Discrete Hammersley's lines with sources and sinks

A-L. Basdevant, N. Enriquez, L. Gerin and J-B. Gouéré

Laboratoire Modal'X, Université Paris Ouest, France.
E-mail address: `anne-laure.basdevant@u-paris10.fr`

Laboratoire Modal'X, Université Paris Ouest, France.
E-mail address: `nenriquez@u-paris10.fr`

CMAP, Ecole Polytechnique, France.
E-mail address: `gerin@cmap.polytechnique.fr`

LMPT, Université de Tours, France.
E-mail address: `jean-baptiste.gouere@lmpt.univ-tours.fr`

Abstract. We consider stationary versions of two discrete variants of Hammersley's process in a finite box. This allows us to recover in a unified and simple way the laws of large numbers proved by T. Seppäläinen for two generalized Ulam's problems. As a by-product we obtain an elementary solution for the original Ulam problem.

We also prove that for the first process defined on $\mathbb{Z}$, Bernoulli product measures are the only extremal and translation-invariant stationary measures.

1. Introduction

In a celebrated paper, J.M.Hammersley used Poissonization to attack the so-called *Ulam problem* of the typical length $\ell(n)$ of the longest increasing subsequence of a uniform permutation of size n. Namely, he reduced this problem to finding the greatest number of points of a Poisson point process inside a square, an increasing path can go through. He proved (Hammersley (1972), Theorem 4) that $\ell(n)/\sqrt{n}$ converges in probability to some constant c, sometimes referred to as the *Ulam constant*, and conjectured that $c = 2$.

The proof of $c = 2$ was achieved independently by Logan and Shepp and by Vershik and Kerov in 1977, using algebraic methods. Various authors were then interested in finding a more probabilistic proof of this result. First, Aldous and Diaconis (1995) gave one, using the properties of what they called Hammersley's process, which was implicitly introduced in Hammersley (1972) (p.358 and following). Hammersley's process is continuous in time and space and Aldous and

Diaconis studied its properties on the infinite line, in particular its stationary distributions. A few years later, Groeneboom (2002) and Cator and Groeneboom (2005) studied Hammersley's process on a quarter plane. By adding what they called Poisson *sinks* and *sources* on the x and y-axis, they also found a stationary version of this process on the quarter plane. Using this point of view, they were able to recover again the value of c.

In this paper, we study two discrete variants of Ulam's problem. Namely, for all p in $[0,1]$, $n, m \geq 1$, we replace the original Poisson point process by the following random set ξ of integer points: each integer point of the rectangle $[1, n] \times [1, m]$ is chosen independently with probability p. We are interested in the two following quantities.

- The length of the longest increasing subsequence through points of ξ:

$$L_{(n,m)}^{(1)} = \max \left\{ \begin{array}{l} L; \text{ there exists } (i_1, j_1), \ldots, (i_L, j_L) \in \xi, i_1 < \cdots < i_L \\ \qquad\qquad\qquad\qquad\qquad \text{and } j_1 < j_2 < \cdots < j_L \end{array} \right\}.$$

- The length of the longest non-decreasing subsequence through points of ξ:

$$L_{(n,m)}^{(2)} = \max \left\{ \begin{array}{l} L; \text{ there exists } (i_1, j_1), \ldots, (i_L, j_L) \in \xi, i_1 < \cdots < i_L \\ \qquad\qquad\qquad\qquad\qquad \text{and } j_1 \leq j_2 \leq \cdots \leq j_L \end{array} \right\}.$$

One aim of the present work is to recover by simple and unified probabilistic arguments the first order asymptotics of $L_{(n,m)}^{(1)}$ and $L_{(n,m)}^{(2)}$ already obtained by T. Seppäläinen in two independent papers, for $L_{(n,m)}^{(1)}$ in Seppäläinen (1997) and for $L_{(n,m)}^{(2)}$ in Seppäläinen (1998). In the following result the variables $\left(L_{(n,m)}^{(i)} \right)_{(n,m)}$ are coupled in the obvious way. Moreover, for any $a, b > 0$, $L_{(an,bm)}^{(i)}$ stands for $L_{(\lfloor an \rfloor, \lfloor bn \rfloor)}^{(i)}$.

Theorem 1.1. *For $a, b > 0$, we have, when n tends to infinity*

$$\frac{L_{(an,bn)}^{(1)}}{n} \quad \xrightarrow{a.s.} \quad \begin{cases} \dfrac{\sqrt{p}(2\sqrt{ab} - (a+b)\sqrt{p})}{1-p} & \text{if } p < \min\{a/b, b/a\}, \\ \min\{a, b\} & \text{otherwise,} \end{cases} \qquad (1.1)$$

$$\frac{L_{(an,bn)}^{(2)}}{n} \quad \xrightarrow{a.s.} \quad \begin{cases} 2\sqrt{abp(1-p)} + (a-b)p & \text{if } p < a/(a+b), \\ a & \text{otherwise.} \end{cases} \qquad (1.2)$$

To prove this, Seppäläinen associates to each problem a particle system on the infinite line $\mathbb{Z}$ and checks that Bernoulli product measures are stationary. The position of particles at a given time is then characterized as the solution of a discrete optimization problem, and the a.s. limit of $L_{(an,bn)}^{(i)}/n$ is identified using convex analysis arguments. Note also that for the second question, Johansson used later (Johansson (2001), Th.5.3) another description of $L_{(an,bn)}^{(2)}/n$ and proved that the rescaled fluctuations converge to the Tracy-Widom distribution.

In Rolla et al. (2010), Rolla-Sidoravicius-Surgailis-Vares considered the closely related model of last-passage percolation with geometric weights. Here is a quick

description of the model. Let $(\tilde{\xi}(i,j))_{1\leq i\leq n, 1\leq j\leq m}$ be a family of independent geometric random variables with parameter p. Set:

$$L^{(3)}_{(n,m)} = \max\left\{\sum_t \tilde{\xi}(\gamma(t))\right\}$$

where the maximum is taken over all paths γ from $(1,1)$ to (n,m) with $(1,0)$ or $(0,1)$ steps. With respect to $L^{(1)}_{(n,m)}$ or $L^{(2)}_{(n,m)}$, we have simply changed the weight of each point (from Bernoulli to geometric) and the set of paths considered. Rolla-Sidoravicius-Surgailis-Vares gave a new proof of the following result:

$$\frac{L^{(3)}_{(n,n)}}{n} \xrightarrow[n\to+\infty]{\text{a.s.}} \frac{2\sqrt{p}}{1-\sqrt{p}}.$$

(The original proof was given by Jockusch-Propp-Shor in Jockusch et al. (1995).)

Our proof of Theorem 1 is inspired by the proof of Rolla-Sidoravicius-Surgailis-Vares of the above result. The strategy is to consider a particle system on a *bounded domain* which turns out to coincide with the restriction of Seppäläinen's particle system on $\mathbb{Z}$. This simplifies the definition of the process (especially, in the case of the second problem). Moreover, it turns out that a *local* balance property around a single site, see Lemma 3.3 and 3.4 below, is enough to check the stationarity of the process. Our proofs are essentially the same for both models. This kind of remarkable local balance property also occurs in last-passage percolation with geometric weights (see Rolla-Sidoravicius-Surgailis-Vares , Section 3.1 in Rolla et al. (2010) or Seppäläinen, Lemma 2.3 in Seppäläinen (2009)). The particle systems studied here provide two other examples where such a property holds. Theorem 1 is then proven by investigating the behaviour of the models under the different stationary measures which have been exhibited.

A nice by-product of our proof is that we obtain non-asymptotic estimates which provide an elementary proof of $c = 2$ for the original Ulam problem (see the discussion at the end of Section 4).

In the last section, which can be read independently, we complete the study of the original *infinite* discrete Hammersley process introduced by Seppäläinen (1997) for the Problem 1. We prove that the only extremal translation-invariant stationary measures of this process are the Bernoulli product measures.
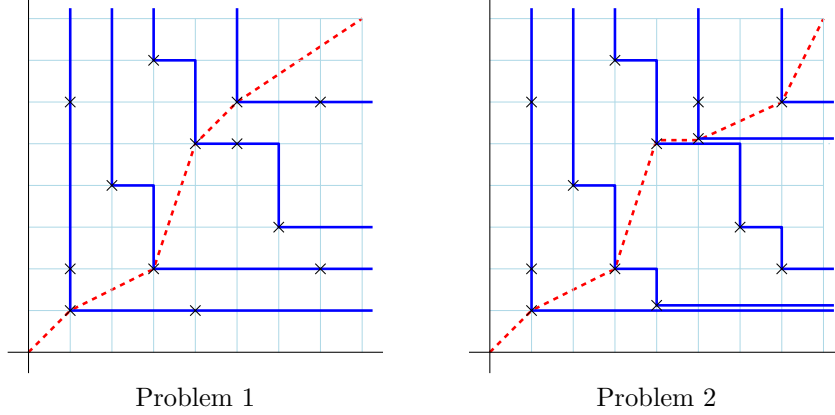
2. Discrete Hammersley's processes

Like Hammersley did, we construct two sets of broken lines whose cardinality is respectively the variable $L^{(1)}_{(n,m)}$ and $L^{(2)}_{(n,m)}$. For each case, we first introduce a partial order:

Case 1:   Case 2:
$(x,y) \prec (x',y')$ iff $(x,y) \prec (x',y')$ iff
$x < x'$ and $y < y'$ $x < x'$ and $y \leq y'$

Now, Hammersley's lines are paths starting from the top side of the rectangle $[1,n]\times[1,m]$, ending at its right side and making only south or east steps. They are constructed recursively. The first line is the highest non-increasing path connecting the minimal points of ξ for $\prec$. We withdraw these points from ξ and connect the new minima to get the second line, and so on. In the below picture, $n = m = 8$

and $L_{(n,m)}^{(1)} = 4$, $L_{(n,m)}^{(2)} = 5$ (crosses denote points of ξ, Hammersley's lines are in blue, one of the longest subsequences is shown in a red dashed line):

Problem 1   Problem 2

We claim that the number $H_{(n,m)}^{(i)}$ of Hammersley's lines is equal to $L_{(n,m)}^{(i)}$. Indeed, note that, in both cases, $L_{(n,m)}^{(i)}$ is equal to the length of the longest increasing subsequence trough points of ξ for the partial order $\prec$. Each Hammersley's line connecting minimal points of a subset of ξ, the $L_{(n,m)}^{(i)}$ points of a maximizing sequence necessarily lie on distinct Hammersley's lines and therefore $H_{(n,m)}^{(i)} \geq L_{(n,m)}^{(i)}$. To prove the opposite inequality, we observe that Hammersley's lines give a construction of an increasing sequence of points $a_1, \ldots a_{H_{(n,m)}^{(i)}}$ of ξ in the following way. For $a_{H_{(n,m)}^{(i)}}$ we take any point of ξ which lies in the top right Hammersley's line. Then, for each ℓ, $a_{\ell-1}$ is taken as one of the points of ξ which lies in the $\ell-1$-th line and such that $a_{\ell-1} \prec a_\ell$. The existence of such $a_{\ell-1}$ is granted. Indeed, if no such point existed, a_ℓ would have belonged to the $(\ell-1)$-th line. This proves $L_{(n,m)}^{(i)} \geq H_{(n,m)}^{(i)}$.

Let us note that the construction of lines above is consistent in n and m, *i.e.* for $n' < n$ and $m' < m$ the restriction of Hammersley's lines to the rectangle $[1, n'] \times [1, m']$ coincides with Hammersley's lines constructed from the points of ξ in this smaller rectangle. In particular, for $n \geq 1$, we can define the two discrete time processes $(X_t^1)_{t \geq 0} := (X_t^1(x), \ x \in [\![1, n]\!])_{t \geq 0}$ and $(X_t^2)_{t \geq 0} := (X_t^2(x), \ x \in [\![1, n]\!])_{t \geq 0}$ defined by

$$X_t^i(x) = \begin{cases} 1 & \text{if there is a Hammersley's line on the vertical edge } \{(x,t), (x, t+1)\} \\ 0 & \text{otherwise.} \end{cases}$$

(2.1)

Note that on each vertical edge there is at most one Hammersley's line. Moreover, in Model 1, there is also at most one Hammersley's line on any horizontal edge and Hammersley's lines do not intersect each other. We will say that there is a *particle* at time t at position x in Model i if $X_t^i(x) = 1$. Hence, for both models we start with the initial condition $X_0^i \equiv 0$ and one can check that both processes are Markovian. Besides, since Hammersley's lines start all from the top side of the rectangle,

for both models, the number of particles at time t is equal to $L_{(n,t)}^{(i)}$.

In both models, the dynamics of the particle systems are quite simple. As the description of these dynamics is not needed in the proof of Theorem 1.1, we do not give such a description here. However, we refer the interested reader to Appendix A.

3. Sinks and sources

Since, in both models, the number of particles can only increase as time goes on, both processes converge to their unique stationary measure $X_\infty^i :\equiv 1$. In this section, we are going to modify a little bit the problems we are looking at, such that the particle systems associated to the new problems admit less trivial stationary measures. To do so, we must first introduce some notation.

From now on, for any integer $x \geq 1$ and $t \geq 1$, we define the random variable $\xi_t(x)$ by

$$\xi_t(x) = \begin{cases} 1 & \text{if } (x, t) \in \xi \\ 0 & \text{otherwise.} \end{cases} \tag{3.1}$$

With this notation, we can write the quantities $L_{(n,m)}^{(i)}$ as

$$L_{(n,m)}^{(i)} = \max \left\{ \sum_s \xi_{t(s)}(x(s)) \right\} \tag{3.2}$$

where the maximum is taken over all the discrete paths $(x(\cdot), t(\cdot))$ in $[1, n] \times [1, m]$ starting from $(1, 1)$ which make only steps in $\mathbb{N} \times \mathbb{N}$ for Model 1 and in $\mathbb{N} \times \mathbb{Z}_+$ for Model 2. Given two sequences of integers $(\xi_t(0))_{t \in [\![1,m]\!]}$ and $(\xi_0(x))_{x \in [\![1,n]\!]}$, we now consider the quantities
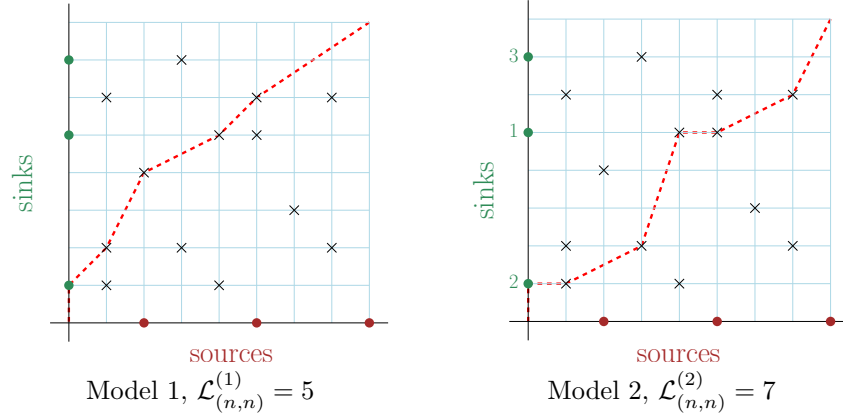
$$\mathcal{L}_{(n,m)}^{(i)} = \max \left\{ \sum_s \xi_{t(s)}(x(s)) \right\} \tag{3.3}$$

where the maximum is taken over all the discrete paths $(x(\cdot), t(\cdot))$ in $[0, n] \times [0, m]$ starting from $(0, 0)$ such that:
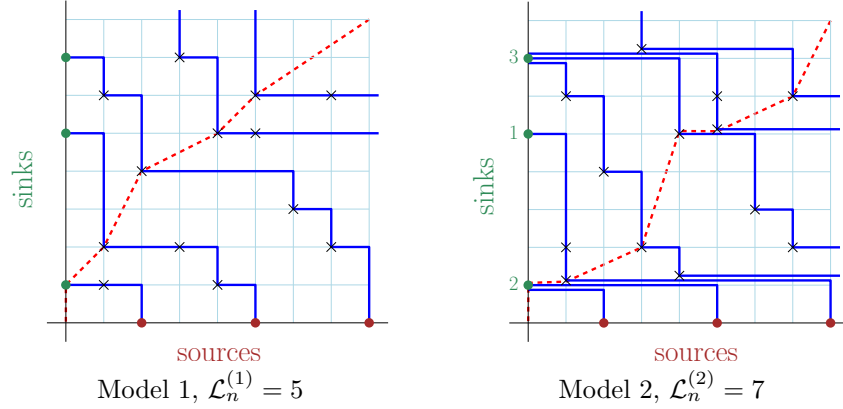
- There exists an integer $u \in \mathbb{Z}_+$ such that: either they first make u $(0, 1)$ steps or they first make u $(1, 0)$ steps.
- Then, they only make steps in $\mathbb{N} \times \mathbb{N}$ for Model 1 and in $\mathbb{N} \times \mathbb{Z}_+$ for Model 2.

Using the same terminology as in Cator and Groeneboom (2005), we say that there are k *sinks* (resp. *sources*) at site $(0, t)$ (resp. $(x, 0)$) if $\xi_t(0) = k$ (resp. $\xi_0(x) = k$). The sources will be drawn on the x-axis and the sinks on the y-axis. Hence, in other words, $\mathcal{L}_{(n,m)}^{(i)}$ is equal to the maximum number of points an increasing (resp. non-decreasing) path can go trough when the path can pick either sources or either sinks before going into $\mathbb{N} \times \mathbb{N}$. The law of the sources and sinks will be specified later, let us just say that the sources will take values in $\{0, 1\}^n$ whereas the sinks will take values in $\{0, 1\}^m$ for Model 1 and in $\mathbb{Z}_+^m$ for Model 2.

As before, we can construct a set of broken lines whose cardinality is equal to $\mathcal{L}_{(n,m)}^{(i)}$. Now, Hammersley's lines are paths starting from the top side or the left side of the rectangle $[0, n] \times [0, m]$, ending at its bottom side or its right side and making only south or east steps. They are still constructed recursively. Assume that the sink of minimal height is located at $(0, t_1)$ and the leftmost source at $(x_1, 0)$. Then, the first line is the highest non-increasing path starting from $(0, t_1)$,

Model 1, $\mathcal{L}_{(n,n)}^{(1)} = 5$ Model 2, $\mathcal{L}_{(n,n)}^{(2)} = 7$

ending at $(x_1, 0)$ and connecting minimal points (x, t) of ξ for the partial order $\prec$ introduced in the previous section such that $(x_1, 0) \not\prec (x, t)$ and $(0, t_1) \not\prec (x, t)$. If there is no sink (resp. source) at all, we do the same construction except that the line starts from the top of the box (resp. ends on the right of the box). Then, we withdraw these points from ξ, one of the sink located at $(0, t_1)$ and the source located at $(x_1, 0)$ and we do the same procedure with this new set of sources, sinks, and set of points in $[\![1, n]\!] \times [\![1, m]\!]$ to obtain the second line. The algorithm goes on until there is no more points, sinks and sources. Using the same argument as in the previous section, one can easily prove that the number of lines obtained with this procedure is equal to $\mathcal{L}_{(n,m)}^{(i)}$.

Model 1, $\mathcal{L}_n^{(1)} = 5$ Model 2, $\mathcal{L}_n^{(2)} = 7$

With a slight abuse of notation, we still denote $(X_t^1)_{t \geq 0} := (X_t^1(x), \ x \in [\![1, n]\!])_{t \geq 0}$ and $(X_t^2)_{t \geq 0} := (X_t^2(x), \ x \in [\![1, n]\!])_{t \geq 0}$ the two discrete time system of particles defined by

$$X_t^i(x) = \begin{cases} 1 & \text{if there is a Hammersley line (with sources and sinks)} \\ & \text{on the edge } \{(x, t), (x, t+1)\} \\ 0 & \text{otherwise.} \end{cases}$$

Hence, the processes now start from the configuration given by the sources *i.e* $X_0^i(x) = \xi_0(x)$.

Theorem 3.1 (Stationarity of Hammersley's processes on a bounded interval, with sources and sinks)**.** *For all n,*

- **Model 1.** *For all $p, \alpha \in (0,1)$, the process $(X_t^1(x), x \in [\![1,n]\!])_{t\geq 0}$ is stationary if sources are i.i.d. $\mathrm{Ber}(\alpha)$ and sinks are i.i.d. $\mathrm{Ber}(\alpha^\star)$ (and sinks independent from sources) with*

$$\alpha^\star = \frac{p(1-\alpha)}{\alpha + p(1-\alpha)}. \tag{3.4}$$

- **Model 2.** *For all $p, \alpha \in (0,1)$ such that $\alpha > p$, the process $(X_t^2(x), x \in [\![1,n]\!])_{t\geq 0}$ is stationary if sources are i.i.d. $\mathrm{Ber}(\alpha)$ and sinks are i.i.d. $(\mathrm{Geo}(\alpha^\star) - 1)$ (i.e. $\mathbb{P}(\xi(0,t) = k) = \alpha^\star(1-\alpha^\star)^k$ for $k = 0, 1, \dots$) (and sinks independent from sources) with*

$$\alpha^\star = \frac{\alpha - p}{\alpha(1-p)}. \tag{3.5}$$

(Note that $\alpha^\star \in (0,1)$ if $\alpha > p$.)

Remark 3.2. In Seppäläinen (1997) (resp. in Seppäläinen (1998)), Seppäläinen associates to Model 1 (resp. Model 2) a particle system defined on the whole line $\mathbb{Z}$. One can check that the restrictions of these systems to $[\![1,n]\!]$ have respectively the same transition probabilities that our processes X_t^i, $i = 1, 2$ when sinks and sources are distributed as in Theorem 3.1. Hence,

- For Model 1 (resp. Model 2), Theorem 3.1 should be compared to Lemma 2.1 in Seppäläinen (1997) (resp. Proposition 1 in Seppäläinen (1998)) which states that Bernoulli product measures are stationary for the analogous model on the infinite line.
- The condition $\alpha > p$ in Model 2 reminds the condition on the density of the initial configuration needed to define the process on $\mathbb{Z}$ in Seppäläinen (1998).
- The theorem states stationarity in time (*i.e.* from bottom to top in our figures). In fact, the proof also shows stationarity from left to right.

Let us collect a few consequences for further use. Let $\mathcal{L}_{(n,m)}^{(i),\alpha,\beta}$ be the number of Hammersley's lines in Model i with parameters α (sources), β (sinks) in the box $[0, n] \times [0, m]$ and let $\mathcal{T}_{(n,m)}^{(i),\alpha,\beta}$ be the number of Hammersley's lines that leave the same box from the top. Then, for $i = 1, 2$ and every $\alpha, \beta \in [0, 1]$

$$\mathcal{L}_{(n,m)}^{(i),0,0} = \mathcal{T}_{(n,m)}^{(i),0,0} \stackrel{\mathrm{d}}{=} L_{(n,m)}^{(i)},$$
$$\mathcal{L}_{(n,m)}^{(i),\alpha,\beta} = \mathcal{T}_{(n,m)}^{(i),\alpha,\beta} + \# \left\{ \text{sinks between } 1 \text{ and } m \right\}. \tag{3.6}$$

This relies on the fact that any lines has to exit by the top or by a sink. Note that in the right-hand side of (3.6) the two terms are not independent. Moreover, using the interpretation of $L_{(n,m)}^{(i)}$ and $\mathcal{L}_{(n,m)}^{(i),\alpha,\beta}$ as the length of a longest subsequence (*cf.* (3.2) and (3.3)), we have

$$L_{(n,m)}^{(i)} \leq \mathcal{L}_{(n,m)}^{(i),\alpha,\beta}. \tag{3.7}$$

Besides, $\mathcal{T}_{(n,m)}^{(i),\alpha,\beta}$ being by definition equal to the number of particles at time m, Theorem 3.1 implies that

$$\mathcal{T}_{(n,m)}^{(i),\alpha,\alpha^\star} \stackrel{\mathrm{law}}{=} \mathrm{Binom}(n, \alpha). \tag{3.8}$$

Proof: We aim to prove Theorem 3.1. We first prove for both models a local balance property, in Lemmas 1 and 2. These lemmas are elementary but do not seem to be written elsewhere, and form the heart of our proof.
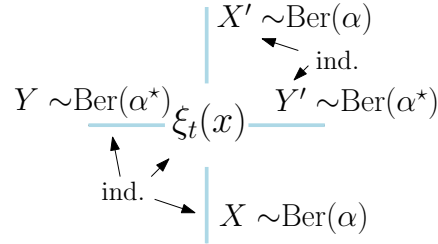
Model 1.

We first focus on what happens around a single point (x,t). Recall that $\xi_t(x)$ is a Bernoulli(p). Denote by X (resp. Y, X', Y') be the indicator that a Hammersley's line hits (x,t) from the bottom (resp. left,top,right).

Lemma 3.3 (Local balance for Model 1). *The random variables (X', Y') are measurable with respect to $(X, Y, \xi_t(x))$ and we have*

$$\left((X,Y,\xi_t(x)) \overset{(d)}{=} \mathrm{Ber}(\alpha) \otimes \mathrm{Ber}(\alpha^\star) \otimes \mathrm{Ber}(p) \right) \Rightarrow \left((X',Y') \overset{(d)}{=} \mathrm{Ber}(\alpha) \otimes \mathrm{Ber}(\alpha^\star) \right)$$

where $\alpha^\star$ is as in (3.4).

Proof of Lemma 3.3: Recalling that, in Model 1, Hammersley's lines are non increasing lines which do not touch each other and noticing that $X' = Y' = 1$ iff $X = Y = 0$ and $\xi_t(x) = 1$, we get that

$$(X',Y') = \begin{cases} (X,Y) & \text{if } X \neq Y \\ (1,1) & \text{if } X = Y = 0 \text{ and } \xi_t(x) = 1. \\ (0,0) & \text{otherwise.} \end{cases}$$

Hence, we see that $(X',Y') \in \sigma(X,Y,\xi_t(x))$ and one can easily check that if $(X',Y') \sim \mathrm{Ber}(\alpha) \otimes \mathrm{Ber}(\alpha^\star)$ with $\alpha^\star$ is as in (3.4), then $(X',Y') \sim \mathrm{Ber}(\alpha) \otimes \mathrm{Ber}(\alpha^\star)$. $\square$

We now explain how Lemma 3.3 proves Theorem 3.1 for Model 1. For every $(x,1)$ define the corresponding $(X'(x), Y'(x))$ be the *output* of $(x,1)$. Independence of sources and sinks and Lemma 3.3 ensure that the output of $(1,1)$ is independent of sources at $(2,0), (3,0), \dots$ and sinks at $(0,2), (0,3), \dots$. Then a simple induction on x proves that the random variables $(X'(x), x \in [\![1,n]\!])$ are distributed as $\mathrm{Ber}(\alpha)$.

Model 2. The strategy is similar:

Lemma 3.4 (Local balance for Model 2). *The random variables (X', Y') are measurable with respect to $(X, Y, \xi_t(x))$ and we have*

$$\left((X,Y,\xi_t(x)) \overset{(d)}{=} \mathrm{Ber}(\alpha) \otimes (\mathrm{Geo}(\alpha^\star) - 1) \otimes \mathrm{Ber}(p) \right) \Rightarrow$$

$$\left((X',Y') \overset{(d)}{=} \mathrm{Ber}(\alpha) \otimes (\mathrm{Geo}(\alpha^\star) - 1) \right)$$

where $\alpha^\star$ is as in (3.5).

$$
\begin{array}{c}
\left| \begin{array}{c} X' \sim \mathrm{Ber}(\alpha) \\ \end{array} \right. \\
\text{ind.} \\
Y \sim \mathrm{Geo}(\alpha^\star) - 1 \qquad Y' \sim \mathrm{Geo}(\alpha^\star) - 1 \\
\xi_t(x) \\
\text{ind.} \\
\left| \begin{array}{c} X \sim \mathrm{Ber}(\alpha) \end{array} \right.
\end{array}
$$

Proof of Lemma 3.4: Recall that $X, X' \in \{0, 1\}$ whereas $Y, Y' \in \mathbb{Z}^+$. Besides, using that Hammersley's lines are non-increasing, the balance between incoming and outcoming lines at a site yields $X' + Y = X + Y'$. Moreover, $X' = 1$ iff $(\xi_t(x) = 1$ or $(X, Y) = (1, 0))$.

We observe that the different cases can be summed up in

$$
(X', Y') = \begin{cases} (1, 0) & \text{if } (X, Y) = (1, 0) \\ (\xi_t(x), Y - X + \xi_t(x)) & \text{otherwise.} \end{cases}
$$

One can easily check that this equality implies Lemma 3.4. For instance, for $k \geq 1$, $\mathbb{P}(X = 1, Y = k) = \alpha \alpha^\star (1 - \alpha^\star)^k$ should be equal to

$$
\mathbb{P}(X' = 1, Y' = k) = \mathbb{P}(\xi_t(x) = 1, X = 0, Y = k - 1) + \mathbb{P}(\xi_t(x) = 1, X = 1, Y = k)
$$
$$
= p(1 - \alpha)\alpha^\star (1 - \alpha^\star)^{k-1} + p\alpha\alpha^\star (1 - \alpha^\star)^k
$$
$$
= p(1 - \alpha^\star)^{k-1}\alpha^\star \left[(1 - \alpha) + \alpha(1 - \alpha^\star) \right],
$$

which requires (3.5). $\qquad \square$

Theorem 3.1 for Model 2 then follows from Lemma 3.4 in the same way as before Theorem 3.1 for Model 1 follows from Lemma 3.3. $\qquad \square$

4. Law of Large Numbers in Hammersley's processes

In this section, we explain how Theorem 3.1 on the stationary measure of the processes with sources and sinks implies Theorem 1.1.

Proof of Theorem 1.1: Trivial case. We first consider Model 1 in the case $p \geq \min\{a/b, b/a\}$. The upper bound is straightforward. The lower bound can be proven as follows. Let us consider for example the case $p \geq ab^{-1}$. We have to prove:

$$
\liminf_{n \to \infty} \frac{L^{(1)}_{(an, bn)}}{n} \geq a. \tag{4.1}
$$

One can build an increasing path as follows. The first point $(1, y_1)$ of the path is the lowest point of ξ having first coordinate equal to 1. The second point $(2, y_2)$ of the path is the lowest point of ξ having first coordinate equal to 2 and second coordinate strictly larger than y_1. The other points are defined in a similar fashion. As $pb \geq a$, a study of this path easily provides (4.1). The case $p \geq ba^{-1}$ can be handled in a similar way.

The same strategy treats the case $p \geq a/(a + b)$ in Model 2.

Non trivial case. We only need to compute $\lim \frac{1}{n}\mathbb{E}[L^{(i)}_{(an, bn)}]$ since almost sure convergence follows from superadditivity.

Model 1. Upper bound. We assume that $p \geq \min\{a/b, b/a\}$. For any $\alpha \in (0, 1)$, taking $\alpha^\star$ as in (3.4) and using (3.6), (3.7) and (3.8), we get

$$
\begin{aligned}
\frac{1}{n}\mathbb{E}[L_{(an,bn)}^{(1)}] &\leq \frac{1}{n}\mathbb{E}[\mathcal{L}_{(an,bn)}^{(1),\alpha,\alpha^\star}] \\
&\leq \frac{1}{n}\mathbb{E}[\mathcal{T}_{(an,bn)}^{(1),\alpha,\alpha^\star}] + \frac{1}{n}\mathbb{E}[\# \{\text{sinks between 1 and } bn\}] \\
&= a\alpha + b\alpha^\star = a\alpha + b\frac{p(1-\alpha)}{\alpha + p(1-\alpha)} =: \phi^{(1)}(a, b, \alpha). \quad (4.2)
\end{aligned}
$$

The latter is minimized for

$$
\alpha(a, b) := \frac{\sqrt{p}\sqrt{\frac{b}{a}} - p}{1 - p}, \qquad \alpha^\star(a, b) := \frac{\sqrt{p}\sqrt{\frac{a}{b}} - p}{1 - p}, \qquad (4.3)
$$

which both are in $(0, 1)$ if $p < \min\{a/b, b/a\}$. This yields

$$
\frac{1}{n}\mathbb{E}[L_{(an,bn)}^{(1)}] \leq \phi^{(1)}(a, b, \alpha(a, b)) = \frac{\sqrt{p}(2\sqrt{ab} - (a+b)\sqrt{p})}{1 - p}.
$$

Model 2. Upper bound. We assume that $p \geq a/(a+b)$. Using (3.8) with $\alpha^\star = \frac{\alpha - p}{\alpha(1-p)}$ and that $\mathbb{E}[\text{Geo}(\alpha^\star) - 1] = \frac{1}{\alpha^\star} - 1$, we get

$$
\begin{aligned}
\frac{1}{n}\mathbb{E}[L_{(an,bn)}^{(2)}] &\leq \frac{1}{n}\mathbb{E}[\mathcal{T}_{(an,bn)}^{(i),\alpha,\alpha^\star}] + \frac{1}{n}\mathbb{E}[\# \{\text{sinks between 1 and } bn\}] \\
&= a\alpha + b(\frac{1}{\alpha^\star} - 1) = a\alpha + b\frac{p(1-\alpha)}{\alpha - p}.
\end{aligned}
$$

We minimize the latter by taking

$$
\alpha = p + \sqrt{\frac{b}{a}p(1-p)}.
$$

Note that this choice is allowed since for $p \in (0, a/(a+b))$ then $0 < p < \alpha < 1$.
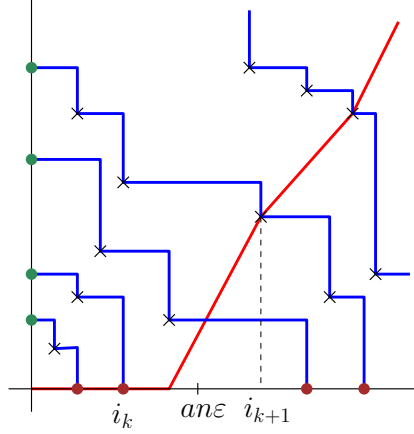
Model 1 and 2. Lower bound.

In Seppäläinen (1997, 1998) the lower bound was obtained with a convexity argument on the scaling limit. Instead, we adapt a more probabilist argument due to Rolla-Sidoravicius-Surgailis-Vares (proof of Theorem 4.1 in Rolla et al. (2010)) for the closely related model of last passage percolation with geometric weights. Here, we only give the details for Model 1 but the same argument applies for Model 2 (there are only minor modifications due mainly to the fact that sinks have no more a Bernoulli distribution but a geometric distribution).

Let us consider Model 1 on the rectangle $[0, an] \times [0, bn]$ with sinks and sources with optimal source intensity $\alpha = \alpha(a, b)$ and $\alpha^\star = \alpha^\star(a, b)$ (see (4.3)). For $\varepsilon \in [0, 1]$, denote by $L_{(an,bn)}^{(1)}(\varepsilon)$ the length of the largest non-decreasing subsequence defined by

$$
L_{(an,bn)}^{(1)}(\varepsilon) = \max\{L; (i_1, 0), \ldots, (i_k, 0), (i_{k+1}, j_{k+1}), \ldots, (i_L, j_L) \in \xi,
$$
$$
0 < i_1 < \cdots < i_k \leq an\varepsilon < i_{k+1} < \ldots < i_L \leq an \text{ and } 0 < j_{k+1} < \cdots < j_L \leq bn\}.
$$

Here is an example where $L_{(an,bn)}^{(1)}(\varepsilon) = 4$. Note that because of the constraint $L_{(an,bn)}^{(1)}(\varepsilon)$ is smaller than the number of Hammersley's lines $\mathcal{L}_{(an,bn)}^{(1),\alpha,\alpha^\star}$.

Recall that $\phi^{(1)}$ has been defined in (4.2).

Lemma 4.1 (Largest subsequence using some sources). *There exist positive and non-decreasing functions f, g on $(0,1]$, which depend on a, b, such that*

$$\mathbb{P}\left(L^{(1)}_{(an,bn)}(\varepsilon) \geq n\left(\phi^{(1)}(a,b,\alpha(a,b)) - f(\varepsilon)\right)\right) \leq \exp(-ng(\varepsilon)).$$

Proof of Lemma 4.1: We write $L^{(1)}_{(an,bn)}(\varepsilon) = I_1 + I_2$ with

$$
\begin{aligned}
I_1 &:= \max\left\{k; (i_1,0),\ldots,(i_k,0) \in \xi, i_1 < \cdots < i_k \leq an\varepsilon\right\} \\
I_2 &:= \max\left\{
\begin{array}{l}
k; (i_1,j_1),\ldots,(i_k,j_k) \in \xi, an\varepsilon < i_1 < \cdots < i_k \leq an \text{ and} \\
0 < j_1 < \cdots < j_k \leq bn
\end{array}
\right\}.
\end{aligned}
$$

The random variables I_1 and I_2 are independent, I_1 is binomially distributed with parameters $(\lfloor an\varepsilon \rfloor, \alpha(a,b))$. The random variable I_2 has the law of $L^{(1)}_{(an(1-\varepsilon),bn)}$, thus is dominated by a random variable I_3 with law $\mathcal{L}^{(1),\alpha(a(1-\varepsilon),b),\alpha^\star(a(1-\varepsilon),b)}_{(an(1-\varepsilon),bn)}$. Using (4.2), we get:

$$
\begin{aligned}
\frac{1}{n}\mathbb{E}\left[L^{(1)}_{(an,bn)}(\varepsilon)\right] &\leq \frac{1}{n}\mathbb{E}\left[I_1 + I_3\right] \\
&\leq a\varepsilon\alpha(a,b) + \phi^{(1)}\left(a(1-\varepsilon),b,\alpha(a(1-\varepsilon),b)\right),
\end{aligned}
$$

where, by convention, $\alpha(a',b') = 1$ when $p \geq a'/b'$ and $\alpha(a',b') = 0$ when $p \geq b'/a'$. One can check that for $p < \min\{a/b, b/a\}$ and $a', b' > 0$ such that $a/b \neq a'/b'$ we have

$$\phi^{(1)}(a',b',\alpha(a,b)) > \phi^{(1)}(a',b',\alpha(a',b')).$$

This is due to the fact that $\alpha(a,b) \neq \alpha(a',b')$ and that there is a unique $\alpha \in [0,1]$ which minimizes $\phi(a',b',\alpha)$. Thus,

$$
\begin{aligned}
a\varepsilon\alpha(a,b) + \phi^{(1)}\Big(a(1-\varepsilon),b,\alpha(a(1-\varepsilon),b)\Big) \;\; &< \;\; a\varepsilon\alpha(a,b) \\
&\quad + \phi^{(1)}\Big(a(1-\varepsilon),b,\alpha(a,b)\Big) \\
&= \;\; a\varepsilon\alpha(a,b) + a(1-\varepsilon)\alpha(a,b) \\
&\quad + b\alpha^{\star}(a,b) \\
&= \;\; a\alpha(a,b) + b\alpha^{\star}(a,b) \\
&= \;\; \phi^{(1)}(a,b,\alpha(a,b)).
\end{aligned}
$$

Denote by $d(\varepsilon)$ the difference between the right and left hand side of the inequality. The function d is positive on $(0,1]$ and continuous. We set $f(\varepsilon) = \frac{1}{2}\min\{d(\delta),\delta \in [\varepsilon,1]\}$ which is non-decreasing and positive. Then, we have

$$
\begin{aligned}
\mathbb{P}\left(L^{(1)}_{(an,bn)}(\varepsilon) \geq n\left(\phi^{(1)}(a,b,\alpha(a,b)) - f(\varepsilon)\right)\right) & \\
\leq \;\; \mathbb{P}\left(I_1 + I_3 \geq \mathbb{E}(I_1 + I_3) + n\big(d(\varepsilon) - f(\varepsilon)\big)\right) & \\
\leq \;\; \mathbb{P}\left(I_1 + I_3 \geq \mathbb{E}(I_1 + I_3) + nf(\varepsilon)\right). &
\end{aligned}
$$

Recalling that I_1 is binomially distributed and I_3 is the sum of two binomial random variables, function g is obtained by applying Hoeffding's inequality to I_1 and I_3.   □

We still consider Model 1 on the rectangle $[0,an] \times [0,bn]$ with sinks and sources with optimal source intensity $\alpha = \alpha(a,b)$ and $\alpha^{\star} = \alpha^{\star}(a,b)$. Let π_n be an optimal path for $\mathcal{L}^{(1),\alpha,\alpha^{\star}}_{(an,bn)}$. If there are several optimal paths, we choose one of them in an arbitrary way. Define D_n as the number of sources $(i,j) \in \pi_n$ with $j = 0$.

Lemma 4.2 (Optimal paths do not take many sources). *There exists a positive function h on $(0,1]$ such that for and any $\delta \in (0,1]$ and for any n large enough*

$$
\mathbb{P}\left(D_n > an\delta\right) \leq \exp(-nh(\delta)).
$$

In particular D_n/n converges to 0 in L^1.

Proof of Lemma 4.2: By definition, if $D_n > an\delta$ then for some $\varepsilon \geq \delta$ such that $\varepsilon an \in \mathbb{N}$ and $\epsilon \leq 1$

- there are more than $an\delta$ sources in $\{1,\ldots,\varepsilon an\}$;
- $L^{(1)}_{(an,bn)}(\varepsilon) = \mathcal{L}^{(1),\alpha,\alpha^{\star}}_{(an,bn)} = \mathcal{T}^{(1),\alpha,\alpha^{\star}}_{(an,bn)} + \#\{\text{sinks between 1 and } bn\}$.

Therefore

$$
\begin{aligned}
\mathbb{P}\left\{D_n > an\delta\right\} \ \leq\ & \mathbb{P}\left(\mathcal{T}_{(an,bn)}^{(1),\alpha,\alpha^\star} \leq n(a\alpha - \tfrac{1}{2}f(\delta))\right) \\
& + \mathbb{P}\left(\#\left\{\text{sinks}\right\} \leq n(b\alpha^\star - \tfrac{1}{2}f(\delta))\right) \\
& + \mathbb{P}\left(L_{(an,bn)}^{(1)}(\varepsilon) > n\left(a\alpha + b\alpha^\star - f(\delta)\right) \text{ for some } \varepsilon \text{ as above}\right) \\
\leq\ & \exp(-n\tilde{g}(\delta)) \\
& + \sum_{1\geq\varepsilon\geq\delta;\ \varepsilon an\in\mathbb{N}} \mathbb{P}\left(L_{(an,bn)}^{(1)}(\varepsilon) > n\left(\phi^{(1)}(a,b,\alpha(a,b)) - f(\delta)\right)\right) \\
\leq\ & \exp(-n\tilde{g}(\delta)) \\
& + \sum_{1\geq\varepsilon\geq\delta;\ \varepsilon an\in\mathbb{N}} \mathbb{P}\left(L_{(an,bn)}^{(1)}(\varepsilon) > n\left(\phi^{(1)}(a,b,\alpha(a,b)) - f(\varepsilon)\right)\right) \\
\leq\ & \exp(-n\tilde{g}(\delta)) + \sum_{1\geq\varepsilon\geq\delta;\ \varepsilon an\in\mathbb{N}} \exp(-ng(\varepsilon)) \\
\leq\ & \exp(-n\tilde{g}(\delta)) + an\exp(-ng(\delta))
\end{aligned}
$$

for large n, where we used Lemma 4.1 and where $\tilde{g}$ is some positive function. This implies the first part of the lemma. The convergence of D_n/n to 0 in probability and in L^1 follows, since the sequence is bounded. $\qquad\square$

We now conclude the proof of the lower bound noticing that we have

$$
\mathcal{L}_{(an,nb)}^{(1),\alpha,\alpha^\star} - D_n - D_n' \leq L_{(an,bn)}^{(1)}
$$

where D_n' is the number of sinks $(i,j) \in \pi_n$ with $i = 0$. By the previous lemma we know that D_n/n tends to 0 in L^1. By symmetry, the same result holds for D_n'/n. Taking expectations of both sides in the previous inequality we get

$$
\liminf_{n\to+\infty} \frac{1}{n}\mathbb{E}[L_{(an,bn)}^{(1)}] \geq a\alpha + b\alpha^\star = \phi^{(1)}(a,b,\alpha(a,b)).
$$

$\qquad\square$

Back to Ulam's constant. We observe that if we take $p = 1/n$, Theorem 1 for Model 1 suggests $L_n^{(1)} \approx 2\sqrt{n}$, which is consistent with the asymptotics of Ulam's problem, since ξ is then close, after renormalization, to a Poisson point process with intensity n.

In fact, one can rigorously recover that $c = 2$ using our proof of Theorem 1. To do so, consider a Poisson point process Ξ with intensity n in the unit square and denote by $\ell(n)$ the greatest number of points of Ξ an increasing path can go through. To get a lower bound and an upper bound of $\ell(n)$, we divide the square $[0,1]^2$ in two different ways.

First, we fix some $k \geq 1$ and we divide it into small squares of length side $1/(k\sqrt{n})$. Say that $\xi_j(i) = 1$ if at least one point of Ξ is in the square with top-right corner $(i/(k\sqrt{n}), j/(k\sqrt{n}))$ and consider the quantity $L_{k\sqrt{n}}^{(1)} := L_{(k\sqrt{n},k\sqrt{n})}^{(1)}$ associated to the family $(\xi_j(i))_{i,j\leq k\sqrt{n}}$. It is clear that

$$
\ell(n) \geq L_{k\sqrt{n}}^{(1)}. \tag{4.4}
$$

Denoting
$$p_k = \mathbb{P}(\xi_j(i) = 1) = \mathbb{P}\left(\text{Poiss}(1/k^2) \geq 1\right) = 1 - e^{-1/k^2},$$
Theorem 1.1 implies
$$\frac{L_{k\sqrt{n}}^{(1)}}{\sqrt{n}} \xrightarrow[n \to +\infty]{\text{a.s.}} \frac{2k\sqrt{p_k}}{\sqrt{p_k} + 1}.$$

Using (4.4) and letting k tend to infinity, we get $c := \lim \ell(n)/\sqrt{n} \geq 2$.

To prove the upper bound, we divide now the square into small squares of length side $1/n^4$. Say that $\xi_j(i) = 1$ if at least one point of Ξ is in the square with top-right corner $(i/n^4, j/n^4)$ and consider the quantity $L_{n^4}^{(1)} := L_{(n^4,n^4)}^{(1)}$ associated to the family $(\xi_j(i))_{i,j \leq n^4}$. The parameter of these Bernoulli random variables is now
$$\tilde{p}_n = \mathbb{P}\left(\text{Poiss}(1/n^7) \geq 1\right) = 1 - e^{-1/n^7}.$$

With probability higher than $1 - n^{-2}$, all the columns and lines of width $1/n^4$ contain at most one point of Ξ. On this event that we denote F_n, $L_{n^4}^{(1)}$ coincides with $\ell(n)$. We now use some intermediate results of the proof of the upper bound for $a = b = 1$. Inequality (3.7) still holds despite the dependence of $\tilde{p}_n$ on n *i.e.*
$$L_{n^4}^{(1)} \leq \mathcal{L}_{(n^4,n^4)}^{(1),\alpha_n,\alpha_n^\star},$$

with $\alpha_n = \alpha_n^\star = \frac{\sqrt{\tilde{p}_n} - \tilde{p}_n}{1 - \tilde{p}_n} \sim n^{-7/2}$. Using (4.2), we get

$$\mathbb{E}[\ell(n)] \leq \mathbb{E}[|\Xi| \mathbf{1}_{\bar{F}_n}] + \mathbb{E}[\mathcal{L}_{(n^4,n^4)}^{(1),\alpha_n,\alpha_n^\star}]$$
$$\leq \sqrt{\mathbb{E}[\text{Poiss}(n)^2]/n^2} + n^4(\alpha_n + \alpha_n^\star),$$

where we used the Cauchy-Schwarz inequality. Dividing by n and letting n go to infinity, we obtain $c \leq 2$.

5. Stationary measures of Hammersley's process on $\mathbb{Z}$

In this section, which can be read independently from the rest of the article, we study the analogous on the infinite line $\mathbb{Z}$ of the process $(X_t^1)_{t \geq 0}$ defined in (2.1). This infinite Hammersley process was introduced by T.Seppäläinen to prove the Law of Large Numbers for Problem 1.

In order to generalize $(X_t)_{t \geq 0}$ to a process taking its values in $\{0,1\}^{\mathbb{Z}}$ we need a construction that does not rely on Hammersley lines. As before, $(\xi_t(i))_{i \in \mathbb{Z}, t \in \mathbb{N}}$ denote i.i.d. Bernoulli(p) random variables we say that there is a *cross* at time t located at x if $\xi_t(x) = 1$. Informally, the infinite Hammersley process is defined as follows:

At time t, if there is a particle at x and if the particle immediately on its left is at y, then the particle at x jumps at time $t + 1$ at the leftmost cross of ξ_{t+1} in interval (y, x) (if any, otherwise it stays still).
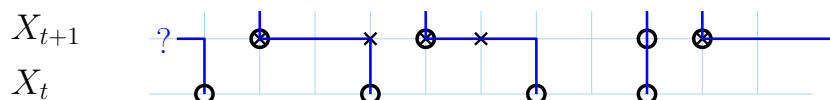
For our purpose, we need a construction of (X_t) in which we let the crosses *act* one by one on configurations. We first draw $X_0 \in \{0,1\}^{\mathbb{Z}}$ at random according to some distribution μ such that μ−almost surely, for all i there is $j < i$ such that $X_0(j) = 1$.

We now explain how to construct (X_{t+1}) from (X_t). A cross will *act* on a given configuration of particles in the following way: if a cross is located at x, then the

leftmost particle of (X_t) in the interval $[\![x, +\infty[\![$ (if any) moves to x; if there is no such particle then a particle is created at x.

With this definition, we can now construct the value of (X_{t+1}) as a function of (X_t) and the crosses at time $t+1$: We define X_{t+1} as the result of the *successive* actions on X_t of all crosses at time $t+1$ *from the right to the left*. Note that by definition, there only a finite number of crosses in ξ_{t+1} that can modify $X_t(i)$: those between $X_t(i)$ and the first particle on its left.

An example is drawn in this figure where the circles represent the particles and the crosses the locations where ξ_{t+1} equals 1. Here, we have $X^1_{t+1} = (1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 0, ...)$ (as in the previous sections, time goes from bottom to top in our pictures):

Theorem 5.1. *The only extremal translation-invariant stationary measures of Hammersley's process on $\mathbb{Z}$ are measures $\mathrm{Ber}(\alpha)^{\otimes \mathbb{Z}}$ for all $\alpha \in (0, 1]$.*

This theorem is precisely the discrete counterpart of (Aldous and Diaconis (1995), Lemma 7). The fact that Bernoulli product measures are stationary was proved by T.Seppäläinen (Seppäläinen (1997) Lemma 2.1). The reason for which we only focus on Model 1 on the infinite line is that the analogos of Model 2 is much more complicated to analyze. Indeed, the evolution of a particle in Model 2 depends on the whole configuration on its left. We don't know if a similar statement to Theorem 5.1 holds for Model 2.

In order to prove that there are no other extremal measures we adapt the proof of Jockush-Propp-Shor (Jockusch et al. (1995) p.17-20) for the discrete-time TASEP (see also Liggett (1985) Ch.VIII). The main difference here with Jockush-Propp-Shor is our Lemma 5.2 below that replaces their *block argument*.

Proof of extremality: Let $\alpha \in (0, 1]$ and μ be an extremal translation-invariant stationary measure on $\{0, 1\}^{\mathbb{Z}}$ with marginals $\mu(X(i) = 1) = \alpha$, we want to prove that $\mu = \mathrm{Ber}(\alpha)^{\otimes \mathbb{Z}}$.

The proof is made of the following steps:

(1) we introduce a measure $\underline{\pi}$ on $\{0, 1\}^{\mathbb{Z}} \times \{0, 1\}^{\mathbb{Z}}$ which is a *minimal* coupling between $\mathrm{Ber}(\alpha)^{\otimes \mathbb{Z}}$ and μ;

(2) we prove (Lemma 5.3) that some patterns that would decrease card $\{i; x_i \neq y_i\}$, have $\underline{\pi}$-probability zero;

(3) we conclude: $\mu = \mathrm{Ber}(\alpha)^{\otimes \mathbb{Z}}$.

Let $\mathcal{M}_\star$ be the set of translation-invariant measures on $\{0, 1\}^{\mathbb{Z}} \times \{0, 1\}^{\mathbb{Z}}$ that are a coupling of $\mathrm{Ber}(\alpha)^{\otimes \mathbb{Z}}$ and μ. The set $\mathcal{M}_\star$ is non-empty (it contains $\mathrm{Ber}(\alpha)^{\otimes \mathbb{Z}} \otimes \mu$) and compact for the weak topology.

We will prove that in $\mathcal{M}_\star$ there is a coupling $((X(i))_{i \in \mathbb{Z}}, (Y(i))_{i \in \mathbb{Z}})$ such that $X = Y$ a.s. Set

$$D: \begin{array}{ccc} \mathcal{M}_\star & \to & [0, 1] \\ \pi & \mapsto & \pi(X(0) \neq Y(0)). \end{array}$$

The function D is continuous on the compact set $\mathcal{M}_\star$ and thus attains its minimum $\delta \geq 0$ at some coupling $\underline{\pi}$. If $\delta = 0$, then the theorem is proved: by translation-invariance $X = Y$ $\underline{\pi}$-a.s.

Let $(X_0, Y_0) \sim \underline{\pi}$, we use the same random variables $(\xi_t(i))_{i \in \mathbb{Z}, t \in \mathbb{N}}$ to define the dynamics of X and Y. We denote by (X_t, Y_t) the pair of configurations at time t, its joint distribution is denoted by $\underline{\pi}_t$. By construction, for every t, $\underline{\pi}_t$ is translation-invariant. For every $k \geq 1$ we denote by

$$\Delta_k(X_t, Y_t)$$

the number of $i \in \{1, \ldots, k\}$ such that $X_t(i) \neq Y_t(i)$.

Let $n \geq 2$ and $t \geq 0$. We say that there is an $n-$forbidden pattern at location $x \in \mathbb{Z}$ at time t if the configuration (X_t, Y_t) between locations x and $x + n - 1$ is either

$$\begin{pmatrix} & & x & & & x+n-1 & \\ X_t = & (& 1 & 0 & \ldots 0 & 0 &) \\ Y_t = & (& 0 & 0 & \ldots 0 & 1 &) \end{pmatrix}$$

or

$$\begin{pmatrix} & & x & & & x+n-1 & \\ X_t = & (& 0 & 0 & \ldots 0 & 1 &) \\ Y_t = & (& 1 & 0 & \ldots 0 & 0 &) \end{pmatrix}.$$

Lemma 5.2. *For $n \geq 2$, let $E_{x,n}$ be the event $\xi_1(x) = 1$, $\xi_1(x+1) = \xi_1(x+2) = \cdots = \xi_1(x+n-1) = 0$. Let $j \geq 1$, denote by $F(n, j)$ the subset of locations of $\{0, n, 2n, \ldots, (j-1)n\}$ at which is located a forbidden pattern at time 0. Let*

$$A(n, j) = \mathrm{card}\{x \in F(n, j), \ E_{x,n} \ occurs\}.$$

Then:

$$\Delta_{jn}(X_1, Y_1) \leq \Delta_{jn}(X_0, Y_0) + 1 - 2A(n, j).$$

Proof of Lemma 5.2: We let the crosses at time 1 act one by one from the right to the left. We consider the impact of each cross action on the discrepancy Δ_{jn} between the two configurations. Let us note that, when a cross located at i acts on some configuration of particles, it changes the value of at most two sites of the configuration: i and the leftmost 1 of the configuration in the interval $[\![i, \infty)$. Thus, we deduce the following facts:

- The crosses strictly to the right of jn have no impact on the discrepancy.
- A cross located at a point x in $\{1, \ldots, jn\}$ cannot increase the discrepancy. To check this fact, one can study all the cases. They are all shown below, up to symmetries between X and Y.

$$\begin{pmatrix} & x & \\ (& 1 &) \\ (& 1 &) \end{pmatrix}, \begin{pmatrix} & x & & & & \\ (& 1 & ? & \cdots & ? & 1 &) \\ (& 0 & 0 & \cdots & 0 & 1 &) \end{pmatrix},$$

$$\begin{pmatrix} & x & & & & \\ (& 1 & ? & \cdots & ? & 0 &) \\ (& 0 & 0 & \cdots & 0 & 1 &) \end{pmatrix}, \begin{pmatrix} & x & & & \\ (& 0 & \cdots & 0 & 1 &) \\ (& 0 & \cdots & 0 & 1 &) \end{pmatrix},$$

$$\begin{pmatrix} & x & & & & & & & \\ (& 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 &) \\ (& 0 & \cdots & 0 & 1 & ? & \cdots & ? & 1 &) \end{pmatrix}, \begin{pmatrix} & x & & & & & & & \\ (& 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 &) \\ (& 0 & \cdots & 0 & 1 & ? & \cdots & ? & 0 &) \end{pmatrix}.$$

- For a forbidden pattern at some in, if $E_{in,n}$ occurs, the cross at in decreases the discrepancy by 2.
- In the worst case, the right-most cross which is strictly to the left of 1 increases the discrepancy by 1.
- The crosses to the left of the previous one have no impact on the discrepancy.

The result follows. □

Lemma 5.3 (Forbidden patterns). *Let $n \geq 2$. The $\underline{\pi}_0$ probability of an n-forbidden pattern at any location is 0.*

Proof of Lemma 5.3: Let $j \geq 1$. By Lemma 5.2 we have:

$$\Delta_{jn}(X_1, Y_1) \leq \Delta_{jn}(X_0, Y_0) + 1 - 2A(n, j).$$

Note that the probability that, at some location in, there is a n-forbidden pattern and that $E_{in,n}$ occurs is $\zeta p(1-p)^{n-1}$ where ζ is the probability of a n-forbidden pattern at a given location. We aim to prove $\zeta = 0$. Taking expectation in the previous display we get

$$jnD(\underline{\pi}_1) \leq jnD(\underline{\pi}_0) + 1 - 2j\zeta p(1-p)^{n-1}.$$

Dividing by nj and letting $j \to \infty$ we get:

$$D(\underline{\pi}_1) \leq D(\underline{\pi}_0) - \frac{2\zeta p(1-p)^{n-1}}{n}.$$

But the minimality of $\underline{\pi}$ yields $D(\underline{\pi}_1) \geq D(\underline{\pi}_0)$ and we get $\zeta = 0$. □

Lemma 5.4 (There is a forbidden pattern somewhere). *Assume that $\delta > 0$, there exists n such that*

$$\underline{\pi} \left(\begin{array}{cccccccc} & & 1 & & & n & \\ X = & (& 1 & ? & \ldots & ? & 0 &) \\ Y = & (& 0 & ? & \ldots & ? & 1 &) \end{array} \right) > 0 \quad or$$

$$\underline{\pi} \left(\begin{array}{cccccccc} & & 1 & & & n & \\ X = & (& 0 & ? & \ldots & ? & 1 &) \\ Y = & (& 1 & ? & \ldots & ? & 0 &) \end{array} \right) > 0$$

Proof: Let $(X, Y) \sim \underline{\pi}$,

$$\delta = \underline{\pi}(X(0) \neq Y(0)) = \underline{\pi}(X(0) = 1, Y(0) = 0) + \underline{\pi}(X(0) = 0, Y(0) = 1).$$

The two terms on the right side are equal:

$$\underline{\pi}(X(0) = 1, Y(0) = 0)$$
$$= \underline{\pi}(X(0) = 1) - \underline{\pi}(X(0) = 1, Y(0) = 1)$$
$$= \underline{\pi}(Y(0) = 1) - \underline{\pi}(X(0) = 1, Y(0) = 1) \quad (X, Y \text{ have same marginals})$$
$$= \underline{\pi}(X(0) = 0, Y(0) = 1).$$

Then $0 < \delta/2 = \underline{\pi}(X(0) = 1, Y(0) = 0)$. Assume that the lemma is false (for every n), we have $\underline{\pi}(A \cup B) = 1$, where

$$A = \{(x, y), x_i \leq y_i \text{ for all } i \in \mathbb{Z}\},$$
$$B = \{(x, y), x_i \geq y_i \text{ for all } i \in \mathbb{Z}\}.$$

As seen above, $\delta > 0$ implies then that $\underline{\pi}(A), \underline{\pi}(B \backslash A) > 0$. Besides, it is easy to check that A and $B \backslash A$ are preserved by the dynamics. Hence, considering the restriction of $\underline{\pi}$ on these two subsets, $\underline{\pi}_A := (\nu_A, \mu_A)$ and $\underline{\pi}_{B \backslash A} := (\nu_{B \backslash A}, \mu_{B \backslash A})$, we see that ν_A (resp. μ_A) and $\nu_{B \backslash A}$ (resp. $\mu_{B \backslash A}$) define two translation-invariant stationary measures such that

$$
\begin{aligned}
\mathrm{Ber}(\alpha)^{\otimes \mathbb{Z}} &= \underline{\pi}(A)\nu_A + \underline{\pi}(B \backslash A)\nu_{B \backslash A}. \\
\mu &= \underline{\pi}(A)\mu_A + \underline{\pi}(B \backslash A)\mu_{B \backslash A}.
\end{aligned}
$$

The extremality of $\mathrm{Ber}(\alpha)^{\otimes \mathbb{Z}}$ and μ implies that $\mathrm{Ber}(\alpha)^{\otimes \mathbb{Z}} = \nu_A = \nu_{B \backslash A}$ and $\mu = \mu_A = \mu_{B \backslash A}$. But, by definition of A and B, $\nu_A \preccurlyeq \mu_A$ and $\mu_{B \backslash A} \preccurlyeq \nu_{B \backslash A}$. Thus, necessarily, $\mu = \mathrm{Ber}(\alpha)^{\otimes \mathbb{Z}}$. $\qquad\square$

Now we can obtain our contradiction. Indeed, there exists an integer n such that

$$
\underline{\pi}\left(
\begin{array}{lccccccc}
 & & & 1 & & & n & \\
X = & (& 1 & ? & \ldots & ? & 0 &) \\
Y = & (& 0 & ? & \ldots & ? & 1 &)
\end{array}
\right) > 0 \quad \text{or}
$$

$$
\underline{\pi}\left(
\begin{array}{lccccccc}
 & & & 1 & & & n & \\
X = & (& 0 & ? & \ldots & ? & 1 &) \\
Y = & (& 1 & ? & \ldots & ? & 0 &)
\end{array}
\right) > 0.
$$

Without loss of generality we do the first case. We can assume that X, Y coincide between 2 and $n-1$, otherwise we can decrease n. Let $1 < k < n$ be the leftmost position at which $\binom{X(k)}{Y(k)} = \binom{1}{1}$.

With positive probability we can turn this $\binom{1}{1}$ into $\binom{1}{0}$:

$$
\left(
\begin{array}{lccccccccc}
 & & \times & - & - & - & - & - & - & \\
 & & 1 & & & & k & & n & \\
X_t = & (& 1 & 0 & \ldots & 0 & 1 & ? & 0 &) \\
Y_t = & (& 0 & 0 & \ldots & 0 & 1 & ? & 1 &)
\end{array}
\right) \xrightarrow{t+1}
$$

$$
\left(
\begin{array}{lcccccccc}
 & & 1 & & & & k & & n \\
X_{t+1} = & (& ? & 0 & \ldots & 0 & 1 & ? & 0 &) \\
Y_{t+1} = & (& ? & 0 & \ldots & 0 & 0 & ? & 1 &)
\end{array}
\right)
$$

We repeat this process between positions k and n and finally for some t, i, j we have

$$
\underline{\pi}\left(
\begin{array}{lcccccccc}
 & & & i & & & & j & \\
X_t = & (& 1 & 0 & \ldots & 0 & 0 &) \\
Y_t = & (& 0 & 0 & \ldots & 0 & 1 &)
\end{array}
\right) > 0,
$$

which contradicts Lemma 5.3. $\qquad\square$

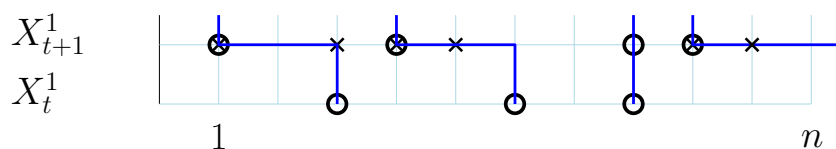## Appendix A. Dynamic of the underlying interacting particle system

In this appendix we explicit the dynamic of the underlying interacting particle system in Models 1 and 2. Consider the processes $(X_t^i)_{t \geq 0}$ defined by (2.1) (without sources and sinks). One can explicitly construct the dynamic of these processes in term of the family of i.i.d. random variables $\{\xi_t(x), x \in [\![1, n]\!], t \in \mathbb{N}\}$ with law Bernoulli(p) (see (3.1)). To do so, let say that there is a *cross* at time t located at x if $\xi_t(x) = 1$. A cross will *act* on a given configuration of particles in the following

way : if a cross is located at x, then the left-most particle in the interval $[\![x, n]\!]$ (if any) moves to x ; if there is no such particle then a particle is created at x.

With this definition, we can now construct the value of X_{t+1}^i as a function of X_t^i and the crosses at time $t + 1$:
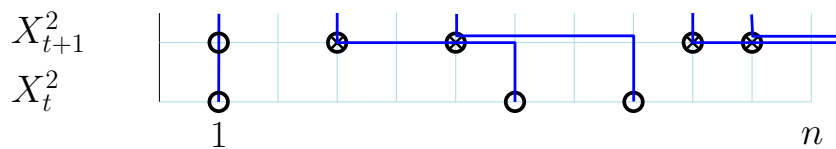
Model 1. We define X_{t+1}^1 as the result of the *successive* actions on X_t^1 of all crosses at time $t + 1$ *from the right to the left*.

An example is drawn in this figure where the circles represent the particles and the crosses the locations where ξ_{t+1} equals 1. Here, we have $X_{t+1}^1 = (1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 0)$ (as it is usual in the literature on Hammersley's processes, time goes from bottom to top in our pictures):

Let us note that there is an alternative way to construct the model 1: At time t, if there is a particle at x and if the particle immediately on its left is at y, then the particle at x jumps at time $t + 1$ at the leftmost cross of ξ_{t+1} in interval (y, x) (if any, otherwise it stays still). If there are some points of ξ_{t+1} at the right of the rightmost particle of X_t^1 then a new particle appears at the leftmost point among them.

Model 2. We define X_{t+1}^2 as the result of the *successive* actions on X_t^2 of all crosses at time $t + 1$ *from the left to the right*.

We let the reader convince themselves that this construction coincides with the definition given in (2.1).

Although the definition of both models seems, at first glance, very close, the nature of the two processes is in fact quite different. Indeed, in the first model, to find the location at time $t + 1$ of a particle located at time t at x, one just need to know the location y of the particle immediately on his left at time t and the position of the crosses of ξ_{t+1} in the interval (y, x). In particular, with no difficulty, one can define a process with similar transitions on the whole line $\mathbb{Z}$ as it is done in Seppäläinen (1997). For Model 2, the dependances are more intricate. Indeed, to determine the location at time $t + 1$ of a particle located at time t at x, one need to know the whole configuration of X_t^2 and ξ_{t+1} on the interval $[\![1, x]\!]$. In particular, the definition on the whole line $\mathbb{Z}$ of a similar process is more delicate and requires a condition between the density of crosses and particles (see Seppäläinen (1998)).

Finally, let us note that the construction can be extended to the models with sources and sinks. The set of sources simply give the initial configuration of the particle system. The sinks acts as the crosses. In Model 1 sinks act after the other crosses while in Model 2 sinks act before the other crosses.

References

D. Aldous and P. Diaconis. Hammersley's interacting particle process and longest increasing subsequences. *Probab. Theory Related Fields* **103** (2), 199–213 (1995). MR1355056.

E. Cator and P. Groeneboom. Hammersley's process with sources and sinks. *Ann. Probab.* **33** (3), 879–903 (2005). MR2135307.

P. Groeneboom. Hydrodynamical methods for analyzing longest increasing subsequences. *J. Comput. Appl. Math.* **142** (1), 83–105 (2002). Probabilistic methods in combinatorics and combinatorial optimization. MR1910520.

J.M. Hammersley. A few seedlings of research. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. I: Theory of statistics*, pages 345–394. Univ. California Press, Berkeley, Calif. (1972). MR0405665.

W. Jockusch, J. Propp and P. Shor. Random domino tilings and the arctic circle theorem. *ArXiv Mathematics e-prints* (1995). arXiv: 9801068.

K. Johansson. Discrete orthogonal polynomial ensembles and the Plancherel measure. *Ann. of Math. (2)* **153** (1), 259–296 (2001). MR1826414.

T.M. Liggett. *Interacting particle systems*, volume 276 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York (1985). ISBN 0-387-96069-4. MR776231.

L.T. Rolla, V. Sidoravicius, D. Surgailis and M.E. Vares. The discrete and continuum broken line process. *Markov Process. Related Fields* **16** (1), 79–116 (2010). MR2664337.

T. Seppäläinen. Increasing sequences of independent points on the planar lattice. *Ann. Appl. Probab.* **7** (4), 886–898 (1997). MR1484789.

T. Seppäläinen. Exact limiting shape for a simplified model of first-passage percolation on the plane. *Ann. Probab.* **26** (3), 1232–1250 (1998). MR1640344.

T. Seppäläinen. Lecture notes on the corner growth model (2009). Available at http://www.math.wisc.edu/~seppalai.