The α-Ford algebraic measure trees

Josué NUSSBAUMER[1] and Anita WINTER[2]

[1]Université Gustave Eiffel
[2]Universität Duisburg-Essen

June 17, 2021

The α-Ford model [Ford '05]

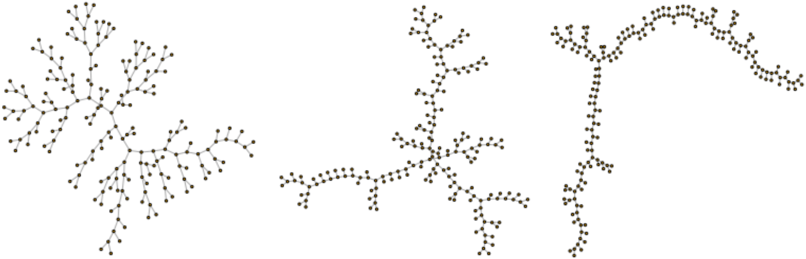


Figure: $\alpha = 0$ on the left; $\alpha = 0.5$ in the middle; $\alpha = 0.9$ on the right.

The α-Ford model [Ford '05]

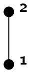
Fix $\alpha \in [0, 1]$. The α-*Ford tree* is a random binary unrooted tree with *n* leaves:

- Start with one edge (yielding 2 leaves).
- Given the α-Ford tree with *k*-leaves, assign weight $1 - \alpha$ to each external and weight α to each internal edge.
- Pick an edge according to its weight, and insert there another leaf.
- Stop when the current binary combinatorial tree has *n* leaves.

The α-Ford model [Ford '05]

Fix $\alpha \in [0,1]$. The α-*Ford tree* is a random binary unrooted tree with n leaves:

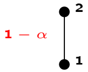
- Start with one edge (yielding 2 leaves).
- Given the α-Ford tree with k-leaves, assign weight $1-\alpha$ to each external and weight α to each internal edge.
- Pick an edge according to its weight, and insert there another leaf.
- Stop when the current binary combinatorial tree has n leaves.



The α-Ford model [Ford '05]

Fix $\alpha \in [0, 1]$. The α-*Ford tree* is a random binary unrooted tree with n leaves:

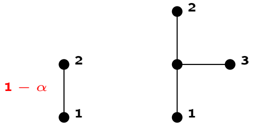
- Start with one edge (yielding 2 leaves).
- Given the α-Ford tree with k-leaves, assign weight $1 - \alpha$ to each external and weight α to each internal edge.
- Pick an edge according to its weight, and insert there another leaf.
- Stop when the current binary combinatorial tree has n leaves.



The α-Ford model [Ford '05]

Fix $\alpha \in [0, 1]$. The *α-Ford tree* is a random binary unrooted tree with n leaves:

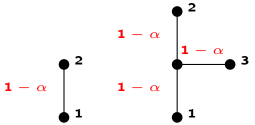
- Start with one edge (yielding 2 leaves).
- Given the α-Ford tree with k-leaves, assign weight $1 - \alpha$ to each external and weight α to each internal edge.
- Pick an edge according to its weight, and insert there another leaf.
- Stop when the current binary combinatorial tree has n leaves.



The α-Ford model [Ford '05]

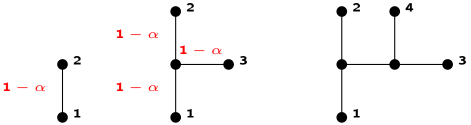
Fix $\alpha \in [0,1]$. The *α-Ford tree* is a random binary unrooted tree with n leaves:

- Start with one edge (yielding 2 leaves).
- Given the α-Ford tree with k-leaves, assign weight $1-\alpha$ to each external and weight α to each internal edge.
- Pick an edge according to its weight, and insert there another leaf.
- Stop when the current binary combinatorial tree has n leaves.



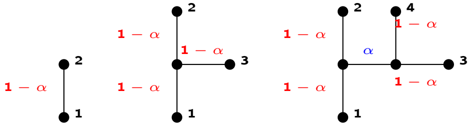
Fix $\alpha \in [0,1]$. The α-*Ford tree* is a random binary unrooted tree with n leaves:

- Start with one edge (yielding 2 leaves).
- Given the α-Ford tree with k-leaves, assign weight $1-\alpha$ to each external and weight α to each internal edge.
- Pick an edge according to its weight, and insert there another leaf.
- Stop when the current binary combinatorial tree has n leaves.



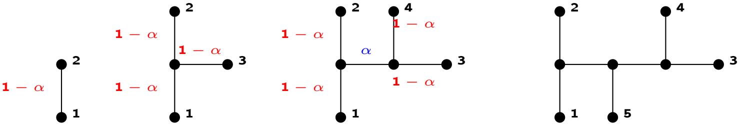
Fix $\alpha \in [0, 1]$. The *α-Ford tree* is a random binary unrooted tree with n leaves:

- Start with one edge (yielding 2 leaves).
- Given the α-Ford tree with k-leaves, assign weight $1 - \alpha$ to each external and weight α to each internal edge.
- Pick an edge according to its weight, and insert there another leaf.
- Stop when the current binary combinatorial tree has n leaves.



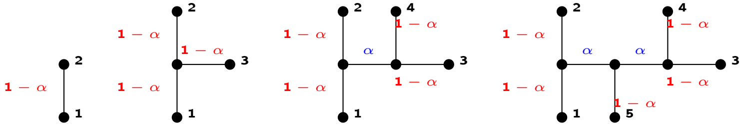
Fix $\alpha \in [0, 1]$. The *α-Ford tree* is a random binary unrooted tree with n leaves:

- Start with one edge (yielding 2 leaves).
- Given the α-Ford tree with k-leaves, assign weight $1 - \alpha$ to each external and weight α to each internal edge.
- Pick an edge according to its weight, and insert there another leaf.
- Stop when the current binary combinatorial tree has n leaves.



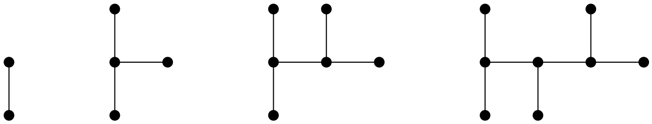
Fix $\alpha \in [0, 1]$. The α-*Ford tree* is a random binary unrooted tree with n leaves:

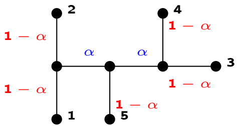
- Start with one edge (yielding 2 leaves).
- Given the α-Ford tree with k-leaves, assign weight $1 - \alpha$ to each external and weight α to each internal edge.
- Pick an edge according to its weight, and insert there another leaf.
- Stop when the current binary combinatorial tree has n leaves.



Fix $\alpha \in [0, 1]$. The α-*Ford tree* is a random binary unrooted tree with n leaves:

- Start with one edge (yielding 2 leaves).
- Given the α-Ford tree with k-leaves, assign weight $1 - \alpha$ to each external and weight α to each internal edge.
- Pick an edge according to its weight, and insert there another leaf.
- Stop when the current binary combinatorial tree has n leaves.

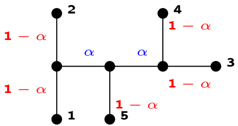




Special cases:

- $\alpha = 0$: *Kingman coalescent tree* whithout edge lengths or *Yule tree*.

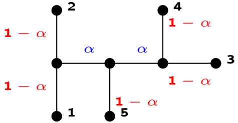
The α-Ford model [Ford '05]



Special cases:

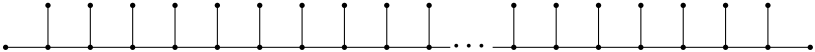
- $\alpha = 0$: *Kingman coalescent tree* whithout edge lengths or *Yule tree*.
- $\alpha = \frac{1}{2}$: *uniform tree* [Aldous '93].

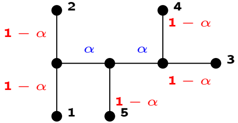
The α-Ford model [Ford '05]



Special cases:

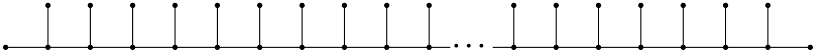
- $\alpha = 0$: *Kingman coalescent tree* whithout edge lengths or *Yule tree*.
- $\alpha = \frac{1}{2}$: *uniform tree* [Aldous '93].
- $\alpha = 1$: *comb tree*.





Special cases:

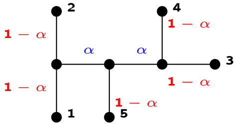
- $\alpha = 0$: *Kingman coalescent tree* whithout edge lengths or *Yule tree*.
- $\alpha = \frac{1}{2}$: *uniform tree* [Aldous '93].
- $\alpha = 1$: *comb tree*.



Goal

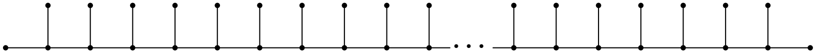
What if the tree gets large?

The α-Ford model [Ford '05]



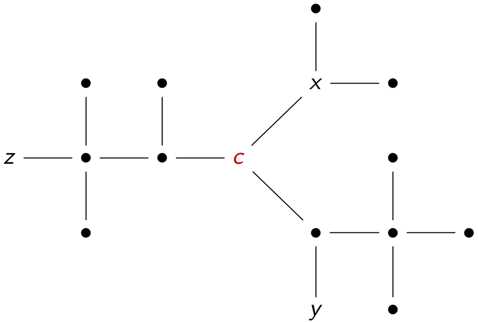
Special cases:

- $\alpha = 0$: *Kingman coalescent tree* whithout edge lengths or *Yule tree*.
- $\alpha = \frac{1}{2}$: *uniform tree* [Aldous '93].
- $\alpha = 1$: *comb tree*.



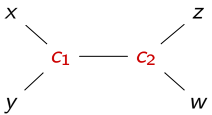
Goal

What if the tree gets large? Need a *state space* of *(infinite) trees* with a "nice" *topology*.



For $x, y, z \in T$ there exists a unique *branch point* $c = c(x, y, z) \in T$ with

$$[x, y] \cap [y, z] \cap [z, x] = \{c\}.$$



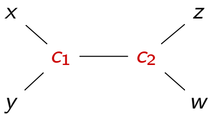
Definition: algebraic tree

An algebraic tree is a set T together with a symmetric map $c\colon T^3 \to T$ such that:

- (2-point condition) For $x, y \in T$, $c(x, y, y) = y$.
- (3-point condition) For $x, y, z \in T$, $c(x, y, c(x, y, z)) = c(x, y, z)$.
- (4-point condition) For $x, y, z, w \in T$,

$$c(x, y, z) \in \{c(x, y, w), c(x, z, w), c(y, z, w)\}.$$

The space of algebraic measure trees [Löhr, Winter '18]



Definition: algebraic tree

An algebraic tree is a set T together with a symmetric map $c\colon T^3 \to T$ such that:

- (2-point condition) For $x, y \in T$, $c(x, y, y) = y$.
- (3-point condition) For $x, y, z \in T$, $c(x, y, c(x, y, z)) = c(x, y, z)$.
- (4-point condition) For $x, y, z, w \in T$,

$$c(x, y, z) \in \{c(x, y, w), c(x, z, w), c(y, z, w)\}.$$

Definition

An algebraic measure tree (T, c, μ) consists of an order separable algebraic tree (T, c) together with a probability measure μ on $(T, \mathcal{B}(T, c))$.

We consider the subset of *binary algebraic measure trees*:

$$\mathbb{T}_2 := \{(T, c, \mu) : \deg(v) \leqslant 3 \ \forall v \in T, \ \operatorname{atoms}(\mu) \subseteq \operatorname{leaves}(T)\}.$$

We consider the subset of *binary algebraic measure trees*:

$$\mathbb{T}_2 := \{(T, c, \mu) : \deg(v) \leqslant 3 \ \forall v \in T, \ \operatorname{atoms}(\mu) \subseteq \operatorname{leaves}(T)\}.$$

Idea: "Gromov-weak" topology

A sequence of trees converges to a limit tree if and only if all randomly sampled finite subtrees converge to the corresponding limit subtrees.

When we sample n points according to μ, we can look at the tree shape they span.

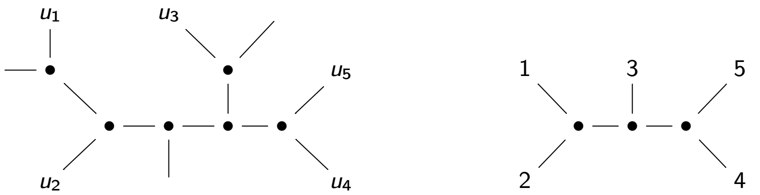


Figure: A tree T and the shape spanned by 5 points.

When we sample n points according to μ, we can look at the tree shape they span.

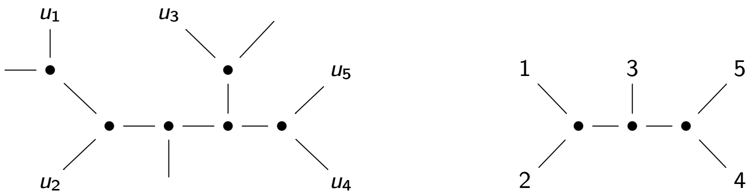


Figure: A tree T and the shape spanned by 5 points.

A sequence of algebraic measure trees is sample-shape convergent if for all n, the distributions of the n-tree shapes converge.

Let $\alpha \in [0, 1]$. For all $n \in \mathbb{N}$, the α-Ford tree with n leaves defines a random tree in $\mathbb{T}_2$ (we put the uniform distribution on the leaves).

Let $\alpha \in [0, 1]$. For all $n \in \mathbb{N}$, the α-Ford tree with n leaves defines a random tree in $\mathbb{T}_2$ (we put the uniform distribution on the leaves).

Proposition [N., Winter '20]

As the number of leaves goes to infinity, the α-Ford model converges in distribution in $\mathbb{T}_2$ to a random *continuum* algebraic measure tree.

Let $\alpha \in [0,1]$. For all $n \in \mathbb{N}$, the α-Ford tree with n leaves defines a random tree in $\mathbb{T}_2$ (we put the uniform distribution on the leaves).

Proposition [N., Winter '20]

As the number of leaves goes to infinity, the α-Ford model converges in distribution in $\mathbb{T}_2$ to a random *continuum* algebraic measure tree.

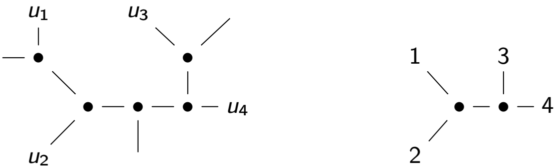
We call the limit the α-Ford algebraic measure tree.

- $\alpha = 0$: the Kingman algebraic measure tree.
- $\alpha = \frac{1}{2}$: the algebraic measure Brownian Continuum Random Tree [Löhr, Mytnik, Winter '18].

We ignore the branch lengths, so we can not look at total length.

We ignore the branch lengths, so we can not look at total length.

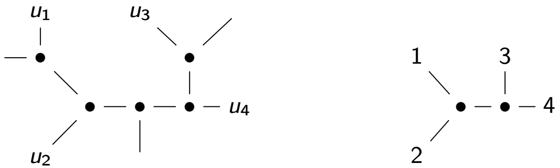
- **Sample shapes:** distribution of the shape spanned by m points sampled according to μ.



Statistics on the limit trees

We ignore the branch lengths, so we can not look at total length.

- **Sample shapes:** distribution of the shape spanned by m points sampled according to μ.



- **Subtree masses**: sample 3 points according to μ and look at how the mass is distributed around their branch point.

For $\underline{u} = u_1, u_2, u_3$ three leaves of T, we consider the vector of the three masses of the components connected to $c(\underline{u})$:

$$\underline{\eta}(u_1, u_2, u_3) = (\eta_1(\underline{u}), \eta_2(\underline{u}), \eta_3(\underline{u})).$$

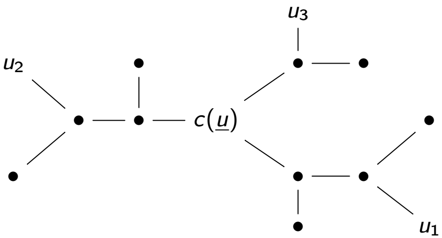


Figure: A tree with 8 leaves; $\underline{\eta}(\underline{u}) = \left(\frac{3}{8}, \frac{3}{8}, \frac{2}{8}\right)$.

We look at polynomials of the form

$$\Phi^f(\chi) := \int_{T^3} \mu^{\otimes 3}(\mathrm{d}\underline{u}) f(\underline{\eta}(\underline{u})).$$

Proposition [Aldous '94]

Let $\mathbb{P}_{\mathrm{CRT}}$ be the law of the Brownian algebraic CRT. Then for all $f\colon \Delta_2 \to \mathbb{R}$ continuous bounded,

$$\mathbb{E}_{\mathrm{CRT}}\left[\int_{T^3} \mu^{\otimes 3}(\mathrm{d}\underline{u}) f(\underline{\eta}(\underline{u}))\right] = \int_{\Delta_2} f(\underline{x}) \mathrm{Dir}\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)(\mathrm{d}\underline{x}),$$

where $\mathrm{Dir}\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$ is the Dirichlet distribution.

Proposition [N., Winter '20]

Let $\mathbb{P}_{\mathrm{Kin}}$ be the law of the Kingman algebraic measure tree. Let $B_{1,2}$ and $B_{2,2}$ be two independent beta random variables, such that $B_{1,2}$ has law $\mathrm{Beta}(1,2)$ and $B_{2,2}$ has law $\mathrm{Beta}(2,2)$. Then for all $f\colon \Delta_2 \to \mathbb{R}$ continuous bounded,

$$\begin{aligned}\mathbb{E}_{\mathrm{Kin}}&\left[\int_{T^3} \mu^{\otimes 3}(\mathrm{d}\underline{u}) f(\underline{\eta}(\underline{u}))\right] \\ &= \frac{1}{6} \sum_{\pi \in \mathcal{S}_3} \mathbb{E}\big[f \circ \pi^*(B_{1,2}B_{2,2}, B_{1,2}(1-B_{2,2}), 1-B_{1,2})\big],\end{aligned}$$

where $\mathcal{S}_3$ is the set of permutations $\{1,2,3\}$, and for $\pi \in \mathcal{S}_3$, $\pi^*\colon \Delta_2 \to \Delta_2$ is the induced map $\pi^*(\underline{x}) = (x_{\pi(1)}, x_{\pi(2)}, x_{\pi(3)})$.

Subtree masses

$\rightarrow$ What about the other $\alpha \in [0, 1]$?

$\rightarrow$ What about the other $\alpha \in [0,1]$?

We can calculate all the moments through a recurrence relation.

$$\mathbb{E}_\alpha\left[\int_{T^3} \mu^{\otimes 3}(\mathrm{d}\underline{u})\eta_1(\underline{u})\right] = \frac{1}{3},$$

$$\mathbb{E}_\alpha\left[\int_{T^3} \mu^{\otimes 3}(\mathrm{d}\underline{u})(\eta_1(\underline{u}))^2\right] = \frac{1}{5},$$

$$\mathbb{E}_\alpha\left[\int_{T^3} \mu^{\otimes 3}(\mathrm{d}\underline{u})(\eta_1(\underline{u}))^3\right] = \frac{11-7\alpha}{15(5-3\alpha)},$$

$$\mathbb{E}_\alpha\left[\int_{T^3} \mu^{\otimes 3}(\mathrm{d}\underline{u})(\eta_1(\underline{u}))^4\right] = \frac{37-25\alpha}{63(5-3\alpha)},$$

$$\mathbb{E}_\alpha\left[\int_{T^3} \mu^{\otimes 3}(\mathrm{d}\underline{u})(\eta_1(\underline{u}))^5\right] = \frac{145-165\alpha+44\alpha^2}{42(5-3\alpha)(7-3\alpha)}.$$

D. Aldous.
The Continuum Random Tree III.
The Annals of Probability, *Vol. 21, No. 1, 248-289*, 1993.

D. Aldous.
Recursive Self-Similarity for Random Trees, Random Triangulations and Brownian Excursion.
The Annals of Probability, *Vol. 22, No. 2, 527-545.*, 1994.

D.J. Ford.
Probabilities on cladograms: introduction to the alpha model.
arXiv:math/0511246, 2005.

W. Löhr, L. Mytnik and A. Winter.
The Aldous chain on cladograms in the diffusion limit.
Annals of Probability, Volume 48, Number 5, 2565-2590, 2020.

W. Löhr and A. Winter.
Spaces of algebraic measure trees and triangulations of the circle.
Bulletin de la Société Mathématique de France, 149 (1), p. 1-63, 2021.

J. Nussbaumer and A. Winter.
The algebraic α-Ford tree under evolution.
arXiv:2006.09316, 2020.