



**HAL**  
open science

# Bidimensional stationarity of random models in the plane

Alexandre Boyer

► **To cite this version:**

Alexandre Boyer. Bidimensional stationarity of random models in the plane. Probability [math.PR]. Université Paris-Saclay, 2022. English. NNT : 2022UPASM011 . tel-03783603

**HAL Id: tel-03783603**

**<https://theses.hal.science/tel-03783603v1>**

Submitted on 22 Sep 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Bidimensional stationarity of random  
models in the plane  
*Stationnarité bidimensionnelle de modèles aléatoires du  
plan*

**Thèse de doctorat de l'université Paris-Saclay**

École doctorale n°574 : mathématiques Hadamard (EDMH)  
Spécialité de doctorat : Mathématiques appliquées  
Graduate School : Mathématiques, Référent : Faculté des sciences  
d'Orsay

Thèse préparée dans le Laboratoire de Mathématiques d'Orsay (Université  
Paris-Saclay, CNRS), sous la direction d'Arvind SINGH, chargé de recherche, la  
co-direction de Nathanaël ENRIQUEZ, professeur

**Thèse soutenue à Paris-Saclay, le 1<sup>er</sup> juillet 2022, par**

**Alexandre BOYER**

**Composition du jury**

<b>Jean-François Le Gall</b> Professeur, Université Paris-Saclay	Président
<b>Quentin Berger</b> Maître de conférences, Sorbonne Université	Rapporteur & Examineur
<b>Irène Marcovici</b> Maîtresse de conférences, Université de Lorraine	Rapporteur & Examinatrice
<b>Nicolas Curien</b> Professeur, Université Paris-Saclay	Examineur
<b>Jean-Baptiste Gouéré</b> Professeur, Université de Tours	Examineur
<b>Arvind Singh</b> Chargé de recherche, Université Paris-Saclay	Directeur de thèse

**Titre :** Stationnarité bidimensionnelle de modèles aléatoires du plan.

**Mots clés :** Réversibilité markovienne, Percolation de dernier passage, Processus de Hammersley, Plus longues sous-suites croissantes.

**Résumé :** Dans le cadre de cette thèse, trois modèles ont été étudiés indépendamment. Ils ont en commun d'être des modèles aléatoires définis dans le plan et possédant une propriété de stationnarité bidimensionnelle. Le premier est le modèle de Hammersley stationnaire dans le quart de plan, introduit et étudié par Cator et Groeneboom. Nous présentons ici une preuve probabiliste des fluctuations gaussiennes dans le cas non critique. Le deuxième modèle peut être vu comme une version stationnaire du problème

d'O'Connell-Yor. La preuve de sa stationnarité est obtenue en introduisant une discrétisation de ce modèle dont nous montrons la stationnarité, puis en observant que cette stationnarité est préservée à la limite. Enfin, le troisième modèle est une classe générale de systèmes aléatoires de lignes brisées dans le quart de plan, dont on montre la réversibilité. Cette classe contient de nombreux processus classiques comme des modèles de percolation de dernier passage. La nouveauté ici est qu'un poids est associé à chaque ligne.

**Title :** Bidimensional stationarity of random models in the plane.

**Keywords :** Markov reversibility, Last Passage Percolation, Hammersley's process, Longest increasing subsequences.

**Abstract :** In this PhD thesis, three models have been independently studied. They all have in common to be random models defined in the plane and having a two-dimensional stationarity property. The first one is Hammersley's stationary model in the quarter plane, introduced and studied by Cator and Groeneboom. We present here a probabilistic proof the Gaussian fluctuations in the non-critical case. The second model can be seen as a stationary modification of O'Connell-Yor's problem. The proof of its

stationarity is obtained by introducing a discretisation of this model, by proving its stationarity and then by observing that this stationarity is preserved in the limit. Finally, the third model is a general class of random systems of horizontal and vertical weighted broken lines on the quarter plane whose distribution are proved to be reversible. This class of systems generalizes several classical processes of the same kind. The novelty comes here from the introduction of a weight associated with each line.

# Remerciements

Entamons ces remerciements par de loin les plus importants, ceux à mes directeurs de thèse. Merci infiniment à tous les deux Arvind et Nathanaël pour m'avoir fourni un encadrement idéal pendant ces quatre années. Jamais je ne serais arrivé à un tel résultat sans votre disponibilité, votre accompagnement à tous les niveaux, vos précieux conseils tant pratiques que mathématiques et bien sûr vos nombreuses idées aussi originales qu'efficaces pour surmonter les différents obstacles rencontrés en recherche et dans tout ce qui tourne autour. Plus spécifiquement, merci à toi Arvind pour ton encadrement de mon mémoire de M2 et à toi Nathanaël pour ton cours de M2 sur les Modèles solubles, qui m'ont tous les deux donnés un bon aperçu de ce qui m'attendait dans mon doctorat à venir, et m'ont conforté dans le choix du sujet de thèse qui m'avait été proposé.

Je tiens à adresser des remerciements tous particuliers à chacun des membres du jury. Merci beaucoup à Quentin Berger et à Irène Marcovici pour avoir accepté de rapporter mon manuscrit, et pour avoir fourni des rapports aussi positifs. Merci beaucoup à Nicolas Curien, à Jean-François Le Gall et à Jean-Baptiste Gouéré pour avoir accepté de faire partie du jury, j'en suis très honoré. Merci également à Nicolas et à Jean-François pour leurs cours de M2 dont la réputation n'est plus à faire, et à Jean-Baptiste pour son implication dans les différentes journées de l'ANR PPPP auxquelles j'ai eu la chance de participer.

Un grand merci également à toi Jérôme, ex-postdoc fraîchement promu MCF, pour m'avoir autant épaulé et accompagné pendant mon doctorat. Nos échanges étaient toujours très riches et m'ont beaucoup appris.

Merci à tous les collègues et amis pour qui et avec qui j'ai pu enseigner durant ces quatre années, et qui m'ont beaucoup aidé à faire mes armes dans ce domaine, notamment Marie-Anne, Olivier et Fabien. Merci également à Jean-Baptiste L, Nathalie, Pascal, Thomas, Florian, Pierre B, Simeng, Timothée, Jérémie, Anthony, Perrine, Zacharie, Aurélien, Baptiste...

En parlant d'enseignement, je remercie aussi les différents professeurs et chargés de TD que j'ai eus pendant ma scolarité et mes études, et qui ont entretenu l'appétence que j'avais pour les mathématiques puis les probabilités. Je pense surtout à Richard Revret en sup, Bernard Lemaire en spé, Alain Trouvé en L3 et Justin Salez en M1, en plus de ceux cités ci-dessus en M2. Je remercie également Olivier Goubet et Frédéric Paccaut pour leur encadrement en stage de M1 au LAMFA dont je garde des très bons souvenirs.

Un merci spécial à toi Timothée pour ces très nombreux moments passés ensemble, que ce soit quotidiennement à l'IMO pendant deux ans ou à distance pendant la pandémie ou depuis que nos chemins se sont séparés. On n'avait pas imaginé pas qu'on se retrouverait à faire une thèse dans le même bureau trois ans après notre stage en binôme au CMLA ! Heureusement

---

que tu étais là pour me remettre à niveau en statistiques et pour résoudre tous mes problèmes informatiques... Et histoire de ne pas faire très original, je te retourne les remerciements que tu m'avais adressés : "merci pour les cookies et les chocolats mais aussi plus généralement pour la compagnie, les discussions et les jeux du midi". Merci également à tous mes autres co-bureaux avec qui il était toujours très agréable de discuter : Pierre R, Raphaël, Wojciech, Étienne, Perrine, Louis et Jérémie.

Merci aussi à tous mes amis qui m'ont accompagné pendant ces quatre années bien qu'extérieurs à l'IMO, que ce soit à distance quotidiennement ou de temps en temps IRL. Merci Jean-Baptiste, Julien B, Céline, Julien F, Louis, Paul, Quentin, Nicolas, Sacha, Sylvain et Tristan, pour tous les innombrables bons moments passés ensemble. Merci également à mes ex-compatriotes cachanais Antoine, Janine, Nicolas, Noémie, Shmuel et Vianney, j'espère que l'on pourra continuer à se voir malgré mon départ d'Île-de-France. Je n'oublie pas non plus mes camarades de prépa : Alexandre, Antonin, Florian, Hugo, Nicolas (oui ça fait beaucoup de Nicolas)... Merci également aux franciliens qui m'ont ouvert leurs portes dès que le provincial que je suis a débarqué dans la région : Denise, Jean, Solenne, Cyril, Laureline et Nils.

Enfin, merci à toute ma famille pour m'avoir supporté dans tous les sens du terme, en premier lieu mes parents bien sûr. Merci pour votre soutien et pour avoir cru en moi pendant toute ma scolarité et mes études supérieures. Merci à mon frère pour tout ce qu'on a partagé ensemble. Merci à mes grands-parents et à mes grands-tantes dont j'ai toujours été très proche. Merci à mes tantes, oncles, cousins, cousines, qui m'ont eux aussi ouvert leurs portes à de nombreuses reprises, ainsi qu'à tout le reste de ma famille.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>7</b>
1.1	Problème d’Ulam . . . . .	7
1.1.1	Approche de Hammersley . . . . .	8
1.1.2	Première résolution : analyse combinatoire . . . . .	13
1.2	Les principaux modèles de percolation de dernier passage dirigée . . . . .	16
1.2.1	Espace continu . . . . .	17
1.2.2	Espace discret . . . . .	19
1.3	Des modèles stationnaires comme outils . . . . .	29
1.3.1	Exemple en espace continu . . . . .	29
1.3.2	Exemple en espace discret . . . . .	34
1.4	Fluctuations . . . . .	39
1.4.1	Approche analytique . . . . .	40
1.4.2	Approche probabiliste . . . . .	42
1.5	Résultats de la thèse . . . . .	45
1.5.1	Fluctuations gaussiennes pour les processus de Hammersley : une preuve probabiliste . . . . .	46
1.5.2	Une version stationnaire du problème d’O’Connell-Yor . . . . .	48
1.5.3	Systèmes de Poisson–Kirchhoff réversibles . . . . .	53
<b>2</b>	<b>Gaussian fluctuations of stationary Hammersley’s processes: a probabilistic proof</b>	<b>61</b>
2.1	Introduction and main result . . . . .	61
2.2	Goedesics do not leave the $x$ -axis too early . . . . .	64
2.3	Proof of Theorem 2.1 and of Corollary 2.1 . . . . .	69

<b>3</b>	<b>A stationary modification of O’Connell–Yor’s problem</b>	<b>73</b>
3.1	Introduction . . . . .	73
3.2	Main results . . . . .	75
3.3	A discrete model . . . . .	78
3.3.1	Hammersley’s lines and longest non-decreasing subsequences . . . . .	78
3.3.2	A stationary system . . . . .	81
3.4	From discrete to continuous: proof of Theorem 3.1 and Proposition 3.1 . . . . .	88
3.5	Localisation of the optimal path when $c = \sqrt{n}$ : proof of Proposition 3.2 . . . . .	94
3.6	Recovering the asymptotics for the largest eigenvalue of a GUE matrix . . . . .	103
<b>4</b>	<b>Reversible Poisson-Kirchhoff Systems</b>	<b>105</b>
4.1	Introduction . . . . .	106
4.2	Poisson-Kirchhoff systems . . . . .	107
4.3	Reversible Poisson-Kirchhoff systems . . . . .	111
4.3.1	Lebesgue case . . . . .	113
4.3.2	Discrete case . . . . .	116
4.3.3	General case . . . . .	118
4.4	Heuristic . . . . .	119
4.4.1	Crossing . . . . .	120
4.4.2	Horizontal turn versus vertical turn . . . . .	121
4.4.3	Spontaneous creation versus annihilation . . . . .	121
4.4.4	Split versus coalescence . . . . .	122
4.5	Proof of reversibility . . . . .	123
4.5.1	Proof of Theorem 4.1 with uniformly bounded rates . . . . .	123
4.5.2	Proof of Theorem 4.2 with uniformly bounded rates . . . . .	132
4.5.3	From the uniformly bounded rate case to the unbounded case . . . . .	135
4.6	Examples . . . . .	136
4.6.1	Potential function of a PKS . . . . .	136
4.6.2	List of examples . . . . .	137
4.6.3	Models with monotone potential (LPP) . . . . .	140
4.6.4	Bullet models . . . . .	141
4.6.5	Six-vertex model . . . . .	142
4.6.6	Gaussian and Poisson models . . . . .	143
4.7	Statistical properties of the tessellation . . . . .	143
4.8	Perspectives . . . . .	147
	<b>Bibliographie</b>	<b>148</b>

# Chapitre 1

## Introduction

### 1.1 Problème d’Ulam

Commençons par décrire un problème historiquement attribué à Ulam [Ula61], qui a motivé l’introduction des outils utilisés dans les modèles dits de *percolation de dernier passage* (PDP) qui seront présentés en section 1.2.

Pour  $n$  un entier naturel non nul, notons  $\mathfrak{S}_n$  l’ensemble des permutations de l’ensemble  $\llbracket 1, n \rrbracket$ . On s’intéresse aux sous-suites croissantes d’une permutation  $\sigma \in \mathfrak{S}_n$ , c’est-à-dire une sous-suite  $(\sigma(i_1), \dots, \sigma(i_k))$  telle que  $i_1 < \dots < i_k$  et  $\sigma(i_1) < \dots < \sigma(i_k)$ . L’entier  $k$  est appelé *longueur* de cette sous-suite. On notera  $\ell(\sigma)$  la longueur de la plus longue sous-suite croissante de la permutation  $\sigma$ . On définit de même une sous-suite décroissante comme vérifiant  $\sigma(i_1) > \dots > \sigma(i_k)$ .

Par exemple, pour  $n = 8$  et

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 5 & 8 & 4 & 2 & 6 & 1 & 7 \end{pmatrix}, \quad (1.1)$$

alors  $(3, 4, 6, 7)$  et  $(3, 5, 6, 7)$  sont deux sous-suites croissantes de  $\sigma$  de longueur 4. Comme il n’en existe pas de longueur 5, on a  $\ell(\sigma) = 4$ .

Le théorème suivant donne une minoration de la longueur de la plus longue sous-suite monotone :

**Théorème 1.1** (Théorème d’Erdős–Szekeres [ES35]). *Toute permutation de  $\mathfrak{S}_n$  admet une sous-suite croissante de longueur au moins  $\sqrt{n}$  ou une sous-suite décroissante de longueur au moins  $\sqrt{n}$ .*

Dans leur papier [ES35], Erdős et Szekeres prouvent une version légèrement plus forte : toute suite réelle de  $n^2 + 1$  éléments admet une sous-suite monotone de longueur au moins  $n + 1$ . Ils procèdent en exhibant une relation de récurrence sur la longueur minimale de la suite à considérer pour pouvoir toujours en extraire une sous-suite de longueur au moins  $n + 1$ . Nous détaillerons une autre preuve du Théorème 1.1 dans la sous-section suivante.

Le problème qui va nous intéresser est le suivant, et est comme mentionné ci-avant attribué à Ulam [Ula61] :



**Problème d’Ulam.** On se donne  $\sigma$  une permutation de  $\llbracket 1, n \rrbracket$  choisie uniformément au hasard. Quelle est la longueur typique de la plus longue sous-suite croissante de  $\sigma$  ?

Avant de nous intéresser à son étude, mentionnons une situation où ce problème apparaît. Dans un de leurs articles [AD99], Aldous et Diaconis font le lien avec le jeu de *patience sorting*, vu comme un type de solitaire ou de réussite (*patience* en anglais), basé sur le tri d’un jeu de cartes mélangé. Les  $n$  cartes ont des valeurs différentes, entre 1 et  $n$ . Le joueur prend les cartes dans l’ordre du paquet et doit former des piles sur la table en respectant la règle suivante :

- Une carte doit être placée sur une carte plus grande, ou doit être placée sur la table à droite des piles existantes pour former une nouvelle pile.

Le but du jeu est de faire le moins de piles possible. Aldous et Diaconis présentent l’algorithme glouton, qui consiste à placer chaque carte sur la pile la plus à gauche possible lorsque l’on a le choix. Cela revient à révéler les cartes au fur et à mesure, dans l’ordre du paquet, et donc à optimiser les décisions à chaque étape. Ils montrent que, de manière surprenante, cet algorithme est en fait optimal parmi tous les algorithmes, y compris parmi ceux s’autorisant à regarder l’intégralité du paquet. De plus, si on voit un paquet de cartes mélangé comme une permutation  $\sigma \in \mathfrak{S}_n$ , c’est-à-dire que  $\sigma(i)$  est le numéro de la carte en  $i$ -ème position, alors ils remarquent qu’avec cette stratégie il y aura exactement  $\ell(\sigma)$  piles, et se ramènent ainsi au problème d’Ulam.

Nous allons présenter dans cette section l’approche qui a été introduite par Hammersley dans son célèbre papier [Ham72] pour étudier ce problème, puis la première résolution par Vershik et Kerov [VK77] et Logan et Shepp [LS77] cinq ans plus tard.

### 1.1.1 Approche de Hammersley

Cette approche repose sur la représentation graphique d’une permutation  $\sigma$  dans le quart de plan  $(\mathbb{R}_+)^2$ . Pour cela, on le munit de la relation d’ordre partielle  $\prec$  définie par

$$(x, y) \prec (x', y') \iff (x < x' \text{ et } y < y'). \quad (1.2)$$

En d’autres termes, un point est supérieur à un autre pour cette relation d’ordre si le second est situé au nord-est du premier, si l’on assimile respectivement les quatre directions (haut, bas, gauche, droite) aux quatre points cardinaux (nord, sud, ouest, est).

Fixons une permutation  $\sigma \in \mathfrak{S}_n$ , et posons  $\mathcal{G} = \{(i, \sigma(i)) \mid i \in \llbracket 1, n \rrbracket\}$  son graphe que l’on représente dans le quart de plan, voir Figure 1.1a. Ainsi, une sous-suite croissante de  $\sigma$  correspond à un sous-ensemble de  $\mathcal{G}$  totalement ordonné pour la relation  $\prec$ . En rangeant ces points dans l’ordre croissant, on obtient un chemin (strictement) nord-est. On confondra les deux points de vue par la suite, et on parlera par abus de langage de “chemin croissant”.

Pour déterminer la ou les plus longues sous-suites croissantes de la permutation  $\sigma$ , on pose dans un premier temps

$$G_1 = \{(x, y) \in \mathcal{G} \mid \forall (x', y') \in \mathcal{G}, (x', y') \not\prec (x, y)\},$$

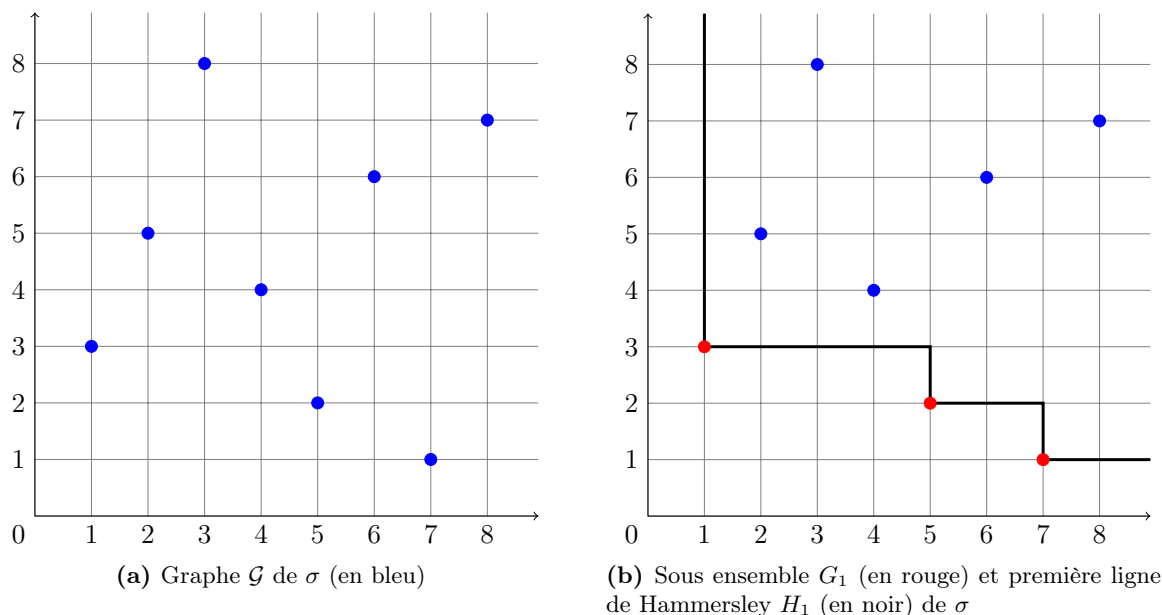


FIGURE 1.1 : Représentation graphique de la permutation  $\sigma$  définie par l'équation (1.1)

c'est-à-dire l'ensemble des points du graphe  $\mathcal{G}$  tel qu'il n'existe aucun point de  $\mathcal{G}$  (strictement) à leur sud-ouest, voir Figure 1.1b. Ainsi, une plus longue sous-suite croissante de  $\sigma$  commence nécessairement par un point de l'ensemble  $G_1$ . En effet, si l'on considère une telle sous-suite, il ne peut y avoir aucun point au sud-ouest du premier point, sinon on aurait pu l'ajouter à la sous-suite pour obtenir une sous-suite croissante de longueur strictement plus importante.

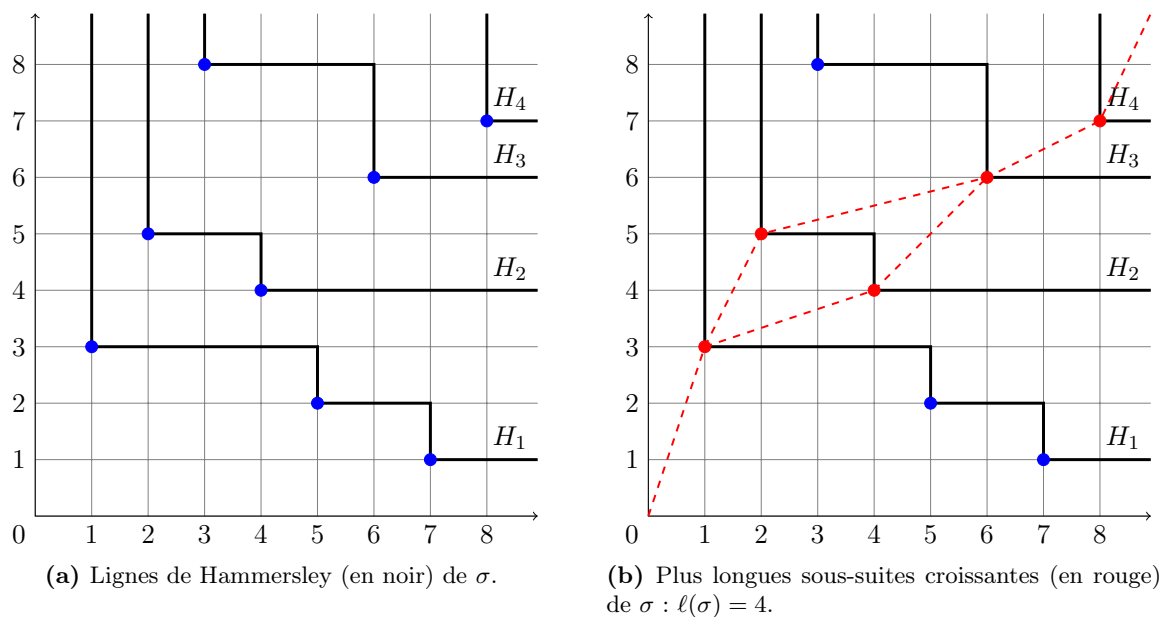
Si un point n'est pas dans  $G_1$ , alors il admet un point de  $\mathcal{G}$  à son sud-ouest. Si ce point n'est pas dans  $G_1$ , alors il admet lui aussi un point de  $\mathcal{G}$  à son sud-ouest. Comme  $\mathcal{G}$  est fini, alors en itérant ce procédé, on finira par arriver sur un point de  $G_1$ . Par transitivité, on conclut que si un point n'est pas dans  $G_1$ , alors il admet un point de  $G_1$  à son sud-ouest.

Ainsi, il devient naturel de considérer l'ensemble des points du quart de plan  $(\mathbb{R}_+)^2$  qui admettent un point de  $G_1$  situé à leur sud-ouest. Cet ensemble est l'union des quarts de plan nord-est ancrés en chacun des points de  $G_1$ . Il est caractérisé par sa frontière  $H_1$ , qui est donc une ligne brisée constituée uniquement de demi-droites et segments verticaux et horizontaux. Une telle ligne est appelée *ligne de Hammersley*, également représentée dans l'exemple de la Figure 1.1b. Le sous-ensemble des points de  $\mathcal{G}$  qui sont strictement au nord-est de cette ligne correspond exactement à l'ensemble  $\mathcal{G} \setminus G_1$ .

On itère ce procédé en remplaçant l'ensemble  $\mathcal{G}$  par l'ensemble  $\mathcal{G} \setminus G_1$  pour définir toutes les lignes de Hammersley. Plus précisément, on applique l'algorithme récursif suivant :

On pose  $\mathcal{G}_0 = \mathcal{G}$ , et pour  $k \in \mathbb{N}$  :

- Si  $\mathcal{G}_k = \emptyset$ , alors l'algorithme termine et renvoie l'entier  $k$  et les  $k$  ensembles  $(G_1, \dots, G_k)$ .



**FIGURE 1.2 :** Lignes de Hammersley et plus longues sous-suites croissantes de la permutation  $\sigma$  définie par l'équation (1.1).

- Sinon, on pose alors

$$G_{k+1} = \{(x, y) \in \mathcal{G}_k \mid \forall (x', y') \in \mathcal{G}_k, (x', y') \not\prec (x, y)\},$$

$$G_{k+1} = \mathcal{G}_k \setminus G_{k+1}.$$

Considérons l'entier  $k$  et les  $k$  sous-ensembles de  $\mathcal{G}$  ( $G_1, \dots, G_k$ ) renvoyés par l'algorithme, et remarquons qu'ils définissent par construction une partition du graphe  $\mathcal{G}$ . Pour chaque sous-ensemble  $G_i$ ,  $i \in \llbracket 1, k \rrbracket$ , on définit sa ligne de Hammersley associée  $H_i$  comme ci-dessus, c'est-à-dire comme la frontière de l'ensemble des points du quart de plan  $(\mathbb{R}_+)^2$  qui admettent un point de  $G_i$  situé à leur sud-ouest. L'ensemble des lignes de Hammersley relatives à la permutation  $\sigma$  définie par l'équation (1.1) est représenté en Figure 1.2a.

On peut alors établir le résultat crucial suivant qui justifie l'intérêt des lignes de Hammersley :

**Lemme 1.1** ([Ham72]). *La longueur  $\ell(\sigma)$  de la plus longue sous-suite croissante de la permutation  $\sigma$  est égale au nombre  $k$  de lignes de Hammersley.*

*Démonstration.* Montrons que cette longueur est plus petite et plus grande que  $k$ .

- Les termes d'une sous-suite croissante de  $\sigma$  doivent appartenir à des lignes de Hammersley différentes. En effet, par construction les lignes de Hammersley définissent des sous-suites décroissantes de  $\sigma$ . Une sous-suite croissante est donc de longueur au plus  $k$ .
- Montrons qu'il existe une sous-suite croissante de longueur  $k$ , en sélectionnant un point dans chacun des ensembles  $G_j$ . De manière récursive, en commençant par la fin :

- Choisissons un point  $(i_k, \sigma(i_k))$  quelconque de  $G_k$ .
- Pour  $1 < j \leq k$ , on suppose construits les points  $(i_j, \sigma(i_j)), \dots, (i_k, \sigma(i_k))$  définissant une sous-suite croissante, et appartenant à  $G_j, \dots, G_k$  respectivement. Par construction, il existe un point de  $G_{j-1}$  qui est situé au sud-ouest du point  $(i_j, \sigma(i_j))$ . On choisit  $(i_{j-1}, \sigma(i_{j-1}))$  un tel point, qui prolonge ainsi par la gauche la sous-suite croissante.

Ainsi,  $(\sigma(i_1), \dots, \sigma(i_k))$  est une sous-suite croissante. La Figure 1.2b illustre les sous-suites croissantes possibles dans l'exemple de la Figure 1.2a, obtenues grâce à ce procédé.  $\square$

On peut en déduire une preuve simple du Théorème 1.1 :

*Démonstration du Théorème 1.1.* On considère toujours une permutation  $\sigma \in \mathfrak{S}_n$ , et l'entier  $k$  et les ensembles  $(G_1, \dots, G_k)$  renvoyés par l'algorithme de Hammersley. Deux cas de figure peuvent se présenter.

- Si  $k \geq \sqrt{n}$ , alors il existe une sous-suite croissante de longueur  $k$  d'après le Lemme 1.1, donc de longueur au moins  $\sqrt{n}$ .
- Si  $k \leq \sqrt{n}$ , montrons qu'il existe une sous-suite décroissante de longueur au moins  $\sqrt{n}$ . Comme les  $(G_i)$  définissent une partition de  $\mathcal{G}$ , on a  $\sum_{i=1}^k |G_i| = n$ . D'après le principe des tiroirs, il existe un entier  $j \in \llbracket 1, n \rrbracket$  tel que  $|G_j| \geq \sqrt{n}$ . En considérant l'ensemble des points de  $G_j$ , on définit une sous-suite décroissante de  $\sigma$ , de longueur au moins  $\sqrt{n}$ .  $\square$

Revenons au problème d'Ulam et intéressons-nous maintenant à la partie aléatoire. Une manière de choisir  $\sigma \in \mathfrak{S}_n$  uniformément au hasard est de considérer  $n$  points uniformément au hasard dans le carré  $[0, 1]^2$ . Ainsi, si l'on étiquette les  $n$  points  $1, 2, \dots, n$  par abscisse croissante, les ordonnées respectives définissent  $\sigma(1), \sigma(2), \dots, \sigma(n)$  en considérant leurs positions relatives, de 1 à  $n$  également. Remarquons qu'il est toujours possible en procédant exactement de la même manière de construire les lignes de Hammersley. Il n'y a d'ailleurs pas d'ambiguïté sur le caractère strict ou large des inégalités étant donné que presque sûrement tous les points ont des abscisses et ordonnées deux à deux distinctes.

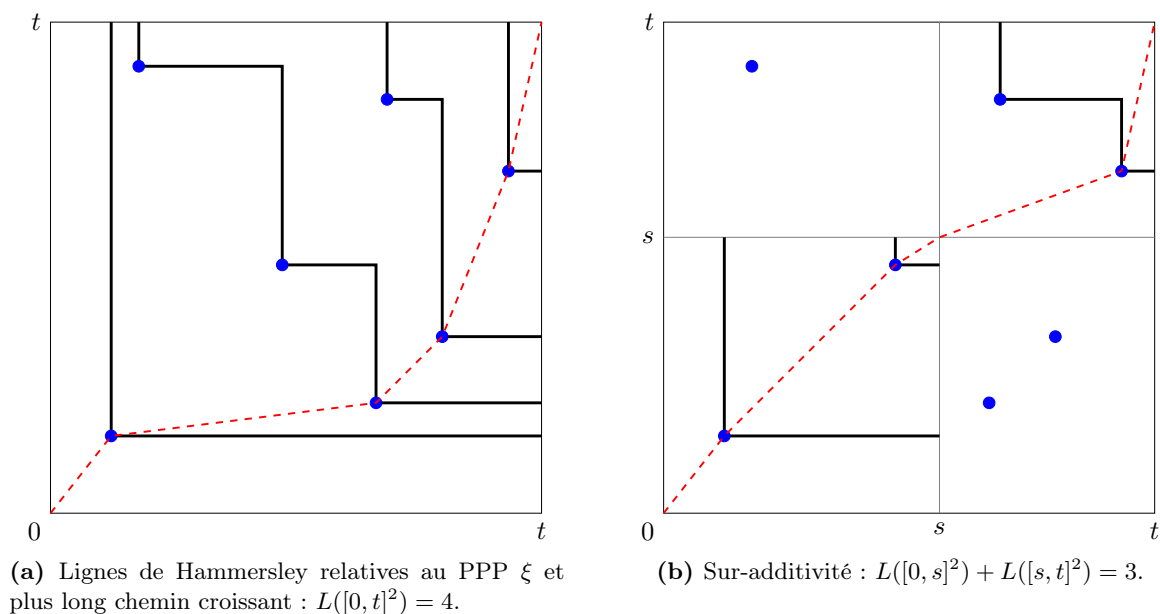
Cependant, ce procédé ne fournit pas d'indépendance spatiale. Par exemple, si  $n$  points sont situés sur une partie du carré, il est certain que l'autre partie n'en contient pas. Une des principales idées de Hammersley [Ham72] est d'utiliser un processus ponctuel de Poisson (PPP) dans ce carré, ce qui revient finalement à rendre également aléatoire le nombre total de points définissant la permutation.

Plus précisément, considérons un PPP  $\xi$  d'intensité 1 dans le quart de plan  $(\mathbb{R}_+)^2$ , et pour un rectangle  $R \subset (\mathbb{R}_+)^2$ , notons  $L(R)$  la variable aléatoire correspondant à la longueur du plus long chemin croissant constitué de points de l'ensemble  $\xi \cap R$ . Là encore, on peut utiliser la construction des lignes de Hammersley pour déterminer cette longueur, voir Figure 1.3a.

Par ailleurs, ce processus vérifie une propriété de sur-additivité : de manière déterministe,

$$\forall 0 \leq s \leq t, \quad L([0, t]^2) \geq L([0, s]^2) + L([s, t]^2).$$

En effet, la concaténation d'un chemin croissant de  $[0, s]^2$  et d'un chemin croissant de  $[s, t]^2$  est un chemin croissant de  $[0, t]^2$ , voir Figure 1.3b. De plus les propriétés du processus de Poisson


**FIGURE 1.3** : Exemple de PPP  $\xi$  dans  $[0, t]^2$ 

impliquent que les variables  $L([0, s]^2)$  et  $L([s, t]^2)$  sont indépendantes, et que le processus est invariant par translation :

$$\forall a > 0, \quad (L([s, t]^2))_{0 \leq s \leq t} \stackrel{(d)}{=} (L([s+a, t+a]^2))_{0 \leq s \leq t},$$

où le symbole  $\stackrel{(d)}{=}$  signifie “égal en loi”. Ainsi, en appliquant le théorème sous-additif de Kingman [Kin68] au processus  $(-L([0, t]^2))_{t \geq 0}$ , on obtient le théorème suivant.

**Théorème 1.2** ([Ham72]). *Il existe une constante  $c > 0$ , éventuellement infinie, telle que*

$$\frac{L([0, t]^2)}{t} \xrightarrow[t \rightarrow \infty]{p.s., L^1} c.$$

Cette application du théorème sous-additif de Kingman est présentée dans [Kin73] et a été reprise entre autres dans [Dur19, section 6.5]. Ils détaillent également comment ce modèle poissonnisé permet de répondre au problème original d’Ulam. Si l’on note  $\tau(n)$  le plus petit réel  $t > 0$  tel que le carré  $[0, t]^2$  contienne  $n$  points du PPP  $\xi$ , alors d’une part la variable  $L([0, \tau(n)]^2)$  a la même loi que la variable  $\ell(\sigma)$  pour  $\sigma \sim \text{Unif}(\mathfrak{S}_n)$ , et d’autre part la loi forte des grands nombres garantit l’asymptotique presque sûre  $\tau(n) \sim \sqrt{n}$  quand  $n \rightarrow \infty$ . Ainsi, on en déduit que pour  $n$  grand, la longueur de la plus longue sous-suite croissante d’une permutation typique de taille  $n$  est  $c\sqrt{n}$  où  $c$  est la constante du Théorème 1.2.

Cette constante  $c$  est longtemps restée inconnue. Toujours dans [Ham72], Hammersley montre qu’elle est finie et obtient même l’encadrement suivant :

$$\frac{\pi}{2} \leq c \leq e \quad \text{d’où} \quad 1.57 < c < 2.72.$$

Il obtient la minoration en construisant un chemin croissant depuis l'origine en considérant successivement le point admissible (c'est-à-dire au nord-est du précédent) le plus proche du précédent. La majoration s'obtient en majorant la probabilité qu'il existe une sous-suite croissante de longueur  $k$  dans  $\sigma$  par l'espérance du nombre de sous-suites croissantes de longueur  $k$  dans  $\sigma$ , qui vaut  $\frac{1}{k!} \binom{n}{k}$ . La constante  $e$  apparaît en prenant  $k$  de l'ordre de  $\sqrt{n}$  et en appliquant entre autres la formule de Stirling.

En utilisant des arguments similaires, Kingman [Kin73, section 2.4] affine les deux bornes obtenues par Hammersley pour la constante  $c$ . Si on note  $\delta$  l'unique solution strictement positive de

$$\log(1 + \delta) = \frac{2\delta}{1 + \delta}, \quad \text{d'où} \quad \delta \approx 3.92,$$

il obtient l'encadrement suivant :

$$\frac{2\sqrt{2}}{\pi} \leq c \leq \sqrt{\delta} + \frac{1}{\sqrt{\delta}} \quad \text{d'où} \quad 1.59 < c < 2.49.$$

Grâce à des calculs numériques sur ordinateur, Baer et Brock [BB68] ont conjecturé que  $c = 2$ . Hammersley [Ham72] a quant à lui proposé trois approches heuristiques pour déterminer la valeur exacte de  $c$  : deux d'entre elles donnent  $c = 2$  également, la troisième donne  $c = 2^{1/4}e^{1/2} \approx 1.961$ . Finalement, Logan et Shepp [LS77] montrent quelques années plus tard que  $c \geq 2$ , et indépendamment Vershik et Kerov [VK77] montrent la même année que  $c = 2$ , c'est l'objet de la sous-section suivante.

### 1.1.2 Première résolution : analyse combinatoire

Dans cette sous-section, nous présentons une heuristique de la preuve de  $c \geq 2$  par Logan et Shepp [LS77], qui repose sur le lien établi par Schensted [Sch61] entre tableau de Young et plus longue sous-suite croissante d'une permutation. Cette heuristique est également détaillée dans [AD99]. Vershik et Kerov [VK77] utilisent des arguments similaires pour montrer que  $c = 2$ .

Commençons par définir les notions de diagramme de Young et de tableau de Young standard.

**Définition 1.1.** • Un *diagramme de Young* (ou *diagramme de Ferrer*) d'ordre  $n$  est un arrangement de  $n$  cases en lignes et en colonnes, toutes justifiées en haut à gauche, et telles que la taille des lignes est décroissante de haut en bas.

- Un *tableau de Young standard* d'ordre  $n$  est un arrangement des entiers de 1 à  $n$  dans les cases d'un diagramme de Young d'ordre  $n$ , de telle sorte que les sous-suites relatives à chaque ligne et à chaque colonne soient croissantes de gauche à droite et de haut en bas respectivement. Le diagramme de Young sous-jacent s'appelle la *forme* du tableau de Young.

Un exemple de diagramme de Young et de tableau de Young standard est présenté en Figure 1.4. Il est également possible d'interpréter un diagramme de Young d'ordre  $n$  comme une partition de l'entier  $n$  en considérant les longueurs des lignes. Ainsi, le diagramme de Young de la Figure 1.4a est associé à la partition  $(3, 2, 2)$  de l'entier 7.

Le lien entre les tableaux de Young et les plus longues sous-suites croissantes d'une permutation mentionné plus haut est le suivant :

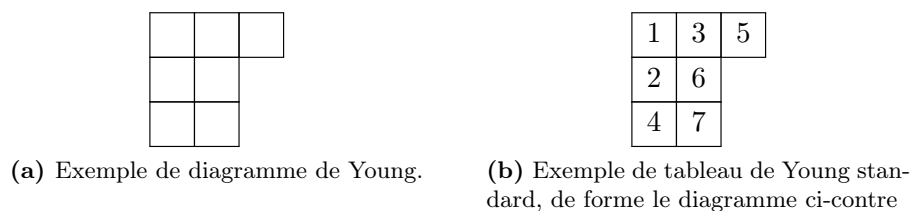


FIGURE 1.4 : Exemple de diagramme de Young et de tableau de Young standard d'ordre  $n = 7$

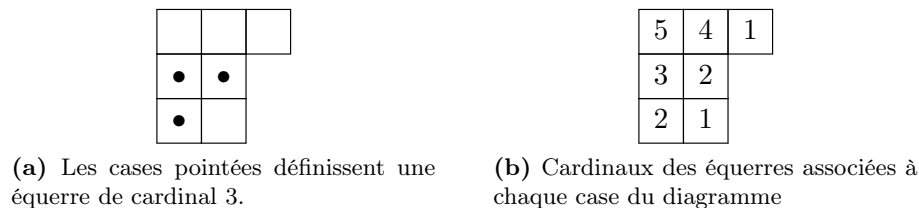


FIGURE 1.5 : Équerres du diagramme de Young de la Figure 1.4a.

**Théorème 1.3** ([Sch61]). *Le nombre de permutations de  $\mathfrak{S}_n$  ayant une plus longue sous-suite croissante de taille  $k \leq n$  est la somme, sur les diagrammes de Young  $\lambda$  d'ordre  $n$  comportant exactement  $k$  colonnes, des carrés du nombre de tableaux de Young standard de forme  $\lambda$ .*

Cette égalité est basée sur la bijection, appelée correspondance RSK (Robinson–Schentsed–Knuth) [Rob38, Sch61, Knu70], entre l'ensemble  $\mathfrak{S}_n$  et les couples de tableaux de Young standards d'ordre  $n$  de même forme. On montre alors que les deux tableaux de Young correspondants à une permutation  $\sigma$  donnée ont exactement  $\ell(\sigma)$  colonnes.

On souhaite donc dénombrer les tableaux de Young de forme donnée. Pour cela, on introduit la notion d'équerre :

**Définition 1.2.** Soit  $\lambda$  un diagramme de Young d'ordre  $n$ .

- On appelle *équerre* un sous-ensemble de cases du diagramme  $\lambda$  contenant les cases situées en dessous ou à droite d'une case donnée ( $y$  compris elle-même), le “ou” étant exclusif.
- On définit l'entier  $\pi(\lambda)$  comme étant le produit des cardinaux des équerres du diagramme  $\lambda$ .

Un exemple d'équerre du diagramme de Young  $\lambda$  de la Figure 1.4a est représenté en Figure 1.5a. Chaque équerre pouvant être associée à sa case en haut à gauche, on peut référencer les cardinaux des équerres dans les cases du diagramme comme effectué dans la Figure 1.5b. Ainsi, pour cet exemple,

$$\pi(\lambda) = 5 \times 4 \times 1 \times 3 \times 2 \times 2 \times 1 = 240.$$

Ce produit  $\pi(\lambda)$  joue un rôle essentiel dans le dénombrement des tableaux de Young de forme donnée  $\lambda$ , comme démontré par Frame, Robinson et Thrall [FRT54] :

**Théorème 1.4** (Formule des équerres [FRT54]). *Soit  $\lambda$  un diagramme de Young d'ordre  $n$ . Le nombre de tableaux de Young standard de forme  $\lambda$  est  $\frac{n!}{\pi(\lambda)}$ .*

Ainsi, si  $\lambda$  désigne le diagramme de Young d'ordre  $n = 7$  de la Figure 1.4a, le nombre de tableaux de Young de forme  $\lambda$  est

$$\frac{n!}{\pi(\lambda)} = \frac{7!}{240} = 21.$$

Pour un diagramme de Young  $\lambda$ , notons  $\lambda(0)$  la longueur de sa première ligne (ou autrement dit son nombre de colonnes). Les Théorèmes 1.3 et 1.4 donnent alors l'égalité

$$|\{\sigma \in \mathfrak{S}_n \mid \ell(\sigma) = k\}| = \sum_{\lambda(0)=k} \left( \frac{n!}{\pi(\lambda)} \right)^2,$$

où la somme est effectuée sur tous les diagrammes de Young  $\lambda$  d'ordre  $n$  et comportant  $\lambda(0) = k$  colonnes. Ainsi, si  $\sigma$  est choisi uniformément au hasard dans  $\mathfrak{S}_n$  de cardinal  $n!$ ,

$$\mathbb{P}(\ell(\sigma) = k) = \sum_{\lambda(0)=k} \frac{n!}{\pi(\lambda)^2}.$$

Par conséquent, si un diagramme de Young  $\Lambda$  d'ordre  $n$  est tiré aléatoirement selon la loi

$$\mathbb{P}(\Lambda = \lambda) = \frac{n!}{\pi(\lambda)^2},$$

appelée *mesure de Plancherel*, alors la longueur de sa première ligne  $\Lambda(0)$  est une variable aléatoire de même loi que  $\ell(\sigma)$ .

Pour étudier un diagramme de Young  $\lambda$  “typique” pour cette loi, il est naturel de chercher ceux qui minimisent  $\pi(\lambda)$ . Pour cela, il est plus commode de voir  $\lambda$  comme une fonction de  $\mathbb{R}_+$  dans  $\mathbb{R}_+$ . Si on écrit la partition correspondant au diagramme de Young comme un vecteur, pour  $x \in \mathbb{R}_+$ ,  $\lambda(x)$  correspondra à sa  $[x]$ -ème coordonnée si on les numérote à partir de 0. De même, pour  $y \in \mathbb{R}_+$ , on définit  $\lambda^{-1}(y)$  comme étant la  $[y]$ -ème coordonnée de la partition duale, c'est-à-dire en considérant cette fois les longueurs des colonnes. Comme pour les matrices, on place l'origine des coordonnées en haut à gauche du diagramme, l'axe des abscisses est vertical et pointe vers le bas, l'axe des ordonnées est horizontal et pointe vers la droite. La Figure 1.6 illustre cette représentation pour l'exemple du diagramme de Young de la Figure 1.4a.

Avec ce point de vue, le cardinal de l'équerre associée à la case de centre  $(x, y) \in (\mathbb{N}^2 + (\frac{1}{2}, \frac{1}{2}))$  est alors

$$\lambda(x) - y + \lambda^{-1}(y) - x.$$

Pour minimiser  $\pi(\lambda)$  défini comme un produit, on cherche en pratique à minimiser son logarithme. On définit la fonction  $f_\lambda$  de  $\mathbb{R}_+$  dans  $\mathbb{R}_+$  comme étant la fonction  $\lambda$  renormalisée comme suit :

$$\forall x \geq 0, \quad f_\lambda(x) = \frac{1}{\sqrt{n}} \lambda(x\sqrt{n}).$$



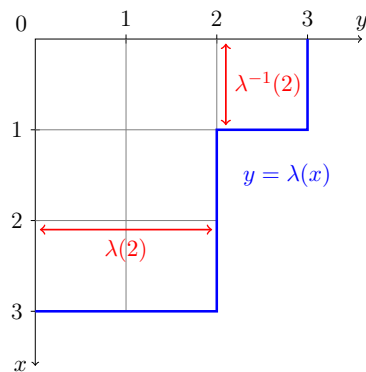


FIGURE 1.6 : Diagramme de Young de la Figure 1.4a vu comme une fonction de  $\mathbb{R}_+$  dans  $\mathbb{R}_+$ .

On définit également la fonctionnelle  $H$  : pour toute fonction  $f \in \mathcal{F}$ , où  $\mathcal{F}$  est l'ensemble des fonctions croissantes de  $\mathbb{R}_+$  dans  $\mathbb{R}_+$  d'intégrale 1,

$$H(f) = \int_0^\infty \int_0^{f(x)} \log(f(x) - y + f^{-1}(y) - x) dy dx.$$

Logan et Shepp [LS77] montrent alors la borne suivante pour  $n \rightarrow \infty$  :

$$\mathbb{P}(\Lambda = \lambda) \leq \frac{n!}{n^n} e^{-2nH(f_\lambda)} \sim \sqrt{2\pi n} e^{-2n(H(f_\lambda) + \frac{1}{2})}. \quad (1.3)$$

Cette borne suggère que la mesure de Plancherel charge davantage les diagrammes de Young  $\lambda$  pour lesquels  $H(f_\lambda)$  est petit. Il est donc naturel de chercher à minimiser la fonctionnelle  $H$  sur l'ensemble  $\mathcal{F}$ . Il se trouve que la fonction argument minimum  $f_*$  a pour support  $[0, 2]$  et vérifie

$$f_*(x) = \frac{2}{\pi}(\sin(\theta) - \theta \cos(\theta)), \quad x = f_*(x) + 2 \cos \theta, \quad 0 \leq \theta \leq \pi.$$

En particulier,  $f_*(0) = 2$ . Un diagramme de Young  $\lambda$  typique vérifierait donc  $f_\lambda(0) \approx 2$ , et donc  $\lambda(0) \approx 2\sqrt{n}$ . Comme la variable aléatoire  $\Lambda(0)$  a même loi que  $\ell(\sigma)$ , cet argument suggère que  $c = 2$ .

## 1.2 Les principaux modèles de percolation de dernier passage dirigée

Le modèle poissonnisé de Hammersley [Ham72] présenté en section 1.1.1 est en fait un cas particulier d'une classe de modèles plus générale, appelés modèles de *percolation de dernier passage dirigée*. Dans cette section nous allons présenter d'autres modèles de cette classe, séparés en deux catégories : ceux pour lesquels l'espace est continu, comme celui de Hammersley puisque les points sont un PPP dans tout le rectangle  $[0, 1]^2$ , et ceux pour lesquels l'espace est discret.

### 1.2.1 Espace continu

Comme évoqué ci-avant, le principal modèle de percolation de dernier passage en espace continu est celui de Hammersley [Ham72]. En 1995, Aldous et Diaconis [AD95] définissent le *processus de Hammersley* et prouvent en considérant sa limite hydrodynamique que  $c = 2$  par une méthode plus “légère” que celle de [VK77, LS77]. Dans cette sous-section nous allons étudier ce processus qui représente l’évolution de particules sur un segment donné.

Plus précisément, en écho à ce qui a été présenté ci-avant, on interprète la première coordonnée des points comme étant une coordonnée spatiale et la deuxième comme étant une coordonnée temporelle. On les notera généralement  $x$  et  $t$  respectivement. On souhaite définir l’ensemble des lignes de Hammersley pour le modèle poissonnisé de la sous-section 1.1.1 dans le rectangle  $[0, x^*] \times [0, t^*]$  pour  $(x^*, t^*) \in (\mathbb{R}_+)^2$ , et on notera  $L(x, t) = L([0, x] \times [0, t])$  pour tout point  $(x, t)$  du rectangle  $[0, x^*] \times [0, t^*]$  la longueur du plus long chemin entre les points  $(0, 0)$  et  $(x, t)$  avec les notations précédentes.

On interprète les lignes de Hammersley comme les “chemins espace-temps” de particules. En d’autres termes, l’intersection des lignes de Hammersley avec un segment horizontal  $[0, x^*] \times \{t\}$  définit la position des particules à l’instant  $t$  sur le segment  $[0, x^*]$ . Pour ne pas qu’il y ait d’ambiguïté dans la définition, on voit ici une ligne de Hammersley comme une union disjointe de segments verticaux semi-ouverts de la forme  $\{x\} \times [t, t'[,$  Les instants  $t$  où la ligne est constituée d’un segment horizontal vont correspondre aux instants de saut d’une particule existante ou d’arrivée d’une nouvelle particule.

De manière formelle, on peut définir ce processus de la manière suivante :

- Au temps  $t = 0$ , il n’y a aucune particule sur le segment  $[0, x^*]$ .
- À taux  $x^*$ , un point  $U$  est tiré uniformément au hasard sur le segment  $[0, x^*]$  indépendamment du passé.
  - S’il y a au moins une particule à droite de la position  $U$ , celle la plus proche est déplacée au point  $U$ .
  - Sinon, une nouvelle particule est créée au point  $U$ .

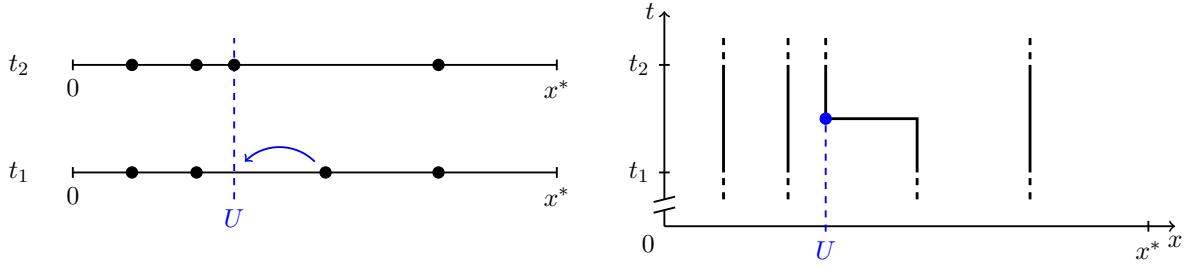
Un exemple de visualisation de ce processus est présenté en Figure 1.7.

De manière plus informelle, la règle d’évolution peut s’interpréter comme suit : pour chaque intervalle  $[x, x + dx]$  au temps  $t$ , avec probabilité  $dx dt$ , la particule la plus proche à droite de  $x$  est déplacée en  $x$  au temps  $t + dt$  s’il y en a, une nouvelle est créée sinon.

On peut ainsi représenter graphiquement les trajectoires des particules dans  $[0, x^*] \times [0, t^*]$ , en les complétant par des segments horizontaux aux instants de sauts (on considère que lorsqu’une nouvelle particule est créée, elle arrive de la droite) pour retrouver exactement les lignes de Hammersley. Ainsi, le nombre de lignes de Hammersley  $L(x, t)$  correspond au nombre de particules dans le segment  $[0, x]$  à l’instant  $t$ .

Par ailleurs, on peut souligner deux propriétés importantes du processus  $(L(x, t))_{x, t \geq 0}$  qui découlent de celles du processus de Poisson :

- **(Propriété de symétrie)**  $(L(x, t))_{x, t \geq 0} \stackrel{(d)}{=} (L(t, x))_{x, t \geq 0}$ ,



**FIGURE 1.7** : À gauche : exemple de deux configurations du processus de Hammersley aux instants  $t_1$  et  $t_2 > t_1$ , où exactement un saut a lieu entre ces deux temps. À droite : situation correspondante en termes de lignes de Hammersley.

- **(Propriété d'échelle)**  $\forall \kappa \in \mathbb{R}_+^*$ ,  $(L(x, t))_{x, t \geq 0} \stackrel{(d)}{=} (L(\kappa x, t/\kappa))_{x, t \geq 0}$ .

En particulier, la propriété d'échelle assure que la loi de la variable  $L(x, t)$  ne dépend que du produit  $tx$ . De plus, celle-ci permet d'obtenir un analogue de la convergence du Théorème 1.2 valable pour les rectangles :

$$\frac{L(x, t)}{\sqrt{tx}} \xrightarrow[t \rightarrow \infty]{\mathbb{P}} c.$$

Aldous et Diaconis [AD95] donnent une heuristique pour justifier l'égalité  $c = 2$ . Si pour  $x, t \geq 0$  on note  $D_{x,t}$  la distance entre l'abscisse  $x$  et la particule la plus proche à sa gauche (ou la distance entre l'abscisse  $x$  et 0 s'il n'y a aucune particule à gauche de  $x$ ) à l'instant  $t$ , alors dans l'intervalle de temps  $[t, t + dt]$  une nouvelle particule est créée dans le segment  $[0, x]$  à taux  $D_{x,t}$ . Plus précisément,

$$\frac{d}{dt} \mathbb{E}[L(x, t)] = \mathbb{E}[D_{x,t}].$$

Par ailleurs, supposons que le processus spatial au voisinage du point  $x$  au temps  $t$  peut être approché par un PPP d'une intensité  $\lambda(x, t)$ . Ainsi,

$$\mathbb{E}[D_{x,t}] \approx \lambda(x, t)^{-1} \approx \left( \frac{d}{dx} \mathbb{E}[L(x, t)] \right)^{-1}.$$

Finalement, la fonction  $(x, t) \mapsto \mathbb{E}[L(x, t)]$  satisfait approximativement cette équation aux dérivées partielles d'inconnue  $w : (\mathbb{R}_+)^2 \rightarrow \mathbb{R}$  :

$$\frac{dw}{dt} = \left( \frac{dw}{dx} \right)^{-1} \quad \text{avec} \quad w(0, t) = w(x, 0) = 0,$$

dont l'unique solution est  $w : (x, t) \mapsto 2\sqrt{tx}$ . D'où  $\mathbb{E}[L(x, t)] \approx 2\sqrt{tx}$  et  $c = 2$ .

Ils détaillent ensuite la preuve de ce résultat de manière rigoureuse, basée sur la définition d'un modèle de Hammersley sur toute la droite  $\mathbb{R}$  au lieu de la demi-droite  $\mathbb{R}_+$ . Dans le modèle actuel le nombre de particules ne peut qu'augmenter, on observe une accumulation autour de l'origine. L'intérêt de prolonger la demi-droite en une droite est de permettre aux particules situées les plus à gauche d'un intervalle de disparaître pour pouvoir rendre le processus **stationnaire**. Ils montrent que les mesures stationnaires invariantes par translation sont les

(mélanges de) PPP, et en déduisent que  $c \leq 2$  puis  $c \geq 2$ . Les arguments utilisés sont en fait proches de ceux que nous présenterons dans la section 1.3, basés sur la méthode utilisée dans [Gro02, CG05].

Il est également possible d'étudier les fluctuations de la variable  $\frac{L(t,t)}{t}$ , qui de manière surprenante ne sont pas de l'ordre de  $t^{1/2}$  comme souvent mais de l'ordre de  $t^{1/3}$ . Ce résultat a été prouvé initialement par Baik-Deift-Johansson [BDJ99] pour ce modèle, et rentre dans la classe d'universalité dite KPZ (Kardar Parisi et Zhang [KPZ86]). L'étude des modèles de cette classe a été en un certain sens révolutionnaire puisqu'elle donne des exemples de théorèmes de fluctuations qui diffèrent du théorème central limite. Les lois limites obtenues peuvent être souvent reliées à des modèles de matrices aléatoires, et sont dites de Tracy-Widom [TW94]. Ces lois sont devenues aussi importantes que la loi normale dans ce domaine, en raison de leur universalité. Plus de détails sur ces résultats de fluctuations ainsi que des éléments de preuve sont présentés en section 1.4, avec d'une part un point de vue analytique comme dans [BDJ99, BR00] et d'autre part un point de vue davantage probabiliste et géométrique comme dans [CG06].

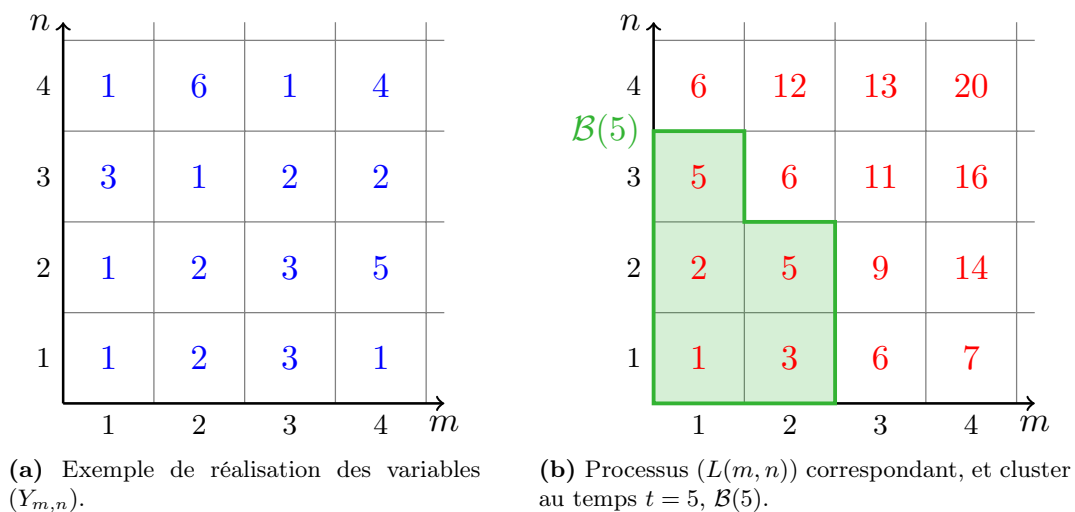
Par ailleurs, il existe de nombreuses généralisations de ce modèle qui ont été étudiées. Citons par exemple celle de Bollobás-Winkler [BW88], où l'on s'intéresse au plus long chemin croissant d'un sous-ensemble uniforme de cardinal  $n$  de l'hypercube  $[0, 1]^d$  : ils montrent que la longueur est équivalente en probabilité à  $c_d n^{1/d}$  quand  $n \rightarrow \infty$ , avec  $c_d \rightarrow e$  quand  $d \rightarrow \infty$ . On peut d'autre part imposer une contrainte géométrique sur le plus long chemin tout en gardant la solubilité du problème : une contrainte de convexité [AB09], de localisation autour de la diagonale [DJP18], d'aire sous la courbe minorée [BGH18], d'espacement minoré entre deux points consécutifs [BG19], de constante de Lipschitz bornée [BG22], ou encore d'entropie ou de distance de Hölder majorée [BT21]. Citons enfin la généralisation introduite et étudiée dans [BHMZ11, IB15, BGG18] sur les arbres de Hammersley en lieu et place des lignes, grâce au point de vue de *patience sorting* décrit au début de la section 1.1.

Dans la section suivante nous présentons d'autres modèles similaires mais cette fois en espace discret, c'est-à-dire que le processus  $L$  ne sera plus défini sur  $(\mathbb{R}_+)^2$  mais sur  $\mathbb{N}^2$  ou  $(\mathbb{N}^*)^2$  par exemple.

## 1.2.2 Espace discret

De manière générale, une percolation de dernier passage dirigée peut être vue comme un modèle de croissance aléatoire dans le quart de plan. Nous présentons ce modèle de manière similaire aux notes de cours de Seppäläinen [Sep09] et de Gerin [Ger13].

On considère dans cette section que l'espace est discret, disons  $(\mathbb{N}^*)^2$  que l'on voit comme une grille, et on cherche à définir un processus  $(L(m, n))_{(m,n) \in (\mathbb{N}^*)^2}$ . À chaque case  $(m, n) \in (\mathbb{N}^*)^2$  on associe une variable aléatoire positive  $Y_{m,n}$  qui peut être vue comme un poids ou un temps d'attente. Dans le second cas, cette notion de temps ne doit pas être confondue avec le temps d'évolution d'un processus comme présenté dans la section précédente, qui est à nouveau interprété ici comme une des coordonnées spatiales. Les variables  $(Y_{m,n})_{(m,n) \in (\mathbb{N}^*)^2}$  sont supposées i.i.d, en pratique souvent de support  $\mathbb{N}$ ,  $\mathbb{N}^*$  ou  $\mathbb{R}_+$ . Un exemple est présenté en Figure 1.8a dans le cas où elles sont à valeurs dans  $\mathbb{N}^*$ .



**FIGURE 1.8** : Exemple de percolation de dernier passage en espace discret

On note  $\Pi(m,n)$  l'ensemble des chemins nord-est de la case  $(1,1)$  à la case  $(m,n)$ , c'est-à-dire l'ensemble des chemins  $\pi$  de la forme

$$\pi = \{(1,1) = (i_1, j_1), (i_2, j_2), \dots, (i_{m+n-1}, j_{m+n-1}) = (m,n)\}$$

tels que :  $\forall s \in \llbracket 1, m+n-2 \rrbracket, (i_{s+1}, j_{s+1}) - (i_s, j_s) \in \{(1,0), (0,1)\}$ . On définit alors

$$L(m,n) = \max_{\pi \in \Pi(m,n)} \sum_{(i,j) \in \pi} Y_{i,j}.$$

$L(m,n)$  correspond donc au poids du plus long chemin nord-est entre les cases  $(1,1)$  et  $(m,n)$ , où le poids d'un chemin est défini comme la somme des poids des cases traversées. Un chemin de poids maximal est appelé *géodésique*. De manière équivalente, on peut également définir le processus  $L$  de manière récurrente. On pose  $L(m,n) = 0$  si  $m = 0$  ou  $n = 0$ , et

$$L(m,n) = L(m-1,n) \vee L(m,n-1) + Y_{m,n}. \quad (1.4)$$

Cette formule peut-être obtenue en partitionnant l'ensemble des chemins de  $\Pi(m,n)$  selon leur avant-dernière case : soit  $(m-1,n)$ , soit  $(m,n-1)$ . La dernière case est nécessairement  $(m,n)$  et son poids  $Y_{m,n}$  apparaît nécessairement dans la somme. Le processus  $(L(m,n))_{(m,n) \in (\mathbb{N}^*)^2}$  correspondant à l'exemple de la Figure 1.8a est présenté en Figure 1.8b.

Une interprétation physique de la percolation de dernier passage est proposée dans [Cas19]. La quantité  $L(m,n)$  est vue comme le temps nécessaire à une parcelle de terre  $(m,n)$  pour être sèche s'il s'arrête de pleuvoir au temps 0. Pour pouvoir sécher, une parcelle de terre  $(m,n)$  doit avoir ses deux parcelles voisines  $(m-1,n)$  et  $(m,n-1)$  sèches également, et ensuite attendre un temps aléatoire  $Y_{m,n}$  avant de devenir elle-même sèche.

Par ailleurs, à tout temps  $t \geq 0$  on définit le **cluster**  $\mathcal{B}(t)$  comme étant l'ensemble des cases atteintes en un temps au plus  $t$ , c'est-à-dire toutes les parcelles sèches au temps  $t$  avec l'interprétation précédente :

$$\mathcal{B}(t) = \{(m,n) \in (\mathbb{N}^*)^2 \mid L(m,n) \leq t\}.$$

Remarquons que le processus  $(L(m, n))_{(m, n) \in (\mathbb{N}^*)^2}$  peut-être retrouvé grâce à la donnée du processus  $(\mathcal{B}(t))_{t \in \mathbb{N}}$  grâce à la relation

$$L(m, n) = \min\{t \geq 0 \mid (m, n) \in \mathcal{B}(t)\}. \quad (1.5)$$

Cette notion de cluster est utile pour établir des liens avec d'autres modèles présentés ci-après. Le cluster au temps  $t = 5$ ,  $\mathcal{B}(5)$ , est représenté en vert dans l'exemple de la Figure 1.8b.

### Corner-Growth Model et TASEP

Nous allons en particulier nous intéresser au cas où les variables  $(Y_{m, n})$  suivent une loi géométrique ou une loi exponentielle. Un modèle équivalent est alors celui dit du *Corner-Growth Model* (CGM).

On s'intéresse à l'évolution d'un cluster  $(\mathcal{B}(t))_{t \in \mathbb{N}}$  dans  $(\mathbb{N}^*)^2$ , croissant pour l'inclusion, et tel que  $\mathcal{B}(0) = \emptyset$ . Par commodité on peut voir ce cluster dans  $\mathbb{N}^2$  en supposant que les cases  $(i, j)$  tels que  $i = 0$  ou  $j = 0$  sont toujours remplies. À chaque étape, chaque case vide dont les voisins de gauche et du bas sont pleins, ce qu'on appelle un *coin*, se remplit avec probabilité  $p \in ]0, 1]$ . Le cluster  $\mathcal{B}(t)$  est alors l'ensemble des cases pleines de  $(\mathbb{N}^*)^2$  au temps  $t$ . Entre deux étapes, les seules cases qui peuvent être ajoutées au cluster sont donc celles qui se trouvent dans un coin extérieur de celui-ci. Un exemple d'évolution est représenté en Figure 1.9.

Si on note  $Y_{m, n}$  le nombre de tentatives pour remplir la case  $(m, n)$  à partir du moment où ses voisins de gauche et du bas sont remplies, alors ces variables aléatoires sont i.i.d. de loi  $\text{Geom}_{\mathbb{N}^*}(p)$ . L'équation (1.4) est alors vérifiée par définition, et on retrouve bien un modèle de percolation de dernier passage comme défini ci-dessus. De plus, le processus  $(\mathcal{B}(t))_{t \in \mathbb{N}}$  est alors une chaîne de Markov sur l'ensemble des clusters finis possibles de  $(\mathbb{N}^*)^2$ .

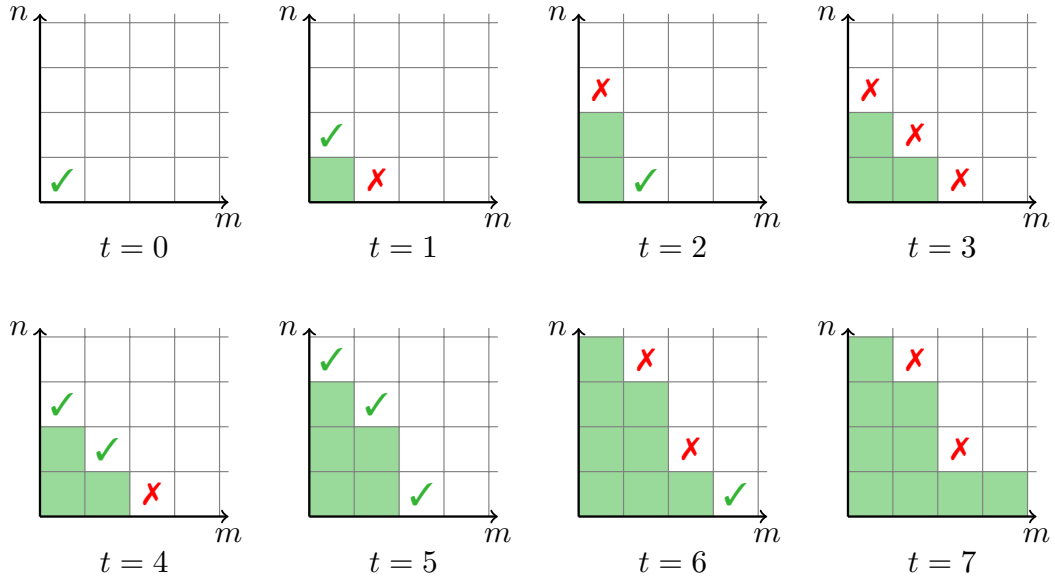
Notons que l'on peut également définir une version à temps continu de ce modèle, en supposant que les cases dont les deux voisines sont pleines se remplissent à taux 1. Cela correspond à considérer que les variables  $(Y_{m, n})$  sont i.i.d. de loi  $\text{Exp}(1)$ , le processus  $(\mathcal{B}(t))_{t \in \mathbb{R}_+}$  est également markovien.

Introduisons désormais le modèle de *Totally Asymmetric Simple Exclusion Process* (TASEP) et exhibons son lien avec le CGM. Ce système de particules a été introduit et étudié notamment par Spitzer et Liggett [Spi70, Lig75], mais c'est Rost [Ros81] qui a établi ce lien, dans le cas continu. Ce modèle avait auparavant été introduit en biologie [MGP68, MG69] indépendamment.

Présentons ici le modèle en temps discret à nouveau, comme effectué dans [Sep09, Ger13] par exemple. Le TASEP est une chaîne de Markov décrivant le déplacement de particules sur la droite  $\mathbb{Z}$ . Chaque site de  $\mathbb{Z}$  ne peut être occupé que par au plus une particule à chaque instant  $t \in \mathbb{N}$ .

À l'instant initial, tous les sites de  $\mathbb{Z}_-$  sont occupés par une particule et tous les sites de  $\mathbb{N}$  sont libres. Numérotions les particules de droite à gauche avec les entiers strictement positifs, et notons  $X_j(t)$  la position de la particule  $j \in \mathbb{N}^*$  à l'instant  $t$ , de sorte que  $X_j(0) = -j$ . La règle d'évolution est alors la suivante, pour  $t \in \mathbb{N}^*$  :

- Si le site  $X_j(t - 1) + 1$  est occupé au temps  $t - 1$ , alors  $X_j(t - 1) = X_j(t)$ .



**FIGURE 1.9 :** Exemple des premières étapes du CGM. Les cases vertes représentent les éléments du cluster  $\mathcal{B}(t)$  à l'instant  $t$ . Les cases susceptibles d'être remplies à l'étape suivante, c'est-à-dire les coins, contiennent un symbole :  $\checkmark$  en cas de remplissage effectif (arrive avec probabilité  $p$ ),  $\times$  sinon.

- Si le site  $X_j(t-1) + 1$  est libre au temps  $t-1$ , alors

$$X_i(t-1) = \begin{cases} X_i(t-1) + 1 & \text{avec probabilité } p, \\ X_i(t-1) & \text{avec probabilité } 1-p. \end{cases}$$

Les déplacements sont tirés indépendamment pour chaque site concerné.

De manière informelle, cela revient à considérer qu'entre deux instants consécutifs, chaque particule peut sauter sur le site situé immédiatement à sa droite seulement si celui-ci est libre. Si tel est le cas, elle saute effectivement avec probabilité  $p$  et libère ainsi le site qu'elle occupait précédemment. Dans les autres cas, elle reste sur place. Un exemple d'évolution est représenté en Figure 1.10.

Le lien entre TASEP et CGM peut être fait en définissant le sous-ensemble de  $(\mathbb{N}^*)^2$  suivant, vu à nouveau comme un ensemble de cases, à chaque instant  $t \in \mathbb{N}$  :

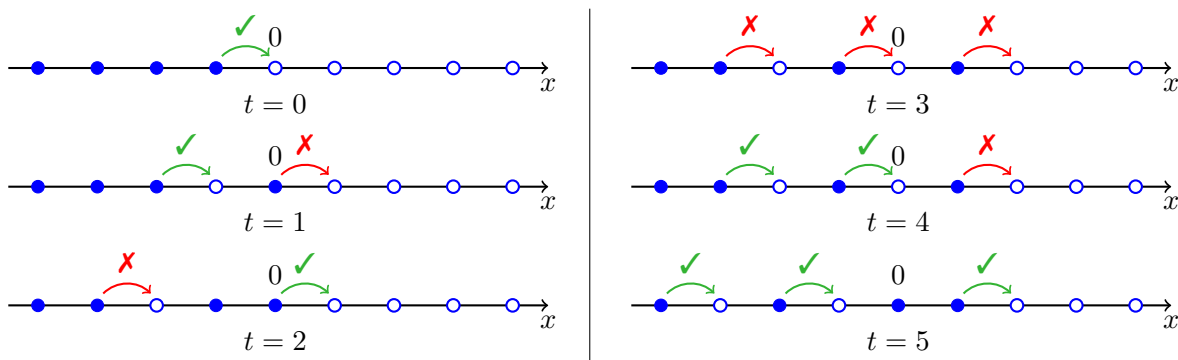
$$\mathcal{A}(t) = \{(i, j) \in (\mathbb{N}^*)^2 \mid i \leq X_j(t) + j\}.$$

En d'autres termes, la longueur de la ligne d'ordonnée  $j$  dans l'ensemble  $\mathcal{A}(t)$  est  $X_j(t) - X_j(0)$ , c'est-à-dire le nombre de sauts effectués par la particule  $j$  jusqu'au temps  $t$ . Lorsque la particule  $j$  saute, cela revient donc à remplir une case supplémentaire sur la ligne d'ordonnée  $j$ . On montre alors l'égalité en loi

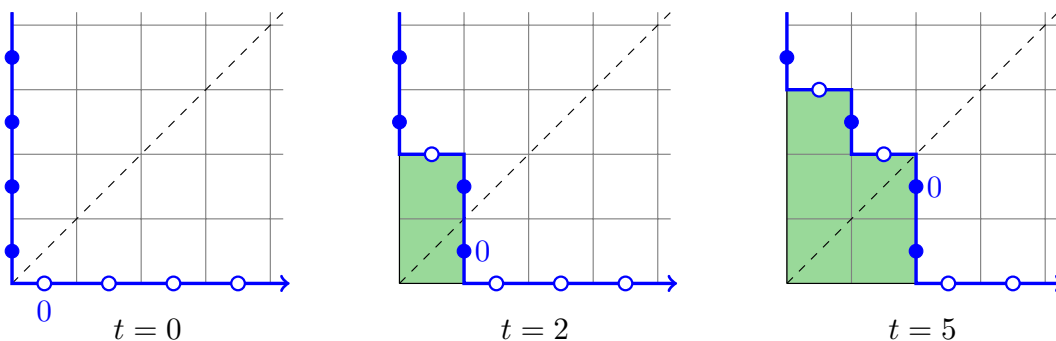
$$(\mathcal{A}(t))_{t \in \mathbb{N}} \stackrel{(d)}{=} (\mathcal{B}(t))_{t \in \mathbb{N}}, \quad (1.6)$$

où  $\mathcal{B}$  est le cluster du CGM.

De manière plus visuelle, le lien entre les deux modèles peut être fait de la façon suivante. Considérons la frontière d'un cluster  $\mathcal{B}(t)$  à un instant  $t$  dans le quart de plan  $\mathbb{N}^2$ , que l'on



**FIGURE 1.10 :** Exemple des premières étapes du TASEP. Les sites occupés sont représentés par un  $\bullet$ , les sites vacants par un  $\circ$ . Les flèches indiquent les sauts possibles pour l'étape suivante, et sont marquées d'un  $\checkmark$  en cas de saut effectif (arrive avec probabilité  $p$ ) ou d'un  $\times$  sinon.

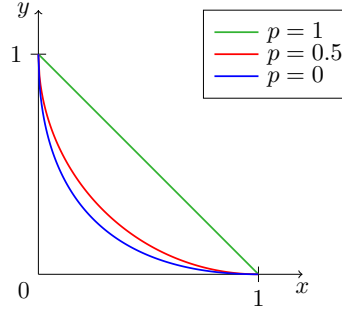


**FIGURE 1.11 :** Correspondance entre CGM et TASEP en reprenant les exemples des Figures 1.9 et 1.10. Le cluster  $\mathcal{B}(t)$  du CGM est représenté en vert. L'axe du TASEP est représenté en bleu, les sites occupés sont représentés par un  $\bullet$  et les sites vacants par un  $\circ$ .

voit comme une succession de segments unitaires verticaux et horizontaux. Plaçons l'origine de cette frontière à son intersection avec la première bissectrice, et numérotons les segments dans  $\mathbb{Z}$  à partir de celle-ci. Plus précisément, le segment 0 est celui qui a pour extrémité gauche ou haute l'intersection avec la première bissectrice, celui immédiatement à sa droite ou en dessous est le numéro 1, alors que celui immédiatement à sa gauche ou au-dessus est le numéro  $-1$ , et ainsi de suite. Cette suite de segments unitaires est assimilée à  $\mathbb{Z}$ , et l'on encode un segment vertical (resp. horizontal) par un site occupé (resp. libre). Cette correspondance est illustrée en Figure 1.11.

Ainsi, les configurations initiales des deux modèles correspondent : la frontière de  $\mathcal{B}(0)$  correspond aux deux axes, c'est-à-dire une succession de segments verticaux strictement avant l'origine et horizontaux après, et est donc associée à la configuration de  $\mathbb{Z}$  où tous les sites de  $\mathbb{Z}_-$  sont occupés par une particule et tous les sites de  $\mathbb{N}$  sont libres. De plus, les carrés qui peuvent être remplis à l'étape suivante dans le CGM sont ceux adjacents à la frontière aux endroits où un segment horizontal suit un segment vertical (de manière à former un coin). Ils correspondent donc exactement aux endroits où un site occupé est suivi par un site libre dans le TASEP. De plus, remplir un tel carré revient à faire sauter la particule correspondante puisque





**FIGURE 1.12** : Courbe d'équation  $x + y + 2\sqrt{(1-p)xy} = 1$  pour  $p \in \{0, 0.5, 1\}$ .

la frontière est localement modifiée, les deux segments changent de type.

Le lien entre TASEP et percolation de dernier passage peut également être énoncé directement en utilisant la relation (1.5) et l'égalité en loi (1.6) :

$$L(m, n) \stackrel{(d)}{=} \min\{t \in \mathbb{N} \mid X_m(t) \geq m - n\}.$$

L'asymptotique de ce modèle a été prouvée par plusieurs approches différentes :

**Théorème 1.5** ([JPS95, Joh00a]). *Pour tout  $p \in (0, 1)$ .*

$$\frac{L(nx, ny)}{n} \xrightarrow[n \rightarrow \infty]{p.s.} \frac{x + y + 2\sqrt{(1-p)xy}}{p}$$

Plus précisément, Jockusch–Propp–Schor [JPS95] obtiennent le cas  $p \leq 1/2$  en introduisant un système stationnaire invariant par translation, puis Johansson [Joh00a] obtient le cas général en utilisant le gaz de Coulomb et des estimées de grandes déviations. Seppäläinen [Sep09] et Rolla–Sidoravicius–Surgailis–Vares [RSSV10] retrouvent eux ce résultat grâce à un argument de stationnarité locale, de manière similaire au cas que nous présenterons en section 1.3.2.

On peut interpréter le Théorème 1.5 comme un résultat de forme limite pour la frontière du cluster en temps long. Si on la renormalise dans le carré  $[0, 1]^2$ , celle-ci régie par la courbe définie implicitement par l'équation

$$x + y + 2\sqrt{(1-p)xy} = 1, \tag{1.7}$$

c'est-à-dire un arc d'ellipse si  $p \in ]0, 1[$ . Le cas  $p = 0.5$  est représenté en Figure 1.12. Notons que pour  $p = 1$  le modèle est déterministe, pour tout couple  $(m, n)$  on a  $L(m, n) = m + n$  et pour tout entier  $t$  le cluster  $\mathcal{B}(t)$  est un triangle. On retrouve bien que la courbe limite est le segment  $[(0, 1), (1, 0)]$ , également représenté en Figure 1.12.

Notons qu'un résultat similaire existe si les  $(Y_{m,n})$  sont i.i.d. de loi  $\text{Geom}_{\mathbb{N}}(p)$ , et peut-être trouvé dans les références ci-dessus :

$$\frac{L(nx, ny)}{n} \xrightarrow[n \rightarrow \infty]{p.s.} \frac{(1-p)(x+y) + 2\sqrt{(1-p)xy}}{p}.$$

Dans le cas où les variables  $(Y_{m,n})$  sont i.i.d. de loi  $\text{Exp}(1)$ , le TASEP peut-être défini de manière analogue, où chaque particule ayant un site vacant à sa droite  $y$  saute à taux 1. Par commodité, on utilise une formalisation différente de celle proposée en discret : on peut définir un processus de Markov  $(Z(t))_{t \in \mathbb{R}_+}$  sur l'espace  $\{0, 1\}^{\mathbb{Z}}$ . Si on note  $Z(t) = (Z_k(t))_{k \in \mathbb{Z}}$ , on a  $Z_k(t) = 1$  si une particule se trouve au site  $k \in \mathbb{Z}$  au temps  $t \in \mathbb{R}_+$ , 0 sinon. Le processus  $(Z(t))_{t \in \mathbb{R}_+}$  est alors markovien, de condition initiale  $Z_k(0) = \mathbf{1}_{k < 0}$  et de générateur infinitésimal

$$Ag(z) = \sum_{k \in \mathbb{Z}} (g(\tau_{k,k+1}z) - g(z)) \mathbf{1}_{z_k=1, z_{k+1}=0},$$

où  $g$  est une fonction bornée de  $\{0, 1\}^{\mathbb{Z}}$  dans  $\mathbb{R}$ ,  $z$  est un vecteur de  $\{0, 1\}^{\mathbb{Z}}$  et  $\tau_{k,\ell}z$  est le vecteur obtenu en transposant les coordonnées  $k$  et  $\ell$  du vecteur  $z$ .

Rost [Ros81] montre alors qu'asymptotiquement, au temps  $t$  la densité de particules décroît linéairement entre  $-t$  et  $t$ , de 1 à 0. Par intégration, il montre la convergence suivante

$$\frac{L(nx, ny)}{n} \xrightarrow[n \rightarrow \infty]{p.s.} (\sqrt{x} + \sqrt{y})^2.$$

et en déduit que le profil limite de la frontière du cluster renormalisé est la courbe définie implicitement par l'équation

$$\sqrt{x} + \sqrt{y} = 1,$$

c'est-à-dire un arc de parabole. Notons que l'équation obtenue coïncide avec l'équation (1.7) pour  $p = 0$ , et est également représentée en Figure 1.12. En effet, ce modèle exponentiel peut-être retrouvé comme étant le cas limite du modèle géométrique défini ci-avant pour le paramètre  $p$  tendant vers 0. Qualitativement, plus le paramètre  $p$  est grand, plus la probabilité que plusieurs cases se remplissent simultanément est importante, et plus l'interface limite est proche d'un segment. A contrario, plus le paramètre  $p$  est petit, plus cette probabilité est faible et plus l'interface limite est proche de l'arc de parabole défini ci-dessus.

Le TASEP est en fait un cas particulier du *Simple Exclusion Process* introduit par Spitzze [Spi70], où l'on autorise également les particules à se déplacer vers la gauche si le site visé est libre, et ce avec probabilité  $q \geq 0$ . Si  $q = p$  on parle de *Symmetric Exclusion Process* (SEP), sinon d'*Asymmetric Exclusion Process* (ASEP), le cas du TASEP correspondant à  $q = 0$ . D'autres généralisations du TASEP ont également été étudiées, comme le TASEP à deux vitesses où deux sous-ensembles de particules ont des probabilités de saut différentes [BFS09], ou le TASEP en espace dit *inhomogène* où la probabilité de saut dépend de la position actuelle de la particule [CR97, Bah98, GKS10, KPS19].

Nous avons montré le lien entre percolation de dernier passage dirigée, CGM et TASEP. Évoquons brièvement une quatrième classe de modèles équivalents en un sens aux précédents : celle des modèles de croissance à nucléation multiple (*PolyNuclear Growth* model, abrégé en PNG), ici dans le cas dit *en gouttelettes*. Il a été introduit et étudié principalement par Prähofer et Spohn [PS00, PS02] en espace continu et par Johansson [Joh00a] en espace discret. Cela revient à considérer que les variables  $(Y_{m,n})$  sont i.i.d. de loi exponentielle ou géométrique respectivement.

Une autre interprétation possible de la percolation de dernier passage dirigée en espace discret s'inscrit dans le cadre d'un modèle de files d'attente, c'est ce que nous présentons dans la section suivante.

### Files d'attentes et matrices aléatoires

Nous détaillons ici le cadre du modèle de files d'attente présenté dans [GW91] et [Sep09], puis le lien qui apparaît avec les matrices aléatoires.

On considère  $n$  clients numérotés de 1 à  $n$  et  $m$  serveurs numérotés de 1 à  $m$ . Chaque client doit visiter chaque serveur de 1 à  $m$  dans l'ordre, puis quitte le système. Les serveurs servent les clients dans leur ordre d'arrivée, c'est ce qu'on appelle une discipline FIFO (*First In, First Out*). On suppose que le temps de service du client  $j$  au serveur  $i$  est égal à la variable aléatoire  $Y_{i,j}$ . Au départ, tous les clients se trouvent dans la file du premier serveur, rangés dans l'ordre croissant. À l'instant  $t = 0$ , le client 1 commence à être servi par le serveur 1. À l'instant  $t = Y_{1,1}$ , le client 1 quitte le serveur 1 pour le serveur 2, et en parallèle le client 2 commence à être servi par le serveur 1, et ainsi de suite.

On remarque alors que le temps auquel le client  $m$  quitte le serveur  $n$  est exactement la variable aléatoire  $L(m, n)$ . En effet, le client  $m$  commence à être servi lorsque d'une part il a quitté le serveur  $n - 1$ , et d'autre part le client  $m - 1$  a quitté le serveur  $n$ . Il doit alors attendre un temps  $Y_{m,n}$  avant de pouvoir quitter le dernier serveur : on retrouve la formule récurrente (1.4) qui caractérise le processus  $(L(m, n))$ .

Glynn et Whitt [GW91] montrent un théorème central limite fonctionnel pour le processus  $L$  dans un cadre très général. Nous présentons ici le cas où les  $(Y_{m,n})$  sont i.i.d. de loi  $\text{Exp}(1)$ . Dans ce cas, le processus limite  $(L(k, \lfloor nt \rfloor))_{t \geq 0}$  peut, après renormalisation, s'écrire comme une fonctionnelle en termes de mouvements browniens. Plus précisément, si on pose pour tout couple d'entiers  $(k, n) \in (\mathbb{N}^*)^2$  et tout réel  $t \in \mathbb{R}_+$ ,

$$L_{k,n}(t) = \frac{L(k, \lfloor nt \rfloor) - nt}{\sqrt{n}} \quad \text{et} \quad L_n = (L_{1,n}, L_{2,n}, \dots),$$

alors le processus  $(L_n)$  converge en loi vers un processus  $\widehat{L}$  quand  $n \rightarrow \infty$ . Cette convergence a lieu dans l'espace des suites de fonction réelles càd-làg définies sur  $\mathbb{R}_+$ , muni de la topologie de Skorohod produit. Le processus limite  $\widehat{L} = (\widehat{L}_1, \widehat{L}_2, \dots)$  peut être exprimé comme suit : pour  $k \in \mathbb{N}^*$ ,

$$\widehat{L}_k(t) = \sup_{0=t_0 \leq t_1 \leq \dots \leq t_k=t} (B_i(t_i) - B_i(t_{i-1})), \quad (1.8)$$

où  $(B_i)_{i \in \mathbb{N}^*}$  est une famille de mouvement browniens standards indépendants. Remarquons que la relation d'échelle des mouvements browniens induit celle du processus  $\widehat{L}$  :

$$\widehat{L}(t) \stackrel{(d)}{=} \sqrt{t} \widehat{L}(1).$$

Baryshnikov [Bar01] et Gravner-Tracy-Widom [GTW01] ont montré que la loi de la variable  $\widehat{L}_k(1)$  était la même que celle de la plus grande valeur propre dans un modèle de matrices aléatoires, que nous allons introduire maintenant.

La loi *Gaussian Unitary Ensemble* est la loi sur l'ensemble  $\mathcal{H}_n$  des matrices Hermitiennes de taille  $n \times n$ , notée  $\text{GUE}(n)$ , dont la densité par rapport à la mesure de Lebesgue est proportionnelle à la fonction

$$H \mapsto \exp\left(-\frac{1}{2} \text{Tr}(H^2)\right).$$

De manière équivalente, on peut construire une matrice aléatoire  $H = (H_{i,j})_{1 \leq i,j \leq n}$  de loi GUE sur  $\mathcal{H}_n$  de la manière suivante. On considère  $(\xi_{i,j})_{1 \leq i \leq j \leq n}$  une famille de variables aléatoires indépendantes telles que les variables sur la diagonale  $(\xi_{i,i})_{1 \leq i \leq n}$  (resp. au-dessus de la diagonale  $(\xi_{i,j})_{1 \leq i < j \leq n}$ ) soient de loi  $\mathcal{N}_{\mathbb{R}}(0, 1)$  (resp. de loi  $\mathcal{N}_{\mathbb{C}}(0, 1)$ ). Ainsi, si on définit les coefficients d'une matrice  $H$  par

$$H_{i,j} = \begin{cases} \xi_{i,j} & \text{si } i \leq j, \\ \overline{\xi_{j,i}} & \text{si } i > j, \end{cases}$$

alors la matrice  $H$  est de loi GUE sur  $\mathcal{H}_n$ .

Ainsi, le lien avec la loi de la variable  $\widehat{L}_n(1)$  définie par l'équation (1.8) est le suivant :

**Théorème 1.6** ([Bar01, GTW01]). *Si  $\lambda_n$  est la plus grande valeur propre d'une matrice de loi GUE sur  $\mathcal{H}_n$ , alors*

$$\lambda_n \stackrel{(d)}{=} \sup_{0=t_0 \leq t_1 \leq \dots \leq t_k=1} (B_i(t_i) - B_i(t_{i-1})).$$

La preuve du Théorème 1.6 repose essentiellement sur une analyse combinatoire des tableaux de Young. Une autre interprétation est donnée dans [Dou05]. Elle repose sur la loi *Laguerre Unitary Ensemble* (LUE) de paramètres  $(n, m)$  avec  $m \geq n$  sur l'ensemble  $\mathcal{H}_n$  que l'on définit comme étant la loi dont la densité est proportionnelle à la fonction

$$H \mapsto (\det H)^{m-n} \exp(-\text{Tr}(H)) \mathbf{1}_{H \geq 0}.$$

Johansson [Joh00a] a établi que la variable aléatoire  $L(m, n)$  dans le cas où les  $(Y_{m,n})$  sont i.i.d. de loi  $\text{Exp}(1)$  avait même loi que la plus grande valeur propre d'une matrice de loi LUE $(n, m)$ . Par ailleurs, si  $(H^{(n,m)})_{m \geq n}$  est une suite de matrices de loi LUE $(n, m)$ , alors on a la convergence en loi suivante (prouvée par Jonsson [Jon82]) :

$$\frac{H^{(n,m)} - mI_n}{\sqrt{m}} \xrightarrow[m \rightarrow \infty]{(d)} \text{GUE}(n).$$

Ainsi, le Théorème 1.6 peut être établi en revenant à l'expression (1.8) et à la définition du processus  $\widehat{L}$  comme processus limite.

Par ailleurs, O'Connell et Yor ont obtenu un résultat plus général que ce Théorème 1.6 dans [OY02], en établissant un lien entre des fonctionnelles similaires appliqués à des mouvements browniens indépendants et l'ensemble des valeurs propres d'une matrice GUE $(n)$ . Ils ont également étudié en détail dans [OY01] le lien entre un modèle de file d'attente browniennes et le modèle de percolation de dernier passage dans le cas où les poids sont i.i.d. de loi exponentielle, notamment la réversibilité. Seppäläinen [Sep12] obtient un résultat similaire dans le cas où les poids sont i.i.d. de loi log-gamma.

D'autre part, mentionnons qu'un lien direct peut être établi entre le problème d'Ulam et les matrices aléatoires, comme détaillé dans [AD99, Section 4] et initialement découvert par Rains [Rai98].

## D'autres modèles en espace discret

Nous présentons ici deux modèles étudiés dans [BEGG16] et initialement étudiés par Seppäläinen dans deux papiers différents [Sep97, Sep98a], que nous nommerons respectivement modèles 1 et 2. Dans les deux cas les variables aléatoires  $(Y_{m,n})$  sont i.i.d. de loi Ber $(p)$ , mais le

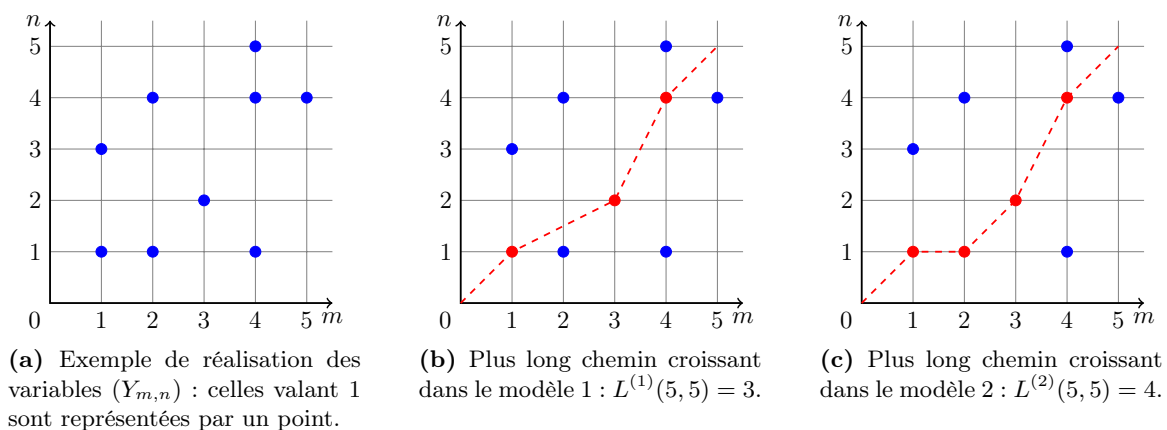


FIGURE 1.13 : Les deux modèles de PDP présentés en section 1.2.2.

changement provient de l'ensemble des chemins autorisés entre les points  $(1, 1)$  et  $(m, n)$ , que l'on notera respectivement  $\Pi^{(1)}(m, n)$  et  $\Pi^{(2)}(m, n)$  pour ces modèles. Plus précisément, pour  $k \in \{1, 2\}$  on note  $\Pi^{(k)}(m, n)$  l'ensemble des chemins  $\pi$  de la case  $(1, 1)$  à la case  $(m, n)$  de la forme

$$\pi = \{(i_1, j_1), (i_2, j_2), \dots, (i_\ell, j_\ell)\}, \ell \in \mathbb{N}^*,$$

tels que :

- si  $k = 1$ ,  $1 \leq i_1 < \dots < i_\ell \leq m$  et  $1 \leq j_1 < \dots < j_\ell \leq n$ ,
- si  $k = 2$ ,  $1 \leq i_1 < \dots < i_\ell \leq m$  et  $1 \leq j_1 \leq \dots \leq j_\ell \leq n$ .

Ainsi, les éléments de ces ensembles sont les chemins nord-est du rectangle  $\llbracket 1, m \rrbracket \times \llbracket 1, n \rrbracket$ , où les deux directions sont considérées au sens strict pour  $\Pi^{(1)}(m, n)$ , et *nord* au sens strict mais *est* au sens large pour  $\Pi^{(2)}(m, n)$ .

On définit alors les processus  $L^{(k)}$  par

$$L^{(k)}(m, n) = \max_{\pi \in \Pi^{(k)}(m, n)} \sum_{(i,j) \in \pi} Y_{i,j}.$$

Un exemple est présenté en Figure 1.13 et permet de visualiser la différence entre les deux modèles. Comme les variables  $(Y_{m,n})$  sont à valeurs dans l'ensemble  $\{0, 1\}$ , nous avons choisi d'adopter une représentation similaire à celle présentée en section 1.1.1. On voit alors  $(\mathbb{N}^*)^2$  comme les sommets d'un réseau, et les variables  $(Y_{m,n})$  valant 1 comme des points de ce réseau. Notons que l'on peut adapter la méthode des lignes de Hammersley décrite en section 1.1.1 pour déterminer les longueurs  $L^{(1)}(5, 5)$  et  $L^{(2)}(5, 5)$  dans cet exemple. Le cas du premier modèle sera détaillé en section 1.3.2

Pour ces deux modèles, on a les asymptotiques suivantes :

**Théorème 1.7** ([Sep97, Sep98a, BEGG16]). *Pour  $a, b > 0$ ,*

$$\frac{L^{(1)}(an, bn)}{n} \xrightarrow[n \rightarrow \infty]{p.s.} \begin{cases} \frac{\sqrt{p}(2\sqrt{ab} - (a+b)\sqrt{p})}{1-p} & \text{si } p < \min\left(\frac{a}{b}, \frac{b}{a}\right), \\ \min(a, b) & \text{sinon.} \end{cases}$$

$$\frac{L^{(2)}(an, bn)}{n} \xrightarrow[n \rightarrow \infty]{p.s.} \begin{cases} 2\sqrt{abp(1-p)} + (a-b)p & \text{si } p < \frac{a}{a+b}, \\ a & \text{sinon.} \end{cases}$$

Comme évoqué ci-avant, le premier modèle sera détaillé dans la section 1.3.2. Il est en effet propice pour bien comprendre le procédé de stationnarisation et son utilisation pour obtenir une asymptotique au premier ordre.

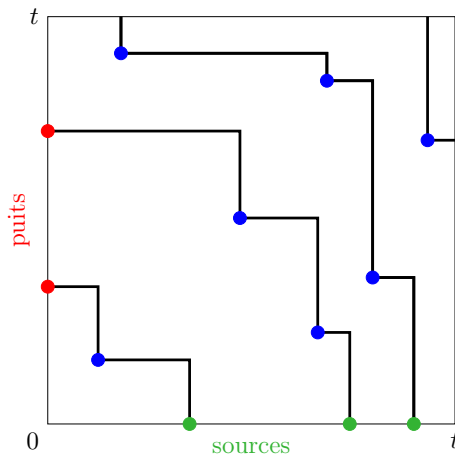
Le cas général de percolation de dernier passage présenté en début de section 1.2.2 a également fait l'objet d'études, notamment par Martin [Mar06] qui donne des conditions sur la loi des variables  $(Y_{i,j})$  pour avoir une loi des grands nombres ou une forme limite. Ont également été étudiées les géodésiques (c'est-à-dire les chemins de poids maximal) et l'interface de compétition (qui sépare les points pour qui la géodésique pour les atteindre depuis l'origine commence par la case  $(0, 1)$  ou par la case  $(1, 0)$ ), notamment dans [Wüt02, FP05, CP11, GRAS17], ou plus récemment la structure géométrique des temps de passage pour différentes lois [AC21]. Par ailleurs, mentionnons également un modèle de percolation de dernier passage généralisé introduit et étudié par Casse [Cas19], où les variables  $(Y_{i,j})$  ne sont pas i.i.d. : la loi dépend alors de la différence entre les temps d'arrivée de la cellule du bas et de la cellule de gauche.

### 1.3 Des modèles stationnaires comme outils

Dans cette section, nous allons détailler dans le cadre de deux exemples différents, l'un en espace continu et l'autre en espace discret, comment la stationnarisation d'un processus permet d'obtenir l'asymptotique de premier ordre du problème de percolation de dernier passage. Cette idée a été utilisée pour la première fois par Aldous et Diaconis [AD95] puis Seppäläinen [Sep96] comme évoqué en section 1.2.1 en prolongeant le processus de Hammersley initialement sur  $\mathbb{R}_+$  sur tout  $\mathbb{R}$ . Ici, nous allons d'une part adopter un point de vue géométrique en voyant le temps comme une dimension au même titre que l'espace, et d'autre part présenter des modèles rendus stationnaires en restant dans le quart de plan  $(\mathbb{R}_+)^2$  ou  $\mathbb{N}^2$ .

#### 1.3.1 Exemple en espace continu

Le modèle considéré ici est celui de Hammersley poissonnisé présenté en section 1.1.1. Par analogie avec la section précédente, on notera  $L(x, t) = L([0, x] \times [0, t])$  la longueur du plus long chemin croissant du point  $(0, 0)$  au point  $(x, t)$ , où le rectangle  $[0, x] \times [0, t]$  contient un PPP d'intensité 1. On appellera les points de ce processus les *points*  $\alpha$ , pour suivre la terminologie introduite par Groeneboom [Gro01]. L'idée générale est d'ajouter des processus sur les axes et a été introduite par Groeneboom seul [Gro02] et avec Cator [CG05]. On veut voir le modèle défini dans le quart de plan  $(\mathbb{R}_+)^2$  comme la restriction d'un modèle défini sur le plan entier  $\mathbb{R}^2$ . Ces processus ajoutés sur les axes pour le modèle dans le quart de plan simulent en quelque sorte l'influence du reste du plan sur cette portion  $(\mathbb{R}_+)^2$ .



**FIGURE 1.14 :** Exemple de construction de lignes de Hammersley avec points  $\alpha$  (en bleu), sources (en vert) et puits (en rouge).

On s'autorise donc à avoir des points sur l'axe des abscisses  $x = 0$  et des ordonnées  $t = 0$  avec probabilité non nulle. Les points sur l'axe des abscisses sont appelés *sources*, et sont l'analogie de particules à l'instant initial pour le processus de Hammersley défini par Aldous et Diaconis [AD95] et présenté en section 1.2.1. Les points sur l'axe des ordonnées sont appelés *puits* et suppriment systématiquement la particule la plus à gauche au temps  $t$ , s'il y en a une. Ces points se comportent de la même manière que les points  $\alpha$  pour la construction des lignes de Hammersley, à la différence qu'on ne représente pas les accumulations de lignes sur les axes. Si on construit les lignes de bas en haut, de chaque source commence verticalement une ligne de Hammersley. Si une ligne de Hammersley rencontre un puits, elle meurt. Un exemple de construction de lignes de Hammersley avec sources et puits est présenté en Figure 1.14.

De manière probabiliste, on considère que les sources (resp. puits) sont positionnées selon un PPP d'intensité  $\lambda > 0$  sur  $[0, x] \times \{0\}$  (resp.  $\mu > 0$  sur  $\{0\} \times [0, t]$ ). Ces deux PPP ainsi que celui générant les points  $\alpha$  sont supposés indépendants.

Cette construction induit un processus de Markov à temps continu sur l'ensemble  $\mathcal{P}_x$  des parties finies de  $]0, x]$ , pour  $x > 0$ . Sa condition initiale est l'ensemble des atomes d'un PPP d'intensité  $\lambda$  sur  $]0, x]$ , et son générateur infinitésimal s'écrit alors, si  $g$  est une fonction bornée de  $\mathcal{P}_x$  dans  $\mathbb{R}$  et si  $X \in \mathcal{P}_x$  est une configuration :

$$Ag(X) = \int_0^x (g(X^u) - g(X))du + \mu(g(X^0) - g(X)),$$

où  $X^u$  correspond à la configuration obtenue en ajoutant un point  $\alpha$  ou un puits à l'abscisse  $u$ . Plus précisément, si  $X$  s'écrit  $X = \{x_1, \dots, x_n\}$  où  $x_1 < \dots < x_n$ , alors

1. si  $u = 0$ ,  $X^0 = X \setminus \{x_1\}$ . Un puits est apparu, on a retiré la particule située la plus à gauche (s'il n'y en avait pas, par convention  $X^0 = X = \emptyset$ ).
2. si  $0 < u < x_n$ , notons  $x_j$  l'abscisse de la particule la plus proche à droite de  $u$ . On a alors  $X^u = X \setminus \{x_j\} \cup \{u\}$ . Un point  $\alpha$  est apparu et a déplacé une particule.

- 3.** si  $u > x_n$ ,  $X^u = X \cup \{u\}$ . Un point  $\alpha$  est apparu et a créé une nouvelle particule comme il n'y en avait pas à sa droite.

Groeneboom [Gro02] montre alors que si les paramètres  $\lambda$  et  $\mu$  vérifient l'égalité  $\mu = \frac{1}{\lambda}$ , le processus est stationnaire. Plus précisément :

**Théorème 1.8** ([Gro02, CG05]). *Si on fixe  $x > 0, t > 0$ , et qu'on note  $N$  (resp.  $S, E, W$ ) le nombre de lignes de Hammersley traversant le nord (resp. sud, est, ouest) de la boîte  $[0, x] \times [0, t]$  dans la modèle avec sources et puits pour  $\mu = \frac{1}{\lambda}$ , alors*

$$N \stackrel{(d)}{=} S, \quad E \stackrel{(d)}{=} W \quad \text{et} \quad N \perp E.$$

*De plus, le processus est réversible dans le sens où la configuration des lignes obtenues dans la boîte  $[0, x] \times [0, t]$  a même loi que celle obtenue en la retournant de  $180^\circ$ .*

Les sources (resp. puits) du processus retourné sont les points de sortie des lignes de Hammersley par le haut (resp. la droite) de la boîte considérée dans le modèle original, et les points  $\alpha$  du processus retourné sont les points où les lignes de Hammersley du processus original tournent et qui ne sont pas des points  $\alpha$  (ce qu'ils appellent des *points  $\beta$* ).

On peut donner une approche de l'heuristique de la relation  $\mu = \frac{1}{\lambda}$  pour obtenir un régime stationnaire de la manière suivante. Parmi les trois cas de figure évoqués plus haut dans l'écriture du générateur infinitésimal, le cas **2.** préserve le nombre de particules. Il faut donc que les cas **1.** et **3.** se produisent à la même fréquence. Le cas **1.** correspond exactement à l'apparition d'un puits, ce qui arrive à taux  $\mu$ . De plus, si le système est stationnaire, à tout instant les particules sont distribuées selon un PPP d'intensité  $\lambda$ , puisque c'est le cas à l'instant initial par définition. La longueur du segment entre la particule la plus à droite et le bord droit de la boîte suit donc la loi exponentielle de paramètre  $\frac{1}{\lambda}$ . Le cas **3.** correspondant à l'arrivée d'un point  $\alpha$  dans ce segment, il se produit typiquement à taux  $\frac{1}{\lambda}$  puisque les points  $\alpha$  arrivent à taux 1. D'où la relation pressentie  $\mu = \frac{1}{\lambda}$ .

Notons  $L_\lambda(x, t)$  la longueur du plus long chemin croissant entre les points  $(0, 0)$  et  $(x, t)$  obtenu dans ce cas, c'est-à-dire le nombre total de lignes de Hammersley dans le rectangle  $[0, x] \times [0, t]$ . Ainsi, avec les notations du théorème, on a de manière déterministe

$$L_\lambda(x, t) = W + N = S + E, \tag{1.9}$$

puisque chaque ligne de Hammersley "entre" dans la boîte par l'ouest ou le nord et en "sort" par le sud ou l'est.

Prenons maintenant  $x = t$ , on cherche à utiliser le Théorème 1.8 pour montrer que la constante  $c$  définie par le Théorème 1.2 est égale à 2. On peut obtenir l'inégalité  $c \leq 2$  assez facilement, en remarquant que de manière déterministe

$$\forall \lambda > 0, \quad L(t, t) \leq L_\lambda(t, t), \tag{1.10}$$

étant donné que l'ensemble des points considérés dans le modèle sans sources ni puits (que nous appellerons *modèle pur* par la suite) est inclus dans celui considéré avec sources et puits (que nous appellerons *modèle stationnaire*). De plus, toujours avec les notations du Théorème 1.8,

$$\mathbb{E}[L_\lambda(t, t)] = \mathbb{E}[W] + \mathbb{E}[N] = \frac{t}{\lambda} + \lambda t.$$



Comme l'inégalité (1.10) est valable quel que soit le paramètre  $\lambda$ , la meilleure majoration de  $\mathbb{E}[L(t, t)]$  sera obtenue en prenant celui qui minimise la fonction  $\lambda \mapsto \lambda + \frac{1}{\lambda}$ , c'est-à-dire  $\lambda = 1$ . On en déduit alors

$$\mathbb{E}[L(t, t)] \leq \mathbb{E}[L_1(t, t)] = 2t.$$

Comme la convergence du Théorème 1.2 a lieu dans  $L^1$ , cela permet de conclure que  $c \leq 2$ .

Pour obtenir la minoration  $c \geq 2$ , on peut poursuivre la comparaison entre le modèle pur et le modèle stationnaire pour  $\lambda = 1$ , que nous appellerons modèle stationnaire *critique*, en montrant que les plus longs chemins dans ce cas prennent un nombre négligeable de sources ou de puits. Pour cela, Cator et Groeneboom [CG05] introduisent la notion de *particule de seconde classe* dans le cadre de ce processus, comme cela avait déjà été fait par Ferrari [Fer92] et Liggett [Lig99, Chapitre 3] pour le TASEP. Ces particules de seconde classe jouent un rôle central dans l'étude des fluctuations par Cator et Groeneboom [CG06] que nous présenterons dans la section 1.4.2.

Considérons une réalisation du système stationnaire  $L_\lambda$ , puis une deuxième réalisation avec les mêmes sources, puits et points  $\alpha$  où l'on a ajouté une source au point  $(0, 0)$  (ce qui revient à supprimer le premier puits). La différence symétrique entre les deux collections de lignes de Hammersley obtenue est une ligne brisée ne faisant que des déplacements vers le nord et vers l'est, et correspond à la trajectoire d'une particule dite *de seconde classe*. En pratique, cette particule commence à l'abscisse  $x = 0$  jusqu'au premier puits rencontré où elle saute à la position de la particule (de première classe) qui a été absorbée par ce puits. Elle continue ainsi en sautant successivement aux positions précédentes des particules situées immédiatement à sa droite aux moments où ces particules passent à gauche de la particule de seconde classe.

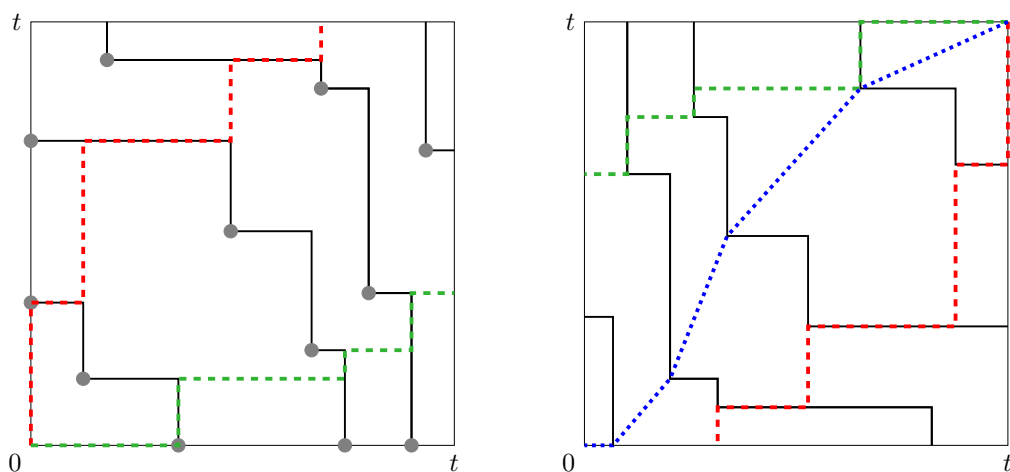
De manière analogue, on définit la trajectoire d'une particule de seconde classe dite *duale* comme étant celle obtenue en procédant de même que ci-dessus mais en ajoutant un puits à l'origine au lieu d'une source. De manière équivalente, cela revient à considérer une particule de seconde classe comme obtenue ci-dessus pour le processus symétrisé, c'est à dire en inversant le rôle des deux coordonnées. Un exemple de trajectoire d'une particule de seconde classe et d'une particule de seconde classe duale est représenté en Figure 1.15a.

L'observation clé est la suivante (utilisée dans la preuve de [CG05, Corollaire 3.1] et davantage détaillée dans [CG06, Section 3]), présentée ici de manière informelle.

**Lemme 1.2** ([CG05, CG06]). *Un plus long chemin croissant pour un système donné est compris entre la trajectoire d'une particule de seconde classe et d'une particule de seconde classe duale pour le système retourné.*

Ce résultat est illustré par la Figure 1.15b. Notons que seules les lignes de Hammersley sont représentées ici, mais déterminent complètement la position des points  $\alpha$ , sources et puits.

*Démonstration.* En section 1.1.1, nous avons présenté une méthode constructive pour obtenir un plus long chemin à l'aide des lignes de Hammersley, et réciproquement tous les plus longs chemins peuvent s'obtenir de cette manière. Une fois les lignes de Hammersley établies, l'algorithme construit un tel chemin en commençant par sa fin, c'est-à-dire en se plaçant en haut à droite de la boîte considérée. Parfois, plusieurs points sont possibles et engendrent des chemins différents. Considérons d'une part la convention qui consiste à choisir le point le plus à gauche (ou de manière équivalente le plus haut) possible à chaque étape, et d'autre part la convention



(a) Trajectoire d'une particule de seconde classe (en rouge) et d'une particule de seconde classe duale (en vert) dans l'exemple de la Figure 1.14.

(b) Un exemple de plus long chemin (en bleu) pour un système donné, et trajectoires de particules de seconde classe (en rouge et vert) pour le système retourné, correspondant au système de la Figure 1.15a.

**FIGURE 1.15** : Trajectoires des particules de seconde classe et plus long chemin croissant.

opposée où l'on choisit systématiquement le point le plus à droite (ou le plus bas). Ainsi, tous les plus longs chemins possibles sont situés entre les deux obtenus en suivant ces deux conventions (il y a unicité si et seulement si les deux conventions fournissent le même chemin). Par construction, les points  $\alpha$  de ces deux chemins sont exactement les points  $\alpha$  (devenus des points  $\beta$ ) présents sur la trajectoire d'une particule de seconde classe et d'une particule de seconde classe duale pour le système retourné.  $\square$

Ainsi, Cator et Groeneboom ramènent l'étude des plus long chemins à l'étude des trajectoires des particules de seconde classe dans le système retourné, qui ont la même loi que celles dans le système d'origine par réversibilité (Théorème 1.8). Le théorème principal de leur article est le suivant :

**Théorème 1.9** ([CG05]). *Dans le modèle stationnaire  $L_\lambda$  pour  $\lambda > 0$ , si on note  $X_\lambda(t)$  l'abscisse de la particule de seconde classe au temps  $t$ , et  $T_\lambda(x)$  l'ordonnée de la particule de seconde classe à l'abscisse  $x$  duale, alors on a les asymptotiques presque sûres suivantes :*

$$X_\lambda(t) \underset{t \rightarrow \infty}{\sim} \frac{t}{\lambda^2} \quad \text{et} \quad T_\lambda(x) \underset{x \rightarrow \infty}{\sim} \lambda^2 x \quad p.s.$$

Au premier ordre les particules de seconde classe évoluent donc linéairement. Leur preuve repose en partie sur le procédé de *thinning* (affinage) et *thickening* (épaississement), qui est un couplage entre un système  $L_\lambda$  et un autre  $L_\delta$  avec un autre paramètre  $\delta > \lambda$  où l'on a gardé les mêmes points  $\alpha$  mais modifié légèrement les intensités des sources et des puits (comme  $\delta > \lambda$ , on a ajouté des sources et retiré des puits). Ils étudient d'abord la trajectoire de la particule de seconde classe correspondant au premier puits présent dans le premier système mais absent dans le second, avant de revenir à celle partant de l'origine dans  $L_\lambda$  en faisant tendre  $\delta$  vers  $\lambda$ .

Ce théorème permet d'obtenir, grâce à la réversibilité et au Lemme 1.2, qu'un plus long chemin entre  $(0,0)$  et  $(t,t)$  dans le modèle stationnaire pour  $\lambda = 1$  passe une proportion de temps négligeable sur les axes pour  $t \rightarrow \infty$ . Cela permet de conclure qu'au premier ordre, le comportement asymptotique de  $L_1(t,t)$  est le même que celui de  $L(t,t)$ , et donc que  $c = 2$ .

### 1.3.2 Exemple en espace discret

Dans cette section, nous allons présenter un autre exemple de stationnarisation d'un modèle, cette fois en espace discret. Pour cela, nous nous plaçons dans le cadre du modèle 1 de [BEGG16] et initialement introduit dans [Sep97], présenté en section 1.2.2.

Notons  $L = L^{(1)}$  dans cette section. Nous allons montrer ici le premier point du Théorème 1.7 dans le cas  $a = b = 1$ , c'est-à-dire la convergence presque sûre

$$\frac{L(n,n)}{n} \xrightarrow[n \rightarrow \infty]{p.s.} \frac{2\sqrt{p}}{1 + \sqrt{p}}. \quad (1.11)$$

Notons que, comme dans la section 1.1.1, le théorème sous-additif de Kingman s'applique et assure que ici aussi, la suite  $\left(\frac{L(n,n)}{n}\right)_n$  tend presque sûrement et dans  $L^1$  vers une constante, fonction de  $p$ . Le but est donc à nouveau d'explicitier cette constante.

Les variables aléatoires  $(Y_{i,j})_{i \geq 1, j \geq 1}$  sont i.i.d. de loi  $\text{Ber}(p)$  où  $p \in ]0,1[$ . Ainsi, en chaque sommet  $(i,j)$  il y a ou non un point, qui joue finalement le même rôle que les points du PPP dans les modèles en espace continu. Ainsi, si on pose

$$\xi = \{(i,j) \in (\mathbb{N}^*)^2 \mid Y_{i,j} = 1\},$$

on peut construire des lignes de Hammersley comme décrit en section 1.1.1 pour une permutation, en remplaçant son graphe  $\mathcal{G}$  par l'ensemble  $\xi$ . Notons que la relation d'ordre partielle définie sur  $(\mathbb{N}^*)^2$  par l'équation (1.2) est bien adaptée au modèle considéré ici, puisque les chemins admissibles sont ceux nord-est où les deux directions sont considérées au sens strict. Ainsi, deux points situés à une même abscisse ou à une même ordonnée peuvent appartenir à une même ligne de Hammersley. Un exemple de construction de lignes de Hammersley dans ce modèle est présenté en Figure 1.16a, reprenant les points de la Figure 1.13a.

Comme dans la section 1.3.1, on souhaite ajouter des points sur les axes pour stationnariser le modèle. Ainsi, on considère deux familles supplémentaires de variables aléatoires :  $(Y_{i,0})_{i \geq 1}$  i.i.d. de loi  $\text{Ber}(\alpha)$  et  $(Y_{0,j})_{j \geq 1}$  i.i.d. de loi  $\text{Ber}(\beta)$ , où toutes les variables  $(Y_{i,j})_{i \geq 0, j \geq 0}$  sont supposées indépendantes, et où  $\alpha, \beta \in ]0,1[$  sont deux paramètres à déterminer en fonction de  $p$ . On pose alors

$$\xi^* = \{(i,j) \in (\mathbb{N})^2 \mid Y_{i,j} = 1\}.$$

Les points de  $\xi^*$  sur l'axe des abscisses sont appelés des *sources*, et ceux sur l'axe des ordonnées des *puits*.

Les chemins considérés comme croissants entre les points  $(0,0)$  et  $(m,n)$  sont ceux qui prennent un certain nombre de sources ou un certain nombre de puits (et ne font donc que des pas nord ou des pas est dans un premier temps) avant de quitter les axes et de ne faire que des

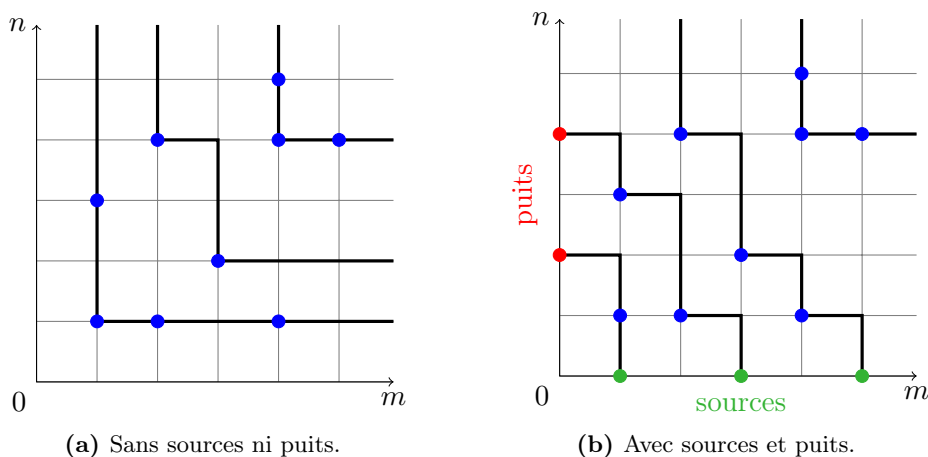


FIGURE 1.16 : Exemple de construction de lignes de Hammersley sans et avec sources et puits.

pas strictement nord-est pour rejoindre  $(m, n)$ . Plus précisément, on note  $\Pi^*(m, n)$  l'ensemble des chemins  $\pi$  du point  $(0, 0)$  au point  $(m, n)$  de la forme

$$\pi = \{(i_1, j_1), (i_2, j_2), \dots, (i_\ell, j_\ell)\}, \ell \in \mathbb{N}^*,$$

tels qu'il existe un entier  $k$  tel que l'un des deux cas de figure suivant est vérifié :

- $1 \leq i_1 < \dots < i_\ell \leq m$  et  $0 = j_1 = \dots = j_k < j_{k+1} < \dots < j_\ell \leq n$ , ou
- $0 = i_1 = \dots = i_k < i_{k+1} < \dots < i_\ell \leq m$  et  $1 \leq j_1 < \dots < j_\ell \leq n$ .

Dans le premier cas le chemin contient des sources, dans le deuxième il contient des puits. De manière équivalente, cela revient à généraliser la définition de la relation d'ordre partielle  $\prec$  définie par l'équation (1.2) sur  $\mathbb{N}^2$  en modifiant légèrement son comportement comme suit :

$$(x, y) \prec (x', y') \iff (x < x' \text{ et } y < y') \text{ ou } (x < x' \text{ et } 0 = y \leq y') \text{ ou } (0 = x \leq x' \text{ et } y < y').$$

Ainsi, la longueur du plus long chemin croissant entre les points  $(0, 0)$  et  $(m, n)$

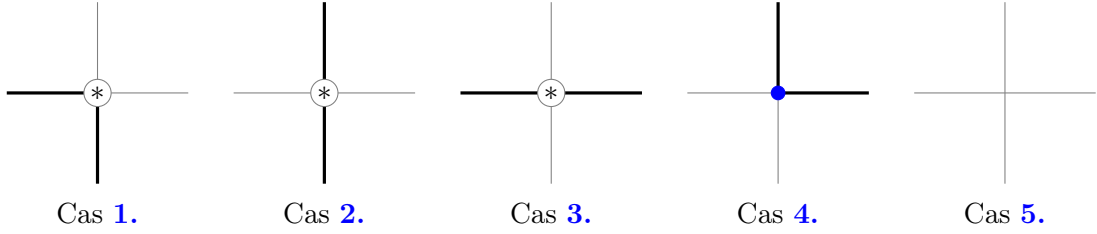
$$L_{\alpha, \beta}(m, n) = \max_{\pi \in \Pi^*(m, n)} \sum_{(i, j) \in \pi} Y_{i, j}$$

correspond au nombre de lignes de Hammersley dans le rectangle  $\llbracket 0, m \rrbracket \times \llbracket 0, n \rrbracket$  construites avec les points de l'ensemble  $\xi^*$  selon la relation  $\prec$ . Un exemple de construction de lignes de Hammersley avec sources et puits est présenté en Figure 1.16b.

Intéressons-nous désormais à la stationnarité. De manière formelle, pour  $x \in \llbracket 1, m \rrbracket$  et  $t \in \mathbb{N}$  on note  $N_t(x)$  la variable aléatoire égale à 1 s'il y a une ligne de Hammersley (dans le modèle avec sources et puits) sur l'arête  $\{(x, t), (x, t+1)\}$  et 0 sinon. Le théorème est alors le suivant :

**Théorème 1.10** ([BEGG16]). *Si  $p, \alpha, \beta \in ]0, 1[$ , le processus  $(N_t(x), x \in \llbracket 1, m \rrbracket)_{t \in \mathbb{N}}$  est stationnaire si les paramètres vérifient la relation*

$$\alpha\beta = (1 - \alpha)(1 - \beta)p. \quad (1.12)$$

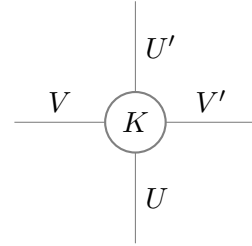


**FIGURE 1.17 :** Les cinq configurations locales possibles autour d'un site. Un astérisque signifie que la présence ou non d'un point de  $\xi$  en ce site ne change pas la configuration des lignes incidentes.

*Démonstration.* Dans ce cas de figure on peut regarder dans un premier temps la stationnarité locale du modèle, comme cela a par exemple été fait par Seppäläinen [Sep09, Lemma 2.3] dans le cas du CGM présenté en section 1.2.2. Considérons un point  $(i, j)$  de  $(\mathbb{N}^*)^2$  et ses quatre arêtes incidentes. On note  $U$  (resp.  $V, U', V'$ ) la variable aléatoire égale à 1 si une ligne de Hammersley est présente sur l'arête incidente du bas (resp. de gauche, du haut, de droite) du point considéré, 0 sinon. Remarquons que dans ce modèle, sur chaque arête de  $\mathbb{N}^2$  il ne peut exister qu'au plus une ligne de Hammersley (ce qui n'est pas le cas dans le Modèle 2 de [BEGG16] présenté en section 1.2.2 par exemple).

Ainsi, on peut exprimer de manière déterministe les variables  $U'$  et  $V'$  en fonction des variables  $U, V$  et  $K = Y_{i,j}$  selon les cas de figure possibles :

1.  $(U, V) = (1, 1) \Rightarrow (U', V') = (0, 0)$ .
2.  $(U, V) = (1, 0) \Rightarrow (U', V') = (1, 0)$ .
3.  $(U, V) = (0, 1) \Rightarrow (U', V') = (0, 1)$ .
4.  $(U, V) = (0, 0)$  et  $K = 1 \Rightarrow (U', V') = (1, 1)$ .
5.  $(U, V) = (0, 0)$  et  $K = 0 \Rightarrow (U', V') = (0, 0)$ .



Ces cinq cas sont illustrés en Figure 1.17. Les cas 1., 2. et 3. proviennent essentiellement du fait que les lignes de Hammersley ne peuvent pas se croiser. Dans les cas 4. et 5. il n'y a pas de ligne en entrée, donc il y aura une ligne de Hammersley au voisinage du point considéré si et seulement si il y a un point à cet endroit, auquel cas elle aura localement une forme de L.

Par hypothèse, si le modèle est stationnaire,  $U$  et  $V$  sont indépendantes de lois respectives  $\text{Ber}(\alpha)$  et  $\text{Ber}(\beta)$ , puisque c'est le cas au voisinage des axes. De plus, ces variables sont également indépendantes de la variable  $Y_{i,j}$  puisque les lignes de Hammersley peuvent être construites progressivement en révélant les points de gauche à droite ou de bas en haut.

Pour obtenir une stationnarité locale, on cherche à ce que les couples  $(U, V)$  et  $(U', V')$  aient la même loi. On peut expliciter la loi du couple  $(U', V')$  grâce aux probabilités totales selon les cinq cas de figure détaillés plus haut. En particulier, on a directement

$$\mathbb{P}((U', V') = (1, 0)) = \mathbb{P}((U, V) = (1, 0)) \quad \text{et} \quad \mathbb{P}((U', V') = (0, 1)) = \mathbb{P}((U, V) = (0, 1)).$$

L'ensemble  $\{0, 1\}^2$  dans lequel ces couples prennent leurs valeurs étant de cardinal 4, il suffit de vérifier par exemple que  $\mathbb{P}((U', V') = (1, 1)) = \mathbb{P}((U, V) = (1, 1))$  pour conclure. Or, par

indépendance de  $U$ ,  $V$  et  $K$ ,

$$\mathbb{P}((U', V') = (1, 1)) = \mathbb{P}((U, V) = (0, 0) \text{ et } K = 1) = (1 - \alpha)(1 - \beta)p.$$

Comme  $\mathbb{P}((U, V) = (1, 1)) = \alpha\beta$ , la condition nécessaire et suffisante de stationnarité locale est donc la relation (1.12).

La stationnarité locale implique la stationnarité globale énoncée dans le théorème 1.10 par récurrence sur les sommets, en voyant localement le couple  $(U, V)$  comme l'entrée et le couple  $(U', V')$  comme la sortie. Les lois en jeu et l'indépendance sont préservées à chaque étape.  $\square$

Pour passer du modèle stationnaire au modèle pur, et ainsi obtenir la convergence (1.11) on procède de manière similaire à la section précédente. On suppose  $p \in ]0, 1[$  fixé et on note pour  $\alpha \in ]0, 1[$  :  $L_\alpha(m, n) = L_{\alpha, \beta}(m, n)$  la longueur du plus long chemin croissant entre les points  $(0, 0)$  et  $(m, n)$  dans le cas stationnaire, c'est-à-dire lorsque la relation (1.12) est vérifiée, ce qui revient à prendre

$$\beta = \frac{p(1 - \alpha)}{\alpha + p(1 - \alpha)} \in ]0, 1[.$$

Comme dans la section 1.3.1, on a la relation déterministe

$$\forall \alpha \in ]0, 1[, \quad L(n, n) \leq L_\alpha(n, n).$$

La stationnarité permet de calculer explicitement l'espérance de la variable aléatoire  $L_\alpha(n, n)$  de la même manière que ce qui a été présenté en espace continu dans la section 1.3.1. Comme elle s'écrit comme une somme de deux variables aléatoires de loi de Bernoulli de paramètres  $\alpha n$  et  $\beta n$  respectivement, on a

$$\mathbb{E}[L_\alpha(n, n)] = (\alpha + \beta)n.$$

Le minimum de la fonction  $(\alpha, \beta) \mapsto \alpha + \beta$  sous la contrainte (1.12) est réalisé pour  $\alpha = \beta$ , ce qui implique  $\alpha = \alpha^*$  où l'on a posé

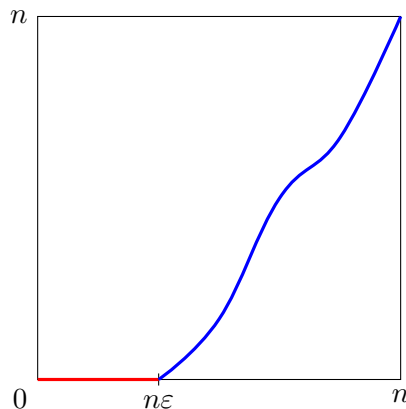
$$\alpha^* = \frac{\sqrt{p}}{1 + \sqrt{p}} \in ]0, 1[.$$

On en déduit alors la majoration

$$\mathbb{E}[L(n, n)] \leq \mathbb{E}[L_{\alpha^*}(n, n)] = \frac{2\sqrt{p}}{1 + \sqrt{p}}n. \tag{1.13}$$

Pour obtenir la minoration, Basdevant-Enriquez-Gerin-Gouéré [BEGG16] procèdent de manière un peu différente de ce qu'ont fait Cator et Groeneboom [CG05], sans utiliser de notion de particule de seconde classe, mais en adaptant une méthode utilisée par Rolla-Sidorovicius-Surgailis-Vares [RSSV10] dans un modèle analogue.

On souhaite montrer que les plus longs chemins dans le cas stationnaire critique, c'est-à-dire avec  $\alpha = \alpha^*$ , prennent un nombre négligeable de sources ou de puits. Pour cela, pour tout réel  $\varepsilon > 0$  on considère la longueur du plus long chemin  $L_{\alpha^*}^{(\varepsilon)}(n, n)$  contraint à rester sur l'axe des abscisses jusqu'au point  $(n\varepsilon, 0)$  exactement. Une représentation schématique d'un tel chemin



**FIGURE 1.18** : Représentation schématique d'un plus long chemin contraint à rester sur l'axe des abscisses jusqu'au point  $(n\varepsilon, 0)$  exactement. La longueur de la partie rouge suit la loi  $\text{Bin}(n\varepsilon, \alpha^*)$ , tandis que celle de la partie bleue a la même loi que la variable  $L(n(1 - \varepsilon), n)$ .

est proposée en Figure 1.18. On a alors l'égalité en loi suivante, en supposant  $n\varepsilon \in \mathbb{N}$  pour simplifier l'écriture,

$$L_{\alpha^*}^{(\varepsilon)}(n, n) \stackrel{(d)}{=} \text{Bin}(n\varepsilon, \alpha^*) + L(n(1 - \varepsilon), n).$$

On peut préciser que les deux variables du membre de droite sont indépendantes, mais ce n'est pas utile ici puisque l'on s'intéresse aux espérances, dans le but de majorer celle de la variable  $L_{\alpha^*}^{(\varepsilon)}(n, n)$ . En particulier, on majore l'espérance de la variable  $L(n(1 - \varepsilon), n)$  en procédant de même que ci-dessus, en introduisant des sources et des puits et en optimisant cette fois-ci dans un rectangle (non-carré). Tous les calculs se font explicitement, et on montre alors que

$$\mathbb{E}\left[L_{\alpha^*}^{(\varepsilon)}(n, n)\right] = n\varepsilon\alpha^* + \mathbb{E}[L(n(1 - \varepsilon), n)] < \frac{2\sqrt{p}}{1 + \sqrt{p}}n.$$

En reprenant cet argument de manière plus fine, avec des inégalités de concentration notamment, on montre alors qu'un plus long chemin dans le modèle stationnaire ne peut pas prendre un nombre macroscopique de sources, et par symétrie ne peut pas prendre un nombre macroscopique de puits non plus. Cela permet de conclure qu'au premier ordre, le modèle pur a la même asymptotique que le modèle stationnaire critique, ce qui prouve la convergence (1.11).

Pour finir, remarquons que ce modèle permet de retrouver l'égalité  $c = 2$  dans le cadre du problème d'Ulam en prenant  $p = \frac{1}{n}$ , de sorte que l'ensemble  $\xi$  approche un PPP d'intensité  $n$  après renormalisation. De manière informelle, la convergence (1.11) suggère dans ce cas l'asymptotique

$$\frac{L(n, n)}{n} \approx \frac{2}{\sqrt{n}} \quad \text{d'où} \quad L(n, n) \approx 2\sqrt{n}.$$

La preuve rigoureuse est établie dans [BEGG16, Section 4].

## 1.4 Fluctuations

Après avoir donné des éléments de preuve pour obtenir l'asymptotique du premier ordre dans des modèles de percolation de dernier passage dans la section précédente, nous nous intéressons ici à l'asymptotique du deuxième ordre, c'est-à-dire les fluctuations. Nous commençons cette section par un bref aparté sur la classe d'universalité dite de Kardar-Parisi-Zhang (KPZ) [KPZ86], présentée notamment dans [Cor16, Occ19].

Le théorème de fluctuation le plus connu est le théorème central limite, généralisation du Théorème de De Moivre-Laplace, qui stipule qu'une somme de  $n$  variables aléatoires d'espérance  $\mu$  et de variance  $\sigma^2$  fluctuent autour de leur moyenne  $\mu n$  à l'échelle  $n^{1/2}$  selon une loi normale centrée de variance  $\sigma^2$ . Cependant, il existe des modèles qui fluctuent autour de leur moyenne à l'échelle  $n^{1/3}$ , selon une loi qui n'est pas une loi normale. Les lois qui apparaissent dans ce cadre peuvent être souvent reliées à des modèles de matrices aléatoires, et sont dites de Tracy-Widom [TW94]. Deux de ces lois seront présentées en section 1.4.1. Cette propriété est typique des modèles de cette classe KPZ.

Historiquement, Kardar Parisi et Zhang [KPZ86] étudient des modèles de croissance d'interfaces (comme le TASEP présenté en section 1.2.2). Il proposent une équation aux dérivées partielles stochastique pour décrire le modèle. Ainsi, si on note  $h(t, x)$  la hauteur de l'interface au temps  $t \in \mathbb{R}_+$  et à la position  $x \in \mathbb{R}$ ,

$$\frac{\partial h}{\partial t}(t, x) = \frac{\partial^2 h}{\partial x^2}(t, x) + \left( \frac{\partial h}{\partial x}(t, x) \right)^2 + \xi(t, x),$$

où  $\xi(t, x)$  est un bruit blanc gaussien espace-temps.

Plusieurs caractéristiques sont typiques de cette classe d'universalité KPZ, toujours avec le point de vue d'évolution d'interface :

- Localisation : l'évolution de l'interface en un point donné ne dépend que de son voisinage.
- Régularisation : de grands trous dans l'interface sont rapidement comblés.
- Dépendance non-linéaire en la pente : la croissance effective de l'interface dépend de la pente locale de manière non-linéaire.
- Bruit espace-temps indépendant : la croissance est gouvernée par un bruit qui décorrèle rapidement en espace et en temps, et est à queue légère.

Au cours de ces trente dernières années, de nombreux modèles ont été conjecturés comme appartenant à cette classe d'universalité, comme par exemple le modèle de déposition balistique introduit par Vold [Vol59] et étudié entre autres par Seppäläinen [Sep00], mais dans la plupart des cas la preuve rigoureuse reste à ce jour un problème ouvert. Dans cette section, nous allons présenter des résultats établis et des éléments de preuves pour le modèle de Hammersley poissonnisé présenté en section 1.1.1, et sa version stationnaire introduit en section 1.3.2. Baik-Deift-Johansson [BDJ99] et Baik et Rains [BR00] les ont étudiés de manière analytique, Cator et Groeneboom [CG06] de manière davantage probabiliste.

Par ailleurs, les fluctuations de la plupart des modèles introduits en section 1.2 ont également été étudiées. Par exemple, Prähofer-Spohn [PS00, PS02] et Johansson [Joh00a] ont étudié



celles de la percolation de dernier passage géométrique (les premiers avec le point de vue du polynuclear growth model), Balázs-Cator-Seppäläinen celles de la percolation de dernier passage exponentielle, Tracy et Widom [TW09] celle de l'ASEP, Ciech et Georgiou [CG19] celles du modèle 1 de [BEGG16] et initialement introduit dans [Sep97]. D'autre part, les fluctuations transverses des géodésiques, c'est-à-dire de leur écart à la diagonale, ont également été étudiées, notamment par Johansson [Joh00b] dans le modèle de Hammersley poissonnisé, et interprété plus récemment par Hammond et Sarka [HS20] en termes de module de continuité de polymères (correspondant aux géodésiques après échelonnage).

### 1.4.1 Approche analytique

Cette section porte principalement sur la première étude des fluctuations des modèles de Hammersley par Baik-Deift-Johansson [BDJ99] et Baik et Rains [BR00]. Ce sont dans ces papiers que des lois de Tracy-Widom sont apparues pour la première fois dans le cadre de problèmes combinatoires. Commençons par définir ces lois rigoureusement.

La fonction d'Airy  $Ai$  est définie par

$$Ai(x) = \frac{1}{\pi} \int_0^\infty \cos\left(\frac{t^3}{3} + xt\right) dt,$$

et est notamment solution de l'équation différentielle ordinaire  $y'' - xy = 0$ . Notons  $u$  la solution de l'équation Painlevé II

$$u''(x) = 2u^3 + xu,$$

avec condition de bord

$$u(x) \underset{x \rightarrow \infty}{\sim} -Ai(x),$$

dont l'existence et l'unicité ont été prouvées par [HM80]. Définissons également les fonctions

$$\begin{aligned} v(x) &= \int_\infty^x (u(s))^2 ds = u(x)^4 + xu(x)^2 - (u'(x))^2, \\ F(x) &= \exp\left(\frac{1}{2} \int_x^\infty v(s) ds\right) = \exp\left(-\frac{1}{2} \int_x^\infty (s-x)(u(s))^2 ds\right), \\ E(x) &= \exp\left(\frac{1}{2} \int_x^\infty u(s) ds\right), \end{aligned}$$

et enfin

$$\begin{aligned} F_{\text{GUE}}(x) &= F(x)^2 = \exp\left(-\int_x^\infty (s-x)(u(s))^2 ds\right), \\ F_0(x) &= \left(1 - (x + 2u'(x) + 2u(x)^2)v(x)\right)(E(x))^4 F_{\text{GUE}}(x). \end{aligned}$$

Ces deux fonctions sont des fonctions de répartition de variables aléatoires réelles. Tracy et Widom [TW94] ont prouvé que la fonction  $F_{\text{GUE}}$  était la fonction de répartition de la loi limite (après centrage et normalisation) de la plus grande valeur propre d'une matrice de loi GUE, définie en section 1.2.2, de taille tendant vers l'infini :

**Théorème 1.11** ([TW94]). *Si  $\lambda_n$  est la plus grande valeur propre d'une matrice de loi GUE sur l'ensemble  $\mathcal{H}_n$  des matrices hermitiennes de taille  $n$ ,*

$$\mathbb{P}\left(\sqrt{2}n^{1/6}(\lambda_n - \sqrt{2n}) \leq s\right) \xrightarrow[n \rightarrow \infty]{} F_{\text{GUE}}(s).$$

En revanche, la fonction  $F_0$  n'a pas d'interprétation connue dans un modèle de matrices aléatoires.

Ce sont ces lois qui apparaissent dans l'écriture des fluctuations dans le modèle considéré. Dans le cas pur, les fluctuations sont d'ordre  $t^{1/3}$  et la loi limite est  $F_{\text{GUE}}$ . Dans le cas stationnaire, le comportement est différent selon s'il est critique (c'est-à-dire  $\lambda = 1$ ) ou non. Dans le premier cas les fluctuations sont d'ordre  $t^{1/3}$  et la loi limite est  $F_0$ , dans le second elles sont gaussiennes d'ordre  $t^{1/2}$ . Plus formellement, le théorème montré par Baik et Rains [BR00], dont le premier point avait été montré par Baik-Deift-Johansson [BDJ99], stipule :

**Théorème 1.12** ([BDJ99, BR00]). *Pour tout  $\lambda > 0$ , on a les convergences en loi suivantes.*

(i) *Pour le modèle pur,*

$$\frac{L(t, t) - 2t}{t^{1/3}} \xrightarrow[t \rightarrow \infty]{(d)} F_{\text{GUE}}.$$

(ii) *Si  $\lambda = 1$ ,*

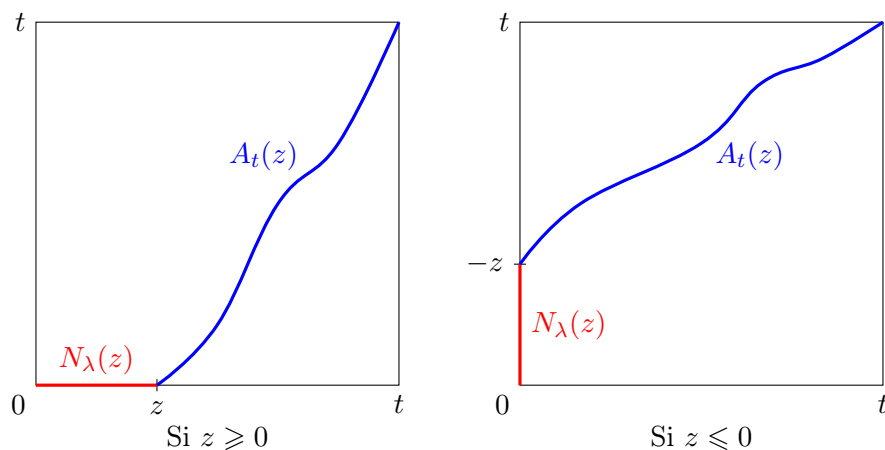
$$\frac{L_1(t, t) - 2t}{t^{1/3}} \xrightarrow[t \rightarrow \infty]{(d)} F_0.$$

(iii) *Si  $\lambda \neq 1$ ,*

$$\frac{L_\lambda(t, t) - (\lambda + \frac{1}{\lambda})t}{\sqrt{t}} \xrightarrow[t \rightarrow \infty]{(d)} \mathcal{N}\left(0, \left|\lambda - \frac{1}{\lambda}\right|\right).$$

Plus généralement, ils détaillent également les fluctuations pour tous les modèles avec sources et puits tel qu'introduit en section 1.3.1. Les sources et les puits sont placés selon un PPP d'intensité  $\lambda$  et  $\mu$  respectivement, sans avoir nécessairement  $\mu = \frac{1}{\lambda}$ , c'est-à-dire sans contrainte de stationnarité. En tout cinq cas se présentent selon les valeurs du couple  $(\lambda, \mu)$ , parmi lesquels les trois énoncés ici dans le Théorème 1.12. Un des deux autres cas donne des fluctuations d'ordre  $t^{1/3}$  avec une loi limite différente de  $F_{\text{GUE}}$  et  $F_0$  mais relié à un modèle de matrices aléatoires, le dernier donne des fluctuations d'ordre  $t^{1/2}$  avec une loi limite correspondant à celle du maximum de deux loi normales centrées réduites indépendantes.

Leurs preuves reposent sur une approche analytique relativement lourde, impliquant notamment des déterminants de Toeplitz et la méthode de descente de Deift et Zhou pour un problème de Riemann–Hilbert. Nous ne détaillerons pas ces éléments ici comme ils s'éloignent significativement du propos du reste de la thèse d'une part, et comme l'approche probabiliste de Cator et Groeneboom [CG06] quelques années plus tard que nous allons présenter en section suivante est plus facile à appréhender d'autre part.



**FIGURE 1.19** : Décomposition d'un plus long chemin entre les points  $(0, 0)$  et  $(t, t)$  contraint à quitter les axes à distance algébrique  $z \in [-t, t]$  de l'origine.

## 1.4.2 Approche probabiliste

Cette section s'inscrit dans la continuité de la section 1.3.1, de la même manière que l'article [CG06] est la continuité de l'article [CG05]. Le résultat principal de Cator et Groeneboom est que la variance de la longueur du plus long chemin entre  $(0, 0)$  et  $(t, t)$  est d'ordre  $t^{2/3}$ , aussi bien dans le cas stationnaire critique que dans le cas pur :

**Théorème 1.13** ([CG06]). *On a les asymptotiques suivantes<sup>1</sup> :*

$$\text{Var}(L_1(t, t)) \underset{t \rightarrow \infty}{\asymp} t^{2/3} \quad \text{et} \quad \text{Var}(L(t, t)) \underset{t \rightarrow \infty}{\asymp} t^{2/3}.$$

Ce théorème est plus faible que le Théorème 1.12 énoncé ci-avant, le principal intérêt du papier [CG06] étant la technique de preuve, basée sur une approche probabiliste utilisant notamment les particules de seconde classe et les points où les plus longs chemins quittent les axes.

Intéressons-nous plus particulièrement à ces points. On les repère par leur distance algébrique au point  $(0, 0)$  : comptée positivement s'ils sont situés sur l'axe des abscisses et négativement sur l'axe des ordonnées. Ainsi, pour  $t > 0$  et  $z \in [-t, t]$ , on note dans le cadre du modèle stationnaire de paramètre  $\lambda > 0$  :

- $N_\lambda(z)$  le nombre de sources sur le segment  $[0, z] \times \{0\}$  si  $z \geq 0$ , ou le nombre de puits sur le segment  $\{0\} \times [0, -z]$  si  $z \leq 0$ .
- $A_t(z)$  la longueur du plus long chemin croissant jusqu'au point  $(t, t)$ , depuis le point  $(z, 0)$  si  $z \geq 0$  ou depuis le point  $(0, -z)$  si  $z \leq 0$ .

Par conséquent, un plus long chemin entre les points  $(0, 0)$  et  $(t, t)$  contraint à quitter les axes à distance algébrique  $z \in [-t, t]$  de l'origine a pour longueur  $N_\lambda(z) + A_t(z)$ . Cette

<sup>1</sup>On utilise la notation  $f(t) \asymp g(t)$  pour signifier " $f(t) = O(g(t))$  et  $g(t) = O(f(t))$ ".

décomposition est illustrée par la Figure 1.19. En optimisant sur le paramètre  $z$ , on obtient l'expression suivante pour la variable  $L_\lambda(t, t)$  :

$$L_\lambda(t, t) = \max_{z \in [-t, t]} (N_\lambda(z) + A_t(z)) \quad (1.14)$$

De plus, on note  $Z_\lambda(t)$  la variable aléatoire correspondant au plus grand réel  $z$  tel que le maximum de l'équation (1.14) est satisfait, c'est-à-dire la distance algébrique à l'origine du point où le plus long chemin croissant entre les points  $(0, 0)$  et  $(t, t)$  prenant le plus de sources et le moins de puits possible quitte les axes.

Le principal résultat nouveau de Cator et Groeneboom [CG06] est le suivant, dans le cas critique  $\lambda = 1$  :

**Théorème 1.14** ([CG06]). *Pour tout réel  $t > 0$ ,*

$$\text{Var}(L_1(t, t)) = 2\mathbb{E}[Z_1(t)_+].$$

Grâce à cette formule, ils ramènent l'étude de la variance recherchée à l'étude du point le plus à droite où un plus long chemin quitte l'axe des abscisses. Pour établir ce théorème, ils montrent d'abord l'égalité suivante

$$\text{Var}(L_1(t, t)) = 2\mathbb{E}[(t - X_1(t))_+], \quad (1.15)$$

où  $X_\lambda(t)$  désigne l'abscisse de la particule de seconde classe partant de  $(0, 0)$  au temps  $t$  comme dans le Théorème 1.9. La preuve de cette égalité utilise notamment la réversibilité énoncée dans le Théorème 1.8, qui permet d'obtenir directement, en utilisant l'égalité (1.9) et la bilinéarité de la covariance :

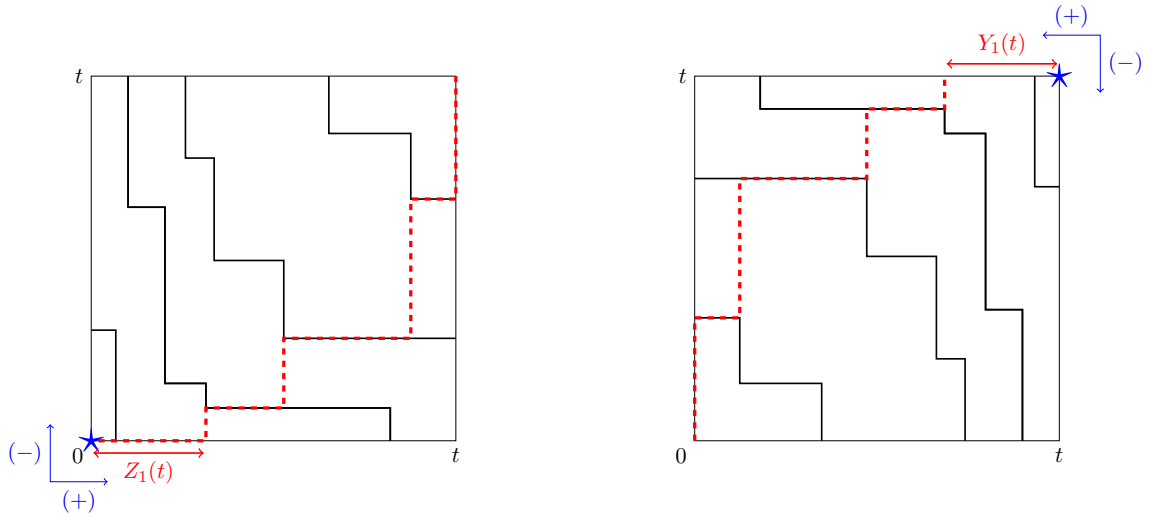
$$\text{Var}(L_1(t, t)) = 2\text{Cov}(S, N),$$

avec les notations des points cardinaux introduites par le Théorème 1.9, pour  $\lambda = 1$  et  $x = t$ . Cette covariance peut s'exprimer en considérant un système où l'on a augmenté l'intensité des sources d'un réel  $\varepsilon > 0$ , en étudiant le premier ordre pour  $\varepsilon$  tendant vers 0 de l'espérance de la variable nord  $N$  dans ce nouveau modèle. Cette espérance fait intervenir la variable  $X_1(t)$ , en remarquant que si on augmente la variable sud  $S$  d'une unité en ajoutant une source au point  $(z, 0)$ , alors la variable nord  $N$  augmentera aussi si et seulement si la particule de seconde classe partant de  $(z, 0)$  sort par le haut du rectangle  $[0, t]^2$ . On obtient finalement l'équation (1.15).

Le Théorème 1.14 se déduit alors en utilisant un argument similaire au Lemme 1.2 : la distance algébrique entre l'origine et le point de sortie de l'axe des abscisses du plus long chemin prenant le plus de sources,  $Z_1(t)$ , est la même que la distance algébrique entre le point  $(t, t)$  et le point de sortie de la boîte  $[0, t]^2$  d'une particule de seconde classe pour le système retourné, qu'ils appellent  $Y_1(t)$ . Cette correspondance est illustrée en Figure 1.20. En particulier, celle-ci est positive si et seulement si la particule de seconde classe sort de la boîte  $[0, t]^2$  par le haut, auquel cas  $Y_1(t) = t - X_1(t)$ . Par réversibilité (Théorème 1.8), on déduit finalement l'égalité en loi

$$(t - X_1(t))_+ \stackrel{(d)}{=} Z_1(t)_+,$$

puis le Théorème 1.14.



(a) Plus long chemin prenant le plus de sources et le moins de puits pour un système donné.

(b) Trajectoire d'une particule de seconde classe pour le système retourné jusqu'à ce qu'elle quitte la boîte  $[0, t]^2$ .

**FIGURE 1.20** : Lien entre plus long chemin et particule de seconde classe. Les distances algébriques  $Z_1(t)$  et  $Y_1(t)$  sont mesurées à partir du point  $\star$ , comptées positivement horizontalement et négativement verticalement.

Donnons désormais une heuristique du Théorème 1.13 utilisant le Théorème 1.14. On s'attend à ce que la variable  $Z_1(t)$  soit typiquement de l'ordre de  $ut^\beta$ , où  $u$  est une constante universelle et où  $\beta \in ]0, 1[$ . Le but est d'obtenir  $\beta = 2/3$ .

Reprenons les notations de l'équation (1.14), et étudions les deux parties d'un plus long chemin pour  $z = ut^\beta$ , dans le cas  $\lambda = 1$ .

- D'une part, comme  $N_1$  est un PPP d'intensité 1, le théorème de Donsker fournit l'asymptotique au second ordre :

$$N_1(ut^\beta) \approx ut^\beta + t^{\beta/2} B_u, \quad (1.16)$$

où  $B$  est un mouvement brownien standard.

- D'autre part, par définition de l'équation (1.14) comme un maximum, on a pour tout  $\lambda > 0$ ,

$$A_t(ut^\beta) \leq L_\lambda(t, t) - N_\lambda(ut^\beta). \quad (1.17)$$

On procède de même que pour obtenir l'équation (1.13) dans la section 1.3.2. L'espérance du membre de droite se calcule explicitement :

$$\mathbb{E} \left[ L_\lambda(t, t) - N_\lambda(ut^\beta) \right] = \left( \lambda + \frac{1}{\lambda} \right) t - \lambda ut^\beta.$$

Le minimum en  $\lambda$  de cette quantité est atteint pour  $\lambda = \lambda^*$ , avec

$$\lambda^* = \frac{1}{\sqrt{1 - ut^{\beta-1}}} \approx 1 + \frac{u}{2} t^{\beta-1}.$$

En injectant dans l'équation (1.17), on obtient alors une majoration pour  $\mathbb{E}[A_t(ut^\beta)]$ . On peut alors procéder comme dans la section 1.3.2 pour montrer que cette majoration pour  $\lambda = \lambda^*$  est optimale, dans le sens où elle donne le bon ordre de grandeur pour la variable  $A_t(ut^\beta)$ , ainsi, on obtient l'approximation suivante au premier ordre

$$A_t(ut^\beta) \approx 2t - \frac{u^2}{4}t^{2\beta-1} - ut^\beta, \quad (1.18)$$

où l'on a supposé que ses fluctuations sont bien négligeables devant ces deux termes.

Grâce aux asymptotiques (1.16) et (1.18), on en déduit celle de la longueur d'un plus long chemin contraint à quitter les axes à l'abscisse  $ut^\beta$  :

$$N_1(ut^\beta) + A_t(ut^\beta) \approx 2t + t^{\beta/2}B_u - \frac{u^2}{4}t^{2\beta-1}.$$

Dans le cas où ce chemin quitte les axes au point optimal, c'est-à-dire à l'endroit où il l'aurait quitté sans contrainte, on s'attend à ce que les deuxièmes et troisièmes termes soient du même ordre de grandeur. On doit donc dans ce cas avoir l'égalité

$$\frac{\beta}{2} = 2\beta - 1 \quad \text{d'où} \quad \beta = \frac{2}{3}.$$

Dans leur papier [CG06], Cator et Groeneboom montrent une majoration puis une minoration de la quantité  $\mathbb{E}[Z(t)_+]$  par un terme de l'ordre de  $t^{2/3}$ . Pour la majoration, ils utilisent comme dans leur précédent article [CG05] évoqué en section 1.3.1 le procédé de thinning et de thickening, puisqu'ils couplent le modèle  $L_1$  avec un modèle  $L_\lambda$  pour  $\lambda > 1$ . Pour la minoration, ils utilisent à nouveau les particules de seconde classe, en remarquant notamment que leur trajectoire délimite la zone d'influence des puits. Autrement dit, si on se donne la trajectoire d'une particule de seconde classe depuis l'origine, alors les lignes de Hammersley situées en bas à droite de celle-ci ne sont pas modifiées si on retire les puits. Il en est de même de manière symétrique avec les particules de seconde classe duales qui définissent la zone d'influence des sources.

Pour obtenir le deuxième point du théorème, c'est-à-dire obtenir les fluctuations de la variable  $L(t, t) = A_t(0)$ , ils montrent la borne suivante pour  $L \geq 1$

$$\mathbb{P}\left(L_1(t, t) - L(t, t) \geq Lt^{1/3}\right) = O(L^{-5/4}),$$

en utilisant là encore les particules de seconde classe.

## 1.5 Résultats de la thèse

Dans cette section nous exposons les différents résultats obtenus au cours de la thèse. Trois modèles ont été étudiés indépendamment et sont présentés successivement ci-après. Ils ont en commun d'être des modèles aléatoires définis dans le plan et possédant une propriété de stationnarité bidimensionnelle.

Le premier est le modèle de Hammersley stationnaire dans le quart de plan, introduit et étudié par Groeneboom [Gro02] et Cator et Groeneboom [CG05], et présenté ici en section 1.3.1. Nous présentons une preuve probabiliste des fluctuations gaussiennes dans le cas non critique, c'est-à-dire le cas (iii) du Théorème 1.12 énoncé en section 1.4.1.

Le deuxième modèle peut être vu comme une version stationnaire du problème d'O'Connell-Yor [OY01]. La preuve de sa stationnarité est obtenue en introduisant une discrétisation de ce modèle dont nous montrons la stationnarité, puis en observant que cette stationnarité est préservée à la limite.

Enfin, le troisième modèle est une classe générale de systèmes aléatoires de lignes brisées dans le quart de plan, dont on montre la réversibilité. Cette classe contient de nombreux processus classiques, comme celui de Hammersley présenté en section 1.1.1 ou la plupart des modèles discrets de percolation de dernier passage présentés en section 1.3.2. La nouveauté ici est qu'un poids est associé à chaque ligne. Les lignes sont initialement générées par un PPP et leur évolution est régie par une dynamique markovienne qui préserve la loi des nœuds de Kirchhoff pour le poids des lignes à chaque intersection.

### 1.5.1 Fluctuations gaussiennes pour les processus de Hammersley : une preuve probabiliste

Ce chapitre se place dans le cadre du modèle de Hammersley stationnaire dans le quart de plan, introduit et étudié par Groeneboom [Gro02] et Cator et Groeneboom [CG05, CG06], et présenté ici en section 1.3.1. On rappelle que la variable  $L_\lambda(x, t)$  est définie comme étant la longueur d'une géodésique entre les points  $(0, 0)$  et  $(x, t)$  dans le modèle où les points  $\alpha$  sont distribués selon un PPP d'intensité 1 dans le quart de plan, les sources selon un PPP d'intensité  $\lambda$  sur l'axe des abscisses et les puits selon un PPP d'intensité  $\lambda^{-1}$  sur l'axe des ordonnées.

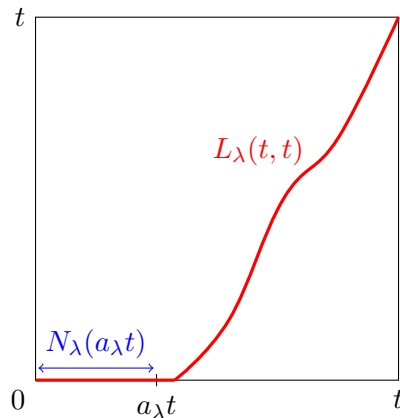
Les fluctuations du processus  $(L_\lambda(t, t))_{t>0}$  ont été établies par Baik et Rains [BR00] dans le Théorème 1.12 comme présenté en section 1.4.1. Deux comportements différents ont été exhibés selon si  $\lambda = 1$ , cas que nous appelons *critique*, ou non.

Cator et Groeneboom [CG06] ont étudié les fluctuations dans le cas critique en utilisant des arguments probabilistes "légers", comme présenté en section 1.4.2. [CG06] fournit une preuve d'une version plus faible du point (ii) du Théorème 1.12, avec l'ordre des fluctuations mais pas la loi limite. Plus précisément, ils prouvent que

$$\text{Var}(L_1(t, t)) \underset{t \rightarrow \infty}{\asymp} t^{2/3}.$$

Dans le papier précédent [CG05] présenté en section 1.3.1, toujours pour  $\lambda = 1$ , ils prouvent qu'une géodésique entre les points  $(0, 0)$  et  $(t, t)$  passe une proportion de temps négligeable sur les axes, plus précisément de l'ordre de  $t^{2/3}$  comme prouvé dans [CG06]. Dans un certain sens, cela suggère que le processus  $(L_1(t, t))_{t>0}$  a les mêmes fluctuations que le modèle pur  $(L(t, t))_{t>0}$ , qui ont été prouvées comme étant en  $t^{1/3}$  par Baik-Deift-Johannson [BDJ99] (mais avec une loi limite différente, c.f. Théorème 1.12 (i)).

On étudie ici le cas non critique  $\lambda \neq 1$ . Supposons sans perte de généralité  $\lambda > 1$  quitte à considérer le modèle symétrisé par rapport à la première bissectrice. On montre que dans ce modèle, les géodésiques entre les points  $(0, 0)$  et  $(t, t)$  passent une proportion macroscopique



**FIGURE 1.21 :** Représentation schématique d'un plus long chemin typique pour  $\lambda > 1$ , avec  $a_\lambda = 1 - \lambda^{-2}$ . Sa longueur totale  $L_\lambda(t, t)$  a les mêmes fluctuations que le nombre de sources sur le segment  $[0, a_\lambda t] \times \{0\}$ ,  $N_\lambda(a_\lambda t)$ .

du temps sur les axes, et que le point où elles quittent les axes est concentré autour du point  $(0, (1 - \lambda^{-2})t)$ . Notre principal résultat est que les fluctuations du processus  $(L_\lambda(t, t))$  sont les mêmes que celles du nombre de sources situées jusqu'à ce point. Ainsi, si on note comme dans la section 1.4.2  $N_\lambda(z) = L_\lambda(z, 0)$  le nombre de sources sur le segment  $[0, z] \times \{0\}$ , pour  $z \geq 0$ , qui suit une loi de Poisson de paramètre  $\lambda z$  :

**Théorème 1.15.** *Pour tout  $\lambda > 1$ , on a la convergence en probabilité suivante :*

$$\frac{L_\lambda(t, t) - (\lambda + \frac{1}{\lambda})t}{\sqrt{t}} - \frac{N_\lambda((1 - \frac{1}{\lambda^2})t) - (\lambda - \frac{1}{\lambda})t}{\sqrt{t}} \xrightarrow[t \rightarrow \infty]{\mathbb{P}} 0.$$

Ce phénomène est illustré par la Figure 1.21. On en déduit directement comme corollaire les fluctuations gaussiennes énoncées par le point (iii) du Théorème 1.12 :

**Corollaire 1.1.** *Pour tout  $\lambda \neq 1$ ,*

$$\frac{L_\lambda(t, t) - (\lambda + \frac{1}{\lambda})t}{\sqrt{t}} \xrightarrow[t \rightarrow \infty]{(d)} \mathcal{N}\left(0, \left|\lambda - \frac{1}{\lambda}\right|\right).$$

Nous donnons ici une heuristique de la preuve du Théorème 1.15. Intéressons-nous au point où une géodésique entre les points  $(0, 0)$  et  $(t, t)$  quitte les axes. Notons  $L_\lambda^{(a)}(t, t)$  la longueur d'une géodésique contrainte à quitter l'axe des abscisses au point  $(at, 0)$ , de sorte que

$$L_\lambda(t, t) = \max_{a \in [0, 1]} L_\lambda^{(a)}(t, t).$$



On cherche désormais à trouver le paramètre  $a$  optimal. Remarquons que comme  $\lambda > 1$ , l'intensité des sources  $\lambda$  est plus importante que celle des puits  $\lambda^{-1}$ . Par le même argument de symétrie que celui évoqué ci-avant, si une géodésique commence verticalement est quitte l'axe des ordonnées en un point  $(0, at)$  pour  $a \in (0, 1]$ , alors sa longueur est strictement stochastiquement dominée par celle d'une géodésique contrainte à quitter l'axe des abscisses au point  $(at, 0)$ . Cela implique donc qu'une géodésique typique commence sur l'axe des abscisses et non sur l'axe des ordonnées, ce qui suggère  $a > 0$ . D'autre part, la variable  $L_\lambda^{(a)}(t, t)$  peut s'écrire comme une somme de deux variables aléatoires (indépendantes). La première correspond à la variable  $N_\lambda(a\lambda t)$  et suit donc la loi  $\text{Poi}(a\lambda t)$ , la deuxième correspond à la longueur d'une géodésique entre les points  $(at, 0)$  et  $(t, t)$  contrainte à ne prendre aucune source (et aucun puits comme  $a > 0$ ). Par invariance par translation, cette deuxième variable a la même loi que la variable  $L((1-a)t, t)$ . En utilisant l'asymptotique  $\mathbb{E}[L(x, t)] \underset{tx \rightarrow \infty}{\sim} 2\sqrt{xt}$  présenté en section 1.2.1, on a alors

$$\mathbb{E}\left[L_\lambda^{(a)}(t, t)\right] = \mathbb{E}[N_\lambda(a\lambda t)] + \mathbb{E}[L((1-a)t, t)] \approx a\lambda t + 2\sqrt{1-a}t.$$

Or, la fonction  $a \mapsto a\lambda + 2\sqrt{1-a}$  admet un unique maximum global atteint au point  $a = a_\lambda$ , avec

$$a_\lambda = 1 - \frac{1}{\lambda^2}.$$

On conclut donc qu'une géodésique entre les points  $(0, 0)$  et  $(t, t)$  quitte l'axe des abscisses au voisinage du point  $(a_\lambda t, 0)$ .

Comme les fluctuations de la variable  $N_\lambda(a_\lambda t)$  sont d'ordre  $t^{1/2}$  et celle de la variable  $L((1-a_\lambda)t, t)$  sont d'ordre  $t^{1/3}$ , et donc négligeables devant les premières, on en déduit que les fluctuations de la variable  $L_\lambda(t, t)$  sont identiques à celles de  $N_\lambda(a_\lambda t)$ , et donc en particulier d'ordre  $t^{1/2}$ .

### 1.5.2 Une version stationnaire du problème d'O'Connell-Yor

Soit  $B = (B_k)_{k \geq 1}$  une suite de mouvements browniens standards indépendants,  $n$  un entier strictement positif et  $u$  un réel strictement positif. Considérons la variable aléatoire suivante :

$$L_{n,u}(B) := \sup_{0=t_0 < t_1 < \dots < t_{n-1} < t_n = u} \left( \sum_{k=1}^n [B_k(t_k) - B_k(t_{k-1})] \right). \quad (1.19)$$

Cette quantité apparaît dans des modèles de files d'attente et de percolation de dernier passage : des formules similaires peuvent être trouvées dans [Mut79, SK90, Gan98]. Cette variable peut aussi être interprétée comme l'énergie maximale d'un chemin dans un modèle semi-discret de polymères introduit par O'Connell et Yor [OY01]. La question de trouver la loi de la variable  $L_{n,u}(B)$  a été posée par Glynn et Whitt [GW91]. Remarquons que par un argument d'échelle il suffit d'étudier le cas  $u = 1$ , et on notera par la suite  $L_n(B) = L_{n,1}(B)$ . Il a été observé indépendamment par Baryshnikov dans [Bar01] et par Gravner–Tracy–Widom dans [GTW01] que la loi de cette variable coïncide avec celle de la plus grande valeur propre d'une matrice GUE de taille  $n \times n$ . Par la suite, O'Connell et Yor in [OY02] ont prouvé que cette correspondance pouvait être étendue à tout le spectre d'une matrice GUE en considérant des extensions de la fonctionnelle  $L_n$ .

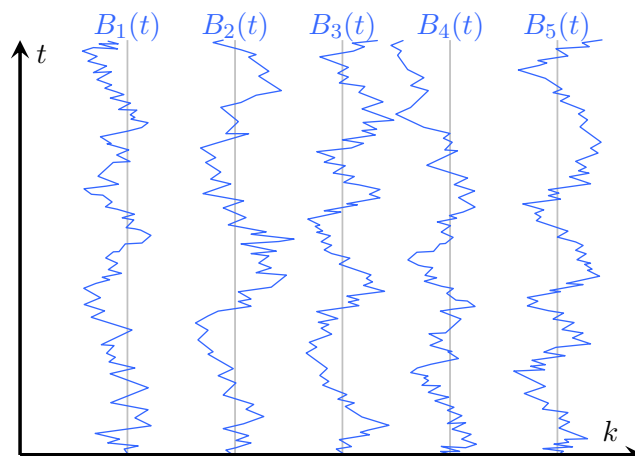


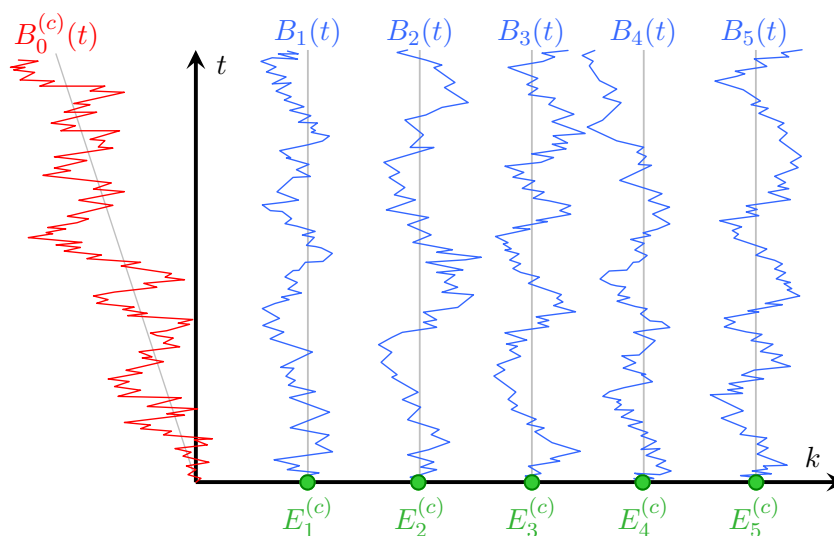
FIGURE 1.22 : Un exemple de la représentation verticale des mouvements browniens

De façon peu surprenante, la loi de la variable  $L_n(B)$  ne peut pas être calculée explicitement. Ici, nous exhibons une modification de la fonctionnelle  $L_n$  pour laquelle un processus sous-jacent à deux paramètres est stationnaire à la fois en le paramètre spatial  $n$  (ce que nous appellerons la stationnarité *horizontale*) et en le paramètre temporel  $u$  (stationnarité *verticale*). Ce processus peut être vu comme un modèle de files d'attente browniennes avec  $n$  serveurs jusqu'au temps  $u$  comme défini par O'Connell et Yor dans [OY01]. Dans ce papier, la stationnarité verticale est en fait implicite : les processus sont définis à partir de  $-\infty$ , de telle sorte qu'au temps  $t = 0$  le système a atteint sa mesure stationnaire par ergodicité. Ce processus a également été étudié par Konstantopoulos dans [Kon16], où la stationnarité horizontale est prouvée dans le cas d'un seul serveur.

Ici, nous reprenons et synthétisons les deux notions de stationnarité dans un seul résultat. Nous les prouvons simultanément, en utilisant une discrétisation du processus qui peut être alors vu comme un modèle de percolation de dernier passage similaire à ceux présentés en section 1.2, cette fois dans le quart de plan  $\mathbb{N} \times \mathbb{R}_+$ , discret dans une direction et continu dans l'autre. La stationnarité de ce processus discrétisé découle d'un argument de stationnarité locale dans l'esprit de ce qui a été fait dans [Sep09, Lemme 2.3] et [BEGG16, Section 3] dans des modèles discrets de percolation de dernier passage, comme celui présenté en section 1.3.2. Le théorème principal s'en déduit alors, et se formule de manière concise.

Une des conséquences de cette stationnarité est que la loi de la version modifiée de  $L_n(B)$  peut être écrite comme la somme de deux variables aléatoires (négativement corrélées) dont la loi est explicite.

Comme mentionné ci-avant, la variable décrite dans l'équation (1.19) représente la longueur du plus long chemin dans un problème de percolation de dernier passage. En effet, cela revient à considérer des chemins croissants depuis le point sud-ouest  $(1, 0)$  jusqu'au point nord-est  $(n, u)$  dans le rectangle  $\llbracket 1, n \rrbracket \times [0, u]$ , où la longueur d'un segment de la forme  $\{k\} \times [s, t]$  est définie par l'incrément du mouvement brownien  $B_k$  entre les temps  $s$  et  $t$ . En d'autres termes,



**FIGURE 1.23 :** Un exemple de la représentation verticale des mouvements browniens (en bleu) avec les sources (en vert) et le processus des puits (en rouge)

on saute successivement d'un mouvement brownien au suivant dans le but de maximiser la somme des incréments des mouvements browniens observés entre les temps de saut. Dans le modèle initial décrit ci-dessus, on considère uniquement des mouvements browniens standards indépendants. Ils peuvent être vus comme évoluant sur chaque ligne verticale  $\{k\} \times \mathbb{R}_+$ , avec  $k \in \llbracket 1, n \rrbracket$ . La Figure 1.22 montre un exemple de cette représentation. La modification évoquée ci-dessus consiste à ajouter deux processus qui peuvent être vus comme évoluant sur chacun des deux axes. Sur l'axe des ordonnées, on considère également un mouvement brownien standard indépendant, cette fois avec dérive, appelé *processus des puits*. Comparé aux autres mouvements browniens, on obtient un bonus déterministe proportionnel au temps passé sur l'axe. D'autre part, une suite de variables aléatoires i.i.d. de loi exponentielle est ajoutée sur les points à coordonnées entières de l'axe des abscisses, appelés *sources*. Ces variables aléatoires peuvent être vues comme des récompenses : à chaque fois que l'on décide d'ignorer un mouvement brownien (avec dérive ou non), on récupère la valeur de la variable aléatoire exponentielle correspondante. La Figure 1.23 montre un exemple de cette représentation. Notons que pour avoir un chemin croissant, les récompenses ne peuvent être gagnées qu'au début du chemin. Autrement dit, il est impossible de tirer profit à la fois des puits et des sources.

Plus formellement, soit  $n \in \mathbb{N}, c > 0$ . Considérons la collection suivante de variables aléatoires indépendantes :

- $B = (B_k)_{1 \leq k \leq n}$  une suite de mouvements browniens standards indépendants,
- $B_0^{(c)}$  un mouvement brownien standard de dérive  $c$ , c'est-à-dire  $B_0^{(c)}(t) = B_0(t) + ct$  où  $B_0$  est un mouvement brownien standard,
- $E^{(c)} = (E_k^{(c)})_{1 \leq k \leq n}$  une suite de variables aléatoires i.i.d. de loi exponentielle d'espérance  $1/c$ .

Nous définissons alors notre principale variable d'intérêt, pour tout  $u > 0$ ,

$$L_{n,u}^{(c)} = \sup_{i,t_0,\dots,t_{n-1}} \left( \sum_{k=1}^i E_k^{(c)} + B_0^{(c)}(t_0) + \sum_{k=1}^n [B_k(t_k) - B_k(t_{k-1})] \right), \quad (1.20)$$

où le supremum est considéré sur l'ensemble

$$\{(i, t_0, \dots, t_{n-1}) : i \in \llbracket 0, n \rrbracket ; \forall j \leq i-1, t_j = 0 ; 0 < t_i < t_{i+1} < \dots < t_n = u\}.$$

Comme pour la variable aléatoire  $L_{n,u}(B)$ , la variable  $L_{n,u}^{(c)}$  représente la longueur du plus long chemin dans un problème de percolation de dernier passage, mais maintenant avec des conditions aux bords données par les variables  $E^{(c)}$  et par le processus  $B_0^{(c)}$ . Autrement dit, le chemin est autorisé à obtenir soit les récompenses de la forme  $E_i^{(c)}$  en traversant le point  $(i, 0)$  de l'axe des abscisses, soit les incréments du mouvement brownien avec dérive le long de l'axe des ordonnées jusqu'au temps  $t_0$ . La variable  $i$  correspond au nombre de récompenses exponentielles obtenues, c'est-à-dire que le second cas a lieu si et seulement si  $i = 0$ , auquel cas  $t_0$  est non nul tout comme  $B_0^{(c)}(t_0)$ .

Le résultat principal est que cette modification rend le système stationnaire sans le sens suivant :

**Théorème 1.16.** *Pour tout couple  $(n, s) \in \mathbb{N} \times \mathbb{R}_+$ , la loi du processus*

$$\left( L_{n+k,s+t}^{(c)} - L_{n,s}^{(c)} \right)_{(k,t) \in \mathbb{N} \times \mathbb{R}_+}$$

*ne dépend pas du couple  $(n, s)$ .*

En particulier, cette stationnarité fournit de l'information sur la loi de la variable  $L_{n,u}^{(c)}$  : elle peut être écrite comme la somme d'une variable aléatoire gaussienne et une variable aléatoire de loi Gamma.

**Proposition 1.1.** *Il existe deux variables aléatoires (dépendantes)  $Z_u^{(cu)}$  et  $\Gamma_n^{(c)}$  de lois marginales respectives  $\mathcal{N}(cu, u)$  et  $\text{Gamma}(n, c)$  telles que*

$$L_{n,u}^{(c)} \stackrel{(d)}{=} Z_u^{(cu)} + \Gamma_n^{(c)}. \quad (1.21)$$

La Proposition 1.1 permet également d'obtenir des informations sur le problème original sans sources ni puits. Avec les notations de l'équation (1.19), on a la domination stochastique suivante :

$$L_n(B) \leq L_{n,1}^{(c)}.$$

Pour avoir la majoration la plus fine, on peut chercher à optimiser en le paramètre  $c$  le membre de droite. Une heuristique est de le faire en espérance, qui est égale à  $c + \frac{n}{c}$  et minimale pour  $c = \sqrt{n}$ . On utilise le modèle stationnaire avec cet échelonnage particulier entre les paramètres  $c$  et  $n$  dans notre prochain résultat. Nous prouvons qu'un chemin optimal<sup>2</sup> dans ce cas est asymptotiquement proche du segment qui relie les points  $(0, 0)$  et  $(n, 1)$ .

Formellement, prenons  $u = 1$  et notons  $P_n^{(\sqrt{n})}$  le chemin optimal dans le cas stationnaire avec  $c = \sqrt{n}$  qui a été rééchelonné dans le carré  $[0, 1]^2$ , c'est-à-dire

$$P_n^{(\sqrt{n})} := \bigcup_{j=0}^n \left( \left\{ \frac{j}{n} \right\} \times [\mathbf{t}_{j-1}^*, \mathbf{t}_j^*] \right), \quad (1.22)$$

où la subdivision  $(\mathbf{t}_0^*, \dots, \mathbf{t}_{n-1}^*)$  réalise le supremum dans l'équation (1.20) et où l'on a posé  $\mathbf{t}_{-1}^* := 0$  et  $\mathbf{t}_n^* := 1$ . De plus, notons  $\Delta$  la première diagonale du carré  $[0, 1]^2$ , c'est-à-dire

$$\Delta := \{(x, y) \in [0, 1]^2 : x = y\},$$

et pour tout sous-ensemble non vide  $E \subset \mathbb{R}^2$ , définissons sa distance à la diagonale  $\Delta$  par

$$d(E, \Delta) = \sup_{(x,y) \in E} |x - y|.$$

**Proposition 1.2.** *Le chemin optimal rééchelonné  $P_n^{(\sqrt{n})}$  est asymptotiquement proche de la diagonale  $\Delta$ , dans le sens où pour tout réel  $\beta > 4$ ,*

$$\mathbb{P} \left( d \left( P_n^{(\sqrt{n})}, \Delta \right) \geq \frac{1}{n^{1/\beta}} \right) \xrightarrow{n \rightarrow \infty} 0.$$

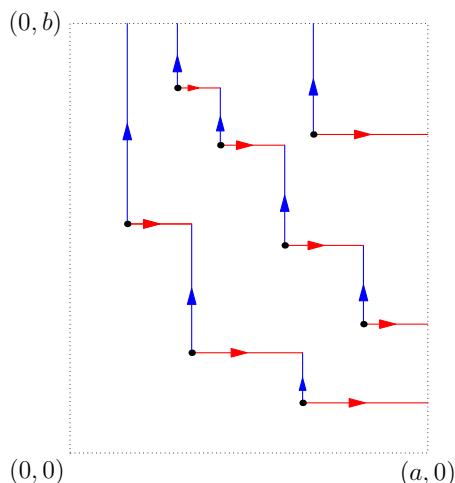
Cette concentration du chemin optimal dans le modèle stationnaire autour de la diagonale nous indique qu'il utilise peu les sources et les puits sur les axes. Par conséquent, cela permet de retrouver le résultat classique suivant, étant donné que la variable  $L_n(B)$  a la même loi que la plus grande valeur propre d'une matrice GUE de taille  $n \times n$  :

**Corollaire 1.2.** *On a la convergence presque sûre suivante :*

$$\frac{L_n(B)}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{p.s.} 2.$$

Une autre preuve de ce résultat utilisant le processus  $L_n(B)$  a été faite dans [HMO02] en utilisant des arguments de [OY01] basés sur une technique introduite dans [Sep98b] qui exploitent une sorte de dualité convexe entre la densité et la vitesse dans des modèles microscopiques de systèmes hydrodynamiques.

<sup>2</sup>Le supremum est en fait un maximum, et il existe un unique chemin optimal.



**FIGURE 1.24** : Une réalisation du processus de lignes brisées de Hammersley dans le rectangle  $[0, a] \times [0, b]$ . Les traces des particules de charge  $+1$  sont représentées en rouge et celles des antiparticules de charge  $-1$  sont représentées en bleu.

### 1.5.3 Systèmes de Poisson–Kirchhoff réversibles

*Travail en commun avec Jérôme Casse<sup>3</sup>, Nathanaël Enriquez<sup>4</sup> et Arvind Singh<sup>5</sup>.*

Dans son papier fondateur [Ham72], Hammersley a introduit son désormais célèbre *processus de lignes brisées* comme un moyen d’étudier la longueur de la plus longue sous-suite croissante dans une permutation aléatoire. Ce modèle de percolation de dernier passage possède de nombreuses propriétés remarquables et a depuis été analysé en profondeur [Ros81, Sep09]. Une construction possible du processus de Hammersley dans le quart de plan  $[0, \infty)^2$  est la suivante : considérons un PPP d’intensité 1 sur  $[0, \infty)^2$ . Chaque atome de ce processus ponctuel “émet” une paire de particule/anti-particule avec la particule de charge  $+1$  se déplaçant horizontalement vers la droite et l’antiparticule de charge  $-1$  se déplaçant verticalement vers le haut. Quand les traces de deux particules de charges opposées se rencontrent, elles disparaissent toutes les deux. Ainsi, la collection de toutes les traces obtenues avec cette procédure est exactement le processus de lignes brisées de Hammersley dans le quart de plan défini en section 1.1.1. Une illustration de cette construction est présentée en Figure 1.24. Remarquons que, au vu de cette construction, le système peut être appelé “conservatif” dans le sens où la charge totale du système reste nulle comme les particules et antiparticules apparaissent et disparaissent simultanément.

Dans ce chapitre, nous introduisons une nouvelle classe de processus aléatoires que nous appelons *Systèmes de Poisson–Kirchhoff* (*Poisson–Kirchhoff Systems* en anglais, PKS) qui généralise la construction décrite ci-dessus. Ces processus consistent à nouveau en des collections aléatoires de lignes brisées horizontales et verticales *pondérées*, dans le sens où elles sont chacune associée à un poids, dans le quart de plan  $[0, \infty)^2$ . Comme pour le processus de lignes brisées de Hammersley, on peut voir ces lignes comme étant les traces de particules “chargées” se déplaçant

<sup>3</sup>LMO, Université Paris-Saclay. Email : [jerome.casse.math@gmail.com](mailto:jerome.casse.math@gmail.com)

<sup>4</sup>LMO, Université Paris-Saclay. Email : [nathanael.enriquez@universite-paris-saclay.fr](mailto:nathanael.enriquez@universite-paris-saclay.fr)

<sup>5</sup>CNRS. LMO, Université Paris-Saclay. Email : [arvind.singh@universite-paris-saclay.fr](mailto:arvind.singh@universite-paris-saclay.fr)

soit horizontalement (i.e. augmentant leur abscisse  $x$ ) ou verticalement (i.e. augmentant leur ordonnée  $y$ ). Cependant, dans cette nouvelle classe de processus, des particules peuvent avoir des charges arbitraires et peuvent de manière aléatoire tourner, se diviser ou coalescer selon une dynamique markovienne spéciale. Cette dynamique préserve la loi des nœuds de Kirchhoff pour le poids des lignes à chaque intersection, et est donc conservative dans le sens où la charge totale des particules reste constante. Nous montrons ici que, quand les paramètres de la dynamique s'écrivent sous une forme particulière, le processus PKS est réversible en espace. Il est alors possible de construire un processus PKS invariant par translation sur le plan entier dont les lois marginales selon les lignes horizontales et verticales sont des PPP (que nous qualifierons de *marqués*, comme ils régissent à la fois la position et le poids des lignes).

Dans ce chapitre, nous définissons le processus PKS dans un cadre général et prouvons son existence quand les paramètres sont uniformément bornés. Nous introduisons une notion de réversibilité pour les processus PKS qui signifie essentiellement que la loi d'un PKS restreinte à n'importe quelle boîte rectangulaire est invariante par rotation de 180 degrés. Ensuite nous présentons, dans nos résultats principaux, des conditions propices qui garantissent la réversibilité et donc la stationnarité des processus PKS. Nous le faisons dans trois cadres différents selon si la loi des poids des lignes est absolument continue par rapport à la mesure de Lebesgue, discrète ou arbitraire.

Dans le cadre de cette introduction, nous avons choisi de ne présenter les processus PKS que dans le cas où les lois en jeu sont absolument continues par rapport à la mesure de Lebesgue, et uniquement dans le cas où le processus est effectivement réversible. Nous commençons par en introduire les paramètres, avant de présenter la dynamique et d'énoncer les résultats de réversibilité.

### Paramètres du modèle

Soient  $\nu_V$  et  $\nu_H$  deux mesures positives sur  $\mathbb{R}$  admettant des densités  $g_V$  et  $g_H$  par rapport à la mesure de Lebesgue. Ces mesures sont appelées les *mesures d'intensité des lignes verticales et horizontales* du processus. Soit  $q : \mathbb{R} \rightarrow \mathbb{R}_+$  une fonction mesurable appelée *fonction de virage*. Soient également  $p_V : \mathbb{R} \rightarrow [0, 1]$  et  $p_H : \mathbb{R} \rightarrow [0, 1]$  deux fonctions mesurables appelées *fonctions de séparation verticale et horizontale*. On suppose que les fonctions  $q$ ,  $p_V$  et  $p_H$  satisfont les conditions suivantes : pour tout réel  $s$ , on a

$$g_V(s) = 0 \Rightarrow p_V(s) = q(s) = 0 \quad \text{et} \quad g_H(s) = 0 \Rightarrow p_H(s) = q(s) = 0, \quad (1.23)$$

et

$$p_V(s) + p_H(s) \leq 1. \quad (1.24)$$

On définit deux fonctions de taux de saut  $\lambda_V$  et  $\lambda_H$  comme suit : pour tout réel  $s$ ,

$$\lambda_V(s) := p_V(s) \frac{h(s)}{g_V(s)} \quad \text{et} \quad \lambda_H(s) := p_H(s) \frac{h(s)}{g_H(s)}, \quad (1.25)$$

où

$$h(s) := (g_V * g_H)(s) = \int_{\mathbb{R}} g_V(s-t) g_H(t) dt. \quad (1.26)$$

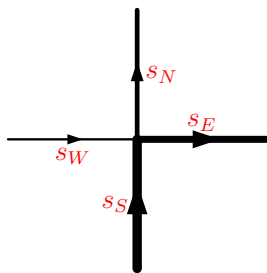


FIGURE 1.25 : Loi des nœuds de Kirchhoff à une intersection :  $s_S + s_W = s_N + s_E$ .

avec la convention  $0/0 = 0$  dans la formule ci-dessus <sup>6</sup>. On définit également un noyau de densité de probabilité  $f(s, \cdot)$  par, pour tous réels  $s, t$ ,

$$f(s, t) := \frac{g_V(s - t)g_H(t)}{h(s)}. \tag{1.27}$$

Remarquons que la fonction  $f$  peut être facilement interprétée de manière probabiliste. Si  $X_V$  et  $X_H$  sont deux variables aléatoires indépendantes à densité, de densités proportionnelles aux fonctions  $g_V$  et  $g_H$  respectivement, alors  $f(s, \cdot)$  est la densité de la variable  $X_H$  conditionnellement à l'événement  $\{X_V + X_H = s\}$ .

### Dynamique

Soit  $\mathcal{C}_X$  un PPP sur  $\mathbb{R}_+ \times \mathbb{R}$  d'intensité  $dx g_V(s)ds$  dont les atomes décrivent la loi initiale des particules verticales commençant sur l'axe des abscisses. Plus précisément, un atome  $(x, s)$  de  $\mathcal{C}_X$  correspond au point  $(x, 0)$  de l'axe des abscisses auquel on place initialement une particule de charge  $s$  (appelée également *intensité algébrique*) se déplaçant verticalement. De même, soit  $\mathcal{C}_Y$  un PPP sur  $\mathbb{R}_+ \times \mathbb{R}$ , indépendant de  $\mathcal{C}_X$  et d'intensité  $dy g_H(s)ds$ , qui, de manière similaire, est à interpréter comme un processus ponctuel marqué mais cette fois sur l'axe des ordonnées.

On définit maintenant une dynamique régissant un système de lignes pondérées verticales et horizontales qui vérifie la loi des nœuds de Kirchhoff à chaque intersection (par rapport à leurs intensités), comme décrit en Figure 1.25. Les lignes de notre processus commencent depuis les atomes du processus ponctuel des deux axes de la façon suivante :

- 1<sub>V</sub>.** pour chaque point  $(x, s)$  de  $\mathcal{C}_X$ , une ligne verticale commence au point  $(x, 0)$  et continue vers le haut avec une intensité  $s$ .
- 1<sub>H</sub>.** pour chaque point  $(y, s)$  de  $\mathcal{C}_Y$ , une ligne horizontale commence au point  $(0, y)$  et continue vers la droite avec une intensité  $s$ .

Il y a deux types d'événements qui peuvent survenir pendant la dynamique. Le premier, correspondant au cas **2** ci-dessous, consiste en des "séparations" et des "virages" de lignes :

<sup>6</sup>En fait, la convention  $0/0 = 0$  n'importe pas puisque, comme nous le verrons, nous n'appliquons la fonction de taux  $\lambda_V(s)$  (resp.  $\lambda_H(s)$ ) que pour  $s$  dans le support de  $\nu_V$  (resp.  $\nu_H$ ) de telle sorte que le quotient est toujours bien défini.



**2<sub>V</sub>**. Le long d'une ligne verticale d'intensité  $s$  :

- a. Une séparation arrive à taux  $\lambda_V(s) = p_V(s) \frac{h(s)}{g_V(s)}$ . Quand un tel événement survient, on tire une variable aléatoire  $T$  de densité  $f(s, \cdot)$  indépendante de tout le reste. À la suite de cette séparation, la ligne verticale continue vers le haut avec une nouvelle intensité  $s - T$ , et une ligne horizontale d'intensité  $T$  commence vers la droite à partir du point où la séparation a eu lieu.
- b. La ligne tourne à sa droite (i.e. vers l'est) à taux  $q(s) \sqrt{\frac{g_H(s)}{g_V(s)}}$  en gardant la même intensité, et devient de fait une ligne horizontale.

**2<sub>H</sub>**. Le long d'une ligne horizontale d'intensité  $s$  :

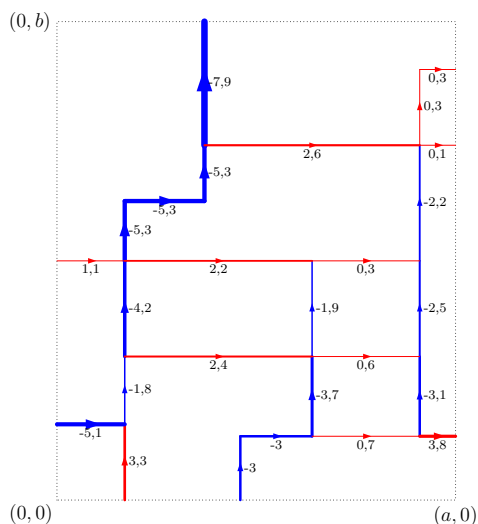
- a. Une séparation arrive à taux  $\lambda_H(s) = p_H(s) \frac{h(s)}{g_H(s)}$ . Quand un tel événement survient, on tire une variable aléatoire  $T$  de densité  $f(s, \cdot)$  indépendante de tout le reste. À la suite de cette séparation, la ligne horizontale continue vers la droite avec une nouvelle intensité  $T$ , et une ligne verticale d'intensité  $s - T$  commence vers le haut à partir du point où la séparation a eu lieu.
- b. La ligne tourne à sa gauche (i.e. vers le nord) à taux  $q(s) \sqrt{\frac{g_V(s)}{g_H(s)}}$  en gardant la même intensité, et devient de fait une ligne verticale.

Le second type d'événements à considérer correspond à l'intersection de lignes. Quand une ligne horizontale allant vers la droite (i.e. venant de l'ouest) d'intensité  $s_W$  rencontre une ligne verticale allant vers le haut (i.e. venant du sud) d'intensité  $s_S$ , on applique les règles suivantes :

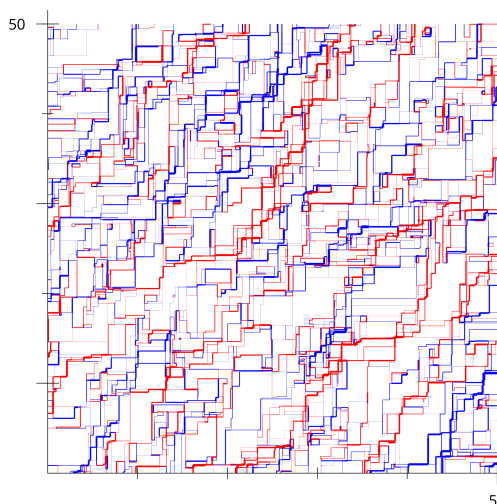
- 3.**
  - a. avec probabilité  $p_V(s_S + s_W)$ , la ligne horizontale s'arrête et la ligne verticale continue avec une nouvelle intensité  $s_N := s_S + s_W$  ;
  - b. avec probabilité  $p_H(s_S + s_W)$ , la ligne verticale s'arrête et la ligne horizontale continue avec une nouvelle intensité  $s_E := s_S + s_W$  ;
  - c. sur l'événement complémentaire, c'est-à-dire avec probabilité  $1 - p_V(s_S + s_W) - p_H(s_S + s_W)$ , on tire une variable aléatoire  $T$  de densité  $f(s_S + s_W, \cdot)$  indépendante de tout le reste. À la suite de ce croisement, l'intensité de la ligne horizontale devient  $s_E := T$  et celle de la ligne verticale devient  $s_N := s_S + s_W - T$ .

Les règles **1**, **2** et **3** ainsi que les processus définissant les points et intensités de départ  $\mathcal{C}_X$  et  $\mathcal{C}_Y$  définissent de manière unique un système aléatoire de lignes algébriquement pondérées sur tout le quart de plan  $\mathbb{R}_+^2$ , que l'on appelle **Système de Poisson-Kirchhoff** (PKS pour *Poisson-Kirchhoff System*) de paramètres  $(g_V, g_H, p_V, p_H, q)$ . Une illustration d'un tel système est donnée en Figure 1.26. Remarquons que, grâce à la dynamique ci-dessus, la système est conservatif : il vérifie la loi de Kirchhoff (comme en Figure 1.25) à chaque intersection.

Comme évoqué plus haut, un PKS peut être défini dans un cadre plus général que celui-ci. Citons par exemple le cas où les mesures  $\nu_V$  et  $\nu_H$  sont discrètes : l'essentiel de la dynamique décrite ci-avant s'applique de la même manière, à la différence que lorsque deux lignes se rencontrent (cas **3**), celles-ci peuvent avoir des intensités opposées. On autorise donc qu'avec probabilité non nulle, ces deux lignes s'arrêtent en ce point de croisement. Pour qu'une réversibilité puisse être établie, on introduit également un PPP sur le quart de plan tel que de chacun



(a) Réalisation typique d'un PKS.



(b) Simulation d'un PKS sur  $[0, 50] \times [0, 50]$  avec les paramètres  $\nu_V = \nu_H = \mathcal{N}(0, 1)$ ,  $p_V(s) = 0.4$ ,  $p_H(s) = 0.4$ ,  $q(s) = 0.1$ .

**FIGURE 1.26 :** Exemple de dynamique. Les lignes d'intensité positive sont en rouge et celles d'intensité négative en bleu. L'épaisseur d'une ligne est proportionnelle à la valeur absolue de son intensité.

de ses atomes partent une ligne verticale et une ligne horizontale d'intensités opposées. Ces deux événements sont exactement ceux qui interviennent dans le cas du processus de Hammersley présenté en Figure 1.24 et détaillé dans les sections 1.1.1 et 1.2.1.

### Réversibilité

Même si le processus PKS est défini dans tout le quart de plan, il est pratique de considérer et d'étudier sa restriction à une boîte de la forme  $[0, a] \times [0, b]$  pour  $a, b \in \mathbb{R}_+^*$ . On note  $\mathcal{D}_{a,b}$  l'espace image de notre processus restreint à cette boîte, qui consiste en une collection finie de *segments pondérés verticaux et horizontaux* dans le rectangle et qui de plus vérifient la loi de Kirchhoff à chaque intersection. On appelle les éléments de l'ensemble  $\mathcal{D}_{a,b}$  des *dessins*, et on note  $\mathbf{D}$  un dessin aléatoire qui a la loi du processus PKS défini ci-avant.

Étant donné un dessin  $D \in \mathcal{D}_{a,b}$ , on définit le dessin retourné de  $D$ , noté  $\widehat{D} \in \mathcal{D}_{a,b}$ , obtenu en appliquant à  $D$  une rotation de 180 degrés de centre  $(a/2, b/2)$ . Remarquons que l'image obtenue est encore un dessin valide. Un exemple de dessin  $D$  et de son retourné  $\widehat{D}$  est donné en Figure 1.27. Notre résultat principal établit que le processus PKS est réversible en "espace-temps" dans le sens suivant :

**Théorème 1.17 (Réversibilité).** *Les variables aléatoires  $\mathbf{D}$  et  $\widehat{\mathbf{D}}$  ont la même loi.*

Une conséquence importante de ce théorème est la stationnarité de la dynamique :

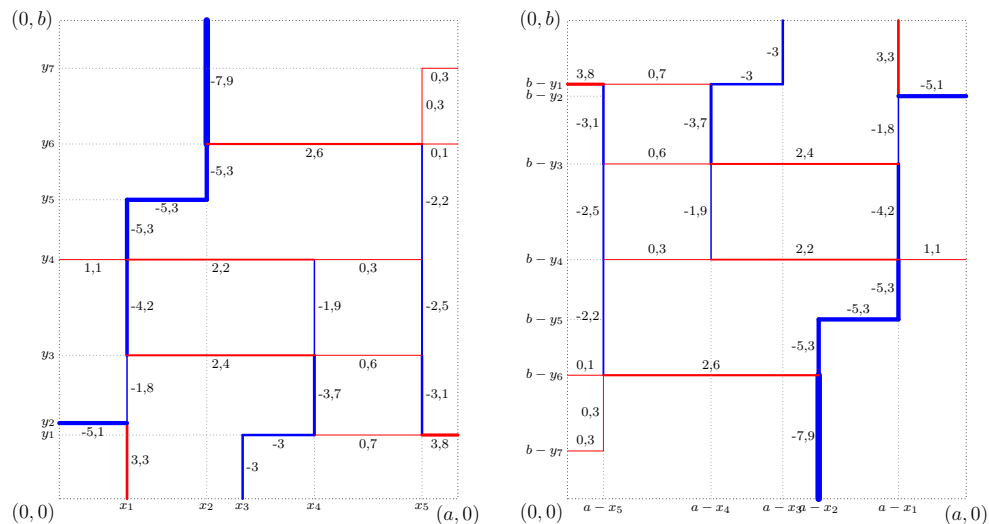


FIGURE 1.27 : Un exemple de dessin (à gauche) et de son retourné (à droite)

**Corollaire 1.3** (Stationnarité verticale/horizontale). *La restriction du dessin  $\mathbf{D}$  au segment  $[0, a] \times \{b\}$  est un  $\text{Leb} \otimes \nu_V$ -PPP. Cela signifie que l'ensemble des points de sortie de la boîte des segments verticaux a la même loi que l'ensemble de leurs points d'entrée  $\mathcal{C}_X$ . De même, la restriction du dessin  $\mathbf{D}$  au segment  $\{a\} \times [0, b]$  est un  $\text{Leb} \otimes \nu_H$ -PPP. Cela signifie que l'ensemble des points de sortie de la boîte des segments horizontaux a la même loi que l'ensemble de leurs points d'entrée  $\mathcal{C}_Y$ . De plus, ces deux PPP sont indépendants.*

Pour prouver le Théorème 1.17, nous introduisons une famille de paramétrisations du système, étant donné que l'espace d'état des processus PKS est assez compliqué en tant que tel. Il s'avère que des paramétrisations différentes de cette famille définissent toujours la même forme. Nous appliquons ce résultat à deux paramétrisations spécifiques : la première associée à la dynamique du PKS et la seconde associée à sa dynamique retournée. Une analyse soignée montre ensuite que les densités associées à la dynamique et à celle retournée dans leurs paramétrisations respectives coïncident.

Dans la suite, nous montrons d'abord comment la loi de Kirchhoff permet de définir une notion de fonction potentiel associée aux faces de la tessellation définie par un PKS. Ce potentiel correspond aux temps de derniers passages en termes de PDP. Nous recueillons ensuite plusieurs modèles de PDP qui peuvent être associés à des processus PKS, notamment la plupart de ceux présentés en section 1.2. Ensuite, nous fournissons une liste (non-exhaustive) de processus PKS obtenus pour des lois des poids des lignes spécifiques. À partir de cette liste, nous retrouvons plusieurs autres modèles classiques de physique statistique comme les modèles balistiques [KRL95, BM20, HST21] ou les modèles à six sommets [Pau35, Bax72]. En particulier, nous exhibons des nouvelles mesures invariantes explicites pour des problèmes balistiques de même que de nouvelles propriétés de réversibilité pour des modèles à six sommets avec

un champ électromagnétique externe. Par ailleurs, les cas particuliers où les poids des lignes suivent une loi normale ou de Poisson fournissent des nouveaux modèles avec des dynamiques explicites qui pourraient mériter une étude plus approfondie.

Enfin, nous nous intéressons aux propriétés géométriques de base de la tessellation aléatoire du quart de plan induite par un PKS, telles que le nombre moyen de composantes connexes à l'intérieur d'un rectangle, et le nombre moyen de nœuds d'une composante connexe typique.



## Chapitre 2

# Gaussian fluctuations of stationary Hammersley's processes: a probabilistic proof

### Abstract

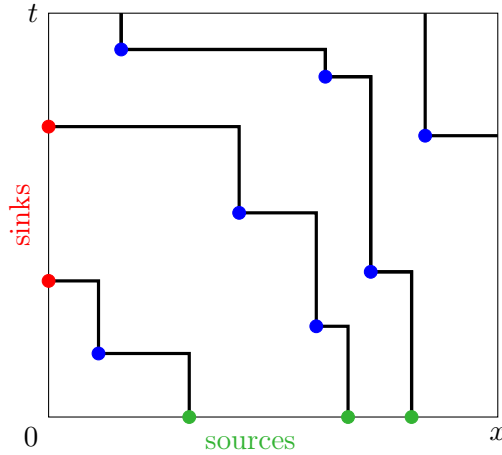
In [CG05, CG06], Cator and Groeneboom introduce a family of stationary versions of Hammersley's processes in the quarter plane which can be seen as last passage percolation models (LPP) with sources and sinks and they prove the famous cube root asymptotics for the fluctuation of the passage time in the critical regime. Here, we study instead the non-critical case and prove by analogous “soft” probabilistic arguments the Gaussian fluctuations for the length of the longest path in this regime. This recovers results obtained by Baik and Rains [BR00] using “hard” analytic arguments.

## 2.1 Introduction and main result

In this section we recall the definition of the stationary version of Hammersley's process in the quarter plane introduced and studied by Groeneboom [Gro02] and Cator and Groeneboom [CG05], and then state our main result.

Consider three independent Poisson point processes (PPPs): one in the quarter plane  $\{(x, t) \in (\mathbb{R}_+)^2\}$  with unit intensity, called  $\alpha$ -point process, one on the positive  $x$ -axis with intensity  $\lambda > 0$ , called *source* process, and one on the positive  $t$ -axis with intensity  $\lambda^{-1}$ , called *sink* process. We call an *increasing path* between two points a path made of points of the PPPs which only makes non-negative steps in both coordinates, and its *length* the number of points of the PPPs belonging to it. We call a *geodesic* between two points an increasing path whose length is maximal. Finally we denote by  $L_\lambda(x, t)$  the length of a geodesic between the points  $(0, 0)$  and  $(x, t)$  in the model with the three PPPs, and by  $L(x, t)$  the same length in the model where we do not consider sources nor sinks.

To compute these lengths, one can use Hammersley's lines as defined in [Ham72] which here can be seen as space-time paths of particles as it is done by Aldous and Diaconis [AD95]. The particles start from sources and evolves vertically. When a sink or an  $\alpha$ -point appears, the



**Figure 2.1:** An example of construction of Hammersley's lines with sources and sinks. The  $\alpha$ -points are drawn in blue.

leftmost particle to the right of it jumps to its location. If it is a sink, the particle also dies, and its space-time path ends. An example of construction of Hammersley's lines with sources and sinks is given in Figure 2.1.

It is known that for a given point  $(x, t) \in (\mathbb{R}_+^*)^2$ , the length of a geodesic between the points  $(0, 0)$  and  $(x, t)$  is equal to the number of different Hammersley's lines living in the box  $[0, x] \times [0, t]$ . Indeed, on the one hand two points of a given increasing path necessarily belong to distinct Hammersley's lines, and on the other hand one can construct in an effective way an increasing path made of one point per Hammersley's line. Denote by  $N$  (resp.  $S$ ,  $E$ ,  $W$ ) the number of space-time paths crossing the north (resp. south, east, west) side of the box  $[0, x] \times [0, t]$  in the model with sources and sinks with intensity  $\lambda > 0$ . Then, by noticing that each Hammersley's line leaves the box either by the north side or by the west side, and enter the box either by the south side or the east side, we can write the following decomposition:

$$L_\lambda(x, t) = N + W = S + E.$$

The main benefit of introducing sources and sinks is that they make the system **stationary** (indeed reversible) in the following sense [Gro02, CG05]:

$$N \stackrel{(d)}{=} S, \quad E \stackrel{(d)}{=} W \quad \text{and} \quad N \text{ and } E \text{ are independent,}$$

where the symbol  $\stackrel{(d)}{=}$  means "equal in distribution". Recall that by definition, the variables  $S$  and  $W$  follow Poisson distributions with parameters  $\lambda x$  and  $\lambda^{-1}t$  respectively. This property yields the first order asymptotics for the process  $(L_\lambda(t, t))_{t>0}$ :

$$\frac{L_\lambda(t, t)}{t} \xrightarrow[t \rightarrow \infty]{a.s.} \lambda + \frac{1}{\lambda}.$$

The almost sure convergence is indeed guaranteed by Kingman's subadditive ergodic theorem [Kin68].

Baik and Rains [BR00] studied the second order asymptotics using “hard” analysis arguments, and proved that the process  $(L_\lambda(t, t))_{t>0}$  has two very different behaviors depending on whether  $\lambda = 1$ , the case we call *critical*, or not:

(i) If  $\lambda = 1$ ,

$$\frac{L_1(t, t) - 2t}{t^{1/3}} \xrightarrow[t \rightarrow \infty]{(d)} F_0,$$

where the distribution  $F_0$  is related to Tracy-Widom [TW94] distribution. We do not give its exact definition here as it is quite cumbersome, see [BR00] for more details.

(ii) If  $\lambda \neq 1$ ,

$$\frac{L_\lambda(t, t) - (\lambda + \frac{1}{\lambda})t}{\sqrt{t}} \xrightarrow[t \rightarrow \infty]{(d)} \mathcal{N}\left(0, \left|\lambda - \frac{1}{\lambda}\right|\right).$$

Cator and Groeneboom [CG06] studied the second order asymptotics in the critical case using “soft” probabilistic and geometric arguments. [CG06] provides a proof of a weaker version of the above-mentioned point (i): the order of the asymptotics is obtained but not the limiting distribution. Specifically, it is proved that

$$\text{Var}(L_1(t, t)) \underset{t \rightarrow \infty}{\asymp} t^{2/3}.$$

In the previous paper [CG05], Cator and Groenboom proved that always when  $\lambda = 1$ , a geodesic between the points  $(0, 0)$  and  $(t, t)$  spends a vanishing proportion of time on the axes (which is proved to be of order  $t^{1/3}$  in [CG06]). In a certain sense, this suggests that the process  $(L_\lambda(t, t))_{t>0}$  has the same second order asymptotics as the one without sources and sinks  $(L(t, t))_{t>0}$ , which has been proved to be of order  $t^{1/3}$  by Baik-Deift-Johannson [BDJ99] (but has a different limiting distribution).

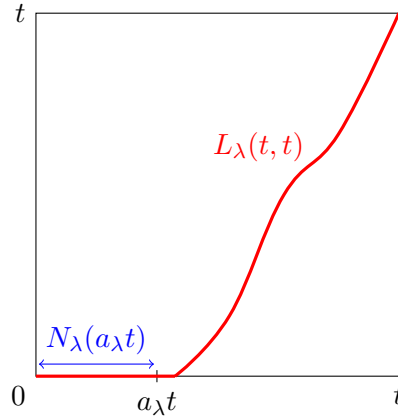
In this chapter, we study the non-critical case  $\lambda \neq 1$ . Without loss of generality, we assume  $\lambda > 1$  by a symmetry argument. Our main result relies on the fact that in this model, the geodesics between the points  $(0, 0)$  and  $(t, t)$  spends a non-vanishing proportion of time on the axes. The point where they leave the axes, which we call the *takeoff point*, is indeed concentrated around the point  $(0, a_\lambda t)$  for  $a_\lambda = 1 - \frac{1}{\lambda^2}$ .

Heuristically, in this model a geodesic can be seen as being split into two independent parts: the first one between the points  $(0, 0)$  and  $(a_\lambda t, 0)$  and the second between the points  $(a_\lambda t, 0)$  and  $(t, t)$ . The fluctuations of the first part are Gaussian and of order  $t^{1/2}$  since it is a Poisson process, those of the second one are of order  $t^{1/3}$  since it behaves like in the model  $L$  without sources nor sinks and are therefore negligible compared to the ones of the first part. More precisely, if for a positive  $z$  we denote by  $N_\lambda(z) = L_\lambda(z, 0)$  the number of sources on the segment  $[0, z] \times \{0\}$ , which is a Poisson random variable with mean  $\lambda z$ , we have:

**Theorem 2.1.** *For any  $\lambda > 1$ , the following convergence in probability holds:*

$$\frac{L_\lambda(t, t) - (\lambda + \frac{1}{\lambda})t}{\sqrt{t}} - \frac{N_\lambda((1 - \frac{1}{\lambda^2})t) - (\lambda - \frac{1}{\lambda})t}{\sqrt{t}} \xrightarrow[t \rightarrow \infty]{\mathbb{P}} 0.$$





**Figure 2.2:** A schematic representation of a typical longest path when  $\lambda > 1$ , with  $a_\lambda = 1 - \lambda^{-2}$ . Its total length  $L_\lambda(t, t)$  has the same fluctuations as the number of sources on the segment  $[0, a_\lambda t] \times \{0\}$ ,  $N_\lambda(a_\lambda t)$ .

In other words, we prove that the fluctuations of the variable  $L_\lambda(t, t)$  **are only driven by the horizontal part of the geodesic**, on the  $x$ -axis. This phenomenon is illustrated in Figure 2.2. This result makes it possible to recover the convergence (ii) as a corollary:

**Corollary 2.1.** *For any  $\lambda \neq 1$ ,*

$$\frac{L_\lambda(t, t) - (\lambda + \frac{1}{\lambda})t}{\sqrt{t}} \xrightarrow[t \rightarrow \infty]{(d)} \mathcal{N}\left(0, \left|\lambda - \frac{1}{\lambda}\right|\right).$$

We start by proving that the geodesics spend a lower-bounded amount of time on the  $x$ -axis using concentration arguments, this is the object of section 2.2. We then deduce Theorem 2.1 and Corollary 2.1 in section 2.3, by writing our quantity of interest as a suitable sum of several variables for which the convergence is easier to study.

## 2.2 Geodesics do not leave the $x$ -axis too early

In this section we prove that the geodesics take a lower-bounded amount of sources with high probability in the following sense:

**Lemma 2.1.** *Let  $\lambda > 1$  and set  $a_\lambda = 1 - \frac{1}{\lambda^2}$ . For any reals  $t > 0$  and  $\eta \in (0, a_\lambda)$ , denote by  $E_{\eta, t}$  the following event: “in the stationary model, all optimal paths from  $(0, 0)$  to  $(t, t)$  take all the sources on the interval  $(0, (a_\lambda - \eta)t] \times \{0\}$ . Then there exists a constant  $c$  which only depends on  $\lambda$ , such that*

$$\forall \eta \in (0, c), \forall t > 0, \quad \mathbb{P}(\overline{E_{\eta, t}}) \leq \frac{1}{c\eta^2} \exp(-c\eta^4 t).$$

The proof is split into three parts. We first give concentration inequalities for Poisson random variables in Lemma 2.2 below. We then show that a geodesic compelled to leave the

$x$ -axis “too early” cannot have a length of order  $(\lambda + \lambda^{-1})t$ , which is the purpose of Lemma 2.3. We eventually prove Lemma 2.1.

As announced, let us start by giving concentration inequalities for Poisson random variables:

**Lemma 2.2.** *With local notation, the following Chernoff bounds for Poisson distribution hold.*

$$\text{If } 0 \leq t \leq \lambda, \quad \begin{cases} \mathbb{P}(\text{Poi}(\lambda) \geq \lambda + t) \leq \exp\left(-\frac{t^2}{4\lambda}\right), \\ \mathbb{P}(\text{Poi}(\lambda) \leq \lambda - t) \leq \exp\left(-\frac{t^2}{4\lambda}\right). \end{cases}$$

*Proof.* Define the function  $h$  on  $[-1, +\infty)$  by  $h : x \mapsto (1+x)\log(1+x) - x$ . [BLM13, section 2.2] provides the following bounds for any  $t \in [0, \lambda]$ :

$$\begin{cases} \mathbb{P}(\text{Poi}(\lambda) \geq \lambda + t) \leq \exp\left(-\lambda h\left(\frac{t}{\lambda}\right)\right), \\ \mathbb{P}(\text{Poi}(\lambda) \leq \lambda - t) \leq \exp\left(-\lambda h\left(-\frac{t}{\lambda}\right)\right). \end{cases}$$

The wanted inequalities arise from the fact that

$$\forall x \in [-1, 1], \quad h(x) \geq \frac{x^2}{4}. \quad \square$$

We use this lemma to prove that a geodesic compelled to leave the  $x$ -axis “too early” cannot have a length of order  $(\lambda + \lambda^{-1})t$ . To state this result properly, we introduce some notation.

For a sub-interval  $(a, b] \subset [0, 1]$ , denote by  $L_\lambda^{\uparrow(a,b]}(t, t)$  (resp.  $L_\lambda^{\rightarrow(a,b]}(t, t)$ ) the length of a geodesic from  $(0, 0)$  to  $(t, t)$  compelled to leave the  $y$ -axis (resp. the  $x$ -axis) at a point located on the half-opened vertical segment  $\{0\} \times (at, bt]$  (resp. horizontal segment  $(at, bt] \times \{0\}$ ) in the stationary model with source intensity  $\lambda > 0$ . Then,

**Lemma 2.3.** *Let  $t > 0$ ,  $\lambda > 1$  and set  $a_\lambda = 1 - \frac{1}{\lambda^2}$ . The following inequalities hold.*

(i)  $\exists \delta > 0, C > 0$ ,

$$\mathbb{P}\left(L_\lambda^{\uparrow(0,1]}(t, t) \geq \left(\lambda + \frac{1}{\lambda} - \delta\right)t\right) \leq C \exp\left(-\frac{t}{C}\right).$$

(ii)  $\forall \eta \in (0, a_\lambda), \exists \delta > 0, C > 0$ ,

$$\mathbb{P}\left(L_\lambda^{\rightarrow(0, a_\lambda - \eta]}(t, t) \geq \left(\lambda + \frac{1}{\lambda} - \delta\right)t\right) \leq C \exp\left(-\frac{t}{C}\right).$$

In words, if a path is compelled to take some sinks or to take significantly less than the segment  $[0, a_\lambda t]$  of sources, then its length is significantly less than  $\lambda + \lambda^{-1}$  with high probability.

*Proof.* Both points are proved in a similar way, let us start with the first one.

- (i) We start by proving the wanted result for the variable  $L_\lambda^{\uparrow(\alpha, \alpha + \varepsilon]}(t, t)$  for any positive  $\alpha$  and for a well-chosen positive  $\varepsilon$ . The result will follow for the variable  $L_\lambda^{\uparrow(0, 1]}(t, t)$  by considering a regular subdivision of the interval  $(0, 1]$ .

Formally, let  $\delta := \frac{1}{4}(\lambda + \frac{1}{\lambda} - 2) > 0$ , and  $\varepsilon := 2\lambda\delta = \frac{1}{2}(\lambda - 1)^2$ . These expressions have been chosen so that several inequalities below actually hold.

Let  $\alpha \in [0, (1 - \varepsilon)_+]$ . The variable  $L_\lambda^{\uparrow(\alpha, \alpha + \varepsilon]}(t, t)$  is stochastically dominated by a sum of two random variables which have the marginal distribution  $\text{Poi}((\alpha + \varepsilon)\frac{t}{\lambda})$  and  $L(t, (1 - \alpha)t)$ , with mean  $(\alpha + \varepsilon)\frac{t}{\lambda}$  and  $2t\sqrt{1 - \alpha}$  respectively. The parameter  $\delta$  has been defined in this way since it makes it possible to lower bound the quantity  $\lambda + \frac{1}{\lambda} - \delta$  by a sum of four terms which are greater than the two expectations of these variables. Indeed, using  $\lambda > 1$ ,

$$\begin{aligned} \lambda + \frac{1}{\lambda} - \delta &= 2 + \frac{\varepsilon}{\lambda} + \delta \geq 2 - \alpha \left(1 - \frac{1}{\lambda}\right) + \frac{\varepsilon}{\lambda} + \delta = \frac{\alpha + \varepsilon}{\lambda} + \frac{\delta}{2} + 2\left(1 - \frac{\alpha}{2}\right) + \frac{\delta}{2} \\ &\geq \frac{\alpha + \varepsilon}{\lambda} + \frac{\delta}{2} + 2\sqrt{1 - \alpha} + \frac{\delta}{2} \quad \text{since } \sqrt{1 - x} \leq 1 - \frac{x}{2}. \end{aligned}$$

Consequently,

$$\begin{aligned} \mathbb{P}\left(L_\lambda^{\uparrow(\alpha, \alpha + \varepsilon]}(t, t) \geq \left(\lambda + \frac{1}{\lambda} - \delta\right)t\right) &\leq \mathbb{P}\left(L_\lambda^{\uparrow(\alpha, \alpha + \varepsilon]}(t, t) \geq \left(\frac{\alpha + \varepsilon}{\lambda} + \frac{\delta}{2} + 2\sqrt{1 - \alpha} + \frac{\delta}{2}\right)t\right) \\ &\leq \mathbb{P}\left(\text{Poi}\left((\alpha + \varepsilon)\frac{t}{\lambda}\right) \geq (\alpha + \varepsilon)\frac{t}{\lambda} + \frac{\delta t}{2}\right) + \mathbb{P}\left(L(t, (1 - \alpha)t) \geq 2t\sqrt{1 - \alpha} + \frac{\delta t}{2}\right). \end{aligned}$$

Let us deal with the two terms of the right-hand side separately. For the first one, we use Lemma 2.2 after checking that for any positive  $t$ ,  $0 \leq \frac{\delta t}{2} \leq (\alpha + \varepsilon)\frac{t}{\lambda}$  since  $\varepsilon = 2\lambda\delta$ :

$$\mathbb{P}\left(\text{Poi}\left((\alpha + \varepsilon)\frac{t}{\lambda}\right) \geq (\alpha + \varepsilon)\frac{t}{\lambda} + \frac{\delta t}{2}\right) \leq \exp\left(-\frac{\lambda\delta^2 t}{16(\alpha + \varepsilon)}\right) \leq \exp\left(-\frac{\delta^2 t}{16}\right),$$

since  $\lambda \geq 1$  and  $\alpha + \varepsilon \leq 1$ . For the second one, we recall the following scaling argument first proved by Aldous and Diaconis [AD95]: for any positive  $x, t$ ,

$$L(x, t) \stackrel{(d)}{=} L(\sqrt{xt}, \sqrt{xt}). \quad (2.1)$$

Moreover, the variable  $L(\sqrt{xt}, \sqrt{xt})$  is stochastically dominated by the variable  $L_1(\sqrt{xt}, \sqrt{xt})$ . As a result,

$$\begin{aligned} \mathbb{P}\left(L(t, (1 - \alpha)t) \geq 2t\sqrt{1 - \alpha} + \frac{\delta t}{2}\right) &\leq \mathbb{P}\left(L_1(t, (1 - \alpha)t) \geq 2t\sqrt{1 - \alpha} + \frac{\delta t}{2}\right) \\ &\leq 2\mathbb{P}\left(\text{Poi}(t\sqrt{1 - \alpha}) \geq t\sqrt{1 - \alpha} + \frac{\delta t}{4}\right) \\ &\leq 2\exp\left(-\frac{\delta^2 t}{64\sqrt{1 - \alpha}}\right) \leq 2\exp\left(-\frac{\delta^2 t}{64}\right), \end{aligned}$$

where we used Lemma 2.2 again, after checking that  $\frac{\delta t}{4} \leq t\sqrt{1 - \alpha}$ , since  $\sqrt{1 - \alpha} \geq 1 - \alpha \geq \varepsilon$ . Finally,

$$\mathbb{P}\left(L_\lambda^{\uparrow(\alpha, \alpha + \varepsilon]}(t, t) \geq \left(\lambda + \frac{1}{\lambda} - \delta\right)t\right) \leq 3\exp\left(-\frac{\delta^2 t}{64}\right).$$

As mentioned above, we use this concentration inequality for the variable  $L_\lambda^{\uparrow(\alpha, \alpha + \varepsilon]}(t, t)$  to deduce the wanted one for the variable  $L_\lambda^{\uparrow(0, 1]}(t, t)$  by using a subdivision of the interval  $(0, 1]$  in subintervals of the form  $(\alpha, \alpha + \varepsilon]$ .

Set  $K := \lceil \varepsilon^{-1} \rceil$  and for all integer  $k \in \llbracket 1, K \rrbracket$ ,  $\alpha_k := (1 - k\varepsilon) \vee 0$ , such that  $0 = \alpha_K < \alpha_{K-1} < \dots < \alpha_1 < 1$  defines a finite subdivision of the interval  $(0, 1]$  in subintervals of width at most  $\varepsilon$ . Consequently,

$$\begin{aligned} \mathbb{P}\left(L_\lambda^{\uparrow(0, 1]}(t, t) \geq \left(\lambda + \frac{1}{\lambda} - \delta\right)t\right) &\leq \sum_{k=1}^K \mathbb{P}\left(L_\lambda^{\uparrow(\alpha_k, \alpha_k + \varepsilon]}(t, t) \geq \left(\lambda + \frac{1}{\lambda} - \delta\right)t\right) \\ &\leq 3K \exp\left(-\frac{\delta^2 t}{64}\right) \leq 3\left(\frac{1}{2\lambda\delta} + 1\right) \exp\left(-\frac{\delta^2 t}{64}\right), \end{aligned}$$

since  $K \leq \frac{1}{\varepsilon} + 1 = \frac{1}{2\lambda\delta} + 1$ . The wanted inequality follows by taking  $C = (3(\frac{1}{2\lambda\delta} + 1)) \vee \frac{64}{\delta^2}$ .

(ii) The second inequality is proved in a very similar way.

Set  $f : [0, 1] \rightarrow \mathbb{R}_+$ ,  $\alpha \mapsto \lambda\alpha + 2\sqrt{1 - \alpha}$ . The function  $f$  admits a unique strict maximum point  $a_\lambda = 1 - \frac{1}{\lambda^2}$ , with  $f(a_\lambda) = \lambda + \frac{1}{\lambda}$ . Set  $\delta := \frac{4}{\lambda} \wedge \frac{1}{4}(f(a_\lambda) - f(a_\lambda - \eta)) > 0$ , and  $\varepsilon := \frac{2\delta}{\lambda}$ . Let  $\alpha \in [0, (a_\lambda - \eta - \varepsilon)_+]$ , such that  $0 \leq \alpha \leq \alpha + \varepsilon \leq a_\lambda - \eta$ .

The variable  $L_\lambda^{\rightarrow(\alpha, \alpha + \varepsilon]}(t, t)$  is stochastically dominated by a sum of two random variables which have the marginal distribution  $\text{Poi}((\alpha + \varepsilon)\lambda t)$  and  $L((1 - \alpha)t, t)$ , with mean  $(\alpha + \varepsilon)\lambda t$  and  $2t\sqrt{1 - \alpha}$  respectively. Moreover, since  $\delta \leq \frac{1}{4}(f(a_\lambda) - f(a_\lambda - \eta))$  and since the function  $f$  is increasing on the interval  $[0, a_\lambda]$ ,

$$\lambda + \frac{1}{\lambda} - \delta \geq f(a_\lambda - \eta) + \varepsilon\lambda + \delta \geq f(\alpha) + \varepsilon\lambda + \delta = (\alpha + \varepsilon)\lambda + \frac{\delta}{2} + 2\sqrt{1 - \alpha} + \frac{\delta}{2}.$$

Consequently, as above,

$$\begin{aligned} \mathbb{P}\left(L_\lambda^{\rightarrow(\alpha, \alpha + \varepsilon]}(t, t) \geq \left(\lambda + \frac{1}{\lambda} - \delta\right)t\right) &\leq \mathbb{P}\left(L_\lambda^{\rightarrow(\alpha, \alpha + \varepsilon]}(t, t) \geq \left((\alpha + \varepsilon)\lambda + \frac{\delta}{2} + 2\sqrt{1 - \alpha} + \frac{\delta}{2}\right)t\right) \\ &\leq \mathbb{P}\left(\text{Poi}((\alpha + \varepsilon)\lambda t) \geq (\alpha + \varepsilon)\lambda t + \frac{\delta t}{2}\right) + \mathbb{P}\left(L((1 - \alpha)t, t) \geq 2t\sqrt{1 - \alpha} + \frac{\delta t}{2}\right). \end{aligned}$$

For the first term, we use Lemma 2.2 after checking that for any positive  $t$ ,  $0 \leq \frac{\delta t}{2} \leq (\alpha + \varepsilon)\lambda t$  since  $\varepsilon = \frac{2\delta}{\lambda}$ :

$$\mathbb{P}\left(\text{Poi}((\alpha + \varepsilon)\lambda t) \geq (\alpha + \varepsilon)\lambda t + \frac{\delta t}{2}\right) \leq \exp\left(-\frac{\delta^2 t}{16\lambda(\alpha + \varepsilon)}\right) \leq \exp\left(-\frac{\delta^2 t}{16\lambda}\right),$$

since  $\alpha + \varepsilon \leq 1$ . For the second one, by reasoning exactly as above,

$$\mathbb{P}\left(L(t, (1 - \alpha)t) \geq 2t\sqrt{1 - \alpha} + \frac{\delta t}{2}\right) \leq 2 \exp\left(-\frac{\delta^2 t}{64}\right),$$

where we have checked the hypothesis of Lemma 2.2, that is to say  $\frac{\delta t}{4} \leq t\sqrt{1-\alpha}$ . This is the case since  $\sqrt{1-\alpha} \geq \sqrt{1-a_\lambda} = \frac{1}{\lambda} \geq \frac{\delta}{4}$  by definition of  $\delta$ . Finally,

$$\mathbb{P}\left(L_\lambda^{\rightarrow(\alpha, \alpha+\varepsilon]}(t, t) \geq \left(\lambda + \frac{1}{\lambda} - \delta\right)t\right) \leq 3 \exp\left(-\frac{\delta^2 t}{64\lambda}\right).$$

By subdividing the interval  $(0, a_\lambda - \eta]$  as above, we eventually get

$$\mathbb{P}\left(L_\lambda^{\uparrow(0, a_\lambda - \eta]}(t, t) \geq \left(\lambda + \frac{1}{\lambda} - \delta\right)t\right) \leq 3\left(\frac{\lambda}{2\delta} + 1\right) \exp\left(-\frac{\delta^2 t}{64\lambda}\right),$$

and the wanted inequality follows.  $\square$

We can then use Lemma 2.3 to deduce Lemma 2.1.

*Proof of Lemma 2.1.* The goal is to prove that a path compelled to leave the  $x$ -axis too early cannot be an (unconstrained) optimal path. Lemma 2.3 tells us that the length of such a path is significantly less than  $\lambda + \lambda^{-1}$  with high probability, Lemma 2.1 will follow by using the fact that the length of the (unconstrained) optimal path is close to  $\lambda + \lambda^{-1}$  with high probability.

For a fixed  $\eta > 0$ , take  $\delta > 0$  as the minimum of the two parameters  $\delta$  of the previous Lemma, that is to say

$$\delta := \delta(\eta) = \min\left(\frac{f(a_\lambda) - f(a_\lambda - \eta)}{4}, \frac{\lambda + \frac{1}{\lambda} - 2}{4}, \frac{4}{\lambda}\right).$$

Then, decompose the probability into two parts, whether the length of the geodesic is greater than  $(\lambda + \frac{1}{\lambda} - \delta)t$  or not:

$$\mathbb{P}(\overline{E_{\eta, t}}) \leq \mathbb{P}\left(\overline{E_{\eta, t}}, L_\lambda(t, t) \geq \left(\lambda + \frac{1}{\lambda} - \delta\right)t\right) + \mathbb{P}\left(L_\lambda(t, t) < \left(\lambda + \frac{1}{\lambda} - \delta\right)t\right).$$

The second term can be bounded using Lemma 2.2:

$$\begin{aligned} \mathbb{P}\left(L_\lambda(t, t) < \left(\lambda + \frac{1}{\lambda} - \delta\right)t\right) &\leq \mathbb{P}\left(\text{Poi}(\lambda t) < \lambda t - \frac{\delta t}{2}\right) + \mathbb{P}\left(\text{Poi}\left(\frac{t}{\lambda}\right) < \frac{t}{\lambda} - \frac{\delta t}{2}\right) \\ &\leq \exp\left(-\frac{\delta^2 t}{16\lambda}\right) + \exp\left(-\frac{\delta^2 \lambda t}{16}\right) \\ &\leq 2 \exp\left(-\frac{\delta^2 t}{16\lambda}\right) \quad \text{since } \lambda > 1 \end{aligned}$$

The first term can be bounded using the previous Lemma:

$$\begin{aligned} \mathbb{P}\left(\overline{E_{\eta, t}}, L_\lambda(t, t) \geq \left(\lambda + \frac{1}{\lambda} - \delta\right)t\right) &\leq \mathbb{P}\left(L_\lambda^{\uparrow(0, 1]}(t, t) \geq \left(\lambda + \frac{1}{\lambda} - \delta\right)t\right) + \mathbb{P}\left(L_\lambda^{\rightarrow(0, a_\lambda - \eta]}(t, t) \geq \left(\lambda + \frac{1}{\lambda} - \delta\right)t\right) \\ &\leq 3\left(\frac{1}{2\lambda\delta} + 1\right) \exp\left(-\frac{\delta^2 t}{64}\right) + 3\left(\frac{\lambda}{2\delta} + 1\right) \exp\left(-\frac{\delta^2 t}{64\lambda}\right) \\ &\leq \frac{6\lambda}{\delta} \exp\left(-\frac{\delta^2 t}{64\lambda}\right) \quad \text{since } \delta < \frac{\lambda}{2}. \end{aligned}$$

Finally,

$$\mathbb{P}(\overline{E_{\eta,t}}) \leq 2 \exp\left(-\frac{\delta^2 t}{16\lambda}\right) + \frac{6\lambda}{\delta} \exp\left(-\frac{\delta^2 t}{64\lambda}\right) \leq \frac{8\lambda}{\delta} \exp\left(-\frac{\delta^2 t}{64\lambda}\right)$$

Moreover, the Taylor series at  $a_\lambda$  of the function  $f$  is

$$f(a_\lambda - \eta) = f(a_\lambda) - \frac{\lambda^3 \eta^2}{4} + o(\eta^2), \quad \text{so} \quad \delta(\eta) \underset{\eta \rightarrow 0}{\sim} \frac{\lambda^3 \eta^2}{16} \quad \text{and} \quad \frac{\delta(\eta)^2}{64\lambda} \underset{\eta \rightarrow 0}{\sim} \frac{\lambda^5 \eta^4}{2^{14}}.$$

Consequently, there exists a constant  $c$  which only depends on  $\lambda$ , such that  $\forall \eta \in (0, c)$ ,

$$\mathbb{P}(\overline{E_{\eta,t}}) \leq \frac{1}{c\eta^2} \exp(-c\eta^4 t). \quad \square$$

### 2.3 Proof of Theorem 2.1 and of Corollary 2.1

In this section we first show how Lemma 2.1 makes it possible to get Theorem 2.1, and then prove Corollary 2.1.

*Proof of Theorem 2.1.* As mentioned in section 2.1, the main idea of this proof is to decompose the variable to study in a suitable way such that the convergence of each term can be studied. This decomposition involves indicator functions of events whose probability tends to zero or one, and a decomposition of the variable  $L_\lambda(t, t)$  for a positive  $t$  governed by a decomposition of a geodesic into two part.

Formally, take  $\lambda > 1$ ,  $t > 0$ , and as above set  $a_\lambda = 1 - \lambda^{-2}$  and  $\eta_t := t^{-1/5} \wedge \frac{a_\lambda}{2}$ . We consider the stationary process with intensity  $\lambda$  in the rectangle  $[0, t]^2$  denoted by  $L_\lambda$ . denote by  $\Delta_t$  the variable to study

$$\Delta_t = \frac{L_\lambda(t, t) - (\lambda + \lambda^{-1})t}{\sqrt{t}} - \frac{N_\lambda(a_\lambda t) - \lambda a_\lambda t}{\sqrt{t}},$$

and decompose it on the event  $E_{\eta_t, t}$  defined above and its complementary:

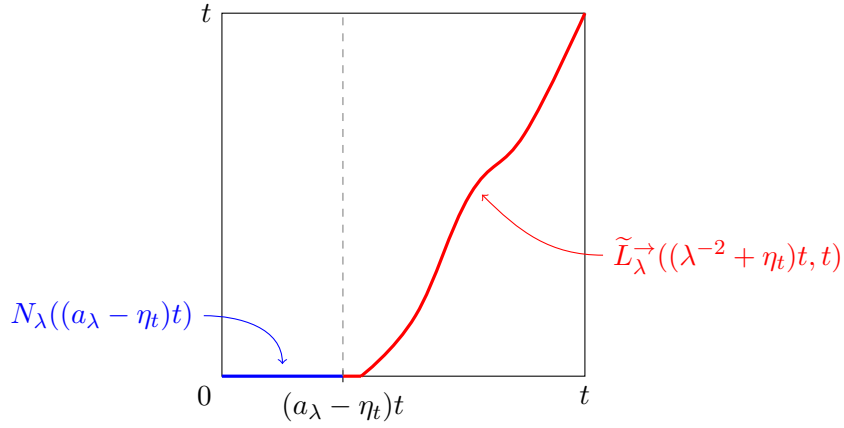
$$\Delta_t = \Delta_t \mathbf{1}_{\overline{E_{\eta_t, t}}} + \Delta_t \mathbf{1}_{E_{\eta_t, t}}. \quad (2.2)$$

According to Lemma 2.1, the first term converges to 0 in probability when  $t$  tends to  $+\infty$ , since  $\eta_t^4 t \asymp t^{1/5}$  tends to  $+\infty$  with the chosen  $\eta_t$ . It remains to prove that so does the second term.

We couple the process  $L_\lambda$  with another stationary process in the rectangle  $[0, (\lambda^{-2} + \eta_t)t] \times [0, t]$  such that the sources and the  $\alpha$ -points correspond to the first one in the rectangle  $[(a_\lambda - \eta_t)t, t] \times [0, t]$  translated by the horizontal vector  $(-(a_\lambda - \eta_t)t, 0)$ , and the sinks are drawn independently of everything else with intensity  $\lambda^{-1}$ . This second process is denoted by  $\tilde{L}_\lambda$ . Consequently,

$$L_\lambda(t, t) \mathbf{1}_{E_{\eta_t, t}} \stackrel{(d)}{=} \left[ N_\lambda((a_\lambda - \eta_t)t) + \tilde{L}_\lambda^\rightarrow((\lambda^{-2} + \eta_t)t, t) \right] \mathbf{1}_{E_{\eta_t, t}},$$

where  $\tilde{L}_\lambda^\rightarrow(x, t)$  denotes the length of a geodesic from  $(0, 0)$  to  $(x, t)$  in the second stationary process compelled to take no sinks. This decomposition is illustrated in Figure 2.3.



**Figure 2.3:** Decomposition of the longest path on the event  $E_{\eta t, t}$ , that is to say when it passes through the point  $((a_\lambda - \eta_t)t, 0)$ .

Moreover, the expectation of the variable  $L_\lambda(t, t)$  can be written as follows:

$$\mathbb{E}[L_\lambda(t, t)] = (\lambda + \lambda^{-1})t = (\lambda a_\lambda - \eta_t \lambda)t + (2\lambda^{-1} + \eta_t \lambda)t.$$

Consequently, if we set

$$\begin{aligned} \Delta_t^{(1)} &= \frac{N_\lambda((a_\lambda - \eta_t)t) - (\lambda a_\lambda - \eta_t \lambda)t}{\sqrt{t}} - \frac{N_\lambda(a_\lambda t) - \lambda a_\lambda t}{\sqrt{t}} \\ \text{and } \Delta_t^{(2)} &= \frac{\tilde{L}_\lambda^{\rightarrow}((\lambda^{-2} + \eta_t)t, t) - (2\lambda^{-1} + \eta_t \lambda)t}{\sqrt{t}} \end{aligned}$$

then

$$\Delta_t \mathbf{1}_{E_{\eta t, t}} = \left[ \Delta_t^{(1)} + \Delta_t^{(2)} \right] \mathbf{1}_{E_{\eta t, t}}. \quad (2.3)$$

Since  $N_\lambda$  is a PPP with intensity  $\lambda$ ,

$$\Delta_t^{(1)} = \frac{\lambda \eta_t t - (N_\lambda(a_\lambda t) - N_\lambda((a_\lambda - \eta_t)t))}{\sqrt{t}} \stackrel{(d)}{=} \frac{\lambda \eta_t t - N_\lambda(\eta_t t)}{\sqrt{t}}.$$

Consequently,

$$\mathbb{E}[\Delta_t^{(1)}] = 0 \quad \text{and} \quad \text{Var}(\Delta_t^{(1)}) = \frac{\text{Var}(N_\lambda(\eta_t t))}{t} = \lambda \eta_t \xrightarrow{t \rightarrow \infty} 0.$$

So the variable  $\Delta_t^{(1)}$  converges to 0 in  $L^2$ , hence in probability.

Now, let us prove that the variable  $\Delta_t^{(2)}$  converges to 0 in probability. To do that we introduce as above an event whose probability tends to 1. Denote by  $\tilde{E}_{\eta t, t}$  the event ‘‘in the second stationary model, all geodesics from  $(0, 0)$  to  $((\lambda^{-2} + \eta_t)t, t)$  take no sink’’. In particular,

$$\tilde{L}_\lambda^{\rightarrow}((\lambda^{-2} + \eta_t)t, t) \mathbf{1}_{\tilde{E}_{\eta t, t}} = \tilde{L}_\lambda((\lambda^{-2} + \eta_t)t, t) \mathbf{1}_{\tilde{E}_{\eta t, t}}.$$

It can be proved as in Lemma 2.1 that there exists a constant  $c'$  which only depends on  $\lambda$ , such that

$$\forall \eta \in (0, c'), \forall t > 0, \quad \mathbb{P}(\tilde{E}_{\eta t, t}) \leq \frac{1}{c\eta^2} \exp(-c\eta^4 t). \quad (2.4)$$

The expectation of the variable  $\tilde{L}_\lambda((\lambda^{-2} + \eta_t)t, t)$  can be computed explicitly:

$$\mathbb{E} \left[ \tilde{L}_\lambda((\lambda^{-2} + \eta_t)t, t) \right] = \lambda(\lambda^{-2} + \eta_t)t + t\lambda^{-1} = (2\lambda^{-1} + \eta_t\lambda)t.$$

Let us prove that  $\text{Var} \left( \tilde{L}_\lambda((\lambda^{-2} + \eta_t)t, t) \right) = o(t)$ , in order to get the following  $L^2$  convergence:

$$\frac{\tilde{L}_\lambda((\lambda^{-2} + \eta_t)t, t) - (2\lambda^{-1} + \eta_t\lambda)t}{\sqrt{t}} \xrightarrow[t \rightarrow \infty]{L^2} 0. \quad (2.5)$$

Set  $\gamma_t = \sqrt{\lambda^{-2} + \eta_t}$ . Using a scaling argument similar to the one of equation (2.1), we get

$$\tilde{L}_\lambda((\lambda^{-2} + \eta_t)t, t) = \tilde{L}_\lambda(\gamma_t^2 t, t) \stackrel{(d)}{=} L_{\lambda\gamma_t}(\gamma_t t, \gamma_t t).$$

Furthermore, [CG06, equation (2.7)] provides the following inequality :

$$\text{Var}(L_\lambda(t, t)) \leq \left( \lambda - \frac{1}{\lambda} \right) t + \text{Var}(L_1(t, t)).$$

Consequently,

$$\text{Var} \left( \tilde{L}_\lambda(\gamma_t^2 t, t) \right) = \text{Var}(L_{\lambda\gamma_t}(\gamma_t t, \gamma_t t)) \leq \left( \lambda\gamma_t - \frac{1}{\lambda\gamma_t} \right) \gamma_t t + \text{Var}(L_1(\gamma_t t, \gamma_t t)).$$

Since  $\gamma_t \xrightarrow[t \rightarrow \infty]{} \lambda^{-1}$ ,  $(\lambda\gamma_t - \frac{1}{\lambda\gamma_t})\gamma_t t = o(t)$ . Moreover, again according to [CG06],  $\text{Var}(L_1(t, t)) \asymp t^{2/3} = o(t)$ . Consequently,  $\text{Var}(L_1(\gamma_t t, \gamma_t t)) = o(t)$ , so  $\text{Var} \left( \tilde{L}_\lambda((\lambda^{-2} + \eta_t)t, t) \right) = o(t)$  and the  $L^2$  convergence (2.5) holds. Finally, in the following decomposition

$$\begin{aligned} \Delta_t^{(2)} &= \Delta_t^{(2)} \mathbf{1}_{\tilde{E}_{\eta_t, t}} + \Delta_t^{(2)} \mathbf{1}_{\tilde{E}_{\eta_t, t}^c} \\ &= \Delta_t^{(2)} \mathbf{1}_{\tilde{E}_{\eta_t, t}} + \frac{\tilde{L}_\lambda((\lambda^{-2} + \eta_t)t, t) - (2\lambda^{-1} + \eta_t\lambda)t}{\sqrt{t}} \mathbf{1}_{\tilde{E}_{\eta_t, t}^c}, \end{aligned}$$

the first term converges to 0 in probability thanks to equation (2.4) since  $\eta_t^4 t \underset{t \rightarrow \infty}{\asymp} t^{1/5}$ , and so does the second as it converges to 0 in  $L^2$ . So we have shown the wanted convergence :

$$\Delta_t^{(2)} \xrightarrow[t \rightarrow \infty]{\mathbb{P}} 0.$$

Using decompositions (2.3) and then (2.2), we conclude that

$$\Delta_t = \frac{L_\lambda(t, t) - (\lambda + \lambda^{-1})t}{\sqrt{t}} - \frac{N_\lambda(a_\lambda t) - \lambda a_\lambda t}{\sqrt{t}} \xrightarrow[t \rightarrow \infty]{\mathbb{P}} 0.$$

□

Corollary 2.1 can now be deduced.



*Proof of Corollary 2.1.* As said in Section 2.1, one can assume  $\lambda > 1$  without loss of generality by a symmetry argument.

Since  $N_\lambda((1 - \frac{1}{\lambda^2})t)$  follows a Poisson distribution with parameter  $(\lambda - \frac{1}{\lambda})t$ , the central limit theorem provides the following convergence:

$$\frac{N_\lambda((1 - \frac{1}{\lambda^2})t) - (\lambda - \frac{1}{\lambda})t}{\sqrt{t}} \xrightarrow[t \rightarrow \infty]{(d)} \mathcal{N}\left(0, \lambda - \frac{1}{\lambda}\right).$$

By summing this convergence in distribution with the convergence in probability stated in Theorem 2.1, Slutsky's Lemma ensures that

$$\frac{L_\lambda(t, t) - (\lambda + \lambda^{-1})t}{\sqrt{t}} \xrightarrow[t \rightarrow \infty]{(d)} \mathcal{N}\left(0, \lambda - \frac{1}{\lambda}\right),$$

and the corollary is proved. □

## Chapitre 3

# A stationary modification of O’Connell–Yor’s problem

### Abstract

We introduce an optimisation problem which can be seen as a stationary modification of O’Connell–Yor’s problem. In order to prove the stationarity of this model, we introduce a system of Hammersley’s lines which is discrete in space and continuous in time, and prove that it is stationary in both directions. Rescaling this system to transform among others Poisson processes into Brownian motions, we prove the stationarity of our original model. This stationarity makes it possible to write the solution of the optimisation problem as a sum of two (correlated) random variables with Gaussian and Gamma distribution respectively. This fact, together with a concentration property of the optimiser gives the well-known first order asymptotics of the largest eigenvalue of a  $n \times n$  GUE matrix as  $2\sqrt{n}$ .

### 3.1 Introduction

Let  $B = (B_k)_{k \geq 1}$  be a sequence of independent standard Brownian motions,  $n$  a positive integer and  $u$  a positive real. Consider the following random variable:

$$L_{n,u}(B) := \sup_{0=t_0 < t_1 < \dots < t_{n-1} < t_n = u} \left( \sum_{k=1}^n [B_k(t_k) - B_k(t_{k-1})] \right). \quad (3.1)$$

This quantity appears in queuing processes and last passage percolation (LPP) models: related formulas can be found in [Mut79, SK90, Gan98]. This variable can also be interpreted as the maximal energy of a path (defined by the indexes  $t_0, \dots, t_n$ ) in a semi-discrete polymer model introduced by O’Connell and Yor [OY01]. The problem of finding the distribution of the variable  $L_{n,u}(B)$  was posed by Glynn and Whitt [GW91]. Notice that by a scaling argument it suffices to study the case  $u = 1$ , and in the sequel we will denote  $L_n(B) = L_{n,1}(B)$ . It was observed independently by Baryshnikov in [Bar01] and Gravner–Tracy–Widom in [GTW01] that the distribution of the variable  $L_n(B)$  coincides with the distribution of the largest eigenvalue of a Gaussian Unitary Ensemble (GUE)  $n \times n$  matrix. Subsequently, O’Connell and Yor in [OY02] showed that this correspondence can be extended to the whole spectrum of a GUE matrix by considering extensions of the functional  $L_n$ .

Not very surprisingly, the distribution of the variable  $L_n(B)$  cannot be explicitly computed. In this chapter, we exhibit a modification of the functional  $L_{n,u}$  for which an underlying two-parameter process is stationary both in the space parameter  $n$  (what we will refer as *horizontal* stationarity) and in the time parameter  $u$  (*vertical* stationarity). This process can be seen as a Brownian queuing process with  $n$  servers up to time  $u$  as defined by O’Connell and Yor in [OY01]. In this chapter, the vertical stationarity is implicit as the process starts from  $-\infty$ , so that at time  $t = 0$  the system has already reached its stationary measure by ergodicity. This process has also been studied by Konstantopoulos in [Kon16] where horizontal stationarity is proved in the case of a single server. Here, we combine both notions of stationarity into a single result, and prove them simultaneously.

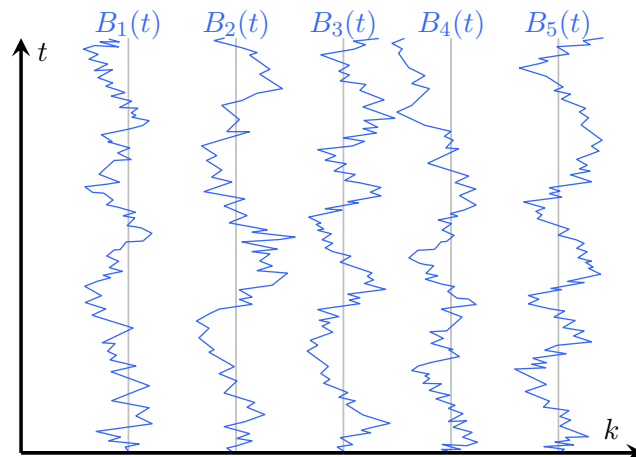
The chapter is organised as follows. In section 3.2, we introduce the notation and give the precise definition of our process of interest, which can also be interpreted as a LPP process in the semi-discrete quarter plane  $\mathbb{N} \times \mathbb{R}_+$ . We then state the main results, with first of all the stationarity theorem which ensures that this process is translation invariant in a certain sense. We then state two propositions. The first one establishes that the length of the longest path in this LPP model can be written as a sum of two (correlated) random variables with Gaussian and Gamma distribution respectively. The second one states that the optimal path is close to the straight line when the parameters are properly scaled together. Finally, we recover as a corollary of these propositions the first order asymptotics of the largest eigenvalue of a  $n \times n$  GUE matrix as  $2\sqrt{n}$ .

In section 3.3, we introduce a discretisation of the process which is a model of Hammersley’s lines which is discrete in space and continuous in time, that is to say a LPP model in the quarter plane  $\mathbb{N} \times \mathbb{R}_+$ . This process is indeed easier to study. We prove that it is stationary in both directions, using a local balance approach in the spirit of what is done in [Sep09, Lemma 2.3] or [BEGG16, Section 3] in discrete LPP models.

In section 3.4, we rescale this system to transform among others Poisson processes into Brownian motions. We then show that the stationarity property is preserved, by showing that the involved functional is continuous for a suitable topology. We prove that as a consequence of this stationarity, the distribution of the modified version of  $L_n(B)$  can be written as a sum of two random variables with explicit distributions.

In section 3.5, we prove that when the parameters are properly scaled, an optimal path is concentrated around the diagonal of the rectangle. This is done by showing that a longest path compelled to be quite far from this diagonal is too short to be an unconstrained longest path, and by using concentration inequalities of the optimal length given by the previous proposition.

In section 3.6, we prove that the above-mentioned scaling between the parameters makes (almost) coincide the original model and the stationary model. This makes it possible to recover as a corollary of these propositions the first order asymptotics of the largest eigenvalue of a  $n \times n$  GUE matrix. Another proof of this result using the process  $L_n(B)$  has been done in [HMO02] using arguments of [OY01] based on a technique introduced in [Sep98b] which exploits a kind of convex duality between density and speed in microscopic models for hydrodynamic systems.



**Figure 3.1:** An example of the vertical representation of Brownian motions

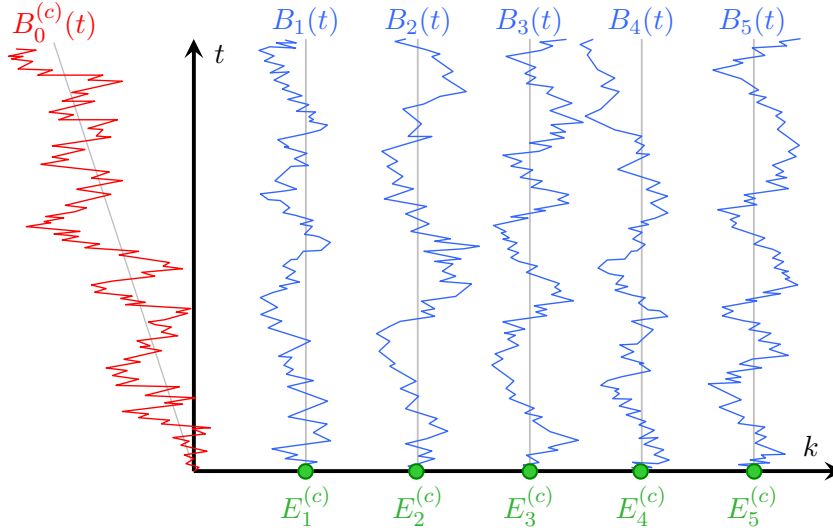
## 3.2 Main results

As mentioned above, in this section we start by introducing the notation and by defining the process which will be proved to be stationary, before stating our main results.

The variable described in equation (3.1) can be interpreted as the length of a longest path in a LPP problem. Indeed, consider increasing paths from the south-west point  $(1, 0)$  to the north-east point  $(n, u)$  in the rectangle  $\llbracket 1, n \rrbracket \times [0, u]$ , where the length of a segment of the form  $\{k\} \times [s, t]$  is defined by the increment of the Brownian motion  $B_k$  between time  $s$  and time  $t$ . In other terms, we jump from a Brownian motion to the next Brownian motion successively, in order to maximise the sum of the increments of the Brownian motions we observed between jump times.

In the initial model described above, we only consider standard Brownian motions. They can be seen as evolving on each vertical line  $\{k\} \times \mathbb{R}_+$ , with  $k \in \llbracket 1, n \rrbracket$ . An example of this representation is shown in Figure 3.1. The above-mentioned modification consists of adding two processes which can be seen as evolving on both axis, as it has been done in LPP models by Groeneboom [Gro01] and Cator and Groeneboom [CG05]. On the  $y$ -axis, consider also an independent drifted standard Brownian motion, called “sink process”. Compared to the other Brownian motions, you get a deterministic bonus proportional to the time spent on the axis. On the other hand, a sequence of i.i.d. exponential random variables is added on each integer point on the  $x$ -axis, called “sources”. These variables can be seen as rewards: each time we decide to ignore a (drifted or not) Brownian motion, we collect the value of the corresponding exponential random variable. An example of the vertical representation is shown in Figure 3.2. Note that to have an increasing path, the rewards can only be earned at the beginning of the path. In other words, it is impossible to benefit from both the sinks and the sources.

Formally, let  $n \in \mathbb{N}^*, c > 0$ . We consider the following collection of independent random variables:



**Figure 3.2:** An example of the vertical representation of Brownian motions (in blue) with sources (in green) and sink process (in red)

- $B = (B_k)_{k \geq 1}$  a sequence of independent standard Brownian motions,
- $B_0^{(c)}$  a standard Brownian motion with drift  $c$ , namely, for non-negative  $t$ ,  $B_0^{(c)}(t) = B_0(t) + ct$  where  $B_0$  is a standard Brownian motion,
- $E^{(c)} = (E_k^{(c)})_{k \geq 1}$  a sequence of independent exponential random variables with mean  $1/c$ .

We define our main quantity of interest: for any positive  $u$ ,

$$L_{n,u}^{(c)} = L_{n,u}^* \left( E^{(c)}, B_0^{(c)}, B \right) := \sup_{i, t_0, \dots, t_{n-1}} \left( \sum_{k=1}^i E_k^{(c)} + B_0^{(c)}(t_0) + \sum_{k=1}^n [B_k(t_k) - B_k(t_{k-1})] \right), \quad (3.2)$$

where the supremum is taken over the set

$$\{(i, t_0, \dots, t_{n-1}) : i \in \llbracket 0, n \rrbracket ; \forall j \leq i-1, t_j = 0 ; 0 < t_i < t_{i+1} < \dots < t_n = u\}.$$

When  $u = 1$ , we will denote  $L_n^* = L_{n,1}^*$  and  $L_n^{(c)} = L_{n,1}^{(c)}$ . As for the random variable  $L_n(B)$ , the variable  $L_n^{(c)}$  represents the length of the longest path in a LPP problem, but now with boundary conditions given by the variables  $E^{(c)}$  and the process  $B_0^{(c)}$ . Namely, the path is either allowed to collect rewards of the form  $E_i^{(c)}$  when going through the point of the  $x$ -axis  $(i, 0)$ , or collecting the increment of the drifted Brownian motion along the  $y$ -axis up to time  $t_0$ . The variable  $i$  corresponds to the number of exponential rewards taken, that is to say the second option occurs if and only if  $i$  is zero, in which case  $t_0$  and  $B_0^{(c)}(t_0)$  are positive.

We can now state our main results. First of all, this modification makes the system stationary in the following sense:

**Theorem 3.1.** *For any  $(n, s) \in \mathbb{N} \times \mathbb{R}_+$ , the distribution of the process*

$$\left( L_{n+k, s+t}^{(c)} - L_{n, s}^{(c)} \right)_{(k, t) \in \mathbb{N} \times \mathbb{R}_+}$$

*does not depend on the couple  $(n, s)$ .*

This stationarity provides information on the distribution of the variable  $L_{n, u}^{(c)}$ : it can be written as a sum of a Gaussian random variable and a Gamma random variable.

**Proposition 3.1.** *For any  $(n, s, u) \in \mathbb{N} \times (\mathbb{R}_+)^2$ , there exist two (dependent) random variables  $Z_u^{(cu)}$  and  $\Gamma_n^{(c)}$  with marginal distribution  $\mathcal{N}(cu, u)$  and  $\text{Gamma}(n, c)$  such that*

$$L_{n, u}^{(c)} \stackrel{(d)}{=} Z_u^{(cu)} + \Gamma_n^{(c)}, \quad (3.3)$$

where the symbol  $\stackrel{(d)}{=}$  means “equal in distribution”.

*Remark.* • It can be proved that the random variables  $Z_u^{(cu)}$  and  $\Gamma_n^{(c)}$  are negatively correlated, that is to say  $\text{Cov}(\Gamma_n^{(c)}, Z_u^{(cu)}) \leq 0$ . However, the joint distribution of the couple  $(\Gamma_n^{(c)}, Z_u^{(cu)})$  is unknown.

- Let us point out that the variable  $Z_u^{(cu)}$  has the same distribution as the increment of the drifted Brownian motion which is on  $y$ -axis between times 0 and  $u$ , and the variable  $\Gamma_n^{(c)}$  has the same distribution as the sum of the  $n$  rewards on the  $x$ -axis. If we could take a path only made of the two axes (which is not an admissible path of course), we would have got a length which would have been the sum of two random variables with the same marginal distribution as in equation (3.3), but they would have been independent.

Proposition 3.1 also makes it possible to recover information on the original problem without sinks nor sources. The following stochastic domination holds, with the notation of equation (3.1),

$$L_n(B) \leq L_n^{(c)}, \quad (3.4)$$

since an admissible path for  $L_n$  is also an admissible path for  $L_n^{(c)}$ , taking exactly one exponential reward  $E_1^{(c)}$  which is positive.

To get the tightest possible bound, one may want to optimise the right-hand side in the parameter  $c$ . Heuristically, one can first do it for the expectation. It is equal to  $c + \frac{n}{c}$  by Proposition 3.1, which is minimal for  $c = \sqrt{n}$ . We use the stationary model with this particular scaling between the parameters  $c$  and  $n$  in our next result. We prove that an optimal path<sup>1</sup> in this case is asymptotically close to the segment with ends the points  $(0, 0)$  and  $(n, 1)$ .

Formally, take  $u = 1$  and denote by  $P_n^{(\sqrt{n})}$  the optimal path in the stationary case where  $c = \sqrt{n}$  which has been rescaled in the square  $[0, 1]^2$ , that is to say

$$P_n^{(\sqrt{n})} := \bigcup_{j=0}^n \left( \left\{ \frac{j}{n} \right\} \times [\mathbf{t}_{j-1}^*, \mathbf{t}_j^*] \right), \quad (3.5)$$

<sup>1</sup>The supremum is indeed a maximum, and there exists a unique optimal path.

where the subdivision  $(\mathbf{t}_0^*, \dots, \mathbf{t}_{n-1}^*)$  satisfies the supremum in equation (3.2) and where we have set  $\mathbf{t}_{-1}^* := 0$  and  $\mathbf{t}_n^* := 1$ . Moreover, denote by  $\Delta$  the first diagonal of the square  $[0, 1]^2$ , that is to say

$$\Delta := \{(x, y) \in [0, 1]^2 : x = y\},$$

and for any non-empty subset  $E \subset \mathbb{R}^2$ , defines its distance to the diagonal  $\Delta$  by

$$d(E, \Delta) = \sup_{(x,y) \in E} |x - y|.$$

**Proposition 3.2.** *The rescaled optimal path  $P_n^{(\sqrt{n})}$  is concentrated around the diagonal  $\Delta$ , in the sense that for any real  $\beta > 4$ ,*

$$\mathbb{P}\left(d\left(P_n^{(\sqrt{n})}, \Delta\right) \geq \frac{1}{n^{1/\beta}}\right) \xrightarrow{n \rightarrow \infty} 0.$$

This concentration of the optimal path in this stationary model around the diagonal tells us that it does not collect many sources and sinks on the axes. Consequently, it makes it possible to recover the following classical result for the original model (3.1):

**Corollary 3.1.** *The following almost sure limit holds:*

$$\frac{L_n(B)}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{a.s.} 2.$$

Let us mention that the order of magnitude of the next term in the asymptotics of  $L_n(B)$  is  $n^{1/6}$ , see for example [Bar01, GTW01].

The main idea of the proof of Theorem 3.1 is to introduce a Hammersley’s lines model which is discrete in space and continuous in time. We prove that it is stationary in both directions, then rescale this system to transform among others Poisson processes into Brownian motions, and show that the stationarity property is preserved. Let us start by defining this discrete model.

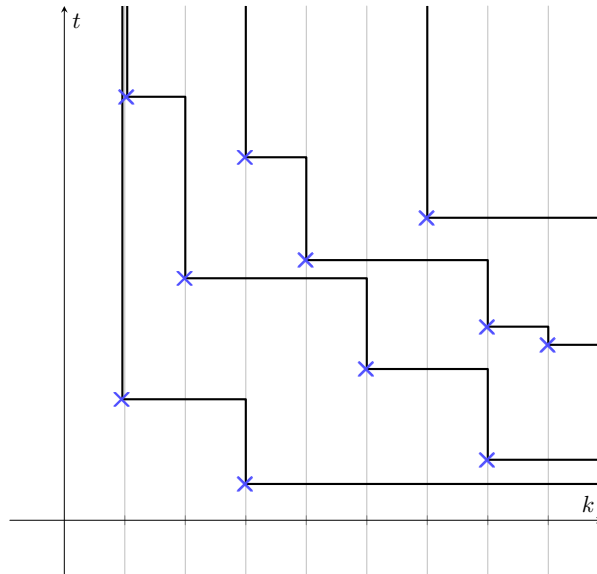
### 3.3 A discrete model

Before introducing the stationary stochastic process, we define a deterministic system and explain the link between Hammersley’s lines and the length of a longest non-decreasing subsequence.

#### 3.3.1 Hammersley’s lines and longest non-decreasing subsequences

Let us work on the semi-discrete quarter plane  $\mathbb{N} \times \mathbb{R}_+$ . The  $x$ -axis represents space (discrete here) while the  $y$ -axis represents time (continuous here) and will also be called  $t$ -axis. No object is assumed to be random in this subsection.

Fix  $b > 0$ , and let  $\xi_{n,b}$  be a discrete subset of the box  $\llbracket 1, n \rrbracket \times (0, b]$  whose elements have distinct second coordinates. It can be thought as a realisation of a Poisson point process, as it



**Figure 3.3:** Example of construction of Hammersley’s lines without sinks and sources

will be explained in the next subsection. We call the elements of this subset  $\alpha$ -points, following a terminology introduced by Groeneboom [Gro01] and used by Cator and Groeneboom [CG05].

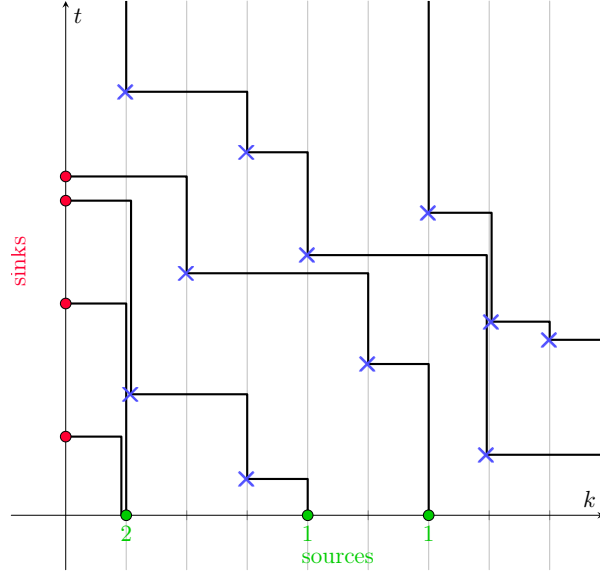
Let us define what are Hammersley’s lines here. They are broken lines running only upward and to the left, and can be constructed from the bottom up. As soon as an  $\alpha$ -point appears, the closest line among those located strictly at the right of the  $\alpha$ -point (if any) bifurcates: it turns left until it reaches the  $\alpha$ -point, then goes on vertically. If there is no such line, a new line appears horizontally from the right side, up to the  $\alpha$ -point, and then goes on vertically.

Several Hammersley’s lines may live on the same vertical line. They are always supposed to be distinct, and ordered from left to right, according to the order they arrived the first time on this vertical. With this precaution, two Hammersley’s lines never cross. See Figure 3.3 for an example.

To study this model, it is useful to add points on the axes which behaves like  $\alpha$ -points for the construction. The set of the points on the  $t$ -axis is also supposed to be discrete, and its points are called *sinks*. As soon as a line enters a sink it can never leaves the axis, so we consider that the line “dies” and we will no longer represent it. On the  $x$ -axis (supposed to be discrete) there may be several points at the same site, because with our construction several lines can live on the same vertical. These points are called *sources*, since a Hammersley’s line starts from each of them. Denote by  $\xi_{n,b}^*$  the set of  $\alpha$ -points, sinks and sources in  $\llbracket 0, n \rrbracket \times [0, b]$ . See Figure 3.4 for an example.

We relate the number of Hammersley’s lines to the length of longest non-drecreasing subsequences in a similar way to what is done by Hammersley [Ham72]. Let us introduce a partial order  $\prec$  on  $\mathbb{N} \times \mathbb{R}_+$ :





**Figure 3.4:** Example of construction of Hammersley's lines with sinks and sources

$$(k, t) \prec (k', t') \Leftrightarrow \begin{cases} k \leq k' \text{ and } t < t' \\ \text{or} \\ k < k' \text{ and } t = t'. \end{cases}$$

In words, a point is smaller than another one if it is located to its southwest, widely speaking. Moreover, set

$$L_{n,b} := \max\{L \mid \exists (k_i, t_i)_{1 \leq i \leq L} \in (\xi_{n,b})^L, \text{ non-decreasing for } \prec\},$$

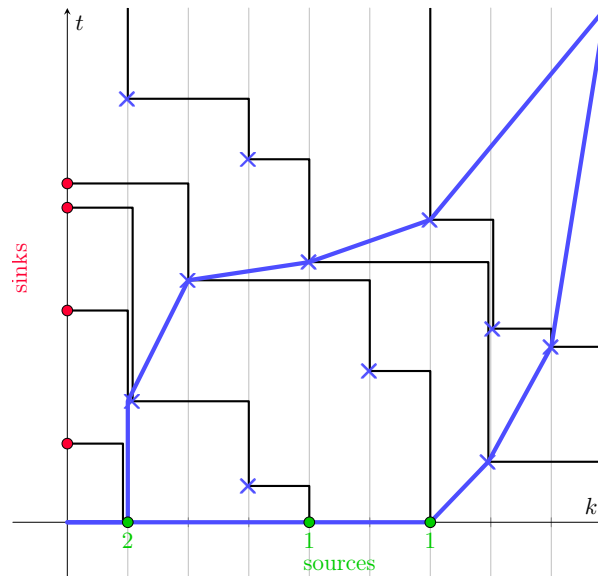
$$L_{n,b}^* := \max\{L \mid \exists (k_i, t_i)_{1 \leq i \leq L} \in (\xi_{n,b}^*)^L, \text{ non-decreasing for } \prec\}.$$

In words, the quantity  $L_{n,b}$  (resp.  $L_{n,b}^*$ ) represents the maximum number of  $\alpha$ -points (resp.  $\alpha$ -points, sinks and sources) a northeast path matching the points  $(0, 0)$  to  $(n, b)$  can collect. Notice that in the environment  $\xi_{n,b}^*$  it is impossible to collect both some sinks and some sources on such a path. Also, this notation is consistent with the one defined in the introduction, where the Brownian motions  $B$  are replaced by step functions which increase by one at the position of the  $\alpha$ -points.

As intended, these two lengths are related to the number of Hammersley's lines in the rectangle:

**Lemma 3.1.** *The variables  $L_{n,b}$  and  $L_{n,b}^*$  correspond to the number of Hammersley's lines in  $\llbracket 0, n \rrbracket \times [0, b]$ , respectively without and with sources and sinks.*

*Proof.* The proof is the same in both cases, let us deal here with the case without sources and sinks. Let  $H_{n,b}$  be the number of Hammersley's lines in  $\llbracket 0, n \rrbracket \times [0, b]$  in this case, and let us prove that  $L_{n,b} \leq H_{n,b}$  and  $L_{n,b} \geq H_{n,b}$ .



**Figure 3.5:** Example of two longest increasing subsequences (length 6) with sinks and sources

Since all the Hammersley's lines are decreasing, the points of any longest non-decreasing subsequence belong to distinct Hammersley's lines. Hence  $L_{n,b} \leq H_{n,b}$ .

Conversely, it is possible to construct a non-decreasing subsequence such as each point belongs to a different Hammersley's line, by recursion. To do so, we order Hammersley's lines from bottom left to top right, and we start constructing the path from the top-right corner. We choose any point from the last Hammersley's line. Then choose step by step the previous point among those of the previous Hammersley's line such that the subsequence remains non-decreasing which is always possible by construction. Hence  $L_{n,b} \geq H_{n,b}$ , and the wanted equality is proved.  $\square$

See Figure 3.5 for an example of longest non-decreasing subsequences with sinks and sources.

### 3.3.2 A stationary system

In the sequel, we are going to randomise the set  $\xi$  and its extension  $\xi^*$  such that the system becomes stationary in both dimensions in the following sense: the distribution of the process defined by the intersection of Hammersley's lines relating to  $\xi^*$  and the horizontal line with second coordinate  $t > 0$  (resp. the vertical line with first coordinate  $x > 0$ ) does not depend on  $t$  (resp. on  $x$ ).

Namely, let  $\xi$  be a Poisson point process (PPP) on  $\mathbb{N}^* \times \mathbb{R}_+$  with intensity 1. Take  $\lambda > 1$  and define the set  $\xi^*$  by the union of the set  $\xi$  with the sets of sinks and sources, which are defines as follows.

- The set of sinks is a Poisson point process on the half-line  $\{0\} \times \mathbb{R}_+$  with intensity  $\lambda$ .

- The set of sources is the union for all  $k \in \mathbb{N}^*$  of a random number  $G_k^{[\lambda]}$  of points at site  $(k, 0)$ , where the random variables  $(G_k^{[\lambda]})_{k \geq 1}$  are independent geometric random variables with parameter  $(1 - \lambda^{-1})$ , with this distribution denoted by  $\text{Geom}(1 - \lambda^{-1})$  and defined by:

$$\forall i \geq 0, \quad \mathbb{P}(G_k^{[\lambda]} = i) = (1 - \lambda^{-1})\lambda^{-i}.$$

Since the distribution of the set  $\xi^*$  depends on the parameter  $\lambda$ , we will denote it by  $\xi^{[\lambda]}$  in the sequel.

Let us add that all the above random quantities are chosen independently. For any integer  $k \in \mathbb{N}^*$  denote by  $(X_i)$  the counting process of the Poisson process  $\xi$  restricted to the half-line  $\{i\} \times \mathbb{R}_+$ , and similarly denote by  $X_0^{[\lambda]}$  the counting process of the Poisson process defining the sinks, i.e.

$$\forall t \in \mathbb{R}_+^*, \quad \begin{cases} \forall k \in \mathbb{N}^*, & X_k(t) := \#(\xi \cap \{k\} \times (0, t]), \\ & X_0^{[\lambda]}(t) := \#(\xi_{n,b}^{[\lambda]} \cap \{0\} \times (0, t]). \end{cases}$$

Consequently, with this notation we can rewrite the quantities  $L_{n,b}$  and  $L_{n,b}^*$ , which can be seen as functionals of the processes:

$$L_{n,b} := L_{n,b}(X) = \sup_{0=t_0 < t_1 < \dots < t_{n-1} < t_n=b} \left( \sum_{k=1}^n [X_k(t_k) - X_k(t_{k-1})] \right),$$

$$L_{n,b}^{[\lambda]} := L_{n,b}^*(G, X_0^{[\lambda]}, X) = \sup_{i, t_0, \dots, t_{n-1}} \left( \sum_{k=1}^i G_k^{[\lambda]} + X_0^{[\lambda]}(t_0) + \sum_{k=1}^n [X_k(t_k) - X_k(t_{k-1})] \right),$$

where the second supremum is taken over the set

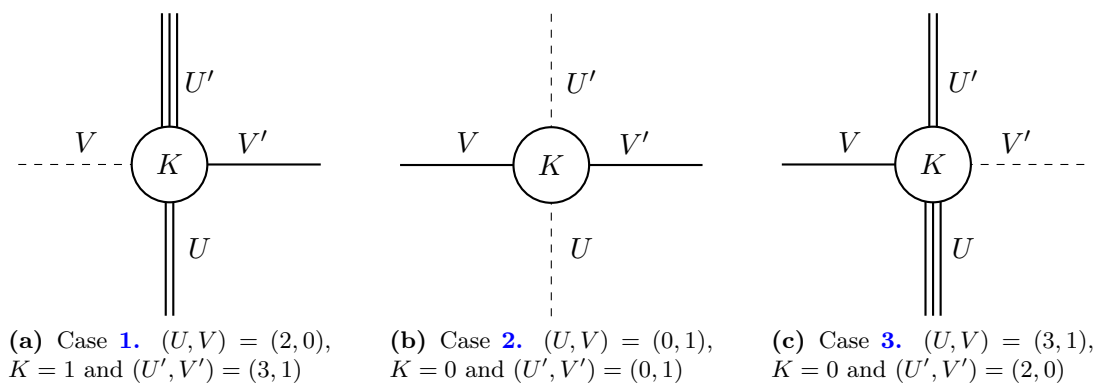
$$\{(i, t_0, \dots, t_{n-1}) : i \in \llbracket 0, n \rrbracket ; \forall j \leq i-1, t_j = 0 ; 0 < t_i < \dots < t_n = b\}.$$

Notice that we recover the same functional as in equation (3.2). To ensure consistent notation, let us also take the convention that if  $b$  is not specified in subscript, it is assumed to be equal to 1.

We start by checking that the chosen distributions make the model stationary in both directions. For  $(k, t) \in \mathbb{N}^* \times \mathbb{R}_+$ , set  $N_k^{[\lambda]}(t)$  the number of Hammersley's lines running (vertically) through the point  $(k, t)$ , and  $\mathbf{N}_k^{[\lambda]}(t) = (N_1^{[\lambda]}(t), \dots, N_k^{[\lambda]}(t))$ . Denote also by  $X_k^{[\lambda]}(t)$  the number of Hammersley's lines which crosses (horizontally) the segment  $\{k + \frac{1}{2}\} \times [0, t]$ . Notice that this notation is consistent with the one of the process  $X_0^{[\lambda]}$ , since Hammersley's lines cross the segment  $\{\frac{1}{2}\} \times [0, t]$  exactly at sinks second coordinate.

**Lemma 3.2.** *The system is stationary in both directions, that is to say: for any integer  $k \in \mathbb{N}^*$  and positive real  $b \in \mathbb{R}_+^*$ ,*

- the process  $(X_k^{[\lambda]}(t))_{t \in \mathbb{R}_+}$  is a Poisson point process with intensity  $\lambda$  ;
- the random vector  $\mathbf{N}_k^{[\lambda]}(b) = (N_1^{[\lambda]}(b), \dots, N_k^{[\lambda]}(b))$  has independent coordinates of marginal distribution  $\text{Geom}(1 - \lambda^{-1})$  ;



**Figure 3.6:** Examples of the three different local configurations around a point  $(1, y_i)$ .

- the process  $(X_k^{[\lambda]}(t))_{0 \leq t \leq b}$  and the vector  $\mathbf{N}_k^{[\lambda]}(b)$  are independent.

*Proof of Lemma 3.2.* As mentioned in section 3.1, the proof of this proposition relies on a local balance argument, similar to those used in [Sep09, Lemma 2.3] or [BEGG16, Section 3] for discrete LPP models.

Let us first show the result for  $k = 1$ .

Consider the subset  $E$  of  $\mathbb{R}_+$  defined by the (disjoint) union of the atoms of the PPPs  $X_0^{[\lambda]}$  and  $X_1$ . Since these two PPPs are independent with respective intensity  $\lambda$  and 1, the set  $E$  is a PPP with intensity  $1 + \lambda$ . Let us label its elements by increasing order :

$$E = \{y_1, y_2, y_3, \dots\}.$$

Notice that the line process on the left of the half line  $\{\frac{3}{2}\} \times \mathbb{R}_+$  do not evolve inside strips of the form  $[\frac{1}{2}, \frac{3}{2}] \times (y_i, y_{i+1})$ , and let us prove a local balance around a point of the form  $(1, y_i)$ , in a fashion similar to what is done in [BEGG16, Section 3].

We introduce some local notation for the sake of clarity. Denote by  $U$  (resp.  $V, U', V'$ ) the number of Hammersley's lines hitting the point  $(1, y_i)$  from the bottom (resp. left, top, right) and  $K$  the indicator that there is an  $\alpha$ -point at the point  $(1, y_i)$ . Notice that, since  $X_0^{[\lambda]}$  and  $X_1$  are independent PPPs, the variables  $V$  and  $K$  satisfy the almost sure relation  $V + K = 1$ .

To get a local balance, we want to prove that the couples  $(U, V)$  and  $(U', V')$  have the same distribution. To do so, one can express deterministically the variables  $U'$  and  $V'$  in terms of  $U, V$  and  $K$  depending on the following cases: for any  $u \in \mathbf{N}^*$ ,

1.  $(U, V) = (u, 0) \Rightarrow (U', V') = (u + 1, 1)$ .
2.  $(U, V) = (0, 1) \Rightarrow (U', V') = (0, 1)$ .
3.  $(U, V) = (u + 1, 1) \Rightarrow (U', V') = (u, 0)$ .

These three different cases are illustrated in Figure 3.6, with  $u = 2$  in these examples.

In case **1.**,  $V = 0$  so  $K = 1$ . There is an  $\alpha$ -point on the site  $(1, y_i)$ , so a Hammersley's line hits this point from the right and  $V = 1$ . This line turns and continues upwards, and increments the number of Hammersley's lines living on the half-line  $\{1\} \times \mathbb{R}_+$  by one:  $U' = U + 1$ .

In cases **2.** and **3.**,  $V = 1$  so  $K = 0$ . There is a Hammersley's line hitting the point  $(1, y_i)$  from the left, that is to say a sink at position  $(0, y_i)$ . If there is no vertical line on the half line  $\{1\} \times \mathbb{R}_+$  at time  $y_i$ , that is to say  $U = 0$ , the Hammersley's line continues to the right:  $V' = 1$ , and  $U' = 0$  still (case **2.**). Otherwise, that is to say  $U \geq 1$ , one of the vertical lines turns to the left towards the sink:  $U' = U - 1$  and  $V' = 0$ .

By construction of the set  $E$ , and since  $X_0^{[\lambda]}$  and  $X_1$  are independent PPPs with respective intensity  $\lambda$  and 1, the variable  $V$  follows a Bernoulli distribution with parameter  $\frac{\lambda}{1+\lambda}$ . In addition, assume that the variable  $U$  follows a geometric distribution of parameter  $1 - \lambda^{-1}$  and is independent from the variable  $V$ .

As mentioned above, we want to prove that the couples  $(U, V)$  and  $(U', V')$  have the same distribution. One can write the distribution of the couple  $(U', V')$  using the law of total probability according to the above-mentioned cases. In particular, for any integer  $u \in \mathbb{N}$ ,

$$\begin{aligned} \mathbb{P}((U', V') = (u, 0)) &= \mathbb{P}((U, V) = (u + 1, 1)) = (1 - \lambda^{-1})\lambda^{-u-1} \times \frac{\lambda}{1 + \lambda}, \\ \mathbb{P}((U', V') = (u + 1, 1)) &= \mathbb{P}((U, V) = (u, 0)) = (1 - \lambda^{-1})\lambda^{-u} \times \frac{1}{1 + \lambda}. \end{aligned}$$

Notice that, for any  $u \in \mathbb{N}$ ,

$$\begin{aligned} \mathbb{P}((U', V') = (u, 0)) &= \mathbb{P}((U, V) = (u, 0)), \\ \text{and } \mathbb{P}((U', V') = (u + 1, 1)) &= \mathbb{P}((U, V) = (u + 1, 1)) \end{aligned}$$

Moreover,

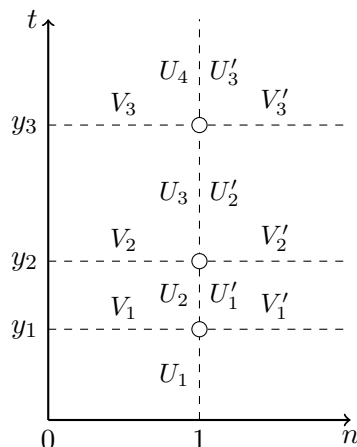
$$\mathbb{P}((U', V') = (0, 1)) = \mathbb{P}((U, V) = (0, 1)).$$

We have proved that the local balance is achieved: if  $(U, V)$  follows the product distribution  $\text{Geom}(1 - \lambda^{-1}) \otimes \text{Ber}\left(\frac{\lambda}{1+\lambda}\right)$ , so does  $(U', V')$ .

Let us use this local balance to prove that the number  $N_1^{[\lambda]}(b)$  of Hammerley's lines crossing vertically the point  $(1, b)$  follows a geometric distribution with parameter  $1 - \lambda^{-1}$ , and is independent of the process  $(X_1^{[\lambda]}(t))_{0 \leq t \leq b}$  describing the Hammerley's lines crossing horizontally the segment  $\{\frac{3}{2}\} \times [0, b]$ . For any integer  $i \in \mathbb{N}^*$ , denote by  $U_i, V_i, U'_i, V'_i$  the number of Hammersley's lines hitting the point  $(1, y_i)$  from the four directions as above. In particular,  $U'_i = U_{i+1}$ . A visualisation of this notation is presented in Figure 3.7.

Set  $i \in \mathbb{N}^*$ . We prove by induction on  $i$  that the variables  $(U'_i, V'_1, \dots, V'_i)$  follow the product distribution  $\text{Geom}(1 - \lambda^{-1}) \otimes \text{Ber}\left(\frac{\lambda}{1+\lambda}\right)^{\otimes i}$ . The case  $i = 1$  is proved above, as the couple  $(U_1, V_1)$  follows the distribution  $\text{Geom}(1 - \lambda^{-1}) \otimes \text{Ber}\left(\frac{\lambda}{1+\lambda}\right)$ .

Assume the result is proved for an integer  $i \in \mathbb{N}^*$ . The couple  $(U_{i+1}, V_{i+1}) = (U'_i, V_{i+1})$  follows the distribution  $\text{Geom}(1 - \lambda^{-1}) \otimes \text{Ber}\left(\frac{\lambda}{1+\lambda}\right)$ , since  $V_{i+1}$  is independent of everything below the line  $y = y_{i+1}$ . So the couple  $(U'_{i+1}, V'_{i+1})$  also follows the distribution  $\text{Geom}(1 - \lambda^{-1}) \otimes$



**Figure 3.7:** Variables used to count the number of Hammersley's lines living on each dashed segment.

$\text{Ber}\left(\frac{\lambda}{1+\lambda}\right)$ . Moreover, notice that this couple only depends on the quantities below the line  $y = y_{i+1}$  through the variable  $U'_i$ , which is assumed to be independent of the variables  $(V'_1, \dots, V'_i)$ . Consequently, the variables  $(U'_{i+1}, V'_1, \dots, V'_{i+1})$  are independent and follow the wanted product distribution.

Let  $b$  be a positive real. The cardinal  $I$  of the set  $E \cap [0, b]$  is equal to  $X_0^{[\lambda]}(b) + X_1(b)$  and follows a  $\text{Poi}(b(1+\lambda))$  distribution. Notice that the random variables  $(U'_I, V'_1, \dots, V'_I)$  do not depend on the values of the variables  $(y_i)_{i \in [1, I]}$ . We have proved that conditionally on  $I$ , those variables  $(U'_I, V'_1, \dots, V'_I)$  are independent. Since the variable  $I$  can also be drawn independently of everything else, we can conclude that the random variables  $(U'_I, V'_1, \dots, V'_I)$  are independent.

Notice that the number  $N_1^{[\lambda]}(b)$  of Hammerley's lines crossing vertically the point  $(1, b)$  is equal to the variable  $U'_I$ , so it follows a geometric distribution with parameter  $1 - \lambda^{-1}$ . Moreover, the random collection  $(V'_1, \dots, V'_I)$  defines the process  $(X_k^{[\lambda]}(t))_{0 \leq t \leq b}$  describing the Hammerley's lines which crosses horizontally the segment  $\{\frac{3}{2}\} \times [0, b]$  as a **thinning** of the PPP defined by the set  $E$ : each point is kept independently with probability  $\frac{\lambda}{1+\lambda}$  or is removed otherwise. Since the intensity of the process defined by  $E$  is  $1 + \lambda$ , we conclude that the process  $X_1^{[\lambda]}$  is a PPP with intensity  $(1 + \lambda) \frac{\lambda}{1+\lambda} = \lambda$ . The previous paragraph allows us to conclude that the variable  $N_1^{[\lambda]}(b)$  and the process  $X_1^{[\lambda]}$  are independent.

The result of the proposition can be deduced by induction on  $k$ . Assume the result is proved for a given  $k$ , and let us show that it is still true for  $k + 1$ .

Since  $X_1^{[\lambda]}$  is a PPP with intensity  $\lambda$ , and is independent of the process  $X_{k+1}$  and of the variable  $G_{k+1}^{[\lambda]}$ , the system behaves exactly the same way in the strip  $[k + \frac{1}{2}, k + \frac{3}{2}] \times \mathbb{R}_+$  than it does in the strip  $[\frac{1}{2}, \frac{3}{2}] \times \mathbb{R}_+$ . From what has been proved above, we deduce that  $N_{k+1}^{[\lambda]}(b)$  follows a  $\text{Geom}(1 - \lambda^{-1})$  distribution, and is independent of  $X_{k+1}^{[\lambda]}$  which is a PPP with intensity  $\lambda$ . Moreover, the variable  $N_{k+1}^{[\lambda]}(b)$  and the process  $X_{k+1}^{[\lambda]}$  only depends on the system in the strip  $[\frac{1}{2}, k + \frac{1}{2}] \times \mathbb{R}_+$  through the process  $X_k^{[\lambda]}$ , which is assumed to be independent of the

random vector  $\mathbf{N}_k^{[\lambda]}(b)$ . So the wanted result is proved for  $k + 1$ , and therefore for any positive integer.  $\square$

We highlight a key equality which makes a link between the processes defined above and the number of Hammersley's lines in a given finite box.

**Lemma 3.3.** *The total number of Hammersley's lines in the rectangle  $\llbracket 0, n \rrbracket \times [0, b]$  in the stationary model breaks down in the following way:*

$$L_{n,b}^{[\lambda]} = X_0^{[\lambda]}(b) + \sum_{k=1}^n N_k^{[\lambda]}(b) = X_n^{[\lambda]}(b) + \sum_{k=1}^n G_k^{[\lambda]}. \quad (3.6)$$

For both equalities, the first term of the sum follows a Poisson distribution of parameter  $\lambda b$ , and the second term follows a negative binomial distribution of parameters  $n$  and  $1 - \lambda^{-1}$ .

*Proof.* We reason in a manner similar to that of the proof of Lemma 3.1: each Hammersley's line exits  $\llbracket 1, n \rrbracket \times (0, b]$  either by a sink, either by the side  $\llbracket 1, n \rrbracket \times \{b\}$ . This important remark is the key point to allow easy counting of the Hammersley's lines. There are as many lines coming out of the box on the left as there are sinks, that is to say  $X_0^{[\lambda]}(b)$ , and there are as many lines coming out of the box on the top as there are lines running through one of the points  $(1, b), \dots, (n, b)$ , that is to say  $\sum_{k=1}^n N_k^{[\lambda]}(b)$ . Other way of counting: the number of lines coming out of the box on the right is  $X_n^{[\lambda]}(b)$  by definition, and there are as many lines coming out of the box on the bottom as there are sources, that is to say  $\sum_{k=1}^n G_k^{[\lambda]}$ .

For the first equality, the process  $X_0^{[\lambda]}(b)$  is a Poisson process with intensity  $\lambda$ , hence the variable  $X_0^{[\lambda]}(b)$  follows a Poisson distribution of parameter  $\lambda b$ . Moreover, Lemma 3.2 ensures that the variables  $(N_1^{[\lambda]}(b), \dots, N_n^{[\lambda]}(b))$  are independent geometric random variables with parameter  $(1 - \lambda^{-1})$ , so the variable  $\sum_{k=1}^n N_k^{[\lambda]}(b)$  follows a negative binomial distribution of parameters  $n$  and  $1 - \lambda^{-1}$ . For the second equality, Lemma 3.2 ensures that the variable  $X_0^{[\lambda]}(b)$  follows a Poisson distribution of parameter  $\lambda b$ . Since the random variables  $(G_k^{[\lambda]})_{k \geq 1}$  are independent geometric random variables with parameter  $(1 - \lambda^{-1})$ , the variable  $\sum_{k=1}^n G_k^{[\lambda]}$  also follows a negative binomial distribution of parameters  $n$  and  $1 - \lambda^{-1}$ .  $\square$

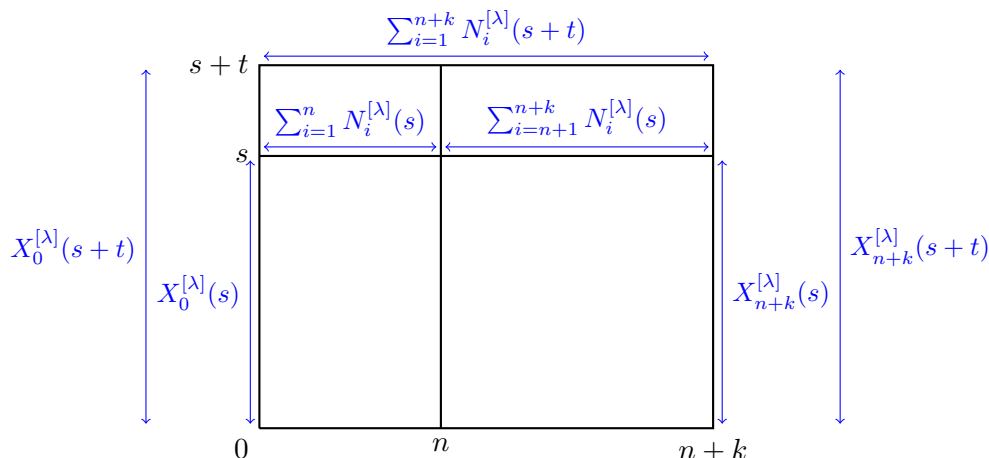
This link makes it possible to get the following stationarity statement, which is the analogue of Theorem 3.1 for this discrete model:

**Lemma 3.4.** *For any  $(n, s) \in \mathbb{N} \times \mathbb{R}_+$ , the distribution of the process*

$$\left( L_{n+k, s+t}^{[\lambda]} - L_{n, s}^{[\lambda]} \right)_{(k, t) \in \mathbb{N} \times \mathbb{R}_+}$$

*does not depend on the couple  $(n, s)$ .*

*Proof.* The proof of this proposition heavily relies the two ways to count Hammersley's lines in a given box stated in Lemma 3.3. We show that this process indeed represents the number of Hammersley's lines crossing the box  $\llbracket n, n+k \rrbracket \times [s, s+t]$  for  $(k, t) \in \mathbb{N} \times \mathbb{R}_+$ , whose dimensions



**Figure 3.8:** The random variables used to express the variables  $L_{n,s}^{[\lambda]}$  and  $L_{n+k,s+t}^{[\lambda]}$ .

do not depend on the couple  $(n, s)$ . The random variables at stake here are summarized in Figure 3.8

Let  $(k, t) \in \mathbb{N} \times \mathbb{R}_+$ . By using Lemma 3.3 twice, we get:

$$L_{n+k,s+t}^{[\lambda]} - L_{n,s}^{[\lambda]} = \left( X_0^{[\lambda]}(s+t) + \sum_{i=1}^{n+k} N_i^{[\lambda]}(s+t) \right) - \left( X_0^{[\lambda]}(s) + \sum_{i=1}^n N_i^{[\lambda]}(s) \right).$$

But, by counting the number of Hammersley's lines exiting the box  $\llbracket 0, n+k \rrbracket \times \llbracket s, s+t \rrbracket$  in two different ways as in the proof of Lemma 3.3, we get

$$\left( X_0^{[\lambda]}(s+t) - X_0^{[\lambda]}(s) \right) + \sum_{i=1}^{n+k} N_i^{[\lambda]}(s+t) = \sum_{i=1}^{n+k} N_i^{[\lambda]}(s) + \left( X_{n+k}^{[\lambda]}(s+t) - X_{n+k}^{[\lambda]}(s) \right).$$

Consequently,

$$L_{n+k,s+t}^{[\lambda]} - L_{n,s}^{[\lambda]} = \sum_{i=n+1}^{n+k} N_i^{[\lambda]}(s) + \left( X_{n+k}^{[\lambda]}(s+t) - X_{n+k}^{[\lambda]}(s) \right).$$

In other words, the variable  $L_{n+k,s+t}^{[\lambda]} - L_{n,s}^{[\lambda]}$  is equal to the number of Hammersley's lines crossing the box  $\llbracket n, n+k \rrbracket \times \llbracket s, s+t \rrbracket$ .

Moreover, by Lemma 3.2, the process  $\left( \sum_{i=n+1}^{n+k} N_i^{[\lambda]}(s) \right)_{k \in \mathbb{N}^*}$  has the same distribution as the process  $\left( \sum_{i=1}^k G_i^{[\lambda]} \right)_{k \in \mathbb{N}^*}$ . Similarly, by Lemma 3.2, the process  $\left( X_n^{[\lambda]}(s+t) - X_n^{[\lambda]}(s) \right)_{t \in \mathbb{R}_+}$  has the same distribution as the process  $\left( X_0^{[\lambda]}(s+t) - X_0^{[\lambda]}(s) \right)_{t \in \mathbb{R}_+}$ . Besides, by construction, the independence of the process  $(X_n^{[\lambda]}(u))_{0 \leq u \leq s}$  and the variable  $\mathbf{N}_n^{[\lambda]}(s)$  ensures that the processes  $\left( X_n^{[\lambda]}(s+t) - X_n^{[\lambda]}(s) \right)_{t \in \mathbb{R}_+}$  and  $\left( X_n^{[\lambda]}(s+t) - X_n^{[\lambda]}(s) \right)_{t \in \mathbb{R}_+}$  are independent.



Consequently, the distribution of the process  $\left(L_{n+k,s+t}^{[\lambda]} - L_{n,s}^{[\lambda]}\right)_{(k,t) \in \mathbb{N} \times \mathbb{R}_+}$  which represents the number of Hammersley's lines crossing the box  $\llbracket n, n+k \rrbracket \times [s, s+t]$  for  $(k, t) \in \mathbb{N} \times \mathbb{R}_+$  is the same as the one of the process  $\left(L_{k,t}^{[\lambda]}\right)_{(k,t) \in \mathbb{N} \times \mathbb{R}_+}$  which represents the number of Hammersley's lines crossing the box  $\llbracket 0, k \rrbracket \times [0, t]$ , and thus does not depend on the couple  $(n, s)$ .  $\square$

The goal is now to transpose this Proposition to the continuous model, this is the purpose of the following section.

### 3.4 From discrete to continuous: proof of Theorem 3.1 and Proposition 3.1

This section makes the link between the discrete model introduced in section 3.3 and the continuous model introduced in section 3.2. Informally, the time argument will be sent to infinity such that, with a suitable scaling, the Poisson processes will converge to Brownian motions. The key argument which then guarantees that the stationarity theorem still holds after taking the limit is the **continuity** of the functional  $L_{n,u}^*$  defined by equation (3.2). Notice that again with scaling arguments, it suffices to study the functional  $L_n^* = L_{n,1}^*$  to get information about the functional  $L_{n,u}^*$  for a positive  $u$ , that is why we take  $u = 1$  in the first place.

Let us start by properly scaling the processes. Consider a positive real  $b$ , the variable which will tend to infinity. The Poisson processes  $(X_k)_{k \in \llbracket 1, n \rrbracket}$  will provide Brownian motions, and we want the Poisson process  $X_0^{[\lambda]}$  to provide a drifted Brownian motion with a given drift  $c > 0$ . To do so, we need to scale the parameter  $\lambda$  with the variable  $b$ , and it will be proved that this is done by setting

$$\lambda = \lambda(c, b) := 1 + \frac{c}{\sqrt{b}}. \quad (3.7)$$

Formally, for any integer  $k \in \llbracket 1, n \rrbracket$  and reals  $c \in \mathbb{R}_+^*$ ,  $b \in \mathbb{R}_+^*$ ,  $t \in \mathbb{R}_+$ , we set

$$Y_{k,b}(t) := \frac{X_k(tb) - tb}{\sqrt{b}} \quad \text{and} \quad Y_{0,b}^{(c)}(t) := \frac{X_0^{[\lambda(c,b)]}(tb) - tb}{\sqrt{b}}.$$

We will prove that the processes  $Y_{k,b}$  and  $Y_{0,b}^{(c)}$  converge in distribution as  $b$  tends to infinity to standard Brownian motions without and with drift  $c > 0$  respectively. Similarly, we also need to rescale the source process  $(G_k^{[\lambda(c,b)]})$  by the same amount  $\sqrt{b}$  as the other processes. So we set for any integer  $k \in \llbracket 1, n \rrbracket$  and reals  $c \in \mathbb{R}_+^*$ ,  $b \in \mathbb{R}_+^*$ ,

$$H_{k,b}^{(c)} := \frac{G_k^{[\lambda(c,b)]}}{\sqrt{b}}.$$

We will prove that the family of variables  $(H_{k,b}^{(c)})_k$  converges in distribution as  $b$  tends to infinity to independent exponential random variables with mean  $c^{-1}$ .

To prove Theorem 3.1 we introduce the following functional, for  $(n, k, s, t) \in \mathbb{N}^2 \times \mathbb{R}_+^2$ ,

$$\Delta_{n,k,s,t} \left( E^{(c)}, B_0^{(c)}, B \right) = L_{n+k,s+t}^* \left( E^{(c)}, B_0^{(c)}, B \right) - L_{n,s}^* \left( E^{(c)}, B_0^{(c)}, B \right).$$

This theorem will follow from Lemma 3.4, using Lemma 3.5 below which states a convergence result which makes the link between the two results. The functionals  $L_n^*$  and  $\Delta_{n,k,s,t}$  are indeed considered with the scaled processes defined above, and the parameter  $b$  is sent to infinity.

**Lemma 3.5.** *Set  $H_{\cdot,b}^{(c)} := (H_{k,b}^{(c)})_{1 \leq k \leq n}$  and  $Y_{\cdot,b} := (Y_{k,b})_{1 \leq k \leq n}$ . As  $b$  goes to infinity, the following convergences in distribution hold:*

$$L_n^*(H_{\cdot,b}^{(c)}, Y_{0,b}^{(c)}, Y_{\cdot,b}) \xrightarrow{b \rightarrow \infty} L_n^*(E^{(c)}, B_0^{(c)}, B) \quad (3.8)$$

$$\Delta_{n,k,s,t}(H_{\cdot,b}^{(c)}, Y_{0,b}^{(c)}, Y_{\cdot,b}) \xrightarrow{b \rightarrow \infty} \Delta_{n,k,s,t}(E^{(c)}, B_0^{(c)}, B) \quad (3.9)$$

*Proof.* We start by showing the following joint convergence in distribution of the processes:

$$(H_{\cdot,b}^{(c)}, Y_{0,b}^{(c)}, Y_{\cdot,b}) \xrightarrow{b \rightarrow \infty} (E^{(c)}, B_0^{(c)}, B), \quad (3.10)$$

and then we will show that the the functionals  $L_n^*$  and therefore  $\Delta_{n,k,s,t}$  are continuous.

To prove the joint convergence (3.10), it suffices to show the convergence of each marginal process because all the processes are independent.

On the one hand, it follows from relation (3.7) that the sequence  $(H_{k,b}^{(c)})_k$  converges almost surely to (independent) exponential random variables with mean  $c^{-1}$  as  $b$  goes to infinity.

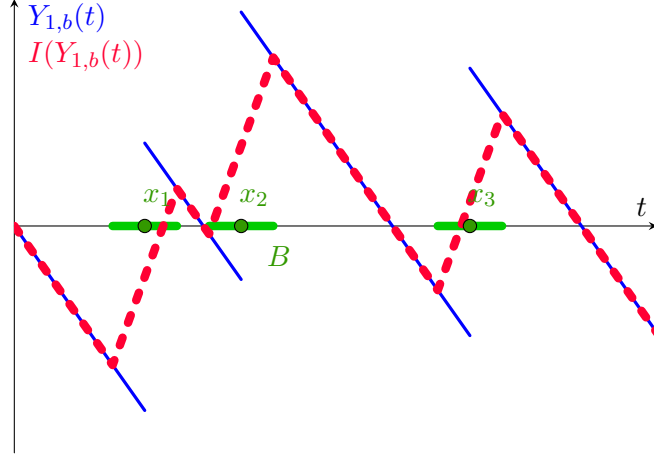
On the other hand, the processes  $(X_k)$  are independent unit intensity Poisson processes so the processes  $(Y_{k,b})$  converge in Skorohod space to independent standard Brownian motions. Furthermore, we can decompose the process  $X_0^{[\lambda(b,c)]}$  which is a Poisson process with intensity  $1 + \frac{c}{\sqrt{b}}$  as a sum of two independent Poisson processes  $X_0$  and  $X'_0$  with intensity 1 and  $\frac{c}{\sqrt{b}}$  respectively. The first one rescales as a standard Brownian motion  $B_0$ , and the second term provides the linear drift almost surely. We eventually get that the process  $Y_{0,b}^{(c)}$  converges in distribution to the drifted Brownian motion  $B_0^{(c)}$ , and the joint convergence (3.10) is proved.

As mentioned above, the joint convergence of the processes makes it possible to get the convergence of the functional, using a continuity argument.

Notice that the processes  $Y_{0,b}^{(c)}, Y_{1,b}, \dots$  are not continuous almost surely. Since it is more convenient to work with continuous functions, we start by defining an interpolation functional. Even if they are not continuous, the above processes are almost surely càdlàg functions with only finitely many discontinuities. For such a function  $f$ , we define another function  $I(f)$  which is continuous and close enough to  $f$ .

More precisely, we can define the function  $I(f)$  in the following way. If we denote by  $(x_1, \dots, x_k)$  the jump points of  $f$ , set

$$\delta := \min_{i \neq j} |x_i - x_j| \quad \text{and} \quad B := \bigcup_{i=1}^n \left[ x_i \pm \frac{\delta}{3} \right].$$



**Figure 3.9:** Example of interpolation of a typical function  $Y_{1,b}$

Define the function  $I(f)$  as being equal to  $f$  on  $\mathbb{R} \setminus B$ , and to the linear interpolation of the points of  $\partial B \times f(\partial B)$  on each connected component of  $B$ . An example of this interpolation of a typical function  $Y_{1,b}$  is shown in Figure 3.9.

Note that the processes  $Y_{0,b}^{(c)}$  and  $Y_{1,b}, \dots, Y_{n,b}$  are almost surely piecewise linear functions, with always the same slope  $-\sqrt{b}$ , and which can only make positive jumps of height  $\frac{1}{\sqrt{b}}$ . Consequently we have almost surely

$$\forall f \in \{Y_{0,b}^{(c)}, Y_{1,b}, \dots, Y_{n,b}\}, \forall t \in [0, 1], \quad |f(t) - I(f(t))| \leq \frac{1}{2\sqrt{b}}.$$

so

$$\left| L_n^* \left( H_{\cdot,b}^{(c)}, Y_{0,b}^{(c)}, Y_{\cdot,b} \right) - L_n^* \left( H_{\cdot,b}^{(c)}, I(Y_{0,b}^{(c)}), I(Y_{\cdot,b}) \right) \right| \leq \frac{n}{2\sqrt{b}} \xrightarrow{b \rightarrow \infty} 0, \quad (3.11)$$

where  $I(Y_{\cdot,b}) := (I(Y_{1,b}), \dots, I(Y_{n,b}))$ .

Moreover, notice that because the processes  $B_0^{(c)}, B_1, \dots, B_n$  are almost surely continuous, the joint convergence (3.10) implies the following convergence in distribution:

$$\left( H_{\cdot,b}^{(c)}, I(Y_{0,b}^{(c)}), I(Y_{\cdot,b}) \right) \xrightarrow{b \rightarrow \infty} \left( E^{(c)}, B_0^{(c)}, B \right).$$

Now let us show that the functional  $L_n^*$  is continuous on the product space  $\mathbb{R}^n \times \mathcal{C} \times \mathcal{C}^n$  equipped with the uniform product topology, where  $\mathcal{C}$  denotes the set of all continuous real-valued functions defined on  $\mathbb{R}_+$ . Let  $\varepsilon > 0$ ,  $(S_1, \dots, S_n, T_1, \dots, T_n) \in (\mathbb{R}^n)^2$ ,  $(f_0, f_1, \dots, f_n, g_0, g_1, \dots, g_n) \in (\mathcal{C} \times \mathcal{C}^n)^2$ , such that

$$\|(S, f_0, f) - (T, g_0, g)\| \leq \varepsilon.$$

Let  $(i, t_0, \dots, t_{n-1})$  be the index defining the longest path in the definition of  $L_n^*(S, f_0, f)$ , that is to say

$$L_n^*(S, f_0, f) = \sum_{k=1}^i S_k + f_0(t_0) + \sum_{k=1}^n [f_k(t_k) - f_k(t_{k-1})].$$

Consequently,

$$\begin{aligned}
 L_n^*(T, g_0, g) &\geq \sum_{k=1}^i T_k + g_0(t_0) + \sum_{k=1}^n [g_k(t_k) - g_k(t_{k-1})] \\
 &\geq \sum_{k=1}^i (S_k - \varepsilon) + f_0(t_0) - \varepsilon + \sum_{k=1}^n [f_k(t_k) - \varepsilon - (g_k(t_{k-1}) + \varepsilon)] \\
 &= L_n^*(S, f_0, f) - (3n + 1)\varepsilon.
 \end{aligned}$$

Conversely, for all  $(j, s_0, \dots, s_{n-1})$  as above,

$$\begin{aligned}
 L_n^*(S, f_0, f) &\geq \sum_{k=1}^j S_k + f_0(s_0) + \sum_{k=1}^n [f_k(s_k) - f_k(s_{k-1})] \\
 &\geq \sum_{k=1}^j (T_k - \varepsilon) + g_0(s_0) - \varepsilon + \sum_{k=1}^n [g_k(s_k) - \varepsilon - (g_k(s_{k-1}) + \varepsilon)] \\
 &= \sum_{k=1}^j T_k + g_0(s_0) + \sum_{k=1}^n [g_k(s_k) - g_k(s_{k-1})] - (3n + 1)\varepsilon.
 \end{aligned}$$

Hence  $L_n^*(T, g_0, g) \leq L_n^*(S, f_0, f) + (3n + 1)\varepsilon$ , so we have

$$|L_n^*(S, f_0, f) - L_n^*(T, g_0, g)| \leq (3n + 1)\varepsilon.$$

So  $L_n^*$  is continuous on  $\mathbb{R}^n \times \mathcal{C} \times \mathcal{C}^n$ . Finally, by writing

$$L_n^*(H_{\cdot,b}^{(c)}, Y_{0,b}^{(c)}, Y_{\cdot,b}) = \left[ L_n^*(H_{\cdot,b}^{(c)}, Y_{0,b}^{(c)}, Y_{\cdot,b}) - L_n^*(H_{\cdot,b}^{(c)}, I(Y_{0,b}^{(c)}), I(Y_{\cdot,b})) \right] + L_n^*(H_{\cdot,b}^{(c)}, I(Y_{0,b}^{(c)}), I(Y_{\cdot,b})),$$

convergence (3.8) can be deduced using equation (3.11) and the continuity of  $L_n^*$ . Convergence (3.9) follows similarly, using the continuity of  $\Delta_{n,k,s,t}$  and the same interpolation arguments.  $\square$

We can now use Lemma 3.5 to transpose the stationarity result stated by Lemma 3.4 in the discrete model into the continuous model, that is to say to get Theorem 3.1.

*Proof of Theorem 3.1.* With the notation introduced above, we have

$$L_{n+k,s+t}^{(c)} - L_{n,s}^{(c)} = \Delta_{n,k,s,t}(E^{(c)}, B_0^{(c)}, B).$$

Moreover, the following scaling relation holds for  $L_{n,t}^{[\lambda]}$ ,

$$L_{n,t}^*(H_{\cdot,b}^{(c)}, Y_{0,b}^{(c)}, Y_{\cdot,b}) = \frac{L_{n,tb}^{[\lambda]} - tb}{\sqrt{b}},$$

and induces a similar one for the difference  $L_{n+k,s+t}^{[\lambda]} - L_{n,s}^{[\lambda]}$ :

$$\Delta_{n,k,s,t}(H_{\cdot,b}^{(c)}, Y_{0,b}^{(c)}, Y_{\cdot,b}) = \frac{L_{n+k, sb+tb}^{[\lambda]} - L_{n, sb}^{[\lambda]} - tb}{\sqrt{b}}.$$

By Lemma 3.4 the distribution of the variable  $L_{n+k, sb+tb}^{[\lambda]} - L_{n, sb}^{[\lambda]}$  does not depend on the couple  $(n, s)$ , thus convergence (3.9) makes it possible to get Theorem 3.1.  $\square$

On the other hand, Proposition 3.1 is deduced from Lemma 3.3, also using Lemma 3.5.

*Proof of Proposition 3.1.* Recall that in a first step, we prove the proposition for  $u = 1$  and then deduce the general result for any positive  $u$ .

The main idea of this proof is to compute the limit of the quantity  $L_n^*(H_{\cdot, b}^{(c)}, Y_0^{(c)}, Y_{\cdot, b})$  as  $b$  goes to infinity in two different ways. The first one is presented in Lemma 3.5 above. The second one is detailed below, and uses Lemma 3.3 to write this quantity as a sum of two random variables which converge.

According to Lemma 3.3, we can rewrite the quantity  $L_n^*(H_{\cdot, b}^{(c)}, Y_0^{(c)}, Y_{\cdot, b})$  as a sum of two terms :

$$L_n^*(H_{\cdot, b}^{(c)}, Y_{0, b}^{(c)}, Y_{\cdot, b}) = Y_{0, b}^{(c)}(1) + \sum_{k=1}^n \frac{N_k^{[\lambda]}(b)}{\sqrt{b}}. \quad (3.12)$$

We have to ensure that the couple made of the two terms of the sum of the right hand side converges in distribution. To do so, let us consider the following functional with the same notation as above

$$F_n : (S, f_0, f) \mapsto (f_0(1), L_n^*(S, f_0, f) - f_0(1)).$$

Notice that by equation (3.12),

$$F_n(H_{\cdot, b}^{(c)}, Y_0^{(c)}, Y_{\cdot, b}) = \left( Y_{0, b}^{(c)}(1), L_n^*(H_{\cdot, b}^{(c)}, Y_0^{(c)}, Y_{\cdot, b}) - Y_{0, b}^{(c)}(1) \right) = \left( Y_{0, b}^{(c)}(1), \sum_{k=1}^n \frac{N_k^{[\lambda]}(b)}{\sqrt{b}} \right).$$

It follows from the proof of Lemma 3.5 that the functional  $F_n$  is continuous on  $\mathbb{R}^n \times \mathcal{C} \times \mathcal{C}^n$ . If we denote  $Z_1^{(c)} = B_0^{(c)}(1)$ , a random variable with distribution  $\mathcal{N}(c, 1)$ , and if we set

$$\Gamma_n^{(c)} := L_n^*(E^{(c)}, B_0^{(c)}, B) - Z_1^{(c)},$$

we have the following convergence in distribution as  $b$  goes to infinity again using similar interpolation arguments:

$$\left( Y_{0, b}^{(c)}(1), \sum_{k=1}^n \frac{N_k^{[\lambda]}(b)}{\sqrt{b}} \right) \xrightarrow{b \rightarrow \infty} (Z_1^{(c)}, \Gamma_n^{(c)}). \quad (3.13)$$

As mentioned in Lemma 3.3, the sum  $\sum_{k=1}^n \frac{N_k^{[\lambda]}(b)}{\sqrt{b}}$  is a sum of  $n$  i.i.d. geometric random variables with mean  $\frac{1}{\lambda-1} = \frac{\sqrt{b}}{c}$ , using equation (3.7). Hence the variable  $\sum_{k=1}^n \frac{N_k^{[\lambda]}(b)}{\sqrt{b}}$  converges to a sum of  $n$  i.i.d. geometric exponential variables with mean  $c^{-1}$ , that is to say a variable with a  $\Gamma(n, c)$  distribution. Consequently, this distribution is the one of the variable  $\Gamma_n^{(c)}$  by convergence (3.13).

Using Lemma 3.5 and equation (3.12), we get the following convergence in distribution:

$$Y_{0,b}^{(c)}(1) + \sum_{k=1}^n \frac{N_k^{[\lambda]}(b)}{\sqrt{b}} \xrightarrow{b \rightarrow \infty} L_n^*(E^{(c)}, B_0^{(c)}, B).$$

But using equation (3.13) we get the following one:

$$Y_{0,b}^{(c)}(1) + \sum_{k=1}^n \frac{N_k^{[\lambda]}(b)}{\sqrt{b}} \xrightarrow{b \rightarrow \infty} Z_1^{(c)} + \Gamma_n^{(c)}.$$

Hence the two right-hand-sides have the same distribution, that is to say

$$L_n^{(c)} \stackrel{(d)}{=} Z_1^{(c)} + \Gamma_n^{(c)}. \quad (3.14)$$

Let us now deduce Proposition 3.1 for any positive  $u$ . We prove that the distribution of the variable  $L_{n,u}^{(c)}$  can be linked to the one of the variable  $L_{n,1}^{(c')}$  for a well-chosen parameter  $c'$ . More precisely,

$$L_{n,u}^{(c)} \stackrel{(d)}{=} \sqrt{u} L_{n,1}^{(c\sqrt{u})}. \quad (3.15)$$

Indeed, if we set for all integer  $k \in \llbracket 0, n \rrbracket$ ,

$$\widehat{B}_k(t) := \frac{1}{\sqrt{u}} B_k(ut) \quad ; \quad \widehat{B}_0^{(c)}(t) := \frac{1}{\sqrt{u}} B_0^{(c)}(ut) = \widehat{B}_0(t) + c\sqrt{ut} \quad ; \quad \widehat{E}_k^{(c)} = \frac{E_k^{(c)}}{\sqrt{u}},$$

then  $(\widehat{B}_k)$  are standard Brownian motions,  $\widehat{B}_0^{(c)}$  is a standard Brownian motion with drift  $c\sqrt{u}$ , and  $(\widehat{E}_k^{(c)})$  are exponential random variables with mean  $1/(c\sqrt{u})$ . The independence is of course preserved. Therefore, almost surely,

$$L_{n,u}^*(E^{(c)}, B_0^{(c)}, B) = \sqrt{u} L_{n,1}^*(\widehat{E}^{(c)}, \widehat{B}_0^{(c)}, \widehat{B}),$$

and the equality in distribution (3.15) actually holds.

Consequently, equation (3.14) provides information on the distribution of the variable  $L_{n,u}^{(c)}$ : there exists a random variable  $\Gamma_n^{(c)}$  with distribution  $\Gamma(n, c)$  such that

$$L_{n,u}^{(c)} \stackrel{(d)}{=} B_0^{(cu)}(u) + \Gamma_n^{(c)}. \quad \square$$

Theorem 3.1 and Proposition 3.1 are now proved for any drift parameter  $c > 0$ . In the next section, we are interested in the particular case when this parameter scales with  $n$ , in fact  $c = \sqrt{n}$ .

### 3.5 Localisation of the optimal path when $c = \sqrt{n}$ : proof of Proposition 3.2

We start by recalling how the particular case  $c = \sqrt{n}$  arises. The original motivation is to exploit the results about the stationary model to get information on the original model defined by equation (3.1). As explained in section 3.2, the stochastic domination given by equation (3.4) holds.

Proposition 3.1 ensures that the variable  $L_n^{(c)} = L_n^*(E^{(c)}, B_0^{(c)}, B)$  can be written as a sum of a normal random variable and a Gamma random variables. Both are concentrated around their expectation which are respectively  $c$  and  $\frac{n}{c}$ . Hence, to get the tightest bound possible, one can consider the particular case when the expectation of their sum is minimal, that is to say when  $c = \sqrt{n}$ . Notice that in this case, this expectation is  $2\sqrt{n}$ .

The goal is now to prove that the optimal path in this case are concentrated around the diagonal of the rectangle  $[0, 1] \times [0, n]$ , which is stated in Proposition 3.2. The proof is split into two main parts. The first parts consists in showing that a longest path compelled to pass far from the diagonal cannot have a length close to  $2\sqrt{n}$  with high probability, which is the purpose of Lemma 3.6 below. The second part consists in proving that always with high probability, such a path cannot be the (unconstrained) optimal path, using Proposition 3.1 to prove that the length of the optimal path is close to  $2\sqrt{n}$ .

In spite  $c = \sqrt{n}$  in this case, we keep the parameter  $c$  explicit in the notation since we will use the result of Proposition 3.1 for other values.

We recall the notation

$$L_{n,u}^{(c)} := L_{n,u}^*(E^{(c)}, B_0^{(c)}, B) = \sup_{i, t_0, \dots, t_{n-1}} \left( \sum_{k=1}^i E_k^{(c)} + B_0^{(c)}(t_0) + \sum_{k=1}^n [B_k(t_k) - B_k(t_{k-1})] \right), \quad (3.16)$$

where the supremum is taken over the set

$$\{(i, t_0, \dots, t_{n-1}) : i \in \llbracket 0, n \rrbracket ; \forall j \leq i-1, t_j = 0 ; 0 < t_i < \dots < t_n = u\}.$$

Also recall that with this expression the integer  $i$  corresponds to the number of exponential rewards  $E_k^{(c)}$  gathered by the longest path, and that the parameter  $b$  is taken to be 1 by default as above:  $L_n^{(c)} = L_{n,1}^{(c)}$ .

Let us introduce some notation. For three real numbers  $(a, t, u) \in [0, 1]^3$ ,  $t \leq u$ , define the length  $L_n^{(c)}(a; t, u)$  of the longest path compelled to intersect the vertical segment  $\lfloor an \rfloor \times [t, u]$ , i.e. the same quantity as in equation (3.16) but with the supremum taken over the set

$$\{(i, t_0, \dots, t_{n-1}) : i \in \llbracket 0, n \rrbracket ; \forall j \leq i-1, t_j = 0 ; 0 < t_i < \dots < t_n = 1 ; [t_{\lfloor an \rfloor - 1}, t_{\lfloor an \rfloor}] \cap [t, u] \neq \emptyset\}.$$

Define similarly the length  $L_n^{(c)}(a; t^+, u)$  of the longest path compelled to intersect the half-open vertical segment  $\lfloor an \rfloor \times (t, u]$ , and the length  $L_n^{(c)}(a; t) := L_n^{(c)}(a; t, t)$  of the longest path

compelled to pass exactly through the point  $(\lfloor an \rfloor, t)$ . Define also the length  $L_n^{(c)}(a; 0^+)$  of the longest path compelled to intersect every half-open vertical segment of the form  $\{\lfloor an \rfloor\} \times (0, \varepsilon]$  for a positive  $\varepsilon$ , i.e. the same quantity as in equation (3.16) but with the supremum taken over the set

$$\{(i, t_0, \dots, t_{n-1}) : i = \lfloor an \rfloor ; \forall j \leq i-1, t_j = 0 ; 0 < t_i < \dots < t_n = 1\}.$$

As mentioned above, the following lemma states that a longest path compelled to pass far from the diagonal cannot have a length close to  $2\sqrt{n}$  with high probability. The result is split into two concentration inequalities, with a path compelled to pass far above or far below the diagonal.

**Lemma 3.6.** *There exists a positive constant  $C$  such that for any real  $\beta > 4$  and for any integer  $n$ , if we set  $\eta_n = \frac{1}{n^{1/\beta}}$ , then*

$$\begin{cases} \sup_{a \in [0, 1 - \eta_n)} \mathbb{P}\left(L_n^{(\sqrt{n})}(a; a + \eta_n, 1) \geq (2 - C\eta_n^2)\sqrt{n}\right) = o\left(\frac{1}{n}\right) \\ \sup_{a \in (\eta_n, 1]} \mathbb{P}\left(L_n^{(\sqrt{n})}(a; 0, a - \eta_n) \geq (2 - C\eta_n^2)\sqrt{n}\right) = o\left(\frac{1}{n}\right) \end{cases}$$

To prove Lemma 3.6 we need some concentration arguments. Let us state Chernoff bounds for the Gamma distribution in the following tool lemma.

**Lemma 3.7.** *If  $0 \leq t \leq \frac{p}{\lambda}$ ,*

$$\begin{cases} \mathbb{P}\left(\Gamma(p, \lambda) \geq \frac{p}{\lambda} + t\right) \leq \exp\left(-\frac{\lambda^2 t^2}{4p}\right), \\ \mathbb{P}\left(\Gamma(p, \lambda) \leq \frac{p}{\lambda} - t\right) \leq \exp\left(-\frac{\lambda^2 t^2}{4p}\right). \end{cases}$$

*Proof.* We proceed similarly to [BLM13, Sections 2.2 and 2.4].

Let  $(X_1, \dots, X_p)$  be i.i.d exponential random variables with mean  $\lambda^{-1}$ , so that the sum  $\sum_{i=1}^p X_i$  follows the distribution  $\Gamma(p, \lambda)$ .

Set  $Z = \sum_{i=1}^p X_i - \frac{p}{\lambda}$ , and denote by  $\Psi_X$  and  $\Psi_Z$  the Laplace transform of  $X_1$  and  $Z$  respectively. We will apply Chernoff inequality to the variable  $Z$ , so we start by upper-bounding  $\Psi_Z$ :

$$\forall \alpha, \quad \Psi_Z(\alpha) := \mathbb{E}[e^{\alpha Z}] = \mathbb{E}\left[\exp\left(\alpha X_1 - \frac{\alpha}{\lambda}\right)\right]^p = \Psi_X(\alpha)^p \exp\left(-\frac{p\alpha}{\lambda}\right).$$

Recall that for any real  $\alpha < \lambda$ ,  $\Psi_X(\alpha) = (1 - \frac{\alpha}{\lambda})^{-1}$ , and notice that for any real  $u \leq \frac{1}{2}$ ,  $-\ln(1 - u) \leq u + u^2$ . Then for  $\alpha \leq \frac{\lambda}{2}$ ,  $\Psi_X(\alpha) \leq \exp\left(\frac{\alpha}{\lambda} + \frac{\alpha^2}{\lambda^2}\right)$ . So,

$$\forall \alpha \leq \frac{\lambda}{2}, \quad \Psi_Z(\alpha) = \Psi_X(\alpha)^p \exp\left(-\frac{p\alpha}{\lambda}\right) \leq \exp\left(\frac{p\alpha^2}{\lambda^2}\right).$$

We use this upper bound to show both inequalities.



- For  $\alpha \in [0, \frac{\lambda}{2}]$ , we apply Markov's inequality and the upper-bound above:

$$\mathbb{P}(Z \geq t) \leq \Psi_Z(\alpha)e^{-\alpha t} \leq \exp\left(\frac{p\alpha^2}{\lambda^2} - \alpha t\right).$$

By taking  $\alpha = \frac{\lambda^2 t}{2p} \in [0, \frac{\lambda}{2}]$  since  $t \in [0, \frac{p}{\lambda}]$ , we eventually get  $\mathbb{P}(Z \geq t) \leq \exp\left(-\frac{\lambda^2 t^2}{4p}\right)$ .

- Similarly, for  $\alpha \geq 0$ :

$$\mathbb{P}(Z \leq -t) \leq \Psi_Z(-\alpha)e^{-\alpha t} \leq \exp\left(\frac{p\alpha^2}{\lambda^2} - \alpha t\right).$$

We conclude by taking the same value for  $\alpha$ . □

We can now prove Lemma 3.6:

*Proof of Lemma 3.6.* The two parts are proved in a similar way, and rely on a subdivision of the vertical segment  $\{\lfloor an \rfloor\} \times [a + \eta_n, 1]$  or  $\{\lfloor an \rfloor\} \times [0, a - \eta_n]$  into a finite number of segments, in order to control the length of a path which passes through one of them. This is done thanks to a suitable stochastic domination which will be presented below, see equation (3.17).

- Let us determine an upperbound for the first supremum for a fixed integer  $n$ , and assume without loss of generality that the real number  $an$  is an integer to simplify the writing. First, consider the supremum on the interval  $[\frac{1}{n}, 1 - \eta_n)$ , that is to say assume the integer  $an$  is non-zero. Set  $t = a + \eta_n < 1$  and let  $u \in [t, 1]$ . The optimal path compelled to intersect the segment  $\{an\} \times [t, u]$  passes through at least one of the points of this segment, say  $(an, u^*)$ . We divide the path from  $(0, 0)$  to  $(n, 1)$  into two parts : the first from  $(0, 0)$  to  $(an, u^*)$  and the second from  $(an, u^*)$  to  $(n, 1)$ . Notice that since  $an > 0$  by assumption and  $t > 0$ , the second part of the path passes neither through sources nor through sinks. We can deduce that the following stochastic domination holds:

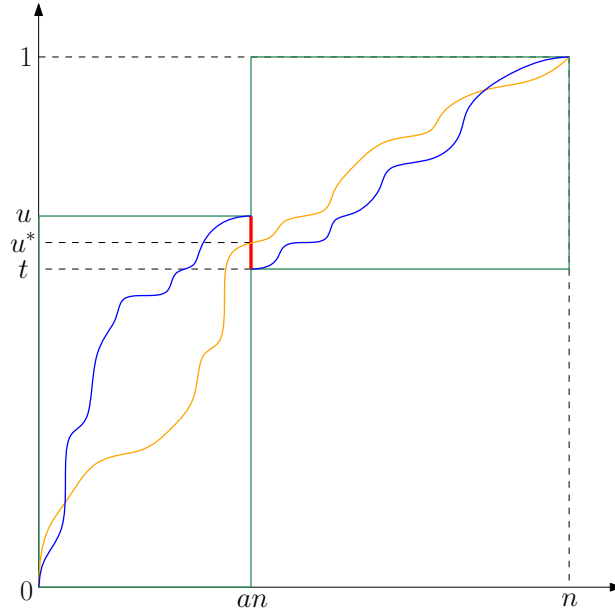
$$L_n^{(\sqrt{n})}(a; t, u) \leq L_{an, u^*}^{(\sqrt{n})} + L_{(1-a)n, 1-u^*}. \quad (3.17)$$

This inequality is indeed the heart of this proof and is illustrated in Figure 3.10. Consequently, the following stochastic dominations also hold:

$$\begin{aligned} L_n^{(\sqrt{n})}(a; t, u) &\leq L_{an, u}^{(\sqrt{n})} + L_{(1-a)n, 1-t} \quad \text{by monotonicity, since } t \leq u^* \leq u \\ &\leq L_{an, u}^{(\sqrt{n})} + L_{(1-a)n, 1-t}^{(\alpha\sqrt{n})} \quad \text{for all positive } \alpha. \end{aligned}$$

From now on, let us take  $\alpha = \sqrt{\frac{1-a}{1-t}}$  in order to minimise the right-hand side. By Proposition 3.1, this right-hand side is a sum of four random variables with explicit distribution. The goal is to lower bound the quantity  $(2 - \varepsilon)\sqrt{n}$  for a well-chosen positive  $\varepsilon$  by a sum of four terms which are greater than the four expectations of these variables.

Formally, set  $f : x \mapsto \frac{x^2}{12}$  and take  $\varepsilon = f(t - a) > 0$  since  $a \neq t$ . Now take  $u = (t + \varepsilon) \wedge 1$ , and notice that



**Figure 3.10:** A schematic illustration of equation (3.17). The optimal path from  $(0,0)$  to  $(n,1)$  compelled to pass through the segment  $\{an\} \times [t,u]$  (in red) is drawn in orange. The optimal paths in the rectangles  $[0,an] \times [0,u]$  and  $[an,n] \times [t,1]$  are drawn in blue.

$$\begin{aligned}
 2 - \varepsilon &\geq 2 - \frac{(t-a)^2}{12(1-a)} = 2(1-a) \left[ 1 - \frac{(t-a)^2}{8(1-a)^2} \right] + 2a + \frac{(t-a)^2}{6(1-a)} \\
 &\geq 2(1-a) \left[ 1 - \frac{(t-a)^2}{8(1-a)^2} \right] + 2a + 2\varepsilon \\
 &\geq 2(1-a) \left[ \sqrt{1 - \frac{t-a}{1-a}} + \frac{t-a}{2(1-a)} \right] + 2a + 2\varepsilon \quad \text{since } 1 - \frac{x^2}{8} \geq \sqrt{1-x} + \frac{x}{2} \text{ for } x \in [0,1] \\
 &= 2\sqrt{1-a}\sqrt{1-t} + t + a + 2\varepsilon \geq 2\sqrt{1-a}\sqrt{1-t} + u + a + \varepsilon \quad \text{since } u \leq t + \varepsilon \\
 &= (u + \delta) + (a + \delta) + (\alpha(1-t) + \delta) + \left( \frac{1-a}{\alpha} + \delta \right) \quad \text{with } \delta := \frac{\varepsilon}{4}.
 \end{aligned}$$

Hence,

$$\begin{aligned}
 \mathbb{P}\left(L_n^{(\sqrt{n})}(a; t, u) \geq (2 - \varepsilon)\sqrt{n}\right) &\leq \mathbb{P}\left(L_{an,u}^{(\sqrt{n})} + L_{(1-a)n,1-t}^{(\alpha\sqrt{n})} \geq (2 - \varepsilon)\sqrt{n}\right) \\
 &\leq \textcircled{1} + \textcircled{2} + \textcircled{3} + \textcircled{4},
 \end{aligned}$$

with, using Proposition 3.1:

$$\textcircled{1} = \mathbb{P}\left(\mathcal{N}(u\sqrt{n}, u) \geq (u + \delta)\sqrt{n}\right) \leq \exp\left(-\frac{\delta^2}{2u}n\right)$$

$$\textcircled{2} = \mathbb{P}\left(\Gamma(an, \sqrt{n}) \geq (a + \delta)\sqrt{n}\right) \leq \exp\left(-\frac{\delta^2}{4a}n\right)$$

$$\textcircled{3} = \mathbb{P}\left(\mathcal{N}(\alpha(1-t)\sqrt{n}, 1-t) \geq (\alpha(1-t)\delta)\sqrt{n}\right) \leq \exp\left(-\frac{\delta^2}{2(1-t)}n\right)$$

$$\textcircled{4} = \mathbb{P}\left(\Gamma((1-a)n, \alpha\sqrt{n}) \geq \left(\frac{1-a}{\alpha} + \delta\right)\sqrt{n}\right) \leq \exp\left(-\frac{\delta^2}{4(1-t)}n\right),$$

where we have used Chernoff bounds provided by Lemma 3.7 for Gamma Distribution, and by [BLM13, p22] for Normal distribution. These Chernoff bounds will be reused several times afterwards in a similar way.

Finally, since the parameters  $a$ ,  $t$  and  $u$  live in the interval  $[0, 1]$ ,

$$\mathbb{P}\left(L_n^{(\sqrt{n})}(a; t, (t+\varepsilon) \wedge 1) \geq (2-\varepsilon)\sqrt{n}\right) \leq 4 \exp\left(-\frac{\delta^2}{4}n\right),$$

that is to say

$$\mathbb{P}\left(L_n^{(\sqrt{n})}(a; t, (t+f(t-a)) \wedge 1) \geq (2-f(t-a))\sqrt{n}\right) \leq 4 \exp\left(-\frac{f(t-a)^2}{64}n\right). \quad (3.18)$$

Now, set  $u_0 := t$ , and for all integer  $k \geq 0$ ,

$$\varepsilon_k := f(u_k - a) \quad \text{and} \quad u_{k+1} := (u_k + \varepsilon_k) \wedge 1.$$

Equation (3.18) applied to  $u_k$  gives

$$\mathbb{P}\left(L_n^{(\sqrt{n})}(a; u_k, u_{k+1}) \geq (2-\varepsilon_k)\sqrt{n}\right) \leq 4 \exp\left(-\frac{\varepsilon_k^2}{64}n\right).$$

Notice that since the function  $f(\cdot - a)$  is non-decreasing on the interval  $(a, +\infty)$ ,  $\varepsilon_k \geq \varepsilon$  and  $(t+k\varepsilon) \wedge 1 \leq u_k \leq 1$  for all non-negative  $k$ . Hence, the lower bound  $K := \min\{k \geq 0 : u_k = 1\}$  is well defined and  $K \leq \frac{1}{\varepsilon}$ . In other words,  $t = u_0 < u_1 < \dots < u_{K-1} < u_K = 1$  defines a finite subdivision of the segment  $[t, 1]$ . Consequently,

$$L_n^{(\sqrt{n})}(a; t, 1) = \max_{k \in \llbracket 0, K-1 \rrbracket} L_n^{(\sqrt{n})}(a; u_k, u_{k+1}).$$

We can deduce the wanted concentration inequality:

$$\begin{aligned} \mathbb{P}\left(L_n^{(\sqrt{n})}(a; t, 1) \geq (2-\varepsilon)\sqrt{n}\right) &\leq \sum_{k=0}^{K-1} \mathbb{P}\left(L_n^{(\sqrt{n})}(a; u_k, u_{k+1}) \geq (2-\varepsilon)\sqrt{n}\right) \\ &\leq \sum_{k=0}^{K-1} \mathbb{P}\left(L_n^{(\sqrt{n})}(a; u_k, u_{k+1}) \geq (2-\varepsilon_k)\sqrt{n}\right) \\ &\leq 4 \sum_{k=0}^{K-1} \exp\left(-\frac{\varepsilon_k^2}{64}n\right) \leq 4K \exp\left(-\frac{\varepsilon^2}{64}n\right). \end{aligned}$$

Using the bound  $K \leq \frac{1}{\varepsilon}$ , we finally get

$$\mathbb{P}\left(L_n^{(\sqrt{n})}(a; t, 1) \geq (2-f(t-a))\sqrt{n}\right) \leq \frac{4}{f(t-a)} \exp\left(-\frac{f(t-a)^2}{64}n\right).$$

By plugging  $t = a + \eta_n$  and  $f(x) = \frac{x^2}{12}$ , we get the following upper bound:

$$\sup_{a \in [\frac{1}{n}, 1 - \eta_n]} \mathbb{P} \left( L_n^{(\sqrt{n})}(a; a + \eta_n, 1) \geq \left(2 - \frac{\eta_n^2}{12}\right) \sqrt{n} \right) \leq \frac{48}{\eta_n^2} \exp \left( -\frac{\eta_n^4}{9216} n \right). \quad (3.19)$$

Now, let us deal with the case  $\lfloor an \rfloor = 0$ . For  $n$  large enough,  $\lfloor \frac{\eta_n}{2} n \rfloor \geq 1$ . Notice that a non-decreasing path from  $(0, 0)$  to  $(n, 1)$  compelled to intersect the vertical segment  $\{0\} \times [\eta_n, 1]$  intersects necessarily the vertical segment  $\lfloor \frac{\eta_n}{2} n \rfloor \times [\eta_n, 1]$ . Consequently, the following stochastic domination holds:

$$L_n^{(\sqrt{n})}(0; \eta_n, 1) \leq L_n^{(\sqrt{n})}(\frac{\eta_n}{2}; \eta_n, 1),$$

and we can deduce the following concentration inequality in this case by using equation (3.19):

$$\mathbb{P} \left( L_n^{(\sqrt{n})}(0; \eta_n, 1) \geq \left(2 - \frac{\eta_n^2}{48}\right) \sqrt{n} \right) \leq \frac{192}{\eta_n^2} \exp \left( -\frac{\eta_n^4}{147456} n \right).$$

Putting this new bound together with the one of equation (3.19), we eventually get this one for  $n$  large enough:

$$\sup_{a \in [0, 1 - \eta_n]} \mathbb{P} \left( L_n^{(\sqrt{n})}(a; a + \eta_n, 1) \geq \left(2 - \frac{\eta_n^2}{48}\right) \sqrt{n} \right) \leq \frac{192}{\eta_n^2} \exp \left( -\frac{\eta_n^4}{147456} n \right).$$

Recall that  $\eta_n = \frac{1}{n^{1/\beta}}$ , so  $\eta_n^4 n = n^{1-4/\beta}$ . Since  $\beta > 4$ , this supremum is indeed asymptotically dominated by the sequence  $\frac{1}{n}$ .

- Now, assume that  $a > \eta_n$ , set  $t = a - \eta_n$  and let us proceed as above. Let  $u \in [0, t]$ , we are going to establish a concentration inequality on the longest path  $L_n^{(\sqrt{n})}(a; u^+, t)$ . The optimal path compelled to intersect the half-opened segment  $\lfloor an \rfloor \times (u, t]$  passes through at least one of the points of this segment. We divide the path from  $(0, 0)$  to  $(n, 1)$  as above, and we notice that the second part passes neither through sources nor through sinks since the imposed half-opened segment does not contain any point of the axes, even if  $u = 0$ . We get the following stochastic domination for all positive  $\alpha$ :

$$L_n^{(\sqrt{n})}(a; u^+, t) \leq L_{an, t}^{(\sqrt{n})} + L_{(1-a)n, 1-u}^{(\alpha\sqrt{n})}$$

From now on, let us take  $\alpha = \sqrt{\frac{1-a}{1-u}}$  in order to minimise the right-hand side, and set  $\varepsilon = f(a-t) = \frac{(a-t)^2}{12} > 0$  since  $a \neq t$ . Now take  $u = (t - \frac{\varepsilon}{2}) \vee 0$ , and notice that, as above,

$$\begin{aligned} 2 - \varepsilon &\geq 2\sqrt{1-a}\sqrt{1-t} + t + a + 2\varepsilon \\ &\geq 2\sqrt{1-a} \left( \sqrt{1-u} - \frac{t-u}{\sqrt{1-u}} \right) + t + a + 2\varepsilon \quad \text{since } \sqrt{x} \geq \sqrt{y} - \frac{y-x}{\sqrt{y}} \text{ for } y > x \geq 0 \\ &\geq t + a + 2\sqrt{1-a}\sqrt{1-u} - \frac{\sqrt{1-a}}{\sqrt{1-u}} \varepsilon + 2\varepsilon \quad \text{since } 2(t-u) \leq \varepsilon \\ &\geq t + a + 2\sqrt{1-a}\sqrt{1-u} + \varepsilon \quad \text{since } u < a \\ &= (t + \delta) + (a + \delta) + (\alpha(1-u) + \delta) + \left( \frac{1-a}{\alpha} + \delta \right) \quad \text{with } \delta := \frac{\varepsilon}{4}. \end{aligned}$$

Hence,

$$\begin{aligned} \mathbb{P}\left(L_n^{(\sqrt{n})}(a; u^+, t) \geq (2 - \varepsilon)\sqrt{n}\right) &\leq \mathbb{P}\left(L_{an,t}^{(\sqrt{n})} + L_{(1-a)n,1-u}^{(\alpha\sqrt{n})} \geq (2 - \varepsilon)\sqrt{n}\right) \\ &\leq \textcircled{1} + \textcircled{2} + \textcircled{3} + \textcircled{4}, \end{aligned}$$

with (using Proposition 3.1 and Chernoff bounds to upper-bound each quantity):

$$\begin{aligned} \textcircled{1} &= \mathbb{P}\left(\mathcal{N}(t\sqrt{n}, t) \geq (t + \delta)\sqrt{n}\right) \leq \exp\left(-\frac{\delta^2}{2t}n\right) \\ \textcircled{2} &= \mathbb{P}\left(\Gamma(an, \sqrt{n}) \geq (a + \delta)\sqrt{n}\right) \leq \exp\left(-\frac{\delta^2}{4a}n\right) \\ \textcircled{3} &= \mathbb{P}\left(\mathcal{N}(\alpha(1-u)\sqrt{n}, 1-u) \geq (\alpha(1-u) + \delta)\sqrt{n}\right) \leq \exp\left(-\frac{\delta^2}{2(1-u)}n\right) \\ \textcircled{4} &= \mathbb{P}\left(\Gamma((1-a)n, \alpha\sqrt{n}) \geq \left(\frac{1-a}{\alpha} + \delta\right)\sqrt{n}\right) \leq \exp\left(-\frac{\delta^2}{4(1-u)}n\right) \end{aligned}$$

Finally, since the parameters  $a$ ,  $t$  and  $u$  live in the interval  $[0, 1]$ ,

$$\mathbb{P}\left(L_n^{(\sqrt{n})}(a; ((t - \frac{\varepsilon}{2}) \vee 0)^+, t) \geq (2 - \varepsilon)\sqrt{n}\right) \leq 4 \exp\left(-\frac{\delta^2}{4}n\right) = 4 \exp\left(-\frac{\varepsilon^2}{64}n\right). \quad (3.20)$$

Now, set  $u_0 := t$ , and for all integer  $k \geq 0$ ,

$$\varepsilon_k := f(a - u_k) \quad \text{and} \quad u_{k+1} := \left(u_k - \frac{\varepsilon_k}{2}\right) \vee 0.$$

Equation (3.20) applied to  $u_k$  gives

$$\mathbb{P}\left(L_n^{(\sqrt{n})}(a; u_{k+1}, u_k) \geq (2 - \varepsilon_k)\sqrt{n}\right) \leq 4 \exp\left(-\frac{\varepsilon_k^2}{64}n\right).$$

Notice that since the function  $f(a - \cdot)$  is non-increasing on the interval  $(-\infty, a)$ ,  $\varepsilon_k \geq \varepsilon$  and  $0 \leq u_k \leq (t - k\frac{\varepsilon}{2}) \vee 0$  for all non-negative  $k$ . Hence, the lower bound  $K := \min\{k \geq 0 : u_k = 0\}$  is well defined and  $K \leq \frac{2}{\varepsilon}$ . In other words,  $0 = u_K < u_{K-1} < u_{K-2} < \dots < u_0 = t$  defines a finite subdivision of the interval  $(0, t]$ . Consequently,

$$\begin{aligned} \mathbb{P}\left(L_n^{(\sqrt{n})}(a; 0^+, t) \geq (2 - \varepsilon)\sqrt{n}\right) &\leq \sum_{k=0}^{K-1} \mathbb{P}\left(L_n^{(\sqrt{n})}(a; u_{k+1}^+, u_k) \geq (2 - \varepsilon)\sqrt{n}\right) \\ &\leq \sum_{k=0}^{K-1} \mathbb{P}\left(L_n^{(\sqrt{n})}(a; u_{k+1}^+, u_k) \geq (2 - \varepsilon_k)\sqrt{n}\right) \\ &\leq 4 \sum_{k=0}^{K-1} \exp\left(-\frac{\varepsilon_k^2}{64}n\right) \leq 4K \exp\left(-\frac{\varepsilon^2}{64}n\right). \end{aligned}$$

Using the bound  $K \leq \frac{2}{\varepsilon}$ , we finally get

$$\mathbb{P}\left(L_n^{(\sqrt{n})}(a; 0^+, t) \geq (2 - f(a-t))\sqrt{n}\right) \leq \frac{8}{f(a-t)} \exp\left(-\frac{f(a-t)^2}{64}n\right).$$

To get a concentration inequality on the length  $L_n^{(\sqrt{n})}(a; 0, t)$ , we also need some information on the length  $L_n^{(\sqrt{n})}(a; 0)$  of the longest path compelled to pass through the point  $(\lfloor an \rfloor, 0)$  of the  $x$ -axis. To do this, let us establish a concentration inequality on  $L_n^{(\sqrt{n})}(a; 0^+)$ .

As above, assume the number  $an$  is an integer. The following stochastic domination holds:

$$L_n^{(\sqrt{n})}(a; 0^+) \leq L_{an,0}^{(\sqrt{n})} + L_{(1-a)n,1}^{(\alpha\sqrt{n})}.$$

Notice that the random variable  $L_{an,0}^{(\sqrt{n})}$  corresponds to the length of a longest path on the  $x$ -axis only, and follows a Gamma distribution of parameters  $(an, \sqrt{n})$ .

By reasoning in the same way as for the variable  $L_n^{(\sqrt{n})}(a; u^+, t)$  for  $u = t = 0$ , we get the following concentration inequality with  $f(a) = \frac{a^2}{12}$ :

$$\mathbb{P}\left(L_n^{(\sqrt{n})}(a; 0^+) \geq (2 - f(a))\sqrt{n}\right) \leq 3 \exp\left(-\frac{f(a)^2}{36}n\right). \quad (3.21)$$

Exploiting the fact that a path which passes through the point  $(an, 0)$  of the  $x$ -axis leaves this axis at a point  $(i, 0)$  for some integer  $i \in \llbracket an, n \rrbracket$ , we get that

$$L_n^{(\sqrt{n})}(a; 0, t) = \max_{i \in \llbracket an, n \rrbracket} L_n^{(\sqrt{n})}\left(\frac{i}{n}; 0^+\right)$$

Consequently, using equation (3.21) and the fact that the function  $f$  is non-decreasing,

$$\begin{aligned} \mathbb{P}\left(L_n^{(\sqrt{n})}(a; 0) \geq (2 - f(a))\sqrt{n}\right) &\leq \sum_{i=an}^n \mathbb{P}\left(L_n^{(\sqrt{n})}\left(\frac{i}{n}; 0^+\right) \geq (2 - f(a))\sqrt{n}\right) \\ &\leq \sum_{i=an}^n \mathbb{P}\left(L_n^{(\sqrt{n})}\left(\frac{i}{n}; 0^+\right) \geq (2 - f\left(\frac{i}{n}\right))\sqrt{n}\right) \\ &\leq \sum_{i=an}^n 3 \exp\left(-\frac{f\left(\frac{i}{n}\right)^2}{36}n\right) \leq 3n \exp\left(-\frac{f(a)^2}{36}n\right) \end{aligned} \quad (3.22)$$

Finally, noticing that  $L_n^{(\sqrt{n})}(a; 0, t) = L_n^{(\sqrt{n})}(a; 0^+, t) \vee L_n^{(\sqrt{n})}(a; 0)$ , we get that, for  $n$  large enough,

$$\begin{aligned} &\mathbb{P}\left(L_n^{(\sqrt{n})}(a; 0, t) \geq (2 - f(a-t))\sqrt{n}\right) \\ &\leq \mathbb{P}\left(L_n^{(\sqrt{n})}(a; 0^+, t) \geq (2 - f(a-t))\sqrt{n}\right) + \mathbb{P}\left(L_n^{(\sqrt{n})}(a; 0) \geq (2 - f(a-t))\sqrt{n}\right) \\ &\leq \mathbb{P}\left(L_n^{(\sqrt{n})}(a; 0^+, t) \geq (2 - f(a-t))\sqrt{n}\right) + \mathbb{P}\left(L_n^{(\sqrt{n})}(a; 0) \geq (2 - f(a))\sqrt{n}\right) \\ &\leq \frac{8}{f(a-t)} \exp\left(-\frac{f(a-t)^2}{64}n\right) + 3n \exp\left(-\frac{f(a)^2}{36}n\right). \end{aligned}$$

Finally using the lower bound  $f(a) \geq f(a-t)$ , we get the following inequality by replacing the parameter  $t$  and the function  $f$  by their expression:

$$\sup_{a \in (\eta_n, 1]} \mathbb{P} \left( L_n^{(\sqrt{n})}(a; 0, a - \eta_n) \geq \left(2 - \frac{\eta_n^2}{12}\right) \sqrt{n} \right) \leq \left(3n + \frac{96}{\eta_n^2}\right) \exp\left(-\frac{\eta_n^4}{9216} n\right)$$

We conclude as above, by writing  $\eta_n^4 n = n^{1-4/\beta}$  and by using the assumption  $\beta > 4$ .  $\square$

Now we can deduce Proposition 3.2:

*Proof of Proposition 3.2.* As mentioned above, it remains to prove that with high probability, a longest path compelled to pass far from the diagonal cannot be the (unconstrained) optimal path. To do this, we use Lemma 3.6 which states that such a path cannot have a length close to  $2\sqrt{n}$  with high probability, and Proposition 3.1 to prove that the length of the optimal path is on the contrary close to  $2\sqrt{n}$ .

Set  $\eta = \eta_n = \frac{1}{n^{1/\beta}}$  as in Lemma 3.6,  $\varepsilon_n := \frac{\eta_n^2}{48}$  and let us decompose the probability into two parts, whether the length of the longest path is greater than  $(2 - \varepsilon_n)\sqrt{n}$  or not:

$$\mathbb{P} \left( d \left( P_n^{(\sqrt{n})}, \Delta \right) \geq \eta \right) \leq \mathbb{P} \left( d \left( P_n^{(\sqrt{n})}, \Delta \right) \geq \eta, L_n^{(\sqrt{n})} \geq (2 - \varepsilon_n)\sqrt{n} \right) + \mathbb{P} \left( L_n^{(\sqrt{n})} < (2 - \varepsilon_n)\sqrt{n} \right) \quad (3.23)$$

The second term can be bounded from above by Proposition 3.1 and Chernoff bounds:

$$\begin{aligned} \mathbb{P} \left( L_n^{(\sqrt{n})} < (2 - \varepsilon_n)\sqrt{n} \right) &\leq \mathbb{P} \left( \mathcal{N}(\sqrt{n}, 1) \leq \left(1 - \frac{\varepsilon_n}{2}\right) \sqrt{n} \right) + \mathbb{P} \left( \Gamma(n, \sqrt{n}) \leq \left(1 - \frac{\varepsilon_n}{2}\right) \sqrt{n} \right) \\ &\leq \exp\left(-\frac{\varepsilon_n^2}{8} n\right) + \exp\left(-\frac{\varepsilon_n^2}{16} n\right) \leq 2 \exp\left(-\frac{\varepsilon_n^2}{16} n\right) \\ &\leq 2 \exp\left(-\frac{n^{1-4/\beta}}{36864}\right) = o(1) \quad \text{since } \beta > 4. \end{aligned}$$

Now, we bound from above the first term of the right hand side of equation (3.23), by exhibiting a first coordinate  $j/n$  where the optimal path  $P_n^{(\sqrt{n})}$  is at a distance  $\eta$  away from the diagonal  $\Delta$ .

$$\begin{aligned} &\mathbb{P} \left( d \left( P_n^{(\sqrt{n})}, \Delta \right) \geq \eta, L_n^{(\sqrt{n})} \geq (2 - \varepsilon_n)\sqrt{n} \right) \\ &= \mathbb{P} \left( \exists j \in \llbracket 0, n \rrbracket, [\mathbf{t}_{j-1}^*, \mathbf{t}_j^*] \cap \left( \left[0, \frac{j}{n} - \eta\right] \cup \left[\frac{j}{n} + \eta, 1\right] \right) \neq \emptyset, L_n^{(\sqrt{n})} \geq (2 - \varepsilon_n)\sqrt{n} \right) \\ &\leq \sum_{j=0}^n \left[ \mathbb{P} \left( L_n^{(\sqrt{n})} \left( \frac{j}{n}; \frac{j}{n} + \eta, 1 \right) \geq (2 - \varepsilon_n)\sqrt{n} \right) \mathbf{1}_{j \leq n(1-\eta)} + \mathbb{P} \left( L_n^{(\sqrt{n})} \left( \frac{j}{n}; 0, \frac{j}{n} - \eta \right) \geq (2 - \varepsilon_n)\sqrt{n} \right) \mathbf{1}_{j \geq n\eta} \right]. \end{aligned}$$

Recall that  $\varepsilon_n = f(\eta) = f\left(\frac{j}{n} + \eta - \frac{j}{n}\right)$ . Hence, for all integer  $j \leq n(1 - \eta)$ , one can apply Lemma 3.6 to get

$$\mathbb{P} \left( L_n^{(\sqrt{n})} \left( \frac{j}{n}; \frac{j}{n} + \eta, 1 \right) \geq (2 - \varepsilon_n)\sqrt{n} \right) = o\left(\frac{1}{n}\right).$$

By bounding the second term similarly, we finally get

$$\mathbb{P}\left(d\left(P_n^{(\sqrt{n})}, \Delta\right) \geq \eta, L_n^{(\sqrt{n})} \geq (2 - \varepsilon_n)\sqrt{n}\right) = o(1).$$

We have shown that the two terms of the right hand side of equation (3.23) tend to zero, hence the Proposition is proved.  $\square$

### 3.6 Recovering the asymptotics for the largest eigenvalue of a GUE matrix : proof of Corollary 3.1

The aim of this section is to prove Corollary 3.1. We prove that the above-mentioned scaling between the parameters  $c$  and  $n$ , that is to say  $c = \sqrt{n}$ , makes the boundary effects disappear, which makes (almost) coincide both functionals  $L_n$  and  $L_n^*$ .

We use the same strategy of proof as the one of [RSSV10, Theorem 4.1] and [BEGG16, Section 4]. We first observe that the sequence  $(L_n)$  is superadditive. Hence, the convergence of the expectation  $\mathbb{E}\left[\frac{L_n}{\sqrt{n}}\right]$  to a constant is enough to prove almost sure convergence, by Kingman's subadditive ergodic theorem [Kin68].

To get this constant, recall the stochastic domination (3.4) which we use with the optimal drift parameter  $c = \sqrt{n}$ :

$$L_n \leq L_n^{(\sqrt{n})}.$$

We know that the first order asymptotics of the variable  $L_n^{(\sqrt{n})}$  is  $2\sqrt{n}$  thanks to Proposition 3.1, it suffices to establish a lower bound with the same asymptotics. This lower bound will be provided by Proposition 3.2, since it has a consequence that the longest path in the stationary model with  $c = \sqrt{n}$  do not collect many sources and sinks.

Let us now study the number of sources of sinks taken by an optimal path. Denote by  $D_n^{(\sqrt{n})}(\rightarrow)$  and  $D_n^{(\sqrt{n})}(\uparrow)$  the amount of sources and sinks respectively taken by a longest path in the stationary model. Notice that almost surely one of these two quantities is zero. We have the following bounds on  $L_n = L_n(B)$ :

$$L_n^{(\sqrt{n})} - D_n^{(\sqrt{n})}(\rightarrow) - D_n^{(\sqrt{n})}(\uparrow) \leq L_n \leq L_n^{(\sqrt{n})} \tag{3.24}$$

The aim is to show the following lemma, which is assumed to be accepted for the moment:

**Lemma 3.8.** *The following convergences in  $\mathbb{L}^1$  hold:*

$$\frac{D_n^{(\sqrt{n})}(\rightarrow)}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{\mathbb{L}^1} 0 \quad \text{and} \quad \frac{D_n^{(\sqrt{n})}(\uparrow)}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{\mathbb{L}^1} 0.$$

*Proof of Corollary 3.1 assuming Lemma 3.8.* Once this lemma is shown we can deduce Corollary 3.1. Proposition 3.1 gives

$$\frac{L_n^{(\sqrt{n})}}{\sqrt{n}} \stackrel{(d)}{=} \frac{1}{\sqrt{n}} \left( B_0^{(\sqrt{n})}(1) + \Gamma_n^{(\sqrt{n})} \right),$$



so we can compute its expectation:

$$\mathbb{E} \left[ \frac{L_n^{(\sqrt{n})}}{\sqrt{n}} \right] = \frac{1}{\sqrt{n}} \left( \mathbb{E} \left[ B_0^{(\sqrt{n})}(1) \right] + \mathbb{E} \left[ \Gamma_n^{(\sqrt{n})} \right] \right) = 2.$$

By Cauchy–Schwarz inequality we can also bound from above its variance

$$\text{Var} \left( \frac{L_n^{(\sqrt{n})}}{\sqrt{n}} \right) \leq \frac{2}{n} \left( \text{Var} \left( B_0^{(\sqrt{n})}(1) \right) + \text{Var} \left( \Gamma_n^{(\sqrt{n})} \right) \right) = \frac{4}{n} \xrightarrow{n \rightarrow \infty} 0,$$

so the following convergence in  $\mathbb{L}^2$  follows:

$$\frac{L_n^{(\sqrt{n})}}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{\mathbb{L}^2} 2.$$

Hence with Lemma 3.8 and the bounds given by equation (3.24), we get the following convergence of the expectation:

$$\mathbb{E} \left[ \frac{L_n}{\sqrt{n}} \right] \xrightarrow[n \rightarrow \infty]{} 2.$$

Thus, the wanted almost sure convergence holds using Kingman's subadditive ergodic theorem as mentioned above.  $\square$

Now, let us use the concentration inequalities of Proposition 3.2 to prove Lemma 3.8.

*Proof of Lemma 3.8.* Let  $\eta \in (0, 1]$ . The aim here is to get an upper bound on  $\mathbb{P}(D_n^{(\sqrt{n})}(\rightarrow) > 2\sqrt{n}\eta)$ . Roughly speaking, either there are many sources close to 0, either the rescaled longest path  $P_n^{(\sqrt{n})}$  stays on the  $x$ -axis during a time at least  $\eta$  which implies it is at least as distance  $\eta$  from the diagonal  $\Delta$ . Indeed,

$$\begin{aligned} \mathbb{P} \left( D_n^{(\sqrt{n})}(\rightarrow) > 2\sqrt{n}\eta \right) &\leq \mathbb{P} \left( \sum_{k=1}^{\eta n} E_k^{(\sqrt{n})} > 2\sqrt{n}\eta \right) + \mathbb{P} \left( (0, \eta) \in P_n^{(\sqrt{n})} \right) \\ &\leq \mathbb{P}(\Gamma(\eta n, \sqrt{n}) > 2\sqrt{n}\eta) + \mathbb{P} \left( d \left( P_n^{(\sqrt{n})} - \Delta \right) \geq \eta \right) \\ &\leq \exp \left( -\frac{\eta}{4} n \right) + \exp(-h(\eta)n). \end{aligned}$$

This implies that  $\frac{D_n^{(c)}(\rightarrow)}{\sqrt{n}}$  converges almost surely to 0. Since the sequence is bounded in  $\mathbb{L}^2$  it is uniformly integrable, hence the convergence in  $\mathbb{L}^1$  follows. The convergence of the sequence  $\frac{D_n^{(c)}(\uparrow)}{\sqrt{n}}$  is proved in the same way.  $\square$

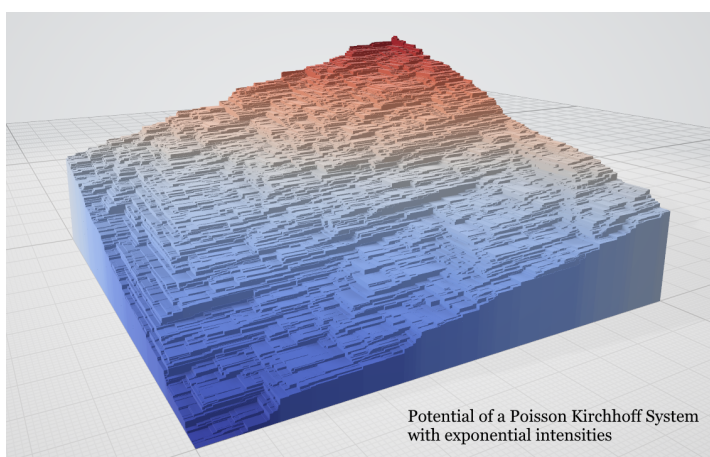
## Chapitre 4

# Reversible Poisson-Kirchhoff Systems

Joint work with Jérôme Casse<sup>1</sup>, Nathanaël Enriquez<sup>2</sup> and Arvind Singh<sup>3</sup>.

### Abstract

We define a general class of random systems of horizontal and vertical weighted broken lines on the quarter plane whose distribution are proved to be translation invariant. This invariance stems from a reversibility property of the model. This class of systems generalizes several classical processes of the same kind, such as Hammersley's broken line processes involved in Last Passage Percolation theory or such as the six-vertex model for some special sets of parameters. The novelty comes here from the introduction of a weight associated with each line. The lines are initially generated by spatially homogeneous weighted Poisson Point Process and their evolution (turn, split, crossing) are ruled by a Markovian dynamics which preserves Kirchhoff's node law for the line weights at each intersection. Among others, we derive some new explicit invariant measures for some bullet models as well as new reversible properties for some six-vertex models with an external electromagnetic field.

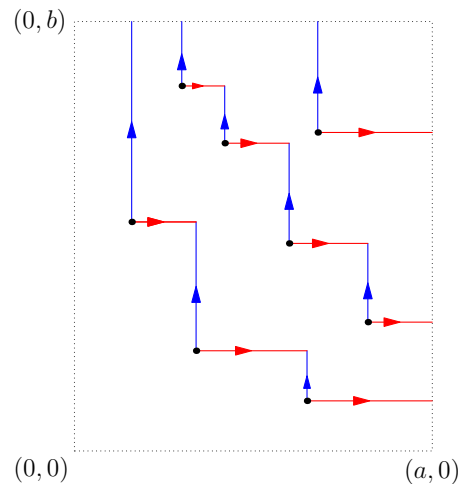


---

<sup>1</sup>LMO, Université Paris-Saclay. Email: [jerome.casse.math@gmail.com](mailto:jerome.casse.math@gmail.com)

<sup>2</sup>LMO, Université Paris-Saclay. Email: [nathanael.enriquez@universite-paris-saclay.fr](mailto:nathanael.enriquez@universite-paris-saclay.fr)

<sup>3</sup>CNRS. LMO, Université Paris-Saclay. Email: [arvind.singh@universite-paris-saclay.fr](mailto:arvind.singh@universite-paris-saclay.fr)



**Figure 4.1:** A realisation of Hammersley’s broken line process in the rectangle  $[0, a] \times [0, b]$ . The traces of particles with charge  $+1$  are represented in red and those of the antiparticles with charge  $-1$  are represented in blue.

## 4.1 Introduction

In his seminal work [Ham72], Hammersley introduced its now famous *broken line process* as a mean to study the length of the longest increasing sequence in a random permutation. This model of Last Passage Percolation (LPP) enjoys many remarkable properties and has since been thoroughly scrutinized [Ros81, Sep09]. One possible construction of Hammersley’s process on the quarter plane  $[0, \infty)^2$  goes as follow: consider a unit intensity Poisson Point Process (PPP) on  $[0, \infty)^2$ . Each atom of the point process “emits” a pair of particle/anti-particle with the particle of charge  $+1$  moving horizontally to the right and the antiparticle with charge  $-1$  moving upward. When the traces of two particles of opposite charge meet, they both disappear. Then, the collection of all traces obtained with this procedure is exactly the Hammersley’s broken line process (see Figure 4.1 for an illustration of the construction). Let us note that, in view of this construction, the system may be called “conservative” in the sense that the total charge of the system remains null since particles and antiparticles appear and disappear simultaneously.

In this chapter, we introduce a new class of random processes which we call *Poisson-Kirchhoff Systems* (PKS) that generalize the construction described above. Those processes consist again of random collections of weighted horizontal and vertical broken lines living on the quarter plane  $[0, \infty)^2$ . As for the Hammersley’s broken line process, one may think of these lines as being the traces of “charged” particles moving either horizontally (i.e. increasing their  $x$ -coordinate) or vertically (i.e. increasing their  $y$ -coordinate). However, in this new class of processes, particles may hold arbitrary charges and may randomly turn, split or coalesce according to a special Markovian dynamics which is still conservative in the sense that the total charge remains constant. We show in this chapter that, when the parameters of the dynamics take a particular form, the PKS process is spatially reversible. Then, it is possible to construct a translation invariant PKS process on the whole plane whose marginal distribution along vertical and horizontal lines are (weighted) PPPs.

The chapter is organized as follows. In Section 4.2 we define the PKS process in a general setting and prove its existence under a uniform boundedness assumption on the parameters.

In Section 4.3, we introduce a notion of reversibility for PKS processes which essentially says that the distribution of a PKS restricted to any rectangular box is invariant by a rotation of 180 degrees. Then we present, in our main results, suitable conditions that guarantee the reversibility and therefore the invariance of PKS processes. We do it in three different frameworks according to whether the distribution of the line weights is absolutely continuous with respect to Lebesgue measure, discrete or arbitrary.

The proof of this reversibility property is carried out in Section 4.5. The state space of PKS processes is quite complicated, and in order to deal with it, we introduce a family of parametrizations. It turns out that two different parametrizations of this family define the same volume form. We apply this result to two specific parametrizations: a first one associated to the dynamics of the PKS and the second one associated to its reverse dynamics. Once we have done it, a careful analysis shows that the densities associated to the dynamics and to the reversed one in their respective parametrizations coincide under the above-mentioned conditions. Interestingly, one can exploit this invariance result in order to extend the proof of the existence of the PKS to unbounded parameters.

In Section 4.6, we first show how Kirchhoff's node law makes it possible to define a notion of potential function associated with the faces of the tessellation defined by a PKS. This potential function corresponds to the last passage times in LPP. We then collect several LPP models which can be mapped to PKS processes. In the sequel, we provide a (non-exhaustive) list of PKS processes obtained for specific distributions of the line weights. From this list, we recover several other classical models of statistical physics like bullet models [KRL95, BM20, HST21] or six-vertex models [Pau35, Bax72]. In particular, we exhibit some new explicit invariant measures for some bullet models as well as new reversible properties for some six-vertex models with an external electromagnetic field. Furthermore, the special cases of Gaussian or Poisson distributions for the line weights provide new models with explicit dynamics which might be worthy of further study.

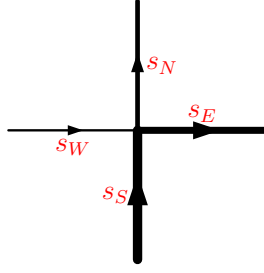
Finally, in Section 4.7, we look at basic geometric properties of the random tessellation of the quarter plane induced by a PKS, such as the mean number of connected components inside a rectangle, and the mean number of nodes of a typical connected component.

## 4.2 Poisson-Kirchhoff systems

The definition of a generic Poisson-Kirchhoff process relies on 9 parameters. First, let  $\lambda_0$  be a non-negative number which will be referred to as the *spontaneous creation rate*. Let  $\lambda_V$  and  $\lambda_H$  be two functions from  $\mathbb{R}$  to  $\mathbb{R}_+$ , called *vertical and horizontal split rate functions*. Let  $\tau_V$  and  $\tau_H$  be two functions from  $\mathbb{R}$  to  $\mathbb{R}_+$ , called *vertical and horizontal turn rate functions*. Let  $p_0 \in [0, 1]$  called the *annihilation probability*, let also  $p_V$  and  $p_H$  be two functions from  $\mathbb{R}$  to  $[0, 1]$ , called respectively *vertical and horizontal coalescence probability functions* that satisfy, for any  $s \in \mathbb{R}$ ,

$$p_V(s) + p_H(s) + p_0 \mathbf{1}_{s=0} \leq 1. \tag{4.1}$$

Finally, let  $F = (F(s, \cdot) : s \in \mathbb{R})$  be a probability transition kernel on  $\mathbb{R}$ , called the *division kernel*, which satisfies:



**Figure 4.2:** Kirchhoff's node law at a crossing: two lines are coming from the south and the west directions with respective weights  $s_S$  and  $s_W$ . Lines exiting the intersection to the north and west direction have respective weights  $s_N$  and  $s_E$ . The sum of weights entering and exiting the intersection is conserved:  $s_S + s_W = s_N + s_E$ .

- The map  $s \mapsto F(s, B)$  is  $\mathcal{B}(\mathbb{R})$ -measurable for any Borel set  $B \in \mathcal{B}(\mathbb{R})$ .
- $B \mapsto F(s, B)$  is a probability measure on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  for any  $s \in \mathbb{R}$ .

The collection  $(\lambda_0, \lambda_V, \lambda_H, p_0, p_V, p_H, \tau_V, \tau_H, F)$  represents the parameters of the model. The three parameters  $(\lambda_0, \lambda_V, \lambda_H)$  can be seen as splitting rates whereas  $(p_0, p_V, p_H)$  can be seen as merging probabilities. We will see that these two sets of parameters play a dual role. The two parameters  $(\tau_V, \tau_H)$  have a symmetric role and describe how often vertical and horizontal lines turn. Finally, the kernel  $F$  describes the distribution of the weights when a line splits or when two lines meet and split again.

We now define a random system of horizontal and vertical algebraic weighted lines inside the quarter plane  $[0, \infty)^2$  which preserves Kirchhoff's node law at every intersection (w.r.t. their weights), as prescribed in Figure 4.2. As in the description of Hammersley's process in Section 4.1, one can think of those lines as the traces of charged particles moving either to the right or upwards. Let us emphasize that, in our setting, the weight (i.e. charge) of a line may be positive, negative or even null.

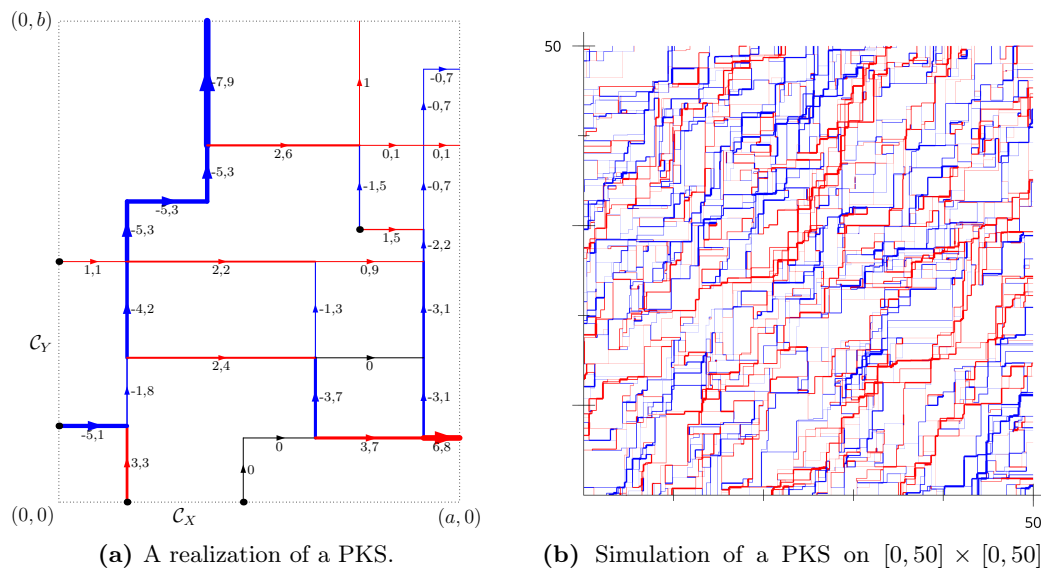
We define the *initial condition* of our process by specifying the positions and weights of the vertical (resp. horizontal) lines that start from the  $x$ -axis (resp.  $y$ -axis). To this end, we fix two sets of weighted points:  $\mathcal{C}_X$  on the positive  $x$ -axis and  $\mathcal{C}_Y$  on the positive  $y$ -axis. More precisely, an element of  $\mathcal{C}_X$  is of the form  $((x, 0), s) \in (\mathbb{R}_+ \times \{0\}) \times \mathbb{R}$ . Similarly, an element  $\mathcal{C}_Y$  is of the form  $((0, y), s) \in (\{0\} \times \mathbb{R}_+) \times \mathbb{R}$ . The two sets  $\mathcal{C}_X$  and  $\mathcal{C}_Y$  can be taken randomly. In order to avoid degeneracy, we will always assume that

$$\text{the sets of points in } \mathcal{C}_X \text{ and } \mathcal{C}_Y \text{ are locally finite a.s.,} \quad (\text{LF})$$

i.e. there is no accumulation point on either axis.

We also take a PPP  $\Xi_0$  on  $(0, \infty)^2 \times \mathbb{R}$  with intensity  $\lambda_0 dx dy F(0, ds)$ . From the initial conditions  $\mathcal{C}_X$  and  $\mathcal{C}_Y$ , and the parameters  $(\lambda_0, \lambda_V, \lambda_H, p_0, p_V, p_H, \tau_V, \tau_H, F)$ , we construct a system of lines with the following rules:

- 1<sub>V</sub>. From each element  $((x, 0), s) \in \mathcal{C}_X$ , we start a vertical line from the point  $(x, 0)$  going up with weight  $s$ .



(a) A realization of a PKS.

(b) Simulation of a PKS on  $[0, 50] \times [0, 50]$  according to Model 16 of Table 4.3 (see Section 4.6) with  $p_V(s) = p_H(s) = 0.4$  and  $\tau_V(s) = \tau_H(s) = 0.1$  whose initial condition  $(C_X, C_Y)$  are given by two independent PPPs.

**Figure 4.3:** Example of dynamics. Lines with positive weights are in red and those with negative weights in blue. The thickness of a line is proportional to the absolute value of its weight.

- 1<sub>H</sub>. From each element  $((0, y), s) \in \mathcal{C}_Y$ , we start an horizontal line from the point  $(0, y)$  going right with weight  $s$ .
- 1<sub>0</sub>. From each element  $((x, y), s) \in \Xi_0$ , we start an horizontal line from the point  $(x, y)$  going right with weight  $s$  and a vertical line going up with weight  $-s$ .

There are two kinds of events that occur during the dynamics. The first kind concerns what happens to a single line which may *turn* or *split* into two lines.

2<sub>V</sub>. Along a vertical line of weight  $s$ :

- a. A split occurs at rate  $\lambda_V(s)$ . When such an event happens, we pick a random variable  $T \sim F(s, \cdot)$ , independent of everything else. As a result of this split, the vertical line continues going up with new weight  $s - T$ , and a horizontal line with weight  $T$  starts going right from the point where the split occurs.
- b. The line turns to its right (i.e. to the east) at rate  $\tau_V(s)$  keeping the same weight and becoming a horizontal line.

2<sub>H</sub>. Along a horizontal line of weight  $s$ :

- a. A split occurs at rate  $\lambda_H(s)$ . When such an event happens, we pick a random variable  $T \sim F(s, \cdot)$ , independent of everything else. As a result of this split, the horizontal line continues going right with new weight  $T$ , and a vertical line with weight  $s - T$  starts going up from the point where the split occurs.

- b. The line turns to its left (i.e. to the north) at rate  $\tau_H(s)$  keeping the same weight and becoming a vertical line.

The second kind of event corresponds to intersections of lines (which we shall refer to as *crossing* events) when a horizontal line going right (i.e. coming from the west) with weight  $s_W$  meets a vertical line going up (i.e. coming from the south) with weight  $s_S$ . We apply the following rules:

3.
  - a. with probability  $p_V(s_S + s_W)$ , the horizontal line stops and the vertical line continues with weight  $s_N := s_S + s_W$ ;
  - b. with probability  $p_H(s_S + s_W)$ , the vertical line stops and the horizontal line continues with weight  $s_E := s_S + s_W$ ;
  - c. with probability  $p_0 \mathbf{1}_{s_S + s_W = 0}$ , both lines disappear;
  - d. on the complementary event, which happens with probability  $1 - p_V(s_S + s_W) - p_H(s_S + s_W) - p_0 \mathbf{1}_{s_S + s_W = 0}$ , we pick a random variable  $T \sim F(s_S + s_W, \cdot)$ , independent of everything else. Then, after meeting each other, the weight of the horizontal line becomes  $s_E := T$  and the weight of the vertical one becomes  $s_N := s_S + s_W - T$ .

Rules 1, 2 and 3 together with the initial set of weighted starting points  $\mathcal{C}_X$  and  $\mathcal{C}_Y$  define a random system of algebraic weighted lines which we call **Poisson-Kirchhoff System** (PKS) with parameters  $(\lambda_0, \lambda_V, \lambda_H, p_0, p_V, p_H, \tau_V, \tau_H, F)$  under the initial condition  $(\mathcal{C}_X, \mathcal{C}_Y)$ . Let us note that, according to the rules of the dynamics, the system is conservative: it satisfies Kirchhoff's node law (as in Figure 4.2) at every intersection, be it a split, a turn or a crossing. An illustration of a PKS process is given in Figure 4.3.

**Are PKS well defined?** Without further assumptions, the process constructed with the above procedure could be not well defined on the whole quarter plane. Indeed, the previous construction can fail (i.e. blow up) if an accumulation of lines appears and prevents us from defining the process any further. From now on, we will say that the PKS is *well defined* if, a.s., the construction above has no accumulation points on the whole quarter plane (or equivalently, there is only a finite number of lines intersecting any bounded region a.s.). The following trivial example illustrates the problem.

**Example 4.1.** Set  $\lambda_0 = \lambda_V(s) = \tau_V(s) = \tau_H(s) = p_0 = p_V(s) = p_H(s) = 0$ . Set  $\lambda_H(s) = s^2$  and  $F(s, \cdot) = \delta_{s+1}$ . Fix  $\mathcal{C}_X = \emptyset$  and  $\mathcal{C}_Y = \{(0, 1), 1\}$ . Then, the PKS starts from of a single horizontal line beginning at point  $(0, 1)$  on the  $y$ -axis and with initial weight 1. This horizontal line never disappears and splits infinitely many times, creating at each split a new vertical line with weight  $-1$  while its own weight increases by 1. Thus, the splitting rate of the horizontal line is equal to  $(n + 1)^2$  after the  $n$ th split. This means that the  $x$  coordinate of the  $n$ -th split is equal to  $\sum_{i=1}^n \frac{\xi_i}{i^2}$  where the  $(\xi_i)$  are i.i.d. exponential random variables with mean 1. The previous sum converges a.s. which shows that the PKS blows up almost surely.

Deciding whether a generic PKS is well defined seems tricky. However, the following elementary result ensures that the PKS is well defined a.s. whenever its jump rates are bounded. Later on, the main results in Section 4.3 will provide examples of well-defined PKS with unbounded jump rates.



**Proposition 4.1.** *Assume that  $\mathcal{C}_X$  and  $\mathcal{C}_Y$  satisfy assumption (LF), and that*

$$\sup_{s \in \mathbb{R}} (\lambda_V(s), \lambda_H(s), \tau_V(s), \tau_H(s)) < \infty. \quad (4.2)$$

*Then, the PKS is well defined on the whole quarter plane  $[0, \infty)^2$  a.s..*

*Proof.* Let us first note that

$$\begin{aligned} & \{\text{the PKS is well defined on the whole quarter plane}\} \\ &= \bigcap_{a,b \in \mathbb{N}} \{\text{the PKS is well defined inside the rectangle } [0, a] \times [0, b]\} \end{aligned}$$

Thus, we just need to prove that the PKS does not blow up inside any box  $[0, a] \times [0, b]$  a.s.. Let us fix such a box  $[0, a] \times [0, b]$ . Let  $y_1 < y_2 < \dots < y_{N_0}$  denote the  $y$ -coordinates of the points in  $\mathcal{C}_Y$  located on the segment  $\{0\} \times [0, b]$ . We just need to prove that the PKS is well defined a.s. inside  $[0, a] \times [0, y_1]$  and then we can repeat the same argument, starting now from height  $y_1$ , and conclude, after  $N_0$  steps that the process is a.s. well defined on the whole box.

Let  $r := \sup_{s \in \mathbb{R}} (\lambda_V(s), \lambda_H(s), \tau_V(s), \tau_H(s)) < \infty$  and  $M_0$  denote the number of weighted points of  $\mathcal{C}_X$  located on the segment  $[0, a] \times \{0\}$ . We follow the dynamics starting from the bottom side of the box and moving upward.

Initially, we start with  $M_0$  vertical lines going upward. The first split/turn event occurs at some random height  $H_1$  which is stochastically larger than an exponential random variable with mean  $1/(2M_0r)$  (since all rates are bounded by  $r$ ). At height  $H_1$ , a new horizontal line is created. This line creates  $U_1$  new vertical lines (by splitting and at most one by turning) that will grow upward, and stops  $V_1 \geq 0$  vertical lines coming from the bottom (including itself in case of a turn event), see Figure 4.4. Hence, after height  $H_1$ , the process continues to grow upward with  $M_1 = M_0 + U_1 - V_1$  vertical lines. Similarly, after the  $n$ th split/turn event that occurs at height  $H_n$ , the process grows up with  $M_n = M_{n-1} + U_n - V_n$  vertical lines.

Now, the height  $H_n$  of the  $n$ th split/turn event is stochastically larger than

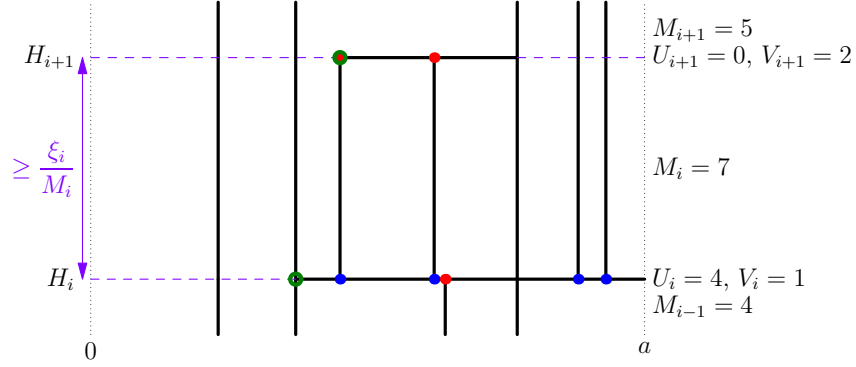
$$\sum_{i=0}^{n-1} \frac{\xi_i}{M_i}, \quad (4.3)$$

where  $(\xi_i)$  are i.i.d. exponential random variables with mean  $1/(2r)$  which are independent of  $(M_i)$ . But, remark that since the width of the box is equal to  $a$ , and that the split and turn rate functions are bounded by  $r$ , the sequence  $(U_i)_{i \geq 1}$  is stochastically dominated by a sequence  $(W_i)_{i \geq 1}$  of independent Poisson random variables with parameter  $2ar$ . This implies that the sequence of variables  $M_n$ , which are individually bounded by  $M_0 + \sum_{i=1}^n W_i$ , grows at most linearly with  $n$ , and so the sum given in (4.3) goes a.s. to infinity. Hence, the PKS process cannot blow up before reaching height  $y_1$ , as requested.  $\square$

### 4.3 Reversible Poisson-Kirchhoff systems

Although the PKS is defined on the whole quarter plane, it is convenient to consider its restriction to a box of the form  $[0, a] \times [0, b]$  for  $a, b \in \mathbb{R}_+^*$ . We denote by  $\mathcal{D}_{a,b}$  the image space of the





**Figure 4.4:** In this example, the  $i$ th event (green circle) occurs at height  $H_i$  and is a split. The  $(i+1)$ th event (also green circle) occurs at height  $H_{i+1}$  and is a turn. On the  $i$ th event, the horizontal line creates  $U_i = 4$  new lines (blue dots) and stops  $V_i = 1$  line (red dot). The  $(i+1)$ th event stops 1 line additionally to the one that turns (the two red dots) and creates no line. The time between the  $i$ th event and the  $(i+1)$ th event is greater than  $\xi_i/M_i$  that is an exponential random variable with mean  $1/(2M_i r)$ .

PKS process restricted to the box  $[0, a] \times [0, b]$ . An element  $D \in \mathcal{D}_{a,b}$  is called a *drawing* in the box  $[0, a] \times [0, b]$ . It consists of a finite collection of *weighted vertical* and *horizontal segments* inside this rectangle and which furthermore satisfy the Kirchhoff node law at every intersection (in the sense of Figure 4.2).

Given a drawing  $D \in \mathcal{D}_{a,b}$ , we define its reverse drawing  $\widehat{D} \in \mathcal{D}_{a,b}$ , obtained by rotating  $D$  by 180 degrees around the center point  $(a/2, b/2)$ . Let us note that this rotation yields a valid drawing. An example of a drawing  $D$  and its reverse  $\widehat{D}$  is given in Figure 4.5. From now on, we shall denote by  $\mathbf{D}_{a,b}$  (or simply  $\mathbf{D}$  when the box considered is obvious) a random drawing which has the law of the PKS process defined in Section 4.2.

**Definition 4.1** (Reversibility). A PKS is said to be *reversible* if there exists a random initial condition  $(\mathcal{C}_X, \mathcal{C}_Y)$  such that for any  $a, b$ ,  $\mathbf{D}_{a,b} \stackrel{(d)}{=} \widehat{\mathbf{D}}_{a,b}$ .

Another related notion is that of *stationarity* of the PKS, which will be implied by the reversibility property in all the cases we shall consider.

**Definition 4.2** (Stationarity). A PKS is said to be *stationary* if there exists a random initial condition  $(\mathcal{C}_X, \mathcal{C}_Y)$  such that it is translation invariant. Equivalently, this means that the law of a drawing does not depend on the position of the box inside the quarter plane but only on its size. In that case, the law of the initial condition  $(\mathcal{C}_X, \mathcal{C}_Y)$  is called an *invariant probability measure* for the PKS.

We now give sufficient conditions on the parameters of a PKS to be reversible under initial conditions  $(\mathcal{C}_X, \mathcal{C}_Y)$  taken as independent weighted PPPs. Thus, from now on, we consider two non-zero finite measures  $\nu_V$  and  $\nu_H$  on  $\mathbb{R}$  and we will always assume that

$$\begin{cases} \mathcal{C}_X \text{ is a PPP on } (\mathbb{R}_+ \times \{0\}) \times \mathbb{R} \text{ with intensity } dx d\nu_V(s). \\ \mathcal{C}_Y \text{ is a PPP on } (\{0\} \times \mathbb{R}_+) \times \mathbb{R}, \text{ independent of } \mathcal{C}_X, \text{ and with intensity } dy d\nu_H(s). \end{cases} \quad (4.4)$$

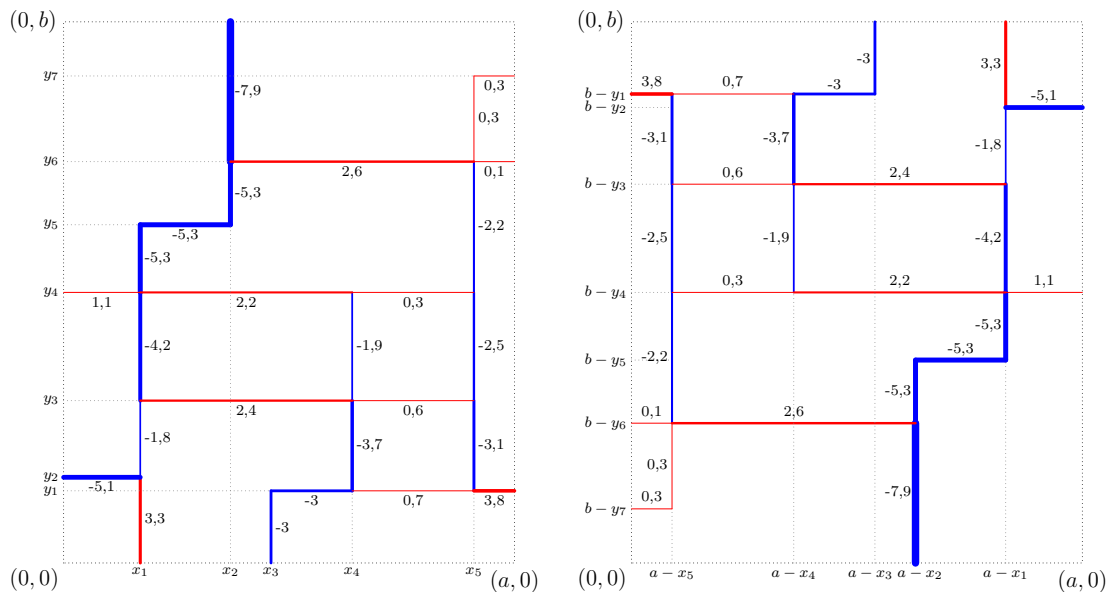


Figure 4.5: An example of a drawing and, on its right, its reverse

Let us stress that  $\nu_V$  and  $\nu_H$  are not necessarily probability measures. In particular, the positions of the vertical lines starting from the  $x$ -axis is a PPP with intensity  $\nu_V(\mathbb{R})dx$ . Similarly the positions of the horizontal lines starting from the  $y$ -axis is a PPP with intensity  $\nu_H(\mathbb{R})dy$ .

We call the measures  $\nu_V$  (resp.  $\nu_H$ ) the *vertical* (resp. *horizontal*) *line weight measures*. For technical reasons, we distinguish the following three cases depending on their properties:

- when  $\nu_V$  and  $\nu_H$  are both absolutely continuous w.r.t. the Lebesgue measure (Section 4.3.1);
- when  $\nu_V$  and  $\nu_H$  are discrete measures with support included in  $\mathbb{Z}$  (Section 4.3.2);
- finally, we discuss the extension of the previous results to arbitrary measures (Section 4.3.3).

### 4.3.1 Lebesgue case

We assume here that the line weight measures  $\nu_V$  and  $\nu_H$  are two non-zero finite measures on  $\mathbb{R}$  with Lebesgue densities  $g_V$  and  $g_H$ . Thus, the initial conditions (4.4) now take the form:

$$\begin{cases} \mathcal{C}_X \text{ is a PPP on } (\mathbb{R}_+ \times \{0\}) \times \mathbb{R} \text{ with intensity } dx g_V(s)ds. \\ \mathcal{C}_Y \text{ is a PPP on } (\{0\} \times \mathbb{R}_+) \times \mathbb{R}, \text{ independent of } \mathcal{C}_X, \text{ and with intensity } dy g_H(s)ds. \end{cases} \quad (4.5)$$

Consider a PKS with parameters  $(\lambda_0, \lambda_V, \lambda_H, p_0, p_V, p_H, \tau_V, \tau_H, F)$ . The following assumptions ensure the existence of a reversible measure for the PKS process:

(L1) The *spontaneous creation rate* is zero, i.e.

$$\lambda_0 = 0. \quad (4.6)$$

Indeed, since we are here in a continuous setting, case 3(c) of the dynamics (in Section 4.2) never occurs so lines never annihilate. Therefore, in order for the system to be reversible, there must be no spontaneous creation of lines. Consequently, the *annihilation probability*  $p_0$  also does not matter here (and can be taken to be zero).

- (L2) The *coalescence probability functions*  $p_V$  and  $p_H$  satisfy the two following conditions with respect to the support of the measures  $\nu_V$  and  $\nu_H$ : for any  $s \in \mathbb{R}$ , we have

$$g_V(s) = 0 \Rightarrow p_V(s) = 0 \quad \text{and} \quad g_H(s) = 0 \Rightarrow p_H(s) = 0. \quad (4.7)$$

- (L3) The two *turn rate functions*  $\tau_V$  and  $\tau_H$  satisfy, for any  $s \in \mathbb{R}$ ,

$$\tau_V(s)g_V(s) = \tau_H(s)g_H(s). \quad (4.8)$$

- (L4) The two *splitting rate functions*  $\lambda_V$  and  $\lambda_H$  satisfy, for any  $s \in \mathbb{R}$ ,

$$\lambda_V(s) = p_V(s) \frac{h(s)}{g_V(s)} \quad \text{and} \quad \lambda_H(s) = p_H(s) \frac{h(s)}{g_H(s)}, \quad (4.9)$$

where  $h$  is defined by

$$h(s) := (g_V * g_H)(s) = \int_{\mathbb{R}} g_V(s-t)g_H(t) dt.$$

- (L5) The *division kernel*  $F$  satisfies that, for any  $s \in \mathbb{R}$ , the measure  $F(s, \cdot)$  is absolutely continuous with respect to the Lebesgue measure and its density  $f(s, \cdot)$  is such that, for any  $t \in \mathbb{R}$ ,

$$f(s, t) = \frac{g_V(s-t)g_H(t)}{h(s)} \quad (4.10)$$

provided that  $h(s) > 0$ . If  $h(s) = 0$ , then  $f(s, \cdot)$  can be any probability density<sup>4</sup> on  $\mathbb{R}$ . Notice that this density has a simple probabilistic interpretation: let  $X_V$  and  $X_H$  be two independent random variables with density proportional to  $g_V$  and  $g_H$  respectively, then  $f(s, \cdot)$  is the density of the variable  $X_H$  conditionally on the event  $\{X_V + X_H = s\}$ .

We can now state our main result.

**Theorem 4.1** (Reversibility in Lebesgue case). *Consider a PKS with parameters  $(\lambda_0, \lambda_V, \lambda_H, p_0, p_V, p_H, \tau_V, \tau_H, F)$ . If there exist two non-zero finite measures  $\nu_V$  and  $\nu_H$  on  $\mathbb{R}$  with densities (according to Lebesgue measure)  $g_V$  and  $g_H$  such that the previous conditions (L1), (L2), (L3), (L4) and (L5) hold, then this PKS under the initial condition  $(\mathcal{C}_X, \mathcal{C}_Y)$  as defined in equation (4.5) is well defined and reversible in the sense of Definition 4.1.*

*Remark.* A PKS satisfying the conditions of Theorem 4.1 may have several reversible measures. Indeed, suppose that there exists a positive constant  $r$  such that  $\tilde{g}_V(s) = r^s g_V(s)$  and  $\tilde{g}_H(s) = r^s g_H(s)$  are still the densities of finite measures. Then Theorem 4.1 still applies by replacing  $g_V$  and  $g_H$  by  $\tilde{g}_V$  and  $\tilde{g}_H$ . Consequently, the PKS admits another reversible distribution, given by the law of two independent PPPs: one with intensity  $\text{Leb} \otimes \tilde{\nu}_V$  and the other one with intensity  $\text{Leb} \otimes \tilde{\nu}_H$ . In that case, the PKS admits a family of reversible measures parameterised by  $r$  in an open subinterval of  $\mathbb{R}_+$ .

<sup>4</sup>The distribution of  $f(s, \cdot)$  when  $h(s) = 0$  do not matter. This is just to insure that  $f$  is well defined for all  $s \in \mathbb{R}$ .

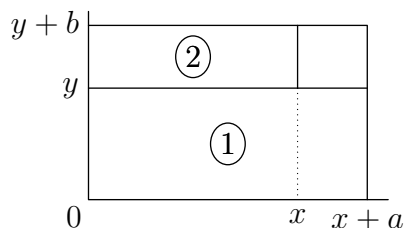
This theorem is proved in Section 4.5.1. As stated in the next corollary, the reversibility in this case implies the stationarity. Moreover, it also characterizes the law of the restriction of the process along any fixed decreasing curve.

**Corollary 4.1.** *Under the conditions of Theorem 4.1, the PKS is stationary as defined in Definition 4.2. In particular, the following properties hold:*

- (i) *One of its invariant distribution is the law of two independent PPPs  $(\mathcal{C}_X, \mathcal{C}_Y)$  where  $\mathcal{C}_X$  has intensity  $\text{Leb} \otimes \nu_V$  and  $\mathcal{C}_Y$  has intensity  $\text{Leb} \otimes \nu_H$ .*
- (ii) *Let  $L$  be any broken line of  $[0, a] \times [0, b]$  consisting only of eastern and southern steps. Then,  $\mathbf{D}$  restricted to  $L$  on its eastern steps is a  $\text{Leb} \otimes \nu_V$ -PPP, and  $\mathbf{D}$  restricted to  $L$  on its southern steps is a  $\text{Leb} \otimes \nu_H$ -PPP. These two PPPs are independent.*
- (iii) *Let  $L$  be any straight segment of  $\mathbb{R}^2$ :  $y = -\alpha x + \beta$  with  $x \in [c, d]$  and  $\alpha \in (0, +\infty)$ . The restriction of  $\mathbf{D}$  to its vertical (resp. horizontal) lines is a  $\text{Leb} \otimes \frac{1}{\sqrt{1+\alpha^2}}\nu_V$ -PPP (resp.  $\text{Leb} \otimes \frac{\alpha}{\sqrt{1+\alpha^2}}\nu_H$ -PPP). Moreover, these two PPPs are independent.*

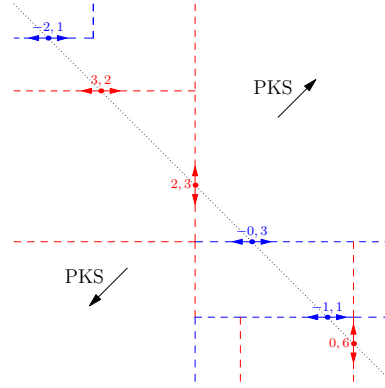
*Remark.* The last result (iii) can still be generalised to any rectifiable curve  $\gamma(t) = (x(t), y(t))$  from  $[0, 1]$  to  $[0, a] \times [0, b]$  which is “decreasing” in the sense that  $x' \geq 0$  and  $y' \leq 0$ . Then, again, the restrictions of the horizontal and vertical lines of the random drawing to this curve form independent inhomogeneous PPPs whose respective intensities with respect to  $d\lambda \otimes \nu_V$  and  $d\lambda \otimes \nu_H$ , where  $d\lambda$  denotes the length measure on the curve, at the point of parameter  $t$ , are given by the formulas of Corollary 4.1 (iii) taking  $\alpha = -\frac{y'(t)}{x'(t)}$ .

*Proof of Corollary 4.1.* We prove that the restriction of the process to any box  $[x, x + a] \times [y, y + b]$  has the same distribution as the one to the box  $[0, a] \times [0, b]$ , by showing that they are similarly distributed on their left and down boundaries. For this purpose, we first apply Theorem 4.1 to the box  $[0, x + a] \times [0, y]$  and then to the box  $[0, x] \times [y, y + b]$  as illustrated below.



Indeed, the first application implies that the restriction of the process to the segment  $[0, x + a] \times \{y\}$  is distributed as a  $\text{Leb} \otimes \nu_V$ -PPP. Moreover, it is independent of the restriction of the process to the segment  $\{0\} \times [y, y + b]$ .

Consequently, the second application applies and permits to prove that the restriction of the process to the segment  $\{x\} \times [y, y + b]$  is distributed as a  $\text{Leb} \otimes \nu_H$ -PPP. Moreover, this restriction is independent of the one to the segment  $[x, x + a] \times \{y\}$  since it only depends on the restrictions to the segments  $\{0\} \times [y, y + b]$  and  $[0, x] \times \{y\}$  as well as on the dynamics of the process above  $y$ .  $\square$



**Figure 4.6:** Schematic construction of a reversible PKS process on the full plane.

Another nice consequence of Theorem 4.1 is that the reversibility property makes it straightforward to extend the stationary PKS process defined on the quarter plane  $\mathbb{R}_+^2$  to a stationary process defined on the full plane  $\mathbb{R}^2$ . There are several ways to do so. For instance, we can use (iii) of Corollary 4.1 and start by choosing initially two independent  $\text{Leb} \otimes \frac{1}{\sqrt{2}}\nu_V$ -PPP and  $\text{Leb} \otimes \frac{1}{\sqrt{2}}\nu_H$ -PPP on the anti-diagonal line  $y = -x$ . We start weighted vertical lines from the atoms of the first PPP (with lines propagating in both top and bottom direction). Similarly, we start weighted horizontal lines from the atoms of the second PPP (with lines propagating in both left and right direction). Then, conditionally on these initial lines, we construct independent processes on the upper region  $x > -y$  and the lower region  $x < -y$  following the PKS dynamic (*c.f.* Figure 4.6). The resulting process defined on the whole plane  $\mathbb{R}^2$  is translation invariant and its restriction to any box (or quarter plane) coincide with the reversible PKS defined above.

### 4.3.2 Discrete case

We call *discrete case* when all the line weights are integers. Let us start by noticing that, contrary to the Lebesgue setting, case 3(c) of the dynamics can now occur since two lines with exactly opposite weights can meet. This happens for example with Hammersley's broken line process (schematized in Figure 4.1), see [Ham72, AD95, Gro02, CG05, CG06] for additional details. Therefore, in this section, the value of  $p_0$  matters and  $\lambda_0$  may be non-zero.

Let  $\nu_V$  and  $\nu_H$  be two non-zero finite measures taking values in  $\mathbb{Z}$ . We consider the initial conditions:

$$\begin{cases} \mathcal{C}_X \text{ is a PPP on } (\mathbb{R}_+ \times \{0\}) \times \mathbb{Z} \text{ with intensity } dx \nu_V(ds). \\ \mathcal{C}_Y \text{ is a PPP on } (\{0\} \times \mathbb{R}_+) \times \mathbb{Z}, \text{ independent of } \mathcal{C}_X, \text{ and with intensity } dy \nu_H(ds). \end{cases} \quad (4.11)$$

We consider a PKS with parameters  $(\lambda_0, \lambda_V, \lambda_H, p_0, p_V, p_H, \tau_V, \tau_H, F)$ . the following conditions ensure the reversibility of the process.

(D1) The *spontaneous creation rate*  $\lambda_0$  is related to the *annihilation probability*  $p_0$  as follows:

$$\lambda_0 = p_0 \sum_{s \in \mathbb{Z}} \nu_V(-s) \nu_H(s). \quad (4.12)$$

(D2) The *coalescence probability functions*  $p_V$  and  $p_H$  satisfy two conditions with respect to  $\nu_V$  and  $\nu_H$ : for any  $s \in \mathbb{Z}$ ,

$$\nu_V(s) = 0 \Rightarrow p_V(s) = 0 \quad \text{and} \quad \nu_H(s) = 0 \Rightarrow p_H(s) = 0. \quad (4.13)$$

(D3) The two *turn rate functions*  $\tau_V$  and  $\tau_H$  satisfy, for any  $s \in \mathbb{Z}$ ,

$$\tau_V(s) \nu_V(s) = \tau_H(s) \nu_H(s). \quad (4.14)$$

(D4) The two *splitting rate functions*  $\lambda_V$  and  $\lambda_H$  satisfy, for any  $s \in \mathbb{Z}$ ,

$$\lambda_V(s) := p_V(s) \frac{h(s)}{\nu_V(s)} \quad \text{and} \quad \lambda_H(s) := p_H(s) \frac{h(s)}{\nu_H(s)}, \quad (4.15)$$

where

$$h(s) = (\nu_V * \nu_H)(s) = \sum_{t \in \mathbb{Z}} \nu_V(s-t) \nu_H(t).$$

(D5) The *division kernel*  $F$  satisfies, for any  $s \in \mathbb{Z}$ , for any  $t \in \mathbb{Z}$ ,

$$F(s, t) = \frac{\nu_V(s-t) \nu_H(t)}{h(s)} \quad (4.16)$$

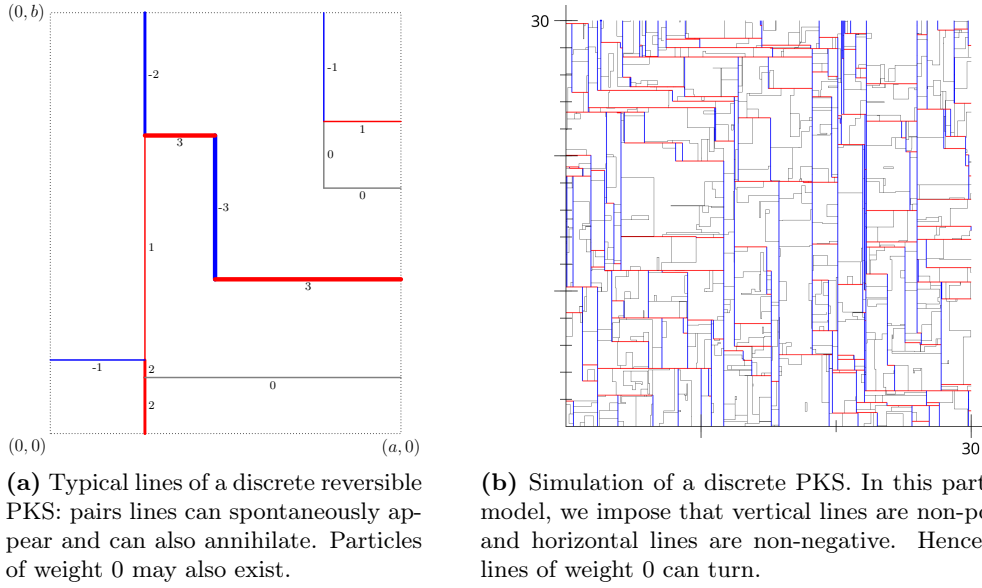
provided that  $h(s) > 0$ . If  $h(s) = 0$ , then  $F(s, \cdot)$  can be chosen to be any probability measure.

We can now state the theorem in the discrete case.

**Theorem 4.2** (Reversibility in the discrete case). *Consider a PKS with parameters  $(\lambda_0, \lambda_V, \lambda_H, p_0, p_V, p_H, \tau_V, \tau_H, F)$ . If there exist two non-zero finite measures  $\nu_V$  and  $\nu_H$  on  $\mathbb{Z}$  such that the previous conditions (D1), (D2), (D3), (D4) and (D5) hold, then this PKS under the initial condition  $(\mathcal{C}_X, \mathcal{C}_Y)$  as defined in equation (4.11) is well defined and reversible.*

Corollary 4.1 also holds in this case.

*Remark.* Let us note that Theorems 4.1 and 4.2 insure that a reversible PKS is well defined when the initial condition is given by one of its invariant measures. It follows that the PKS is also well defined starting from any initial condition that is absolutely continuous w.r.t. this invariant measure. In the discrete case, one may check the stronger result that the reversible PKS is, in fact, well defined for any deterministic initial conditions  $(\mathcal{C}_X, \mathcal{C}_Y)$ . This may not necessarily be true in the Lebesgue case as it is possible to construct forbidden “pathological” initial conditions, for example with lines having opposite weights.



**Figure 4.7:** Examples of a discrete reversible PKS. Lines with positive weight are drawn in red and those with negative weight in blue. Line with weight 0 are drawn in black.

### 4.3.3 General case

The Lebesgue and discrete cases described previously represent the most natural settings for PKS yet is it still possible to generalize the reversibility result to a more general framework as explained below.

Let us denote by  $\mathcal{A}_V$  (resp.  $\mathcal{A}_H$ ) the set of atoms of  $\nu_V$  (resp.  $\nu_H$ ), and set

$$\mathcal{A} := \mathcal{A}_H \cap (-\mathcal{A}_V) = \{s \in \mathbb{R} : \nu_H(\{s\})\nu_V(\{-s\}) \neq 0\}.$$

In the absence of a common measure against which both  $\nu_V$  and  $\nu_H$  are absolutely continuous, we shall make use of the Radon–Nikodym derivatives with respect to  $\nu_V$  and to  $\nu_H$  in order to define the appropriate rate functions that guarantee reversibility.

We recall that, according to the Radon–Nikodym theorem, given two arbitrary finite measures  $\mu$  and  $m$  there exists a unique decomposition of  $\mu = \mu_{\parallel} + \mu_{\perp}$  such that  $\mu_{\parallel}$  is abs. continuous w.r.t.  $m$  and  $\mu_{\perp}$  is singular w.r.t.  $m$ . Henceforth, we define the Radon–Nikodym derivative of the measure  $\mu$  according to the measure  $m$  as  $\frac{d\mu_{\parallel}}{dm}$  and denote it as  $\frac{d\mu}{dm}$ .

Consider a PKS with parameters  $(\lambda_0, \lambda_V, \lambda_H, p_0, p_V, p_H, \tau_V, \tau_H, F)$ . We present some sufficient conditions for reversibility.

**(G1)** The *spontaneous creation rate*  $\lambda_0$  satisfies

$$\lambda_0 = p_0 \sum_{s \in \mathcal{A}} \nu_V(\{-s\})\nu_H(\{s\}). \quad (4.17)$$

**(G2)** The *coalescence probability functions*  $p_V$  and  $p_H$  satisfy

$$p_V \in L^\infty(\mathbb{R}, \mathcal{B}(\mathbb{R}), \nu_V) \text{ and } p_H \in L^\infty(\mathbb{R}, \mathcal{B}(\mathbb{R}), \nu_H). \quad (4.18)$$

**(G3)** The two *turn rate functions*  $\tau_V$  and  $\tau_H$  satisfy, for any  $s \in \mathbb{R}$ ,

$$\tau_V(s) = \tau_H(s) \frac{d\nu_H}{d\nu_V}(s), \text{ or equivalently } \tau_H(s) = \tau_V(s) \frac{d\nu_V}{d\nu_H}(s). \quad (4.19)$$

**(G4)** The two *splitting rate functions*  $\lambda_V$  and  $\lambda_H$  satisfy, for any  $s \in \mathbb{R}$ ,

$$\lambda_V(s) = p_V(s) \frac{d\eta}{d\nu_V}(s) \text{ and } \lambda_H(s) = p_H(s) \frac{d\eta}{d\nu_H}(s), \quad (4.20)$$

where  $\eta = \nu_V * \nu_H$  is the convolution product of  $\nu_V$  and  $\nu_H$ , i.e. for any  $A \in \mathcal{B}(\mathbb{R})$ ,

$$\eta(A) = \int_{\mathbb{R}^2} \mathbf{1}_{t+s \in A} d\nu_V(t) d\nu_H(s).$$

**(G5)** The *division kernel*  $F$  satisfies, for any  $s \in \mathbb{R}$ , for any  $t \in \mathbb{R}$ ,

$$F(s, A) = \int_A \frac{d\nu_V^{(t)}}{d\eta}(s) d\nu_H(t), \quad (4.21)$$

where  $\nu_V^{(t)}$  the  $t$ -translated measure<sup>5</sup> of  $\nu_V$ . The probability kernel  $F$  can be also seen as the regular conditional probability of  $X \sim \nu_H$  with respect to  $\sigma(X + Y)$  where  $Y \sim \nu_V$  and  $Y$  is independent of  $X$  as defined in [Dur19, Section 4.1.3].

As in Theorems 4.1 and 4.2, a PKS with parameters  $(\lambda_0, \lambda_V, \lambda_H, p_0, p_V, p_H, \tau_V, \tau_H, F)$  such that there exist two non-zero finite measures  $\nu_V$  and  $\nu_H$  on  $\mathbb{R}$  that satisfy the conditions **(G1)**, **(G2)**, **(G3)**, **(G4)** and **(G5)** is well defined and reversible under the initial condition  $(\mathcal{C}_X, \mathcal{C}_Y)$  defined in equation (4.4).

The proof of the reversibility in that case is much more technical than it is for the Lebesgue and discrete cases because we do not have access to a reference translation invariant measure against which both weight measures  $\nu_V$  and  $\nu_H$  are absolutely continuous. Thus, the proof of this statement is omitted in the chapter.

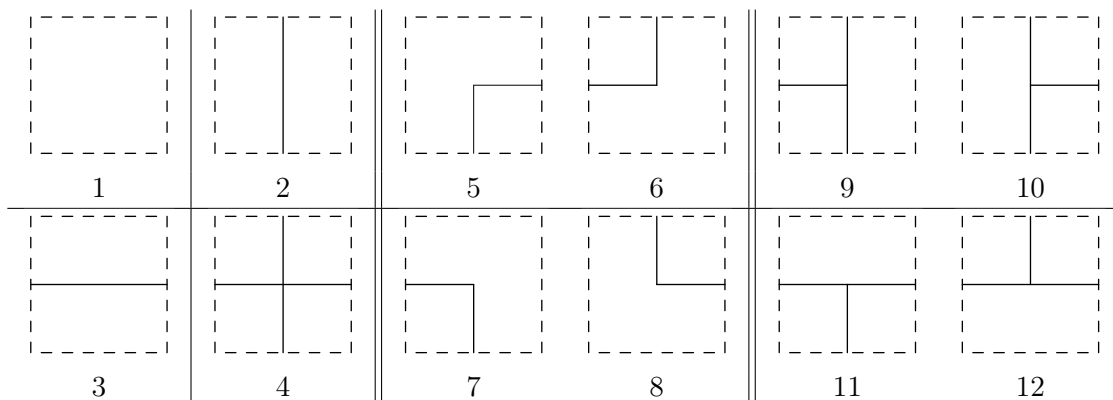
## 4.4 Heuristic

Conditions **(L1-L5)**, **(D1-D5)** and **(G1-G5)** seem technical and somewhat ad hoc at first glance. However, they appear naturally when studying the PKS “locally”. Before providing the rigorous (and technical) proof of the main theorems in the next section, we give below a heuristic argument that hopefully shed some light on the necessity of the assumptions.

---

<sup>5</sup>The  $t$ -translated measure  $\nu_V^{(t)}$  of  $\nu_V$  is the measure defined by, for any  $A \in \mathcal{B}(\mathbb{R})$ ,  $\nu_V^{(t)}(A) = \nu_V(\{x - t \in \mathbb{R} : x \in A\})$



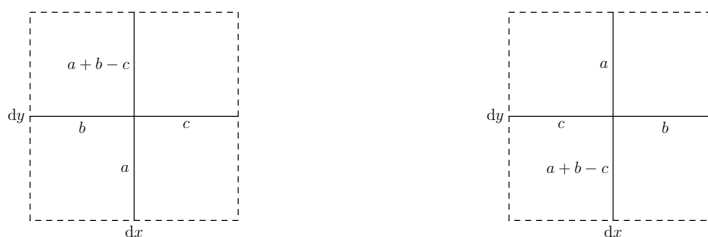


**Figure 4.8:** The twelve local configurations: 4 self-dual configurations (1-4) and 4 pairs (5-12).

We look at the PKS process inside an “infinitely small” rectangle  $dx \times dy$  so that at most one event (split, turn, ...) can occur inside this region simultaneously. In order for the PKS process to be reversible, the probability of any elementary event must be equal to the probability of the corresponding event when the  $dx \times dy$  rectangle is rotated by 180 degrees. Thus, we can consider, in turn, each of the 12 possible elementary events pictured in Figure 4.8 and check the relations that they entail on the parameters of the process. The first 3 elementary events (empty square, single vertical line and single horizontal line) are symmetric by rotation of 180 degrees so they entail no condition.

### 4.4.1 Crossing

Let us consider the elementary crossing event 4 of Figure 4.8 and its rotation by 180 degrees.



In order for reversibility to hold true, both configuration should appear with equal probability. This means that, given three test functions  $u, v$  and  $w$ , the expectation of  $u(a)v(b)w(c)$  should be the same on both events. In view of the PKS dynamics, this entails that:

$$\iiint u(a)v(b)w(c)d\nu_V(a)d\nu_H(b)F(a+b, dc) = \iiint u(a)v(b)w(c)d\nu_V(a+b-c)d\nu_H(c)F(a+b, db). \tag{4.22}$$

In particular, in the case of discrete measures, the above equation implies the equality

$$\nu_V(a)\nu_H(b)F(a+b, c) = \nu_V(a+b-c)\nu_H(c)F(a+b, b) \tag{4.23}$$

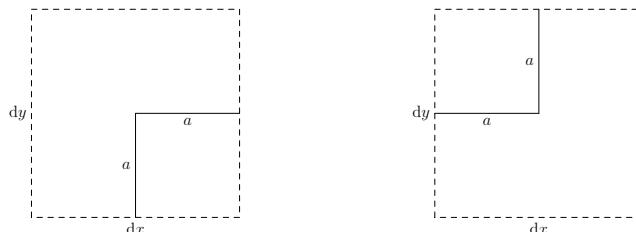
whereas, in the Lebesgue case, it implies the relation on the densities:

$$g_V(a)g_H(b)f(a+b, c) = g_V(a+b-c)g_H(c)f(a+b, b). \quad (4.24)$$

It is not difficult to check (excluding possible degenerates cases) that equation (4.24) (resp. (4.23) and (4.22)) is equivalent to **(L5)** (resp. **(D5)** and **(G5)**).

### 4.4.2 Horizontal turn versus vertical turn

Let us now look at the complementary events 5 and 6 of Figure 4.8.



Just as in the previous section, the reversibility property implies that, for any test functions  $u$ , it must hold that

$$\int u(a)\tau_V(a)d\nu_V(a) = \int u(a)\tau_H(a)d\nu_H(a).$$

Therefore, in the discrete case, we have

$$\nu_V(a)\tau_V(a) = \nu_H(a)\tau_H(a).$$

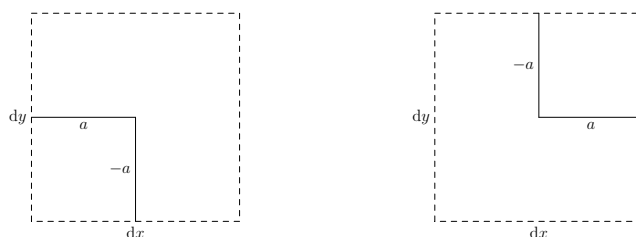
and in the Lebesgue case

$$g_V(a)\tau_V(a) = g_H(a)\tau_H(a).$$

The previous three equations are equivalent to **(G3)**, **(D3)** and **(L3)** respectively.

### 4.4.3 Spontaneous creation versus annihilation

We consider the complementary events 7 and 8 of Figure 4.8.



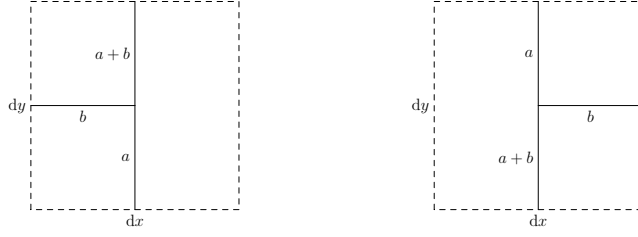
In order for reversibility to hold true, they should have the same probability. In particular,  $a$  in the picture above must be an atom of the measure  $\nu_H$  and  $-a$  an atom of  $\nu_V$ . In the Lebesgue case, the probability of the left event is zero which implies that the creation rate  $\lambda_0 = 0$  which is exactly **(L1)**. In the general case, we find that, for any such atom  $a$ , we must have

$$\nu_V(-a)\nu_H(a)p_0 = \lambda_0 F(0, a)$$

which is, already assuming **(G5)** (resp. **(D5)**), equivalent to **(G1)** (resp. **(D1)**).

#### 4.4.4 Split versus coalescence

Let finally consider the two elementary events 9 and 10 of Figure 4.8:



Using the same argument as before, we now find that, for any two test functions  $u$  and  $v$ , it must hold that

$$\iint u(a)v(b)p_V(a+b)d\nu_V(a)d\nu_H(b) = \iint u(a)v(b)d\nu_V(a+b)\lambda_V(a+b)F(a+b, db). \quad (4.25)$$

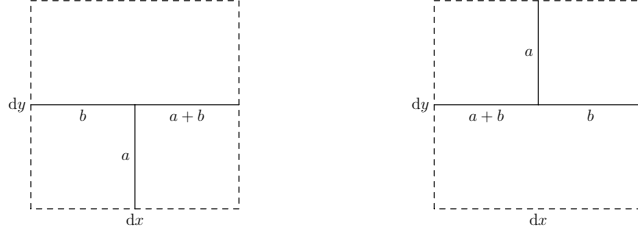
In the case of discrete measures, the above equation implies the equality

$$\nu_V(a)\nu_H(b)p_V(a+b) = \nu_V(a+b)\lambda_V(a+b)F(a+b, b) \quad (4.26)$$

whereas, in the Lebesgue case, it implies the relation on the densities:

$$g_V(a)g_H(b)p_V(a+b) = g_V(a+b)\lambda_V(a+b)f(a+b, b). \quad (4.27)$$

Similarly, considering now events 11 and 12 of Figure 4.8:



We find that

$$\iint u(a)v(b)p_H(a+b)d\nu_V(a)d\nu_H(b) = \iint u(a)v(b)d\nu_H(a+b)\lambda_H(a+b)F(a+b, db) \quad (4.28)$$

which translate, in the discrete case to

$$\nu_V(a)\nu_H(b)p_H(a+b) = \nu_H(a+b)\lambda_H(a+b)F(a+b, b) \quad (4.29)$$

and in the Lebesgue case to

$$g_V(a)g_H(b)p_H(a+b) = g_H(a+b)\lambda_H(a+b)f(a+b, b). \quad (4.30)$$

Under the assumption that **(L5)** holds, then (4.27) and (4.30) are equivalent to **(L2)** and **(L4)**. The same holds true for the discrete and general cases.

The previous analysis shows that the five conditions of the previous section are indeed necessary (excluding maybe some degenerate cases) to have reversibility of the PKS process on the microscopic scale. In the next section, we prove that those conditions are actually sufficient and imply, in fact, reversibility on the macroscopic scale.

## 4.5 Proof of reversibility

In this section, we prove Theorems 4.1 and 4.2 of this article. In the two following sections, we first prove these results assuming further that the rates are uniformly bounded. Then, we explain in the last section how we can bootstrap the argument to obtain the general results.

### 4.5.1 Proof of Theorem 4.1 with uniformly bounded rates

Consider a PKS in the Lebesgue case with parameters  $(\lambda_0, \lambda_V, \lambda_H, p_0, p_V, p_H, \tau_V, \tau_H, F)$ , such that there exist two non-zero finite measures on  $\mathbb{R}$  with densities  $g_V$  and  $g_H$  such that the conditions (4.6), (4.7), (4.8), (4.9) and (4.10) hold. We start this PKS process with the initial condition  $(\mathcal{C}_X, \mathcal{C}_Y)$  as defined in equation (4.5).

In this section, we assume that the rates of the PKS are uniformly bounded, that is to say they satisfy condition (4.2), which implies that the PKS is well defined a.s. by Proposition 4.1. Hence, we just need to show the reversibility of the PKS. The uniformly bounded rates assumption will be relaxed in Section 4.5.3.

Recalling the definition of a drawing and of reversibility in Section 4.3, we want to prove that, for any non-negative measurable function  $\Phi : \mathcal{D}_{a,b} \rightarrow \mathbb{R}_+$ ,

$$\mathbb{E}[\Phi(\mathbf{D})] = \mathbb{E}[\Phi(\widehat{\mathbf{D}})] \quad (4.31)$$

which exactly states that  $\mathbf{D}$  and  $\widehat{\mathbf{D}}$  have the same law. However, the set of all drawings which is infinite dimensional is not a very convenient space to work with. To overcome this difficulty, we partition the set of drawings according to their combinatorial nature which will enable us to rewrite the expectation above as a sum of expectations over finite dimensional spaces. Before doing so, we introduce some notation and definitions that will be helpful to understand the combinatorial structure of a drawing.

It will be convenient to represent a *weighted vertical* (resp. *horizontal*) *segment*  $\sigma$  as a triplet  $(\sigma_-, \sigma_+, s)$  where the endpoints are  $\sigma_- = (x, y_-)$  (resp.  $(x_-, y)$ ) and  $\sigma_+ = (x, y_+)$  (resp.  $(x_+, y)$ ) with  $y_- < y_+$  (resp.  $x_- < x_+$ ) and the weight is  $s \in \mathbb{R}$ .

**Types of nodes.** We can define eleven types of nodes that correspond to events in the dynamics occurring inside the box as well as events on the boundary of the domain. For each type of node, we introduce a notation as a pictogram for the set of all nodes of this type.

- **Vertical entry:** a vertical entry is a boundary point  $(x, 0)$  on the bottom side of the box which has an outgoing segment  $\sigma \in D$ ,  $\sigma = ((x, 0), (x, \cdot), \cdot)$ . We denote this set by  $\square_{\downarrow}$ . Remark that  $\square_{\downarrow} = \{(x, 0) : \exists s \in \mathbb{R}, ((x, 0), s) \in \mathcal{C}_X\}$ .
- **Vertical exit:** a vertical exit is a boundary point  $(x, b)$  on the top side of the box which has an outgoing segment  $\sigma \in D$ ,  $\sigma = ((x, \cdot), (x, b), \cdot)$ . We denote this set by  $\square_{\uparrow}$ .

- **Vertical split:** a vertical split of  $D$  is a point  $(x, y)$  where 3 segments are meeting from the south, north and east, i.e. there exist  $\sigma_S, \sigma_N, \sigma_E \in D$  such that  $\sigma_S = ((x, \cdot), (x, y), \cdot)$ ,  $\sigma_N = ((x, y), (x, \cdot), \cdot)$  and  $\sigma_E = ((x, y), (\cdot, y), \cdot)$ . This corresponds to case  $2_V(a)$  in the dynamics defined in Section 4.2. We denote this set by  $\square$ .
- **Vertical turn:** a vertical turn of  $D$  is a point  $(x, y)$  where 2 segments are meeting from the south and east, i.e. there exist  $\sigma_S, \sigma_E \in D$  such that  $\sigma_S = ((x, \cdot), (x, y), \cdot)$  and  $\sigma_E = ((x, y), (\cdot, y), \cdot)$ . This corresponds to case  $2_V(b)$  in the dynamics. We denote this set by  $\square$ .
- **Vertical coalescence:** a vertical coalescence is a point  $(x, y)$  where 3 segments are meeting from the west, south and north, i.e. there exist  $\sigma_W, \sigma_S, \sigma_N \in D$  such that  $\sigma_W = ((\cdot, y), (x, y), \cdot)$ ,  $\sigma_S = ((x, \cdot), (x, y), \cdot)$  and  $\sigma_N = ((x, y), (x, \cdot), \cdot)$ . This corresponds to case 3(a) in the dynamics. We denote this set by  $\square$ .

For all these kinds of nodes, we also define their obvious horizontal counterpart: **horizontal entry**  $\square$ , **horizontal exit**  $\square$ , **horizontal split**  $\square$ , **horizontal turn**  $\square$  and **horizontal coalescence**  $\square$ . Finally, we define a last kind of nodes:

- **Crossing:** a crossing is a point  $(x, y)$  where 4 segments are meeting. This corresponds to case 3(d) of the dynamics. Alternatively, this event can be interpreted as a coalescence immediately followed by a split. We denote this set by  $\square$ .

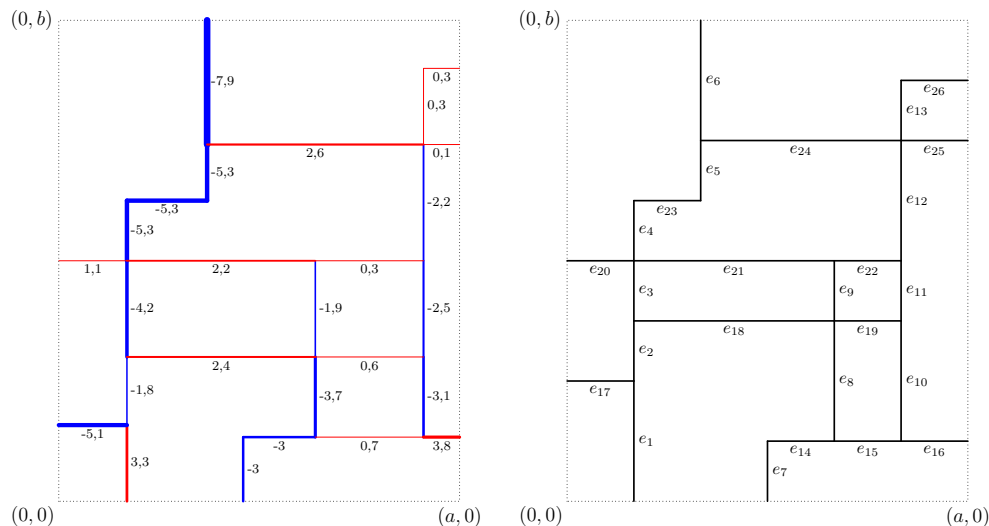
**Skeleton and parametrization of a drawing.** We introduce the notion of *skeleton* of a drawing which will be instrumental in the rest of the proof. We say that two drawings  $D, D' \in \mathcal{D}$  have the same skeleton, and denote it by  $D \sim D'$ , if there exist two increasing functions  $\psi_X$  from  $[0, a]$  to  $[0, a]$  and  $\psi_Y$  from  $[0, b]$  to  $[0, b]$  such that for any weighted segment  $\sigma = ((x_-, y_-), (x_+, y_+), s) \in D$ , there exists a unique  $s' \in \mathbb{R}$  such that  $\psi(\sigma) := ((\psi_X(x_-), \psi_Y(y_-)), (\psi_X(x_+), \psi_Y(y_+)), s') \in D'$ .

In other words, the skeleton represents the “combinatorial” structure of a drawing where we forget about the exact positions and weights of segments, so that two drawings with the same skeleton can be mapped from one to the other by changes of space and weight. Thus, two drawings with the same skeleton  $S$  have the same numbers of segments  $\ell = \ell(S)$  as well as the same number of nodes of each type. Furthermore, a skeleton induces a graph whose edges will be denoted  $(e_1, \dots, e_\ell)$  (for some arbitrary ordering) in the following. An illustration of a drawing and its skeleton is given on Figure 4.9.

Let us note that a drawing  $D$ , given its skeleton  $S$ , is uniquely determined once we specify the spatial positions of its segments together with their weights. Thus, we shall now identify the set of all drawings  $D$  with skeleton  $S$  as a subset of  $\mathbb{R}^{m+n+\ell}$  and we shall represent a drawing  $D$  by a vector

$$(x_1, \dots, x_m, y_1, \dots, y_n, s_1, \dots, s_\ell) \in (0, a)^m \times (0, b)^n \times \mathbb{R}^\ell \quad (4.32)$$

where the  $(x_i)$ 's are the  $m := |\square| + |\square| + |\square|$  horizontal coordinates of the points in  $\square \cup \square \cup \square$  ordered increasingly, the  $(y_i)$ 's are the  $n := |\square| + |\square| + |\square|$  vertical coordinates of the points in  $\square \cup \square \cup \square$  ordered increasingly and the  $(s_i)$ 's are the weights of the segments corresponding to the edges  $(e_i)$  of the skeleton  $S$ .



**Figure 4.9:** An example of a drawing and, on its right, its skeleton. In this example,  $\ell = 26$ .

However, not all such vectors represent a valid drawing since the Kirchhoff's node law induces relations between segment weights, so the dimension of the space generated by all valid vectors is smaller than  $m + n + \ell$ . More precisely, its dimension is  $m + n + d$  where

$$d = d(S) = \ell - (|\square| + |\square| + |\square| + |\square| + |\square| + |\square| + |\square|). \quad (4.33)$$

Indeed, we notice that each internal node (i.e. a node belonging to  $\square \cup \square \cup \square \cup \square \cup \square \cup \square \cup \square$ ) adds an independent linear constraint, coming from Kirchhoff's node law, which decreases the space dimension by 1. We can now derive the following lemma:

**Lemma 4.1.** *For any skeleton  $S$ , the dimension of the set of admissible weights of a drawing  $D$  with a given skeleton  $S$  is equal to:*

$$d(S) = |\square| + |\square| + |\square| + |\square| + |\square|.$$

*Proof.* By counting the number of half-edges of  $S$ , which is equal to  $2\ell(S)$ , we get:

$$2\ell(S) = (|\square| + |\square| + |\square| + |\square|) + 2(|\square| + |\square|) + 3(|\square| + |\square| + |\square| + |\square|) + 4|\square|.$$

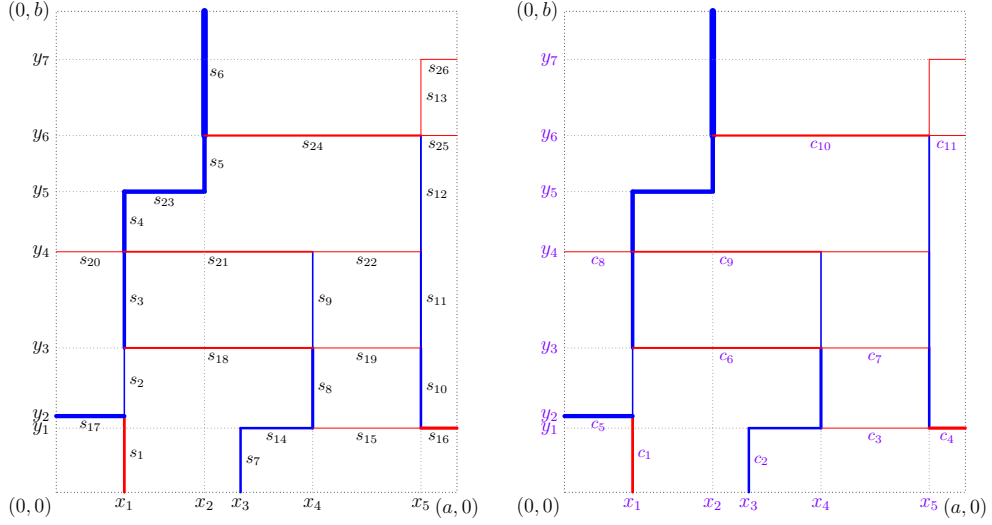
Indeed, each node in  $\square \cup \square \cup \square \cup \square$  contributes for 1 half-edge, each node of  $\square \cup \square$  for 2 half-edges, each node in  $\square \cup \square \cup \square \cup \square$  for 3 half-edges and each node in  $\square$  for 4 half-edges.

Moreover, remark that

$$\begin{cases} |\square| + |\square| + |\square| = |\square| + |\square| + |\square| & (\text{because both are equal to } m), \\ |\square| + |\square| + |\square| = |\square| + |\square| + |\square| & (\text{because both are equal to } n). \end{cases}$$

Consequently,

$$2\ell(S) = 2(|\square| + |\square| + |\square| + |\square|) + 2|\square| + |\square| + 2|\square| + |\square| + 2|\square|.$$



**Figure 4.10:** On the left, a drawing  $D$  with all its coordinates in  $\mathbb{R}^{m+n+\ell}$  and, on the right, the same drawing with its free  $m+n+d$  coordinates chosen as in Section 4.5.1. In this example,  $\ell = 26$ ,  $m = 5$ ,  $n = 7$  and  $d = 11$ .

Then, by using equation (4.33),

$$\begin{aligned} d(S) &= (|\square| + |\square| + |\square| + |\square| + 2|\square| + |\square| + 2|\square| + |\square| + 2|\square|) \\ &\quad - (|\square| + |\square| + |\square| + |\square| + |\square| + |\square| + |\square|) \\ &= |\square| + |\square| + |\square| + |\square| + |\square|. \end{aligned} \quad \square$$

Define a *parametrization* of a skeleton  $S$  by selecting  $d$  edges  $(e_{\rho(1)}, \dots, e_{\rho(d)})$  where  $\rho$  is an injective mapping from  $\{1, \dots, d\}$  to  $\{1, \dots, \ell\}$  such that the knowledge of the weights on the edges  $e_{\rho(1)}, \dots, e_{\rho(d)}$  together with Kirchhoff's node law entirely defines the weights of all edges in the skeleton. In particular, a parametrization defines an injective linear mapping  $\mathfrak{D}_{S,\rho} : \mathbb{R}^{d(S)} \rightarrow \mathbb{R}^{\ell(S)}$  whose image is the vector space generated by valid drawing vectors (i.e. satisfying Kirchhoff's law at each node), and where the  $j$ th coordinate corresponds to the weight  $s_{\rho(j)}$  on the edge  $e_{\rho(j)}$ , i.e.

$$\mathfrak{D}_{S,\rho}((c_j)_{j=1..d}) = (s_i)_{i=1..\ell} \quad (4.34)$$

with  $s_{\rho(j)} = c_j$  for any  $j$ .

**A parametrization related to the dynamics.** A particular parametrization  $\rho_S$  related to the dynamics of the PKS defined in Section 4.2 is obtained by selecting only the vertical (resp. horizontal) edges whose starting point belongs to  $\square$  (resp.  $\square \cup \square \cup \square \cup \square$ ). In term of the dynamics, this means that we keep track of the weights of the entry points and of the weights of the eastern edges when split or crossing events occur.

It is clear that this subset of edges yields a valid parametrization of a drawing since this family has the correct cardinal  $d(S)$  and since all weights in the drawing can be reconstructed iteratively by following the dynamics of the process. See Figure 4.10 for an illustration.

Using the parametrization  $\rho_S$ , we can decompose the expectation  $\mathbb{E}[\Phi(\mathbf{D})]$  in equation (4.31) with the following formula:

$$\mathbb{E}[\Phi(\mathbf{D})] = \sum_{S \in \mathcal{D}/\sim} \int_{\mathbb{R}^{m+n+d}} dx dy dc \Phi((x, y, \mathfrak{D}_{S, \rho_S}(c))) \alpha_S((x, y, \mathfrak{D}_{S, \rho_S}(c))), \quad (4.35)$$

where  $\alpha_S$  is to be thought of as the “density” of the drawing  $\mathbf{D}$  on the event that its skeleton is  $S$  (and when using the parametrization  $\rho_S$  described previously).

Before expressing  $\alpha_S$ , we introduce the function  $q : \mathbb{R} \rightarrow \mathbb{R}_+$ , we will refer as the *turn function* defined, for any  $s \in \mathbb{R}$ , by

$$q(s) := \tau_V(s) \sqrt{\frac{g_V(s)}{g_H(s)}} = \tau_H(s) \sqrt{\frac{g_H(s)}{g_V(s)}} \quad \text{by equation (4.8),}$$

with the convention  $0/0 = 0$  in the formula above. Beyond simplifying the expression of  $\alpha_S$ , the introduction of the additional function  $q$  will be of great help in the proof of the invariance of  $\alpha_S$  by the reverse operation  $\hat{\cdot}$  (further Lemma 4.4), since it will turn out to be itself invariant by this operation.

**Lemma 4.2.** *For any skeleton  $S$ , and any drawing  $D$  whose skeleton is  $S$  and identified to*

$$(x_1, \dots, x_m, y_1, \dots, y_n, s_1, \dots, s_\ell),$$

see equation (4.32), we have

$$\begin{aligned} \alpha_S(D) &= (\mathbf{1}_{0 < x_1 < x_2 < \dots < x_m < a}) e^{-\left(\int_{\mathbb{R}} g_V(s) ds\right)a} (\mathbf{1}_{0 < y_1 < y_2 < \dots < y_n < b}) e^{-\left(\int_{\mathbb{R}} g_H(s) ds\right)b} \\ &\quad \left( \prod_{\sigma = ((x_-, y_-), (x_+, y_+), s) \in D} \left[ \mathbf{1}_{x_- = x_+} g_V(s) \left( \mathbf{1}_{(x_-, y_-) \in \square \cup \square \cup \square \cup \square} \mathbf{1}_{(x_+, y_+) \in \square \cup \square} \right) \right. \right. \\ &\quad \left. \sqrt{q(s)g_V(s)} \mathbf{1}_{(x_-, y_-) \in \square} \mathbf{1}_{(x_+, y_+) \in \square} p_V(s) \left( \mathbf{1}_{(x_-, y_-) \in \square} \mathbf{1}_{(x_+, y_+) \in \square} \right) e^{-(\tau_V(s) + \lambda_V(t))(y_+ - y_-)} \right. \\ &\quad \left. \left. + \mathbf{1}_{y_- = y_+} g_H(s) \left( \mathbf{1}_{(x_-, y_-) \in \square \cup \square \cup \square \cup \square} \mathbf{1}_{(x_+, y_+) \in \square \cup \square} \right) \right. \right. \\ &\quad \left. \left. \sqrt{q(s)g_H(s)} \mathbf{1}_{(x_-, y_-) \in \square} \mathbf{1}_{(x_+, y_+) \in \square} p_H(s) \left( \mathbf{1}_{(x_-, y_-) \in \square} \mathbf{1}_{(x_+, y_+) \in \square} \right) e^{-(\tau_H(s) + \lambda_H(t))(x_+ - x_-)} \right] \right) \\ &\quad \left( \prod_{(x, y) \in \square} \frac{1 - p_V(s_W + s_S) - p_H(s_W + s_S)}{h(s_W + s_S)} \mathbf{1}_{((x, \cdot), (x, y), s_S) \in D} \mathbf{1}_{((\cdot, y), (x, y), s_W) \in D} \right). \end{aligned}$$

*Proof.* The formula above is nothing more than a rearrangement of a product of terms where each one represents the probability of a local event which, put together, ensures that  $D$  is indeed a drawing with skeleton  $S$  chosen according to the Poisson-Kirchhoff dynamics. Let us analyse each term separately.

First, the indicator functions  $\mathbf{1}_{0 < x_1 < \dots < x_{|\square| + |\square| + |\square|} < a}$  and  $\mathbf{1}_{0 < y_1 < \dots < y_{|\square| + |\square| + |\square|} < b}$  ensure that the  $(x_i)_i$ 's and  $(y_j)_j$ 's are correctly ordered.



Secondly, the terms  $e^{-\left(\int_{\mathbb{R}} g_V(s) ds\right)a}$  and  $e^{-\left(\int_{\mathbb{R}} g_H(s) ds\right)b}$  are respectively equal to the probabilities that there is no other entry on the bottom and left boundaries  $[0, a] \times \{0\}$  and  $\{0\} \times [0, b]$ .

Thirdly, each segment  $\sigma = ((x_-, y_-), (x_+, y_+), s) \in D$  contributes to the product through the terms

$$e^{-(\tau_V(s) + \lambda_V(t))(y_+ - y_-)} \text{ or } e^{-(\tau_H(s) + \lambda_H(t))(x_+ - x_-)},$$

which represents the probability of non-splitting and non-turning along the segment  $\sigma$  depending on whether it is vertical or horizontal.

Finally, we look at the contribution to the density of each node  $(x, y)$  and show how it can be decomposed into factors associated to each segment adjacent to the node, and to the node itself when it is a crossing, i.e. when  $(x, y)$  belongs to  $\boxplus$ . We distinguish the following cases with respect to the node type, using the notation  $s_N, s_E, s_S$  and  $s_W$  for the sizes of the northern, eastern, southern and western segments which are adjacent to  $(x, y)$ :

- if  $(x, y) \in \boxminus$ , its northern segment  $\sigma_N$ , which is its only adjacent segment, gets the contribution  $g_V(s_N) = g_V(s_N)^{\mathbf{1}(\sigma_N) - \in \boxminus}$  coming from the vertical entry of the boundary PPP.
- if  $(x, y) \in \boxplus$ , no contribution is assigned to the southern segment, which is its only adjacent segment, since it is an exit point.
- if  $(x, y) \in \boxtimes$ , the term  $\lambda_V(s_S) f(s_S, s_E) = p_V(s_S) \frac{g_H(s_E) g_V(s_S - s_E)}{g_V(s_S)}$  splits into three terms which are distributed on the three adjacent segments to  $(x, y)$  as follows:
  - the term  $g_H(s_E) = g_H(s_E)^{\mathbf{1}(\sigma_E) - \in \boxtimes}$  on the eastern segment  $\sigma_E$ ,
  - the term  $g_V(s_S - s_E) = g_V(s_N) = g_V(s_N)^{\mathbf{1}(\sigma_N) - \in \boxtimes}$  on the northern segment  $\sigma_N$ ,
  - the term  $p_V(s_S) / g_V(s_S) = p_V(s_S)^{\mathbf{1}(\sigma_S) + \in \boxtimes} g_V(s_S)^{-\mathbf{1}(\sigma_S) + \in \boxtimes}$  on the western segment  $\sigma_S$ .
- if  $(x, y) \in \boxdot$ , the term

$$\tau_V(s_S) = q(s_S) \sqrt{\frac{g_H(s_S)}{g_V(s_S)}} = \sqrt{q(s_E) g_H(s_E)} \frac{\sqrt{q(s_S) g_V(s_S)}}{g_V(s_S)}$$

splits into two terms which are distributed on the two segments adjacent to  $(x, y)$  as follows:

- the term  $\sqrt{q(s_E) g_H(s_E)} = \sqrt{q(s_E) g_H(s_E)}^{\mathbf{1}(\sigma_E) - \in \boxdot}$  on the eastern segment  $\sigma_E$ ,
- the term  $\sqrt{q(s_S) g_V(s_S)} / g_V(s_S) = \sqrt{q(s_S) g_V(s_S)}^{\mathbf{1}(\sigma_S) + \in \boxdot} g_V(s_S)^{-\mathbf{1}(\sigma_S) + \in \boxdot}$  on the southern segment  $\sigma_S$ ;
- if  $(x, y) \in \boxdot$ , the term  $p_H(s_S + s_W) = p_H(s_N) = p_H(s_N)^{\mathbf{1}(\sigma_N) - \in \boxdot}$  is assigned to its northern segment  $\sigma_N$ . Its southern and western adjacent segments get no contribution.
- The contributions of nodes of horizontal type  $\boxleftarrow$ ,  $\boxrightarrow$ ,  $\boxup$ ,  $\boxdown$  and  $\boxtimes$  are decomposed analogously as the nodes of vertical type above.
- if  $(x, y) \in \boxplus$ , the term  $(1 - p_V(s_S + s_W) - p_H(s_S + s_W)) \frac{g_V(s_N) g_H(s_E)}{h(s_S + s_W)}$  splits into three terms which are distributed as follows:

- the term  $g_V(s_N) = g_V(s_N)^{\mathbf{1}^{(\sigma_N) - \epsilon_{\boxplus}}}$  on the northern segment  $\sigma_N$ ,
- the term  $g_H(s_E) = g_H(s_E)^{\mathbf{1}^{(\sigma_E) - \epsilon_{\boxplus}}}$  on the eastern segment  $\sigma_E$ ,
- the term  $\frac{1 - p_H(s_S + s_W) - p_H(s_S + s_W)}{h(s_S + s_W)}$  is attached to the node itself,
- its southern and western adjacent segments get no contribution.  $\square$

**Change of parametrization.** Formula (4.35) presents a decomposition of the expectation of  $\Phi(\mathbf{D})$  in terms of the special parametrization  $\rho_S$  defined above. However, this formula is in fact valid for any parametrization  $\rho$  thanks to the following lemma:

**Lemma 4.3.** *Let  $\rho$  and  $\rho'$  denote two parametrizations of a skeleton  $S$  with respective linear mappings  $\mathfrak{D}_{S,\rho}$  and  $\mathfrak{D}_{S,\rho'}$  from  $\mathbb{R}^d$  to  $\mathbb{R}^\ell$ . We have*

$$\left| \det\left(\mathfrak{D}_{S,\rho}^{-1} \circ \mathfrak{D}_{S,\rho'}\right) \right| = 1.$$

*Proof.* Notice that if  $\rho$  and  $\rho'$  have the same image, then the application  $\mathfrak{D}_{S,\rho}^{-1} \circ \mathfrak{D}_{S,\rho'}$  is just a permutation and the result is trivial. We will prove the lemma in the case where  $\rho$  and  $\rho'$  differ only by one coordinate. Then, the general case will follow by choosing a finite sequence of parametrizations where two consecutive parametrizations differ by exactly one coordinate.

Take now  $\rho$  and  $\rho'$  such that they differ only by one coordinate. Without loss of generality, we can assume that for all  $i \leq d-1$ ,  $\rho(i) = \rho'(i)$ . Consider the set of edges  $e$  in  $S$  such that  $(e_{\rho(1)}, \dots, e_{\rho(d-1)}, e)$  is a parametrization of  $S$ . This set is necessarily connected. Indeed, if this was not the case then we could pick an edge in each connected component and add it to the parametrization since, according to the Kirchhoff's node law, setting the weight of an edge can only constrain the weight of edges in the same connected component. But this would yield a parametrization with more than  $d$  edges, which is absurd.

Consequently, there exists a path  $(e^{(1)} = \rho(d), \dots, e^{(k)} = \rho'(d))$  such that, for any  $i$ ,  $e^{(i)}$  and  $e^{(i+1)}$  are adjacent in  $S$ . Now, for any  $i$ , let  $\mathfrak{D}^{(i)} = (e_{\rho(1)}, \dots, e_{\rho(d-1)}, e^{(i)})$ . Finally, we just need to check that  $|\det((\mathfrak{D}^{(i)})^{-1} \circ \mathfrak{D}^{(i+1)})| = 1$ . This is clearly the case because, according to Kirchhoff's law around the node shared by  $e^{(i)}$  and  $e^{(i+1)}$ , we have  $s(e^{(i)}) = \pm s(e^{(i+1)}) + \sum_{i=1}^{d-1} \lambda_i s(e_{\rho(i)})$  for some fixed  $(\lambda_i)$ .  $\square$

**Corollary 4.2.** *The formula (4.35) still holds true when replacing  $\rho_S$  by any parametrization  $\rho$ .*

*Proof.* Let  $\rho$  be any parametrization of  $S$ . Doing the change of variable  $c' = \mathfrak{D}_{S,\rho}^{-1} \circ \mathfrak{D}_{S,\rho_S}(c)$  in equation (4.35) and applying Lemma 4.3, we get

$$\begin{aligned} \mathbb{E}[\Phi(\mathbf{D})] &= \sum_{S \in \mathcal{D}/\sim} \int_{\mathbb{R}^{m+n+d}} dx \, dy \, dc \, \Phi((x, y, \mathfrak{D}_{S,\rho_S}(c))) \, \alpha_S((x, y, \mathfrak{D}_{S,\rho_S}(c))) \\ &= \sum_{S \in \mathcal{D}/\sim} \int_{\mathbb{R}^{m+n+d}} dx \, dy \, dc' \, \Phi((x, y, \mathfrak{D}_{S,\rho}(c'))) \, \alpha_S((x, y, \mathfrak{D}_{S,\rho}(c'))). \end{aligned} \quad \square$$

$(x, y) \in D$	$\begin{array}{ c } \hline \square \\ \hline \end{array}$	$\begin{array}{ c } \hline \top \\ \hline \end{array}$	$\begin{array}{ c } \hline \vdash \\ \hline \end{array}$	$\begin{array}{ c } \hline \dashv \\ \hline \end{array}$	$\begin{array}{ c } \hline \sqcup \\ \hline \end{array}$	$\begin{array}{ c } \hline \sqcap \\ \hline \end{array}$
$(a-x, b-y) = (\hat{x}, \hat{y}) \in \hat{D}$	$\begin{array}{ c } \hline \hat{\square} \\ \hline \end{array}$	$\begin{array}{ c } \hline \hat{\top} \\ \hline \end{array}$	$\begin{array}{ c } \hline \hat{\vdash} \\ \hline \end{array}$	$\begin{array}{ c } \hline \hat{\dashv} \\ \hline \end{array}$	$\begin{array}{ c } \hline \hat{\sqcup} \\ \hline \end{array}$	$\begin{array}{ c } \hline \hat{\sqcap} \\ \hline \end{array}$

**Table 4.1:** Correspondence between each type of vertical node, and of crossing nodes as viewed in  $D$  or in  $\hat{D}$ . Similar correspondences hold for horizontal nodes.

**Reversibility.** The last ingredient we need to prove the reversibility of the model is the invariance of the density  $\alpha_S$  by the rotation of 180 degrees.

**Lemma 4.4.** *For any skeleton  $S$ , for any drawing  $D$  with skeleton  $S$ , we have  $\alpha_{\hat{S}}(\hat{D}) = \alpha_S(D)$ .*

*Proof.* The function  $\alpha_S(D)$  only depends on the length and weight of the segments and on the crossings of the drawing  $D$ . To any segment  $\sigma = ((x_-, y_-), (x_+, y_+), s) \in D$ , we associate its reverse segment  $\hat{\sigma} = ((\hat{x}_-, \hat{y}_-), (\hat{x}_+, \hat{y}_+), \hat{s}) \in \hat{D}$  where  $\hat{x}_- = a - x_+$ ,  $\hat{x}_+ = a - x_-$ ,  $\hat{y}_- = b - y_+$ ,  $\hat{y}_+ = b - y_-$ , and  $\hat{s} = s$  (obviously, the weight of a segment does not change by a rotation of 180 degrees).

In particular, for any  $i$ ,  $\hat{x}_i = a - x_{m+1-i}$  and  $\hat{y}_i = b - y_{n+1-i}$ . Hence, the terms in the first line of the expression of  $\alpha_S(D)$  in Lemma 4.2 and its  $\alpha_{\hat{S}}(\hat{D})$ -counterpart coincide.

Consider now a vertical segment  $\sigma$ . Its contribution in  $\alpha_S(D)$  equals

$$g_V(s) \left( \mathbf{1}_{(x_-, y_-) \in \square \cup \top \cup \vdash \cup \dashv} - \mathbf{1}_{(x_+, y_+) \in \square \cup \top} \right) \sqrt{q(s)g_V(s)} \mathbf{1}_{(x_-, y_-) \in \square} + \mathbf{1}_{(x_+, y_+) \in \square} p_V(s) \left( \mathbf{1}_{(x_-, y_-) \in \square} + \mathbf{1}_{(x_+, y_+) \in \square} \right) e^{-(\tau_V(s) + \lambda_V(s))(y_+ - y_-)}.$$

The contribution of the reverse segment  $\hat{\sigma}$  to  $\alpha_{\hat{S}}(\hat{D})$  is equal to

$$g_V(\hat{s}) \left( \mathbf{1}_{(\hat{x}_-, \hat{y}_-) \in \hat{\square} \cup \hat{\top} \cup \hat{\vdash} \cup \hat{\dashv}} - \mathbf{1}_{(\hat{x}_+, \hat{y}_+) \in \hat{\square} \cup \hat{\top}} \right) \sqrt{q(\hat{s})g_V(\hat{s})} \mathbf{1}_{(\hat{x}_-, \hat{y}_-) \in \hat{\square}} + \mathbf{1}_{(\hat{x}_+, \hat{y}_+) \in \hat{\square}} p_V(\hat{s}) \left( \mathbf{1}_{(\hat{x}_-, \hat{y}_-) \in \hat{\square}} + \mathbf{1}_{(\hat{x}_+, \hat{y}_+) \in \hat{\square}} \right) e^{-(\tau_V(\hat{s}) + \lambda_V(\hat{s}))(\hat{y}_+ - \hat{y}_-)}.$$

Let us show that both contributions are equal. Indeed, their fourth terms are equal because  $s = \hat{s}$  and  $\hat{y}_+ - \hat{y}_- = y_+ - y_-$ . Their third terms are also equal since, by Table 4.1,

$$p_V(\hat{s}) \left( \mathbf{1}_{(\hat{x}_-, \hat{y}_-) \in \hat{\square}} + \mathbf{1}_{(\hat{x}_+, \hat{y}_+) \in \hat{\square}} \right) = p_V(s) \left( \mathbf{1}_{(x_+, y_+) \in \square} + \mathbf{1}_{(x_-, y_-) \in \square} \right).$$

Similarly, their second terms are equal since

$$\sqrt{q(\hat{s})g_V(\hat{s})} \mathbf{1}_{(\hat{x}_-, \hat{y}_-) \in \hat{\square}} + \mathbf{1}_{(\hat{x}_+, \hat{y}_+) \in \hat{\square}} = \sqrt{q(s)g_V(s)} \mathbf{1}_{(x_+, y_+) \in \square} + \mathbf{1}_{(x_-, y_-) \in \square}.$$

Finally, their first terms are equal since

$$\begin{aligned} & \mathbf{1}_{(\hat{x}_-, \hat{y}_-) \in \hat{\square} \cup \hat{\top} \cup \hat{\vdash} \cup \hat{\dashv}} - \mathbf{1}_{(\hat{x}_+, \hat{y}_+) \in \hat{\square} \cup \hat{\top}} = \mathbf{1}_{(x_+, y_+) \in \square \cup \top \cup \vdash \cup \dashv} - \mathbf{1}_{(x_-, y_-) \in \square \cup \top} \\ & = \mathbf{1}_{(x_-, y_-) \in \square \cup \top \cup \vdash \cup \dashv} - \mathbf{1}_{(x_+, y_+) \in \square \cup \top} + \underbrace{\mathbf{1}_{(x_+, y_+) \in \square \cup \top \cup \vdash \cup \dashv}}_{=1} - \underbrace{\mathbf{1}_{(x_-, y_-) \in \square \cup \top \cup \vdash \cup \dashv}}_{=1}, \end{aligned} \tag{4.36}$$

where we used that the set  $\boxed{\uparrow} \cup \boxed{\downarrow} \cup \boxed{\leftarrow} \cup \boxed{\rightarrow} \cup \boxed{\nwarrow} \cup \boxed{\nearrow}$  collects all the nodes ending a vertical segment and, similarly, the set  $\boxed{\leftarrow} \cup \boxed{\rightarrow} \cup \boxed{\uparrow} \cup \boxed{\downarrow} \cup \boxed{\nwarrow} \cup \boxed{\nearrow}$  collects all the nodes beginning a vertical segment. The same considerations holds for horizontal segments.

Finally, the last terms contributing to  $\alpha_S$  are those that concern crossings in  $\boxed{\times}$ . Let us consider a crossing  $(x, y) \in \boxed{\times}$  whose weights of its adjacent edges are denoted by  $s_S, s_W, s_N, s_E$ . Its contribution to  $\alpha_S(D)$  equals

$$\frac{1 - p_V(s_W + s_S) - p_H(s_W + s_S)}{h(s_W + s_S)}.$$

Similarly, the contribution of  $(\hat{x}, \hat{y}) \in \widehat{\boxed{\times}}$  to  $\alpha_{\widehat{S}}(\widehat{D})$ , is equal to

$$\frac{1 - p_V(\widehat{s}_W + \widehat{s}_S) - p_H(\widehat{s}_W + \widehat{s}_S)}{h(\widehat{s}_W + \widehat{s}_S)}.$$

But,  $\widehat{s}_S = s_N$  and  $\widehat{s}_W = s_E$  and, by Kirchhoff's node law,  $s_E + s_N = s_W + s_S$ . Hence, both contributions coincide again.  $\square$

We can now deduce Theorem 4.1 when  $\lambda_V, \lambda_H, \tau_V$  and  $\tau_H$  are uniformly bounded.

*Proof of Theorem 4.1 (uniformly bounded rates).* Let  $\Phi : \mathcal{D}_{a,b} \rightarrow \mathbb{R}_+$  be a non-negative measurable function. For any skeleton  $S$ , let  $(e_{\rho(1)}, \dots, e_{\rho(d)})$  be a parametrization of  $S$ . Choose any order on the set of edges of  $\widehat{S}$ . Now, for any  $i$ , the edge  $e_{\rho(i)} \in S$  has a reverse edge in  $\widehat{S}$  whose index in  $\widehat{S}$  is denoted by  $\widehat{\rho}(i)$ . The set  $(e_{\widehat{\rho}(1)}, \dots, e_{\widehat{\rho}(d)})$  is a parametrization of  $\widehat{S}$ . In the next formula, for a drawing  $D$ , we write indifferently  $\wedge(D)$  or  $\widehat{D}$ .

$$\begin{aligned} \mathbb{E} \left[ \Phi(\widehat{\mathbf{D}}) \right] &= \sum_{S \in \mathcal{D}/\sim} \int_{\mathbb{R}^{m+n+d}} dx dy dc \Phi(\wedge(x, y, \mathfrak{D}_{S, \rho}(c))) \alpha_S((x, y, \mathfrak{D}_{S, \rho}(c))) \\ &= \sum_{S \in \mathcal{D}/\sim} \int_{\mathbb{R}^{m+n+d}} dx dy dc \Phi(\widehat{\wedge}(\widehat{x}, \widehat{y}, \mathfrak{D}_{\widehat{S}, \widehat{\rho}}(c))) \alpha_S((x, y, \mathfrak{D}_{S, \rho}(c))). \end{aligned}$$

Now, we apply the change of variable from  $(x, y)$  to  $(\widehat{x}, \widehat{y})$ . Recalling that  $\widehat{x}_i = a - x_{m+1-i}$  and  $\widehat{y}_i = b - y_{n+1-i}$ , it follows that the absolute value of the Jacobian is equal to 1, hence

$$\mathbb{E} \left[ \Phi(\widehat{\mathbf{D}}) \right] = \sum_{S \in \mathcal{D}/\sim} \int_{\mathbb{R}^{m+n+d}} dx dy dc \Phi((x, y, \mathfrak{D}_{\widehat{S}, \widehat{\rho}}(c))) \alpha_S((\widehat{x}, \widehat{y}, \mathfrak{D}_{S, \rho}(c))) \quad (4.37)$$

(by Lemma 4.4)

$$= \sum_{S \in \mathcal{D}/\sim} \int_{\mathbb{R}^{m+n+d}} dx dy dc \Phi((x, y, \mathfrak{D}_{\widehat{S}, \widehat{\rho}}(c))) \alpha_{\widehat{S}}((x, y, \mathfrak{D}_{\widehat{S}, \widehat{\rho}}(c))) \quad (4.38)$$

(by Corollary 4.2)

$$= \sum_{S \in \mathcal{D}/\sim} \int_{\mathbb{R}^{m+n+d}} dx dy dc \Phi((x, y, \mathfrak{D}_{\widehat{S}, \rho_{\widehat{S}}}(c))) \alpha_{\widehat{S}}((x, y, \mathfrak{D}_{\widehat{S}, \rho_{\widehat{S}}}(c))) \quad (4.39)$$

(by re-indexation of  $\widehat{S}$  into  $S$ )

$$= \sum_{S \in \mathcal{D}/\sim} \int_{\mathbb{R}^{m+n+d}} dx dy dc \Phi((x, y, \mathfrak{D}_{S, \rho_S}(c))) \alpha_S((x, y, \mathfrak{D}_{S, \rho_S}(c))) \quad (4.40)$$

$$= \mathbb{E}[\Phi(\mathbf{D})]. \quad \square$$

#### 4.5.2 Proof of Theorem 4.2 with uniformly bounded rates

As above, we prove the theorem under the assumption that the rates of the PKS are uniformly bounded. The proof is mostly identical to the one of Theorem 4.1 (the Lebesgue case), so we shall only point out the changes needed to deal with the rules 1<sub>0</sub> and 3(c) of the dynamics. In particular, two new types of nodes need to be considered:

- **Spontaneous split:** a spontaneous split is a point where 2 segments are meeting coming from the north and east i.e. there exist  $\sigma_N, \sigma_E \in D$  such that  $\sigma_N = ((x, y), (x, \cdot), \cdot)$  and  $\sigma_E = ((x, y), (\cdot, y), \cdot)$ . This corresponds to case 1<sub>0</sub> of the dynamics. We denote this set by  $\square$ . We remark that  $\square = \{(x, y) : \exists s \in \mathbb{Z}, ((x, y), s) \in \mathcal{C}_0\}$ .
- **Double coalescence:** a double coalescence is a point  $(x, y)$  where 2 segments are meeting, coming from the west and south i.e. there exist  $\sigma_W, \sigma_S \in D$  such that  $\sigma_W = ((\cdot, y), (x, y), \cdot)$  and  $\sigma_S = ((x, \cdot), (x, y), \cdot)$ . This corresponds to case 3(c) of the dynamics. We denote this set by  $\square$ .

The notion of skeleton is the same as the one defined in the previous Section 4.5.1. But, now, the number of free horizontal coordinates is  $m = |\square| + |\square| + |\square| + |\square|$ , the number of free vertical coordinates is  $n = |\square| + |\square| + |\square| + |\square|$ , and the number of free weight coordinates is

$$d := \ell - (|\square| + |\square| + |\square| + |\square| + |\square| + |\square| + |\square| + |\square| + |\square|).$$

We can now derive the following lemma instead of Lemma 4.1:

**Lemma 4.5.** *For any skeleton  $S$ , the dimension of the set of admissible weights of a drawing  $D$  with a given skeleton  $S$  is equal to:*

$$d(S) = |\square| + |\square| + |\square| + |\square| + |\square| + |\square| - |\square|.$$

*Proof.* The proof is the same as the one of Lemma 4.1, except that now the number of half-edges of  $S$  is

$$2\ell(S) = (|\square| + |\square| + |\square| + |\square|) + 2(|\square| + |\square| + |\square| + |\square|) + 3(|\square| + |\square| + |\square| + |\square|) + 4|\square|,$$

and the spatial dimensions expressions give

$$\begin{cases} |\square| + |\square| + |\square| + |\square| = |\square| + |\square| + |\square| + |\square| & (\text{because both are equal to } m), \\ |\square| + |\square| + |\square| + |\square| = |\square| + |\square| + |\square| + |\square| & (\text{because both are equal to } n). \quad \square \end{cases}$$

As before, the set of all drawings  $D$  with skeleton  $S$  is identified as a subset of  $\mathbb{R}^{m+n+\ell}$ , and a drawing  $D$  is represented by a vector as in equation (4.32). A parametrization  $\rho$  of a skeleton  $S$  is the selection of  $d$  edges  $(e_{\rho(1)}, \dots, e_{\rho(d)})$  that permits to define the weights of all edges.

As before, we define  $\alpha_S$  as the density of the drawing  $D$  on the event that its skeleton is  $S$ , and the *turn function*  $q : \mathbb{Z} \rightarrow \mathbb{R}_+$  by the following formula: for any  $s \in \mathbb{Z}$ ,

$$q(s) := \tau_V(s) \sqrt{\frac{\nu_V(s)}{\nu_H(s)}} = \tau_H(s) \sqrt{\frac{\nu_H(s)}{\nu_V(s)}},$$

with the convention  $0/0 = 0$ . As in Corollary 4.2, for any parametrization  $\rho$ ,

$$\mathbb{E}[\Phi(\mathbf{D})] = \sum_{S \in \mathcal{D}/\sim} \int_{\mathbb{R}^{m+n}} dx dy \sum_{c \in \mathbb{Z}^d} \Phi((x, y, \mathfrak{D}_{S,\rho}(c))) \alpha_S((x, y, \mathfrak{D}_{S,\rho}(c))),$$

where  $\alpha_S$  is given by the following lemma (instead of Lemma 4.2):

**Lemma 4.6.** *For any skeleton  $S$ , and any drawing  $D$  whose skeleton is  $S$  and identified to*

$$(x_1, \dots, x_m, y_1, \dots, y_n, s_1, \dots, s_\ell),$$

see equation (4.32), we have  $\alpha_S(D) =$

$$\begin{aligned} & (\mathbf{1}_{0 < x_1 < x_2 < \dots < x_m < a}) e^{-(\sum_{s \in \mathbb{Z}} \nu_V(s) ds) a} (\mathbf{1}_{0 < y_1 < y_2 < \dots < y_n < b}) e^{-(\sum_{s \in \mathbb{Z}} \nu_H(s) s) b} e^{-p_0 h(0) ab} p_0^{|\square| + |\square|} \\ & \left( \prod_{\sigma = ((x_-, y_-), (x_+, y_+), s) \in D} \left[ \mathbf{1}_{x_- = x_+} \nu_V(s) \left( \mathbf{1}_{(x_-, y_-) \in \square \cup \square \cup \square \cup \square \cup \square} \mathbf{1}_{(x_+, y_+) \in \square \cup \square} \right) \right. \right. \\ & \quad \left. \sqrt{q(s) \nu_V(s)} \mathbf{1}_{(x_-, y_-) \in \square} \mathbf{1}_{(x_+, y_+) \in \square} p_V(s) \left( \mathbf{1}_{(x_-, y_-) \in \square} \mathbf{1}_{(x_+, y_+) \in \square} \right) e^{-(\tau_V(s) + \lambda_V(s))(y_+ - y_-)} \right. \\ & \quad \left. + \mathbf{1}_{y_- = y_+} \nu_H(s) \left( \mathbf{1}_{(x_-, y_-) \in \square \cup \square \cup \square \cup \square \cup \square} \mathbf{1}_{(x_+, y_+) \in \square \cup \square} \right) \right. \\ & \quad \left. \left. \sqrt{q(s) \nu_H(s)} \mathbf{1}_{(x_-, y_-) \in \square} \mathbf{1}_{(x_+, y_+) \in \square} p_H(s) \left( \mathbf{1}_{(x_-, y_-) \in \square} \mathbf{1}_{(x_+, y_+) \in \square} \right) e^{-(\tau_H(s) + \lambda_H(s))(x_+ - x_-)} \right] \right) \\ & \left( \prod_{(x, y) \in \square} \frac{1 - p_V(s_W + s_S) - p_H(s_W + s_S) - p_0 \mathbf{1}_{s_W + s_S = 0}}{h(s_W + s_S)} \mathbf{1}_{((x, \cdot), (x, y), s_S) \in D} \mathbf{1}_{((\cdot, y), (x, y), s_W) \in D} \right). \end{aligned}$$

*Proof.* This argument is the same as the one of Lemma 4.2, with some additional terms. First, the term  $e^{-p_0 h(0) ab}$  is equal to the probability that there is no other spontaneous split in the rectangle  $[0, a] \times [0, b]$ . As before, we look at the contribution of each node and distribute it to its adjacent edges or to the node itself. We detail what happens for the three new kinds of nodes:

- if  $(x, y) \in \square$ , the term  $p_0 \nu_V(-s) \nu_H(s)$  splits into three terms:
  - the term  $\nu_V(-s) = \nu_V(-s) \mathbf{1}_{(\sigma_N) - \in \square}$  is assigned to the northern segment  $\sigma_N$ ,

- the term  $\nu_H(s) = \nu_H(s)^{\mathbf{1}(\sigma_E) - \epsilon_{\square}}$  is assigned to the eastern segment  $\sigma_E$ ,
- the term  $p_0$  is attached to the node itself. All of these contributions are found in the term  $p_0^{\square}$  in  $\alpha_S$ .
- if  $(x, y) \in \square$ , the term  $p_0$  is assigned to the node itself. All these contributions are found in the term  $p_0^{\square}$  in  $\alpha_S(D)$ . Its adjacent segments get no contribution.
- if  $(x, y) \in \boxplus$  such that  $s_S = -s_W$ , the term  $(1 - p_H(0) - p_V(0) - p_0) \frac{\nu_V(-s)\nu_H(s)}{h(0)}$  splits into three terms:
  - the term  $\nu_V(-s) = \nu_V(-s)^{\mathbf{1}(\sigma_N) - \epsilon_{\boxplus}}$  is assigned to the northern segment  $\sigma_N$ ,
  - the term  $\nu_H(s) = \nu_H(s)^{\mathbf{1}(\sigma_E) - \epsilon_{\boxplus}}$  is assigned to the eastern segment  $\sigma_E$ ,
  - the term  $\frac{1 - p_V(0) - p_H(0) - p_0}{h(0)}$  is attached to the node itself. All of these contributions are found in the terms  $\frac{1 - p_V(s_W + s_S) - p_H(s_W + s_S) - p_0 \mathbf{1}_{s_W + s_S = 0}}{h(s_W + s_S)}$  where  $s_S = -s_W$  in  $\alpha_S$ . □

Finally, to end the proof of the invariance by rotation of 180 degrees, we need to prove that a lemma similar to Lemma 4.4 holds in our new case.

**Lemma 4.7.** *For any skeleton  $S$ , for any drawing  $D$  with skeleton  $S$ , we have  $\alpha_{\widehat{S}}(\widehat{D}) = \alpha_S(D)$ .*

*Proof.* This proof is the same as the one of Lemma 4.4, with some new terms to check. First, notice that  $\square = \widehat{\square}$  and  $\boxplus = \widehat{\boxplus}$ .

Hence, the contribution of the following factor in  $\alpha_S(D)$

$$e^{-p_0 h(0) ab} p_0^{|\square| + |\boxplus|} \prod_{(x, y) \in \boxplus} \frac{1 - p_V(s_W + s_S) - p_H(s_W + s_S) - p_0 \mathbf{1}_{s_W + s_S = 0}}{h(s_W + s_S)}$$

coincides with the same in  $\alpha_{\widehat{S}}(\widehat{D})$

$$e^{-p_0 h(0) ab} p_0^{|\widehat{\square}| + |\widehat{\boxplus}|} \prod_{(\widehat{x}, \widehat{y}) \in \widehat{\boxplus}} \frac{1 - p_V(\widehat{s}_W + \widehat{s}_S) - p_H(\widehat{s}_W + \widehat{s}_S) - p_0 \mathbf{1}_{\widehat{s}_W + \widehat{s}_S = 0}}{h(\widehat{s}_W + \widehat{s}_S)}.$$

The first two terms are obviously equal, the last one is equal for the same reason as in the proof of Lemma 4.4 remarking that if  $s_S + s_W = 0$ , then  $\widehat{s}_S + \widehat{s}_W = s_N + s_E = 0$ .

The last point to see is that the following indicator function has changed

$$\mathbf{1}_{(x_-, y_-) \in \square \cup \boxplus \cup \boxminus \cup \boxtimes \cup \boxplus \cup \boxminus \cup \boxtimes}.$$

Nevertheless, that does not change the proof of the equation (4.36) because, now, the set that collects all the ending nodes of a vertical segments is  $\square \cup \boxplus \cup \boxminus \cup \boxtimes \cup \boxplus \cup \boxminus \cup \boxtimes$ , and the one that collects all the beginning nodes of a vertical segments is  $\boxplus \cup \square \cup \boxminus \cup \boxtimes \cup \boxplus \cup \boxminus \cup \boxtimes$ . □

Finally, the end of the proof of Theorem 4.2 with uniformly bounded rates is the same as the one of Theorem 4.1 with uniformly bounded rates, since equations (4.37), (4.38), (4.39) and (4.40) are unchanged.

### 4.5.3 From the uniformly bounded rate case to the unbounded case

The framework of this section is the general case which contains Lebesgue and discrete cases. We assume that the reversibility stated in the main theorems holds true under uniformly bounded rates assumption (proved in Lebesgue and discrete cases in Sections 4.5.1 and 4.5.2, and whose proof is omitted in the general case). The purpose of this section is to show how to extend the previous proofs to the case where the rate functions are unbounded. The idea is to see a PKS as a limit of reversible PKSs with uniformly bounded rates.

Let  $L$  be a PKS with parameters  $(\lambda_0, \lambda_V, \lambda_H, p_0, p_V, p_H, \tau_V, \tau_H, F)$  which satisfies the assumptions of the main theorems, under the initial condition  $(\mathcal{C}_X, \mathcal{C}_Y)$  as defined in equation (4.4).

For any  $n \geq 1$ , we define the set

$$\mathcal{S}(n) := \{s \in \mathbb{R}, \sup(\lambda_V(s), \lambda_H(s), \tau_V(s), \tau_H(s)) > n\}.$$

In words, the set  $\mathcal{S}(n)$  is the set of weights for which the split and turn rates are larger than  $n$ . Remark that

$$\lim_{n \rightarrow \infty} \downarrow \mathcal{S}(n) = \bigcap_{n \geq 1} \mathcal{S}(n) = \emptyset.$$

This is due to the fact that the rate functions  $\lambda_V$ ,  $\lambda_H$ ,  $\tau_V$  and  $\tau_H$  are never equal to  $+\infty$  by the hypotheses given by equations (4.18), (4.19) and (4.20). Now, we define the following notation: for any  $f : \mathbb{R} \rightarrow \mathbb{R}_+$ ,

$$f^{(n)}(s) := f(s) \mathbf{1}_{s \notin \mathcal{S}_n},$$

We denote by  $L^{(n)}$  the PKS with parameters  $(\lambda_0, \lambda_V^{(n)}, \lambda_H^{(n)}, p_0, p_V^{(n)}, p_H^{(n)}, \tau_V^{(n)}, \tau_H^{(n)}, F)$  with initial condition  $(\mathcal{C}_X, \mathcal{C}_Y)$ . The PKS  $L^{(n)}$  has rates uniformly bounded by  $n$  and satisfies the hypotheses of the theorem *for the same measures  $\nu_V$  and  $\nu_H$* .

Hence, by the version of Theorem 4.1 we have proved, which assumes the boundedness of the rate functions, the PKS  $L^{(n)}$  is well defined a.s. and reversible for the line weight measures  $\nu_V$  and  $\nu_H$ .

Now, we consider the process  $L^{(n)}$  in the box  $[0, a] \times [0, b]$ . We want to estimate the number of lines of  $L^{(n)}$  with weight  $s \in \mathcal{S}(n)$  in this box, i.e. the number of lines that, without truncation, would have split or turn rates greater than  $n$ . For that, we will count the mean number of nodes in the box which have at least one edge with a weight in  $\mathcal{S}(n)$ .

As the pair of independent PPPs with intensities  $\text{Leb} \otimes \tilde{\nu}_V$  and  $\text{Leb} \otimes \tilde{\nu}_H$  is a stationary probability measure of  $L^{(n)}$ , for any small element of size  $dx \times dy$ , the probability to see a node such that the weight of its south edge or its west edge is an element of  $\mathcal{S}(n)$  is

$$\left( \nu_V(\mathcal{S}(n)) \nu_H(\mathbb{R}) + \nu_V(\mathcal{S}(n)) \nu_H(\mathbb{R}) - \nu_V(\mathcal{S}(n)) \nu_H(\mathcal{S}(n)) \right) dx dy.$$

By integration on the box  $[0, a] \times [0, b]$ , we find that the mean number of lines with weight in  $\mathcal{S}(n)$  and which are a south or west edge of a node is

$$\left( \nu_V(\mathcal{S}(n)) \nu_H(\mathbb{R}) + \nu_H(\mathcal{S}(n)) \nu_V(\mathbb{R}) \right) ab.$$



Moreover, because  $L^{(n)}$  is reversible, this quantity is also equal to the mean number of lines with weight in  $\mathcal{S}(n)$  and which are a north or east edge of a node. By summing these two means and adding the mean number of lines that are entering in the box, we deduce the following upper bound

$$\begin{aligned} & \mathbb{E} \left[ \text{number of segments of } L^{(n)} \text{ inside } [0, a] \times [0, b] \text{ whose weight } s \in \mathcal{S}(n) \right] \\ & \leq \varepsilon(n) := 2 \left( \nu_V(\mathcal{S}(n)) \nu_H(\mathbb{R}) + \nu_H(\mathcal{S}(n)) \nu_V(\mathbb{R}) \right) ab + \left( a \nu_V(\mathcal{S}(n)) + b \nu_H(\mathcal{S}(n)) \right). \end{aligned}$$

In particular, because it is an integer-valued random variable:

$$\mathbb{P} \left( \text{there exists a segment in } L^{(n)} \text{ of weight } s \in \mathcal{S}(n) \right) \leq \varepsilon(n).$$

As  $\mathcal{S}(n) \rightarrow \emptyset$  when  $n \rightarrow \infty$ , we have that  $\varepsilon(n) \rightarrow 0$ . But, if  $L^{(n)}$  does not contain any line whose weight is in  $\mathcal{S}(n)$ , then  $L^{(n)}$  and  $L$  coincide for the trivial coupling. Then we deduce that

$$\begin{aligned} & \mathbb{P}(\text{the PKS } L \text{ is well defined, i.e. it does not explode, inside the box } [0, a] \times [0, b]) \\ & \geq \mathbb{P} \left( \text{the PKS } L^{(n)} \text{ does not have any segment of weight } s \in \mathcal{S}(n) \right) \geq 1 - \varepsilon(n). \end{aligned}$$

This is true for all  $n$ , so the PKS  $L$  is well defined a.s.. Moreover, the process  $L$  is reversible with line weight measures  $\nu_V$  and  $\nu_H$  since it coincides with probability converging to 1 with the reversible process  $L^{(n)}$ .

## 4.6 Examples

### 4.6.1 Potential function of a PKS

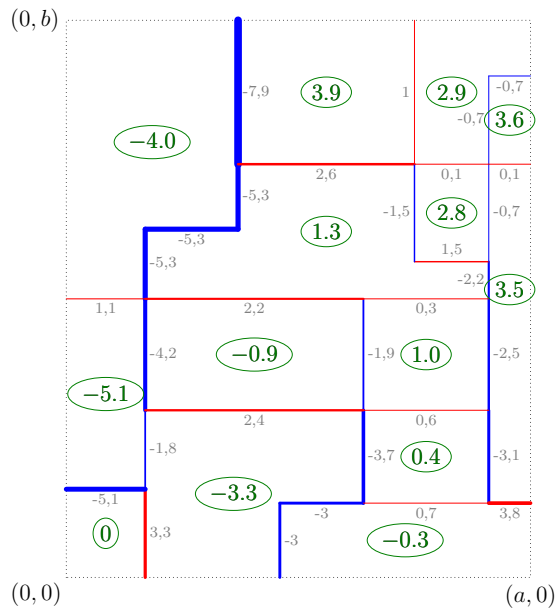
By construction, a PKS induces a random tessellation of the quarter plane into polygonal regions (which are the connected components obtained after removing the lines of the process). We call these connected components the *faces* of the tessellation. The fact that a PKS satisfies Kirchhoff's node law at every intersection is equivalent to the existence of a potential function associated with the faces of the random tessellation. More precisely, we can associate to each face  $F$  a scalar value  $v(F)$  in such way that the following holds true:

- Let  $\sigma$  denote a horizontal segment in the PKS with weight  $s(\sigma)$ . This segment separates two faces of the tessellation. Let  $F$  denote the face *below*  $\sigma$  and let  $F'$  denote the face *above*  $\sigma$ . Then, it holds that

$$v(F') - v(F) = s(\sigma). \tag{4.41}$$

- Let  $\sigma$  denote a vertical segment in the PKS with weight  $s(\sigma)$ . This segment separates two faces of the tessellation. Let us  $F$  denote the face *on the left* of  $\sigma$  and let  $F'$  denote the face *on the right* of  $\sigma$ . Then, it holds that

$$v(F') - v(F) = -s(\sigma). \tag{4.42}$$



**Figure 4.11:** An example of the potential on each connected component of a drawing.

In other words, equation (4.41) states that crossing a horizontal segment by moving upward *increases* the potential by the value of the weight of the segment. On the other hand, equation (4.42) states that crossing a horizontal segment while moving to the right *decreases* the potential by the value of the weight of this segment. See Figure 4.11 for an illustration.

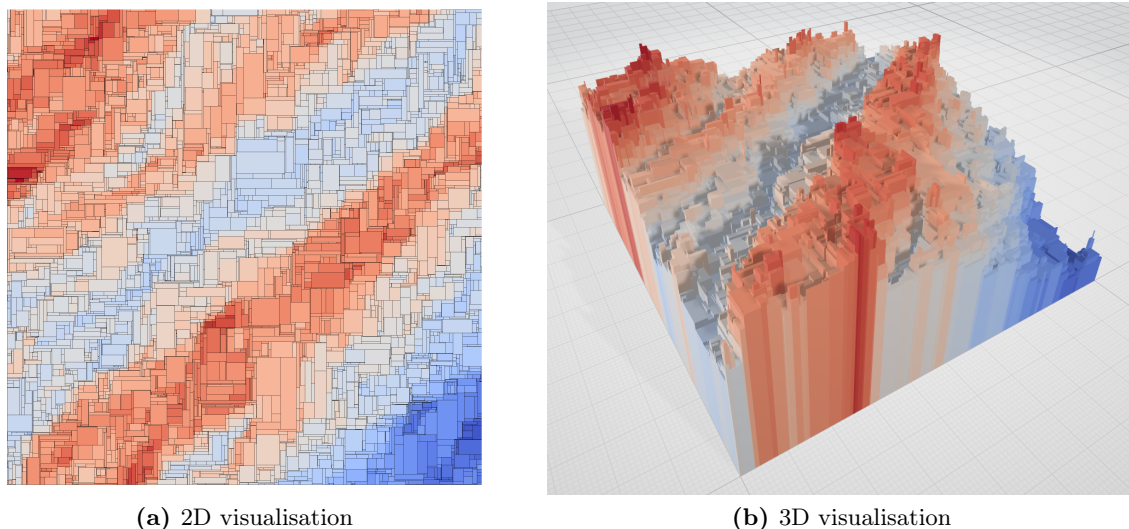
The consistency of equations (4.41) and (4.42) for any segment is straightforward thanks to Kirchhoff's node law: looking at Figure 4.2, we simply check that the sum of the potential differences when going (say clockwise) around a node is  $s_W - s_N - s_E + s_S = 0$ . Furthermore, it is clear that the potential function  $v$  is unique up to an additive constant. By convention, we choose it to be 0 for the bottom left face containing the origin. Figure 4.12 shows 2D and 3D representation of the potential function for a PKS process obtained by simulation with Gaussian line weights.

## 4.6.2 List of examples

As explained above, a PKS can be seen either as a weighted line process or as a potential function on faces of a random tessellation. These dual points of view make it possible to recover several well-known models appearing in the statistical physics literature, in particular classical models related to Last Passage Percolation (LPP) as explained in Section 4.6.3.

In the rest of the section, we compute the parameter of the reversible PKS with line weight measures  $\nu_V$  and  $\nu_H$  for many usual continuous and discrete distribution. A list of examples of reversible PKS is presented in Table 4.3. Subsequently, we discuss further some of the models in this list that enjoy special properties and that are connected to well-known models.

The usual distributions and their parametrization considered here are summarized in Table 4.2.



**Figure 4.12:** Simulation of a reversible PKS on  $[0, 80] \times [0, 80]$  with parameters  $p_V(s) = 0.5$ ,  $p_H(s) = 0.5$ ,  $\tau_V(s) = \tau_H(s) = 0$  and with line weight measures  $\nu_V = \nu_H = \mathcal{N}(0, 1)$  (Model 16 in Table 4.3). Colors represent potential values: blue for negative ones and red for positive ones.

Name	Parameters	Notation	Support	Density
Dirac	$a \in \mathbb{R}$	$\delta_a$	$\{a\}$	$\mathbf{1}_{x=a}$
Discrete Uniform	$A \subset \mathbb{Z}$	$\text{Unif}(A)$	$A$	$\frac{1}{ A } \mathbf{1}_{x \in A}$
Bernoulli	$p \in [0, 1]$	$\text{Ber}(p)$	$\{0, 1\}$	$(1-p)\mathbf{1}_{x=0} + p\mathbf{1}_{x=1}$
Binomial	$(n, p) \in \mathbb{N} \times [0, 1]$	$\text{Bin}(n, p)$	$\llbracket 0, n \rrbracket$	$\binom{n}{x} p^x (1-p)^{n-x}$
Geometric	$p \in [0, 1]$	$\text{Geom}(p)$	$\mathbb{N}$	$p(1-p)^x$
Poisson	$\lambda \in \mathbb{R}_+$	$\text{Poi}(\lambda)$	$\mathbb{N}$	$e^{-\lambda} \frac{\lambda^x}{x!}$
Uniform	$[a, b] \subset \mathbb{R}$	$\text{Unif}([a, b])$	$[a, b]$	$\frac{1}{b-a} \mathbf{1}_{a \leq x \leq b}$
Exponential	$\lambda \in \mathbb{R}_+^*$	$\text{Exp}(\lambda)$	$\mathbb{R}_+$	$\lambda e^{-\lambda x} \mathbf{1}_{x \geq 0}$
Gamma	$(k, \theta) \in \mathbb{R} \times \mathbb{R}_+^*$	$\text{Gamma}(k, \theta)$	$\mathbb{R}_+$	$\frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-x/\theta} \mathbf{1}_{x \geq 0}$
Normal (or Gaussian)	$(\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+$	$\mathcal{N}(\alpha, \beta)$	$\mathbb{R}$	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$
Beta	$(\alpha, \beta) \in (\mathbb{R}_+^*)^2$	$\text{Beta}(\alpha, \beta)$	$[0, 1]$	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}$
The opposite of the measure	$\mu$ , a real measure	$-\mu$	$-\text{Support}(\mu)$	$d\mu(-x)$

**Table 4.2:** Parametrization of the usual distributions. Here,  $\mathbb{N}$  denotes the set  $\{0, 1, \dots\}$ .

Model	$\nu_V/\nu_H(\mathbb{R})$	$\nu_H/\nu_H(\mathbb{R})$	$\lambda_V(s)/(\nu_H(\mathbb{R})p_V(s))$	$\lambda_H(s)/(\nu_V(\mathbb{R})p_H(s))$	$F(s, \cdot)$
1	$\delta_0$	$\delta_0$	1	1	$\delta_0$
2	$\delta_a$ with $a \neq 0$	$\delta_b$ with $b \neq 0$	0	0	$\delta_b$
3	$\delta_a$ with $a \neq 0$	$\delta_0$	1	0	$\delta_0$
4	Ber( $q_V$ )	Ber( $q_H$ )	$(1 - q_H)\mathbf{1}_{s=0} + (1 + q_H(q_V^{-1} - 2))\mathbf{1}_{s=1}$	$(1 - q_V)\mathbf{1}_{s=0} + (1 + q_V(q_H^{-1} - 2))\mathbf{1}_{s=1}$	$\text{Ber}\left(\frac{q_H(1 - q_V)}{q_V + q_H - 2q_Vq_H}\right)\mathbf{1}_{s=1} + \delta_0\mathbf{1}_{s=0} + \delta_1\mathbf{1}_{s=2}$
5	-Ber( $q_V$ )	Ber( $q_H$ )	$(1 - q_H)\mathbf{1}_{s=-1} + (1 + q_H\frac{2q_V - 1}{1 - q_V})\mathbf{1}_{s=0}$	$(1 - q_V)\mathbf{1}_{s=1} + (1 + q_V\frac{2q_H - 1}{1 - q_H})\mathbf{1}_{s=0}$	$\text{Ber}\left(\frac{q_Vq_H}{1 - q_V - q_H + 2q_Vq_H}\right)\mathbf{1}_{s=0} + \delta_0\mathbf{1}_{s=-1} + \delta_1\mathbf{1}_{s=1}$
6	Unif( $[0, a]$ )	Unif( $[0, b]$ )	$\min\left(\frac{s}{b}, 1\right)$	$\min\left(\frac{s}{a}, 1\right)$	Unif( $[\max(0, s - a), \min(b, s)]$ )
7	-Unif( $[0, a]$ )	Unif( $[0, b]$ )	$\min\left(\frac{a + s}{b}, 1\right)$	$\min\left(\frac{b - s}{a}, 1\right)$	Unif( $[\max(0, s), \min(b, s + a)]$ )
8	Geom( $q$ )	Geom( $q$ )	$(s + 1)q$	$(s + 1)q$	Unif( $\{0, \dots, s\}$ )
9	Exp( $\gamma$ )	Exp( $\gamma$ )	$\gamma s$	$\gamma s$	Unif( $[0, s]$ )
10	Exp( $\gamma$ )	Geom( $q$ )	$q \sum_{t=0}^{[s]} ((1 - q)e^\gamma)^t$	0	$\propto \sum_{t=0}^{[s]} ((1 - q)e^\gamma)^t \delta_t$
11	Gamma( $k_V, \theta$ )	Gamma( $k_H, \theta$ )	$\frac{\Gamma(k_V)\left(\frac{s}{\theta}\right)^{k_V}}{\Gamma(k_V + k_H)\left(\frac{s}{\theta}\right)^{k_H}}$	$\frac{\Gamma(k_H)\left(\frac{s}{\theta}\right)^{k_V}}{\Gamma(k_V + k_H)\left(\frac{s}{\theta}\right)^{k_V}}$	$s \text{Beta}(k_V, k_H)$
12	-Ber( $q_V$ )	Geom( $q_H$ )	$\frac{q_H(1 - q_Vq_H)}{(1 - q_V)}\mathbf{1}_{s=0} + q_H\mathbf{1}_{s=-1}$	$(1 - q_Vq_H)$	$s + \text{Ber}\left(\frac{q_V(1 - q_H)}{1 - q_Vq_H}\right)\mathbf{1}_{s \geq 0} + \delta_1\mathbf{1}_{s=-1}$
13	-Geom( $q_V$ )	Geom( $q_H$ )	$\frac{q_V + q_H - q_Vq_H}{\gamma_H}$	$\frac{q_V}{q_V + q_H - q_Vq_H}$	$\max(0, s) + \text{Geom}(q_V + q_H - q_Vq_H)$
14	-Exp( $\gamma_V$ )	Exp( $\gamma_H$ )	$\frac{\gamma_V + \gamma_H}{\gamma}$	$\frac{\gamma_V}{\gamma_V + \gamma_H}$	$\max(0, s) + \text{Exp}(\gamma_V + \gamma_H)$
15	-Exp( $\gamma$ )	Geom( $q$ )	$\frac{1}{1 - (1 - q)e^{-\gamma}}$	0	$\max(0, [s]) + \text{Geom}(1 - (1 - q)e^{-\gamma})$
16	$\mathcal{N}(0, 1)$	$\mathcal{N}(0, 1)$	$\frac{1}{\sqrt{2}} \exp\left(\frac{s^2}{4}\right)$	$\frac{1}{\sqrt{2}} \exp\left(\frac{s^2}{4}\right)$	$\mathcal{N}\left(\frac{s}{2}, \frac{1}{2}\right)$
17	Poi( $\gamma_V$ )	Poi( $\gamma_H$ )	$e^{-\gamma_H} \left(1 + \frac{\gamma_H}{\gamma_V}\right)^s$	$e^{-\gamma_V} \left(1 + \frac{\gamma_V}{\gamma_H}\right)^s$	$\text{Bin}\left(s, \frac{\gamma_H}{\gamma_V + \gamma_H}\right)$

Table 4.3: Some remarkable PKS.

### 4.6.3 Models with monotone potential (LPP)

For a PKS which satisfies the hypotheses of the main theorems, the weights of the vertical lines (resp. horizontal lines) take their values in the support of  $\nu_V$  (resp.  $\nu_H$ ) a.s.. It follows that the potential function  $v$  is monotone in its both coordinates when the measures  $\nu_V$  and  $\nu_H$  have their support included in  $\mathbb{R}_+$  and  $\mathbb{R}_-$  respectively, or the opposite. For instance, if  $\text{Support}(\nu_V) \subset \mathbb{R}_-$  and  $\text{Support}(\nu_H) \subset \mathbb{R}_+$ , then the potential  $v$  is non-decreasing in both its coordinates. This is the case for Model 13 in Table 4.3 simulated in Figure 4.7b. Another example is that of Model 14 whose associated potential is represented in the simulation on the front page of this chapter.

As it turns out, such PKS with monotone potentials can often be mapped to LPP models.

**Standard Hammersley's model.** The standard Hammersley's broken line process described in the introduction of the chapter and illustrated in Figure 4.1 (and studied for instance in [Ham72, AD95, Gro02, CG05, CG06]) is clearly a PKS whose dynamics is the one of Model 2 with  $a = -1$ ,  $b = 1$  and  $p_0 = 1$ .

**Hammersley interacting fluid system.** In [CP12, CPS12], a generalisation of the standard Hammersley's model is introduced. As for the standard Hammersley's model, it starts with a unit intensity PPP on  $[0, \infty)^2$  but here, to each atom of the PPP is also associated a random positive number, chosen in i.i.d fashion, with common probability distribution  $\mathcal{F}$  on the positive real numbers. At each atom of the PPP, a particle of positive weight equal to this number goes to the right and a particle of opposite weight goes up. When two particles of opposite weights collide, they both disappear; otherwise the particle of maximal absolute weight continues with a weight equal to the sum of the weights before the encounter and the other one disappears. When  $\mathcal{F}$  is the Dirac measure at 1, we recover the classical Hammersley process.

The Hammersley fluid model is a PKS with parameters  $\lambda_0 = 1$ ,  $\lambda_V = \lambda_H = 0$ ,  $\tau_V = \tau_H = 0$ ,  $p_0 = 1$ ,  $p_V(s) = \mathbf{1}_{s < 0}$ ,  $p_H(s) = \mathbf{1}_{s > 0}$  and  $F(0; \cdot) = \mathcal{F}(\cdot)$ . Note that we do not need to specify  $F(s, \cdot)$  for  $s \neq 0$  because  $p_V(s) + p_H(s) = 1$  and  $\lambda_V(s) = \lambda_H(s) = 0$ . However, in view of Theorems 4.1 and 4.2, the only model in this class that is reversible corresponds to the usual Hammersley's process when the law  $\mathcal{F}$  is a Dirac law.

We can also recover, within the framework of PKS, other LPP systems defined on the discrete grid  $\mathbb{N}^2$  by embedding these models on  $\mathbb{R}_+^2$ . In order to do so, we impose the functions  $p_V$ ,  $p_H$  as well as the constant  $p_0$  to be all identically 0. This assumption ensures that splits, annihilations and creation events can never happen during the dynamics. Thus, each line on the initial PPPs on the  $x$ - and  $y$ -axis survives forever and the trace of all those lines define a 2-dimensional discrete grid (embedded in  $\mathbb{R}_+^2$  with exponential spacing).

**Exponential Last Passage Percolation.** Model 14 where  $-\nu_V$  and  $\nu_H$  are proportional to exponential distributions corresponds to the Exponential LPP studied by [Ros81].

**Geometric Last Passage Percolation.** There are two cases of Geometric LPP depending on whether the geometric distributions start from 0 or 1. Both cases have been studied [CPE96,

Mar06, Sep09] and both can be seen as special PKS models. In the Table 4.3, we have only detailed the case of the geometric starting from 0, see Model 13.

**Discrete Hammersley’s processes.** In [BEGG16], two models are defined. Their second model corresponds to Model 12, mixing a discrete Bernoulli distribution with a Geometric distribution. In that case, with the notation of [BEGG16], the probability that a site contains a “cross” is equal to  $p = \frac{q_V(1 - q_H)}{1 - q_V q_H}$ . Let us note that their first model however *cannot* be mapped to a PKS. Indeed, the model is still conservative (i.e. it obeys Kirchhoff law), but the transition kernel at a crossing of lines depends not only on the total incoming weights but also on the horizontal/vertical division of the global weight. Thus, in order to encompass this first model, one would need to significantly generalize the definition of a PKS process. This is doable but it lies outside the scope of this chapter.

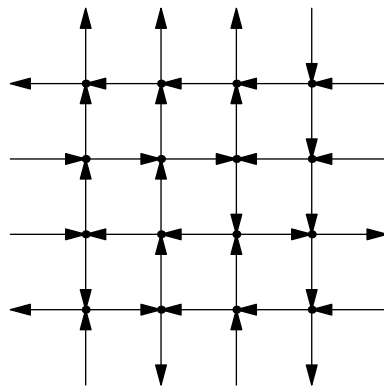
**Generalised Last Passage Percolation.** Let  $\mu_0$  be a probability measure on  $\mathbb{N}^*$  (resp. on  $\mathbb{R}_+$  with density  $f_0$ ). Taking, for any  $A \in \mathcal{B}(\mathbb{R}_+)$ ,  $\nu_V(-A) = \nu_H(A) = \sum_{i \in A} \sqrt{\mu_0(i)}/Z$  where  $Z = \sum_{i \in \mathbb{N}^*} \sqrt{\mu_0(i)}$  (resp.  $\nu_V(-A) = \nu_H(A) = \int_A \sqrt{f_0(s)} ds/Z$  where  $Z = \int_{\mathbb{R}^*} \sqrt{f_0(s)} ds$ ), we recover the generalised LPP defined in [Cas19].

#### 4.6.4 Bullet models

We call *bullet models* the family of models where the weight of the lines plays no role. These models can be obtained by taking  $\nu_V$  and  $\nu_H$  proportional to  $\delta_0$  (see Model 1 of the Table 4.3), and the turn rate functions  $\tau_V = \tau_H = 0$ . Some of them have been lingering in the scientific community for a few years and are notoriously difficult to study out of equilibrium, see for instance [BM20, HST21] and references therein. In [BCG<sup>+</sup>15], some bullet models are proved to be stationary. Some of them correspond to PKS processes with specific parameters. For instance, setting  $\nu_V$  and  $\nu_H$  to be proportional to  $\delta_0$  and assuming that  $p_0 = 0$ , the dynamics of the bullets can be formulated as follows. When an horizontal and a vertical bullets meet:

- with probability  $p_V(0)$ , the horizontal bullet is destroyed and the vertically one continues its course.
- with probability  $p_H(0)$ , the vertical bullet is destroyed and the horizontal one continues its course.
- with probability  $1 - p_V(0) - p_H(0)$ , both bullets continue their course (passing through each other).

Of course, taking the Dirac at 0 for the weight measure may appear as cheating a somewhat since the potential associated with the PKS is then constant to 0. However, it is also possible to define bullet models by choosing  $\nu_V$  and  $\nu_H$  to be Geometric or Exponential measures instead of a Dirac at 0. Indeed, for all these measures, it follows from Models 13, 14 and 15, that the splitting rate of lines remains constant (i.e. does not depend on the weight of the line). Thus, interestingly enough, this shows that some bullet model can be interpreted as the trace of more complex, non-trivial, potential models.



**Figure 4.13:** A configuration of the six-vertex model in the grid  $4 \times 4$ . Its probability associated is  $\frac{1}{Z} w_1^2 w_2^3 w_3^1 w_4^5 w_5^3 w_6^2$ .

### 4.6.5 Six-vertex model

The six-vertex model is a standard model in statistical physics introduced first by Pauling in 1935 [Pau35] to study the ice in two dimensions. From a mathematical point of view, it is a family of probability measures on the set of orientations of the grid  $N \times N$ , such that there are only two incoming edges around each node. Hence, there are only six possible local configurations allowed. To each type  $i$  of a local configuration, we associate a weight (an “energy”)  $w_i$ , see Table 4.4. From these weights, we can define a probability measure on the set of orientations of the grid  $N^2$  via the following formula: for any orientation  $O$ ,

$$\mathbb{P}(O) = \frac{1}{Z} \prod_{(x,y) \in [0,N]^2} w_{\text{type}_O(x,y)}$$

where  $Z = \sum_O \prod_{(x,y) \in [0,N]^2} w_{\text{type}_O(x,y)}$  and where  $\text{type}_O(x,y) \in \{1, \dots, 6\}$  denote the type of the local configuration seen around the point  $(x,y)$  in the orientation  $O$ , see Figure 4.13.

Usually, the model is studied with the assumption that *there does not exist an external electromagnetic field* that implies that  $w_1 = w_2 = a$ ,  $w_3 = w_4 = b$  and  $w_5 = w_6 = c$ . Such models of six and also eight-vertex models have been deeply studied, and we refer the interested reader to [FW70, Sut70, Bax72, Bax82, KDN90, BCG16, Cas18, DCGH<sup>+</sup>18, Mel21] and references therein.

Some six-vertex models *with an external electromagnetic field* turn out to be special cases of PKS from Models 4 and 5 where  $p_V = p_H = \tau_V = \tau_H = 0$  (to get only crossings). Namely, we can construct a six-vertex model from PKSs of type 4 and 5, in the following way:

- In the case of Model 4, to any horizontal segment with weight 0 (resp. 1) of the PKS, we associate an oriented segment to the west (resp. to the east) in the six-vertex configuration; and similarly, to any vertical segment with weight 0 (resp. 1) of the PKS, we associate an oriented segment to the south (resp. to the north) in the six-vertex model.
- In the case of Model 5, to any horizontal segment with weight 0 (resp.  $-1$ ) of the PKS, we associate an oriented segment to the east (resp. to the west) in the six-vertex configuration;

Type	1	2	3	4	5	6
Six-vertex model						
“Energy”	$w_1$	$w_2$	$w_3$	$w_4$	$w_5$	$w_6$
PKS Model 4						
“Energy” Model 4	$q_V + q_H - 2q_V q_H$	$q_V + q_H - 2q_V q_H$	$q_H(1 - q_V)$	$q_V(1 - q_H)$	$q_V(1 - q_H)$	$q_H(1 - q_V)$
PKS Model 5						
“Energy” Model 5	$1 - q_H - q_V + 2q_V q_H$	$1 - q_H - q_V + 2q_V q_H$	$q_V q_H$	$(1 - q_V) \cdot (1 - q_H)$	$(1 - q_V) \cdot (1 - q_H)$	$q_V q_H$

**Table 4.4:** On the first line, the six local configurations allowed. On the second line, their local “energy”. On the third line, their correspondence with a configuration of Model 4: a plain line stands for a line with weight 1 and a dotted line for a line with weight 0. On the fourth line, their correspondence with a configuration of Model 5: a plain line stands for a line with weight 1 or  $-1$  and a dotted line for a line with weight 0.

and similarly, to any vertical segment with weight 0 (resp. 1) of the PKS, we associate an oriented segment to the south (resp. to the north) in the six-vertex model.

See Figure 4.4 to get an illustration of these correspondences.

#### 4.6.6 Gaussian and Poisson models

Models 16 and 17 correspond to models with Gaussian and Poisson marginals respectively. Both models have a particularly nice explicit dynamics. These models are new to the best of our knowledge and look interesting to study further. In model 17, one can indeed realize that the potential has the same distribution as the difference of two independent Poisson variables, which gives rise to a law of large number with Gaussian fluctuations.

### 4.7 Statistical properties of the tessellation

In this section, we look at the basic geometric properties of the system of lines generated by a reversible PKS satisfying the assumptions of Theorem 4.1 or 4.2. We focus here our attention on the case where  $p_0 = 0$ . Indeed, if  $p_0 \neq 0$ , the number of faces could be sub-quadratic according to the length  $a$  of a square  $[0, a] \times [0, a]$  as it is the case for the Hammersley broken line process (presented in the Introduction) where the number of faces is linear in  $a$ .



**Number of faces and nodes.** Let  $D \in \mathcal{D}_{a,b}$  be a drawing, we can associate to this drawing a *tessellation* as the set of segments of  $D$  without notifying their weight. We call a *face* of a tessellation  $T$  a connected component of  $([0, a] \times [0, b]) \setminus T$ .

**Proposition 4.2.** *Consider a reversible PKS such that its initial condition  $(\mathcal{C}_X, \mathcal{C}_Y)$  is distributed according to two independent PPPs respectively on  $(\mathbb{R}_+ \times \{0\}) \times \mathbb{R}$  with intensity  $dx d\nu_V(s)$  and on  $(\{0\} \times \mathbb{R}_+) \times \mathbb{R}$  with intensity  $dy d\nu_H(s)$  where  $\nu_V$  and  $\nu_H$  are two non-zero finite measures on  $\mathbb{R}$  satisfying conditions (4.17), (4.18), (4.19), (4.20) and (4.21). Then, the law of the tessellation associated to this PKS is translation-invariant and satisfies:*

- (i) *The mean number of faces that do not touch the northern or eastern sides of  $[0, a] \times [0, b]$ , is  $ab\nu_V(\mathbb{R})\nu_H(\mathbb{R})$ .*
- (ii) *The mean number of nodes of each type is summarized in the table below:*

Type	Mean number
$\square \perp$	$a\nu_V(\mathbb{R})$
$\square \top$	
$\square \vdash$	$ab \int_{\mathbb{R}} p_V(s) (\nu_V * \nu_H)(ds)$
$\square \dashv$	
$\square \ulcorner$	$ab \int_{\mathbb{R}} \tau_V(s) d\nu_V(s) = ab \int_{\mathbb{R}} \tau_H(s) d\nu_H(s)$
$\square \llcorner$	$ab \int_{\mathbb{R}} (1 - p_V(s) - p_H(s)) (\nu_V * \nu_H)(ds)$

The mean number of horizontal nodes can be found by swapping  $H$  and  $V$ , and  $a$  and  $b$ .

- (iii) *When  $a$  or  $b$  goes to infinity, the number of each type of node is almost surely asymptotically equal to their rescaled mean.*

*Proof.* We will first determine the mean number of each node, and we will deduce the mean number of faces, so that we will first prove (ii) and then (i) and (iii).

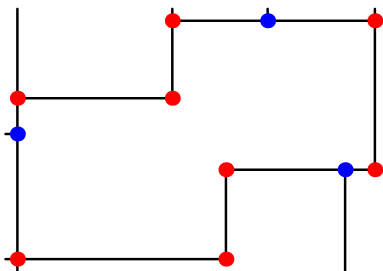
- (ii) • Nodes of type  $\square \perp$  or  $\square \top$ : by definition, the nodes  $\square \perp$  are distributed according to a PPP of intensity  $\nu_V(\mathbb{R})$  on the  $x$ -axis. Thus, the mean number of such nodes on the segment  $(0, a) \times \{0\}$  is  $a\nu_V(\mathbb{R})$ . By Theorem 4.1, the same holds for the mean number of nodes of type  $\square \top$  on the segment  $(0, a) \times \{b\}$ .
- Nodes of type  $\square \vdash$  or  $\square \dashv$ : we do the proof for  $\square \vdash$ . For any  $(x, y)$ , the probability to see in the box  $[x, x+dx] \times [y, y+dy]$ , a vertical line, a horizontal line and a vertical coalescence is

$$\begin{aligned} dx dy \int_{\mathbb{R}} \int_{\mathbb{R}} p_V(s+t) d\nu_V(s) d\nu_H(t) &= dx dy \int_{\mathbb{R}} \int_{\mathbb{R}} p_V(u) d\nu_V(s) d\nu_H(u-s) \\ &= dx dy \int_{\mathbb{R}} p_V(u) (\nu_V * \nu_H)(du) \end{aligned}$$

Hence, the mean number of nodes on type  $\square \vdash$  in  $[0, a] \times [0, b]$  is

$$\int_{[0,a]} dx \int_{[0,b]} dy \int_{\mathbb{R}} p_V(s) (\nu_V * \nu_H)(ds) = ab \int_{\mathbb{R}} p_V(s) (\nu_V * \nu_H)(ds).$$

By Theorem 4.1, the same holds for the mean number of nodes of type  $\square \dashv$ .



**Figure 4.14:** A connected component with 11 nodes (in blue and red) and 8 corners (in red).

- Nodes of type  $\square$ : similarly, the probability to see in the box  $[x, x + dx] \times [y, y + dy]$  a vertical line of size  $[s, s + ds]$  and a vertical turn (that happens at rate  $\tau_V(s)$ ) is

$$dx dy \tau_V(s) d\nu_V(s) = dx dy \tau_H(s) d\nu_H(s).$$

We conclude by integration on  $\mathbb{R}$  for  $s$ ,  $[0, a]$  for  $x$  and  $[0, b]$  for  $y$ .

- Nodes of type  $\boxplus$ : similar to the case of  $\boxminus$  where we multiply by  $1 - p_V(s+t) - p_H(s+t)$  instead of  $p_V(s+t)$  since we are in the case 3(d) of the dynamics of Section 4.2.
- (i) Now, we have two ways to prove the mean number of connected components that do not touch the northern or eastern sides of the rectangle, which is the same as the one which do not touch the southern or western sides of the rectangle by Theorem 4.1. Just remark that any connected component has only one north-east corner and one south-west corner. Hence, the mean number of connected components is both equal to the mean number of nodes of types  $\boxminus \cup \boxplus \cup \boxtimes$ , and to the mean number of nodes of types  $\boxplus \cup \boxminus \cup \boxtimes$ .
- (iii) We treat the case where  $a$  is fixed and  $b \rightarrow \infty$ . By monotonicity, we can assume that  $b$  is integer-valued. Let  $\mathbf{D}$  be a random drawing on  $[0, a] \times \mathbb{R}_+$ . For any integer  $n \geq 0$ , we denote by  $M_n$ , the restriction of  $\mathbf{D}$  to the segment  $[0, a] \times \{n\}$ . The process  $(M_n)$  is a Markov chain and, according to Corollary 4.1 (i), it admits an invariant measure. Moreover, this chain is irreducible since the empty set (no line going up) can be reached with positive probability from any starting configuration. Thus, according to classical results on Markov chains having an accessible atom (see [MT93, Section 15.1] and [AG11, Theorem 1]), this chain is Harris recurrent and the law of large numbers applies.  $\square$

**Mean number of nodes and corners around a face.** For any positive  $a, b$ , let  $\mathbf{D}$  be a random drawing of  $\mathcal{D}_{a,b}$ . Denote by  $\mathcal{F}(\mathbf{D})$  the set of faces of the tessellation of  $\mathbf{D}$ . For a given face  $F$  we denote by  $s_F$  and  $c_F$  the number of nodes and of corners (that are the nodes on the boundary of  $F$  whose angle is  $\pi/2$  or  $3\pi/2$ ) around the face  $F$ . This is illustrated on Figure 4.14. Here, we are interested in  $\mathbf{s}_{a,b}$  (resp.  $\mathbf{c}_{a,b}$ ) the mean number of nodes (resp. corners) of  $\mathcal{F}(\mathbf{D})$ , namely

$$\mathbf{s}_{a,b} = \frac{\sum_{F \in \mathcal{F}(\mathbf{D})} s_F}{|\mathcal{F}(\mathbf{D})|} \quad \left( \text{resp. } \mathbf{c}_{a,b} = \frac{\sum_{F \in \mathcal{F}(\mathbf{D})} c_F}{|\mathcal{F}(\mathbf{D})|} \right). \quad (4.43)$$

**Corollary 4.3.** *The following almost sure limits hold:*

$$(i) \lim_{b \rightarrow \infty} \lim_{a \rightarrow \infty} \mathbf{s}_{a,b} = \lim_{a \rightarrow \infty} \lim_{b \rightarrow \infty} \mathbf{s}_{a,b} = 4 + 2 \frac{\int_{\mathbb{R}} (p_V(u) + p_H(u)) (\nu_V * \nu_H)(du) + 2 \int_{\mathbb{R}} \tau_V(u) d\nu_V(u)}{\nu_V(\mathbb{R}) \nu_H(\mathbb{R})}.$$

$$(ii) \lim_{b \rightarrow \infty} \lim_{a \rightarrow \infty} \mathbf{c}_{a,b} = \lim_{a \rightarrow \infty} \lim_{b \rightarrow \infty} \mathbf{c}_{a,b} = 4 + 4 \frac{\int_{\mathbb{R}} \tau_V(u) d\nu_V(u)}{\nu_V(\mathbb{R}) \nu_H(\mathbb{R})}.$$

*Remark.* When  $p_V + p_H = 1$  and  $\tau_V = \tau_H = 0$ , all the nodes are of degree 3 and all faces are rectangles. Hence, our result recover the well-known fact that the mean number of nodes per faces is equal to 6 and the number of corners is obviously 4.

*Proof.* First, we use point (iii) of Proposition 4.2 to go back and forth between a.s. convergence and convergence of mean. Notice that since  $a$  and  $b$  go to infinity, we do not care about counting the faces that touch the boundary of a finite rectangle  $[0, a] \times [0, b]$  or the nodes on it, because their proportion, compared to the total number of node in the box, goes to zero as the box gets larger.

(i) By definition:

$$\mathbf{s}_{a,b} = \frac{\sum_{F \in \mathcal{F}(\mathbf{D})} s_F}{|\mathcal{F}(\mathbf{D})|} = \frac{\sum_{F \in \mathcal{F}(\mathbf{D})} s_F/a}{|\mathcal{F}(\mathbf{D})|/a}.$$

And so by (iii) of Proposition 4.2,

$$\lim_{a \rightarrow \infty} \mathbf{s}_{a,b} = \frac{\lim_{a \rightarrow \infty} \mathbb{E} \left[ \sum_{F \in \mathcal{F}(\mathbf{D})} s_F \right] / a}{\lim_{a \rightarrow \infty} \mathbb{E} [|\mathcal{F}(\mathbf{D})|] / a}.$$

But, by (ii) of Proposition 4.2,  $\mathbb{E}[|\mathcal{F}(\mathbf{D})|] = ab \nu_V(\mathbb{R}) \nu_H(\mathbb{R})$  and

$$\begin{aligned} \mathbb{E} \left[ \sum_{F \in \mathcal{F}(\mathbf{D})} s_F \right] &= \mathbb{E}[\text{number of nodes with multiplicity}] \\ &= \mathbb{E}[2(|\square| + |\boxminus| + |\boxplus| + |\square|) + 3(|\boxplus| + |\boxminus| + |\boxplus| + |\boxminus|) + 4|\boxplus|] \\ &= 2a \nu_V(\mathbb{R}) + 2b \nu_H(\mathbb{R}) + 6ab \int_{\mathbb{R}} (p_V(u) + p_H(u)) (\nu_V * \nu_H)(du) \\ &\quad + 4ab \int_{\mathbb{R}} (1 - p_V(u) - p_H(u)) (\nu_V * \nu_H)(du) + 4ab \int_{\mathbb{R}} \tau_V d\nu_V(u) \\ &= 4ab \int_{\mathbb{R}} (\nu_V * \nu_H)(du) + 2ab \int_{\mathbb{R}} (p_V(u) + p_H(u)) (\nu_V * \nu_H)(du) \\ &\quad + 2a \nu_V(\mathbb{R}) + 2b \nu_H(\mathbb{R}) + 4ab \int_{\mathbb{R}} \tau_V(u) d\nu_V(u). \end{aligned}$$

Consequently,

$$\begin{aligned} \lim_{a \rightarrow \infty} \mathbf{s}_{a,b} &= \lim_{a \rightarrow \infty} \left[ 2 \left( \frac{\int_{\mathbb{R}} (p_V(u) + p_H(u)) (\nu_V * \nu_H)(du) + 2 \int_{\mathbb{R}} \tau_V(u) d\nu_V(u)}{\nu_V(\mathbb{R}) \nu_H(\mathbb{R})} \right) \right. \\ &\quad \left. + 4 \underbrace{\frac{\int_{\mathbb{R}} (\nu_V * \nu_H)(du)}{\nu_V(\mathbb{R}) \nu_H(\mathbb{R})}}_{=1} + 2(b \nu_H(\mathbb{R}))^{-1} + 2(a \nu_V(\mathbb{R}))^{-1} \right]. \end{aligned}$$

Now, taking the limit when  $b \rightarrow \infty$  leads to the wanted results for  $\lim_{b \rightarrow \infty} \lim_{a \rightarrow \infty} \mathbf{s}_{a,b}$ . The same holds for  $\lim_{a \rightarrow \infty} \lim_{b \rightarrow \infty} \mathbf{s}_{a,b}$ .

(ii) Similarly for  $\lim_{b \rightarrow \infty} \lim_{a \rightarrow \infty} \mathbf{c}_{a,b}$  and  $\lim_{a \rightarrow \infty} \lim_{b \rightarrow \infty} \mathbf{c}_{a,b}$ . The difference is the term in the numerator that becomes

$$\mathbb{E}[2(|\square| + |\boxminus| + |\boxplus| + |\boxtimes| + |\boxdot| + |\boxminus| + |\boxplus| + |\boxtimes|) + 4|\boxplus|]. \quad \square$$

The presence of double limits in this corollary is difficult to avoid. One could wish to get results about averages taken on boxes  $[0, a] \times [0, b]$  with  $a$  and  $b$  having the same order of magnitude. But this type of results is related to decorrelation properties of the process which are more and more difficult to prove as lines get closer to the diagonal, as we can see on Figure 4.12.

## 4.8 Perspectives

In this chapter we defined the Poisson-Kirchhoff model as a system of vertical and horizontal broken weighted lines with a Markovian reversible dynamic that preserves Kirchhoff's node law. In doing so, we made several assumptions, some of which could be relaxed, yet might still lead to tractable (and still reversible) dynamics. For instance one could look at:

- Models where the distribution of weights on crossing events depends on the value of the two entries and not only through their sums. By relaxing this condition, we would recover the first model of [BEGG16] as explained in Section 4.6.3.
- Models where the measures  $\nu_V$  and  $\nu_H$  may have infinite mass. Relaxing the finite mass assumption should make it be possible to construct systems which are self similar, i.e. invariant by re-scaling of both space and weights simultaneously. Such models will be obtained by choosing  $g_V$  and  $g_H$  of the form  $s \mapsto s^{-\alpha}$ .
- More generally, one could consider models whose lines are not vertical and horizontal anymore, but can instead make an angle with the axis. A trivial example, with deterministic dynamics, is the Crofton model also called Poisson Line Process [Cro68].

Moreover, we could also be interested by the fluctuations of these models. For now, we know that the Hammersley processes that are in the KPZ family class have fluctuations in  $n^{1/3}$  [CG05, CG06]. We could also prove that the Model 17 of Table 4.3 has Gaussian fluctuations as said in Section 4.6.6. For now, we think that this should be the only two regimes of fluctuations, and that fluctuations depends on the support of  $\nu_V$  and  $\nu_H$ . In particular, we conjecture that if the supports of both  $\nu_V$  and  $\nu_H$  are respectively included, either in  $\mathbb{R}_+$  and  $\mathbb{R}_-$ , or in  $\mathbb{R}_-$  and  $\mathbb{R}_+$ , then the fluctuations should be Tracy-Widom and should be Gaussian otherwise.



# Bibliographie

- [AB09] G Ambrus and I Bárány. Longest convex chains. *Random Structures & Algorithms*, 35(2) :137–162, 2009.
- [AC21] T Alberts and E Cator. On the passage time geometry of the last passage percolation problem. *Latin American Journal of Probability and Mathematical Statistics*, 18 :211, 01 2021.
- [AD95] D Aldous and P Diaconis. Hammersley’s interacting particle process and longest increasing subsequences. *Probability Theory and Related Fields*, 103(2) :199–213, 1995.
- [AD99] D Aldous and P Diaconis. Longest increasing subsequences : from patience sorting to the Baik-Deift-Johansson theorem. *Bulletin of the American Mathematical Society*, 36(4) :413–432, 1999.
- [AG11] S Asmussen and PW Glynn. A new proof of convergence of MCMC via the ergodic theorem. *Statistics & Probability Letters*, 81(10) :1482–1485, 2011.
- [Bah98] C Bahadoran. Hydrodynamical limit for spatially heterogeneous simple exclusion processes. *Probability Theory and Related Fields*, 110(3) :287–331, 1998.
- [Bar01] Y Baryshnikov. GUEs and queues. *Probability Theory and Related Fields*, 119(2) :256–274, 2001.
- [Bax72] RJ Baxter. Partition function of the eight-vertex lattice model. *The Annals of Physics*, 70(1) :193–228, 1972.
- [Bax82] RJ Baxter. *Exactly solved models in statistical mechanics*. London : Academic Press Inc., 1982.
- [BB68] RM Baer and P Brock. Natural sorting over permutation spaces. *Mathematics of Computation*, 22(102) :385–410, 1968.
- [BCG<sup>+</sup>15] A Briquet, P Chassaing, L Gerin, M Krikun, and S Popov. Processus de branchement avec annihilation ballistique. Presentation given in *journées ALÉA 2015*, 2015.
- [BCG16] A Borodin, I Corwin, and V Gorin. Stochastic six-vertex model. *Duke Mathematical Journal*, 165(3) :563 – 624, 2016.

- [BDJ99] J Baik, P Deift, and K Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *Journal of the American Mathematical Society*, 12(4) :1119–1178, 1999.
- [BEGG16] A-L Basdevant, N Enriquez, L Gerin, and J-B Gou  r  . Discrete Hammersley’s lines with sources and sinks. *Latin American Journal of Probability and Mathematical Statistics*, 13, 03 2016.
- [BFS09] A Borodin, P L Ferrari, and T Sasamoto. Two speed TASEP. *Journal of Statistical Physics*, 137(5) :936–977, 2009.
- [BG19] A-L Basdevant and L Gerin. Longest increasing paths with gaps. *Latin American Journal of Probability and Mathematical Statistics*, 16 :1141, 01 2019.
- [BG22] A-L Basdevant and L Gerin. Longest increasing paths with Lipschitz constraints. To appear in *Annales de l’Institut Henri Poincar  *, 2022.
- [BGG18] A-L Basdevant, L Gerin, J-B Gou  r  , and A Singh. From Hammersley’s lines to Hammersley’s trees. *Probability Theory and Related Fields*, 171(1) :1–51, 2018.
- [BGH18] R Basu, S Ganguly, and A Hammond. The competition of roughness and curvature in area-constrained polymer models. *Communications in Mathematical Physics*, 364(3) :1121–1161, 2018.
- [BHMZ11] J Byers, B Heeringa, M Mitzenmacher, and G Zervas. Heapable sequences and subsequences. In *Proceedings of the Eighth Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, pages 33–44. SIAM, 2011.
- [BLM13] S Boucheron, G Lugosi, and P Massart. *Concentration inequalities. A nonasymptotic theory of independence*. Oxford University Press, 2013.
- [BM20] N Broutin and J-F Marckert. The combinatorics of the colliding bullets. *Random Structures & Algorithms*, 56(2) :401–431, 2020.
- [BR00] J Baik and EM Rains. Limiting distributions for a polynuclear growth model with external sources. *Journal of Statistical Physics*, 100(3) :523–541, 2000.
- [BT21] Q Berger and N Torri. Beyond Hammersley’s last-passage percolation : a discussion on possible local and global constraints. *Annales de l’Institut Henri Poincar  *, 8(2) :213–241, 2021.
- [BW88] B Bollob  s and P Winkler. The longest chain among random points in euclidean space. *Proceedings of the American Mathematical Society*, 103(2) :347–353, 1988.
- [Cas18] J Casse. Edge correlation function of the 8-vertex model when  $a + c = b + d$ . *Annales de l’Institut Henri Poincar  *, 5(4) :557–619, 2018.
- [Cas19] J Casse. Generalised last passage percolation : invariant laws on the cylinders. *arXiv preprint arXiv :1912.00645*, 2019.
- [CG05] E Cator and P Groeneboom. Hammersley’s process with sources and sinks. *The Annals of Probability*, 33(3) :879–903, 2005.

- [CG06] E Cator and P Groeneboom. Second class particles and cube root asymptotics for Hammersley’s process. *The Annals of Probability*, 34(4) :1273–1295, 2006.
- [CG19] F Ciech and N Georgiou. Order of the variance in the discrete Hammersley process with boundaries. *Journal of Statistical Physics*, 176(3) :591–638, 2019.
- [Cor16] I Corwin. Kardar-Parisi-Zhang universality. *Notices of the American Mathematical Society*, 63 :230–239, 03 2016.
- [CP11] E Cator and LPR Pimentel. A shape theorem and semi-infinite geodesics for the Hammersley model with random weights. *Latin American journal of probability and mathematical statistics*, 8 :163–175, 2011.
- [CP12] E Cator and LPR Pimentel. Busemann functions and equilibrium measures in last passage percolation models. *Probability Theory and Related Fields*, 154(1) :89–125, 2012.
- [CPE96] H Cohn, JG Propp, and ND Elkies. Local statistics for random domino tilings of the Aztec diamond. *Duke Mathematical Journal*, 85(math. CO/0008243) :117–166, 1996.
- [CPS12] E Cator, LPR Pimentel, and M Souza. Influence of the initial condition in equilibrium last-passage percolation models. *Electronic Communications in Probability*, 17 :1–7, 2012.
- [CR97] P Covert and F Rezakhanlou. Hydrodynamic limit for particle systems with nonconstant speed parameter. *Journal of Statistical Physics*, 88(1) :383–426, 1997.
- [Cro68] MW Crofton. VII. on the theory of local probability, applied to straight lines drawn at random in a plane; the methods used being also extended to the proof of certain new theorems in the integral calculus. *Philosophical Transactions of the Royal Society of London*, (158) :181–199, 1868.
- [DCGH<sup>+</sup>18] H Duminil-Copin, M Gagnebin, M Harel, I Manolescu, and V Tassion. The Bethe ansatz for the six-vertex and XXZ models : An exposition. *Probability Surveys*, 15 :102–130, 2018.
- [DJP18] P Dey, M Joseph, and R Peled. Longest increasing path within the critical strip. *arXiv preprint arXiv :1808.08407*, 2018.
- [Dou05] Y Doumerc. *Matrices aléatoires, processus stochastiques et groupes de réflexions*. PhD thesis, Toulouse 3, 2005.
- [Dur19] R Durrett. *Probability : theory and examples*, volume 49. Cambridge university press, 5th edition, 2019.
- [ES35] P Erdős and G Szekeres. A combinatorial problem in geometry. *Compositio mathematica*, 2 :463–470, 1935.
- [Fer92] PA Ferrari. Shocks in the Burgers equation and the asymmetric simple exclusion process. In *Statistical Physics, Automata Networks and Dynamical Systems*, pages 25–64. Springer, 1992.



- [FP05] PA Ferrari and LPR Pimentel. Competition interfaces and second class particles. *The Annals of Probability*, 33(4) :1235–1254, 2005.
- [FRT54] JS Frame, GB Robinson, and RM Thrall. The hook graphs of the symmetric group. *Canadian Journal of Mathematics*, 6 :316–324, 1954.
- [FW70] C Fan and FY Wu. General lattice model of phase transitions. *Physical Review B*, 2(3) :723, 1970.
- [Gan98] AJ Ganesh. Large deviations of the sojourn time for queues in series. *Annals of Operations Research*, 79 :3–26, 1998.
- [Ger13] L Gerin. Percolation et processus de croissance aléatoires. Lecture, 2013.
- [GKS10] N Georgiou, R Kumar, and T Seppäläinen. TASEP with discontinuous jump rates. *Latin American Journal of Probability and Mathematical Statistics*, 7 :293–318, 2010.
- [GRAS17] N Georgiou, F Rassoul-Agha, and T Seppäläinen. Geodesics and the competition interface for the corner growth model. *Probability Theory and Related Fields*, 169(1) :223–255, 2017.
- [Gro01] P Groeneboom. Ulam’s problem andammersley’s process. *The Annals of Probability*, 29(2) :683–690, 2001.
- [Gro02] P Groeneboom. Hydrodynamical methods for analyzing longest increasing subsequences. *Journal of Computational and Applied Mathematics*, 142(1) :83–105, 2002. Probabilistic Methods in Combinatorics and Combinatorial Optimization.
- [GTW01] J Gravner, C Tracy, and H Widom. Limit theorems for height fluctuations in a class of discrete space and time growth models. *Journal of Statistical Physics*, 102, 05 2001.
- [GW91] PW Glynn and W Whitt. Departures from many queues in series. *The Annals of Applied Probability*, pages 546–572, 1991.
- [Ham72] JM Hammersley. A few seedlings of research. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 345–394, 1972.
- [HM80] SP Hastings and JB McLeod. A boundary value problem associated with the second Painlevé transcendent and the Korteweg-de Vries equation. *Archive for Rational Mechanics and Analysis*, 73(1) :31–51, 1980.
- [HMO02] BM Hambly, JB Martin, and N O’Connell. Concentration results for a Brownian directed percolation problem. *Stochastic processes and their applications*, 102(2) :207–220, 2002.
- [HS20] A Hammond and S Sarkar. Modulus of continuity for polymer fluctuations and weight profiles in Poissonian last passage percolation. *Electronic Journal of Probability*, 25 :1–38, 2020.

- 
- [HST21] J Haslegrave, V Sidoravicius, and L Tournier. Three-speed ballistic annihilation : phase transition and universality. *Selecta Mathematica*, 27(5) :1–38, 2021.
- [IB15] G Istrate and C Bonchiş. Partition into heapable sequences, heap tableaux and a multiset extension of Hammersley’s process. In *Annual Symposium on Combinatorial Pattern Matching*, pages 261–271. Springer, 2015.
- [Joh00a] K Johansson. Shape fluctuations and random matrices. *Communications in Mathematical Physics*, 209(2) :437–476, 2000.
- [Joh00b] K Johansson. Transversal fluctuations for increasing subsequences on the plane. *Probability Theory and Related Fields*, 116(4) :445–456, 2000.
- [Jon82] D Jonsson. Some limit theorems for the eigenvalues of a sample covariance matrix. *Journal of Multivariate Analysis*, 12(1) :1–38, 1982.
- [JPS95] W Jockusch, JG Propp, and P Shor. Random domino tilings and the arctic circle theorem. *arXiv preprint math/9801068*, 1995.
- [KDN90] D Kandel, E Domany, and B Nienhuis. A six-vertex model as a diffusion problem : derivation of correlation functions. *Journal of Physics A : Mathematical and General*, 23(15) :L755, 1990.
- [Kin68] JFC Kingman. The ergodic theory of subadditive stochastic processes. *Journal of the Royal Statistical Society : Series B (Methodological)*, 30(3) :499–510, 1968.
- [Kin73] JFC Kingman. Subadditive ergodic theory. *The Annals of Probability*, pages 883–899, 1973.
- [Knu70] D Knuth. Permutations, matrices, and generalized Young tableaux. *Pacific journal of mathematics*, 34(3) :709–727, 1970.
- [Kon16] T Konstantopoulos. A review of Burke’s theorem for Brownian motion. *Queueing systems*, 83(1) :1–12, 2016.
- [KPS19] A Knizel, L Petrov, and A Saenz. Generalizations of TASEP in discrete and continuous inhomogeneous space. *Communications in Mathematical Physics*, 372(3) :797–864, 2019.
- [KPZ86] M Kardar, G Parisi, and Y-C Zhang. Dynamic scaling of growing interfaces. *Physical Review Letters*, 56(9) :889, 1986.
- [KRL95] PL Krapivsky, S Redner, and F Leyvraz. Ballistic annihilation kinetics : The case of discrete velocity distributions. *Physical Review E*, 51(5) :3977, 1995.
- [Lig75] TM Liggett. Ergodic theorems for the asymmetric simple exclusion process. *Transactions of the American Mathematical Society*, 213 :237–261, 1975.
- [Lig99] TM Liggett. *Stochastic interacting systems : contact, voter and exclusion processes*, volume 324. Springer science & Business Media, 1999.
- [LS77] BF Logan and LA Shepp. A variational problem for random Young tableaux. *Advances in Mathematics*, 26(2) :206–222, 1977.

- [Mar06] James B. Martin. Last-passage percolation with general weight distribution. *Markov Processes and Related Fields*, 12(2) :273–299, 2006.
- [Mel21] P Melotti. The free-fermion eight-vertex model : couplings, bipartite dimers and Z-invariance. *Communications in Mathematical Physics*, 381(1) :33–82, 2021.
- [MG69] CT MacDonald and JH Gibbs. Concerning the kinetics of polypeptide synthesis on polyribosomes. *Biopolymers : Original Research on Biomolecules*, 7(5) :707–725, 1969.
- [MGP68] CT MacDonald, JH Gibbs, and AC Pipkin. Kinetics of biopolymerization on nucleic acid templates. *Biopolymers : Original Research on Biomolecules*, 6(1) :1–25, 1968.
- [MT93] SP Meyn and RL Tweedie. *Markov chains and stochastic stability*. Springer Science & Business Media, 1993.
- [Mut79] EJ Muth. The reversibility property of production lines. *Management science*, 25(2) :152–158, 1979.
- [Occ19] A Occelli. *KPZ universality for last passage percolation models*. PhD thesis, Rheinische Friedrich-Wilhelms-Universität Bonn, 2019.
- [OY01] N O’Connell and M Yor. Brownian analogues of Burke’s theorem. *Stochastic Processes and their Applications*, 96(2) :285–304, 2001.
- [OY02] N O’Connell and M Yor. A representation for non-colliding random walks. *Electronic Communications in Probability*, 7 :1–12, 2002.
- [Pau35] L Pauling. The structure and entropy of ice and of other crystals with some randomness of atomic arrangement. *Journal of the American Chemical Society*, 57(12) :2680–2684, 1935.
- [PS00] M Prähofer and H Spohn. Universal distributions for growth processes in  $1+1$  dimensions and random matrices. *Physical review letters*, 84(21) :4882, 2000.
- [PS02] M Prähofer and H Spohn. Scale invariance of the PNG droplet and the Airy process. *Journal of Statistical Physics*, 108(5) :1071–1106, 2002.
- [Rai98] EM Rains. Increasing subsequences and the classical groups. *Electronic Journal of Combinatorics*, 5 :Art–No, 1998.
- [Rob38] GB Robinson. On the representations of the symmetric group. *American Journal of Mathematics*, pages 745–760, 1938.
- [Ros81] H Rost. Non-equilibrium behaviour of a many particle process : density profile and local equilibria. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 58(1) :41–53, 1981.
- [RSSV10] L Rolla, V Sidoravicius, D Surgailis, and ME Vares. The discrete and continuum broken line process. *Markov Processes and Related Fields*, 16(1) :79–116, 2010.

- 
- [Sch61] C Schensted. Longest increasing and decreasing subsequences. *Canadian Journal of mathematics*, 13 :179–191, 1961.
- [Sep96] T Seppäläinen. A microscopic model for the Burgers equation and longest increasing subsequences. *Electronic Journal of Probability*, 1 :1–51, 1996.
- [Sep97] T Seppäläinen. Increasing sequences of independent points on the planar lattice. *The Annals of Applied Probability*, 7(4) :886–898, 1997.
- [Sep98a] T Seppäläinen. Exact limiting shape for a simplified model of first-passage percolation on the plane. *The Annals of Probability*, pages 1232–1250, 1998.
- [Sep98b] T Seppäläinen. Hydrodynamic scaling, convex duality, and asymptotic shapes of growth models. *Markov Processes and Related Fields*, 01 1998.
- [Sep00] T Seppäläinen. Strong law of large numbers for the interface in ballistic deposition. *Annales de l'Institut Henri Poincaré*, 36 :691–736, 11 2000.
- [Sep09] T Seppäläinen. Lecture notes on the corner growth model. Unpublished notes, 2009.
- [Sep12] T Seppäläinen. Scaling for a one-dimensional directed polymer with boundary conditions. *The Annals of Probability*, 40(1) :19–73, 2012.
- [SK90] W Szczotka and FP Kelly. Asymptotic stationarity of queues in series and the heavy traffic approximation. *The Annals of Probability*, pages 1232–1248, 1990.
- [Spi70] F Spitzer. Interaction of Markov processes. *Advances in Mathematics*, 5(2) :246–290, 1970.
- [Sut70] B Sutherland. Two-dimensional hydrogen bonded crystals without the ice rule. *Journal of Mathematical Physics*, 11(11) :3183–3186, 1970.
- [TW94] CA Tracy and H Widom. Level-spacing distributions and the Airy kernel. *Communications in Mathematical Physics*, 159(1) :151–174, 1994.
- [TW09] CA Tracy and H Widom. Asymptotics in ASEP with step initial condition. *Communications in Mathematical Physics*, 290(1) :129–154, 2009.
- [Ula61] SM Ulam. Monte-Carlo calculations in problems of mathematical physics. *Modern Mathematics for the Engineers*, pages 261–281, 1961.
- [VK77] AM Vershik and SV Kerov. Asymptotics of the plancherel measure of the symmetric group and the limiting form of young tableaux. In *Doklady Akademii Nauk*, volume 233, pages 1024–1027. Russian Academy of Sciences, 1977.
- [Vol59] MJ Vold. A numerical approach to the problem of sediment volume. *Journal of colloid science*, 14(2) :168–174, 1959.
- [Wüt02] MV Wüthrich. Asymptotic behaviour of semi-infinite geodesics for maximal increasing subsequences in the plane. In *In and out of equilibrium*, pages 205–226. Springer, 2002.