



HAL
open science

Grandes déviations pour des modèles de percolation dirigée et des matrices aléatoires.

Jean-Paul Ibrahim

► **To cite this version:**

Jean-Paul Ibrahim. Grandes déviations pour des modèles de percolation dirigée et des matrices aléatoires.. Mathématiques [math]. Université Paul Sabatier - Toulouse III, 2010. Français. NNT : . tel-00577242

HAL Id: tel-00577242

<https://theses.hal.science/tel-00577242>

Submitted on 16 Mar 2011

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.

THÈSE

présentée en vue de l'obtention du

DOCTORAT

DE

L'UNIVERSITÉ DE TOULOUSE

délivré par : *Université Paul Sabatier, Toulouse III*

Discipline : Mathématiques

Spécialité : Probabilités

par

Jean-Paul IBRAHIM

Grandes Déviations pour des Modèles de Percolation Dirigée et pour des Matrices Aléatoires

Soutenue le xx xxx 2010 devant le jury composé de

Bernard BERCU	Université Bordeaux 1	Rapporteur
Thierry BODINEAU	ENS Paris	Rapporteur
Mireille CAPITAINE	Université Toulouse III	Examinateur
Fabrice GAMBOA	Université Toulouse III	Examinateur
Michel LEDOUX	Université Toulouse III	Directeur de thèse
Alain ROUAULT	Université de Versailles	Examinateur

Institut Mathématique de Toulouse

UMR CNRS 5219, Université de Toulouse, 31062 Toulouse, France

Grandes déviations pour des modèles de percolation dirigées et pour des matrices aléatoires

Durant cette thèse, on a étudié essentiellement deux modèles aléatoires qui, malgré leur différence apparente, cachent un intérêt commun et mettent en évidence des phénomènes mathématiques et physiques communs.

Le modèle de percolation de dernier passage dans le plan (last-passage directed percolation model ou LPP) est un modèle de percolation orientée bidimensionnel. Il fait partie d'une vaste liste de modèles de croissance et sert à modéliser des phénomènes dans des domaines variés. Dans la première partie de cette thèse, on s'est intéressé essentiellement aux propriétés de grandes déviations de ce modèle. On a également examiné les fluctuations transversales du même modèle. Toute cette étude a été faite dans le cadre d'un rectangle fin.

Parallèlement aux travaux sur les modèles de croissance, on a étudié un autre sujet qui émerge également du monde de la Physique : celui des matrices aléatoires. Ces matrices se divisent en deux catégories principales introduites à une vingtaine d'années d'intervalle : les matrices de covariance empirique et les matrices de Wigner. L'étendue du champ d'application de ces matrices est tellement vaste qu'on peut les rencontrer presque dans toutes les filières scientifiques : probabilité, combinatoire, physique atomique, statistique multivariée, télécommunication théorie des représentations, etc. Parmi les objets mathématiques les plus étudiés, on cite la loi jointe des valeurs propres, la densité spectrale, l'espacement des valeurs propres, la plus grande valeur propre et les vecteurs propres associés. En mécanique quantique par exemple, les valeurs propres d'une matrice du GUE modélisent les niveaux d'énergie d'un électron autour du noyau tandis que le vecteur propre associé à la plus grande valeur propre d'une matrice de covariance empirique indique la direction ou l'axe principal en analyse de données.

Comme pour le modèle de percolation dirigée, on s'est intéressé en particulier aux propriétés de grandes déviations de la valeur propre maximale d'un certain type de matrices de covariance empirique. Cette étude pourrait avoir des applications en statistique et notamment en analyse en composantes principales.

Malgré l'apparente différence, la théorie des matrices aléatoires est strictement liée au modèle de percolation dirigée. Leurs structures de corrélation se ressemblent dans certains cas d'une manière troublante. La convergence des fluctuations, dans les deux cas, vers la célèbre loi de Tracy-Widom en est un bon exemple.

Mots clés : Percolation dirigée, temps de dernier passage, percolation Brownienne, grandes déviations, fluctuations longitudinales, fluctuations transversales, approximation K-M-T, matrices aléatoires, matrice de covariance empirique, valeur propre maximale.

Remerciements

Tout d'abord, je tiens à exprimer mes sincères remerciements à mon directeur de thèse Michel Ledoux. Je le remercie entre autres pour sa disponibilité, son sérieux, sa rigueur et son dévouement pour la recherche. Durant ces quatre années de thèse, j'ai appris, en travaillant avec lui, à apprécier la recherche dans toutes ses étapes et ses « aléas ». Toutes les discussions que nous avons pu avoir ensemble m'ont enrichi aussi bien sur le plan scientifique que personnel. Pour tout dire, ça a été un grand plaisir de travailler sous sa direction.

Ma reconnaissance va également vers mes rapporteurs de thèse :

-Thierry Bodineau que je connais plus à travers ses travaux de recherche qui ont inspiré les nôtres et que je remercie d'avoir accepté cette tâche.

-Bernard Bercu que j'ai eu la chance d'avoir en cours de DEA et que je remercie particulièrement d'avoir annoté soigneusement le manuscrit de la thèse.

Je remercie aussi les membres du jury : Mireille Capitaine et Fabrice Gamboa que je connais depuis mon DEA et qui sont remarquables de gentillesse et de sympathie et Alain Rouault que je remercie infiniment d'avoir accepté de faire partie de mon jury de thèse.

Je remercie aussi tous les professeurs que j'ai eus tout au long de mon cursus universitaire. Je pense notamment à Monique Pontier qui m'a toujours encouragé. Merci à tous les matheux que j'ai rencontrés durant cette thèse. Je pense surtout à Gregorio avec qui j'ai eu des échanges intéressants.

Un grand merci à tous mes amis et mes collègues de travail. Je pense aux nombreux co-bureaux rencontrés grâce aux moult déménagements : Matthieu, Maxime, Myriam, Michel, Erwan et Nolwen. Je pense aussi à mes partenaires de tennis et tous les membres du club BTR. Merci à François, Romain, Vincent, Damien, Fabrice, Sarah, Pierre-François, Aurélie et Hervé pour tous les bons moments passés ensemble.

Un merci spécial à Marion, ma meilleure amie et adepte de Bach et du gâteau basque. J'espère qu'elle rejoindra au plus vite la « cour des grands ».

Un merci du fond du cœur à mon grand frère Marc qui fut pendant mon enfance mon idéal et qui depuis tout petit m'a transmis son besoin d'apprendre et d'analyser. Malgré l'éloignement, nous avons pu avoir énormément de discussions scientifiques durant lesquelles il a partagé avec moi ses talents en informatique. Sans lui, la majeure partie des simulations n'auraient pas pu voir le jour.

Un autre merci à mon cousin Elie grâce à qui j'ai débarqué dans cette ville (Toulouse) et qui m'est venu en aide aux moments où j'en avais le plus besoin. Il m'a fait profiter également de ses talents de « bricoleur du numérique ». Je lui souhaite bonne chance dans sa vie privée et professionnelle.

Je n'oublie pas mes parents qui se sont dévoués pour nous et qui m'ont toujours soutenu pendant ce périple. Je n'oublie pas non plus mes deux petits frères, Pierre et Raphaël à qui je souhaite

bon courage avec le rasoir. Je pense aussi à tous mes cousins et cousines avec qui j'ai passé d'agréables étés.

Pour terminer, je dédie ce mémoire de thèse aux deux êtres qui me sont le plus chers : ma petite fille Mélina et ma femme Céline.

Table des matières

1	Processus déterminantaux et Ensemble Gaussien Unitaire	13
1.1	Ensemble Gaussien Unitaire	16
1.1.1	Loi jointe et mesure spectrale	17
1.1.2	Grandes déviations	20
1.1.3	Représentation déterminantale et fluctuations	22
1.1.4	GUE et percolation dirigée	26
2	Percolation dirigée de dernier passage	27
2.1	Définition du modèle	27
2.2	Percolation dirigée et applications	29
2.3	Résultats connus	32
2.3.1	Résultats antérieurs aux travaux de Johansson : loi des grands nombres et théorème central limite	32
2.3.2	Processus déterminantal : poids géométriques et exponentiels	33
2.4	Modèles de percolation planaires voisins	38
2.4.1	Modèle de percolation de premier passage de Seppäläinen	38
2.4.2	Percolation dirigée sur un triangle	40
3	Percolation dirigée sur un rectangle fin et universalité	47
3.1	Principes d'invariance et LPP	48
3.1.1	Un couplage simple	48
3.1.2	Le théorème de plongement de Skorokhod et l'approximation K-M-T	51
3.1.3	Un théorème d'invariance faible de Chatterjee	53
3.2	Fluctuations longitudinales et transversales	53
3.2.1	Arguments heuristiques	54
3.2.2	Fluctuations longitudinales et inégalités exponentielles	55
3.2.3	Fluctuations transversales sur un rectangle fin	60
3.2.4	Convergence des moments	65
3.3	Grandes déviations	66
3.3.1	Grandes déviations du premier ordre (N).	67
3.3.2	Asymptotiques des grandes déviations du second ordre (\sqrt{Nk}).	69
3.3.3	Inégalités optimales non-asymptotiques.	72
4	Large deviations for directed percolation on a thin rectangle.	75
4.1	Results	76
4.2	Proof of Theorem 4.1.	81
4.3	Exponential-tailed distribution and the K-M-T approximation	85

4.4	Subexponential weights and the Skorokhod embedding	87
4.5	Small and large deviations inequalities	90
4.5.1	Standard normal variables	90
4.5.2	Finite moment-generating function case	91
5	Matrices aléatoires : matrices de covariance empirique	93
5.1	Matrices de covariance empirique et applications	94
5.2	Résultats connus : modèle blanc	96
5.3	Résultats connus : modèle non-blanc	101
5.4	Grandes déviations et matrices de covariance empirique.	105
5.4.1	Grandes déviations lorsque k est négligeable devant N	106
5.5	Théorèmes d'invariance et matrices de covariance empirique	117
	Bibliographie	121

Introduction

Introduction en Français

Durant cette thèse, on a étudié essentiellement deux modèles aléatoires qui, malgré leur différence apparente, cachent un intérêt commun et mettent en évidence des phénomènes mathématiques et physiques communs.

Le modèle de percolation de dernier passage dans le plan (last-passage directed percolation model ou LPP) est un modèle de percolation orientée bidimensionnel. Il fait partie d'une vaste liste de modèles de croissance et sert à modéliser des phénomènes dans des domaines variés. Bien que simple à introduire, la version discrète la plus connue du modèle, et celle étudiée ici, est très riche aussi bien sur le plan mathématique qu'au niveau des applications. Les premières études sur ce modèle ont été réalisées dans les années 70 au sujet des files d'attente en série. Au début des années 80, et dans le cadre d'une étude sur le comportement des particules sur la droite discrétisée \mathbb{Z} , Rost [Ros81] fait le lien entre une version processus du modèle, connue sous le nom de "corner growth model", et le processus d'exclusion totalement asymétrique issu de la physique des particules. Glynn et Whitt [GW91] poursuivent au début des années 90 l'étude sur les files d'attente en série et réussissent à dégager quelques propriétés de convergence du premier ordre (loi des grands nombres) mais sans pouvoir pour autant expliciter la limite. Il faudra attendre le début du siècle pour que Johansson [Joh00a] obtienne des résultats précis de convergence du premier et second ordre et des formules exactes de grandes déviations.

Le modèle existe principalement sous deux versions : continue et discrète. La version continue ne sera pas détaillée ici mais notons qu'elle a servi essentiellement dans l'étude de la plus longue sous-suite croissante d'une permutation aléatoire. C'est la version discrète qui nous intéresse le plus et qui sera analysée en détail dans cette première partie de la thèse. D'autres modèles planaires voisins seront également présentés pour compléter le panorama.

Parallèlement aux travaux sur les modèles de croissance, on a étudié un autre sujet qui émerge également du monde de la Physique : celui des matrices aléatoires. Ces matrices se divisent en deux catégories principales introduites à une vingtaine d'années d'intervalle. En 1928, Wishart [Wis28] étudie la loi d'une matrice de covariance empirique à entrées Gaussiennes. Ce type de matrices, connu maintenant sous le nom de matrices de Wishart, aura un impact direct en statistique multivariée. Vingt sept ans plus tard, Wigner [Wig55] introduit en 1955 les matrices Hermitiennes à entrées indépendantes connues sous son nom. Ces matrices sont utilisées pour résoudre des problèmes de physique atomique. Wigner fait des approximations d'opérateurs Hamiltonien auto-adjoint dans un espace de Hilbert de dimension infinie par un ensemble de matrices aléatoires ; les valeurs propres de telles matrices représentent les niveaux d'énergie

d'un électron. Depuis Wigner, cette branche des mathématiques a connu un progrès considérable. Voici quelques dates importantes qui ont marqué l'évolution de cette théorie. En 1955, Wigner [Wig55] démontre que la mesure spectrale d'une matrice de l'ensemble Gaussien unitaire (GUE) converge vers la loi du demi-cercle. Plus tard, en 1967, Marchenko et Pastur [MP67] démontrent un résultat similaire pour les matrices de Wishart. Dans les années 70, les travaux de Dyson et Mehta ont permis de transformer les matrices aléatoires d'un simple outil mathématique en une théorie mathématique. En 1994, Tracy et Widom [TW94] découvrent la loi universelle qui porte leur nom. Celle-ci est la loi limite des fluctuations de la plus grande valeur propre de plusieurs modèles matriciels. Durant les dix dernières années, de nombreux auteurs ont étendu les résultats les plus importants pour couvrir une large classe de matrices aléatoires.

L'étendue du champ d'application de ces matrices est tellement vaste qu'on peut les rencontrer presque dans toutes les filières scientifiques : probabilité, combinatoire, physique atomique, statistique multivariée, télécommunication théorie des représentations, etc. Parmi les objets mathématiques les plus étudiés, on cite la loi jointe des valeurs propres, la densité spectrale, l'espacement des valeurs propres, la plus grande valeur propre et les vecteurs propres associés. En mécanique quantique par exemple, les valeurs propres d'une matrice du GUE modélisent les niveaux d'énergie d'un électron autour du noyau tandis que le vecteur propre associé à la plus grande valeur propre d'une matrice de covariance empirique indique la direction ou l'axe principal en analyse des données.

Malgré l'apparente différence, la théorie des matrices aléatoires est strictement liée au modèle de percolation dirigée. Leurs structures de corrélation se ressemblent dans certains cas d'une manière troublante. Par exemple, Johansson [Joh00a] a démontré que le temps de dernier passage dans le cas de variables exponentielles est égale en loi à la plus grande valeur propre d'une matrice de Wishart. De plus, la loi limite (Tracy-Widom) des fluctuations est la même pour certains modèles de percolation et d'autres de matrices aléatoires.

Les travaux effectués pendant cette thèse portent essentiellement sur les propriétés asymptotiques des grandes déviations du modèle de percolation dirigée ainsi que sur celles d'un modèle particulier de matrices de covariance empirique. Les grandes déviations d'un phénomène sont des événements rares qui peuvent se produire et qui constituent une exception par rapport à la stabilité globale du modèle. En général, on cherche à estimer la probabilité qu'une variable ou un vecteur aléatoire s'éloignent de leur espérance limite.

Dans le premier chapitre, on parle d'abord des processus ponctuels déterminantaux dans un cadre général. Ensuite, on fournit les résultats les plus importants dans le cas particulier de l'ensemble Gaussien Unitaire. Ce chapitre ne résume en aucun cas nos contributions mais offre aux lecteurs une vision globale des différents modèles étudiés dans cette thèse.

Dans le deuxième chapitre, on traite du modèle de percolation dirigée. On introduit le modèle en donnant quelques applications, puis on effectue un rappel des résultats connus avec un survol des preuves les plus importantes. A la fin, on expose quelques modèles bidimensionnels proches du modèle étudié.

Le troisième et le quatrième chapitres constituent le noyau de notre contribution en modèle percolation dirigée sur un rectangle fin. Après un rappel des résultats connus sur les fluctuations, on

expose les formules de grandes déviations et de fluctuations transversales obtenus durant cette thèse. Pour éviter toutes formes de redondance, les preuves des formules de grandes déviations seront détaillées uniquement dans le quatrième chapitre sous forme d'une publication. Par ailleurs, une preuve du théorème sur les fluctuations transversales est donnée dans le troisième chapitre.

Le cinquième chapitre traite des matrices aléatoires et notamment celles de covariance empirique. Après un résumé bibliographique assez détaillé, on énonce et on démontre notre résultat principal sur les grandes déviations de la plus grande valeur propre d'une matrice de covariance empirique perturbée. On clôt ce chapitre et l'ensemble du mémoire de thèse par une discussion sur l'application des principes d'invariance aux matrices aléatoires.

Chapitre 1

Processus déterminantaux et Ensemble Gaussien Unitaire

Dans ce chapitre, on introduit un type de processus ponctuels déterminantaux. Bien que les propriétés de ces processus n'ont pas été exploitées directement durant cette thèse, leur introduction offre une vision d'ensemble permettant aux lecteurs de comprendre les liens entre les deux modèles principalement étudiés. Dans la deuxième partie de ce chapitre, on parle de l'ensemble Gaussien unitaire qui constitue le pilier d'un certain nombre de résultats obtenus durant cette thèse.

Dans le cadre de notre étude, ces processus ponctuels continus (resp. discret) sont souvent assimilés à des vecteur aléatoires de \mathbb{R}^k (resp. \mathbb{Z}^k). Ils sont issus de problèmes divers liés souvent à la physique et décrivent le comportement de k objets aléatoires (particules) se trouvant sur la droite réelle (resp. discrète) et interagissant entre eux suivant une corrélation bien définie. La particularité des processus étudiés ici est que, dans certain cas, ils mettent en évidence une fonction de corrélation faisant intervenir le déterminant de Vandermonde au carré, [Kön05]. De plus, Cette fonction de corrélation admet, sous des conditions bien précises, ce qu'on appelle une représentation déterminantale utilisant un noyau de polynômes orthogonaux, [Dei99]. Cette représentation facilite l'étude du comportement de l particules en sachant la position des $k - l$ particules restantes. De plus, elle permet l'analyse asymptotique de certaines quantités (lorsque le nombre de particules tend vers l'infini).

On commence d'abord par exposer la forme générale de la fonction de corrélation des processus en question. Ensuite, on fournit quelques exemples ou cas particuliers de processus ponctuels déterminantaux. A la fin ce chapitre, on parcourt les résultats connus les plus importants dans le cas particulier de l'Ensemble Gaussien Unitaire (GUE) : le modèle qui a révélé ce type de corrélation et qui a permis de faire évoluer toute une branche des mathématiques.

On désigne par $X = (X_1, \dots, X_k)$ un vecteur aléatoire de \mathbb{R}^k (resp. \mathbb{Z}^k) et on suppose que les X_i sont ordonnés *i.e.* $X_1 < X_2 < \dots < X_k$. La mesure de probabilité annoncée là-dessus et décrivant la loi de X est de la forme

$$\mathbb{P}_k(dx) = \frac{1}{Z_k} \prod_{1 \leq i < j \leq k} (x_i - x_j)^2 \prod_{i=1}^k d\nu(x_i), \quad (1.1)$$

où x appartient à la chambre de Weyl de \mathbb{R}^k (resp. \mathbb{Z}^k) définie par

$$W_k(\mathbb{R}) := \{x \in \mathbb{R}^k; x_1 < x_2 < \dots < x_k\}.$$

Ici ν est une mesure positive définie sur \mathbb{R} (resp. \mathbb{Z}). La formule (1.1) est composée de trois termes : une constante de normalisation Z_k , le déterminant de Vandermonde au carré $\Delta_k(x)^2 = \det [(x_i^{j-1})_{i,j=1,\dots,k}]^2$ et le produit tensoriel k fois de ν . L'interaction entre les particules ou bien la corrélation entre les différentes coordonnées de X est décrite par le déterminant de Vandermonde au carré. Sans ce terme, les particules se déplacent indépendamment les unes des autres. C'est en quelque sorte une interaction répulsive à cause des termes en $(x_i - x_j)^2$. D'un point de vue physique, on peut imaginer qu'on a k particules sur la droite \mathbb{R} , qui se repoussent entre elles suivant un potentiel en $\log|x_i - x_j|$ et qui sont également soumises à un champs extérieur représenté par le logarithme de l'inverse de la densité de la mesure ν . Il est clair que \mathbb{P}_k est symétrique par rapport aux coordonnées de x . On peut alors étendre la probabilité \mathbb{P}_k en une mesure de probabilité \mathbb{Q}_k sur \mathbb{R}^k (resp. \mathbb{Z}^k) en supposant que $\mathbb{Q}_k(dx) = \mathbb{Q}_k(dy)$ lorsque x et y sont égales à une permutation près. Ainsi, pour tout $x \in \mathbb{R}^k$ (resp. $x \in \mathbb{Z}^k$), on définit la mesure

$$\mathbb{Q}_k(dx) = \frac{1}{k!} \mathbb{P}_k(dx).$$

Lorsque ν admet une densité par rapport à la mesure de Lebesgue, la mesure de probabilité \mathbb{Q}_k est plus adéquate pour exprimer la loi marginale des l premières variables ou bien ce qu'on appelle aussi la fonction de corrélation de X . Cette fonction de corrélation, notée $R_l^{(k)}$, permet de mieux comprendre le comportement des différentes composantes du processus ponctuels X et permet surtout de calculer la loi du $\max_{i=1,\dots,k} \{X_i\}$.

On s'intéresse surtout au cas où ν est σ -finie et admet une densité par rapport à la mesure de Lebesgue (resp. à la mesure de comptage). Dans le cas contraire, il est difficile d'explicitement les calculs. On suppose également que ν admet des moments finis de tout ordre. Cette dernière hypothèse est essentielle pour la construction des suites de polynômes orthogonaux. Dans ce cas, la mesure de probabilité \mathbb{P}_k est souvent appelée « Ensemble de polynômes orthogonaux ». Cette nomenclature provient de l'utilisation des polynômes orthogonaux pour réinterpréter \mathbb{P}_k . Plus précisément, on utilise la suite de polynômes orthogonaux construite sur $L^2(\mathbb{R}, \nu)$ (resp. $l^2(\mathbb{Z}, \nu)$) afin de représenter $R_l^{(k)}$ sous forme d'un déterminant d'un noyau de polynômes orthogonaux, ([Kön05], Sec 2). Cette représentation permet de calculer la loi marginale des l premières variables comme étant le déterminant du mineur d'ordre l d'une matrice de taille $k \times k$. Le processus ponctuel en question est ainsi appelé « **Processus Ponctuel Déterminantal** ».

En théorie, on peut remplacer ν dans (1.1) par n'importe quelle mesure positive de \mathbb{R} (resp. \mathbb{Z}) admettant des moments finis de tout ordre et ensuite modifier convenablement Z_k pour que \mathbb{P}_k reste une mesure de probabilité. On pourra alors dégager une représentation déterminantale de \mathbb{P}_k à condition de pouvoir expliciter les coefficients des polynômes orthogonaux de $L^2(\mathbb{R}, \nu)$ (resp. $l^2(\mathbb{Z}, \nu)$). On obtient ainsi un processus ponctuel déterminantal associé à un ensemble de polynômes orthogonaux, mais qui n'a pas forcément une application dans la pratique. Voici quelques exemples d'ensembles de polynômes orthogonaux \mathbb{P}_k continus et discrets qui interviennent dans des problèmes de physiques, statistiques ou combinatoires.

Les ensembles polynomiaux continus sont souvent attribués à des modèles de matrices aléatoires. En voici les plus célèbres :

L'ensemble d'Hermite : Il correspond à la mesure de Gauss $d\nu(x_i) = e^{-x_i^2} dx_i$. Cet ensemble est lié à un modèle de physique atomique introduit par Wigner, [Wig55]. $\mathbb{P}_k(dx)$ représente la densité jointe des valeurs propres ordonnées d'une matrice Hermitienne à entrées Gaussiennes. Le paragraphe suivant est entièrement consacré à l'étude de ce modèle.

L'ensemble de Laguerre : Il correspond à la mesure Gamma $d\nu(x_i) = x_i^{N-k} e^{-x_i} dx_i$. Dans ce cas, $\mathbb{P}_k(dx)$ est la loi jointe des valeurs propres d'une matrice de Wishart $W = XX^*$ où X est une matrice à entrées Gaussiennes indépendantes de taille $k \times N$. Ce modèle est souvent utilisé en analyse de données statistiques et sera détaillé dans le Chapitre 5. Cet ensemble décrit également la loi du temps de dernier passage dans le modèle de percolation de dernier passage avec des poids exponentiels, Chapitre 2.

L'ensemble de Jacobi : Il est associé à la loi Beta et décrit la loi du spectre d'une certaine matrice aléatoire $\pi_k \tilde{\pi}_k \pi_k$. π_k est une matrice déterministe qui représente une projection orthogonale sur \mathbb{C}^k de rang q_k et $\tilde{\pi}_k$ est une projection orthogonale sur \mathbb{C}^k de rang \tilde{q}_k uniformément choisie, voir [Col05].

Les ensembles polynomiaux discrets sont souvent liés à des problèmes de probabilités combinatoires :

L'ensemble de Charlier : On considère l'ensemble $W_{N,k}$ des mots de N lettres choisis parmi k . En d'autres termes, $W_{N,k}$ n'est autre que l'ensemble des applications $w : \{1, \dots, N\} \rightarrow \{1, \dots, k\}$. On munit $W_{N,k}$, qui est de cardinal k^N , de la probabilité uniforme. On choisit un mot w au hasard et on regarde la probabilité que la longueur de la plus longue sous-suite croissante, notée $L_{N,k}$, soit inférieure à t , *i.e.* $\mathbb{P}[L_{N,k} \leq t]$. Une sous-suite croissante de w est de la forme $w(i_1), \dots, w(i_k)$ tel que $i_1 < i_2 < \dots < i_k$ et $w(i_1) \leq \dots \leq w(i_k)$. Pour calculer cette probabilité, on identifie chaque mot w à une permutation généralisée à laquelle on attribue via la correspondance RSK, un couple unique (P, Q) de tableaux de Young. P est un tableau semi-standard dont les éléments sont choisis dans $\{1, \dots, k\}$ tandis que Q est un tableau standard dont les éléments sont pris dans $\{1, \dots, N\}$. De plus, P et Q sont de même forme $\lambda \in \{\lambda \in \mathbb{Z}_+^k; \lambda_1 \geq \dots \geq \lambda_k \text{ et } \sum_i \lambda_i = N\}$. Comme la longueur de la première ligne λ_1 de P (ou Q) est égale à $L_{N,k}$, il suffit de connaître f^λ , le nombre de tableaux de Young standards de forme λ fixée et $d_\lambda(k)$ le nombre de tableaux de Young semi-standards de forme λ et dont les éléments sont dans $\{1, \dots, k\}$. On a que

$$f^\lambda = N! \prod_{1 \leq i < j \leq k} (\lambda_i - \lambda_j + j - i) \prod_{i=1}^N \frac{1}{(\lambda_i + N - i)!} \quad \text{et} \quad d_\lambda(k) = \prod_{1 \leq i < j \leq k} \frac{\lambda_i - \lambda_j + j - i}{j - i} \quad (1.2)$$

Pour plus de détails concernant la correspondance RSK et le dénombrement des tableaux de Young, le lecteur peut consulter les deux ouvrages [Ful97, Sag91].

La version Poissonisée $L_k(\alpha)$ de $L_{N,k}$ consiste à supposer que N est aléatoire et suit une loi de Poisson de paramètre α . Par suite, on a que

$$\mathbb{P}[L_k(\alpha) = t] = \sum_{N=0}^{\infty} e^{-\alpha} \frac{\alpha^N}{N!} \mathbb{P}[L_{N,k} = t].$$

En utilisant (1.2), on peut démontrer, [Joh01a], que $L_k + k - 1 = \max_{i=1, \dots, k} \{h_i\}$ où $h = (h_i)_{i=1}^k := (\lambda_i + k - i)_{i=1}^k$ est un processus déterminantal de loi \mathbb{P}_k avec ν la mesure de **Poisson**

de paramètre α/k .

L'ensemble de Charlier ne sera pas évoqué dans la suite de cette thèse et on envoie les lecteurs intéressés vers [Joh01a] pour une étude approfondie du sujet. Notons juste qu'avec les bonnes normalisations, la mesure \mathbb{P}_k qui correspond à l'ensemble de Charlier converge vers la mesure de Plancherel. La mesure de Plancherel agit sur l'ensemble des couples de tableaux de Young standards. Elle permet par conséquent, de traiter la question de la plus longue sous-suite croissante d'une permutation aléatoire uniformément choisie dans le groupe symétrique S_k , [BDJ99, AD99].

L'ensemble de Meixner : C'est un ensemble polynomial discret intervenant en percolation dirigée de dernier passage sur un rectangle de taille $N \times k$ avec des poids géométriques. Les polynômes de Meixner sont obtenus par rapport à la mesure binomiale négative de paramètre $(q, K = N - k + 1)$:

$$\nu(n) = \omega_k^q(n) = \binom{n + K - 1}{n} q^n.$$

Le chapitre suivant est entièrement consacré à la percolation dirigée de dernier passage et notamment au cas avec des poids géométriques.

L'ensemble de Krawtchouk et l'ensemble de Hahn : Ils sont associés respectivement à la mesure binomiale et à la mesure hypergéométrique. Ils interviennent dans les problèmes de pavage aléatoire de certaines surfaces par des dominos ou bien par des losanges. L'étude de ces modèles est détaillée dans ([Joh01a], Sec 5) et [Joh02]. L'ensemble de Krawtchouk intervient également dans le modèle de percolation dirigée de Seppäläinen, voir Paragraphe 2.4.1.

Dans le paragraphe suivant, on regarde de plus près l'exemple particulier où ν est la mesure de Gauss. C'est l'ensemble Gaussien unitaire (l'ensemble polynomial d'Hermite).

1.1 Ensemble Gaussien Unitaire

Bien que le modèle principalement étudié dans cette thèse soit celui de la percolation de dernier passage, on commence par introduire un modèle de matrices aléatoires connu sous le nom de l'Ensemble Gaussien Unitaire (Gaussian Unitary Ensemble). Ce modèle, qui met en avant un processus déterminantal, est très intéressant tant du point de vue mathématiques que du point de vue historique. Tous les calculs effectués, notamment par Wigner, Mehta, Tracy et Widom, sur ce modèle ont été généralisés plus tard pour couvrir d'autres processus déterminantaux.

Le premier système à k particules mettant en évidence une interaction répulsive du type déterminant de Vandermonde au carré (formule (1.1)), est celui du spectre d'une matrice aléatoire appartenant à l'ensemble Gaussien unitaire (GUE). Ce modèle matriciel a été introduit par Wigner [Wig55] dans les années cinquante dans l'étude des niveaux d'énergie des atomes à noyau lourd. Wigner modélisa le comportement des niveaux d'énergie par les valeurs propres d'une matrice aléatoire.

Le GUE est évoqué dans cette première partie de la thèse en tant que modèle de base ou de référence. Le Chapitre 5 sera dédié entièrement aux matrices aléatoires.

1.1.1 Loi jointe et mesure spectrale

L'ensemble Gaussien unitaire n'est autre que l'ensemble des matrices Hermitiennes de taille $k \times k$ munis d'une certaine mesure de probabilité et soumis à des conditions d'invariance ou de symétrie. Plus précisément, soit $\mathbb{H}_k(\mathbb{C})$ des matrices Hermitiennes de taille $k \times k$ muni d'une mesure de probabilité de la forme

$$dP_k(H) = ce^{-f(H)}dH,$$

où c est une constante de normalisation et f une fonction de \mathbb{R}^{k^2} dans \mathbb{R} , [Meh91, Dei99]. On identifie une matrice de $\mathbb{H}_k(\mathbb{C})$ à un vecteur de \mathbb{R}^{k^2} et on note

$$dH = \prod_{i=1}^k dH_{i,i} \prod_{i < j} (dH_{i,j}^R dH_{i,j}^I)$$

la mesure de Lebesgue sur \mathbb{R}^{k^2} . Les $H_{i,j}^R$ et $H_{i,j}^I$ sont respectivement les parties réelles et complexes des entrées de la matrice H et les $H_{i,i}$ sont les éléments diagonaux de la matrice H . Comme toute matrice $H \in \mathbb{H}_k(\mathbb{C})$ représente un opérateur auto-adjoint indépendamment de la base orthonormée utilisée comme repère, on peut supposer que la probabilité d'une matrice est indépendante de la base orthonormée dans laquelle la matrice est écrite. Ainsi, pour toute matrice unitaire $U \in \mathcal{U}_k(\mathbb{R})$, on suppose que

$$e^{-f(H)}dH = e^{-f(UHU^*)}d(UHU^*). \quad (1.3)$$

Or, on peut facilement démontrer [Dei99] que pour toute matrice unitaire U , $dH = d(UHU^*)$. Par suite, (1.3) nous donne l'égalité des densités $e^{-f(H)} = e^{-f(UHU^*)}$ et donc f est une fonction symétrique des valeurs propres de H . Dans le cas particulier où f est la trace d'une fonction polynomiale de degré pair de H , i.e. $f(H) = \text{Tr}(Q(H))$ où Q est un polynôme de degré pair, on parle d'**Ensemble Unitaire**.

Définition 1.1. *Un ensemble unitaire est dit **Gaussien Unitaire** si la fonction Q est un polynôme de la forme*

$$Q(H) = aH^2 + bH + cI_k, \quad a > 0.$$

Comme H est Hermitienne, cette hypothèse supplémentaire signifie que les entrées $(H_{i,i})_{i=1}^k$, $(H_{i,j}^R)_{i < j}$ et $(H_{i,j}^I)_{i < j}$ sont indépendantes. De plus, les éléments diagonaux $(H_{i,i})_{i=1}^k$ suivent une loi Gaussienne $\mathcal{N}(-\frac{b}{2a}, \frac{1}{2a})$ et les éléments sous-diagonaux $(H_{i,j})_{i < j}$ sont de loi Gaussienne complexe avec $(H_{i,j}^R)_{i < j}$ et $(H_{i,j}^I)_{i < j} \sim \mathcal{N}(0, \frac{1}{4a})$. Dans l'ouvrage de Mehta [Meh91], l'auteur considère l'hypothèse d'invariance par transformation unitaire et celle de l'indépendance des entrées pour démontrer que les entrées sont des Gaussiennes. Notons aussi que les entrées d'une matrice de l'Ensemble Unitaire ne sont pas indépendantes (sauf cas Gaussien).

Pour calculer la densité jointe des valeurs propres $\lambda_1 < \dots < \lambda_k$ de H , on démontre d'abord que la probabilité d'avoir une valeur propre multiple est nulle. Ensuite, on calcule le Jacobien de l'application ϕ qui associe à chaque matrice H le couple (D, U) . D est la matrice diagonale formée par les valeurs propres $\lambda = (\lambda_1, \dots, \lambda_k)$ de H et U la matrice unitaire telle que $H = UDU^*$, [Dei99]. Tout calcul fait, on obtient que

$$\mathbb{P}_k(\lambda) = \frac{1}{Z_k} \prod_{1 \leq i < j \leq k} (\lambda_i - \lambda_j)^2 \prod_{i=1}^k e^{-Q(\lambda_i)} d\lambda_1 \dots d\lambda_k,$$

où $\lambda = (\lambda_1, \dots, \lambda_k) \in W_k(\mathbb{R})$. Dans la suite, on considère l'extension par symétrie sur \mathbb{R}^k de la loi jointe \mathbb{P}_k puisque qu'il est toujours plus simple d'intégrer sur \mathbb{R}^k que sur la chambre de Weyl. Pour tout $x \in \mathbb{R}^k$, on considère alors la mesure de probabilité

$$\mathbb{Q}_k(dx) = \frac{1}{k!Z_k} \prod_{1 \leq i < j \leq k} (x_i - x_j)^2 \prod_{i=1}^k e^{-Q(x_i)} dx_1 \dots dx_k := P_k(x)dx. \quad (1.4)$$

Le lecteur pourra regarder en détail le calcul de la loi du spectre de H dans [Meh91] ou ([Dei99], Chap 5). Avant d'expliciter la représentation déterminantale de (1.4), on donne quelques résultats, notamment de grandes déviations, concernant les valeurs propres d'une matrice du GUE. Ces résultats ne dépendent pas de la représentation déterminantale de la distribution jointe. On suppose désormais que $\mathbf{Q}(\mathbf{t}) = \frac{1}{2}\mathbf{t}^2$ pour des raisons de simplification.

Le premier résultat obtenu par Wigner [Wig55, Wig58], est la convergence de la mesure spectrale vers la fameuse loi du demi-cercle. Etant donnée un vecteur $x \in \mathbb{R}^k$, on définit la mesure empirique sur \mathbb{R}

$$\mu_x = \frac{1}{k} \sum_{i=1}^k \delta_{x_i}.$$

Soit H_k une matrice $k \times k$ du GUE et $\hat{\lambda}$ le spectre normalisé de H_k , *i.e.* $\hat{\lambda}_i = \lambda_i/(2\sqrt{k})$ pour tout $i = 1, \dots, k$. $\mu_{\hat{\lambda}}$ est appelée mesure spectrale empirique de H_k et permet de calculer la densité de valeurs propres normalisées dans un Borélien donné. C'est une mesure (variable) aléatoire sur $\mathcal{M}^1(\mathbb{R})$ qui est l'ensemble des mesures de probabilité sur \mathbb{R} . Ici $\mathcal{M}^1(\mathbb{R})$ est muni de la topologie faible engendrée par la distance de Lévy-Prokhorov. On a le théorème suivant :

Théorème 1.2 (Wigner [Wig55]). *Soit H_k une matrice $k \times k$ du GUE et $\hat{\lambda}$ le spectre normalisé de H_k . On a que*

$$\mu_{\hat{\lambda}} \xrightarrow{P} \mu^*,$$

où μ^* est la loi du demi-cercle définie sur \mathbb{R} par

$$d\mu^*(t) = \frac{2}{\pi} \sqrt{1-t^2} \mathbf{1}_{[-1,1]}(t) dt.$$

Si on suppose que les matrices H_k sont des mineurs d'une seule matrice infinie dont les entrées sont définies sur le même espace de probabilité, alors on a la convergence presque sûre dans le théorème précédent. Wigner a démontré ce résultat pour des matrices de Wigner non-Gaussiennes, *i.e.* des matrices Hermitiennes avec entrées indépendantes mais pas forcément de même loi. Plusieurs preuves du théorème de Wigner existent dans la littérature. Parmi ces preuves, on cite la méthode des moments faites par Wigner lui même [Wig55, Wig58]. Voiculescu [Voi91] a étudié la limite de la mesure spectrale pour plusieurs type de matrices aléatoires en utilisant la transformée de Stieltjes et des techniques de la théorie des probabilités libres. Ici, on parlera brièvement de la méthode de la mesure d'équilibre qui nous servira également pour les résultats de grandes

déviations. Pour comprendre l'idée de cette preuve, on exprime, à partir de (1.4), la densité jointe des valeurs propres normalisées sous forme d'une mesure de Gibbs

$$\widehat{P}_k(x) := (2\sqrt{k})^k P_k(2\sqrt{k}x) = \frac{1}{\widehat{Z}_k} e^{-k^2 \mathcal{H}_k(x)}$$

et le Hamiltonien \mathcal{H}_k vaut pour tout $x \in \mathbb{R}^k$,

$$\mathcal{H}_k(x) = \frac{2}{k} \sum_{i=1}^k x_i^2 - \frac{2}{k^2} \sum_{1 \leq i < j \leq k} \log(x_j - x_i).$$

La normalisation des valeurs propres en $1/(2\sqrt{k})$ fait en sorte que les deux termes du Hamiltonien se compensent et par conséquent la mesure d'équilibre est à support compact. $\mathcal{H}_k(x)$ peut être réécrit en utilisant la mesure spectrale μ_x , de la manière suivante :

$$\mathcal{H}_k(x) = E(\mu_x) := 2 \int_{\mathbb{R}} s^2 d\mu_x(s) + \int \int_{s \neq t} \log|t - s|^{-1} d\mu_x(s) d\mu_x(t). \quad (1.5)$$

D'un point de vue physique, $\widehat{P}_k(x) dx$ exprimé de la sorte est une mesure de Gibbs sur l'ensemble des configurations des k particules (valeurs propres normalisées) en interaction. Le Hamiltonien \mathcal{H}_k représente l'énergie d'une configuration donnée, k^2 est l'inverse de la température et \widehat{Z}_k n'est autre que la fonction de partition

$$\widehat{Z}_k = \int_{\mathbb{R}^k} e^{-k^2 \mathcal{H}_k(x)} dx.$$

On peut voir facilement que, lorsque k est grand,

$$\frac{1}{k^2} \log \widehat{Z}_k = - \inf_{x \in \mathbb{R}^k} \mathcal{H}_k(x) + \frac{1}{k^2} \log \int_{\mathbb{R}^k} \exp(-(\mathcal{H}_k(x) - \inf_{x \in \mathbb{R}^k} \mathcal{H}_k(x))) dx \approx - \inf_{x \in \mathbb{R}^k} \mathcal{H}_k(x).$$

Heuristiquement, on peut dire que lorsque $k \rightarrow \infty$, $\frac{1}{k^2} \log \widehat{Z}_k$ converge vers une énergie libre contrôlée par la configuration qui minimise le Hamiltonien. Ainsi, on se ramène à un problème variationnel de la forme

$$E := \inf_{\mu \in \mathcal{M}^1(\mathbb{R})} E(\mu). \quad (1.6)$$

La limite μ^* de la mesure spectrale n'est autre que la mesure de probabilité qui réalise l'infimum du potentiel logarithmique dans (1.6). L'existence et l'unicité de μ^* relèvent de la théorie des potentiels logarithmiques [ST97]. En électrostatique, $\mathcal{H}_k(x)$ est l'énergie potentielle de k charges négatives placées sur la droite réelle suivant la distribution μ_x . Lorsque k est très grand, l'énergie potentielle est minimale à l'équilibre et la densité des charges correspond à la mesure d'équilibre recherchée. La convergence de $\mu_{\widehat{\lambda}}$ vers μ^* et la convergence de la fonction de partition vers l'énergie libre E est détaillée dans la Section 4 de [Joh98].

1.1.2 Grandes déviations

Avant de parler des grandes déviations du GUE, faisons quelques rappels sur le principe de grandes déviations. Soit $(\mathcal{X}, \mathcal{B})$ un espace topologique muni de la tribu Borélienne et soit $(\mu_\varepsilon)_{\varepsilon>0}$ une famille de mesure de probabilité sur \mathcal{X} .

Définition 1.3. *On appelle fonction de taux toute fonction $I : \mathcal{X} \rightarrow [0, \infty]$ qui est semi-continue inférieurement. On dit que I est une bonne fonction de taux si les ensembles $\{x; I(x) \leq \alpha\}$ sont compacts pour tout $\alpha \in [0, \infty)$.*

Définition 1.4. *On dit que (μ_ε) satisfait un principe de grandes déviations de vitesse $1/\varepsilon$ et de fonction de taux I si pour tout $\Gamma \in \mathcal{B}$,*

$$-\inf_{x \in \Gamma^0} I(x) \leq \liminf_{\varepsilon \rightarrow 0} \varepsilon \log \mu_\varepsilon(\Gamma) \leq \limsup_{\varepsilon \rightarrow 0} \varepsilon \log \mu_\varepsilon(\Gamma) \leq -\inf_{x \in \bar{\Gamma}} I(x)$$

Le lecteur peut consulter le livre de Dembo et Zeitouni [DZ98], notamment le Chapitre 1, pour une discussion détaillée sur le principe de grandes déviations.

Revenons à notre problème de matrices aléatoires. **Afin d'alléger les notations, on remplace $\hat{\lambda}$ par λ qui désignera dans cette partie le spectre normalisé.** L'étude de la convergence de la mesure spectrale μ_λ vers la loi du demi-cercle permet d'obtenir des formules asymptotiques de grandes déviations pour la plus grande valeur propre normalisée λ_k . Benarous et Guionnet [BAG97] ont prouvé que μ_λ converge faiblement vers μ^* avec une vitesse exponentielle d'ordre k^2 au sens des grandes déviations. Plus précisément, ils ont démontré que la loi de μ_λ satisfait un principe de grandes déviations sur $\mathcal{M}^1(\mathbb{R})$ de bonne fonction de taux

$$I(\mu) = E(\mu) - E.$$

D'après la théorie des potentiels logarithmiques [ST97], $E = 3/4 + \log 2$ est l'énergie à l'équilibre *i.e.* l'énergie qui correspond à une infinité de charges réparties suivant la distribution du demi-cercle. Rappelons que $\mathcal{M}^1(\mathbb{R})$ est muni de la topologie de la convergence faible engendrée par la distance de Lévy-Prokhorov. On a donc le théorème suivant [BAG97] :

Théorème 1.5 (Ben Arous-Guionnet [BAG97]). *La loi de la mesure spectrale μ_λ satisfait un principe de grandes déviations de vitesse k^2 et de fonction de taux I , *i.e.**

pour tout fermé F de $\mathcal{M}^1(\mathbb{R})$,

$$\limsup_{k \rightarrow \infty} \frac{1}{k^2} \log \mathbb{P}[\mu_\lambda \in F] \leq -\inf_{\mu \in F} I(\mu),$$

pour tout ouvert O de $\mathcal{M}^1(\mathbb{R})$,

$$\liminf_{k \rightarrow \infty} \frac{1}{k^2} \log \mathbb{P}[\mu_\lambda \in O] \geq -\inf_{\mu \in O} I(\mu).$$

La fonction I est bien définie et est convexe sur $\mathcal{M}^1(\mathbb{R})$. Le lecteur pourra regarder [ST97] ou [BAG97] pour mieux comprendre le comportement de la fonction I .

On s'intéresse maintenant aux propriétés de grandes déviations de la plus grande valeur propre normalisée λ_k . Ces propriétés sont des conséquences plus ou moins immédiates du Théorème 1.5 sur la mesure spectrale. En fait, on a pour tout $s \in \mathbb{R}$,

$$\mathbb{P}[\{\lambda_k \leq s\}] = \mathbb{P}[\{\max_i x_i \leq s\}] = \mathbb{P}[\mu_x \in \mathcal{M}^1((-\infty, s])].$$

où $\mathcal{M}^1((-\infty, s])$ est l'ensemble de mesure de probabilité à support dans $(-\infty, s]$. Comme $\mathcal{M}^1((-\infty, s])$ est un fermé de $\mathcal{M}^1(\mathbb{R})$, on a, d'après le Théorème 1.5, que

$$\limsup_{k \rightarrow \infty} \frac{1}{k^2} \log \mathbb{P}[\lambda_{\max} \leq (1 - \varepsilon)] \leq - \inf_{\mu \in \mathcal{M}^1((-\infty, 1 - \varepsilon])} I(\mu).$$

L'inégalité dans l'autre sens est moins évidente et nécessite plus de calcul, (voir [Fér08]). On a le théorème suivant :

Théorème 1.6 (Ben Arous-Dembo-Guionnet [BADG01], Féral [Fér08]). *Soit H une matrice du GUE de taille $k \times k$ tel que $Q(t) = (1/2)t^2$. On désigne par $\lambda_{\max} = \lambda_k$ la plus grande valeur propre normalisée de H . Alors on a*

$$\lambda_{\max} \xrightarrow[k \rightarrow \infty]{p.s.} 1$$

De plus, λ_{\max} satisfait des principes de grandes déviations à droite et à gauche de la moyenne. Pour tout $\varepsilon > 0$, on a

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathbb{P}[\lambda_{\max} \geq (1 + \varepsilon)] = -J_{\text{GUE}}(\varepsilon)$$

et

$$\lim_{k \rightarrow \infty} \frac{1}{k^2} \log \mathbb{P}[\lambda_{\max} \leq (1 - \varepsilon)] = -I_{\text{GUE}}(\varepsilon),$$

où

$$\begin{aligned} J_{\text{GUE}}(\varepsilon) &= 4 \int_0^\varepsilon \sqrt{s(s+2)} \, ds \\ &= 2(1 + \varepsilon) \sqrt{(1 + \varepsilon)^2 - 1} - 2 \log(1 + \varepsilon + \sqrt{(1 + \varepsilon)^2 - 1}). \end{aligned}$$

et

$$I_{\text{GUE}}(\varepsilon) = \inf_{\mu \in \mathcal{M}^1((-\infty, 1 - \varepsilon])} I(\mu)$$

La formule des grandes déviations à gauche est une conséquence du Théorème 1.5, (voir [Fér08]). La preuve de la formule des grandes déviations à droite de la moyenne est établie dans ([BADG01], Section 6). Les grandes déviations à droite de la moyenne sont en e^{-k} tandis que celles à gauche de la moyenne sont en e^{-k^2} . Pour comprendre intuitivement ce fait, on remarque que λ_{\max} est grande lorsqu'il existe une direction $x^* \in \mathbb{R}^k$ telle que $(\langle x^*, Hx^* \rangle / \|x^*\|_2)$ soit grand. Cependant, pour que λ_{\max} soit petite, il faut que $(\langle x, Hx \rangle / \|x\|_2)$ soit petit pour toute les direction $x \in \mathbb{R}^k$. Cela affecte donc toutes les variables de la matrice H .

Remarque 1.7. Notons ici que les Théorèmes 1.5 et 1.6 restent valables en remplaçant $Q(t) = (1/2)t^2$ par un potentiel quelconque $V_k(t)$ vérifiant des conditions de régularité et de convergence uniforme. Plus précisément, on suppose que X admet une densité d'un gaz de Coulomb *i.e.*

$$P_k^{\beta, V}(x) = \frac{1}{Z_k} \prod_{1 \leq i < j \leq k} |x_i - x_j|^\beta \prod_{i=1}^k \exp\left(-\frac{\beta k}{2} V_k(x_i)\right) \quad (1.7)$$

où $\beta > 0$ et $V_k(t)$ est un potentiel tel que

- (i) Pour tout $k \geq 1$, $V_k : \mathbb{R} \rightarrow \mathbb{R}$ est continue,
- (ii) il existe deux constantes $\xi > 0$ et $T > 0$ telles que pour tout $t \geq T$,

$$V_k(t) \geq (1 + \xi) \log(1 + t^2),$$

- (iii) Il existe un potentiel V tel que V_k converge uniformément vers V sur tout compact de \mathbb{R} .

Alors μ_x et $\max_i \{x_i\}$ satisfont des principes de grandes déviations similaires à ceux des Théorèmes 1.5 et 1.6. Les bonnes fonctions de taux correspondantes sont contrôlées par les mesures d'équilibre respectivement sur $\mathcal{M}^1(\mathbb{R})$ et $\mathcal{M}^1((-\infty, 1 - \varepsilon])$. Notons que les potentiels logarithmiques en question dépendent de V_k et de β et par conséquence les mesures d'équilibre ne sont pas forcément les mêmes que ceux du GUE, [ST97]. L'étude dans le cadre d'un potentiel V_k discret est faite dans [Joh00a] et dans [HP00b] pour les potentiels continus. Le lecteur pourra regarder également [Fér08] pour une étude récente du sujet.

1.1.3 Représentation déterminantale et fluctuations

Dans cette partie, on établit une représentation déterminantale de la mesure de probabilité $\mathbb{Q}_k(dx)$. Cette représentation est indispensable dans l'étude des fluctuations d'un ensemble de polynômes orthogonaux. Bien que l'étude ci-dessous soit valable pour une multitude de mesure de probabilité de la forme $\mathbb{Q}_k(dx)$ avec ν admettant des moments finis de tout ordre, on ne fera les calculs que dans le cadre du GUE. On suppose alors que

$$d\nu(t) = \frac{1}{2} t^2 dt.$$

Il est plus convenable de travailler sur la mesure $\mathbb{Q}_k(dx)$ définie sur \mathbb{R} que sur la mesure $\mathbb{P}_k(dx)$ définie sur la chambre de Weyl. On rappelle que P_k est la densité de la mesure \mathbb{Q}_k . Notons ici qu'on travaille sur le spectre **non normalisé**. On définit pour tout $x = (x_1, \dots, x_k) \in \mathbb{R}^k$ la fonction de corrélation de la mesure $\mathbb{Q}_k(dx)$ par

$$R_l^{(k)}(x_1, \dots, x_l) = \frac{k!}{(k-l)!} \int_{\mathbb{R}^{k-l}} P_k(x) dx_{l+1} \cdots dx_k. \quad (1.8)$$

La fonction $(k-l)!R_l^{(k)}/k!$ est une densité de probabilité sur \mathbb{R}^l qui n'est autre que la loi marginale des l premières particules. La multiplication par la constante $k!/(k-l)!$ est justifiée dans les calculs qui suivent. On définit également la suite $(\pi_i)_{i \geq 0}$ des polynômes unitaires orthogonaux dans $L^2(\mathbb{R}, \nu)$. On désigne par $(c_i)_{i \geq 0}$ la suite des réels telle que la famille (π_i/c_i) soit orthonormée. Notons que π_i est de degré i .

Théorème 1.8 (Mehta [Meh91]).

$$R_l^{(k)}(x_1, \dots, x_l) = \det (K_k(x_i, x_j))_{1 \leq i, j \leq l}, \quad (1.9)$$

où

$$K_k(s, t) = \sum_{i=0}^{k-1} \phi_i(s) \phi_i(t) \quad \text{avec} \quad \phi_i(t) = c_i^{-1} e^{-\frac{t^2}{4}} \pi_i(t).$$

Démonstration : En effectuant les bonnes opérations sur les lignes successives de la matrice de Vandermonde, on obtient facilement que

$$\prod_{i < j} (x_j - x_i) = \det \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_k \\ \vdots & \vdots & \vdots \\ x_1^{k-1} & \cdots & x_k^{k-1} \end{pmatrix} = \det \begin{pmatrix} \pi_0(x_1) & \cdots & \pi_0(x_k) \\ \vdots & \vdots & \vdots \\ \pi_{k-1}(x_1) & \cdots & \pi_{k-1}(x_k) \end{pmatrix}$$

Par conséquent,

$$\begin{aligned} \left(\prod_{i=0}^{k-1} \frac{1}{c_i^2} \right) \prod_{i=1}^k e^{-\frac{x_i^2}{2}} \prod_{i < j} (x_j - x_i)^2 &= \left(\det (\phi_{i-1}(x_j))_{1 \leq i, j \leq k} \right)^2 \\ &= \det \left((\phi_{i-1}(x_j))_{1 \leq i, j \leq k} \times (\phi_{j-1}(x_i))_{1 \leq i, j \leq k} \right) \\ &= \det \left(\sum_{r=1}^k \phi_{r-1}(x_i) \phi_{r-1}(x_j) \right)_{1 \leq i, j \leq k} \\ &= \det (K_k(x_i, x_j))_{1 \leq i, j \leq k}. \end{aligned} \quad (1.10)$$

Le lemme suivant, [Meh91], décrit la relation entre la fonction de corrélation de l particules et le déterminant du mineur d'ordre l de la matrice $((K_k(x_i, x_j))_{1 \leq i, j \leq k})$. La preuve de ce lemme est également donnée dans ([Dei99], Lemma 5.27).

Lemme 1.9. *Soit la matrice $J_k(x) = (f(x_i, x_j))_{1 \leq i, j \leq k}$ avec $x \in \mathbb{R}^k$ et f est une fonction mesurable de $\mathbb{R}^2 \rightarrow \mathbb{C}$. On suppose qu'il existe une mesure μ sur \mathbb{R} telle que*

$$\int_{\mathbb{R}} f(s, t) f(t, w) d\mu(t) = f(s, w).$$

Alors

$$\int_{\mathbb{R}} \det J_k(x) d\mu(x_k) = (d - k + 1) \det J_{k-1},$$

où $J_{k-1} = (J_{ij})_{1 \leq i, j \leq k-1}$ est le mineur d'ordre $k-1$ de $J_k(x)$ et $d = \int_{\mathbb{R}} f(t, t) d\mu(t)$.

Il est facile de voir que $(K_k(x_i, x_j))_{1 \leq i, j \leq k}$ vérifie les hypothèses du lemme précédent et que $d = k$. Par suite,

$$\int_{\mathbb{R}^k} \det (K_k(x_i, x_j))_{1 \leq i, j \leq k} d^k x = k!$$

et donc

$$Z_k = \prod_{i=0}^{k-1} c_i^2.$$

Par conséquent, (1.10) nous donne que

$$P_k(x) = \frac{1}{k!} \det (K_k(x_i, x_j))_{1 \leq i, j \leq k}.$$

En appliquant encore une fois le lemme d'intégration précédent, on obtient que

$$R_l^{(k)} = \frac{k!}{(k-l)! k!} \int_{\mathbb{R}^{k-l}} \det (K_k(x_i, x_j))_{1 \leq i, j \leq k} dx_{l+1} \cdots dx_k = \det (K_k(x_i, x_j))_{1 \leq i, j \leq l}.$$

□

La représentation déterminantale donnée dans le Théorème 1.8 a notamment été utilisée dans le calcul asymptotique des fluctuations de la plus grande valeur propre d'une matrice du GUE. Sans rentrer dans les détails techniques de calcul, on décrit les étapes importantes qui ont permis à Tracy et Widom d'établir la convergence en loi des fluctuations de λ_k vers la loi dite de Tracy-Widom. Pour commencer, on remarque que pour tout $u \in \mathbb{R}$,

$$\mathbb{P}[\lambda_k \leq u] = \mathbb{P}[\max_i x_i \leq u] = 1 - \mathbb{P}\left[\bigcup_i \{x_i > u\}\right].$$

En combinant la formule du crible de Poincaré et le Théorème 1.8 et en utilisant le fait que $R_l^{(k)}$ est invariante par permutation des coordonnées, on démontre que

$$\mathbb{P}[\lambda_k \leq u] = \sum_{r=0}^k \frac{(-1)^r}{r!} \int_u^\infty \cdots \int_u^\infty \det (K_k(x_i, x_j))_{1 \leq i, j \leq r} dx_1 \cdots dx_r. \quad (1.11)$$

D'après la théorie des opérateurs de Fredholm, l'expression (1.11) n'est autre que le déterminant d'un opérateur de Fredholm \mathcal{K}_k de noyau K_k agissant sur $L^2([u, \infty), du)$, ([RS78], Vol IV). On peut alors écrire

$$\mathbb{P}[\lambda_k \leq u] = \det (I - \mathcal{K}_k) \Big|_{L^2([u, \infty), du)}.$$

En choisissant la bonne normalisation de u , Tracy et Widom [TW94] ont démontré que le noyau K_k de l'opérateur \mathcal{K}_k converge uniformément sur tout compact vers le noyau d'Airy K_{Ai} . Plus précisément, ils ont démontré que

$$\frac{1}{k^{1/6}} K_k \left(2\sqrt{k} + \frac{s}{k^{1/6}}, 2\sqrt{k} + \frac{t}{k^{1/6}} \right) \xrightarrow{k \rightarrow \infty} K_{\text{Ai}}(s, t), \quad (1.12)$$

où

$$K_{\text{Ai}}(s, t) = \int_0^\infty \text{Ai}(s+u) \text{Ai}(t+u) du.$$

La convergence est uniforme en s et t sur tout compact de \mathbb{R}^2 . La fonction Ai est l'unique solution de l'équation différentielle $f''(t) = tf(t)$ équivalente en $+\infty$ à $(4\pi\sqrt{t})^{1/2}e^{-\frac{2}{3}t^{3/2}}$. En tenant compte de (1.12), on choisit

$$u = 2\sqrt{k} + \frac{v}{k^{1/6}}$$

et on obtient

$$\begin{aligned} \lim_{k \rightarrow \infty} \mathbb{P}\left[\lambda_k \leq 2\sqrt{k} + \frac{v}{k^{1/6}}\right] &= \det(I - \mathcal{K}_{\text{Ai}})|_{L^2([v, \infty), dv)} & (1.13) \\ &= \sum_{r=0}^{\infty} \frac{(-1)^r}{r!} \int_v^{\infty} \cdots \int_v^{\infty} \det(K_{\text{Ai}}(x_i, x_j))_{1 \leq i, j \leq r} dx_1 \cdots dx_r. \end{aligned}$$

Ici \mathcal{K}_{Ai} est l'opérateur d'Airy agissant sur $L^2([v, \infty), dv)$ et admettant comme noyau K_{Ai} . Le théorème suivant, démontré par Tracy et Widom [TW94], résume les calculs faits là-dessus. Il fournit une forme quasi explicite du déterminant de l'opérateur d'Airy et par conséquent identifie la fonction de répartition de la fameuse loi limite des fluctuations de λ_k .

Théorème 1.10 (Tracy-Widom [TW94]). *Soit H une matrice $k \times k$ du GUE telle que $Q(t) = \frac{1}{2}t^2$. On note λ_{\max} la valeur propre maximale de H . Pour tout $v \in \mathbb{R}$, on a que*

$$\lim_{k \rightarrow \infty} \mathbb{P}\left[k^{1/6}(\lambda_{\max} - 2\sqrt{k}) \leq v\right] = \det(I - \mathcal{K}_{\text{Ai}})|_{L^2([v, \infty), dv)}.$$

De plus, on a que

$$\det(I - \mathcal{K}_{\text{Ai}})|_{L^2([v, \infty), dv)} = F_2^{\text{TW}}(v) := \exp\left\{-\int_v^{\infty} (t-v)q^2(t)dt\right\},$$

où q est l'unique solution, satisfaisant $q(t) > 0$ pour tout $t > 0$, de l'équation différentielle de Painlevé II

$$q''(t) = tq(t) + 2q(t)^3.$$

Ce résultat peut se généraliser à d'autres systèmes à k -particules admettant une représentation déterminantale à l'image du Théorème 1.8. Johansson [Joh00a] a démontré un résultat similaire pour l'ensemble polynomial discret de Meixner. La clé de la preuve réside dans l'expression (1.12) qui nous fournit l'asymptotique du noyau polynomial (ici d'Hermite) comme étant le noyau d'Airy. Le calcul des asymptotiques des noyaux polynomiaux orthogonaux peut passer par la résolution de ce qu'on appelle un problème de Riemann-Hilbert, (voir [Dei99]).

Remarque 1.11. Le Théorème 1.10 reste valable pour la plus grande valeur propre d'une matrice de l'Ensemble Unitaire, (voir [DG07]). Dans ce cas, la plus grande valeur propre normalisée converge vers la borne supérieure du support de la mesure spectrale limite. Cette mesure est la solution du problème de minimisation du potentiel logarithmique (1.5) en remplaçant t^2 par $Q(t)$.

1.1.4 GUE et percolation dirigée

Ce paragraphe est une transition entre le Chapitre 1 et le Chapitre 2. Il constitue un lien entre le modèle du GUE et le modèle de percolation dirigée du chapitre suivant.

Le modèle du GUE possède une importance particulière dans cette thèse. Mise à part le fait qu'il soit le modèle de référence en ce qui concerne la convergence vers la loi de Tracy-Widom, ce modèle est étroitement lié à la percolation dirigée dans le plan. Plusieurs résultats obtenus durant cette thèse, notamment dans le Chapitre 3, reposent essentiellement sur des propriétés de la plus grande valeur propre d'une matrice du GUE. Afin de préparer le lecteur aux relations entre GUE et percolation dirigée dans le plan, on énonce le théorème suivant démontré par Baryshnikov dans [Bar01].

Théorème 1.12 (Barishnikov [Bar01]). *Soit $B_t^{(1)}, \dots, B_t^{(k)}$ des mouvements Browniens indépendants. Si on note λ_{\max} la valeur propre maximale d'une matrice du GUE de taille $k \times k$, alors on a que*

$$\lambda_{\max} \stackrel{\mathcal{L}}{=} L(1, k),$$

où

$$L(N, k) := \sup_{0=u_0 < u_1 < \dots < u_{k-1} < u_k=N} \sum_{r=1}^k (B_{u_r}^{(r)} - B_{u_{r-1}}^{(r)}). \quad (1.14)$$

$L(N, k)$ est le temps de dernier passage d'une percolation dirigée Brownienne dans le plan. Il intervient également [GW91] en tant que limite en loi du temps de dernier passage du modèle de percolation planaire principalement étudié dans cette thèse. Le théorème précédent est à l'origine de plusieurs résultats obtenus dans les deux chapitres qui suivent.

Chapitre 2

Percolation dirigée de dernier passage

Dans ce chapitre, on commence par la définition de modèle de percolation de dernier passage dans le plan. Avant de rappeler les principaux résultats, on expose les différents domaines dans lesquels apparaît ce modèle. A la fin du chapitre, on parle de deux modèles de percolation planaire qui diffèrent de notre modèle par la règle d'évolution des chemins ou la géométrie de la boîte dans laquelle on travaille. Dans la suite, l'abréviation **LPP** (Last-Passage Percolation) sera souvent utilisée pour désigner le modèle de percolation de dernier passage.

2.1 Définition du modèle

On se place dans le plan discrétisé et plus précisément dans le premier quadrant $\mathbb{Z}_+^* \times \mathbb{Z}_+^*$. On attribue à chaque point (i, j) de cet espace un poids aléatoire noté $X_i^{(j)}$ et on considère l'ensemble $\Pi(N, k)$ des chemins qui partent du point $(1, 1)$ pour atteindre le point (N, k) en ne permettant que les pas vers la droite et vers le haut. Un chemin π appartenant à $\Pi(N, k)$ est donc un ensemble de points

$$\pi = \left\{ (i_1, j_1) = (1, 1), \dots, (i_{N+k-1}, j_{N+k-1}) = (N, k) \right\},$$

tel que $(i_{l+1}, j_{l+1}) - (i_l, j_l)$ vaut $(1, 0)$ ou $(0, 1)$ pour tout l entre 1 et $N + k - 2$. Le terme percolation dirigée vient du fait qu'on oblige les chemins d'emprunter des directions sans d'autres, (voir Figure 3.2). Dans ce cadre là, on définit le temps de dernier passage par la variable aléatoire réelle

$$G(N, k) = \max_{\pi \in \Pi(N, k)} \left\{ \sum_{(i, j) \in \pi} X_i^{(j)} \right\}. \quad (2.1)$$

Glynn et Whitt [GW91] utilisent cette version du modèle dans leurs travaux sur les files d'attente en série. En tenant compte du Théorème 1.12, il convient de bénéficier d'une autre interprétation de ces chemins dirigés, et par suite d'une autre façon d'exprimer le temps du dernier de passage. Plus précisément, un chemin $\pi \in \Pi(N, k)$ est caractérisé par un $(k - 1)$ -uplet de $(\mathbb{Z}_+^*)^{k-1}$ qui représente les abscisses des $k - 1$ sauts vers le haut de ce chemin. On a donc une bijection entre $\Pi(N, k)$ et l'ensemble $U(N, k) \cap \mathbb{Z}_+^{k+1}$ où

$$U(N, k) := \{u = (u_0, u_1, \dots, u_k) \in \mathbb{R}_+^{k+1}; 0 = u_0 \leq u_1 \leq \dots \leq u_k = N\}. \quad (2.2)$$

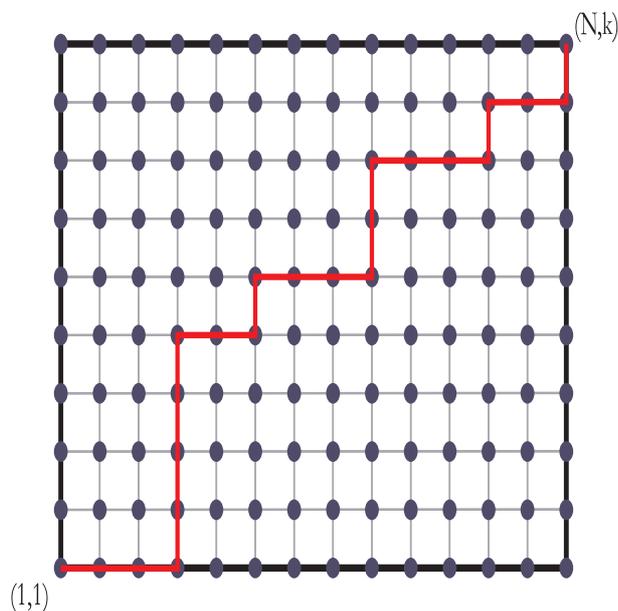


FIG. 2.1 – Percolation dirigée dans le plan

On réécrit alors le temps de dernier passage de la manière suivante :

$$G(N, k) = \sup_{u \in U(N, k)} \left\{ \sum_{r=1}^k [S_{[u_r]}^{(r)} - S_{[u_{r-1}] - 1}^{(r)}] \right\}, \quad (2.3)$$

où $S_m^{(r)} = \sum_{i=1}^m X_i^{(r)}$, avec la convention $S_{-1}^{(r)} = S_0^{(r)} = 0$.

D'une manière générale, on peut définir le maximum sur les chemins entre deux points quelconques. On introduit ainsi la variable

$$G((N_1, k_1), (N_2, k_2)) = \max_{\pi \in \Pi((N_1, k_1), (N_2, k_2))} \left\{ \sum_{(i, j) \in \pi} X_i^{(j)} \right\}$$

où $\Pi((N_1, k_1), (N_2, k_2))$ est l'ensemble des chemins dirigés qui partent du point (N_1, k_1) jusqu'au point (N_2, k_2) . On a donc que $G(N, k) = G((1, 1), (N, k))$. Sous cette forme là, on peut vérifier facilement que G est superadditive. On a en effet que

$$G((N_1, k_1), (N_2, k_2)) \geq G((N_1, k_1), (N_0, k_0)) + G((N_0, k_0), (N_2, k_2)), \quad (2.4)$$

pour tout $N_1 \leq N_0 \leq N_2$ et $k_1 \leq k_0 \leq k_2$. Cette propriété a des conséquences importantes sur le comportement limite de la variable G . On constate également que $G(N, k)$ satisfait la formule de récurrence

$$G(N, k) = G((N-1, k) \vee G(N, k-1) + X_N^{(k)}, \quad (2.5)$$

avec les conditions aux bords $G(N, 0) = G(0, k) = 0$ pour tous N et k entiers positifs. Cette caractérisation de $G(N, k)$ permet de faire le lien avec les files d'attente en série et le processus

d'exclusion simple. En faisant la marche inversée, on retrouve la forme initiale (2.1). L'intérêt de réécrire $G(N, k)$ sous la forme (2.5) est de pouvoir se passer des chemins orientés. Ainsi, tout modèle, dont la variable principale est indexée par $\mathbb{Z}_+^* \times \mathbb{Z}_+^*$ et vérifiant (2.5), est équivalent en quelque sorte à notre modèle de percolation. Dans le paragraphe suivant, on expose trois situations dans lesquelles apparaît la variable G .

2.2 Percolation dirigée et applications

Comme on l'a mentionné auparavant, le modèle de percolation dirigée en dimension 2 suscite un intérêt particulier du fait de son lien avec la physique des particules et la théorie des files d'attente en série. On renvoie le lecteur vers le résumé de Seppäläinen [Sep09] et celui de König [Kön05] pour plus de détails et de références bibliographiques.

Files d'attente en série : la variable G est apparue pour la première fois sous la forme (2.5) dans [TW74] dans l'étude des files d'attente en série et plus tard dans [GW91]. On considère la situation suivante : N clients font la queue avant de passer devant k guichets numérotés de 1 à k . Au temps $t = 0$, tout le monde attend devant le premier guichet et le premier client passe au guichet 1. Une fois servi, ce même client passe au guichet 2 et laisse sa place au deuxième client au guichet 1 et ainsi de suite. Dès que le client i quitte le guichet j , il se dirige vers le guichet $j + 1$. Si ce guichet est vacant, il l'occupe, sinon il fait la queue derrière le client $i - 1$. On note $Y_i^{(j)}$ le temps de service du client i au guichet j . Après un certain temps, la scène est constituée de k guichets avec une file d'attente plus ou moins grande devant chacun d'eux. On suppose que les clients sont tous là depuis le début. Cependant, considérer des arrivées aléatoires au guichet 1 revient à rajouter un guichet avec un temps de service de même loi que le temps d'arrivée des clients. En notant $T(n, k)$ le temps passé avant que le client n ne quitte le guichet k , on remarque facilement que

$$T(n, k) = T(n - 1, k) \vee T(n, k - 1) + Y_n^{(k)},$$

avec les conditions aux bords $T(n, 0) = T(0, k) = 0$. La variable T vérifie alors la même formule de récurrence que G . Par suite, lorsque les $(X_i^{(j)})_{i,j=1}^\infty$ et les $(Y_i^{(j)})_{i,j=1}^\infty$ sont *i.i.d.* et $X_1^{(1)} \stackrel{\mathcal{L}}{=} Y_1^{(1)}$, alors $T(N, k)$ et $G(N, k)$ sont de même loi. On comprend ainsi l'appellation : temps de dernier passage. Glynn et Whitt [GW91] se servent de la propriété de superadditivité pour démontrer que

$$\lim_{N \rightarrow \infty} \frac{T(\lfloor \gamma N \rfloor, N)}{N} = g(\gamma),$$

où $g(\gamma)$ est une fonction déterministe. Comme la variable $Y_i^{(j)}$ représente un temps de service, il convient de choisir une loi exponentielle pour un temps continu et une loi géométrique pour un temps discret. Ce sont d'ailleurs les deux cas où l'on connaît le plus de résultats.

Dans un cadre plus général, on peut supposer que chaque guichet fournit un service différent des autres. Ainsi, le temps de service d'un client i va dépendre du guichet où il se trouve. Par suite, les $Y_i^{(j)}$ sont toujours considérées indépendantes mais de loi exponentielle de paramètre λ_j qui varie avec les guichets. Là aussi, on connaît plein de résultats sur la variable $T(N, k)$, (voir [Joh08]).

Si les $X_i^{(j)}$ suivent des lois à densité, on sait que le chemin maximisant la somme dans (2.1), celui donc qui correspond au temps de dernier passage, est unique avec probabilité 1. Pour construire ce chemin tout en connaissant la configuration des files d'attente en série en tout instant, on commence par la fin. On trace une ligne verticale qui lie les points (N, k) et le point (N, j_1) où j_1 correspond au numéro du guichet à partir duquel le client N ne croise plus aucun des clients se situant devant lui dans la file de départ. Ensuite, on fait un pas à gauche vers le point $(N-1, j_1)$ et on trace une ligne verticale entre le point $(N-1, j_1)$ et le point $(N-1, j_2)$. j_2 est le premier guichet tel que le client $N-1$ ne croise plus aucun des clients se situant devant lui entre j_2 et j_1-1 . On réitère cette construction N fois pour obtenir le chemin maximisant.

The corner-growth model : C'est la version processus du modèle de percolation de dernier passage. Il s'agit d'un amas (cluster) de points dans le plan qui croît avec le temps. On se place encore une fois dans le premier cadran du plan discrétisé et on suppose qu'au temps $t = 0$ le point $(1, 1)$ constitue l'amas. Le processus évolue suivant la règle qui suit : pour qu'un point (i, j) rejoigne l'amas, il faut attendre que les points $(i-1, j)$ et $(i, j-1)$, situés respectivement à sa gauche et en dessous, aient rejoint l'amas à leur tour. Une fois ceci réalisé, le point (i, j) pourra rejoindre l'amas après un certain temps aléatoire $X_i^{(j)}$. Si le point est de la forme $(i, 1)$ (resp. $(1, j)$) (un point du bord), il suffit que le point situé à sa gauche (resp. en dessous) ait rejoint l'amas pour lancer le compteur aléatoire. Encore une fois, on remarque, en utilisant la propriété (2.5), que $G(n, k)$ n'est autre que le temps passé avant que le point (n, k) ne rejoigne l'amas. Le processus, noté $A(t)$, décrit alors le nombre de point qui forme l'amas au temps t . On a alors que

$$A(t) = \{(i, j) \in \mathbb{Z}_+^* \times \mathbb{Z}_+^*; G(i, j) \leq t\}.$$

Dans le cas où les $X_i^{(j)}$ sont *i.i.d.* de loi exponentielle, le processus $A(t)$ est Markovien. Dans [Ros81], Rost fait le lien entre le « corner-growth model » et le processus d'exclusion simple totalement asymétrique.

Remarque 2.1. Le « corner-growth model » est une version orientée du processus d'Eden-Richardson où l'amas de point évolue dans \mathbb{Z}^2 en entier. Le compteur d'un point (i, j) débute lorsque l'un de ses voisins rejoint l'amas. L'évolution se fait alors dans tous les sens en laissant éventuellement des trous ce qui complique considérablement l'étude théorique, (voir [Kön05]).

Le processus d'exclusion simple totalement asymétrique : C'est un modèle simplifié qui relève de la physique des particules et plus précisément, des systèmes d'interaction des particules. On considère la droite discrétisée \mathbb{Z} dont chaque point représente un site qui peut héberger au plus une particule. Dans le cas général, le processus démarre avec une répartition aléatoire des particules. Ici on suppose qu'au temps $t = 0$, tous les sites à gauche de l'origine sont occupés et ceux à sa droite, y compris 0, sont vides. Les particules sont numérotées de la façon suivante : la particule i se trouve au départ sur le site $-i$. Au temps $t = 1$, chaque particule décide avec probabilité p et indépendamment des autres particules de faire un pas vers la droite. Si le site est vacant, elle l'occupe, sinon elle reste sur place. C'est comme si on attribuait à chaque particule une suite de variables de Bernoulli indépendantes et de même paramètre qui lui indique les moments où elle peut faire les pas à droite. Les suites de variables de Bernoulli sont indépendantes entre elles et de même paramètre.

Si on considère le temps que met la particule i pour faire j pas vers la droite, on se rend vite

compte que ce temps aléatoire vérifie (2.5) du fait qu'une loi géométrique est sans mémoire. En fait, si on note $Z_i^{(j)}$ le temps qui sépare le moment où la particule $i - 1$ a fait son j ème pas et le moment où la particule i a effectué le sien, $Z_i^{(j)}$ suit une loi géométrique de même paramètre que celui des variables de Bernoulli. Ceci découle de la perte de mémoire de la loi géométrique. Lorsque les poids du modèle de percolation dirigée sont *i.i.d.* de loi géométrique, celui-ci est alors équivalent au processus d'exclusion simple totalement asymétrique.

Une version à temps continue du processus d'exclusion existe également. Cela consiste à remplacer les suites de variables de Bernoulli par des processus de Poisson indépendants et de même paramètre. Dans ce cas, le processus d'exclusion est équivalent au modèle de percolation muni de poids exponentiels.

Polymères dirigés dans un environnement aléatoire : Le modèle de polymères dirigés est issu de la physique des particules. Il sert à modéliser les interfaces dans le modèle d'Ising soumis à des impuretés. On assimile les points de la droite discrète \mathbb{Z} à un ensemble de particules ferromagnétiques. En sachant qu'une particule ferromagnétique prend deux polarisations, on considère $G := \{-1, 1\}^{\mathbb{Z}}$ l'ensemble des configurations possibles des particules. A une température donnée $1/\beta$, on considère la mesure de probabilité (mesure de Gibbs) sur l'ensemble des configurations G définie par sa densité

$$f(s) = \frac{1}{Z_\beta} \exp(\beta H(s)),$$

où $H(s)$ est l'énergie de la configuration s et Z_β est la fonction de partition :

$$Z_\beta = \int_G \exp(\beta H(s)) ds.$$

Lorsque $H(s)$ est aléatoire, on parle de polymères dans un environnement aléatoire. Ici, on ne fait pas une étude approfondie des polymères dirigés ni de leurs applications en physique. Dans la suite, on s'intéresse à un cas particulier étroitement lié au LPP.

On se place sur \mathbb{Z}^{1+1} et on attribue à chaque point un poids aléatoire $X_i^{(j)}$. On munit $\Pi(N, k)$ de la mesure uniforme et on définit l'énergie d'un chemin $\pi \in \Pi(N, k)$ par

$$H(\pi) = \sum_{(i,j) \in \pi} X_i^{(j)}.$$

La mesure de Gibbs correspondante est alors donnée par

$$P_{N,k}(\pi) = \frac{1}{Z_\beta(N, k)} \exp(\beta H(\pi))$$

où

$$Z_\beta(N, k) = \sum_{\pi \in \Pi(N, k)} \exp(\beta H(\pi)).$$

$P_{N,k}$ est une mesure de probabilité aléatoire car elle dépend des $X_i^{(j)}$. On s'intéresse au comportement de la fonction de partition aléatoire $Z_\beta(N, k)$. En considérant le maximum des énergies sur l'ensemble des chemins $\Pi(N, k)$, on démontre facilement que

$$G(N, k) \leq \frac{1}{\beta} \log Z_\beta(N, k) \leq G(N, k) + \frac{1}{\beta} \log \#\Pi(N, k). \quad (2.6)$$

Une conséquence immédiate de l'inégalité (2.6) est que $\frac{1}{\beta} \log Z_\beta(N, k)$ converge vers $G(N, k)$ quand la température converge vers 0 *i.e.* lorsque $\beta \rightarrow \infty$. De plus, $G(N, k)$ et $\frac{1}{\beta} \log Z_\beta(N, k)$ se comportent, asymptotiquement en N et k , de la même manière dans tout régime faisant disparaître le terme $\frac{1}{\beta} \log \#\Pi(N, k)$.

2.3 Résultats connus

Dans ce paragraphe, on fait un rappel des résultats importants concernant le modèle de percolation de dernier passage (LPP). On s'intéresse à la variable $G(N, k)$ particulièrement sous la forme (2.1). D'un côté, c'est la version la plus répandue dans la littérature et d'un autre côté, on peut se poser des questions sur la nature du chemin π qui maximise la somme dans (2.1). **Désormais, les variables $X_i^{(j)}$ sont supposées *i.i.d.* de même loi qu'une variable générique X d'espérance μ et de variance σ^2 .**

On présente les résultats sur le LPP en deux parties. D'abord, on expose brièvement quelques résultats importants obtenus avant les travaux de Johansson [Joh00a]. Ensuite, on évoque l'identification fondamentale du temps de dernier passage dans le cas géométrique ou exponentiel à un processus déterminantal de la forme (1.1). Ce résultat essentiel obtenu par Johansson [Joh00a] entraîne plusieurs théorèmes liés à la structure déterminantale détaillée dans le chapitre précédent.

2.3.1 Résultats antérieurs aux travaux de Johansson : loi des grands nombres et théorème central limite

Les premiers résultats obtenus sont des théorèmes de convergence du premier ordre ou bien ce qu'on appelle lois des grands nombres. Rost [Ros81] est le premier à démontrer que si les $X_i^{(j)}$ sont *i.i.d.* de loi exponentielle de paramètre 1, alors on a que

$$\lim_{N \rightarrow \infty} \frac{G(\lfloor xN \rfloor, N)}{N} = (1 + \sqrt{x})^2 \quad p.s. \quad (2.7)$$

pour tout $x > 0$. Cette limite est souvent appelée « limite hydrodynamique ». Plus tard, Glynn et Whitt [GW91] ont démontré que si les $X_i^{(j)}$ sont *i.i.d.* admettant des moments exponentiels finis, alors on a pour tout $\gamma > 0$, que

$$\lim_{N \rightarrow \infty} \frac{G(N, \lfloor \gamma N \rfloor)}{N} = g(\gamma) \quad p.s.$$

La fonction g est déterministe, continue, concave et strictement croissante. L'existence d'une telle fonction repose essentiellement sur une version améliorée du théorème sous-additive ergodique de Kingman démontrée par Liggett, (voir [Lig85]). La condition de moment exponentiel fini est largement suffisante mais n'est pas nécessaire pour que g soit finie partout. Martin [Mar04] donne une condition plus faible sur la queue de X pour que la fonction g soit finie pour tout $\gamma \geq 0$, et continue partout y compris sur les bords. Si on note F la fonction de répartition de la variable X , Martin suppose que

$$\mathbb{E}|X| < \infty \quad \text{et} \quad \int_0^\infty (1 - F(s))^{\frac{1}{2}} ds < \infty. \quad (2.8)$$

Cette condition est légèrement plus forte que $\mathbb{E}X_+^2 < \infty$ et s'étend au modèle de percolation en dimension d en remplaçant le $1/2$ dans l'intégrale par $1/d$. En dimension 2, si en plus de la condition (2.8), on a que

$$\int_{-\infty}^0 F(s)^{\frac{1}{2}} ds < \infty,$$

alors Martin [Mar04] démontrent que pour tout γ proche de 0,

$$g(\gamma) = \mu + 2\sigma\sqrt{\gamma} + o(\sqrt{\gamma}). \quad (2.9)$$

On connaît une forme explicite de g lorsque X suit une loi géométrique de paramètre $q \in (0, 1)$. Dans [KS91], les auteurs démontrent en effet que

$$g(\gamma) = \frac{\gamma q + 2\sqrt{\gamma q} + q}{1 - q} \quad (2.10)$$

Plus tard, Johansson [Joh00a] identifie la loi de $G(N, k)$ dans le cas géométrique et exponentielle à un gaz de Coulomb et retrouve donc les limites hydrodynamiques (2.7) et (2.10) comme étant la borne supérieure du support d'une certaine mesure d'équilibre.

En ce qui concerne les théorèmes de convergence du second ordre, Glynn et Whitt [GW91] ont établi un premier théorème central limite à k fixé. Si les $X_i^{(j)}$ sont *i.i.d.* de moyenne 0 et de variance σ^2 , on a que

$$\lim_{N \rightarrow \infty} \frac{G(N, k) - N}{\sigma\sqrt{N}} = L(1, k), \quad (2.11)$$

où $L(1, k)$ est la même fonctionnelle du mouvement Brownien k -dimensionnel définie dans (1.14). En observant (2.11) et le Théorème 1.12, on comprend que $L(N, k)$ joue le rôle du pont entre le LPP et le GUE.

2.3.2 Processus déterminantal : poids géométriques et exponentiels

Après les travaux de Glynn et Whitt, il faudra attendre une dizaine d'année pour avoir des résultats précis sur la distribution du temps de dernier passage. Dans son célèbre article [Joh00a], Johansson étudie le comportement du temps de dernier passage lorsque les poids suivent une loi géométrique ou exponentielle. Il démontre que la loi de $G(N, k)$ admet dans ces cas une représentation déterminantale de la forme (1.1). Cette représentation avait déjà été observée dans l'étude de la plus grande valeur du GUE [Meh91] et de la plus longue sous-suite croissante d'une permutation aléatoire [BDJ99].

Théorème 2.1 (Johansson [Joh00a]). *On suppose que X suit une loi géométrique de paramètre q . Alors, pour $N \geq k$ et pour tout $t \geq 0$,*

$$\mathbb{P}[G(N, k) \leq t] = \frac{1}{Z_{N,k}} \sum_{\substack{h \in \mathbb{N}^k \\ \max\{h_i\} \leq t+k-1}} \prod_{1 \leq i < j \leq k} (h_i - h_j)^2 \prod_{i=1}^k \binom{h_i + N - k}{h_i} q^{h_i}, \quad (2.12)$$

où $Z_{N,k}$ est une constante de normalisation que l'on peut calculer explicitement.

Etant donné qu'une loi exponentielle est une limite en loi d'une loi géométrique convenablement normalisée, il obtient un résultat similaire pour des poids exponentiels.

Corollaire 2.2. *On suppose que X suit une loi exponentielle de paramètre 1. Alors, pour $N \geq k$ et pour tout $t \geq 0$,*

$$\mathbb{P}[G(N, k) \leq t] = \frac{1}{Z'_{N,k}} \int_{[0,t]^k} \prod_{1 \leq i < j \leq k} (x_i - x_j)^2 \prod_{i=1}^k x_i^{N-k} e^{-x_i} d^k x. \quad (2.13)$$

$Z_{N,k}$ et $Z'_{N,k}$ sont souvent appelées fonctions de partition. Dans l'étude de la plus longue sous-suite croissante d'une permutation aléatoire, Baik, Deift et Johansson utilise la correspondance R-S (Robinson-Schensted) qui associe à chaque permutation du groupe symétrique S_N , deux tableaux de Young de taille N et de même forme, (voir [Ful97, Sag91]). La plus longue sous-suite croissante d'une permutation correspond alors à la longueur de la première ligne de l'un des deux tableaux de Young. La preuve de (2.12) repose essentiellement sur une version généralisée de la correspondance R-S, à savoir, la correspondance R-S-K (Robinson-Schensted-Knuth).

Éléments de preuve du Théorème 2.1 : On considère l'ensemble \mathcal{A} des matrices de taille $N \times k$ dont les entrées sont des entiers naturels. On remarque d'abord que la probabilité d'une configuration donnée des $X_i^{(j)}$ vaut $(1-q)^{Nk} q^M$ où M est la somme des valeurs de cette configuration. Par conséquent, la probabilité d'une matrice quelconque de \mathcal{A} ne dépend que de la somme de ses entrées. Ensuite, on associe à chaque matrice $(a_{ij}) \in \mathcal{A}$ ce qu'on appelle une permutation généralisée. Cette permutation n'est autre que l'ensemble des couples (i, j) répété chacun $a_{i,j}$ fois. Une dernière étape consiste à associer à chaque permutation généralisée deux tableaux de Young semi-standards de même forme $\lambda \vdash M$, (voir [Ful97, Sag91]). Un calcul direct montre que

$$\mathbb{P}[G(N, k) \leq t] = \sum_{M=0}^{\infty} (1-q)^{Nk} q^M \sum_{\lambda \vdash M, \lambda_1 \leq t} L(\lambda, N, k), \quad (2.14)$$

où $L(\lambda, N, k)$ est le nombre de paires $(P(\lambda), Q(\lambda))$ de tableaux de Young semi-standards de forme λ et dont les éléments sont choisis respectivement dans $\{1, \dots, N\}$ et $\{1, \dots, k\}$. λ_1 est la longueur de la première ligne du tableau de Young et correspond à $G(N, k)$. En utilisant les fonctions de Schur [Sag91], on démontre que

$$L(\lambda, N, k) = \prod_{1 \leq i < j \leq N} \frac{\lambda_i - \lambda_j + j - i}{j - i} \prod_{1 \leq i < j \leq k} \frac{\lambda_i - \lambda_j + j - i}{j - i}. \quad (2.15)$$

Pour terminer, on pose $h_i = \lambda_i + k - i$ pour tout $i = 1, \dots, k$ et on insère (2.15) dans (2.14). \square

Il est clair que $G(N, k) = \max_i \{h_i\}$ et que le processus ponctuel $h = (h_i)_{i=1}^k$ suit une loi de la forme (1.1) avec ν la loi binomiale négative. En utilisant les mêmes techniques de la Partie 1.1.3,

on peut démontrer que la fonction de corrélation de h satisfait l'équation (1.9) du Théorème 1.8. En effet, pour tout $l \geq 1$, on a

$$R_l^{(k)}(h_1, \dots, h_l) = \det (K_k(h_i, h_j))_{1 \leq i, j \leq l} \quad (2.16)$$

Le noyau polynomial $K_k(h_i, h_j)$ est le noyau discret de Meixner et il est construit via les polynômes orthogonaux de Meixner de $l^2(\mathbb{N}, \nu)$. Dans le cas exponentiel, on parle du noyau continu de Laguerre. Cette représentation déterminantale de la loi de h est essentielle pour démontrer la convergence des fluctuations de $G(N, k)$ vers la loi de Tracy-Widom.

Remarque 2.3. Pour démontrer (2.13), il suffit de remarquer qu'une variable de loi exponentielle de paramètre 1 est la limite en loi lorsque $l \rightarrow \infty$ de X_l/l où X_l est une variable de loi géométrique de paramètre $q = 1 - 1/l$. Si on suppose maintenant que $q = \frac{\alpha}{N^2}$, alors on a que $G(N, N)$ converge en loi vers le processus déterminantal de Hammersley $L_k(\alpha)$ [Bai05]. $L_k(\alpha)$ est la version Poissonnisée du problème de la plus longue sous-suite croissante l_n d'une permutation uniformément choisie dans S_n :

$$\mathbb{P}[L(\alpha) \leq t] = \sum_{m=0}^{\infty} \frac{e^{-\alpha} \alpha^m}{N!} \mathbb{P}[l_m \leq t].$$

Remarque 2.4. Il est très important de noter que la fonction de répartition (2.13) est la même que celle de la plus grande valeur propre d'une matrice de Wishart complexe $k \times k$. Si on perturbe les paramètres de la loi exponentielle suivant les colonnes, on a la même égalité entre la fonction de répartition de $G(N, k)$ et celle de la plus grande valeur propre d'une matrice de Wishart complexe perturbée, (voir Chapitre 5). Ce fait, ainsi que l'égalité en loi entre $L(1, k)$ et la plus grande valeur propre du GUE, illustre la ressemblance structurelle entre les modèles matriciels aléatoires et ceux de percolation dirigée planaire.

Une preuve alternative de (2.12) est donnée plus tard par Johansson dans [Joh10]. Il regarde la chaîne de Markov multidimensionnelle

$$\mathbf{G}(N) = (G(N, 1), \dots, G(N, k)),$$

et remarque qu'on peut réécrire $G(N, k)$ de la manière suivante :

$$G(N, k) = \max_{1 \leq j \leq N} \left\{ G(j, k-1) + \sum_{i=1}^N X_i^{(k)} - \sum_{i=1}^{j-1} X_i^{(k)} \right\}.$$

L'étude est basée sur la connaissance de la probabilité de transition de la chaîne de Markov $\mathbf{G}(N)$. Sans aller plus loin dans les détails, on introduit les notations nécessaires pour l'énoncé du résultat de Johansson. On note $W_k = \{x \in \mathbb{Z}^k; x_1 \leq \dots \leq x_k\}$ et $\omega_N(\cdot)$ la densité d'une loi binomiale négative de paramètre (N, q) . On définit les opérateurs discrets de dérivation et d'anti-dérivation

$$\Delta f(x) = f(x+1) - f(x) \quad \text{et} \quad (\Delta^{-1}f)(x) = \sum_{y=-\infty}^{x-1} f(y).$$

On a alors le théorème suivant :

Théorème 2.5 (Johansson [Joh10]). *Pour $x, y \in W_k$, on a pour tout $m \geq n$, que*

$$\mathbb{P}[\mathbf{G}(m) = y | \mathbf{G}(n) = x] = \det(\Delta^{j-i} \omega_{m-n}(y_i - x_i))_{1 \leq i, j \leq k}.$$

Par conséquent, pour tout $t \in \mathbb{N}$,

$$\mathbb{P}[G(N, k) \leq t] = \det(\Delta^{j-i-1} \omega_N(t))_{1 \leq i, j \leq k}. \quad (2.17)$$

A partir de (2.17), on peut retrouver la forme déterminantale (2.12) sans utiliser la correspondance R-S-K, (voir [Joh10]). Cette approche est motivée par un travail de Dieker et Warren [DW08] sur les valeurs propres de certaines matrices Hermitiennes. Si $(B_n)_{n \geq 0}$ est une suite de mouvements Brownien indépendants, on considère le processus multidimensionnel $\mathbf{X}(t) = (X_1(t), \dots, X_k(t))$ où

$$X_k(t) = \sup_{0 \leq s \leq t} \{X_{k-1}(s) + B_k(t) - B_k(s)\}.$$

Dieker et Warren démontrent que $\mathbf{X}(\frac{1}{2}) \stackrel{\mathcal{L}}{=} (\lambda_{\max}(H_1), \dots, \lambda_{\max}(H_k))$, où H_j est le mineur d'ordre j d'une matrice H_k $k \times k$ du GUE.

On continue la présentation des résultats sur $G(N, k)$ obtenus dans le cas géométrique grâce à la formule (2.12) et à sa représentation déterminantale. Dans le cas exponentiel, les résultats sont identiques à quelques modifications près et seront donnés dans le Chapitre 5 dans le cadre des matrices de Wishart complexes. Contrairement à l'ensemble d'Hermite, l'ensemble de Meixner est discret. Ceci constitue une difficulté technique supplémentaire par rapport à l'étude du noyau d'Hermite présentée dans le chapitre précédent. De même, le gaz de coulomb correspondant à $G(N, k)$, *i.e.* la fonction de répartition (2.12), est discret et donc nécessite un effort supplémentaire pour établir des formules asymptotiques de grandes déviations.

Les grandes déviations à droite de la moyenne ont d'abord été étudiées par Seppäläinen [Sep98c] dans le cadre des sous-suites croissantes dans le plan. Sa méthode s'applique également pour le LPP. Pour les grandes déviations à gauche, Johansson fait les calculs dans le cadre général d'un gaz de Coulomb discret, (voir Remarque 1.1.2). En modifiant le potentiel V_k dans (1.7), on modifie la mesure d'équilibre recherchée. La limite hydrodynamique, qui existe grâce à la superadditivité, apparaît comme étant la borne supérieure du support de cette mesure d'équilibre.

Théorème 2.6 (Johansson [Joh00a], Seppäläinen [Sep98a]). *On suppose que X suit une loi géométrique de paramètre $q \in (0, 1)$. Pour tout $\gamma \geq 1$, on a que*

$$\omega(\gamma, q) := \lim_{k \rightarrow \infty} \frac{1}{k} \mathbb{E}[G([\gamma k], k)] = \frac{(1 + \sqrt{\gamma q})^2}{1 - q} - 1.$$

De plus, pour tout $\varepsilon > 0$,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathbb{P}[G([\gamma k], k) \geq k(\omega(\gamma, q) + \varepsilon)] = -i(\varepsilon),$$

et

$$\lim_{k \rightarrow \infty} \frac{1}{k^2} \log \mathbb{P}[G([\gamma k], k) \leq k(\omega(\gamma, q) - \varepsilon)] = -l(\varepsilon).$$

Les bonnes fonctions de taux $l(\varepsilon)$ et $i(\varepsilon)$ sont positives pour tout $\varepsilon > 0$.

Les calculs détaillés sont donnés dans les sections 2 et 4 de [Joh00a]. Le lecteur pourra consulter aussi les travaux de Hiai et Petz [HP00a] et de Féral [Fér08].

L'étude de la convergence des fluctuations de $G(N, k)$ dans le cas géométrique montre encore une fois l'importance et l'universalité de la loi de Tracy-Widom. En analysant le noyau discret de Meixner, Johansson obtient le théorème suivant :

Théorème 2.7 (Johansson [Joh00a]). *On suppose que X suit une loi géométrique de paramètre $q \in (0, 1)$. Pour tout $\gamma \geq 1$ et $s \in \mathbb{R}$,*

$$\lim_{k \rightarrow \infty} \mathbb{P} \left[\Theta_k := \frac{G([\gamma k], k) - k\omega(\gamma, q)}{\sigma(\gamma, q)k^{1/3}} \leq s \right] = F_2^{\text{TW}}(s),$$

où

$$\sigma(\gamma, q) = \frac{q^{1/6}\gamma^{-1/6}}{1-q} (\sqrt{\gamma} + \sqrt{q})^{2/3} (1 + \sqrt{\gamma q})^{2/3}$$

F_2^{TW} est la même fonction de répartition définie dans le Théorème 1.10. La preuve du théorème précédent passe par plusieurs étapes. D'abord, on établit, à partir de (2.12), une représentation déterminantale (2.16) de la fonction de corrélation avec le noyau de Meixner, ([Joh00a], Section 3). Ensuite, on étudie les asymptotiques des polynômes de Meixner pour obtenir, avec la bonne normalisation, la convergence du noyau de Meixner vers celui d'Airy, ([Joh00a], Section 5). Cette étape nous donne une idée sur l'ordre de convergence des fluctuations (ici $k^{1/3}$). A la fin, on démontre que le déterminant de l'opérateur de Fredholm correspondant au noyau d'Airy coïncide avec la fonction de répartition de la loi de Tracy-Widom, (voir [TW94]). Des détails sont donnés dans le Paragraphe 1.1.3 pour l'ensemble continu d'Hermite.

Pour terminer l'exposition, on énonce deux inégalités optimales de déviations. La convergence en loi de Θ_k vers la loi de Tracy-Widom n'implique pas forcément la convergence des moments de Θ_k . Dans le cas géométrique et exponentiel, on peut espérer que Θ_k aura des queues similaires à celles d'une loi de Tracy-Widom. En effet, Baik et al. [BDM⁺01] ont établi, en utilisant ce qu'on appelle le problème de Riemann-Hilbert, des inégalités optimales mettant en évidence les queues de la loi de Tracy-Widom.

Théorème 2.8 (Baik et al. [BDM⁺01]). *On suppose que X est de loi géométrique de paramètre $q \in (0, 1)$. Pour tout $\gamma \geq 1$ fixés, il existe deux constantes $L > 0$ (grande) et $\delta > 0$ (petite) telles que pour tout k suffisamment grand et pour tout $L \leq \varepsilon \leq \delta N^{\frac{2}{3}}$, on a que*

$$\mathbb{P}[\Theta_k \geq \varepsilon] \leq e^{-c\varepsilon^{3/2}},$$

et

$$\mathbb{P}[\Theta_k \leq -\varepsilon] \leq e^{-c'\varepsilon^3}.$$

Ici c et c' sont deux constantes numériques.

L'inégalité à droite de la moyenne peut être obtenue en utilisant les résultats asymptotiques de grandes déviations associés à un argument de superadditivité, (voir Section 2 de [Joh00a]). Cependant, la preuve de l'inégalité à gauche de la moyenne est beaucoup plus délicate et passe par la solution d'un problème de Riemann-Hilbert associé à la fonction de répartition de $G(N, k)$. Cette

méthode, appelée « steepest descent method for Riemann-Hilbert problem », a été introduite par Deift et Zhou dans [DZ92], (voir également [Dei99]).

Le Théorème 2.8, associés au Théorème 2.7, a deux conséquences importantes : d'abord, on peut démontrer, en utilisant une simple intégration par partie que les moments de tout ordre de Θ_k convergent vers les moments de la loi de Tracy-Widom. En effet, pour tout m entier positif,

$$\lim_{k \rightarrow \infty} \mathbb{E}[\Theta_k^m] = \mathbb{E}[\Theta^m].$$

Une autre conséquence du Théorème 2.8 est la détermination de l'ordre des fluctuations transversales. En utilisant la méthode développée par Johansson dans [Joh02], on démontre sans grandes difficultés, qu'avec probabilité allant vers 1, un chemin maximisant est inclus dans un cylindre de largeur $k^{2/3}$ autour de la diagonale. Cela confirme la conjecture selon laquelle $2\xi - 1 = \chi$, où ξ désigne l'ordre des fluctuations transversales (ici $\xi = 2/3$) et χ désigne l'ordre de convergence des fluctuations longitudinales (ici $\chi = 1/3$). Dans le chapitre suivant, on fournit des preuves complètes de ces deux conséquences dans le cadre d'une percolation sur un rectangle fin.

Les mathématiciens croient fortement que les Théorèmes 2.7 et 2.8, ainsi que leurs conséquences, restent valables si on remplace les poids géométriques par d'autres poids satisfaisant des conditions de moments. Cependant, aucune preuve d'universalité n'a été encore donnée. En l'absence de forme déterminantale, la tâche semble assez difficile. Dans le cas d'une matrice de Wigner, plusieurs théorèmes d'universalité ont été établis grâce à la structure algébrique très connue des matrices. La structure du LPP y compris les chemins maximisant n'est pas aussi favorable. Dans le chapitre suivant, toujours en quête d'universalité, on contourne ces difficultés en considérant un rectangle de largeur négligeable par rapport à sa longueur.

2.4 Modèles de percolation planaires voisins

Nous regardons dans ce paragraphe deux modèles de percolation planaire discrets mettant en évidence des propriétés analogues à celles du modèle de percolation de dernier passage (LPP) introduit précédemment. Ces modèles, comparés au LPP, résultent souvent d'une simple modification des trajectoires des chemins ou bien de la géométrie de la boîte dans laquelle ces chemins évoluent. Par exemple, si on remplace les poids géométriques par d'autres poids de loi de Bernoulli, un changement spécifique des trajectoires des chemins dirigés permet de récupérer un processus déterminantale lié à l'ensemble polynomial de Krawtchouk. Si on impose que les chemins dirigés s'arrêtent sur la diagonale du carré (à mi-chemin), on obtient un modèle de percolation dirigée sur un triangle pour lequel il existe des résultats dans la littérature. On commence d'abord par le modèle de percolation de premier passage de Seppäläinen.

2.4.1 Modèle de percolation de premier passage de Seppäläinen

Ce modèle est une version simplifiée du LPP défini plus haut. Il a été introduit par Seppäläinen [Sep98b] qui en a explicité la limite hydrodynamique. Plus tard, Johansson [Joh01a] a établi une représentation déterminantale de la loi du temps de premier passage via l'ensemble polynomial de Krawtchouk.

On considère la boîte rectangulaire incluse dans \mathbb{Z}^2 et limitée par les axes et les points $(0,0)$ et (l,k) . On attribue à chaque côté e reliant deux sites voisins un poids $\tau(e)$. Pour les côtés verticaux, $\tau(e)$ est déterministe et vaut $\tau_0 > 0$ tandis que $\tau(e)$ suit une loi de Bernoulli de paramètre p pour les côtés horizontaux *i.e.* $\mathbb{P}[\tau(e) = \lambda] = p = 1 - q = 1 - \mathbb{P}[\tau(e) = \kappa]$ avec $\kappa > \lambda \geq 0$. On définit alors le temps de premier passage

$$T(l, k) = \min_{\pi \in \Pi(l, k)} \sum_{e \in \pi} \tau(e).$$

Ici un chemin π est l'ensemble des arêtes qui le constituent et non pas des sommets comme dans le LPP. Seppäläinen a calculé dans [Sep98b] la limite hydrodynamique de $T(\lfloor nx \rfloor, \lfloor ny \rfloor)$. En effet, il a obtenu que

$$\lim_{n \rightarrow \infty} \frac{T(\lfloor nx \rfloor, \lfloor ny \rfloor)}{n} := u(x, y) = \begin{cases} \lambda x + \tau_0 y & \text{si } py > qx \\ \lambda x + \tau_0 y + (\kappa - \lambda)(\sqrt{qx} - \sqrt{py})^2 & \text{si } py \leq qx \end{cases}$$

Cette limite a été calculé sans la connaissance de la loi de T . Cependant, si on parvient à établir une représentation en gaz de Coulomb de la loi de T , on peut retrouver cette limite hydrodynamique en tant que borne supérieure du support de la mesure d'équilibre de ce gaz, (voir [Joh00a, Joh01a]).

On considère maintenant un modèle de percolation dirigée sur la même grille rectangulaire du LPP avec des poids $X_i^{(j)}$ de loi de Bernoulli indépendantes et de même paramètre. On effectue une légère modification sur les chemins dirigés en imposant un sommet par colonne. Un chemin est alors de la forme $\pi = \{(i, j_i)\}_{i=1}^N$ avec $1 \leq j_1 \leq \dots \leq j_N \leq k$. On définit la variable

$$L(N, k) = \max_{\pi} \sum_{(i, j) \in \pi} X_i^{(j)}.$$

On remarque d'abord que

$$T(N, k - 1) = k\tau_0 + N\kappa - (\kappa - \lambda)L(N, k). \quad (2.18)$$

Pour déterminer la loi de $L(N, k)$, Johansson [Joh01a] procède de la même manière que pour la loi de $G(N, k)$ en utilisant cette fois une variante de la correspondance R-S-K. Cette version associe à chaque matrice A à valeurs dans $\{0, 1\}$ un couple de tableaux de Young semi-standards (P, Q) . Dans le cas de variables géométriques, on attribuait à chaque matrice à valeurs dans \mathbb{N} une permutation généralisée à laquelle on attribuait un couple de tableaux de Young semi-standards de même forme. Ici, la permutation généralisée qui correspond à une matrice à valeurs dans $\{0, 1\}$ est sans répétition *i.e.* chaque terme (couple) apparaît au plus une fois. On attribue à ce type de permutation, deux tableaux de Young semi-standards de forme conjuguée, (voir [Ful97]). La preuve de la formule suivante repose également sur le dénombrement des tableaux de Young via les polynômes de Schur. On a ainsi, pour tout $t \geq 0$,

$$\mathbb{P}[T(N, k) \leq t] = \frac{1}{Z_{N, k}} \sum_{\substack{h \in \mathbb{N}^k \\ \max\{h_i\} \leq t+k-1}} \prod_{1 \leq i < j \leq k} (h_i - h_j)^2 \prod_{i=1}^k \binom{N+k-1}{h_i} p^{h_i} (1-p)^{N+k-1-h_i}. \quad (2.19)$$

Cette fonction de répartition suit le prototype (1.1). On distingue clairement le déterminant de Vandermonde au carré multiplié par le produit tensoriel de la loi binomiale. On parle alors de l'ensemble polynomial de Krawtchouk. Grâce à la formule (2.19), Johansson [Joh01a] démontre la convergence des fluctuations de L , et donc de T , vers la loi de Tracy-Widom. L'ordre de convergence est encore une fois $k^{1/3}$. (2.19) est également un gaz de Coulomb discret et donc, on peut établir des formules asymptotiques de grandes déviations.

2.4.2 Percolation dirigée sur un triangle

On introduit une autre variante du modèle de percolation défini précédemment qui émerge également de la physique des particules. Le modèle est présenté dans [CGM09] sous une forme multidimensionnelle mais seul le cas planaire sera traité ici. Dans ce cas, la seule différence avec le LPP est que les chemins dirigés s'arrêtent à la diagonale qui ne traverse pas l'origine. Plus précisément, on considère la boîte carré de \mathbb{Z}^2 ayant pour sommet les points $(1, 1)$ et (N, N) et dont les côtés sont parallèles aux axes. On suppose que les chemins dirigés, notés p , évoluent à partir du point $(1, 1)$ de la même manière que les chemins du LPP mais n'effectuent que la moitié du trajet, soit N pas, et s'arrêtent à la diagonale. On définit alors la variable

$$Z(N) = \max_p E(p),$$

avec $E(p) = \sum_{(i,j) \in p} X_i^{(j)}$. On connaît très peu de résultats sur $Z(N)$ malgré la ressemblance avec le LPP. Ceci est dû à l'absence d'une représentation déterminantale de la loi de $Z(N)$. Tous les résultats obtenus sont donc approximatifs et donnent rarement lieu à des formules explicites. On suppose dans la suite que les $X_i^{(j)}$ sont des variables Gaussiennes indépendantes. On analysera alors deux propriétés traitées dans la littérature : les grandes déviations étudiées dans [CGM09] et la majoration de la variance [Cha08].

Grandes déviations : Cranston, Gauthier et Mountford ont introduit dans [CGM09] une méthode combinatoire qui permet de construire un chemin p_0 dont l'énergie $E(p_0)$ dépasse une certaine valeur. Ceci leur permet de contrôler les grandes déviations à gauche de la moyenne sans pour autant pouvoir calculer les fonctions de taux. La technique utilisée est assez générale dans le sens où elle dépend très peu de la dimension de l'espace et de la loi des $X_i^{(j)}$. Par contre, elle est spécifique à la géométrie de la boîte dans laquelle évoluent les chemins. Les calculs fait dans [CGM09] sont techniques et nécessitent l'introduction de nombreuses notations et indices. Cependant, par souci de simplification, l'étape cruciale de la preuve sera rapportée ci-dessous d'une façon allégée en notation mais sera appuyée par des figures.

Pour démontrer la superadditivité de $Z(N)$, on définit la variable

$$Z(M, N) = \min_{s \in \mathcal{D}_M} \max_{p \in \mathcal{P}(s, N)} E(p),$$

où \mathcal{D}_M est la diagonale se situant à M pas de l'origine et $\mathcal{P}(s, N)$ est l'ensemble des chemins qui croissent du point s vers la diagonale située à N pas de l'origine. On a immédiatement que $Z(M, N)$ est superadditive et que $Z(N) = Z(1, N)$. D'après la version de Liggett [Lig85] du théorème ergodique de sous-additivité de Kingman, on a que

$$\lim_{N \rightarrow \infty} \frac{Z(N)}{N} = \mu \quad \text{p.s.}$$

où μ est une constante positive. D'après le même théorème, on a l'ordre des grandes déviations à droite de la moyenne :

$$-\infty < \lim_N \frac{1}{N} \log \mathbb{P}[Z(N) \geq N(\mu + \varepsilon)] = \inf_N \frac{1}{N} \log \mathbb{P}[Z(N) \geq (\mu + \varepsilon)N] < 0.$$

Pour les grandes déviations à gauche de la moyenne, on a le théorème suivant de Cranston, Gauthier et Mountford.

Théorème 2.1 (Cranston-Gauthier-Mountford [CGM09]). *On suppose que les $X_i^{(j)}$ sont indépendantes de loi Gaussienne $\mathcal{N}(0, 1)$. Alors pour tout $\varepsilon > 0$ suffisamment petit, on a*

$$-\infty < \lim_N \inf \frac{\log N}{N^2} \log \mathbb{P}[Z(N) \leq N(\mu - \varepsilon)] \leq \lim_N \sup \frac{\log N}{N^2} \log \mathbb{P}[Z(N) \leq N(\mu - \varepsilon)] < 0.$$

Ce théorème nous fournit l'ordre de convergence des grandes déviations à gauche de la moyenne mais ne permet pas d'explicitier la fonction de taux correspondante. L'ordre de convergence en $N^2/\log N$ est assez surprenant surtout que dans le cas rectangulaire, on a plutôt du N^2 . On pense que le $\log N$ est dû à la forme triangulaire de la boîte.

Le but n'étant pas de faire une démonstration complète mais de faire connaître aux lecteurs cette technique combinatoire de décomposition en blocs, la plupart des étapes techniques de la preuve sont omises. Les lecteurs curieux des détails calculatoires sont invités à consulter [CGM09].

Éléments de preuve du Théorème 2.1 : On considère la variable $Z(2^N)$ et on trace le segment $\mathcal{D}_{\delta 2^N}$ pour $\delta \in (0, 1)$ et la diagonale \mathcal{D}_{2^N} , (voir Figure 2.2). La première étape de la preuve consiste à étudier la partie des chemins qui se situe loin de l'origine entre $\mathcal{D}_{\delta 2^N}$ et \mathcal{D}_{2^N} . Pour cela, on décompose cette partie du triangle en petits blocs carrés de taille l , (voir Figure 2.2). On démontre qu'il existe une constante c telle que pour un ε donné, il existe l_0 tel que pour tout $l \geq l_0$, on peut construire à partir de n'importe quel point du côté gauche d'un carré donné, un chemin p aboutissant au côté droit du même carré et vérifiant

$$\sum_{(i,j) \in p} X_i^{(j)} \geq l(\mu - \varepsilon), \quad (*)$$

et ceci avec probabilité supérieure à $1 - c\varepsilon$. Pour une configuration donnée des $X_i^{(j)}$, on dit qu'un bloc est « bon » s'il vérifie la condition (*) et « mauvais » sinon. Par conséquence, pour parvenir à construire un chemin p reliant $\mathcal{D}_{\delta 2^N}$ à \mathcal{D}_{2^N} et vérifiant $\sum_{(i,j) \in p} X_i^{(j)} \geq (1 - \delta)2^N(\mu - \varepsilon)$, il nous faut une succession de $(1 - \delta)2^N/l$ « bon » blocs adjacents et reliant $\mathcal{D}_{\delta 2^N}$ à \mathcal{D}_{2^N} . Le nombre de « bon » blocs sur une ligne de blocs suit une loi binomiale de paramètre $((1 - \delta)2^N/l, 1 - c\varepsilon)$. On a ainsi $\delta 2^N/l$ variables binomiales indépendantes et de même paramètre. De cette manière, on parvient à construire à partir de 9/10 des points de $\mathcal{D}_{\delta 2^N}$, un chemin p tel que $\sum_{(i,j) \in p} X_i^{(j)} \geq (1 - \delta)2^N(\mu - \varepsilon)$ et ceci avec probabilité $> 1 - e^{-c\varepsilon 2^{2N}}$.

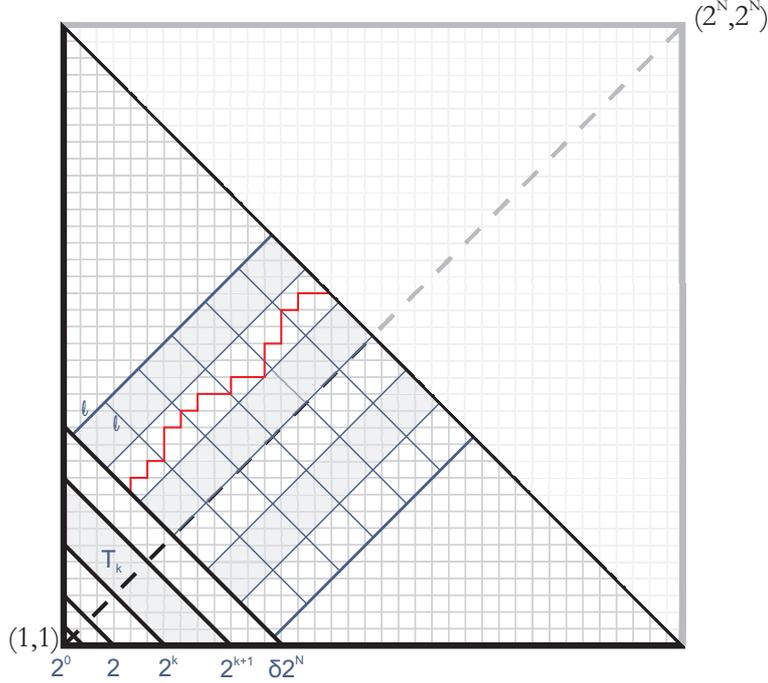


FIG. 2.2 – Décomposition du triangle en blocs

La seconde étape de la preuve consiste à analyser la partie proche de l'origine des chemins. On sait qu'une forte concentration de petits $X_i^{(j)}$ proche de l'origine peut engendrer un $Z(N)$ petit. On définit, Figure 2.2, pour tout $k = 0, \dots, N + \log \delta$, les blocs

$$T_k = \{(i, j) \in \mathbb{Z}^2; (i, j) \text{ est entre } \mathcal{D}_{2^k} \text{ à } \mathcal{D}_{2^{k+1}}\}$$

et les variables

$$V_k^- = \sum_{(i,j) \in T_k} (X_i^{(j)})^-,$$

où $(X_i^{(j)})^-$ représente la partie négative de $X_i^{(j)}$, i.e. $X_i^{(j)} = (X_i^{(j)})^+ - (X_i^{(j)})^-$. La variable V_k^- contrôle la masse négative dans le bloc T_k . On démontre, pour N assez grand, que

$$\mathbb{P} \left[\sum_{k=0}^{N+\log \delta} \frac{V_k^-}{2^k} \geq \epsilon 2^N \right] \leq \exp \left(- \frac{c\epsilon^2 2^{2N}}{N} \right). \quad (2.20)$$

En combinant (2.20) et les résultats obtenus loin de l'origine, on construit, avec probabilité supérieure à $1 - e^{-c'2^{2N}/N}$, un chemin p tel que $E(p) \geq 2^N(\mu - \epsilon)$. Ceci nous donne une majoration de la probabilité. Sans faire la preuve, on peut obtenir une minoration de la même probabilité qui vaut $e^{-c'2^{2N}/N}$. Le Théorème 2.1 en découle. □

Remarque 2.2. On a l'impression qu'un chemin maximisant a plus de liberté de fluctuer transversalement du fait qu'il n'est pas contraint à revenir à un point donné. Par suite, on s'attend à

un coefficient de super-diffusivité (fluctuation transversale) supérieur à $2/3$. Dans le cas du LPP, le chemin maximisant ne doit pas trop s'éloigner de la diagonale pour conserver une certaine « marge de manœuvre » avant d'aboutir au point (N, N) . De plus, on pense que le chemin maximisant dans le cas triangulaire aboutit sur un segment de la diagonale d'une largeur maximale de l'ordre de $N^{3/4}$. Si c'est le cas, on doit réussir à comparer facilement la limite hydrodynamique dans le cas triangulaire avec celle du LPP sur un carré. Si on décompose le carré de taille N en deux moitié séparées par la diagonale, on sait que $G(N, N) \leq Z(N) + Z'(N)$, (voir Figure 2.3). En reliant les chemins maximisants P et P' par le plus court chemin, on garantie que la différence entre $G(N, N)$ et $Z(N) + Z'(N)$ est $o(N)$.

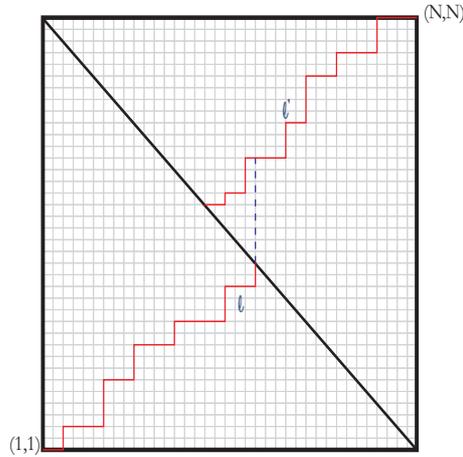


FIG. 2.3 – Passage du triangle au carré

Variance et chaos : Chatterjee a fait une première tentative pour mieux comprendre le modèle triangulaire et le comportement du chemin maximisant. Son article [Cha08] constitue une étude générale des champs Gaussiens et du chaos généré par une légère perturbation de ce champ. Il démontre dans un cadre général qu'une variance suffisamment petite est équivalente au chaos. Plus précisément, dans le cadre du modèle de percolation dans un triangle par exemple, Chatterjee considère la variable $Z(N)$ et une autre version perturbée notée $Z'(N)$. Il cherche une borne supérieure de la variance de $Z(N)$ pour en déduire que le nombre de points communs aux deux chemins maximisants P et P' est $o(N)$.

On note $\mathcal{P}(N)$ l'ensemble des chemins dirigés du modèle de percolation dans le triangle et on considère le champ (vecteur) Gaussien $\mathbf{G} = (E(p); p \in \mathcal{P}(N))$. On définit également le champ Gaussien perturbé $\mathbf{G}^t = (E^t(p); p \in \mathcal{P}(N))$ où $E^t(p) = \sum_{(i,j) \in p} Z_i^{(j)}(t)$. Les $Z_i^{(j)}(t)$ sont des variables Gaussiennes, indépendantes qu'on construit de la façon suivante :

$$Z_i^{(j)}(t) = e^{-t} X_i^{(j)} + \sqrt{1 - e^{-2t}} Y_i^{(j)},$$

où $t > 0$ est le réel perturbateur et les $Y_i^{(j)}$ sont des copies indépendantes des $X_i^{(j)}$. Si on note P le chemin maximisant $E(p)$ et P^t celui maximisant $E^t(p)$, on cherche à évaluer la quantité

$$\mathbb{E}[|P \cap P^t|],$$

où $|P \cap P^t|$ correspond au nombre de sites communs aux deux chemins P et P^t . On dit que la perturbation engendre le chaos ssi $\mathbb{E}[|P \cap P^t|] = o(N)$. Comme la perturbation t atteint tous les sites du triangle, il est naturel de supposer que t converge vers 0 quand $N \rightarrow \infty$. D'ailleurs, Chatterjee affirme qu'une petite perturbation sur l'ensemble des poids est équivalente à une grande perturbation sur un nombre restreint de poids. Le théorème suivant est le résultat principal concernant notre modèle obtenu dans [Cha08].

Théorème 2.3 (Chatterjee [Cha08]). *Pour $N \geq 1$, $\forall t \geq t_0 = (\log N)^{-\frac{1}{2}}$, on a*

$$\mathbb{E}[|P \cap P^t|] \leq \frac{CN}{\sqrt{\log N}},$$

où C est une constante universelle.

Sans faire la preuve complète de ce théorème, on attire l'attention du lecteur sur les deux étapes importantes permettant de le démontrer.

La première étape de la preuve concerne les champs Gaussiens en général et elle est prouvée dans la section 3 de [Cha08]. Soit $\mathbf{G} = (G_i)_{i \in S}$ un vecteur Gaussien centré indexé par un ensemble S fini. On note

$$R(i, j) := \text{Cov}(G_i, G_j) \quad \text{et} \quad \sigma^2 := \max_{i \in S} \text{Var} G_i.$$

On définit respectivement le maximum sur les indices et l'indice maximisant

$$M := \max_{i \in S} G_i \quad \text{et} \quad I := \operatorname{argmax}_{i \in S} G_i.$$

Pour être sûr que l'indice maximisant I est unique, on suppose que $G_i = G_j$ avec probabilité nulle pour tout $i \neq j$. On note finalement

$$v := \text{Var} M.$$

Remarque 2.4. Notons ici qu'on peut affirmer, sans tenir compte de la corrélation entre les G_i , que $v \leq \sigma^2$. Cela signifie que l'ordre des fluctuations de M ne peut pas dépasser l'ordre des fluctuations des G_i . Mais cette borne s'avère être grossière dans la plus part des cas où l'on a une structure de corrélation de \mathbf{G} .

A partir de \mathbf{G} , on définit le champ Gaussien perturbé $\mathbf{G}^t := e^{-t}\mathbf{G} + \sqrt{1 - e^{-2t}}\mathbf{G}$ et on note I^t l'indice maximisant G_i^t .

Théorème 2.5 (Chatterjee [Cha08]). *Pour tout t , on a*

$$0 \leq \mathbb{E}[R(I, I^t)] \leq \frac{v}{1 - e^{-t}}$$

et

$$v \leq \sigma^2(1 - e^{-t}) + e^{-t}\mathbb{E}[R(I, I^t)].$$

De plus $\mathbb{E}[R(I, I^t)]$ est une fonction décroissante de t .

Si on appelle superconcentration le fait que $v = o(\sigma^2)$ et chaos le fait que $\mathbb{E}[R(I, I^t)] = o(\sigma^2)$, le Théorème 2.5 nous dit que pour t tendant vers 0 avec une vitesse convenablement choisie, on a que

$$\text{superconcentration} \Leftrightarrow \text{chaos}.$$

Remarque 2.6. Pour démontrer le Théorème 2.5, on utilise le fait que

$$\text{Var}(Z(N)) = \mathbb{E}[|P \cap P^\tau|],$$

où τ est une variable de loi exponentielle indépendante de toutes les autres variables du modèle.

Pour faire l'analogie avec la percolation dirigée dans le triangle, il suffit de remarquer que $S = \mathcal{P}(N)$, $M = Z(N)$, $I = P$, $\sigma^2 = N$ et $v = v_N := \text{Var} Z(N)$. De plus, la structure de corrélation de \mathbf{G} , dans le cas d'une percolation dirigée sur le triangle, est complètement régie par les intersections des chemins dirigés. On a en effet que $R(p, p') = |p \cap p'|$. Si on parvient à démontrer que $v = o(\sigma^2)$, alors on a, d'après le Théorème 2.5 que $\mathbb{E}[R(I, I^t)] = o(\sigma^2)$.

La deuxième étape de la preuve, développée dans la partie 8 de [Cha08], consiste à trouver une majoration de la variance de $Z(N)$. Pour cela, Chatterjee applique une méthode introduite par Benjamini, Kalai et Schramm [BKS03] pour un résultat similaire pour un modèle planaire de percolation (non dirigée) de premier passage. Il démontre alors en se servant d'arguments d'hypercontractivité de Talagrand, que

$$\text{Var}(Z(N)) \leq \frac{CN}{\log N}, \quad (2.21)$$

où C est une constante universelle. En insérant la borne sur la variance (2.21) dans le Théorème 2.5, on obtient que

$$\mathbb{E}[|P \cap P^t|] \leq \frac{CN}{(1 - e^{-t}) \log N} \quad (2.22)$$

Or t est supposé converger vers 0 avec N et $(1 - e^{-t})$ se comporte comme t quand t est proche de 0. Par suite, pour obtenir du chaos, il faut que t converge vers 0 moins vite que $1/\log N$. Un bon compromis entre la vitesse de t et la majoration dans (2.22) serait de choisir t de l'ordre de $(\log N)^{-\frac{1}{2}}$. La majoration dans (2.22) serait donc de l'ordre de $N(\log N)^{\frac{1}{2}}$ et la perturbation engendre le chaos.

Chapitre 3

Percolation dirigée sur un rectangle fin et universalité

Dans le chapitre précédent, tous les résultats asymptotiques du temps de dernier passage $G(N, k)$, que ce soit la limite hydrodynamique, les fluctuations transversales et longitudinales ou bien les grandes déviations, supposent que N et k sont du même ordre, *i.e* $k/N = \gamma$. Cependant, ces résultats ont été obtenus pour des variables géométriques et exponentielles. Dans ce chapitre, on contourne les obstacles liés à l'universalisation des résultats du Chapitre 2 en suivant les travaux de Bodineau et Martin [BM05]. L'idée consiste à travailler sur un rectangle fin où le rapport k/N converge vers 0 lorsque $k \rightarrow \infty$. Cela permet de compenser l'erreur engendrée par le remplacement des poids géométriques par d'autres poids quelconques. Notons que, par symétrie, l'étude du cas où le rapport converge vers l'infini revient au même.

Dans la pratique, on suppose souvent que N et k sont du même ordre de grandeur aussi bien pour les files d'attente que pour le modèle d'exclusion simple. Le fait de considérer que k est petit devant N est une perte de généralité et non pas le contraire. Dans l'état actuel des recherches, on ne sait pas étendre les formules du chapitre 2 au cas de variables *i.i.d.* quelconques à moins de travailler sur un rectangle fin. C'est le prix à payer dans la quête de l'universalité.

Ce chapitre contient essentiellement notre contribution aux modèles de percolation de dernier passage sur un rectangle fin. On démontre des propriétés asymptotiques et non-asymptotiques de grandes déviations ainsi qu'un résultat sur les fluctuations transversales. L'étude ci-dessous a été motivée par les travaux de Bodineau et Martin [BM05] et Baik et Suidan [BS05] sur les fluctuations du temps de dernier passage sur un rectangle fin. Les auteurs de [BM05] et de [BS05] utilisent le résultat célèbre de Tracy-Widom [TW94] sur les fluctuations de plus grande valeur propre d'une matrice du GUE (Théorème 1.10) et celui de Barishnikov [Bar01] sur la percolation dirigée Brownienne (Théorème 1.12). Ils démontrent alors que sur un rectangle fin, les fluctuations de $G(N, k)$ convergent vers la loi de Tracy-Widom et ceci en remplaçant les variables géométriques (ou exponentielles) par des variables *i.i.d.* admettant des moments finis.

Théorème 3.7 (Bodineau-Martin [BM05] Baik-Suidan [BS05]). *On suppose que les $X_i^{(j)}$ sont *i.i.d.* et que $\mathbb{E}X = 0$ et $\mathbb{E}X^2 = 1$. De plus, on suppose qu'il existe $p > 2$ tel que $\mathbb{E}|X|^p < \infty$. Si $k = o(N^\alpha)$ avec $0 < \alpha < \frac{6}{7}(\frac{1}{2} - \frac{1}{p})$, on a que*

$$\Theta_k := \frac{G(N, k) - 2\sqrt{Nk}}{N^{1/2}k^{-1/6}} \xrightarrow[N, k \rightarrow \infty]{\mathcal{L}} \Theta,$$

où Θ est une variable de loi de Tracy-Widom de fonction de répartition F_2^{TW} .

Dans les deux articles, les auteurs comparent le temps du dernier passage $G(N, k)$ au temps du dernier passage Brownien $L(1, k)$. Cette comparaison engendre une erreur qui dépend de p et de k et qui est compensée par le fait que k est négligeable devant N .

3.1 Principes d'invariance et LPP

Les principes d'invariance, et en particulier l'approximation K-M-T (voir ci-dessous), sont au cœur de la preuve du Théorème 3.3 et de la plupart des résultats dans ce chapitre. On évoque dans ce paragraphe des techniques d'approximation (ou de couplage) qui seront utilisées dans la suite dans le cadre du LPP.

Pour illustrer l'utilité des techniques d'approximation, on attire l'attention du lecteur sur la preuve donnée par Lindeberg [Lin22] du théorème centrale limite classique. Celle-ci consiste à remplacer progressivement les variables Gaussiennes par des variables d'une loi différente et à majorer ensuite l'erreur cumulée due à ces remplacements. Dans le cadre du LPP, on note, à titre d'exemple, $G(N, k)$ le temps de dernier passage avec des poids exponentiels et $G'(N, k)$ celui avec des poids quelconques notés $X_i^{(j)'}$. Le but est de reconstruire les $X_i^{(j)}$ et les $X_i^{(j)'}$ d'une façon à ce que $G(N, k)$ et $G'(N, k)$ soient les plus proches possible. Ainsi, des résultats valables pour G pourront éventuellement être étendus pour G' .

Parmi les techniques de couplage les plus courantes, on cite le théorème de plongement de Skorokhod qui fut pendant longtemps la meilleure façon d'approximer, via le mouvement Brownien, une somme de variables *i.i.d.*, (voir [Sko65]). En 1975, Komlos, Major et Tusnady ont introduit dans [KMT76] la fameuse approximation K-M-T. Etant donné une mesure de probabilité μ satisfaisant des conditions de moments, cette approximation permet de construire une suite de variables *i.i.d.* de loi μ et dont la somme partielle est très proche d'une somme partielle de variables Gaussiennes indépendantes. Ce résultat est important d'autant plus que la majoration de l'erreur est optimale. On commence d'abord par l'exemple d'un couplage simple lié au LPP.

3.1.1 Un couplage simple

Ce paragraphe permet au lecteur de comprendre, par un exemple très simple, le fonctionnement et l'utilité des méthodes de couplage et d'approximation dans le cadre du LPP. On cherche à comparer le temps de dernier passage avec des variables de loi Gamma et celui avec des variables exponentielles. L'étude est également valable pour comparer le cas avec des poids binomiaux négatifs et celui avec des poids géométriques.

On considère donc le modèle de percolation de dernier passage muni des poids indépendants $(X_i^{(j)})_{i,j=1}^\infty$ de loi Gamma de paramètre $(m, 1)$ avec $m \in \mathbb{N}^*$. On appelle ce modèle le **modèle 1**. On note $G_1(N, k)$ le temps de dernier passage correspondant. On appelle **modèle 2** la même chose mais avec des variables $Y_i^{(j)}$ de loi exponentielle de paramètre 1. On note également $G_2(N, k)$ le temps de dernier passage correspondant. On désigne par X et Y deux variables génériques respectivement de même loi que les $X_i^{(j)}$ et les $Y_i^{(j)}$. Rappelons qu'une loi Gamma de paramètre

$(m, 1)$ est la convolée additive m fois d'une loi exponentielle de paramètre 1. On peut supposer ainsi que

$$X_i^{(j)} = \sum_{l=1}^m X_{i,l}^{(j)}, \quad (3.1)$$

où les $X_{i,l}^{(j)}$ sont *i.i.d.* de loi exponentielle de paramètre 1. Rappelons que la densité d'une loi Gamma de paramètre $(m, 1)$ est donnée, pour tout $x \in \mathbb{R}_+$, par

$$f(x) = \frac{x^{m-1} e^{-x}}{\Gamma(m)} \quad \text{où} \quad \Gamma(m) = (m-1)!.$$

Comme on l'a déjà vu dans le chapitre 2, le comportement du **modèle 2** est assez connu. Afin de valider des résultats du **modèle 2** pour le **modèle 1**, on effectue le couplage qui consiste à choisir, pour tous i et j entiers non nuls,

$$Y_i^{(j)} = X_{\lfloor \frac{i}{m} \rfloor, i - \lfloor \frac{i-1}{m} \rfloor m}^{(j)}. \quad (3.2)$$

On étale ainsi chaque variable du **modèle 1** sur les lignes dans le **modèle 2**. On compare maintenant $G_1(N, k)$ et $G_2(mN, k)$ en utilisant l'interprétation (2.3) de G . Soit $u^* \in U(N, k) \cap \mathbb{Z}_+^{k+1}$ un chemin maximisant la somme dans (2.3) pour le **modèle 1** et $v^* \in U(mN, k) \cap \mathbb{Z}_+^{k+1}$ un chemin maximisant de $G_2(mN, k)$ du **modèle 2**. D'après (2.3), on a que

$$G_1(N, k) = \sum_{r=1}^k [S_{u_r^*}^{(r)} - S_{u_{(r-1)-1}^*}^{(r)}] \quad (3.3)$$

et

$$G_2(mN, k) = \sum_{r=1}^k [T_{v_r^*}^{(r)} - T_{v_{(r-1)-1}^*}^{(r)}], \quad (3.4)$$

où S et T sont les sommes respectives sur les $X_i^{(j)}$ et les $Y_i^{(j)}$.

On attribue à u^* le chemin $v' \in U(mN, k) \cap \mathbb{Z}_+^{k+1}$ et à v^* le chemin $u' \in U(N, k) \cap \mathbb{Z}_+^{k+1}$ tels que

$$u' = \left(v_0^*, \left\lfloor \frac{v_1^*}{m} \right\rfloor, \left\lfloor \frac{v_2^*}{m} \right\rfloor, \dots, \left\lfloor \frac{v_{k-1}^*}{m} \right\rfloor, \frac{v_k^*}{m} = N \right)$$

et

$$v' = (mu_0^*, mu_1^*, mu_2^*, \dots, mu_{k-1}^*, mu_k^* = mN).$$

En comparant les deux modèles, (3.3) et (3.4) montrent que

$$G_1(N, k) - G_2(mN, k) \leq \sum_{r=1}^k [S_{u_r^*}^{(r)} - S_{u_{(r-1)-1}^*}^{(r)}] - \sum_{r=1}^k [T_{v_r'}^{(r)} - T_{v_{(r-1)-1}'}^{(r)}]$$

et que

$$G_2(N, k) - G_1(mN, k) \leq \sum_{r=1}^k [T_{v_r^*}^{(r)} - T_{v_{(r-1)-1}^*}^{(r)}] - \sum_{r=1}^k [S_{u_r'}^{(r)} - S_{u_{(r-1)-1}'}^{(r)}].$$

u' et v' sont construits respectivement de manière à effectuer des trajets similaires à v^* et u^* . D'après le couplage (3.1), les sommes S et T se compensent mutuellement à quelques variables près. Plus précisément, on obtient que

$$Z(N, k) := |G_1(N, k) - G_2(mN, k)| \leq \sum_{j=1}^k \max_{i=1, \dots, N} \{X_i^{(j)}\}. \quad (3.5)$$

le calcul qui suit permet de déterminer le comportement asymptotique de $Z(N, k)$ en probabilité. D'abord, on a que $\mathbb{E}e^{\lambda X} = (\mathbb{E}e^{\lambda Y})^m$ pour tout $\lambda \in \mathbb{R}$. Par suite, $\mathbb{E}e^{\lambda X} < \infty$ pour tout $\lambda < 1$. L'inégalité de Markov appliquée à (3.5) montre que, pour tout $0 < \lambda < 1$, $\varepsilon > 0$ et $0 < \beta \leq 1$,

$$\mathbb{P}[Z(N, k) > \varepsilon N^\beta] \leq \frac{\mathbb{E} \left[\exp \left(\lambda \max_{i=1, \dots, N} \{X_i^{(1)}\} \right) \right]^k}{\exp(\lambda \varepsilon N^\beta)} \leq CN \exp(-\lambda \varepsilon N^\beta + ck). \quad (3.6)$$

où c est une constante qui ne dépend que de m . La majoration dans (3.6) repose sur le calcul élémentaire suivant : pour $\lambda < \lambda' < 1$ convenablement choisi et N suffisamment grand,

$$\begin{aligned} \mathbb{E} \left[\exp \left(\lambda \max_{i=1, \dots, N} \{X_i^{(1)}\} \right) \right] &= 1 + \int_1^\infty \mathbb{P} \left[\exp \left(\lambda \max_{i=1, \dots, N} \{X_i^{(1)}\} \right) \geq s \right] ds \\ &\leq 1 + \int_{\log(N^{-\frac{1}{\lambda'}})}^\infty \lambda N^{\frac{\lambda}{\lambda'}} \exp(\lambda t) (1 - (1 - \mathbb{P}[X > t + \log(N^{\frac{1}{\lambda'}})])^N) dt. \end{aligned}$$

En coupant l'intégrale en deux parties, l'une sur $[-\log(N^{\frac{1}{\lambda'}}), 0]$ et l'autre sur $(0, +\infty)$, on obtient que

$$\begin{aligned} \mathbb{E} \left[\exp \left(\lambda \max_{i=1, \dots, N} \{X_i^{(1)}\} \right) \right] &\leq N^{\frac{\lambda}{\lambda'}} + \int_0^\infty \lambda N^{\frac{\lambda}{\lambda'}} \exp(\lambda t) \left(1 - \left(1 - \frac{\mathbb{E}[e^{\lambda' X}]}{N \exp(\lambda' t)} \right)^N \right) dt \\ &\leq N^{\frac{\lambda}{\lambda'}} + \int_0^\infty \lambda N^{\frac{\lambda}{\lambda'}} \exp(\lambda t) \left(1 - \exp \left(-\frac{\mathbb{E}[\exp(\lambda' X)]}{\exp(\lambda' t)} \right) \right) dt \\ &\leq N^{\frac{\lambda}{\lambda'}} \left(1 + \frac{1}{(1 - \lambda')^m} \int_0^\infty \lambda \exp(-(\lambda' - \lambda)t) dt \right) \\ &\leq CN, \end{aligned}$$

où C est une constante ne dépendant que de m . Si on suppose maintenant que $k = o(N^\alpha)$ pour $\alpha \in (0, 1]$, alors, d'après (3.6), $Z(N, k)/N^\beta$ converge vers 0 en probabilité dès que $\alpha < \beta$. Par suite, la limite en loi des fluctuations de $G_2(N, k)$ établie dans [Joh00a] est valable $G_1(N, k)$ lorsque $\alpha < 3/7$. En effet, si on note μ et σ^2 la moyenne et la variance de Y , d'après (3.6) montre que

$$\left| \frac{G_1(N, k) - (N + k - 1)m\mu - 2\sqrt{m}\sigma\sqrt{Nk}}{\sqrt{m}\sigma N^{1/2}k^{-1/6}} - \frac{G_2(mN, k) - (mN + k - 1)\mu - 2\sigma\sqrt{mNk}}{\sigma(mN)^{1/2}k^{-1/6}} \right|$$

converge en loi vers 0 dès que $\frac{1}{2} - \frac{\alpha}{6} > \alpha$.

Ce couplage est présenté dans ce chapitre à titre illustratif. Pour les fluctuations, il n'apporte aucun résultat nouveau ni sur la taille du rectangle ni sur la nature des poids. En effet, dans [BM05], les auteurs ont obtenu la majoration $\alpha < 3/7$ pour le LPP avec n'importe quel poids admettant des moments exponentiels finis.

De la même manière, on peut établir des propriétés asymptotiques et non-asymptotiques de grandes déviations de $G_1(N, k)$. A titre d'exemple, on regarde les grandes déviations à droite de la moyenne. Pour tout $\varepsilon_1 < \varepsilon/\sqrt{m}$, on a que

$$\begin{aligned} \mathbb{P}[A] &:= \mathbb{P}[G_1(N, k) - m\mu(N + k - 1) \geq 2\sqrt{Nk}(\sqrt{m}\sigma + \varepsilon)] \\ &\leq \mathbb{P}[G_2(mN, k) - \mu(mN + k - 1) \geq 2\sqrt{mNk}(\sigma + \varepsilon_1)] \\ &\quad + \mathbb{P}[|G_1(N, k) - G_2(N, k)| \geq (m - 1)k\mu + 2\sqrt{Nk}(\varepsilon - \sqrt{m}\varepsilon_1)], \end{aligned}$$

et que

$$\begin{aligned} \mathbb{P}[A] &\geq \mathbb{P}[G_2(mN, k) - \mu(mN + k - 1) \geq 2\sqrt{mNk}(\sigma + 2\frac{\varepsilon}{\sqrt{m}} - \varepsilon_1)] \\ &\quad - \mathbb{P}[|G_1(N, k) - G_2(N, k)| \geq (m - 1)k\mu + 2\sqrt{Nk}(\frac{\varepsilon}{\sqrt{m}} - \varepsilon_1)]. \end{aligned}$$

On applique de nouveau (3.6) avec cette fois-ci $\beta = (1 + \alpha)/2$ et on obtient, pour tout $\alpha < 1$,

$$\lim_{N \rightarrow \infty} \frac{1}{k} \log \mathbb{P}[G_1(N, k) - m\mu(N + k - 1) \geq 2\sqrt{Nk}(\sqrt{m}\sigma + \varepsilon)] = J_{\text{GUE}}(\varepsilon).$$

Plus tard dans ce chapitre, on reparlera des propriétés de grandes déviations pour le LPP avec des poids quelconques. Dans ce cas, il est difficile de construire un couplage direct comme c'était le cas avec les poids Gamma. On utilisera à la place le plongement de Skorokhod et l'approximation K-M-T qu'on introduit dans le paragraphe qui suit.

3.1.2 Le théorème de plongement de Skorokhod et l'approximation K-M-T

Le théorème de plongement de Skorokhod fut pendant près d'une décennie le meilleur outil pour approximer une somme de variables aléatoires *i.i.d.* par l'intermédiaire du Mouvement Brownien. Cette méthode a débuté avec les travaux de Shorokhod [Sko65] et de Strassen [Str67]. Etant donnée une loi μ de variance finie, on est à la recherche d'un mouvement Brownien B_t et d'un temps d'arrêt τ , construits sur le même espace de probabilité que celui de μ , tels que μ soit égale en loi à B_τ . Au fil des années, les constructions du temps d'arrêt τ se sont succédées dans le but d'améliorer ses conditions de moment. On cite notamment les constructions de Root, Monroe, Azema-Yor etc. D'autres auteurs ont étendu le résultat en utilisant des processus de Markov autres que le Brownien. Dans [AY79] par exemple, Azema et Yor ont réussi à établir un plongement pour n'importe quelle martingale continue. Le lecteur est invité à consulter l'aperçu historique complet de Obló, [Obł04].

Theorem 3.1 (Skorokhod). *Soit $(B_t)_{t \geq 0}$ un mouvement Brownien et X une variable aléatoire réelle telle que $\mathbb{E}X = 0$ et $\mathbb{E}X^2 = 1$. Alors il existe un temps d'arrêt T tel que $B_T \stackrel{d}{=} X$ et $\mathbb{E}T = 1$.*

On considère maintenant la suite $(X_n)_{n \geq 1}$ de variables aléatoires *i.i.d.* de variance finie et on note $S_k = \sum_{i=1}^k X_i$ la somme partielle correspondante. On cherche à contruire sur le même espace de probabilité, une copie de la suite (X_n) et un mouvement Brownien $(B_t)_{t \geq 0}$ tels que

$$\max_{1 \leq i \leq n} |S_i - B_i| \quad (3.7)$$

soit le plus petit possible au sens des probabilités. Avec la technique du plongement de Skorokhod, si X_1 admet un moment fini d'ordre $p \geq 4$, on démontre que dans les meilleurs des cas, (3.7) se comporte à l'infini comme $n^{1/4}(\log n)^{1/2}(\log \log n)^{1/4}$.

En fait, il existe, d'après le théorème de Skorokhod, une suite de temps d'arrêt τ_1, \dots, τ_n telle que

$$(3.7) \stackrel{\mathcal{L}}{=} \max_{k \leq n} |B_{\tau_1 + \dots + \tau_k} - B_k|. \quad (3.8)$$

L'évaluation de (3.8) revient à comparer la somme partielle des τ_i et k . Pour cela, on a besoin d'une ample connaissance des temps d'arrêt et en particulier, du lien entre les moments des τ_i et ceux des X_i . Le théorème suivant de Sawyer [Saw72] nous fournit cette information lorsque X_1 admet des moments de type exponentiels.

Théorème 3.2 (Sawyer [Saw72]). *Soit X une variable aléatoire centrée vérifiant*

$$\mathbb{E} \exp(\mu|X|^\gamma) < +\infty$$

pour un certain réel $\gamma > 0$ et $\mu > 0$. On note τ le temps d'arrêt correspondant à la représentation de Skorokhod. Si on note $\theta = \gamma/(2 + \gamma)$ et $\nu = \mu^{1-\theta}$, alors on a que

$$\mathbb{E} \exp(\nu\tau^\theta) \leq \Phi_\gamma \mathbb{E} \exp(\mu|X|^\gamma),$$

où Φ_γ est une constante positive qui ne dépend que de γ .

Ce théorème montre une perte éventuelle des moments finis en passant de la variable X au temps d'arrêt correspondant. Supposons par exemple que $\gamma = 1$, i.e. X_1 admet un moment exponentiel fini au voisinage de 0. D'après le Théorème 3.2, le temps d'arrêt τ_1 correspondant admet un moment exponentiel d'ordre $1/3$. En utilisant la loi du supremum du mouvement Brownien, l'inégalité de Fuk-Nagaev [FN71] et l'inégalité de Markov, on obtient pour tous x, y et z réels positifs, que

$$\mathbb{P}\left[\max_{k \leq n} |S_k - B_k| \geq x\right] \leq CNe^{-x^2/y} + Ce^{-y^2/(n+yz)} + CNe^{-z^{1/3}}.$$

Pour que cette probabilité converge vers 0, il faut au moins que y^2 converge vers l'infini plus vite que n et que x^2/y converge vers l'infini. Par suite, on est contraint de choisir x de l'ordre de $n^{1/4}(\log n)^{1/2}(\log \log n)^{1/4}$ (au minimum). Cette analyse très concise nous donne une idée sur l'ordre de grandeur de (3.7) en utilisant la représentation de Skorokhod.

En 1975, quelques années après l'introduction par Skorokhod du théorème de plongement, Komlos, Major et Tusnady [KMT76] ont introduit la célèbre approximation qui porte leur nom : l'approximation K-M-T. Cette nouvelle approche n'utilise pas les temps d'arrêt mais plutôt une construction progressive d'une suite de variables Gaussiennes. Pour des X_i admettant une transformée de Laplace au voisinage de 0, les auteurs ont démontré qu'on peut reconstruire la suite (X_n) de manière à ce que (3.7) soit de l'ordre de $\log n$. De plus, ils ont démontré qu'on ne peut pas faire mieux. Plus tard, ces résultats ont été généralisés pour couvrir des variables non indépendantes et des processus empiriques.

Théorème 3.3 (Komlós-Major-Tusnady [KMT76]). *Si $\mathbb{E}X_1 = 0$ et $\mathbb{E}X_1^2 = 1$. On suppose qu'il existe $\mu_0 > 0$ tel que pour tout $\mu < \mu_0$,*

$$\mathbb{E} \exp(\mu|X_1|) < +\infty.$$

Alors, pour tout $N \geq 1$, il existe un couplage de la suite $(X_i)_{i \geq 1}$ et du mouvement Brownien $(B_t)_{t \geq 0}$ tel que pour tout $x > 0$,

$$\mathbb{P}\left[\max_{i=1,\dots,N} |S_i - B_i| > \Theta \log N + x\right] \leq C \exp(-\theta x).$$

Les constantes positives Θ , C and θ ne dépendent que de la loi de X_1 . De plus, on peut choisir θ aussi grande que l'on veut en choisissant un Θ aussi grand que nécessaire.

Dans le même article, les auteurs établissent un résultat similaire lorsque X admet des moments finis d'ordre $p > 2$. La décroissance de l'erreur est dans ce cas polynomial.

3.1.3 Un théorème d'invariance faible de Chatterjee

L'approximation K-M-T est dite forte car elle offre une majoration en probabilité de l'erreur. Cependant, cette approximation ne concerne que les sommes partielles et s'adapte difficilement aux changements de situation. Il existe dans la littérature des approximations faibles (en loi) qui fournissent des bornes supérieures de l'erreur moins attractives certes mais qui sont plus faciles à manipuler. On cite en particulier un théorème d'approximation du à Chatterjee, [Cha05, Sui06]. Bien que la majoration de l'erreur soit assez importante, ce résultat offre un champ d'application plus vaste puisqu'il ne concerne pas que le maximum sur les sommes partielles de variables *i.i.d.* mais plutôt le maximum (sur un ensemble fini) de fonctions (des X_i) trois fois différentiables. Suidan utilise cette approximation faible dans [Sui06] pour démontrer la convergence vers la loi de Tracy-Widom des fluctuations du temps de dernier passage sur un rectangle fin. Il fournit ainsi une preuve alternative à celles données par Baik-Suidan et Bodineau-Martin. On reviendra en détails sur le principe d'invariance de Chatterjee et la méthode de Suidan dans le Paragraphe 5.5 où il sera question de matrices aléatoires.

3.2 Fluctuations longitudinales et transversales

On s'intéresse dans ce paragraphe à deux propriétés essentielles des modèles de percolation dirigée de dernier passage : la convergence des fluctuations longitudinales et celle des fluctuations transversales. Pour le LPP, les fluctuations longitudinales sont en quelque sorte les déviations standards du temps de dernier passage $G(N, k)$ tandis que les fluctuations transversales décrivent les déviations des chemins maximisants par rapport à un chemin médian.

Ces deux types de fluctuations sont souvent représentées par deux réels positifs χ et ξ . χ indique l'ordre de grandeur des fluctuations longitudinales tandis que ξ fournit la largeur minimale d'un cylindre contenant tous les chemins maximisants. On considère par exemple le LPP avec des variables géométriques. D'après le Théorème 2.7, l'écart entre $G(k, k)$ est sa limite hydrodynamique est de l'ordre de $k^{1/3}$ et donc $\chi = 1/3$. D'autre part, une conséquence des Théorèmes 2.8 et 2.7 est qu'un chemin maximisant de $G(k, k)$ est inclus, avec probabilité allant vers 1, dans un cylindre de largeur $2k^{2/3}$ et donc $\xi = 2/3$, (voir [Joh00b, BDM⁺01]). Ces résultats restent-ils vrais en dehors des poids géométriques ?

Il est conjecturé que pour les modèles de percolation dirigée en dimension d , les deux réels χ et ξ satisfont la relation $2\xi - 1 = \chi$. Cette conjecture, avancée par Krug et Spohn dans [KS91], est justifiée par des arguments liés à la géométrie de la boîte dans laquelle évoluent les chemins dirigés, et est valable lorsque les dimensions de la boîte sont du même ordre de grandeur. Lorsque les dimensions de la boîte ne sont pas du même ordre, cette relation n'est plus exacte. En dimension 2 par exemple, on démontre que si $k = o(N^\alpha)$ avec α convenablement choisi, alors χ et ξ satisfont l'égalité

$$2\xi - \frac{3\alpha}{2} + \frac{1}{2} = \chi. \quad (3.9)$$

La preuve de (3.9) repose essentiellement sur les arguments avancés par Johansson dans [Joh00b]. Elle sera développée dans le Paragraphe 3.2.3. Dans le paragraphe suivant, on donne quelques arguments heuristiques pour comprendre l'égalité dans (3.9).

3.2.1 Arguments heuristiques

On se place dans le cadre du LPP muni de poids *i.i.d.* quelconques et on suppose que $k = N^\alpha$ avec $0 < \alpha < 1$. De plus, on suppose que les fluctuations longitudinales de $G(N, k)$ sont de l'ordre de N^ξ et qu'elles convergent en loi vers une variable Θ , *i.e.*

$$\frac{G(N, k) - (N + k)\mu - 2\sigma\sqrt{Nk}}{N^\chi} \xrightarrow{\mathcal{L}} \Theta. \quad (3.10)$$

Lorsque N est suffisamment grand, (3.10) montre que

$$G(N, k) \approx (N + k)\mu + 2\sigma\sqrt{Nk} + N^\chi\Theta \quad \text{et} \quad \mathbb{E}G(N, k) \approx (N + k)\mu + 2\sigma\sqrt{Nk}. \quad (3.11)$$

On trace autour de la diagonale du rectangle, un cylindre de largeur $2N^\delta$ et on cherche le plus petit δ tel que tous les chemins maximisants soient inclus dedans. On suppose qu'il existe un chemin maximisant qui sort du cylindre par un point quelconque noté S , (voir Figure 3.1). On suppose que S se trouve du côté supérieur du cylindre. Le cas contraire est traité de la même manière par symétrie. Le point S admet alors pour coordonnées $(N^b - N^{\delta+1-\alpha}, N^b + N^\delta)$ pour un certain $0 < b < 1$. On note $G_S(N, k)$ le temps de dernier passage restreint au chemin passant par S . Par conséquent, on a que $\mathbb{E}G_S(N, k)$ est de l'ordre de

$$(N + k)\mu + 2\sigma\left(\sqrt{(N^b - N^{\delta+1-\alpha})(N^b + N^\delta)} + \sqrt{(1 - N^b + N^{\delta+1-\alpha})(1 - N^b + N^\delta)}\right).$$

D'après (3.11), l'ordre de $|\mathbb{E}G_S(N, k) - \mathbb{E}G(N, k)|$ ne doit pas excéder N^χ . Cette différence vaut

$$2\sigma\sqrt{Nk} - 2\sigma\left(\sqrt{(N^b - N^{\delta+1-\alpha})(N^b + N^\delta)} + \sqrt{(1 - N^b + N^{\delta+1-\alpha})(1 - N^b + N^\delta)}\right).$$

Il se trouve que cette quantité est également la différence entre la racine carré de la surface du grand rectangle et la somme des racines carrés des aires des deux petits rectangles, (voir Figure 3.1). Cette différence est de l'ordre de $N^{2\delta-3\alpha/2+1/2}$ d'où le résultat.

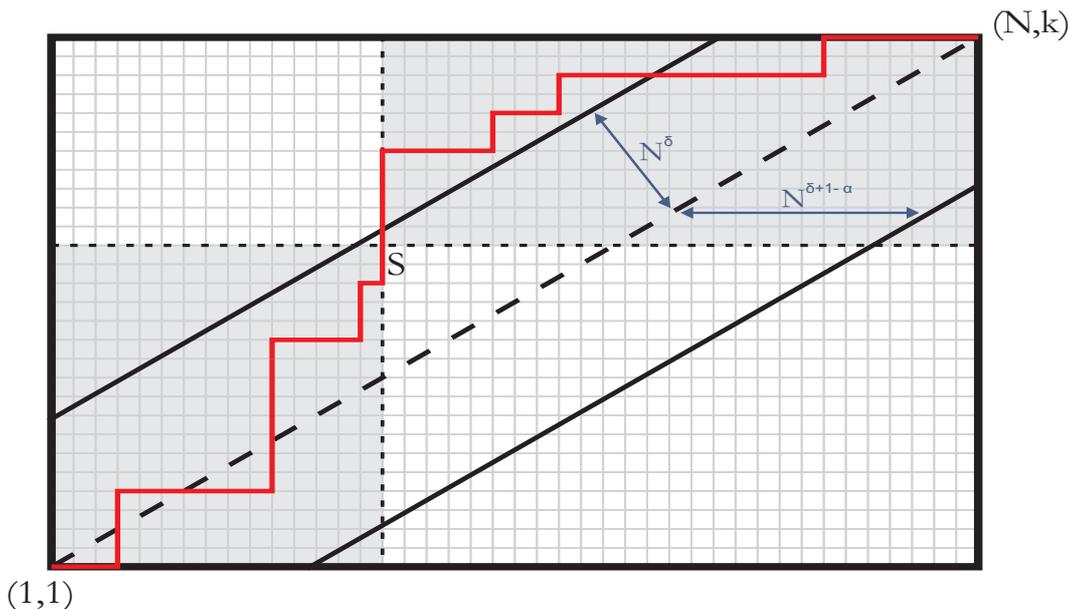


FIG. 3.1 – *Chemin sortant du cylindre*

Remarque 3.1. Ces arguments ne constituent pas une preuve rigoureuse de (3.9). Cependant, si on rajoute des hypothèses sur la variable Θ et si $G(N, k)$ satisfait des inégalités exponentielles de grandes déviations, on peut établir une preuve formelle de (3.9) sans pour autant connaître les valeurs des coefficients χ et ξ .

Remarque 3.2. Notons ici que pour des variables admettant des queues lourdes, on ne peut pas espérer contenir les chemins maximisants dans un cylindre fin autour de la diagonale. Le lecteur pourra consulter l'article de Hambly et Martin [HM07] pour une étude complète du sujet.

3.2.2 Fluctuations longitudinales et inégalités exponentielles

Dans cette partie, on reprend dans un premier temps les grandes lignes de la démonstration du Théorème 3.3 donnée dans [BM05]. Les techniques avancées dans cet article sont au cœur de toutes les preuves de ce chapitre et du chapitre suivant. Ensuite, on utilise ces techniques pour démontrer des inégalités exponentielles à droite et à gauche de la moyenne de $G(N, k)$. On aura ainsi les outils nécessaires pour aborder d'une manière rigoureuse la question des fluctuations transversales.

On se place alors dans le cadre du LPP sur un rectangle fin et on suppose donc que k tend vers l'infini et que $k = o(N^\alpha)$ avec $0 < \alpha < 1$. On rappelle le théorème de Bodineau-Martin et Baik-Suidan et on le démontre dans la suite.

Théorème 3.3 (Bodineau-Martin [BM05] Baik-Suidan [BS05]). *On suppose que les $X_i^{(j)}$ sont i.i.d. et que $\mathbb{E}X = 0$ et $\mathbb{E}X^2 = 1$. De plus, on suppose qu'il existe $p > 2$ tel que $\mathbb{E}|X|^p < \infty$. Si $k = o(N^\alpha)$ avec $0 < \alpha < \frac{6}{7}(\frac{1}{2} - \frac{1}{p})$, on a que*

$$\Theta_k := \frac{G(N, k) - 2\sqrt{Nk}}{N^{1/2}k^{-1/6}} \xrightarrow[N, k \rightarrow \infty]{\mathcal{L}} \Theta,$$

où Θ est une variable de loi de Tracy-Widom de fonction de répartition F_2^{TW} .

L'idée de la preuve est similaire à ce qu'on a présenté dans le Paragraphe (3.1.1) pour les variables de loi Gamma. Néanmoins, le couplage ici n'est pas immédiat et nécessite l'utilisation de nouveaux outils de comparaison. On donne dans la suite les grandes lignes de cette preuve et on envoie le lecteur vers le chapitre suivant pour les calculs détaillés.

Démonstration : Vu le Théorème 1.12, on cherche à démontrer que

$$\mathbb{P}[|G(N, k) - L(N, k)| \geq \varepsilon N^{\frac{1}{2}} k^{-\frac{1}{6}}] \longrightarrow 0$$

pour tout $\varepsilon > 0$. D'après les calculs faits dans [BM05] et repris dans le paragraphe 4.2, on a pour tout $\varepsilon > 0, \beta > 0$ et $N, k \geq 1$,

$$\mathbb{P}[|G(N, k) - L(N, k)| \geq \varepsilon N^\beta] \leq \mathbb{P}[Y_k \geq \varepsilon N^\beta / 2] + \mathbb{P}[Z_k \geq \varepsilon N^\beta / 2], \quad (3.12)$$

où

$$Y_k = 2 \sum_{r=1}^k \left(\max_{i=1, \dots, N} |S_i^{(r)} - B_i^{(r)}| \right)$$

et

$$Z_k = 2 \sum_{r=1}^k \left(\sup_{\substack{0 \leq s, t \leq N \\ |s-t| < 2}} |B_s^{(r)} - B_t^{(r)}| \right).$$

Grâce au principe de réflexion du mouvement Brownien, on démontre (voir (4.17), (4.18) et (4.19)) qu'il existe deux constantes $c > 0$ et $C > 0$ telles que, pour tout $\varepsilon > 0$ et tout $N, k \geq 1$,

$$\mathbb{P}[Z_k \geq \varepsilon N^\beta / 2] \leq CN^{1+\alpha} e^{-(c\varepsilon N^\beta - k \log N)} \quad (3.13)$$

Pour contrôler Y_k , on utilise une version améliorée de l'approximation K-M-T pour des variables i.i.d. admettant des moments finis d'ordre $p \geq 2$.

Théorème 3.4 (Sakhanenko [Sak91]). *Soit X_1, \dots, X_N, \dots une suite de variables aléatoires i.i.d., centrées et réduites telle que $\mathbb{E}|X_1|^p < \infty$ avec $p \geq 2$. Soit le mouvement Brownien $(B_t)_{t \geq 0}$ défini sur le même espace probabilisé que $(X_N)_N$. Alors on peut trouver des copies de (X_N) et de (B_t) telles que pour tout $m \leq p$,*

$$\mathbb{E} \left[\max_{i=1, \dots, N} |S_i - B_i|^m \right] \leq (Cm)^m \mathbb{E}[|X_1|^m] N,$$

où C est une constante qui ne dépend pas de N .

Ce théorème, associé à l'inégalité de Markov, montre que

$$\begin{aligned}
\mathbb{P}[Y_k \geq \varepsilon N^\beta / 2] &\leq \frac{4^p \mathbb{E} \left[\left(\sum_{r=1}^k \max_{i=1, \dots, N} |S_i^{(1)} - B_i^{(1)}| \right)^p \right]}{\varepsilon^p N^{p\beta}} \\
&\leq \frac{4^p N^{p\alpha} \mathbb{E} \left[\max_{i=1, \dots, N} |S_i^{(1)} - B_i^{(1)}|^p \right]}{\varepsilon^p N^{p\beta}} \\
&\leq \frac{CN^{p\alpha+1-p\beta}}{\varepsilon^p}, \tag{3.14}
\end{aligned}$$

où C est une constante positive qui dépend de p . Pour terminer, on choisit $\beta = \frac{1}{2} - \frac{\alpha}{6}$. Par suite, pour que le terme de droite de l'inégalité (3.14) converge vers 0, il faut que $\alpha < \frac{6}{7}(\frac{1}{2} - \frac{1}{p})$ \square

Il est conjecturé que ce résultat tient pour $\alpha = 1$. Cependant, l'erreur due au plongement dans les trajectoires Browniennes est au mieux d'ordre k . Il faudra donc trouver d'autres approches pour démontrer cette conjecture. On pense notamment aux méthodes de Tao et Vu [TV10, TV] pour démontrer l'universalité des fluctuations pour la plus grande valeur propre d'une matrice de Wigner. Tao et Vu utilisent les formules variationnelles de Hadamard comme outil de comparaison. Bien que la structure du LPP soit différente de celle d'une matrice aléatoire, il est légitime de se demander s'il n'existe pas des formules variationnelles analogues pour le LPP.

Venons-en maintenant aux inégalités exponentielles à gauche et à droite de la moyenne. On dit que ces inégalités sont optimales si les bornes obtenues sont les mêmes que les queues de la loi de Tracy-Widom. Dans cette partie, on ne cherche pas à démontrer des inégalités optimales mais plutôt des inégalités suffisamment décroissantes pour établir la convergence des moments de Θ_k et pour pouvoir calculer le coefficient ξ . Dans le Paragraphe 3.3.3, on refait une étude similaire dans le but d'améliorer les inégalités obtenues ici et atteindre donc les bornes des queues de la loi de Tracy-Widom.

Grâce au plongement dans les trajectoires Browniennes et aux liens avec le GUE, le théorème suivant sera très utile dans la suite. Il fournit deux inégalités optimales (à gauche et à droite de la moyenne) pour la plus grande valeur propre d'une matrice du GUE.

Théorème 3.5 (Ledoux-Rider [LR10]). *Pour tout $k \geq 1$ et $0 < \varepsilon \leq 2k^{2/3}$,*

$$\mathbb{P}[\lambda_{max} - 2\sqrt{k} \geq \varepsilon k^{-1/6}] \leq C e^{-\frac{\varepsilon^{3/2}}{\sqrt{2C}}}$$

et

$$\mathbb{P}[\lambda_{max} - 2\sqrt{k} \leq -\varepsilon k^{-1/6}] \leq C^2 e^{-\frac{\varepsilon^3}{4C}},$$

où C est une constante qui ne dépend ni de k ni de ε .

Notons ici que pour $\varepsilon > 2k^{2/3}$, on peut toujours obtenir des inégalités non-asymptotiques avec des bornes exponentielles comme dans le Théorème 3.5. Par contre, on n'a plus les exposants des queues de la loi de Tracy-Widom ε^3 et $\varepsilon^{3/2}$ mais plutôt un exposant en ε^2 . En effet, pour tout $r \geq 0$ et $N \geq 1$,

$$\mathbb{P}[|\lambda_{\max} - \mathbb{E}\lambda_{\max}| > 2\sqrt{kr}] \leq 2e^{-2kr^2}. \quad (3.15)$$

Ce résultat ne provient pas particulièrement de la structure du GUE, mais plutôt du fait que λ_{\max} est une fonction Lipschitz de variables Gaussiennes complexes indépendantes. En effet, on a que

$$\lambda_{\max} = \sup_{|u|=1} uXu^*,$$

où X est une matrice $k \times k$ du GUE. Le lecteur pourra consulter le chapitre 4 de [Led07]. Voici les arguments qui démontrent (3.15). A droite de la moyenne, on sait d'après [Led07] que pour $r \leq 1$,

$$\mathbb{P}[\lambda_{\max} - 2\sqrt{k} > 2\sqrt{kr}] \leq Ce^{-ckr^2}.$$

Ici $r \geq 1$ correspond à $\varepsilon \geq 2k^{2/3}$ dans le théorème 3.5. A gauche de la moyenne, il suffit d'utiliser (3.15) en remarquant que pour tout $r \geq 0$,

$$\begin{aligned} \mathbb{P}[|\lambda_{\max} - \mathbb{E}\lambda_{\max}| > 2\sqrt{kr}] &\geq \mathbb{P}[\lambda_{\max} - 2\sqrt{k} < -2\sqrt{kr} - 2\sqrt{k} + \mathbb{E}\lambda_{\max}] \\ &\geq \mathbb{P}[\lambda_{\max} - 2\sqrt{k} < -2\sqrt{k}(r+1)]. \end{aligned}$$

Par suite, pour tout $r \geq 1$,

$$\mathbb{P}[\lambda_{\max} - 2\sqrt{k} < -2\sqrt{kr}] \leq Ce^{-ckr^2}. \quad (3.16)$$

On remarque alors que pour $\varepsilon \geq 2k^{2/3}$, on a des inégalités avec des bornes Gaussiennes plutôt que celles de Tracy-Widom.

Dans la suite, on reprend les grandes lignes du calcul fait dans le Paragraphe 4.5. On suppose d'abord que $\mathbb{E}X = 0$ et $\mathbb{E}X^2 = 1$, alors pour tout $\varepsilon > 0$, on a que

$$\begin{aligned} &\mathbb{P}[G(N, k) - 2\sqrt{Nk} \geq \varepsilon N^{1/2}k^{-1/6}] \\ &\leq \mathbb{P}[L(1, k) - 2\sqrt{k} \geq \frac{\varepsilon}{2}k^{-1/6}] + \mathbb{P}[|G(N, k) - L(N, k)| \geq \frac{\varepsilon}{2}N^{1/2}k^{-1/6}] \end{aligned} \quad (3.17)$$

et

$$\begin{aligned} &\mathbb{P}[G(N, k) - 2\sqrt{Nk} \leq -\varepsilon N^{1/2}k^{-1/6}] \\ &\leq \mathbb{P}[L(1, k) - 2\sqrt{k} \leq -\frac{\varepsilon}{2}k^{-1/6}] + \mathbb{P}[|G(N, k) - L(N, k)| \geq \frac{\varepsilon}{2}N^{1/2}k^{-1/6}] \end{aligned} \quad (3.18)$$

Pour des inégalités non-asymptotiques, il est préférable de prendre $k = \lfloor N^\alpha \rfloor$ au lieu de $k = o(N^\alpha)$. Sinon, les éventuelles constantes multipliant la borne supérieure dépendront de l'allure de $k = k(N)$ ce qui est à éviter. Ici on distinguera le cas où les poids admettent des moments exponentiels et le cas où on a simplement des moments finis d'ordre $p > 2$. Dans le premier cas, on a la proposition suivante :

Proposition 3.6. *On suppose que $\mathbb{E}X = 0$, $\mathbb{E}X^2 = 1$ et que $\mathbb{E}[\exp(\mu|X|)] < +\infty$ pour un réel $\mu > 0$. Si de plus $k = \lfloor N^\alpha \rfloor$ avec $\alpha < \frac{3}{7}$, alors pour tout $N > 0$ et pour tout $\varepsilon > 1$, on a que*

$$\mathbb{P}[G(N, k) - 2\sqrt{Nk} \leq -\varepsilon N^{1/2}k^{-1/6}] \leq Ce^{-c\varepsilon},$$

et

$$\mathbb{P}[G(N, k) - 2\sqrt{Nk} \geq \varepsilon N^{1/2}k^{-1/6}] \leq Ce^{-c\varepsilon},$$

où c et C sont des constantes positives indépendantes de k et ε .

Démonstration : Grâce aux inégalités (3.17) et (3.18), démontrer la Proposition 3.6 revient à contrôler suffisamment la taille de $|G(N, k) - L(N, k)|$. Avec les mêmes notations de la preuve du Théorème 3.3, l'inégalité (3.13) nous donne une borne exponentielle de Z_k . Pour Y_k , on utilise la version de l'approximation K-M-T du Théorème 3.3 et on refait les mêmes calculs que ceux du Paragraphe 4.3 (voir (4.33)). Par suite, pour tout $\varepsilon > 0$, il existe une constante $C > 0$, telle que

$$\mathbb{P}[Y_k \geq \frac{\varepsilon}{4}N^{1/2}k^{-1/6}] \leq C \exp\left(-\frac{\varepsilon N^{1/2}k^{-1/6} - k \log N}{C}\right).$$

Comme $\alpha < 3/7$ et $\varepsilon > 1$, la constante C peut être choisie indépendamment de ε , de N et de k . Le résultat découle donc immédiatement du Théorème 3.5 et l'inégalité (3.16). \square

On a un résultat similaire lorsque les variables admettent seulement des moments finis d'ordre $p > 2$. La preuve est presque identique à celle d'avant. Cependant, on utilise la version de l'approximation K-M-T donnée par le Théorème 3.4 de Sakhanenko.

Proposition 3.7. *On suppose que $\mathbb{E}X = 0$ et que $\mathbb{E}X^2 = 1$. Si $k = \lfloor N^\alpha \rfloor$, alors pour tout $N > 0$ et pour tout $\varepsilon > 0$, on a pour $\alpha < \frac{6}{7}(\frac{1}{2} - \frac{1}{p})$,*

$$\mathbb{P}[G(N, k) - 2\sqrt{Nk} \leq -\varepsilon N^{1/2}k^{-1/6}] \leq \frac{C}{\varepsilon^p},$$

et

$$\mathbb{P}[G(N, k) - 2\sqrt{Nk} \geq \varepsilon N^{1/2}k^{-1/6}] \leq \frac{C}{\varepsilon^p},$$

où C est une constante positive indépendante de N et ε .

La preuve de cette proposition est immédiate au vue de l'inégalité (3.14).

Maintenant, on a les outils nécessaires pour attaquer la question des fluctuations transversales.

3.2.3 Fluctuations transversales sur un rectangle fin

Dans [Joh00b], Johansson détermine l'ordre des fluctuations transversales pour la version Poissonisée de la plus longue sous suite croissante d'une permutation aléatoire. Dans cette partie, on suit les étapes de la preuve de Johansson pour réécrire rigoureusement les arguments avancés dans la section 3.2.1. La preuve repose essentiellement sur le théorème 3.5 et la proposition 3.6.

On se place donc dans le cadre du LPP sur un rectangle fin avec $k = \lfloor N^\alpha \rfloor$. Dans la suite, on considère que N^α est un entier et on remplace systématiquement k par N^α . On définit d'abord le cylindre de largeur N^δ autour de la diagonale d'équation $y = N^{\alpha-1}x$, par l'ensemble des points situés à une distance inférieure ou égale à N^δ de la diagonale :

$$C^\delta(N) = \{(x, y) \in \mathbb{R}^2; -N^\delta \leq y - N^{\alpha-1}x \leq N^\delta\}. \quad (3.19)$$

On peut supposer sans perte de généralité que la variable X est sans atome et donc on peut dire avec probabilité 1 qu'il existe un unique chemin qui maximise la somme dans (2.1). On note ce chemin $\pi_N^* = \pi^*$ et on définit le coefficient des fluctuations transversales ξ par le réel positif

$$\xi = \inf \{ \delta > 0; \liminf_N \mathbb{P}[(i, j) \in C^\delta(N); \forall (i, j) \in \pi^*] = 1 \}. \quad (3.20)$$

Les techniques utilisées ici nous permettent de déterminer ξ dans le cas où X admet une transformée de Laplace au voisinage de 0. Par contre, dans le cas où X admet seulement des moments d'ordre $p > 2$, on parvient juste à trouver une borne inférieure et une borne supérieure de ξ qui ne sont pas les mêmes. Cela est dû à l'écart polynomial entre $G(N, k)$ et $L(N, k)$ dans ce cas

On démontre alors les deux théorèmes suivants qui sont analogues au théorème 1 de [Joh00b].

Théorème 3.8. *On suppose que $\mathbb{E}|X|^p < \infty$ pour $p > 2$. Alors, si $k = N^\alpha$ avec $\alpha < \frac{6}{7}(\frac{1}{2} - \frac{1}{p})$, on a que*

$$\frac{2\alpha}{3} \leq \xi \leq \frac{2\alpha}{3} + \frac{\alpha}{2p}.$$

Il est probable que lorsque X admet seulement des moments d'ordre $p > 2$, on a l'égalité $\xi = 2\alpha/3$. Cependant, la méthode de comparaison qui repose sur le plongement dans le Brownien ne nous fournit pas les bornes nécessaires à cette fin. D'autres techniques sont donc recherchées. En examinant le Théorème 3.8, on voit bien que lorsque p tend vers l'infini, les deux bornes convergent vers la même valeur. Le théorème suivant concrétise ce fait.

Théorème 3.9. *On suppose qu'il existe $c_0 > 0$ tel que $\mathbb{E}e^{c|X|} < \infty$ pour tout $0 < c < c_0$. Alors, si $k = N^\alpha$ avec $\alpha < \frac{3}{7}$, on a que*

$$\xi = \frac{2\alpha}{3}.$$

On peut étudier les fluctuations transversales d'une manière plus « faible ». Au lieu de considérer l'éloignement du chemin maximisant de la diagonale en tout point, on cherche à évaluer l'écart entre $\min \{j; (\lfloor N/2 \rfloor, j) \in \pi^*\}$ et $k/2$. Donc au lieu de contrôler le chemin maximisant en tout point, on voudrait simplement connaître son comportement à mi-parcours.

Corollaire 3.10. *On suppose qu'il existe $c_0 > 0$ tel que $\mathbb{E}e^{c|X|} < \infty$ pour tout $0 < c < c_0$. Alors, si $k = N^\alpha$ avec $\alpha < \frac{3}{7}$, on a que*

$$\xi' := \lim_{N \rightarrow \infty} \frac{\log \mathbb{E}[(R_N - \lfloor N^\alpha/2 \rfloor)^2]}{2 \log N} \leq \frac{2\alpha}{3},$$

où $R_N = \min \{j; (\lfloor N/2 \rfloor, j) \in \pi^*\}$.

Dans [BM05], les auteurs introduisent le coefficient ξ' et conjecturent qu'il est égal à $2\alpha/3$. Ici, on parviendra, en utilisant les techniques de la preuve de Johansson, à montrer que $\xi' \leq 2\alpha/3$. Par contre, on ne sait pas prouver l'inégalité dans l'autre sens.

Démonstration des Théorèmes 3.8 et 3.9 : On démontre l'égalité $\xi = 2\alpha/3$ en démontrant successivement les deux inégalités. Commençons d'abord par la preuve de l'inégalité $\xi \leq 2\alpha/3$. C'est cette inégalité là qui est la plus « importante » car elle permet de préciser la borne supérieure sur la largeur du cylindre.

Dire que $\xi \leq 2\alpha/3$ revient à dire que pour tout $\delta > 2\alpha/3$,

$$\lim_{N \rightarrow \infty} \mathbb{P}[\pi^* \subset C^\delta(N)] = 1.$$

On choisit alors $2\alpha/3 \leq \delta < \alpha$ et on essaie de démontrer que la probabilité pour que le chemin maximisant sorte du cylindre $C^\delta(N)$ converge vers 0. Si π^* sort du cylindre $C^\delta(N)$ par le bord supérieur (Figure 3.1), il est contraint à revenir dans le cylindre par ce même bord. On considère ainsi les points $s_j = (\lfloor jN^{1-\alpha} - N^{\delta+1-\alpha} \rfloor, j)$ pour $1 \leq j \leq N^\alpha$. Les s_j correspondent au dernier point de π^* avant qu'il ne revienne définitivement dans le cylindre. Si π^* sort maintenant par le bord inférieur, alors π^* croise celui-ci au point de coordonnées $(jN^{1-\alpha} + N^{\delta+1-\alpha}, j)$. On considère alors les points $t_j = (\lceil jN^{1-\alpha} + N^{\delta+1-\alpha} \rceil, j)$ qui correspondent au dernier point de π^* à appartenir au cylindre avant une première sortie.

Les s_j et les t_j sont situés sur le bord du cylindre et tout chemin qui sort du cylindre doit forcément passer par l'un de ces points. Ainsi, on peut affirmer que

$$\begin{aligned} \mathbb{P}[\pi^* \not\subset C^\delta(N)] &= \mathbb{P}\left[\bigcup_{j=1}^{N^\alpha} \{s_j \in \pi^*\} \cup \{t_j \in \pi^*\}\right] \\ &\leq \sum_{j=1}^{N^\alpha} (\mathbb{P}[s_j \in \pi^*] + \mathbb{P}[t_j \in \pi^*]) \\ &\leq 2 \sum_{j=1}^{N^\alpha} \mathbb{P}[s_j \in \pi^*]. \end{aligned} \tag{3.21}$$

Maintenant, on fait intervenir les inégalités non-asymptotiques pour montrer que la somme sur un chemin qui passe loin de la diagonale est assez petite. D'après (3.21), on a que

$$\begin{aligned}
\mathbb{P}[\pi^* \not\subseteq C^\delta(N)] &= 2 \sum_{j=1}^{N^\alpha} \mathbb{P}[\{s_j \in \pi^*\} \cap \{G(N, N^\alpha) - 2N^{\frac{1+\alpha}{2}} \leq -\varepsilon N^{\frac{1}{2}-\frac{\alpha}{6}}\}] \\
&+ \\
&2 \sum_{j=1}^{N^\alpha} \mathbb{P}[\{s_j \in \pi^*\} \cap \{G(N, N^\alpha) - 2N^{\frac{1+\alpha}{2}} \geq -\varepsilon N^{\frac{1}{2}-\frac{\alpha}{6}}\}].
\end{aligned} \tag{3.22}$$

En sachant que π^* passe par le point s_j , on a que $G(N, N^\alpha) = G((1, 1), s_j) + G(s_j, (N, N^\alpha))$. Alors (3.22) devient

$$\begin{aligned}
\mathbb{P}[\pi^* \not\subseteq C^\delta(N)] &= 2 \sum_{j=1}^{N^\alpha} \mathbb{P}[G(N, N^\alpha) - 2N^{\frac{1+\alpha}{2}} \leq -\varepsilon N^{\frac{1}{2}-\frac{\alpha}{6}}] \\
&+ \\
&2 \sum_{j=1}^{N^\alpha} \mathbb{P}[G((1, 1), s_j) + G(s_j, (N, N^\alpha)) - 2N^{\frac{1+\alpha}{2}} \geq -\varepsilon N^{\frac{1}{2}-\frac{\alpha}{6}}].
\end{aligned} \tag{3.23}$$

Comme les poids sont d'espérance nulle et de variance 1, on sait que $G((1, 1), s_j)$ se comporte en moyenne comme deux fois la racine carré de la surface du rectangle ayant pour sommets les points $(1, 1)$ et s_j . Donc il serait convenable de comparer $G((1, 1), s_j)$ et $G(s_j, (N, N^\alpha))$ à leur moyenne respective et non pas à $2N^{(1+\alpha)/2}$ qui représente deux fois la racine carré de la surface totale du rectangle. Ainsi, le calcul de la probabilité dans la deuxième partie du côté droit de l'inégalité (3.23) dépend fortement de la différence entre la surface du grand rectangle et les surfaces des deux petits rectangles comme le montre la Figure 3.1. Pour simplifier le calcul, on désigne par a_j la surface du rectangle dont les côtés sont parallèles aux axes et admettant $(1, 1)$ et s_j comme sommets opposés. De la même manière, on désigne par b_j le rectangle ayant s_j et (N, N^α) comme sommets opposés. Ainsi, on a que

$$a_j = j \lfloor jN^{1-\alpha} - N^{\delta+1-\alpha} \rfloor$$

et

$$b_j = (N^\alpha - j)(N - \lfloor jN^{1-\alpha} - N^{\delta+1-\alpha} \rfloor).$$

Le lemme géométrique suivant donne une borne supérieure de la différence entre les moyennes présumées et justifie l'égalité (3.9).

Lemme 3.11. *Pour tout entier $j \in [N^\delta, N^\alpha]$, on a que*

$$\sqrt{a_j} + \sqrt{b_j} - \frac{1}{2}\sqrt{N^{1+\alpha}} \leq -N^{2\delta-\frac{3\alpha}{2}+\frac{1}{2}}.$$

La preuve du Lemme 3.11 est similaire à la preuve du lemme 3.2 dans [Joh00b]. Elle consiste à trouver le maximum de la fonction

$$f(y) = \sqrt{y(yN^{1-\alpha} - N^{\delta+1-\alpha})} + \sqrt{(N^\alpha - y)(N - (yN^{1-\alpha} - N^{\delta+1-\alpha}))},$$

pour $y \in [N^\delta, N^\alpha]$. On démontre facilement en dérivant f que le maximum est atteint pour $y_0 = (N^\alpha + N^\delta)/2$. En insérant cette valeur dans la fonction f , on obtient que

$$f(y_0) - \sqrt{N^{1+\alpha}} = \sqrt{N^{1+\alpha}} \left(\sqrt{1 - N^{2\delta-2\alpha}} - 1 \right) \leq -\frac{1}{2} N^{2\delta - \frac{3\alpha}{2} + \frac{1}{2}}.$$

La dernière inégalité est due au fait que $\delta \leq \alpha$ et que $\sqrt{1-x} \leq 1 - \frac{x}{2}$ pour $x \in [0, 1]$. A partir de ce moment de la preuve, on distinguera deux cas : celui avec des variables admettant une transformée de Laplace finie et celui avec des variables admettant seulement des moments finis d'ordre $p > 2$. Commençons d'abord par le cas où la transformée de Laplace des poids est finie au voisinage de 0.

Dans ce cas, on choisit $\varepsilon = N^\eta$ avec $\eta \in (0, \delta - \frac{2\alpha}{3})$. En retranchant et en rajoutant les bonnes moyennes de $G((1, 1), s_j)$ et $G(s_j, (N, N^\alpha))$ dans (3.23) et en appliquant la Proposition 3.6 et le Lemme 3.11, on a pour tout N que

$$\begin{aligned} \mathbb{P}[\pi^* \notin C^\delta(N)] &\leq 2N^\alpha e^{-cN^\eta} \\ &+ \\ &2 \sum_{j=1}^{N^\alpha} \mathbb{P}[G((1, 1), s_j) - 2\sqrt{a_j} \geq \frac{1}{2} N^{\frac{1}{2} - \frac{\alpha}{6}} (N^\tau - N^\eta)] \\ &+ \\ &2 \sum_{j=1}^{N^\alpha} \mathbb{P}[G(s_j, (N, N^\alpha)) - 2\sqrt{b_j} \geq \frac{1}{2} N^{\frac{1}{2} - \frac{\alpha}{6}} (N^\tau - N^\eta)], \end{aligned} \tag{3.24}$$

où $\tau = 2\delta - \frac{3\alpha}{2} + \frac{1}{2} - (\frac{1}{2} - \frac{\alpha}{6}) = 2(\delta - \frac{2\alpha}{3})$. Pour pouvoir appliquer la Proposition 3.6 à $G((1, 1), s_j)$, il faut trouver l'ordre de grandeur des coordonnées du point s_j afin d'obtenir l'ordre des fluctuations. Pour cela, on pose $j = tN^\alpha = N^{-\nu}N^\alpha$ avec $\nu \in [0, \alpha - \delta]$. Ainsi, pour N assez grand et $\nu \in [0, \alpha - \delta - \rho]$ avec $\rho > 0$ assez petit, $s_j = (cN^{1-\nu}, N^{\alpha-\nu})$ où c est une constante assez petite. Donc, les fluctuations de $G((1, 1), s_j)$ sont plutôt de l'ordre de $N^{-\frac{\mu}{3}} N^{\frac{1}{2} - \frac{\alpha}{6}}$. Comme $\tau > \eta$, on obtient finalement, pour N suffisamment grand et pour $\nu \in [0, \alpha - \delta - \rho]$, que

$$\begin{aligned} &\mathbb{P}[G((1, 1), s_j) - 2\sqrt{a_j} \geq \frac{1}{2} N^{\frac{1}{2} - \frac{\alpha}{6}} (N^\tau - N^\eta)] \\ &\leq \mathbb{P}[G((1, 1), s_j) - 2\sqrt{a_j} \geq cN^{\tau + \frac{\mu}{3}} N^{\frac{1}{2} - \frac{\alpha}{6} - \frac{\mu}{3}}] \\ &\leq e^{-cN^\tau}. \end{aligned} \tag{3.25}$$

Si maintenant $\nu \in [\alpha - \delta - \rho, \alpha]$ avec ρ très petit, l'inégalité (3.25) reste valable. Pour cela, il suffit de considérer la somme de toutes les variables dans la surface a_j et d'appliquer les inégalités

classiques pour des sommes de variables indépendantes. Le raisonnement pour $G(s_j, (N, N^\alpha))$ est identique. Ainsi, (3.23) et (3.25) nous donnent que

$$\mathbb{P}[\pi^* \not\subseteq C^\delta(N)] \leq CN^\alpha e^{-cN^\eta} + CN^\alpha e^{-cN^\tau} \quad (3.26)$$

pour η et τ strictement positifs. Ainsi, on démontre que $\xi \leq 2\alpha/3$.

Venons maintenant au cas où les poids admettent simplement un moment fini d'ordre $p > 2$. La preuve est presque identique au cas précédent mise à part la modification suivante. Du fait de la convergence polynomiale vers 0 dans la proposition 3.7, on est contraint de choisir $\varepsilon = N^\eta$ avec $\eta \in (\frac{\alpha}{p}, 2(\delta - \frac{2\alpha}{3}))$ pour garantir la convergence vers 0 dans (3.26). Par suite, le côté droit de (3.26) converge vers 0 pour tout $\delta > \frac{2\alpha}{3} + \frac{\alpha}{2p}$.

La preuve de $\xi \geq \frac{2\alpha}{3}$ est identique à celle proposée par Johansson pour la plus longue sous-suite croissante. Elle est valable pour les deux cas de variables traitées ici et repose sur des arguments similaires à ceux avancés plus haut. Le lecteur est donc invité à consulter [Joh00b] pour la preuve complète. □

Démonstration du Corollaire 3.10 : D'abord, on écrit

$$\mathbb{E}[(R_N - \lfloor N^\alpha/2 \rfloor)^2] = \sum_{i=1}^{N^\alpha} (i - \lfloor N^\alpha/2 \rfloor)^2 \mathbb{P}[R_N = i].$$

On suit les mêmes arguments de la preuve précédente et notamment l'inégalité (3.26). On obtient alors que $\mathbb{P}[R_N = i] \leq \mathbb{P}[\pi^* \not\subseteq C^\delta(N)]$ et donc, décroît exponentiellement vers 0 pour tout $\delta \geq 2\alpha/3$ et tout i entre 1 et $(N^\alpha - N^\delta)/2$ et tout i entre $(N^\alpha + N^\delta)/2$ et N^α . Par conséquent,

$$\mathbb{E}[(R_N - \lfloor N^\alpha/2 \rfloor)^2] \leq CN^{4\alpha/3},$$

d'où le résultat. □

Remarque 3.12. il est important de noter que le logarithme du nombre de chemins dirigés inclus dans le cylindre de largeur $N^{2\alpha/3}$ est presque du même ordre que le logarithme du nombre total de chemins dirigés dans le rectangle. En fait, pour N assez grand et $0 < \alpha < 1$, la formule de Stirling nous donne

$$\log \#\Pi(N, N^\alpha) \sim N^\alpha \log(N^{1-\alpha}).$$

Dénombrons maintenant les chemins dirigés inclus dans $C^{2\alpha/3}(N)$. En observant la Figure 3.2, on remarque que le nombre de chemins contenus dans $C^{2\alpha/3}(N)$ est strictement supérieur au nombre de chemins qui sont contraints à rester dans tous les triangles rectangles isocèles (en vert). On note B_N ce nombre et on dit que B_N est égal au nombre de chemins contenus dans l'un des triangles rectangles isocèles élevé à la puissance $N^{2\alpha/3}$. Or, le nombre de chemins dirigés dans un triangle rectangle isocèle de côté n est égal au nombre de Catalan $C_n = \frac{1}{n+1} \binom{2n}{n}$. Par conséquent,

$$\log B_N = N^{\frac{\alpha}{3}} \log C_{N^{\frac{2\alpha}{3}}} = N^{\frac{\alpha}{3}} \log \frac{(2N^{\frac{2\alpha}{3}})!}{(N^{\frac{2\alpha}{3}})!(N^{\frac{2\alpha}{3}} + 1)!} \stackrel{N}{\sim} 2N^\alpha \log 2.$$

Cela veut dire que la restriction sur $C^{2\alpha/3}(N)$ de la maximisation sur les chemins ne facilite pas vraiment l'étude de $G(N, k)$.

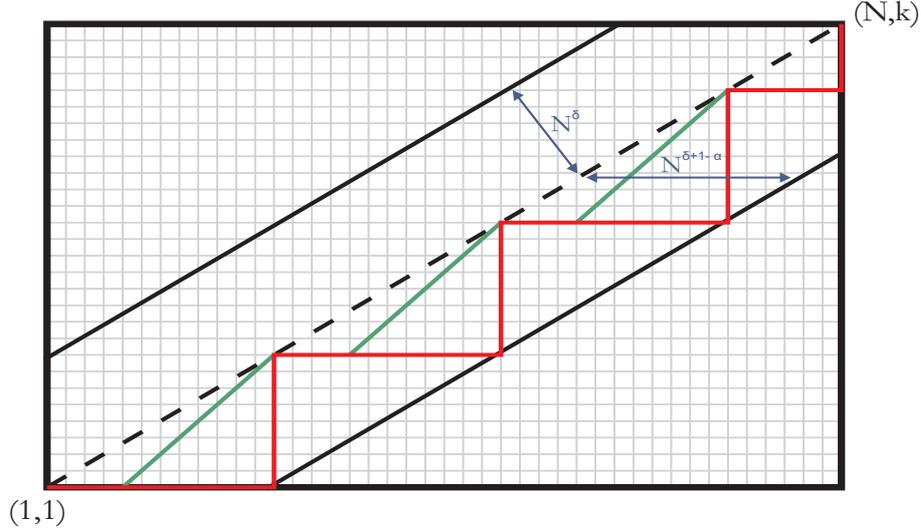


FIG. 3.2 – Chemin sortant du cylindre

3.2.4 Convergence des moments

Dans le paragraphe précédent, on a énoncé le résultat de [BM05, BS05] sur la convergence en loi des fluctuations longitudinales vers la loi de Tracy-Widom. Mais qu'est-ce qu'il en est de la convergence des moments des fluctuations? En fait, le Théorème 3.3 ne suffit pas pour affirmer que les moments de ces fluctuations convergent vers ceux de la loi de Tracy-Widom. Pour démontrer la convergence des moments, on a besoin de contrôler suffisamment les queues des fluctuations à N et k fixés. La Proposition 3.7 a été démontrée dans ce but.

Baik, Deift et Johansson ont démontré dans [BDJ99] la convergence des moments des fluctuations de la plus longue sous-suite croissante d'une permutation aléatoire. Plus tard, les auteurs de [BDM⁺01] ont démontré un résultat similaire pour le temps de dernier passage avec des poids géométriques et pour N et k proportionnelles. Grâce aux inégalités non-asymptotiques de Ledoux et Rider (Théorème 3.5), on peut obtenir un résultat similaire pour la valeur propre maximale de plusieurs types de matrices aléatoires. L'obtention d'inégalités optimales pour les variables citées plus haut nécessite une analyse plus ou moins compliquée de la loi de ces variables. Les techniques utilisées dans ce but n'ont pas fait l'objet d'étude durant cette thèse.

On se place toujours dans le cadre d'un LPP sur un rectangle fin avec des poids de moyenne 0 et de variance 1. La proposition suivante découle d'une simple intégration par parties et de l'application de la Proposition 3.7.

Proposition 3.13. *On suppose que $\mathbb{E}X = \mu$, $\mathbb{E}X^2 = \sigma^2$ et que $\mathbb{E}[|X|^p] < +\infty$ pour $p > 2$. Alors, si $k = N^\alpha$, on a pour tout entier positif $m < p$,*

$$\lim_{N \rightarrow \infty} \mathbb{E}[\Theta_N^m] := \lim_{N \rightarrow \infty} \mathbb{E} \left[\left(\frac{G(N, k) - 2\sqrt{Nk}}{N^{\frac{1}{2}} k^{-\frac{1}{6}}} \right)^m \right] = \mathbb{E}[\Theta^m]$$

pour tout $\alpha < \frac{6}{7}(\frac{1}{2} - \frac{1}{p})$, où Θ suit une loi de Tracy-Widom de fonction de répartition F_2^{TW} .

Démonstration : En appliquant la Proposition 3.7, on a par simple intégration par parties, que

$$\begin{aligned} \mathbb{E}[|\Theta_N^m|] &\leq \int_0^\infty \mathbb{P}[\Theta_N^m \geq s] ds + \int_0^\infty \mathbb{P}[\Theta_N^m \leq -s] ds \\ &\leq \int_0^\infty mt^{m-1} \mathbb{P}[\Theta_N \geq t] dt + \int_0^\infty mt^{m-1} \mathbb{P}[\Theta_N \leq -t] dt \\ &\leq 2 + 2 \int_1^\infty mt^{m-1} \frac{C}{t^p} dt. \end{aligned}$$

Par suite, $\mathbb{E}[|\Theta_N^m|]$ est uniformément bornée lorsque $m < p$. Le théorème de convergence dominée nous permet de conclure. □

Si on suppose que X admet une transformée de Laplace finie au voisinage de 0, on peut démontrer de la même manière que la transformée de Laplace de Θ_k converge dans un voisinage de 0 vers celle de la loi de Tracy-Widom. On utilise pour cela la Proposition 3.6.

3.3 Grandes déviations

Dans ce paragraphe, on étudie les propriétés des grandes déviations liées au LPP sur un rectangle fin. On ne fournit pas de preuves complètes puisque tous les calculs sont détaillés, sous forme d'une publication, dans le chapitre suivant. Rappelons encore une fois que $N = N_k$ dépend uniquement de k et que $k = o(N_k)$. On écrit N à la place de N_k pour simplifier les notations.

L'étude des grandes déviations consiste à calculer la probabilité des événements rares ou des valeurs extrêmes que peut prendre une suite ergodique. On distingue souvent dans la littérature deux types de propriétés de déviations : les propriétés non-asymptotiques qui donnent une borne supérieure de la probabilité des événements rares à n fixé et les propriétés asymptotiques qui fournissent un « équivalent » à l'infini de cette probabilité. Ces propriétés sont souvent utilisées en statistiques. Elles permettent d'établir la performance de certains tests [BDMN10] ou de construire les intervalles de confiance des tests de vraisemblance [BGR97, BGL00, DC79].

Pour le modèle de percolation dirigée sur un rectangle fin, plusieurs résultats d'ergodicité sont connus et ont été énoncés dans le chapitre précédent. Ces résultats de convergence permettent d'avoir une idée sur l'ordre des grandes déviations à considérer. En voici quelques exemples.

Comme $G(N, k)$ est superadditive, on sait d'une part, d'après le théorème de Kingman [Lig85], que $G(N, k)/N$ converge presque sûrement vers une constante déterministe γ . Quand on connaît

γ , on peut se poser la question d'évaluer les grandes déviations d'ordre N . D'autre part, un résultat de Seppalaïnen [Sep97] affirme que, lorsque $k = o(N)$,

$$\frac{G(N, k) - (N + k - 1)\mu}{2\sqrt{Nk}} \xrightarrow{\mathbb{P}} \sigma,$$

et ceci lorsque les poids sont *i.i.d.* admettant des moments exponentiels. Notons que ce résultat a été amélioré par Martin [Mar04], (voir (2.9)). Etant donnée cette dernière convergence, il est légitime de calculer les grandes déviations d'ordre \sqrt{Nk} . Notons que lorsque N est proportionnel à k , cette question ne se pose plus puisque N et \sqrt{Nk} seront du même ordre.

Finalement, on peut regarder les déviations qui correspondent à l'ordre des fluctuations. C'est le plus petit ordre de convergence que l'on puisse considérer. Ce sont les déviations optimales. D'après [BM05, BS05], on a que

$$\Theta_k := \frac{G(N, k) - (N + k - 1)\mu - 2\sigma\sqrt{Nk}}{\sigma N^{1/2}k^{-1/6}} \xrightarrow{\mathcal{L}} \Theta,$$

où Θ est une variable qui suit une loi de Tracy-Widom. Par suite, pour tout $s \in \mathbb{R}$,

$$\mathbb{P}[\Theta_k \leq s] \longrightarrow \mathbb{P}[\Theta \leq s].$$

Par ailleurs, peut-on dire que Θ_k et Θ ont les mêmes queues de distribution à k fixé ? Dans le paragraphe 3.2.4, on a démontré que Θ_k admet des queues de distribution au moins exponentielles. Mais cela n'est pas suffisant pour dire qu'on a les mêmes queues d'une loi de Tracy-Widom. Ces inégalités seront améliorées dans la suite.

Pour résumer la situation, on peut dire que pour N et k assez grands,

$$G(N, k) \approx \mu(N + k - 1) + 2\sigma\sqrt{Nk} + \sigma N^{1/2}k^{-1/6}\Theta.$$

Cela nous pousse donc à regarder les propriétés de grandes déviations à trois niveaux différents.

3.3.1 Grandes déviations du premier ordre (N).

Le calcul des grandes déviations d'ordre N est motivé par la loi forte des grands nombres que satisfait $G(N, k)$. Dans le cas où $k = O(N)$, celles-ci seront confondues avec les grandes déviations en \sqrt{Nk} et ne sont connues que dans le cas de variables géométriques ou exponentielles.

Dans le cas d'un rectangle fin, on démontre que les fonctions de taux de $G(N, k)$ sont les mêmes que celles d'une somme partielle de variables *i.i.d.*. Ceci est dû au fait que le nombre de chemin sur un rectangle fin est négligeable devant N .

Théorème 3.1. *On suppose que $\mathbb{E}X = \mu$ et que $\text{Var}X = \sigma^2$. On suppose également qu'il existe $c_0 > 0$ tel que $\mathbb{E}e^{c|X|} < \infty$ pour tout $c < c_0$. Si $k = o(N_k)$, on a pour tout $\varepsilon > 0$, que*

$$\lim_{k \rightarrow \infty} \frac{1}{N_k} \log \mathbb{P}[G(N_k, k) \geq (N_k + k - 1)(\mu + \varepsilon)] = -I(\varepsilon)$$

La fonction $I(\varepsilon)$ est strictement positive pour tout $\varepsilon > 0$. C'est la transformée de Fenchel-Legendre de X qui apparaît dans le théorème de Cramér classique. On a que

$$I(\varepsilon) = \sup_{t \geq 0} \{t(\mu + \varepsilon) - \log \mathbb{E}[e^{tX}]\}$$

Démonstration : La preuve du Théorème 3.1 repose essentiellement sur le dénombrement des chemins dirigés dans le rectangle fin de dimension N et k . On fixe un chemin $\pi_0 \in \Pi(N, k)$ et on remarque que

$$\mathbb{P}[G(N, k) \geq (N + k - 1)(\mu + \varepsilon)] \geq \mathbb{P}\left[\sum_{(i,j) \in \pi_0} X_i^{(j)} \geq (N + k - 1)(\mu + \varepsilon)\right].$$

Par suite, d'après le théorème de Cramér [DZ98],

$$\liminf_{k \rightarrow \infty} \frac{1}{N} \log \mathbb{P}[G(N, k) \geq (N + k - 1)(\mu + \varepsilon)] \geq -I(\varepsilon).$$

D'autre part, on sait que

$$\begin{aligned} \mathbb{P}[G(N, k) \geq (N + k - 1)(\mu + \varepsilon)] &\leq \sum_{\pi \in \Pi(N, k)} \mathbb{P}\left[\sum_{(i,j) \in \pi} X_i^{(j)} \geq (N + k - 1)(\mu + \varepsilon)\right] \\ &\leq \#\Pi(N, k) \mathbb{P}\left[\sum_{(i,j) \in \pi_0} X_i^{(j)} \geq (N + k - 1)(\mu + \varepsilon)\right] \\ &\leq \binom{N + k - 1}{k - 1} \mathbb{P}\left[\sum_{(i,j) \in \pi_0} X_i^{(j)} \geq (N + k - 1)(\mu + \varepsilon)\right] \end{aligned}$$

Par conséquent,

$$\limsup_{k \rightarrow \infty} \frac{1}{N} \log \mathbb{P}[G(N, k) \geq (N + k - 1)(\mu + \varepsilon)] \leq -I(\varepsilon) + \lim_{k \rightarrow \infty} \frac{1}{N} \log \binom{N + k - 1}{k - 1}.$$

En utilisant la formule de Stirling, on a, pour N et k assez grands et $k = o(N)$, que

$$\log \#\Pi(N, k) \leq \log \left(\frac{N}{k}\right)^k \left(1 + \frac{N}{k}\right)^{N+k} \leq 2k \log \frac{N}{k}.$$

Pour terminer, il suffit de remarquer que $\frac{k}{N} \log \left(\frac{N}{k}\right) = -\frac{k}{N} \log \left(\frac{k}{N}\right)$ converge vers 0 quand $k = o(N)$.

□

Pour les grandes déviations à gauche de la moyenne, cette méthode ne fonctionne plus. On pense même que la probabilité de ces grandes déviations décroît en e^{-Nk} au lieu de e^{-N} . En fait, pour que $G(N, k)$ soit grand, il suffit qu'il existe un chemin qui vérifie cela. Cependant, pour que $G(N, k)$ soit petit, il faut que la majorité des Nk variables soient petites.

3.3.2 Asymptotiques des grandes déviations du second ordre (\sqrt{Nk}).

Cette partie contient les principaux résultats obtenus durant cette thèse pour le LPP sur un rectangle fin. Ces résultats mettent en évidence encore une fois le lien entre le LPP et les modèles de matrices aléatoires et en particulier, celui du GUE. L'étude des grandes déviations d'ordre \sqrt{Nk} est motivée par la formule (2.9) de Martin [Mar04]. Cette propriété asymptotique est une sorte d'étape intermédiaire entre la convergence presque sûre dans la LGN forte et la convergence en loi des fluctuations dans le TCL (Tracy-Widom).

On se contente ici d'énoncer et d'interpréter les résultats obtenus dans [Ibr07]. Les preuves sont données en détail dans le chapitre suivant entièrement consacré à cette fin. On suppose pour des raisons de simplification et sans perte de généralité que $\mu = 0$ et $\sigma^2 = 1$. La suite repose essentiellement sur deux propriétés. D'une part, on a une égalité en loi entre la plus grande valeur propre d'une matrice du GUE et le temps de dernier passage Brownien $L(1, k)$ [O'03]. D'autre part, on est capable, en utilisant l'approximation K-M-T, de reconstruire les poids du LPP de sorte que le temps de dernier passage et le temps de dernier passage Brownien soient le plus proche possible en probabilité. Cela se traduit de la manière suivante : si on note

$$A = \{G(N, k) \geq 2\sqrt{Nk}(1 + \varepsilon)\},$$

alors pour tout $\varepsilon > 0$ et tout $0 < \varepsilon_1 < \varepsilon$, on a que

$$\mathbb{P}[A] \leq \mathbb{P}[L(N, k) \geq 2\sqrt{Nk}(1 + \varepsilon_1)] + \mathbb{P}[|G(N, k) - L(N, k)| \geq 2\sqrt{Nk}(\varepsilon - \varepsilon_1)] \quad (3.27)$$

et

$$\mathbb{P}[A] \geq \mathbb{P}[L(N, k) \geq 2\sqrt{Nk}(1 + 2\varepsilon - \varepsilon_1)] - \mathbb{P}[|G(N, k) - L(N, k)| \geq 2\sqrt{Nk}(\varepsilon - \varepsilon_1)] \quad (3.28)$$

Connaissant le comportement de $L(N, k)$, il nous reste à contrôler le second terme de la partie droite de l'inégalité (3.27). On reconstruit les variables $X_i^{(j)}$ de manière à ce que $G(N, k)$ et $L(N, k)$ soient proches. On obtient donc que

$$|G(N, k) - L(N, k)| \approx b_k.$$

On peut dire alors que tout régime d'ordre a_k admet un comportement similaire à celui de $L(N, k)/a_k$ du moment où a_k l'emporterait sur b_k . Nous traitons maintenant, avec ces outils, les différentes possibilités de poids aléatoires. En modifiant les conditions sur les poids, la largeur du rectangle risque de changer pour s'adapter à la précision du couplage utilisé.

Le cas Gaussien : Comme la comparaison se fait avec des trajectoires Browniennes, il est naturel de penser que le cas de poids Gaussiens est le plus adapté à cette technique. En tenant compte de l'interprétation (2.3) de $G(N, k)$ et de la définition (1.14) de $L(N, k)$, on construit les $X_i^{(j)}$ de la façon suivant :

$$X_i^{(j)} = B_i^{(j)} - B_{i-1}^{(j)}.$$

Par conséquent,

$$|G(N, k) - L(N, k)| \leq \sup_{0 \leq t_1 \leq t_2 \leq \dots \leq t_{k-1} \leq N} \left\{ \sum_{r=1}^k \left(\sup_{\substack{t_{r-1} \leq s, t \leq t_r \\ |s-t| < 2}} |B_s^{(r)} - B_t^{(r)}| \right) \right\} \quad (3.29)$$

En intervertissant le premier supremum et la somme, on obtient un terme qui est de l'ordre de $k\sqrt{\log N}$. Tout calcul fait, on obtient pour $k = o(N/\log N)$ que

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathbb{P}[G(N_k, k) \geq 2\sqrt{N_k k}(1 + \varepsilon)] = -J_{\text{GUE}}(\varepsilon). \quad (3.30)$$

Les grandes déviations à gauche de la moyenne de la plus grande valeur propre du GUE sont plutôt de l'ordre de e^{-k^2} . En effet, si $k = o(N^{1/2})$,

$$\lim_{k \rightarrow \infty} \frac{1}{k^2} \log \mathbb{P}[G(N_k, k) \leq 2\sqrt{N_k k}(1 - \varepsilon)] = -I_{\text{GUE}}(\varepsilon). \quad (3.31)$$

Remarque 3.2. Ici le couplage est quasi parfait à l'image de celui fait au début du chapitre entre les poids exponentiels et les poids Gamma. Pourtant, nous pensons que la largeur du rectangle pourrait être améliorée en $k = o(N)$. Cette largeur est la meilleure qu'on puisse espérer pour obtenir les mêmes fonctions de taux que celle GUE. La perte se produit au moment où l'on intervertit le supremum et la somme dans (3.29). Sans pouvoir le démontrer, nous pensons que si on garde (3.29) telle qu'elle est, on doit pouvoir obtenir une borne supérieure de l'ordre de $k \log(N/k)$. En fait, en effectuant l'interversion, les maxima des variables Gaussiennes sont choisis dans des échantillons de taille N , d'où le $\log N$. En gardant (3.29) telle qu'elle est, les maxima sont choisis dans des échantillons de longueur moyenne N/k . La difficulté réside dans le fait que les longueurs respectives de ces échantillons sont aléatoires. A part le fait que la somme des longueurs des échantillons est N , on ne connaît pas leur comportement quand N est grand.

Transformée de Laplace finie au voisinage de 0 : On considère maintenant des poids *i.i.d.* quelconques dont la transformée de Laplace au voisinage de 0 est finie. Physiquement parlant, c'est le cas le plus intéressant à examiner du fait des liens étroits avec les modèles de polymères dirigés en milieu aléatoire. $G(N, k)$ est vu comme la limite de la fonction de partition d'une certaine mesure de Gibbs quand la température décroît vers 0. Cela justifie la condition de finitude des moments exponentiels.

D'après les calculs effectués par Bodineau et Martin dans [BM05], on peut obtenir des formules similaires à (3.30) et (3.31) si on parvient à contrôler suffisamment

$$\sum_{r=1}^k \left(\max_{i=1, \dots, N} |S_i^{(r)} - B_i^{(r)}| \right),$$

où $S_m^{(r)} = \sum_{i=1}^m X_i^{(r)}$. Pour cela, on reconstruit les $X_i^{(j)}$ dans le sens de l'approximation K-M-T afin d'obtenir que

$$\max_{i=1, \dots, N} |S_i^{(r)} - B_i^{(r)}| = O(\log N).$$

Notons ici que la construction des $X_i^{(j)}$ change avec N , mais cela n'a aucune influence sur notre preuve. Ce couplage montre que $b_k = O(k \log N)$. Pour terminer, on compare b_k à \sqrt{Nk} et on

déduit que (3.30) reste vraie si $k = o(N/(\log N)^2)$. A gauche de la moyenne, on a également une formule identique à (3.31) dès que $k = o(N^{1/3})$.

Remarque 3.3. On peut penser que si les poids ont une queue exponentielle à droite et une queue lourde à gauche, la formule des grandes déviations à droite de la moyenne (3.30) tiendra toujours. Le problème est que l'approximation K-M-T utilisée ici ne fonctionne que pour des variables à double queue exponentielle.

Remarque 3.4. On examine maintenant la fonction de partition du modèle de polymères dirigés en milieu aléatoire introduit dans le Paragraphe 2.2. L'inégalité (2.6) montre que $\frac{1}{\beta_k} \log Z_{\beta_k}(N, k)$ satisfait les mêmes propriétés asymptotiques de grandes déviations que $G(N, k)$ lorsque $\beta_k = O(k^{-\gamma})$ avec $\gamma < \frac{1-\alpha}{2\alpha}$. Cependant, Quand $\gamma = \frac{1-\alpha}{2\alpha}$, $\frac{1}{\beta_k} \log Z_{\beta_k}(N, k)$ n'est plus comparable à $G(N, k)$ dans le régime \sqrt{Nk} . Moreno [MF10] a étudié ce cas de plus près en effectuant un plongement dans le modèle de polymères dirigés dans un environnement Brownien. L'énergie libre de ce dernier modèle étant connue suite aux travaux de Moriarty et O'Connell [MO07], Moreno démontre, via l'approximation K-M-T que lorsque les $X_i^{(j)}$ admettent des moments exponentiels, on a pour $\beta_k = \beta k^{-\gamma}$, que

$$\lim_{k \rightarrow \infty} \frac{1}{\beta_k k^{\frac{1+\alpha}{2\alpha}}} \log Z_{\beta_k}(N, k) = \frac{p(\beta)}{\beta}. \quad (3.32)$$

La fonction $p(\beta)$ est l'énergie libre du modèle de polymères dirigés dans un environnement Brownien. Elle est donnée explicitement dans [MO07].

Le cas de poids sous-exponentiels : les variables sous-exponentielles font partie de la catégorie des distributions à queues lourdes *i.e.* la transformée de Laplace n'est finie en aucun voisinage de 0. Soit (X_n) une suite de variables *i.i.d.* de loi μ . On dit que μ est sous-exponentielle si et seulement si pour tout $n \geq 1$,

$$\mathbb{P}\left[\sum_{i=1}^n X_i \geq x\right] \underset{x}{\sim} \mathbb{P}\left[\max_{i=1, \dots, n} \{X_i\} \geq x\right].$$

Il est clair que cette définition n'engage que la queue de droite. Ainsi, une loi est sous-exponentielle si et seulement si sa queue droite est lourde. Dans la littérature, on distingue plusieurs types de distributions sous-exponentielles. Ici, on n'évoquera que deux cas : les variables de type Weibull, *i.e.* il existe $c > 0$ et $0 < \gamma < 1$ tels que $\mathbb{E}e^{c|X|^\gamma} < \infty$, et les variables admettant des moments finis d'ordre $p > 2$.

Dans le cas de variables admettant des moments finis d'ordre p , on utilise le Théorème 3.4 de Sakhanenko [Sak91]. On a donc que

$$\mathbb{P}\left[\sum_{r=1}^k \left(\max_{i=1, \dots, N} |S_i^{(r)} - B_i^{(r)}|\right) \geq \sqrt{Nk}(\varepsilon - \varepsilon_1)\right] \leq \frac{Ck^{p/2}}{(\varepsilon - \varepsilon_1)^p N^{\frac{p}{2}-1}}.$$

A cause de la décroissance polynomial, on a besoin d'un rectangle très fin, soit $k = o(\log N^{p-1})$. Notons que l'étude dans ce cas n'est pas très intéressante vue la taille du rectangle. En plus, on ne peut pas établir des inégalités optimales à N et k fixés.

Le cas des variables du type Weibull est plus délicat à démontrer. On ne connaît pas de version de l'approximation K-M-T adaptée à ce type de variables. On utilise à la place un couplage moins efficace mais qui répond à notre condition de moments : le plongement de Skorokhod. Les résultats concernant ce plongement sont donnés dans le paragraphe 3.1.2. La démonstration des propriétés similaires à (3.30) et (3.31) est détaillée dans la quatrième partie du Chapitre 4 et repose essentiellement sur des arguments de [BS05]. En voici les étapes essentielles :

1. D'abord, on reconstruit les $X_i^{(j)}$ en utilisant le Corollaire 4.2 du théorème de Skorokhod de la manière suivante :

$$X_i^{(r)} = B_{\tau_1 + \dots + \tau_i}^{(r)} - B_{\tau_1 + \dots + \tau_{i-1}}^{(r)}.$$

Par conséquent,

$$S_i^{(r)} = B_{\tau_1 + \dots + \tau_i}^{(r)} \quad p.s.$$

2. La deuxième étape consiste à trouver une relation entre les moments de X et les moments du temps d'arrêt τ_1 . Cela est possible grâce au Théorème 4.3 de Sawyer [Saw72].
3. Il est possible maintenant de contrôler la somme partielle des temps d'arrêt en utilisant l'inégalité de Fuk-Nagaev (Théorème 4.4) qui est une version de l'inégalité de Bernstein mais avec des conditions de moments plus faibles.
4. Maintenant qu'on contrôle la somme partielle des temps d'arrêts, on sait que $B_{\tau_1 + \dots + \tau_i}^{(r)}$ et $B_i^{(r)}$ sont très proches. On refait alors les mêmes calculs que dans les cas précédents pour conclure.

Finalement, on a une formule asymptotique de grandes déviations à droite (resp. à gauche) de la moyenne si $k = o(N^\alpha)$ avec $\alpha < \frac{\gamma}{2\gamma+2}$ (resp. $\alpha < \frac{\gamma}{5\gamma+4}$).

3.3.3 Inégalités optimales non-asymptotiques.

Dans ce paragraphe, on établit des inégalités optimales à droite et à gauche de la moyenne du temps de dernier passage sur un rectangle fin. Dans la suite, on suppose que les poids sont *i.i.d.* admettant une transformée de Laplace finie au voisinage de 0. Une légère amélioration de la largeur du rectangle sera signalée dans le cas Gaussien. Les preuves exactes de ces inégalités sont dans la dernière partie du chapitre suivant.

Malgré leurs intérêts scientifiques, les résultats asymptotiques de grandes déviations obtenus pour $G(N, k)$ fournissent très peu d'informations sur le comportement et la variabilité de la probabilité des grandes déviations lorsque N et k sont fixés. Les bornes optimales non-asymptotiques donnent une idée claire d'une part sur la décroissance de la probabilité des déviations optimales en N et k et d'autre part sur la décroissance de cette même probabilité en ε tout en fixant N et k .

L'étude des inégalités optimales est motivée par le théorème de convergence des fluctuations vers la loi de Tracy-Widom [BM05]. L'ordre de convergence des fluctuations (ici $N^{1/2}k^{-1/6}$) est le plus petit ordre possible dans l'étude des déviations. En dessous de cet ordre, on ne parle plus d'événements rares car la suite n'est plus ergodique. Bien qu'on ne connaît pas de forme explicite de la densité de la loi de Tracy-Widom, on peut obtenir des équivalents des queues de droite et de gauche de cette distribution. On a que

$$F_2^{\text{TW}}(t) \sim e^{t^3/12} \quad \text{quand } t \rightarrow -\infty$$

et

$$1 - F_2^{\text{TW}} \sim \frac{e^{-(4/3)t^{3/2}}}{16\pi t^{3/2}} \quad \text{quand } t \rightarrow \infty.$$

On se pose naturellement la question de savoir si Θ_k possède les mêmes queues de distribution que celles de la loi de Tracy-Widom. Dans le paragraphe 3.2.4, on a démontré la Proposition 3.6 qui fournit une borne exponentielle des déviations optimales. Le but n'était pas de trouver une borne optimale mais plutôt d'avoir une majoration approximative exponentiellement décroissante en ε et qui soit valable sur le rectangle le plus large possible. Cette proposition nous a servi ensuite pour calculer l'ordre des fluctuations transversales (Théorème 3.8) et prouver la convergence des moments de Θ_k (Proposition 3.13).

Ici, on ne se préoccupe pas de la largeur du rectangle. Le but est d'atteindre une majoration optimale du même ordre que les queues de distribution de la loi de Tracy-Widom. On a le résultat suivant :

Théorème 3.5. *On suppose que $\mathbb{E}X = 0$ et $\mathbb{E}X^2 = 1$. On suppose de plus qu'il existe $c_0 > 0$ tel que $\mathbb{E}[e^{c|X|}] < +\infty$ pour tout $0 \leq c < c_0$. Si de plus $k = N^\alpha$ avec $\alpha < \frac{3}{7}$, alors pour tous N et k et pour tout $0 < \varepsilon \leq k^{2/3}$, on a que*

$$\mathbb{P}[G(N, k) - 2\sqrt{Nk} \geq \varepsilon N^{1/2} k^{-1/6}] \leq C_\alpha \exp\left(-\frac{\varepsilon^{3/2}}{C_\alpha}\right)$$

Si $\alpha < \frac{1}{3}$, alors

$$\mathbb{P}[G(N, k) - 2\sqrt{Nk} \leq -\varepsilon N^{1/2} k^{-1/6}] \leq C_\alpha \exp\left(-\frac{\varepsilon^3}{C_\alpha}\right).$$

où C_α est une constante positive ne dépendant que de α .

Un résultat similaire peut être obtenu dans le cas de variables de du type Weibull en changeant la largeur du rectangle. Les calculs sont laissés au lecteur. Dans les inégalités non-asymptotiques, il est préférable de prendre $k = N^\alpha$ au lieu de $k = o(N^\alpha)$. Sinon, la constante C_α risque de dépendre de la fonction $N = N(k)$.

Remarque 3.6. Dans [BDM⁺01], les auteurs ont établi des formules asymptotiques de déviations modérées ainsi que des inégalités optimales dans le cas géométrique avec N et k proportionnelles. L'analyse de la fonction de répartition de $G(N, k)$, bien identifiée dans ce cas, leur permet de connaître le comportement des déviations d'ordre quelconque compris entre celui des fluctuations et celui de la loi forte des grands nombres. Un autre résultat similaire a été donné dans [BDJ99] pour le problème de la plus longue sous-suite croissante d'une permutation aléatoire. On cite enfin [LR10] pour des inégalités similaires dans le cadre général des β -ensembles polynomiaux.

Pour démontrer le Théorème 3.5, il suffit de combiner (3.28) et le Théorème 3.5 de Ledoux et Rider [LR10]. Pour la preuve détaillée, le lecteur est invité à regarder la dernière partie du Chapitre 4.

Chapter 4

Large deviations for directed percolation on a thin rectangle.

Following the recent investigations of Baik and Suidan in [BS05] and Bodineau and Martin in [BM05], we prove large deviation properties for a last-passage percolation model in \mathbb{Z}_+^2 whose paths are close to the axis. The results are mainly obtained when the random weights are Gaussian or have a finite moment-generating function and rely, as in [BS05] and [BM05], on an embedding in Brownian paths and the K-M-T approximation. The study of the subexponential case completes the exposition.

Dans cet article, on démontre des propriétés de grandes déviations pour un modèle de percolation dirigée dans le plan dont les chemins sont proches de l'axe. Les résultats sont obtenus dans le cas de variables aléatoires suivant une loi Gaussienne. La preuve s'appuie essentiellement sur l'approximation K-M-T exploitée dans ce cadre dans l'article de Bodineau et Martin [BM05]. Des résultats similaires sont obtenus pour des variables aléatoires admettant une transformée de Laplace au voisinage de 0.

Introduction

Random matrix theory has developed extensively in the last several decades following the pioneering results by E. Wigner in the fifties. Gaussian models attracted a lot of attention, among them the Gaussian Unitary Ensemble (GUE). In this example, the knowledge of the joint distribution of the eigenvalues allowed for a rather complete understanding of both their global and local behaviors. In particular, the limiting behavior of the largest eigenvalue gave rise to the famous Tracy-Widom distribution [Joh98, Meh91, TW94]. Several random growth models, such as the longest increasing subsequence and the corner growth models, have been shown to develop a similar behavior relying on a common determinantal structure [Joh00a, Kön05]. In particular, the last-passage percolation or so-called corner growth model (see below for a precise description) has been deeply studied by Johansson in [Joh00a]. For geometric or exponential random variables (the only cases leading to a determinantal description), Johansson established both fluctuations and large deviation asymptotics similar to the ones for the GUE random matrix model. Following the recent investigations by Baik and Suidan [BS05] and Bodineau and Martin [BM05] at the level of fluctuations, the present paper deals with large deviations for the random growth model for more general random variables but on rectangles such that one side is asymptotically negligible with respect to the other at a given rate. The main results concern

Gaussian random variables and random weights having a finite moment-generating function. Somewhat surprisingly, the rate may be shown to be larger than the one for the fluctuations. The comparison method used in this work is basically inspired from [BS05] and [BM05] and relies similarly on an embedding in Brownian paths.

Recall first the basic corner growth model under study. It can be described as directed paths in the lattice \mathbb{Z}_+^2 going from $(1, 1)$ to $(N, k) \in \mathbb{Z}_+^2$ where only up and right steps are allowed. More precisely, denoting by $\Pi(N, k)$ the set of all such paths, a path $\pi \in \Pi(N, k)$ is called an *up/right* path and is defined as a collection of sites $\{(i_l, j_l)\}_{l=1}^{N+k-1}$ satisfying $(i_1, j_1) = (1, 1)$, $(i_{N+k-1}, j_{N+k-1}) = (N, k)$ and $(i_{l+1}, j_{l+1}) - (i_l, j_l)$ is either $(1, 0)$ or $(0, 1)$. The main random variable under consideration is the last-passage time defined by

$$G(N, k) = \max_{\pi \in \Pi(N, k)} \left\{ \sum_{(i, j) \in \pi} X_i^{(j)} \right\}$$

where the $X_i^{(j)}$'s are *i.i.d.* random variables. As an alternate description, set $U(N, k)$ as the subset of \mathbb{R}_+^{k+1} given by

$$U(N, k) = \{u = (u_0, u_1, \dots, u_k) \in \mathbb{R}_+^{k+1}; 0 = u_0 \leq u_1 \leq \dots \leq u_k = N\}.$$

Then

$$G(N, k) = \sup_{u \in U(N, k)} \left\{ \sum_{r=1}^k [S_{\lfloor u_r \rfloor}^{(r)} - S_{\lfloor u_{r-1} \rfloor - 1}^{(r)}] \right\}$$

where $S_m^{(r)} = \sum_{i=1}^m X_i^{(r)}$ with the convention $S_{-1}^{(r)} = S_0^{(r)} = 0$. This follows from the fact that $U(N, k) \equiv \Pi(N, k)$ when $u \in \mathbb{Z}_+^{k+1}$. Actually, every $u \in U(N, k) \cap \mathbb{Z}_+^{k+1}$ maps to a unique path $\pi \in \Pi(N, k)$ whose i^{th} up-jump occurs on u_i . On the other hand, each path π is characterized by its up-step sites. It will be more appropriate to adopt the second form of $G(N, k)$ in order to compare it later with the Brownian last-passage percolation model.

4.1 Results

Having introduced the model, recall briefly some of its properties established by Johansson in [Joh00a] in the particular case of geometric or exponential distributions. The key of the results in this case relies on the explicit description of the last-passage time distribution $G(N, k)$. When weights are *i.i.d.* geometric with parameter $q \in (0, 1)$, we have

$$\mathbb{P}[G(N, k) \leq t] = \frac{1}{Z_{N, k}} \sum_{\substack{h \in \mathbb{N}^k \\ \max\{h_i\} \leq t+k-1}} \prod_{1 \leq i < j \leq k} (h_i - h_j)^2 \prod_{i=1}^k \binom{h_i + N - k}{h_i} q^{h_i},$$

where $N \geq k$ and $Z_{N, k}$ is the normalizing constant. Using results from logarithmic potential theory, Johansson described in [Joh00a] the large deviation behaviors of $G(N, k)$ when k and N are of the same order. Namely, he obtained that, for $\gamma \geq 1$, there exist two rate functions $i(\varepsilon)$ and $l(\varepsilon)$ such that for any $\varepsilon > 0$,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathbb{P}[G(\lceil \gamma k \rceil, k) \geq k(\omega(\gamma, q) + \varepsilon)] = -i(\varepsilon) \quad (4.1)$$

and

$$\lim_{k \rightarrow \infty} \frac{1}{k^2} \log \mathbb{P}[G([\gamma k], k) \leq k(\omega(\gamma, q) - \varepsilon)] = -l(\varepsilon). \quad (4.2)$$

The functions $l(x)$ and $i(x)$ are positive for every $x > 0$ and

$$\omega(\gamma, q) := \lim_{k \rightarrow \infty} \frac{1}{k} \mathbb{E}[G([\gamma k], k)] = \frac{(1 + \sqrt{\gamma q})^2}{1 - q} - 1.$$

Using the asymptotics of the Meixner orthogonal polynomial ensemble, Johansson [Joh00a] further established the fluctuations of $G(N, k)$ at the Tracy-Widom GUE rate. He proved that for $\gamma \geq 1$ and $s \in \mathbb{R}$,

$$\lim_{k \rightarrow \infty} \mathbb{P}\left[\frac{G([\gamma k], k) - k\omega(\gamma, q)}{\sigma(\gamma, q)k^{1/3}} \leq s\right] = F_2^{\text{TW}}(s),$$

where

$$\sigma(\gamma, q) = \frac{q^{1/6}\gamma^{-1/6}}{1 - q} (\sqrt{\gamma} + \sqrt{q})^{2/3} (1 + \sqrt{\gamma q})^{2/3}$$

and $F_2^{\text{TW}}(s)$ is the distribution function of the Tracy-Widom law (see [TW94]). Replacing geometric weights with exponential ones gives similar results since an exponential distribution can be seen as the limit of a rescaled geometric one. See [Joh00a] for the precise formulas.

Recently, Bodineau and Martin [BM05] and Baik and Suidan [BS05] studied the same model when paths are close to the axis, *i.e.* $k = o(N^\alpha)$ for some $\alpha < 1$ but allowing more general distributions. The authors used a coupling with the Brownian trajectories through the following Brownian last-passage percolation. Letting $(B_t^{(r)})_{r \geq 1}$ be a sequence of independent Brownian motions, set

$$L(N, k) = \sup_{u \in U(N, k)} \left\{ \sum_{r=1}^k [B_{u_r}^{(r)} - B_{u_{r-1}}^{(r)}] \right\}.$$

It has been proved in [Bar01, GTW01, OY02] that $L(1, k)$ has the same distribution as the largest eigenvalue of a $k \times k$ rescaled GUE random matrix. As a consequence of the fluctuation result for the GUE model, it follows that

$$k^{1/6} [L(1, k) - 2\sqrt{k}] \xrightarrow{d} \Theta,$$

where $\Theta \sim \text{TW}_2$. Using this result and a comparison between the continuous model with Brownian paths and the discrete one with random weights, the authors of [BS05] and [BM05] deduced fluctuation properties of the corner growth model for rather general random variables. However, the embedding in the Brownian paths requires to restrict the paths on small rectangles. For example, in [BM05], the discrete and the continuous models were coupled using the Komlós-Major-Tusnády (K-M-T) approximation which couples random walks with Brownian motion. The authors proved that if the weights satisfy $\mathbb{E}|X_i^{(j)}|^p < \infty$ for some $p > 2$, setting $\mu = \mathbb{E}X_i^{(j)}$ and $\sigma^2 = \text{var}(X_i^{(j)})$, then for all $\alpha < \frac{6}{7}(\frac{1}{2} - \frac{1}{p})$,

$$\frac{G(N, \lfloor N^\alpha \rfloor) - N\mu - 2\sigma N^{\frac{1+\alpha}{2}}}{\sigma N^{\frac{1}{2} - \frac{\alpha}{6}}} \xrightarrow{d} \Theta.$$

If the random variables $X_i^{(j)}$ have all moments, *i.e.* $p = \infty$, then α is lower than $3/7$. This is true when the weights are Gaussian or are bounded for example. It is not known how optimal this

rate could be: the authors in [BM05] think that such a result might hold, for some independence reasons, when $\alpha < 3/4$. However, they do not give a complete proof. In [BS05], the authors compared the discrete and continuous model via the Skorokhod embedding theorem in order to obtain almost the same results. Lately, Suidan in [Sui06] produced another proof of the last theorem when the variables have a third moment. He compared two discrete directed percolation models using a theorem of Chatterjee [Cha05]. The fluctuation properties for the first (with geometric distribution) lead him to similar ones for the second.

In this paper, we follow the comparison methods of [BM05] and [BS05] to establish large deviations limit theorems for directed percolation models on thin rectangles. The results mainly concern Gaussian weights and random weights with finite moment-generating function around zero. We rely similarly on the corresponding results for the Brownian percolation model. Namely, as a consequence of the GUE random matrix interpretation [GTW01, OY02], for all $\varepsilon > 0$,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathbb{P}[L(1, k) \geq 2\sqrt{k}(1 + \varepsilon)] = -J_{\text{GUE}}(\varepsilon) \quad (4.3)$$

and

$$\lim_{k \rightarrow \infty} \frac{1}{k^2} \log \mathbb{P}[L(1, k) \leq 2\sqrt{k}(1 - \varepsilon)] = -I_{\text{GUE}}(\varepsilon). \quad (4.4)$$

The two functions $J_{\text{GUE}}(x)$ and $I_{\text{GUE}}(x)$ are both positive for every positive x . J_{GUE} can be computed explicitly (see [BADG01]) as

$$J_{\text{GUE}}(\varepsilon) = 4 \int_0^\varepsilon \sqrt{x(x+2)} \, dx.$$

To the best of our knowledge, there is no explicit form for I_{GUE} . This function appears in the logarithmic potential theory and it represents physically the minimal potential energy of charges on a one-dimension conductor exposed to an external field (see [ST97]). In this work, we do not need the explicit form of I_{GUE} . However, its continuity, proved at the end of Section 4.2, will be necessary for the proof.

The following three theorems are the main results of this paper. Despite some similarity in their proofs, the second one requires more work. One surprising feature is that the rate α in the Gaussian case can be taken arbitrary less than 1 (compared to $\alpha < 3/7$ for the fluctuation result).

Throughout the article, k and N are two integers and N depends on k , *i.e.* $N = N(k) := N_k$. We assume further that $k = o(N_k)$ and we let k goes to infinity. For simplicity, we write N instead of N_k throughout the proofs.

Theorem 4.1. *Assume that the variables $(X_i^{(j)})_{i,j=1}^\infty$ are i.i.d. standard normal random variables. Assume further that $k = o(\frac{N_k}{\log N_k})$. Then, for all $\varepsilon > 0$,*

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathbb{P}[G(N_k, k) \geq 2\sqrt{N_k k}(1 + \varepsilon)] = -J_{\text{GUE}}(\varepsilon).$$

On the left of the mean, we have for $k = o(N_k^{\frac{1}{2}})$,

$$\lim_{k \rightarrow \infty} \frac{1}{k^2} \log \mathbb{P}[G(N_k, k) \leq 2\sqrt{N_k k}(1 - \varepsilon)] = -I_{\text{GUE}}(\varepsilon).$$

In the second statement, we replace Gaussian variables with weights having finite exponential moments. Loosing the Gaussian assumption will complicate the coupling and reduce the size of the rectangles. We denote by X a random variable having the common law of the *i.i.d.* variables in the sequence $(X_i^{(j)})_{i,j=1}^\infty$.

Theorem 4.2. *Assume that the variables $(X_i^{(j)})_{i,j=1}^\infty$ are *i.i.d.* random variables such that $\mathbb{E}X = 0$ and $\mathbb{E}X^2 = 1$. Assume further that there exists $\mu_0 > 0$ such that for all $\mu < \mu_0$,*

$$\mathbb{E} \exp(\mu|X|) < +\infty. \quad (4.5)$$

If $k = o(\frac{N_k}{(\log N_k)^2})$, then for all $\varepsilon > 0$,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathbb{P}[G(N_k, k) \geq 2\sqrt{N_k k}(1 + \varepsilon)] = -J_{\text{GUE}}(\varepsilon).$$

Similarly, if $k = o(N_k^{\frac{1}{3}})$, for all $\varepsilon > 0$,

$$\lim_{k \rightarrow \infty} \frac{1}{k^2} \log \mathbb{P}[G(N_k, k) \leq 2\sqrt{N_k k}(1 - \varepsilon)] = -I_{\text{GUE}}(\varepsilon).$$

The proof of Theorem 4.2 relies on the Komlós-Major-Tusnády approximation for sums of *i.i.d.* centered random variables with finite exponential moments, see [KMT76]. The following theorem deals with a particular class of subexponential weights. We make use of the Skorokhod embedding theorem to obtain this result.

Theorem 4.3. *Assume that the variables $(X_i^{(j)})_{i,j=1}^\infty$ are *i.i.d.* random variables satisfying $\mathbb{E}X = 0$ and $\mathbb{E}X^2 = 1$. Furthermore, assume that there exist $\mu > 0$ and $0 < \gamma < 1$ such that*

$$\mathbb{E} \exp(\mu|X|^\gamma) < +\infty. \quad (4.6)$$

If $k = o(N_k^\alpha)$ with $\alpha < \frac{\gamma}{2\gamma+2}$, then, for all $\varepsilon > 0$,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathbb{P}[G(N_k, k) \geq 2\sqrt{N_k k}(1 + \varepsilon)] = -J_{\text{GUE}}(\varepsilon).$$

Similarly, if $k = o(N_k^\alpha)$ with $\alpha < \frac{\gamma}{5\gamma+4}$, for all $\varepsilon > 0$,

$$\lim_{k \rightarrow \infty} \frac{1}{k^2} \log \mathbb{P}[G(N_k, k) \leq 2\sqrt{N_k k}(1 - \varepsilon)] = -I_{\text{GUE}}(\varepsilon).$$

The results in Theorem 4.3 cover in particular the examples of Weibull and Lévy distributions. Notice that in Theorem 4.3 we can take $\gamma \geq 1$. However, the result is worthless because of Theorem 4.2. Actually, the K-M-T approximation is more efficient than the Skorokhod embedding theorem for *i.i.d.* random variables with finite exponential moments.

Let us notice here that if the weights have a finite p -th moment with $p \geq 2$, we still obtain the same deviation results but for $k = o(\log N^{p-1})$. In this case, we use again the K-M-T approximation for random weights with finite p -th moment, see [KMT76, Sak91].

It should be noted also that the previous results hold for geometric and exponential weights as soon as $k = o(N_k)$, with the same GUE rate functions. This is a consequence of the large

deviations (4.1) and (4.2) and the fact that the Laguerre ensemble converges to the GUE on the scaling $k = o(N_k)$. The reader can see [Joh00a] and [Led07] for rigorous results.

Non-asymptotic bounds for the preceding models can be deduced from the previous theorems proofs. The rectangle width for small deviations matches in this case the fluctuation results. For this we use analogous deviation inequalities to the right of the mean obtained for the largest eigenvalue of the GUE, see [Led07]. To the left of the mean, we use recent deviation results for the largest eigenvalue of the GUE obtained by Ledoux and Rider [LR10].

Theorem 4.4. *Assume that the variables $(X_i^{(j)})_{i,j=1}^\infty$ are i.i.d. standard normal random variables, and that $k = N_k^\alpha$ with $\alpha < \frac{3}{7}$. Then, there exists a positive constant C_α depending only on α such that, for all $0 < \varepsilon < 1$,*

$$\mathbb{P}[G(N, k) \geq 2\sqrt{Nk}(1 + \varepsilon)] \leq C_\alpha \exp\left(-\frac{k\varepsilon^{3/2}}{C_\alpha}\right). \quad (4.7)$$

On the left of the mean, we have

$$\mathbb{P}[G(N, k) \leq 2\sqrt{Nk}(1 - \varepsilon)] \leq C_\alpha \exp\left(-\frac{k^2\varepsilon^3}{C_\alpha}\right). \quad (4.8)$$

In the Gaussian case, large deviation inequalities for large $\varepsilon > 1$ hold for the optimal rate $\alpha = 1$ using some concentration arguments. We refer to [Led07] for the results and the proofs. For random weights with finite moment-generating function, we have similar results.

Theorem 4.5. *Assume that the variables $(X_i^{(j)})_{i,j=1}^\infty$ are i.i.d. random variables such that $\mathbb{E}X = 0$ and $\mathbb{E}X^2 = 1$. Assume further that there exist $\mu_0 > 0$ such that for all $\mu < \mu_0$,*

$$\mathbb{E} \exp(\mu|X|) < +\infty. \quad (4.9)$$

If $k = N_k^\alpha$ with $\alpha < \frac{3}{7}$, then there exists a positive constant C_α depending on α and the distribution of X such that, for all $0 < \varepsilon < 1$,

$$\mathbb{P}[G(N_k, k) \geq 2\sqrt{N_k k}(1 + \varepsilon)] \leq C_\alpha \exp\left(-\frac{k\varepsilon^{3/2}}{C_\alpha}\right). \quad (4.10)$$

Similarly, if $k = N_k^\alpha$ with $\alpha < \frac{1}{3}$, then

$$\mathbb{P}[G(N_k, k) \leq 2\sqrt{N_k k}(1 - \varepsilon)] \leq C_\alpha \exp\left(-\frac{k^2\varepsilon^3}{C_\alpha}\right). \quad (4.11)$$

As the reader can notice, if X satisfies the condition (4.6), the same exponential inequalities as in Theorem 4.5 can be obtained. In this case, the positive constant will depend on α , μ and γ . A smaller α will be also necessary. The precise calculations are left to the reader.

Non-asymptotic bounds in the two theorems above can be used to compute the transversal fluctuations order. For this, we use the same method introduced by Johansson in [Joh00b]. The convergence of the moments to those of the Tracy-Widom law can also be deduced.

We strongly believe that the preceding results hold on a wider rectangle. However, the method used here does not allow us to improve the rectangle width. Theorem 4.1 will be proved in Section 4.2 while Theorem 4.2 will be proved in Section 4.3. The case of the subexponential weights given in Theorem 4.3 will be discussed in section 4.4. Theorem 4.4 and Theorem 4.5 will be addressed in Section 4.5 on the basis of the preceding results and proofs.

4.2 Proof of Theorem 4.1.

Throughout the rest of the paper, we write N instead of N_k for simplicity. As claimed before, to prove Theorem 4.1, we compare $G(N, k)$ and $L(N, k)$. To do so, let for any $\varepsilon > 0$,

$$A = \{G(N, k) \geq 2\sqrt{Nk}(1 + \varepsilon)\}$$

and

$$B = \{|G(N, k) - L(N, k)| \geq 2\sqrt{Nk}(\varepsilon - \varepsilon_1)\}$$

where $0 < \varepsilon_1 < \varepsilon$. Clearly,

$$\mathbb{P}[A] \leq \mathbb{P}[L(N, k) \geq 2\sqrt{Nk}(1 + \varepsilon_1)] + \mathbb{P}[B] \quad (4.12)$$

and

$$\mathbb{P}[A] \geq \mathbb{P}[L(N, k) \geq 2\sqrt{Nk}(1 + 2\varepsilon - \varepsilon_1)] - \mathbb{P}[B]. \quad (4.13)$$

Moreover, for every $\eta > 0$,

$$\mathbb{P}[L(N, k) \geq 2\sqrt{Nk}(1 + \eta)] = \mathbb{P}[L(1, k) \geq 2\sqrt{k}(1 + \eta)] \quad (4.14)$$

as a consequence of the Brownian scaling $\sqrt{N}L(1, k) \stackrel{d}{=} L(N, k)$. To evaluate $\mathbb{P}[B]$, we couple $G(N, k)$ and $L(N, k)$ by letting $X_i^{(j)} = B_i^{(j)} - B_{i-1}^{(j)}$ for all $i, j \geq 1$ so that the sequence $(X_i^{(j)})_{i,j=1}^\infty$ is *i.i.d.* with standard normal distribution. When comparing $G(N, k)$ and $L(N, k)$, it is obvious that most of the variables will vanish. More precisely, repeating the computation done by Bodineau and Martin in Section 2 of [BM05], we get, by letting $B_{-1}^{(r)} = 0$,

$$\begin{aligned} & |G(N, k) - L(N, k)| \\ &= \left| \sup_{u \in U(N, k)} \sum_{r=1}^k [S_{\lfloor u_r \rfloor}^{(r)} - S_{\lfloor u_{r-1} \rfloor - 1}^{(r)}] - \sup_{u' \in U(N, k)} \sum_{r=1}^k [B_{u'_r}^{(r)} - B_{u'_{r-1}}^{(r)}] \right| \\ &\leq \sup_{u \in U(N, k)} \sum_{r=1}^k \left[\left| S_{\lfloor u_r \rfloor}^{(r)} - B_{\lfloor u_r \rfloor}^{(r)} \right| + \left| S_{\lfloor u_{r-1} \rfloor - 1}^{(r)} - B_{\lfloor u_{r-1} \rfloor - 1}^{(r)} \right| \right. \\ &\quad \left. + \left| B_{\lfloor u_r \rfloor}^{(r)} - B_{u_r}^{(r)} \right| + \left| B_{\lfloor u_{r-1} \rfloor - 1}^{(r)} - B_{u_{r-1}}^{(r)} \right| \right] \\ &\leq 2 \sum_{r=1}^k \left(\max_{i=1, \dots, N} |S_i^{(r)} - B_i^{(r)}| \right) + 2 \sum_{r=1}^k \left(\sup_{\substack{0 \leq s, t \leq N \\ |s-t| < 2}} |B_s^{(r)} - B_t^{(r)}| \right). \end{aligned} \quad (4.15)$$

Denote by Y_k and Z_k respectively the two terms on the right-hand side of the last line in (4.15). Then we have,

$$\begin{aligned} \mathbb{P}[B] &\leq \mathbb{P}[Y_k + Z_k \geq 2\sqrt{Nk}(\varepsilon - \varepsilon_1)] \\ &\leq \mathbb{P}[Y_k \geq \sqrt{Nk}(\varepsilon - \varepsilon_1)] + \mathbb{P}[Z_k \geq \sqrt{Nk}(\varepsilon - \varepsilon_1)]. \end{aligned} \quad (4.16)$$

Of course, $Y_k = 0$ in this Gaussian example, but (4.16) will be used below for more general variables. Applying the Markov inequality gives for all $\lambda > 0$,

$$\begin{aligned}
\mathbb{P}[Z_k \geq \sqrt{Nk}(\varepsilon - \varepsilon_1)] &\leq \mathbb{E}[\exp(\lambda Z_k^2)] \cdot \exp(-\lambda(\varepsilon - \varepsilon_1)^2 Nk) \\
&\leq \mathbb{E}\left[\exp\left(4\lambda k \left(\sup_{\substack{0 \leq s, t \leq N \\ |s-t| < 2}} |B_s^{(1)} - B_t^{(1)}|\right)^2\right)\right]^k \cdot \exp(-\lambda(\varepsilon - \varepsilon_1)^2 Nk) \\
&\leq \frac{\left(\int_0^\infty 8\lambda k u \cdot \exp(4\lambda k u^2) \cdot \mathbb{P}\left[\sup_{\substack{0 \leq s, t \leq N \\ |s-t| < 2}} |B_s^{(1)} - B_t^{(1)}| \geq u\right] du\right)^k}{\exp(\lambda(\varepsilon - \varepsilon_1)^2 Nk)}.
\end{aligned} \tag{4.17}$$

However, for every $u \geq 0$,

$$\begin{aligned}
\mathbb{P}\left[\sup_{\substack{0 \leq s, t \leq N \\ |s-t| < 2}} |B_s^{(1)} - B_t^{(1)}| \geq u\right] &\leq \sum_{i=0}^{N-3} \mathbb{P}\left[\sup_{i \leq t \leq i+3} B_t - \inf_{i \leq t \leq i+3} B_t \geq u\right] \\
&\leq N \mathbb{P}\left[\sup_{0 \leq t \leq 3} |B_t| \geq u/2\right].
\end{aligned}$$

By the Brownian motion reflection principle (see for example [RY99]), $\sup_{0 \leq t \leq a} B_t \stackrel{d}{=} |B_a|$. Thus,

$$\begin{aligned}
\mathbb{P}\left[\sup_{\substack{0 \leq s, t \leq N \\ |s-t| < 2}} |B_s^{(1)} - B_t^{(1)}| \geq u\right] &\leq 4N \mathbb{P}[B_3 \geq u/2] \\
&\leq C_1 \frac{N}{u} \exp\left(-\frac{u^2}{C_1}\right).
\end{aligned} \tag{4.18}$$

where C_1 is a numerical positive constant. Now, insert (4.18) in the integral in (4.17) and choose $\lambda = \frac{c}{k}$ where c is a positive constant smaller than $\frac{1}{4C_1}$. Then we get for some constant $C_2 > 0$,

$$\mathbb{P}[Z_k \geq \sqrt{Nk}(\varepsilon - \varepsilon_1)] \leq C_2 \exp(k \log(C_2 N)) \cdot \exp\left(-\frac{(\varepsilon - \varepsilon_1)^2 N}{C_2}\right). \tag{4.19}$$

The last bound leads to the condition $k = o(N/\log N)$. Combining (4.12), (4.13), (4.14) and (4.19) then leads to

$$\mathbb{P}[A] \leq \mathbb{P}[L(1, k) \geq 2\sqrt{k}(1 + \varepsilon_1)] + C_2 \exp\left(k \log(C_2 N) - \frac{(\varepsilon - \varepsilon_1)^2 N}{C_2}\right) \tag{4.20}$$

and

$$\mathbb{P}[A] \geq \mathbb{P}[L(1, k) \geq 2\sqrt{k}(1 + 2\varepsilon - \varepsilon_1)] - C_2 \exp\left(k \log(C_2 N) - \frac{(\varepsilon - \varepsilon_1)^2 N}{C_2}\right). \tag{4.21}$$

Dividing (4.20) by $e^{-kJ_{\text{GUE}}(\varepsilon_1)}$ and (4.21) by $e^{-kJ_{\text{GUE}}(2\varepsilon - \varepsilon_1)}$, taking their logarithm and then dividing the results by k , we get for $\alpha < \frac{1}{2}$,

$$\frac{1}{k} \log \left(\frac{\mathbb{P}[A]}{e^{-kJ_{\text{GUE}}(\varepsilon_1)}} \right) \leq \frac{1}{k} \log \left(\frac{\mathbb{P}[L(1, k) \geq 2\sqrt{k}(1 + \varepsilon_1)]}{e^{-kJ_{\text{GUE}}(\varepsilon_1)}} + g_k(\varepsilon_1, \varepsilon) \right) \quad (4.22)$$

and

$$\frac{1}{k} \log \left(\frac{\mathbb{P}[A]}{e^{-kJ_{\text{GUE}}(2\varepsilon - \varepsilon_1)}} \right) \geq \frac{1}{k} \log \left(\frac{\mathbb{P}[L(1, k) \geq 2\sqrt{k}(1 + 2\varepsilon - \varepsilon_1)]}{e^{-kJ_{\text{GUE}}(2\varepsilon - \varepsilon_1)}} - g'_k(\varepsilon_1, \varepsilon) \right) \quad (4.23)$$

where

$$g_k(\varepsilon_1, \varepsilon) = C_2 \exp \left(-\frac{(\varepsilon - \varepsilon_1)^2 N}{C_2} + k(\log(C_2 N) + J_{\text{GUE}}(\varepsilon_1)) \right)$$

and

$$g'_k(\varepsilon_1, \varepsilon) = C_2 \exp \left(-\frac{(\varepsilon - \varepsilon_1)^2 N}{C_2} + k(\log(C_2 N) + J_{\text{GUE}}(2\varepsilon - \varepsilon_1)) \right),$$

are two positive functions. Moreover, for k large enough, $g_k(\varepsilon_1, \varepsilon)$ and $g'_k(\varepsilon_1, \varepsilon)$ are negligible with respect to $e^{-\eta k}$ for every $\eta > 0$ since $k = o(N/\log N)$. Thus, using (4.3), a straightforward computation shows that the right-hand sides of (4.22) and (4.23) both converge to zero when $k \rightarrow \infty$. In other words, for $k = o(N/\log N)$ and $\varepsilon_1 < \varepsilon$,

$$\limsup_{k \rightarrow \infty} \frac{1}{k} \log \mathbb{P}[A] \leq -J_{\text{GUE}}(\varepsilon_1) \quad (4.24)$$

and

$$\liminf_{k \rightarrow \infty} \frac{1}{k} \log \mathbb{P}[A] \geq -J_{\text{GUE}}(2\varepsilon - \varepsilon_1). \quad (4.25)$$

Finally, notice that $J_{\text{GUE}}(\varepsilon)$ is a continuous function of $\varepsilon > 0$. It therefore follows from (4.24) and (4.25) that for every $\varepsilon > 0$,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathbb{P}[G(N, k) \geq 2\sqrt{Nk}(1 + \varepsilon)] = -J_{\text{GUE}}(\varepsilon).$$

The proof of the large deviation formula on the left of the mean is similar. Set now, for all $\varepsilon > 0$ and $\varepsilon_1 < \varepsilon$,

$$E = \{G(N, k) \leq 2\sqrt{Nk}(1 - \varepsilon)\}.$$

By the same arguments as before, we get

$$\mathbb{P}[E] \leq \mathbb{P}[L(N, k) \leq 2\sqrt{Nk}(1 - \varepsilon_1)] + \mathbb{P}[B] \quad (4.26)$$

and

$$\mathbb{P}[E] \geq \mathbb{P}[L(N, k) \leq 2\sqrt{Nk}(1 - 2\varepsilon + \varepsilon_1)] - \mathbb{P}[B]. \quad (4.27)$$

Furthermore, by (4.4),

$$\lim_{k \rightarrow \infty} \frac{1}{k^2} \log \left(\frac{\mathbb{P}[L(1, k) \leq 2\sqrt{k}(1 - \varepsilon_1)]}{e^{-k^2 J_{\text{GUE}}(\varepsilon_1)}} \right) = 0 \quad (4.28)$$

and

$$\lim_{k \rightarrow \infty} \frac{1}{k^2} \log \left(\frac{\mathbb{P}[L(1, k) \leq 2\sqrt{k}(1 - 2\varepsilon + \varepsilon_1)]}{e^{-k^2 I_{\text{GUE}}(2\varepsilon - \varepsilon_1)}} \right) = 0. \quad (4.29)$$

Using the same upper bound on $\mathbb{P}[B]$ and combining (4.26), (4.27), (4.28) and (4.29), one can easily deduce that, for $k = o(N^{1/2})$ and $\varepsilon_1 < \varepsilon$,

$$\limsup_{k \rightarrow \infty} \frac{1}{k^2} \log \mathbb{P}[E] \leq -I_{\text{GUE}}(\varepsilon_1)$$

and

$$\liminf_{k \rightarrow \infty} \frac{1}{k^2} \log \mathbb{P}[E] \geq -I_{\text{GUE}}(2\varepsilon - \varepsilon_1).$$

At this stage, let us assume that $I_{\text{GUE}}(\varepsilon)$ is a continuous function of ε . Then, for $k = o(N^{1/2})$,

$$\lim_{N \rightarrow \infty} \frac{1}{k^2} \log \mathbb{P}[G(N, k) \leq 2\sqrt{N}k(1 - \varepsilon)] = -I_{\text{GUE}}(\varepsilon),$$

which is the result.

We are left with the proof of the continuity of $I_{\text{GUE}}(\varepsilon)$. Set $\mathcal{M}((-\infty, t])$, the set of all probability measures on $(-\infty, t]$ when $t \in \mathbb{R}$. For a given distribution $\mu \in \mathcal{M}((-\infty, t])$, define the corresponding potential energy, as in [ST97], by

$$I_\mu(t) = 2 \int_{-\infty}^t x^2 d\mu(x) - \int_{-\infty}^t \int_{-\infty}^t \log |x - y| d\mu(x) d\mu(y).$$

The minimal energy

$$I(t) = \inf_{\mu \in \mathcal{M}((-\infty, t])} I_\mu(t)$$

precisely allows us to compute the rate function I_{GUE} via the formula $I_{\text{GUE}}(\varepsilon) = I(1 - \varepsilon) - I(\infty)$. The last equality could be found in [Fér08]. For $t \geq 1$, $I(t)$ is a constant function, the extremal measure is the so-called semi-circular law supported on $[-1, 1]$ and the energy $I(t) = \log(2) + 3/4$, (cf. [BAG97, ST97]). For each $t \in \mathbb{R}$, there is a unique measure $\nu_t \in \mathcal{M}((-\infty, t])$, with no mass point, achieving the infimum (cf. [ST97]). Furthermore, ν_t is compactly supported and the corresponding energy is finite. Since $I(t)$ is an infimum and a non-increasing function of t , for any $\eta > 0$,

$$I(t) \leq I(t - \eta) \leq \frac{I_{\nu_t}(t - \eta)}{\nu_t^2((-\infty, t - \eta])}. \quad (4.30)$$

It is obvious that the right-hand side of (4.30) converges to $I(t)$ when η converges to zero. This proves the left-continuity of $I(t)$.

To show the right-continuity, notice that by a simple change of variable,

$$I(t) = \inf_{\mu \in \mathcal{M}((-\infty, t + \eta])} I_\mu^\eta(t)$$

where

$$I_\mu^\eta(t) = 2 \int_{-\infty}^{t+\eta} (x - \eta)^2 d\mu(x) - \int_{-\infty}^{t+\eta} \int_{-\infty}^{t+\eta} \log |x - y| d\mu(x) d\mu(y).$$

Consequently,

$$\begin{aligned} I(t) - I(t + \eta) &\leq I_{\nu_{t+\eta}}^\eta(t) - I(t + \eta) \\ &\leq 2\eta^2 + 4\eta \int_{-\infty}^{t+\eta} |x| d\nu_{t+\eta}(x). \end{aligned} \quad (4.31)$$

For $|x| \geq |y|$, we have $\log|x - y| \leq \log|2x|$. Moreover, there is a positive constant C_3 such that $|x| \leq C_3(2x^2 - 2\log|2x|)$. In view of (4.31), a straightforward calculation leads to

$$I(t) - I(t + \eta) \leq 2\eta^2 + 4\eta C_3 I(t). \quad (4.32)$$

Since $I(t)$ is finite, the right-hand side of (4.32) converges to zero when $\eta \rightarrow 0$. Thus, the continuity of $I(t)$ is proved, and that of $I_{\text{GUE}}(\varepsilon)$ as well. The proof of Theorem 4.1 is now complete. □

4.3 Exponential-tailed distribution and the K-M-T approximation

In this section, we replace the standard normal variables with weights having a finite moment-generating function around zero. When comparing $G(N, k)$ to the Brownian last-passage percolation model, Y_k will not vanish as in the Gaussian case where the coupling was “perfect”. In the non-Gaussian case, a new coupling is required, and the things become more complicated. Following [BM05], we make use of the K-M-T approximation: a powerful tool to couple a partial sum of *i.i.d.* random variables and a Wiener process, both constructed on the same probability space. The K-M-T approximation was first introduced in 1975 by Komlós, Major and Tusnády in their famous work [KMT76]. The basic version deals with a partial sum of *i.i.d.* random variables reconstructed in a way to be “close” to a partial sum of *i.i.d.* standard normal random variables. Later versions of this strong approximation do not require a common distribution, see [Sak91]. The reader can also consult [CH84] for a complete survey.

Let $(X_i)_{i \geq 1}$ be a sequence of independent random variables and denote by S_N the corresponding partial sum. Let $(B_t)_{t \geq 0}$ be a Brownian motion built on the same probability space. The following theorem is an immediate consequence of Theorem 1 in [KMT76].

Theorem 4.1 (Komlós-Major-Tusnády). *Assume that $\mathbb{E}X_1 = 0$ and $\mathbb{E}X_1^2 = 1$. Assume further that there exist $\mu_0 > 0$ such that for all $\mu < \mu_0$,*

$$\mathbb{E} \exp(\mu|X_1|) < +\infty.$$

Then for every $N \geq 1$, the sequence $(X_i)_{i \geq 1}$ and the Brownian motion $(B_t)_{t \geq 0}$ can be constructed in such a way that for all $x > 0$,

$$\mathbb{P}\left[\max_{i=1, \dots, N} |S_i - B_i| > \Theta \log N + x\right] \leq C \exp(-\theta x).$$

The positive constants Θ , C and θ depend only on the distribution of X_1 and θ can be taken as large as desired by choosing Θ large enough.

We now establish, with this tool, Theorem 4.2. According to the notation and the steps in Section 4.2, recall from (4.15) that

$$|G(N, k) - L(N, k)| \leq Y_k + Z_k.$$

We control Z_k as in (4.19) and we want Y_k to be as small as possible. To this purpose, we construct the sequence $(X_i^{(j)})_{i,j=1}^\infty$ and the independent Brownian motions $(B_t^{(r)})_{t \geq 0}$ in the sense of Theorem 4.1. By the Markov inequality, we have for all $0 < \varepsilon_1 < \varepsilon$ and $\lambda > 0$,

$$\begin{aligned} \mathbb{P}[Y_k \geq \sqrt{Nk}(\varepsilon - \varepsilon_1)] &\leq \mathbb{E}[\exp(\lambda Y_k)] \cdot \exp(-\lambda(\varepsilon - \varepsilon_1)\sqrt{Nk}) \\ &\leq \mathbb{E}\left[\exp\left(2\lambda \max_{i=1, \dots, N} |S_i^{(1)} - B_i^{(1)}|\right)\right]^k \cdot \exp(-\lambda(\varepsilon - \varepsilon_1)\sqrt{Nk}) \\ &\leq \frac{\left(\int_0^\infty 2\lambda \exp(2\lambda u) \cdot \mathbb{P}\left[\max_{i=1, \dots, N} |S_i^{(1)} - B_i^{(1)}| \geq u\right] du\right)^k}{\exp(\lambda(\varepsilon - \varepsilon_1)\sqrt{Nk})}. \end{aligned}$$

In order to apply Theorem 4.1, we make the simple variable change $t = u - \Theta \log N$ and we choose $\lambda < \theta/2$. Therefore, there exists as constant $C_4 > 0$ such that,

$$\mathbb{P}[Y_k \geq \sqrt{Nk}(\varepsilon - \varepsilon_1)] \leq C_4 \exp\left(k \log(C_4 N) - \frac{(\varepsilon - \varepsilon_1)\sqrt{Nk}}{C_4}\right). \quad (4.33)$$

Now, putting (4.12), (4.13), (4.19) and (4.33) together gives,

$$\begin{aligned} \mathbb{P}[A] &\leq \mathbb{P}[L(N, k) \geq 2\sqrt{Nk}(1 + \varepsilon_1)] + C_4 \exp\left(k \log(C_4 N) - \frac{(\varepsilon - \varepsilon_1)^2 N}{C_4}\right) \\ &\quad + C_4 \exp\left(k \log(C_4 N) - \frac{(\varepsilon - \varepsilon_1)\sqrt{Nk}}{C_4}\right) \end{aligned} \quad (4.34)$$

and

$$\begin{aligned} \mathbb{P}[A] &\geq \mathbb{P}[L(N, k) \geq 2\sqrt{Nk}(1 + 2\varepsilon - \varepsilon_1)] - C_4 \exp\left(k \log(C_4 N) - \frac{(\varepsilon - \varepsilon_1)^2 N}{C_4}\right) \\ &\quad - C_4 \exp\left(k \log(C_4 N) - \frac{(\varepsilon - \varepsilon_1)\sqrt{Nk}}{C_4}\right). \end{aligned} \quad (4.35)$$

On the left of the mean, we have

$$\begin{aligned} \mathbb{P}[E] &\leq \mathbb{P}[L(N, k) \leq 2\sqrt{Nk}(1 - \varepsilon_1)] + C_4 \exp\left(k \log(C_4 N) - \frac{(\varepsilon - \varepsilon_1)^2 N}{C_4}\right) \\ &\quad + C_4 \exp\left(k \log(C_4 N) - \frac{(\varepsilon - \varepsilon_1)\sqrt{Nk}}{C_4}\right) \end{aligned} \quad (4.36)$$

and

$$\begin{aligned} \mathbb{P}[E] &\geq \mathbb{P}[L(N, k) \leq 2\sqrt{Nk}(1 - 2\varepsilon + \varepsilon_1)] - C_4 \exp\left(k \log(C_4 N) - \frac{(\varepsilon - \varepsilon_1)^2 N}{C_4}\right) \\ &\quad - C_4 \exp\left(k \log(C_4 N) - \frac{(\varepsilon - \varepsilon_1)\sqrt{Nk}}{C_4}\right). \end{aligned} \quad (4.37)$$

Proceeding like in Section 4.2, we divide (4.34) by $e^{-kJ_{\text{GUE}}(\varepsilon_1)}$, (4.35) by $e^{-kJ_{\text{GUE}}(2\varepsilon-\varepsilon_1)}$, (4.36) by $e^{-k^2I_{\text{GUE}}(\varepsilon_1)}$ and (4.37) by $e^{-k^2I_{\text{GUE}}(2\varepsilon-\varepsilon_1)}$. To handle the remaining parts when $k \rightarrow \infty$, we take k negligible with respect to N . On the right of the mean, we need $k = o\left(\frac{N}{(\log N)^2}\right)$ and on the left, we take $k = o(N^{1/3})$. Finally we conclude as in section 4.2 using the continuity of J_{GUE} and I_{GUE} .

□

4.4 Subexponential weights and the Skorokhod embedding

In this section, we consider a particular category of subexponential weights such that $\mathbb{E} \exp(\mu|X|^\gamma)$ is finite for some $\mu > 0$ and $\gamma \in (0, 1)$. We say that these variables have a Weibull-like tails because of the similarity with the right-tail of the Weibull distribution with a shape parameter lower than 1. Such weights are considered to be heavy-tailed and then, do not have finite moment-generating functions. However, $G(N, k)$ still satisfies an LDP principle on a very thin rectangle. The precise definition of a subexponential distribution and large deviations for a partial sum of *i.i.d.* such weights can be found in [MN98].

We do not know a K-M-T strong approximation version for random variables satisfying the moment condition above. We therefore follow [BS05] by using the Skorokhod embedding theorem instead. The Skorokhod embedding theorem is another tool to couple a sum of *i.i.d.* random variables with a Brownian motion, see (cf. [Bre92, Saw72, Sko65]) for more details concerning this theorem.

Theorem 4.1 (Skorokhod). *Let $(B_t)_{t \geq 0}$ be a standard one-dimensional Brownian motion and X a real valued random variable satisfying $\mathbb{E}X = 0$ and $\mathbb{E}X^2 = 1$. Then, there is a stopping time T for the Brownian motion such that $B_T \stackrel{d}{=} X$ and $\mathbb{E}T = 1$.*

An immediate consequence of this theorem allows to embed sums of real independent random variables into the Brownian motion. Applying the strong Markov property to the Brownian motion, Theorem 4.1 yields the following classical corollary.

Corollary 4.2. *Let $X_1, X_2, \dots, X_N, \dots$ be *i.i.d.* satisfying $\mathbb{E}X_1 = 0$, $\mathbb{E}X_1^2 = 1$ and set $S_N = X_1 + X_2 + \dots + X_N$, $N \geq 1$. There is a sequence of *i.i.d.* stopping times $\tau_0 = 0, \tau_1, \dots, \tau_N, \dots$ such that*

$$S_N \stackrel{d}{=} B_{\tau_1 + \dots + \tau_N}$$

*and $(B_{\tau_1 + \dots + \tau_{N+1}} - B_{\tau_1 + \dots + \tau_N})_{N \geq 0}$ is a sequence of *i.i.d.* random variables having the same distribution as X_1 .*

In our context, an application of the last corollary allows to claim that there exists *i.i.d.* stopping times for the Brownian motion $\tau_0 = 0, \tau_1, \dots, \tau_N, \dots$ such that $\mathbb{E}\tau_1 = 1$ and $S_i^{(r)} \stackrel{d}{=} B_{\tau_1 + \dots + \tau_i}^{(r)}$ for $i \geq 1$ and $r \geq 1$. Consequently, choose

$$X_i^{(r)} = B_{\tau_1 + \dots + \tau_i}^{(r)} - B_{\tau_1 + \dots + \tau_{i-1}}^{(r)},$$

in order to have

$$S_i^{(r)} = B_{\tau_1 + \dots + \tau_i}^{(r)} \quad a.s.$$

Thus, for any $u \geq 0$,

$$\begin{aligned} & \mathbb{P} \left[\max_{i=1, \dots, N} |S_i^{(1)} - B_i^{(1)}| \geq u \right] \\ &= \mathbb{P} \left[\max_{i=1, \dots, N} |B_{\tau_1 + \dots + \tau_i}^{(1)} - B_i^{(1)}| \geq u; \max_{i=1, \dots, N} \left| \sum_{l=1}^i (\tau_l - 1) \right| \geq u^\beta \right] \\ &+ \mathbb{P} \left[\max_{i=1, \dots, N} |B_{\tau_1 + \dots + \tau_i}^{(1)} - B_i^{(1)}| \geq u; \max_{i=1, \dots, N} \left| \sum_{l=1}^i (\tau_l - 1) \right| < u^\beta \right]. \end{aligned}$$

Hence

$$\begin{aligned} \mathbb{P} \left[\max_{i=1, \dots, N} |S_i^{(1)} - B_i^{(1)}| \geq u \right] &\leq \mathbb{P} \left[\sup_{\substack{0 \leq s, t \leq N \\ |s-t| < u^\beta}} |B_s^{(1)} - B_t^{(1)}| \geq u \right] \\ &+ \mathbb{P} \left[\max_{i=1, \dots, N} \left| \sum_{l=1}^i (\tau_l - 1) \right| \geq u^\beta \right]. \end{aligned} \quad (4.38)$$

We evaluate each term of (4.38) separately. First,

$$\begin{aligned} & \mathbb{P} \left[\sup_{\substack{0 \leq s, t \leq N \\ |s-t| < u^\beta}} |B_s^{(1)} - B_t^{(1)}| \geq u \right] \\ &\leq \sum_{i=0}^{N-u^\beta} \mathbb{P} \left[\sup_{i \leq t \leq i+u^\beta+1} B_t - \inf_{i \leq t \leq i+u^\beta+1} B_t \geq u \right] \\ &\leq N \mathbb{P} \left[\sup_{0 \leq t \leq u^\beta+1} |B_t| \geq u/2 \right]. \end{aligned}$$

Applying the reflection principle as in Section 4.2, we get for some constant $C_5 > 0$,

$$\begin{aligned} \mathbb{P} \left[\sup_{\substack{0 \leq s, t \leq N \\ |s-t| < u^\beta}} |B_s^{(1)} - B_t^{(1)}| \geq u \right] &\leq 4N \mathbb{P} \left[B_{u^\beta+1} \geq t/2 \right] \\ &\leq C_5 N u^{\frac{\beta}{2}-1} \exp \left(- \frac{u^{2-\beta}}{C_5} \right). \end{aligned} \quad (4.39)$$

To find an upper bound for the second term on the right-hand side of (4.38), we need a connection between the weight moments and those of the stopping times obtained by the Skorokhod embedding. Furthermore, we need to control the sum of the independent stopping times to reach an exponentially decaying inequality. When the weights are bounded for exemple, we can construct a stopping time with finite exponential moments. The sum is then controlled by the Bernstein inequality.

However, when X only satisfies (4.6), the Skorokhod stopping time does not necessarily have a finite exponential moment and thus the Bernstein inequality can not be applied. For example, in [Dav76], Davis found the best universal constant connecting the stopping time moments to those of the stopped Brownian motion. More precisely, if $(B_t)_{t>0}$ is a Brownian motion and τ is a stopping time, then there is a universal constant a_p such that, when $1 < p < \infty$ and $\mathbb{E}\tau^{p/2} < +\infty$,

$$a_p \mathbb{E}\tau^{p/2} \leq \mathbb{E}|B_\tau|^p. \quad (4.40)$$

Moreover, Davis proved that the best constant for $p = 2n$ ($n \in \mathbb{N}^*$) is z_{2n}^* which is the smallest positive zero of the Hermite polynomial of order $2n$. In [Bur73], this constant is shown to be $O((2n)^{-n})$. So unless $X = B_\tau$ is a bounded variable, τ can not have finite exponential moments. The constant above is universal but it could be sharpened for some particular stopping times. For example, considering the stopping time of the Skorokhod representation [Bre92, Sko65], Sawyer improved the constant a_p and established, in [Saw72], the following inequality.

Theorem 4.3. *Let X be a centered random variable such that*

$$\mathbb{E} \exp(\mu|X|^\gamma) < +\infty$$

for some $\gamma > 0$ and $\mu > 0$, and let τ be the corresponding stopping time of the Skorokhod representation. Set $\theta = \frac{\gamma}{2+\gamma}$ and $\nu = \mu^{1-\theta}$. Then,

$$\mathbb{E} \exp(\nu\tau^\theta) \leq \Phi_\gamma \mathbb{E} \exp(\mu|X|^\gamma),$$

for some positive constant Φ_γ depending only on γ .

Note that a similar exponential bound may be obtained from (4.40). However the cost is a worse constant μ .

Under the assumption (4.6) and in view of Theorem 4.3, the Bernstein inequality can not be applied to the sum of the independent stopping times because $\theta = \gamma/(2+\gamma) < 1$. To avoid this obstacle, we introduce the Fuk-Nagaev inequality [FN71] which requires less restrictive assumptions.

Theorem 4.4 (Fuk-Nagaev). *Let X_1, \dots, X_N be a sequence of real i.i.d. random variables satisfying $\mathbb{E}X_1 = 0$ and $\mathbb{E}X_1^2 = \sigma^2$. Then, for all $x > 0$ and $y > 0$,*

$$\mathbb{P} \left[\max_{1 \leq i \leq N} \left| \sum_{l=1}^i X_l \right| \geq x \right] \leq N \mathbb{P} \left[|X_1| > y \right] + 2 \exp \left(- \frac{x^2}{2(N\sigma^2 + xy/3)} \right). \quad (4.41)$$

We refer to [Fuk73] for more detail on this inequality. Recall now the second term of the right-hand side of (4.38) and consider the stopping times of the Skorokhod representation. Choosing $x = u^\beta$ and $y = u^\delta$ in Theorem 4.4 and applying Markov inequality to the first term of the right-hand side of (4.41), one has

$$\mathbb{P} \left[\max_{i=1, \dots, N} \left| \sum_{l=1}^i (\tau_l - 1) \right| \geq u^\beta \right] \leq C_6 N \exp \left(- \frac{\nu u^{\theta\delta}}{C_6} \right) + C_6 \exp \left(- \frac{u^{2\beta}}{C_6 \max \{N, u^{\beta+\delta}\}} \right). \quad (4.42)$$

Now, we apply Markov inequality to Y_k as in Section 4.3 and we get for some $\lambda > 0$ and $0 < \eta < 1$,

$$\begin{aligned} \mathbb{P}[Y_k \geq \sqrt{Nk}(\varepsilon - \varepsilon_1)] &\leq \mathbb{E}[\exp(\lambda Y_k^\eta)] \cdot \exp(-\lambda(\varepsilon - \varepsilon_1)^\eta (Nk)^{\eta/2}) \\ &\leq \frac{\left(\int_0^\infty 2^\eta \eta \lambda u^{\eta-1} \exp(2^\eta \lambda u^\eta) \mathbb{P}\left[\max_{i=1, \dots, N} |S_i^{(1)} - B_i^{(1)}| \geq u\right] du \right)^k}{\exp(\lambda(\varepsilon - \varepsilon_1)^\eta (Nk)^{\eta/2})}. \end{aligned} \quad (4.43)$$

Inserting (4.38) and (4.42) in (4.43) and choosing λ very small yield the following constraints on η , β , δ and θ .

$$\begin{cases} \eta < 2 - \beta \\ \eta < \beta - \delta \\ \eta < \theta \delta. \end{cases}$$

Straightforward computations lead us to choose $\beta = \frac{4\gamma+4}{3\gamma+2}$ and $\delta = \frac{2\gamma+4}{3\gamma+2}$ since $\theta = \frac{\gamma}{\gamma+2}$. Consequently, we obtain $\eta < \frac{2\gamma}{3\gamma+2}$. To get large deviation asymptotic formulas on the right and the left of the mean, we respectively need $\alpha < \frac{\eta}{2-\eta}$ and $\alpha < \frac{\eta}{4-\eta}$. This completes the proof of Theorem 4.3. □

4.5 Small and large deviations inequalities

Non-asymptotic bound on the right and the left of the mean is an immediate consequence of the corresponding bound for the GUE and the arguments developed in Sections 2 and 3. In particular, we use (cf.[Led07]) that there exists a positive constant C_7 such that, for any $\varepsilon > 0$,

$$\mathbb{P}[L(1, k) \geq 2\sqrt{k}(1 + \varepsilon)] \leq \exp(-k J_{\text{GUE}}(\varepsilon)) \leq C_7 \exp\left(-\frac{k \max(\varepsilon^2, \varepsilon^{3/2})}{C_7}\right). \quad (4.44)$$

On the left of the mean, deviation inequalities for the largest eigenvalue of the GUE for a given k are quite more complicated to prove. Ledoux and Rider obtained in a recent paper, [LR10], that the leftmost charge of the largest eigenvalue of a large set of random matrices behaves like the left tail of the corresponding Tracy-Widom law. More precisely, they get for all $0 < \varepsilon \leq 1$,

$$\mathbb{P}[L(1, k) \leq 2\sqrt{k}(1 - \varepsilon)] \leq C_7 \exp\left(-\frac{k^2 \varepsilon^3}{C_7}\right). \quad (4.45)$$

As we mentioned before, when $\varepsilon > 1$, we have Gaussian behavior for both left and right tails. This follows from concentration arguments dealing with Lipschitz functions of independent standard normal variables. Once more, two cases will be tackled: Standard normal weights and finite exponential moments ones.

4.5.1 Standard normal variables

We now prove Theorem 4.4. Following the proof of Theorem 4.1 in Section 4.2, choose $\varepsilon_1 = \frac{\varepsilon}{2}$. Then, combining (4.20) and (4.44), for any $\varepsilon > 0$,

$$\mathbb{P}[A] \leq C_8 \exp\left(-\frac{k \max(\varepsilon^{3/2}, \varepsilon^2)}{C_8}\right) + C_8 \exp\left(-\frac{\varepsilon^2 N - k \log N}{C_8}\right)$$

where $C_8 > 0$. In order to reach (4.10) when $0 < \varepsilon < 1$, we need a positive constant $C(\alpha) > C_8$, depending only on α , such that

$$C_\alpha \exp\left(-\frac{k\varepsilon^{3/2}}{C_\alpha}\right) \geq C_8 \exp\left(-\frac{\varepsilon^2 N - k \log N}{C_8}\right). \quad (4.46)$$

Taking the logarithm of (4.46), C_α has to satisfy

$$\log \frac{C_8}{C_\alpha} - k\varepsilon^{3/2} \left(\frac{N\varepsilon^{1/2}}{C_8 k} - \frac{\log N}{C_8 \varepsilon^{3/2}} - \frac{1}{C_\alpha} \right) \leq 0. \quad (4.47)$$

However, since $\mathbb{P}[A] \leq 1$, ε has to satisfy

$$k\varepsilon^{3/2} \geq 1. \quad (4.48)$$

Combining now (4.47) and (4.48), we finally get that C_α has to satisfy

$$\log \frac{C_8}{C_\alpha} + \frac{1}{C_\alpha} - \frac{N^{1-\frac{4\alpha}{3}} - N^\alpha \log N}{C_8} \leq 0. \quad (4.49)$$

Hence C_α exists and satisfies (4.49) only if $\alpha < \frac{3}{7}$. In that case, we make the reverse computation to conclude that

$$\mathbb{P}[G(N, k) \geq 2\sqrt{Nk}(1 + \varepsilon)] \leq 2C_\alpha \exp\left(\frac{-k\varepsilon^{\frac{3}{2}}}{C_\alpha}\right).$$

We make the same computations for the left-tail upper bound. Here, C_α has to satisfy

$$\log \frac{C_8}{C_\alpha} - k^2 \varepsilon^2 \left(\frac{N}{C_8 k^2} - \frac{\log N}{C_8 k \varepsilon^2} - \frac{\varepsilon}{C_\alpha} \right) \leq 0,$$

which finally gives

$$\log \frac{C_8}{C_\alpha} + \frac{1}{C_\alpha} - \frac{N^{1-2\alpha} - N^{\alpha/3} \log N}{C_8} \leq 0.$$

This proves Theorem 4.4. □

4.5.2 Finite moment-generating function case

We finally prove Theorem 4.5. Choosing $\varepsilon_1 = \frac{\varepsilon}{2}$ in (4.34) and taking into consideration (4.48), the inequalities (4.44) and (4.45) imply that there exists a positive constant C_9 depending on α and the distribution of X such that, for all $\varepsilon > 0$,

$$\begin{aligned} \mathbb{P}[A] \leq & C_9 \exp\left(-\frac{k\varepsilon^{3/2}}{C_9}\right) \left(1 + \exp\left(-\frac{N^{\frac{1}{2}-\frac{\alpha}{3}} - N^\alpha \log N - N^\alpha}{C_9}\right) \right. \\ & \left. + \exp\left(-\frac{N^{1-4\alpha/3} - N^\alpha \log N - N^\alpha}{C_9}\right) \right) \end{aligned}$$

and

$$\mathbb{P}[E] \leq C_9 \exp\left(-\frac{k^2 \varepsilon^3}{C_9}\right) \left(1 + \exp\left(-\frac{N^{\frac{1-3\alpha}{2}} - 1}{C_9}\right) + \exp\left(-\frac{N^{1-2\alpha} - N^{\frac{\alpha}{3}} - 1}{C_9}\right) \right).$$

This means that we have a right-tail bound for $\alpha < 3/7$ and a left-tail bound for $\alpha < 1/3$. The proof is complete and thus Theorem 4.5 is proved. □

Chapitre 5

Matrices aléatoires : matrices de covariance empirique

La théorie des matrices aléatoires est devenue au fil des années une branche incontournable des mathématiques. Ces matrices sont utilisées dans de nombreux domaines de la science, notamment la physique atomique et les statistiques multivariées. La théorie s'est développé considérablement durant les cinq dernières décennies et la recherche liée au sujet est très active. Mise à part le nombre considérable d'applications des matrices aléatoires, le développement de cette théorie a permis l'introduction d'outils mathématiques très puissants, tel que le problème de Riemann-Hilbert par exemple. Dans la littérature, on distingue deux grandes catégories de matrices aléatoires : les matrices de Wigner et les matrices de covariance empirique. Ces deux classes de matrices aléatoires ont évolué parallèlement en profitant chacune des techniques utilisées pour l'autre. Pourtant, c'est un quart de siècle qui sépare l'introduction de chacun de ces modèles.

En 1928, Wishart [Wis28] a explicité la densité jointe d'une matrice de covariance empirique à entrées Gaussiennes (matrice de Wishart) d'un échantillon k -varié de taille N . Bien qu'il y ait eu plusieurs travaux antérieurs liés au sujet, Wishart est le premier à considérer une matrice de taille quelconque. Ces travaux ont permis plus tard le développement de l'analyse en composante principale (ACP). Dans un contexte tout à fait différent, Wigner [Wig55] a introduit en 1955 une matrice aléatoire Hermitienne dont les entrées sont des variables aléatoires indépendantes. Les valeurs propres d'une telle matrice modélisaient les niveaux d'énergie de certains atomes à noyau lourd. Contrairement à Wishart, Wigner s'est intéressé aux matrices de grandes tailles et a cherché alors des propriétés asymptotiques du spectre. Après Wigner, la théorie des matrices aléatoires a progressé notamment avec les travaux de Dyson, Mehta et plein d'autres auteurs. Pour un aperçu historique et une bibliographie complète, on renvoie le lecteur vers les deux ouvrages majeurs qui traitent de la théorie des matrices aléatoires : le fameux livre de Mehta [Meh91] et le livre plus récent de Anderson, Guionnet et Zeitouni [AGZ10]. Ces deux livres sont plutôt orientés vers les matrices du GUE et plus généralement vers les matrices de Wigner mais parlent peu des matrices de covariance empirique vue la ressemblance dans les résultats. Pour ces dernières, il n'existe pas à notre connaissance un ouvrage résumant l'ensemble des résultats les concernant. Cependant, le lecteur pourra consulter le livre d'Anderson [And03] pour les applications en analyse de données, celui de Bai et Silverstein [BS10] pour les outils théoriques, le résumé de El Karoui [EK05] sur les applications des matrices de covariance aléatoires et l'habilitation à diriger des recherches de Pécé [Péc] pour les théorèmes les plus récents.

Dans cette thèse, on s'intéresse de plus près aux matrices de covariance empirique et à leurs applications. Ce chapitre sera entièrement dédié à ce type de matrice et on ne parlera que très peu des matrices de Wigner. La bibliographie et les résultats les plus importants concernant les matrices de covariance seront donnés au fur et à mesure. Dans le paragraphe suivant, on expose l'intérêt derrière l'étude de telles matrices ainsi que les résultats les plus célèbres. A noter qu'il existe souvent des résultats analogues pour les matrices de Wigner. Dans la deuxième partie, on examine en particulier les propriétés de grandes déviations liées à la plus grande valeur propre d'une matrice de covariance empirique. A la fin de ce chapitre, on parlera des difficultés rencontrées pour appliquer les techniques de comparaison du Chapitre 3 aux matrices aléatoires.

5.1 Matrices de covariance empirique et applications

Dans ce paragraphe, on étudie un type particulier de matrices aléatoires : les matrices de covariance empirique. Ce type de matrices aléatoires admet des applications multiples en statistiques (Analyse de données), en télécommunication, en finance, en génétique etc, [Joh01c, BNMD09, PPR06]. Wishart, [Wis28], était le premier à regarder ces matrices en dimensions quelconques. Etant donné un échantillon multivarié (k variables réelles) de taille N , il cherchait à estimer la variance de chacune des variables ainsi que les corrélations entre elles. Wishart parvient à expliciter la loi jointe de la matrice de covariance empirique en supposant que l'échantillon est formé de N Gaussiennes k -variées. Lorsque N est très grande par rapport à k fixée, la matrice de covariance empirique converge vers la matrice de covariance théorique des k variables en question. Par conséquent, cela permet d'avoir une bonne estimation des différentes variances et corrélations, [And03].

A l'époque, il était courant de supposer k petit par rapport à N considérant, qu'une expérience peut être répétée le nombre de fois que l'on souhaite. Mais ceci n'est pas vrai pour tous les phénomènes étudiés. Une expérience liant un très grand nombre de paramètre peut être difficilement réalisable car très coûteuse par exemple ou tout simplement à cause de la durée entre deux réalisations. Pour cette raison, les scientifiques de nos jours s'intéressent encore plus au cas où le nombre de variables et la taille de l'échantillon sont du même ordre, (voir [Joh01c, EK05]). Dans ce cas, la matrice de covariance empirique n'est guère une bonne approximation de la matrice de covariance. La convergence de la mesure spectrale vers la loi de Marchenko-Pastur en est une bonne preuve. Pour cela, les mathématiciens ont cherché à connaître le comportement du spectre d'une matrice de covariance empirique lorsque N et k sont proportionnelles. Les données récoltées sur le comportement asymptotique de ce spectre fournissent de multiples informations sur la matrice de covariance théorique.

Soit $\tilde{C} = (\tilde{C}_{ij})$ une matrice $k \times N$ dont les entrées sont des variables aléatoires réelles (resp. complexes) centrées et réduites. Pour $j = 1, \dots, N$, les vecteurs aléatoires $\tilde{C}_j = (\tilde{C}_{1j}, \dots, \tilde{C}_{kj})$ représentent N réalisations indépendantes d'une expérience k -variée donnée. Il est donc naturel de supposer les \tilde{C}_j indépendants et de même loi. On note alors Σ_k , $k \times k$ symétrique (resp. Hermittienne), la matrice déterministe de covariance de C_1 et $l = (l_1, \dots, l_k)$ les valeurs propres de Σ_k . On considère la matrice $W = \frac{1}{N} \tilde{C} \tilde{C}^*$ où \tilde{C}^* représente la transposée (resp. transconjugée) de \tilde{C} et on désigne par $\lambda = (\lambda_1 > \dots > \lambda_k)$ le spectre de W . W est la matrice de covariance empirique de C_1 . Dans la pratique, les expérimentateurs connaissent W et espèrent, en l'analysant, tirer le

plus d'informations possibles sur Σ_k .

Lorsque \widetilde{C}_1 est un vecteur Gaussien, il existe une matrice $k \times N$ $C = (C_{ij})_{i,j}$ dont les entrées sont des Gaussiennes centrées, réduites et indépendantes telle que $W = \frac{1}{N} \Sigma_k^{\frac{1}{2}} C C^* \Sigma_k^{\frac{1}{2}}$. Ceci est dû au fait que la loi d'un vecteur Gaussien est entièrement déterminée par sa matrice de covariance. Mais cela ne reste pas forcément vrai pour des entrées non Gaussiennes et par suite, l'étude de la matrice W peut s'avérer très compliquée. Pour cette raison, on suppose que les variables du vecteur \widetilde{C}_1 dépendent linéairement d'une famille de k variables indépendantes. Plus précisément, on suppose pour tout $j = 1, \dots, N$, que $\widetilde{C}_j = \Sigma_k^{\frac{1}{2}} C_j$ où $C_j = (C_{1j}, \dots, C_{kj})$ est un vecteur aléatoire formé par des variables aléatoires indépendantes, centrées et réduites. Cette contrainte réduit considérablement le choix de la loi du vecteur \widetilde{C}_1 puisque qu'on impose que la décorrélation de \widetilde{C}_1 entraîne l'indépendance de ces composantes. Σ_k est bien évidemment la matrice de covariance de \widetilde{C}_1 et on a que

$$W = \frac{1}{N} \Sigma_k^{\frac{1}{2}} C C^* \Sigma_k^{\frac{1}{2}}$$

avec $C = (C_j)_{j=1}^N$. Lorsque Σ_k vaut l'identité, on parle d'une matrice de covariance non-perturbée (null case ou modèle blanc), sinon on parle de matrice perturbée (spiked population ou modèle non-blanc). **Dans la suite, on suppose que $N = N(k)$ dépend de k et on note $\gamma_k = k/N$.**

Définition 5.1. *Dans la suite, on appelle matrice de covariance empirique réelle (resp. complexe) la matrice*

$$W = \frac{1}{N} \Sigma_k^{\frac{1}{2}} C C^* \Sigma_k^{\frac{1}{2}}$$

où C est une matrice de taille $k \times N$ dont les colonnes C_j sont des vecteurs aléatoires de \mathbb{R}^k (resp. \mathbb{C}^k) i.i.d. centrées et réduites. Σ_k est une matrice symétrique (resp. Hermitienne) définie positive.

Définition 5.2. *On appelle matrice de Wishart réelle (resp. complexe) toute matrice de covariance empirique dont les entrées sont des Gaussiennes réelle (resp. complexe).*

L'étude des matrices de covariance empirique est notamment en rapport avec l'Analyse en Composante Principale (ACP). Les résultats obtenus sur le spectre et surtout sur les valeurs propres extrémales de W ont permis d'établir des tests statistiques concernant la matrice Σ_k . Dans la pratique, le but est de savoir si Σ_k vaut l'identité ou sinon, de connaître les directions (composantes) qui résument au mieux l'information contenue dans la matrice des données \widetilde{C} . Ces directions principales ne sont que les vecteurs propres correspondant aux valeurs propres maximales (celles qui se détachent du lot) de W . Quand k est fixée, il suffit de choisir N suffisamment grand pour que W soit une bonne estimation de Σ_k du point de vue spectral, [And03]. Mais lorsqu'on réalise des expériences qui lient un très grand nombre de paramètres, k devient très proche voire plus grand que N et l'étude devient plus délicate. Ce cas de figure existe dans l'étude climatique par exemple. Le lecteur pourra consulter [Joh01c] et [EK05] et les références de ces articles pour d'autres exemples d'applications. Des applications en télécommunications

et en génétique sont respectivement présentées dans [BNMD09, PPR06]. Plusieurs questions se posent alors naturellement : est-ce qu'on est capable, en regardant le spectre de W , de savoir si Σ_k est l'identité ou pas ? De plus, lorsque Σ_k n'est pas l'identité, est-ce possible de dégager les directions principales ? En d'autres termes, existe-t-il des liens explicites entre les λ_i et les l_i ? Des réponses complètes ou partielles à ces questions sont données dans les deux parties qui suivent.

Remarque 5.3. En réalité, les statisticiens s'intéressent à la matrice empirique de la forme

$$W = \frac{1}{N-1}(\tilde{C} - \bar{\tilde{C}})(\tilde{C}^* - \bar{\tilde{C}}^*).$$

Cependant, l'étude d'une telle matrice aléatoire est très compliquée ce qui nous oblige à supposer les entrées centrées.

Comme pour les matrices de Wigner, les travaux sur les matrices de covariance empirique ont duré plus d'un demi siècle et continuent encore. Ces travaux ont engendré une immense bibliographie qui couvre un très grand nombre de résultats signifiants. Les résultats obtenus concernent notamment la loi jointe des valeurs propres, la limite de la mesure spectrale, l'espacement des valeurs propres ainsi que les fluctuations et les grandes déviations de la plus grande valeur propre.

Dans le paragraphe suivant, on énonce les résultats les plus importants sur les matrices de covariance empirique tout en essayant de rester concis vu le nombre d'articles qui touchent au sujet. Notons que la plupart des théorèmes sont énoncés avec l'hypothèse $\gamma_k \rightarrow \gamma \leq 1$. Le cas $\gamma > 1$ en découle immédiatement du fait que $\Sigma_k^{1/2} C C^* \Sigma_k^{1/2}$ et $C^* \Sigma_k C$ ont les mêmes valeurs propres non nulles.

5.2 Résultats connus : modèle blanc

Les résultats sur les matrices aléatoires concernent essentiellement deux objets : l'intégralité du spectre représenté par ce qu'on appelle la mesure spectrale empirique et les valeurs propres extrémales. L'étude asymptotique de la mesure spectrale fournit une information très importante sur le comportement de la matrice aléatoire W . Néanmoins, une partie non négligeable de l'information est contenue aussi dans les valeurs propres extrémales. Ces valeurs sont quelque part mises à l'écart lorsqu'on examine la mesure spectrale. On fera donc un résumé des résultats en tenant compte de ces deux points de vue.

On distingue dans la littérature plusieurs types de matrices de covariance empirique qui correspondent à des types différents de données : on distingue notamment les matrices à entrées réelles et celles à entrées complexes, les matrices non-perturbées (modèle blanc) et celles avec perturbations (modèle non blanc) et enfin les matrices Gaussiennes et celles avec des entrées quelconques. On est donc face à huit cas distincts au minimum avec des résultats pour chacun d'entre eux. Afin de faciliter l'exposition, on énoncera les théorèmes dans le cas réel Gaussien ou complexe Gaussien et ensuite on donnera l'état des lieux de l'universalité. Le modèle non-blanc sera traité dans un paragraphe à part.

On suppose dans cette partie que $\Sigma_k = I_k$.

Loi jointe des valeurs propres : La densité jointe des valeurs propres d'une matrice de Wishart complexe (resp. réelle) a été donnée par James [Jam64]. Cette densité fournit une structure déterminantale aux matrices de Wishart et met en évidence l'ensemble polynomial orthogonal de Laguerre. Dans le cas réel, on parle de l'Ensemble Orthogonal de Laguerre (LOE) et de l'Ensemble Unitaire de Laguerre (LUE) dans le cas complexe. La formule suivante est le résultat de l'invariance par transformation unitaire de la densité de la matrice W de Wishart, [Wis28]

Proposition 5.1. *Soit W une matrice de Wishart complexe (resp. réelle) dont les entrées sont des Gaussiennes centrées et réduites. Si $k \leq N$, la densité jointe des valeurs propres est de la forme*

$$P_k(\lambda_1, \dots, \lambda_k) = \frac{1}{Z_k} \prod_{1 \leq i < j \leq k} |\lambda_i - \lambda_j|^\beta \prod_{i=1}^k \lambda_i^\alpha \exp\left(-k \frac{\beta}{2} \lambda_i\right), \quad (5.1)$$

avec $\beta = 2$ (resp. $\beta = 1$) et $\alpha = N - k$ (resp. $\alpha = (N - k - 1)/2$).

Notons que lorsque $\beta = 4$, l'expression 5.1 est la loi jointe des valeurs propres d'une matrice de Wishart symplectique (Laguerre Symplectic Ensemble LSE). Le cas symplectique n'est pas étudié dans cette thèse. Comme on l'a déjà expliqué dans le Chapitre 1, la représentation déterminantale de la loi jointe des valeurs propres est à l'origine de la preuve de la convergence des fluctuations de λ_{\max} vers la loi de Tracy-Widom. La forme de la loi jointe (en gaz de Coulomb) est également derrière l'étude des grandes déviations de la mesure spectrale. Cette forme n'a été établie pour le moment que dans le cas d'une Wishart réelle, complexe et symplectique.

Loi de Marchenko-Pastur : Lorsque N et k sont proportionnels, le spectre de W ne converge pas vers le singleton 1. Marchenko et Pastur ont démontré en 1967 dans le célèbre article [MP67] la convergence de la mesure spectrale d'une matrice de covariance empirique vers une loi à densité connue sous leur nom. Le résultat de Marchenko et Pastur est d'autant plus important qu'il a été démontré dans un cadre très général. Des améliorations sur les conditions de moments ont été apportées ultérieurement. Voici le théorème dans le cas d'une matrice de Wishart Complexe.

Théorème 5.2 (Marchenko-Pastur [MP67]). *Soit W une matrice de Wishart complexe et soit μ_p la mesure spectrale correspondante. On suppose que $0 < \lim_{k \rightarrow \infty} \gamma_k = \gamma \leq 1$. Alors*

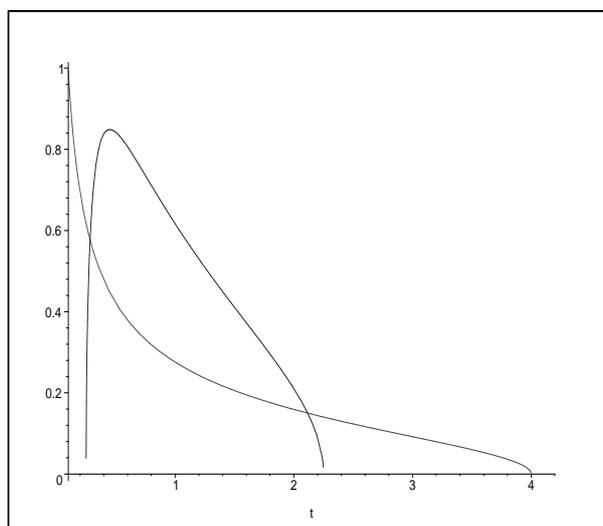
$$\mu_k \xrightarrow{P} \mu_{\text{mp}},$$

où μ_{mp} est la loi de Marchenko-Pastur admettant pour densité sur \mathbb{R}

$$f(t) = \frac{1}{2\gamma\pi t} \sqrt{(b-t)(t-a)} \mathbf{1}_{[a,b]}(t)$$

avec $a = (1 - \sqrt{\gamma})^2$ et $b = (1 + \sqrt{\gamma})^2$.

On remarque que le spectre est de plus en plus étalé quand γ s'approche de 1. Quand γ tend vers 0, le spectre se condense autour du singleton 1, (voir Figure 5.1). Lorsque $\gamma > 1$ ($k > N$), on a $k - N$ valeurs propres nulles. Dans ce cas, une partie de la masse est au point 0 (Dirac en 0) et une autre est répartie suivant la densité f . Comme pour la convergence vers la loi du demi-cercle, il existe plusieurs preuves du Théorème 5.2 : la méthode des moments, la convergence de la transformée de Stieltjes et la méthode de la mesure d'équilibre, voir Chapitre 1.

FIG. 5.1 – $\gamma = 1$ et $\gamma = 1/4$.

Universalité : Le Théorème 5.2 a été démontré dans [MP67] pour des variables aléatoires quelconques. Marchenko et Pastur ont supposé que les entrées C_{ij} sont *i.i.d.* avec $\mathbb{E}C_{11} = 0$, $\mathbb{E}C_{11}^2 = 1$ et $\mathbb{E}C_{11}^4 < \infty$. Plus tard, Yin [Yin86] a amélioré les conditions de moments en ne supposant que $\mathbb{E}C_{11} = 0$ et $\mathbb{E}C_{11}^2 = 1$. Wachter [Wac78] a étudié le cas des entrées indépendantes mais pas forcément identiquement distribuées. Il a imposé pour tout i, j que $\mathbb{E}C_{ij} = 0$, $\mathbb{E}C_{ij}^2 = 1$ et $\mathbb{E}|C_{11}|^{2+\varepsilon} \leq C < \infty$ pour un certain $\varepsilon > 0$.

Remarque 5.3. La condition de finitude du second moment des entrées est nécessaire pour la convergence de la mesure spectrale vers la loi de Marchenko-Pastur. Dans le cas d'une matrice de Wigner à entrées *i.i.d.*, Ben Arous et Guionnet [BAG08] ont démontré la convergence de la mesure spectrale vers une mesure à support non borné.

On se pose maintenant la question de savoir si la plus grande valeur λ_{\max} de W converge vers une limite fini. Yin, Bai et Krishnaiah [YBK88] ont démontré que la finitude du quatrième moment des entrées est une condition nécessaire et suffisante pour que λ_{\max} converge presque sûrement vers le bord droit du support de la loi de Marchenko-Pastur, *i.e.*

$$\mathbb{E}C_{11}^4 < \infty \Leftrightarrow \lambda_{\max} \xrightarrow{p.s.} (1 + \sqrt{\gamma})^2. \quad (5.2)$$

Dans le cas contraire, les auteurs ont démontré que la limite supérieure de λ_{\max} vaut l'infini. Le lecteur pourra consulter le résumé de Bai [Bai99] et ces références pour plus de détails.

Fluctuations de la plus grande valeur propre : Tracy et Widom ont introduit dans [TW94] la loi connue sous leur nom comme étant la limite des fluctuations de la plus grande valeur propre d'une matrice du GUE. Les auteurs de [BDJ99] ont établi un résultat similaire pour la plus longue sous-suite croissante d'une permutation aléatoire. Dans le même esprit, Johansson [Joh00a] a démontré, en utilisant le noyau de Meixner, la convergence des fluctuations du temps de dernier passage vers la loi de Tracy-Widom et ceci pour des variables géométriques. Par simple passage à la limite, il obtient un résultat similaire dans le cas exponentiel. Or, dans le

cas exponentiel, le temps de dernier passage est égal en loi à la plus grande valeur propre d'une matrice de Wishart complexe.

Théorème 5.4. *Soit W une matrice de Wishart complexe où $C_{11} \sim \mathcal{N}(0, \sigma^2)$. On note λ_{\max} la valeur propre maximale de W . Si $\gamma_k \rightarrow \gamma > 0$, alors on a que*

$$\frac{k^{\frac{2}{3}}(\lambda_{\max} - \mu_{Nk})}{\sigma_{Nk}} \xrightarrow{\mathcal{L}} \text{TW}_2,$$

où

$$\mu_{Nk} = \sigma^2(1 + \sqrt{\gamma_k})^2 \quad \text{et} \quad \sigma_{Nk} = \sigma^2(1 + \sqrt{\gamma_k})\left(1 + \frac{1}{\sqrt{\gamma_k}}\right)^{\frac{1}{3}}.$$

Johnstone [Joh01c] a démontré le même résultat pour les matrices de Wishart réelles alors que El Karoui [EK03] a étendu le résultat pour $\gamma \in [0, \infty]$. Dans le cas réel, la convergence à lieu vers la loi de TW_1 , voir [Joh01c] pour la définition exacte. Notons qu'il existe une version multidimensionnelle du Théorème 5.4 qui concerne les l plus grandes valeurs propres de W avec l fixée, voir [TW94].

Universalité : Comme pour les matrices de Wigner, les mathématiciens pensent que la condition des moments finis d'ordre 4 est nécessaire et suffisante pour la convergence vers la loi de Tracy-Widom dans le cas d'une matrice de covariance à entrées *i.i.d.*. Cette condition est nécessaire au vue de l'équivalence (5.2). Par contre, jusqu'au moment de la rédaction de cette thèse, la suffisance de cette condition n'a pas été démontrée. Plusieurs auteurs ont travaillé et travaillent encore pour améliorer les conditions sur les entrées de la matrice W . Voici, par ordre chronologique, les résultats d'universalité concernant la convergence vers la loi de Tracy-Widom :

Soshnikov était le premier à démontrer, dans un cadre universel, le Théorème 5.4 d'abord pour des matrices de Wigner et ensuite pour des matrices de covariance empirique, [Sos99, Sos02]. Son résultat est le suivant :

Théorème 5.5 (Soshnikov [Sos02]). *Soit W une matrice de covariance empirique complexe (resp. réelle). On suppose que les C_{ij} sont *i.i.d.*, centrées, réduites et de loi symétrique. De plus, on suppose que les C_{ij} sont sous-Gaussiennes, i.e. il existe une constante κ telle que pour tout l ,*

$$\mathbb{E}|C_{11}|^{2l} \leq (\kappa l)^l.$$

Si $N - k = O(k^{1/3})$, les fluctuations de λ_{\max} proprement normalisées convergent en loi vers TW_2 (resp. TW_1).

Ce théorème a été amélioré par Pécché [Péc09] qui a supprimé la condition de différence entre N et k . Dans le même article, Pécché s'est inspiré du travail de Ruzmaikina [Ruz06] sur les matrices de Wigner pour démontrer un résultat similaire dans le cas d'une matrice de covariance empirique :

Théorème 5.6 (Ruzmaikina [Ruz06], Pécché [Péc09]). *Soit W une matrice de covariance empirique complexe (resp. réelle). On suppose que les C_{ij} sont *i.i.d.*, centrées, réduites et de loi symétrique. De plus, on suppose qu'il existe $p > 36$ tel que*

$$\mathbb{P}[C_{11} \geq t] \leq \frac{1}{t^p}.$$

Alors les fluctuations de λ_{\max} proprement normalisées convergent en loi vers TW_2 (resp. TW_1).

Plus récemment, Tao et Vu [TV10, TV] ont investi les formules variationnelles de Hadamard pour démontrer l'universalité de plusieurs propriétés concernant les matrices de Wigner et de covariance empirique. Les formules de Hadamard sont des formules de dérivation qui décrivent explicitement les variations des valeurs propres d'une matrice Hermitienne en faisant varier l'une des entrées. A l'image de la preuve de Lindeberg du TCL classique, Tao et Vu [TV10] ont démontré une version universelle du Théorème 5.4 pour les matrices de Wigner. Tao et Vu démontrent un théorème général duquel découlent une multitude de propriétés universelles. En ce qui concerne les fluctuations de la plus grande valeur propre, à part la non-symétrie des entrées, leur résultat n'améliore en rien les conditions sur les entrées imposées par Soshnikov, Péché ou Ruzmaikina.

Grandes déviations de la plus grande valeur propre : L'étude des formules asymptotiques de grandes déviations pour la mesure spectrale d'une matrice de Wishart réelle ou complexe dépend essentiellement de la représentation sous forme d'un gaz de Coulomb de la densité du spectre. A l'image de ce qu'on fait les auteurs de [BAG97, BADG01] dans le cas d'une matrice du GUE, Hiai et Petz [HP98] ont d'abord fait l'étude pour les matrices de Wishart réelle et complexe ($k < N$ et $k \geq N$). Ils ont ensuite généralisé l'étude pour des gaz de Coulomb avec un potentiel continu quelconque [HP00a], voir Remarque 1.1.2. Pour un gaz de Coulomb continu, il est possible, dans certains cas encore mal identifiés, de déduire du principe de grandes déviations de la mesure spectrale des propriétés de grandes déviations pour la valeur propre maximale, (voir [Fér08]).

Dans un autre contexte, Johansson [Joh00a] a obtenu des formules de grandes déviations pour le temps de dernier passage avec des poids géométriques (gaz de Coulomb discret). Par simple passage à la limite, il a établi des propriétés similaires pour le cas exponentiel, et donc pour λ_{\max} .

Théorème 5.7. *Soit W une matrice de Wishart complexe. On suppose que $\gamma_k = \gamma \leq 1$. Alors pour tout $\varepsilon > 0$, on a*

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathbb{P}[\lambda_{\max} \geq (1 + \sqrt{\gamma})^2(1 + \varepsilon)] = -J_{\text{LUE}}(\varepsilon)$$

et

$$\lim_{k \rightarrow \infty} \frac{1}{k^2} \log \mathbb{P}[\lambda_{\max} \leq (1 + \sqrt{\gamma})^2(1 - \varepsilon)] = -I_{\text{LUE}}(\varepsilon),$$

où

$$J_{\text{LUE}}(\varepsilon) = \int_1^s (s - t) \frac{(1 + \gamma)t + 2\sqrt{\gamma}}{(t + B)^2 \sqrt{t^2 - 1}} dt$$

$$\text{avec } s = 1 + \frac{\varepsilon}{2\sqrt{\gamma}} \text{ et } B = \frac{1 + \gamma}{2\sqrt{\gamma}},$$

et

$$I_{\text{GUE}}(\varepsilon) = \inf_{\mu \in \mathcal{M}^1((0, (1 + \sqrt{\gamma})^2(1 - \varepsilon)))} I(\mu) - I(\mu_{mp})$$

$$\text{avec } I(\mu) = \int_{\mathbb{R}} \left[s - \left(\frac{1}{\gamma} - 1 \right) \log s \right] d\mu_x(s) + \int \int_{s \neq t} \log |t - s|^{-1} d\mu_x(s) d\mu_x(t).$$

A N et k fixés, Ledoux et Rider [LR10] ont obtenu des inégalités optimales avec des bornes supérieures du même ordre des queues de Tracy-Widom et qui vont aussi dans le sens des formules de grandes déviations précédentes.

Théorème 5.8 (Ledoux-Rider [LR10]). *Soit W une matrice de Wishart complexe telle que $N + 1 > k \geq 1$. Alors pour tout $0 < \varepsilon \leq 1$*

$$\mathbb{P}[\lambda_{\max} \geq (1 + \sqrt{\gamma k})^2 (1 + \varepsilon)] \leq C \exp \left(\frac{-2\sqrt{Nk} \varepsilon^{3/2} \left(\frac{1}{\sqrt{\varepsilon}} \wedge \left(\frac{N}{k} \right)^{\frac{1}{4}} \right)}{C} \right)$$

et

$$\mathbb{P}[\lambda_{\max} \leq (1 + \sqrt{\gamma k})^2 (1 - \varepsilon)] \leq C^2 \exp \left(\frac{-2Nk \varepsilon^3 \left(\frac{1}{\varepsilon} \wedge \left(\frac{N}{k} \right)^{\frac{1}{2}} \right)}{C} \right),$$

où C est une constante numérique qui ne dépend ni de k ni de ε .

Notons que lorsque $\varepsilon > \sqrt{\gamma k}$, les bornes supérieures sont respectivement en ε et ε^2 au lieu $\varepsilon^{3/2}$ et ε^3 . Ces inégalités permettent de démontrer la convergence des moments de λ_{\max} vers ceux de la loi de Tracy-Widom et d'avoir notamment des bornes sur la variance de λ_{\max} , voir ([LR10], Corollaire 3). En statistique, ce type d'inégalité est plus utile que les formules limites de grandes déviations du fait que N et k soient fixées et que les constantes numériques n'en dépendent pas.

Universalité : Si on remplace les entrées Gaussiennes par d'autres de loi quelconque, on perd la représentation en gaz de Coulomb de la loi jointe des valeurs propres. Par suite, lorsque N et k sont du même ordre, il semble difficile d'explicitier les fonctions de taux. Pour les inégalités à N et k fixées, la généralisation du Théorème 5.8 nécessite des outils de comparaison très puissants fournissant des erreurs exponentiellement petites et cela n'est pas à portée de mains. Ceci dit, si on suppose que $k = o(N)$, on pense sans pouvoir le démontrer, qu'on peut établir des formules de grandes déviations d'ordre \sqrt{Nk} avec les fonctions de taux du GUE à l'image de l'étude faite sur le LPP dans le Chapitre 3. Dans la Partie 5.5, on expose rapidement les obstacles qui empêchent une telle généralisation.

5.3 Résultats connus : modèle non-blanc

Lorsque les scientifiques regardent simultanément un certain nombre de variables, ils estiment que ces variables sont liées entre elles et cherchent à en dégager la corrélation. Pour cette raison, on s'attend souvent à ce que Σ_k ne soit pas l'identité. Dans cette partie, on résume les résultats connus sur les valeurs propres de W lorsque $\Sigma_k \neq I$. Dans un premier temps, le but est de pouvoir dire, en analysant W , si Σ_k vaut l'identité ou pas. Cela constitue un test d'indépendance des variables en étude. Ensuite, on cherche à calculer les plus grandes valeurs propres de Σ_k en fonction de celles de W . Ces valeurs propres correspondent aux directions principales ; celles qui résument une grande partie de l'information sur les variables.

Johnstone a proposé un modèle où les valeurs propres de Σ_k valent 1 sauf un nombre fini fixé, *i.e.* $\Sigma_k = \text{diag}(l_1, l_2, \dots, l_r, l_{r+1} = 1, \dots, l_k = 1)$. Cela veut dire que l'information sur les différentes corrélations se résume à un nombre fini de directions. On parle alors d'un modèle perturbé de rang r . Ce modèle, au vaste champ d'applications, a été soigneusement étudié durant la dernière décennie. Les mathématiciens ont examiné, comme pour le modèle blanc, la limite de la mesure spectrale, la limite des valeurs propres extrémales, les fluctuations et les grandes déviations de celles-ci. Je cite notamment les travaux de Bai et Silverstein [BS98], Baik, Ben Arous et Péché [BBAP05], Baik et Silverstein [BS06], El Karoui [EK07] et Bai et Yao [BY08].

Loi jointe des valeurs propres : On connaît la loi jointe d'une matrice de Wishart avec un Σ_k quelconques, [Wis28]. Ceci permet d'obtenir la loi jointe des valeurs propres sous forme appropriée, [Jam64].

Proposition 5.1 (James [Jam64]). *Soit W une matrice de Wishart ($k \leq N$) complexe centrée et Σ_k la matrice de covariance correspondante. On note $l = (l_1, \dots, l_k)$ le spectre de Σ_k , $D = \text{diag}(l)$ et $L = \text{diag}(\lambda)$. La densité jointe des valeurs propres de W s'écrit sous la forme*

$$P(\lambda) = \frac{1}{Z_k} V(\lambda)^2 \prod_{i=1}^k \lambda_i^{N-k} \int_{Q \in U(k)} \exp(-N \text{Tr}(D^{-1}QL^{-1}Q^{-1})) dQ, \quad (5.3)$$

où Z_k est une constante de normalisation, V le déterminant de Vandermonde et $U(k)$ l'ensemble des matrices unitaires de taille $k \times k$.

Dans le cas d'une matrice Wishart complexe, l'intégrale dans (5.3) peut être explicitée pour obtenir une formule déterminantale exploitable dans le calcul asymptotique des fluctuations, voir ([BBAP05], Section 2). Le lien avec la percolation de dernier passage avec des poids exponentiels est également expliqué dans la dernière partie du même article et dans [Joh08].

Loi limite de la mesure spectrale : On suppose dans cette partie que W est une matrice de covariance empirique avec des entrées *i.i.d.* centrées et réduites et que $\gamma_k \rightarrow \gamma \leq 1$. On suppose également que la mesure spectrale empirique de Σ_k converge vers une mesure de densité H sur \mathbb{R}_+ . Etant donnée une mesure μ sur \mathbb{R} , on définit la transformée de Stieltjes de μ

$$S_\mu(z) = \int_{\mathbb{R}} \frac{1}{t-z} d\mu(t), \quad z \in \mathbb{C}^+ := \{z \in \mathbb{C}; \text{Im } z > 0\}.$$

Marchenko et Pastur [MP67] et puis Bai et Silverstein [SB95] ont démontré le théorème suivant :

Théorème 5.2 (Marchenko-Pastur [MP67], Silverstein-Bai [SB95]). *La mesure spectrale μ_k de W converge p.s. vers la mesure déterministe à densité μ_{mp}^* définie de la manière suivante : $s = S_{\mu_{mp}^*}(z)$ est l'unique solution de l'équation*

$$s = \int_{\mathbb{R}_+} \frac{1}{t(1-\gamma-\gamma zt) - z} dH(t)$$

se trouvant dans l'ensemble $\{s \in \mathbb{C}; -\frac{1-\gamma}{z} + \gamma s \in \mathbb{C}^+\}$.

Bien que l'étude de la mesure spectrale est importante, celle-ci n'apporte pas l'information nécessaire pour effectuer des tests sur la nature de Σ_k . Les statisticiens cherchent alors du côté des valeurs propres extrémales. Ils ont remarqué que pour certaines données numériques, plusieurs

valeurs propres quittent le support de la mesure spectrale limite contrairement à ce qu'on a obtenu pour le modèle blanc. Plus précisément, on cherche à connaître le comportement des valeurs propres de W lorsque les valeurs propres correspondantes de Σ_k ne sont pas dans le support de H . Une réponse partielle a été donnée dans [BS98]. Le théorème suivant de Baik et Silverstein [BS06] complète l'analyse.

Théorème 5.3 (Baik-Silverstein [BS06]). *Soit W une matrice de covariance dont les entrées sont i.i.d. avec $\mathbb{E}C_{11} = 0$, $\mathbb{E}C_{11}^2 = 1$ et $\mathbb{E}C_{11}^4 < \infty$. On suppose que $\gamma_k \rightarrow \gamma \leq 1$ et que*

$$l = (l_1 \geq l_2 \geq \dots \geq l_r > l_{r+1} = 1, \dots, l_k = 1).$$

On note r_1 le nombre de $l_i > 1 + \sqrt{\gamma}$.

- Pour $1 \leq i \leq r_1$, on a

$$\lambda_i \xrightarrow{p.s.} l_i + \frac{\gamma l_i}{l_i - 1}.$$

- Pour $r_1 < i \leq k$, on a

$$\lambda_i \xrightarrow{p.s.} (1 + \sqrt{\gamma})^2.$$

Le théorème précédent nous dit en quelque sorte que si N et k sont du même ordre, il est difficile de détecter une faible corrélation entre les variables étudiées en regardant la limite des valeurs propres extrémales. Cependant, lorsque cette corrélation dépasse un certain point critique, certaines valeurs propres de W commencent à migrer loin du support rendant la corrélation détectable. Dans un article récent de Bai et Yao [BY08], les auteurs examinent la limite des valeurs propres maximales lorsque H est quelconque et le nombre de valeurs propres isolées est fini. Ils introduisent un critère de proximité entre les valeurs propres isolées de Σ_k et le support de H . Ce critère permet de savoir s'il va y avoir des valeurs propres isolées de W et vers quelles valeurs elles vont converger. Le théorème précédent est un cas particulier de ce résultat en prenant $H = \delta_1$.

Fluctuations de la plus grande valeur propre : Lorsque la limite de λ_{\max} sort du support de la loi de Marchenko-Pastur, les fluctuations de celle-ci sont de l'ordre de \sqrt{k} au lieu de $k^{1/3}$. La loi limite n'est plus celle de Tracy-Widom. Ce phénomène de transition de phase des fluctuations a été démontré pour la première fois par Baik, Ben Arous et Péché [BBAP05] dans le cas d'une matrice de Wishart complexe et puis par Paul [Pau07] pour une matrice de Wishart réelle.

Théorème 5.4 (Baik-Ben Arous-Péché [BBAP05], Paul [Pau07]). *Soit W une matrice de Wishart complexe et Σ_k la matrice de covariance correspondante. On fixe un entier $r \geq 0$ et on suppose que*

$$l_{r+1} = l_{r+2} = \dots = l_k = 1.$$

Lorsque $\gamma_k \rightarrow \gamma \leq 1$, les deux propositions suivantes sont vraies pour tout $x \in \mathbb{R}$.

1. Si pour tout $1 \leq i \leq r$, $l_i < 1 + \sqrt{\gamma}$, on a que

$$\mathbb{P}\left[\frac{k^{\frac{2}{3}}}{\sqrt{\gamma}(1 + \sqrt{\gamma})^{\frac{4}{3}}}(\lambda_{\max} - (1 + \sqrt{\gamma})^2) \leq t\right] \longrightarrow F_2^{\text{TW}}(t).$$

2. S'il existe un entier $1 \leq s \leq r$ tel que

$$l_1 = \dots = l_s = 1 + \sqrt{\gamma} \quad \text{et} \quad l_{s+1}, \dots, l_k \in (0, 1 + \sqrt{k}),$$

alors

$$\mathbb{P} \left[\frac{k^{\frac{2}{3}}}{\sqrt{\gamma}(1 + \sqrt{\gamma})^{\frac{4}{3}}} (\lambda_{\max} - (1 + \sqrt{\gamma})^2) \leq t \right] \longrightarrow F_{s+2}^{\text{TW}}(t).$$

3. S'il existe un entier $1 \leq s \leq r$ tel que

$$l_1 = \dots = l_s \in (1 + \sqrt{\gamma}, \infty) \quad \text{et} \quad l_{s+1}, \dots, l_k \in (0, l_1),$$

alors

$$\mathbb{P} \left[\frac{\sqrt{k}}{\sqrt{l_1^2 - \frac{\gamma l_1^2}{(l_1-1)^2}}} (\lambda_{\max} - (l_1 + \frac{\gamma l_1}{l_1-1})) \leq t \right] \longrightarrow G_s(t).$$

Ici F_{s+2}^{TW} est la fonction de répartition d'une variante de la loi de Tracy-Widom et $G_s(t)$ est la fonction de répartition de la plus grande valeur propre d'une matrice $s \times s$ du GUE.

La définition exacte de F_{s+2}^{TW} est présentée dans [BBAP05] et un résumé complet des résultats dans [FP09]. Notons que $G_1(t)$ est la fonction de répartition d'une variable Gaussienne. Les fluctuations « Gaussiennes » au-delà du point critique ont été également démontrées par Paul [Pau07] pour les matrices de Wishart réelles. Les techniques utilisées sont différentes de celles utilisées dans le cas complexe à cause de l'absence d'une représentation explicite de la loi jointe des valeurs propres. Les travaux de Féral et Pécché [FP09] sont les derniers en date sur le sujet. Elles ont démontré le théorème précédent pour des entrées complexes quelconques *i.i.d.* admettant des moments finis d'ordre 4. Dans le cas d'entrées réelles, elles ont démontré l'universalité des fluctuations au-delà et au dessous du point critique (le 1) et le 3) du théorème précédent).

Bai et Yao [BY08] ont étudié le cas des matrices de covariance avec des entrées quelconques et un H quelconque (nombre fini de valeurs propres isolées). Ils ont établi les fluctuations Gaussiennes au-delà d'un certain seuil critique qu'ils ont introduit. El Karoui [EK07] a étudié les matrices de Wishart complexe perturbées avec Σ_k quelconque. Il a établi un théorème de Tracy-Widom pour la plus grande valeur propre en imposant une borne supérieure sur la plus grande valeur propre de Σ_k .

Encore une fois, on est incapable de détecter les faibles corrélations en regardant le comportement asymptotique des fluctuations de la plus grande valeur propre. Dans la partie suivante, on étudie les grandes déviations de la plus grande valeur propre d'une matrice de covariance empirique perturbée.

Remarque 5.5. La liste des théorèmes donnés ici n'est pas exhaustive mais contient les résultats les plus importants. Dans cette thèse, on regarde surtout le comportement des valeurs propres extrêmes mais des résultats existent notamment sur l'espacement des valeurs propres et leur

comportement à l'intérieur du spectre, [Dei99, Joh01b, Gus05]. Il existe aussi des théorèmes du type Berry-Essen qui donnent la vitesse de convergence de la mesure spectrale vers la loi de Marchenko-Pastur, [GT04]. Des études ont également été faites sur la nature des vecteurs propres (localisés ou délocalisés) correspondant aux valeurs propres extrémales surtout dans le cas de matrice de Wigner perturbée.

5.4 Grandes déviations et matrices de covariance empirique.

Dans cette partie de la thèse, on fait le point sur les résultats obtenus sur les grandes déviations de la valeur propre maximale d'une matrice de covariance empirique perturbée. On expose d'abord deux résultats connus, le premier sur les matrices de Wishart non-blanches et un autre sur les matrices de covariances empirique dans un cadre général. Une fois le terrain préparé, on expose notre contribution.

La connaissance de formules explicites de grandes déviations d'une suite de variables aléatoires peut être utile dans l'étude de la performance des tests statistiques. Une illustration de ce fait est l'article récent de Bianchi et al [BDMN10] sur les tests d'indépendance effectués sur des matrices de covariances empirique. Les auteurs regardent une matrice de Wishart dont ils ignorent la matrice de covariance Σ_k correspondante. La matrice de Wishart en question contient des valeurs représentant des signaux reçus par k antennes à N fractions de temps successives. La moindre corrélation détectée par les tests utilisés indique la présence d'un signal primaire (signal attendu en dehors des bruits blancs). La plus grande valeur propre de la matrice de Wishart, notamment la transition de phase dont elle est sujette, est au cœur des tests proposés. En l'absence d'inégalités exponentielles à N et k fixés, les formules asymptotiques de grandes déviations de la plus grande valeur propre sont utilisées pour comparer les tests.

Le théorème suivant de Bianchi et al. repose sur des calculs faits par Maïda [Maï07] dans le cadre du GUE perturbé.

Théorème 5.1 (Bianchi et al. [BDMN10]). *Soit W une matrice de Wishart complexe. On suppose que $\gamma_k \rightarrow \gamma$ avec $0 < \gamma < 1$ et que $l_1 > 1 + \sqrt{\gamma}$ et $l_2 = \dots = l_k = 1$ (perturbation de rang 1). Alors λ_{\max} satisfait un principe de grandes déviations de vitesse N avec comme bonne fonction de taux I_γ définie sur $((1 + \sqrt{\gamma})^2, +\infty)$ de la façon suivante :*

$$I_\gamma(t) = \frac{t - \lambda_{\text{spk}}}{l_1} - (1 - \gamma) \log \left(\frac{t}{\lambda_{\text{spk}}} \right) - \gamma(F(t) - F(\lambda_{\text{spk}})),$$

avec

$$F(t) = \int \log(t - s) \mu_{\text{mp}}(dy).$$

$\lambda_{\text{spk}} = l_1 + \frac{\gamma l_1}{l_1 - 1}$ est la limite de λ_{\max} se trouvant à droite du support de la loi de Marchenko-Pastur.

Ce théorème permet de calculer les grandes déviations à droite de la limite λ_{spk} et à gauche de λ_{spk} jusqu'au bord $(1 + \sqrt{\gamma})^2$ du support de la loi de Marchenko-Pastur. On remarque que, contrairement au cas non-perturbé, les grandes déviations à gauche et à droite ont la même vitesse N . En regardant ce fait en plus du phénomène de transition de phase des fluctuations, on

a l'impression que lorsque $l_1 > 1 + \sqrt{\gamma}$, λ_{\max} n'interagit plus avec le reste du spectre et devient asymptotiquement indépendante des autres valeurs propres. Néanmoins, cette observation n'est pas appuyée par un calcul.

L'étude dans [BDMN10] permet aux auteurs de construire, via le Théorème 5.4 (Baik-Benarous-Péché), un intervalle de confiance dans lequel on choisit l'hypothèse d'indépendance (absence de signal primaire) et en dehors duquel on réfute cette hypothèse. Lorsque $1 < l_1 \leq 1 + \sqrt{\gamma}$, leurs tests s'avèrent inefficaces à cause du comportement de λ_{\max} similaire à celui du cas non-perturbé. On ne connaît pas à ce jour de moyens efficaces de détecter ces faibles perturbations.

5.4.1 Grandes déviations lorsque k est négligeable devant N .

Dans ce paragraphe, on étudie les formules de grandes déviations de la valeur propre maximale d'une matrice de covariance perturbée lorsque k est négligeable devant N . Ce travail a été motivé par un article de Fey, Van Der Hofstad et Klok [FVDHK08] dans lequel les auteurs démontrent un résultat d'universalité des grandes déviations en supposant les entrées sous-Gaussiennes.

Théorème 5.2 (Fey-Van Der Hofstad-Klok). *Soit W une matrice de covariance empirique dont les entrées sont des variables aléatoires réelles *i.i.d.* tel que $\mathbb{E}C_{11} = 0$ et $\mathbb{E}C_{11}^2 = 1$. On suppose que $\Sigma_k = I_k$ et que pour tout $\lambda \in \mathbb{R}$,*

$$\mathbb{E}e^{\lambda C_{11}} \leq e^{\frac{\lambda^2}{2}}. \quad (5.4)$$

Si $N, k \rightarrow \infty$ avec $k = o\left(\frac{N}{\log \log N}\right)$, alors pour tout $a > 1$,

$$\lim_{k \rightarrow \infty} \frac{1}{N} \log \mathbb{P}[\lambda_{\max} \geq a] = -\frac{1}{2}(a - 1 - \log a).$$

La condition (5.4) veut dire en quelque sorte que les entrées sont des sous-Gaussiennes. La fonction de taux obtenues est la même que celle d'une somme de carrés de Gaussiennes indépendantes. La preuve du Théorème 5.2 est donnée dans [FVDHK08] et sera détaillée dans la suite dans le cadre des matrices perturbées. Dans le même article, les auteurs démontrent d'autres résultats sur la plus grande et la plus petite valeur propre et en fournissent des applications en communication des mobiles.

Dans la même optique de l'article [FVDHK08] et en particulier du théorème précédent, on exploite les matrices de covariance perturbée de la forme

$$W = \frac{1}{N} \Sigma_k^{\frac{1}{2}} C P_N C^* \Sigma_k^{\frac{1}{2}}.$$

C est une matrice de taille $k \times N$ dont les colonnes $(C_j)_{j=1}^N$ sont des vecteurs aléatoires *i.i.d.* de \mathbb{R}^k dont les composantes sont des variables indépendantes, centrées et réduites mais pas forcément de même loi. Rappelons aussi que Σ_k est symétrique définie positive et Σ_k est la matrice de covariance du vecteur aléatoire $\widetilde{C}_1 = \Sigma_k^{\frac{1}{2}} C_1$. Dans le cas Gaussien, on peut supposer sans perte de généralité que Σ_k est diagonale. Ceci est dû au fait que $\Sigma_k^{\frac{1}{2}}$ est symétrique et donc diagonalisable dans une base orthonormée et que $O_k^* C \stackrel{\mathcal{L}}{=} C$ pour toute matrice O_k orthogonale. Dans le cas non-Gaussien, cela n'est plus vrai. Remplacer Σ_k par sa réduite D_k nous contraint à remplacer

C par O^*C où $\Sigma_k = OD_kO^*$. Or les entrées de la matrice O^*C sont réelles décorrélées centrées et réduites mais pas forcément indépendantes. La matrice P est accessoire et son introduction n'est pas indispensable dans la pratique. Lorsque P_N est diagonale, les racines carrées de ces éléments diagonaux peuvent être vues comme une sorte de pondérations ou de poids attribués respectivement aux N réalisations de l'échantillon. Voici les hypothèses sur les matrices qui nous seront utiles dans la suite :

A_0 $N = N_k = N(k)$ dépend uniquement de k et de plus $k = o(N)$.

A_1 Les colonnes $(C_j)_{j=1}^N$ de C sont des vecteurs aléatoires *i.i.d.* de \mathbb{R}^k dont les composantes $(C_{ij})_{i=1}^k$ sont des variables aléatoires indépendantes, centrées et réduites (pas forcément identiquement distribuées).

A_2 $\Sigma_k = ((\sigma_{ij})_{i,j=1}^k)^2$ est symétrique définie positive dont le spectre $l = (l_1^{(k)}, \dots, l_k^{(k)})$ avec $l_1^{(k)} \geq \dots \geq l_k^{(k)} > 0$. On suppose que les $l_i^{(k)}$ sont déterministes et que

$$\frac{1}{k} \sum_{i=1}^k \delta_{l_i^{(k)}} \implies H \quad \text{et} \quad l_1^{(k)} \longrightarrow l^*$$

où H est une mesure de probabilité à support compact dans \mathbb{R}_+ et $l^* \in \mathbb{R}_+$. On note également $x^{(*)}$ l'un des vecteurs propres associés à $l_1^{(k)}$

A_3 $P_N = \text{diag}(p_1^{(k)}, \dots, p_N^{(k)})$ avec $p_1^{(k)} \geq \dots \geq p_N^{(k)} \geq 0$. On suppose que les $p_j^{(k)}$ sont déterministes et que

$$\frac{1}{N} \sum_{j=1}^N \delta_{p_j^{(k)}} \implies \nu \quad \text{et} \quad p_1^{(k)} \longrightarrow p^*$$

où ν est une mesure de probabilité à support compact dans \mathbb{R}_+ et $p^* \in \mathbb{R}_+$. On note

$$p = \int_{\mathbb{R}} s d\nu(s).$$

Remarque 5.3. Ici, on n'impose pas que les k variables constituant chaque réalisation soient de même loi. Ceci est important car dans la pratique, il arrive souvent d'avoir des expériences multivariées liant des paramètres de lois différentes. Cependant, on a besoin dans la suite que les variables satisfassent la condition de Lindeberg nécessaire et suffisante pour avoir un TCL classique.

Afin d'alléger les notations, on remplace les $l_i^{(k)}$ et les $p_j^{(k)}$ respectivement par l_i et p_j pour tout $i = 1, \dots, k$ et $j = 1, \dots, N$. On commence d'abord par considérer le cas Gaussien et ensuite on fera une tentative de généralisation. Le théorème suivant repose essentiellement sur les résultats concernant les grandes déviations des processus Gaussiens stationnaires obtenus par Bercu, Gamboa et Rouault [BGR97].

Théorème 5.4. *On suppose que $A_0 - A_3$ sont vérifiées. On suppose également que $C_{i1} \sim \mathcal{N}(0, 1)$ pour tout $i = 1, \dots, k$. Si $k = o(\frac{N}{\log \log N})$, on a que*

$$\lim_{k \rightarrow \infty} \frac{1}{N} \log \mathbb{P}[\lambda_{\max} \geq a] = -I(a).$$

La bonne fonction de taux $I(a)$ est convexe et vaut

$$I(a) = \sup_{\lambda \in \mathbb{R}} \{\lambda a - \Lambda(\lambda)\} = \begin{cases} L(a) := \sup_{\lambda \in \mathbb{R}} \{\lambda a - \Lambda(\lambda)\} & \text{pour } pl^* < a < a_0 \\ I(a_0) + \frac{1}{2p^*l^*}(a - a_0) & \text{pour } a \geq a_0 \end{cases}$$

où

$$\Lambda(\lambda) = \begin{cases} -\frac{1}{2} \int_{\mathbb{R}} \log(1 - 2\lambda l^* s) d\nu(s) & \text{pour } \lambda \in [0, 1/(2p^*l^*)) \\ \infty & \text{ailleurs} \end{cases}$$

et

$$a_0 = \lim_{\lambda \rightarrow 1/(2p^*l^*)} \Lambda'(\lambda) = \lim_{\lambda \rightarrow 1/(2p^*l^*)^-} l^* \int_{\mathbb{R}} \frac{s}{1 - 2\lambda l^* s} d\nu(s) \geq l^* p.$$

Ce résultat montre une sorte de transition de phase que subit la fonction de taux $I(a)$. Celle-ci semble ne pas être affectée particulièrement par la plus grande pondération p^* pour les valeurs de a proche de la moyenne. A partir d'un certain seuil (dépendant de p^*), la fonction de taux subit une transition de phase et devient linéaire avec une pente qui dépend de p^* , (voir Figure 5.2). La preuve du théorème précédent combine des arguments de [FVDHK08] et les résultats sur les grandes déviations des processus Gaussiens stationnaires [BGR97]. On utilise également des outils de grandes déviations du Chapitre 2 de l'ouvrage [DZ98].

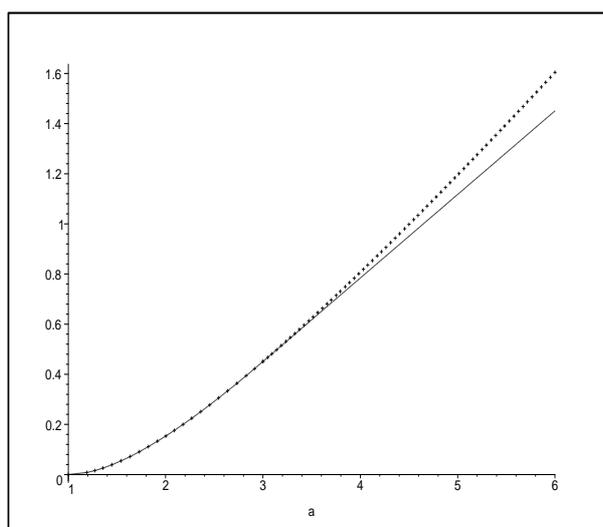


FIG. 5.2 – Comparaison entre $I(a)$ (ligne) et $L(a)$ (pointillée) pour $l^* = 1$ et $p^* = 3/2$ et $\nu = \delta_1$. Cette figure montre la transition de phase que subit $I(a)$ à partir du point $a_0 = 3$.

Exemple 5.5. On considère maintenant le cas particulier où $P_N = \text{diag}(p_1, p_2, \dots, p_r > 1, p_{r+1} = 1, \dots, p_N = 1)$ pour r fixé. Dans ce cas, on a que $\nu = \delta_1$ et par suite, on peut facilement expliciter la fonction de taux et le seuil de transition a_0 . On a que

$$I(a) = \begin{cases} L(a) = \frac{1}{2} \left(\frac{a}{l^*} - 1 - \log \left(\frac{a}{l^*} \right) \right) & \text{pour } l^* < a < a_0 = \frac{l^* p^*}{p^* - 1} \\ \frac{1}{2} \left(\frac{a}{l^* p^*} - \log \left(\frac{p^*}{p^* - 1} \right) \right) & \text{pour } a \geq \frac{l^* p^*}{p^* - 1} \end{cases}$$

Cet exemple montre clairement que pour les a proches de $p l^*$, la fonction de taux est la même que celle d'une somme de carrés de Gaussiennes $\mathcal{N}(0, l^*)$ (notée $L(a)$). Lorsque a dépasse le seuil a_0 , $I(a)$ devient linéaire, (voir Figure 5.2).

En suivant les arguments avancés dans ([FVDHK08], Section 4), on prouve l'universalité des formules de grandes déviations obtenues dans le Théorème 5.4.

Théorème 5.6. *On suppose que $A_0 - A_3$ sont vérifiées. De plus, on suppose que $\Sigma_k = I_k$ et que les entrées sont des sous-Gaussiennes, i.e. pour tout $i = 1, \dots, k$,*

$$\mathbb{E} e^{\lambda C_{i1}} \leq e^{\frac{\lambda^2}{2}}. \quad (5.5)$$

Si $k = o\left(\frac{N}{\log \log N}\right)$, alors on a :

a) Pour tout $p < a < a_0$,

$$\lim_{k \rightarrow \infty} \frac{1}{N} \log \mathbb{P}[\lambda_{\max} \geq a] = -I(a)$$

b) Pour tout $a \geq a_0$,

$$\limsup_{k \rightarrow \infty} \frac{1}{N} \log \mathbb{P}[\lambda_{\max} \geq a] \leq -I(a).$$

$I(a)$ est la même bonne fonction de taux du Théorème 5.4 dans laquelle on choisit $l^* = 1$.

L'hypothèse $\Sigma_k = I_k$ est malheureusement cruciale pour obtenir la même fonction de taux que celle du cas Gaussien. Ceci est expliqué dans la Remarque 5.4.1. Avant de démontrer les deux théorèmes précédents, on donne un tout dernier résultat dans lequel on suppose que $p_1^{(k)}$ converge vers l'infini avec une certaine vitesse inférieure à N .

Théorème 5.7. *On suppose A_0, A_1, A_2 et que $C_{i1} \sim \mathcal{N}(0, 1)$ pour tout $i = 1, \dots, k$. On suppose en plus que $P_N = \text{diag}(p_1^{(k)}, p_2^{(k)} = 1, \dots, p_N^{(k)} = 1)$ et qu'il existe $K > 0$ et $0 < \beta < 1$ telles que $\lim_{k \rightarrow \infty} p_1^{(k)} / N^\beta = 1/K$. Si $k = o\left(\frac{N^{1-\beta}}{\log \log N}\right)$, on a que*

$$\lim_{k \rightarrow \infty} \frac{1}{N^{1-\beta}} \log \mathbb{P}[\lambda_{\max} \geq a] = -K(a - l^*)/2.$$

La preuve de ce théorème est donnée à la fin de ce paragraphe.

Preuve du Théorème 5.4 : Rappelons d'abord que

$$\lambda_{\max} = \sup_{x \in \mathbb{S}_k} \langle x, Wx \rangle \quad \text{où } \mathbb{S}_k = \{x \in \mathbb{R}^k : \|x\|_2^2 = 1\} \quad (5.6)$$

et que le maximum est atteint. D'autre part, pour tout $x \in \mathbb{S}_k$, on a que

$$\langle x, Wx \rangle = \| P_N^{\frac{1}{2}} C^* \Sigma_k^{\frac{1}{2}} x \|_2^2 = \frac{1}{N} \sum_{j=1}^N S_{x,j}^2, \quad (5.7)$$

où

$$S_{x,j} = \sqrt{p_j} \sum_{i=1}^k \left(\sum_{m=1}^k \sigma_{im} x_m \right) C_{ij}.$$

Par conséquent, comme les C_{ij} sont des Gaussiennes indépendantes, on a que

$$S_{x,j} \sim \mathcal{N}(0, p_j \sigma_x^2) \quad \text{où} \quad \sigma_x^2 = \sum_{i=1}^k \left(\sum_{m=1}^k \sigma_{im} x_m \right)^2.$$

Pour simplifier les notations, \mathbb{P}_a désigne la probabilité que λ_{\max} soit supérieure à a . Pour trouver une minoration de \mathbb{P}_a , il suffit de remarquer que pour tout $a > 0$ et tout $x \in \mathbb{S}_k$,

$$\mathbb{P}_a \geq \mathbb{P}[\langle x, Wx \rangle \geq a] = \mathbb{P}\left[\frac{1}{N} \sum_{j=1}^N S_{x,j}^2 \geq a\right] \quad \text{pour tout } x \in \mathbb{S}_k. \quad (5.8)$$

D'un autre côté, d'après (5.6),

$$\mathbb{P}_a = \mathbb{P}[\{\exists x \in \mathbb{S}_k; \langle x, Wx \rangle \geq a\}].$$

Afin de trouver une majoration adéquate de \mathbb{P}_a , on se sert de la technique de discrétisation utilisée dans [FVDHK08]. Cela consiste à recouvrir \mathbb{S}_k par des boules de rayon $1/R$ et dont les centres se situent sur \mathbb{S}_k . On note alors $\mathbb{S}_k(R)$ l'ensemble des centres de ces boules. On note également $M_R := \#\mathbb{S}_k(R)$. De cette manière, on est sûr qu'un point de la sphère unité est toujours à une distance au plus $1/R$ d'au moins un point de $\mathbb{S}_k(R)$. Pour profiter de la proximité des points de $\mathbb{S}_k(R)$, on rappelle la formule élémentaire suivante :

$$|\langle x, Wx \rangle - \langle y, Wy \rangle| \leq 2 \|x - y\|_2 \lambda_{\max} \quad \forall x, y \in \mathbb{S}_k. \quad (5.9)$$

En combinant (5.6), (5.7) et (5.9), on obtient pour tout $b_k > a$,

$$\begin{aligned} \mathbb{P}_a &= \mathbb{P}[a \leq \lambda_{\max} \leq b_k] + \mathbb{P}_{b_k} \\ &\leq \mathbb{P}\left[\{\exists x \in \mathbb{S}_k(R_k^{(1)}); \langle x, Wx \rangle \geq a - \frac{2b_k}{R_k^{(1)}}\}\right] + \mathbb{P}_{b_k} \\ &\leq M_{R_k^{(1)}} \sup_{x \in \mathbb{S}_k} \mathbb{P}\left[\frac{1}{N} \sum_{j=1}^N S_{x,j}^2 \geq a - \frac{2b_k}{R_k^{(1)}}\right] + \mathbb{P}_{b_k}. \end{aligned} \quad (5.10)$$

On applique le même raisonnement pour contrôler \mathbb{P}_{b_k} tout en utilisant une discrétisation différente de la sphère unité. On obtient

$$\mathbb{P}_{b_k} \leq M_{R_k^{(2)}} \sup_{x \in \mathbb{S}_k} \mathbb{P} \left[\frac{1}{N} \sum_{j=1}^N S_{x,j}^2 \geq b_k - \frac{2c_k}{R_k^{(2)}} \right] + \mathbb{P}[Tr(W) \geq c_k]. \quad (5.11)$$

Comme les $S_{x,j} \sim \mathcal{N}(0, p_j \sigma_x^2)$, on peut démontrer facilement que le supremum dans (5.10) et dans (5.11) est atteint lorsque x réalise le maximum de σ_x^2 . Or, on sait que

$$\sup_{x \in \mathbb{S}_k} \sigma_x^2 = \sup_{x \in \mathbb{S}_k} \sum_{i=1}^k \left(\sum_{m=1}^k \sigma_{im} x_m \right)^2 = l_1 \quad \text{et} \quad x^{(*)} = \arg \sup_{x \in \mathbb{S}_k} \sigma_x^2.$$

On a donc que $S_{x,j} \sim \mathcal{N}(0, p_j l_1)$. Avant de choisir les bons b_k , $R_k^{(1)}$, c_k et $R_k^{(2)}$, on rappelle un résultat de Rogers [Rog63] sur la taille de M_R en fonction de R et de k . On a pour tout $R > \sqrt{\frac{k}{k-1}}$, que

$$M_R = 4k\sqrt{k}R^k (\log k + \log \log k + \log R)(1 + O(1/\log k)) := f(R, k)R^k. \quad (5.12)$$

Ce résultat nous dit que $\log M_R = O(k \log R)$. Afin d'obtenir la bonne majoration de \mathbb{P}_a , on choisit

$$b_k = \log k, \quad R_k^{(1)} = \log N, \quad c_k = Kk \quad \text{et} \quad R_k^{(2)} = k.$$

La constante K est positive et est destinée à être aussi grande que l'on veut. Commençons d'abord par majorer la trace de W . On a que

$$Tr(W) = \sum_{j=1}^N p_j \sum_{i=1}^k T_{ij}^2 \quad \text{où} \quad T_{ij} = \sum_{m=1}^k \sigma_{im} C_{mj}.$$

Chaque T_{ij} suit une loi Gaussienne $\mathcal{N}(0, \tau_i^2 \sum_{m=1}^k \sigma_{im}^2)$. Les $(\tau_i^2)_{i=1}^k$ sont les éléments diagonaux de Σ_k . En diagonalisant Σ_k dans une base orthonormée, on obtient immédiatement que $\tau_i^2 \leq l_1$ pour tout $i = 1, \dots, k$. Par suite, si on considère la suite $(Z_{ij})_{i,j=1}^\infty$ de variables aléatoires indépendantes de loi $\mathcal{N}(0, 1)$, on a que

$$\begin{aligned} \mathbb{P}[Tr(W) \geq Kk] &\leq \mathbb{P} \left[\frac{l_1 p_1}{Nk} \sum_{i=j}^N \sum_{i=1}^k Z_{ij}^2 \geq K \right] \\ &\leq \exp \left(-Nk \sup_{\lambda \in \mathbb{R}} \left(\lambda \frac{K}{l_1 p_1} + \frac{1}{2} \log(1 - 2\lambda) \right) \right) \\ &\leq \exp \left(-Nk \left(\frac{K}{l_1 p_1} - 1 - \log \frac{K}{l_1 p_1} \right) \right) \end{aligned} \quad (5.13)$$

Lorsque $K \gg l^* p^*$, le côté droit de (5.13) décroît exponentiellement vers 0. En insérant (5.12) et (5.13) dans (5.11), on obtient à partir d'un certain k assez grand, que

$$\mathbb{P}_{b_k} \leq f(R_k^{(2)}, k) \exp(k \log k - N \log k(1/(l_1 p_1) - o(1))) + \exp(-Nk(1 - o(1))) \quad (5.14)$$

Revenons maintenant à \mathbb{P}_a . On a que $R_k^{(1)} = \log N$ et par suite, d'après (5.12), $\log M_{R_k^{(1)}}/N \rightarrow 0$ quand $k \rightarrow \infty$ puisque $k = o(N/\log \log N)$. En insérant (5.14) dans (5.10), on obtient pour $\varepsilon > 0$ assez petit, que

$$\limsup_{k \rightarrow \infty} \frac{1}{N} \log \mathbb{P}_a \leq \limsup_{k \rightarrow \infty} \frac{1}{N} \log \mathbb{P} \left[\frac{1}{N} \sum_{j=1}^N S_{x^{(*)},j}^2 \geq a - \varepsilon \right]. \quad (5.15)$$

D'autre part, d'après (5.8), on a que

$$\liminf_{k \rightarrow \infty} \frac{1}{N} \log \mathbb{P}_a \geq \liminf_{k \rightarrow \infty} \frac{1}{N} \log \mathbb{P} \left[\frac{1}{N} \sum_{j=1}^N S_{x^{(*)},j}^2 \geq a \right]. \quad (5.16)$$

D'après (5.15) et (5.16), les grandes déviations de λ_{\max} se comportent comme celles d'une somme partielle des carrés de variables Gaussiennes indépendantes. Le lemme suivant est un cas particulier du résultat de Bercu, Gamboa et Rouault [BGR97] sur les grandes déviations des processus Gaussiens stationnaires.

Lemme 5.8 (Bercu-Gamboa-Rouault [BGR97]). *Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes de loi $\mathcal{N}(0, l^*)$ et soit $p_1^{(n)}, \dots, p_n^{(n)}$ des réelles positives dépendant de n . On suppose que*

$$\frac{1}{n} \sum_{j=1}^n \delta_{p_j^{(n)}} \implies \nu \quad \text{et} \quad \max_j p_j^{(n)} \longrightarrow p^* \in \mathbb{R}_+$$

où ν est une mesure de probabilité à support compact dans \mathbb{R}_+ . Alors, pour tout $a > l^* p := l^* \int s d\nu(s)$, on a

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} \left[\frac{1}{n} \sum_{j=1}^n p_j^{(n)} X_j^2 \geq a \right] = -I(a).$$

La fonction de taux $I(a)$ la même que celle du Théorème 5.4.

On construit dans la suite une preuve simple du lemme précédent. Le lecteur pourra également consulter la démonstration complète donnée dans [BGR97] dans le cadre des processus Gaussiens stationnaires. Une version fonctionnelle du lemme précédent et du Théorème 3 de [BGR97] est donnée dans [GRZ99].

Preuve du Lemme 5.8 : La preuve du Lemme 5.8 repose en partie sur le théorème de Gärtner-Ellis. Celui-ci est une généralisation du théorème de Cramér, (voir le Chapitre 2 de [DZ98]).

Théorème 5.9 (Gärtner-Ellis). *Soit $(Z_n)_{n \geq 1}$ une suite de variables aléatoires réelles. On note μ_n la loi de Z_n et $\Lambda_n(\lambda)$ le logarithme de sa transformée de Laplace. On suppose que pour tout $\lambda \in \mathbb{R}$, $\frac{1}{n} \Lambda_n(\lambda)$ converge vers une fonction $\Lambda(\lambda) \in \overline{\mathbb{R}}$ et on note $\mathcal{D}_\Lambda := \{\lambda \in \mathbb{R}; \Lambda(\lambda) < \infty\}$. Alors, on a les deux inégalités asymptotiques suivantes :*

a) Pour tout fermé $F \subset \mathbb{R}$,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(F) \leq - \inf_{a \in F} \sup_{\lambda \in \mathcal{D}_\Lambda} (\lambda a - \Lambda(\lambda)).$$

b) Pour tout ouvert $O \subset \mathbb{R}$,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(O) \geq - \inf_{a \in O \cap \mathcal{G}} \sup_{\lambda \in \mathcal{D}_\Lambda} (\lambda a - \Lambda(\lambda)),$$

où $\mathcal{G} := \{a \in \mathbb{R}; \exists \lambda \in \mathcal{D}_\Lambda \text{ tel que } a = \Lambda'(\lambda)\}$.

Une version multidimensionnelle de ce théorème est donnée dans la Section 2.3 de [DZ98]. Le lecteur pourra y en trouver une preuve complète. Ici, $Z_n = \frac{1}{n} \sum_{j=1}^n p_j^{(n)} X_j^2$. Par suite,

$$\Lambda_n(\lambda) = -\frac{1}{2n} \sum_{j=1}^n \log(1 - 2p_j l^* \lambda/n)$$

On voit facilement que $\mathcal{D}_\Lambda = (-\infty, \frac{1}{2p^* l^*})$ et que pour tout $\lambda \in \mathcal{D}_\Lambda$,

$$\frac{1}{n} \Lambda_n(n\lambda) \xrightarrow{k \rightarrow \infty} \frac{1}{2} \int_{\mathbb{R}} \log(1 - 2sl^* \lambda) d\nu(s) := \Lambda(\lambda).$$

La majoration de la formule de grandes déviations dans le Lemme 5.8 est donc une conséquence immédiate du théorème de Gärtner-Ellis. En effet, on a pour tout $a > pl^*$,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}[Z_n \geq a] \leq -I(a).$$

La fonction $\Lambda(\lambda)$ est convexe et est dérivable pour tout $\lambda \in \mathcal{D}_\Lambda^0$. Par suite, $\Lambda'(\lambda)$ est croissante et admet pour maximum $a_0 := \lim_{\lambda \rightarrow \frac{1}{2p^*}} \Lambda'(\lambda)$. Cela veut dire que l'ensemble $\mathcal{G} = (-\infty, a_0)$. Par conséquent, le b) du théorème de Gärtner-Ellis montre que pour $pl^* < a < a_0$,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}[Z_n \geq a] \geq -I(a). \quad (5.17)$$

Notons qu'en effectuant une dérivation sous le signe intégrale, on peut démontrer que

$$a_0 = \lim_{\lambda \rightarrow 1/(2p^* l^*)^-} l^* \int_{\mathbb{R}} \frac{s}{1 - 2\lambda l^* s} d\nu(s) \geq l^* p.$$

Lorsque $a \geq a_0$, le théorème de Gärtner-Ellis ne permet plus d'avoir une borne inférieure des formules de grandes déviations. Les calculs sont donc effectués à la main. La fonction $\lambda(a - \Lambda(\lambda))$ est concave et atteint son maximum lorsque $a = \Lambda'(\lambda)$. Or, quand $a \geq a_0$, $(a - \Lambda'(\lambda))$ est strictement positive pour tout $\lambda \in [0, \frac{1}{2p^* l^*})$. Par conséquent, le maximum se trouve sur le bord et donc,

$$\sup_{\lambda \in [0, \frac{1}{2p^* l^*})} (\lambda a - \Lambda(\lambda)) = \frac{a}{2p^* l^*} - \Lambda(1/(2p^* l^*)).$$

Vérifions alors que cette même fonction est une borne inférieure des grandes déviations. Pour cela, on remarque que pour tout $\delta > 0$,

$$\mathbb{P}[Z_n \geq a] \geq \mathbb{P}[Z_n - p_1^{(n)} X_1^2/n \geq a_0 - \delta] \cdot \mathbb{P}[p_1^{(n)} X_1^2/n \geq a - a_0 + \delta].$$

On peut voir sans difficultés que $(Z_n - p_1^{(n)} X_1^2/n)$ satisfait (5.17). De plus, X_1 est de loi Gaussienne. Par conséquent, pour $a \geq a_0$,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}[Z_n \geq a] \geq -\frac{a_0 - \delta}{2p^* l^*} + \Lambda(1/(2p^* l^*)) - \frac{a - a_0 + \delta}{2p^* l^*}. \quad (5.18)$$

Le lemme 5.8 est donc établi. □

Revenons à la preuve du Théorème 5.4. $I(a)$ est convexe donc continue. Par conséquent, les inégalités (5.15) et (5.16) et le Lemme 5.8 permettent de conclure. \square

Venons maintenant à la preuve du Théorème 5.6. Celle-ci est inspirée essentiellement des arguments donnés dans [FVDHK08] pour démontrer le Théorème 5.2. Elle repose sur le fait que lorsque k est très grand, le x qui réalise le supremum de la probabilité dans (5.10) est le vecteur isotrope $\frac{1}{\sqrt{k}}(1, \dots, 1)$.

Preuve du Théorème 5.6 : On note $\tilde{x} = \frac{1}{\sqrt{k}}(1, \dots, 1)$ le vecteur isotrope de \mathbb{S}_k et on désigne par Z une variable de loi $\mathcal{N}(0, 1)$ indépendante des C_{ij} . Afin de pouvoir utiliser les mêmes arguments que la preuve précédente, on démontre, pour tout $i = 1$ allant de 1 à k et tout $\lambda \in \mathbb{R}$, que

$$\mathbb{E}e^{\lambda C_{i1}^2} \leq \frac{1}{\sqrt{1 - 2\lambda}}. \quad (5.19)$$

Pour cela, il suffit de remarquer que

$$\mathbb{E}e^{\lambda C_{i1}^2} = \mathbb{E}\mathbb{E}_Z e^{\sqrt{2\lambda}ZC_{i1}} = \mathbb{E}_Z \mathbb{E}e^{\sqrt{2\lambda}ZC_{i1}}.$$

Par conséquent, (5.19) découle immédiatement de l'hypothèse (5.5). On reprend maintenant les mêmes étapes de la preuve précédente avec les mêmes discrétisations de la sphère unité mais en supposant des entrées sous-Gaussiennes et $\Sigma_k = I_k$. L'inégalité (5.19) nous dit que (5.14) tient toujours. Par conséquent, on a les deux inégalités suivantes : pour tout $a > p$ et pour tout $\varepsilon > 0$ très petit,

$$\limsup_{k \rightarrow \infty} \frac{1}{N} \log \mathbb{P}_a \leq \limsup_{k \rightarrow \infty} \sup_{x \in \mathbb{S}_k} \frac{1}{N} \log \mathbb{P} \left[\frac{1}{N} \sum_{j=1}^N S_{x,j}^2 \geq a - \varepsilon \right], \quad (5.20)$$

et pour tout $x \in \mathbb{S}_k$,

$$\liminf_{k \rightarrow \infty} \frac{1}{N} \log \mathbb{P}_a \geq \liminf_{k \rightarrow \infty} \frac{1}{N} \log \mathbb{P} \left[\frac{1}{N} \sum_{j=1}^N S_{x,j}^2 \geq a \right]. \quad (5.21)$$

La clé de la preuve est de démontrer que la transformée de Laplace de $S_{x,j}^2$ est inférieure à celle d'une Gaussienne et que la transformée de Laplace de $S_{x,j}^2$ converge vers celle d'une Gaussienne. On suit alors les arguments de [FVDHK08] et on affirme que pour tout $j = 1, \dots, N$, pour tout $x \in \mathbb{S}_k$ et tout $0 \leq \lambda < \frac{1}{2p_j}$,

$$\begin{aligned} \mathbb{E}e^{\lambda S_{x,j}^2} &= \mathbb{E}\mathbb{E}_Z e^{\sqrt{2\lambda}ZS_{x,j}} = \mathbb{E}_Z \mathbb{E}e^{\sqrt{2\lambda}ZS_{x,j}} = \mathbb{E}_Z \prod_{i=1}^k \mathbb{E}e^{\sqrt{2\lambda p_j} x_i Z C_{ij}} \stackrel{(5.5)}{\leq} \mathbb{E}_Z \prod_{i=1}^k e^{\lambda p_j x_i^2 Z^2} \\ &\leq \mathbb{E}_Z e^{\lambda p_j Z^2} \\ &= \frac{1}{\sqrt{1 - 2\lambda p_j}}. \end{aligned} \quad (5.22)$$

D'autre part, les C_{ij} sont des sous-Gaussiennes, elles satisfont donc la condition de Lindeberg du TCL pour des variables indépendantes, centrées, réduites mais pas forcément de même loi. Par suite, $S_{\bar{x},j}$ converge vers une $\mathcal{N}(0, 1)$ pour tout $j = 1, \dots, N$. L'inégalité exponentielle de Bernstein [dlP99] garantit la convergence des moments de $S_{\bar{x},j}$ y compris la convergence de la transformée de Laplace pour tout $\lambda < \frac{1}{2p_j}$. On a donc que

$$\lim_{k \rightarrow \infty} \mathbb{E} e^{\lambda S_{\bar{x},j}^2} = \frac{1}{\sqrt{1 - 2\lambda p_j}}. \quad (5.23)$$

En examinant (5.20) et (5.21) d'un côté et (5.22) et (5.23) de l'autre, on constate que les grandes déviations de λ_{\max} sont contrôlées par celles de la somme partielle des $S_{\bar{x},j}^2$. En combinant (5.20) et (5.22), on a pour tout $a > p$ et $\varepsilon > 0$, que

$$\limsup_{k \rightarrow \infty} \frac{1}{N} \log \mathbb{P}_a \leq - \liminf_{k \rightarrow \infty} \sup_{\lambda \in \mathbb{R}} \left(\lambda(a - \varepsilon) - \frac{1}{2N} \sum_{j=1}^N \log(1 - 2p_j \lambda) \right). \quad (5.24)$$

Par convexité de la transformée de Laplace, on intervertit la limite et le supremum dans l'inégalité précédente. Or, par simple passage à la limite, on a pour tout $\lambda \in (-\infty, \frac{1}{2p^*})$ et k suffisamment grand, que

$$\frac{1}{2N} \sum_{j=1}^N \log(1 - 2p_j \lambda) \longrightarrow \frac{1}{2} \int_{\mathbb{R}} \log(1 - 2s\lambda) d\nu(s). \quad (5.25)$$

Par suite, (5.24) et (5.25) montre que pour tout $a > p$ et tout $\varepsilon > 0$ suffisamment petit,

$$\limsup_{k \rightarrow \infty} \frac{1}{N} \log \mathbb{P}_a \leq -I(a - \varepsilon). \quad (5.26)$$

On établit maintenant une borne inférieure des grandes déviations. Avec les mêmes notations du théorème de Gärtner-Ellis, on choisit $Z_N = \frac{1}{N} \sum_{j=1}^N S_{\bar{x},j}^2$. Par suite,

$$\Lambda_N(\lambda) = -\frac{1}{2} \sum_{j=1}^N \mathbb{E} e^{\lambda S_{\bar{x},j}^2 / N}.$$

La convergence dans (5.23) est uniforme en λ sur tout compact de $(-\infty, \frac{1}{2p_j})$. En faisant le simple changement de variables $\lambda' = p_j \lambda / p^*$, on démontre que cette convergence est également uniforme en j pour tout $\lambda \in (-\infty, \frac{1}{2p^*})$ puisque $p_1 \rightarrow p^*$. Par conséquent, pour tout $\lambda \in (-\infty, \frac{1}{2p^*})$, on a que

$$\left| \frac{1}{2N} \sum_{j=1}^N \mathbb{E} e^{\lambda S_{\bar{x},j}^2} - \frac{1}{2N} \sum_{j=1}^N \log(1 - 2p_j \lambda) \right| \longrightarrow 0. \quad (5.27)$$

(5.25) et (5.27) montrent alors que pour tout $\lambda \in (-\infty, \frac{1}{2p^*})$,

$$\frac{1}{N} \Lambda_N(N\lambda) \xrightarrow{k \rightarrow \infty} \frac{1}{2} \int_{\mathbb{R}} \log(1 - 2s\lambda) d\nu(s) = \Lambda(\lambda). \quad (5.28)$$

D'après le théorème de Gärtner-Ellis, pour tout $a < a_0 := \lim_{\lambda \rightarrow \frac{1}{2p^*}} \Lambda'(\lambda)$,

$$\liminf_{k \rightarrow \infty} \frac{1}{N} \log \mathbb{P}_a \geq -I(a).$$

La continuité de $I(a)$ permet de conclure. □

Remarque 5.10. Les Théorèmes 5.7 et 5.8 permettent d'obtenir la convergence en probabilité de λ_{\max} vers une limite finie. Par exemple, sous les hypothèses de Théorèmes 5.7, on a systématiquement que

$$\lambda_{\max} \xrightarrow{P} pl^*.$$

Pour cela, il suffit de remarquer que pour tout $a < pl^*$,

$$\mathbb{P}[\lambda_{\max} \leq a] \leq \mathbb{P}[\langle x^{(*)}, Wx^{(*)} \rangle \leq a].$$

Remarque 5.11. Dans la preuve précédente, le vecteur isotrope \tilde{x} joue un rôle crucial pour obtenir l'universalité. En effet, \tilde{x} réalise asymptotiquement le supremum de la probabilité dans (5.20) et permet ensuite d'avoir un TCL. Lorsque $\Sigma_k \neq I_k$, c'est plutôt $x^{(*)}$ qui réalise le supremum mais ne permet pas d'avoir le bon TCL. Donc il est difficile de s'attendre à une fonction de taux universelle dans le cas d'une matrice de covariance non-blanche. Cela peut se comprendre par le fait que dans le cas non-blanc λ_{\max} converge vers pl^* et par suite toutes les variables de C ne participent pas à pied d'égalité à la construction de λ_{\max} .

Avant d'aborder les principes d'invariance, on clôt cette partie par la preuve du Théorème 5.7.

Preuve du Théorème 5.7 : Pour démontrer le Théorème 5.7, on reprend les mêmes arguments de la démonstration du Théorème 5.4 et on obtient pour tout $a > l^*$ et tout $\varepsilon > 0$ suffisamment petit, que

$$\limsup_{k \rightarrow \infty} \frac{1}{N^{1-\beta}} \log \mathbb{P}_a \leq \limsup_{k \rightarrow \infty} \frac{1}{N^{1-\beta}} \log \mathbb{P} \left[\frac{1}{N} \sum_{j=1}^N S_{x^{(*)},j}^2 \geq a - \varepsilon \right]. \quad (5.29)$$

et

$$\liminf_{k \rightarrow \infty} \frac{1}{N^{1-\beta}} \log \mathbb{P}_a \geq \liminf_{k \rightarrow \infty} \frac{1}{N^{1-\beta}} \log \mathbb{P} \left[\frac{1}{N} \sum_{j=1}^N S_{x^{(*)},j}^2 \geq a \right]. \quad (5.30)$$

Dans la suite, on note

$$Z_N = \frac{1}{N} \sum_{j=1}^N S_{x^{(*)},j}^2,$$

et on rappelle que $S_{x^{(*)},j}^2 \sim \mathcal{N}(0, p_j l_1)$ pour tout $j = 1, \dots, N$. Tout d'abord, on cherche une minoration de (5.30). Pour cela, il suffit de remarquer

$$\mathbb{P}[Z_N \geq a] \geq \mathbb{P}[Z_N - S_{x^{(*)},1}^2/N \geq l^*] \cdot \mathbb{P}[S_{x^{(*)},1}^2/N \geq a - l^*].$$

Par conséquent, on a que

$$\liminf_{k \rightarrow \infty} \frac{1}{N^{1-\beta}} \log \mathbb{P}[Z_N \geq a] \geq \frac{K}{2} \left(\frac{a}{l^*} - 1 \right). \quad (5.31)$$

Pour obtenir une borne supérieure de (5.29), on utilise les techniques classiques de majoration des grandes déviations. L'inégalité de Markov montre que

$$\limsup_{k \rightarrow \infty} \frac{1}{N^{1-\beta}} \log \mathbb{P}[Z_N \geq a] \leq - \liminf_{k \rightarrow \infty} \sup_{\lambda \in \mathbb{R}} \left(\frac{\lambda N^\beta}{l_1} a - \Lambda_N(\lambda) \right), \quad (5.32)$$

où

$$\Lambda_N(\lambda) := -\frac{N-1}{2N^{1-\beta}} \log(1-2\lambda l_1) - \frac{1}{2N^{1-\beta}} \log(1-2\lambda p_1 l_1).$$

On calcule le supremum en dérivant la fonction concave $\left(\frac{\lambda N^\beta}{l_1} a - \Lambda_N(\lambda)\right)$ et on trouve que, lorsque N et k sont suffisamment grand, le maximiseur

$$\lambda_N = \frac{1}{4p_1 a} \left(a + p_1 a - p_1 l_1 - \sqrt{((p_1 - 1)a - p_1 l_1)^2 + (4p_1^2 l_1 a - 4p_1 l_1 a)/N} \right). \quad (5.33)$$

Lorsque N est suffisamment grand, a est supérieur à $l_1 p_1 / (p_1 - 1)$. Par suite, en simplifiant (5.33), on obtient que

$$\lambda_N \stackrel{N}{\sim} \frac{1}{2p_1} (1 - O(1/N)).$$

Pour terminer, on remplace λ par λ_N dans (5.32) et on obtient le résultat. \square

5.5 Théorèmes d'invariance et matrices de covariance empirique

En examinant le Chapitre 3, on peut s'interroger sur l'efficacité des méthodes de comparaison et de plongement appliquées à des matrices de covariance aléatoires. A l'image du modèle de percolation dirigée dans le plan, une matrice de covariance admet une structure rectangulaire. En choisissant $k = o(N)$, on peut se poser la question si des résultats obtenus dans le cas d'une matrice de Wishart ne peuvent être étendus à des matrices de covariances empirique quelconques et ceci en utilisant les mêmes méthodes d'approximations que celles du Chapitre 3. Dans ce paragraphe, on expose la différence structurelle entre la plus grande valeur propre d'une matrice de covariance et le temps de dernier passage ainsi que les difficultés que cette différence entraîne dans l'application de l'approximation KMT.

En utilisant les mêmes notations des paragraphes précédents, on considère une matrice de covariance empirique réelle W avec $\Sigma_k = I_k$ et $P_N = I_N$. On suppose que les colonnes de C sont *i.i.d.* et que les C_{i1} sont des variables aléatoires réelles indépendantes de moyenne nulle et de variance 1. On rappelle que

$$\lambda_{\max} = \sup_{x \in \mathbb{S}_k} \frac{1}{N} \sum_{j=1}^N \left(\sum_{i=1}^k x_i C_{ij} \right)^2.$$

En faisant l'analogie avec les résultats du Chapitre 4, on conjecture que, lorsque $k = o(N)$ et lorsque les C_{i1} sont *i.i.d.* admettant des moments exponentiels,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathbb{P}[\lambda_{\max} - 1 \geq 2\sqrt{Nk}(1 + \varepsilon)] = J_{\text{GUE}}, \quad (5.34)$$

pour tout $\varepsilon > 0$. Pour essayer de comprendre les raisons pour lesquelles les techniques de comparaison ne fonctionnent pas ici, on considère une matrice de Wishart réelle W' avec des entrées $G_{ij} \sim \mathcal{N}(0, 1)$ et on note λ'_{\max} sa valeur propre maximale. On pense que λ'_{\max} satisfait (5.34) puisque J_{LUE} converge vers J_{GUE} lorsque $k = o(N)$ (ce n'est pas une preuve). Étendre ce résultat à λ_{\max} en utilisant les techniques de comparaison du Chapitre 3 revient à contrôler la quantité

$$|\lambda_{\max} - \lambda'_{\max}|.$$

Or, on a que

$$|\lambda_{\max} - \lambda'_{\max}| \leq \frac{1}{N} \sum_{j=1}^N \sup_{x \in \mathbb{S}_k} \left| \left(\sum_{i=1}^k x_i C_{ij} \right)^2 - \left(\sum_{i=1}^k x_i G_{ij} \right)^2 \right|. \quad (5.35)$$

Peut-on reconstruire les C_{ij} de manière à contrôler suffisamment (5.35) pour établir (5.34)? Contrairement à la structure du temps de dernier passage, l'expression de la plus grande valeur propre de W n'est pas adaptée ni au plongement de Skorokhod ni à l'approximation KMT. Cette dernière ne concerne que les sommes partielles de variables *i.i.d.* et est difficilement généralisable pour couvrir d'autres situations. En effet, on est devant deux obstacles : la comparaison des carrés des sommes partielles d'une part et la maximisation sur \mathbb{S}_k de l'autre. Par conséquent, on est incapable pour le moment de démontrer (5.34).

Dans le cadre des fluctuations, une solution possible à ces deux problèmes serait d'utiliser les approximations faibles. Bien qu'elles fournissent des bornes supérieures modestes (insuffisantes pour les grandes déviations), ces méthodes peuvent s'adapter facilement à différents types de problèmes. Le principe d'invariance faible suivant a été établi par Chatterjee dans [Cha05] et s'adapte bien notre situation.

Théorème 5.1 (Chatterjee [Cha05]). *Soit $X = (X_1, \dots, X_n)$ et $Y = (Y_1, \dots, Y_n)$ deux vecteurs aléatoires formés de v.a.r indépendantes. Pour $i = 1, \dots, n$, on suppose que $\mathbb{E}X_i = \mathbb{E}Y_i$, $\mathbb{E}X_i^2 = \mathbb{E}Y_i^2$ et $\gamma = \max \{\mathbb{E}X_i^3, \mathbb{E}Y_i^3; i = 1, \dots, n\} < \infty$. Soit \mathcal{F} une famille finie de fonctions de $\mathbb{R}^n \rightarrow \mathbb{R}$ trois fois différentiables, $U = \max_{f \in \mathcal{F}} f(X)$ et $V = \max_{f \in \mathcal{F}} f(Y)$. Alors, pour toute fonction réelle g trois fois différentiable, on a que*

$$|\mathbb{E}g(U) - \mathbb{E}g(V)| \leq K(g) [(\gamma n \lambda_3(\mathcal{F}))^{1/3} (\log |\mathcal{F}|)^{2/3} + \gamma n \lambda_3(\mathcal{F})],$$

où

$$K(g) = \frac{19}{3} \|g'\|_{\infty} + 13 \|g''\|_{\infty} + \frac{13}{3} \|g'''\|_{\infty}$$

et

$$\lambda_3(\mathcal{F}) = \sup_{f \in \mathcal{F}} \left\{ \sup \left\{ \left| \frac{\partial^p}{\partial x_i} f(x) \right|^{\frac{3}{p}} : 1 \leq i \leq n, 1 \leq p \leq 3 \right\} \right\}.$$

Dans le même article, Chatterjee propose des applications de ce principe dans différents domaines des mathématiques y compris les matrices aléatoires. Ce même principe a été également utilisé par Suidan pour démontrer l'universalité des fluctuations du temps de dernier passage, [Sui06].

Afin d'exploiter le Théorème 5.1 dans notre contexte, on discrétise \mathbb{S}_k et on l'assimile à l'ensemble des points $\mathbb{S}_k(R_k)$, (voir le paragraphe précédent). En tenant compte de cette discrétisation, on choisit

$$\mathcal{F} = \{f_x : \mathbb{R}^{Nk} \rightarrow \mathbb{R}; x \in \mathbb{S}_k(R_k)\} \quad \text{où} \quad f_x(W) = \frac{\sum_{j=1}^N \left(\sum_{i=1}^k x_i C_{ij} \right)^2 - (\sqrt{N} + \sqrt{k})^2}{N^{\frac{1}{2}} k^{-\frac{1}{6}}}.$$

Notons que $\#\mathcal{F}$ est donné dans (5.12). On suppose maintenant que les C_{ij} sont indépendantes de moyenne 0, de variance 1 et qu'elles admettent des moments finis d'ordre $p > 6$. On suppose également que tous les moments jusqu'à l'ordre p de toutes les variables sont uniformément bornés (indépendamment de k et de N). En choisissant $k = N^\alpha$ et $R_k = N$, on a pour toute fonction g trois fois différentiable, que

$$|\mathbb{E}g(\max_x f_x(W)) - \mathbb{E}g(\max_x f_x(W'))| \leq c N^{\frac{1}{6} - \frac{7\alpha}{6} - \frac{\alpha}{p} - \frac{1}{p}} \quad (5.36)$$

où c est une constante positive. D'autre part, on sait d'après (5.9) qu'il existe $x \in \mathbb{S}_k(R_k)$ tel que

$$|\lambda'_{\max} N - f_x(W')| / N^{\frac{1}{2} - \frac{\alpha}{6}} \leq 2\lambda'_{\max} / N^{\frac{1}{2} - \frac{\alpha}{6}}. \quad (5.37)$$

En combinant (5.36) et (5.37), on démontre que

$$\lim_{N, k \rightarrow \infty} \mathbb{P} \left[\frac{N\lambda_{\max} - (\sqrt{N} + \sqrt{k})^2}{N^{\frac{1}{2}} k^{-\frac{1}{6}}} \leq s \right] = F_2^{\text{TW}},$$

et ceci lorsque

$$\alpha < \frac{6p}{6 + 7p} \left(\frac{1}{6} - \frac{1}{p} \right).$$

Remarque 5.2. Cette méthode n'est pas très performante pour démontrer l'universalité des fluctuations. L'erreur due à la comparaison entre λ_{\max} et λ'_{\max} est si grande qu'on est contraint de choisir une matrice rectangulaire C très fine. Dans le meilleur des cas *i.e.* lorsque les entrées admettent des moments finis de tout ordre, α est strictement inférieure à $1/7$ ce qui n'est pas très satisfaisant. De nouvelles méthodes de comparaison aussi élégantes que techniques ont été développées récemment et ont permis de démontrer, en rajoutant des conditions de moments sur les entrées, l'universalité des fluctuations pour N et k du même ordre. Ces méthodes n'ont pas fait l'objet d'étude durant cette thèse et ne seront pas développées ici. On cite notamment les travaux de Soshnikov [Sos99, Sos02] et ceux de Tao et Vu [TV10].

Remarque 5.3. Bien qu'elle soit inefficace pour les fluctuations, la méthode de Chatterjee s'avère être très rentable pour démontrer l'invariance de la limite de la mesure spectrale μ_k de $\frac{1}{N}W$. La preuve repose essentiellement sur une version modifiée du Théorème 5.1. Dans ce cas, l'ensemble \mathcal{F} est réduit à une seule fonction $f(x) = (1/k)\text{Tr}(\frac{1}{N}W(x) - zI)^{-1}$ de $\mathbb{R}^{Nk} \rightarrow \mathbb{R}$ qui n'est autre que la transformée de Stieltjes de μ_k . Les dérivées successives de f convergent assez vite vers 0 en k pour pouvoir choisir N et k du même ordre. On a ainsi la convergence de la transformée de Stieltjes de μ_k vers celle de la loi du demi-cercle. La preuve détaillée dans le cas d'une matrice de Wigner est dans ([Cha05], Sec 2).

Bibliographie

- [AD99] David Aldous and Persi Diaconis. Longest increasing subsequences : from patience sorting to the Baik-Deift-Johansson theorem. *Bull. Amer. Math. Soc. (N.S.)*, 36(4) :413–432, 1999.
- [AGZ10] G.W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*. Cambridge University Press, 2010.
- [And03] T. W. Anderson. *An introduction to multivariate statistical analysis*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, third edition, 2003.
- [AY79] Jacques Azéma and Marc Yor. Une solution simple au problème de Skorokhod. In *Séminaire de Probabilités, XIII (Univ. Strasbourg, Strasbourg, 1977/78)*, volume 721 of *Lecture Notes in Math.*, pages 90–115. Springer, Berlin, 1979.
- [BADG01] Gérard Ben Arous, Amir Dembo, and Alice Guionnet. Aging of spherical spin glasses. *Probab. Theory Related Fields*, 120(1) :1–67, 2001.
- [BAG97] Gerard Ben Arous and Alice Guionnet. Large deviations for Wigner’s law and Voiculescu’s non-commutative entropy. *Probab. Theory Related Fields*, 108(4) :517–542, 1997.
- [BAG08] Gérard Ben Arous and Alice Guionnet. The spectrum of heavy tailed random matrices. *Comm. Math. Phys.*, 278(3) :715–751, 2008.
- [Bai99] Z. D. Bai. Methodologies in spectral analysis of large-dimensional random matrices, a review. *Statist. Sinica*, 9(3) :611–677, 1999. With comments by G. J. Rodgers and Jack W. Silverstein ; and a rejoinder by the author.
- [Bai05] Jinho Baik. Limiting distribution of last passage percolation models. In *XIVth International Congress on Mathematical Physics*, pages 339–346. World Sci. Publ., Hackensack, NJ, 2005.
- [Bar01] Yu. Baryshnikov. GUEs and queues. *Probab. Theory Related Fields*, 119(2) :256–274, 2001.
- [BBAP05] Jinho Baik, Gérard Ben Arous, and Sandrine Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.*, 33(5) :1643–1697, 2005.
- [BDJ99] Jinho Baik, Percy Deift, and Kurt Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.*, 12(4) :1119–1178, 1999.
- [BDM⁺01] Jinho Baik, Percy Deift, Ken T.-R. McLaughlin, Peter Miller, and Xin Zhou. Optimal tail estimates for directed last passage site percolation with geometric random variables. *Adv. Theor. Math. Phys.*, 5(6) :1207–1250, 2001.

- [BDMN10] P. Bianchi, M. Debbah, M. Maida, and J. Najim. Performance of Statistical Tests for Single Source Detection using Random Matrix Theory. 2010.
- [BGL00] Bernard Bercu, Fabrice Gamboa, and Marc Lavielle. Sharp large deviations for Gaussian quadratic forms with applications. *ESAIM Probab. Statist.*, 4 :1–24 (electronic), 2000.
- [BGR97] Bernard Bercu, Fabrice Gamboa, and Alain Rouault. Large deviations for quadratic forms of stationary Gaussian processes. *Stochastic Process. Appl.*, 71(1) :75–90, 1997.
- [BKS03] Itai Benjamini, Gil Kalai, and Oded Schramm. First passage percolation has sub-linear distance variance. *Ann. Probab.*, 31(4) :1970–1978, 2003.
- [BM05] Thierry Bodineau and James Martin. A universality property for last-passage percolation paths close to the axis. *Electron. Comm. Probab.*, 10 :105–112 (electronic), 2005.
- [BNMD09] Pascal Bianchi, Jamal Najim, Mylène Maida, and Mérouane Debbah. Performance analysis of some eigen-based hypothesis tests for collaborative sensing. 2009.
- [Bre92] Leo Breiman. *Probability*, volume 7 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992. Corrected reprint of the 1968 original.
- [BS98] Z. D. Bai and Jack W. Silverstein. No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.*, 26(1) :316–345, 1998.
- [BS05] Jinho Baik and Toufic M. Suidan. A GUE central limit theorem and universality of directed first and last passage site percolation. *Int. Math. Res. Not.*, (6) :325–337, 2005.
- [BS06] Jinho Baik and Jack W. Silverstein. Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.*, 97(6) :1382–1408, 2006.
- [BS10] Zhidong Bai and Jack W. Silverstein. *Spectral analysis of large dimensional random matrices*. Springer Series in Statistics. Springer, New York, second edition, 2010.
- [Bur73] D. L. Burkholder. Distribution function inequalities for martingales. *Ann. Probability*, 1 :19–42, 1973.
- [BY08] Z. Bai and J.F. Yao. Limit theorems for sample eigenvalues in a generalized spiked population model. *Arxiv preprint arXiv :0806.1141*, 2008.
- [CGM09] M. Cranston, D. Gauthier, and T. S. Mountford. On large deviation regimes for random media models. *Ann. Appl. Probab.*, 19(2) :826–862, 2009.
- [CH84] Sándor Csörgő and Peter Hall. The Komlós-Major-Tusnády approximations and their applications. *Austral. J. Statist.*, 26(2) :189–218, 1984.
- [Cha05] Sourav Chatterjee. A simple invariance theorem. *Arxiv preprint math.PR/0508213*, 2005.
- [Cha08] S. Chatterjee. Chaos, concentration, and multiple valleys. *Arxiv preprint arXiv :0810.4221*, 2008.
- [Col05] Benoît Collins. Product of random projections, Jacobi ensembles and universality problems arising from free probability. *Probab. Theory Related Fields*, 133(3) :315–344, 2005.

- [Dav76] Burgess Davis. On the L^p norms of stochastic integrals and other martingales. *Duke Math. J.*, 43(4) :697–704, 1976.
- [DC79] Didier Dacunha-Castelle. Remarque sur l'étude asymptotique du rapport de vraisemblance de deux processus gaussiens stationnaires. *C. R. Acad. Sci. Paris Sér. A-B*, 288(3) :A225–A228, 1979.
- [Dei99] Percy A. Deift. *Orthogonal polynomials and random matrices : a Riemann-Hilbert approach*, volume 3 of *Courant Lecture Notes in Mathematics*. New York University Courant Institute of Mathematical Sciences, New York, 1999.
- [DG07] Percy Deift and Dimitri Gioev. Universality at the edge of the spectrum for unitary, orthogonal, and symplectic ensembles of random matrices. *Comm. Pure Appl. Math.*, 60(6) :867–910, 2007.
- [dlP99] Victor H. de la Peña. A general class of exponential inequalities for martingales and ratios. *Ann. Probab.*, 27(1) :537–564, 1999.
- [DW08] A. B. Dieker and J. Warren. Determinantal transition kernels for some interacting particles on the line. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(6) :1162–1172, 2008.
- [DZ92] P. Deift and X. Zhou. A steepest descent method for oscillatory Riemann-Hilbert problems. *Bull. Amer. Math. Soc. (N.S.)*, 26(1) :119–123, 1992.
- [DZ98] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 1998.
- [EK03] N. El Karoui. On the largest eigenvalue of Wishart matrices with identity covariance when n , p and p/n tend to infinity. *Arxiv preprint math/0309355*, 2003.
- [EK05] Noureddine El Karoui. Recent results about the largest eigenvalue of random covariance matrices and statistical application. *Acta Phys. Polon. B*, 36(9) :2681–2697, 2005.
- [EK07] Noureddine El Karoui. Tracy-Widom limit for the largest eigenvalue of a large class of complex sample covariance matrices. *Ann. Probab.*, 35(2) :663–714, 2007.
- [Fér08] Déphine Féral. On large deviations for the spectral measure of discrete coulomb gas. In *Séminaire de Probabilités, XLI*, volume 1934 of *Lecture Notes in Math.*, pages 19–50. Springer, Berlin, 2008.
- [FN71] D. H. Fuk and S. V. Nagaev. Probabilistic inequalities for sums of independent random variables. *Teor. Veroyatnost. i Primenen.*, 16 :660–675, 1971.
- [FP09] Delphine Féral and Sandrine Péché. The largest eigenvalues of sample covariance matrices for a spiked population : diagonal case. *J. Math. Phys.*, 50(7) :073302, 33, 2009.
- [Fuk73] D. H. Fuk. Certain probabilistic inequalities for martingales. *Sibirsk. Mat. Ž.*, 14 :185–193, 239, 1973.
- [Ful97] William Fulton. *Young tableaux*, volume 35 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1997. With applications to representation theory and geometry.
- [FVDHK08] Anne Fey, Remco Van Der Hofstad, and Marten J. Klok. Large deviations for eigenvalues of sample covariance matrices, with applications to mobile communication systems. *Adv. in Appl. Probab.*, 40(4) :1048–1071, 2008.

- [GRZ99] F. Gamboa, A. Rouault, and M. Zani. A functional large deviations principle for quadratic forms of Gaussian stationary processes. *Statist. Probab. Lett.*, 43(3) :299–308, 1999.
- [GT04] Friedrich Götze and Alexander Tikhomirov. Rate of convergence in probability to the Marchenko-Pastur law. *Bernoulli*, 10(3) :503–548, 2004.
- [GTW01] Janko Gravner, Craig A. Tracy, and Harold Widom. Limit theorems for height fluctuations in a class of discrete space and time growth models. *J. Statist. Phys.*, 102(5-6) :1085–1132, 2001.
- [Gus05] Jonas Gustavsson. Gaussian fluctuations of eigenvalues in the GUE. *Ann. Inst. H. Poincaré Probab. Statist.*, 41(2) :151–178, 2005.
- [GW91] Peter W. Glynn and Ward Whitt. Departures from many queues in series. *Ann. Appl. Probab.*, 1(4) :546–572, 1991.
- [HM07] Ben Hambly and James B. Martin. Heavy tails in last-passage percolation. *Probab. Theory Related Fields*, 137(1-2) :227–275, 2007.
- [HP98] Fumio Hiai and Dénes Petz. Eigenvalue density of the Wishart matrix and large deviations. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 1(4) :633–646, 1998.
- [HP00a] Fumio Hiai and Dénes Petz. A large deviation theorem for the empirical eigenvalue distribution of random unitary matrices. *Ann. Inst. H. Poincaré Probab. Statist.*, 36(1) :71–85, 2000.
- [HP00b] Fumio Hiai and Dénes Petz. *The semicircle law, free random variables and entropy*, volume 77 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000.
- [Ibr07] J.P. Ibrahim. Large deviations for directed percolation on a thin rectangle. 2007.
- [Jam64] Alan T. James. Distributions of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.*, 35 :475–501, 1964.
- [Joh98] Kurt Johansson. On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.*, 91(1) :151–204, 1998.
- [Joh00a] Kurt Johansson. Shape fluctuations and random matrices. *Comm. Math. Phys.*, 209(2) :437–476, 2000.
- [Joh00b] Kurt Johansson. Transversal fluctuations for increasing subsequences on the plane. *Probab. Theory Related Fields*, 116(4) :445–456, 2000.
- [Joh01a] Kurt Johansson. Discrete orthogonal polynomial ensembles and the Plancherel measure. *Ann. of Math. (2)*, 153(1) :259–296, 2001.
- [Joh01b] Kurt Johansson. Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. *Comm. Math. Phys.*, 215(3) :683–705, 2001.
- [Joh01c] Iain M. Johnstone. On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.*, 29(2) :295–327, 2001.
- [Joh02] Kurt Johansson. Non-intersecting paths, random tilings and random matrices. *Probab. Theory Related Fields*, 123(2) :225–280, 2002.
- [Joh08] Kurt Johansson. On some special directed last-passage percolation models. In *Integrable systems and random matrices*, volume 458 of *Contemp. Math.*, pages 333–346. Amer. Math. Soc., Providence, RI, 2008.

- [Joh10] Kurt Johansson. A multi-dimensional Markov chain and the Meixner ensemble. *Ark. Mat.*, 48(1) :79–95, 2010.
- [KMT76] J. Komlós, P. Major, and G. Tusnády. An approximation of partial sums of independent RV's, and the sample DF. II. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 34(1) :33–58, 1976.
- [Kön05] Wolfgang König. Orthogonal polynomial ensembles in probability theory. *Probab. Surv.*, 2 :385–447 (electronic), 2005.
- [KS91] J. Krug and H. Spohn. Kinetic roughening of growing surfaces. *Solids far from equilibrium*, pages 479–582, 1991.
- [Led07] Michel Ledoux. Deviation inequalities on largest eigenvalues. In *Geometric aspects of functional analysis*, volume 1910 of *Lecture Notes in Math.*, pages 167–219. Springer, Berlin, 2007.
- [Lig85] Thomas M. Liggett. An improved subadditive ergodic theorem. *Ann. Probab.*, 13(4) :1279–1285, 1985.
- [Lin22] J. W. Lindeberg. Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Math. Z.*, 15(1) :211–225, 1922.
- [LR10] Michel Ledoux and Brian Rider. Small deviations for beta ensembles. *Préprint*, 2010.
- [Mai07] Mylène Maïda. Large deviations for the largest eigenvalue of rank one deformations of Gaussian ensembles. *Electron. J. Probab.*, 12 :1131–1150 (electronic), 2007.
- [Mar04] James B. Martin. Limiting shape for directed percolation models. *Ann. Probab.*, 32(4) :2908–2937, 2004.
- [Meh91] Madan Lal Mehta. *Random matrices*. Academic Press Inc., Boston, MA, second edition, 1991.
- [MF10] Gregorio Moreno Flores. Asymmetric directed polymers in random environments. *Arxiv preprint arXiv :1009.5576*, 2010.
- [MN98] T. Mikosch and A. V. Nagaev. Large deviations of heavy-tailed sums with applications in insurance. *Extremes*, 1(1) :81–110, 1998.
- [MO07] J. Moriarty and N. O'Connell. On the free energy of a directed polymer in a Brownian environment. *Markov Process. Related Fields*, 13(2) :251–266, 2007.
- [MP67] Vladimir Marchenko and Lionid. Pastur. The eigenvalue distribution in some ensembles of random matrices. *Math. USSR Sbornik*, 1 :457–483, 1967.
- [Obł04] Jan Obłój. The Skorokhod embedding problem and its offspring. *Probab. Surv.*, 1 :321–390 (electronic), 2004.
- [O'C03] Neil O'Connell. Random matrices, non-colliding processes and queues. In *Séminaire de Probabilités, XXXVI*, volume 1801 of *Lecture Notes in Math.*, pages 165–182. Springer, Berlin, 2003.
- [OY02] Neil O'Connell and Marc Yor. A representation for non-colliding random walks. *Electron. Comm. Probab.*, 7 :1–12 (electronic), 2002.
- [Pau07] Debashis Paul. Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statist. Sinica*, 17(4) :1617–1642, 2007.
- [Péc] S. Péché. The edge of the spectrum of random matrices. *Habilitations dirigées des recherches, Université Joseph Fourier Grenoble I, to be submitted*.

- [Péc09] Sandrine Péché. Universality results for the largest eigenvalues of some sample covariance matrix ensembles. *Probab. Theory Related Fields*, 143(3-4) :481–516, 2009.
- [PPR06] Nick Patterson, Alkes L. Price, and David Reich. Population structure and eigenanalysis. *PLoS Genet*, 2(12) :e190, 2006.
- [Rog63] C. A. Rogers. Covering a sphere with spheres. *Mathematika*, 10 :157–164, 1963.
- [Ros81] H. Rost. Nonequilibrium behaviour of a many particle process : density profile and local equilibria. *Z. Wahrsch. Verw. Gebiete*, 58(1) :41–53, 1981.
- [RS78] Michael Reed and Barry Simon. *Methods of modern mathematical physics. IV. Analysis of operators*. Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1978.
- [Ruz06] A. Ruzmaikina. Universality of the edge distribution of eigenvalues of Wigner random matrices with polynomially decaying distributions of entries. *Comm. Math. Phys.*, 261(2) :277–296, 2006.
- [RY99] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.
- [Sag91] Bruce E. Sagan. *The symmetric group*. The Wadsworth & Brooks/Cole Mathematics Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1991. Representations, combinatorial algorithms, and symmetric functions.
- [Sak91] A. I. Sakhanenko. On the accuracy of normal approximation in the invariance principle [translation of Trudy Inst. Mat. (Novosibirsk) **13** (1989), Asimptot. Analiz Raspred. Sluch. Protsess., 40–66 ; MR1037248 (91d :60082)]. *Siberian Adv. Math.*, 1(4) :58–91, 1991. Siberian Advances in Mathematics.
- [Saw72] Stanley Sawyer. A remark on the Skorohod representation. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 23 :67–74, 1972.
- [SB95] Jack W. Silverstein and Z. D. Bai. On the empirical distribution of eigenvalues of a class of large-dimensional random matrices. *J. Multivariate Anal.*, 54(2) :175–192, 1995.
- [Sep97] Timo Seppäläinen. A scaling limit for queues in series. *Ann. Appl. Probab.*, 7(4) :855–872, 1997.
- [Sep98a] T. Seppäläinen. Coupling the totally asymmetric simple exclusion process with a moving interface. *Markov Process. Related Fields*, 4(4) :593–628, 1998. I Brazilian School in Probability (Rio de Janeiro, 1997).
- [Sep98b] Timo Seppäläinen. Exact limiting shape for a simplified model of first-passage percolation on the plane. *Ann. Probab.*, 26(3) :1232–1250, 1998.
- [Sep98c] Timo Seppäläinen. Large deviations for increasing sequences on the plane. *Probab. Theory Related Fields*, 112(2) :221–244, 1998.
- [Sep09] Timo Seppäläinen. Lecture Notes on the Corner Growth Model. 2009.
- [Sko65] A. V. Skorokhod. *Studies in the theory of random processes*. Translated from the Russian by Scripta Technica, Inc. Addison-Wesley Publishing Co., Inc., Reading, Mass., 1965.

- [Sos99] Alexander Soshnikov. Universality at the edge of the spectrum in Wigner random matrices. *Comm. Math. Phys.*, 207(3) :697–733, 1999.
- [Sos02] Alexander Soshnikov. A note on universality of the distribution of the largest eigenvalues in certain sample covariance matrices. *J. Statist. Phys.*, 108(5-6) :1033–1056, 2002. Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays.
- [ST97] Edward B. Saff and Vilmos Totik. *Logarithmic potentials with external fields*, volume 316 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1997. Appendix B by Thomas Bloom.
- [Str67] Volker Strassen. Almost sure behavior of sums of independent random variables and martingales. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66)*, pages Vol. II : Contributions to Probability Theory, Part 1, pp. 315–343. Univ. California Press, Berkeley, Calif., 1967.
- [Sui06] Toufic Suidan. A remark on a theorem of Chatterjee and last passage percolation. *J. Phys. A*, 39(28) :8977–8981, 2006.
- [TV] T. Tao and V. Vu. Random covariance matrices : Universality of local statistics of eigenvalues. *Preprint. arXiv*, 912.
- [TV10] Terence Tao and Van Vu. Random matrices : universality of local eigenvalues statistics. *Arxiv preprint math.PR/0906.0510v10*, 2010.
- [TW74] Shantanu V. Tembe and Ronald W. Wolff. The optimal order of service in tandem queues. *Operations Res.*, 22(4) :824–832, 1974.
- [TW94] Craig A. Tracy and Harold Widom. Level-spacing distributions and the Airy kernel. *Comm. Math. Phys.*, 159(1) :151–174, 1994.
- [Voi91] Dan Voiculescu. Limit laws for random matrices and free products. *Invent. Math.*, 104(1) :201–220, 1991.
- [Wac78] Kenneth W. Wachter. The strong limits of random matrix spectra for sample matrices of independent elements. *Ann. Probability*, 6(1) :1–18, 1978.
- [Wig55] Eugene P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math. (2)*, 62 :548–564, 1955.
- [Wig58] Eugene P. Wigner. On the distribution of the roots of certain symmetric matrices. *Ann. of Math. (2)*, 67 :325–327, 1958.
- [Wis28] J. Wishart. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, 20(1-2) :32, 1928.
- [YBK88] Y. Q. Yin, Z. D. Bai, and P. R. Krishnaiah. On the limit of the largest eigenvalue of the large-dimensional sample covariance matrix. *Probab. Theory Related Fields*, 78(4) :509–521, 1988.
- [Yin86] Y. Q. Yin. Limiting spectral distribution for a class of random matrices. *J. Multivariate Anal.*, 20(1) :50–68, 1986.