = x\}$, and let $k := \gcd\{q \geqslant 1\colon \ell^q(y) = y\}$. Without loss of generality, $j \leqslant k$. Then $\ell^j(x) = \ell^j(y) = x$. Applying $\ell^{k-j}$ to both sides gives $\ell^k(x) = \ell^k(y) = y = \ell^{k-j}(x) = \ell^{k-j}(y)$, implying that $k = j$ by minimality of $k$, so that $x = y$. That is, $\ell$ is injective. If $x \in \mathcal{M}$, then $\ell(f^{j-1}(x)) = x$, so that $\ell$ is surjective. (We set $f^0(x) := x$.) $\square$

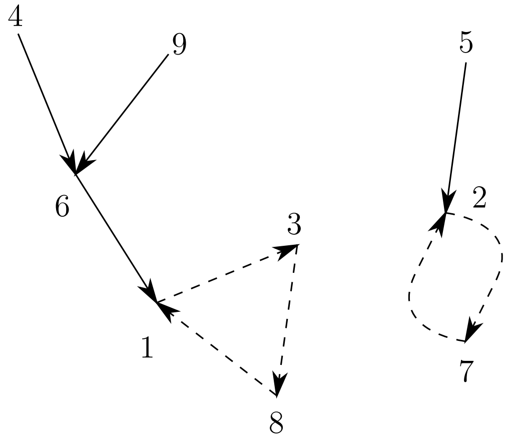


Figure 1: Example of the mapping directed graph of $f$. In this example, $f(1) = 3$, $f(2) = 7$, $f(3) = 8$, $f(4) = 6$, $f(5) = 2$, $f(6) = 1$, $f(7) = 2$, $f(8) = 1$ and $f(9) = 6$. Also $\mathcal{M} = \{1, 2, 3, 7, 8\}$.

Since the original Joyal bijection was described in French [15], we present it below for completeness. This bijection will then be improved in Lemma 9 below. (A partial translation and commentary of [15] is available at http://ozark.hendrix.edu/~yorgey/pub/series-formelles.pdf)

**Lemma 6** (**Joyal Bijection**, [15])**.** *There exists a bijection $J$ from the set of mappings $\{f\colon \{1,\ldots,n\} \to \{1,\ldots,n\}\}$ to the set of doubly rooted trees on $n$ vertices such that*

$$|E(f)\,\Delta\, E(J(f))| \leqslant 2\,|\mathcal{M}(f)| - 1.$$

*(When we compute this symmetric difference, we remove the directions on the edges $E(f)$, and we count multiple edges from $E(f)$ as distinct.)*

*Proof.* Let $f\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$. Let $\mathcal{M}$ be the core of $f$, as in Definition 3. Let $m := |\mathcal{M}|$. Denote $\mathcal{M} =: \{s_1, s_2, \ldots, s_m\}$ such that $s_1 < s_2 < \cdots < s_m$. Consider the undirected graph on the vertices $V := \{1,\ldots,n\}$ with edge set

$$E := \{\{x, f(x)\}\colon x \in \mathcal{M}^c\}. \tag{1}$$

(By definition of $\mathcal{M}^c$ in Definition 3, these edges are all distinct.) Since $\mathcal{M}^c = \{x \in \{1,\ldots,n\}\colon \forall\, j \geqslant 1\ f^j(x) \neq x\}$, $\forall\, x \in \mathcal{M}^c$, $\exists\, y \in \mathcal{M}, j \geqslant 1$ such that $f^j(x) = y$. For any $x \in \mathcal{M}^c$, let $j(x)$ denote the smallest positive integer $j$ such that there exists $y \in \mathcal{M}$ with $f^j(x) = y$. For any $y \in \mathcal{M}$, let $T_y := \{x \in \mathcal{M}^c\colon f^{j(x)}(x) = y\}$. Then $\mathcal{M}^c$ is a disjoint union $\cup_{y\in\mathcal{M}} T_y$. For any $y \in \mathcal{M}$, the edge set $\{\{x, f(x)\}\colon x \in T_y\}$ forms a (possibly empty) tree. That is, $\cup_{y\in\mathcal{M}} T_y$ is a disjoint union of $m$ trees. Consider now the edge set

$$E' := E \cup \bigcup_{i=1}^{m-1} \{f(s_i), f(s_{i+1})\}. \tag{2}$$

(If $m=1$, let $E' := E$.) (Recall that $E$ defined in (1) are disjoint edges, and by definition of $\mathcal{M}$ and Lemma 5, all edges in (2) are distinct.) The graph $T=(V,E')$ is then a (connected) tree with $n-1$ edges, and $n$ vertices. More specifically, $T$ is $m$ trees $\cup_{y\in\mathcal{M}}T_y$ connected along the path $\{f(s_1),\ldots,f(s_m)\}$ of length $m-1$. We define the **Joyal bijection** $J\colon$ {mappings} $\to$ {doubly rooted trees} by

$$J(f) := (T, f(s_1), f(s_m)).$$

It remains to show that $J$ is in fact a bijection.

**Proof of injectivity of $J$.** Let $f,g\colon\{1,\ldots,n\}\to\{1,\ldots,n\}$. Let $\mathcal{M},\mathcal{N}$ be the cores of $f$ and $g$, respectively. Suppose $J(f)=J(g)$. By definition of $J$ (i.e (1)), we have $\mathcal{M}=\mathcal{N}$, $\mathcal{M}^c=\mathcal{N}^c$, and $f|_{\mathcal{M}^c}=g|_{\mathcal{N}^c}$. Also by definition of $J$ (i.e. (2)), $f|_{\mathcal{M}}=g|_{\mathcal{M}}$. That is, $f=g$.

**Proof of surjectivity of $J$.** Let $(T,r_1,r_2)$ be a doubly rooted tree. Form the unique path $p$ of vertices in $T$ starting at $r_1$ and ending at $r_2$. Let $T' := T$. Repeat the following procedure until $T'\setminus p=\varnothing$:

- Choose one $x\in T'\setminus p$ of degree 1, and define $f(x)$ to be the label of the unique vertex connected to $x$.
- Re-define $T'$ by removing from $T'$ the vertex $x$ and the edge emanating from $x$.

In this way, $f(x)$ is defined for all $x\in T\setminus p$. Now, we define $f$ on $p$. Label the elements of $p$ in the order they appear in the path as $r_1=x_1,x_2,x_3,\ldots,x_m=r_2$. Let $\ell\colon\{x_1,\ldots,x_m\}\to\{x_1,\ldots,x_m\}$ be the permutation defined so that $x_{\ell(1)}<x_{\ell(2)}<\cdots<x_{\ell(m)}$. Then define $f$ so that

$$f(x_{\ell(i)}) := x_i, \qquad \forall\, 1\leqslant i\leqslant m.$$

Then $J(f)=(T,r_1,r_2)$ by (2), so that $J$ is surjective.

Finally, comparing Definition 4 with (2) proves the desired inequality. □

**Example 7.** Let $f\colon\{1,\ldots,n\}\to\{1,\ldots,n\}$ be such that $f(i)=i$ for all $1\leqslant i\leqslant n$. Then $\mathcal{M}=\{1,\ldots,n\}$, and $J(f)$ is the path that respects the ordering on $\{1,2,3,\ldots,n\}$. The roots are $r_1=\{1\}$ and $r_2=\{n\}$.

**Example 8.** Let $f\colon\{1,\ldots,n\}\to\{1,\ldots,n\}$ be such that $f(i)=1$ for all $1\leqslant i\leqslant n$. Then $\mathcal{M}=\{1\}$, and $J(f)$ is a star graph with a single vertex of degree $n-1$. The roots are $r_1=r_2=\{1\}$.

### 2.2 Rényi-Joyal Bijection

For the Joyal bijection of Lemma 6, $|E(f)\,\Delta\,E(J(f))|$ is approximately $\sqrt{n}$ with high probability (with respect to a uniformly random choice of $f\colon\{1,\ldots,n\}\to\{1,\ldots,n\}$), since the core of a random mapping is of size approximately $\sqrt{n}$ with high probability

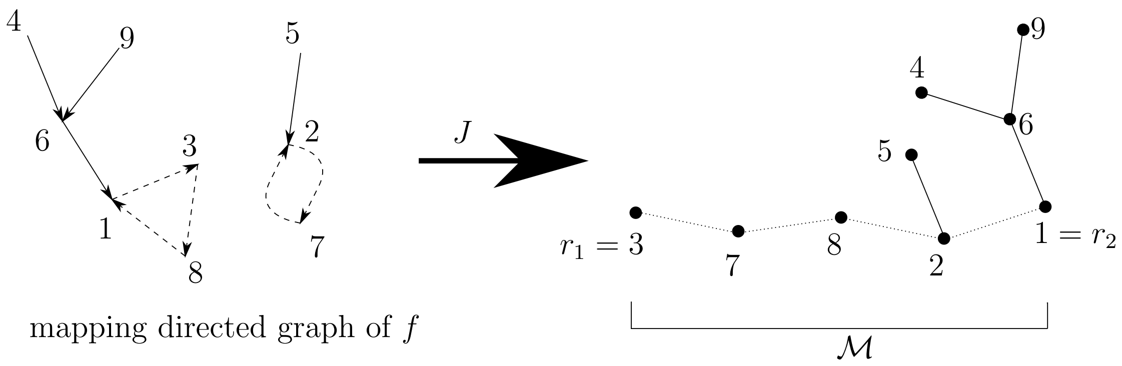


Figure 2: Example of the Joyal bijection. In this example, $f(1) = 3$, $f(2) = 7$, $f(3) = 8$, $f(7) = 2$, and $f(8) = 1$. So, the chosen order of the core $\mathcal{M} = \{1, 2, 3, 7, 8\}$ in $J(f)$ is $3, 7, 8, 2, 1$. Also, $r_1 = f(1) = 3$, $r_2 = f(8) = 1$.

[10]. Our ultimate goal is to prove a concentration inequality for a random mapping $f$, and then transfer it to a concentration inequality for a random tree. So, it is most desirable to have a bijection $R$ such that $|E(f) \,\Delta\, E(R(f))|$ as small as possible. Using Rényi's bijection within Joyal's bijection, it is possible to design a bijection $R$ satisfying $|E(f) \,\Delta\, E(R(f))| \leqslant 3 \log n$ with high probability, as we describe in Lemma 9 below. The idea is: the number of cycles of a random mapping is of size about $\log n$, so removing one edge from each cycle of the random mapping only changes about $\log n$ edges. In contrast, the Joyal bijection could change essentially all edges in the core, resulting in about $\sqrt{n}$ edge changes.

**Lemma 9** (**Rényi-Joyal Bijection**). *There exists a bijection $R$ from the set*

$$\{f \colon \{1, \ldots, n\} \to \{1, \ldots, n\}\}$$

*to the set of doubly rooted trees on $n$ vertices such that for all maps $f$,*

$$|E(f) \,\Delta\, E(R(f))| = 2c(f) - 1.$$

*(When we compute this symmetric difference, we remove the directions on the edges $E(f)$, and we count multiple edges from $E(f)$ as distinct.)*

*Proof.* We first repeat the definitions and reasoning from the proof of Lemma 6 before (2), including the definitions of $f, \mathcal{M}, m, E$ and $j(x)$.

From Lemma 5, recall that $f|_{\mathcal{M}}$ is a permutation on $\mathcal{M}$. We can then write $\mathcal{M} = \cup_{i=1}^{c(f)} \mathcal{M}_i$, where $\mathcal{M}_1, \ldots, \mathcal{M}_{c(f)}$ are subsets of vertices corresponding to the disjoint cycles of $f|_{\mathcal{M}}$. For each $1 \leqslant i \leqslant c(f)$, denote $\mathcal{M}_i = \{m_{i1}, \ldots, m_{ik(i)}\}$, where $k(i) := |\mathcal{M}_i|$, $m_{i1}$ is the smallest element of $\mathcal{M}_i$, and $m_{i(j+1)} = f(m_{ij})$ for all $1 \leqslant j < k(i)$. [That is, we can write $\mathcal{M}_i$ in cycle notation as $(m_{i1} \cdots m_{ik(i)})$, for all $1 \leqslant i \leqslant c(f)$.] We also choose the ordering on $\mathcal{M}_1, \ldots, \mathcal{M}_{c(f)}$ such that $m_{i1} > m_{(i+1)1}$ for all $1 \leqslant i \leqslant c(f) - 1$. [That is, we order the cycles in the reverse order of their smallest elements.] Consider now the edge set

$$E' := E \cup \Big( \bigcup_{i=1}^{c(f)} \bigcup_{j=1}^{k(i)-1} \{m_{ij}, m_{i(j+1)}\} \Big) \cup \bigcup_{i=1}^{c(f)-1} \{m_{ik(i)}, m_{(i+1)1}\}. \tag{3}$$

(If $m = 1$, let $E' := E$.) (Recall that $E$ defined in (1) are disjoint edges, and by definition of $\mathcal{M}$ and Lemma 5, all edges in (1) are distinct.) In words, we write each $\mathcal{M}_i$ in cycle notation with the lowest number in each cycle appearing first, we remove the edge connecting the first and last endpoints of the cycle, and we connect the last vertex of the $i^{th}$ cycle to the first vertex of the $(i+1)^{st}$ cycle, for all $1 \leqslant i \leqslant c(f) - 1$. The graph $T = (V, E')$ is then a (connected) tree with $n - 1$ edges, and $n$ vertices. More specifically, it is $m$ trees connected along a path of length $m - 1$. The first element in the path is $m_{11}$ and the last element in the path is $m_{c(f)k(c(f))}$. We define the **Rényi-Joyal bijection** $R\colon \{\text{mappings}\} \to \{\text{doubly rooted trees}\}$ by

$$R(f) := (T, m_{11}, m_{c(f)k(c(f))}).$$

It remains to show that $R$ is one-to-one.

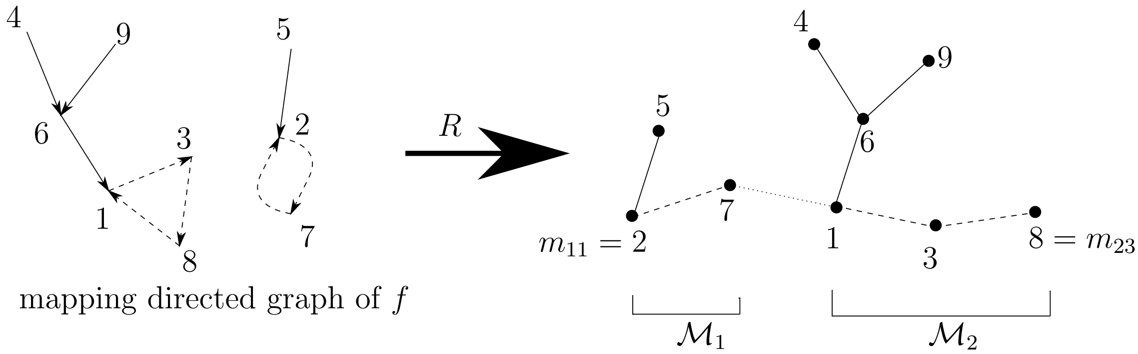


Figure 3: Example of the Rényi-Joyal bijection $R$. In this example, $f(1) = 3$, $f(3) = 8$, and so on. Also, $\mathcal{M} = \{1, 2, 3, 7, 8\}$, $\mathcal{M}_1 = \{2, 7\}$, $\mathcal{M}_2 = \{1, 3, 8\}$.

**Proof of injectivity**. Let $f, g\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$. Let $\mathcal{M}, \mathcal{N}$ be the cores of $f$ and $g$, respectively. Suppose $R(f) = R(g)$. By definition of $R$ (i.e (1)), we have $\mathcal{M} = \mathcal{N}$, $\mathcal{M}^c = \mathcal{N}^c$. Also, $f|_{\mathcal{M}^c} = g|_{\mathcal{M}^c}$. It remains to show that $f|_{\mathcal{M}} = g|_{\mathcal{M}}$. Since $R(f) = R(g)$, both $f$ and $g$ have the same ordered path of their cores in $R(f), R(g)$ respectively. (In Figure 3, this ordering would be $(2, 7, 1, 3, 8)$.) We can then e.g. recover the action of $f$ on the core (i.e. the permutation $f|_{\mathcal{M}}$) by creating cycles at the smallest elements of this ordering, read from left to right. So, if the ordered path of the core is $(s_1, \ldots, s_m)$, let $k(1) \geqslant 1$ be the largest integer $k$ so that $s_1 < s_2, s_1 < s_3, \ldots, s_1 < s_k$, and inductively define $k(i+1)$ to be the largest integer $k \leqslant m$ such that $s_{k(i)+1} < s_{k(i)+2}, \ldots, s_{k(i)+1} < s_k$. It then follows by definition of $R$ that $f|_{\mathcal{M}}$ is a permutation in the cycle notation

$$(s_1 \cdots s_{k(1)})(s_{k(1)+1} \cdots s_{k(2)}) \cdots (s_{k(c(f)-1)+1} \cdots s_{k(c(f))}).$$

Since $R(f) = R(g)$, $g|_{\mathcal{M}}$ is also a permutation with this same cycle notation. That is, $f|_{\mathcal{M}} = g|_{\mathcal{M}}$. In conclusion, $f = g$.

**Proof of surjectivity of *R*.** Let $(T, r_1, r_2)$ be a doubly rooted tree. Form the unique path $p$ of vertices in $T$ starting at $r_1$ and ending at $r_2$. Let $T' := T$. Repeat the following procedure until $T' \setminus p = \varnothing$:

- Choose one $x \in T'$ of degree 1, and define $f(x)$ to be the label of the vertex connected to $x$.
- Re-define $T'$ by removing from $T'$ the vertex $x$ and the edge emanating from $x$.

In this way, $f(x)$ is defined for all $x \in T \backslash p$. Now, we define $f$ on $p$. Label the elements of $p$ in the order they appear in the path as $r_1 = s_1, s_2, s_3, \ldots, s_m = r_2$. Let $\mathcal{M} := \{s_1, \ldots, s_m\}$. Let $k(1) \geqslant 1$ be the largest integer $k$ so that $s_1 < s_2, s_1 < s_3, \ldots, s_1 < s_k$, and inductively define $k(i+1)$ to be the largest integer $k \leqslant m$ such that $s_{k(i)+1} < s_{k(i)+2}, \ldots, s_{k(i)+1} < s_k$. Define $f|_{\mathcal{M}}$ to be the following permutation (written in cycle notation)

$$(s_1 \cdots s_{k(1)})(s_{k(1)+1} \cdots s_{k(2)}) \cdots (s_{k(c(f)-1)+1} \cdots s_{k(c(f))}).$$

Then $R(f) = (T, r_1, r_2)$, so that $J$ is surjective.

Finally, comparing Definition 4 with (3) proves the desired edge inequality. □

By removing the roots of the tree from the definition of $R$ in Lemma 9, we arrive at the following.

**Corollary 10** (**Rényi-Joyal Bijection with Roots Removed**). *There exists an $n^2$-to-one function $R$ from the set*

$$\{f\colon \{1, \ldots, n\} \to \{1, \ldots, n\}\}$$

*to the set of trees on $n$ vertices such that for all maps $f$,*

$$|E(f) \,\Delta\, E(R(f))| \leqslant 2c(f).$$

*(When we compute this symmetric difference, we remove the directions on the edges $E(f)$, and we count multiple edges from $E(f)$ as distinct.)*

### 2.3 Rényi-Joyal Bijection, Restricted

For the main application, Lemma 2, we also need to restrict the bijection in Lemma 9 to a specific class of mappings.

**Lemma 11** (**Rényi-Joyal Bijection, Restricted**). *Let $1 \leqslant k < n$. Denote $S := \{1, \ldots, k\}$ and $S^c := \{k+1, \ldots, n\}$. There exists a bijection $\widetilde{R}$ from the set*

$$\{f\colon \{1, \ldots, n\} \to \{1, \ldots, n\},\ f(S) \subseteq S^c\} \tag{4}$$

*to the set*

$$\{\text{doubly rooted trees on } n \text{ vertices such that } S \text{ is an independent set in the tree and the second root is in } S^c\}, \tag{5}$$

*such that for all maps $f$ with $f(S) \subseteq S^c$,*

$$|E(f) \,\Delta\, E(\widetilde{R}(f))| \leqslant 2c(f).$$

*(When we compute this symmetric difference, we remove the directions on the edges $E(f)$, and we count multiple edges from $E(f)$ as distinct.)*

*Moreover, if $N_S$ denotes the number of vertices in the graph that do not belong to any edge touching $S$, we have*

$$|N_S(\widetilde{R}(f)) - N_S(f)| \leqslant 1. \tag{6}$$

*Proof.* Let $\widetilde{R}$ be $R$ from Lemma 9, restricted to the set (4). Since $R$ itself is a bijection by Lemma 9, $\widetilde{R}$ is also injective. We therefore show that $\widetilde{R}$ is surjective. Let $(T, r_1, r_2)$ be in the set (5), so that $T$ is a tree, $r_1, r_2 \in \{1, \ldots, n\}$ and $r_2 \in S^c$. From Lemma 9, there exists a unique $f\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ such that $R(f) = (T, r_1, r_2)$. As in Lemma 9, we write $\mathcal{M} = \cup_{i=1}^{c(f)} \mathcal{M}_i$, where $\mathcal{M}_1, \ldots, \mathcal{M}_{c(f)}$ are subsets of vertices corresponding to the disjoint cycles of $f|_{\mathcal{M}}$. For each $1 \leqslant i \leqslant c(f)$, denote $\mathcal{M}_i = \{m_{i1}, \ldots, m_{ik(i)}\}$, where $k(i) := |\mathcal{M}_i|$, $m_{i1}$ is the smallest element of $\mathcal{M}_i$, and $m_{i(j+1)} = f(m_{ij})$ for all $1 \leqslant j < k(i)$. [That is, we can write $\mathcal{M}_i$ in cycle notation as $(m_{i1} \cdots m_{ik(i)})$, for all $1 \leqslant i \leqslant c(f)$.] We also choose the ordering on $\mathcal{M}_1, \ldots, \mathcal{M}_{c(f)}$ such that $m_{i1} > m_{(i+1)1}$ for all $1 \leqslant i \leqslant c(f)-1$. [That is, we order the cycles in the reverse order of their smallest elements.]

Let $1 \leqslant i \leqslant c(f)$. Since $S = \{1, \ldots, k\}$, if $\mathcal{M}_i \cap S \neq \varnothing$, it follows that $m_{i1} \in S$, since $m_{i1}$ is the smallest element of $\mathcal{M}_i$. Since $f(S) \subseteq S^c$, if $m_{i1} \in S$, then $m_{ik(i)} \notin S$. And if $\mathcal{M}_i \cap S = \varnothing$, then $m_{i1} \notin S$ and $m_{ik(i)} \notin S$. In either case, we have

$$m_{ik(i)} \notin S, \qquad \forall\, 1 \leqslant i \leqslant c(f). \tag{7}$$

In particular $r_2 := m_{c(f)k(c(f))} \notin S$. Since $S = \{1, \ldots, k\}$ and $m_{11} > m_{21} > \cdots > m_{c(f)1}$, there exists some integer $z$ satisfying $0 \leqslant z \leqslant c(f)$ such that

$$m_{i1} \in S \ \forall\, z < i \leqslant c(f), \qquad \text{and} \qquad m_{i1} \notin S \ \forall\, 1 \leqslant i \leqslant z. \tag{8}$$

(In the case $z = 0$, we have $m_{i1} \in S$ for all $1 \leqslant i \leqslant c(f)$, and in the case $z = c(f)$ we have $m_{i1} \notin S$ for all $1 \leqslant i \leqslant c(f)$.) That is, $z$ is the number of cycles whose smallest element is an element of $S^c$.

Since $f(S) \subseteq S^c$, $S$ is an independent set in the mapping directed graph of $f$. So, in order for $S$ to be an independent set in $R(f)$, we only need to check that $R$ does not add any edges from $S$ to itself. That is, we need a guarantee that (3) does not add an edge from $S$ to itself. The only new edges added to $R(f)$ are those specified in the right-most term of (3), and none of these edges go from $S$ to itself by (7). Therefore, $S$ is an independent set in $R(f)$. We have already shown that $r_2 := m_{c(f)k(c(f))} \notin S$, so that

$R(f)$ is a doubly rooted tree whose second root is in $S^c$. It then remains to show that (6) holds, but this again follows from (3) and (8). These equations imply that $R$ deletes exactly $c(f)-z$ edges from $S$ to $S^c$ (one for each of the $c(f)-z$ cycles $\mathcal{M}_{z+1},\ldots,\mathcal{M}_{c(f)}$), and it then adds exactly $c(f)-\max(z,1)$ edges from $S$ to $S^c$ in the right-most term of (3) (one for each term $m_{(\max(z,1))1},\ldots,m_{c(f)1}$). $\square$

*Remark* 12. It follows from the matrix-tree theorem that the number of labelled trees on $n>k$ vertices where vertices $\{1,\ldots,k\}$ form an independent set is

$$(n-k)^{k-1}n^{n-k-1}.$$

This fact also follows from Lemma 11. It is unclear if other consequences of the matrix-tree theorem can also follow from Lemma 11.

## 3 Inequalities From Mappings to Trees

In this section we demonstrate that the (randomized) Joyal bijection can find the distribution of some quantities on random trees.

As above, let $S\subseteq\{1,\ldots,n\}$, and let $k:=|S|$. Let $\alpha:=k/n$. The following proposition is a corollary of the matrix-tree theorem, but it also follows from Lemma 9.

**Proposition 13.** *Assume that $|S|=o((n-|S|)^2)$. Let $T$ be a uniformly random tree on $n$ vertices, conditioned on $S$ being an independent set. Let $N$ be the number of vertices in $S^c$ not connected to $S$. Then*

$$\mathbb{E}N=n(1-\alpha)^2e^{-\alpha/(1-\alpha)}(1+o_n(1)).$$

*Proof.* Let $f$ be a uniformly random mapping from $\{1,\ldots,n\}\to\{1,\ldots,n\}$ conditioned on $f(S)\subseteq S^c$. For any $x\in S^c$, let $N_x$ be 1 if $x$ is not connected to $S$, and let $N_x$ be 0 otherwise. Then

$$\mathbb{P}(N_x=1)=\Big(1-\frac{1}{n-k}\Big)^k(1-k/n),\qquad\forall\, x\in S^c.$$

So, using also $|S|=o((n-|S|)^2)$,

$$\begin{aligned}\mathbb{E}N=\mathbb{E}\sum_{x\in S^c}N_x&=(n-k)\Big(1-\frac{1}{n-k}\Big)^k(1-k/n)\\&=(1+o(1))e^{-k/(n-k)}n(1-\alpha)^2=n(1-\alpha)^2e^{-\alpha/(1-\alpha)}(1+o(1)).\end{aligned}$$

Lemma 11 then completes the proof. $\square$

*Remark* 14. With no constraints on $f$, we have

$$\mathbb{E}N=(n-k)\Big(1-\frac{1}{n}\Big)^k(1-k/n)=n(1-\alpha)^2e^{-\alpha}(1+o(1)).$$

*Remark* 15. It follows from the Azuma-Hoeffding Inequality, Lemma 18 (using e.g. the Aldous-Broder algorithm [6, 1] to construct a tree one edge at a time, as a martingale) and Proposition 13 that $\forall\ t > 0$,

$$\mathbb{P}(|N - \mathbb{E}N| > t\mathbb{E}N \mid S \text{ is an independent set}) \leqslant 2e^{-\frac{n^2}{n-1}\frac{t^2(1-\alpha)^4 e^{-2\alpha/(1-\alpha)}}{2}(1+o_n(1))}. \tag{9}$$

However, this inequality can be improved to Lemma 2, and this is important for our application to independent sets in uniformly random labelled trees. Since Lemma 2 improves on (9), we will not prove (9).

Below we will prove Lemma 2. We will use several properties of negatively associated (NA) random variables from [13]. A function $f\colon \mathbb{R}^k \to \mathbb{R}$ is said to be increasing if for any $1 \leqslant i \leqslant k$, and for any $x_1, \ldots, x_k, x_i' \in \mathbb{R}$ with $x_i \leqslant x_i'$, we have $f(x_1, \ldots, x_k) \leqslant f(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_k)$. Real-valued random variables $X_1, \ldots, X_k$ are said to be **negatively associated**, denoted **NA**, if for any disjoint subsets $A, B \subseteq \{1, \ldots, k\}$, and for any increasing functions $f\colon \mathbb{R}^{|A|} \to \mathbb{R}$, $g\colon \mathbb{R}^{|B|} \to \mathbb{R}$ such that the following expression is well-defined,

$$\mathbb{E}f(\{X_i\}_{i\in A})g(\{X_i\}_{i\in B}) - \mathbb{E}f(\{X_i\}_{i\in A}) \cdot \mathbb{E}g(\{X_i\}_{i\in B}) \leqslant 0.$$

An equivalent definition can be made by requiring both $f$ and $g$ to be decreasing.

Here are some properties of NA random variables, listed in [13, Page 288].

(i) A set of independent random variables is NA.

(ii) Increasing functions defined on disjoint subsets of a set of NA random variables are NA. (Similarly, decreasing functions defined on disjoint subsets of a set of NA random variables are NA.)

(iii) The disjoint union of independent families of NA random variables is NA.

*Proof of Lemma 2.* Let $f\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ be a uniformly random mapping, conditioned on the event $f(S) \subseteq S^c$. Let $V' := \{1, \ldots, n\} \setminus (S \cup f(S))$ to be the vertices in $S^c$ that are not in the image of $f(S)$. For any $x \in S^c$, let $N_x := 1_{x\in V'}$. The distribution of $N' := |V'|$ is well-known as the classical occupancy problem. It is also well known, that the random variables $\{N_x\}_{x\in S^c}$ are NA [7, 5]. Now, for any $x \in S^c$, let $M_x := 1_{f(x)\notin S}$. Since $f$ is a uniformly random mapping, the random variables $\{M_x\}_{x\in S^c}$ are independent of each other, and independent of $\{N_x\}_{x\in S^c}$. So, the random variables $\{M_x\}_{x\in S^c}$ are NA by Property (i) and the union of the random variables $\{N_x\}_{x\in S^c} \cup \{M_x\}_{x\in S^c}$ also is NA by Property (iii). Since the minimum function is monotone decreasing, the random variables $\{\min(N_x, M_x)\}_{x\in S^c}$ are also NA by Property (ii). Finally, define

$$\widetilde{N} := \sum_{x\in S^c} \min(N_x, M_x).$$

(If we started the proof with a random tree instead of a uniformly random mapping, then $N$ would be equal to $\widetilde{N}$.) Since $\widetilde{N}$ is the sum of NA random variables, $\widetilde{N}$ satisfies Chernoff bounds, i.e. the bounds of Lemma 17, by repeating the standard proof of Chernoff bounds, as noted e.g. in [7, Proposition 29]. We then transfer this inequality to the random tree by the last part of Lemma 11. (The computation of $\mathbb{E}N$ was done in Proposition 13.) $\square$

*Remark* 16. It is not obvious to the author how to apply the negative association property directly to random trees. That is, we are not aware of a proof of Lemma 2 that uses the negative association property for random variables on trees, the main difficulty being lack of any obvious independence. So, at present it seems necessary to use the bijection from Lemma 11 to prove Lemma 2.

## 4 Concentration Inequalities

These concentration inequalities are referenced elsewhere in the paper. We include them here for the reader's convenience.

**Lemma 17** (**Chernoff Bounds**). *Let $N$ be a binomial random variable with parameters $n$ and $p$, or a sum of NA random variables [7, Proposition 29]. Then*

$$\begin{aligned}
&\mathbb{P}(|N-\mathbb{E}N| > s\mathbb{E}N) \leqslant e^{-\min(s,s^2)\mathbb{E}N/3}, && \forall\, s>0,\\
&\mathbb{P}(N<(1-s)\mathbb{E}N) \leqslant e^{-s^2\mathbb{E}N/2}, && \forall\, 0<s<1,\\
&\mathbb{P}(N>(1+s)\mathbb{E}N) \leqslant e^{-s^2\mathbb{E}N/(2+s)}, && \forall\, s\geqslant 0.
\end{aligned}$$

**Lemma 18** (**Azuma-Hoeffding Inequality**[20]). *Let $c>0$. Let $Y_0,\ldots,Y_n$ be a real-valued martingale with $Y_0$ constant and $|Y_{m+1}-Y_m|\leqslant c$ for all $0\leqslant m\leqslant n-1$. Then*

$$\mathbb{P}(|Y_n-Y_0|>t)\leqslant 2e^{-\frac{t^2}{2c^2n}}, \qquad \forall\, t>0.$$

## 5 Algorithmic Interpretation of Bijection

The bijection $R$ from Lemma 9 also gives an algorithm for sampling from uniformly random trees on $n$ vertices. Prüfer codes themselves give an elementary way to generate random trees in $O(n\log n)$ time, though $O(n)$ time is possible with a less elementary implementation [24]. Algorithm 19 can generate a random tree in $O(n)$ time. The Aldous-Broder algorithm [6, 1] is perhaps the most elementary way to generate a uniformly random labelled tree on $n$ vertices, though its run time is $O(n^2)$ with its most naive implementation. (However, a less naive implementation of the Aldous-Broder algorithm has $O(n)$ run time [12, Section 5].)

*Algorithm* 19 (**Sampling a Uniformly Random Labelled Tree on $n$ vertices**).

- The input of the algorithm is a random mapping $f\colon\{1,\ldots,n\}\to\{1,\ldots,n\}$, presented as a list $(f(1),\ldots,f(n))$ of $n$ independent identically distributed random variables, each uniformly distributed in $\{1,\ldots,n\}$.

- The output of the algorithm is a uniformly random labelled tree on $n$ vertices.

The algorithm proceeds as follows.

(i) Compute the core of $f$, via a standard algorithm such as Algorithm 20. The core is the set of $1 \leqslant k \leqslant n$ cycles of $f$, written as $C_1, \ldots, C_k$.

(ii) For each $1 \leqslant i \leqslant k$, the smallest element of $C_i$ is given the left-most position in the cycle notation for $C_i$. (so e.g. the cycle (365) is written rather than (653)).

(iii) Arrange the cycles in reverse age order, according to their smallest elements (so e.g. we write two cycles in the ordering (365)(289) rather than (289)(365).)

(iv) Let $C$ be the set of vertices in the core of $f$. Output the tree formed by the edges

$$\{\{x, f(x)\} \colon x \in \{1, \ldots, n\} \setminus C\}$$

together with the path that passes through the cycles in the order specified by (iii).

The proof of Theorem 1, or more specifically the definition of the bijection $R$ defined in the proof, implies that the output of Algorithm 19 is a uniformly random labelled tree on $n$ vertices with run time $O(n)$.

*Algorithm* 20 (**Computing the Core of a Mapping**).

- The input of the algorithm is a mapping $f \colon \{1, \ldots, n\} \to \{1, \ldots, n\}$, presented as a list $(f(1), \ldots, f(n))$.
- The output of the algorithm is the set of $1 \leqslant k \leqslant n$ cycles of $f$ presented in cycle notation as $C_1, \ldots, C_k$ (so e.g. $C_1 = (245)$ indicates that $f(2) = 4, f(4) = 5$ and $f(5) = 2$.)

The algorithm proceeds as follows. Let $B := \varnothing$, $k := 0$. While $B \neq \{1, \ldots, n\}$, repeat the following procedure.

- Let $x \in \{1, \ldots, n\} \setminus B$. Compute the sequence $x, f(x), f(f(x)), f(f(f(x))), \ldots$ until one element of the sequence is repeated (so that $f^j(x) = f^k(x)$ for some $0 \leqslant j < k \leqslant n$). (We can find a repeated element in this sequence in $O(k)$ time using either a separate array or hash table to keep track of which elements of $\{1, \ldots, n\}$ have appeared in the sequence.)
- Define $C_{k+1} := (f^j(x), f^{j+1}(x), \ldots, f^{k-1}(x))$. This is the $(k+1)^{st}$ cycle of $f$.
- Re-define $k$ to be one more than its previous value. Also re-define $B$ to be $B$ union with the set $\{x, f(x), f^2(x), \ldots, f^k(x)\}$.

## 6 Appendix: Cycle Distributions

Let $C_n(f)$ be the number of cycles in a mapping $f\colon \{1,\dots,n\} \to \{1,\dots,n\}$. The following Lemma is sketched in [10]. We give a detailed proof.

**Lemma 21** (**Cycle Distribution of a Random Mapping**)**.**

$$\mathbb{P}(C_n > (1+t)\log n) \leqslant (1+o_n(1))e^{-\frac{t^2}{2+t}(\log n)/4}, \qquad \forall\, t>0.$$

*Proof.* Let $\Pi$ be a uniformly random element of the group $S_n$ of permutations on $n$ elements. As in [8, Lemma 2.2.9], let $X_{n,k}(\pi) := 1$ if a right parenthesis occurs after entry $k$ in the standard cycle notation of the permutation $\pi \in S_n$, and $X_{n,k}(\pi) := 0$ otherwise. Then $X_{n,1},\dots,X_{n,n}$ are independent random variables with $\mathbb{P}(X_{n,k}=1) = 1/(n-k+1)$. Let $C(\pi)$ be the number of cycles in $\pi \in S_n$. Then $C = C_n = \sum_{k=1}^n X_{n,k}$, $\mathbb{E}C_n = \sum_{k=1}^n \mathbb{E}X_{n,k}$, and from Chernoff's bound 17,

$$\mathbb{P}(C_n(\Pi) \geqslant (1+t)\mathbb{E}C_n(\Pi)) \leqslant e^{-\frac{t^2}{2+t}\mathbb{E}C_n(\Pi)}, \qquad \forall\, t>0. \tag{10}$$

When we condition a random mapping $f\colon \{1,\dots,n\} \to \{1,\dots,n\}$ such that its core size $|\mathcal{M}|$ is constant, then $f|_{\mathcal{M}}$ is a uniformly random permutation on $|\mathcal{M}|$ elements.

Let $Y_n(f)$ be the total number of vertices in all cycles of a uniformly random mapping $f\colon \{1,\dots,n\} \to \{1,\dots,n\}$. As in [10], $\mathbb{P}(Y_n = k) = \frac{k(n-1)!}{n^k(n-k)}$ for all $0 \leqslant k \leqslant n$. Then, $\forall$ $0<a<b<\sqrt{n}$,

$$\begin{aligned}
\mathbb{P}(a\sqrt{n} \leqslant Y_n \leqslant b\sqrt{n}) &= \mathbb{P}(a < \frac{Y_n}{\sqrt{n}} < b) = \sum_{k=a\sqrt{n}}^{b\sqrt{n}} \frac{k(n-1)!}{n^k(n-k)} \\
&= \sum_{j=a,a+1/\sqrt{n},\dots,b} \frac{j\sqrt{n}(n-1)!}{n^{j\sqrt{n}}(n-j\sqrt{n})!} \\
&= (1+o(1)) \sum_{j=a,a+1/\sqrt{n},\dots,b} j\sqrt{n}\sqrt{\frac{n-1}{n-j\sqrt{n}}}\frac{(n-1)^{n-1}e^{-n+1}}{n^{j\sqrt{n}}(n-j\sqrt{n})^{n-j\sqrt{n}}e^{-n+j\sqrt{n}}} \\
&= (1+o(1)) \sum_{j=a,a+1/\sqrt{n},\dots,b} je^1\sqrt{n}\Big(\frac{n-1}{n}\Big)^{n-1/2}\Big(\frac{n-j\sqrt{n}}{n}\Big)^{-n+j\sqrt{n}-1/2} e^{-j\sqrt{n}} \\
&= (1+o(1)) \sum_{j=a,a+1/\sqrt{n},\dots,b} je^1\sqrt{n}\Big(1-\frac{1}{n}\Big)^{n-1/2}\Big(1-\frac{j}{\sqrt{n}}\Big)^{-n+j\sqrt{n}-1/2} e^{-j\sqrt{n}}.
\end{aligned}$$

We write

$$\begin{aligned}
\Big(1-\frac{j}{\sqrt{n}}\Big)^{-n+j\sqrt{n}-1/2} &= e^{[\log(1-j/\sqrt{n})](-n+j\sqrt{n}-1/2)} = e^{[-j/\sqrt{n}-j^2/(2n)+O(j/\sqrt{n})^{3/2}](-n+j\sqrt{n}-1/2)} \\
&= e^{j\sqrt{n}+j^2/2-j^2+O(1/\sqrt{n})} = e^{j\sqrt{n}}e^{-j^2/2+O(1/\sqrt{n})}.
\end{aligned}$$

Therefore

$$\begin{aligned}\mathbb{P}(a\sqrt{n} \leqslant Y_n \leqslant b\sqrt{n}) &= (1+o(1)) \sum_{j=a,a+1/\sqrt{n},\ldots,b} jn^{-1/2}e^{-j^2/2} \\ &= (1+o(1)) \int_a^b xe^{-x^2/2}dx = (1+o(1))(e^{-a^2/2} - e^{-b^2/2}).\end{aligned}$$

If $\varepsilon > 0$, then

$$\mathbb{P}(Y_n \notin \sqrt{n}[\sqrt{\varepsilon}, \sqrt{\log(1/\varepsilon)}]) = (1+o(1))(1 - e^{-\varepsilon/2} + e^{-\log(1/\varepsilon)/2}) = (1+o(1))\varepsilon. \tag{11}$$

So, recalling that $C_n(f)$ is the number of cycles in a mapping $f\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$,

$$\begin{aligned}\mathbb{P}(C_n > (1+t)\log n) &\leqslant (1+o(1))(1+\varepsilon)\mathbb{P}\big(C_n > (1+t)\log n \mid Y_n \in \sqrt{n}[\sqrt{\varepsilon}, \sqrt{\log(1/\varepsilon)}]\big) \\ &\overset{(10)}{\leqslant} (1+o(1))(1+\varepsilon) \sup_{m\in\sqrt{n}\left[\sqrt{\varepsilon},\sqrt{\log(1/\varepsilon)}\right]} e^{-\frac{t^2}{2+t}\log m} \leqslant e^{-\frac{t^2}{2+t}[\log n - \log(1/\varepsilon)]/2}.\end{aligned}$$

Choosing $\varepsilon := 1/\sqrt{n}$,

$$\mathbb{P}(C_n > (1+t)\log n) \leqslant (1+o_n(1))e^{-\frac{t^2}{2+t}(\log n)/4}.$$

□

*Remark* 22. It is tempting to try to apply Talagrand's convex distance inequality [14, Theorem 2.29] to prove Lemma 21 but it is not obvious to the author how to make such an argument.

### 6.1 Independent Set Case

Let $S \subseteq \{1,\ldots,n\}$, and let $k := |S|$. Let $\alpha := k/n$.

Let $C_n(f)$ be the number of cycles in a mapping $f\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$.

**Lemma 23.** *Let $F$ be a uniformly random mapping from $\{1,\ldots,n\}\setminus S \to \{1,\ldots,n\}\setminus S$. Let $G$ be a random mapping, uniformly distributed over all $f\colon \{1,\ldots,n\} \to \{1,\ldots,n\}$ such that $f(S) \subseteq S^c$. Then the random variables*

$$C_{n-k}(F), \qquad C_n(G)$$

*are identically distributed.*

*Proof.* Define $r(G)\colon \{1,\ldots,n\}\setminus S \to \{1,\ldots,n\}\setminus S$ by

$$r(G)(x) = \begin{cases} G(x) & \text{if } G(x) \in S^c \\ G(G(x)) & \text{if } G(x) \in S. \end{cases}$$

Since $G(S) \subseteq S^c$, if $G(x) \in S$ then $G(G(x)) \in S^c$, so that $r(G)$ always takes values in $S^c$.

Note that $G$ and $r(G)$ have the same number of cycles, since $r(G)$ removes all elements of $S$ from all cycles of $G$, but each cycle in $G$ must have at least one element in $S^c$. That is, $C_n(G) = C_n(r(G))$.

Also, $r(G)$ is a uniformly random mapping on $\{1,\ldots,n\} \setminus S$. To see this, denote $S =: \{s_1,\ldots,s_k\}$, let $x_1,\ldots,x_k \in \{1,\ldots,n\} \setminus S$ and let $y_1,\ldots,y_k \in \{1,\ldots,n\}$. Then the conditional probability

$$\mathbb{P}(r(G)(s_i) = x_i \ \forall\, 1 \leqslant i \leqslant k \mid G(x_j) = y_j \ \forall\, 1 \leqslant j \leqslant k)$$

does not depend on $x_1,\ldots,x_k,y_1,\ldots,y_k$. So, we can remove the conditioning and conclude that

$$\mathbb{P}(r(G)(s_i) = x_i \ \forall\, 1 \leqslant i \leqslant k)$$

does not depend on $x_1,\ldots,x_k$. That is, $r(G)$ is a uniformly random element of mappings from $\{1,\ldots,n\} \setminus S$ to itself. Re-labeling $r(G)$ as $F$ completes the proof. □

**Lemma 24** (**Cycle Distribution of a Restricted Random Mapping**).

$$\mathbb{P}(C_n(f) > (1+t)\log(n-k) \,|\, f(S) \subseteq S^c) \leqslant (1+o_n(1)e^{-\frac{t^2}{2+t}(\log(n-k))/4}, \qquad \forall\, t > 0.$$

*Proof.* Combine Lemmas 21 and 23. □


**Acknowledgements**

Thanks to Richard Arratia for explaining to me various things such as the Joyal bijection (Lemma 6) and the improved Rényi-Joyal bijection (Lemma 9), and for suggesting that these bijections could play a role in the main application (Lemma 2). Thanks also to Larry Goldstein for helpful discussions. Thanks to an anonymous reviewer for various comments improving the manuscript and for correcting several typos.


## References


[1] David J. Aldous, *The random walk construction of uniform spanning trees and uniform labelled trees*, SIAM J. Discrete Math. **3** (1990), no. 4, 450–465.

[2] David J. Aldous, Grégory Miermont, and Jim Pitman, *Brownian bridge asymptotics for random p-mappings*, Electron. J. Probab. **9** (2004), no. 3, 37–56.

[3] David J. Aldous, Gregory Miermont, and Jim Pitman, *The exploration process of inhomogeneous continuum random trees, and an extension of Jeulin's local time identity*, Probab. Theory Related Fields **129** (2004), no. 2, 182–218.

[4] David J. Aldous, Gregory Miermont, and Jim Pitman, *Weak convergence of random p-mappings and the exploration process of inhomogeneous continuum random trees*, Probab. Theory Related Fields **133** (2005), no. 1, 1–17.

[5] Jay Bartroff, Larry Goldstein, and Ümit Işlak, *Bounded size biased couplings, log concave distributions and concentration of measure for occupancy models*, Bernoulli **24** (2018), no. 4B, 3283–3317.

[6] Andrei Broder, *Generating random spanning trees*, Proceedings of the 30th Annual Symposium on Foundations of Computer Science (USA), SFCS 89, IEEE Computer Society, 1989, p. 442447.

[7] Devdatt Dubhashi and Desh Ranjan, *Balls and bins: a study in negative dependence*, Random Structures Algorithms **13** (1998), no. 2, 99–124.

[8] Rick Durrett, *Probability—theory and examples*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 49, Cambridge University Press, Cambridge, 2019, Fifth edition of [ MR1068527].

[9] Dominique Foata, *Étude algébrique de certains problèmes d'analyse combinatoire et du calcul des probabilités*, Publ. Inst. Statist. Univ. Paris **14** (1965), 81–241.

[10] Bernard Harris, *Probability distributions related to random mappings*, Ann. Math. Statist. **31** (1960), 1045–1062.

[11] Steven Heilman, *Independent sets of random trees and of sparse random graphs*, Preprint, 2020.

[12] Mikhail Isaev, Angus Southwell, and Maksim Zhukovskii, *Distribution of tree parameters by martingale approach*, *Combinatorics, Probability and Computing* (2022), 1–28.

[13] Kumar Joag-Dev and Frank Proschan, *Negative association of random variables with applications*, The Annals of Statistics **11** (1983), no. 1, 286–295.

[14] Svante Janson, Tomasz Luczak, and Andrzej Rucinski, *Random graphs*, Wiley Series in Discrete Mathematics and Optimization, Wiley, 2011.

[15] André Joyal, *Une théorie combinatoire des séries formelles*, Adv. in Math. **42** (1981), no. 1, 1–82.

[16] Valentin F. Kolchin, *Random mappings*, Translation Series in Mathematics and Engineering, Optimization Software, Inc., Publications Division, New York, 1986, Translated from the Russian, With a foreword by S. R. S. Varadhan.

[17] Colin McDiarmid, *Concentration*, Probabilistic methods for algorithmic discrete mathematics, Algorithms Combin., vol. 16, Springer, Berlin, 1998, pp. 195–248.

[18] Alfréd Rényi, *Théorie des éléments saillants d'une suite d'observations*, Ann. Fac. Sci. Univ. Clermont-Ferrand **8** (1962), 7–13.

[19] Dimbinaina Ralaivaosaona and Stephan Wagner, *A central limit theorem for additive functionals of increasing trees*, Combin. Probab. Comput. **28** (2019), no. 4, 618–637.

[20] Eli Shamir and Joel Spencer, *Sharp concentration of the chromatic number on random graphsgn, p*, Combinatorica **7** (1987), no. 1, 121–129.

[21] Richard P. Stanley, *Enumerative combinatorics. Volume 1*, second ed., Cambridge Studies in Advanced Mathematics, vol. 49, Cambridge University Press, Cambridge, 2012.

[22] V. E. Stepanov, *Limit distributions of certain characteristics of random mappings*, Theor. Probability Appl. **14** (1969), 612–626.

[23] Stephan Wagner, *Central limit theorems for additive tree parameters with small toll functions*, Combin. Probab. Comput. **24** (2015), no. 1, 329–353.

[24] Xiaodong Wang, Lei Wang, and Yingjie Wu, *An optimal algorithm for prufer codes*, JSEA **2** (2009), 111–115.