# Sampling Uniform Tajima Trees with Alternating Pemrutations

Leonardo Bonanno, Julia Palacios

July 27, 2020

**Abstract**

# 1 Introduction

## 1.1 Euler-Zigzag Numbers

An alternating permutation is a permutation such that successive terms alternate in increasing and decreasing order. Letting $A_n$ denote the set of down-up $(\sigma(1) > \sigma(2))$ alternating permutations on $\{1, \ldots n\}$ and $E_n = |A_n|$, then for $n > 0$,

$$2E_{n+1} = \sum_{k=0}^{n} \binom{n}{k} E_k E_{n-k} \tag{1}$$

The $n+1$st Euler-Zigzag number. For $n = 3$, the up-down alternating permutations are $132, 231$ and the down-up alternating permutations are $312, 213$.

## 1.2 Tajima Trees

A tajimaa tree is a rooted, ranked and unlabeled tree, an example of which is presented below.

## 1.3 Differences between Uniform and Tajima Distribution

Let $Tn$ denote the tajima distribution for ranked unlabeled trees and $U_n$ the uniform distribution. We will see that these distributions vary significantly. The plot of $\|U_n - T_n\|_{TV}$ is below. From this, it is clear that these distributions differ significantly, even for relatively small $n$.

**Conjecture 1.1.** $\lim_{n\to\infty} \|U_n - T_n\| = 1$

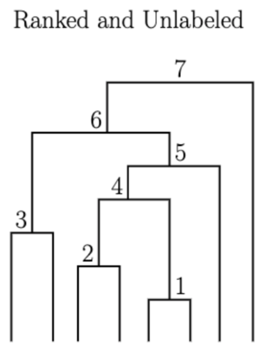



Figure 1: Example of Tajima Tree

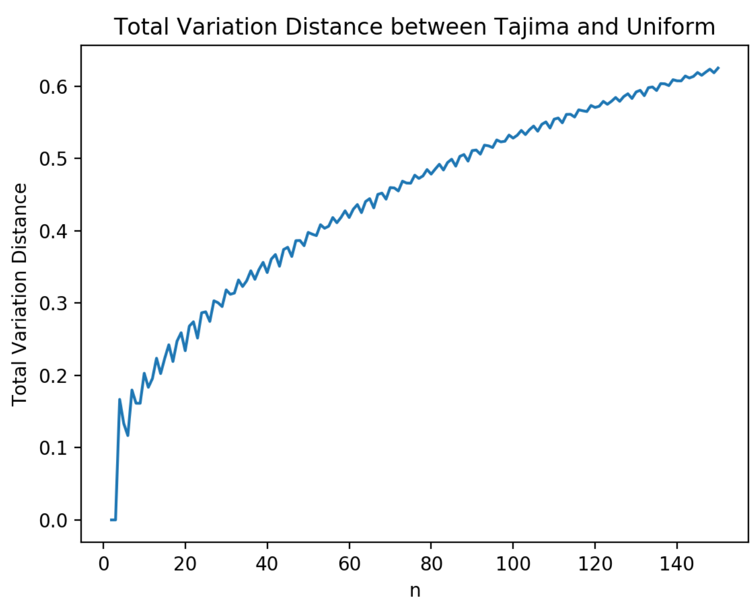



Figure 2: Total Variation Distance between Tajima and Uniform

## 2 Tajima Trees are Alternating Permutations

To recover the relationship between Let $A_n$ denote the set of (down-up) alternating permutations of $[n]$ and $\tau_n$ denote the set of tajima trees with $n$ leaves.

**Theorem 2.1.** *For $n \geq 2$, $|A_{n-1}| = |\tau_n|$*

*Proof.* It is shown in [1] that $|A_n| = E_n$ and , and in [6] that $|\tau_n| = E_{n-1}$. □

With this result in hand, it is clear that their is a bijection between the set of alternating permutations $A_{n-1}$ of $[n-1]$ and the set of tajima trees $\tau_n$ of size $n$. The full bijection is presented in [1] and we present a compressed version here.

### 2.1 Bijection Between $\tau_n$ and $A_{n-1}$

Let $B_n$ denote the set of binary increasing trees of size $n$. We illustrate $B_4$ in Figure 1. We

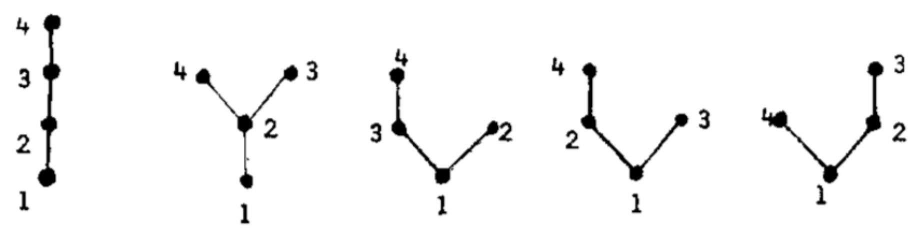



Figure 3: Binary increasing trees of size $n = 4$

will construct a bijection between $B_n$ and $A_{n-1}$ and a bijection between $B_n$ and $\tau_n$. The set $B_n$ for $n = 3$ and $n = 4$ is presented below (from [1]).

**Lemma 2.1.** *There is a bijection between $B_n$ and $\tau_n$.*

*Proof.* Let $B \in B_n$. Now let $\tau$ be the ranked unlabeled tree formed by adding a bifurcation to each leaf branch of $B$ and extending from each vertex with only one upward connecting vertex a singleton (leaf) branch. Then $\tau \in \tau_n$. Similarly, let $\tau \in \tau_n$ and $B \in B_n$ be the binary increasing tree formed by removing all cherries and singleton's from $B$. It is easy to see that these maps are two-sided inverses of each other and therefore they are bijections between $B_n$ and $\tau_n$. The bijection in the case $n = 4$ is presented below. We will use this bijection in the future sections.

□

We now construct a bijection between $B_n$ and $A_{n-1}$, which will be very important in constructing markov chains on $\tau_n$ later. If $a_1 a_2 \dots a_k$ is a sequence of distinct numbers in $[n]$, for $x \in \{a_1, \dots, a_k\}$, let $Ind(x) = i$ if $a_i = x$. Now let $b_1 \dots b_k$ be a sorted version of the sequence $a_1 \dots a_k$ (in increasing order). Then let us define, again for $x \in \{a_1, \dots a_k\}$, define

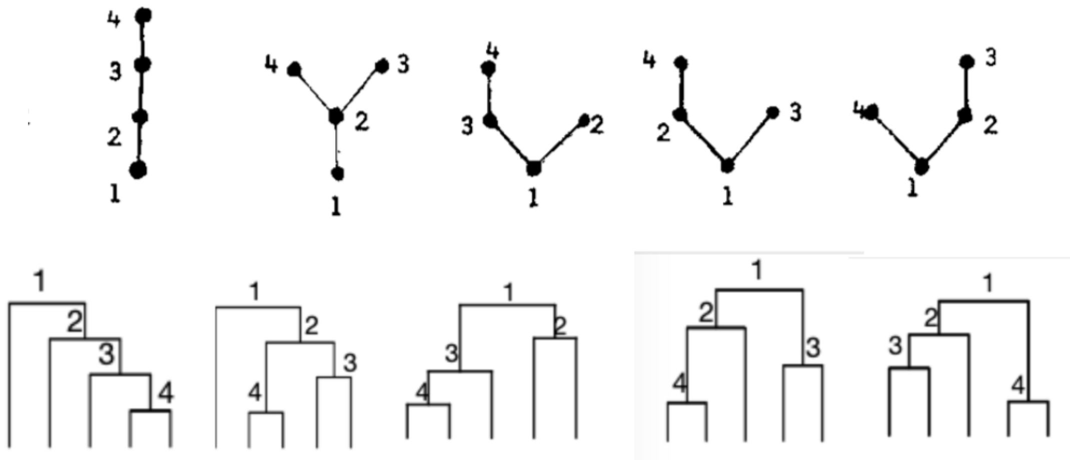



Figure 4: Bijection between $B_4$ and $\tau_4$

$\varphi(x = a_i) = a_{k-i+1}$. Now let $G_n = \{a_1 \dots a_k : k \in [n], a_i \in [n]\ \forall i, a_i \neq a_j \forall i, j\}$. This is the set of sequences of distinct elements of $[n]$. Let us construct a map $G : G_n \to G_n$. Let $S = a_1 \dots a_k \in G_n$. If $k = 1$, then $G(S) = S$. Now let us assume that $k > 1$. Let $i = Ind(max(S))$ and $j = Ind(min(S))$. If $i < j$ then $G_n(S) = G_n(a_1 \dots a_{i-1})\, iG_n(a_{i+1} \dots a_k)$. Otherwise let $b_i = \varphi(a_i)$ for all $a_i$. Then $G_n(S) = G_n(b_1 \dots b_{j-1})\, a_i G_n(b_{j+1} \dots b_k)$. Now let $X_n : A_n \to S_n$ be given by $X_n(A) = G_n(A)$ for $A \in A_n$. We note that by definition $X_n$ is injective.

**Example 2.2.** *INCLUDE EXAMPLE HERE 1.*

**Example 2.3.** *INCLUDE EXAMPLE HERE 2.*

Now let $X = X_n(A_n)$. It is then easy to construct a map from $X$ to $B_n$. We show it by aan example below.

**Example 2.4.** *INCLUDE EXAMPLE HERE 3.*

### 2.2 Case $n = 4$

We present the bijection in the case $n = 4$ here, from ([1]).

## 3 Random Transpositions on Alternating Permutations

Having constructed a bijection between $\tau_n$ and $A_{n-1}$, there a number of immediate markov chains that we can construct on $\tau_n$ by constructing chains on $A_{n-1}$ and using the bijection presented in the previous sections. An obvious markov chain on the space of alternating permutations is the

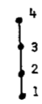
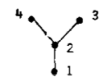
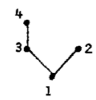
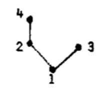
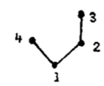
| $P$ | $(P^T)$ | $T$ |
|---|---|---|
| 3241 & 2314 | ((((4) 3) 2) 1) |  |
| 4231 & 1324 | (((4) 2(3)) 1) |  |
| 3412 & 2143 | (((4) 3) 1(2)) |  |
| 2413 & 3142 | (((4) 2) 1(3)) |  |
| 4132 & 1423 | ((4) 1((3) 2)) |  |

Figure 5: Bijection between $B_4$ and $A_4$

random transpositions chain. That is, at each time step, selecting two elements of the permutation and switching them if the new permutation formed is still alternating. This chain is clearly reversible with respect to the uniform distribution on $A_n$, aperiodic, and irreducible (I would like a proof for this, [3] says it is trivial). Experiments show that this chain is fast mixing, with a mixing time on the order of $n^2$. Additionally, experiments indicate that moves are accepted about $\frac{1}{2}$ of the time as $n \to \infty$.

### 3.1 Markov Chain on Alternating Permutations

## 4 Generating Uniform Alternating Permutations using a Method of Quasistationary Distributions

## 5 Generating Uniform Alternating Permutations Recursively

[3] presents an algorithm for recursively generating a single alternating permutation. The method is $O\left(n^3\right)$, partially because it requires computing the first $n$ Euler-Zigzag numbers. We can

# 6 Generating Uniform Alternating Permutations with Gibbs Sampling

## References


[1] J. Donaghey. *Alternating Permutations and Binary Increasing Trees*. Journal of Combinatorial Theory (A), 1975.

[2] P. Marchal. *Generating random alternating permutations in time* $n \log n$. http://hal.archives-ouvertes.fr/docs/00/76/54/32/PDF/altperm.pdf, 2012.

[3] P. Diaconis and Philip Wood. *Random double stochastic tridiagonal matrices*. Random Structures and Algorithms, 2012.

[4] Richard Stanley. *A Survey of Alternating Permutations*. https://arxiv.org/abs/0912.4240, 2009.

[5] L. Capello, A. Veber and J. Palacios *The Tajima heterochronous n-coalescent: inference from heterochronously sampled molecular data*. https://arxiv.org/abs/2004.06826, 2020.

[6] JA Palacios et. al. *Bayesian Estimation of Population Size Changes by Sampling Tajima's Trees*. Genetics, 2019.