Random factorizations of a long cycle into transpositions

Valentin Féray
(joint work with Igor Kortchemski)

Institut für Mathematik, Universität Zürich

Nancy, October 2018

Our model

We consider minimal factorizations of a full cycle into transpositions.

$$\mathfrak{M}_n := \{(\tau_1, \dots, \tau_{n-1}) \text{ transpositions s.t. } \tau_1\tau_2\cdots\tau_{n-1} = (1, 2, \dots, n)\}$$

Well-known: $|\mathfrak{M}_n| = n^{n-2}$ (bijections with Cayley trees, parking functions)

Our model

We consider minimal factorizations of a full cycle into transpositions.

$$\mathfrak{M}_n := \{(\tau_1, \dots, \tau_{n-1}) \text{ transpositions s.t. } \tau_1 \tau_2 \cdots \tau_{n-1} = (1, 2, \dots, n)\}$$

Well-known: $|\mathfrak{M}_n| = n^{n-2}$ (bijections with Cayley trees, parking functions)

Take a uniform random minimal factorization of size n

$$(\boldsymbol{\tau_1}, \dots, \boldsymbol{\tau_{n-1}}).$$

Our model

We consider minimal factorizations of a full cycle into transpositions.

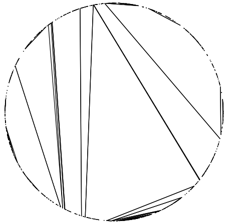
$$\mathfrak{M}_n := \{(\tau_1, \dots, \tau_{n-1}) \text{ transpositions s.t. } \tau_1 \tau_2 \cdots \tau_{n-1} = (1, 2, \dots, n)\}$$

Well-known: $|\mathfrak{M}_n| = n^{n-2}$ (bijections with Cayley trees, parking functions)

Take a uniform random minimal factorization of size n

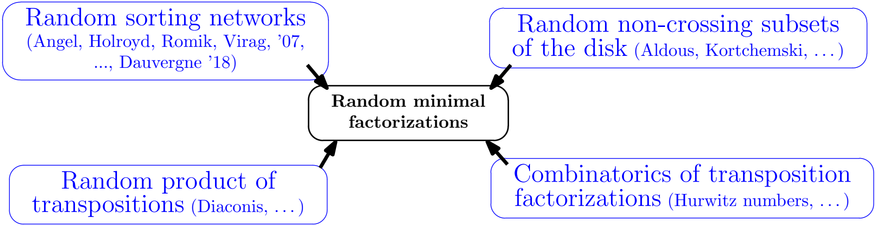
$$(\boldsymbol{\tau_1}, \dots, \boldsymbol{\tau_{n-1}}).$$

We are interested in partial products $\boldsymbol{\tau_1} \cdots \boldsymbol{\tau_k}$. They can be geometrically represented by non-crossing sets of chords.



Goal: find the limit of these random non-crossing objects.

Motivations



Geometric representations as non-crossing objects

With a minimal factorization $(\tau_1, \dots, \tau_{n-1})$ of size n and an integer $k \leq n-1$, we associate two non-crossing objects.

Example: take $n = 12$, $k = 6$, and

$\big((1,3),(6,12),(1,5),(7,12),(9,10),(11,12),(2,3),(4,5),(1,6),(8,11),(9,11)\big) \in \mathfrak{M}_{12}$

Geometric representations as non-crossing objects

With a minimal factorization $(\tau_1, \dots, \tau_{n-1})$ of size n and an integer $k \leq n-1$, we associate two non-crossing objects.

Example: take $n = 12$, $k = 6$, and

$\big((1,3),(6,12),(1,5),(7,12),(9,10),(11,12),(2,3),(4,5),(1,6),(8,11),(9,11)\big) \in \mathfrak{M}_{12}$

For both constructions, we only consider the first k factors.

Geometric representations as non-crossing objects

With a minimal factorization $(\tau_1, \ldots, \tau_{n-1})$ of size n and an integer $k \leq n-1$, we associate two non-crossing objects.

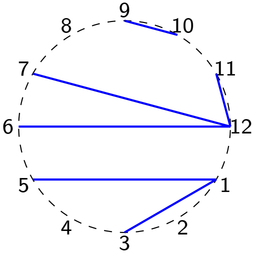
Example: take $n = 12$, $k = 6$, and

$((1,3),(6,12),(1,5),(7,12),(9,10),(11,12),(2,3),(4,5),(1,6),(8,11),(9,11)) \in \mathfrak{M}_{12}$

For both constructions, we only consider the first k factors.

First construction $\boldsymbol{F_k^n}$

Each transposition corresponds to an edge.



Geometric representations as non-crossing objects

With a minimal factorization $(\tau_1, \dots, \tau_{n-1})$ of size n and an integer $k \leq n-1$, we associate two non-crossing objects.

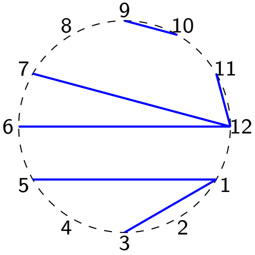
Example: take $n = 12$, $k = 6$, and

$((1,3),(6,12),(1,5),(7,12),(9,10),(11,12),(2,3),(4,5),(1,6),(8,11),(9,11)) \in \mathfrak{M}_{12}$

For both constructions, we only consider the first k factors.

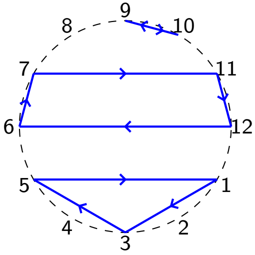
First construction F_k^n

Each transposition corresponds to an edge.



Second construction P_k^n

Draw the cycles of the partial product.



Main result

Theorem (F., Kortchemski, '18)

There exists a family of random compact subset $(\mathbf{L}_c)_{c\in[0,\infty]}$ of the unit disk such that the following holds. (We use Hausdorff metric.)

(i) *Assume that $K_n \to \infty$ and $\frac{K_n}{\sqrt{n}} \to c$, with $c < \infty$. Then*

$$\left(\boldsymbol{F}^{\boldsymbol{n}}_{\boldsymbol{K_n}}, \boldsymbol{P}^{\boldsymbol{n}}_{\boldsymbol{K_n}}\right) \xrightarrow[n\to\infty]{(d)} (\mathbf{L}_c, \mathbf{L}_c).$$

(ii) *Assume that $\frac{K_n}{\sqrt{n}} \to \infty$ and that $\frac{n-K_n}{\sqrt{n}} \to \infty$. Then*

$$\left(\boldsymbol{F}^{\boldsymbol{n}}_{\boldsymbol{K_n}}, \boldsymbol{P}^{\boldsymbol{n}}_{\boldsymbol{K_n}}\right) \xrightarrow[n\to\infty]{(d)} (\mathbf{L}_\infty, \mathbf{L}_\infty).$$

(iii) *Assume that $\frac{n-K_n}{\sqrt{n}} \to c$, with $c < \infty$. Then*

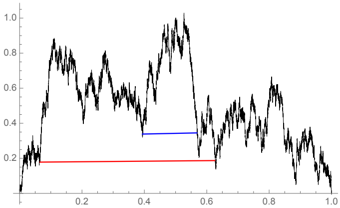
$$\boldsymbol{F}^{\boldsymbol{n}}_{\boldsymbol{K_n}} \xrightarrow[n\to\infty]{(d)} \mathbf{L}_\infty, \qquad \boldsymbol{P}^{\boldsymbol{n}}_{\boldsymbol{K_n}} \xrightarrow[n\to\infty]{(d)} \mathbf{L}_c.$$

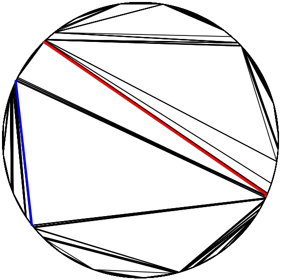
What are the limit objects?

- $\mathbf{L}_0$ is the unit circle.
  i.e., according to our theorem, when $K_n = o(\sqrt{n})$, there is no macroscopic chord in the first K_n factors.

What are the limit objects?

- $\mathbf{L}_0$ is the unit circle.
- $\mathbf{L}_\infty$ is Aldous' Brownian triangulation.
  Start from a Brownian excursion $\boldsymbol{X}_\infty^{\text{exc}}$ and draw a chord $[e^{-2\pi i s}, e^{-2\pi i t}]$ for each tunnel (s,t) in $\boldsymbol{X}_\infty^{\text{exc}}$.

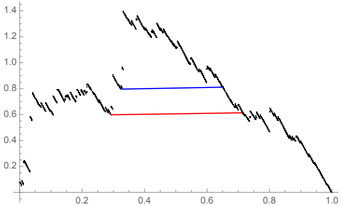


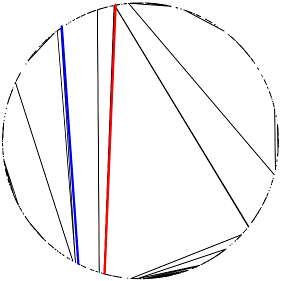


Simulation of $\boldsymbol{X}_\infty^{\text{exc}}$ and $\boldsymbol{L}_\infty$.

What are the limit objects?

- $\mathbf{L}_0$ is the unit circle.
- $\mathbf{L}_\infty$ is Aldous' Brownian triangulation.
- $\mathbf{L}_c$ is obtained in a similar way as $\mathbf{L}_\infty$, except that we start from a certain Lévy excursion process $\boldsymbol{X}_c^{\text{exc}}$ (with discontinuities).





Simulation of $\boldsymbol{X}_5^{\text{exc}}$ and $\boldsymbol{L}_5$.

What are the limit objects?

- $\mathbf{L}_0$ is the unit circle.
- $\mathbf{L}_\infty$ is Aldous' Brownian triangulation.
- $\mathbf{L}_c$ is obtained in a similar way as $\mathbf{L}_\infty$, except that we start from a certain Lévy excursion process $\boldsymbol{X}_c^{\text{exc}}$ (with discontinuities).

About the underlying Lévy process X_t:

$$\mathbb{E}e^{-\lambda X_t} = e^{t\,c^2\left(1-\sqrt{1+\frac{2\lambda}{c}}\right)+t\,\lambda c}, \qquad \lambda \geq 0.$$

It is spectrally positive, makes infinitely many jumps on all intervals, X_t = inverse Gaussian process + a negative drift $(-ct)$.

Illustrating movie

We take a uniform random minimal factorization of size $n = 20,000$.

The following movies displays the non-crossing forest $\boldsymbol{F^n_{K_n}}$ and the non-crossing partition $\boldsymbol{P^n_{K_n}}$ for $K_n = t(n-1)$ (t varying from 0 to 1);

Movie 1

Illustrating movie

We take a uniform random minimal factorization of size $n = 20,000$.

The following movies displays the non-crossing forest $\boldsymbol{F^n_{K_n}}$ and the non-crossing partition $\boldsymbol{P^n_{K_n}}$ for $K_n = t(n-1)$ (t varying from 0 to 1);

Movie 1

Our theorem describes the marginals of the above process.

☹ We have no limit result for the process itself.

A "concrete" corollary

For a transposition $\tau = (i,j)$ with $i < j$, set $w(\tau) := \min(j-i, n+i-j)$.

Corollary

Let $(\boldsymbol{\tau_1}, \dots, \boldsymbol{\tau_{n-1}})$ *be a uniform random factorization of size n. Then the random variable*

$$n^{-1} \max_{1 \le i \le K_n} w(\boldsymbol{\tau_i})$$

A "concrete" corollary

For a transposition $\tau = (i, j)$ with $i < j$, set $w(\tau) := \min(j - i, n + i - j)$.

Corollary

Let $(\boldsymbol{\tau_1}, \dots, \boldsymbol{\tau_{n-1}})$ be a uniform random factorization of size n. Then the random variable

$$n^{-1} \max_{1 \le i \le K_n} w(\boldsymbol{\tau_i})$$

(i) *tends to 0 in probability if $K_n = o(\sqrt{n})$;*

(ii) *tends to some non-trivial random variable ℓ_c if $\lim \frac{K_n}{\sqrt{n}} = c$ (for c in $(0, +\infty]$).*

The law of ℓ_∞ is explicit (computed by Aldous):

$$\frac{1}{\pi} \frac{3\theta - 1}{\theta^2 (1-\theta)^2 \sqrt{1 - 2\theta}} \mathbf{1}_{\frac{1}{3} \le \theta \le \frac{1}{2}} \, d\theta, \text{ where } \ell = 2\sin(\pi\theta).$$

Remarkable that the limit is the same for $K_n = \sqrt{n}\log(n)$ or $K_n = n - 1$!

Structure of the proof (for P_k^n)

1. The distribution of the random non-crossing partition $\mathbf{P_k^n}$ is explicit;

Structure of the proof (for P_k^n)

1. The distribution of the random non-crossing partition $\mathbf{P_k^n}$ is explicit;
2. We use a bijection $\mathcal{T}$ from non-crossing partitions to trees;

Structure of the proof (for P_k^n)

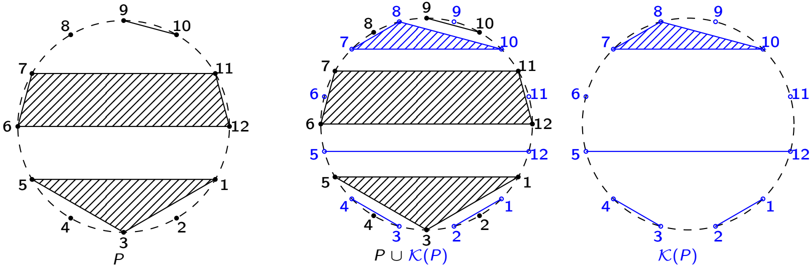
1. The distribution of the random non-crossing partition $\mathbf{P_k^n}$ is explicit;
2. We use a bijection $\mathcal{T}$ from non-crossing partitions to trees;
3. Combining 1 and 2, we see that $\mathcal{T}(\mathbf{P_k^n})$ has the distribution of a conditioned Galton-Watson tree;

Structure of the proof (for P_k^n)

1. The distribution of the random non-crossing partition $\mathbf{P_k^n}$ is explicit;
2. We use a bijection $\mathcal{T}$ from non-crossing partitions to trees;
3. Combining 1 and 2, we see that $\mathcal{T}(\mathbf{P_k^n})$ has the distribution of a conditioned Galton-Watson tree;
4. We analyse this random tree model.

An explicit formula for the distribution of $\mathbf{P^n_k}$

Given a non-crossing partition P, we can construct *its Kreweras complement* $\mathcal{K}(P)$. On an example,



Blocks of $\mathcal{K}(P)$ = white "faces" between blocks of P

Fact: blocks of P $\leftrightarrow$ cycles of σ.
blocks of $\mathcal{K}(P)$ $\leftrightarrow$ cycles of $\sigma^{-1}\,(1\ 2\ \ldots\ n)$

An explicit formula for the distribution of $\mathbf{P_k^n}$

Given a non-crossing partition P, we can construct *its Kreweras complement* $\mathcal{K}(P)$.

Proposition

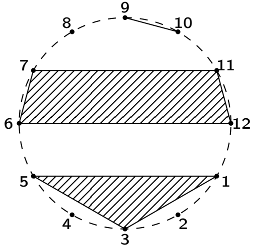
Fix $1 \leq k \leq n-1$ and let P be a non-crossing partition with $n-k$ blocks. Then

$$\mathbb{P}\left(\mathbf{P_k^n} = P\right) = \frac{k!(n-k-1)!}{n^{n-2}} \cdot \left(\prod_{B \in P \uplus \mathcal{K}(P)} \frac{|B|^{|B|-2}}{(|B|-1)!} \right).$$

The proof is easy; it relies on the fact that the number of minimal factorizations of any permutation into transpositions is explicit.

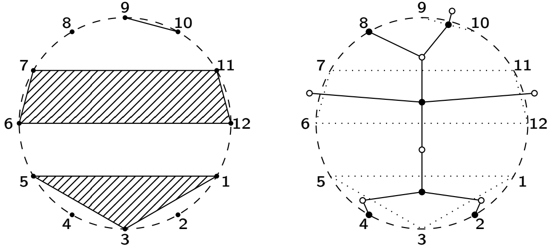
A bijection $\mathcal{T}$ from non-crossing partitions to trees

1. Start from a non-crossing partition;



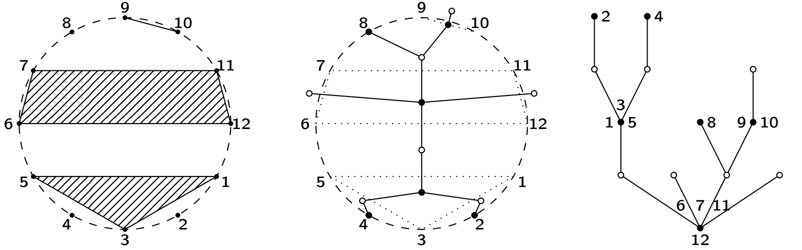
A bijection $\mathcal{T}$ from non-crossing partitions to trees

1. Start from a non-crossing partition;
2. Associate a black vertex to each block, a white vertex to each block of the Kreweras complement, and join neighbour blocks;
   (We get a properly bicolored plane tree with labeled black corners.)



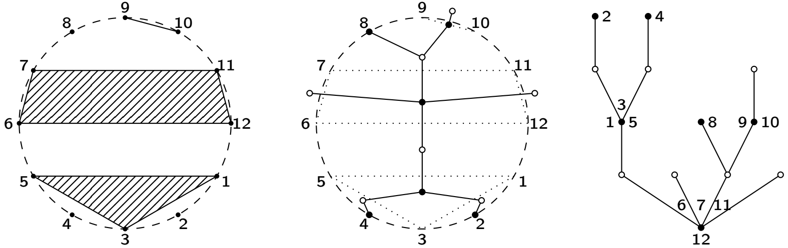
A bijection $\mathcal{T}$ from non-crossing partitions to trees

1. Start from a non-crossing partition;
2. Associate a black vertex to each block, a white vertex to each block of the Kreweras complement, and join neighbour blocks;
   (We get a properly bicolored plane tree with labeled black corners.)
3. Root the resulting tree in the corner n.



A bijection $\mathcal{T}$ from non-crossing partitions to trees

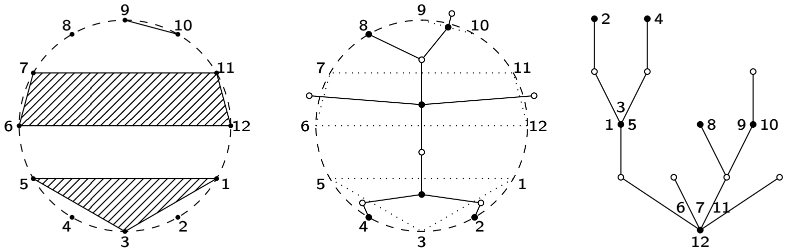
1. Start from a non-crossing partition;
2. Associate a black vertex to each block, a white vertex to each block of the Kreweras complement, and join neighbour blocks;
   (We get a properly bicolored plane tree with labeled black corners.)
3. Root the resulting tree in the corner n.



Inverse: label black corners by the order of visit of the contour of trees. Each block corresponds to the corners of some black vertex.

A bijection $\mathcal{T}$ from non-crossing partitions to trees

1. Start from a non-crossing partition;
2. Associate a black vertex to each block, a white vertex to each block of the Kreweras complement, and join neighbour blocks;
   (We get a properly bicolored plane tree with labeled black corners.)
3. Root the resulting tree in the corner n.



Here, and throughout this talk, black (resp. white) vertices are vertices at even (resp. odd) height.

Conditioned Galton-Watson trees

Reminder:
if P is a non-crossing partition of n with $n-k$ blocks,

$$\mathbb{P}(\mathbf{P_k^n} = P) \propto \prod_{B \in P \uplus \mathcal{K}(P)} \frac{|B|^{|B|-2}}{(|B|-1)!}.$$

Conditioned Galton-Watson trees

Using the bijection:
if T is a tree with $n-k$ black and $k+1$ white vertices, then

$$\mathbb{P}(\mathcal{T}(\mathbf{P_k^n}) = T) \propto \prod_{v \in V_T} \frac{(d_v+1)^{d_v-1}}{d_v!},$$

where d_v is the out-degree of v (small correction needed at the root).

Conditioned Galton-Watson trees

Using the bijection:
if T is a tree with $n-k$ black and $k+1$ white vertices, then

$$\mathbb{P}(\mathcal{T}(\mathbf{P_k^n}) = T) \propto \prod_{v \in V_T} \frac{(d_v+1)^{d_v-1}}{d_v!},$$

where d_v is the out-degree of v (small correction needed at the root).

T has the law of a Galton-Watson tree...
conditioned to have $n-k$ black vertices and $k+1$ white vertices.

Conditioned Galton-Watson trees

Using the bijection:
if T is a tree with $n-k$ black and $k+1$ white vertices, then

$$\mathbb{P}\left(\mathcal{T}(\mathbf{P^n_k}) = T\right) \propto \prod_{v \in V_T} \frac{(d_v+1)^{d_v-1}}{d_v!},$$

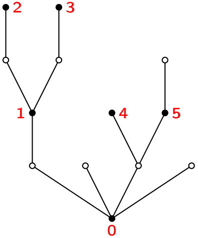
where d_v is the out-degree of v (small correction needed at the root).

T has the law of a Galton-Watson tree...
conditioned to have $n-k$ black vertices and $k+1$ white vertices.

☹ When $k = \Theta(\sqrt{n})$, the conditioning event has exponentially small probability....

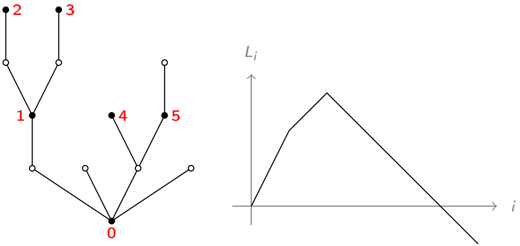
$\rightarrow$ we see T as a conditioned bi-type Galton-Watson tree, with offspring distributions depending on n.

Łukasiewicz paths



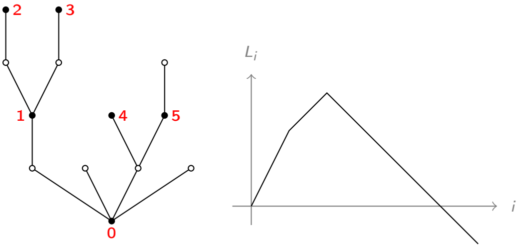
- Start with a tree T and label its black vertices in depth-first search order.

Łukasiewicz paths



- Start with a tree T and label its black vertices in depth-first search order.
- Łukasiewicz path: $L_0 = 0$ and $L_{i+1} = L_i + \mathrm{GC}_i - 1$, where GC_i is the number of (black) grandchildren of i.

Łukasiewicz paths

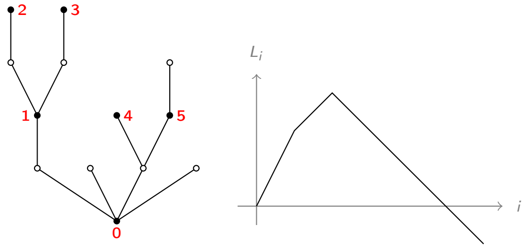


- Start with a tree T and label its black vertices in depth-first search order.
- Łukasiewicz path: $L_0 = 0$ and $L_{i+1} = L_i + \mathrm{GC}_i - 1$, where GC_i is the number of (black) grandchildren of i.

Why Łukasiewicz paths?

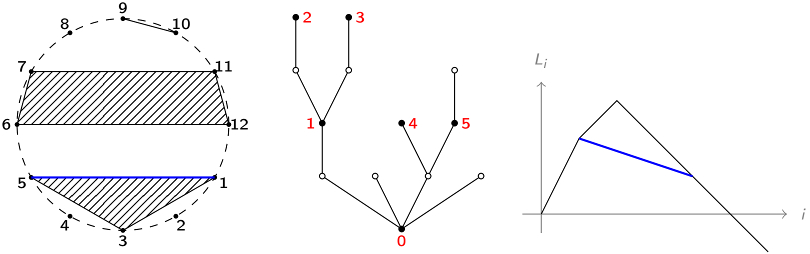
For many models of (bitype) conditioned GW trees, the Łukasiewicz path is a conditioned random walk, which we can analyse asymptotically.

Łukasiewicz paths



In our model, the Łukasiewicz path of $\mathcal{T}(\boldsymbol{P}^{\boldsymbol{n}}_{\boldsymbol{K_n}})$ tends (after normalization) to the excursion process $\boldsymbol{X}_c^{\text{exc}}$.

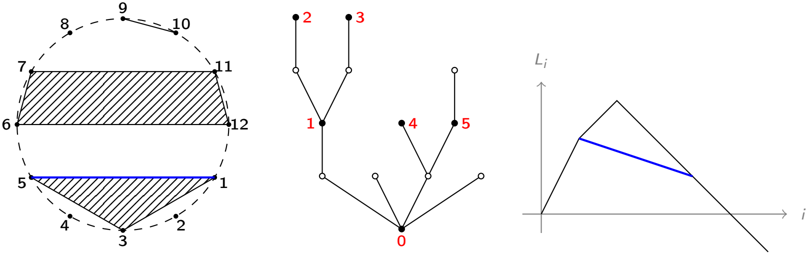
Łukasiewicz paths



In our model, the Łukasiewicz path of $\mathcal{T}(\boldsymbol{P^n_{K_n}})$ tends (after normalization) to the excursion process $\boldsymbol{X}_c^{\mathrm{exc}}$.

Moreover, chords in P **asymptotically** correspond to tunnels in the Łukasiewicz path of $\mathcal{T}(P)$.

Łukasiewicz paths



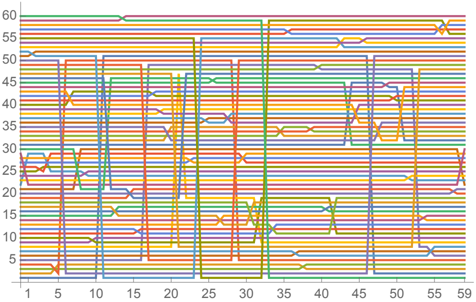
In our model, the Łukasiewicz path of $\mathcal{T}(\boldsymbol{P}^{\boldsymbol{n}}_{\boldsymbol{K_n}})$ tends (after normalization) to the excursion process $\boldsymbol{X}^{\text{exc}}_c$.

Moreover, chords in P **asymptotically** correspond to tunnels in the Łukasiewicz path of $\mathcal{T}(P)$.

Conclusion: $\boldsymbol{P}^{\boldsymbol{n}}_{\boldsymbol{K_n}}$ tends to $\boldsymbol{L}_c$.

Local convergence of trajectories (1/3)

Our "scaling limit" result gives no information on what happens for some fixed elements, *e.g.* what is the trajectory of a given i? The number of transpositions containing it?

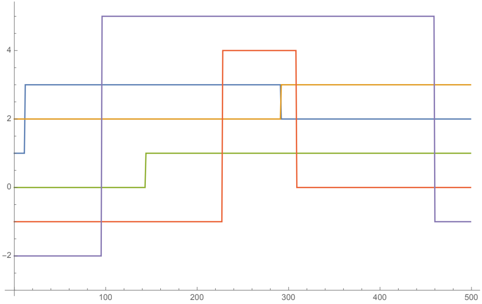


Local convergence of trajectories (2/3)

We work with factorizations of the cycle

$$(-\lfloor (n-1)/2 \rfloor, \dots, 0, 1, \dots, \lfloor n/2 \rfloor)$$

and look around $i = 0$:



Trajectories of all i in $\{-2, -1, 0, 1, 2\}$
in a random minimal factorization of size $n = 500$.

Local convergence of trajectories (2/3)

We work with factorizations of the cycle

$$(-\lfloor (n-1)/2 \rfloor, \dots, 0, 1, \dots, \lfloor n/2 \rfloor)$$

and look around $i = 0$:

Theorem (F., Kortchemski)

There is a family $(X_i)_{i \in \mathbb{Z}}$ of integer-valued step functions on $[0,1]$

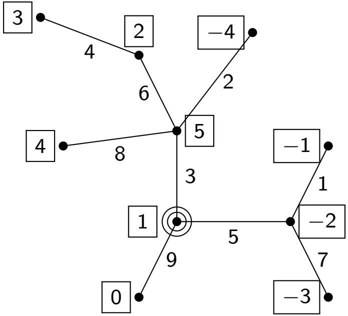
$$\left(X_i^{(n)}(\lfloor nt \rfloor) : 0 \le t \le 1 \right)_{i \in \mathbb{Z}} \xrightarrow[n\to\infty]{(d)} (X_i)_{i \in \mathbb{Z}}$$

holds in distribution (in the sense of convergence of finite dimensional marginals).

Note: there is a rescaling in time, but not in space.

Local convergence of trajectories (3/3)

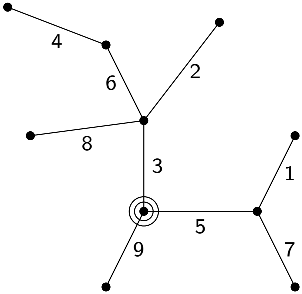
For this we need to study the local convergence around 1 of the edge and vertex-labelled tree of the factorization.



Tree of the factorization
(-1,-2), (-4, 5), (1, 5), (2, 3), (-2, 1), (2, 5), (-3,-2), (4, 5), (0, 1)

Local convergence of trajectories (3/3)

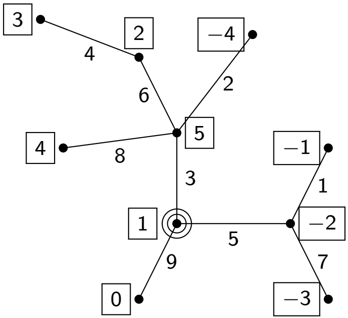
For this we need to study the local convergence around 1 of the edge and vertex-labelled tree of the factorization.



- Standard: the edge-labelled tree converges to "Kesten tree".

Local convergence of trajectories (3/3)

For this we need to study the local convergence around 1 of the edge and vertex-labelled tree of the factorization.



- Standard: the edge-labelled tree converges to "Kesten tree".
- Difficulty: understand the vertex labelling and prove that it's "local".

Conclusion

1. Global and local convergence results for uniform random factorizations of a long cycle into transposition.

2. For the global convergence, we are missing the convergence of the process of non-crossing partitions, we only have convergence of the marginals.

3. Factorization models have a very rich combinatorics (Hurwitz numbers, maps, genomics), but there are almost no probabilistic results. There is work to do...

Conclusion

1. Global and local convergence results for uniform random factorizations of a long cycle into transposition.
2. For the global convergence, we are missing the convergence of the process of non-crossing partitions, we only have convergence of the marginals.
3. Factorization models have a very rich combinatorics (Hurwitz numbers, maps, genomics), but there are almost no probabilistic results. There is work to do...

Thank you for your attention