Patterns in Mallows permutations

Victor Dubach
under the supervision of Valentin Féray

IECL, Université de Lorraine

Workshop L^2, 20/09/2024

Outline

I. Background

II. Regenerative regime (strong bias)

III. Almost regenerative regime (moderately strong bias)

IV. Almost uniform regime (weak bias)

I. Background: Mallows-like models

Context:
A "judge" is asked to rank a list of objects, numbered from 1 to n.
The output is a (random) permutation $\tau = \tau(1)\,\tau(2)\ldots\tau(n)$.

Suppose that there exists "true" ordering σ_0 of the objects, and that τ tends to be close to it.

I. Background: Mallows-like models

Context:
A "judge" is asked to rank a list of objects, numbered from 1 to n.
The output is a (random) permutation $\tau = \tau(1)\,\tau(2)\ldots\tau(n)$.

Suppose that there exists "true" ordering σ_0 of the objects, and that τ tends to be close to it.

Definition: Mallows-like models

Let d be a distance on the symmetric group, and let $\sigma_0 \in \mathfrak{S}_n$.
A random permutation τ_n follows the Mallows-like distribution with parameters $n \in \mathbb{N}^*$ and $q \in [0, 1]$ with respect to the distance d if:

$$\forall \sigma \in \mathfrak{S}_n, \quad \mathrm{P}\left(\tau_n = \sigma\right) = \frac{1}{Z} q^{d(\sigma, \sigma_0)}$$

where $Z > 0$ is a normalizing constant.

I. Background: The Mallows model

Definition: inversions

An inversion of $\sigma \in \mathfrak{S}_n$ is a pair $i < j$ such that $\sigma(i) > \sigma(j)$.
The total number of inversions of σ is denoted by $\mathrm{inv}\,(\sigma)$.

Let $d(\sigma, \sigma_0) := \mathrm{inv}\left(\sigma\sigma_0^{-1}\right)$ (Kendall's distance), and let $\sigma_0 := \mathrm{id}_n$.

I. Background: The Mallows model

Definition: inversions

An inversion of $\sigma \in \mathfrak{S}_n$ is a pair $i < j$ such that $\sigma(i) > \sigma(j)$.
The total number of inversions of σ is denoted by $\mathrm{inv}(\sigma)$.

Let $d(\sigma, \sigma_0) := \mathrm{inv}\left(\sigma\sigma_0^{-1}\right)$ (Kendall's distance), and let $\sigma_0 := \mathrm{id}_n$.

Definition: the Mallows model

$\tau_n \sim \mathrm{Mallows}(n, q)$ if:

$$\forall \sigma \in \mathfrak{S}_n, \quad \mathrm{P}(\tau_n = \sigma) = \frac{1}{Z_{n,q}} q^{\mathrm{inv}(\sigma)}$$

where $Z_{n,q} = \prod_{k=1}^{n} \frac{1-q^k}{1-q}$ is a normalizing constant.

Intuition: for each pair $i < j$ of objects, the judge has a probability $\theta \in [0, 1]$ of ranking them correctly, with $\frac{1-\theta}{\theta} = q$.

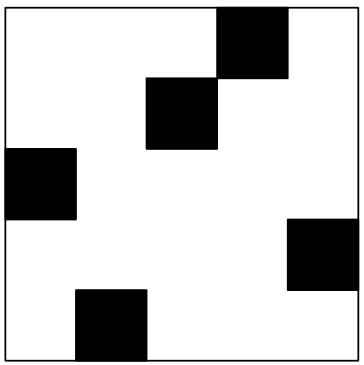
I. Background: Large Mallows permutations

Goal: asymptotic properties of $\tau_n \sim \mathrm{Mallows}\,(n, q_n)$ as $n \to \infty$.

I. Background: Large Mallows permutations

Goal: asymptotic properties of $\tau_n \sim \mathrm{Mallows}\,(n, q_n)$ as $n \to \infty$.

Let's see what they look like.

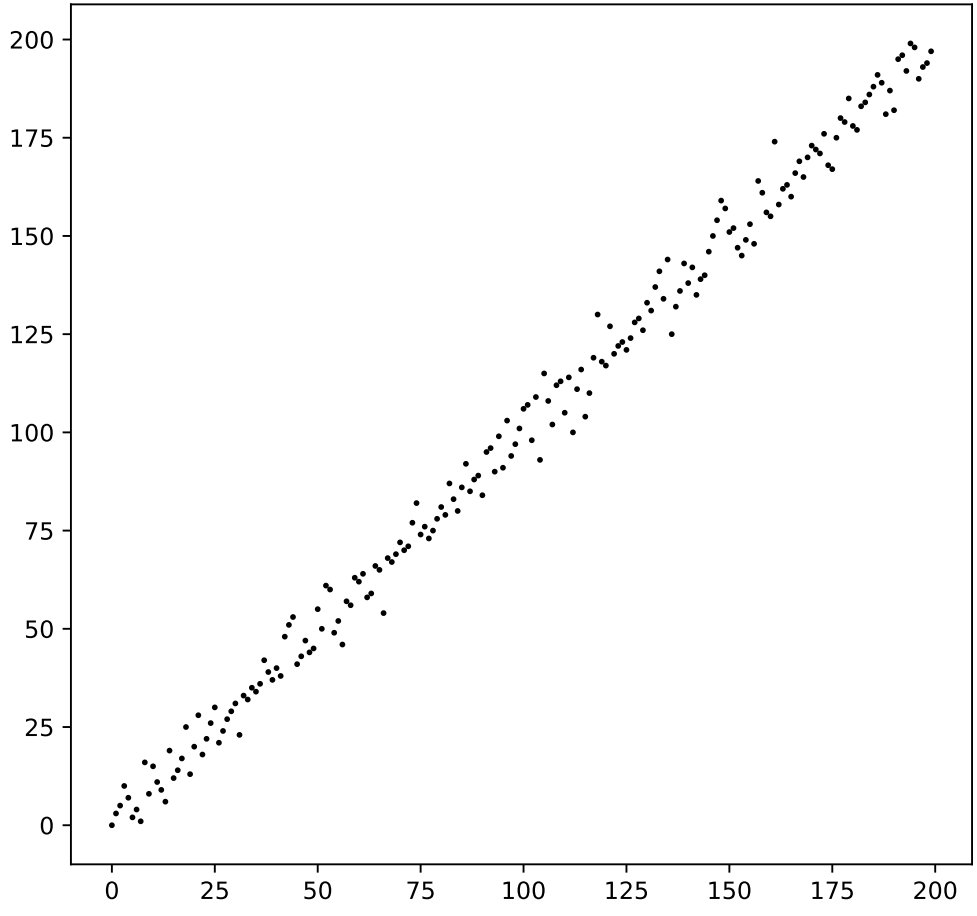
$\sigma = 3\,1\,4\,5\,2$ 

I. Background: Large Mallows permutations

Goal: asymptotic properties of $\tau_n \sim \mathrm{Mallows}\,(n, q_n)$ as $n \to \infty$.

Let's see what they look like.

$q_n = q$ fixed ($= 0.7$ here).



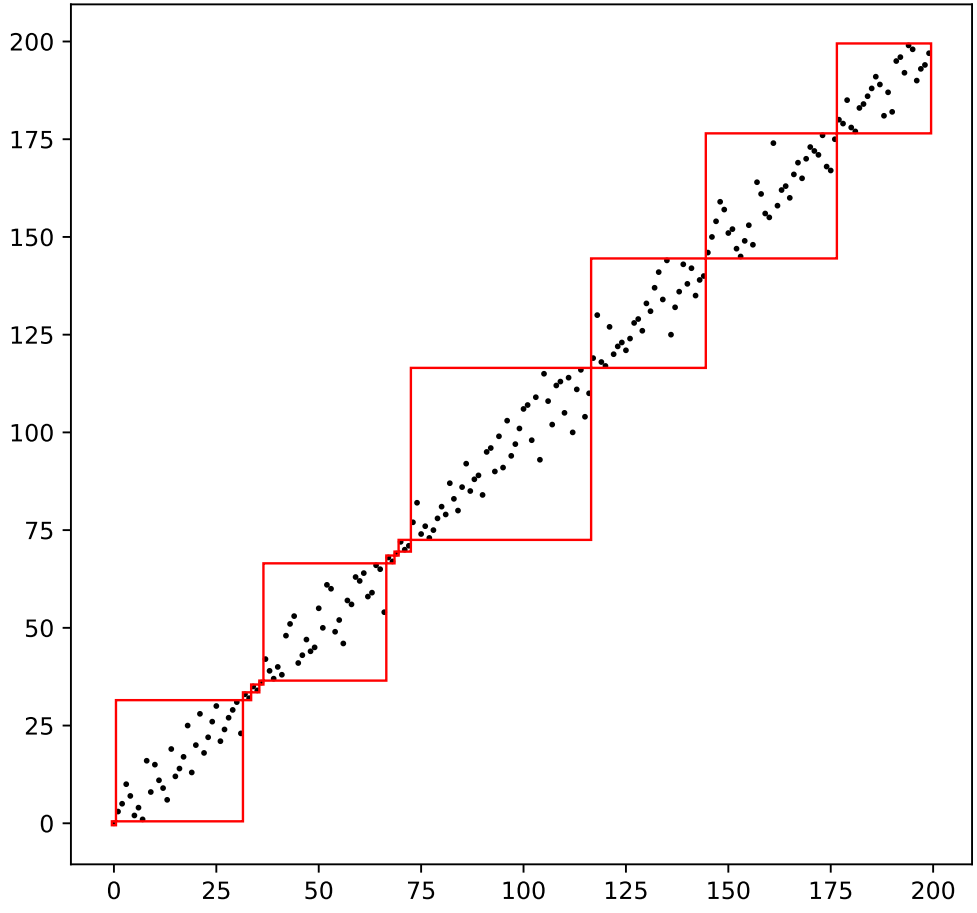
I. Background: Large Mallows permutations

Goal: asymptotic properties of $\tau_n \sim \mathrm{Mallows}\,(n, q_n)$ as $n \to \infty$.

Let's see what they look like.

$q_n = q$ fixed ($= 0.7$ here).

Regenerative property: the permutation decomposes into a sequence of smaller sub-permutations.



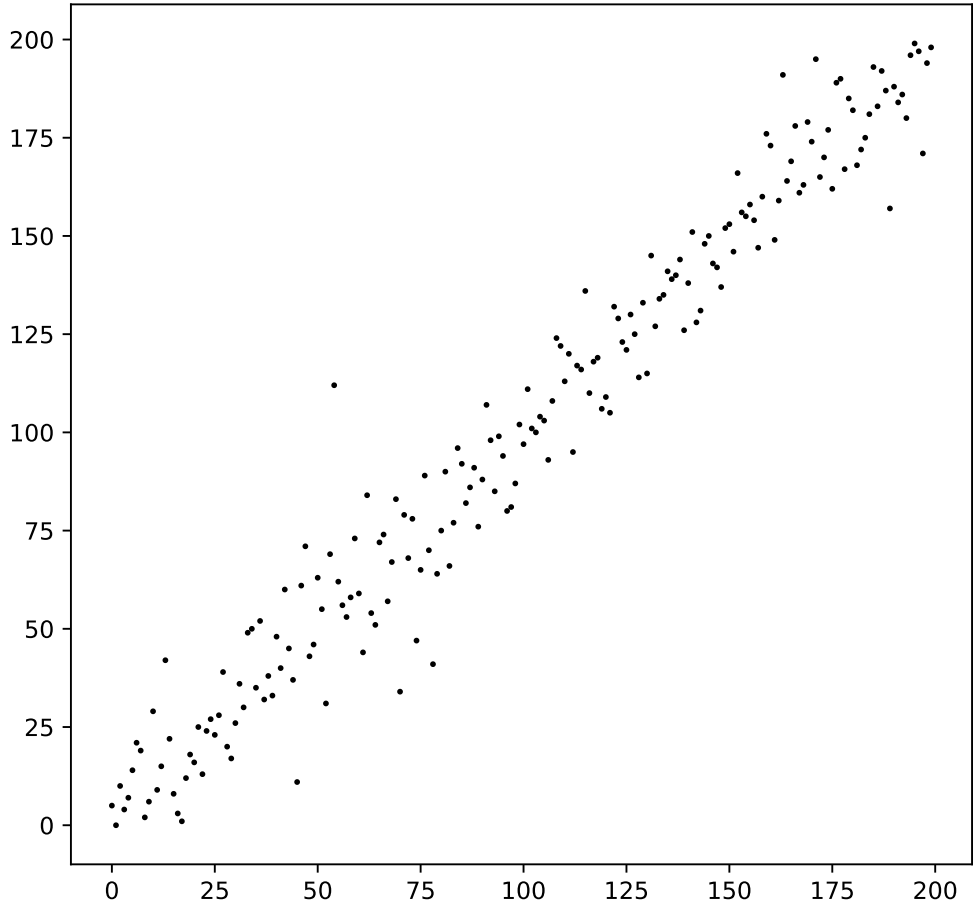
I. Background: Large Mallows permutations

Goal: asymptotic properties of $\tau_n \sim \mathrm{Mallows}\,(n, q_n)$ as $n \to \infty$.

Let's see what they look like.

$q_n \to 1$ and $n(1-q_n) \to \infty$

($q_n = 1 - 2/\sqrt{n}$ here).



I. Background: Large Mallows permutations

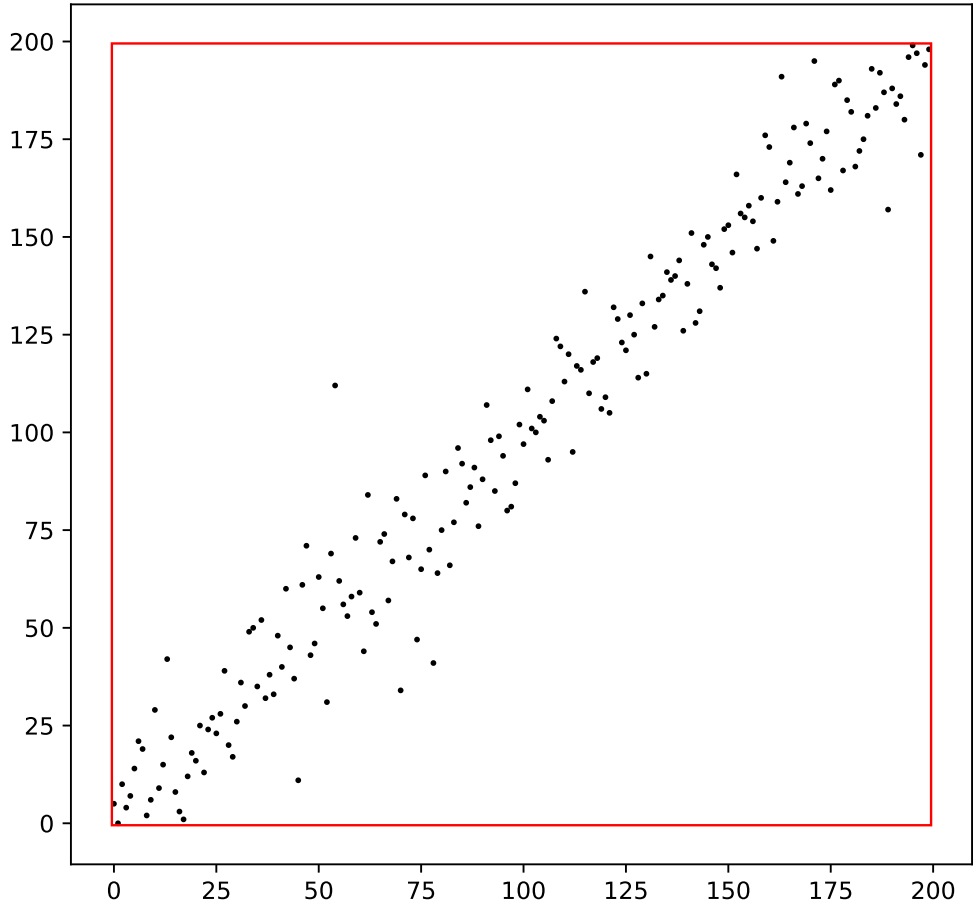
Goal: asymptotic properties of $\tau_n \sim \mathrm{Mallows}(n, q_n)$ as $n \to \infty$.

Let's see what they look like.

$q_n \to 1$ and $n(1-q_n) \to \infty$

($q_n = 1 - 2/\sqrt{n}$ here).

The permutation is not regenerative anymore...



I. Background: Large Mallows permutations

Goal: asymptotic properties of $\tau_n \sim \mathrm{Mallows}(n, q_n)$ as $n \to \infty$.

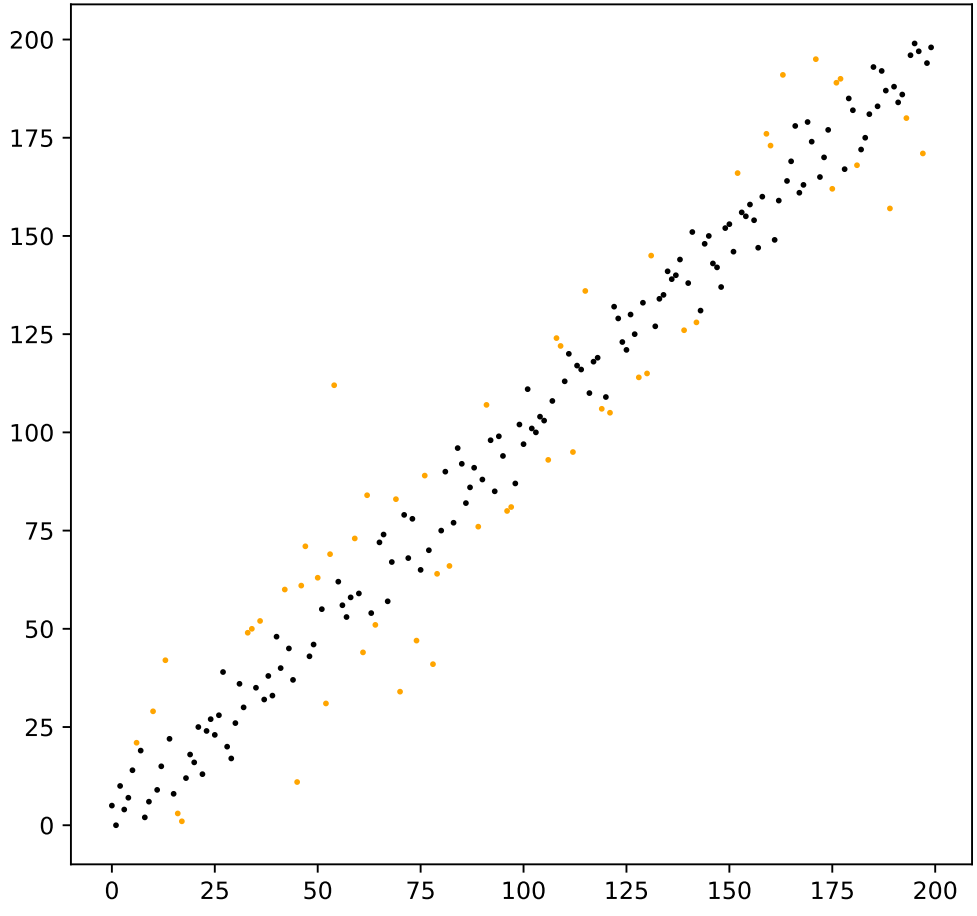
Let's see what they look like.

$q_n \to 1$ and $n(1-q_n) \to \infty$

($q_n = 1 - 2/\sqrt{n}$ here).

The permutation is not regenerative anymore...

... but upon removing a certain fraction of points:



I. Background: Large Mallows permutations

Goal: asymptotic properties of $\tau_n \sim \mathrm{Mallows}(n, q_n)$ as $n \to \infty$.

Let's see what they look like.

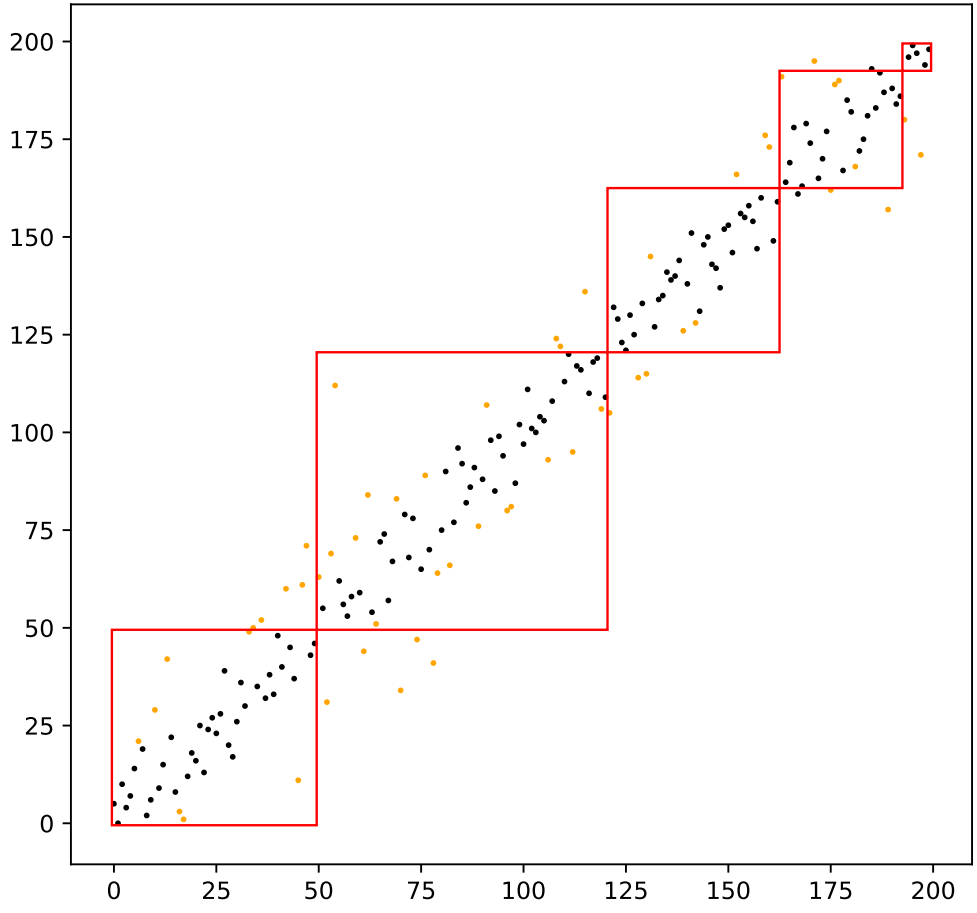
$q_n \to 1$ and $n(1-q_n) \to \infty$

($q_n = 1 - 2/\sqrt{n}$ here).

The permutation is not regenerative anymore...

... but upon removing a certain fraction of points:

we can recover this property!



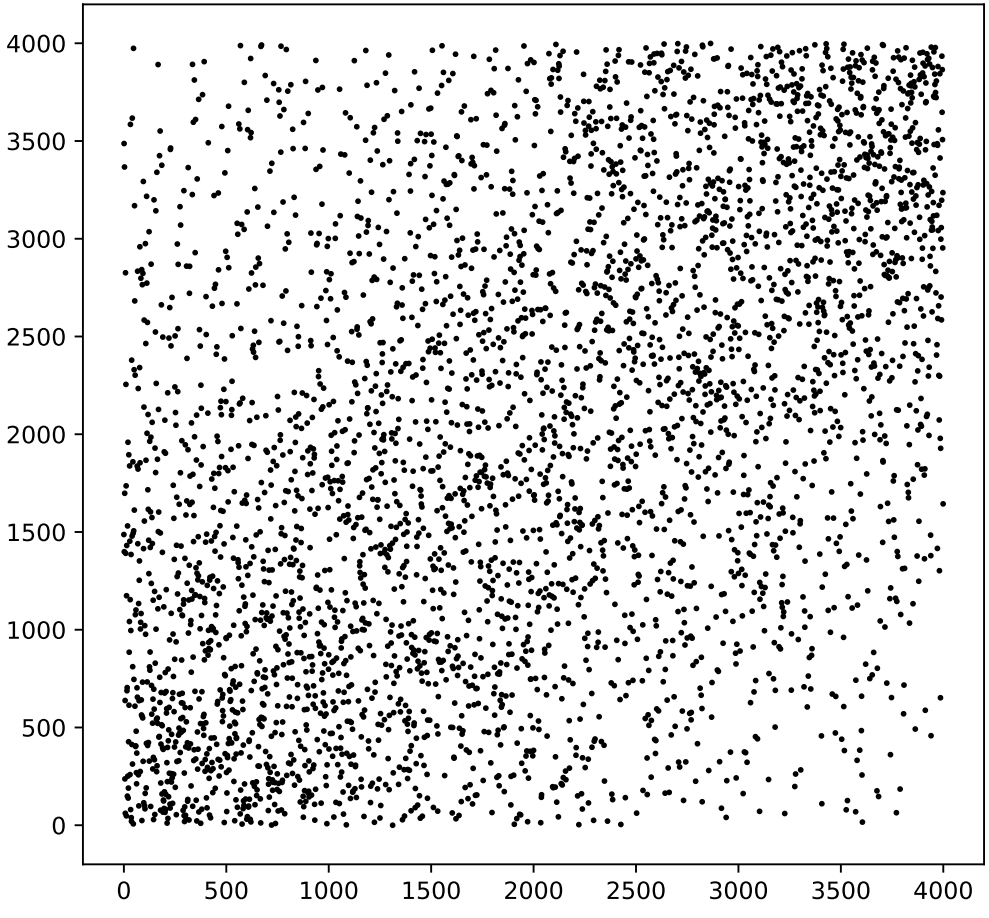
I. Background: Large Mallows permutations

Goal: asymptotic properties of $\tau_n \sim \mathrm{Mallows}\,(n, q_n)$ as $n \to \infty$.

Let's see what they look like.

$q_n = 1 - \frac{\beta}{n}$

"Thermodynamic regime"



I. Background: Large Mallows permutations

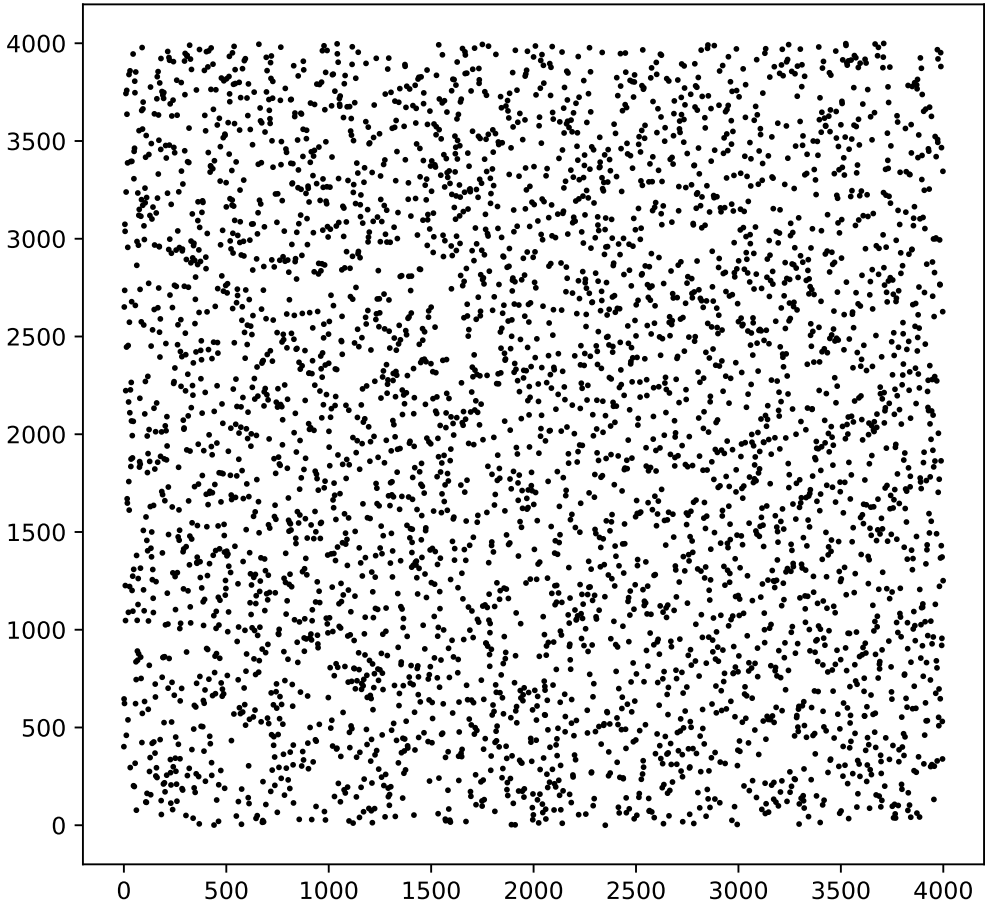
Goal: asymptotic properties of $\tau_n \sim \mathrm{Mallows}\,(n, q_n)$ as $n \to \infty$.

Let's see what they look like.

$n(1 - q_n) \to 0$

($q_n = 1 - 3/n^2$ here).

Almost indistinguishable from uniformity.

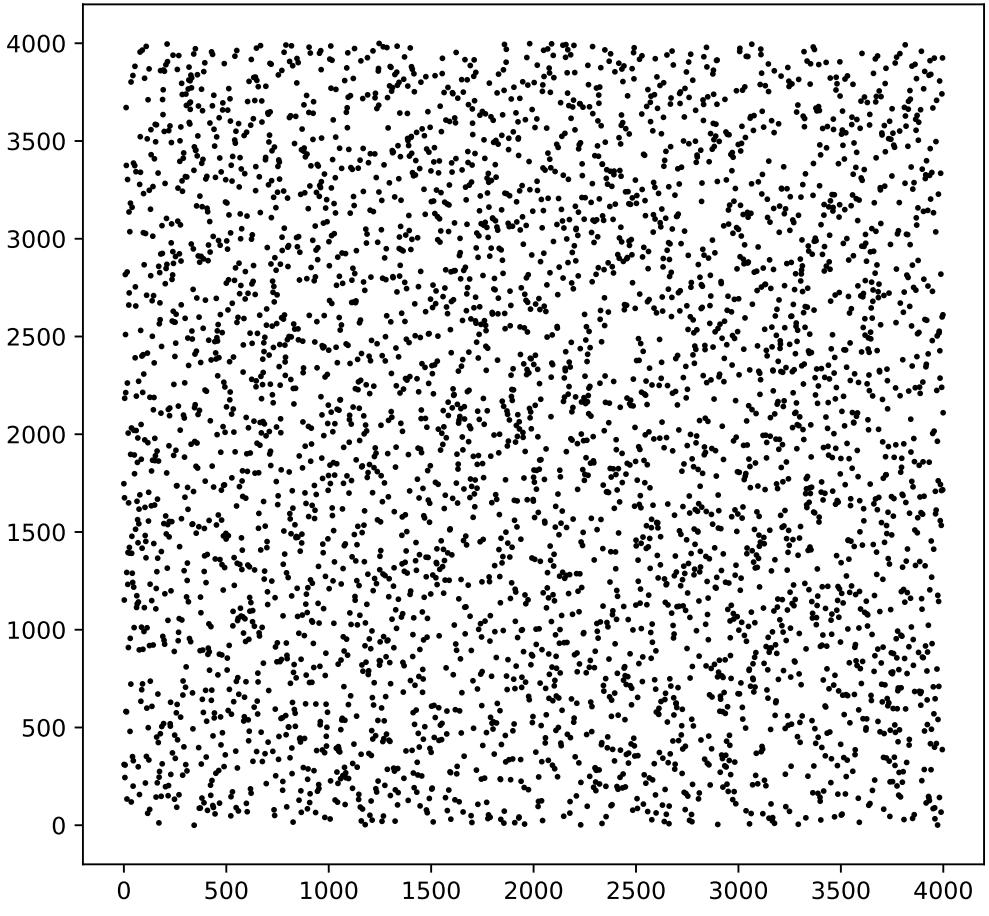


I. Background: Large Mallows permutations

Goal: asymptotic properties of $\tau_n \sim \mathrm{Mallows}\,(n, q_n)$ as $n \to \infty$.

Let's see what they look like.

$q_n = 1$



I. Background: Pattern occurrences

Definition: patterns

Let $\tau = \tau(1)\ \tau(2) \dots \tau(n) \in \mathfrak{S}_n$. The *pattern* induced by τ on $I = (i_1 < \cdots < i_r)$ is the permutation $\mathrm{pat}\,(I, \tau) = \pi \in \mathfrak{S}_r$ with:

$$\forall j, k \in [r], \quad \pi(j) < \pi(k) \Longleftrightarrow \tau(i_j) < \tau(i_k).$$

I. Background: Pattern occurrences

Definition: patterns

Let $\tau = \tau(1)\ \tau(2) \dots \tau(n) \in \mathfrak{S}_n$. The *pattern* induced by τ on $I = (i_1 < \cdots < i_r)$ is the permutation $\mathrm{pat}\,(I, \tau) = \pi \in \mathfrak{S}_r$ with:

$$\forall j, k \in [r], \quad \pi(j) < \pi(k) \Longleftrightarrow \tau(i_j) < \tau(i_k).$$

$\tau =$ 4 2 7 6 3 1 5

I. Background: Pattern occurrences

Definition: patterns

Let $\tau = \tau(1)\ \tau(2) \dots \tau(n) \in \mathfrak{S}_n$. The *pattern* induced by τ on $I = (i_1 < \cdots < i_r)$ is the permutation $\mathrm{pat}\,(I, \tau) = \pi \in \mathfrak{S}_r$ with:

$$\forall j, k \in [r], \quad \pi(j) < \pi(k) \Longleftrightarrow \tau(i_j) < \tau(i_k).$$

$$\tau = \underline{4}\,2\,7\,\underline{6}\,3\,\underline{1}\,5$$

I. Background: Pattern occurrences

Definition: patterns

Let $\tau = \tau(1)\ \tau(2) \dots \tau(n) \in \mathfrak{S}_n$. The *pattern* induced by τ on $I = (i_1 < \cdots < i_r)$ is the permutation $\mathrm{pat}\,(I, \tau) = \pi \in \mathfrak{S}_r$ with:

$$\forall j, k \in [r], \quad \pi(j) < \pi(k) \Longleftrightarrow \tau(i_j) < \tau(i_k).$$

$$\tau = \underline{4}\,2\,7\,\underline{6}\,3\,\underline{1}\,5 \quad \Rightarrow \quad \mathrm{pat}\,(I, \tau) = 2\,3\,1$$

I. Background: Pattern occurrences

Definition: patterns

Let $\tau = \tau(1)\ \tau(2) \dots \tau(n) \in \mathfrak{S}_n$. The *pattern* induced by τ on $I = (i_1 < \cdots < i_r)$ is the permutation $\mathrm{pat}\,(I, \tau) = \pi \in \mathfrak{S}_r$ with:

$$\forall j, k \in [r], \quad \pi(j) < \pi(k) \Longleftrightarrow \tau(i_j) < \tau(i_k).$$

$$\tau = \underline{4}\,2\,7\,\underline{6}\,3\,\underline{1}\,5 \quad \Rightarrow \quad \mathrm{pat}\,(I, \tau) = 2\,3\,1$$

Definition: pattern counts

The number of (classical) occurrences of $\pi \in \mathfrak{S}_r$ in $\tau \in \mathfrak{S}_n$ is:

$$\mathrm{Occ}\,(\pi, \tau) := \sum_{I \in \binom{[n]}{r}} 1_{\mathrm{pat}(I,\tau)=\pi}$$

Example: $\mathrm{Occ}\,(2\,1, \tau) = \mathrm{inv}\,(\tau)$.

I. Background: Pattern occurrences

Definition: patterns

Let $\tau = \tau(1)\ \tau(2) \dots \tau(n) \in \mathfrak{S}_n$. The *pattern* induced by τ on $I = (i_1 < \cdots < i_r)$ is the permutation $\mathrm{pat}\,(I, \tau) = \pi \in \mathfrak{S}_r$ with:

$$\forall j, k \in [r], \quad \pi(j) < \pi(k) \Longleftrightarrow \tau(i_j) < \tau(i_k).$$

$$\tau = \underline{4}\,2\,7\,\underline{6}\,3\,\underline{1}\,5 \quad \Rightarrow \quad \mathrm{pat}\,(I, \tau) = 2\,3\,1$$

Definition: pattern counts

The number of (classical) occurrences of $\pi \in \mathfrak{S}_r$ in $\tau \in \mathfrak{S}_n$ is:

$$\mathrm{Occ}\,(\pi, \tau) := \sum_{I \in \binom{[n]}{r}} 1_{\mathrm{pat}(I,\tau)=\pi}$$

Applications: independence testing and "quasi-randomness"; property and parameter testing...

Question

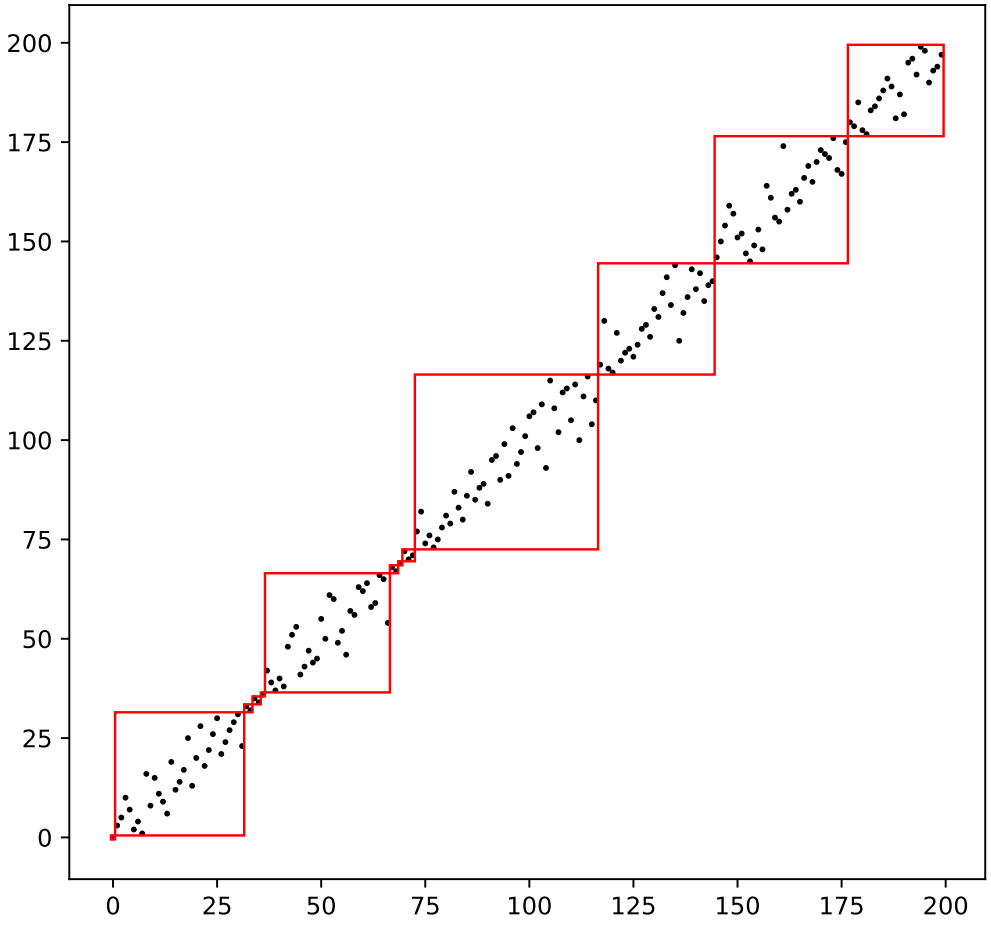
What are the first- and second-order asymptotics of

$$\mathrm{Occ}\left(\pi, \tau_n\right)$$

as $n \to \infty$, for fixed $\pi \in \mathfrak{S}_r$ and $\tau_n \sim \mathrm{Mallows}\left(n, q_n\right)$?

II. Strong bias

$\tau_n \sim \mathrm{Mallows}\,(n, q_n)$ with $q_n = q \in (0, 1)$ fixed.

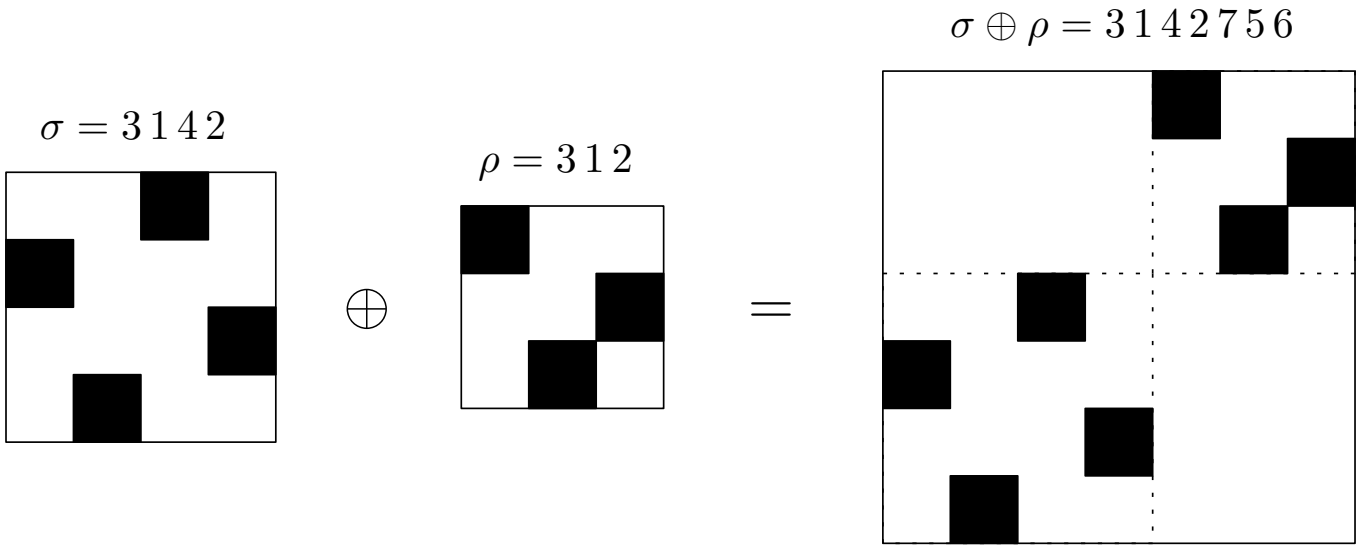


II. Strong bias: decomposing permutations

Definition: direct sum

If $\sigma \in \mathfrak{S}_\ell$ and $\rho \in \mathfrak{S}_m$, define $\sigma \oplus \rho \in \mathfrak{S}_{\ell+m}$ by:

$$\sigma \oplus \rho = \sigma(1)\ \ \sigma(2)\ \ \ldots\ \ \sigma(\ell)\ \ \rho(1)+\ell\ \ \rho(2)+\ell\ \ \ldots\ \ \rho(m)+\ell.$$



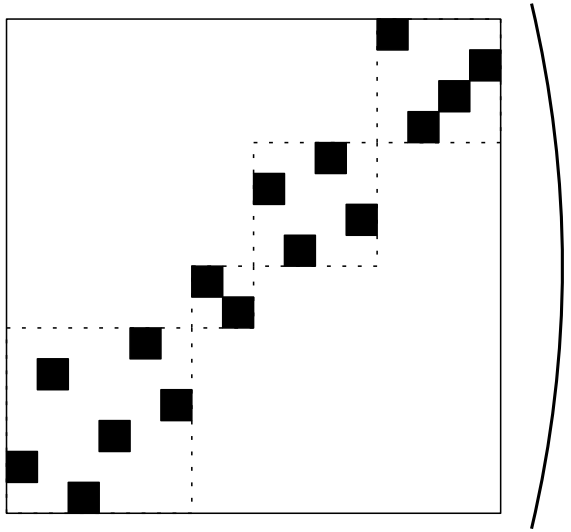
II. Strong bias: decomposing permutations

Definition: decomposability

If $\pi \in \mathfrak{S}_r$ is (uniquely) decomposed into

$$\pi = \pi_1 \oplus \cdots \oplus \pi_d$$

with maximal $d \in \mathbb{N}^*$, then we call $(\pi_i)_{1 \leq i \leq d}$ the blocks of π and $d = d(\pi)$ the number of blocks of π.

$$d\left(\right.$$  $$\left.\right) = 4$$

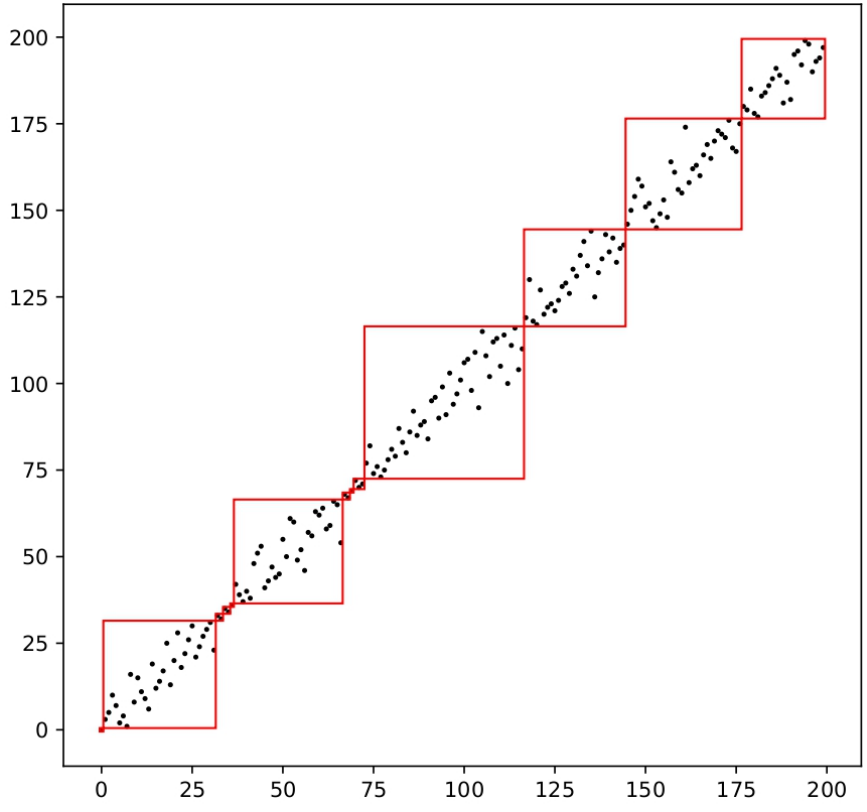
II. Strong bias: placing each block

Regenerative property

$$\tau_n = B_1 \oplus \cdots \oplus B_K \oplus \sigma$$

where $(B_i)_{i \geq 1}$ are i.i.d. random blocks, K is a random stopping time, and σ is a random remainder.

$$\pi = \pi_1 \oplus \cdots \oplus \pi_d$$



II. Strong bias: placing each block

Regenerative property

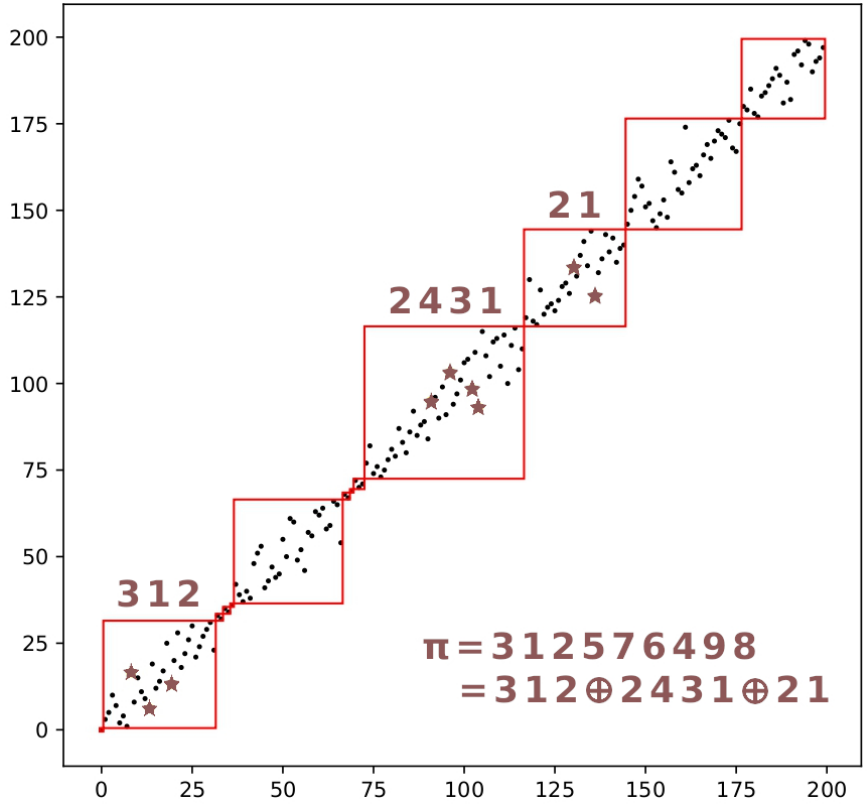
$$\tau_n = B_1 \oplus \cdots \oplus B_K \oplus \sigma$$

where $(B_i)_{i \geq 1}$ are i.i.d. random blocks, K is a random stopping time, and σ is a random remainder.

$$\pi = \pi_1 \oplus \cdots \oplus \pi_d$$

$$\mathrm{Occ}(\pi, \tau_n) \approx \sum_{i_1 < \cdots < i_d} \prod_{1 \leq j \leq d} \mathrm{Occ}(\pi_j, B_{i_j})$$

"U-statistic"



II. Strong bias: conclusion

D. 2024+

Fix $q \in (0,1)$ and let $\tau_n := \mathrm{pat}([n], \Sigma) \sim \mathrm{Mallows}(n, q)$.
For any $\pi = \pi_1 \oplus \cdots \oplus \pi_d$, we have:

$$\frac{\mathrm{Occ}(\pi, \tau_n) - \binom{n}{d} e_{\pi,q}}{n^{d-1/2}} \longrightarrow \mathcal{N}(0, \gamma_{\pi,q})$$

for some "explicit" constants $e_{\pi,q} > 0$ and $\gamma_{\pi,q} \geq 0$.
Furthermore, $\gamma_{\pi,q} > 0$ for all $\pi \neq \mathrm{id}$.

II. Strong bias: conclusion

D. 2024+

Fix $q \in (0,1)$ and let $\tau_n := \mathrm{pat}([n], \Sigma) \sim \mathrm{Mallows}(n, q)$.
For any $\pi = \pi_1 \oplus \cdots \oplus \pi_d$, we have:

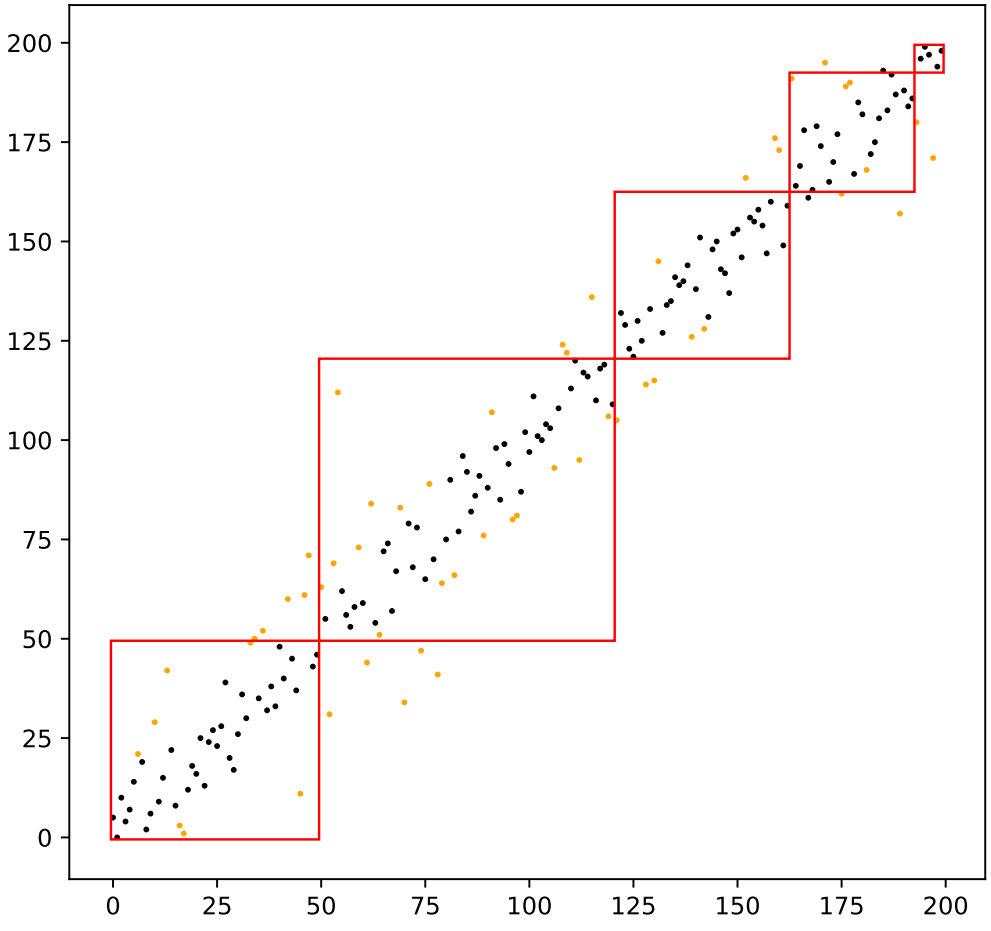
$$\frac{\mathrm{Occ}(\pi, \tau_n) - \binom{n}{d} e_{\pi,q}}{n^{d-1/2}} \longrightarrow \mathcal{N}(0, \gamma_{\pi,q})$$

for some "explicit" constants $e_{\pi,q} > 0$ and $\gamma_{\pi,q} \geq 0$.
Furthermore, $\gamma_{\pi,q} > 0$ for all $\pi \neq \mathrm{id}$.

Also: moment and joint convergences...

III. Moderately strong bias

$\tau_n \sim \mathrm{Mallows}\,(n, q_n)$ with $q_n \to 1$ and $n(1-q_n) \to \infty$.



III. Moderately strong bias

D. 2024+

Let $\tau_n \sim \mathrm{Mallows}\,(n, q)$, where $q_n \to 1$ and $n(1-q_n) \to \infty$.
For any $\pi = \pi_1 \oplus \cdots \oplus \pi_d \in \mathfrak{S}_r$, we have:

$$\mathrm{Occ}\,(\pi, \tau_n) = \widetilde{\Theta}_{\mathsf{P}}\left(\frac{n^d}{(1-q_n)^{r-d}}\right).$$

- This means that

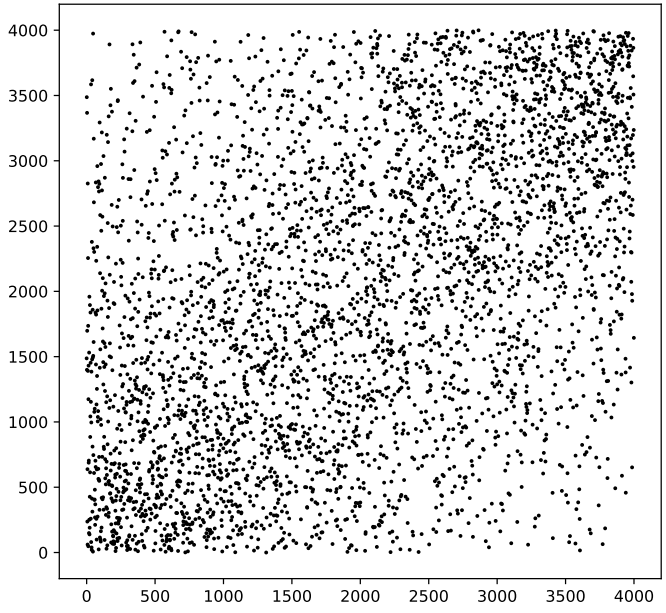
$$\alpha\,\frac{n^d}{(1-q_n)^{r-d}} \;\le\; \mathrm{Occ}\,(\pi, \tau_n) \;\le\; \beta\,(\log n)^{\gamma}\,\frac{n^d}{(1-q_n)^{r-d}}$$

  with high probability, for some constants...

- If $q_n = 1 - cn^{-x}$ for some $c > 0$ and $x \in (0,1)$, then $\mathrm{Occ}\,(\pi, \tau_n)$ has order of magnitude $n^{xr+(1-x)d}$.

Thermodynamic bias

$\tau_n \sim \mathrm{Mallows}\,(n, q_n)$ with $n(1-q_n) \to \beta$.



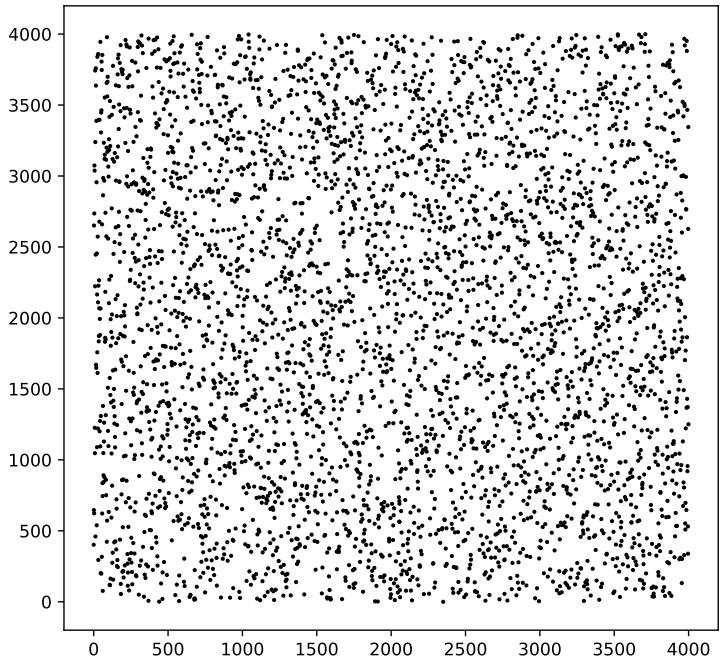
Starr 2009

If $\pi \in \mathfrak{S}_r$ then

$$n^{-r}\mathrm{Occ}\,(\pi, \tau_n) \xrightarrow[\mathsf{P}]{} c_{\beta,\pi} \in (0, \infty).$$

IV. Weak bias

$\tau_n \sim \mathrm{Mallows}\,(n, q_n)$ with $n^x(1-q_n) \to 0$, $x > 1$ well chosen.



Coupling with the uniform case?

IV. Weak bias: online construction

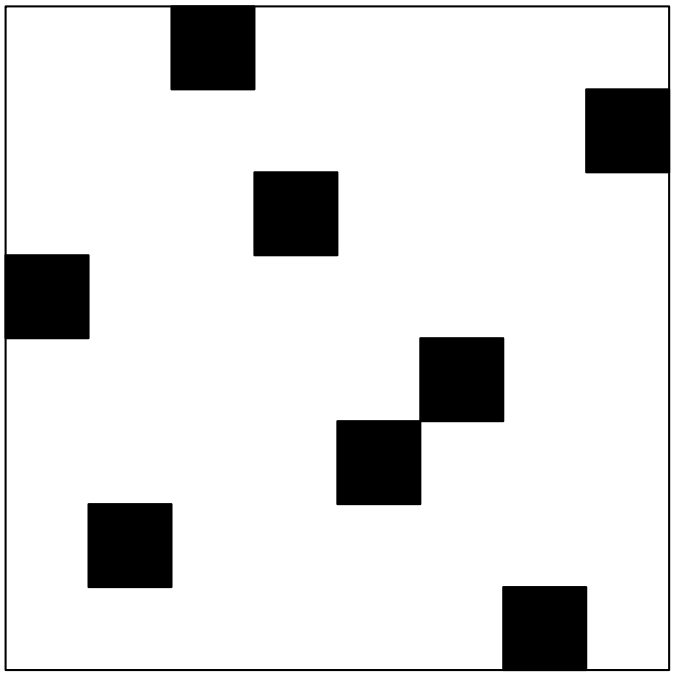
Definition: inversion sequence

Let $\sigma \in \mathfrak{S}_n$, $j \in [n]$. The number of left-inversions at position j is:

$$\ell_j(\sigma) := \left|\left\{ i \in [n] \ : \ i < j \, , \, \sigma(i) > \sigma(j) \right\}\right| .$$

- $\sum_{j=1}^{n} \ell_j(\sigma) = \mathrm{inv}\,(\sigma)$
- $\forall j, \ 0 \leq \ell_j(\sigma) \leq j-1$

$(\ell_j) = \ 0 \quad 1 \quad 0 \quad 1 \quad 3 \quad 3 \quad 6 \quad 1$



IV. Weak bias: online construction

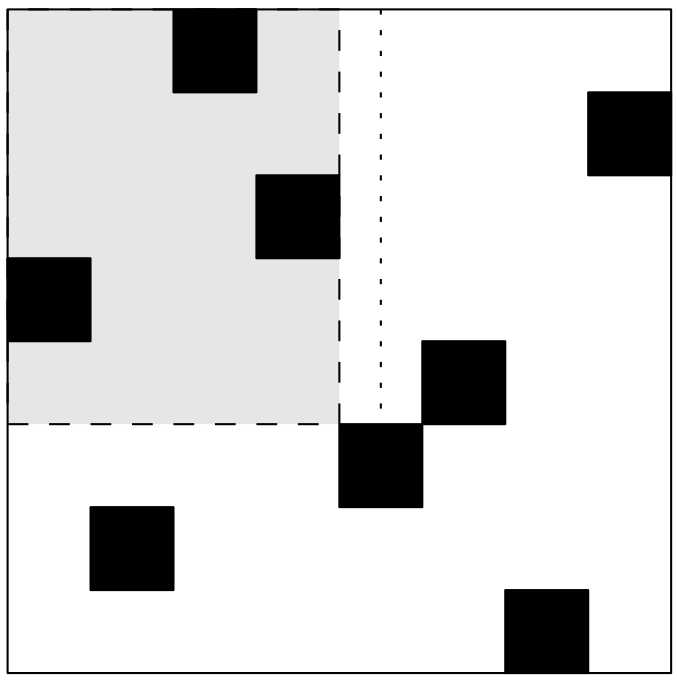
Definition: inversion sequence

Let $\sigma \in \mathfrak{S}_n$, $j \in [n]$. The number of left-inversions at position j is:

$$\ell_j(\sigma) := \left|\left\{ i \in [n] \ : \ i < j \, , \, \sigma(i) > \sigma(j) \right\}\right| .$$

- $\sum_{j=1}^{n} \ell_j(\sigma) = \mathrm{inv}\,(\sigma)$
- $\forall j, \ 0 \leq \ell_j(\sigma) \leq j-1$

$(\ell_j) = \ 0 \quad 1 \quad 0 \quad 1 \quad 3 \quad 3 \quad 6 \quad 1$



IV. Weak bias: online construction

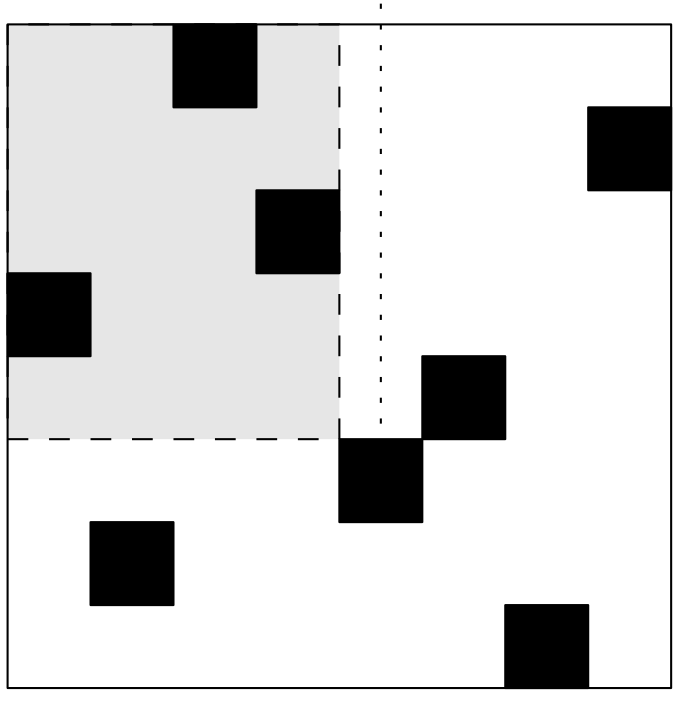
Definition: inversion sequence

Let $\sigma \in \mathfrak{S}_n$, $j \in [n]$. The number of left-inversions at position j is:

$$\ell_j(\sigma) := \left|\left\{ i \in [n] \ : \ i < j \, , \, \sigma(i) > \sigma(j) \right\}\right| .$$

- $\sum_{j=1}^{n} \ell_j(\sigma) = \mathrm{inv}\,(\sigma)$
- $\forall j, \ 0 \leq \ell_j(\sigma) \leq j-1$
- Reciprocally, for any list of numbers $(\ell_j)_{j\in[n]}$ with $0 \leq \ell_j \leq j-1$, there exists a unique $\sigma \in \mathfrak{S}_n$ with these left-inversion numbers.

$(\ell_j) = 0 \ 1 \ 0 \ 1 \ 3 \ 3 \ 6 \ 1$



IV. Weak bias: online construction

Definition: inversion sequence

Let $\sigma \in \mathfrak{S}_n$, $j \in [n]$. The number of left-inversions at position j is:

$$\ell_j(\sigma) := \left|\left\{ i \in [n] \ : \ i < j \, , \, \sigma(i) > \sigma(j) \right\}\right| .$$

$$(\ell_j) = \ 0$$

- $\sum_{j=1}^{n} \ell_j(\sigma) = \mathrm{inv}\,(\sigma)$
- $\forall j, \ 0 \leq \ell_j(\sigma) \leq j-1$
- Reciprocally, for any list of numbers $(\ell_j)_{j\in[n]}$ with $0 \ \leq \ \ell_j \ \leq \ j-1$, there exists a unique $\sigma \in \mathfrak{S}_n$ with these left-inversion numbers.

$$\mathrm{pat}([1], \tau) = 1$$

IV. Weak bias: online construction

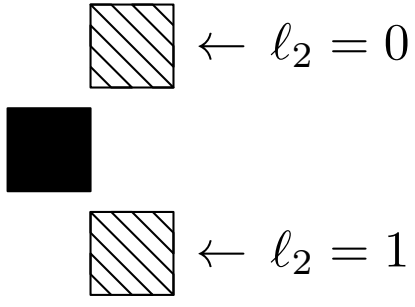
Definition: inversion sequence

Let $\sigma \in \mathfrak{S}_n$, $j \in [n]$. The number of left-inversions at position j is:

$$\ell_j(\sigma) := \left|\left\{ i \in [n] \ : \ i < j \, , \, \sigma(i) > \sigma(j) \right\}\right| .$$

$$(\ell_j) = \ 0$$

- $\sum_{j=1}^{n} \ell_j(\sigma) = \mathrm{inv}\,(\sigma)$
- $\forall j, \ 0 \leq \ell_j(\sigma) \leq j-1$
- Reciprocally, for any list of numbers $(\ell_j)_{j\in[n]}$ with $0 \ \leq \ \ell_j \ \leq \ j-1$, there exists a unique $\sigma \in \mathfrak{S}_n$ with these left-inversion numbers.



$$\mathrm{pat}([1], \tau) = 1$$

IV. Weak bias: online construction

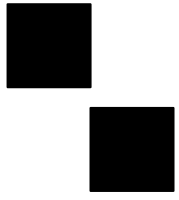
Definition: inversion sequence

Let $\sigma \in \mathfrak{S}_n$, $j \in [n]$. The number of left-inversions at position j is:

$$\ell_j(\sigma) := \left|\left\{ i \in [n] \ : \ i < j \, , \, \sigma(i) > \sigma(j) \right\}\right| .$$

$$(\ell_j) = \ 0 \quad 1$$

- $\sum_{j=1}^{n} \ell_j(\sigma) = \mathrm{inv}\,(\sigma)$
- $\forall j, \ 0 \leq \ell_j(\sigma) \leq j-1$
- Reciprocally, for any list of numbers $(\ell_j)_{j\in[n]}$ with $0 \ \leq \ \ell_j \ \leq \ j-1$, there exists a unique $\sigma \in \mathfrak{S}_n$ with these left-inversion numbers.



$$\mathrm{pat}([2], \tau) = 2\,1$$

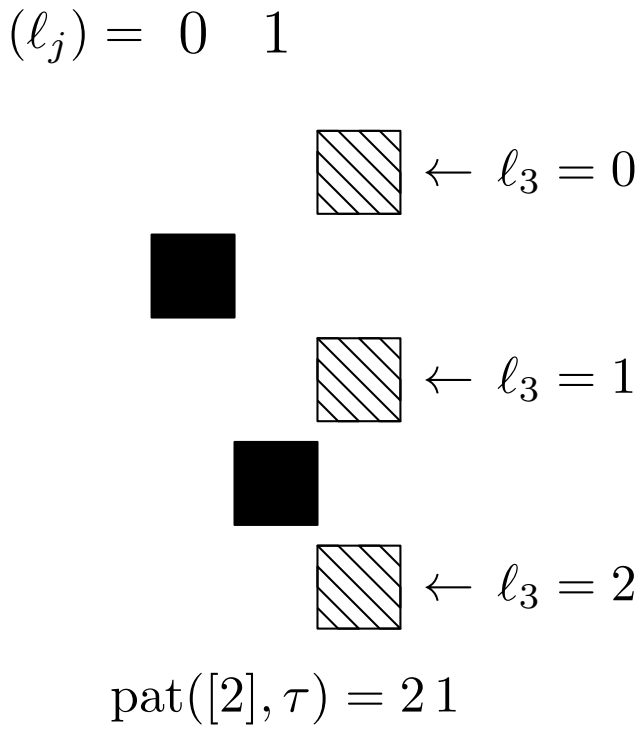
IV. Weak bias: online construction

Definition: inversion sequence

Let $\sigma \in \mathfrak{S}_n$, $j \in [n]$. The number of left-inversions at position j is:

$$\ell_j(\sigma) := \left|\left\{ i \in [n] \ : \ i < j \, , \, \sigma(i) > \sigma(j) \right\}\right| .$$

- $\sum_{j=1}^{n} \ell_j(\sigma) = \mathrm{inv}\,(\sigma)$
- $\forall j,\ 0 \leq \ell_j(\sigma) \leq j-1$
- Reciprocally, for any list of numbers $(\ell_j)_{j\in[n]}$ with $0 \leq \ell_j \leq j-1$, there exists a unique $\sigma \in \mathfrak{S}_n$ with these left-inversion numbers.



IV. Weak bias: online construction

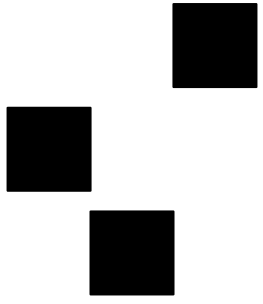
Definition: inversion sequence

Let $\sigma \in \mathfrak{S}_n$, $j \in [n]$. The number of left-inversions at position j is:

$$\ell_j(\sigma) := \left|\left\{ i \in [n] \ : \ i < j \, , \, \sigma(i) > \sigma(j) \right\}\right| .$$

- $\sum_{j=1}^{n} \ell_j(\sigma) = \mathrm{inv}\,(\sigma)$
- $\forall j,\ 0 \leq \ell_j(\sigma) \leq j-1$
- Reciprocally, for any list of numbers $(\ell_j)_{j\in[n]}$ with $0 \ \leq \ \ell_j \ \leq \ j-1$, there exists a unique $\sigma \in \mathfrak{S}_n$ with these left-inversion numbers.

$$(\ell_j) = \ 0 \quad 1 \quad 0$$



$$\mathrm{pat}([3], \tau) = 2\,1\,3$$

IV. Weak bias: online construction

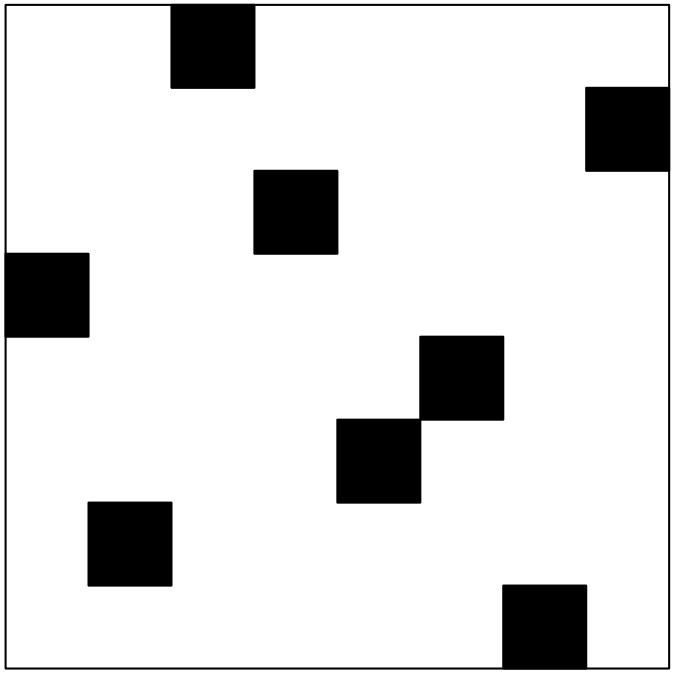
Definition: inversion sequence

Let $\sigma \in \mathfrak{S}_n$, $j \in [n]$. The number of left-inversions at position j is:

$$\ell_j(\sigma) := \left|\left\{i \in [n] \ : \ i < j \, , \, \sigma(i) > \sigma(j)\right\}\right| .$$

- $\sum_{j=1}^{n} \ell_j(\sigma) = \mathrm{inv}\,(\sigma)$
- $\forall j, \ 0 \leq \ell_j(\sigma) \leq j-1$
- Reciprocally, for any list of numbers $(\ell_j)_{j\in[n]}$ with $0 \leq \ell_j \leq j-1$, there exists a unique $\sigma \in \mathfrak{S}_n$ with these left-inversion numbers.

$(\ell_j) =$ 0 1 0 1 3 3 6 1



IV. Weak bias: online construction

Mallows left-inversions

Let $\tau \sim \mathrm{Mallows}(n, q)$. The numbers $\left(\ell_j(\tau)\right)_{j\in[n]}$ are independent random variables, and $\ell_j(\tau)$ follows the $\mathrm{Geom}(1-q)$ distribution conditioned to be in $[0, j-1]$.

$$\forall 0 \leq k \leq j-1, \quad \mathrm{P}\left(\ell_j(\tau) = k\right) = \frac{1-q}{1-q^j} q^k$$

We can thus construct τ using independent geometric variables.

Extremely useful for both simulations and theoretical study!

IV. Weak bias: coupling

Let $(L_j)_{j\in[n]}$ be independent r.v.'s, with $\mathrm{Geom}\,(1-q_n)$ distribution conditioned to be in $[0, j-1]$.
Let $\tau_n \sim \mathrm{Mallows}\,(n, q_n)$ have these left-inversion numbers.

Let $(U_j)_{j\in[n]}$ be independent r.v.'s, with $\mathrm{Unif}\,(\{0, 1, \ldots, j-1\})$ distribution.
Let $u_n \sim \mathrm{Unif}\,(S_n)$ have these left-inversion numbers.

Can we couple (L_j) with (U_j) so that $\tau_n \approx u_n$?

IV. Weak bias: coupling

Let $(L_j)_{j\in[n]}$ be independent r.v.'s, with $\mathrm{Geom}\,(1-q_n)$ distribution conditioned to be in $[0, j-1]$.
Let $\tau_n \sim \mathrm{Mallows}\,(n, q_n)$ have these left-inversion numbers.

Let $(U_j)_{j\in[n]}$ be independent r.v.'s, with $\mathrm{Unif}\,(\{0, 1, \ldots, j-1\})$ distribution.
Let $u_n \sim \mathrm{Unif}\,(S_n)$ have these left-inversion numbers.

Can we couple (L_j) with (U_j) so that $\tau_n \approx u_n$?

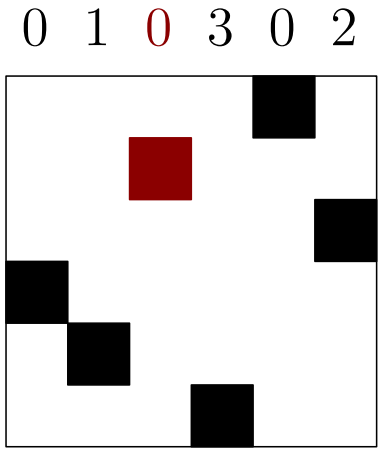
Perfect coupling

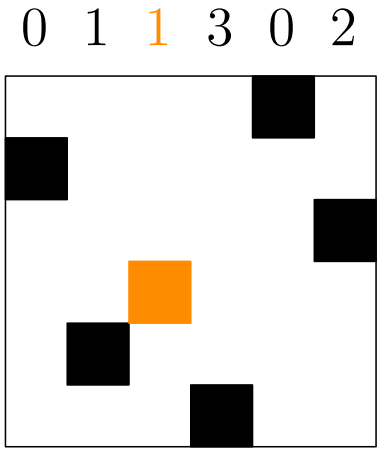
$$\mathrm{d_{TV}}\,(L_j, U_j) \;\leq\; 3j(1-q_n)\,.$$

If $n^2(1-q_n) \to 0$ then $\mathrm{d_{TV}}\,(\tau_n, u_n) \to 0$:
there exists a coupling such that $\tau_n = u_n$ with high probability.

IV. Weak bias: coupling

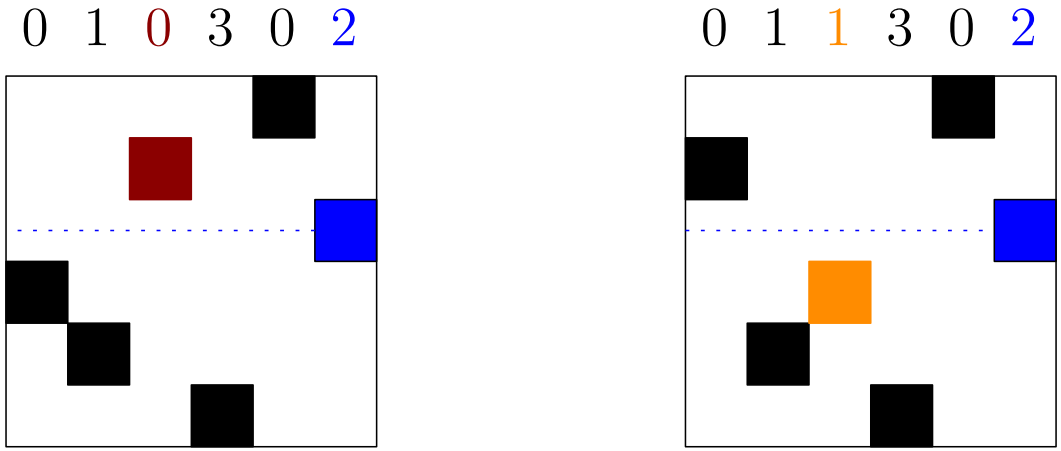
If $n^x(1-q_n) \to 0$ for some $x \in (1,2)$, there might be disagreements $L_j \neq U_j$ for several $j \in [n]$.





IV. Weak bias: coupling

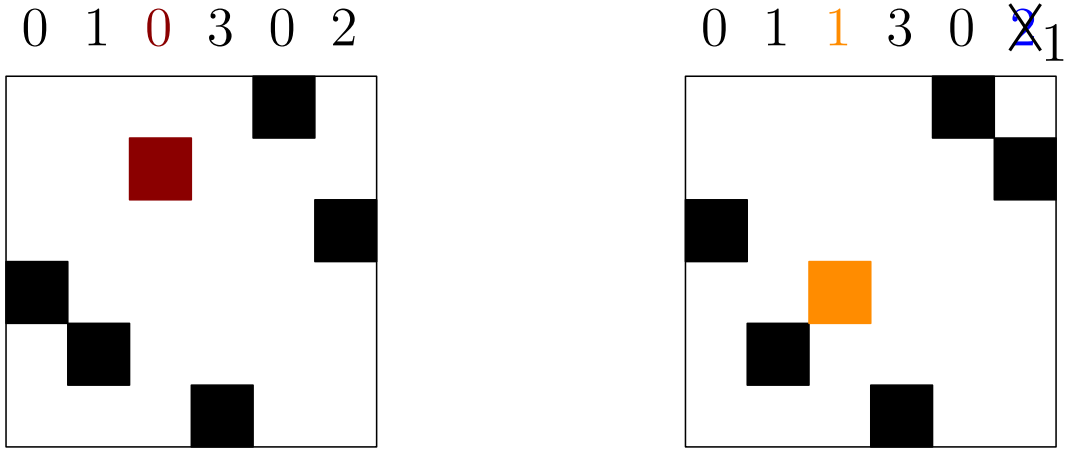
If $n^x(1-q_n) \to 0$ for some $x \in (1,2)$, there might be disagreements $L_j \neq U_j$ for several $j \in [n]$.



Problem: a single coupling mistake can have dire consequences on the following values...

IV. Weak bias: coupling

If $n^x(1-q_n) \to 0$ for some $x \in (1,2)$, there might be disagreements $L_j \neq U_j$ for several $j \in [n]$.



Problem: a single coupling mistake can have dire consequences on the following values...

Solution: at the beginning, maximize the probability of $L_j = U_j$.
After the first mistake, maximize the probability of $L_j = \phi_j(U_j)$ where ϕ_j is an appropriate shift.
After each mistake, adapt the upcoming couplings.

IV. Weak bias: conclusion

With an appropriate coupling, we have $\mathrm{pat}\,(I, \tau_n) = \mathrm{pat}\,(I, u_n)$ for some random $I \subseteq [n]$ of size $|I| = n - \mathcal{O}_{\mathrm{P}}\left(n^2(1 - q_n)\right)$.

IV. Weak bias: conclusion

With an appropriate coupling, we have $\mathrm{pat}\,(I, \tau_n) = \mathrm{pat}\,(I, u_n)$ for some random $I \subseteq [n]$ of size $|I| = n - \mathcal{O}_{\mathrm{P}}\left(n^2(1-q_n)\right)$.

Janson–Nakamura–Zeilberger 2015

Let $u_n \sim \mathrm{Unif}\,(\mathfrak{S}_n)$. For any $\pi \in \mathfrak{S}_r$, we have:

$$\frac{\mathrm{Occ}\,(\pi, u_n) - \frac{1}{r!}\binom{n}{r}}{n^{r-1/2}} \xrightarrow[n\to\infty]{} \mathcal{N}\left(0, \nu_\pi^2\right),$$

for some $\nu_\pi > 0$.

IV. Weak bias: conclusion

With an appropriate coupling, we have $\mathrm{pat}\,(I, \tau_n) = \mathrm{pat}\,(I, u_n)$ for some random $I \subseteq [n]$ of size $|I| = n - \mathcal{O}_{\mathrm{P}}\left(n^2(1-q_n)\right)$.

Janson–Nakamura–Zeilberger 2015

Let $u_n \sim \mathrm{Unif}\,(\mathfrak{S}_n)$. For any $\pi \in \mathfrak{S}_r$, we have:

$$\frac{\mathrm{Occ}\,(\pi, u_n) - \frac{1}{r!}\binom{n}{r}}{n^{r-1/2}} \xrightarrow[n\to\infty]{} \mathcal{N}\left(0, \nu_\pi^2\right),$$

for some $\nu_\pi > 0$.

D. 2024+

The same holds for $\tau_n \sim \mathrm{Mallows}\,(n, q_n)$ if $n^{3/2}(1-q_n) \to 0$.

This hypothesis is optimal.

Thanks

Questions?

Bonus content: infinite Mallows permutations

Definition: the *other* inversion sequence

Let $\sigma \in \mathfrak{S}_n$, $i \in [n]$. The number of right-inversions at position i is:

$$r_i(\sigma) := \left|\left\{j \in [n] \ : \ i < j \, , \, \sigma(i) > \sigma(j)\right\}\right| .$$

This is again one-to-one:

Bonus content: infinite Mallows permutations

Definition: the *other* inversion sequence

Let $\sigma \in \mathfrak{S}_n$, $i \in [n]$. The number of right-inversions at position i is:

$$r_i(\sigma) := \left|\left\{j \in [n] \ : \ i < j \, , \, \sigma(i) > \sigma(j)\right\}\right| .$$

This is again one-to-one:

- $\sigma(1) = r_1 + 1$;

Bonus content: infinite Mallows permutations

Definition: the *other* inversion sequence

Let $\sigma \in \mathfrak{S}_n$, $i \in [n]$. The number of right-inversions at position i is:

$$r_i(\sigma) := \left|\left\{j \in [n] \ : \ i < j \, , \, \sigma(i) > \sigma(j)\right\}\right| .$$

This is again one-to-one:

- $\sigma(1) = r_1 + 1$;
- for all $k \in [n]$, $\sigma(k)$ is the $(r_k + 1)$st lowest number in the set $[n] \setminus \{\sigma(1), \ldots, \sigma(k-1)\}$.

Bonus content: infinite Mallows permutations

Definition: the *other* inversion sequence

Let $\sigma \in \mathfrak{S}_n$, $i \in [n]$. The number of right-inversions at position i is:

$$r_i(\sigma) := \left|\left\{j \in [n] \ : \ i < j \,,\, \sigma(i) > \sigma(j)\right\}\right| .$$

This is again one-to-one:

- $\sigma(1) = r_1 + 1$;
- for all $k \in [n]$, $\sigma(k)$ is the $(r_k + 1)$st lowest number in the set $[n] \setminus \{\sigma(1), \ldots, \sigma(k-1)\}$.

Mallows right-inversions

Let $\tau \sim \mathrm{Mallows}\,(n, q)$. The numbers $\big(r_i(\tau)\big)_{i \in [n]}$ are independent random variables, and $r_i(\tau)$ follows the $\mathrm{Geom}\,(1-q)$ distribution conditioned to be in $[0, n-i]$.

Bonus content: infinite Mallows permutations

Definition: the infinite Mallows permutation

Fix $q \in [0, 1)$, and let $(R_i)_{i \geq 1}$ be a sequence of i.i.d. $\mathrm{Geom}\,(1-q)$ r.v.'s. Almost surely, there exists a (unique) infinite permutation $\Sigma \in \mathfrak{S}_{\mathbb{N}^*}$ with right-inversions $(R_i)_{i \geq 1}$. We call it the infinite Mallows permutation with parameter q.

$\Sigma = \Sigma(1)\ \Sigma(2)\ \Sigma(3) \ldots$ is a random bijection of $\mathbb{N}^*$.

Bonus content: infinite Mallows permutations

Definition: the infinite Mallows permutation

Fix $q \in [0,1)$, and let $(R_i)_{i\geq 1}$ be a sequence of i.i.d. $\mathrm{Geom}\,(1-q)$ r.v.'s. Almost surely, there exists a (unique) infinite permutation $\Sigma \in \mathfrak{S}_{\mathbb{N}^*}$ with right-inversions $(R_i)_{i\geq 1}$. We call it the infinite Mallows permutation with parameter q.

$\Sigma = \Sigma(1)\ \Sigma(2)\ \Sigma(3)\ldots$ is a random bijection of $\mathbb{N}^*$.

Projective limit

$$\forall n \in \mathbb{N}^*, \quad \mathrm{pat}\,(\{1,\ldots,n\},\Sigma) \sim \mathrm{Mallows}\,(n,q)\,.$$

Regenerative property

$$\Sigma = B_1 \oplus B_2 \oplus \ldots$$

where $(B_i)_{i\geq 1}$ is an infinite sequence of i.i.d. random blocks.

Bonus content: Mallows process, functional CLT

Q: continuous coupling of $\mathrm{Mallows}(n, q)$ for $q \in [0, 1)$?

Bonus content: Mallows process, functional CLT

Q: continuous coupling of $\mathrm{Mallows}(n, q)$ for $q \in [0, 1)$?

Geometric birth process

Let $(R_t)_{t\in[0,1)}$ be a birth process with jump rates $p(j, t) = \frac{j+1}{1-t}$, started at $R_0 = 0$.
It is non-explosive, and its marginals are $R_t \sim \mathrm{Geom}(1 - t)$.

Definition: the continuous-time infinite Mallows permutation

Let $(R_{i,t})_{t\in[0,1)}$, $i \geq 1$, be i.i.d. geometric birth processes. Almost surely, there exists a process of infinite permutations $(\Sigma_t)_{t\in[0,1)}$ with these right-inversions.

Bonus content: Mallows process, functional CLT

Q: continuous coupling of $\mathrm{Mallows}(n, q)$ for $q \in [0, 1)$?

Geometric birth process

Let $(R_t)_{t\in[0,1)}$ be a birth process with jump rates $p(j, t) = \frac{j+1}{1-t}$, started at $R_0 = 0$.
It is non-explosive, and its marginals are $R_t \sim \mathrm{Geom}(1 - t)$.

Definition: the continuous-time infinite Mallows permutation

Let $(R_{i,t})_{t\in[0,1)}$, $i \geq 1$, be i.i.d. geometric birth processes. Almost surely, there exists a process of infinite permutations $(\Sigma_t)_{t\in[0,1)}$ with these right-inversions.

Then, we can define $\tau_{n,t} := \mathrm{pat}(\{1, \ldots, n\}, \Sigma_t) \sim \mathrm{Mallows}(n, t)$.
The permutation process $(\tau_{n,t})_{t\in[0,1)}$ is *a* Mallows process of size n.

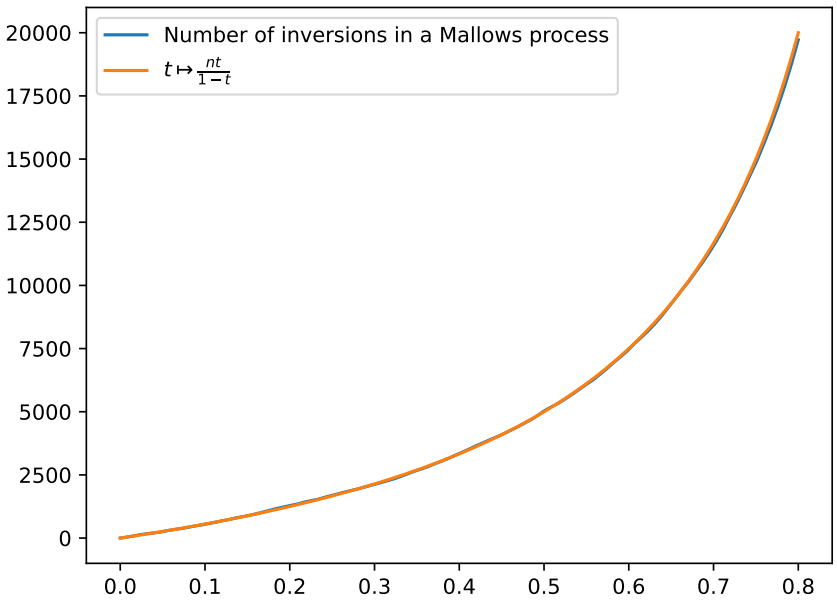
Bonus content: Mallows process, functional CLT

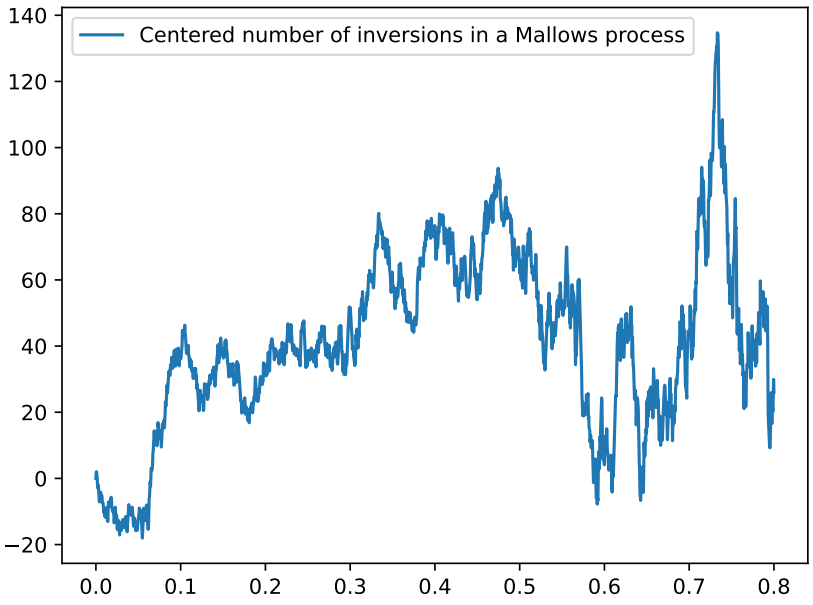
D. 2024+

Let $\tau_{n,t} := \mathrm{pat}([n], \Sigma_t)$. Fix $\pi \in \mathfrak{S}$ with d blocks.
The finite-dimensional distributions of the stochastic process

$$t \in (0,1) \mapsto \frac{\mathrm{Occ}(\pi, \tau_{n,t}) - \binom{n}{d} e_{\pi,t}}{n^{d-1/2}}$$

weakly converge to that of a continuous Gaussian process $(X_{\pi,t})$.





Thanks

Questions?