The Ulam-Hammersley problem for multiset permutations

Lucas Gerin

The Ulam-Hammersley problem for multiset permutations

Lucas Gerin

January 5, 2023

Abstract

We obtain the asymptotic behaviour of the longest increasing/non-decreasing subsequences in a random uniform multiset permutation in which each element in $\{1, \dots, n\}$ occurs k times, where k may depend on n. This generalizes the famous Ulam-Hammersley problem of the case $k = 1$. The proof relies on poissonization and a connection with variants of the Hammersley-Aldous-Diaconis particle system.

1 Introduction

A k-multiset permutation of size n is a word with letters in $\{1, 2, \dots, n\}$ such that each letter appears exactly k times. When this is convenient we identify a multiset permutation $s = (s(1), \dots, s(kn))$ and the set of points $\{(i, s(i)),\ 1 \le i \le kn\}$. For example we say that there is a point at *height* j in s is $s(i) = j$ for some i.

We introduce two partial orders over the quarter-plane $[0, \infty)^2$:

$$(x, y) \prec (x', y') \text{ if } x < x' \text{ and } y < y',$$
$$(x, y) \preccurlyeq (x', y') \text{ if } x < x' \text{ and } y \le y'.$$

(Note that the roles of x, y are not identical in the definition of $\preccurlyeq$.) For a finite set $\mathcal{P}$ of points in the quarter-plane we put

$$\mathcal{L}_<(\mathcal{P}) = \max\{L;\ \text{there exists } P_1 \prec P_2 \prec \cdots \prec P_L,\ \text{where each } P_i \in \mathcal{P}\},$$
$$\mathcal{L}_\le(\mathcal{P}) = \max\{L;\ \text{there exists } P_1 \preccurlyeq P_2 \preccurlyeq \cdots \preccurlyeq P_L,\ \text{where each } P_i \in \mathcal{P}\}.$$

The integer $\mathcal{L}_<(\mathcal{P})$ (resp. $\mathcal{L}_\le(\mathcal{P})$) is the length of the longest increasing (resp. non-decreasing) subsequence of $\mathcal{P}$.

Let $S_{k;n}$ be a k-multiset permutation of size n taken uniformly among the $\binom{kn}{k\ k\ \dots\ k}$ possibilities. In the case $k = 1$ the word $S_{1;n}$ is just a uniform permutation and estimating $\mathcal{L}_<(S_{1;n})$ is known as the Hammersley or Ulam-Hammersley problem. The first order was solved by Veršik and Kerov [VK77] (see [Rom15] for a review of the problem):

$$\mathbb{E}[\mathcal{L}_<(S_{1;n})] = \mathbb{E}[\mathcal{L}_\le(S_{1;n})] \overset{n\to+\infty}{\sim} 2\sqrt{n}.$$

Note that the above limit also holds in probability, in the sense that $\mathbb{P}\left(|\frac{1}{2\sqrt{n}}\mathcal{L}_<(S_{1;n}) - 1| > \varepsilon\right) \to 0$ for every $\varepsilon > 0$, that we shorten into: $\mathcal{L}_<(S_{1;n}) = 2\sqrt{n} + \mathrm{o}_{\mathbb{P}}(\sqrt{n})$.

In the context of card guessing games it is asked in [CDH+22, Question 4.3] the behaviour of $\mathcal{L}_<(S_{k;n})$ for a fixed k. Using known results regarding longest increasing subsequences through independent points we can make an educated guess. Indeed, let $\mathrm{Ber}^{(p)}$ be a field of i.i.d. Bernoulli

random variables with mean p over the square lattice. Seppäläinen [Sep97, Th.1.] proved that for fixed p and k,

$$\mathcal{L}_< \left(\mathrm{Ber}^{(p)} \cap [0, kn] \times [0, n] \right) \overset{n\to\infty}{\sim} n \times \frac{\sqrt{p}}{1-p} \left(2\sqrt{k} - (k+1)\sqrt{p} \right).$$

For $p = 1/n$ there are in average k points of $\mathrm{Ber}^{(p)}$ on each line of $[0, kn] \times [0, n]$. Pretending that we can safely let p depend on n in the above approximation we obtain

$$\mathcal{L}_<(S_{k;n}) \approx \mathcal{L}_< \left(\mathrm{Ber}^{(1/n)} \cap [0, kn] \times [0, n] \right) \approx 2\sqrt{kn}.$$

The goal of the present paper is to make this approximation rigorous (however we are not going to use Seppäläinen's result but rather poissonize the problem). We actually adress this question in the case where k depends on n.

Theorem 1 (Longest increasing subsequences)**.** *Let (k_n) be a sequence of integers such that $k_n \leq n$ for all n. Then*

$$\mathbb{E}[\mathcal{L}_<(S_{k_n;n})] = 2\sqrt{nk_n} - k_n + o(\sqrt{nk_n}). \tag{1}$$

(Of course if $k_n = o(n)$ then the RHS of (1) reduces to $2\sqrt{nk_n} + \mathrm{o}(\sqrt{nk_n})$.) If $k_n \geq n$ for some n then the naive greedy strategy shows very easily that $\mathbb{E}[\mathcal{L}_<(S_{k_n;n})] \geq n - \mathrm{o}(n)$.

Theorem 2 (Longest non-decreasing subsequences)**.** *Let (k_n) be an arbitrary sequence of integers. Then*

$$\mathbb{E}[\mathcal{L}_\leq(S_{k_n;n})] = 2\sqrt{nk_n} + k_n + o(\sqrt{nk_n}). \tag{2}$$

We are not aware of previous results for multiset permutations. However Theorems 1 and 2 in the linear regime $k_n \sim \text{constant} \times n$ should be compared to a result by Biane ([Bia01, Theorem 3]). Indeed he obtains the exact limiting shape of the random Young Tableau induced through the RSK correspondence by a random word $\mathcal{W}_{q;N}$ of q i.i.d. uniform letters in $\{1, 2, \ldots, N\}$ in the regime where $\sqrt{q}/N \to c$ for some constant $c > 0$. (The regime $\sqrt{q}/N \to c$ corresponds to $k_n \sim c^2 n$ with our notation.) Regarding longest increasing subsequences we expect the asymptotics of the word $\mathcal{W}_{q;N}$ to be close to that of the multiset permutation $S_{c^2N;N}$ and Theorems 1 and 2 respectively suggest:

$$L_<(\mathcal{W}_{q;N}) \approx L_<(S_{c^2N;N}) \sim 2Nc - c^2N \sim (2-c)\sqrt{q},$$
$$L_\leq(\mathcal{W}_{q;N}) \approx L_\leq(S_{c^2N;N}) \sim 2Nc + c^2N \sim (2+c)\sqrt{q}.$$

As the length of the first row (resp. the number of rows) in the Young Tableau corresponds to the length of the longest non-decreasing subsequence in $\mathcal{W}_{k;n}$ (resp. the length of the longest decreasing sequence) a weak consequence of ([Bia01, Theorem 3]) is that, in probability,

$$\liminf \frac{1}{\sqrt{q}} L_<(\mathcal{W}_{q;N}) \geq (2-c), \qquad \limsup \frac{1}{\sqrt{q}} L_\leq(\mathcal{W}_{q;N}) \leq (2+c),$$

which is indeed consistent with the above heuristic.

Strategy of proof and organization of the paper. In Section 2 we first provide the proof of Theorems 1 and 2 in the case of a constant or slowly growing sequence (k_n). The proof is elementary (assuming known the Veršik-Kerov Theorem).

For the general case we borrowed a few tools in the literature. In particular we first study poissonized versions of $\mathcal{L}_<(S_{k_n;n}), \mathcal{L}_\leq(S_{k_n;n})$. As already suggested by Hammersley ([Ham72],

Sec.9) and achieved by Aldous-Diaconis [AD95] the case $k = 1$ can be tackled by considering an interacting particle system which is now known as the Hammersley or Hammersley-Aldous-Diaconis particle system.

In Section 3 we introduce and analyze the two variants[1] of the Hammersley process adapted to multiset permutations. The standard path to analyze Hammersley-like models consists in using subadditivity to prove the existence of a limiting shape and then proving that this limiting shape satisfies a variational problem. Typically this variational problem is solved using convex duality (see [Sep97, CG19]). The issue here is that since we allow k_n to have different scales we cannot use this approach and we need to derive non-asymptotic bounds for both processes. This is the purpose of Theorem 8 whose proof is the most technical part of the paper. In Section 4 we detail the multivariate de-poissonization procedure in order to conclude the proof of Theorem 1. De-poissonization is more convoluted for non-decreasing subsequences: see Section 5.

Beyond expectation. In the course of the proof we actually obtain results beyond the estimation of the expectation. We obtain concentration inequalities for the poissonized version of $\mathcal{L}_<(S_{k_n;n}), \mathcal{L}_\leq(S_{k_n;n})$: see Theorem 8 and also the discussion in Section 5.3. Regarding fluctuations a famous result by Baik, Deift and Johansson [BDJ99, Th.1.1] states that

$$\frac{\mathcal{L}_\leq(S_{1;n}) - 2\sqrt{n}}{n^{1/6}} \overset{(d)}{\to} \mathrm{TW}$$

where TW is the Tracy-Widom distribution. For $\mathcal{L}_<(S_{k;n})$ with fixed k we are quite confident that the particle system approach and more precisely the analysis of *second class particles* (see [CG06, CG19]) could be adapted and would yield that the fluctuations are also of order $n^{1/6}$. For general (k_n) the intuition given by the comparison with the Hammersley process would suggest that the fluctuations of $\mathcal{L}_<(S_{k_n;n})$ might be of order $(k_n n)^{1/6}$ as long as (k_n) does not grow too fast. Yet we have no evidence for this.

2 Preliminaries: the case of small k_n

We first prove Theorems 1 and 2 in the case of a *small* sequence (k_n). We say that a sequence (k_n) of integers is *small* if

$$k_n^2 (k_n)! = \mathrm{o}(\sqrt{n}). \tag{3}$$

Note that a sequence of the form $k_n = (\log n)^{1-\varepsilon}$ is small while $k_n = \log n$ is not small.

Proof of Theorems 1 and 2 in the case of a small sequence (k_n). (In order to lighten notation we skip the dependence in n and write $k = k_n$.)
Let σ_{kn} be a random uniform permutation of size kn. We can associate to σ_{kn} a k-multiset permutation $S_{k;n}$ in the following way. For every $1 \leq i \leq kn$ we put

$$S_{k;n}(i) = \lceil \sigma(i)/k \rceil.$$

It is clear that $S_{k;n}$ is uniform and we have

$$\mathcal{L}_<(S_{k;n}) \leq \mathcal{L}_\leq(\sigma_{kn}) \leq \mathcal{L}_\leq(S_{k;n}).$$

[1] The first one is very close to the historical Hammersley process, I discovered during the preparation of this article that the second one had recently appeared in A.Boyer's PhD Thesis [Boy22] with a connection to the O'Connell-Yor Brownian polymer.

The Veršik-Kerov Theorem says that the middle term in the above inequality grows like $2\sqrt{kn}$. Hence we need to show that if (k_n) is small then

$$\mathcal{L}_{\leq}(S_{k;n}) = \mathcal{L}_{<}(S_{k;n}) + o_{\mathbb{P}}(\sqrt{kn}),$$

which is a big step towards the small case of Theorems 1 and 2. For this purpose we introduce for every $\delta > 0$ the event

$$\mathcal{E}_\delta := \left\{ \mathcal{L}_{\leq}(S_{k;n}) \geq \mathcal{L}_{<}(S_{k;n}) + \delta\sqrt{n} \right\}. \tag{4}$$

If $\mathcal{E}_\delta$ occurs then in particular there exists a non-decreasing subsequence with $\delta\sqrt{n}$ ties, *i.e.* points of $S_{k;n}$ which are at the same height as their predecessor in the subsequence. These ties have distinct heights $1 \leq i_1 < \cdots < i_\ell \leq n$ for some $\delta\sqrt{n}/k \leq \ell \leq \delta\sqrt{n}$. Fix

- Integers $m_1, \ldots, m_\ell \geq 2$ such that $(m_1 - 1) + \cdots + (m_\ell - 1) = \delta\sqrt{n}$;
- Column indices $r_{1,1} < \cdots < r_{1,m_1} < r_{2,1} < r_{2,m_1} < \cdots < r_{\ell,1} < \cdots < r_{1,m_\ell}$.

We then introduce the event

$$\begin{aligned} &F = F\left((i_\ell)_\ell, (r_{i,j})_{i\leq \ell, j \leq m_i}\right) \\ &= \left\{ S(r_{1,1}) = \cdots = S(r_{1,m_1}) = i_1, S(r_{2,1}) = \cdots = S(r_{2,m_1}) = i_2, \ldots, S(r_{\ell,1}) = \cdots = S(r_{1,m_\ell}) = i_\ell \right\}. \end{aligned}$$

By the union bound (we skip the integer parts)

$$\mathbb{P}(\mathcal{E}_\delta) \leq \sum_{\delta\sqrt{n}/k \leq \ell \leq \delta\sqrt{n}} \quad \sum_{1 \leq i_1 < \cdots \leq i_\ell \leq n} \quad \sum_{(r_{i,j})_{i\leq \ell, j\leq m_i}} \mathbb{P}\left(F\left((i_\ell)_\ell, (r_{i,j})_{i\leq \ell, j\leq m_i}\right)\right).$$

Using

$$\operatorname{card}\left\{ \sum m_i = \delta\sqrt{n} + \ell;\ \text{each } m_i \geq 2 \right\} = \operatorname{card}\left\{ \sum p_i = \delta\sqrt{n};\ \text{each } p_i \geq 1 \right\} = \binom{\delta\sqrt{n} - 1}{\ell - 1}$$

we obtain the upper bound

$$\begin{aligned} \sum_{(r_{i,j})_{i\leq \ell, j\leq m_i}} \mathbb{P}(F) &= \frac{1}{\binom{kn}{k\ k\ \ldots\ k}} \underbrace{\binom{nk}{\sum m_i}}_{\text{choices of } r\text{'s}} \underbrace{\binom{\delta\sqrt{n} - 1}{\ell - 1}}_{\text{choices of } m_i'\text{s}} \underbrace{\binom{kn - \sum m_i}{(k-m_1)\ (k-m_2)\ \ldots (k-m_\ell) k \ldots k}}_{\text{choices of } kn - \sum m_i \text{ remaining points}} \\ &= \frac{(k!)^\ell (\delta\sqrt{n} - 1)!}{(\delta\sqrt{n} + \ell)!(\delta\sqrt{n} - \ell)!(\ell - 1)!(k - m_1)!(k - m_2)! \times \cdots \times (k - m_\ell)!} \\ &\leq \frac{(k!)^\ell}{(\delta\sqrt{n})^{\ell+1}(\delta\sqrt{n} - \ell)!(\ell - 1)!}. \end{aligned}$$

We now sum over $1 \leq i_1 < \cdots \leq i_\ell \leq n$ and then sum over ℓ:

$$\begin{aligned} \mathbb{P}(\mathcal{E}_\delta) &\leq \sum_{\ell = \delta\sqrt{n}/k}^{\delta\sqrt{n}} \binom{n}{\ell} \frac{(k!)^\ell}{(\delta\sqrt{n})^{\ell+1}(\delta\sqrt{n} - \ell)!(\ell - 1)!} \\ &\leq \sum_{\ell = \delta\sqrt{n}/k}^{\delta\sqrt{n} - 3} \binom{n}{\ell} \frac{(k!)^\ell}{(\delta\sqrt{n})^{\ell+1}(\delta\sqrt{n} - \ell)!(\ell - 1)!} + 3\binom{n}{\delta\sqrt{n}} \frac{(k!)^{\delta\sqrt{n}}}{(\delta\sqrt{n})^{\delta\sqrt{n}-2}(\delta\sqrt{n} - 3)!} \end{aligned} \tag{5}$$

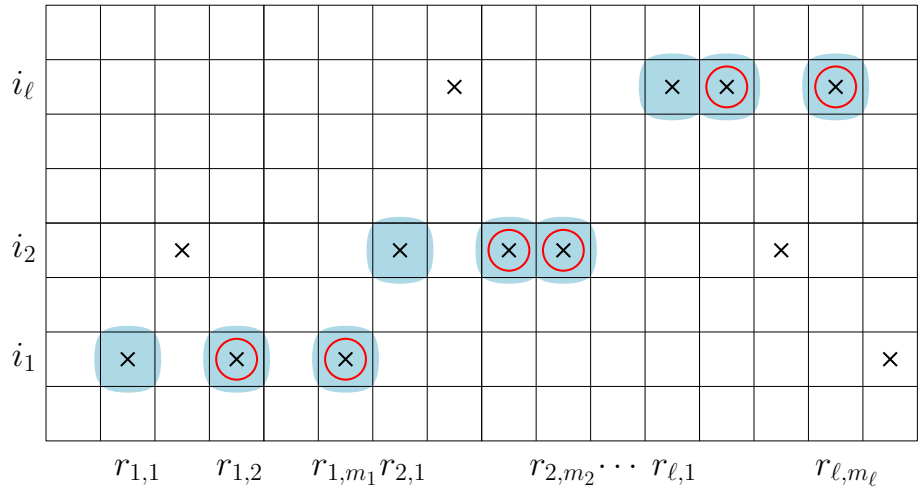


Figure 1: The event F. (Ties are surrounded in red. Points with blue background represent the subsequence with $\delta\sqrt{n}$ ties.)

Using the two following inequalities valid for every $j \le m$ (see *e.g.* [CLRS09, eq.(C.5)])

$$\binom{m}{j} \le \left(\frac{me}{j}\right)^j, \qquad m! \ge m^m \exp(-m)$$

we first obtain that if $k_n! = \mathrm{o}(\sqrt{n})$ (which is the case if (k_n) is small) then the last term of (5) tends to zero. Regarding the sum we write

$$\begin{aligned}
\mathbb{P}(\mathcal{E}_\delta) &\le \sum_{\ell=\delta\sqrt{n}/k}^{\delta\sqrt{n}-3} \left(\frac{ne}{\ell}\right)^\ell \frac{(k!)^\ell}{(\delta\sqrt{n})^{\ell+1}(\delta\sqrt{n}-\ell)^{\delta\sqrt{n}-\ell}e^{-\delta\sqrt{n}+\ell}(\ell-1)^{\ell-1}e^{-\ell+1}} + \mathrm{o}(1) \\
&\le \sum_{\ell=\delta\sqrt{n}/k}^{\delta\sqrt{n}-3} \left(\frac{nek!(\delta\sqrt{n}-\ell)}{\delta\sqrt{n}\ell(\ell-1)}\right)^\ell \frac{(\ell-1)e^{-1}}{\delta\sqrt{n}} \Big(\underbrace{\frac{e}{\delta\sqrt{n}-\ell}}_{\le e/3<1}\Big)^{\delta\sqrt{n}} + \mathrm{o}(1) \\
&\le \sum_{\ell=\delta\sqrt{n}/k}^{\delta\sqrt{n}-3} \left(\frac{\sqrt{n}ek!(\delta\sqrt{n}-\ell)}{\delta\ell(\ell-1)}\right)^\ell \frac{(\ell-1)e^{-1}}{\delta\sqrt{n}} \Big(\frac{e}{\delta\sqrt{n}-\ell}\Big)^{\ell} + \mathrm{o}(1) \\
&\le \sum_{\ell=\delta\sqrt{n}/k}^{\delta\sqrt{n}-3} \left(\frac{\sqrt{n}e^2k!}{\delta\ell(\ell-1)}\right)^\ell \frac{(\ell-1)e^{-1}}{\delta\sqrt{n}} + \mathrm{o}(1), \qquad (6)
\end{aligned}$$

which tends to zero for every $\delta > 0$, as long as (k_n) satisfies (3). This proves that $\mathcal{L}_{\le}(S_{k;n}) = \mathcal{L}_{<}(S_{k;n}) + o_{\mathbb{P}}(\sqrt{kn})$. Using the crude bounds $\mathcal{L}_{<}(S_{k;n}) \le n$ and $\mathcal{L}_{\le}(S_{k;n}) \le kn$, eq.(6) also implies that

$$\mathbb{E}[\mathcal{L}_{\le}(S_{k;n})] = \mathbb{E}[\mathcal{L}_{<}(S_{k;n})] + o(\sqrt{kn}).$$

□

3 Poissonization: variants of the Hammersley process

Remark. *In the sequel,* Poisson(μ) *(resp.* Binomial(n,q)*) stand for generic random variables with Poisson distribution with mean μ (resp. Binomial distribution with parameters n, q).*

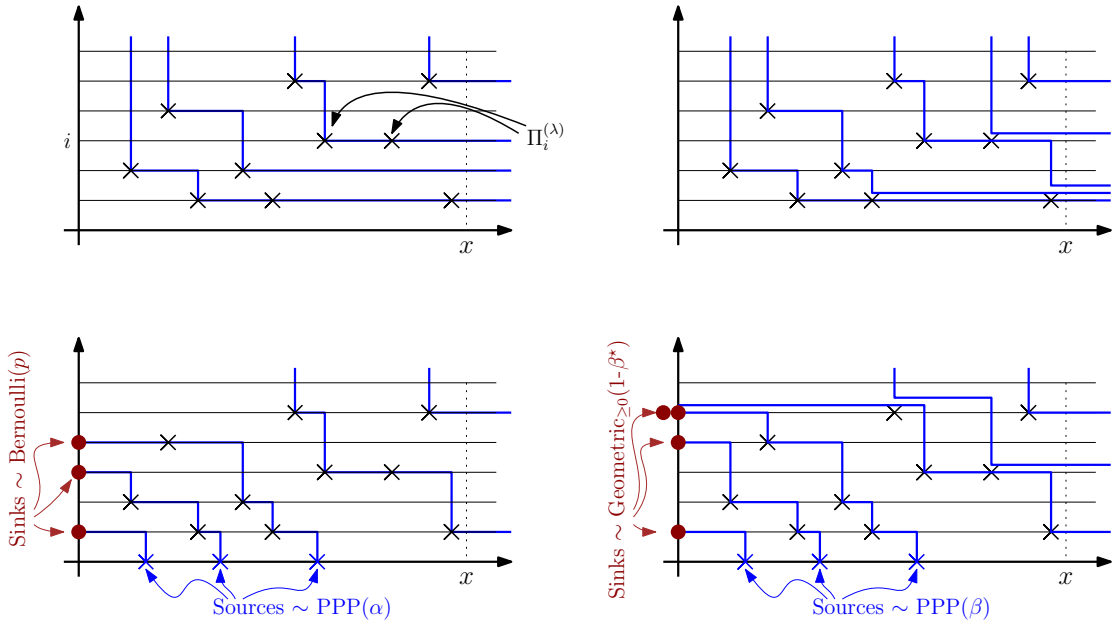


Figure 2: Our four variants of the Hammersley process (time goes from bottom to top, trajectories of particules are indicated in blue). Top left: The process $L_<(t)$. Top right: The process $L_\leq(t)$. Bottom left: The process $L_<^{(\alpha,p)}(t)$. Bottom right: The process $L_\leq^{(\beta,\beta^\star)}(t)$.

Notation $\mathrm{Geometric}_{\geq 0}(1-\beta)$ *stands for a geometric random variable with the convention* $\mathbb{P}(\mathrm{Geometric}_{\geq 0}(1-\beta)=k)=(1-\beta)\beta^k$ *for* $k\geq 0$. *In particular* $\mathbb{E}[\mathrm{Geometric}_{\geq 0}(1-\beta)]=\frac{\beta}{1-\beta}$.

3.1 Definitions of the processes $L_<(t)$ and $L_\leq(t)$

In this Section we define formally and analyze two *semi-discrete* variants of the Hammersley process.

For a parameter $\lambda>0$ let $\Pi^{(\lambda)}$ be the random set $\Pi^{(\lambda)}=\cup_i \Pi_i^{(\lambda)}$ where $\Pi_i^{(\lambda)}$'s are independent and each $\Pi_i^{(\lambda)}$ is a homogeneous Poisson Point Process (PPP) with intensity λ on $(0,\infty)\times\{i\}$. For simplicity set

$$\Pi_{x,t}^{(\lambda)}=\Pi^{(\lambda)}\cap([0,x]\times\{1,\dots,t\})\,.$$

The goal of the present section is to obtain non-asymptotic bounds for $\mathcal{L}_<\left(\Pi_{x,t}^{(\lambda)}\right)$ and $\mathcal{L}_\leq\left(\Pi_{x,t}^{(\lambda)}\right)$.

Fix $x>0$ throughout the section. For every $t\in\{0,1,2,\dots\}$ the function $y\in[0,x]\mapsto\mathcal{L}_<(y,t)$ (resp. $\mathcal{L}_\leq(y,t)$) is a non-decreasing integer-valued function whose all steps are equal to $+1$. Therefore this function is completely determined by the finite set

$$L_<(t):=\left\{y\leq x,\mathcal{L}_<(y,t)=\mathcal{L}_<(y,t^-)+1\right\}.$$

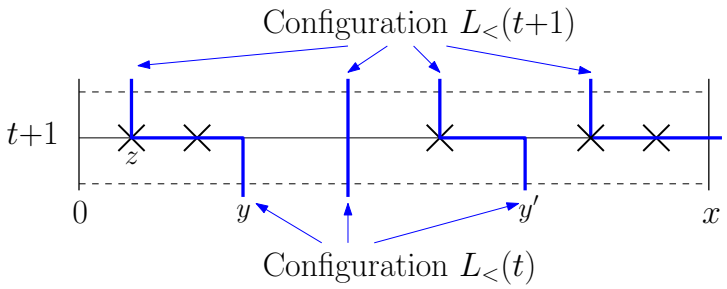
(Respectively:

$$L_\leq(t):=\left\{y\leq x,\mathcal{L}_\leq(y,t)=\mathcal{L}_\leq(y,t^-)+1\right\}.)$$

Sets $L_<(t)$ and $L_\leq(t)$ are finite subsets of $[0,x]$ whose elements are considered as particles. It is easy to see that for fixed $x>0$ both processes $(L_<(t))_t$ and $(L_\leq(t))_t$ are Markov processes taking their values in the family of point processes of $[0,x]$.

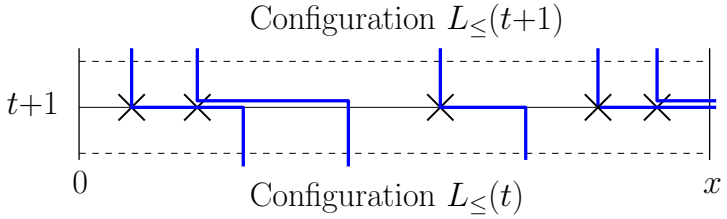
Exactly the same way as for the classical Hammersley process ([Ham72, Sec.9], [AD95]) the individual dynamic of particles is very easy to describe:

- **The process $L_<$.** We put $L_<(0) = \emptyset$. In order to define $L_<(t+1)$ from $L_<(t)$ we consider particles from left to right. A particle at y in $L_<(t)$ moves at time $t+1$ at the location of the leftmost available point z in $\Pi^{(\lambda)}_{t+1} \cap (0, y)$ (if any, otherwise it stays at y). This point z is not available anymore for subsequent particles, as well as every other point of $\Pi^{(\lambda)}_t \cap (0, y)$.



  If there is a point in $\Pi^{(\lambda)}_{t+1}$ which is on the right of $y' := \max\{L_<(t)\}$ then a new particle is created in $L_<(t+1)$, located at the leftmost point in $\Pi^{(\lambda)}_{t+1} \cap (y', x)$. (In pictures this new particle comes from the right.)
  A realization of $L_<$ is shown on top-left of Fig.2.

- **The process $L_\leq$.** We put $L_\leq(0) = \emptyset$. In order to define $L_\leq(t+1)$ from $L_\leq(t)$ we also consider particles from left to right. A particle at y in $L_\leq(t)$ moves at time $t+1$ at the location of the leftmost available point z in $\Pi^{(\lambda)}_t \cap (0, y)$. This point z is not available anymore for subsequent particles, **other points in (z, y) remain available**.



  If there is a point in $\Pi^{(\lambda)}_{t+1}$ which is on the right of $y' := \max\{L_<(t)\}$ then new particles are created in $L_<(t+1)$, one for each point in $\Pi^{(\lambda)}_{t+1} \cap (y', x)$.
  A realization of $L_\leq$ is shown on top-right of Fig.2.

Processes $L_<(t)$ and $L_\leq(t)$ are designed in such a way that they record the length of longest increasing/non-decreasing paths in Π. In fact particles trajectories correspond to the level sets of the functions $(x,t) \mapsto \mathcal{L}_< \left(\Pi^{(\lambda)}_{x,t}\right)$, $(x,t) \mapsto \mathcal{L}_\leq \left(\Pi^{(\lambda)}_{x,t}\right)$.

Proposition 3. *For every x,*

$$\mathcal{L}_< \left(\Pi^{(\lambda)}_{x,t}\right) = \mathrm{card}(L_<(t)), \qquad \mathcal{L}_\leq \left(\Pi^{(\lambda)}_{x,t}\right) = \mathrm{card}(L_\leq(t)),$$

where on each right-hand side we consider the particle system on $[0, x]$.

Proof. We are merely restating the original construction from Hammersley ([Ham72], Sec.9). We only do the case of $L_<(t)$.

Let us call each particle trajectory a *Hammersley line*. By construction each Hammersley line is a broken line starting from the right of the box $[0,x]\times[0,t]$ and is formed by a succession of north/west line segments. Because of this, two distinct points in a given longest increasing subsequence of $\Pi^{(\lambda)}_{x,t}$ cannot belong to the same Hammersley line. Since there are $L_<(t)$ Hammersley's lines this gives $\mathcal{L}_<\left(\Pi^{(\lambda)}_{x,t}\right)\leq \mathrm{card}(L_<(t))$.

In order to prove the converse inequality we build from this graphical construction a longest increcreasing subsequence of $\Pi^{(\lambda)}_{x,t}$ with exactly one point on each Hammersley line. To do so, we order Hammersley's lines from bottom-left to top-right, and we build our path starting from the top-right corner. We first choose any point of $\Pi^{(\lambda)}_{x,t}$ belonging to the last Hammersley line. We then proceed by induction: we choose the next point among the points of of $\Pi^{(\lambda)}_{x,t}$ lying on the previous Hammersley line such that the subsequence remains increasing. (This is possible since Hammersley's lines only have North/West line segments.) This proves $\mathcal{L}_<\left(\Pi^{(\lambda)}_{x,t}\right)\geq \mathrm{card}(L_<(t))$. □

3.2 Sources and sinks: stationarity

Proposition 3 tells us that in on our way to prove Theorem 1 and Theorem 2 we need to understand the asymptotic behaviour of processes $L_<, L_\leq$. These processes are far from being stationary as particles may appear all the time in the system and never disappear. To solve this issue we use the trick of *sources/sinks* introduced formally and exploited by Cator and Groeneboom [CG05] (following the intuition given by [AD95]):

- *Sources* form a finite subset of $[0,x]\times\{0\}$ which plays the role of the initial configuration $L_<(0), L_\leq(0)$.
- *Sinks* are points of $\{0\}\times[1,t]$ which add up to $\Pi^{(\lambda)}$ when one defines the dynamics of $L_<(t), L_\leq(t)$. For $L_\leq(t)$ it makes sense to add several sinks at the same location $(0,i)$ so sinks may have a multiplicity.

Examples of dynamics of $L_<, L_\leq$ under the influence of sources/sinks is illustrated at the bottom of Fig.2.

Lemma 4. *For every $\lambda,\alpha>0$ let $L^{(\alpha,p)}_<(t)$ be the Hammersley process defined as $L_<(t)$ with:*

- *sources distributed according to a homogeneous PPP with intensity α on $[0,x]\times\{0\}$;*
- *sinks distributed according to i.i.d.* Bernoulli(p) *with*

$$\frac{\lambda}{\lambda+\alpha}=p. \tag{7}$$

If sources, sinks, and $\Pi^{(\lambda)}$ are independent then the process $\left(L^{(\alpha,p)}_<(t)\right)_{t\geq 0}$ is stationary.

Lemma 5. *For every $\beta>\lambda>0$ let $L^{(\beta,\beta^\star)}_\leq(t)$ be the Hammersley process defined as $L_\leq(t)$ with:*

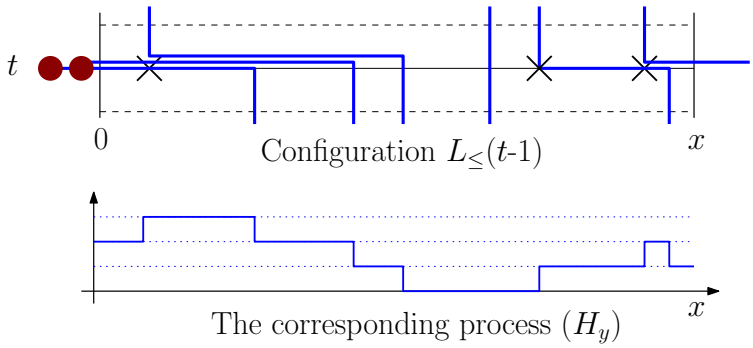
- *sources distributed according to a homogeneous PPP with intensity β on $[0,x]\times\{0\}$;*
- *sinks distributed according to i.i.d.* Geometric$_{\geq 0}(1-\beta^\star)$ *with*

$$\beta^\star\beta=\lambda. \tag{8}$$

If sources, sinks, and $\Pi^{(\lambda)}$ *are independent then the process* $\left(L_{\leq}^{(\beta,\beta^\star)}(t)\right)_{t\geq 0}$ *is stationary.*

Proof of Lemmas 4 and 5. Lemma 5 could be obtain from several minor adjustments of [Boy22, Chap.3, Lemma 3.2]. (Be aware that we have to switch $x \leftrightarrow t$ and sources $\leftrightarrow$ sinks in [Boy22] in order to fit our setup.) For the sake of the reader we however propose the following alternative proof.

Consider for some fixed $t \geq 1$ the process $(H_y)_{0\leq y\leq x}$ given by the number of Hammersley lines passing through the point (y, t).



The initial value H_0 is the number of sinks at $(0, t)$, which is distributed as a Geometric$_{\geq 0}(1-\beta^\star)$. The process (H_y) is a random walk (reflected at zero) with '+1 rate' equal to λ and '−1 rate' equal to β. (Jumps of (H_y) are independent from sinks as sinks are independent from $\Pi^{(\lambda)}$.) The Geometric$_{\geq 0}(1-\beta^\star)$ distribution is stationary for this random walk exactly when (8) holds. The set of points of $L_{\leq}^{(\beta,\beta^\star)}(t)$ is given by the union of $\Pi_t^{(\lambda)}$ and the points of $L_{\leq}^{(\beta,\beta^\star)}(t)$ that do not correspond to a '−1' jump. A elementary computation with exponential and gamma random variables then shows that this is distributed as a homogeneous PPP with intensity β.

Lemma 4 is proved exactly in the same way (alternatively one can mimic the proof of [CG05, Th.3.1.] as the dynamics is the same as in the classical Hammersley process) so we omit details. We simply explain where (7) comes from.

If Lemma 4 is true for some α, p, λ then in particular the number of particles in the system should be constant in expectation. At time t a particle leaves the system (from the left) if and only if there is a sink at height t, which occurs with probability p. On the other hand a new particle appears from the right if there is a point in $\Pi_t^{(\lambda)}$ which is on the right of the right-most particle of $L_<(t)$. This occurs with probability $\lambda/(\alpha+\lambda)$, hence we need (7) to hold so that the expected number of particles remain constant.

□

3.3 Processes $L_<(t)$ and $L_{\leq}(t)$: non-asymptotic bounds

From Lemmas 4 and 5 it is straightforward to derive non-asymptotic upper bounds for $L_<(t), L_{\leq}(t)$.

For $y \leq x$ let $\mathsf{So}_x^{(\alpha)}$ be the random set of sources with intensity α and for $s \leq t$ let $\mathsf{Si}_t^{(p)}$ the random set of sinks with intensity p. It is convenient to use the notation $\mathcal{L}_{=<}(\mathcal{P})$ which is, as before, the length of the longest increasing path taking points in $\mathcal{P}$ but when the path is also allowed to go through several sources (which have however the same y-coordinate) or several sinks (which have the same x-coordinate). Formally,

$$\mathcal{L}_{=<}(\mathcal{P}) = \max\{L;\ \text{there exists } P_1 =\!\prec P_2 =\!\prec \cdots =\!\prec P_L,\ \text{where each } P_i \in \mathcal{P}\},$$

where

$$(x,y) =\prec (x',y') \text{ if } \begin{cases} & x < x' \text{ and } y < y', \\ \text{or} & x = x' = 0 \text{ and } y < y', \\ \text{or} & x < x' \text{ and } y = y' = 0. \end{cases}$$

Proposition 3 can be generalized in

$$\mathcal{L}_{=<}\left(\Pi_{x,t}^{(\lambda)} \cup \mathsf{So}_x^{(\alpha)} \cup \mathsf{Si}_t^{(p)}\right) = L_<^{(\alpha,p)}(t) + \mathrm{card}(\mathsf{Si}_t^{(p)}). \tag{9}$$

Lemma 6 (Upper bound for $\mathcal{L}_<$)**.** *For every $\alpha, p \in (0,1)$ such that* (7) *holds, there is a stochastic domination of the form:*

$$\mathcal{L}_<\left(\Pi_{x,t}^{(\lambda)}\right) \preccurlyeq \mathrm{Poisson}(x\alpha) + \mathrm{Binomial}(t,p). \tag{10}$$

(The Poisson *and* Binomial *random variables involved in* (10) *are not independent.)*

Proof. Adding sources and sinks may not decrease longest increasing paths. Thus,

$$\begin{aligned} \mathcal{L}_<\left(\Pi_{x,t}^{(\lambda)}\right) &\preccurlyeq \mathcal{L}_{=<}\left(\Pi_{x,t}^{(\lambda)} \cup \mathsf{So}_x^{(\alpha)} \cup \mathsf{Si}_t^{(p)}\right) \\ &= L_<^{(\alpha,p)}(t) + \mathrm{card}(\mathsf{Si}^{(p)}) \text{ (using (9))} \\ &\overset{(\mathrm{d})}{=} L_<^{(\alpha,p)}(0) + \mathrm{card}(\mathsf{Si}^{(p)}) \text{ (using stationarity: Lemma 4)} \\ &\overset{(\mathrm{d})}{=} \mathrm{Poisson}(x\alpha) + \mathrm{Binomial}(t,p). \end{aligned}$$

□

Taking expectations in (10) we obtain

$$\mathbb{E}\left[\mathcal{L}_<\left(\Pi_{x,t}^{(\lambda)}\right)\right] \le x\alpha + tp.$$

As the LHS in the above equation does not depend on α, p (provided that α, p satisfies (7)) we will apply (10) with the minimizing choice

$$\bar{\alpha}, \bar{p} := \mathrm{argmin}_{\alpha,p \text{ satisfying (7)}} \{x\alpha + tp\},$$

i.e.

$$\bar{\alpha} = \sqrt{\frac{t\lambda}{x}} - \lambda, \qquad \bar{p} = \sqrt{\frac{x\lambda}{t}}, \qquad x\bar{\alpha} + t\bar{p} = 2\sqrt{xt\lambda} - x\lambda. \tag{11}$$

We have proved

$$\mathbb{E}\left[\mathcal{L}_<\left(\Pi_{x,t}^{(\lambda)}\right)\right] \le 2\sqrt{xt\lambda} - x\lambda.$$

(Compare with (1).) We have a similar statement for non-decreasing subsequences:

Lemma 7 (Upper bound for $\mathcal{L}_\le$)**.** *For every $\beta, \beta^\star \in (0,1)$ such that* (8) *holds, there is a stochastic domination of the form:*

$$\mathcal{L}_\le\left(\Pi_{x,t}^{(\lambda)}\right) \preccurlyeq \mathrm{Poisson}(x\beta) + \mathcal{G}_1^{(\beta^\star)} + \cdots + \mathcal{G}_t^{(\beta^\star)}, \tag{12}$$

where $\mathcal{G}_i^{(\beta^\star)}$'s are i.i.d. $\mathrm{Geometric}_{\ge 0}(1-\beta^\star)$.

We put

$$\bar{\beta}, \bar{\beta}^\star := \mathrm{argmin}_{\beta,\beta^\star \text{ satisfying (8)}} \left\{ x\beta + t\left(\frac{\beta^\star}{1-\beta^\star}\right)\right\}, \tag{13}$$

i.e.

$$\bar{\beta} = \sqrt{\frac{t\lambda}{x}} + \lambda, \qquad \bar{\beta}^\star = \frac{1}{1+\sqrt{t/x\lambda}}, \qquad x\bar{\beta} + t\left(\frac{\bar{\beta}}{1-\bar{\beta}}\right) = 2\sqrt{xt\lambda} + x\lambda. \tag{14}$$

(In particular $\bar{\beta} > \lambda$, as required in Lemma 5.) Eq.(12) yields

$$\mathbb{E}\left[\mathcal{L}_{\leq}\left(\Pi_{x,t}^{(\lambda)}\right)\right] \leq 2\sqrt{xt\lambda} + x\lambda. \tag{15}$$

(Compare with (2).)

Theorem 8 (Concentration for $\mathcal{L}_<$, $\mathcal{L}_\leq$)**.** *There exist strictly positive functions g, h such that for all $\varepsilon > 0$ and for every x, t, λ such that $t \geq x\lambda$*

$$\mathbb{P}(\mathcal{L}_<(\Pi_{x,t}^{(\lambda)}) > (1+\varepsilon)(2\sqrt{xt\lambda} - x\lambda)) \leq \exp(-g(\varepsilon)(\sqrt{xt\lambda} - x\lambda)), \tag{16}$$

$$\mathbb{P}(\mathcal{L}_<(\Pi_{x,t}^{(\lambda)}) < (1-\varepsilon)(2\sqrt{xt\lambda} - x\lambda)) \leq \exp(-h(\varepsilon)(\sqrt{xt\lambda} - x\lambda)). \tag{17}$$

Similarly:

$$\mathbb{P}(\mathcal{L}_\leq(\Pi_{x,t}^{(\lambda)}) > (1+\varepsilon)(2\sqrt{xt\lambda} + x\lambda)) \leq \exp(-g(\varepsilon)\sqrt{xt\lambda}), \tag{18}$$

$$\mathbb{P}(\mathcal{L}_\leq(\Pi_{x,t}^{(\lambda)}) < (1-\varepsilon)(2\sqrt{xt\lambda} + x\lambda)) \leq \exp(-h(\varepsilon)\sqrt{xt\lambda}). \tag{19}$$

For the proof of Theorem 8 we will focus on the case of $\mathcal{L}_<$, *i.e.* eq.(16), (17). When necessary we will give the slight modification needed to prove eq.(18) and (19). The beginning of the proof mimics Lemmas 4.1 and 4.2 in [BEGG16].

We first prove similar bounds for the stationary processes with minimizing sources and sinks.

Lemma 9 (Concentration for $\mathcal{L}_<$ with sources and sinks)**.** *Let $\bar{\alpha}, \bar{p}$ be defined by* (11)*. There exists a strictly positive function g_1 such that for all $\varepsilon > 0$ and for every x, t, λ such that $t \geq x\lambda$*

$$\mathbb{P}(\mathcal{L}_{=<}(\Pi_{x,t}^{(\lambda)} \cup \mathsf{So}_x^{(\bar{\alpha})} \cup \mathsf{Si}_t^{(\bar{p})}) > (1+\varepsilon)(2\sqrt{xt\lambda} - x\lambda)) \leq 2\exp(-g_1(\varepsilon)(\sqrt{xt\lambda} - x\lambda)) \tag{20}$$

$$\mathbb{P}(\mathcal{L}_{=<}(\Pi_{x,t}^{(\lambda)} \cup \mathsf{So}_x^{(\bar{\alpha})} \cup \mathsf{Si}_t^{(\bar{p})}) < (1-\varepsilon)(2\sqrt{xt\lambda} - x\lambda)) \leq 2\exp(-g_1(\varepsilon)(\sqrt{xt\lambda} - x\lambda)). \tag{21}$$

Proof of Lemma 9. By stationarity (Lemma 4) we have

$$\mathcal{L}_{=<}(\Pi_{x,t}^{(\lambda)} \cup \mathsf{So}_x^{(\bar{\alpha})} \cup \mathsf{Si}_t^{(\bar{p})}) \overset{(d)}{=} \mathrm{Poisson}(x\bar{\alpha}) + \mathrm{Binomial}(t, \bar{p}).$$

Then

$$\begin{aligned}\mathbb{P}(\mathcal{L}_{=<}(\Pi_{x,t}^{(\lambda)} \cup \mathsf{So}_x^{(\bar{\alpha})} \cup \mathsf{Si}_t^{(\bar{p})} > (1+\varepsilon)(2\sqrt{xt\lambda} - x\lambda)) &\leq \mathbb{P}\left(\mathrm{Poisson}(x\bar{\alpha}) > (1+\frac{\varepsilon}{2})(\sqrt{xt\lambda} - x\lambda)\right) \\ &\quad + \mathbb{P}\left(\mathrm{Binomial}(t,\bar{p}) > (1+\frac{\varepsilon}{2})\sqrt{xt\lambda}\right).\end{aligned}$$

Recall that $x\bar{\alpha} = \sqrt{xt\lambda} - x\lambda$, $t\bar{p} = \sqrt{xt\lambda}$. Using the tail inequality for the Poisson distribution (Lemma 14):

$$\mathbb{P}\left(\mathrm{Poisson}(x\bar{\alpha}) > (1+\frac{\varepsilon}{2})(\sqrt{xt\lambda} - x\lambda)\right) \leq \exp\left(-(\sqrt{xt\lambda} - x\lambda)\varepsilon^2/4\right)$$

Using the tail inequality for the binomial (Lemma 15) we get

$$\mathbb{P}\left(\mathrm{Binomial}(t,\bar{p}) > (1+\frac{\varepsilon}{2})\sqrt{xt\lambda}\right) \leq \exp(-\tfrac{1}{12}\varepsilon^2\sqrt{xt\lambda}) \leq \exp(-\tfrac{1}{12}\varepsilon^2(\sqrt{xt\lambda} - x\lambda))$$

The proof of (21) is identical. This shows Lemma 9 with $g_1(\varepsilon) = \varepsilon^2/12$. □

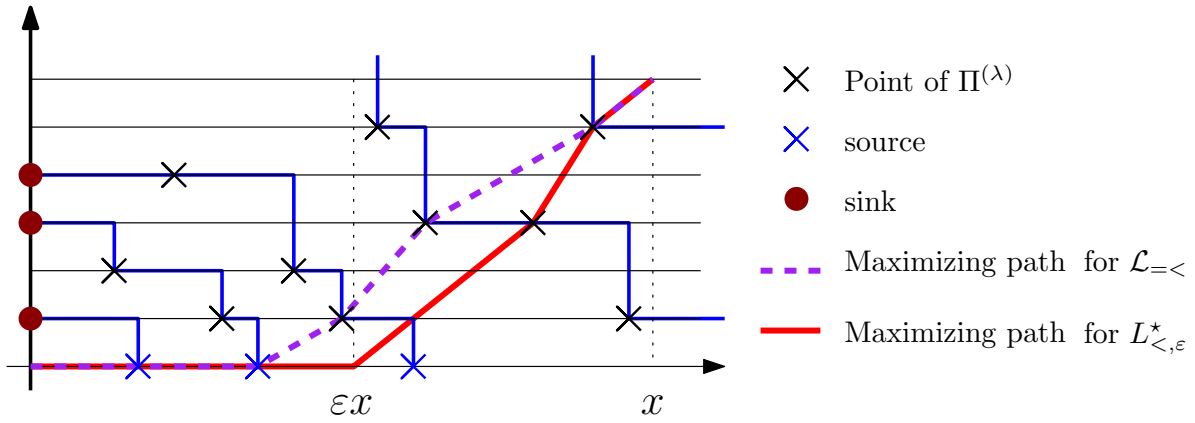


Figure 3: A sample of $\Pi^{(\lambda)}_{x,t}$, sources, sinks, and the corresponding trajectories of particles (in blue). Here $\mathcal{L}_{=<}(\Pi^{(\lambda)}_{x,t} \cup \mathsf{So}^{(\alpha)}_x \cup \mathsf{Si}^{(p)}_t) = 5$ (pink path) and $L^{(\alpha,p)}_<(t) = 2$ (two remaining particles at the top of the box).

For longest non-decreasing subsequences we have a statement similar to Lemma 9. The only modification in the proof is that in order to estimate the number of sinks one has to replace Lemma 15 (tail inequality for the Binomial) by Lemma 16 (tail inequality for a sum of geometric random variables). During the proof we need to bound $\sqrt{xt\lambda} + x\lambda$ by $\sqrt{xt\lambda}$, this explains the form of the right-hand side in eq.(18) and (19).

Proof of Theorem 8. Adding sources/sinks may not decrease $\mathcal{L}_<$ so

$$\mathcal{L}_{=<}(\Pi^{(\lambda)}_{x,t} \cup \mathsf{So}^{(\bar\alpha)}_x \cup \mathsf{Si}^{(\bar p)}_t) \succcurlyeq \mathcal{L}_<(\Pi^{(\lambda)}_{x,t}),$$

thus the upper bound (16) is a direct consequence of Lemma 9.

Let us now prove the lower bound. We consider the length of a maximizing path among those using sources from 0 to εx and then only increasing points of $\Pi^{(\lambda)}_{x,t} \cap ([\varepsilon x, x] \times [0,t])$ (see Fig.3). Formally we set

$$L^{\star}_{=<,\varepsilon} := \mathrm{card}\left(\mathsf{So}^{(\bar\alpha)}_{\varepsilon x}\right) + \mathcal{L}_<\left((\Pi^{(\lambda)}_{x,t} \cap ([\varepsilon x, x] \times [0,t])\right) \overset{(\mathrm{d})}{=} \mathrm{Poisson}(\varepsilon x \bar\alpha) + \mathcal{L}_<\left((\Pi^{(\lambda)}_{x,t} \cap ([\varepsilon x, x] \times [0,t])\right) \tag{22}$$

The idea is that for any fixed ε the paths contributing to $L^{\star}_{=<,\varepsilon}$ will typically not contribute to $\mathcal{L}_{=<}\left(\Pi^{(\lambda)}_{x,t} \cup \mathsf{So}^{(\bar\alpha)}_x \cup \mathsf{Si}^{(\bar p)}_t\right) = L^{(\bar\alpha,p)}_<(t) + \mathrm{card}(\mathsf{Si}^{(\bar p)}_t)$. Indeed eq.(22) suggests that for large x, t

$$\begin{aligned} L^{\star}_{=<,\varepsilon} &\approx \mathbb{E}[\mathrm{Poisson}(\varepsilon x \bar\alpha)] + \mathbb{E}\left[\mathcal{L}_<\left(\Pi^{(\lambda)}_{x,t} \cap ([\varepsilon x, x] \times [0,t])\right)\right] \\ &\approx x\varepsilon\bar\alpha + 2\sqrt{x(1-\varepsilon)\lambda t} - x(1-\varepsilon)\lambda \\ &= 2\sqrt{x\lambda t} - x\lambda - \sqrt{xt\lambda}\delta(\varepsilon), \end{aligned}$$

where $\delta(\varepsilon) = 2 - \varepsilon - 2\sqrt{1-\varepsilon}$ is positive and increasing. In order to make the above approximation rigorous we first write

$$2\sqrt{x\lambda t} - x\lambda - \tfrac{1}{2}\sqrt{xt\lambda}\delta(\varepsilon) = x\varepsilon\bar\alpha + \tfrac{1}{4}\sqrt{xt\lambda}\delta(\varepsilon) + 2\sqrt{x(1-\varepsilon)\lambda t} - x(1-\varepsilon)\lambda + \tfrac{1}{4}\sqrt{xt\lambda}\delta(\varepsilon). \tag{23}$$

Combining (22) and (23) gives

$$
\begin{aligned}
&\mathbb{P}\left(L^{\star}_{<,\varepsilon} \geq 2\sqrt{x\lambda t} - x\lambda - \tfrac{1}{2}\sqrt{xt\lambda}\delta(\varepsilon)\right)\\
&\leq \mathbb{P}\left(\text{Poisson}(x\varepsilon\bar{\alpha}) \geq x\varepsilon\bar{\alpha} + \tfrac{1}{4}\sqrt{xt\lambda}\delta(\varepsilon)\right)\\
&+ \mathbb{P}\left(\mathcal{L}_{<}(\Pi^{(\lambda)}_{x,t} \cap ([\varepsilon x, x] \times [0,t]) \geq 2\sqrt{x(1-\varepsilon)\lambda t} - x(1-\varepsilon)\lambda + \tfrac{1}{4}\sqrt{xt\lambda}\delta(\varepsilon)\right)\\
&\leq \exp\left(-\frac{xt\lambda\delta(\varepsilon)^2}{16\times 4\varepsilon^2(\sqrt{xt\lambda} - x\lambda)}\right) \qquad \text{(using Lemma 14)}\\
&+ \mathbb{P}\left(\mathcal{L}_{<}(\Pi^{(\lambda)}_{x,t} \cap ([\varepsilon x, x] \times [0,t]) \geq 2\sqrt{x(1-\varepsilon)\lambda t} - x(1-\varepsilon)\lambda + \tfrac{1}{8}(2\sqrt{x(1-\varepsilon)\lambda t} - x(1-\varepsilon)\lambda)\delta(\varepsilon)\right)\\
&\leq \exp\left(-\sqrt{xt\lambda}\delta(\varepsilon)^2/(64\varepsilon^2)\right)\\
&+ \mathbb{P}\left(\mathcal{L}_{<}(\Pi^{(\lambda)}_{x,t} \cap ([\varepsilon x, x] \times [0,t]) \geq \left(2\sqrt{x(1-\varepsilon)\lambda t} - x(1-\varepsilon)\lambda\right) \times (1 + \tfrac{1}{8}\delta(\varepsilon))\right)\\
&\leq \exp\left(-\sqrt{xt\lambda}\delta(\varepsilon)^2/(64\varepsilon^2)\right)\\
&\quad + \exp\left(-g(\delta(\varepsilon)/8)(\sqrt{x(1-\varepsilon)t\lambda} - x(1-\varepsilon)\lambda)\right) \text{(using the upper bound (16))}
\end{aligned}
$$

and thus we can find some positive h such that

$$\mathbb{P}\left(L^{\star}_{<,\varepsilon} \geq 2\sqrt{x\lambda t} - x\lambda - \tfrac{1}{2}\sqrt{xt\lambda}\delta(\varepsilon)\right) \leq \exp\left(-h(\varepsilon)(\sqrt{xt\lambda} - x\lambda)\right). \tag{24}$$

One proves exactly in the same way a similar bound for the length of a maximizing path among those using sinks in $\{0\} \times [0, \varepsilon t]$ and then only increasing points of $\Pi^{(\lambda)}_{x,t} \cap ([0,x] \times [\varepsilon t, t])$.

Choose now one of the maximizing paths $\mathcal{P}$ for $\mathcal{L}_{=<}\left(\Pi^{(\lambda)}_{x,t} \cup \mathsf{So}^{(\bar{\alpha})}_{x} \cup \mathsf{Si}^{(\bar{p})}_{t}\right)$ (if there are many of them, choose one arbitrarily in a deterministic way: the lowest, say). Denote by $\mathsf{sources}(\mathcal{P})$ and $\mathsf{sinks}(\mathcal{P})$ the number of sources and sinks in the path $\mathcal{P}$:

$$\mathsf{sources}(\mathcal{P}) = \text{card}\left\{0 \leq y \leq x \text{ such that } (y,0) \in \mathcal{P}\right\}.$$

In Fig.3 the path $\mathcal{P}$ is sketched in pink and $\mathsf{sources}(\mathcal{P}) = 2$, $\mathsf{sinks}(\mathcal{P}) = 0$.

Lemma 10. *There exists a positive function ψ such that for all real $\eta > 0$*

$$\mathbb{P}\left(\mathsf{sources}(\mathcal{P}) + \mathsf{sinks}(\mathcal{P}) \geq \eta\sqrt{x\lambda t}\right) \leq 2\exp(-\psi(\eta)(\sqrt{x\lambda t} - x\lambda)).$$

Proof of Lemma 10. If the event $\left\{\mathsf{sources}(\mathcal{P}) \geq \eta\sqrt{x\lambda t}\right\}$ holds then there exists a (random) ε such that the two following events hold:

- $\mathsf{So}_{\varepsilon x} \geq \eta\sqrt{x\lambda t}$;
- $L^{\star}_{=<,\varepsilon} = \mathcal{L}_{=<}\left(\Pi^{(\lambda)}_{x,t} \cup \mathsf{So}^{(\bar{\alpha})}_{x} \cup \mathsf{Si}^{(\bar{p})}_{t}\right) = L^{(\bar{\alpha},\bar{p})}_{<}(t) + \text{card}(\mathsf{Si}^{(\bar{p})}_{t})$.

This implies that this random ε is larger than $\eta/2 > 0$ unless the number of sources in $[0, x\eta/2]$ is improbably high:

$$
\begin{aligned}
\mathbb{P}(\mathsf{sources}(\mathcal{P}) \geq \eta\sqrt{x\lambda t}) &\leq \mathbb{P}(\mathsf{So}_{\eta x/2} \geq \eta\sqrt{x\lambda t}) + \mathbb{P}(\mathsf{sources}(\mathcal{P}) \geq \eta\sqrt{x\lambda t};\ \mathsf{So}_{\eta x/2} < \eta\sqrt{x\lambda t})\\
&\leq \mathbb{P}(\mathsf{So}_{\eta x/2} \geq \eta\sqrt{x\lambda t})\\
&+ \mathbb{P}(L^{(\bar{\alpha},\bar{p})}_{<}(t) \leq \sqrt{x\lambda t} - x\lambda - \tfrac{1}{4}\delta(\eta/3)\sqrt{x\lambda t})\\
&+ \mathbb{P}(\text{card}(\mathsf{Si}^{(\bar{p})}_{t}) \leq \sqrt{x\lambda t} - \tfrac{1}{4}\delta(\eta/3)\sqrt{x\lambda t})\\
&+ \mathbb{P}(L^{\star}_{=<,\varepsilon} \geq 2\sqrt{x\lambda t} - x\lambda - \tfrac{1}{2}\delta(\eta/3)\sqrt{x\lambda t} \text{ for some } \eta/2 \leq \varepsilon \leq 1).
\end{aligned}
$$

From previous calculations, the three first terms in the above display are less than $\exp(-\phi(\eta)(\sqrt{x\lambda t}-x\lambda))$ for some positive function ϕ. To conclude the proof it remains to bound the fourth term. Let K be an integer larger than $6/\eta^3$, by definition of $L^\star_{=<,\varepsilon}$ we have for every $\varepsilon \in [\frac{k}{K}, \frac{k+1}{K})$

$$L^\star_{=<,\varepsilon} \le L^\star_{=<,k/K} + \mathrm{card}(\mathsf{So}_x^{(\bar\alpha)} \cap [\tfrac{k}{K}, \tfrac{k+1}{K}]).$$

Thus

$$\begin{aligned}
&\mathbb{P}\Big(\bigcup_{\eta/2\le\varepsilon\le 1} L^\star_{=<,\varepsilon} > 2\sqrt{x\lambda t} - x\lambda - \tfrac12\delta(\eta/3)\sqrt{x\lambda t}\Big)\\
&\quad\le \sum_{k\ge\lfloor\eta K/2\rfloor} \mathbb{P}\left(L^\star_{=<,k/K} > 2\sqrt{x\lambda t} - x\lambda - \delta(\eta/3)\sqrt{x\lambda t}\right)\\
&\quad+ \sum_{k\ge\lfloor\eta K/2\rfloor} \mathbb{P}\left(\mathrm{card}(\mathsf{So}_x^{(\bar\alpha)} \cap [\tfrac{k}{K}, \tfrac{k+1}{K}]) > \tfrac12\delta(\eta/3)\sqrt{x\lambda t}\right)\\
&\quad\le \sum_{k\ge\lfloor\eta K/2\rfloor} \mathbb{P}\left(L^\star_{=<,k/K} > 2\sqrt{x\lambda t} - x\lambda - \delta(k/K)\sqrt{x\lambda t}\right) \qquad \text{(since } K > 6/\eta^3 > 6/\eta \text{ and } \delta \text{ is increasing)}\\
&\quad+ \sum_{k\ge\lfloor\eta K/2\rfloor} \mathbb{P}\left(\mathrm{card}(\mathsf{So}_x^{(\bar\alpha)} \cap [\tfrac{k}{K}, \tfrac{k+1}{K}]) > \tfrac12\delta(\eta/3)\sqrt{x\lambda t}\right)\\
&\quad\le \sum_{k\ge\lfloor\eta K/2\rfloor} \exp(-h(k/K)(\sqrt{xt\lambda} - x\lambda)) \qquad \text{(using (24))}\\
&\quad+K\times\mathbb{P}\left(\mathrm{Poisson}(\bar\alpha/K) > \tfrac12\delta(\eta/3)\sqrt{x\lambda t}\right)\\
&\quad\le K\exp(-h(\eta/3)(\sqrt{xt\lambda} - x\lambda)) + K\times\mathbb{P}\left(\mathrm{Poisson}(\bar\alpha/K) > \tfrac12\delta(\eta/3)\sqrt{x\lambda t}\right),\\
&\quad\le K\exp(-h(\eta/3)(\sqrt{xt\lambda} - x\lambda)) + K\times\mathbb{P}\left(\mathrm{Poisson}(\bar\alpha/K) > \bar\alpha/K + \frac{\eta^2}{32}\sqrt{x\lambda t}\right), \qquad \text{(using } K > 6/\eta^3)\\
&\quad\le\frac{12}{\eta}\exp(-\varphi(\eta)(\sqrt{xt\lambda} - x\lambda)),
\end{aligned}$$

for some positive function $\varphi(\eta)$. We can find ψ such that

$$\min\left\{1, \frac{12}{\eta}e^{-\varphi(\eta)(\sqrt{xt\lambda}-x\lambda)} + 3\exp(-\phi(\eta))\right\} \le e^{-\psi(\eta)(\sqrt{xt\lambda}-x\lambda)}$$

and thus $\mathbb{P}(\mathsf{sources}(\mathcal{P}) \ge \eta\sqrt{x\lambda t}) \le \exp(-\psi(\eta)(\sqrt{xt\lambda} - x\lambda))$. With minor modifications one proves the same bound for sinks (possibly by changing ψ): $\mathbb{P}(\mathsf{sinks}(\mathcal{P}) \ge \eta\sqrt{x\lambda t}) \le \exp(-\psi(\eta)(\sqrt{xt\lambda}-x\lambda))$ and Lemma 10 is proved. □

We can conclude the proof of the lower bound in Theorem 8. Let us write

$$L_<(t) \ge \mathcal{L}_{=<}(\Pi^{(\lambda)}_{x,t} \cup \mathsf{So}_x^{(\bar\alpha)} \cup \mathsf{Si}_t^{(\bar p)}) - \mathsf{sources}(\mathcal{P}) - \mathsf{sinks}(\mathcal{P}),$$

we bound the right-hand side using Lemmas 9 and 10. □

4 Proof of Theorem 1 when $k_n \to +\infty$: de-Poissonization

In order to conclude the proof of Theorem 1 it remains to de-Poissonize Theorem 8. We need a few notation. For any integers $i_1, \dots, i_n$ let $\mathcal{S}_{i_1,\dots,i_n}$ be the random set of points given by i_ℓ uniform points on each horizontal line:

$$\mathcal{S}_{i_1,\dots,i_n} = \cup_{\ell=1}^n \cup_{r=1}^{i_\ell} \{U_{\ell,r}\} \times \{\ell\},$$

where $(U_{\ell,r})_{\ell,r}$ is an array of i.i.d. uniform random variables in $[0,1]$. Set also $e_{i_1,\dots,i_n} = \mathbb{E}[\mathcal{L}_<(\mathcal{S}_{i_1,\dots,i_n})]$. By uniformity of U's we have the identity $\mathbb{E}[\mathcal{L}_<(S_{k;n})] = e_{k,\dots,k}$ and therefore our problem reduces to estimating $e_{k,\dots,k}$. On the other hand if $X_1,\dots,X_n$ are i.i.d. Poisson random variables with mean k then

$$\mathbb{E}[e_{X_1,\dots,X_n}] = \mathbb{E}\left[\mathcal{L}_<(\mathcal{P}^{(1/n)}_{nk_n,n})\right] = 2\sqrt{nk_n} - k_n + o(\sqrt{nk_n}). \tag{25}$$

The last equality is obtained by combining Theorem 8 for

$$x = nk_n, \qquad t = n, \qquad \lambda_n = \frac{1}{n}$$

with the trivial bound $\mathcal{L}_<(\mathcal{P}^{(1/n)}_{nk_n,n}) \le n$. In order to exploit (25) we need the following smoothness estimate.

Lemma 11. *For every $i_1,\dots,i_n$ and $j_1,\dots,j_n$*

$$|e_{i_1,\dots,i_n} - e_{j_1,\dots,j_n}| \le 6\sqrt{\sum_{\ell=1}^n |i_\ell - j_\ell|}.$$

Proof. Let $\mathcal{S} = \mathcal{S}_{i_1,\dots,i_n}$ be as above. If we replace in $\mathcal{S}$ the y-coordinate of each point of the form (x,ℓ) by a new y-coordinate uniform in the interval $(\ell,\ell+1)$ (independent from anything else) then this defines a uniform permutation $\sigma_{i_1+\dots+i_n}$ of size $i_1+\dots+i_n$. The longest increasing subsequence in $\mathcal{S}$ is mapped onto an increasing subsequence in $\sigma_{i_1+\dots+i_n}$ and thus this construction shows the stochastic domination $\mathcal{L}_<(\mathcal{S}_{i_1,\dots,i_n}) \preccurlyeq \mathcal{L}_<(\sigma_{i_1+\dots+i_n})$. Thus for every $i_1,\dots,i_n$,

$$e_{i_1,\dots,i_n} \le \mathbb{E}[\mathcal{L}_<(\sigma_{i_1+\dots+i_n})] \le 6\sqrt{i_1+\dots+i_n}. \tag{26}$$

(The last inequality follows from [Ste97, Lemma 1.4.1].) Besides, consider for two n-tuples $i_1,\dots,i_n$ and $j_1,\dots,j_n$ two independent set of points $\mathcal{S}_{i_1,\dots,i_n}$, $\widetilde{\mathcal{S}}_{j_1,\dots,j_n}$ then

$$\mathcal{L}_<(\mathcal{S}_{i_1,\dots,i_n}) \le \mathcal{L}_<(\mathcal{S}_{i_1,\dots,i_n} \cup \widetilde{\mathcal{S}}_{j_1,\dots,j_n}) \le \mathcal{L}_<(\mathcal{S}_{i_1,\dots,i_n}) + \mathcal{L}_<(\widetilde{\mathcal{S}}_{j_1,\dots,j_n}).$$

This proves that

$$e_{i_1,\dots,i_n} \le e_{i_1+j_1,\dots,i_n+j_n} \le e_{i_1,\dots,i_n} + e_{j_1,\dots,j_n}.$$

(In particular $(i_1,\dots,i_n) \mapsto e_{i_1,\dots,i_n}$ is non-decreasing with respect to any of its coordinate.) Therefore

$$\begin{aligned} e_{i_1,\dots,i_n} &\le e_{(i_1-j_1)^+,\dots,(i_n-j_n)^+} + e_{j_1-(i_1-j_1)^-,\dots,j_n-(i_n-j_n)^-} \\ &\le e_{|i_1-j_1|,\dots,|i_n-j_n|} + e_{j_1,\dots,j_n}. \end{aligned}$$

By switching the role of i's and j's:

$$|e_{i_1,\dots,i_n} - e_{j_1,\dots,j_n}| \le e_{|i_1-j_1|,\dots,|i_n-j_n|} \le 6\sqrt{\sum_{\ell=1}^n |i_\ell - j_\ell|},$$

using (26). □

Proof of Theorem 1 for any sequence $(k_n) \to +\infty$. Using smoothness we write

$$|e_{k,\dots,k} - \mathbb{E}[e_{X_1,\dots,X_n}]| \le \mathbb{E}\left[|e_{k,\dots,k} - e_{X_1,\dots,X_n}|\right] \le 6 \times \mathbb{E}\left[\left(\sum_{\ell=1}^n |X_\ell - k|\right)^{1/2}\right]. \tag{27}$$

Using twice the Cauchy-Schwarz inequality:

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{\ell=1}^{n}|X_\ell-k|\right)^{1/2}\right] &\le \sqrt{\mathbb{E}\left[\sum_{\ell=1}^{n}|X_\ell-k|\right]}\\ &\le \sqrt{n\mathbb{E}\left[|X_1-k|\right]}\\ &\le \sqrt{n\mathbb{E}\left[|X_1-k|^2\right]^{1/2}} = \sqrt{n\sqrt{\mathrm{Var}(X_1)}} = \sqrt{n\sqrt{k}}.\end{aligned}$$

If $k=k_n\to\infty$ then the last display is a $o(\sqrt{nk_n})$ and eq.(27) and (25) show that

$$e_{k,\dots,k} = \mathbb{E}[\mathcal{L}_<(S_{k;n})] = 2\sqrt{nk_n} - k_n + o(\sqrt{nk_n}).$$

□

5 Proof of Theorem 2

5.1 Proof for large (k_n)

We now prove Theorem 2 for a *large* sequence (k_n). We say that (k_n) is *large* if

$$n^2k_n\exp(-(k_n)^\alpha) = \mathrm{o}(\sqrt{nk_n}) \tag{28}$$

for some $\alpha\in(0,1)$. Note that $k_n=\log n$ is not large while $k_n=(\log n)^{1+\varepsilon}$ is large.

We first observe that de-Poissonization cannot be applied as in the previous section. We lack smoothness as, for instance, $\mathbb{E}[\mathcal{L}_\le(\mathcal{S}_{i_1,0,0,\dots,0})] = i_1 \ne \mathcal{O}(\sqrt{\sum i_\ell})$. The strategy is to apply Theorem 8 with

$$x = nk_n, \qquad t = n, \qquad \lambda_n \approx \frac{1}{n}.$$

(The exact value of λ_n will be different for the proofs of the lower and upper bounds.)

Proof of the upper bound of (2) for large (k_n).

Choose α such that $n^2k_n\exp(-k_n^\alpha) = \mathrm{o}(\sqrt{nk_n})$. Put

$$\lambda_n = \frac{1}{n} + \frac{\delta_n}{n}, \qquad \text{with } \delta_n = k_n^{-(1-\alpha)/2}.$$

Let $E_n^{\lambda_n}$ be the event

$$E_n^{\lambda_n} = \left\{ \text{ at least } k_n \text{ points in each row of } \mathcal{P}^{(\lambda_n)}_{nk_n,n} \right\}.$$

The event E_n occurs with large probability. Indeed,

$$\begin{aligned}1-\mathbb{P}(E_n^{\lambda_n}) &\le n\mathbb{P}\left(\mathrm{Poisson}(nk_n\lambda_n)\le k_n\right)\\ &\le n\mathbb{P}\left(\mathrm{Poisson}(nk_n\lambda_n)\le nk_n\lambda_n + k_n - nk_n\lambda_n\right)\\ &\le n\mathbb{P}\left(\mathrm{Poisson}(nk_n\lambda_n)\le nk_n\lambda_n - k_n\delta_n\right)\\ &\le n\exp\left(-\frac{k_n^2\delta_n^2}{4nk_n\lambda_n}\right) \le n\exp\left(-\tfrac{1}{8}k_n\delta_n^2\right) = n\exp\left(-\tfrac{1}{8}k_n^\alpha\right).\end{aligned} \tag{29}$$

At the last line we used Lemma 14. The latter probability tends to 0 as (k_n) is large.

Lemma 12. *Random sets $S_{k_n;n}$ and $\Pi^{(\lambda_n)}_{nk_n,n}$ can be defined on the same probability space in such a way that*

$$\mathcal{L}_\le(S_{k_n;n}) \le \mathcal{L}_\le(\Pi^{(\lambda_n)}_{nk_n,n}) + nk_n(1-\mathbf{1}_{E_n^{\lambda_n}}). \tag{30}$$

Proof of Lemma 12. Draw a sample of $\Pi^{(\lambda_n)}_{nk_n,n}$ and let $\tilde{\Pi}^{(\lambda_n)}_{nk_n,n}$ be the subset of $\Pi^{(\lambda_n)}_{nk_n,n}$ obtained by keeping only the k_n leftmost points in each row. If $E^{\lambda_n}_n$ occurs then the relative orders of points in $\tilde{\Pi}^{(\lambda_n)}_{nk_n,n}$ corresponds to a uniform k_n-multiset permutation. If $E^{\lambda_n}_n$ does not hold we bound $\mathcal{L}_{\leq}(S_{k_n;n})$ by the worst case nk_n. □

Taking expectations in (30) and using the upper bound (15) yields

$$\mathbb{E}[\mathcal{L}_{\leq}(S_{k_n;n})] \leq 2\sqrt{nk_n(1+\delta_n)} + k_n(1+\delta_n) + n^2 k_n \exp\left(-k_n^{\alpha}\right),$$

hence the upper bound in (2).

Proof of the lower bound of (2) **for large** (k_n)**.** Choose now $\lambda_n = \frac{1}{n}(1-\delta_n)$ with $\delta_n = k_n^{-(1-\alpha)/2}$. Let F_n be the event

$$F^{\lambda_n}_n = \left\{ \text{ at most } k_n \text{ points in each row of } \mathcal{P}^{(\lambda_n)}_{nk_n,n} \right\}.$$

The event $F^{\lambda_n}_n$ occurs with large probability:

$$1 - \mathbb{P}(F^{\lambda_n}_n) \leq n\mathbb{P}\left(\text{Poisson}(nk_n\lambda_n) \geq k_n\right) \leq n\exp\left(-\tfrac{1}{8}k_n^{\alpha}\right),$$

which tends to zero. Random sets $S_{k_n;n}$ and $\mathcal{P}^{(\lambda_n)}_{nk_n,n}$ can be defined on the same probability space in such a way that

$$\mathcal{L}_{\leq}(S_{k_n;n}) \geq \mathcal{L}_{\leq}(\mathcal{P}^{(\lambda_n)}_{nk_n,n})\mathbf{1}_{F^{\lambda_n}_n}.$$

Therefore

$$\mathbb{P}\left(\mathcal{L}_{\leq}(S_{k_n;n}) < (2\sqrt{nk_n(1-\delta_n)} + k_n(1-\delta_n))(1-\varepsilon)\right) \leq$$
$$\mathbb{P}\left(\mathcal{L}_{\leq}(\mathcal{P}^{(\lambda_n)}_{nk_n,n}) < (2\sqrt{nk_n(1-\delta_n)} + k_n(1-\delta_n))(1-\varepsilon)\right) + \mathbb{P}\left(\text{not } F^{\lambda_n}_n\right).$$

and we conclude with (19).

5.2 The gap between small and large (k_n)

After I circulated a preliminary version of this article, Valentin Féray came up with a simple argument for bridging the gap between small and large (k_n). This allows to prove Theorem 2 for an arbitrary sequence (k_n), I reproduce his argument here with his permission.

Lemma 13. *Let n, k, A be positive integers. Two random uniform multiset permutations $\widetilde{S}_{kA;\lfloor n/A \rfloor}$ and $S_{k;n}$ can be built on the same probability space in such a way that*

$$\mathcal{L}_{\leq}\left(S_{k;n}\right) \leq \mathcal{L}_{\leq}\left(\widetilde{S}_{kA;\lfloor n/A \rfloor}\right) + kA.$$

Proof of Lemma 13 . Draw $S_{k;n}$ uniformly at random, the idea is to group all points of $S_{k;n}$ whose height is between 1 and A, to group all points whose height is between $A+1$ and $2A$, and so on.

Formally, denote by $1 \leq i_1 < i_2 < \cdots < i_{kA\lfloor n/A \rfloor}$ the indices such that $1 \leq i_\ell \leq \lfloor n/A \rfloor$ for every ℓ (see Fig.4). For $1 \leq \ell \leq kA\lfloor n/A \rfloor$ put

$$\widetilde{S}(\ell) = \lceil S(i_\ell)/k \rceil.$$

The word $\widetilde{S}$ is a uniform kA-multiset permutation of size $\lfloor n/A \rfloor$. A longest non-decreasing subsequence in S is mapped onto a non-decreasing subsequence in $\widetilde{S}$, except maybe some points with height $> A\lfloor n/A \rfloor$ (there are no more than kA such points). This shows the Lemma. □

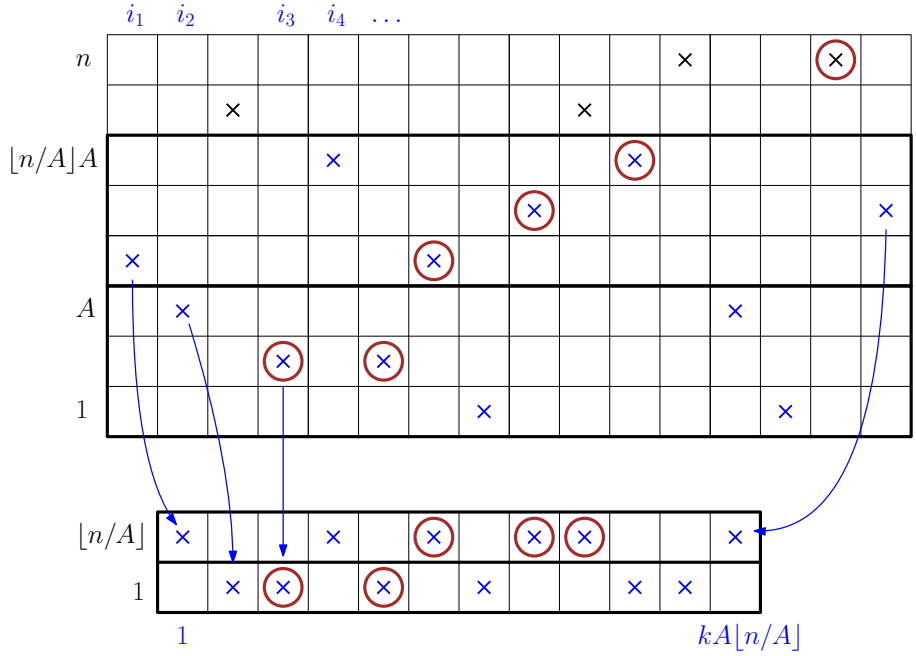


Figure 4: Illustration of the notation of Lemma 13. Top: the multiset permutation $S_{k;n}$. Bottom: the corresponding $\widetilde{S}$. The longest non-decreasing subsequence in $S_{k;n}$ (circled points) is mapped onto a non-decreasing subsequence in $\widetilde{S}$, except one point with height $> A\lfloor n/A \rfloor$.

We conclude the proof of Theorem 2 by an estimation of $\mathbb{E}[\mathcal{L}_{\leq}(S_{k_n;n})]$ in the case where there are infinitely many k_n's such that, say, $(\log n)^{3/4} \leq k_n \leq (\log n)^{5/4}$. For the lower bound the job is already done by Theorem 1 since

$$\mathbb{E}[\mathcal{L}_{\leq}(S_{k_n;n})] \geq \mathbb{E}[\mathcal{L}_{<}(S_{k_n;n})] = 2\sqrt{nk_n} - k_n + \mathrm{o}(\sqrt{nk_n}),$$

which of course also $2\sqrt{nk_n} + \mathrm{o}(nk_n)$ for this range of (k_n). For the upper bound take $A = \lfloor \log n \rfloor$ in Lemma 13:

$$\mathbb{E}[\mathcal{L}_{\leq}(S_{k_n;n})] \leq \mathbb{E}[\mathcal{L}_{\leq}\left(S_{k_n \log n;\lfloor n/\lfloor \log n \rfloor \rfloor}\right)] + k_n \log n \tag{31}$$

and we can apply the large case since

$$(n/\log n)^2 k_n \log n \exp(-(k_n \log n)^{\alpha}) = \mathrm{o}(k_n \log n \times \lfloor n/\lfloor \log n \rfloor \rfloor).$$

Thus the right-hand side of (31) is also $2\sqrt{nk_n} + \mathrm{o}(\sqrt{nk_n})$.

5.3 Deviation inequalities

We briefly explain here how to deduce from previous calculations deviation inequalities for $\mathcal{L}_{<}(S_{k_n;n})$ and $\mathcal{L}_{\leq}(S_{k_n;n})$ in the case where (k_n) is large. We only write the case of an upper bound for $L_{<}(S_{k_n;n})$ and do not aim at optimality. As in Section 5.1 choose $\lambda_n = \frac{1}{n} + \frac{\delta_n}{n}$ where $\delta_n = k_n^{-(1-\alpha)/2}$.

$$\begin{aligned}
\mathbb{P}\left(L_<\left(S_{k_n;n}\right) > (2\sqrt{nk_n} - k_n)(1+\varepsilon)\right) &\le \mathbb{P}\left(E_n^{\lambda_n} \text{ does not occur}\right) \\
&+ \mathbb{P}\left(\mathcal{L}_<\left(\Pi^{(\lambda_n)}_{nk_n,n}\right) > (2\sqrt{nk_n} - k_n)(1+\varepsilon)\right) \\
&\le n\exp\left(-\tfrac{1}{8}k_n^\alpha\right) \qquad \text{(using (29))} \\
&+ \mathbb{P}\left(\mathcal{L}_<\left(\Pi^{(\lambda_n)}_{nk_n,n}\right) > (1+\delta_n)(2\sqrt{nk_n} - k_n)\frac{1+\varepsilon}{1+\delta_n}\right) \\
&\le n\exp\left(-\tfrac{1}{8}k_n^\alpha\right) \\
&+ \mathbb{P}\left(\mathcal{L}_<\left(\Pi^{(\lambda_n)}_{nk_n,n}\right) > (2\sqrt{nk_n(1+\delta_n)} - k_n(1+\delta_n))\frac{1+\varepsilon}{1+\delta_n}\right) \\
&\le n\exp\left(-\tfrac{1}{8}k_n^\alpha\right) + \exp(-\tilde{g}(\varepsilon/2)(\sqrt{nk_n} - k_n)),
\end{aligned}$$

for large enough n and for some positive $\tilde{g}$, using (16).

Appendix: Useful tail inequalities

We collect here for convenience some (non-optimal) tail inequalities.

Lemma 14 ((See Chap.2 in [JŁR00])). *Let* Poisson(λ) *be a Poisson random variable with mean* λ. *For every* $A > 0$

$$\begin{aligned}
\mathbb{P}\left(\text{Poisson}(\lambda) \le \lambda - A\right) &\le \exp(-A^2/4\lambda), \\
\mathbb{P}\left(\text{Poisson}(\lambda) \ge \lambda + A\right) &\le \exp(-A^2/4\lambda).
\end{aligned}$$

Lemma 15 (Th.2.1 in [JŁR00]). *Let* Binomial(n,p) *be a Binomial random variable with parameters* (n,p). *For* $0 < \varepsilon < 1$,

$$\begin{aligned}
\mathbb{P}(\text{Binomial}(n,p) \le np + \varepsilon np) &\le \exp\left(-\varepsilon^2 np/2\right), \\
\mathbb{P}(\text{Binomial}(n,p) \ge np - \varepsilon np) &\le \exp\left(-\varepsilon^2 np/3\right).
\end{aligned}$$

Lemma 16. *Let* $\mathcal{G}_1^{(\alpha)}, \dots, \mathcal{G}_k^{(\alpha)}$ *be i.i.d. random variables with distribution* Geometric$_{\ge 0}(1-\alpha)$. *For* $0 < \varepsilon < 1$,

$$\begin{aligned}
\mathbb{P}\left(\mathcal{G}_1^{(\alpha)} + \cdots + \mathcal{G}_k^{(\alpha)} \ge (1+\varepsilon)k\frac{\alpha}{1-\alpha}\right) &\le \exp\left(-\tfrac{1}{4}\varepsilon^2 k\frac{\alpha}{1-\alpha}\right), \\
\mathbb{P}\left(\mathcal{G}_1^{(\alpha)} + \cdots + \mathcal{G}_k^{(\alpha)} \le (1-\varepsilon)k\frac{\alpha}{1-\alpha}\right) &\le \exp\left(-\tfrac{1}{4}\varepsilon^2 k\frac{\alpha}{1-\alpha}\right).
\end{aligned}$$

Proof of Lemma 16. Using $e^\lambda \le 1 + \lambda + \lambda^2$ for $|\lambda| < 1$ we have

$$\mathbb{E}[e^{\lambda(\mathcal{G}_1^{(\alpha)} - \frac{\alpha}{1-\alpha})}] = \frac{(1-\alpha)e^{-\lambda\frac{\alpha}{1-\alpha}}}{1-\alpha e^\lambda} \le \exp\left(-\lambda^2\frac{\alpha}{1-\alpha}\right).$$

This says that $\mathcal{G}_1^{(\alpha)}$ is subexponential and the Chernov method (see *e.g.* [Wai19, Prop.2.2] for the case of subexponential random variables) implies that

$$\begin{aligned}
\mathbb{P}\left(\mathcal{G}_1^{(\alpha)} + \cdots + \mathcal{G}_k^{(\alpha)} \ge (1+\varepsilon)k\frac{\alpha}{1-\alpha}\right) &\le \max\left\{\exp\left(-\tfrac{1}{4}\varepsilon^2 k^2\frac{\alpha}{1-\alpha}\right), \exp\left(-\tfrac{1}{2}\varepsilon k\frac{\alpha}{1-\alpha}\right)\right\} \\
&\le \exp\left(-\tfrac{1}{4}\varepsilon^2 k\frac{\alpha}{1-\alpha}\right).
\end{aligned}$$

□

Acknowledgements. This work started as a collaboration with Anne-Laure Basdevant, I would like to thank her very warmly. I am also extremely indebted to Valentin Féray for Lemma 13 and for having enlightened me on the links with [Bia01]. Finally, thanks to Sam Spiro for stimulating exchanges.

References

[AD95] David Aldous and Persi Diaconis. Hammersley's interacting particle process and longest increasing subsequences. *Probability Theory and Related Fields*, 103(2):199–213, 1995.

[BDJ99] Jinho Baik, Percy Deift, and Kurt Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.*, 12(4):1119–1178, 1999.

[BEGG16] Anne-Laure Basdevant, Nathanaël Enriquez, Lucas Gerin, and Jean-Baptiste Gouéré. Discrete Hammersley's lines with sources and sinks. *ALEA Lat. Am. J. Probab. Math. Stat.*, 13:33–52, 2016.

[Bia01] Philippe Biane. Approximate factorization and concentration for characters of symmetric groups. *Internat. Math. Res. Notices*, (4):179–192, 2001.

[Boy22] Alexandre Boyer. Chapter 3 (in English) of *Stationnarité bidimensionnelle de modèles aléatoires du plan*, 2022. PhD Thesis, available at `https://tel.archives-ouvertes.fr/tel-03783603/`.

[CDH+22] Alexander Clifton, Bishal Deb, Yifeng Huang, Sam Spiro, and Semin Yoo. Continuously increasing subsequences of random multiset permutations. *Sém. Lothar. Combin.*, 86B:Art. 4, 11, 2022. (Proceedings of FPSAC'22.).

[CG05] Eric Cator and Piet Groeneboom. Hammersley's process with sources and sinks. *Annals of Probability*, 33(3):879–903, 2005.

[CG06] Eric Cator and Piet Groeneboom. Second class particles and cube root asymptotics for Hammersley's process. *Annals of Probability*, 34(4):1273–1295, 2006.

[CG19] Federico Ciech and Nicos Georgiou. Order of the variance in the discrete Hammersley process with boundaries. *Journal of Statistical Physics*, 176(3):591–638, 2019.

[CLRS09] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to algorithms*. MIT Press, Cambridge, MA, third edition, 2009.

[Ham72] John M. Hammersley. A few seedlings of research. In *Proceedings of the 6th Berkeley Symp. Math. Statist. and Probability*, volume 1, pages 345–394, 1972.

[JŁR00] Svante Janson, Tomasz Łuczak, and Andrzej Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[Rom15] Dan Romik. *The surprising mathematics of longest increasing subsequences*, volume 4 of *Institute of Mathematical Statistics Textbooks*. Cambridge University Press, New York, 2015.

[Sep97] Timo Seppäläinen. Increasing sequences of independent points on the planar lattice. *Annals of Applied Probability*, 7(4):886–898, 1997.

[Ste97] John Michael Steele. *Probability theory and combinatorial optimization*. Society for Industrial and Applied Mathematics (SIAM), 1997.

[VK77] Anatoly M. Veršik and Sergei V. Kerov. Asymptotics of Plancherel measure of symmetrical group and limit form of Young tables. *Doklady Akademii Nauk SSSR*, 233.6:1024–1027, 1977.

[Wai19] Martin J. Wainwright. *High-dimensional statistics*, volume 48 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2019. A non-asymptotic viewpoint.

Lucas Gerin gerin@cmap.polytechnique.fr
Cmap, Cnrs, École Polytechnique,
Institut Polytechnique de Paris,
Route de Saclay,
91120 Palaiseau Cedex (France).