Math. Proc. Camb. Phil. Soc. (2024), **177**, 23–48
doi:10.1017/S0305004124000124
First published online 9 May 2024

The Ulam–Hammersley problem for multiset permutations

BY LUCAS GERIN
CMAP, CNRS, École Polytechnique, 91120 *Palaiseau, France.*
e-mail: gerin@cmap.polytechnique.fr

(*Received* 01 *September* 2023*; revised* 21 *March* 2024*; accepted* 20 *March* 2024)

Abstract

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

2020 Mathematics Subject Classification: 60K35 (Primary); 60F15 (Secondary)

1. *Introduction*

A k-multiset permutation of size n is a word with letters in $\{1, 2, \ldots, n\}$ such that each letter appears exactly k times. When this is convenient we identify a multiset permutation $s = (s(1), \ldots, s(kn))$ and the set of points $\{(i, s(i)),\ 1 \leq i \leq kn\}$. We introduce two partial orders over the quarter-plane $[0, \infty)^2$:

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

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

For a finite set $\mathcal{P}$ of points in the quarter-plane we put

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

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

In words the integer $\mathcal{L}_<(\mathcal{P})$ (resp. $\mathcal{L}_\leq(\mathcal{P})$) 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 drawn uniformly among the $(kn)!/k!^n$ possibilities. In the case $k = 1$ the word $S_{1;n}$ is simply a uniform permutation and estimating $\mathcal{L}_<(S_{1;n}) = \mathcal{L}_\leq(S_{1;n})$ is known as the Hammersley or Ulam–Hammersley problem. The first order was solved by Veršik and Kerov [**VK77**] and simultaneously by Logan and Shepp:

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

Note that the above limit also holds in probability: $\mathcal{L}_<(S_{1;n}) = 2\sqrt{n} + o_{\mathbb{P}}(\sqrt{n})$. This problem has a long history and has revealed deep and unexpected connections between

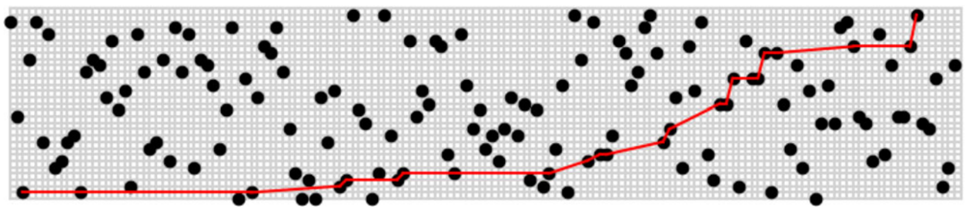


Fig. 1. A uniform 5-multiset permutation $S_{5;30}$ of size $n = 30$ and one of its longest non-decreasing subsequences.

combinatorics, interacting particle systems, calculus of variations, random matrix theory, representation theory. We refer to Romik [**Rom15**] for a very nice description of this problem and some of its ramifications.

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 (see Fig. 1 for an example). Using the Veršik–Kerov Theorem we can make an educated guess. The intuition is that, for fixed k, it is quite unlikely that many points at the same height contribute to the same longest increasing/non-decreasing subsequence. Thus at the first order everything should happen as if the kn points had distinct heights and we expect that

$$\mathcal{L}_<(S_{k;n}) \approx \mathcal{L}_\leq(S_{k;n}) \approx \mathcal{L}_<(S_{1;kn}) \approx 2\sqrt{kn}.$$

The original motivation of this paper was to make this approximation rigorous. 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 $2\sqrt{nk_n} + \mathrm{o}(\sqrt{nk_n})$.)

Remark 1. If $k_n \geq n$ for some n then the following greedy strategy shows that $\mathbb{E}[\mathcal{L}_<(S_{k_n;n})] = n - \mathrm{o}(n)$ so the picture is complete.

Indeed, first choose the leftmost point $(x_1, 1)$ in $S_{k_n;n}$ which has height 1. Then recursively define (x_ℓ, ℓ) at the leftmost point (if any) in $S_{k_n;n}$ with height ℓ such that $x_\ell > x_{\ell-1}$, and so on until you are stuck (either because $\ell = n$ or because there is no point in $S_{k_n;n} \cap (x_{\ell-1}, kn] \times \{\ell\}$). A few elementary computations show that this strategy defines an increasing path of length $n - \mathrm{o}(n)$ with probability tending to one. As $\mathcal{L}_<(S_{k_n;n}) \leq n$ a.s. this yields $\mathbb{E}[\mathcal{L}_<(S_{k_n;n})] = n - \mathrm{o}(n)$.

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

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

Strategy of proof and organisation 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 the Veršik–Kerov Theorem is known).

For the general case we first borrow a few tools in the literature. In particular we introduce and analyse poissonised versions of $\mathcal{L}_<(S_{k_n;n})$, $\mathcal{L}_\leq(S_{k_n;n})$. As already suggested by Hammersley ([**Ham72**, section 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 (HAD) process.

In Section 3 we introduce and analyse the two variants of the Hammersley process adapted to multiset permutations. The first one is the discrete-time HAD process [**Fer96, FM06**], the second one appeared in [**Boy22**] with a connection to the O'Connell–Yor Brownian polymer. The standard path to analyse Hammersley-like processes 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 either using convex duality [**Sep97, CG19**] or through the analysis of *second class particles* [**CG06, 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 9 whose proof is the most technical part of the paper. In Section 4 we detail the multivariate de-poissonisation procedure in order to conclude the proof of Theorem 1. De-poissonisation 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 poissonised version of $\mathcal{L}_<(S_{k_n;n})$, $\mathcal{L}_\leq(S_{k_n;n})$: see Theorem 9 and also the discussion in Section 6. We also obtain the convergence in probability, unfortunately for some technical reasons we miss a small range of scales of (k_n)'s.

PROPOSITION 3. *Let (k_n) be either a* small *or a* large *sequence. Then*

$$\frac{\mathcal{L}_<(S_{k_n;n})}{2\sqrt{nk_n}-k_n} \overset{prob.}{\to} 1, \qquad \frac{\mathcal{L}_\leq(S_{k_n;n})}{2\sqrt{nk_n}+k_n} \overset{prob.}{\to} 1.$$

We refer to (3),(31) below for the formal definitions of small/large sequences. Let us just say that sequences such that $k_n=\mathcal{O}((\log n)^{1-\varepsilon})$ for some $\varepsilon>0$ are small while sequences such that $(\log n)^{1+\varepsilon}=\mathcal{O}(k_n)$ are large. Sequences in-between are neither small nor large so in Proposition 3 we miss scales like $k_n \approx \log(n)$.

Regarding fluctuations a famous result by Baik, Deift and Johansson [**BDJ99**, theorem 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. The intuition given by the comparison with the Hammersley process would suggest that the fluctuations of $\mathcal{L}_<(S_{k_n;n})$, $\mathcal{L}_\leq(S_{k_n;n})$ might be of order $(k_n n)^{1/6}$ as long as (k_n) does not grow too fast. A natural question to explore for furthering this work would involve understanding for which (k_n) the model preserves KPZ scaling exponents. The non-asymptotic estimates of Section 3 could serve as a first step in this direction.

Comparison with previous works. There are only few random sets $\mathcal{P}$ for which the asymptotics of $\mathcal{L}_<(\mathcal{P})$, $\mathcal{L}_\leq(\mathcal{P})$ are known:

(i) as already mentioned, the case of a uniform permutation (and its poissonised version) is very well understood, via different approaches. For proofs close to the spirit of the present paper, we refer to [**AD95**] and [**CG05**];

(ii) the case where $\mathcal{P}$ is given by a field of i.i.d. Bernoulli random variables on the square grid has been solved by Seppäläinen in [**Sep97**] for $\mathcal{L}_<$ and in [**Sep98**] for $\mathcal{L}_\leq$. (See [**BEGG16**] for an elementary proof of both results.)

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]).

We need a few notations to describe his result. Let $\mathcal{W}_{q_N;N}$ be the random word given by of q_N i.i.d. uniform letters in $\{1, 2, \dots, N\}$. The word $\mathcal{W}_{q_N;N}$ is not a multiset permutation but since for large N there are in average q_N/N points on each horizontal line of $\mathcal{W}_{q_N;N}$ we expect that $\mathcal{L}_<(\mathcal{W}_{q_N;N}) \approx \mathcal{L}_<(S_{q_N/N;N})$ and $\mathcal{L}_\leq(\mathcal{W}_{q_N;N}) \approx \mathcal{L}_\leq(S_{q_N/N;N})$.

Biane obtains the exact limiting shape of the random Young Tableau induced through the RSK correspondence by $\mathcal{W}_{q_N;N}$ in the regime where $\sqrt{q_N}/N \to c$ for some constant $c > 0$. 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 consequence of ([**Bia01**, theorem 3]) is that, in probability,

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

For that regime our Theorems 1 and 2 respectively suggest:

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

which is indeed consistent with Biane's result.

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

$$\mathcal{L}_<(S_{k;n}) \leq \mathcal{L}_\leq(\sigma_{kn}) \leq \mathcal{L}_\leq(S_{k;n}). \tag{4}$$

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 proves the small case of Proposition 3 and Theorems 1 and 2. For this purpose we introduce for every $\delta > 0$ the event

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

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 $1 \leq i_1 < \ldots < i_\ell \leq n$ 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 $r_{1,1} < \ldots < r_{1,m_1} < r_{2,1} < r_{2,m_1} < \ldots < r_{\ell,1} < \ldots < r_{1,m_\ell}$.

We then introduce the event (Fig. 2)

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

By the union bound (we skip the integer parts)

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

Using that

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

we obtain

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

Bounding each factor $(k - m_i)!$ by 1 we get

$$\sum_{(r_{i,j})_{i\leq\ell, j\leq m_i}} \mathbb{P}(F) \leq \frac{(k!)^\ell}{(\delta\sqrt{n})^{\ell+1}(\delta\sqrt{n} - \ell)!(\ell - 1)!}.$$

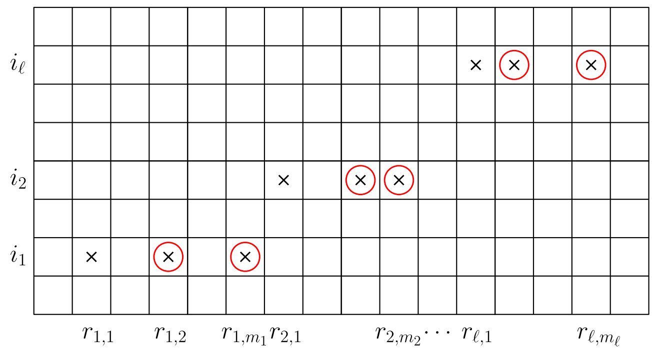


Fig. 2. The event F: a subsequence with $\delta\sqrt{n}$ (ties are surrounded) is depicted with $\times$'s.

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

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

Using the two following inequalities valid for every $j \le m$ (see *e.g.* [**CLRS09**, equation (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

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

which tends to zero for every $\delta > 0$, as long as (k_n) satisfies (3). This proves that $\mathcal{L}_{\le}(S_{k;n}) = \mathcal{L}_{<}(S_{k;n}) + o_{\mathbb{P}}(\sqrt{kn})$. Combining this with (4), this proves that

$$\frac{\mathcal{L}_{<}(S_{k_n;n})}{2\sqrt{nk_n}} \overset{\text{prob.}}{\to} 1, \qquad \frac{\mathcal{L}_{\le}(S_{k_n;n})}{2\sqrt{nk_n}} \overset{\text{prob.}}{\to} 1,$$

which is the "small" case of Proposition 3 since $k_n = \mathrm{o}(\sqrt{nk_n})$.

To conclude the proof of small cases of Theorems 1 and 2 we observe that we have the crude bounds $\mathcal{L}_{<}(S_{k;n}) \le n$ and $\mathcal{L}_{\le}(S_{k;n}) \le nk_n$. This allows us to write

$$\mathbb{E}\left[\left|\mathcal{L}_{\le}(S_{k;n}) - \mathcal{L}_{<}(S_{k;n})\right|\right] \le \delta\sqrt{n} + nk_n \times \mathbb{P}(\text{not } \mathcal{E}_\delta).$$

Together with equation (6) this implies that

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

We use again Veršik–Kerov and (4) to deduce that both sides are $2\sqrt{nk_n} + \mathrm{o}(\sqrt{nk_n})$.

3. *Poissonisation*: *variants of the Hammersley process*

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

Remark 2. 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).

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

3·1. *Definitions of the processes $L_{<}(t)$ and $L_{\le}(t)$*

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

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

The goal of this section is to obtain non-asymptotic bounds for $\mathcal{L}_{<}\left(\Pi_{x,t}^{(\lambda)}\right)$ and $\mathcal{L}_{\le}\left(\Pi_{x,t}^{(\lambda)}\right)$. Indeed if we then choose

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

then there are $kn + \mathcal{O}(\sqrt{kn})$ points on each line of a $\Pi_{x,t}^{(\lambda)}$ and we expect that

$$\mathcal{L}_{<}\left(\Pi_{kn,n}^{(\lambda_n)}\right) \approx \mathcal{L}_{<}(S_{k;n}), \qquad \mathcal{L}_{\le}\left(\Pi_{kn,n}^{(\lambda_n)}\right) \approx \mathcal{L}_{\le}(S_{k;n}).$$

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

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

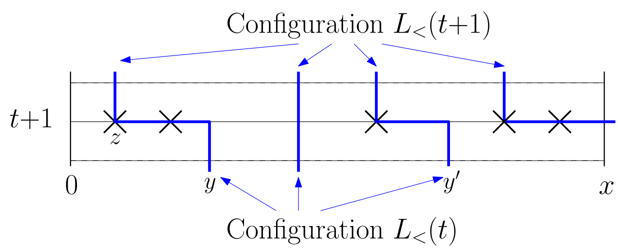
(Respectively:

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

Sets $L_<(t)$ and $L_\le(t)$ 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 $(L_\le(t))_t$ 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**, section 9], [**AD95**]) the individual dynamic of particles is very easy to describe:

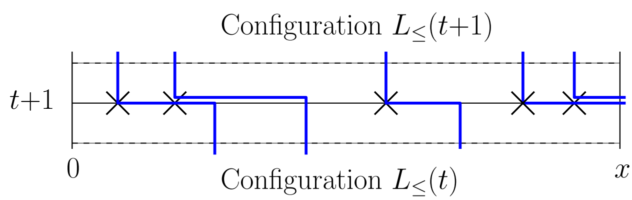
The process $L_<$. We put $L_<(0) = \emptyset$. In order to define $L_<(t+1)$ from $L_<(t)$ we consider particles from left to right. A particle at y in $L_<(t)$ moves at time $t+1$ at the location of the leftmost available point z in $\Pi^{(\lambda)}_{t+1} \cap (0,y)$ (if any, otherwise it stays at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



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. 3.

The process $L_\le$. We put $L_\le(0) = \emptyset$. In order to define $L_\le(t+1)$ from $L_\le(t)$ we also consider particles from left to right. A particle at y in $L_\le(t)$ moves at time $t+1$ at the location of the leftmost available point z in $\Pi^{(\lambda)}_{t+1} \cap (0,y)$. 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 $L_\le$ is shown in top-right of Figure 3.

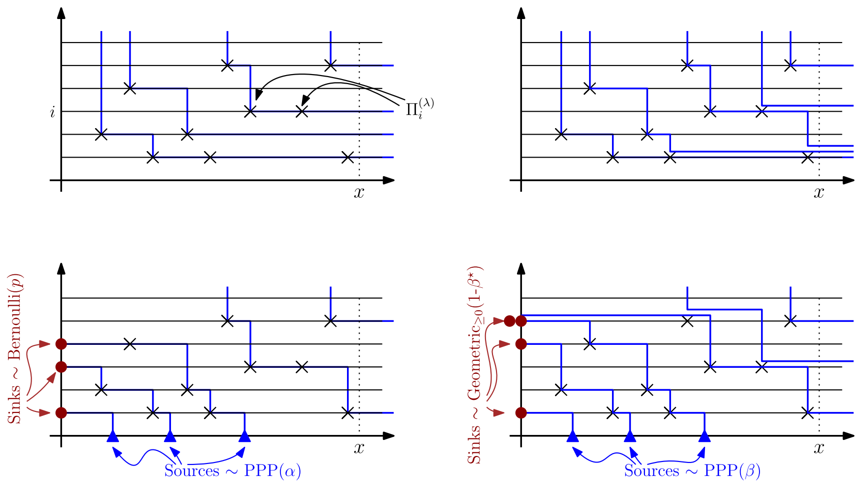


Fig. 3. Our four variants of the Hammersley process (time goes from bottom to top). Top left: the process $L_<(t)$. Top right: the process $L_\leq(t)$. Bottom left: the process $L_<^{(\alpha,p)}(t)$. Bottom right: the process $L_\leq^{(\beta,\beta^\star)}(t)$.

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 $(x,t) \mapsto \mathcal{L}_<\left(\Pi_{x,t}^{(\lambda)}\right)$, $(x,t) \mapsto \mathcal{L}_\leq\left(\Pi_{x,t}^{(\lambda)}\right)$.

PROPOSITION 4. *For every x,*

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

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**, section 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 $\Pi_{x,t}^{(\lambda)}$ cannot belong to the same Hammersley line. Since there are $L_<(t)$ Hammersley's lines this gives $\mathcal{L}_<\left(\Pi_{x,t}^{(\lambda)}\right) \leq \mathrm{card}(L_<(t))$.

In order to prove the converse inequality we build from this graphical construction a longest increcreasing subsequence of $\Pi_{x,t}^{(\lambda)}$ 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 $\Pi_{x,t}^{(\lambda)}$ belonging to the last Hammersley line. We then proceed by induction: we choose the next point among the points of of $\Pi_{x,t}^{(\lambda)}$ 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 $\mathcal{L}_<\left(\Pi_{x,t}^{(\lambda)}\right) \geq \mathrm{card}(L_<(t))$.

3·2. *Sources and sinks: stationarity*

Proposition 4 tells us that in on our way to prove Theorem 1 and Theorem 2 we need to understand the asymptotic behaviour of processes $L_<, L_\leq$.

It is proved in [**FM06**] that the homogeneous PPP with intensity α on $\mathbb{R}$ is stationary for $(L_<(t))_t$. However we need non-asymptotic estimates for $(L_<(t))_t$ (and $(L_\leq(t))_t$) on a given interval $(0, x)$. To solve this issue we use the trick of *sources/sinks* introduced formally and exploited by Cator and Groeneboom [**CG05**] for the continuous HAD process:

> *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 Figure 3.

Here is the discrete-time analogous of [**CG05**, theorem 3·1]:

LEMMA 5. *For every $\lambda, \alpha > 0$ let $L_<^{(\alpha,p)}(t)$ be the Hammersley process defined $L_<(t)$ with*:

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

(ii) *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 6. *For every $\beta > \lambda > 0$, let $L_\leq^{(\beta,\beta^\star)}(t)$ be the Hammersley process defined like $L_\leq(t)$ with additional sources and sinks*:

(i) *sources distributed according to a homogeneous PPP with intensity β on* $[0,x]\times\{0\}$;

(ii) *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 5 and 6. Lemma 6 could be obtained from minor adjustments of [**Boy22**, chapter 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 which explains where (8) come from.

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) (Fig. 4).

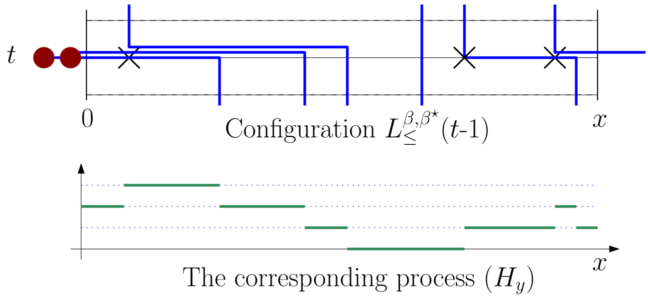


Fig. 4. A sample of the process H.

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. Computations given in Appendix B show that this is distributed as a homogeneous PPP with intensity β.

Lemma 5 is proved exactly in the same way, calculations are even easier. In this case the corresponding process $(H_y)_{0\leq y\leq x}$ takes its values in $\{0, 1\}$ and its stationary distribution is the Bernoulli distribution with mean $\lambda/(\alpha+\lambda)$, hence (7).

3·3. *Processes $L_<(t)$ and $L_\leq(t)$: non-asymptotic bounds*

From Lemmas 5 and 6 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. In particular,

$$\mathrm{card}(\mathsf{So}_x^{(\alpha)}) \stackrel{(\mathrm{d})}{=} \mathrm{Poisson}(\alpha x), \qquad \mathrm{card}(\mathsf{Si}_t^{(p)}) \stackrel{(\mathrm{d})}{=} \mathrm{Binomial}(t, p).$$

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,

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

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 4 generalises easily to the settings of sinks and sources.

Claim 1.

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

Proof of the Claim. By the same reasoning as in the proof of Proposition 4 the LHS is exactly the number of broken lines in the box $[0,x]\times[0,t]$. Each such line escapes the box either through the left (it thus corresponds to a sink) or through the top (and is thus counted by $L^{(\alpha,p)}_{<}(t)$).

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

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

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

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

$$\begin{aligned}
\mathcal{L}_{<}\left(\Pi^{(\lambda)}_{x,t}\right) &\preccurlyeq \mathcal{L}_{=<}\left(\Pi^{(\lambda)}_{x,t}\cup \mathsf{So}^{(\alpha)}_x \cup \mathsf{Si}^{(p)}_t\right)\\
&= L^{(\alpha,p)}_{<}(t) + \mathrm{card}(\mathsf{Si}^{(p)}) \text{ (using (9))}\\
&\overset{(\mathrm{d})}{=} L^{(\alpha,p)}_{<}(0) + \mathrm{card}(\mathsf{Si}^{(p)}) \text{ (using stationarity: Lemma 5)}\\
&\overset{(\mathrm{d})}{=} \mathrm{Poisson}(x\alpha) + \mathrm{Binomial}(t,p).
\end{aligned}$$

Taking expectations in (10) we obtain

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

The LHS in the above equation does not depend on α, p so the idea is to apply (10) with the minimising choice

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

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

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

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

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

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

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

We put

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

i.e.

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

(In particular $\bar{\beta} > \lambda$, as required in Lemma 6.) Equation (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 9 (Concentration for $\mathcal{L}_<$, $\mathcal{L}_\leq$). *There exist strictly positive functions g, h such that for all* $\varepsilon > 0$ *and for every* $x, t \geq 1$, $\lambda > 0$ *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 9 we will focus on the case of $\mathcal{L}_<$, *i.e.* Equations (16), (17). When necessary we will give the slight modification needed to prove Equations (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 minimising sources and sinks.

LEMMA 10 (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* $x, t \geq 1$, $\lambda > 0$ *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 10. By stationarity (Lemma 5) we have

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

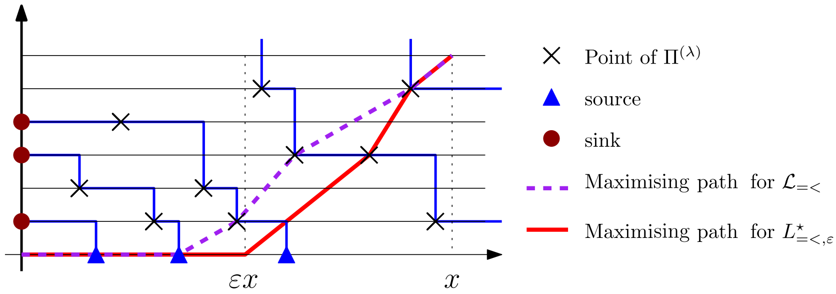


Fig. 5. A sample of $\Pi^{(\lambda)}_{x,t}$, sources, sinks, and the corresponding trajectories of particles. Here $\mathcal{L}_{=<}(\Pi^{(\lambda)}_{x,t} \cup \mathsf{So}^{(\alpha)}_x \cup \mathsf{Si}^{(p)}_t) = 5$ and $L^{(\alpha,p)}_<(t) = 2$ (two remaining particles at the top of the box).

Then

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

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 15):

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

Using the tail inequality for the binomial (Lemma 16) we get

$$\mathbb{P}\Big(\mathrm{Binomial}(t,\bar p) > (1+\frac{\varepsilon}{2})\sqrt{xt\lambda}\Big) \le \exp(-\tfrac{1}{12}\varepsilon^2\sqrt{xt\lambda}) \le \exp(-\tfrac{1}{12}\varepsilon^2(\sqrt{xt\lambda} - x\lambda)) \tag{22}$$

The proof of (21) is identical. This shows Lemma 10 with $g_1(\varepsilon) = \varepsilon^2/12$.

For longest non-decreasing subsequences we have a statement similar to Lemma 10. The only modification in the proof is that in order to estimate the number of sinks one has to replace Lemma 16 (tail inequality for the Binomial) by Lemma 17 (tail inequality for a sum of geometric random variables[1]). 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 Equations (18) and (19).

Proof of Theorem 9. 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 10.

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

[1] Note that it is only stated for $0 < \varepsilon < 1$ but this is enough for our purpose since the left-hand side of (22) is non-increasing in ε.

(see Figure 5). Formally we set

$$L^{\star}_{=<,\varepsilon} := \operatorname{card}\left(\mathsf{So}^{(\bar{\alpha})}_{\varepsilon x}\right) + \mathcal{L}_{<}\left((\Pi^{(\lambda)}_{x,t} \cap ([\varepsilon x, x] \times [0,t])\right)$$
$$\overset{(\mathrm{d})}{=} \mathrm{Poisson}(\varepsilon x\bar{\alpha}) + \mathcal{L}_{<}\left((\Pi^{(\lambda)}_{x,t} \cap ([\varepsilon x, x] \times [0,t])\right). \tag{23}$$

The idea is that for any fixed ε the paths contributing to $L^{\star}_{=<,\varepsilon}$ will typically not contribute to $\mathcal{L}_{=<}\left(\Pi^{(\lambda)}_{x,t} \cup \mathsf{So}^{(\bar{\alpha})}_{x} \cup \mathsf{Si}^{(\bar{p})}_{t}\right) = L^{(\bar{\alpha},p)}_{<}(t) + \operatorname{card}(\mathsf{Si}^{(\bar{p})}_{t})$. Indeed Equation (23) 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

$$2\sqrt{x\lambda t} - x\lambda - \tfrac{1}{2}\sqrt{xt\lambda}\,\delta(\varepsilon) = x\varepsilon\bar{\alpha} + \tfrac{1}{4}\sqrt{xt\lambda}\,\delta(\varepsilon) + 2\sqrt{x(1-\varepsilon)\lambda t} - x(1-\varepsilon)\lambda + \tfrac{1}{4}\sqrt{xt\lambda}\,\delta(\varepsilon). \tag{24}$$

Combining (23) and (24) gives

$$\mathbb{P}\left(L^{\star}_{=<,\varepsilon} \geq 2\sqrt{x\lambda t} - x\lambda - \tfrac{1}{2}\sqrt{xt\lambda}\,\delta(\varepsilon)\right) \leq \mathbb{P}_1 + \mathbb{P}_2,$$

where

$$\mathbb{P}_1 = \mathbb{P}\left(\mathrm{Poisson}(x\varepsilon\bar{\alpha}) \geq x\varepsilon\bar{\alpha} + \tfrac{1}{4}\sqrt{xt\lambda}\,\delta(\varepsilon)\right),$$

$$\mathbb{P}_2 = \mathbb{P}\left(\mathcal{L}_{<}(\Pi^{(\lambda)}_{x,t} \cap ([\varepsilon x, x] \times [0,t]) \geq 2\sqrt{x(1-\varepsilon)\lambda t} - x(1-\varepsilon)\lambda + \tfrac{1}{4}\sqrt{xt\lambda}\,\delta(\varepsilon)\right).$$

Using the tail inequality for the Poisson distribution (see Lemma 15) we have that

$$\mathbb{P}_1 \leq \exp\left(-\frac{xt\lambda\delta(\varepsilon)^2}{16 \times 4\varepsilon^2(\sqrt{xt\lambda} - x\lambda)}\right) \leq \exp\left(-\sqrt{xt\lambda}\,\delta(\varepsilon)^2/64\varepsilon^2\right).$$

Besides

$$\begin{aligned} \mathbb{P}_2 &\leq \mathbb{P}\left(\mathcal{L}_{<}(\Pi^{(\lambda)}_{x,t} \cap ([\varepsilon x, x] \times [0,t]) \geq \left(2\sqrt{x(1-\varepsilon)\lambda t} - x(1-\varepsilon)\lambda\right) \times (1 + \tfrac{1}{8}\delta(\varepsilon))\right) \\ &\leq \exp\left(-g(\delta(\varepsilon)/8)(\sqrt{x(1-\varepsilon)t\lambda} - x(1-\varepsilon)\lambda)\right) \text{ (using the upper bound (16)).} \end{aligned}$$

Finally we can find some positive h such that

$$\mathbb{P}\left(L^{\star}_{=<,\varepsilon} \geq 2\sqrt{x\lambda t} - x\lambda - \tfrac{1}{2}\sqrt{xt\lambda}\,\delta(\varepsilon)\right) \leq \exp\left(-h(\varepsilon)(\sqrt{xt\lambda} - x\lambda)\right). \tag{25}$$

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 $\mathcal{P}$:

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

In Figure 5 the path $\mathcal{P}$ is sketched, in that example $\mathsf{sources}(\mathcal{P}) = 2$, $\mathsf{sinks}(\mathcal{P}) = 0$.

LEMMA 11. *There exists a positive function ψ such that for all real $\eta > 0$*

$$\mathbb{P}\Big(\mathsf{sources}(\mathcal{P}) + \mathsf{sinks}(\mathcal{P}) \ge \eta\sqrt{x\lambda t}\Big) \le 2\exp(-\psi(\eta)(\sqrt{x\lambda t} - x\lambda)).$$

Proof of Lemma 11. (As the left-hand side is non-increasing in η it is enough to prove the lemma for $\eta < 1$.)

If the event $\Big\{\mathsf{sources}(\mathcal{P}) \ge \eta\sqrt{x\lambda t}\Big\}$ holds then there exists a (random) ε such that the two following events occur:

$$\mathsf{So}_{\varepsilon x} \ge \eta\sqrt{x\lambda t};$$

$$L^{\star}_{=<,\varepsilon} = \mathcal{L}_{=<}\Big(\Pi^{(\lambda)}_{x,t} \cup \mathsf{So}^{(\bar\alpha)}_{x} \cup \mathsf{Si}^{(\bar p)}_{t}\Big) = L^{(\bar\alpha,\bar p)}_{<}(t) + \text{card}(\mathsf{Si}^{(\bar p)}_{t}).$$

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:

$$\begin{aligned}\mathbb{P}(\text{sources}(\mathcal{P}) \ge \eta\sqrt{x\lambda t}) \le\ & \mathbb{P}(\mathsf{So}_{\eta x/2} \ge \eta\sqrt{x\lambda t})\\ & +\mathbb{P}(\mathsf{sources}(\mathcal{P}) \ge \eta\sqrt{x\lambda t};\ \mathsf{So}_{\eta x/2} < \eta\sqrt{x\lambda t}).\end{aligned}$$

Therefore

$$\begin{aligned}\mathbb{P}(\mathsf{sources}(\mathcal{P}) \ge \eta\sqrt{x\lambda t}) \le\ & \mathbb{P}(\mathsf{So}_{\eta x/2} \ge \eta\sqrt{x\lambda t})\\ +\ & \mathbb{P}(L^{(\bar\alpha,\bar p)}_{<}(t) \le \sqrt{x\lambda t} - x\lambda - \tfrac14\delta(\eta/3)\sqrt{x\lambda t})\\ +\ & \mathbb{P}(\text{card}(\mathsf{Si}^{(\bar p)}_{t}) \le \sqrt{x\lambda t} - \tfrac14\delta(\eta/3)\sqrt{x\lambda t})\\ +\ & \mathbb{P}(L^{\star}_{=<,\varepsilon} \ge 2\sqrt{x\lambda t} - x\lambda - \tfrac12\delta(\eta/3)\sqrt{x\lambda t} \text{ for some } \eta/2 \le \varepsilon \le 1).\end{aligned}$$

Let us call the four terms in the right-hand of the above display $\mathbb{P}_3, \mathbb{P}_4, \mathbb{P}_5, \mathbb{P}_6$ respectively.

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 see why:

(i) we bound $\mathbb{P}_3$ with Lemma 15 again (recall $\mathsf{So}_{\eta x/2}$ is a Poisson random variable);

(ii) the term $\mathbb{P}_4$ is bounded thanks to Lemma 10 (recall also (9));

(iii) we bound $\mathbb{P}_5$ with Lemma 16 (recall that $\mathsf{Si}^{(\bar p)}_{t}$ is a Binomial).

To conclude the proof it remains to bound $\mathbb{P}_6$. Let K be an integer larger than $144/\eta^3$, by definition of $L^{\star}_{=<,\varepsilon}$ we have for every $1 \le k \le \lceil xK \rceil$ and every $\varepsilon \in [k/K, (k+1)/K)$

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

Thus

$$\begin{aligned}
&\mathbb{P}\Big(\bigcup_{\eta/2\le\varepsilon\le 1}\Big\{L^{\star}_{=<,\varepsilon} > 2\sqrt{x\lambda t} - x\lambda - \tfrac{1}{2}\delta(\eta/3)\sqrt{x\lambda t}\Big\}\Big)\\
&\qquad \le \sum_{k\ge\lfloor\eta K/2\rfloor}\mathbb{P}\Big(L^{\star}_{=<,k/K} > 2\sqrt{x\lambda t} - x\lambda - \delta(\eta/3)\sqrt{x\lambda t}\Big)\\
&\qquad\quad + \sum_{k\ge\lfloor\eta K/2\rfloor}\mathbb{P}\Big(\mathrm{card}(\mathsf{So}_x^{(\bar\alpha)}\cap[\tfrac{k}{K},\tfrac{k+1}{K}]) > \tfrac{1}{2}\delta(\eta/3)\sqrt{x\lambda t}\Big)\\
&\qquad \le \sum_{k\ge\lfloor\eta K/2\rfloor}\mathbb{P}\Big(L^{\star}_{=<,k/K} > 2\sqrt{x\lambda t} - x\lambda - \delta(k/K)\sqrt{x\lambda t}\Big)\\
&\qquad\quad + \sum_{k\ge\lfloor\eta K/2\rfloor}\mathbb{P}\Big(\mathrm{card}(\mathsf{So}_x^{(\bar\alpha)}\cap[\tfrac{k}{K},\tfrac{k+1}{K}]) > \tfrac{1}{2}\delta(\eta/3)\sqrt{x\lambda t}\Big).
\end{aligned}$$

In the last inequality we use the facts that $K > 144/\eta^3 > 6/\eta$ and that δ is increasing. Using now (25) it holds that

$$\begin{aligned}
&\mathbb{P}\Big(\bigcup_{\eta/2\le\varepsilon\le 1}\Big\{L^{\star}_{=<,\varepsilon} > 2\sqrt{x\lambda t} - x\lambda - \tfrac{1}{2}\delta(\eta/3)\sqrt{x\lambda t}\Big\}\Big)\\
&\qquad \le \sum_{k\ge\lfloor\eta K/2\rfloor}\exp(-h(k/K)(\sqrt{xt\lambda}-x\lambda))\\
&\qquad + K\times\mathbb{P}\Big(\mathrm{Poisson}(\bar\alpha/K) > \tfrac{1}{2}\delta(\eta/3)\sqrt{x\lambda t}\Big)\\
&\qquad \le K\exp(-h(\eta/3)(\sqrt{xt\lambda}-x\lambda)) \qquad\qquad (26)\\
&\qquad + K\times\mathbb{P}\Big(\mathrm{Poisson}(\bar\alpha/K) > \tfrac{1}{2}\delta(\eta/3)\sqrt{x\lambda t}\Big).
\end{aligned}$$

We finally bound the last display. First recall from our notation that

$$\bar\alpha < \sqrt{t\lambda/x}, \qquad x\ge 1, \qquad \delta(\varepsilon) = 2-\varepsilon-2\sqrt{1-\varepsilon} \ge \varepsilon^2/4.$$

Then:

$$\begin{aligned}
\mathbb{P}\Big(\mathrm{Poisson}(\bar\alpha/K) > \tfrac{1}{2}\delta(\eta/3)\sqrt{x\lambda t}\Big) &= \mathbb{P}\Big(\mathrm{Poisson}(\bar\alpha/K) > \bar\alpha/K - \bar\alpha/K + \tfrac{1}{2}\delta(\eta/3)\sqrt{x\lambda t}\Big)\\
&\le \mathbb{P}\Big(\mathrm{Poisson}(\bar\alpha/K) > \bar\alpha/K + \sqrt{x\lambda t}\Big(-\frac{1}{xK} + \frac{1}{2}\delta(\eta/3)\Big)\Big)\\
&\le \mathbb{P}\Big(\mathrm{Poisson}(\bar\alpha/K) > \bar\alpha/K + \sqrt{x\lambda t}\Big(-\frac{\eta^3}{144} + \frac{\eta^2}{72}\Big)\Big). \qquad (27)
\end{aligned}$$

We can find a positive function φ such that (26) and (27) are both less than $(144/\eta)e^{-\varphi(\eta)(\sqrt{xt\lambda}-x\lambda)}$. We then choose a positive function ψ such that

$$\min\Big\{1, \frac{288}{\eta}e^{-\varphi(\eta)(\sqrt{xt\lambda}-x\lambda)} + 3e^{-\phi(\eta)(\sqrt{x\lambda t}-x\lambda)}\Big\} \le 2e^{-\psi(\eta)(\sqrt{xt\lambda}-x\lambda)}$$

and thus $\mathbb{P}(\mathsf{sources}(\mathcal{P}) \geq \eta\sqrt{x\lambda t}) \leq \exp(-\psi(\eta)(\sqrt{xt\lambda} - x\lambda))$. With minor modifications one proves the same bound for sinks (possibly by changing ψ): $\mathbb{P}(\mathsf{sinks}(\mathcal{P}) \geq \eta\sqrt{x\lambda t}) \leq \exp(-\psi(\eta)(\sqrt{xt\lambda} - x\lambda))$ and Lemma 11 is proved.

We can conclude the proof of the lower bound in Theorem 9. Let us write

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

we bound the right-hand side using Lemmas 10 and 11.

4. *Proof of Theorem 1 when $k_n \to +\infty$: de-Poissonisation*

In order to conclude the proof of Theorem 1 it remains to de-Poissonise Theorem 9. 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:

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

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

$$\mathbb{E}[e_{X_1,\dots,X_n}] = \mathbb{E}\left[\mathcal{L}_<(\Pi_{nk_n,n}^{(1/n)})\right] = 2\sqrt{nk_n} - k_n + o(\sqrt{nk_n}). \tag{28}$$

The last equality is obtained by combining Theorem 9 for

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

with the trivial bound $\mathcal{L}_<(\Pi_{nk_n,n}^{(1/n)}) \leq n$. In order to exploit (28) we need the following smoothness estimate.

LEMMA 12. *For every $i_1, \dots, i_n$ and $j_1, \dots, j_n$*

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

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 $i_1 + \cdots + i_n$. 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 $i_1, \dots, i_n$,

$$e_{i_1,\dots,i_n} \leq \mathbb{E}[\mathcal{L}_<(\sigma_{i_1+\dots+i_n})] \leq 6\sqrt{i_1 + \cdots + i_n}. \tag{29}$$

(The second inequality follows for example 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 sets of points $\mathcal{S}_{i_1,\dots,i_n}$, $\widetilde{\mathcal{S}}_{j_1,\dots,j_n}$ then

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

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 coordinates.) 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 (29).

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

$$|e_{k,\dots,k} - \mathbb{E}[e_{X_1,\dots,X_n}]| \le \mathbb{E}\left[|e_{k,\dots,k} - e_{X_1,\dots,X_n}|\right] \le 6\times\mathbb{E}\left[\left(\sum_{\ell=1}^{n}|X_\ell - k|\right)^{1/2}\right]. \tag{30}$$

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 Equations (30) and (28) show that

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

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

$$n^2 k_n \exp(-(k_n)^\alpha) = \mathrm{o}(\sqrt{nk_n}) \tag{31}$$

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

We first observe that de-Poissonisation 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 9 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 $n^2 k_n \exp(-k_n^\alpha) = \mathrm{o}(\sqrt{nk_n})$. 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

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

The event E_n occurs with large probability. Indeed,

$$\begin{aligned} 1 - \mathbb{P}(E_n^{\lambda_n}) &\le n\mathbb{P}(\mathrm{Poisson}(nk_n\lambda_n) \le k_n) \\ &\le n\mathbb{P}(\mathrm{Poisson}(nk_n\lambda_n) \le nk_n\lambda_n + k_n - nk_n\lambda_n) \\ &\le n\mathbb{P}(\mathrm{Poisson}(nk_n\lambda_n) \le nk_n\lambda_n - k_n\delta_n) \\ &\le n\exp\left(\frac{k_n^2\delta_n^2}{4nk_n\lambda_n}\right) \le n\exp\left(-\tfrac{1}{8}k_n\delta_n^2\right) = n\exp\left(-\tfrac{1}{8}k_n^\alpha\right). \end{aligned} \tag{32}$$

At the last line we used Lemma 15. The latter probability tends to 0 as (k_n) is large.

LEMMA 13. *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*

$$\mathcal{L}_\le(S_{k_n;n}) \le \mathcal{L}_\le(\Pi^{(\lambda_n)}_{nk_n,n}) + nk_n(1 - \mathbf{1}_{E_n^{\lambda_n}}). \tag{33}$$

Proof of Lemma 13. 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 $E_n^{\lambda_n}$ occurs then the relative orders of points in $\tilde{\Pi}^{(\lambda_n)}_{nk_n,n}$ corresponds to a uniform k_n-multiset permutation. If $E_n^{\lambda_n}$ does not hold we bound $\mathcal{L}_\le(S_{k_n;n})$ by the worst case nk_n.

Taking expectations in (33) and using the upper bound (15) yields

$$\mathbb{E}[\mathcal{L}_\le(S_{k_n;n})] \le 2\sqrt{nk_n(1+\delta_n)} + k_n(1+\delta_n) + n^2 k_n \exp\left(-\tfrac{1}{8}k_n^\alpha\right),$$

hence the upper bound in (2).

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

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

The event $F_n^{\lambda_n}$ occurs with large probability. Indeed

$$1 - \mathbb{P}(F_n^{\lambda_n}) \le n\mathbb{P}(\mathrm{Poisson}(nk_n\lambda_n) \ge k_n) \le n\exp\left(-\tfrac{1}{8}k_n^\alpha\right),$$

which tends to zero. 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

$$\mathcal{L}_\le(S_{k_n;n}) \ge \mathcal{L}_\le(\Pi^{(\lambda_n)}_{nk_n,n})\mathbf{1}_{F_n^{\lambda_n}}.$$

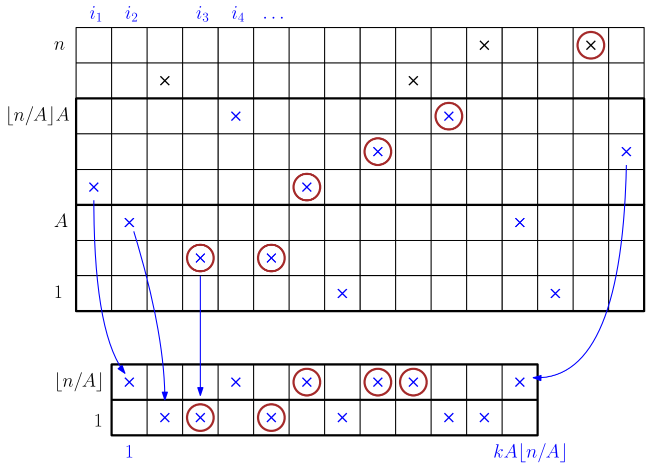


Fig. 6. Illustration of the notation of Lemma 14. Top: the multiset permutation $S_{k;n}$. Bottom: the corresponding $\widetilde{S}$. The longest non-decreasing subsequence in $S_{k;n}$ (circled points) is mapped onto a non-decreasing subsequence in $\widetilde{S}$, except one point with height $> A\lfloor n/A \rfloor$.

Therefore

$$\mathbb{P}\Big(\mathcal{L}_{\leq}(S_{k_n;n}) < (2\sqrt{nk_n(1-\delta_n)} + k_n(1-\delta_n))(1-\varepsilon)\Big) \leq$$
$$\mathbb{P}\Big(\mathcal{L}_{\leq}(\Pi^{(\lambda_n)}_{nk_n,n}) < (2\sqrt{nk_n(1-\delta_n)} + k_n(1-\delta_n))(1-\varepsilon)\Big) + \mathbb{P}\big(\text{not } F_n^{\lambda_n}\big).$$

and we conclude with (19).

5·2. *The gap between small and large* (k_n): *conclusion of the proof of Theorem* 2

After I circulated a preliminary version of this paper, 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 14. *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*

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

Proof of Lemma 14. 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 $1 \leq i_1 < i_2 < \ldots < i_{kA\lfloor n/A \rfloor}$ the indices such that $1 \leq i_\ell \leq \lfloor n/A \rfloor$ for every ℓ (see Fig. 6). 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.

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 is 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 14:

$$\mathbb{E}[\mathcal{L}_{\leq}(S_{k_n;n})\,] \leq \mathbb{E}[\mathcal{L}_{\leq}(S_{k_n \log n; \lfloor n/\lfloor \log n \rfloor \rfloor})\,] + k_n \log n \tag{34}$$

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 (34) is also $2\sqrt{nk_n} + \mathrm{o}(\sqrt{nk_n})$.

6. *Conclusion: Proof of Proposition* 3

In this short section we give the arguments needed to enhance estimates in expectation into convergences in probability. We have to prove that for every $\varepsilon > 0$:

$$\mathbb{P}\Big(L_{<}(S_{k_n;n}) > (2\sqrt{nk_n} - k_n)(1+\varepsilon)\Big) \to 0,$$

$$\mathbb{P}\Big(L_{<}(S_{k_n;n}) < (2\sqrt{nk_n} - k_n)(1-\varepsilon)\Big) \to 0,$$

$$\mathbb{P}\Big(L_{\leq}(S_{k_n;n}) > (2\sqrt{nk_n} + k_n)(1+\varepsilon)\Big) \to 0,$$

$$\mathbb{P}\Big(L_{\leq}(S_{k_n;n}) < (2\sqrt{nk_n} + k_n)(1-\varepsilon)\Big) \to 0$$

We only write the details for the first case, as the three other ones are almost identical.

The case where (k_n) is small has been proved in Section 2 so it remains to prove the case where (k_n) is large. We reuse the event $E_n^{\lambda_n}$ introduced in Section 5·1.

$$\begin{aligned}
&\mathbb{P}\big(L_{<}(S_{k_n;n}) > (2\sqrt{nk_n} - k_n)(1+\varepsilon)\big) \\
&\qquad \leq \mathbb{P}\big(E_n^{\lambda_n} \text{ does not occur}\big) \\
&\qquad\quad + \mathbb{P}\left(\mathcal{L}_{<}\left(\Pi^{(\lambda_n)}_{nk_n,n}\right) > (1+\delta_n)(2\sqrt{nk_n} - k_n)\frac{1+\varepsilon}{1+\delta_n}\right) \\
&\qquad \leq n\exp\left(-\tfrac{1}{8}k_n^{\alpha}\right) \qquad \text{(recall (32))} \\
&\qquad\quad + \mathbb{P}\left(\mathcal{L}_{<}\left(\Pi^{(\lambda_n)}_{nk_n,n}\right) > (2\sqrt{nk_n(1+\delta_n)} - k_n(1+\delta_n))\frac{1+\varepsilon}{1+\delta_n}\right) \\
&\qquad \leq n\exp\left(-\tfrac{1}{8}k_n^{\alpha}\right) + \exp(-\tilde{g}(\varepsilon/2)(\sqrt{nk_n} - k_n)),
\end{aligned}$$

for large enough n and for some positive $\tilde{g}$, using (16). This tends to zero as desired.

The lower bound for $L_{<}(S_{k_n;n})$ is proved in the same way. For the convergence of $L_{\leq}(S_{k_n;n})$ we reuse the event $F_n^{\lambda_n}$ with $\lambda_n = (1/n)(1+\log(n))$.

A. *Useful tail inequalities*

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

LEMMA 15 (See [**JŁR00**, chapter 2]). *Let* Poisson(λ) *be a Poisson random variable with mean* λ*. For every* $A > 0$

$$\mathbb{P}(\text{Poisson}(\lambda) \leq \lambda - A) \leq \exp(-A^2/4\lambda),$$
$$\mathbb{P}(\text{Poisson}(\lambda) \geq \lambda + A) \leq \exp(-A^2/4\lambda).$$

LEMMA 16 ([**JŁR00**, theorem 2·1]). *Let* Binomial(n, p) *be a Binomial random variable with parameters (n,p). For* $0 < \varepsilon < 1$,

$$\mathbb{P}(\text{Binomial}(n, p) \leq np - \varepsilon np) \leq \exp\left(-\varepsilon^2 np/2\right),$$

$$\mathbb{P}(\text{Binomial}(n, p) \geq np + \varepsilon np) \leq \exp\left(-\varepsilon^2 np/3\right).$$

LEMMA 17. *Fix* $\alpha \in (0, 1)$ *and let* $\mathcal{G}_1^{(\alpha)}, \ldots, \mathcal{G}_k^{(\alpha)}$ *be i.i.d. random variables with distribution* $\text{Geometric}_{\geq 0}(1-\alpha)$*. Then* $\mathbb{E}[\mathcal{G}_1^{(\alpha)}] = \alpha/(1-\alpha)$ *and for every* $0 < \varepsilon < 1$,

$$\mathbb{P}\left(\mathcal{G}_1^{(\alpha)} + \cdots + \mathcal{G}_k^{(\alpha)} \geq (1+\varepsilon)k\frac{\alpha}{1-\alpha}\right) \leq \exp\left(-\varepsilon^2 k\alpha/20\right),$$

$$\mathbb{P}\left(\mathcal{G}_1^{(\alpha)} + \cdots + \mathcal{G}_k^{(\alpha)} \leq (1-\varepsilon)k\frac{\alpha}{1-\alpha}\right) \leq \exp\left(-\varepsilon^2 k\alpha/20\right).$$

Proof of Lemma 17. We will use the two inequalities:

$$\exp(z) \leq 1 + z + z^2 \text{ for } |z| < 1, \qquad \frac{1}{1-u} \leq \exp(u + u^2) \text{ for } |u| < 1/2.$$

Fix λ such that $|\lambda| < \min\{1, (1-\alpha)/4\alpha\}$ so that $(\alpha/(1-\alpha))|\lambda + \lambda^2| < 1/2$:

$$\begin{aligned}
\mathbb{E}[e^{\lambda(\mathcal{G}_1^{(\alpha)} - \frac{\alpha}{1-\alpha})}] &= \frac{(1-\alpha)}{1-\alpha e^{\lambda}} e^{-\lambda\frac{\alpha}{1-\alpha}} = \frac{1}{1 - \frac{\alpha}{1-\alpha}(e^{\lambda}-1)} e^{-\lambda\frac{\alpha}{1-\alpha}} \\
&\leq \frac{1}{1 - \frac{\alpha}{1-\alpha}(\lambda + \lambda^2)} e^{-\lambda\frac{\alpha}{1-\alpha}} \\
&\leq \exp\left(\frac{\alpha}{1-\alpha}(\lambda+\lambda^2) + \left(\frac{\alpha}{1-\alpha}\right)^2 (\lambda+\lambda^2)^2 - \lambda\frac{\alpha}{1-\alpha}\right) \\
&\leq \exp\left(\frac{\alpha}{(1-\alpha)^2}\lambda^2\left(2 + \lambda^2 + 2\lambda\right)\right) \leq \exp\left(5\lambda^2\frac{\alpha}{(1-\alpha)^2}\right).
\end{aligned}$$

Thus, for every $|\lambda| < 1/\beta := \min\{1, (1-\alpha)/4\alpha\}$ it holds that $\mathbb{E}[e^{\lambda(\sum_{i=1}^k \mathcal{G}_i^{(\alpha)} - k\frac{\alpha}{1-\alpha})}] \leq \exp\left(\nu^2\lambda^2/2\right)$ where $\nu^2 := 10k\alpha/(1-\alpha)^2$.

This says that for every $k \geq 1$ the random variable $\mathcal{G}_1^{(\alpha)} + \cdots + \mathcal{G}_k^{(\alpha)}$ is subexponential and the Chernov method applies (use *e.g.* [**Wai19**, proposition 2·9] with $t = \varepsilon k\alpha/(1-\alpha)$):

$$\mathbb{P}\left(\mathcal{G}_1^{(\alpha)} + \cdots + \mathcal{G}_k^{(\alpha)} \geq (1+\varepsilon)k\frac{\alpha}{1-\alpha}\right)$$
$$\leq \exp\left(-\frac{t^2}{2\nu^2}\right) = \exp\left(-\varepsilon^2 k^2 \frac{\alpha^2}{(1-\alpha)^2} \times \frac{(1-\alpha)^2}{2\times 10k\alpha}\right)$$
$$= \exp\left(-\varepsilon^2 k\alpha/20\right),$$

as long as

$$\varepsilon k\frac{\alpha}{1-\alpha} \leq \frac{\nu^2}{\beta} = 10k\frac{\alpha}{(1-\alpha)^2}\min\{1, (1-\alpha)/4\alpha\}$$

which is always the case if $\varepsilon < 1$. The similar inequality holds for the left-tail bound (see [**Wai19**, proposition 2·9] again).

B. *An invariance property for the M/M/*1 *queue*

To conclude we state and prove the very simple property of the recurrent M/M/1 queue which allows to prove stationarity in Lemma 6. It is very close to Burke's property of the discrete HAD process [**FM06**].

Let $\beta > \lambda > 0$ be fixed parameters. Consider two independent homogeneous Poisson Point Process (PPP) $\Pi_{\nearrow}, \Pi_{\searrow}$ over $(0, +\infty)$ with respective intensities λ, β. Let $(H_y)_{y\geq 0}$ be the queue whose '+1' steps (customer arrivals) are given by $\Pi_{\nearrow}$ and '-1' steps (service times) are given by $\Pi_{\searrow}$ and whose initial distribution H_0 is drawn (independently from $\Pi_{\nearrow}, \Pi_{\searrow}$) according to a Geometric$_{\geq 0}(1-\beta^\star)$ with $\beta^\star = \lambda/\beta$.

Let Π_0 be the point process given by *unused* service times:

$$\Pi_0 = \left\{y \in \Pi_{\searrow} \text{ such that } H_y = 0\right\}.$$

LEMMA 18. *The process* $\overline{\Pi} := \Pi_{\nearrow} \cup \Pi_0$ *is a homogeneous PPP with intensity* β.

Proof. (The reader is invited to look at Fig. B1 for notation.)

The point process $\Pi_{\nearrow} \cup \Pi_{\searrow}$ is a homogeneous PPP with intensity $\lambda + \beta$, independent from H_0. We claim that $\overline{\Pi}$ is a subset of $\Pi_{\nearrow} \cup \Pi_{\searrow}$ where each point in $\Pi_{\nearrow} \cup \Pi_{\searrow}$ is taken independently with probability $\beta/(\lambda+\beta)$, it is therefore a homogeneous PPP with intensity β.

We need a few notation in order to prove the claim. Set $P_0 = 0$ and for $i \geq 1$ let P_i be the ith point of $\Pi_{\nearrow} \cup \Pi_{\searrow}$ and let $(\tilde{H}_i)_{i\geq 0}$ be the discrete-time embedded chain associated to H, *i.e.* $\tilde{H}_i = H_{P_i}$ for every i.

We will prove by induction that for every $i \geq 1$:

the points P_i belongs to $\overline{\Pi}$ with probability $\beta/(\lambda+\beta)$ independently from the events $\{P_1 \in \overline{\Pi}\}, \dots, \{P_{i-1} \in \overline{\Pi}\}$;
$\tilde{H}_i$ is independent from $\{P_1 \in \overline{\Pi}\}, \dots, \{P_i \in \overline{\Pi}\}$ and is a Geometric$_{\geq 0}(1-\beta^\star)$.

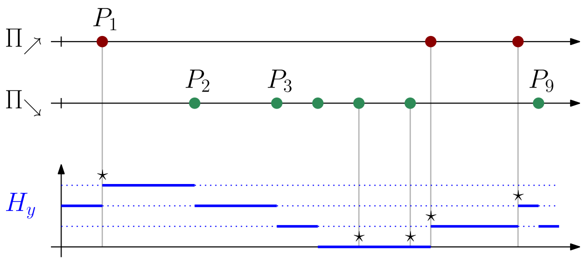


Fig. B1. Notation of Lemma 18. Points of $\overline{\Pi}$ are depicted with $\star$'s.

This implies the claim and proves the Lemma. For the base case:

$$\begin{aligned}\mathbb{P}(P_1 \in \overline{\Pi}, \tilde{H}_1 = k) &= \mathbb{P}(P_1 \in \Pi_{\nearrow}, \tilde{H}_0 = k-1)\mathbf{1}_{k\geq 1} + \mathbb{P}(P_1 \in \Pi_{\searrow}, \tilde{H}_0 = 0)\mathbf{1}_{k=0},\\ &= \frac{\lambda}{\lambda+\beta} \times (1-\beta^\star)(\beta^\star)^{k-1}\mathbf{1}_{k\geq 1} + \frac{\beta}{\lambda+\beta} \times (1-\beta^\star)\mathbf{1}_{k=0},\\ &= \frac{\beta}{\lambda+\beta} \times (1-\beta^\star)(\beta^\star)^{k} \qquad \text{(recall } \beta\beta^\star = \lambda).\end{aligned}$$

More generally let E_j be one of the two events $P_j \in \overline{\Pi}/P_j \notin \overline{\Pi}$:

$$\begin{aligned}\mathbb{P}(P_i \in \overline{\Pi}, \tilde{H}_i = k \mid E_1, \dots, E_{i-1}) &= \mathbb{P}(P_i \in \Pi_{\nearrow}, \tilde{H}_{i-1} = k-1 \mid E_1, \dots, E_{i-1})\mathbf{1}_{k\geq 1}\\ &\quad + \mathbb{P}(P_i \in \Pi_{\searrow}, \tilde{H}_{i-1} = 0 \mid E_1, \dots, E_{i-1})\mathbf{1}_{k=0},\\ &= \mathbb{P}(P_i \in \Pi_{\nearrow}, \tilde{H}_{i-1} = k-1)\mathbf{1}_{k\geq 1}\\ &\quad + \mathbb{P}(P_i \in \Pi_{\searrow}, \tilde{H}_{i-1} = 0)\mathbf{1}_{k=0}, \quad \text{(by induction hypothesis).}\\ &= \frac{\beta}{\lambda+\beta} \times (1-\beta^\star)(\beta^\star)^{k}.\end{aligned}$$

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 14 and for having enlightened me on the links with [**Bia01**]. Finally, thanks to the authors of [**CDH+22**] for their stimulating paper and to anonymous referees for their careful readings.

REFERENCES

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

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

[BEGG16] A-L. Basdevant, N. Enriquez L. Gerin and J-B. Gouéré. Discrete Hammersley's lines with sources and sinks. *ALEA Lat. Am. J. Probab. Math. Stat.* **13** (2016), 33–52.

[Bia01] P. Biane. Approximate factorization and concentration for characters of symmetric groups. *Internat. Math. Res. Notices* (2) (2001), 179–192.

[Boy22] A. BOYER. Chapter 3 (in English) of *Stationnarité bidimensionnelle de modèles aléatoires du plan*. PhD. thesis. Université Paris-Saclay, (2022), Available at https://tel.archives-ouvertes.fr/tel-03783603/.

[CDH+22] A. CLIFTON, B. DEB Y. HUANG, S. SPIRO and S. YOO. Continuously increasing subsequences of random multiset permutations. *Sém. Lothar. Combin.* **86B**(Art. 4, 11) (2022). (Proceedings of FPSAC'22.)

[CG05] E. CATOR and P. GROENEBOOM. Hammersley's process with sources and sinks. *Ann. Probab.* **33**(3) (2005), 879–903.

[CG06] E. CATOR and P. GROENEBOOM. Second class particles and cube root asymptotics for Hammersley's process. *Ann. Probab.* **34**(4) (2006), 1273–1295.

[CG19] F. CIECH and N. GEORGIOU. Order of the variance in the discrete Hammersley process with boundaries. *J. Statist. Phys.* **176**(3) (2019), 501–638.

[CLRS09] T. H. CORMEN, C. E. LEISERSON, R. L. RIVEST and C. STEIN. *Introduction to Algorithms* (MIT Press, Cambridge, MA, third edition, 2009).

[Fer96] P. A. FERRARI. Limit theorems for tagged particles. *Markov Process. Related Fields* **2**(1) (1996), 17–40.

[FM06] P. A. FERRARI and J. B. MARTIN. Multi-class processes, dual points and $M/M/1$ queues. *Markov Process. Related Fields* **12**(2) (2006), 175–201.

[Ham72] J. M. HAMMERSLEY. A few seedlings of research. In *Proceedings of the 6th Berkeley Symp. Math. Statist. and Probability*, vol. 1 (1972), 345–394.

[JŁR00] S. JANSON, T. ŁUCZAK and A. RUCINSKI. *Random Graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization (Wiley-Interscience, New York, 2000).

[Rom15] D. ROMIK. *The surprising mathematics of longest increasing subsequences*. Institute of Mathematical Statistics Textbooks, vol. 4 (Cambridge University Press, New York, 2015).

[Sep97] T. SEPPÄLÄINEN. Increasing sequences of independent points on the planar lattice. *Ann. Appl. Probab.* **7**(4) (1997), 886–898.

[Sep98] T. SEPPÄLÄINEN. Exact limiting shape for a simplified model of first-passage percolation on the plane. *Ann. Probab.* **26**(3) (1998), 1232–1250.

[Ste97] J. M. STEELE. *Probability theory and combinatorial optimization*. Society for Industrial and Applied Mathematics (SIAM), **69** 1997.

[VK77] A. M. VERŠIK and S. V. KEROV. Asymptotics of Plancherel measure of symmetrical group and limit form of Young tables. *Dokl. Akad. Nauk SSSR* **233**(6) (1977), 1024–1027.

[Wai19] M. J. WAINWRIGHT. *High-dimensional statistics*. Cambridge Series in Statistical and Probabilistic Mathematic, vol. 48 (Cambridge University Press, Cambridge, 2019).