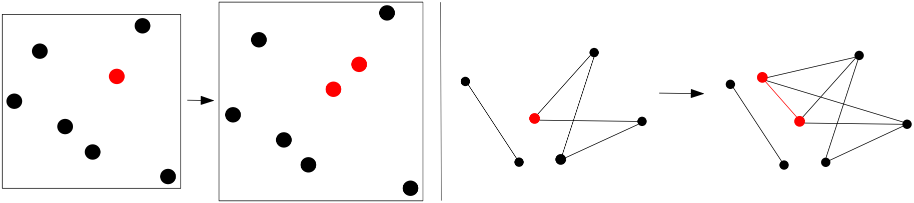
Up-down chains and scaling limits: construction of permuton- and graphon-valued diffusions

Valentin Féray
joint work with Kelvin Rivera-Lopez

CNRS, Institut Élie Cartan de Lorraine (IECL)

5eme rencontre de l'ANR Cortipom
Tours, Oct. 1st, 2024

An updown Markov chain on permutations/graphs

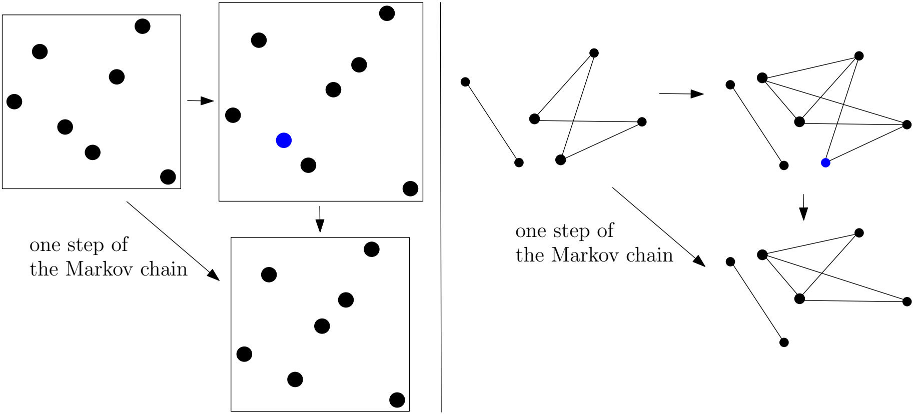


Upstep : duplicate a uniform random element/vertex.

With probability $p \in (0,1)$,

$$\begin{cases} \text{the "twin" elements are in increasing order (permutation case);} \\ \text{the two "twin" vertices are connected with probability } p \text{ (graph case).} \end{cases}$$

An updown Markov chain on permutations/graphs



Downstep: delete a uniform random element/vertex

In this talk: scaling limit (in the sense of permutons or graphons) and its stationary distribution, mixing time (in terms of separation distance).

Motivations

- Inspired from updown chains on partitions, with scaling limit results by Petrov 2009 and 2013, Borodin–Olshanski 2009, Olshanski 2010, ...

Motivations

- Inspired from updown chains on partitions, with scaling limit results by Petrov 2009 and 2013, Borodin–Olshanski 2009, Olshanski 2010, ...
- The duplication operation is standard in combinatorics literature. Permutations/graphs obtained this way are known as separable permutations/cographs.

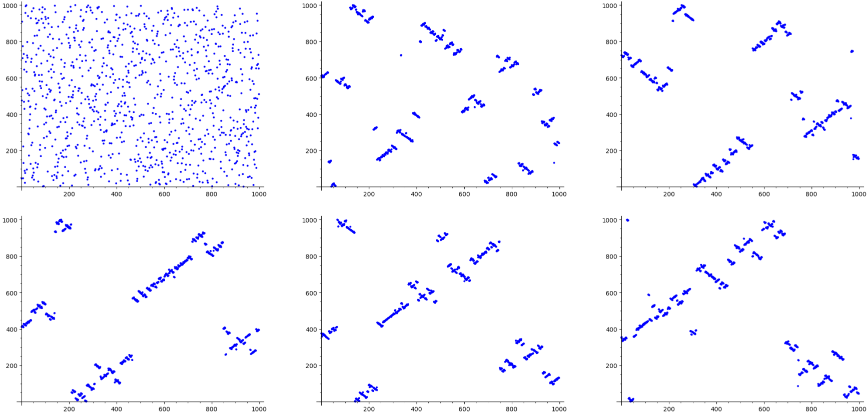
Motivations

- Inspired from updown chains on partitions, with scaling limit results by Petrov 2009 and 2013, Borodin–Olshanski 2009, Olshanski 2010, ...
- The duplication operation is standard in combinatorics literature. Permutations/graphs obtained this way are known as separable permutations/cographs.
- Local edge replacement in graphs leads to a deterministic limiting process in the space of graphons (see Garbe–Hladký–Šileikis–Skerman, 2022); here, we show that local vertex replacement leads to a random diffusion on graphons at the limit.

Motivations

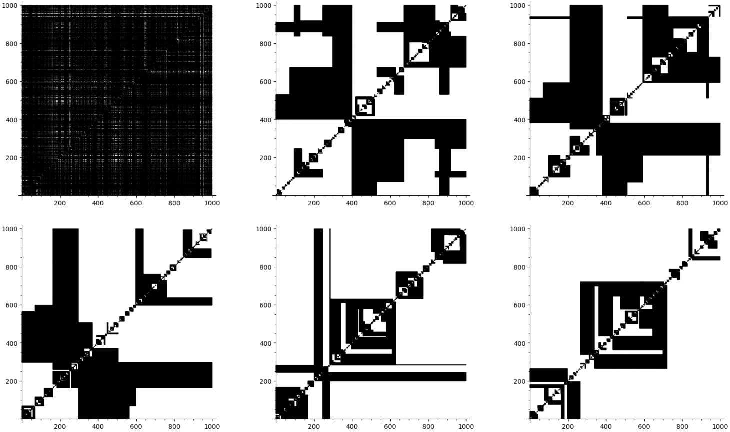
- Inspired from updown chains on partitions, with scaling limit results by Petrov 2009 and 2013, Borodin–Olshanski 2009, Olshanski 2010, ...
- The duplication operation is standard in combinatorics literature. Permutations/graphs obtained this way are known as separable permutations/cographs.
- Local edge replacement in graphs leads to a deterministic limiting process in the space of graphons (see Garbe–Hladký–Šileikis–Skerman, 2022); here, we show that local vertex replacement leads to a random diffusion on graphons at the limit.
- Large literature on mixing time/separation distances of Markov chains on combinatorial objects (related to cutoff). Here, we get exact and asymptotic expressions for the separation distance.

Simulation (permutation case)



Simulation of the up-down chain on permutations. Here, we take $p = 1/2$, $n = 1000$, and we plot the permutation after m steps, where $m \in \{0, 1, 2, 3, 4, 5\} \cdot 50000$.

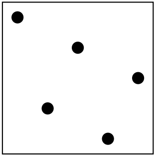
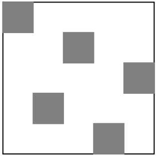
Simulation (graph case)



Simulation of the up-down chain on graphs. Here, we take $p = 1/2$, $n = 1000$, and we plot the adjacency matrix of the graph after m steps, where $m \in \{0,1,2,3,4,5\} \cdot 50000$.

Permutons

A permutation π can be encoded as a probability measure μ_π on $[0,1]^2$.

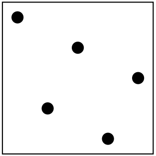
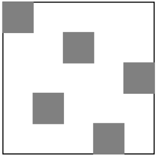
$\pi = 52413 =$  $\mapsto \quad \mu_\pi =$ 

In μ_π, each small square has area $1/n^2$ and weight $1/n$.

We have a natural notion of limit for such objects: the weak convergence.
This defines a compact Polish space $\mathscr{P}$.

Permutons

A permutation π can be encoded as a probability measure μ_π on $[0,1]^2$.

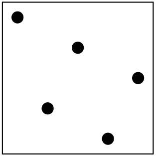
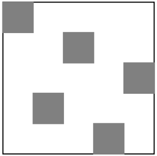
$\pi = 52413 =$  $\mapsto \quad \mu_\pi =$ 

In μ_π, each small square has area $1/n^2$ and weight $1/n$.

Note: the projection on μ_π on each axis is the Lebesgue measure on $[0,1]$ (in other words, μ_π has uniform marginals).
$\rightarrow$ potential limits also have uniform marginals.

Permutons

A permutation π can be encoded as a probability measure μ_π on $[0,1]^2$.

$\pi = 52413 =$  $\mapsto$ $\mu_\pi =$ 

In μ_π, each small square has area $1/n^2$ and weight $1/n$.

Definition

A permuton is a probability measure on $[0,1]^2$ with uniform marginals.

Nice feature: permuton convergence is equivalent to the convergence of substructure densities (here pattern densities);
$\rightarrow$ analogy with the well-developed graphon theory.

Scaling limit (permutation case)

Theorem (F., Rivera-Lopez, '24+)

Let X_n be the above defined Markov chains on permutations of size n, starting at $\sigma_{n,0}$. Assume that $\sigma_{n,0}$ converges to some permuton μ. Then there exists a continuous Feller diffusion $F = F_\mu$ in the space $\mathscr{P}$ of permutons with initial distribution μ such that

$$\big(X_n(\lfloor n^2 t \rfloor)\big)_{t\geq 0} \Longrightarrow \big(F(t)\big)_{t \geq 0},$$

in distribution in the Skorokhod space $D([0,+\infty),\mathscr{P})$.

+ explicit description of the generator on pattern densities.

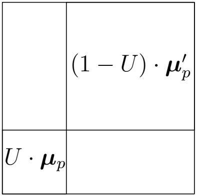
Stationary distribution

Proposition (F., Rivera-Lopez, '24+)

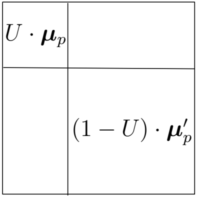
The limiting process F is ergodic and its stationary distribution is the *recursive separable permuton*, *i.e. the unique random permuton $\boldsymbol{\mu}_p$ which satisfies*

$$\boldsymbol{\mu}_p \overset{law}{=} \begin{cases} (U \cdot \boldsymbol{\mu}_p) \oplus ((1-U) \cdot \boldsymbol{\mu}'_p) & \text{with probability } p; \\ (U \cdot \boldsymbol{\mu}_p) \ominus ((1-U) \cdot \boldsymbol{\mu}'_p) & \text{with probability } 1-p. \end{cases},$$

where $\boldsymbol{\mu}'_p$ is an independent copy of $\boldsymbol{\mu}_p$

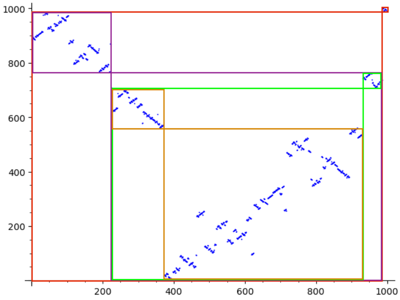


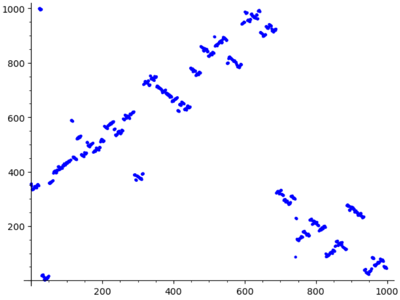
$\oplus$ - case



$\ominus$ - case

Convergence to the stationary distribution - simulation





Left: Simulation of the stationary distribution ($n = 1000$), the colored square emphasizes the recursive structure of the limit.
Right: Simulation of the up-down chain on permutations after 250000 steps ($n = 1000$, $p = 1/2$).

Separation distance

Standard question for Markov chain: how quick does it converge to the stationary distribution?

We use the separation distance (Aldous–Diaconis, '87)

$$\Delta_n(m) = \max_{\substack{x,y \in S_n \\ M_n(y) \neq 0}} 1 - \frac{\mathbb{P}_x\big(X_n(m) = y\big)}{M_n(y)},$$

where M_n is the stationary distribution of the chain on S_n.

Separation distance

Standard question for Markov chain: how quick does it converge to the stationary distribution?

We use the separation distance (Aldous–Diaconis, '87)

$$\Delta_n(m) = \max_{\substack{x,y \in S_n \\ M_n(y) \neq 0}} 1 - \frac{\mathbb{P}_x(X_n(m) = y)}{M_n(y)} = \sup_{x \in S_n, f \in C(S_n, \mathbb{R}_+^*)} \left(1 - \frac{\mathbb{E}_x(f(X_n(m)))}{\int_{S_n} f dM_n}\right),$$

where M_n is the stationary distribution of the chain on S_n.

Continuous analog:

$$\Delta_F(t) = \sup_{\mu \in \mathscr{P}, f \in C(\mathscr{P}, \mathbb{R}_+^*)} 1 - \frac{\mathbb{E}_\mu(f(F_\mu(t)))}{\int_{S_n} f \, d\boldsymbol{\mu}_p}.$$

Asymptotics of the separation distance

Theorem (F., Rivera-Lopez, '24+)

Let $\Delta_n(m)$ be the separation distance for the up-down Markov chain on permutations of size n, and $\Delta_F(t)$ be the one of the limiting process F. We have, for any $t > 0$,

$$\lim \Delta_n(\lfloor n^2 t \rfloor) = \Delta_F(t) = \sum_{j=1}^{+\infty} (-1)^{j-1}(2j+1)e^{-tj(j+1)}.$$

Asymptotics of the separation distance

Theorem (F., Rivera-Lopez, '24+)

Let $\Delta_n(m)$ be the separation distance for the up-down Markov chain on permutations of size n, and $\Delta_F(t)$ be the one of the limiting process F. We have, for any $t > 0$,

$$\lim \Delta_n(\lfloor n^2 t \rfloor) = \Delta_F(t) = \sum_{j=1}^{+\infty} (-1)^{j-1}(2j+1)e^{-tj(j+1)}.$$

Asymptotics:

- as $t \to +\infty$, $\Delta_F(t) \sim 3e^{-2t}$;
- as $t = 0$, not so clear a priori...

Jacobi identity and modular form

A miracle: using an identity of Jacobi

$$\sum_{j=0}^{+\infty} (-1)^j (2j+1) q^{\binom{j}{2}} = \left(\prod_{i=1}^{+\infty} (1-q^i) \right)^3,$$

we can rewrite $\Delta_F(t)$ as

$$\Delta_F(t) = 1 - \prod_{j=1}^{\infty} (1-e^{-2jt})^3.$$

Jacobi identity and modular form

A miracle: using an identity of Jacobi

$$\sum_{j=0}^{+\infty} (-1)^j (2j+1) q^{\binom{j}{2}} = \left(\prod_{i=1}^{+\infty} (1-q^i) \right)^3,$$

we can rewrite $\Delta_F(t)$ as

$$\Delta_F(t) = 1 - \prod_{j=1}^{\infty} (1 - e^{-2jt})^3.$$

But, defining $\eta(\tau) := q^{1/24} \prod_{i=1}^{+\infty} (1-q^i)$ (with $q = e^{2\pi i \tau}$), the function η is a modular form and satisfies $\eta(-1/\tau) = \sqrt{-i\tau}\,\eta(\tau)$.

Jacobi identity and modular form

A miracle: using an identity of Jacobi

$$\sum_{j=0}^{+\infty} (-1)^j (2j+1) q^{\binom{j}{2}} = \left(\prod_{i=1}^{+\infty} (1-q^i) \right)^3,$$

we can rewrite $\Delta_F(t)$ as

$$\Delta_F(t) = 1 - \prod_{j=1}^{\infty} (1-e^{-2jt})^3.$$

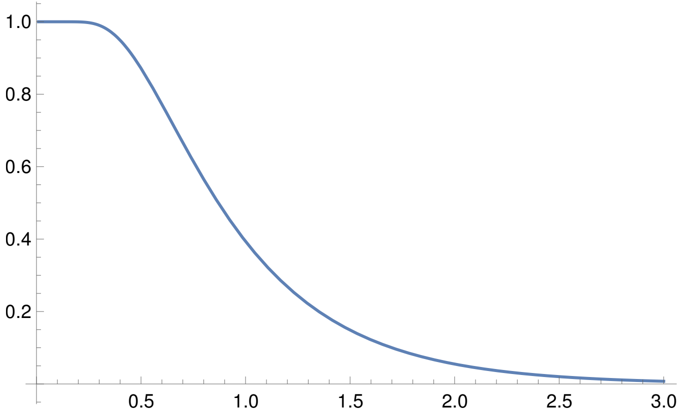
But, defining $\eta(\tau) := q^{1/24} \prod_{i=1}^{+\infty} (1-q^i)$ (with $q = e^{2\pi i \tau}$), the function η is a modular form and satisfies $\eta(-1/\tau) = \sqrt{-i\tau}\eta(\tau)$.

Hence

$$1 - \Delta_F(t) = \exp\left(-\frac{\pi^2}{4t} + \frac{t}{4}\right) \left(\frac{\pi}{t}\right)^{3/2} \left(1 - \Delta_F\left(\frac{\pi^2}{t}\right)\right)$$

and for small t, we have $\Delta_F(t) = 1 - \exp\left(-\frac{\pi^2}{4t}\right) \left(\frac{\pi}{t}\right)^{3/2} \left(1 + O(e^{-2\pi^2/t})\right)$.

Plot of the limiting separation distance



Plot of the function $\Delta_F(t)$. The Markov chain does not exhibit a separation cutoff (which would correspond to $\Delta_F(t) = \mathbf{1}[t \le t_0]$), but the curve is very flat near $t = 0$ and $t = +\infty$.

Transition

Some proof elements

Key identity: the commutation relation

- Let $p_n^{\uparrow} \in \mathcal{M}(S_n \times S_{n+1})$ be the up transition matrix, i.e. $p_n^{\uparrow}(\tau,\sigma)$ is the probability to find σ when duplicating a uniform random point in τ.
- Let $p_{n+1}^{\downarrow} \in \mathcal{M}(S_{n+1} \times S_n)$ be the down transition matrix, i.e. $p_{n+1}^{\downarrow}(\sigma,\tau)$ is the probability to find τ when deleting a uniform random point in σ.

Proposition

For any $n \geq 2$, we have

$$p_n^{\uparrow} p_{n+1}^{\downarrow} = \frac{n-1}{n+1}\, p_n^{\downarrow} p_{n-1}^{\uparrow} + \frac{2}{n+1}\, \mathrm{Id}_{S_n},$$

Our results (scaling limit and computation of the separation distance) hold generally for up-down chains satisfying this kind of commutation relation.

Key identity: the commutation relation

- Let $p_n^{\uparrow} \in \mathcal{M}(S_n \times S_{n+1})$ be the up transition matrix, i.e. $p_n^{\uparrow}(\tau,\sigma)$ is the probability to find σ when duplicating a uniform random point in τ.
- Let $p_{n+1}^{\downarrow} \in \mathcal{M}(S_{n+1} \times S_n)$ be the down transition matrix, i.e. $p_{n+1}^{\downarrow}(\sigma,\tau)$ is the probability to find τ when deleting a uniform random point in σ.

Proposition

For any $n \geq 2$, we have

$$p_n^{\uparrow} p_{n+1}^{\downarrow} = \frac{n-1}{n+1}\, p_n^{\downarrow} p_{n-1}^{\uparrow} + \frac{2}{n+1}\, \mathrm{Id}_{S_n},$$

Corollary (follows from Fulman, '09)

The transition matrix $p_n = p_n^{\uparrow} p_{n+1}^{\downarrow}$ of the up-down chain has eigenvalue $1 - \frac{i(i-1)}{n(n+1)}$, with multiplicity $|S_i| - |S_{i-1}|$.
(for $1 \leq i \leq n$, with the convention $|S_0| = 0$.)

Density functions and right eigenvectors of p_n

For τ in S_k and σ in S_n, with $k \leq n$

$$d_\tau(\sigma) = (p_n^{\downarrow} \dots p_{k+1}^{\downarrow})(\sigma, \tau).$$

In words, $d_\tau(\sigma)$ is the probability to obtain τ when deleting $n-k$ uniform random elements in σ, or the "proportion of τ" in σ.

Density functions and right eigenvectors of p_n

For τ in S_k and σ in S_n, with $k \leq n$

$$d_\tau(\sigma) = (p_n^{\downarrow} \dots p_{k+1}^{\downarrow})(\sigma, \tau).$$

In words, $d_\tau(\sigma)$ is the probability to obtain τ when deleting $n-k$ uniform random elements in σ, or the "proportion of τ" in σ.

$\rightarrow$ d_τ is a central functional in permuton theory; it can be extended to a continuous function on $\mathscr{P}$ and $\mathrm{Span}(d_\tau)$ is dense subalgebra of $C(\mathscr{P})$.

Density functions and right eigenvectors of p_n

For τ in S_k and σ in S_n, with $k \le n$

$$d_\tau(\sigma) = (p_n^\downarrow \dots p_{k+1}^\downarrow)(\sigma, \tau).$$

In words, $d_\tau(\sigma)$ is the probability to obtain τ when deleting $n-k$ uniform random elements in σ, or the "proportion of τ" in σ.

$\to$ d_τ is a central functional in permuton theory; it can be extended to a continuous function on $\mathscr{P}$ and $\mathrm{Span}(d_\tau)$ is dense subalgebra of $C(\mathscr{P})$.

Proposition (F., Rivera-Lopez, '24+)

Define, for τ in S_k,

$$h_\tau = \sum_{j \le k} \left[\left(\prod_{i=j}^{k-1} \frac{(-1)^{j-i}\, i(i+1)}{k(k-1) - i(i-1)} \right) \sum_{\pi \in S_j} (p_j^\uparrow \dots p_{k-1}^\uparrow)(\pi, \tau)\, d_\pi \right].$$

Then, seeing h_τ as a vector in $\mathbb{C}^{S_n}$

$$p_n h_\tau = \left(1 - \frac{k(k-1)}{n(n+1)}\right) h_\tau.$$

Scaling limit

Recall that $p_n h_\tau = \left(1 - \frac{k(k-1)}{n(n+1)}\right) h_\tau$, for $\tau \in S_k$.

Hence $p_n^{\lfloor tn^2 \rfloor} h_\tau = e^{-tk(k-1)}(1 + o(1)) h_\tau$.

Scaling limit

Recall that $p_n h_\tau = \left(1 - \frac{k(k-1)}{n(n+1)}\right) h_\tau$, for $\tau \in S_k$.

Hence $p_n^{\lfloor tn^2 \rfloor} h_\tau = e^{-tk(k-1)}(1+o(1)) h_\tau$.

Since $\mathrm{Span}(h_\tau) = \mathrm{Span}(d_\tau)$ is dense in $C(\mathscr{P})$, this implies (see, e.g., Ethier–Kurtz '05) that

$$X_n(\lfloor tn^2 \rfloor) \to F(t),$$

where $F(t)$ has a transition semi-group $\mathscr{T}(t)$ defined by

$$\mathscr{T}(t) h_\tau := e^{-tk(k-1)} h_\tau, \text{ for } \tau \in S_k.$$

Scaling limit

Recall that $p_n h_\tau = \left(1 - \frac{k(k-1)}{n(n+1)}\right) h_\tau$, for $\tau \in S_k$.

Hence $p_n^{\lfloor tn^2 \rfloor} h_\tau = e^{-tk(k-1)}(1 + o(1)) h_\tau$.

Since $\mathrm{Span}(h_\tau) = \mathrm{Span}(d_\tau)$ is dense in $C(\mathscr{P})$, this implies (see, e.g., Ethier–Kurtz '05) that

$$X_n(\lfloor tn^2 \rfloor) \to F(t),$$

where $F(t)$ has a transition semi-group $\mathscr{T}(t)$ defined by

$$\mathscr{T}(t) h_\tau := e^{-tk(k-1)} h_\tau, \text{ for } \tau \in S_k.$$

Generator A with domain $\mathrm{Span}(d_\tau)$: for τ in S_k,

$$A h_\tau = -k(k-1) h_\tau;$$

$$A d_\tau = -k(k-1)\Big(d_\tau - \sum_{\pi \in S_{k-1}} d_\pi \, p_{k-1}^{\uparrow}(\pi, \tau)\Big).$$

And the separation distance? (inspired from Fulman '07,'09)

We want to compute $\Delta_n(m) = \max_{x,y \in S_n} 1 - \frac{\mathbb{P}_x\left(X_n(m)=y\right)}{M_n(y)}$.

And the separation distance? (inspired from Fulman '07,'09)

We want to compute $\Delta_n(m) = \max_{x,y \in S_n} 1 - \frac{\mathbb{P}_x\big(X_n(m)=y\big)}{M_n(y)}$.

Proposition

$$\Delta_n(m) = \sum_{i=1}^{n-1} \lambda_i^m \prod_{j \neq i} \frac{1-\lambda_j}{\lambda_i - \lambda_j},$$

where $\lambda_1, \dots, \lambda_{n-1}$ are the non-trivial distinct eigenvalues of p_n.

And the separation distance? (inspired from Fulman '07,'09)

We want to compute $\Delta_n(m) = \max_{x,y \in S_n} 1 - \frac{\mathbb{P}_x\big(X_n(m)=y\big)}{M_n(y)}$.

Proposition

$$\Delta_n(m) = \sum_{i=1}^{n-1} \lambda_i^m \prod_{j \neq i} \frac{1-\lambda_j}{\lambda_i - \lambda_j},$$

where $\lambda_1, \dots, \lambda_{n-1}$ are the non-trivial distinct eigenvalues of p_n.

Key point: there are elements $x_0 = \mathrm{id}$ and $y_0 = n(n-1)\cdots 1$, whose distance in the chain is exactly the number of distinct eigenvalues of p_n, minus one.

- first step: prove that the maximum is reached for x_0 and y_0 (use the commutation relation).
- second step: Compute $\mathbb{P}_x\big(X_n(m)=y\big) = p_n^m(x_0,y_0)$ (next slide).

And the separation distance? (inspired from Fulman '07,'09)

Set $\lambda_0 = 1$ ($\lambda_1, \dots, \lambda_{n-1}$ are the other eigenvalues of p_n). The polynomials

$$Z^m \text{ and } \sum_{i=0}^{n-1} \lambda_i^m \prod_{j \neq i} \frac{Z - \lambda_j}{\lambda_i - \lambda_j}$$

coincide on $\{\lambda_0, \dots \lambda_{n-1}\}$. Hence, since p_n is diagonalizable,

$$p_n^m = \sum_{i=0}^{n-1} \lambda_i^m \prod_{j \neq i} \frac{p_n - \lambda_j}{\lambda_i - \lambda_j}.$$

And the separation distance? (inspired from Fulman '07,'09)

Set $\lambda_0 = 1$ ($\lambda_1, \dots, \lambda_{n-1}$ are the other eigenvalues of p_n). The polynomials

$$Z^m \text{ and } \sum_{i=0}^{n-1} \lambda_i^m \prod_{j \neq i} \frac{Z - \lambda_j}{\lambda_i - \lambda_j}$$

coincide on $\{\lambda_0, \dots \lambda_{n-1}\}$. Hence, since p_n is diagonalizable,

$$p_n^m = \sum_{i=0}^{n-1} \lambda_i^m \prod_{j \neq i} \frac{p_n - \lambda_j}{\lambda_i - \lambda_j}.$$

But $p_n^k(x_0, y_0) = 0$ for $k < n-1$, thus

$$p_n^m(x_0, y_0) = \left(\sum_{i=0}^{n-1} \lambda_i^m \prod_{j \neq i} \frac{1}{\lambda_i - \lambda_j} \right) p_n^{n-1}(x_0, y_0).$$

And the separation distance? (inspired from Fulman '07,'09)

Set $\lambda_0 = 1$ ($\lambda_1, \dots, \lambda_{n-1}$ are the other eigenvalues of p_n). The polynomials

$$Z^m \text{ and } \sum_{i=0}^{n-1} \lambda_i^m \prod_{j \neq i} \frac{Z - \lambda_j}{\lambda_i - \lambda_j}$$

coincide on $\{\lambda_0, \dots \lambda_{n-1}\}$. Hence, since p_n is diagonalizable,

$$p_n^m = \sum_{i=0}^{n-1} \lambda_i^m \prod_{j \neq i} \frac{p_n - \lambda_j}{\lambda_i - \lambda_j}.$$

But $p_n^k(x_0, y_0) = 0$ for $k < n-1$, thus

$$p_n^m(x_0, y_0) = \left(\sum_{i=0}^{n-1} \lambda_i^m \prod_{j \neq i} \frac{1}{\lambda_i - \lambda_j} \right) p_n^{n-1}(x_0, y_0).$$

Taking m to $+\infty$ gives

$$M_n(y_0) = \left(\prod_{j \neq 0} \frac{1}{1 - \lambda_j} \right) p_n^{n-1}(x_0, y_0).$$

Conclusion:

$$\frac{p_n^m(x_0, y_0)}{M_n(y_0)} = 1 + \sum_{i=1}^{n-1} \lambda_i^m \prod_{j \neq i} \frac{1 - \lambda_j}{\lambda_i - \lambda_j}.$$

And the separation distance? (inspired from Fulman '07,'09)

Since $\lambda_i = 1 - \frac{i(i+1)}{n(n+1)}$ is explicit, we obtain an explicit expression for $\Delta_n(m)$:

$$\Delta_n(m) = \sum_{j=1}^{n-1} (-1)^{j-1}(2j+1)\frac{(n-1)!n!}{(n-1-j)!(n+j)!}\left(1 - \frac{j(j+1)}{n(n+1)}\right)^m.$$

And the separation distance? (inspired from Fulman '07,'09)

Since $\lambda_i = 1 - \frac{i(i+1)}{n(n+1)}$ is explicit, we obtain an explicit expression for $\Delta_n(m)$:

$$\Delta_n(m) = \sum_{j=1}^{n-1} (-1)^{j-1}(2j+1)\frac{(n-1)!n!}{(n-1-j)!(n+j)!}\left(1 - \frac{j(j+1)}{n(n+1)}\right)^m.$$

Finding the asymptotics for large n and $m = \lfloor tn^2 \rfloor$ is straightforward:

$$\lim_{n \to +\infty} \Delta_n(\lfloor tn^2 \rfloor) = \sum_{j=1}^{+\infty} (-1)^{j-1}(2j+1)e^{tj(j+1)}.$$

Thank you for your attention