

Alternating Permutations and Tajima Trees

Leonardo Bonanno, Julia Palacios

July 27, 2020

1 Introduction/Preliminaries

For the biological and statistical procedures such as those in [5], it is necessary to sample from the space of tajima (ranked unlabeled) trees. This has been achieved previously using a MCMC algorithm. In this paper, we present an $O(n \log n)$ algorithm to sample exactly uniformly from the space of tajima trees using ideas from [1] and [2], and then discuss applications in population genetics. More stuff needs to be added to this section when the applications are added and depending on how much more/what background information is needed because I am currently unsure.

2 Bijection Between Alternating Permutations and Ranked Unlabeled Trees

Let A_n denote the set of alternating permutations on $\{1, \dots, n\}$, and τ_{n+1}^R , the set of tajima trees with $n+1$ leaves (and hence n branchpoints). In this section, we establish a canonical bijection between A_n and τ_{n+1}^R . In [5], the authors establish that the cardinality of the space of tajima trees with n leaves is E_{n-1} , the $n-1$ Euler-Zigzag number, satisfying the recurrence relation.

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

[4] shows that the n -th Euler-Zigzag number is also the cardinality of the set of alternating up-down and down-up permutations of $\{1, \dots, n-1\}$. These identical cardinalities indicate that there must be a bijection between these two groups. [1] establishes a natural bijection between the set of binary increasing trees with n vertices and the set of alternating permutations of $\{1, \dots, n\}$ using a method of nestling and relative complementing. In fact, each binary increasing tree with n vertices is the internal branches of a tajima tree with $n+1$ vertices. Thus, the bijection using in 5 can be easily extended to present a bijection between the tajima trees with $n+1$ leaves and alternating permutations on $\{1, \dots, n-1\}$. With this bijection, to sample uniformly from the space of tajima trees, it is sufficient to sample uniformly from the space of up-down (or equivalently down-up) alternating permutations, an algorithm presented by [2] is presented in section 4.

3 Case $n = 4$ and $n = 5$

In this section we present the bijection between the space of tajima trees, binary increasing trees and alternating permutations. Below we present the bijection for the case $n = 4$. Here we present the bijection for the case $n = 5$. INSERT DIAGRAMS HERE.

4 Generating A Uniformly Random Alternating Permutation

In [2], the author presents an $O(n \log n)$ algorithm to generate a uniform random (on the space of all, both up-down and down-up) permutations using the theory of quasistationary markov chains. It improves on the $O(n^2)$ time that can be achieved using boltzmann sampling, and is almost surely below the unknown mixing time of the alternating permutations markov chain described in the section below. Additionally, it provides a mechanism for exact sampling from this distribution. Additionally, the algorithm can easily be adapted to generating a uniform sample from the space of either up-down or down-up alternating permutations by simply taking the complement, as described in [1] of the end result achieved by the algorithm if it is oriented incorrectly

5 Markov Chain on Alternating Permutations

Using the algorithm from section 4 and the bijection between 3, it is possible to generate a single uniformly random tajima tree in time $n \log n$. To generate multiple samples, the markov chain on alternating permutations presented in [3] can be used. We describe it again here. The chain involves at each step randomly transposing two elements uniformly at random between the positions that are at least 2 apart, accepting the move only if it yields a valid permutation. As established in [3], this chain symmetric and reversible with respect to the uniform distribution on alternating permutations. Thus, by first generating a uniformly random permutation as the starting point for this chain and then running the chain from there, it is possible to sample repeatedly from the uniform distribution on tajima trees. Simulations indicate that the moves this chain makes are accepted slightly above half of the time. We note here, that this chain can also be adopted to be reversible with respect to the tajima coalescent. This can be achieved due to the following lemma.

Proposition 5.1. *Let σ_{n-1} be an alternating permutation. Then the number of cherries in the tree that one attains from the bijection presented in [1] is the number of numbers immediately surrounded by parentheses in the bracketing process.*

Proof. This is immediate from the branching process described in [1] □

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.