Habilitation à diriger les recherches

Spécialité : Mathématiques

présentée et soutenue publiquement par

Mylène MAÏDA

le 13 décembre 2012

Modèles matriciels déformés, mouvement brownien unitaire et transport libre

---

Après avis de :

M.	PHILIPPE BIANE	Université Paris-Est Marne-la-Vallée
M.	JEAN-FRANÇOIS LE GALL	Université Paris-Sud
M.	OFER ZEITOUNI	University of Minnesota et Weizmann Institute

Devant le jury composé de :

M.	PHILIPPE BIANE	Université Paris-Est Marne-la-Vallée
Mme	ANNE BOUTET DE MONVEL	Université Paris-Diderot
M.	DJALIL CHAFAÏ	Université Paris-Est Marne-la-Vallée
Mme	CATHERINE DONATI-MARTIN	Université de Versailles Saint-Quentin-en-Yvelines
Mme	ALICE GUIONNET	Massachussets Institute of Technology
M.	JEAN-FRANÇOIS LE GALL	Université Paris-Sud

Grâce à Chi, pour Elvire

L'habilitation est pour la plupart des jeunes chercheurs l'occasion de faire le bilan sur les quelques années de recherche qui se sont écoulées depuis la fin de leur thèse. Pour moi comme pour presque tout le monde, il y a eu les heures blanches où l'on peine à trouver le temps de faire des "vraies maths", les heures grises où l'on ne sait plus très bien dans quelle direction chercher la solution, les heures noires où les émotions paralysent l'esprit. Mais à l'heure du bilan, une chose me frappe surtout : je me rends compte combien, à chaque étape de cette aventure, j'ai eu de la chance !

Lorsque, fraîchement agrégée, en quête d'un sujet de DEA, j'ai frappé à la porte du directeur des études à l'ENS, j'ai eu la chance de tomber sur Jean-François Le Gall, qui a eu la clairvoyance de m'aiguiller vers Alice Guionnet, alors chargée de recherches à l'ENS. Jean-François, je te remercie pour cela mais j'aurai d'autres occasions de te remercier plus loin.

Ensuite, bien sûr, ma plus grande chance a été d'être encadrée par Alice, sur un si beau sujet que les matrices aléatoires. Ce domaine n'avait pas encore pris l'essor qu'il a connu depuis mais ton enthousiasme et la variété du paysage mathématique que tu me faisais découvrir, m'ont tout de suite convaincue. Tu as été juste parfaite, toujours à la bonne distance, aujourd'hui comme hier.

Avant de te suivre à Lyon, j'ai eu la chance que tu m'envoies une année en Israël, travailler avec Ofer Zeitouni. Ofer, je te remercie pour cette année très formatrice, dont certains bénéfices se sont vus directement dans ma thèse et certains autres ont mis plusieurs années à mûrir. Depuis, j'ai eu la chance de te rendre visite au Weizmann, de te croiser régulièrement en conférences : tu as toujours prêté une oreille attentive à mes maths et répondu avec toute ta sagacité à mes questions, apportant parfois une aide tout à fait cruciale. Il était très naturel que tu sois l'un des rapporteurs de ce mémoire, je te remercie d'avoir accepté, j'en suis heureuse et honorée.

Je ne raconterai pas ici en détail les chances et les péripéties des années lyonnaises et de l'année nanterroise ; les curieux pourront les deviner en filigrane dans les remerciements de ma thèse.

Je voudrais cependant dire que à cette période que j'ai eu la chance de rencontrer Philippe Biane. Philippe, je suis très honorée que tu aies accepté à la fois de rapporter ce mémoire et de faire partie du jury. Je te connais depuis plusieurs années maintenant et je suis toujours aussi impressionnée par la profondeur de tes mathématiques. Il est de tes articles que l'on a décortiqué, repris dix fois, vingt fois en y trouvant sans cesse quelque chose de nouveau. C'est toujours une grande chance de pouvoir discuter avec toi de nos questions mathématiques et merci en particulier pour ton aide dans la simplification de la preuve de l'inégalité T_1 libre.

Je vais aussi vous raconter que, pendant cette période, j'ai subi un rite initiatique par lequel passent beaucoup de jeunes probabilistes français, celui de passer un mois de juillet à Saint-Flour. J'ai eu la chance d'y être un peu adoptée par une partie du groupe des log-Sob de Toulouse, en particulier Katy Paroux, Djalil Chafaï, Florent Malrieu et Pierre Fougères. Votre jolie histoire de maths et d'amitié m'a bien motivée pour continuer à un moment où je doutais beaucoup. Depuis ce temps-là, nous avons eu la chance, Djalil, que tu en viennes aux matrices aléatoires. Ta façon si particulière de poser les questions, ton sens de l'esthétique mathématique et ton culte de l'éclectisme me fascinent depuis longtemps et je suis vraiment très heureuse que tu aies accepté de faire partie de ce jury.

Quelques mois après la fin de ma thèse, j'ai eu une autre grande chance, celle d'être recrutée comme maître de conférences à Orsay. Je remercie l'ensemble des personnes qui ont pris part à la décision pour leur

confiance. J'ai ainsi eu, grâce notamment à Jean-François, la chance, et la lourde resposabilité, d'enseigner à des étudiants incroyables, dont certains sont déjà devenus de brillants chercheurs. J'ai aussi pu bénéficier d'un des meilleurs environnements mathématiques possibles et d'un cadre de travail exceptionnel. Je remercie pour cela tout le département, l'équipe de probabilités et statistiques en particulier, avec en outre une mention particulière pour la chaleureuse équipe d'enseignants à la préparation au CAPES.

Au moment où l'on essaie de voler de ses propres ailes mathématiques, les collaborations que l'on a la chance de nouer sont souvent déterminantes. Je crois que j'ai été plus comblée dans ce domaine puisque j'ai eu la chance de travailler avec des chercheuses et chercheurs tels que Florent Benaych-Georges, Pascal Bianchi, Merouane Debbah, Catherine Donati-Martin, Alice Guionnet, Thierry Lévy, Édouard Maurel-Segala, Jamal Najim, Sandrine Péché. La seule lecture de cette liste vous donne un aperçu de ma chance ! Je vous remercie tous et espère que nous aurons l'occasion de refaire un bout de chemin mathématique ensemble. Je voudrais remercier particulièrement Edouard, Thierry et Catherine dont la présence mathématique et l'enthousiasme m'ont beaucoup aidée aux heures noires. Vous avez tous les trois des manières très différentes de faire des mathématiques, que j'aimerais toutes trois prendre en exemple. Merci à toi, Catherine, d'avoir accepté de faire partie de mon jury.

Je connais Anne Boutet de Monvel depuis plusieurs années mais ce n'est qu'il y a quelques mois que j'ai eu la chance de travailler avec elle pour l'organisation de deux colloques à l'IHP. Merci à toi, Anne, de ton amitié protectrice et réconfortante. Je suis heureuse que tu m'accordes ta confiance.

Anne, Alice, Catherine, j'ai la chance grâce à vous d'avoir un jury strictement paritaire. Merci à vous d'être si inspirante dans vos vies de mathématiciennes.

Lorsqu'après beaucoup d'hésitations, je me suis vraiment décidée à écrire ce mémoire, il était déjà bien tard dans la saison et le temps filait vite entre deux tétées. Heureusement j'ai eu la chance que Jean-François ait tout de suite accepté d'être mon rapporteur local. Je n'ai tenu aucun des délais impartis, mais tu as malgré tout rendu ton rapport à temps. Merci de ton aide, de tes conseils et de tes encouragements.

Merci aussi à Catherine Ardin, Nicole Lhermitte et Valérie Lavigne pour l'aide administrative depuis que je suis à Orsay et en particulier dernièrement.

Le scientifique est souvent ingrat envers ceux qui l'entourent. J'essaie de ne pas oublier la chance que j'ai d'avoir toujours eu le soutien de mes parents, de mon petit frère et de ma grand-mère qui me manque tant maintenant, la chance aussi de partager des amitiés d'une force étonnante. Lumi, Dimitri et Suzanne, Nathaëlle et Gaël, Brigitte et Nicolas, Yann, Alice, Claire, Laure, Lionel, Mireille, Eugénie, vous êtes ma garde rapprochée, merci à vous d'être là quand il le faut.

Enfin il est une chance que je ne risque pas d'oublier tant elle m'éblouit tous les jours : la chance qui a mis Chi sur mon chemin est tout simplement prodigieuse et son résultat miraculeux. Comment dire merci pour ça ?

Liste des travaux

[A1] F. Benaych-Georges, A. Guionnet, et M. Maïda. Fluctuations of the extreme eigenvalues of finite rank deformations of random matrices. *Electron. J. Probab.*, 16(60) :1621–1662, 2011.

[A2] F. Benaych-Georges, A. Guionnet, et M. Maïda. Large deviations of extreme eigenvalues of finite rank deformations of deterministic matrices. *Prob. Theory. Rel. Fields.*, 2011.

[A3] P. Bianchi, M.. Debbah, M. Maïda, et J. Najim. Performance of statistical tests for single-source detection using random matrix theory. *IEEE Transactions on Information Theory*, 57(4), 2011.

[A4] C. Donati-Martin et M. Maïda. Large deviations for the largest eigenvalue of an Hermitian brownian motion. A paraître dans *ALEA*.

[A5] A. Guionnet et M. Maïda. Character expansion method for the first order asymptotics of a matrix integral. *Prob. Theory. Rel. Fields.*, 132(4), 2004.

[A6] A. Guionnet et M. Maïda. An asymptotic log-Fourier interpretation of the R-transform and related asymptotics of the spherical integrals. *Journal of Functional Analysis*, 222(2), 2005.

[A7] T. Lévy et M. Maïda. Central limit theorem for the heat kernel measure on the unitary group. *Journal of Functional Analysis*, 259(12), 2010.

[A8] M. Maïda. Large deviations for the largest eigenvalue of rank one deformations of gaussian ensembles. *Elec. J. Probab.*, 12 :1131–1150, 2007.

[A9] M. Maïda, J. Najim, et S. Péché. Large deviations for weighted empirical mean with outliers. *Stochastic Processes and Their Applications*, 117(10), 2007.

Prépublication

[A10] M. Maïda et É. Maurel-Segala. Free transport-entropy inequalities for non-convex potentials and application to concentration for random matrices. *Preprint*, (arXiv :1204.3208v1), 2012.

Actes de conférences

[C1] P. Bianchi, M.. Debbah, M. Maïda, et J. Najim. Performance analysis of some eigen-based hypothesis tests for collaborative sensing. *SSP 2009, Cardiff.*

[C2] P. Bianchi, L. Cardoso, M. Maïda, et J. Najim. Ecoute Coopérative de Spectre pour la Radio Cognitive. *GRETSI 2009, Dijon, France.*

[C3] M. Maïda. A Log-Fourier Interpretation of the R-Transform and Related Asymptotics of the Spherical Integrals, *Oberwolfach reports, 15/2005.*

[C4] M. Maïda. Fluctuations and large deviations of some perturbed random matrices. *Preprint.*

Thèse

[T] M. Maïda. Etude asymptotique du spectre de grandes matrices aléatoires. applications aux modèles de matrices. *Thèse de doctorat de l'ENS Lyon*, 2004.

Les articles [A5] et [A9] ne sont pas présentés dans ce mémoire. Le papier [C4] est un article de survol qui reprend notamment les résultats de [A1] et [A2].

Table des matières

Introduction

Le but du présent mémoire est de donner une présentation synthétique des travaux que j'ai effectués depuis ma thèse, soutenue en décembre 2004. Mes recherches s'inscrivent dans le cadre général de l'étude du comportement asymptotique de matrices aléatoires, essentiellement par des techniques de grandes déviations et de probabilités libres. La présente introduction a pour objectif de servir de guide de lecture pour le mémoire.

Dans l'ordre chronologique, mon premier projet a consisté en l'étude d'une intégrale à plusieurs matrices en utilisant un résultat de grandes déviations pour des mesures empiriques de tableaux d'Young. Il s'agissait de rendre mathématiquement rigoureuse, dans un cas précis au moins, une technique assez utilisée par les physiciens dite de développement en caractères. Ce résultat [A5] ayant peu de liens avec le reste de mes travaux, je ne le décrirai pas en détail dans le mémoire.

Ensuite, je me suis intéressée dans [A6] à une intégrale matricielle, dite d'Itzykson-Zuber, qui apparaît dans de nombreux modèles et est donnée, pour deux matrices hermitiennes A_n et B_n de taille n, par

$$I_n(A_n, B_n) = \int e^{ntr(UA_nU^*B_n)}dU,$$

où dU est la mesure de Haar sur le groupe unitaire (ou orthogonal) dans le régime où l'une des deux matrices, disons B_n, est de rang fixé (indépendant de la taille n des matrices). L'asymptotique fait intervenir une fonctionnelle bien connue en probabilités libres, la R-transformée. Dans ce mémoire, nous n'insisterons pas sur les liens de ce résultat avec les probabilités libres mais plutôt sur l'utilisation de ces asymptotiques pour l'étude de modèles matriciels déformés, auxquels est consacré le premier chapitre.
En effet, ce que l'on appelle un modèle matriciel déformé est la somme d'une matrice aléatoire dont on connaît bien le spectre à laquelle on rajoute une perturbation, déterministe ou aléatoire (en général indépendante de la partie non perturbée). Ici on s'intéressera au cas où la perturbation est de rang fini (fixé indépendamment de la taille des matrices). On étudie alors comment l'ajout de la perturbation affecte le comportement - notamment les valeurs propres extrêmes - du modèle initial, rejoignant ainsi les questions très à la mode d'universalité ou non universalité en matrices aléatoires.
Or, l'intégrale d'Itzykson-Zuber apparaît notamment lorsqu'on exprime la loi jointe des valeurs propres pour le modèle $W_n + A_n$, avec W_n une matrice de l'Ensemble Unitaire Gaussien (GUE) et A_n une matrice déterministe de rang un. En affinant les résultats obtenus dans ma thèse sur l'intégrale sphérique, j'ai pu montrer dans [A8] un résultat de grandes déviations pour la plus grande valeur propre dans le cadre d'une perturbation de rang un. Ces résultats seront présentés dans la partie 1.3.2.
Outre leur intérêt théorique lié aux questions d'universalité, les modèles déformés apparaissent dans de nombreuses applications (finance, communications numériques) où interviennent des modèles de type signal plus bruit. Une adaptation du résultat de [A8] sur la déformation de rang un nous a permis, dans un travail en collaboration avec P. Bianchi, M. Debbah et J. Najim [A3] de comparer l'efficacité de différents estimateurs qui ont été proposés pour tester la présence d'un signal dans un contexte de détection collaborative en communications numériques. Cette application sera brièvement présentée dans la partie 1.3.3.

L'ensemble de ces travaux a constitué pour moi une première approche des modèles matriciels déformés et nous a permis de donner les premiers résultats de grandes déviations pour ces modèles, dans le cas de déformation de rang un.
Le problème de l'étude des déformations de rang fini est plus délicat. Une première tentative nous a conduit à étudier l'intégrale sphérique en rang supérieur. Dans ce but, en collaboration avec J. Najim et S. Péché [A9], nous avons obtenu dans des résultats théoriques de grandes déviations pour des moyennes empiriques pondérées qui permettent une compréhension (seulement partielle) du comportement de cette intégrale en rang supérieur mais cette compréhension n'est pas suffisante pour en tirer les grandes déviations des valeurs propres extrêmes. C'est pourquoi nous avons choisi de ne pas présenter en détail ce travail dans le mémoire.
Dans deux articles [A1, A2], en collaboration avec F. Benaych-Georges et A. Guionnet, nous avons repris ce problème sous un angle différent. Au lieu d'utiliser sur l'expression explicite de la loi jointe des valeurs propres - qui est propre au cas où la matrice non perturbée a une forme particulière - nous utilisons le fait que les valeurs propres de la matrice déformée sont les zéros d'un déterminant de taille fixée, qui dépend du rang de la déformation plutôt que de la taille des matrices. Ceci nous a permis, dans deux articles jumeaux [A1, A2], d'étudier les fluctuations et les déviations des valeurs propres extrêmes de nombreux modèles déformés. Le papier [A1] qui étudie les fluctuations est détaillé dans la partie 1.2 tandis que le contenu de [A2] est présenté dans les parties 1.3.5 et 1.3.6.
Enfin, un dernier travail [A4], que j'ai effectué avec Catherine Donati-Martin et qui est présenté dans la partie 1.3.4, est apparenté à l'étude des modèles déformés : nous considérons le mouvement brownien hermitien, qui est une version dynamique du GUE et nous étudions les déviations du processus de sa plus grande valeur propre. Comme on s'autorise à partir d'une condition initiale qui est une matrice de rang un, cela correspond à temps fixé à une déformation de rang un d'une matrice du GUE.

Un cousin du mouvement brownien hermitien est le mouvement brownien sur le groupe unitaire $U(N)$. En collaboration avec T. Lévy [A7], je me suis intéressée à son comportement quand la dimension N tend vers l'infini. Il constitue un objet central dans la théorie de Yang-Mills quand le groupe de jauge est $U(N)$. P. Biane a montré la convergence de ce mouvement matriciel vers un processus libre appelé mouvement brownien unitaire libre. Avec T. Lévy, nous établissons un théorème de la limite centrale pour des fonctions $U \mapsto trf(U)$ (assez régulières) sur le groupe unitaire et étudions la forme quadratique qui donne la covariance limite. Elle s'exprime aussi en termes du mouvement brownien unitaire libre et converge en temps grand vers celle qui apparaît pour les fluctuations d'une matrice distribuée selon la mesure de Haar (ce résultat avait été établi par Diaconis et Evans). Le chapitre 2 tout entier est consacré à la présentation de [A7].

Enfin, dans le chapitre 3, je présente un travail récent, actuellement soumis pour publication qui porte sur le transport libre, plus précisément l'équivalent libre des inégalités transport-entropie classiques et en particulier les inégalités de Talagrand.
Talagrand a en particulier montré que la mesure gaussienne dans $\mathbb{R}^n$ vérifiait l'inégalité T_2. On sait que pour diverses raisons, la loi semi-circulaire est l'analogue en probabilités libres de la gaussienne. Par ailleurs, dans les années 80, D. Voiculescu a introduit la notion d'entropie libre. Il est donc naturel de se demander si la loi semi-circulaire vérifie un analogue des inégalités de Talagrand, où la notion d'entropie classique est remplacée par sa version libre. P. Biane et D. Voiculescu ont répondu par l'affirmative à cette question. Hiai, Petz et Ueda ont ensuite généralisé ce résultat aux mesures d'équilibre associées à un potentiel strictement convexe (analogues des mesures log-concaves).
Avec Édouard Maurel-Segala [A10], nous avons cherché à généraliser leur résultat au-delà du cas convexe. Comme les mesures d'équilibre associées ne sont pas nécessairement à support connexe, on ne peut guère espérer une inégalité T_2 mais nous avons réussi à montrer pour ces mesures l'analogue libre de l'inégalité T_1. De là, nous en déduisons une inégalité de concentration pour des modèles de matrices dits unitairement invariants, qui sont très étudiés en ce moment.

Chapitre 1

Sur les modèles matriciels déformés [A1, A2, A3, A4, A6, A8]

L'objet de ce chapitre est de présenter l'ensemble de mes travaux concernant les valeurs propres extrêmes de certains modèles déformés. L'article [A1] en étudie les fluctuations tandis que les autres articles traitent de résultats de grandes déviations et de leurs applications.

1.1 Présentation générale du problème

La question suivante est un problème classique qui a suscité de nombreux travaux : "*soit A et B deux matrices hermitiennes de même taille. Que peut-on dire du spectre de leur somme $A + B$?*".
Le problème est notamment posé au début du vingtième siècle par Weyl dans [Wey12]. Il y donne un ensemble de conditions nécessaires, connues sous le nom d'inégalités d'entrelacement de Weyl : si $\lambda_1(A) \geq \ldots \geq \lambda_n(A)$, $\lambda_1(B) \geq \ldots \geq \lambda_n(B)$ et $\lambda_1(A+B) \geq \ldots \geq \lambda_n(A+B)$ sont les spectres respectifs de A, B et $A+B$, alors

$$\lambda_{i+j+1}(A+B) \leq \lambda_{i+1}(A) + \lambda_{j+1}(B)$$

lorsque $0 \leq i, j, i+j < n$.
Ensuite, cela prit très longtemps pour parvenir à établir des conditions nécessaires et suffisantes. Dans les années 60, Horn énonça la bonne conjecture mais la réponse finale arriva dans les années 90 dans une série de papiers [Kly98, HR95, KT01].

Si on s'intéresse maintenant au problème asymptotique, quand la taille des matrices tend vers l'infini, une avancée importante est venue de la théorie des probabilités libres, avec la notion de liberté asymptotique. Cette propriété peut être résumée ainsi : "*Si A et B sont des matrices de grande taille en position générique l'une par rapport à l'autre, le spectre limite de leur somme dépend seulement de leur spectre respectif et est donné par la convolution libre de leurs deux spectres.*" Nous renvoyons le lecteur à l'appendice pour ce qui concerne les notions de probabilités libres, en particulier les notions de liberté asymptotique et de convolution libre.

Nous nous intéressons au problème du spectre asymptotique de la somme dans le cas où l'une des matrices est de rang fini, fixé indépendamment de la taille des matrices. On peut énoncer le problème de la façon suivante : choisissons notre ensemble de matrices préféré, pour lequel on connaît bien le comportement global et local de son spectre (mesure spectrale asymptotique, convergence et fluctuations des valeurs propres extrêmes etc.) Ajoutons à notre matrice aléatoire une perturbation, ici de rang fini. Comment le spectre est-il affecté par cette perturbation ?

Grâce aux inégalités de Weyl, on peut facilement vérifier que le comportement du spectre reste inchangé au niveau macroscopique. Seules les valeurs propres extrêmes peuvent être affectées de manière significative. Nous allons étudier dans ce chapitre les limites presque sûres, les fluctuations et les déviations de ces valeurs propres extrêmes pour différents modèles de ce type.

Le texte de ce chapitre, en particulier celui de la section 1.2, est en grande partie inspiré par l'article de survol [C4].

Résumé de la partie 1.1 :

Les modèles matriciels déformés que nous allons étudier sont donnés comme somme d'une matrice, aléatoire ou déterministe, dont on connaît bien le comportement du spectre et d'une matrice de rang fini, déterministe ou indépendante de la matrice non perturbée. Le comportement global n'est pas affecté par la perturbation mais on étudie les changements pour les valeurs propres extrêmes.

1.2 Fluctuations pour certains modèles matriciels déformés

1.2.1 La transition BBP

Avant de donner un panorama de la littérature sur ces modèles, nous allons décrire en détails les premiers résultats rigoureux sur le sujet. L'article [BBAP05] est dû à Baik, Ben Arous et Péché et le phénomène que nous allons décrire s'appelle la transition BBP.

On considère le modèle suivant : soit G_n une matrice $n \times m$ dont les vecteurs colonnes sont des vecteurs gaussiens centrés de matrice de covariance Σ et $S_n = \frac{1}{m} G_n G_n^*$ la matrice de covariance empirique associée. On suppose que $\frac{n}{m} \to c \in (0,1)$ et que Σ est une perturbation de l'identité dans le sens où elle a au plus r valeurs propres différentes de un [1]. On note $\ell_1 \geq \ell_2 \geq \ldots \geq \ell_r \geq 1, \ldots, 1$ les valeurs propres de Σ.

Rappelons d'abord les résultats concernant le modèle non-perturbé. Si Σ est la matrice identité, le modèle correspondant s'appelle l'Ensemble de Laguerre Unitaire (LUE) et les résultats suivants sont maintenant bien connus :

0. la mesure spectrale limite est la loi de Marcenko-Pastur [2] (de support $[(1-\sqrt{c})^2, (1+\sqrt{c})^2]$),
1. pour tout k fixé, les k plus grandes valeurs propres convergent vers le bord du support $(1+\sqrt{c})^2$,
2. elles ont des fluctuations d'ordre $n^{-2/3}$ suivant la loi de Tracy-Widom unitaire.

Comme nous l'avons déjà dit plus haut, il est facile de vérifier que le régime global n'est pas affecté par la perturbation : la propriété 0. reste valide quelle que soit l'intensité de la perturbation et nous ne la répèterons pas à chaque fois. Nous nous concentrons sur le comportement des plus grandes valeurs propres.

Avant d'énoncer les résultats de [BBAP05], on rappelle qu'une matrice de l'Ensemble Unitaire Gaussien (GUE) de taille $k \times k$ est une matrice aléatoire hermitienne $(X_{ij})_{i,j=1,\ldots,k}$ telle que les entrées au-dessus de la diagonale $(X_{ij})_{i<j}$ sont des variables gaussiennes complexes centrées de variance $1/k$ et les entrées sur la diagonale $(X_{ii})_{1 \leq i \leq k}$ sont des variables gaussiennes réelles centrées de variance $1/k$. Il s'agit de l'ensemble de matrices aléatoires qui a été le plus étudié et est le mieux compris.

La description suivante résume les résultats de [BBAP05] :
• Si $\ell_1 < 1 + \sqrt{c}$, on est dans le régime sous-critique et les propriétés 1. et 2. restent inchangées.
• Si $\ell_1 = \ldots = \ell_k > 1 + \sqrt{c}$, on est dans le régime sur-critique et

1. la plus grande valeur propre converge en dehors du support vers $\ell_1 + \frac{c\ell_1}{\ell_1 - 1}$

1. en fait plus grandes que un.
2. On appelle loi de Marcenko-Pastur la mesure de probabilité dont la densité sur $\mathbb{R}$ est $\frac{1}{2\pi x}\sqrt{(x-(1-\sqrt{c})^2)((1+\sqrt{c})^2-x)}\mathbf{1}_{[(1-\sqrt{c})^2,(1+\sqrt{c})^2]}$

2. les fluctuations sont dans l'échelle $n^{-1/2}$ et suivent la même loi que la plus grande valeur propre d'une matrice du GUE de taille $k \times k$ (en particulier, si ℓ_1 est valeur propre simple de Σ, la plus grande valeur propre a simplement des fluctuations gaussiennes).

Dans la suite, à chaque fois qu'un modèle aura ce type de comportement, on dira qu'il exhibe une **transition BBP**.
Dans [BBAP05], les auteurs traitent aussi le cas critique lorsque $\ell_1 = \ldots = \ell_k = 1 + \sqrt{c}$, mais nous n'insistons pas sur ce point, qui est spécifique à chaque modèle.

Notons enfin que la perturbation est multiplicative alors que nous allons plutôt étudier des perturbations additives mais cela va donner le même type de comportement (cf par exemple [Péc06]).

1.2.2 Un rapide résumé de la littérature sur les fluctuations

Depuis l'article de Baik, Ben Arous et Péché, la compréhension des modèles déformés a suscité de nombreux travaux.

Une grande partie de cette littérature concerne les applications, notamment statistiques de ces modèles. Le premier article important dans cette direction est probablement [Joh98] qui s'intéresse à l'analyse en composante principale (cf aussi [EK05]). Dans de nombreux papiers, le signal (de quelque nature qu'il soit) est représenté par la matrice de rang fini (avec un nombre fixé de paramètres significatifs) et la matrice non perturbée modélise le bruit. On se demande comment l'observation des valeurs propres de l'ensemble signal plus bruit permet d'accéder aux paramètres significatifs. Les résultats sur les fluctuations que nous allons exposer permettent de construire des tests sur ces paramètres tandis que les résultats de grandes déviations permettent d'évaluer la puissance de ces tests. Nous ne donnerons pas une revue exhaustive de cette littérature mais nous développerons un exemple, tiré de [A3], d'un test statistique pour les communications numériques. Nous résumons pour l'instant les résultats théoriques sur le sujet.

Comme nous l'avons dit plus haut, les travaux [BBAP05, Péc06] traitent de modèles avec des entrées gaussiennes (LUE et GUE perturbés) et il y a eu rapidement des tentatives pour étendre ces résultats au-delà du cas gaussien.

En 2006, Féral and Péché [FP07] ont montré, par des techniques de combinatoire des moments, que le modèle $W_n + A_n$, où W_n est une matrice hermitienne dont les entrées au-dessus de la diagonale sont complexes, indépendantes identiquement distribuées (iid) avec une loi ayant des moments sous-gaussiens et A_n est une perturbation de rang un telle que $(A_n)_{ij} = \frac{\theta}{n}$, pour tout i, j fait apparaître une transition BBP.

En 2008, Bai and Yao dans [BY08] ont étudié en détails les fluctuations des valeurs propres qui convergent hors du support de la mesure limite (bulk) pour des modèles assez généraux et ils ont montré que ces fluctuations, comme dans le cas BBP, sont dans l'échelle $n^{-1/2}$ et suivent les lois des valeurs propres de matrices du GUE (pour des entrées complexes) dont la taille est donnée par la multiplicité des valeurs propres surcritiques de la perturbation.

Ensuite, en 2009, les premiers phénomènes "non-BBP" ont été mis en évidence. Dans [CDMF09], Capitaine, Donati-Martin et Féral ont montré, entre autres, que les fluctuations des valeurs propres convergeant hors du bulk ne sont pas universelles. Si W_n est une matrice hermitienne dont les entrées au-dessus de la diagonale ont une loi symétrique μ vérifiant certaines conditions et A_n est de rang un de la forme $A_n = \mathrm{diag}(\theta, 0, \ldots, 0)$, avec θ assez grand, alors les fluctuations de la plus grande valeur propre de $W_n + A_n$ ne sont plus gaussiennes mais suivent une loi qui est la convolée de μ avec une gaussienne.

Dans [CDMF09], l'explication de ce défaut d'universalité restait un peu mystérieuse mais dans un travail ultérieur [CDMF], les mêmes auteurs ont montré que la forme des vecteurs propres de la perturbation jouait un rôle crucial. S'ils sont délocalisés comme dans [FP07], on observe la transition BBP ; sinon, comme dans [CDMF09], les fluctuations peuvent dépendre de la loi des entrées de la matrice non perturbée.

Signalons que dans ce chapitre, nous travaillerons uniquement dans le cadre de vecteurs propres délocalisés.

1.2.3 Présentation des modèles étudiés

Nous présentons d'abord le cas où le modèle non perturbé est déterrministe puis nous expliquerons comment les résultats peuvent être généralisés à de nombreux ensembles de matrices usuels (Wigner, Wishart etc.)

Soit X_n hermitienne, deterministe de valeurs propres $\lambda_1^n \geq \ldots \geq \lambda_n^n$.
Nous faisons l'hypothèse suivante sur le spectre de X_n : quand n tend vers l'infini,

$$(H1) \quad \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i^n} \longrightarrow \mu_X, \lambda_1^n \longrightarrow a, \lambda_n^n \longrightarrow b,$$

avec μ_X une mesure à support compact et a et b sont respectivement les bords gauche et droit du support de μ_X.

Nous ajoutons alors à X_n une perturbation de rang fini R_n et nous considérons le modèle perturbé $\widetilde{X_n} = X_n + R_n$. On notera $\widetilde{\lambda_1^n} \geq \ldots \geq \widetilde{\lambda_n^n}$ les valeurs propres de $\widetilde{X_n}$. La perturbation de rang fini R_n prend la forme suivante :

$$R_n = \sum_{j=1}^{r} \theta_i G_i^n (G_i^n)^*,$$

avec

$$\theta_1 \geq \cdots \geq \theta_{r_0} > 0 > \theta_{r_0+1} \geq \cdots \geq \theta_r$$

qui sont fixés et indépendants de n et $\sqrt{n} G_i^n$ des vecteurs à entrées iid de loi ν satisfaisant une inégalité de Sobolev logarithmique [3]

On peut aussi prendre R_n de la forme

$$R_n = \sum_{j=1}^{r} \theta_i U_i^n (U_i^n)^*,$$

où les U_i^n sont obtenus à partir des $\sqrt{n} G_i^n$ par un procédé d'orthonormalisation de Gram-Schmidt.

En particulier, les vecteurs propres de cette perturbation sont délocalisés.

1.2.4 Convergence presque sûre des valeurs propres extrêmes

Avant de donner nos résultats sur les fluctuations, précisons la convergence de ces valeurs propres extrêmes.

On note

$$G_{\mu_X}(z) := \int \frac{1}{z-x} d\mu_X(x), \tag{1.2.1}$$

on définit

$$\frac{1}{\underline{\theta}} = \lim_{z \to a^-} G_{\mu_X}(z), \frac{1}{\overline{\theta}} = \lim_{z \to b^+} G_{\mu_X}(z)$$

3. On rappelle qu'une mesure de probabilité μ sur $\mathbb{R}$ vérifie une inégalité de Sobolev logarithmique avec une constante c si pour toute fonction f dérivable et dans $L^2(\mu)$,

$$\int f^2 \log \frac{f^2}{\int f^2 d\mu} d\mu \leq 2c \int |f'|^2 d\mu.$$

Cette hypothèse est technique et nous permet d'utiliser certaines propriétés de concentration.

et

$$\rho_\theta := \begin{cases} G_{\mu_X}^{-1}(1/\theta) & \text{si } \theta \in (-\infty, \underline{\theta}) \cup (\overline{\theta}, +\infty), \\ a & \text{si } \theta \in [\underline{\theta}, 0) \\ b & \text{si } \theta \in (0, \overline{\theta}] \end{cases}$$

Alors, la convergence presque sûre des valeurs propres extrêmes est donnée par un résultat de Benaych et Rao [BGR11].

Théorème 1.2.1. *Soit* $r_0 \in \{0, \ldots, r\}$ *tel que*

$$\theta_1 \geq \cdots \geq \theta_{r_0} > 0 > \theta_{r_0+1} \geq \cdots \geq \theta_r.$$

Les plus grandes[4] *valeurs propres ont le comportement suivant : pour tout* $i \in \{1, \ldots, r_0\}$ *on a*

$$\widetilde{\lambda}_i^n \xrightarrow{p.s.} \rho_{\theta_i}$$

et pour $i > r_0$,

$$\widetilde{\lambda}_i^n \xrightarrow{p.s.} b,$$

où la convergence a lieu presque sûrement (p.s.).

Notre modèle montre donc la première propriété de la transition BBP : si la perturbation est faible, les plus grandes valeurs propres convergent vers le bord du support de la mesure limite ; si elle est forte, elles convergent en dehors.

1.2.5 Universalité des fluctuations gaussiennes hors du bulk

La seconde propriété de la transition BBP est le fait que les fluctuations des valeurs propres qui convergent hors du bulk sont dans l'échelle $n^{-1/2}$ et sont de type gaussien (en fait gouvernées par une petite matrice du GUE)

Sous les hypothèses supplémentaires suivantes

$$(H2) \quad \text{la convergence } \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i^n} \longrightarrow \mu_X \text{ est à vitesse au moins } 1/\sqrt{n},$$

notre modèle fait apparaître cette seconde propriété. Nous avons en effet le résultat suivant. Soit $\alpha_1 > \cdots > \alpha_q > 0$ les différentes valeurs des θ_i telles que $\rho_{\theta_i} > b$.
Pour tout j, soit I_j l'ensemble des indices i tels que $\theta_i = \alpha_j$. Soit $k_j = |I_j|$.

Théorème 1.2.2. *Sous les hypothèses* $(H1), (H2)$ *si le quatrième cumulant de la loi* ν, *noté* $\kappa_4(\nu)$, *est nul*[5], *le vecteur aléatoire*

$$\left(\gamma_i := \sqrt{n}(\widetilde{\lambda}_i^n - \rho_{\theta_i}), i \in I_j\right)_{1 \leq j \leq q}$$

converge en loi vers le vecteur des valeurs propres de $(c_j M_j)_{1 \leq j \leq q}$ *avec les* M_j *des matrices du GUE de taille* $k_j \times k_j$ *indépendantes et* c_j *une constante explicite dépendant seulement de* μ_X *et* α_j.

4. On a un résultat équivalent pour les plus petites valeurs propres. Dans la suite, nous donnons seulement les résultats pour les plus grandes valeurs propres mais il existe à chaque fois des résultats correspondants concernant les plus petites.

5. On a des résultats similaires lorsque $\kappa_4(\nu)$ est non nul. Seules les lois limites sont un peu différentes. On renvoie le lecteur à [A1] pour plus de détails.

1.2.6 Non universalité des fluctuations près du bulk

Dans la transition BBP présentée plus haut, les valeurs propres qui convergent au bord du support de la mesure limite ont les mêmes fluctuations que celles du modèle non perturbé (dans le cas du LUE, il s'agissait de lois de Tracy-Widom).

Dans notre modèle, nous avons aussi étudié la question de ces valeurs propres qui collent au bord. Leur étude est beaucoup plus délicate que celles des valeurs propres hors du bulk.

Comme nous l'avons dit dans l'introduction, nous ne traitons pas le cas des paramètres critiques. Nos résultats sont les suivants :

Théorème 1.2.3. *Sous des hypothèses supplémentaires sur X_n, si aucun des θ_i n'est critique, avec une probabilité exponentiellement proche de un, les valeurs propres de $\widetilde{X}_n$ qui convergent vers a ou b restent à distance $n^{-1+\varepsilon}$ des valeurs propres extrêmes de X_n, pour un $\varepsilon > 0$.*

Les fluctuations des valeurs propres près du bulk ne sont donc pas universelles, dans le sens où elles suivent celles des valeurs propres de X_n, qui peuvent être a priori dans n'importe quelle échelle et suivant n'importe quelle loi.

Avant de donner un énoncé plus précis des hypothèses et du théorème, on peut donner une explication grossière du phénomène : pour des valeurs fixées des θ_i, on a un phénomène de répulsion de la part des valeurs propres de X_n au bord.

- S'il y a beaucoup de valeurs propres de X_n près du bord, la répulsion est très forte et les valeurs propres extrêmes de $\widetilde{X}_n$ convergent hors du bulk
- Si la répulsion est moins forte, les valeurs propres extrêmes de $\widetilde{X}_n$ convergent au bord du bulk
- Si la répulsion est encore moins forte, les valeurs propres extrêmes de $\widetilde{X}_n$ collent à celles de X_n même au niveau des fluctuations.

Pour que la répulsion soit faible, il faut que l'espacement entre les valeurs propres de X_n proches du bord reste suffisamment grand, au sens suivant :

$(H3)[p,\alpha]$ Il existe une suite d'entiers positifs m_n convergeant vers l'infini, telle que $m_n = O(n^\alpha)$, $\eta_2 > 0$ et $\eta_4 > 0$, tels que pour tout $\delta > 0$, pour n assez grand,

$$\sum_{i=m_n+1}^{n} \frac{1}{(\lambda_p^n - \lambda_i^n)^2} \leq n^{2-\eta_2}, \qquad \frac{1}{n}\sum_{i=m_n+1}^{n} \frac{1}{\lambda_p^n - \lambda_i^n} \geq \frac{1}{\underline{\theta}} - \delta.$$

$$\text{et} \sum_{i=m_n+1}^{n} \frac{1}{(\lambda_p^n - \lambda_i^n)^4} \leq n^{4-\eta_4}$$

Le fait que les valeurs propres de la matrice non perturbée près soient suffisamment étalées pour que l'hypothèse ci-dessus soit vérifiée permet que les valeurs propres de la matrice perturbée soient très proches d'elles, comme l'indique le résultat suivant

Théorème 1.2.4. *Soit I_b l'ensemble des indices correspondant aux valeurs propres $\widetilde{\lambda}_i^n$ convergeant vers le bord droit du support de μ_X. On suppose que $(H1) et (H3)$[r, α] sont vérifiées. Alors, pour tout $\alpha' > \alpha$, on a, pour tout $i \in I_b$,*

$$\min_{1\leq k\leq i+r-r_0} |\widetilde{\lambda}_i^n - \lambda_k^n| \leq n^{-1+\alpha'}, \tag{1.2.2}$$

avec une probabilité exponentiellement proche de un.

De plus, lorsque la perturbation est de rang un, on peut décrire plus précisément dans le voisinage de quelle valeur propre de X_n se trouve chacune des valeurs propres du modèle perturbé.

Avant de présenter une application à des ensembles classiques de matrices, nous donnons quelques éléments de preuve.

1.2.7 Les grandes lignes de la preuve

Le point de départ est un calcul de déterminant : on rappelle que si V_n est une matrice $n \times r$ dont les vecteurs colonnes sont $(G_1^n, \dots G_r^n)$ dans le modèle iid ou $(U_1^n, \dots U_r^n)$ dans le modèle orthonormal alors $R_n = V_n \Theta V_n^*$, avec $\Theta = diag(\theta_1, \dots, \theta_r)$. Si z n'est pas une valeur propre de X_n, on a

$$\det(z - \widetilde{X_n}) = \det(z - X_n - V_n \Theta V_n^*) = \det(z - X_n) \det \Theta \det(\Theta^{-1} - V_n^*(z - X_n)^{-1} V_n).$$

Donc les valeurs propres de $\widetilde{X_n}$ qui ne sont pas des valeurs propres de X_n satisfont

$$f_n(z) := \det(\Theta^{-1} - V_n^*(z - X_n)^{-1} V_n) = 0.$$

Le fait que f_n soit le déterminant d'une matrice $r \times r$, donc de taille fixée, va beaucoup nous faciliter la tâche.

A partir de là, les limites presque sûres sont faciles à déterminer : par des arguments de concentration, on peut montrer que

$$(\langle V_i^n, (z - X_n)^{-1} V_j^n \rangle)_{1 \le i,j \le r} \simeq \mathrm{diag}(G_{\mu_X}(z), \dots, G_{\mu_X}(z)),$$

de sorte que les limites possibles sont les solutions de $G_{\mu_X}(z) = \theta_i^{-1}$, pour $i \in \{1, \dots r\}$, où G_{μ_X} a été défini en (1.2.1).

L'analyse des fluctuations hors du bulk consiste à regarder les asymptotiques fines de $\langle V_i^n, (\rho_n - X_n)^{-1} V_j^n \rangle$ autour de sa limite $G_{\mu_X}(\rho_\alpha) = \frac{1}{\alpha}$ pour $\rho_n = \rho_\alpha + \frac{x}{\sqrt{n}}$. Le résultat provient essentiellement d'une analyse précise du procédé d'orthonormalisation et de l'utilisation d'un théorème central limite pour les formes quadratiques développé dans [BY08].

Ensuite, vient l'analyse, plus délicate, des valeurs propres qui collent au bord. La difficulté vient du fait qu'il est difficile de faire la différence entre les valeurs propres de X_n et celles de $\widetilde{X_n}$ qui ne sont pas bien séparées. Le travail consiste à vérifier que $f_n(z)$ ne peut s'annuler que si z est très proche d'une valeur propre de X_n.

1.2.8 Applications aux ensembles classiques de matrices

Si l'on compare avec le modèle de Baik, Ben Arous et Péché ou les modèles présentés dans la partie 1.2.2 le modèle déterministe + rang fini avec vecteurs propres délocalisés peut paraître décevant. En fait, on peut facilement étendre nos résultats à des modèles pour lesquels la partie non perturbée est aléatoire et donc généraliser certains résultats évoqués dans la partie 1.2.2.

Si $(X_n)_{n\ge0}$ est une suite de matrices aléatoires, on dit qu'elle satisfait l'hypothèse (H) en probabilité si la probabilité que $(X_n)_{n\ge0}$ satisfasse (H) tend vers un quand n tend vers l'infini.

Pour étendre nos théorèmes, nous allons utiliser le résultat suivant, qui est facile à montrer :

Théorème 1.2.5. *Soit $(X_n)_{n\ge0}$ une suite de matrices indépendante des vecteurs colonnes G_i^n ou U_i^n de la perturbation*

1. *Si $(H1)$ est satisfaite en probabilité, les conclusions du Théeorème 1.2.1 s'appliquent.*
2. *Si $\kappa_4(\nu) = 0$ et les hypothèses $(H1)$ et $(H2)$ sont satisfaites en probabilité, les conclusions du Théorème 1.2.2 s'appliquent.*
3. *Si aucun des θ_i's n'est critique et les hypothèses $(H1)$ et $(H3)$ sont satisfaites en probabilité, le Théorème 1.2.4 est satisfait avec une probabilité tendant vers un au lieu d'une probabilité exponentiellement proche de un.*

Ce résultat permet d'étudier les perturbations de beaucoup de modèles classiques. Pour chacun de ces modèles, il suffira de vérifier que les différentes hypothèses sont satisfaites en probabilité. On peut ainsi généraliser certains des résultats de [Péc06, FP07, CDMF09]. Détaillons certaines de ces généralisations.

Soit X_n l'un des modèles suivants :

a. X_n est une matrice hermitienne dont les entrées au-dessus de la diagonale sont iid de moyenne nulle, de variance un et de quatrième moment fini

b. X_n est une matrice de Wishart de la forme $X_n = \frac{G_n G_n^*}{m}$, avec G_n une matrice de taille $n \times m$ dont les entrées sont iid de moyenne nulle, de variance un et de quatrième moment fini, avec $n/m \to c \in (0, 1)$.

Alors, dans les deux cas, pour le modèle déformé $\widetilde{X}_n$, on a que

- la convergence presque sûre des valeurs propres extrêmes est donnée par le Théorème 1.2.1
- les valeurs propres qui convergent hors du bulk ont des fluctuations gaussiennes (cf Théorème 1.2.2, où c_j peut être calculée explicitement)
- si de plus les entrées de X_n dans le cas a. et de G_n dans le cas b. ont des queues sous-exponentielles, alors
  - pour une perturbation de rang un : si la perturbation est sur-critique, la plus grande valeur propre converge hors du bulk et a des fluctuations gaussiennes, puis la $p$ième valeur propre suit la loi de Tracy-Widom d'ordre $p-1$; si la perturbation est sous-critique, la $p$ième valeur propre suit la loi de Tracy-Widom d'ordre p.
  - pour une perturbation de rang plus grand que un : les valeurs propres de $\widetilde{X_n}$ qui convergent au bord restent à une distance négligeable par rapport à $n^{-2/3}$ des valeurs propres extrêmes de X_n.

On a aussi le même type de résultats pour une perturbation d'un modèle unitairement invariant [6], c'est-à-dire lorsque la loi jointe des valeurs propres du modèle non perturbé est de la forme

$$dP_n(\lambda_1, \ldots, \lambda_n) = \frac{1}{Z_n} |\Delta(\lambda)|^\beta e^{-n\beta \sum_{i=1}^n V(\lambda_i)} \prod_{i=1}^n d\lambda_i, \tag{1.2.3}$$

avec V un potentiel polynomial strictement convexe et $\beta = 1$ ou 2.

Résumé de la partie 1.2 :

Dans [A1], nous étudions les fluctuations des valeurs propres extrêmes pour des modèles donnés comme somme d'une matrice déterministe vérifiant $(H1)$ ou d'une matrice aléatoire vérifiant $(H1)$ en probabilité et d'une perturbation dont les vecteurs propres sont délocalisés. On montre que ces modèles font apparaître une transition BBP : dans le régime sur-critique, certaines valeurs propres se décollent du bord du spectre et ont des fluctuations de type gaussien ; dans le régime sous-critique, les valeurs propres extrêmes convergent vers le bord du spectre. De plus, si les valeurs propres de la matrice non perturbée sont suffisamment espacées, celles du modèle perturbé collent à elles également dans l'échelle des fluctuations.

1.3 Grandes déviations pour certains modèles matriciels déformés

1.3.1 Introduction

Comme le spectre d'une matrice est fonction très compliquée de ses entrées, les théorèmes usuels de grandes déviations, en grande partie liés à la notion d'indépendance, sont difficilement applicables. Il existe

6. On verra réapparaître ces modèles dans la partie 3.3.

peu de travaux de grandes déviations pour les matrices aléatoires. La première avancée majeure dans cette direction a été faite par Ben Arous et Guionnet dans [BAG97], où ils montrent un principe de grandes déviations (PGD) pour la mesure spectrale empirique d'une matrice du GUE ou plus généralement pour les modèles unitairement invariants introduits plus haut (cf aussi [AGZ10]). Guionnet et Zeitouni [GZ02] ont aussi obtenu un PGD pour la somme d'une matrice du GUE et d'une matrice hermitienne déterministe. Nous pouvons aussi mentionner, dans un contexte un peu différent, le travail de Chatterjee et Varadhan [CV].

On en sait un peu plus sur les grandes déviations pour les valeurs propres extrêmes. Le premier résultat dans ce sens concerne la plus grande valeur propre d'une matrice du GUE et a été montré par Ben Arous, Dembo et Guionnet [BADG01] (cf aussi [AGZ10] pour des généralisations). Les matrices de Wishart de la forme XX^*, avec X une matrice $n \times m$ avec des entrées iid non nécessairement gaussiennes ont été étudiés dans [FvdHK08] lorsque le rapport m/n tend vers zéro.

Pour ce qui concerne les modèles déformés, nous avons obtenu le premier résultat de ce type dans [A8], pour une perturbation de rang un d'une matrice du GUE. Ce travail sera brièvement présenté dans la partie 1.3.2, qui évoquera aussi un de nos travaux précédents [A6], concernant les asymptotiques de l'intégrale sphérique, sur lequel il s'appuie. Nous verrons que nous obtenons dans ce cas un PGD pour la plus grande valeur propre avec une fonction de taux complètement explicite. Ceci nous a permis dans [A3] d'analyser les performances d'un test statistique utilisé pour l'étude coopérative de spectre en communications numériques.

Nous avons ensuite cherché à généraliser ce résultat au cas où la perturbation est de rang supérieur. Le but de [A9] était d'étudier l'intégrale sphérique en rang supérieur. Nous obtenons des résultats partiels dans ce sens, que nous ne détaillerons pas dans ce mémoire mais la limite ayant une forme difficile à analyser, cela ne nous a pas permis d'en tirer un PGD pour la plus grande valeur du modèle GUE plus rang fini.

Nous avons ensuite réaborder dans [A2] le problème sous un angle un peu différent, s'inspirant de ce qui a été expliqué dans la partie 1.2.7 pour obtenir un PGD pour les valeurs propres extrêmes dans des modèles déformés similaires à ceux introduits dans la partie 1.2.3.

Avant de rentrer dans le vif du sujet, nous rappelons, pour le lecteur qui ne serait pas familier de grandes déviations, la définition précise d'un principe de grandes déviations.

Si $(E, \mathcal{E})$ est un espace métrique muni d'une topologie qui contient la tribu borélienne, si $(\mu_\varepsilon)_{\varepsilon \geq 0}$ une famille de mesures de probabilités sur E et I une fonction de taux, c'est-à-dire une application $I : E \to [0, \infty]$, semi-continue inférieurement, on dit que la famille $(\mu_\varepsilon)_{\varepsilon \geq 0}$ satisfait un PGD de fonction de taux I dans l'échelle $\frac{1}{\varepsilon}$ si

$$\forall G \text{ ouvert, } \liminf_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(G) \geq - \inf_G I,$$

$$\forall F \text{ fermé, } \limsup_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(F) \leq - \inf_F I.$$

Pour la plupart des PGD qui apparaîtront dans ce mémoire, on aura $E = \mathbb{R}^d$ et $\varepsilon = 1/N$.

Les fonctions de taux impliquées seront souvent bonnes, dans le sens où leurs ensembles de niveaux seront non seulement fermés (elles sont semi-continues inférieurement) mais aussi compacts.

1.3.2 Grandes déviations pour une déformation de rang un du GOE/GUE et intégrale sphérique

Dans cette section, nous présentons brièvement les résultats de [A8] et [A6].

L'article [A8] représente en effet une première approche du problème des déviations pour les valeurs propres extrêmes d'un modèle déformé. Même s'il concerne un modèle particulier (GOE/GUE avec une déformation de rang un), il a essentiellement deux mérites : constituer le premier résultat dans cette direction et donner une fonction de taux complètement explicite.

On considère donc le modèle suivant : W_n est une matrice du GUE [7] et A_n une matrice hermitienne déterministe de rang un, dont l'unique valeur propre non nulle est notée θ. On étudie le modèle déformé $X_n = W_n + A_n$.

Notre stratégie est de s'appuyer sur le fait que dans ce modèle, on connaît explicitement la loi jointe des valeurs propres. En effet, celle-ci est donnée par :

$$\mathbb{Q}_n^\theta(dx_1, \ldots, dx_n) = \frac{1}{Z_n^\theta} \prod_{i<j} |x_i - x_j|^2 I_n(\theta, X_n) e^{-\frac{n}{2}\sum_{i=1}^N x_i^2} dx_1 \ldots dx_n, \tag{1.3.1}$$

où I_n est l'intégrale sphérique définie par

$$I_n(X_n, \theta) := \int e^{n\,\mathrm{tr}(UX_nU^*A_n)} dU = \int e^{n\theta(UX_nU^*)_{11}} dU,$$

où dU est la mesure de Haar [8] sur le groupe unitaire $\mathcal{U}(N)$.

Le fait que la loi jointe des valeurs propres de X_n ne dépende de la déformation A_n qu'à travers θ est dû à l'invariance de la loi de W_n par conjugaison unitaire.

Dans [A6], issu de ma thèse de doctorat, nous étudions en détail le comportement de l'intégrale sphérique I_n. Par des techniques de concentration gaussienne et également de grandes déviations, nous avons déterminé une expression explicite de la limite dans l'échelle e^n de cette intégrale sphérique.

Théorème 1.3.1. *Soit $(B_n)_{n\in\mathbb{N}}$ une suite de matrice dont le mesure spectrale empirique converge vers une mesure μ à support compact tandis que les plus petite et plus grande valeurs propres de B_n convergent respectivement vers $\lambda_{\min}$ et $\lambda_{\max}$. Soient $H_{\min} := \lim_{z\uparrow\lambda_{\min}} \int \frac{d\mu(t)}{z-t}$ et $H_{\max} := \lim_{z\downarrow\lambda_{\max}} \int \frac{d\mu(t)}{z-t}$. Alors*

$$I_\mu(\lambda_{\min}, \lambda_{\max}, \theta) := \lim_{n\to\infty} \frac{1}{n} \log I_n(\theta, B_n)$$

existe et est donnée par

$$I_\mu(\lambda_{\min}, \lambda_{\max}, \theta) = \theta v(\theta) - \int \log(1 + \theta v(\theta) - \theta\lambda) d\mu(\lambda),$$

avec

$$v(\theta) = \begin{cases} R_\mu(\theta), & \text{si } H_{\min} \le \theta \le H_{\max}, \\ \lambda_{\max} - \frac{1}{\theta}, & \text{si } \theta > H_{\max}, \\ \lambda_{\min} - \frac{1}{\theta}, & \text{si } \theta < H_{\min}, \end{cases}$$

avec R_μ la R-transformée [9] de la mesure μ.

En réalité, si $\theta \geq 0$, la quantité $I_\mu(\lambda_{\min}, \lambda_{\max}, \theta)$ ne dépend que de θ et $\lambda_{\max}$, elle est donc notée $I_\mu(\lambda_{\max}, \theta)$; si $\theta \leq 0$, la quantité $I_\mu(\lambda_{\min}, \lambda_{\max}, \theta)$ ne dépend que de θ et $\lambda_{\min}$, elle est donc notée $I_\mu(\lambda_{\min}, \theta)$.

---

7. On peut aussi considérer une matrcie de l'Ensemble orthogonal Gaussien (GOE) mais pour simplifier l'écriture, nous choisissons de ne présenter que le cas unitaire. La fonction de taux est modifiée de manière simple dans le cas orthogonal.

8. Nous reviendrons sur la notion de mesure de Haar dans le chapitre 2 mais on peut retenir pour l'instant qu'il s'agit de la notion de mesure uniforme adaptée au cas d'un groupe compact.

9. On renvoie le lecteur à l'appendice sur les probabilités libres pour la définition précise de la R-transformée.

L'apparition de la R-transformée dans cette limite a des conséquences sur les liens entre matrices aléatoires et probabilités libres : l'intégrale sphérique peut être considérée comme un modèle matriciel de la R-transformée, qui est une fonctionnelle très importante en probabilités libres. Dans [A6], nous utilisons le Théorème 1.3.1 pour donner une preuve differnte des preuves classiques de l'additivité de la R-transformée par convolution libre. Ces aspects sont largement soulignés dans l'introduction de ma thèse de doctorat [T] et nous n'y revenons pas en détail ici. Nous nous intéressons plutôt à l'utilisation du Théorème 1.3.1 pour l'étude du modèle déformé GUE plus rang un.

Quand on s'intéresse au problème, on voit très vite qu'en réalité, pour obtenir le PGD recherché, nous avons besoin de combiner le Théorème 1.3.1 et un résultat de continuité de l'intégrale sphérique que nous allons énoncer maintenant.

Au vu de la forme de la limite de l'intégrale sphérique et notamment sa dépendance en $\lambda_{\min}$ et $\lambda_{\max}$, on ne peut pas espérer directement une continuité en la mesure spectrale empirique de la matrice de rang plein. Cependant, si on localise aussi la plus grande (ou la plus petite) valeur propre, on obtient la continuité voulue. Plus précisément, si on note d une distance sur l'ensemble des probabilités sur $\mathbb{R}$ compatible avec la topologie de la convergence faible

Proposition 1.3.2. *Pour tout $\theta > 0$ et $\kappa > 0$, il existe une fonction $g_\kappa : \mathbb{R}_+ \to \mathbb{R}_+$ qui converge vers zéro en zéro telle que pour tout $\delta > 0$ et n assez grand, si B_n et B'_n sont deux suites de matrices diagonales telle que leurs mesures spectrales empiriques $\hat{\mu}_{B_n}$ et $\hat{\mu}_{B'_n}$ vérifient $d(\hat{\mu}_{B_n}, \hat{\mu}_{B'_n}) \leq n^{-\kappa}$ et, si $\lambda_1(B_n)$ et $\lambda_1(B'_n)$ sont respectivement les plus grandes valeurs propres de B_n et B'_n, $|\lambda_1(B_n) - \lambda_1(B'_n)| \leq \delta$, avec $\sup \|B_n\|_\infty < \infty$ alors*

$$\left| \frac{1}{n} \log I_n(\theta, B_n) - \frac{1}{n} \log I_n(\theta, B'_n) \right| \leq g_\kappa(\delta).$$

En combinant le théorème 1.3.1, la proposition 1.3.2 et en utilisant le même cheminement de preuve, détaillé ci-dessous, que dans [BADG01], on obtient le résultat suivant

Théorème 1.3.3. *Si $\theta \geq 0$, alors sous $\mathbb{Q}_n^\theta$, la plus grande valeur propre $x_n^* = \max\{x_1, \ldots, x_n\}$ satisfait un principe de grandes déviations dans l'échelle n avec une bonne fonction de taux K^θ telle que*

$$K^\theta(x) := \begin{cases} +\infty, & \text{si } x < 2, \\ -1 - 2\int \log|x-y| d\sigma(y) + \frac{1}{2}x^2 - I_\sigma(x,\theta), & \text{sinon,} \end{cases}$$

avec σ la loi semi-circulaire de densité $\frac{1}{2\pi}\sqrt{4-x^2}\mathbf{1}_{[-2,2]}(x)$.

On peut même calculer explicitement la famille de fonctions K^θ pour les différentes valeurs de θ.

Nous donnons maintenant une rapide esquisse de preuve qui permet de comprendre la forme de la fonction de taux K^θ.

Comme souvent dans les démonstrations de grandes déviations, on commence par montrer la tension exponentielle. On est ensuite ramené à borner des quantités de la forme

$$\mathbb{Q}_n^\theta(x_n^* \in [x, x+\delta], \max_{i=1,\ldots,n} |x_i| \leq M),$$

pour $x \geq 2, \delta \geq 0$ et M assez grand.
Si on pose $\hat{\pi}_n := \frac{1}{n-1}\sum_{i=1}^n \delta_{x_i}$ et $\Phi(x,\mu) = 2\int \log|x-y|d\mu(y) - \frac{1}{2}x^2$, on a

$$\mathbb{Q}_n^\theta(x_n^* \in [x, x+\delta], \max|x_i| \leq M) \\ \leq C_n \int_x^{x+\delta} dx_1 \int_{[-M,M]^{n-1}} e^{(n-1)\Phi(x_1,\hat{\pi}_n)}.I_n(\theta, X_n) d\mathbb{Q}_{n-1}^0(x_2, \ldots, x_n),$$

où C_n est une constante de normalisation.
On utilise ensuite le fait que, d'après [BAG97], les déviations de la mesure empirique $\hat{\pi}_n$ sont dans l'échelle n^2. Comme on regarde les déviations de la plus grande valeur propre dans l'échelle n, la mesure $\hat{\pi}_n$ est bien concentrée autour de la loi semicirculaire σ. Le terme de gauche ci-dessus peut donc être approché par la quantité

$$C_n \int_x^{x+\delta} dx_1 \int_{[-M,M]^{n-1}, \hat{\pi}_n \simeq \sigma} e^{(n-1)\Phi(x_1,\hat{\pi}_n)}.e^{nI_\sigma(x,\hat{\pi}_n)} d\mathbb{Q}^0_{n-1}(x_2,\ldots,x_n),$$

ce qui explique la forme de la fonction de taux K^θ.

1.3.3 Applications aux communications numériques

Nous présentons dans cette partie une application d'un résultat proche de ce que nous avons vu dans le paragraphe précédent.

On se pose le problème suivant, dit d'écoute coopérative du spectre : un réseau sans fil primaire occupe le spectre hertzien en laissant des bandes libres, un système secondaire souhaite utiliser ces bandes inoccupées. Les utilisateurs secondaires partagent l'information pour détecter les bandes libres. On a n capteurs secondaires dont chacun observe m échantillons et on souhaite tester la présence ou l'absence de la source, c'est-à-dire l'hypothèse
H_0 : absence de signal. Chaque capteur secondaire reçoit une série $g_k(\ell) = w_k(\ell)$ pour ℓ de 1 à m où $w_k(\ell)$ est une gaussienne complexe centrée de variance inconnue σ^2
contre H_1 : présence d'un signal. On a $g_k(\ell) = h_k s(\ell) + w_k(\ell)$, où $s(\ell)$ est un signal gausssien primaire (indépendant des $w_k(\ell)$) et h_k le gain du canal entre la source et le capteur numéro k.

On rassemble les observations dans une matrice de taille $n \times m$, $G_n = (g_k(\ell))_{k=1,\ldots,n;\ell=1,\ldots,m}$.

Une idée naturelle pour tester H_0 contre H_1 est d'utiliser le test de vraisemblance de Neyman-Pearson.

Sous H_0, les éléments de G_n sont iid gaussiennes complexes centrées de variance inconnue σ^2 et la vraisemblance s'écrit

$$p_0(G_n,\sigma^2) = \frac{1}{(\pi\sigma^2)^{nm}} \exp\left(-\frac{m}{\sigma^2}\operatorname{Tr} R,\right),$$

où $R = \frac{1}{m} G_n G_n^*$ est la matrice de covariance empirique et Tr la trace usuelle.

Sous H_1, les colonnes de G_n sont des vecteurs gaussiens centrés de matrice de covariance $hh^* + \sigma^2 \mathbf{I}_n$, où $h := [h_1,\ldots,h_n]^t$ représente le gain du canal et $\mathbf{I}_n$ est l'identité de taille $n\times n$. La vraisemblance s'écrit

$$p_1(G_n,\sigma^2,h) = \frac{1}{(\pi^n \det(hh^* + \sigma^2 \mathbf{I}_n))^m} \exp\left(-m \operatorname{Tr}(hh^* + \sigma^2\mathbf{I}_n)^{-1},\right).$$

Comme σ^2 et h sont inconnus, on utilise plutôt le rapport de vraisemblance généralisé

$$\frac{\sup_{\sigma^2,h} p_1(G_n,\sigma^2,h)}{\sup_{\sigma^2} p_0(G_n,\sigma^2)}.$$

Tous calculs faits, cela revient à rejeter H_0 lorsque la statistique $T_n := \frac{\Lambda}{\frac{1}{n}\operatorname{Tr} R}$ est grande, où Λ est la plus grande valeur propre de la matrice R.

La mise en place du test exact est assez compliquée mais on peut faire une analyse précise du test asymptotique, lorsque n et m tendent vers l'infini de sorte que le rapport $\frac{n}{m}$ tende vers une constante c comprise entre 0 et 1, c'est-à-dire lorsque le nombre de capteurs et la taille de la fenêtre d'observation sont du même ordre de grandeur.

On est alors dans le même cadre que dans le paragraphe 1.2.1 : sous H_0, R appartient au LUE et sous H_1, on a une déformation de rang un du LUE. Les résultats suivants sont maintenant bien connus.

Toutes les convergences évoquées ci-dessous ont lieu dans le régime où n et m tendent vers l'infini de sorte que le rapport $\frac{n}{m}$ tende vers une constante c comprise entre 0 et 1. Sous H_0, Λ converge vers le bord droit du support de la loi de Marcenko-Pastur, soit $\sigma^2(1+\sqrt{c})^2$. Par la loi des grands nombres, $\frac{1}{n}\operatorname{Tr} R$ converge vers σ^2 de sorte que T_n converge vers $\lambda^+ := (1+\sqrt{c})^2$.

De plus, sous H_0, T_n a des fluctuations régies par la loi de Tracy-Widom, ce qui permet de calculer le seuil asymptotique du test associé à un niveau donné.

Sous H_1, si la déformation est suffisamment forte, c'est-à-dire ici si le rapport $\rho := \frac{\sum h_i^2}{\sigma^2}$ est assez grand, Λ converge hors du bulk vers $\sigma^2(1+\rho)\left(1+\frac{c}{\rho}\right) := \lambda_{\text{spk}} > \lambda^+$.

On a donc que sous H_0, l'événement $\{T_n > \gamma\}$ est un événement rare pour $\gamma > \lambda^+$ tandis que sous H_1, l'événement $\{T_n < \gamma\}$ est un événement rare pour $\gamma < \lambda_{\text{spk}}$. Dans [A3], on montre le résultat suivant :

Théorème 1.3.4. *Sous H_0 (resp. sous H_1), T_n vérifie un PGD de bonne fonction de taux $\mathcal{E}_0$ (resp. $\mathcal{E}_1$), explicite. On a donc, pour tout $\lambda^+ < \gamma < \lambda_{\text{spk}}$,*

$$\lim_{\frac{n}{m}\to c}\left(\frac{1}{m}\log P_{H_0}(T_n > \gamma), \frac{1}{m}\log P_{H_1}(T_n < \gamma)\right) = -(\mathcal{E}_0(\gamma), \mathcal{E}_1(\gamma)).$$

La courbe $(\mathcal{E}_0(\gamma), \mathcal{E}_1(\gamma))$ est applelée courbe ROC (Receiver Operating Characteristic) et elle est un outil classique dans ce domaine pour évaluer les performances d'un test.

Les courbes ROC permettent de comparer l'efficacité de plusieurs tests et cela nous a notamment permis de montrer que le test que nous venons de présenter, basé sur le rapport de vraisemblance généralisé a de meilleurs performances qu'un test couramment utilisé dans le domaine qui consiste à utiliser comme statistique de test le rapport des plus grande et plus petite valeurs propres de la matrice de covariance empirique R.

En ce qui concerne la preuve du théorème ci-dessus, il s'agit d'une adaptation du résultat de [A8] au cas du LUE. On se base là encore sur l'analyse de l'expression explicite de la loi jointe des valeurs propres, analyse possible grâce à notre connaissance des asymptotiques de l'intégrale sphérique.

1.3.4 Grandes déviations pour la plus grande valeur propre du mouvement brownien hermitien

Lors de l'un de mes exposés sur les résultats de [A8] (présentés dans la partie 1.3.2), Marc Yor m'a demandé si j'avais une idée sur l'équivalent dynamique de ce résultat, c'est-à-dire sur les déviations du processus de la plus grande valeur propre d'un mouvement brownien hermitien ou d'une déformation de rang un de celui-ci. Dans cette partie, nous présentons les résultats de l'article [A4], écrit en collaboration avec Catherine Donati-Martin, qui répond à cette question, effectivement assez naturelle.

Avant d'énoncer le résultat obtenu, il nous faut définir le mouvement brownien hermitien, qui est une version dynamique du GUE. Il a été introduit en 1962 par Dyson.

Soient $(\beta_{ij}, \beta'_{ij})_{1\le i\le j\le n}$ des mouvements browniens réels standards iid. Le mouvement brownien hermitien $(H_n(t))_{t\ge 0}$ est le processus aléatoire à valeurs dans l'espace $\mathcal{H}_n$ des matrices hermitiennes de taille n, dont les entrées $(H_n)_{kl}$ sont données, pour $k \le l$ par

$$(H_n)_{kl} = \begin{cases} \frac{1}{\sqrt{2n}}(\beta_{kl} + i\beta'_{kl}), & \text{si } k < l, \\ \frac{1}{\sqrt{n}}\beta_{kk}, & \text{si } k = l. \end{cases}$$

Dyson [Dei99] a montré que les valeurs propres de $(H_n(t))_{t\geq 0}$ forment un gaz de Coulomb, évoluant selon des mouvements browniens soumis à leur répulsion électrostatique mutuelle. Plus précisément, elles satisfont le système d'équations différentielles stochastiques (EDS)

$$d\lambda_i(t) = \frac{1}{\sqrt{n}} dB_i(t) + \frac{1}{N}\sum_{j\neq i}\frac{1}{\lambda_i(t)-\lambda_j(t)}dt,\ t\geq 0,\ i,j=1,\dots,n \tag{1.3.2}$$

où B_i sont des mouvements browniens réels indépendants.

On peut aussi définir une version dynamique du modèle GUE plus rang un présenté dans la partie 1.3.2 en considérant $H_n^\theta(t) = H_n(t) + H_n^\theta(0)$ le mouvement brownien hermitien issu de $H_n^\theta(0) := \mathrm{diag}(\theta, 0, \dots, 0)$, avec $\theta \geq 0$ et on note $\lambda_1^{\theta,n}(t) \geq \lambda_2^{\theta,n}(t) \geq \dots \geq \lambda_n^{\theta,n}(t)$ les valeurs propres de $H_n^\theta(t)$ ordonnées par ordre croissant. On s'intéresse à sa plus grande valeur propre $(\lambda_1^{\theta,n}(t))_{t\geq 0}$ et on obtient dans [A4] le résultat suivant :

Théorème 1.3.5. *La loi de $(\lambda_1^{\theta,n}(t))_{0\leq t\leq 1}$ satisfait un PGD dans l'échelle n, sur l'ensemble $C_\theta([0,1];\mathbb{R})$ des fonctions continues sur $[0,1]$ valant θ en 0 muni de la topologie de la convergence uniforme, de bonne fonction de taux*

$$I_\theta(\varphi) = \begin{cases} \frac{1}{2}\int_0^1 \left(\dot\varphi(s) - \frac{1}{2s}\left(\varphi(s) - \sqrt{\varphi(s)^2 - 4s}\right)\right)^2 ds, \\ \textit{si } \varphi \textit{ est absolument continue et } \varphi(t) \geq 2\sqrt{t}\ \forall t\in[0,1], \\ +\infty,\ \textit{sinon.}\end{cases} \tag{1.3.3}$$

Par principe de contraction, on peut retrouver à partir de la version processus les PGD à temps fixe. Tout calcul fait, on retrouve le Théorème 1.3.3.

Donnons pour finir une idée de la preuve de cette version dynamique. Dans [BADG01] et [A8], le fait que les déviations de la mesure spectrale empirique des matrices considérées ont lieu dans l'échelle n^2 alors que celle de la plus grande valeur propre ont lieu dans l'échelle n jouait déjà un rôle important. Il en est de même dans la preuve de ce résultat ; dans l'échelle où on regarde la plus grande valeur propre, la mesure empirique de toutes les valeurs propres sauf la première est déjà bien concentrée autour du processus semi-circulaire donné par

$$d\sigma_t(x) = \frac{1}{2\pi t}\mathbf{1}_{[-2\sqrt{t},2\sqrt{t}]}\sqrt{4t-x^2}dx \quad \text{et} \quad \sigma_0 = \delta_0. \tag{1.3.4}$$

Cependant, alors que dans le cas statique, on s'appuyait sur l'expression explicite de la loi jointe des valeurs propres, dans le cas dynamique, on utilise du calcul stochastique en s'appuyant sur le fait que le processus des valeurs propres vérifie le système d'EDS (1.3.2). Grosso modo, la plus grande valeur propre est solution d'une EDS de la forme

$$d\lambda_1^{\theta,n}(t) = \frac{1}{\sqrt{n}}dB(t) + b(\lambda_1^{\theta,n}(t), (\widehat{\nu_n})_t)dt,$$

avec B un mouvement brownien réel standard, $\widehat{\nu_n} := \frac{1}{n-1}\sum_{i=2}^n \delta_{\lambda_i^{\theta,n}(t)}$ la mesure empirique de toutes les valeurs propres sauf la première et la dérive $b(x,\nu) = \int_{-\infty}^x \frac{d\nu(t)}{x-t}$.

Dans l'échelle qui nous intéresse, $\widehat{\nu_n}$ est proche de σ et la fonction de taux I_θ est celle prédite par la théorie de Freidlin-Wentzell (cf [DZ10, Th.5.6.3]) pour l'EDS

$$d\lambda_1(t) = \frac{1}{\sqrt{n}}dB(t) + b(\lambda_1(t), \sigma_t)dt.$$

L'une des principales difficultés techniques est de traiter la singularité du drift b car pour $x \in \mathbb{R}$, $\nu \mapsto b(x, \nu)$ n'est pas continue pour la topologie de la convergence faible des mesures.

1.3.5 Grandes déviations pour des déformations de rang supérieur

Jusqu'ici nous avons exposé divers résultats et leurs applications qui tous concernaient des déformations de rang un. Le vrai enjeu était de les généraliser en rang supérieur. Ceci a été fait dans [A2], qui peut-être considéré comme un papier jumeau de [A1] et considère les mêmes modèles que ceux introduits dans la partie 1.2.3. La perturbation est à peu près de la forme, sauf qu'au lieu de supposer que les vecteurs $\sqrt{n}G_i^n$ ont des entrées iid de loi ν satisfaisant une inégalité de Sobolev logarithmique, on prend $G = (g_1, \ldots, g_r)$ un vecteur aléatoire satisfaisant $E\left(e^{\alpha \sum |g_i|^2}\right) < \infty$ pour un $\alpha > 0$ et on construit $(\sqrt{n}G_1^n, \ldots, \sqrt{n}G_r^n)$ des vecteurs aléatoires dont les entrées sont des copies indépendantes de G ; $(U_1^n, \ldots, U_r^n)$ s'obtiennent là encore à partir de $(\sqrt{n}G_1^n, \ldots, \sqrt{n}G_r^n)$ par un procédé d'orthonormalisation de Gram-Schmidt. On obtient :

Théorème 1.3.6. *Si X_n satisfait $(H1)$, la loi des r_0 plus grandes valeurs propres de $\widetilde{X_n}$ satisfait un PGD dans l'échelle n de bonne fonction de taux L. Celle-ci a un unique minimiseur vers lequel il y a convergence presque sûre.*

En particulier, dans le cas le plus simple, pour le modèle iid quand $X_n = 0$, on est ramenés au modèle suivant : si G_n est une matrice $n \times r$ dont les lignes sont des copies iid de G et $\Theta = diag(\theta_1, \ldots, \theta_r)$, on étudie les déviations de $W_n = \frac{1}{n} G_n^* \Theta G_n$ (cf aussi Fey, van der Hofstad et Klok [FvdHK08] qui traitent le cas où Θ est l'identité).

Evoquons brièvement la stratégie de la preuve du Théorème 1.3.6. Le point de départ est le même que pour la convergence presque sûre et les fluctuations, dont la preuve est évoquée dans la partie 1.2.7 et nous conservons les mêmes notations.

On utilise le fait que les valeurs propres sont les solutions de $f_n(z) = 0$. Si X_n est diagonale, $f_n(z)$ est une fonction polynomiale de

$$\langle G_i^n, (z - X_n)^{-1} G_j^n \rangle = \frac{1}{n} \sum_{i=1}^{n} \frac{g_i(k)\overline{g_j(k)}}{z - \lambda_k^n}$$

et

$$\langle G_i^n, G_j^n \rangle = \frac{1}{n} \sum_{i=1}^{n} g_i(k)\overline{g_j(k)}.$$

Par le théorème de Cramer (ou Cramer pondéré) et des arguments assez standards de la théorie des grandes déviations, on peut montrer un PGD pour ces deux sommes et donc pour f_n.

Comme les valeurs propres sont les zéros de f_n, on pourrait espérer obtenir facilement leurs déviations à partir de là. Mais en réalité, ce ne sont pas des fonctions continues de f_n dans la topologie qui nous intéresse.

Pour finir, nous voulons mentionner une généralisation importante de notre théorème qui va être cruciale dans ce qui suit : si la loi de $\frac{G}{\sqrt{n}}$ satisfait un PGD, on peut relaxer l'hypothèse $(H1)$ dans le sens où on n'a plus besoin de supposer que les valeurs propres extrêmes de X_n convergent vers le bord du support de la mesure limite μ_X. On autorise un nombre fini de valeurs propres à converger hors du support de μ_X. Dans ce cas là encore, on obtient un PGD du même type que le Théorème 1.3.6.

1.3.6 PGD pour des modèles unitairement invariants déformés

Comme dans le cas de l'étude des fluctuations, le modèle déterministe (quand la partie non perturbée est déterministe) peut sembler un peu artificiel. Le vrai enjeu est de comprendre le cas où X_n est aléatoire. Mais on ne peut guère espérer obtenir un PGD pour les valeurs propres du modèle perturbé si on ne connait même pas les déviations dans le modèle initial. Or il existe très peu de modèles pour lesquels ces déviations sont comprises.

Ici, nous considérons le cas où X_n est aléatoire, suivant un modèle unitairement invariant, comme défini par (1.2.3). On suppose par ailleurs que les U_i^n's sont une famille déterministe ou bien indépendante de X_n de vecteurs orthonormés. On a

Théorème 1.3.7. *Si X_n suit un modèle unitairement invariant de potentiel V, sous de bonnes hypothèses sur V, pour tout k fixé, la loi des k plus grandes valeurs propres de $\widetilde{X_n}$ satisfait un PGD avec une bonne fonction de taux.*

L'argument principal de la preuve est le suivant : le PGD pour les valeurs propres de X_n nous indique que seules un nombre fini de ces valeurs propres peuvent dévier ; on conditionne sur ces déviations de sorte que conditionnellement aux positions des valeurs propres de X_n, on peut appliquer la généralisation du Théorème 1.3.6 au cas où il y a un nombre fini de valeurs propres convergeant hors du support.

Résumé de la partie 1.3 :

Dans une série d'articles [A8, A3, A4, A2], nous donnons un certain nombre de PGD pour les valeurs propres extrêmes de modèles déformés. Dans [A8, A3], il s'agit d'un modèle particulier de la forme GUE plus une déformation de rang un, le second papier se concentrant sur les applications statistiques de ces résultats. [A4] étudie une version en processus de ce même modèle. Enfin [A2] étudie le même type de modèles que dans [A1], en particulier les déformations de modèles unitairement invariants.

Chapitre 2

A propos du mouvement brownien unitaire [A7]

L'objet de ce chapitre est de présenter les résultats de l'article [A7] écrit en collaboration avec Thierry Lévy.

2.1 Définition et asymptotiques du mouvement brownien unitaire

2.1.1 Définition du mouvement brownien unitaire

Dans le chapitre précédent, nous avons déjà rencontré le mouvement brownien hermitien [1] et souligné qu'il était la version dynamique du GUE.

Outre ces deux ensembles de matrices, un autre type de matrices a été très étudié (cf par exemple [Dia03] pour une revue) : l'ensemble des matrices unitaires distribuées selon la mesure de Haar. Nous détaillerons plus tard certaines de leurs propriétés mais retenons pour l'instant que cela revient à choisir une matrice unitaire de taille fixée "uniformément".

Nous allons nous intéresser dans ce chapitre à la version dynamique de ces matrices de Haar, à savoir le mouvement brownien unitaire, que nous définissons maintenant.

Soit $\mathcal{U}(N)$ le groupe des matrives unitaires de taille $N \times N$. Son algèbre de Lie, notée $\mathfrak{u}(N)$, est l'espace vectoriel des matrices anti-hermitiennes de taille $N \times N$. De manière analogue à la construction du mouvement brownien hermitien, il est commode de construire un mouvement brownien sur $\mathfrak{u}(N)$, que nous noterons $(K_N(t))_{t\geq 0}$.

En effet, si l'on munit $\mathfrak{u}(N)$ du produit scalaire suivant :

$$\forall A, B \in \mathfrak{u}(N), (A, B) = N\,\mathrm{Tr}(A^*B),$$

où Tr est la trace usuelle, alors $(K_N(t))_{t\geq 0}$ est l'unique processus gaussien centré, indexé par $\mathbb{R}^+$ et à valeurs dans $\mathfrak{u}(N)$, dont la fonction de covariance est donnée par :

$$\mathbb{E}((A, K_N(t))(B, K_N(t))) = \min(s, t)(A, B),$$

pour tous $s, t \in \mathbb{R}^+$ et $A, B \in \mathfrak{u}(N)$.

1. On peut aussi définir le mouvement brownien symétrique, version dynamique du GOE.

Notons que $(K_N(t))_{t\geq 0}$ peut aussi être défini “entrée par entrée” de la façon suivante :

$$(K_N(t))_{kl} = \begin{cases} \frac{1}{\sqrt{2N}}(\beta_{kl} + i\beta'_{kl})(t) & \text{si } k < l \\ \frac{1}{\sqrt{N}} i\beta_{kl} & \text{si } k = l \\ \frac{1}{\sqrt{2N}}(-\beta_{lk} + i\beta'_{lk})(t) & \text{si } k > l, \end{cases}$$

avec $(\beta_{kl}, k \leq l), (\beta'_{kl}, k < l)$ des mouvements browniens réels indépendants.

A partir de là, on peut définir un mouvement brownien $(U_N(t))_{t\geq 0}$ sur le groupe unitaire $\mathcal{U}(N)$ en faisant “rouler sans glisser” le plan tangent $\mathfrak{u}(N)$ autour de la variété $\mathcal{U}(N)$. Autrement dit, $(U_N(t))_{t\geq 0}$ est la solution de l'EDS linéaire suivante (au sens d'Itô) :

$$dU_N(t) = U_N(t)dK_N(t) - \frac{1}{2}U_N(t)dt. \tag{2.1.1}$$

Le premier terme correspond au roulement sans glissement évoqué plus haut et le deuxième est un facteur correctif [2] permettant de rester dans le groupe $\mathcal{U}(N)$.

En effet, la solution de l'équation (2.1.1) est a priori un processus à valeurs dans $\mathcal{M}_N(\mathbb{C})$, algèbre des matrices $N \times N$ à coefficients complexes, mais on peut aisément vérifier en appliquant la formule d'Itô que si la condition initiale $U_N(0)$ est dans $\mathcal{U}(N)$, le processus $(U_N(t))_{t\geq 0}$ tout entier y demeure. On choisira dans la suite la condition initiale $U_N(0) = I_N$, où I_N est l'identité dans $\mathcal{M}_N(\mathbb{C})$.

2.1.2 Définition de sa limite : le mouvement brownien multiplicatif libre

On s'intéresse ici au comportement limite, quand la taille N des matrices tend vers l'infini, du processus $(U_N(t))_{t\geq 0}$. Cette convergence a été étudiée par Biane dans [Bia97] et le processus limite, appelé mouvement brownien multiplicatif libre se définit dans le cadre des probabilités libres [3] comme nous allons l'expliquer ci-après.

De même que le mouvement brownien unitaire se définit à partir du mouvement brownien anti-hermitien, le mouvement brownien multiplicatif libre se définit à partir du mouvement brownien (additif) libre.

Définition 2.1.1. *Soit $(\mathcal{A}, \tau)$ un espace de probabilités non commutatif. Une collection $(x_t)_{t\geq 0}$ d'éléments autoadjoints de $\mathcal{A}$ est appelé mouvement brownien (additif) libre si elle vérifie les trois conditions suivantes :*

1. *le processus $(x_t)_{t\geq 0}$ est à accroissements libres, c'est-à-dire que pour tous $0 \leq t_1 \leq \ldots \leq t_n$, les éléments $x_{t_1}, x_{t_2} - x_{t_1}, \ldots, x_{t_n} - x_{t_{n-1}}$ sont libres entre eux.*
2. *le processus $(x_t)_{t\geq 0}$ est à accroissements stationnaires, c'est-à-dire que pour tous $0 \leq s \leq t$, l'élément $x_t - x_s$ a même loi que x_{t-s}.*
3. *Pour tout $t \geq 0$, la loi de x_t est la loi semi-circulaire σ_t, mesure de probabilités sur $\mathbb{R}$ de densité $\frac{1}{2\pi t}\sqrt{4t - x^2}\mathbf{1}_{[-2\sqrt{t}, 2\sqrt{t}]}(x)$.*

Le mouvement brownien multiplicatif libre $(u_t)_{t\geq 0}$ est alors défini comme la solution de l'EDS libre :

$$du_t = iu_t dx_t - \frac{1}{2}u_t dt. \tag{2.1.2}$$

On peut alors facilement vérifier que $(u_t)_{t\geq 0}$ est une collection d'éléments unitaires de $\mathcal{A}$ et on peut montrer que le mouvement brownien multiplicatif libre répond aussi à la définition suivante :

2. On peut aussi vérifier que $(U_N(t))_{t\geq 0}$ est solution de l'équation $dU_N(t) = U_N(t) \cdot dK_N(t)$, cette fois au sens de Stratonovich (sans facteur correctif).

3. Le lecteur non familier de probabilités libres pourra se reporter à l'appendice qui en introduit les principales notions nécessaires à la compréhension du présent mémoire.

Définition 2.1.2. *Soit $(\mathcal{A}, \tau)$ un espace de probabilités non-commutatif. Une collection $(u_t)_{t\geq 0}$ d'éléments unitaires de $\mathcal{A}$ est appelé mouvement brownien multiplicatif libre si elle vérifie les trois conditions suivantes :*

1. *le processus $(u_t)_{t\geq 0}$ est à accroissements multiplicatifs libres, c'est-à-dire que pour tous $0 \leq t_1 \leq \ldots \leq t_n$, les éléments $u_{t_1}, u_{t_2}^* u_{t_1}, \ldots, u_{t_n}^* u_{t_{n-1}}$ sont libres entre eux.*
2. *le processus $(u_t)_{t\geq 0}$ est à accroissements multiplicatifs stationnaires, c'est-à-dire que pour tous $0 \leq s \leq t$, l'élément $u_t^* u_s$ a même loi que u_{t-s}.*
3. *Pour tout $t \geq 0$, la loi de u_t est mesure de probabilités ν_t sur le cercle unité $\mathbb{U} := \{z \in \mathbb{C}/|z| = 1\}$ qui peut être caractérisée :*
   – *soit par l'identité :*
   $$\int_{\mathbb{U}} \frac{1}{1 - \frac{z}{1+z} e^{tz} e^{t/2} \xi} d\nu_t(\xi) = 1 + z,$$
   pour tout z dans un voisinage de 0,
   – *soit par ses moments : pour tout $n \geq 0$*
   $$\int_{\mathbb{U}} \xi^n d\nu_t(\xi) = \int_{\mathbb{U}} \xi^{-n} d\nu_t(\xi) = e^{-\frac{nt}{2}} \sum_{k=0}^{n-1} \frac{(-t)^k}{k!} \binom{n}{k+1} n^{k-1}.$$

La convergence du mouvement brownien unitaire vers le mouvement brownien multiplicatif libre est donnée par le théorème suivant, montré par Biane dans [Bia97].

Théorème 2.1.3. *(Biane)*

1. *Le processus $(U_N(t))_{t\geq 0}$ converge en loi, quand N tend vers l'infini vers un mouvement brownien multiplicatif libre.*
2. *De plus, si $U_N^{(1)}, \ldots, U_N^{(n)}$ sont des mouvements browniens unitaires indépendants, alors la famille $U_N^{(1)}, \ldots, U_N^{(n)}$ converge en loi quand N tend vers l'infini vers une famille $u^{(1)}, \ldots, u^{(n)}$ de mouvements browniens multiplicatifs libres, libres entre eux.*

Pour la bonne compréhension de ce résultat, nous renvoyons encore une fois le lecteur à l'appendice : on y trouvera notamment détaillée la notion de convergence en loi dans ce cadre.

Le premier point est la partie délicate du résultat, la deuxième assertion s'en déduit ensuite via des arguments relativement standards de la théorie des probabilités libres.

On peut cependant comprendre une partie de la signification de ce résultat, notamment le premier point sans faire appel aux probabilités libres. La proposition 2.1.4 est une conséquence importante du théorème qui s'exprime en termes de probabilités classiques.

Rappelons d'abord que si f est une application de $\mathbb{U}$ dans $\mathbb{R}$, on peut l'étendre en une application de $\mathcal{U}(N)$ dans $\mathbb{R}$ par le calcul fonctionnel : soit $U \in \mathcal{U}(N)$, il existe $z_1, \ldots, z_n \in \mathbb{U}$ et $V \in \mathcal{U}(N)$ telle que $U = V\mathrm{diag}(z_1, \ldots, z_n)V^*$. Alors $f(U) = V\mathrm{diag}(f(z_1), \ldots, f(z_n))V^*$.

Proposition 2.1.4. *Soit $f : \mathbb{U} \to \mathbb{R}$ continue. Alors, pour tout $t \geq 0$, on a*

$$\mathbb{E}\left(\frac{1}{N} \operatorname{Tr} f(U_N(t))\right) \xrightarrow[N\to\infty]{} \int_{\mathbb{U}} f(z) d\nu_t(z).$$

La preuve de cette proposition pour les monômes (convergence des moments) est le point crucial de la démonstration du théorème de Biane, elle fait essentiellement appel à des arguments de théorie des représentations et de combinatoire.

Nous détaillons cette preuve pour rendre moins artificielle la définition du mouvement brownien multiplicatif libre (notamment l'expression des moments de ν_t) mais la lecture de ce paragraphe n'est pas indispensable à la compréhension de la suite : le lecteur pressé pourra s'intéresser directement aux fluctuations.

2.1.3 Aparté combinatoire : convergence des moments du mouvement brownien unitaire vers ceux de ν_t

L'objectif de ce paragraphe est de donner les grandes lignes de la preuve de la proposition suivante :

Proposition 2.1.5. *Pour tout $n \in \mathbb{N}^*$, on a*

$$\mathbb{E}\left(\frac{1}{N}\operatorname{Tr}(U_N(t)^n)\right) \xrightarrow[N\to\infty]{} e^{-\frac{nt}{2}}\sum_{k=0}^{n-1}\frac{(-t)^k}{k!}\binom{n}{k+1}n^{k-1}.$$

Nous suivons pour cela la preuve de Biane dans [Bia97].

On introduit d'abord la notation suivante : si P est un polynôme symétrique et U une matrice unitaire, $P(U)$ désigne le résultat de l'évaluation de ce polynôme sur les valeurs propres de U. Dans toute la démonstration, on prendra N et n des entiers naturels tels que $N \geq n$.

La densité Q_t de la loi de $U_N(t)$ par rapport à la mesure de Haar sur $\mathcal{U}(N)$ possède un développement de Fourier uniformément convergent de la forme suivante :

$$Q_t(U) = \sum_{\alpha\in\mathbb{Z}_\downarrow^N} e^{-\frac{c_2(\alpha)t}{2N}} s_\alpha(I_N)\overline{s_\alpha(U)},$$

où $\mathbb{Z}_\downarrow^N$ désigne l'ensemble des N-uplets décroissants $\alpha = (\alpha_1 \geq \ldots \geq \alpha_N)$ d'éléments de $\mathbb{Z}$, s_α le polynôme de Schur associé au N-uplet α et I_N l'identité dans $\mathcal{U}(N)$. La fonction s_α est fonction propre du laplacien dans $\mathcal{U}(N)$ associée à la valeur propre $c_2(\alpha)$, qui est donnée par la formule suivante :

$$c_2(\alpha) = \sum_{i=1}^N \alpha_i^2 + \sum_{i<j}\alpha_i - \alpha_j. \tag{2.1.3}$$

On peut exprimer

$$\mathbb{E}\left(\frac{1}{N}\operatorname{Tr}(U_N(t)^n)\right) = \frac{1}{N}\mathbb{E}(p_n(U_N(t))),$$

où p_n est le polynôme symétrique donné par $p_n(x_1,\ldots,x_N) = x_1^n + \ldots + x_N^n$.

On a donc

$$\mathbb{E}\left(\frac{1}{N}\operatorname{Tr}(U_N(t)^n)\right) = \frac{1}{N}\sum_{\alpha\in\mathbb{Z}_\downarrow^N} e^{-\frac{c_2(\alpha)t}{2N}} s_\alpha(I_N)\mathbb{E}(\overline{s_\alpha(U)}p_n(U)),$$

où dans le terme de droite $\mathbb{E}$ désigne l'espérance sous la mesure de Haar.

Or les polynômes de Schur possèdent deux propriétés qui vont nous intéresser :

– ils forment une base des polynômes symétriques, de sorte que l'on peut exprimer le polynôme p_n dans la base des polynômes de Schur. Plus précisément, on a la décomposition suivante

$$p_n = \sum_{k=0}^{n-1}(-1)^k s_{\lambda_{n,N,k}},$$

où $\lambda_{n,N,k}$ est le N-uplet $(n-k, 1, \ldots, 1, 0, \ldots, 0)$ avec k fois 1 et $N-k-1$ fois 0 (le tableau correspondant a une forme de crochet).

– ils sont orthonormaux au sens où

$$\mathbb{E}(\overline{s_\alpha(U)}s_\beta(U)) = \delta_{\alpha\beta}\mathbf{1}_{\ell(\alpha)\leq N},$$

où $\ell(\alpha)$ est le nombre de lignes de longueur non nulle dans α.

On obtient donc

$$\mathbb{E}\left(\frac{1}{N}\operatorname{Tr}(U_N(t)^n)\right)=\frac{1}{N}\sum_{k=0}^{n-1}e^{-\frac{c_2(\lambda_{n,N,k})t}{2N}}(-1)^k s_{\lambda_{n,N,k}}(I_N).$$

La quantité $s_{\lambda_{n,N,k}}(I_N)$ est la dimension de la représentation associée à $\lambda_{n,N,k}$. On connaît une formule explicite pour la dimension d'une représentation qui dans le cas simple d'un crochet donne

$$s_{\lambda_{n,N,k}}(I_N)=\frac{(N+n-k-1)!}{(N-k-1)!n(n-k-1)!k!}.$$

De même, grâce à la formule (2.1.3), on obtient

$$c_2(\lambda_{n,N,k})=nN+n(n-1)-2kn.$$

Soit

$$\mathbb{E}\left(\frac{1}{N}\operatorname{Tr}(U_N(t)^n)\right)=\frac{1}{Nn}e^{-\frac{nt}{2}}e^{-\frac{n(n-1)t}{2N}}\sum_{k=0}^{n-1}(-1)^k e^{\frac{ntk}{N}}\frac{(N+n-k-1)!}{(N-k-1)!k!(n-k-1)!}.$$

On écrit $\frac{(N+n-k-1)!}{(N-k-1)!}=\sum_{\ell=0}^{n}A(\ell,N,n)k^{\ell}$, de sorte que

$$\mathbb{E}\left(\frac{1}{N}\operatorname{Tr}(U_N(t))^n\right)=\frac{1}{n!}e^{-\frac{nt}{2}}e^{-\frac{n(n-1)t}{2N}}\sum_{j=0}^{\infty}\frac{(nt)^j}{j!N^{j+1}}\sum_{\ell=0}^{n}A(\ell,N,n)\sum_{k=0}^{n-1}(-1)^k\binom{n-1}{k}k^{j+\ell}.$$

En développant la fonction $x\mapsto(1-e^x)^{n-1}$, on peut observer que

$$\sum_{k=0}^{n-1}(-1)^k\binom{n-1}{k}k^{j+\ell}=\begin{cases}0, & \text{si } j+\ell<n-1,\\ (-1)^{n-1}(n-1)!, & \text{si } j+\ell=n-1.\end{cases}$$

Par ailleurs, on peut vérifier que $A(\ell,N,n)\sim(-1)^{\ell}\binom{n}{\ell}N^{n-\ell}$ donc si $j+\ell>n-1$, le terme correspondant est un $o(1)$. La contribution principale correspond donc à $j+\ell=n-1$ et en remplaçant, on obtient bien l'expression annoncée dans la Proposition 2.1.5.

Résumé de la partie 2.1 :

- Nous appelons mouvement brownien unitaire (de taille N) le processus $(U_N(t))_{t\geq 0}$ à valeurs dans $\mathcal{U}(N)$ de condition initiale $U_N(0)$ et solution de l'EDS $dU_N(t)=U_N(t)dK_N(t)-\frac{1}{2}U_N(t)dt$, où $(K_N(t))_{t\geq 0}$ est un mouvement brownien anti-hermitien.

- Dans un espace de probabilités non-commutatif, nous appelons mouvement brownien multiplicatif libre une famille $(u_t)_{t\geq 0}$ d'éléments unitaires à accroissements mulitplicatifs libres et stationnaires telle que u_t a pour loi ν_t, mesure de probabilités sur $\mathbb{U}$ de moments

$$\int_{\mathbb{U}}\xi^n d\nu_t(\xi)=\int_{\mathbb{U}}\xi^{-n}d\nu_t(\xi)=e^{-\frac{nt}{2}}\sum_{k=0}^{n-1}\frac{(-t)^k}{k!}\binom{n}{k+1}n^{k-1}.$$

- Quand N tend vers l'infini, le mouvement brownien unitaire $(U_N(t))_{t\geq 0}$ converge en loi vers un mouvement brownien multiplicatif libre .

2.2 Etude des fluctuations du mouvement brownien unitaire

L'objet principal de l'article [A7] est d'étudier les fluctuations du mouvement brownien unitaire et d'établir pour ce processus un résultat de type théorème central limite (TCL). Comme nous l'avons souligné au début de la section précédente, le mouvement brownien unitaire peut être considéré comme la version dynamique des matrices de Haar. Il est donc bon avant d'étudier ses fluctuations d'avoir à l'esprit le résultat correspondant pour les matrices de Haar. Il s'agit d'un résultat de Diaconis et Evans, que nous détaillons maintenant.

2.2.1 TCL pour les matrices de Haar

Sur un groupe compact (en fait localement compact), on sait qu'il existe une unique (à normalisation près) mesure borélienne finie sur les compacts, invariante par l'action à gauche du groupe. C'est la mesure de Haar, qui correspond à la notion intuitive de loi uniforme sur le groupe.

Ici U_N désignera une matrice distribuée selon la mesure de Haar sur $\mathcal{U}(N)$. On a alors la proposition suivante :

Proposition 2.2.1. *Soit $f : \mathbb{U} \to \mathbb{R}$ continue. Alors*

$$\mathbb{E}\left(\frac{1}{N}\operatorname{Tr} f(U_N)\right) \xrightarrow[N\to\infty]{} \int_{\mathbb{U}} f(z)d\lambda(z),$$

où λ est la mesure uniforme sur $\mathbb{U}$.

Cette proposition est évidemment l'analogue de la Proposition 2.1.4 où ν_t est remplacée par λ ; elle se vérifie simplement en utilisant les propriétés d'invariance de la mesure de Haar.

Dans [DE01], Diaconis et Evans montrent que pour une fonction f assez régulière , les fluctuations de la fonctionnelle $\frac{1}{N}\operatorname{Tr} f$ sous la mesure de Haar sont asymptotiquement gaussiennes.

Détaillons leur résultat. On définit d'abord l'ensemble des fonctions pour lesquelles le TCL est valable.

Pour $f : \mathbb{U} \to \mathbb{R}$ de carré intégrable et $j \in \mathbb{Z}$, on note $\hat{f}(j) = \frac{1}{2\pi}\int_{\mathbb{U}} f(t)e^{ijt}dt$ son j-ième coefficient de Fourier.

Définition 2.2.2. *$\mathcal{H}^{1/2}(\mathbb{U})$ est l'ensemble des fonctions de $\mathbb{U}$ dans $\mathbb{R}$, de carré intégrable telles que*

$$\|f\|_{1/2}^2 := \sum_{j\in\mathbb{Z}} |j||\hat{f}(j)|^2$$

est fini.

L'application $f \mapsto \|f\|_{1/2}$ est une norme hilbertienne et on note $\langle\cdot,\cdot\rangle_{1/2}$ le produit scalaire associé.

Le résultat est le suivant :

Théorème 2.2.3. *(Diaconis, Evans)*
Soient U_N distribuée selon la mesure de Haar sur $\mathcal{U}(N)$ et $f_1, \dots, f_n \in \mathcal{H}^{1/2}(\mathbb{U})$. On définit la matrice de covariance suivante, de taille $n \times n$:

$$\Sigma(f_1, \dots, f_n) := (\langle f_i, f_j\rangle)_{i,j=1,\dots,n}.$$

Quand N tend vers l'infini, on a

$$(\operatorname{Tr} f_i(U_N) - \mathbb{E}(\operatorname{Tr} f_i(U_N))) \xrightarrow[N\to\infty]{} \mathcal{N}(0, \Sigma(f_1, \dots, f_n)),$$

où le terme de droite désigne la loi normale centrée de covariance $\Sigma(f_1, \dots, f_n)$ et la convergence a lieu au sens de la convergence en loi des vecteurs aléatoires.

Comme le résultat de Biane (Théorème 2.1.4), la preuve utilise des calculs de moments à l'aide d'élégants arguments de théorie des représentations et de combinatoire.

Présentons l'esquisse de preuve pour $n = 1$.

Si $f : \mathbb{U} \to \mathbb{C}$ a pour série de Fourier $f(e^{i\theta}) = \sum_{j\in\mathbb{Z}} \hat{f}_j e^{ij\theta}$ et U_N distribuée selon la mesure de Haar,

$$\operatorname{Tr} f(U_N) = N\hat{f}_0 + \sum_{j=1}^{\infty} \hat{f}_j \operatorname{Tr}(U_N^j) + \sum_{j=1}^{\infty} \hat{f}_{-j} \overline{\operatorname{Tr}(U_N^j)}.$$

Pour montrer que $\operatorname{Tr} f(U_N) - \mathbb{E}(\operatorname{Tr} f(U_N))$ converge vers une variable aléatoire de loi $\mathcal{N}(0, \|f\|_{1/2}^2)$, on est donc amenés à calculer des moments mixtes de la forme

$$\mathbb{E}\left(\prod_{j=1}^{k} (\operatorname{Tr}(U_N^j))^{a_j} \overline{(\operatorname{Tr}(U_N^j))^{b_j}}\right).$$

En particulier, pour le calcul de la variance, qui est le seul que nous détaillerons, on va montrer que

$$\mathbb{E}\left(\operatorname{Tr}(U_N^j)\overline{\operatorname{Tr}(U_N^k)}\right) = \delta_{jk}(j \wedge N).$$

Comme dans la preuve de la Proposition 2.1.5, on va pour cela utiliser encore une fois le fait que le polynôme symétrique p_j s'exprime sous la forme

$$p_j = \sum_{\ell=0}^{(j-1)\wedge(N-1)} (-1)^\ell s_{\lambda_{n,N,\ell}}$$

ainsi que l'orthogonalité des polynômes de Schur.

En développant, on a

$$\mathbb{E}(p_j(U_N)\overline{p_k(U_N)}) = \sum_{\ell=0}^{(j-1)\wedge(N-1)} \sum_{m=0}^{(k-1)\wedge(N-1)} (-1)^{\ell+m} \mathbb{E}(s_{\lambda_{n,N,\ell}} \overline{s_{\lambda_{n,N,m}}}) = \delta_{jk}(j \wedge N).$$

A partir de là, on obtient facilement que la covariance est $\|f\|_{1/2}^2$.

2.2.2 TCL pour le mouvement brownien unitaire

Dans [A7], notre but principal a été d'établir un analogue du résultat de Diaconis et Evans pour le mouvement brownien unitaire.

Une formulation intuitive de notre résultat souligne l'analogie avec leur énoncé : les fluctuations de la fonctionelle $\frac{1}{N} \operatorname{Tr} f$ sous la loi du mouvement brownien unitaire au temps t sont asymptotiquement gaussiennes.

Afin de donner un énoncé plus précis du résultat, il nous faut définir la covariance : pour f, g des fonctions $\mathcal{C}^1$ sur $\mathbb{U}$ à dérivées lipschitziennes et $T > 0$, on définit la forme bilinéaire positive suivante :

$$\sigma_T(f, g) = \int_0^T \tau(f'(u_s v_{T-s}) g'(u_s w_{T-s})) ds, \tag{2.2.1}$$

avec u, v, w trois mouvements browniens multiplicatifs libres, libres entre eux.

Notre résultat principal est le suivant :

Théorème 2.2.4. *Soient $f_1, \dots, f_n$ des fonctions $\mathcal{C}^1$ à dérivées lipschitziennes et $T > 0$, on définit la matrice positive $n \times n$ suivante*

$$\Sigma_T(f_1, \dots, f_n) := (\sigma_T(f_i, f_j))_{i,j=1,\dots,n}.$$

On a alors la convcergence suivante :

$$(\mathrm{Tr}\, f_i(U_N(T)) - \mathbb{E}(\mathrm{Tr}\, f_i(U_N(T)))) \xrightarrow[N\to\infty]{} \mathcal{N}(0, \Sigma_T(f_1, \dots, f_n)),$$

au sens de la convergence en loi des vecteurs aléatoires.

Dans ce résultat, les fonctions pour lesquelles on démontre le TCL sont supposées $\mathcal{C}^1$ à dérivées lipschitziennes. Cette restriction tient à notre technique de preuve et ceci ne constitue sans doute pas l'ensemble optimal pour lequel le résultat est vrai. Des précisions seront données dans le paragraphe 2.3.2.

2.2.3 Liens entre les deux résultats

L'analogie entre les théorèmes 2.2.3 et 2.2.4 est évidente dans leur formulation. Il existe en réalité plus qu'une analogie mais un vrai lien entre ces deux résultat que l'on peut représenter par le diagramme suivant, exprimé pour une seule fonction-test pour plus de commodité : pour f assez régulière,

$$\begin{array}{ccc} \mathrm{Tr}\, f(U_N(t)) - \mathbb{E}(\mathrm{Tr}\, f(U_N(t))) & \xrightarrow{t\to\infty} & \mathrm{Tr}\, f(U_N) - \mathbb{E}(\mathrm{Tr}\, f(U_N)) \\ \Big\downarrow{\scriptstyle N\to\infty} & & \Big\downarrow{\scriptstyle N\to\infty} \\ \mathcal{N}(0, \sigma_t(f,f)) & \xrightarrow{t\to\infty} & \mathcal{N}(0, \|f\|^2_{\mathcal{H}_{1/2}}) \end{array}$$

Comme nous l'avons vu plus haut, les asymptotiques verticales correspondent respectivement aux théorèmes 2.2.3 et 2.2.4 pour $n = 1$.

La convergence du haut correspond simplement au fait que à n fixé la loi du processus de Markov $(U_N(t))_{t\geq 0}$ converge, quand t tend vers l'infini vers sa mesure invariante qui n'est autre que la mesure de Haar sur le groupe $\mathcal{U}(N)$.

Donnons maintenant quelques précisions sur la convergence du bas, c'est-à-dire de notre covariance $\sigma_T(f, f)$ vers celle de Diaconis et Evans $\langle f, f\rangle_{1/2}$. Le résultat correspondant peut être énoncé de la manière suivante.

Proposition 2.2.5. *Si u, v, w sont trois mouvements browniens multiplicatifs libres, libres entre eux, $j, k \in \mathbb{Z}$ et $T > 0$, on pose*

$$\tau_{j,k}(T) := \int_0^T \tau((u_s v_{T-s})^j (u_s w_{T-s})^k) ds.$$

Pour tout T assez grand, pour toutes $f, g \in \mathcal{H}^{1/2}(\mathbb{U})$, la série de terme général $jk\hat{f}(j)\hat{g}(k)\tau_{j,k}(T)$ est convergente et on pose

$$\tilde{\sigma}_T(f, g) := -\sum_{j,k\in\mathbb{Z}} jk\hat{f}(j)\hat{g}(k)\tau_{j,k}(T).$$

Alors pour T assez grand, $\tilde{\sigma}_T(f, g)$ coincide avec $\sigma_T(f, g)$ et

$$\tilde{\sigma}_T(f, g) \xrightarrow[T\to\infty]{} \langle f, g\rangle_{1/2}.$$

Dans l'énoncé de cette proposition, “T assez grand” signifie T supérieur ou égal à un T_0 explicite qui ne dépend pas de f et g (par exemple $T \geq 32$ convient). En revanche l'analyse du comportement des $\tau_{jk}(T)$ pour T petit est très délicate.

La démonstration de cette proposition fait appel à du calcul stochastique libre et notamment au fait que le mouvement brownien multiplicatif libre $(u_t)_{t\geq 0}$ vérifie l'EDS libre 2.1.2.

Résumé de la partie 2.2 :

Si $(U_N(t))_{t\geq 0}$ est un mouvement brownien unitaire et U_N une matrice distribuée selon la mesure de Haar, pour f une fonction suffisamment régulière, les fluctuations des fonctionnelles $\mathrm{Tr}\, f(U_N(t))$ et $\mathrm{Tr}\, f(U_N)$ autour de leur moyenne respective sont asymptotiquement gaussiennes. Il existe de plus un lien étroit entre les deux résultats, donné par le diagramme suivant :

$$\begin{array}{ccc} \mathrm{Tr}\, f(U_N(t)) - \mathbb{E}(\mathrm{Tr}\, f(U_N(t))) & \xrightarrow{t\to\infty} & \mathrm{Tr}\, f(U_N) - \mathbb{E}(\mathrm{Tr}\, f(U_N)) \\ \Big\downarrow{\scriptstyle N\to\infty} & & \Big\downarrow{\scriptstyle N\to\infty} \\ \mathcal{N}(0, \sigma_t(f,f)) & \xrightarrow{t\to\infty} & \mathcal{N}(0, \|f\|^2_{\mathcal{H}_{1/2}}) \end{array}$$

2.3 Eléments de preuve

2.3.1 L'impasse combinatoire partielle

Une idée naturelle pour montrer notre théorème est de suivre le même cheminement que dans la preuve de Diaconis et Evans en calculant des moments mixtes de la forme $\mathbb{E}(\mathrm{Tr}(U_N(t)^j)\overline{\mathrm{Tr}(U_N(t)^k)})$.

Toujours en développant nos polynômes sur la base de Schur et en utilisant l'orthogonalité des polynômes de Schur et en analysant, comme dans la preuve de Biane, quels sont les tableaux qui interviennent, nous avons pu obtenir dans [A7] une expression explicite de ces moments mixtes sous la forme du résultat suivant

Proposition 2.3.1. *Soit $(V_N(t))_{t\geq 0}$ le brownien sur le groupe spécial unitaire*[4] *de taille N associé au produit scalaire $(X,Y) = N\,\mathrm{Tr}(X^*Y)$. Soient n et m des entiers positifs et $N \geq n+m+1$. Alors*

$$\mathbb{E}\left[\mathrm{Tr}(V_N(t)^n)\overline{\mathrm{Tr}(V_N(t)^m)}\right] = n\delta_{n,m} + (-1)^{n+m} e^{-(n+m)\frac{t}{2} - \frac{n(n-1)+m(m-1)}{N}\frac{t}{2} - \frac{(n-m)^2}{N^2}\frac{t}{2}}$$
$$\sum_{r_1=0}^{n-1}\sum_{r_2=0}^{m-1}\left[(-1)^{r_1+r_2} e^{-nr_1\frac{t}{2}}\binom{n-1}{r_1}\binom{N+r_1}{n} e^{-nr_2\frac{t}{2}}\binom{m-1}{r_2}\binom{N+r_2}{m}\right.$$
$$\left.\frac{(N+r_1+r_2+1)(N-n-m+r_1+r_2+1)}{(N-n+r_1+r_2+1)(N-m+r_1+r_2+1)}\right].$$

Cependant, comme ces moments mixtes se présentent sous la forme de sommes alternées de termes de très grande amplitude, l'analyse de leur asymptotique est très délicate et nous n'avons pas pu en établir de façon rigoureuse la convergence assez précisément pour montrer un TCL.

Nous avons donc opté pour une méthode complètement différente basée sur du calcul stochastique que nous présentons dans le paragraphe qui suit.

4. Cette expression est donnée par commodité pour le brownien sur le groupe spécial unitaire mais on sait que les deux browniens ont la même asymptotique.

2.3.2 La preuve par le calcul d'Itô

Sans entrer dans les détails techniques, nous donnons dans ce paragraphe suffisamment d'éléments de preuve du résultat principal de [A7] pour comprendre l'expression de la covariance (2.2.1).

L'idée générale est de considérer les martingales de la forme $M_N^f(t) := \mathbb{E}(\mathrm{Tr}\, f(U_N(t))|\mathcal{F}_{N,t}) = P_{T-t}(\mathrm{Tr}\, f)(U_N(t))$, avec $0 \leq t \leq T$, $(\mathcal{F}_{N,t})_{t\geq 0}$ la filtration canonique et $(P_t)_{t\geq 0}$ le semi-groupe associés au processus $(U_N(t))_{t\geq 0}$. On va montrer que les accroissements de cette martingale $Q_N^f(t) := M_N^f(t) - M_N^f(0)$ sont asymptotiquement gaussiens.

Pour cela, les deux points cruciaux sont l'étude de la convergence de l'espérance du crochet $\mathbb{E}\langle Q_N^f\rangle(t)$ et le contrôle de sa variance. Nous ne détaillons ci-dessous que la convergence du crochet, c'est-à-dire

$$\mathbb{E}\langle Q_N^f\rangle(t) \xrightarrow[N\to\infty]{} \int_0^t \sigma_{T,s}(f,f)ds, \tag{2.3.1}$$

avec

$$\sigma_{T,s}(f,g) = \tau(f'(u_s v_{T-s})g'(u_s w_{T-s})),$$

où u, v, w sont trois mouvements browniens multiplicatifs libres, libres entre eux.

A partir de (2.3.1) et du contrôle de la variance, la preuve du théorème est standard : on pose, pour $\xi \in \mathbb{R}$,

$$R_N(t) = e^{i\xi Q_N^f(t) - \frac{1}{2}\xi^2 \int_0^t \sigma_{T,s}(f,f)ds}.$$

Le processus $(R_N(t))_{t\geq 0}$ vérifie alors l'EDS matricielle :

$$dR_N(t) = i\xi R_N(t) dQ_n^f(t) + \frac{1}{2}\xi^2 R_N(t)[\sigma_{T,t}(f,f)dt - d\langle Q_N\rangle(t)],$$

de sorte que

$$\mathbb{E}(R_N(t)) \xrightarrow[N\to\infty]{} 1.$$

Montrons donc (2.3.1). Il faut pour cela obtenir une formule d'Itô pour les fonctions du brownien unitaire, donné par la proposition suivante

Proposition 2.3.2. *Soit $F : \mathbb{R}\times\mathcal{U}(N) \to \mathbb{R}$ de classe C^2 et $X_1, \ldots, X_{N^2}$ une base orhtonormée de $\mathfrak{u}(N)$. Alors, pour tout $t \geq 0$,*

$$\begin{aligned} F(t, U_N(t)) = F(0, I_N) &+ \sum_{k=1}^{N^2} \int_0^t (\mathcal{L}_{X_k}F)(s, U_N(s))\, d(X_k, K_N)(s) \\ &+ \int_0^t \left(\frac{1}{2}\Delta F + \partial_t F\right)(s, U_N(s))\, ds, \end{aligned} \tag{2.3.2}$$

avec pour $U \in \mathcal{U}(N)$ et $X \in \mathfrak{u}(N)$

$$(\mathcal{L}_X F)(U) = \frac{d}{dt}_{|t=0} F(Ue^{tX}), \quad \Delta = \sum_{k=1}^{N^2} \mathcal{L}_{X_k}^2 \tag{2.3.3}$$

et les processus $\{(X_k, K_N) : k \in \{1, \ldots, N^2\}\}$ sont des mouvements browniens réels standards indépendants.

Pour simplifier un peu l'écriture, on pose $F = \mathrm{Tr}\, f$. La proposition 2.3.2 donne :

$$Q_N^f(T) = \int_0^T \sum_{k=1}^{N^2} \mathcal{L}_{X_k}(P_{T-s}F)(U_N(s))d(X_k, K_N)(s).$$

De là, comme les opérateurs $\mathcal{L}_{X_k}$ et P_t commutent, on obtient directement

$$\begin{aligned} \langle M_N^f(T)\rangle &= \int_0^T \sum_{k=1}^{N^2}(\mathcal{L}_{X_k}(P_{T-s}F)(U_N(s)))^2 ds \qquad (2.3.4) \\ &= \int_0^T \sum_{k=1}^{N^2}((P_{T-s}\mathcal{L}_{X_k}F)(U_N(s)))^2 ds \\ &= \int_0^T \sum_{k=1}^{N^2} \mathbb{E}_{U_N}((\mathcal{L}_{X_k}F)(U_N(s)V_N(T-s)))(\mathcal{L}_{X_k}F)(U_N(s)W_N(T-s)))ds, \end{aligned}$$

avec U_N, V_N, W_N trois mouvements browniens unitaires indépendants et $\mathbb{E}_{U_N}$ l'espérance conditionellement à $\mathcal{F}_{N,t}$.

On voit dans cette partie du calcul pourquoi on exige "trop" de régularité sur la fonction f. Comme le semi-groupe a un effet régularisant, (2.3.4) est défini dès que f est dans $\mathbb{L}^2(\mathbb{U})$. En revanche les opérateurs $\mathcal{L}_{X_k}$ et P_t ne commutent qu'à condition que F soit suffisamment régulière ; en particulier il faut que $\mathcal{L}_{X_k}F$ soit bien définie, donc f différentiable.

Un peu de calcul dans le groupe unitaire permet de voir que pour $Y \in \mathfrak{u}(N)$ et f différentiable,

$$\mathcal{L}_Y(\mathrm{Tr}\, f)(U) = -i\,\mathrm{Tr}(f'(U)Y)$$

et

$$\sum_{k=1}^{N^2} \mathrm{Tr}(AX_k)\,\mathrm{Tr}(BX_k) = -\frac{1}{N}\,\mathrm{Tr}(AB).$$

On obtient alors

$$\mathbb{E}(\langle Q_N^f\rangle(T)) := \int_0^T \mathbb{E}\left(\frac{1}{N}\,\mathrm{Tr}\, f'(U_N(s)V_N(T-s))f'(U_N(s)W_N(T-s))\right) ds.$$

Le théorème 2.1.3 permet d'obtenir la convergence

$$\mathbb{E}\langle Q_N^f\rangle(T) \xrightarrow[N\to\infty]{} \int_0^T \tau(f'(u_s v_{T-s})g'(u_s w_{T-s}))ds,$$

avec u, v, w trois mouvements browniens multiplicatifs libres, libres entre eux.

Le contrôle de la variance du crochet utilise des arguments standards de concentration ; l'idée générale est d'utiliser le fait que $\mathcal{U}(N)$ est une variété à courbure positive ; si $f : \mathbb{U} \to \mathbb{C}$ est 1-Lipschitz alors $\frac{1}{N}\,\mathrm{Tr}\, f : \mathcal{U}(N) \to \mathbb{C}$ est $\frac{1}{N}$–Lipschitz. Finalement, on obtient notre résultat pour des fonctions à dérivées lipschitziennes.

Résumé de la partie 2.3 :

Il est possible d'obtenir une expression explicite des moments mixtes de la forme $\mathbb{E}(\mathrm{Tr}(U_N(t))^n\overline{\mathrm{Tr}(U_N(t))^m})$ mais l'analyse de leur asymptotique paraît difficile. Pour la preuve du Théorème 2.2.4, on préfère utiliser des arguments classiques de TCL pour des martingales en s'appuyant sur la forumule d'Itô pour les fonctionnelles de $(U_N(t))_{t\geq 0}$ donnée par la Proposition 2.3.2.

Chapitre 3

Inégalités de transport libre et matrices aléatoires [A10]

Le but de ce chapitre est de présenter l'article [A10] écrit en collaboration avec Édouard Maurel-Segala.

3.1 Inégalités transport-entropie classiques

Afin de replacer dans un contexte un peu général les inégalités transport-entropie que nous avons cherché à montrer dans [A10], nous commençons dans cette partie par faire quelques rappels sur le pendant classique, maintenant bien connu, de ces inégalités. Dans toute cette partie, la référence principale que nous utilisons est l'article de survol [GL10] écrit par Gozlan et Léonard.

La question du transport optimal est un vieux problème, que l'on fait traditionnellement remonter à Monge, mais aussi un domaine très moderne qui a connu une véritable explosion au cours des dix ou quinze dernières années (cf par exemple [Vil09] pour un état de l'art très complet). Nous ne cherchons pas ici à donner un aperçu global de cette immense littérature mais simplement à poser quelques jalons permettant de mettre nos résultats en perspective.

3.1.1 Transport optimal, inégalités de Talagrand

Le problème de Monge peut être exprimé ainsi : "si transporter une quantité infinitésimale de x en y coûte $c(x,y)$, comment minimiser le coût global pour transporter une masse donnée vers un déblais de forme donnée ?"

Une formulation moderne de cette question est le problème dit de Monge-Kantorovich qui peut être exprimé de la manière suivante : pour $\mathcal{X}$ un espace polonais, si on se donne une fonction de coût $c : \mathcal{X}^2 \to [0,\infty)$, semi-continue inférieurement, nulle sur la diagonale et deux mesures de probabilités μ et ν sur $\mathcal{X}$, on cherche à déterminer

$$\mathcal{T}_c(\mu,\nu) := \inf \mathbb{E}(c(X,Y)),$$

avec X variable aléatoire de loi μ et Y variable aléatoire de loi ν.

Sauf mention contraire, tous les espaces sur lesquels nous considérerons des mesures de probabilités seront des espaces polonais, notés $\mathcal{X}$.

Parmi les nombreuses fonctions de coût possibles, une famille a été particulièrement étudiée : si d est une métrique sur $\mathcal{X}$ et $p \geq 1$, on prend $c = d^p$.

On peut alors vérifier que la puissance $1/p$-ième du coût optimal, soit $W_p := \mathcal{T}_{d^p}^{1/p}$ est une distance sur l'ensemble $\mathcal{P}(\mathcal{X})$ des mesures de probabilités sur $\mathcal{X}$. Elle est appelée distance de Wasserstein d'ordre p.

Une autre façon habituelle de mesurer une “distance” entre deux mesures est l’entropie relative :

$$H(\nu|\mu) := \begin{cases} \int \log \frac{d\nu}{d\mu} d\nu, & \text{si } \nu \text{ est absolument continue par rapport à } \mu, \\ +\infty, & \text{sinon.} \end{cases}$$

On dit que μ vérifie une inégalité transport-entropie (associée au coût c) s’il existe une fonction f : $[0, \infty) \to [0, \infty)$, croissante, telle que $f(0) = 0$, telle que pour tout ν,

$$f(\mathcal{T}(\mu, \nu)) \leq H(\nu|\mu).$$

Une famille très importante d’inégalités de ce type sont les inégalités de Talagrand, notamment T_1 et T_2 dont nous allons beaucoup parler par la suite.

On dit que μ vérifie T_2 avec une constante C (notée $T_2(C)$) s’il existe une constante $C > 0$ telle que pour tout $\nu \in \mathcal{P}(\mathcal{X})$,

$$W_2^2(\nu, \mu) \leq CH(\nu|\mu)$$

(dans ce cas $f = \frac{x}{C}$).

On dit que μ vérifie T_1 avec une constante C (notée $T_1(C)$) s’il existe une constante $C > 0$ telle que pour tout $\nu \in \mathcal{P}(\mathcal{X})$,

$$W_1^2(\nu, \mu) \leq CH(\nu|\mu)$$

(dans ce cas $f(x) = \frac{1}{C}x^2$).

L’inégalité T_2 est plus forte que T_1 puisque cette dernière s’en déduit par l’inégalité de Cauchy-Schwarz. Même si l’inégalité T_2 est en général plus délicate à montrer, les critères connus étant en général difficiles à vérifier, la grande force de l’inégalité T_2 est qu’elle se tensorise facilement. On peut ainsi obtenir de bonnes estimations en grande dimension sans avoir à payer le prix d’une trop grosse dégradation de la constante.

Le résultat fondateur dans le domaine est la démonstration par Talagrand [Tal96] du fait que la mesure gaussienne standard sur $\mathbb{R}^n$ muni de la distance euclidienne vérifie l’inégalité $T_2(2)$, avec en particulier une constante qui ne dépend pas de la dimension n de l’espace.

3.1.2 Inégalités transport-entropie et concentration

Un des intérêts principaux des inégalités transport-entropie est leur lien avec la concentration de la mesure. Là encore, nous n’allons pas tenter de dresser un panorama exhaustif de ces liens (cf [GL10]) mais nous allons donner un des résultats les plus connus dans le domaine, ainsi que sa démonstration, dite “argument de Marton”.

Proposition 3.1.1. *Si μ vérifie $T_2(C)$, μ se concentre comme une gaussienne au sens où*

$$\forall A \subset \mathcal{X} \text{ telle que } \mu(A) \geq 1/2, \mu(A^r) \geq 1 - e^{-\frac{(r-r_0)^2}{C}},$$

pour $r \geq r_0 := \sqrt{c \log 2}$ et $A^r := \{x \in \mathcal{X} / d(x, A) \leq r\}$.

La proposition est aussi vraie si on remplace T_2 par T_1 mais nous donnons la démonstration dans le cas T_2. La démonstration est simple et nous la pensons éclairante pour le lecteur, c’est pourquoi nous la donnons ci-dessous.

Preuve de la Proposition 3.1.1 : Soit μ une mesure de probabilités vérifiant $T_2(C)$. Soit A un borélien tel que $\mu(A) \geq 1/2$, $r \geq r_0$ et $B = \mathcal{X} \setminus A^r$.

On introduit la mesure μ_A qui est la mesure μ restreinte au borélien A soit $\mu_A = \frac{1}{\mu(A)} \mathbf{1}_A . \mu$.

Comme μ vérifie $T_2(C)$,

$$W_2^2(\mu_A,\mu) \leq CH(\mu_A|\mu) = C\log\left(\frac{1}{\mu_A}\right).$$

On a donc

$$\begin{aligned} r \leq W_2(\mu_A,\mu_B) &\leq W_2(\mu_A,\mu)+W_2(\mu,\mu_B) \\ &\leq \sqrt{C}\left(\sqrt{H(\mu_A|\mu)}+\sqrt{H(\mu_B|\mu)}\right) \\ r &\leq \sqrt{C}\left(\log\left(\frac{1}{\mu(A)}\right)+\log\left(\frac{1}{1-\mu(A^r)}\right)\right) \end{aligned}$$

En utilisant que $\mu(A) \geq 1/2$, on obtient à partir de la dernière inégalité que $\mu(A^r) \geq 1-e^{-\frac{(r-\sqrt{C\log 2})^2}{C}}$. □

Mentionnons aussi pour ce qui concerne le lien entre inégalités transport-entropie et concentration le résultat de Bolley, Guillin et Villani. Dans [BGV07], ils montrent comment déduire des inégalités de Talagrand des bornes explicites sur la convergence de la mesure empirique de variables indépendantes et identiquement distribuées vers leur mesure commune. Par exemple, si $X_1,\dots,X_n,\dots$ sont des variables indépendantes dans $\mathbb{R}^d$ de loi μ avec $E_\alpha = \int e^{\alpha|x|}d\mu(x) < +\infty$ et μ qui satisfait $T_2(C)$ alors il existe $v,u>0$ dépendant seulement de d,E_α,C tels que pour tout $d'>d$ et tout $\theta > vN^{-\frac{1}{4+d'}}$,

$$\mathbb{P}\left(W_1\left(\frac{1}{N}\sum_{i=1}^N \delta_{X_i},\mu\right)>\theta\right) < e^{-uN\min(\theta,\theta^2)}.$$

Nous verrons dans la troisième partie de ce chapitre comment appliquer l'inégalité transport-entropie que nous avons obtenue pour montrer des propriétés de concentration pour certains ensembles de matrices aléatoires.

3.1.3 Un critère important pour l'inégalité T_1

Comme nous l'avons déjà mentionné plus haut, l'inégalité T_2 est plus puissante que T_1 car elle se tensorise, ce qui permet automatiquement d'obtenir des inégalités en grande dimension sans dégradation des constantes. La contrepartie est que les critères pour T_2 sont délicats à vérifier alors T_1 est mieux comprise, essentiellement équivalente à des moments sous-gaussiens. Nous donnons ci-dessous un résultat en ce sens de Bobkov et Götze. Nous en donnons aussi la preuve, particulièrement éclairante sur le rapport entre transport et entropie. En effet, le coeur de la preuve de notre théorème principal va s'appuyer sur des arguments relativement similaires.

Théorème 3.1.2. *(Bobkov-Götze)*
Soit μ une mesure de probabilités telle que $\exists\, x_0$ tel que $\int d(x,x_0)d\mu(x) < \infty$. Alors les deux propriétés suivantes sont équivalentes

1. *μ vérifie $T_1(2c)$*
2. *pour toute fonction f 1-lipschitzienne telle que $\int f d\mu = 0, \forall t\in\mathbb{R}, \int e^{tf}d\mu \leq e^{c\frac{t^2}{2}}$.*

Nous présentons ici la preuve de 2. $\Rightarrow$ 1. On pourrait raisonner par équivalence mais c'est le sens 2. $\Rightarrow$ 1. qui nous intéresse et inspirera le cas libre.

Le point crucial de la démonstration est d'utiliser la caractérisation duale de l'entropie : si g est une densité c'est-à-dire une fonction positive telle que $\int g d\mu = 1$, alors son entropie relativement à μ,

$$Ent_\mu(g) := \int g \log g d\mu$$

peut aussi s'exprimer comme

$$Ent_\mu(g) = \sup\left\{ \int ghd\mu \middle/ h \text{ telle que } \int e^h d\mu \leq 1 \right\}. \tag{3.1.1}$$

Soit g une densité, f 1-lipschitzienne telle que $\int f d\mu = 0$ et $t \geq 0$. Si μ vérifie 2., on a

$$\int \left(tf - c\frac{t^2}{2} \right) gd\mu \leq Ent_\mu(g)$$

ou encore

$$\int fgd\mu - \int fd\mu \leq \frac{ct}{2} + \frac{1}{t} Ent_\mu(g).$$

Donc si ν est absolument continue par rapport à μ,

$$\sup\left\{ \int fd\nu - \int fd\mu \middle/ f1 - \text{Lipschitz} \right\} \leq \frac{ct}{2} + \frac{1}{t} H(\nu|\mu)$$

soit

$$W_1(\mu, \nu) \leq \sqrt{2cH(\nu|\mu)},$$

où on a utilisé la dualité de Monge-Kantorovich (cf par exemple [Vil03]) pour le terme gauche et optimisé en t pour celui de droite.

Résumé de la partie 3.1 :

Les inégalités de Talagrand T_1 et T_2 constituent une famille très importante d'inégalités transport-entropie.
La mesure gaussienne dans $\mathbb{R}^n$ muni de la distance euclidienne vérife $T_2(2)$.
L'inégalité T_1 est un peu plus facile à manipuler et les deux inégalités entrainent des propriétés intéressantes de concentration de la mesure.

3.2 Inégalités transport-entropie libres

3.2.1 Inégalité T_2 libre pour la loi semi-circulaire

Depuis l'essor de la théorie des matrices aléatoires, la loi semi-circulaire, notée σ, dont on rappelle qu'elle a pour densité sur $\mathbb{R}$, $\frac{1}{2\pi}\sqrt{4 - x^2}\mathbf{1}_{[-2,2]}(x)dx$, est considérée comme l'analogue en probabilités libres de la gaussienne en probabilités classiques. A nouveau, comme dans le chapitre précédent, nous invitons le lecteur non familier des probabilités libres à se reporter à l'appendice au présent mémoire.

Sans rentrer trop dans les détails de la théorie, mentionnons la principale raison pour laquelle la loi semi-circulaire est considérée comme l'analogue libre de la gaussienne : si $(a_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires libres, de moyenne nulle et de variance 1 alors la suite de variables aléatoires $\frac{a_1+\ldots+a_n}{\sqrt{n}}$ converge en distribution vers σ. Il s'agit évidemment d'une version du théorème central limite adapté au contexte libre.

Il est donc naturel de se demander si σ vérifie une inégalité transport-entropie similaire à T_1 ou T_2. Une réponse positive a été apportée par Biane et Voiculescu dans [BV01] et leur résultat est donné par le théorème suivant :

Théorème 3.2.1. *(Biane, Voiculescu)*
Pour toute mesure ν à support compact[1],

$$W_2^2(\nu,\sigma) \leq 2\Sigma_\sigma(\nu),$$

avec

$$\Sigma_\sigma(\nu) = \frac{1}{2}\int x^2 d\nu(x) - \iint \log|x-y|d\nu(x)d\nu(y) - \frac{3}{4}. \tag{3.2.1}$$

La constante $\frac{3}{4}$ est choisie de sorte que $\Sigma_\sigma(\sigma) = 0$.

La quantité $\Sigma_\sigma(\nu)$ est appelée entropie libre par rapport à σ de la mesure ν. Nous allons maintenant donner quelques éléments sur cette quantité, qui est l'analogue libre de l'entropie relative $H(\nu|\sigma)$.

En effet, pour définir l'entropie libre, Voiculescu (cf par exemple [Voi02]) est reparti du point de vue de Boltzmann sur l'entropie classique, c'est-à-dire le point de vue micro-états, symbolisé par la fameuse formule $S = k\log W$. L'idée est la suivante : pour une variable aléatoire discrète prenant les valeurs $\{1,\dots,n\}$ avec les probabilités respectives $p_1,\dots,p_n$, on regarde le nombre $\Gamma(p_1,\dots,p_n;\varepsilon,N)$ de micro-états $f : \{1,\dots,n\} \mapsto \{1,\dots,N\}$ approchant la distribution discrète, i.e. tels que

$$\left|\frac{|f^{(-1)}(j)|}{N} - p_j\right| < \varepsilon,$$

où $|f^{(-1)}(j)|$ désigne le cardinal de la pré-image de j par f.
L'entropie est alors donnée par

$$\lim_{\varepsilon\downarrow 0}\lim_{N\to\infty}\frac{1}{N}\log\Gamma(p_1,\dots,p_n;\varepsilon,N).$$

Pour l'entropie libre, les micro-états seront matriciels. Si $(X_1,\dots,X_n)$ est un n-uplet d'éléments autoadjoints d'un espace de probabilités non commutatif, on note $\Gamma_R(X_1,\dots,X_n;m,k,\varepsilon)$ l'ensemble des n-uplets de matrices $(A_1,\dots,A_n)$ autoadjointes de taille $k\times k$ dont la norme est bornée par R (dit paramètre de cut-off) et tels que les moments jusqu'à l'ordre m approchent ceux de $(X_1,\dots,X_n)$, c'est-à-dire tels que

$$\left|\tau(X_{i_1}\dots X_{i_p}) - \frac{1}{k}\operatorname{Tr}(X_{i_1}\dots X_{i_p})\right| < \varepsilon, \forall 1 \leq p \leq m, (i_1,\dots,i_p) \in \{1,\dots,n\}^p.$$

Si γ désigne la mesure gaussienne sur l'ensemble des matrices autoadjointes de taille $k\times k$, l'entropie libre $\Sigma_\sigma(X_1,\dots,X_n)$ est définie de la façon suivante :

$$\Sigma_\sigma(X_1,\dots,X_n) := -\sup_{R>0}\inf_{m\in\mathbb{N}}\inf_{\varepsilon>0}\limsup_{N\to\infty}\left(\frac{1}{k^2}\log\gamma\left(\Gamma_R(X_1,\dots,X_n;m,k,\varepsilon)\right) + \frac{n}{2}\log k\right)$$

Dans le cas $n=1$, il est possible de donner une expression explicite de Σ_σ, donnée par (3.2.1).

Une autre raison pour laquelle la loi semicirculaire est considérée comme l'analogue de la gaussienne en probabilités libre est le fait qu'elle maximise l'entropie libre à variance fixée, comme la gaussienne maximise l'entropie classique à variance fixée.

1. On considère les mesures à support compact uniquement. Ceci n'est pas une restriction très importante, on pourrait obtenir l'inégalité pour tout ν, quitte à perdre sur la constante. De plus, en probabilités libres, on considère surtout des variables aléatoires bornées, dont la loi est donc à support compact.

Nous ne reproduisons pas ici la prevu initiale de Biane et Voiculescu qui s'appuie sur des idées dans l'esprit de l'article d'Otto et Villani [OV00], qui a été crucial dans le domaine et qui montre que les mesures vérifiant une inégalité de Sobolev logarithmique (avec une constante C) vérifient aussi une inégalité $T_2(C)$: on voit l'ensemble $\mathcal{P}(\mathcal{X})$ des mesures de probabilités sur $\mathcal{X}$ (polonais) comme une variété riemanienne munie de la distance W_2 et on travaille avec des diffusions (ici libres) sur cette variété. Nous donnerons plus loin une preuve d'une généralisation de ce résultat qui s'appuie sur des approximations matricielles des mesures mises en jeu.

3.2.2 Inégalité T_2 libre pour des mesures associées à des potentiels strictement convexes

Dans la littérature des inégalités transport-entropie classiques, beaucoup d'intérêt a été porté aux mesures dites log-concaves, c'est-à-dire les mesures de probabilités sur $\mathbb{R}$ dont la densité est de la forme $x \mapsto e^{-V(x)}$, avec V un potentiel strictement convexe.

Le fait que de telles mesures vérifient T_2 peut être vu comme une conséquence du résultat d'Otto et Villani mentionné plus haut.

L'analogue côté matrices aléatoires de ces mesures est donné par ce que l'on appelle communément les modèles unitairement invariants, que nous présentons maintenant. Nous avons déjà présenté dans ce mémoire le GUE, qui est sans doute l'ensemble de matrices aléatoires qui a été le plus étudié. Nous l'avons défini en donnant la loi jointe de ses entrées, essentiellement indépendantes identiquement distribuées gaussiennes au-dessus de la diagonale. Avec ce point de vue, la généralisation naturelle du GUE est de considérer des entrées indépendantes identiquement distribuées de loi autre que gaussienne. Ces matrices dites de Wigner ont suscité récemment de très nombreux travaux autour des questions d'universalité.

Mais on peut aussi considérer que le GUE est la mesure, sur l'ensemble $\mathcal{H}_N$ des matrices hermtiennes de taille N, de densité proportionnelle à $e^{-\frac{N}{2}\operatorname{Tr} X^2}$ par rapport à la mesure de Lebesgue sur $\mathcal{H}_N$. Avec ce point de vue, la généralisation naturelle du GUE est un ensemble unitairement invariant c'est-à-dire une mesure de densité $e^{-N \operatorname{Tr} V(X)}$ par rapport à la mesure de Lebesgue sur $\mathcal{H}_N$, pour un potentiel V continu tel que $\liminf_{|x|\to\infty} \frac{V(x)}{x^2} > 0$. En particulier, les entrées d'une matrice suivant une telle loi ne sont plus indépendantes, sauf si le potentiel V est proportionnel à x^2.

Pour un potentiel V vérifiant cette hypothèse de croissance, on notera

$$\mathbb{P}_V^N := \frac{1}{Z_V^N} e^{-N \operatorname{Tr} V(X)} dX,$$

avec dX la mesure de Lebesgue sur $\mathcal{H}_N$ et $Z_N^V := \int_{\mathcal{H}_N} e^{-N \operatorname{Tr} V(X)} dX$ la constante de normalisation.

Si la matrice X_N suit la loi $\mathbb{P}_V^N$, la loi jointe de ses valeurs propres est proportionnelle à

$$\prod_{i<j} |\lambda_i - \lambda_j|^2 e^{-N \sum_{i=1}^N V(\lambda_i)} \prod_{i=1}^N d\lambda_i. \tag{3.2.2}$$

On peut montrer que sous $\mathbb{P}_V^N$, la mesure empirique des valeurs propres converge quand N tend vers l'infini vers une mesure d'équilibre, notée μ_V (la convergence a lieu presque sûrement au sens de la convergence faible des mesures). cette convergence suffirait à définir la mesure μ_V mais il est intéressant et utile pour la suite de définir la mesure μ_V dans des termes de théorie du potentiel. Nous suivons ici l'ouvrage [ST97].

Théorème 3.2.2. *Soit V une fonction continue satisfaisant $\liminf_{|x|\to\infty} \frac{V(x)}{x^2} > 0$. Pour $\mu \in \mathcal{P}(\mathbb{R})$, on définit*

$$J_V(\mu) = \int_{\mathbb{R}} V(x) d\mu(x) - \iint_{\mathbb{R}^2} \ln|x-y| d\mu(x) d\mu(y)$$

avec la convention $J_V(\mu) = +\infty$ *si* $\int V d\mu = +\infty$. *Alors* $c_V = \inf_{\nu \in \mathcal{P}(\mathbb{R})} J_V(\nu)$ *est finie et le minimum de* J_V *est atteint en une unique mesure* μ_V *appelée mesure d'équilibre, qui est à support compact. On pose* $\Sigma_V = J_V - c_V$.

De plus, si on définit le potentiel logarithmique μ_V *comme*

$$U_{\mu_V}(x) = -\int \ln|x-y| d\mu_V(y),$$

pour tout $x \in \mathbb{C}$ *alors* U_{μ_V} *est fini et continu sur* $\mathbb{C}$ *et* μ_V *est l'unique mesure de probabilté sur* $\mathbb{R}$ *telle qu'il existe une constante* C_V *de sorte que*

$$\begin{aligned} -2U_{\mu_V}(x) + C_V \leq V(x) \quad & \textit{pour tout } x \textit{ dans } \mathbb{C}. \\ -2U_{\mu_V}(x) + C_V = V(x) \quad & \textit{pour tout } x \textit{ dans le support de } \mu_V \end{aligned}$$

C_V *est lié à* c_V *par la formule* $C_V = 2c_V - \int V(x) d\mu_V(x)$.

La condition de croissance $\liminf_{|x|\to\infty} \frac{V(x)}{x^2} > 0$ assure notamment l'existence de la mesure d'équilibre μ_V (ainsi que la finitude de la constante de normalisation $Z_N^V := \int_{\mathcal{H}_N} e^{-N \operatorname{Tr} V(X)} dX$). Elle n'est pas nécessaire pour cela mais comme notre résultat est montré sous cette hypothèse, nous nous y tenons.

On peut considérer que les mesures μ_V sont les analogues des mesures log-concaves.

Dans [HPU04], Hiai, Petz et Ueda se sont posé la question de la généralisation du résultat de Biane et Voiculescu (Théorème 3.2.1) aux mesures d'équilibre μ_V, du moins lorsque le potentiel V est strictement convexe. Ils obtiennent le résultat suivant :

Théorème 3.2.3. *(Hiai, Petz, Ueda)*
Si V *est strictement convexe, il existe une constante* $A_V > 0$ *(qui ne dépend que du potentiel* V*) telle que pour tout* $\nu \in \mathcal{P}(\mathcal{X})$,

$$W_2^2(\nu, \mu_V) \leq A_V \Sigma_V(\nu).$$

Pour simplifier la lecture, nous présentons les grandes lignes de la preuve de ce résultat pour $V(x) = \frac{1}{2}x^2$, avec donc $\mu_V = \sigma$ et nous indiquerons à la fin de celle-ci les changements à effectuer pour la généraliser à V strictement convexe, en soulignant où intervient la stricte convexité du potentiel et pourquoi une telle stratégie paraît difficile à étendre au-delà du cas convexe.

Leur preuve, contrairement à celle de Biane et Voiculescu qui travaillent dans les espaces de probabilités limites, s'appuie sur des approximations matricielles des mesures données. On suit la présentation de [GL10].

Par le théorème de Wigner, on sait que $\mathbb{P}_N^{\frac{1}{2}x^2}$ est une loi sur $\mathcal{H}_N$ telle que la mesure spectrale empirique des matrices distribuées selon $\mathbb{P}_N^{\frac{1}{2}x^2}$ converge presque sûrement vers σ. En fait, on peut vérifier que si ν est à support compact, un procédé systématique permet de trouver une mesure sur $\mathcal{H}_N$, que nous appellerons $\mathbb{P}_{N,\nu}$ telle que la mesure spectrale empirique des matrices distribuées selon $\mathbb{P}_{N,\nu}$ converge presque sûrement vers ν et telle que

$$\frac{1}{N^2} H(\mathbb{P}_{N,\nu} | \mathbb{P}_N^{\frac{1}{2}x^2}) \xrightarrow[N\to\infty]{} \Sigma_\sigma(\nu).$$

On construit pour cela un potentiel Q_ν tel que $\mu_{Q_\nu} = \nu$ et Q_ν coincide avec $\frac{1}{2}x^2$ en dehors d'un compact.

Si P_1 et P_2 sont deux mesures de probabilités sur $\mathcal{H}_N$, on considère le transport de l'une à l'autre :

$$\mathcal{T}_2(P_1, P_2) := \inf_{X \sim P_1, Y \sim P_2} \mathbb{E}\left(\sum_{i,j} |X_{ij} - Y_{ij}|^2\right),$$

qui n'est autre que le transport optimal associé à la distance de Frobenius $\|\cdot\|_F$, avec

$$\|X-Y\|_F = \sum_{i,j} |X_{ij}-Y_{ij}|^2.$$

Si on considère $X_{N,\nu}$ et $X_{N,\sigma}$ deux variables aléatoires de lois respectives $\mathbb{P}_{N,\nu}$ et $\mathbb{P}_N^{\frac{1}{2}x^2}$ telles que le couple $(X_{N,\nu}, X_{N,\sigma})$ constitue un couplage optimal entre les deux mesures, on a

$$\begin{aligned} \mathcal{T}_2(\mathbb{P}_{N,\nu}, \mathbb{P}_N^{\frac{1}{2}x^2}) &= \mathbb{E}\left(\|X_{N,\nu}-X_{N,\sigma}\|_F^2\right) \\ &\geq \mathbb{E}\left(\sum_i |\lambda_i(X_{N,\nu})-\lambda_i(X_{N,\nu})|^2\right), \end{aligned}$$

où $\lambda_i(X)$ est la i-ème valeur propre de X (les valeurs propres étant rangées par ordre croissant) et où on a utilisé les inégalités de Wielandt (cf par exemple [HJ90]).

Si on pose $R_N := \frac{1}{N}\sum_{i=1}^N \delta_{(\lambda_i(X_{N,\nu}),\lambda_i(X_{N,\sigma}))}$ et ν_N et σ_N respectivement les première et seconde marginales de la mesure déterministe $\mathbb{E}(R_N)$, on a :

$$\begin{aligned} \mathbb{E}\left(\sum_i |\lambda_i(X_{N,\nu})-\lambda_i(X_{N,\nu})|^2\right) &= N\mathbb{E}\left(\int |x-y|^2 dR_N(x,y)\right) \\ &\geq N\mathcal{T}_2(\nu_N,\sigma_N), \end{aligned}$$

où $\mathcal{T}_2$ désigne cette fois le transport associé au coût x^2 dans $\mathbb{R}$.

Par ailleurs, comme $\mathbb{P}_N^{\frac{1}{2}x^2}$ est un produit de gaussiennes indépendantes de variance $\frac{1}{N}$ qui chacune vérifie donc l'inégalité $T_2(2/N)$, par propriété de tensorisation de cette inégalité, on a

$$\mathcal{T}_2(\mathbb{P}_{N,\nu}, \mathbb{P}_N^{\frac{1}{2}x^2}) \leq \frac{2}{N} H(\mathbb{P}_{N,\nu}|\mathbb{P}_N^{\frac{1}{2}x^2}).$$

Finalement, on a

$$\mathcal{T}_2(\nu_N,\sigma_N) \leq \frac{2}{N^2} H(\mathbb{P}_{N,\nu}|\mathbb{P}_N^{\frac{1}{2}x^2}).$$

Or le théorème de Wigner donne aussi la convergence de σ_N vers σ. On a aussi la convergence de ν_N vers ν. Donc par continuité de la distance W_2, le terme de gauche converge vers $W_2^2(\nu,\sigma)$.

Par ailleurs on rappelle qu'on a choisi $\mathbb{P}_{N,\nu}$ de sorte que $\frac{1}{N^2}H(\mathbb{P}_{N,\nu}|\mathbb{P}_N^{\frac{1}{2}x^2}) \xrightarrow[N\to\infty]{} \Sigma_\sigma(\nu)$. On obtient donc une nouvelle preuve du résultat de Biane et Voiculescu.

La généralisation à V strictement convexe suit la même preuve avec $\mathbb{P}_N^{\frac{1}{2}x^2}$ remplacée par $\mathbb{P}_N^V$. L'argument principal est que si V est strictement convexe, $\mathbb{P}_N^V$ vérifie encore une inégalité $T_2(A_V/N)$ pour une certaine constante A_V (qui ne dépend de V qu'à travers sa dérivée seconde). Ce dernier point n'est plus vrai lorsque le potentiel V n'est plus convexe, ce qui empêche la généralisation au-delà du cas convexe.

Dans [A10], nous nous sommes posé la question du type d'inégalité que l'on pourrait obtenir au-delà du cas convexe.

3.2.3 Inégalité T_1 libre associée à des potentiels non convexes

Notre résultat principal peut s'exprimer de la façon suivante :

Théorème 3.2.4. *Pour V un potentiel qui vérifie* $\liminf_{|x|\to\infty} \frac{V(x)}{x^2} > 0, \exists B_V > 0$ *(qui ne dépend que de V) tel que pour toute mesure de probabilités ν,*

$$W_1^2(\nu, \mu_V) \leq B_V \Sigma_V(\nu).$$

On rappelle que l'entropie libre relativement au potentiel V est définie ci-dessus dans le Théorème 3.2.2.

Avant d'aborder les grandes lignes de la preuve, faisons quelques remarques sur le résultat. Tout d'abord l'hypothèse $\liminf_{|x|\to\infty} \frac{V(x)}{x^2} > 0$, sur le potentiel V est proche de l'optimalité puisque le résultat est trivialement faux si V est négligeable devant x^2. En effet, dans ce cas, si on choisit ν_n la loi uniforme sur $[n, n+1]$, $W_1^2(\nu_n, \mu_V)^2$ est équivalent à n^2 quand n tend vers l'infini alors que $\Sigma_V(\nu_n)$ croît comme $\nu_n(V)$, qui serait alors sous-quadratique.

On peut aussi remarquer que nous avons cherché à montrer une inégalité de type T_1 plutôt que T_2. La raison n'est pas seulement que T_1 est plus facile à montrer mais aussi que si V est non-convexe, μ_V n'a plus nécessairement un support connexe donc ne satisfait pas une inégalité de Poincaré. Or l'inégalité T_2 classique implique celle de Poincaré donc μ_V ne satisfait pas une inégalité T_2 classique - il est alors difficile d'espérer une inégalité T_2 libre.

Enfin nous tenons à souligner que la classe de mesures μ_V est en fait relativement générale, elle contient une grande partie des mesures à support compact. Pour mieux le comprendre, nous donnons une reformulation de notre théorème qui ne fait pas intervenir le potentiel V.

Théorème 3.2.5. *Soit μ une mesure de probabilités sur $\mathbb{R}$ à support compact telle que son potentiel logarithmique U_μ existe et est continu sur $\mathbb{C}$. Alors il existe $C > 0$ tel que pour tout ν,*

$$W_1^2(\nu, \mu) \leq C\Sigma_\mu(\nu),$$

avec

$$\Sigma_\mu(\nu) := \begin{cases} -\iint_{\mathbb{R}^2} \log|x-y| d(\nu-\mu)(x) d(\nu-\mu)(y), & \textit{si le support de } \nu \textit{ est inclus dans celui de } \mu, \\ +\infty, & \textit{sinon.} \end{cases}$$

Cette inégalité est légèrement moins bonne que celle du théorème 3.2.4 puisque $\Sigma_V(\nu) \leq \Sigma_{\mu_V}(\nu)$, avec égalité pour les mesures ν dont le support est inclus dans celui de μ_V.

Nous allons maintenant détailler la preuve d'une "version locale" de notre résultat.

Soit f une fonction u-Lipschitz et bornée, on a

$$\mu_V(f) - \nu(f) - \Sigma_V(\nu) \leq \sup_{\pi \in \mathcal{P}(\mathbb{R})} \{\mu_V(f) - \pi(f) - \Sigma_V(\pi)\}.$$

On va maintenant travailler sur le terme de droite

$$\begin{aligned} \mu_V(f) - \pi(f) - \Sigma_V(\pi) &= \mu_V(f) - \pi(f) - \pi(V) + \iint \log|x-y| d\pi(x) d\pi(y) + J_V(\mu_V) \\ &= J_V(\mu_V) + \mu_V(f) - J_{V+f}(\pi) \\ &\leq J_V(\mu_{V+f}) - J_{V+f}(\mu_{V+f}) + \mu_V(f) \\ &= \mu_V(f) - \mu_{V+f}(f) \leq u W_1(\mu_{V+f}, \mu_V). \end{aligned}$$

Entre la deuxième et la troisième ligne, on a utilisé le fait que J_V atteint son minimum en μ_V.

On est donc ramenés à une question de théorie du potentiel : comment varie la mesure d'équilibre sous une perturbation lipschitzienne du potentiel ?

Par des arguments analytiques que nous ne détaillons pas ici, on obtient le contrôle suivant :

Proposition 3.2.6. *Pour tout $L > 0$, il existe une constante K_L telle que pour tous V, W satisfaisant les conditions du Théorème 3.2.4, si μ_V et μ_W sont des mesures de probabilités sur $[-L; L]$ alors*

$$W_1(\mu_V, \mu_W) \leq K_L \, osc \, (V - W).$$

avec osc $(f) = \sup_{\mathbb{R}} f - \inf_{\mathbb{R}} f$.

Or, si f est $u-$Lipschitz sur le compact $[-L; L]$, on a osc $(f) \leq 2uL$. On obtient la majoration suivante

$$\mu_V(f) - \nu(f) - \Sigma_V(\nu) \leq 2u^2 L K_L.$$

Si on écrit $f = u\varphi$ avec φ 1-lipschitzienne,

$$u(\mu_V(\varphi) - \nu(\varphi)) - \Sigma_V(\nu) \leq 2u^2 L K_L,$$

ce qui donne en optimisant sur u et sur φ la majoration attendue

$$W_1(\mu_V, \nu)^2 \leq 4 L K_L \Sigma_V(\nu),$$

qui peut être considérée comme une version locale de notre résultat, au sens où la constante LK_L dépend de la taille du support des mesures considérées.

Il reste ensuite quelques points techniques pour s'affranchir de la dépendance de la constante en la taille du support. Nous ne détaillons pas ici cette partie de la preuve, qui est un peu technique mais nous tenons à mentionner que c'est dans la démonstration de ce point qu'intervient l'hypothèse sur la vitesse de croissance du potentiel V.

Résumé de la partie 3.2 :

Si V est un potentiel strictement convexe, la mesure d'équilibre correspondante μ_V vérifie un analogue libre de l'inégalité T_2. Pour des potentiels V plus généraux, nous montrons dans [A10] un analogue libre de l'inégalité T_1.

3.3 Applications aux propriétés de concentration pour les matrices aléatoires.

Comme nous l'avons vu dans la partie 3.1.2, les inégalités transport-entropie sont intimement liées à des propriétés de concentration de la mesure. Dans cette partie nous allons expliquer comment utiliser l'inégalité T_1 libre que nous avons obtenue pour montrer des inégalités de concentration pour des modèles de matrices assez généraux.

On définit une légère généralisation des ensembles β-invariants vus en (3.2.2) de la manière suivante : pour $\beta > 0$ et V un potentiel satisfaisant la condition $\liminf_{|x|\to\infty} \frac{V(x)}{x^2} > 0$, l'ensemble β-invariant associé au potentiel V est la famille de lois sur $\mathbb{R}^N$

$$\mathbb{P}^N_{V,\beta}(dx_1, \dots, dx_n) := \prod_{i<j} |x_i - x_j|^\beta \exp\left(-N \sum_{i=1}^N V(x_i)\right) \frac{\prod_{i=1}^N dx_i}{Z^N_{V,\beta}},$$

avec $Z^N_{V,\beta}$ la constante de normalisation.

Pour $\beta = 1, 2, 4$ cela correspond à la loi des valeurs propres d'un modèle de matrices.

Le principal résultat connu dans cette direction concerne les cas $\beta = 1, 2$ lorsque V est strictement convexe : la proposition suivante est tirée de [AGZ10] :

Proposition 3.3.1. *Si V est $\mathcal{C}^\infty$ avec $V'' \geq \kappa > 0$ et V' à croissance polynomiale à l'infini, alors pour tout $\theta \geq 0$ et f 1-lipschitzienne,*

$$\mathbb{P}^N_{V,\beta}\left(\left|\frac{1}{N}\operatorname{Tr} f - \int \frac{1}{N}\operatorname{Tr} f d\mathbb{P}^N_{V,\beta}\right| > \theta\right) < e^{-N^2\frac{\kappa\theta^2}{2}}.$$

Notre motivation est d'obtenir le même type de résultats pour V non nécessairement convexe. On rappelle encore une fois que très peu de résultats existent dans ce cas et que l'une des principales difficultés est que la mesure limite peut avoir un support non connexe. Notre résultat de concentration nécessite des hypothèses sur le potentiel V un peu plus restrictives que pour le résultat limite.

Proposition 3.3.2. *Soit V localement Lipschitz, satisfaisant l'hypothèse du Théorème 3.2.4, différentiable à l'infini et tel qu'il existe $\alpha > 0$ et $d \geq 2$ tel que $|V'(x)| \sim_{|x|\to\infty} \alpha|x|^{d-1}$. On suppose aussi que la mesure d'équilibre $\mu_{\frac{2V}{\beta}}$ a une entropie classique finie [2].*

Alors il existe $u, v > 0$ tels que pour tout $\theta > v\sqrt{\frac{\ln(1+N)}{N}}$,

$$\mathbb{P}^N_{V,\beta}\left(W_1(\hat{\mu}_N, \mu_{\frac{2V}{\beta}}) \geq \theta\right) \leq e^{-uN^2\theta^2}.$$

La force de notre résultat est qu'il est valable pour des β généraux et en dehors de l'hypothèse de convexité pour V mais contrairement au résultat dans le cas convexe, nous n'obtenons pas l'inégalité pour tout $\theta > 0$.

La preuve de la Proposition 3.3.1 combine essentiellement deux arguments : le fait que si V est strictement convexe, la mesure log-concave associée $\frac{1}{Z}e^{-V}$ satisfait une inégalité de Sobolev logarithmique et le fait que si f est une fonction 1-lipschitzienne alors la fonction $X \mapsto f(\lambda_1, \dots, \lambda_n)$, où $(\lambda_1, \dots, \lambda_n)$ sont les valeurs propres de X l'est aussi.

Ici nous adoptons une méthode différente. Sans rentrer dans les détails de la preuve, nous indiquons en quelques lignes en quoi ce résultat de concentration est une conséquence de notre inégalité T_1 libre.

Obtenir une inégalité de concentration, c'est majorer la quantité

$$\mathbb{P}^N_{V,\beta}\left(W_1(\hat{\mu}_N, \mu_V) \geq \theta\right) = \frac{e^{-N^2 c_V}}{Z^N_{V,\beta}} \int_{\{W_1^2(\hat{\mu}_N,\mu_V)\geq\theta\}} e^{-N^2\tilde{\Sigma}_V(\hat{\mu}_N)} d\mathbb{P}^N_{V,\beta}(x_1, \dots, x_N),$$

avec

$$\tilde{\Sigma}_V(\mu) = \int V(x)d\mu(x) - \iint_{x\neq y} \log|x-y| d\mu(x)d\mu(y) - c_V,$$

pour $\mu \in \mathcal{P}(R)$ et $\hat{\mu}_N := \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$ la mesure empirique des x_i.

Pour obtenir la majoration annoncée, on voudrait une inégalité T_1 faisant intervenir $\tilde{\Sigma}_V$ au lieu de Σ_V, cette dernière étant infinie sur les mesures ponctuelles.

Une inégalité de ce type n'est pas vérifiée telle quelle mais nous obtenons une inégalité T_1 approchée, au prix d'une correction en $\frac{\ln N}{N}$ sous la forme suivante :

Proposition 3.3.3. *Soit V satisfaisant les conditions de la Proposition 3.3.2. Alors pour tout $\mathcal{K}$ compact de $\mathbb{R}$, tout $N \in \mathbb{N}^*$ et $(x_1, \dots, x_N) \in \mathcal{K}^N$,*

$$W_1^2(\widehat{\mu}_N, \mu_V) \leq B_V\widetilde{\Sigma}_V(\widehat{\mu}_N) + 3\frac{\|V\mathbf{1}_{\mathcal{K}_1}\|_{Lip} + B + \ln(N)}{N}$$

avec B_V une constante dépendant de V, B une constante universelle, $\mathcal{K}_u$ l'ensemble des réels à distance au plus u de $\mathcal{K}$ et $\|\cdot\|_{Lip}$ la norme Lipschitz.

2. D'après [DKM98], cette hypothèse est vérifiée dès que V est $\mathcal{C}^2$.

Cette inégalité approchée s'obtient en appliquant notre inégalité T_1 à une régularisation ν_N de la mesure empirique $\hat{\mu}_N$ pour laquelle on sait contrôler la distance entre $\Sigma_V(\nu_N)$ et $\tilde{\Sigma}_V(\hat{\mu}_N)$.

Résumé de la partie 3.3 :

A partir de notre inégalité T_1 libre, nous obtenons un résultat de concentration de la mesure de la forme

$$\mathbb{P}_{V,\beta}^N\left(W_1(\widehat{\mu}_N, \mu_{\frac{2V}{\beta}}) \geq \theta\right) \leq e^{-uN^2\theta^2},$$

pour θ grand devant $1/\sqrt{N}$, pour des potentiels V assez généraux, en particulier non nécessairement convexes.

Appendice sur les probabilités libres

Nous rassemblons ici les principales notions de la théorie des probabilités libres que nous utilisons dans le mémoire, en particulier dans le chapitre 2. Cet appendice ne vise pas à l'exhaustivité mais se veut un court résumé des notions utilisées. Nous renvoyons le lecteur par exemple à [Voi00] pour des développements plus complets.

Définitions de base

Les espaces de probabilités non commutatifs que nous allons considérer sont surtout les espaces dans lesquels vivent les limites de suites de matrices aléatoires mais nous allons aussi expliquer comment on peut voir les ensembles de matrices aléatoires de taille finie comme des espaces de ce type.

Définition 3.3.4. *Dans ce mémoire,* ***espace de probabilités non commutatif*** *sera à comprendre au sens suivant : il s'agit d'un couple $(\mathcal{A}, \varphi)$ où*

- *$\mathcal{A}$ est une algèbre de Banach*[3] *sur $\mathbb{C}$, unifère (i.e. portant une unité pour la multiplication) et munie d'une involution $x \mapsto x^*$ antilinéaire telle que $\forall x, y \in \mathcal{A}$, $(xy)^* = y^*x^*$*
- *φ est un état, i.e. une forme linéaire sur $\mathcal{A}$ telle que $\varphi(xy) = \varphi(yx)$, $\varphi(x^*) = \overline{\varphi(x)}$, qui est de plus positif, i.e. $\varphi(xx^*) \geq 0, \forall x, y \in \mathcal{A}$.*

Par exemple, si on considère une algèbre de variables aléatoires définies sur un espace de probabilités classique et à valeurs dans $\mathcal{M}_N(\mathbb{C})$ dont toutes les coordonnées possèdent des moments, muni de l'état $\mathbb{E} \circ \frac{1}{N} \operatorname{Tr}$, on obtient un espace de probabilités non commutatif.

On appelle **variable aléatoire non commutative** un élément de $\mathcal{A}$. Pour simplifier, on considèrera dans la suite seulement des variables aléatoires auto-adjointes.

Dans la suite I désignera un ensemble quelconque.

Définition 3.3.5. *Si $(a_i)_{i \in I}$ une famille de variables aléatoires non commutatives d'un même espace de probabilités $(\mathcal{A}, \varphi)$ alors la* ***distribution (ou loi)*** *de la famille $(a_i)_{i \in I}$ est la forme linéaire*

$$\begin{array}{rcl} \mu : \mathbb{C}\langle X_i, i \in I\rangle & \to & \mathbb{C} \\ P & \mapsto & \varphi(P(a_i, i \in I)), \end{array}$$

où $\mathbb{C}\langle X_i, i \in I\rangle$ est l'algèbre des polynômes à variables non commutatives indexées par I.

3. Cette condition sert à obtenir l'unicité de la distribution d'un élément autoadjoint en tant que mesure sur $\mathbb{R}$.

Remarquons que la distribution d'un seul élément autoadjoint a est donnée par la suite de ses moments $(\varphi(a^k))_{k\geq 0}$, qui est une suite positive (au sens où la matrice de Hankel associée est positive) donc la suite des moments d'une unique mesure de probabilités sur $\mathbb{R}$, notée μ_a. Dans ce cas, on appellera μ_a la distribution (spectrale) de a.

Par exemple, si l'espace de probabilités est $(\mathcal{M}_N(\mathbb{C}), \frac{1}{N}\operatorname{Tr})$ et a est hermitienne, alors μ_a est la mesure uniforme sur le spectre de a, soit $\frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i}$ avec $\{\lambda_1, \dots, \lambda_N\}$ les valeurs propres de a.
Si $\mathcal{A}$ est une algèbre de variables aléatoires définies sur un espace de probabilités classique à valeurs dans $\mathcal{M}_N(\mathbb{C})$ dont tous les coefficients possèdent des moments, muni de de l'état $\mathbb{E} \circ \frac{1}{N}\operatorname{Tr}$, alors pour X_N une matrice hermitienne aléatoire, μ_{X_N} est la mesure d'intensité de la mesure aléatoire $\frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i(X_N)}$, avec $\{\lambda_1(X_N), \dots, \lambda_N(X_N)\}$ les valeurs propres de X_N.

On explique maintenant ce que cela signifie pour une famille de suites de matrices aléatoires de converger en distribution vers une famille de variables aléatoires dans un espace de probabilités non commutatif.

On suppose que que pour tout $N \in \mathbb{N}$, $(X_N^{(i)})_{i\in I}$ est une famille de matrices aléatoires de taille $N \times N$. On dit que la famille **converge en distribution** vers une famille $(a_i)_{i\in I}$ d'éléments de $(\mathcal{A}, \varphi)$ si pour tout $P \in \mathbb{C}\langle X_i, i \in I\rangle$ la variable aléatoire classique donnée par $\frac{1}{N}\operatorname{Tr} P((X_N^{(i)})_{i\in I})$ converge en probabilité vers $\varphi(P((a_i)_{i\in I}))$.

Notion de liberté

Définition 3.3.6. *Si $\mathcal{A}$ est un espace de probabilités non commutatif, une famille $(\mathcal{A}_i)_{i\in I}$ de sous-algèbres de $\mathcal{A}$ est dite **libre** si pour tout $n \geq 1$, $\forall i_1, \dots i_n \in I$ tels que $i_1 \neq i_2, \dots, i_{n-1} \neq i_n$, $x_{i_j} \in \mathcal{A}_{i_j}$,*

$$\varphi(x_1) = \ldots = \varphi(x_n) = 0 \Rightarrow \varphi(x_1 \ldots x_n) = 0.$$

Une famille de variables aléatoires est dite libre si les sous-algèbres engendrées sont libres au sens précédent.

La notion de liberté est le pendant de celle d'indépendance. On sait que la loi jointe d'une famille de variables aléatoires indépendantes ne dépend que de la loi de chacune des variables. De la même façon, si on a une famille de variables libres, la règle énoncée ci-dessus permet de calculer la loi jointe de la famille connaissant les lois individuelles. Les combinatoires impliquées sont simplement différentes au point que deux variables aléatoires d'un espace de probabilités classique qui sont libres ne commutent pas (donc en particulier ne sont pas indépendantes), sauf si l'une des deux est une constante.

Pour comprendre l'intérêt des probabilités libres pour l'étude des suites de matrices aléatoires, un des points importants est la notion de **liberté asymptotique** : l'indépendance au niveau matriciel donne la liberté dans les espaces limites. Plus précisément, on a le résultat suivant [4], dû à Voiculescu :

Théorème 3.3.7. *Soit $(X_1^{(N)}, \dots, X_n^{(N)})_{N\in\mathbb{N}}$ une suite de $n-$uplets de matrices aléatoires hermitiennes ou unitaires de taille $N \times N$ telles que*
- *$\forall N \in \mathbb{N}, (X_1^{(N)}, \dots, X_n^{(N)})$ sont indépendantes,*
- *$\forall i \in \{1, \dots, n\}, N \in \mathbb{N}$, la loi de $X_i^{(N)}$ est invariante par conjugaison par les matrices unitaires,*
- *$\forall i \in \{1, \dots, n\}$, quand N tend vers l'infini, la mesure spectrale empirique de $X_i^{(N)}$ converge vers une loi déterministe à support compact.*

4. Ce résultat a été généralisé de nombreuses manières. On ne cherche pas ici à donner l'énoncé optimal mais simplement un énoncé simple qui donne l'esprit de la notion.

Alors la famille $(X_1^{(N)}, \ldots, X_n^{(N)})_{N\in\mathbb{N}}$ *est* ***asymptotiquement libre*** *au sens suivant : il existe une famille libre* $(a_i)_{i\in I}$ *d'éléments autoadjoints d'un espace de probabilités non commutatif telle que pour tout* $P \in \mathbb{C}\langle X_i, i \in I\rangle$*, la variable aléatoire* $\frac{1}{N}\operatorname{Tr} P((X_i^{(N)})_{i\in I})$ *converge en probabilité, lorsque* N *tend vers l'infini, vers le nombre* $\varphi(P((a_i)_{i\in I}))$.

Convolution libre

Dans le contexte classique, si X et Y sont deux variables indépendantes dont les lois μ_X et μ_Y sont à support compact, la loi de μ_{X+Y} est donnée par la convolution classique $\mu_X * \mu_Y$ et on sait que si $\hat{\mu}$ désigne la transformée de Fourier de la mesure μ, alors $\log \widehat{\mu_X * \mu_Y} = \log \widehat{\mu_X} + \log \widehat{\mu_Y}$. Le logarithme de la transformée de Fourier est additif par convolution classique.

Une conséquence de la définition de la liberté est que si a et b sont deux variables libres, on peut déterminer la loi de la variable $a+b$ et vérifier que cette loi ne dépend que des distributions de a et b. La loi ainsi définie, notée $\mu_a \boxplus \mu_b$ est la **convolution additive libre** de μ_a et μ_b.
Pour la convolution additive libre, il existe aussi une fonctionnelle qui la linéarise [5]. Il s'agit de la **R-transformée**. On peut en donner une définition combinatoire générale mais nous allons nous limiter aux mesures à support compact. Si μ est une mesure à support compact, sa transformée de Cauchy G_μ est donnée par

$$G_\mu(z) = \int_{\mathbb{R}} \frac{d\mu(x)}{z-x} \text{ pour } z \text{ hors du support de } \mu.$$

La fonction G_μ est inversible d'un voisinage de l'infini vers un voisinage de zéro et on définit R_μ sur un voisinage de l'origine par

$$R_\mu(z) = G_\mu^{-1}(z) - \frac{1}{z}.$$

On peut alors vérifier que pour μ, ν deux mesures à support compact sur $\mathbb{R}$, $R_{\mu\boxplus\nu} = R_\mu + R_\nu$.

5. Si a et b sont de plus des variables positives, on peut définir la convolution multiplicative de μ_a et μ_b comme la loi du produit ab. Il existe aussi une fonctionnelle qui linéarise la convolution multiplicative libre mais nous ne la détaillons pas ici.

Bibliographie

[AGZ10] Greg W. ANDERSON, Alice GUIONNET et Ofer ZEITOUNI : *An introduction to random matrices*, volume 118 de *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[BADG01] G. BEN AROUS, A. DEMBO et A. GUIONNET : Aging of spherical spin glasses. *Probab. Theory Related Fields*, 120(1):1–67, 2001.

[BAG97] Gérard BEN AROUS et Alice GUIONNET : Large deviations for Wigner's law and Voiculescu's non-commutative entropy. *Probab. Theory Related Fields*, 108(4):517–542, 1997.

[BBAP05] J. BAIK, G. BEN AROUS et S. PÉCHÉ : Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Prob.*, 33:1643–1697, 2005.

[BGR11] F. BENAYCH-GEORGES et R. RAO : The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Adv. Math.*, 227(1):494–521, 2011.

[BGV07] François BOLLEY, Arnaud GUILLIN et Cédric VILLANI : Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab. Theory Related Fields*, 137(3-4):541–593, 2007.

[Bia97] Philippe BIANE : Free Brownian motion, free stochastic calculus and random matrices. *In Free probability theory (Waterloo, ON, 1995)*, volume 12 de *Fields Inst. Commun.*, pages 1–19. Amer. Math. Soc., Providence, RI, 1997.

[BV01] Philippe BIANE et Dan-Virgil VOICULESCU : A free probability analogue of the Wasserstein metric on the trace-state space. *Geom. Funct. Anal.*, 11(6):1125–1138, 2001.

[BY08] Z. D. BAI et J.-F. YAO : Central limit theorems for eigenvalues in a spiked population model. *Ann. I.H.P.-Prob. et Stat.*, 44(3):447–474, 2008.

[CDMF] M. CAPITAINE, C. DONATI-MARTIN et D. FÉRAL : Central limit theorems for eigenvalues of deformations of Wigner matrices. *Preprint*.

[CDMF09] M. CAPITAINE, C. DONATI-MARTIN et D. FÉRAL : The largest eigenvalues of finite rank deformation of large Wigner matrices : convergence and nonuniversality of the fluctuations. *Ann. Probab.*, 37(1–47), 2009.

[CV] S. CHATTERJEE et S. R. S. VARADHAN : Large deviations for random matrices. *Preprint*.

[DE01] Persi DIACONIS et Steven N. EVANS : Linear functionals of eigenvalues of random matrices. *Trans. Amer. Math. Soc.*, 353(7):2615–2633 (electronic), 2001.

[Dei99] P. A. DEIFT : *Orthogonal polynomials and random matrices : a Riemann-Hilbert approach*, volume 3 de *Courant Lecture Notes in Mathematics*. New York University Courant Institute of Mathematical Sciences, New York, 1999.

[Dia03] Persi DIACONIS : Patterns in eigenvalues : the 70th Josiah Willard Gibbs lecture. *Bull. Amer. Math. Soc.*, 40, 2003.

[DKM98] P. DEIFT, T. KRIECHERBAUER et K. T.-R. MCLAUGHLIN : New results on the equilibrium measure for logarithmic potentials in the presence of an external field. *J. Approx. Theory*, 95(3):388–475, 1998.

[DZ10] Amir DEMBO et Ofer ZEITOUNI : *Large deviations techniques and applications*, volume 38 de *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.

[EK05] N. EL KAROUI : Recent results about the largest eigenvalue of random covariance matrices and statistical application. *Acta Phys. Polon. B*, 36(9):2681–2697, 2005.

[FP07] D. FÉRAL et S. PÉCHÉ : The largest eigenvalue of rank one deformation of large Wigner matrices. *Comm. Math. Phys.*, 272(185–228), 2007.

[FvdHK08] A. FEY, R. van der HOFSTAD et M. KLOK : Large deviations for eigenvalues of sample covariance matrices, with applications to mobile communication systems. *Adv. in Appl. Probab.*, 40:1048–1071, 2008.

[GL10] Nathaël GOZLAN et Christian LÉONARD : Transport inequalities. A survey. *Markov Process. Related Fields. To appear.*, http ://arxiv.org/abs/1003.3852, 2010.

[GZ02] A. GUIONNET et O. ZEITOUNI : Large deviations asymptotics for spherical integrals. *J. Funct. Anal.*, 188:461–515, 2002.

[HJ90] R. HORN et C. R. JOHNSON : *Matrix Analysis*. Cambridge University Press, 1990.

[HPU04] Fumio HIAI, Dénes PETZ et Yoshimichi UEDA : Free transportation cost inequalities via random matrix approximation. *Probab. Theory Related Fields*, 130(2):199–221, 2004.

[HR95] U. HELMKE et J. ROSENTHAL : Eigenvalue inequalities and schubert calculus. *Math. Nachr.*, 171:207–225, 1995.

[Joh98] Kurt JOHANSSON : On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.*, 91(1):151–204, 1998.

[Kly98] A. A. KLYACHKO : Stable vector bundles and Hermitian operators. *Selecta. Math.*, 4:419–445, 1998.

[KT01] A. KNUTSON et T. TAO : Honeycombs and sums of Hermitian matrices. *Notices Amer. Math. Soc.*, 48:175–186, 2001.

[OV00] Felix OTTO et Cédric VILLANI : Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.

[Péc06] S. PÉCHÉ : The largest eigenvalue of small rank perturbations of Hermitian random matrices. *Probab. Theory Related Fields*, 134:127–173, 2006.

[ST97] Edward B. SAFF et Vilmos TOTIK : *Logarithmic potentials with external fields*, volume 316 de *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1997. Appendix B by Thomas Bloom.

[Tal96] Michel TALAGRAND : Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.*, 6(3):587–600, 1996.

[Vil03] Cédric VILLANI : *Topics in optimal transportation*, volume 58 de *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[Vil09] Cédric VILLANI : *Optimal transport, old and new*, volume 338 de *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, 2009.

[Voi00] Dan VOICULESCU : Lectures on free probability theory. *In Lectures on probability theory and statistics (Saint-Flour, 1998)*, volume 1738 de *Lecture Notes in Math.*, pages 279–349. Springer, Berlin, 2000.

[Voi02] Dan VOICULESCU : Free entropy. *Bull. London Math. Soc.*, 34(3):257–278, 2002.

[Wey12] H. WEYL : Das asymptotische Verteilungsdesetz der Eigenwerte lineare partieller Differentialgleichungen. *Math. Ann.*, 71:441–479, 1912.