# THÈSE

**En vue de l'obtention du**
**DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE**

**Délivré par :**
Université Toulouse III Paul Sabatier (UT3 Paul Sabatier)

**Discipline ou spécialité :**
Domaine mathématiques – Mathématiques appliquées

---

**Présentée et soutenue par**
Sandrine Dallaporta
**le :** 20 décembre 2012

**Titre :**
Quelques aspects de l'étude quantitative de la fonction de comptage et des valeurs propres de matrices aléatoires

---

**École doctorale :**
Mathématiques Informatique Télécommunications (MITT)

**Unité de recherche :**
UMR 5219

**Directeur de thèse :**
Michel Ledoux

**Rapporteurs :**
Djalil Chafaï
Sandrine Péché

**Autres membres du jury :**
Charles Bordenave
Mireille Capitaine
Catherine Donati-Martin

*À Guilhem.*

# Remerciements

Mes premiers remerciements vont à Michel Ledoux, qui a accepté de diriger cette thèse. Merci de m'avoir ouvert une porte sur le monde de la recherche et des matrices aléatoires. Merci d'avoir été si disponible et patient, et d'avoir répondu à toutes mes questions.

Je tiens à remercier tout particulièrement Djalil Chafaï et Sandrine Péché pour avoir accepté d'être rapporteurs et pour avoir relu mon manuscrit avec attention. Je n'oublie pas la gentillesse de Djalil Chafaï, en particulier lors d'un workshop à Marne-la-Vallée en première année de thèse dans lequel je me sentais un peu perdue... Catherine Donati-Martin a accepté de présider ce jury, je l'en remercie vivement. Mireille Capitaine et Charles Bordenave me font l'honneur d'être membres du jury, c'est un grand plaisir pour moi et je les en remercie. Merci également à Mireille pour l'organisation du groupe de travail de matrices aléatoires et pour sa gentillesse à toute épreuve.

Ces trois années à l'institut de mathématiques de Toulouse ont été très agréables et riches en rencontres. Je voudrais adresser un grand remerciement au personnel administratif, en particulier à Agnès Requis et Françoise Michel, pour leur grande efficacité et leur gentillesse. Un merci s'adresse tout particulièrement à Marie-Laure Ausset et Sébastien Déjean, organisateurs des « cafés de l'IMT ». L'institut de mathématiques de Toulouse m'a offert un cadre très appréciable (et apprécié !) pendant ces trois années. Merci à tous les membres de l'institut pour leur accueil durant cette thèse.

Les doctorants ont droit à un mot particulier. Sans Alice, ces trois années auraient bien différentes... Merci d'avoir partagé tous ces goûters, ces doutes, ces rêves et d'avoir été sensible à mes idées farfelues ! Et merci à Gilles de supporter que ça déteigne sur son « chez lui ». Merci aux membres successifs du bureau 203 : Maxime, Julie et Daniel, même si ce dernier a refusé avec constance toutes nos propositions de goûter... Merci à Virginie et Jacques pour toutes ces pauses repas et ces soirées vécues ensemble. Merci à Mathieu, Yohann, Emmanuel, Thibault, Claire, Guillaume, Benjamin, Thomas, Sébastien, Steven, Julien, Léo...

De nombreuses amitiés hors des murs de l'institut ont été l'occasion d'apprendre à tenter de donner une réponse à la question fatidique de l'utilité des mathéma-

tiques... Un immense merci à Audrey pour ce temps passé au téléphone, chez l'une ou chez l'autre, pour tous ces moments partagés et son soutien sans faille qui réconforte dans les moments de doute. Merci également à Seb, Élisabeth, Sandrine, Lucie, Marie-Pascale et Réda.

Ce goût pour les mathématiques n'est pas venu par hasard. Sans Jacques Sagné et ses cours de maths en première et terminale, je n'aurais peut-être pas choisi cette voie. Merci à vous et merci à Aurélie, voisine de classe et amie de longue date, pour nos bavardages incessants pendant ces fameux cours... Merci à Alain Soyeur et Michel Gonnord, mes profs de prépa. Ma scolarité à l'ENS Lyon a été déterminante dans mes choix mais surtout très importante par les amitiés nouées. Merci à Aline et Anne-Sandrine pour ces chouettes moments, ainsi qu'à Xavier, Vincent, Romain, Cécile, Michaël, Pierre, Oscar, Anne.

Enfin, je remercie du fond du cœur ma famille. Mes parents, sans qui je ne serais pas là aujourd'hui et qui m'ont toujours soutenue et encouragée, mes frères et sœurs qui volent désormais de leurs propres ailes alors que je les imagine encore petits, mes grands-parents, oncles, tantes, cousins, cousines toujours présents. Merci à ma belle-famille pour son accueil et son soutien.

Le dernier remerciement, sans doute le plus complexe à écrire, est pour Amic. Difficile d'imaginer un tel soutien au quotidien, tout particulièrement sur cette fin de thèse ! Merci de croire en moi et de me pousser à aller de l'avant. Merci d'être là, tout simplement.

# Introduction

Les matrices aléatoires ont été introduites par Wishart dans les années 20, dans le cadre des statistiques multivariées. Dans les années 50, Wigner a utilisé un autre type de matrices aléatoires, appelées matrices de Wigner, dans le but de modéliser les Hamiltoniens qui régissent les évolutions des systèmes en physique statistique. Ces matrices servant de modèles de dimension finie pour des objets de dimension infinie, Wigner s'est intéressé aux propriétés asymptotiques de ces matrices, notamment au comportement de la mesure spectrale empirique ainsi qu'au comportement individuel de chaque valeur propre. Devenues populaires grâce au phénomène d'universalité, i.e. le fait que ces propriétés asymptotiques ne dépendent pas de la loi des coefficients de la matrice, leur étude a intéressé de nombreux domaines comme la combinatoire, les systèmes intégrables, les probabilités libres, les télécommunications, le traitement du signal, l'analyse fonctionnelle géométrique et, plus tard, la finance. Pour certains de ces domaines, comme les télécommunications et le traitement du signal, les propriétés à taille de matrice fixée sont plus exploitables que les propriétés asymptotiques. Nous nous sommes intéressés à ce pan de l'étude des matrices aléatoires, par le biais de la fonction de comptage des valeurs propres, c'est-à-dire du nombre de valeurs propres d'une matrice présentes dans un intervalle $I$.

Les trois premiers chapitres constituent une introduction aux trois chapitres suivants. Le premier chapitre présente les modèles de matrices aléatoires que nous étudions dans cette thèse, les matrices de Wigner et de covariance, ainsi que les principaux résultats asymptotiques en lien avec la fonction de comptage et plus globalement les valeurs propres de ces matrices. Le deuxième chapitre concerne les principaux outils mis en jeu dans ces études asymptotiques. L'accent est mis sur les outils qui peuvent être utilisés dans le cadre de l'étude quantitative, notamment les résultats récents de Erdös *et al* d'une part et de Tao et Vu d'autre part, qui ont permis une avancée spectaculaire des études asymptotiques mais également non asymptotiques. Le troisième chapitre s'attache à décrire les enjeux de l'étude quantitative des matrices aléatoires, ainsi qu'à présenter certains résultats obtenus dans le cadre de cette étude, en lien avec les résultats aymptotiques précédemment décrits dans le chapitre 1. Enfin, ce chapitre nous permet de présenter les résultats obtenus dans cette thèse, ainsi que les grandes lignes des démonstrations, en articulation avec les outils et résultats précédents.

Les chapitres suivants rassemblent les contributions de cette thèse sous la forme d'articles indépendants. Dans le chapitre 4, en collaboration avec V. Vu, nous établissons un théorème central limite pour la fonction de comptage des valeurs propres de matrices de Wigner et de covariance et nous obtenons des estimées quantitatives sur l'espérance et la variance de cette fonction de comptage. Ces résultats se basent sur les résultats précédemment établis par Gustavsson [49] et Su [83] dans le cas de matrices gaussiennes et sont étendus à des familles plus générales de matrices de Wigner et de covariance par le biais de travaux récents de Erdös, Yau et Yin [35], Pillai et Yin [75] et de Tao et Vu [87, 88].

Dans le chapitre 5, nous établissons des bornes quantitatives sur la variance des valeurs propres de matrices de Wigner. En s'appuyant sur les propriétés de la fonction de comptage, nous montrons d'abord une inégalité de déviation pour les valeurs propres individuelles à l'intérieur du spectre dans le cas de matrices gaussiennes et nous en déduisons les bornes dans ce cas. Afin d'étendre ces bornes au cas des matrices de Wigner plus générales, nous combinons de nouveau les outils de Erdös, Yau et Yin [35] et de Tao et Vu [87]. Ce chapitre est conclu par l'annonce des résultats analogues pour les matrices de covariance, qui font l'objet du chapitre 6. Ce dernier chapitre présente en effet de manière indépendante le contexte, les résultats et les démonstrations des résultats précédents pour les matrices de covariance.

# Table des matières

# Chapitre 1

# Quelques aspects asymptotiques de l'étude des matrices aléatoires

Ce chapitre décrit les modèles de matrices aléatoires utilisés dans le cadre de cette thèse et présente certains résultats asymptotiques les concernant. La présentation est volontairement orientée, les résultats sont sélectionnés en articulation avec les travaux effectués durant la thèse. Certains des outils techniques évoqués dans ce premier chapitre, sélectionnés dans le même objectif, seront développés au chapitre 2. Ces résultats et outils, ainsi que bien d'autres peuvent être trouvés par exemple dans les ouvrages récents de Anderson, Guionnet et Zeitouni [1], de Bai et Silverstein [8] et de Pastur et Shcherbina [72].

Les matrices aléatoires sont apparues pour la première fois dans les années 20, introduites par le statisticien Wishart, qui cherchait à modéliser des tables de données aléatoires, dans le cadre des statistiques multivariées. Il s'est intéressé de ce fait aux propriétés des matrices de covariance empiriques. Dans les années 50, le physicien Wigner a introduit un autre type de matrices aléatoires, appelées aujourd'hui matrices de Wigner. Son objectif était alors de modéliser les Hamiltoniens (de dimension infinie) qui régissent les évolutions de systèmes en physique statistique par des matrices aléatoires (de dimension grande mais finie) respectant les mêmes symétries, c'est à dire des matrices hermitiennes. Wigner a conjecturé un phénomène d'universalité, à savoir que les propriétés spectrales asymptotiques des matrices aléatoires dépendent uniquement des propriétés de symétrie du modèle et non des lois des coefficients. Ce phénomène d'universalité a popularisé les matrices aléatoires, qui ont alors intéressé de nombreux domaines, comme la combinatoire, les systèmes intégrables, les probabilités libres, les télécommunications, le traitement du signal, l'analyse fonctionnelle géométrique et, plus tard, la finance (voir par exemple [1] et [8]). Les modèles de matrices aléatoires dont les coefficients suivent une loi gaussienne, dits « modèles gaussiens », se sont révélés plus faciles à étudier, grâce à leur propriété d'invariance par transformation unitaire. La loi jointe des valeurs propres est alors connue et des calculs explicites sont possibles, permettant une étude quasiment complète du spectre. Montrer que les propriétés spectrales sont universelles a été beaucoup plus compliqué. Les techniques utilisées dans le cas gaussien ne fonctionnent plus pour d'autres modèles et de nouvelles méthodes ont dû être mises en œuvre. Si l'universalité du comportement global du spectre, c'est-à-dire des quantités faisant intervenir toutes ou presque toutes les valeurs propres, a été rapidement établie, l'universalité du comportement local, c'est-à-dire des quantités concernant un petit nombre de valeurs propres, n'a été établie que très récemment dans une certaine généralité. Les premiers résultats datent du début des années 2000 avec notamment les travaux de Johansson et de Soshnikov. Des avancées spectaculaires ont eu lieu ces cinq dernières années, avec les travaux de Erdös, Schlein, Yau *et al* d'un côté et ceux de Tao et Vu *et al* de l'autre.

Dans ce chapitre, nous commençons par décrire les modèles de matrices aléatoires auxquels nous nous intéressons dans cette thèse, avant de présenter certains résultats asymptotiques concernant les propriétés spectrales de ces matrices. Les matrices de Wigner considérées ici seront définies de la manière suivante, avec un choix de normalisation des variances.

**Définition 1.1.** *Soit $N \in \mathbb{N}$. $M_N$ est une matrice $N \times N$ hermitienne de Wigner si :*

- *ses coefficients sont indépendants (à la condition de symétrie près),*
- *les parties réelles et imaginaires des coefficients sont indépendantes,*
- *ses coefficients ont une moyenne nulle et une variance égale à* $1$.

*Dans le cas des matrices de Wigner symétriques réelles, les coefficients diagonaux ont une variance égale à* $2$ *tandis que les coefficients non diagonaux ont une variance égale à* $1$.

Le théorème de Wigner dont nous parlerons plus loin dans ce chapitre fournit une loi limite pour le spectre de telles matrices. Le support de cette loi limite dépendant uniquement de la variance des coefficients non diagonaux, nous avons choisi de normaliser les matrices de telle sorte que cette variance soit égale à 1, tant pour des matrices symétriques réelles que pour des matrices hermitiennes complexes.

Plutôt que $M_N$, nous utiliserons fréquemment la matrice renormalisée $W_N$ définie par $W_N = \frac{1}{\sqrt{N}} M_N$. $M_N$ et $W_N$ étant des matrices hermitiennes (ou symétriques réelles), elles ont $N$ valeurs propres réelles que l'on note respectivement $\lambda_1(M_N) \leqslant \cdots \leqslant \lambda_N(M_N)$ et $\lambda_1(W_N) \leqslant \cdots \leqslant \lambda_N(W_N)$. Lorsqu'il n'y a pas de confusion possible, on note simplement $\lambda_1 \leqslant \cdots \leqslant \lambda_N$.

Dans le cas où les coefficients suivent une loi gaussienne, la matrice est dite appartenir à l'ensemble gaussien unitaire (GUE) lorsqu'elle est hermitienne complexe et à l'ensemble gaussien orthogonal (GOE) lorsqu'elle est symétrique réelle. Par convention de normalisation des variances, la densité de probabilité de la loi d'une matrice du GUE sur l'espace $\mathcal{H}_N$ des matrices hermitiennes de taille $N$ peut se réécrire de façon équivalente sous la forme d'une densité gaussienne

$$\mathbb{P}(dM) = \frac{1}{Z_N} \exp\Big( - \frac{1}{2} \operatorname{Tr} M^2 \Big) dM,$$

où $dM = \prod_{j=1}^N dM_{jj} \prod_{1 \leqslant j < k \leqslant N} d\Re(M_{jk}) d\Im(M_{jk})$ et $Z_N$ est une constante de normalisation. L'expression similaire dans le cas réel est de la forme, sur l'ensemble $\mathcal{S}_N$ des matrices symétriques réelles de taille $N$,

$$\mathbb{P}(dM) = \frac{1}{Z'_N} \exp\Big( - \frac{1}{4} \operatorname{Tr} M^2 \Big) dM,$$

où $dM = \prod_{j=1}^N dM_{jj} \prod_{1 \leqslant j < k \leqslant N} dM_{jk}$. De manière synthétique

$$\mathbb{P}_\beta(d_\beta M) = \frac{1}{Z_{N,\beta}} \exp\Big( - \frac{\beta}{4} \operatorname{Tr} M^2 \Big) d_\beta M, \tag{1.1}$$

avec $\beta = 1$ si la matrice est issue du GOE et $\beta = 2$ si elle est issue du GUE. Les propriétés de ces matrices sont détaillées dans les ouvrages [67], [1], [8] et [72]. On observe alors que la loi de la matrice est invariante par transformation unitaire. En effet, la densité de probabilité par rapport à $d_\beta M$ ne dépend que des valeurs propres de la matrice, à travers le terme $\operatorname{Tr} M^2$, et est donc invariante par transformation unitaire. De plus, $d_\beta M$ est également invariante par transformation unitaire. La loi d'une matrice du GUE ou du GOE est donc invariante par transformation unitaire. Par conséquent, deux matrices ayant les mêmes valeurs propres auront même probabilité d'apparaître lors du tirage d'une matrice du GUE ou du GOE. Il est alors possible de calculer la loi jointe des valeurs propres non ordonnées.

$$C_{N,\beta} \exp\Big( - \frac{\beta}{4} \sum_{j=1}^N x_j^2 \Big) \prod_{1 \leqslant j < k \leqslant N} |x_k - x_j|^\beta, \tag{1.2}$$

où $C_{N,\beta}$ est la constante de normalisation. C'est le seul modèle de matrices de Wigner qui vérifie cette propriété d'invariance. Il est en effet possible de montrer que, si une matrice de Wigner a une loi invariante par transformation unitaire, alors les coefficients de cette matrice suivent une loi gaussienne. À la normalisation des variances près, cette matrice appartient donc au GUE ou au GOE. Les techniques utilisées dans le cas gaussien reposent majoritairement sur cette loi jointe des valeurs propres. Elles sont donc inutilisables dans le cas non-gaussien.

Dans le cas du GUE, cette loi jointe montre une propriété très utile sur les valeurs propres. En effet, à partir de (1.2), il est aisé d'étudier la loi de $p$ valeurs propres non ordonnées, notée $\mathcal{P}_{p,N}$ pour $p \leqslant N$. Elle admet une densité $\rho_{p,N}$ de la forme suivante.

$$\rho_{p,N}(x_1, \ldots, x_p) = \frac{(N-p)!}{N!} \det \Big( K_N(x_j, x_k) \Big)_{j,k=1}^{p}, \tag{1.3}$$

où $K_N$ est le noyau

$$K_N(x, y) = \sum_{j=0}^{N-1} \psi_j(x) \psi_j(y). \tag{1.4}$$

Les fonctions $\psi_j$ sont les fonctions harmoniques associées aux polynômes d'Hermite,

$$\psi_j(x) = \frac{1}{(2\pi)^{1/4}\sqrt{j!}} e^{-x^2/4} P_j(x),$$

avec

$$P_j(x) = (-1)^j e^{x^2/2} \frac{d^j}{dx^j} e^{-x^2/2},$$

le $j^{\text{ème}}$ polynôme d'Hermite. On dit alors que les valeurs propres forment un processus déterminantal de noyau associé $K_N$ et de nombreuses propriétés en découlent. De ce fait, l'étude complète des propriétés asymptotiques du spectre devient possible, tant au niveau global qu'au niveau local. Certains de ces résultats sont décrits dans ce chapitre : le comportement de la mesure spectrale empirique, la fonction de comptage des valeurs propres, les statistiques linéaires, l'espacement entre les valeurs propres à l'intérieur du spectre et le comportement des valeurs propres individuelles, en particulier celles au bord du spectre.

S'inspirant du modèle de Wigner, et plus particulièrement des matrices du GUE, d'autres modèles de matrices aléatoires sont apparus : les modèles invariants. Ils consistent à remplacer dans l'expression de la densité (1.1) le terme $\operatorname{Tr} M^2$ par $\operatorname{Tr} V(M)$ où $V : \mathbb{R} \mapsto \mathbb{R}$ est un potentiel satisfaisant certaines hypothèses. La loi de ces matrices reste donc invariante par transformation unitaire mais les coefficients de la matrice ne sont plus indépendants (ce ne sont plus des matrices de Wigner). Seul le GUE (ou le GOE dans le cas réel) appartient aux deux modèles. L'étude des matrices aléatoires s'est alors scindée en deux grandes parties, l'une portant sur les modèles invariants et l'autre sur les matrices de Wigner ou autres modèles proches, qui conservent de l'indépendance. Les techniques utilisées dans ces deux parties sont très différentes. Toutefois, les travaux récents de Erdös, Schlein, Yau *et al* pour les matrices de Wigner introduisent des méthodes qui s'appliquent également aux modèles invariants, comme le montre l'article de Bourgade, Erdös et Yau [17].

La valeur du paramètre $\beta$ qui apparait dans la loi jointe des valeurs propres est 1 pour les matrices réelles, 2 pour les matrices hermitiennes et 4 pour les matrices symplectiques dont nous ne parlerons pas dans ce manuscrit. Une autre manière de généraliser le modèle du GUE est de considérer des processus ponctuels dont les fonctions de corrélation sont données par la formule (1.1) pour $\beta > 0$ quelconque. Ces modèles sont appelés $\beta$-ensembles. Il n'est pas évident a priori de trouver un modèle de matrice aléatoire dont les valeurs propres ont pour loi jointe la formule précédente, hormis pour les valeurs « classiques » de $\beta$, 1, 2 et 4. Dans [27], Dumitriu et Edelman ont répondu à cette question en construisant un modèle tridiagonal. Il est inspiré des matrices des $\beta$-ensembles pour les $\beta$ classiques ($\beta = 1$, 2 ou 4). En effet, pour ces valeurs de $\beta$, il est possible d'appliquer un algorithme de tridiagonalisation aux matrices issues des modèles correspondants (GOE pour $\beta = 1$ et GUE pour $\beta = 2$). Un modèle tridiagonal est donc obtenu pour ces valeurs de $\beta$. Il est alors possible de définir ce modèle pour tout $\beta > 0$ et de montrer que les valeurs propres ont bien la loi jointe recherchée (1.1). Le modèle obtenu est le suivant : pour tout $\beta > 0$,

$$H_\beta = \frac{1}{\sqrt{\beta}} \begin{pmatrix} g_1 & \chi_{(N-1)\beta} & & & \\ \chi_{(N-1)\beta} & g_2 & \chi_{(N-2)\beta} & & \\ & \ddots & \ddots & \ddots & \\ & & \chi_{2\beta} & g_{N-1} & \chi_\beta \\ & & & \chi_\beta & g_N \end{pmatrix}. \tag{1.5}$$

Les $g_i$ sont ici des variables aléatoires indépendantes gaussiennes centrées de variance 2 et les $\chi_\beta, \ldots, \chi_{(N-1)\beta}$ sont des variables aléatoires indépendantes de loi $\chi$ du paramètre indiqué. Les variables $g$ et $\chi$ sont également indépendantes. Ces ensembles ont donné lieu à de nombreux développements récents, voir par exemple [17] et [59]. Les $\beta$-ensembles et les modèles invariants sont traités par exemple dans les ouvrages [1] et [72].

Dans cette thèse, nous nous concentrons uniquement sur les matrices de Wigner et les matrices de covariance. Ces dernières sont définies ainsi, avec un choix de normalisation des variances.

**Définition 1.2.** *Soit $X$ une matrice aléatoire de taille $m \times n$ dont les coefficients sont indépendants, de moyenne nulle et de variance 1. On suppose également, dans le cas où les coefficients sont complexes, que leurs parties réelles et imaginaires sont indépendantes. La matrice $S_{m,n} = \frac{1}{m} X^* X$, de taille $n \times n$, est une matrice de covariance empirique.*

La matrice $S_{m,n}$ étant semi-définie positive, ses $n$ valeurs propres sont réelles et positives ou nulles et elles sont notées $0 \leqslant \lambda_1 \leqslant \cdots \leqslant \lambda_n$. Si $m < n$, les $m - n$ plus petites valeurs propres $\lambda_1, \ldots, \lambda_{m-n}$ sont nulles. On étudie de manière équivalente les valeurs singulières de $\frac{1}{\sqrt{m}} X$, $0 \leqslant s_1 \leqslant \cdots \leqslant s_n$, où $s_j = \sqrt{\lambda_j}$.

Dans le cas où les coefficients de $X$ suivent une loi gaussienne, la matrice est dite appartenir à l'ensemble unitaire de Laguerre (LUE) ou à l'ensemble orthogonal de Laguerre (LOE) selon si elle est complexe ou réelle. De même que pour les ensembles gaussiens unitaire GUE et orthogonal GOE, la densité de probabilité de la loi de $X$ sur $\mathcal{M}_{n,m}(\mathbb{C})$ pour une matrice du LUE ou du LOE s'écrit, du fait du choix de normalisation des variances, sous la forme suivante :

$$\mathbb{P}(d_\beta X) = \frac{1}{Z_{m,n}^{(\beta)}} \exp\Big(-\frac{\beta}{2}\operatorname{Tr}(X^*X)\Big) d_\beta X, \tag{1.6}$$

avec $\beta = 1$ si la matrice est issue du LOE et $\beta = 2$ si elle est issue du LUE. $d_\beta X$ est défini par $d_1 X = \prod_{1\leqslant j,k\leqslant N} dX_{j,k}$ et $d_2 X = \prod_{1\leqslant j,k\leqslant N} d\Re(X_{jk})d\Im(X_{jk})$. Ces matrices présentent les mêmes propriétés d'invariance par transformation unitaire que les matrices du GUE et du GOE. La loi jointe des valeurs propres de ces matrices est donc également connue,

$$C_{m,n;\beta}\prod_{j=1}^{n} x_j^{\beta/2(m-n+1)-1}\exp\Big(-\frac{n\beta}{2}\sum_{j=1}^{n} x_j\Big)\prod_{1\leqslant j<k\leqslant n}|x_k - x_j|^\beta. \tag{1.7}$$

Dans le cas du LUE, comme pour le GUE, il est possible de montrer que les valeurs propres forment un processus déterminantal, de noyau

$$K_{m,n}(x,y) = \sum_{j=0}^{n-1}\varphi_j(x)\varphi_j(y), \tag{1.8}$$

avec $\varphi(x) = x^{(m-n)/2}e^{-nx/2}p_j(x)$ et $p_j(x)$ est le $j^{\text{ème}}$ polynôme de Laguerre, c'est à dire que $(p_j)_{j\in\mathbb{N}}$ est la famille de polynômes orthogonaux associée au poids $x^{m-n}e^{-nx}$ sur $[0,+\infty[$. Des techniques similaires permettront donc d'établir le comportement de diverses quantités reliées aux valeurs propres des matrices du LUE.

De même que pour les matrices de Wigner, les propriétés asymptotiques des valeurs propres de $S_{m,n}$ ont été conjecturées comme étant universelles. Cependant, dans les domaines dans lesquels les matrices de covariance sont utiles, les résultats asymptotiques sont souvent insuffisants et des résultats non asymptotiques sont nécessaires. On s'intéresse alors par exemple à obtenir des bornes sur la probabilité qu'un certain évènement se réalise. Ce type d'approche, également utile en ce qui concerne les matrices de Wigner, est présenté dans le chapitre 3. Le chapitre 1 décrit quelques résultats asymptotiques obtenus pour les matrices de Wigner et de covariance qui seront ensuite examinés d'un point de vue non asymptotique dans le chapitre 3 qui rassemble les résultats principaux de cette thèse dans ce cadre.

La suite du chapitre est organisée de la manière suivante. Les deux premières sections concernent l'étude asymptotique de différentes quantités liées aux valeurs propres des matrices gaussiennes, c'est-à-dire pour des matrices du GUE ou GOE et du LUE ou LOE. Nous commençons par les propriétés globales du spectre. La convergence globale du spectre retient d'abord notre attention, avant de passer à l'étude de la fonction de comptage et des statistiques linéaires. Nous nous penchons ensuite sur le comportement local du spectre, au travers des espacements entre valeurs propres à l'intérieur et au bord du spectre et le comportement des valeurs propres individuelles. La troisième section est consacrée aux résultats d'universalité, à savoir l'extension des résultats précédents au cas non gaussien.

## 1.1 Étude globale du spectre des matrices gaussiennes

Comme annoncé, nous présentons tout d'abord des résultats asymptotiques globaux, i.e. concernant toutes ou presque toutes les valeurs propres, pour les matrices gaussiennes. Ces résultats peuvent être trouvés dans les ouvrages de référence [67], [1], [8] et [72].

### 1.1.1 Convergence de la mesure spectrale empirique

Les résultats de ce paragraphe ont été établis directement pour des classes larges de matrices de Wigner et de covariance. Contrairement à ce qui a été annoncé précédemment, nous ne nous focalisons donc pas ici sur les modèles gaussiens et nous présentons ces résultats dans un cadre général. En revanche, les paragraphes suivants concernent uniquement les matrices gaussiennes.

Le premier résultat concerne le comportement global du spectre des matrices de Wigner. Wigner [102] a montré en 1958 le théorème suivant, amélioré par Arnold [3].

**Théorème 1.3** (Wigner [102])**.** *Soit $W_N$ une matrice de Wigner réelle ou complexe renormalisée. Alors*

$$L_N = \frac{1}{N}\sum_{j=1}^{N} \delta_{\lambda_j(W_N)} \underset{N\to\infty}{\to} \rho_{sc} \text{ presque sûrement,}$$

*où $\rho_{sc}$ est une loi déterministe, appelée loi du demi-cercle, définie par*

$$d\rho_{sc}(x) = \frac{1}{2\pi}\sqrt{4-x^2}\,\mathbb{1}_{[-2,2]}(x)\,dx. \tag{1.9}$$

Ce théorème a d'abord été montré en utilisant la méthode des moments, c'est-à-dire en montrant que les moments de la loi spectrale empirique convergent vers

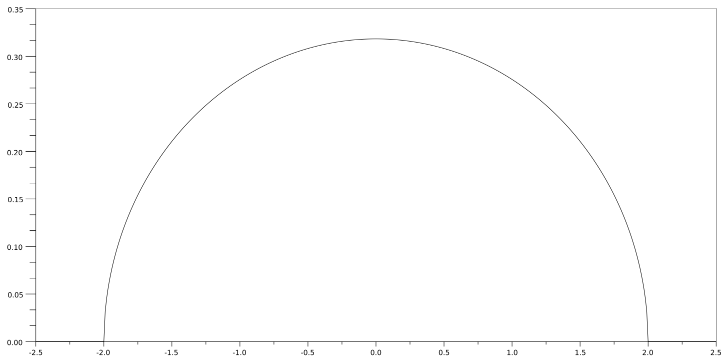


FIGURE 1.1 – Densité de la loi du demi-cercle

ceux de la loi du demi-cercle. Ici, cette propriété suffit pour établir la convergence des mesures. La convergence des moments est montrée par des techniques de combinatoire, faisant intervenir les partitions non croisées et les nombres de Catalan. On peut trouver la démonstration de ce théorème dans [1] et [8]. L'introduction de l'outil de la transformée de Stieltjes dans l'étude des matrices aléatoires par Pastur *et al* a fourni une nouvelle démonstration de ce théorème, qu'on trouvera par exemple dans [8].

De même que pour les matrices de Wigner, Marchenko et Pastur [64] ont montré en 1967 le théorème suivant, concernant le comportement global du spectre des matrices de covariance. La démonstration peut se faire de manière combinatoire ou par transformée de Stieltjes, comme pour le théorème de Wigner.

**Théorème 1.4** (Marchenko-Pastur [64])**.** *Soit $S_{m,n}$ une matrice de covariance réelle ou complexe. On suppose que $\frac{m}{n} \underset{n\to\infty}{\to} \rho \in (0,\infty)$. Alors*

$$L_{m,n} = \frac{1}{n}\sum_{j=1}^{n} \delta_{\lambda_j} \underset{n\to\infty}{\to} \mu_{MP}(\rho) \text{ presque sûrement,}$$

*où $\mu_{MP}(\rho)$ est une loi déterministe, appelée loi de Marchenko-Pastur de paramètre $\rho$, définie par*

$$d\mu_{MP}(\rho)(x) = (1-\rho)_+\delta_0(x)dx + \frac{1}{2\pi x}\sqrt{(b_\rho - x)(x-a_\rho)}\mathbb{1}_{[a_\rho,b_\rho]}(x)\,dx, \qquad (1.10)$$

*où $a_\rho = (1-\sqrt{\rho})^2$ et $b_\rho = (1+\sqrt{\rho})^2$.*

Dans le cas où $\frac{n}{m} \to 0$ quand $n$ tend vers $+\infty$, la matrice $\frac{1}{\sqrt{mn}}X^*X - \sqrt{\frac{m}{n}}I_n$ se comporte « à l'infini » comme une matrice de Wigner et son spectre converge

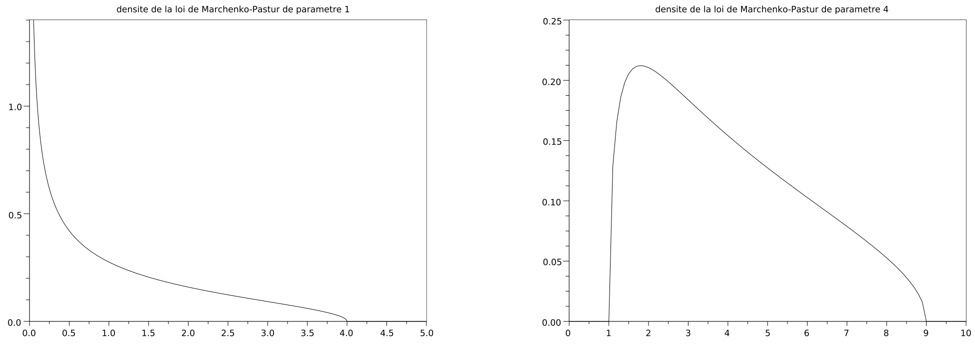


FIGURE 1.2 – Densité de la loi de Marchenko-Pastur pour $\rho = 1$ et $\rho = 4$

vers la loi du demi-cercle [10]. Dans la suite de ce manuscrit, nous nous plaçons dans le cas où $m \geqslant n$ et $\frac{m}{n} \to \rho$, ce qui implique que $\rho \in [1, \infty)$. Dans ce cas, les $n$ valeurs propres sont non triviales et la loi limite n'a pas de masse en zéro. Le cas où $m \leqslant n$ s'en déduit par transposition, les matrices $X^*X$ et $XX^*$ ayant les mêmes valeurs propres non nulles. La loi limite ici dépend du rapport entre le nombre de colonnes et le nombre de lignes. On observe en effet un comportement différent selon que $\rho$ est égal à 1 ou non. Lorsque $\rho > 1$, la borne inférieure du support de la loi limite $a_\rho$ est strictement positive. C'est donc un bord « mou » : les valeurs propres peuvent se trouver de part et d'autre de cette borne inférieure. La densité de la loi de Marchenko-Pastur associée est alors bornée, comme on peut le voir sur la figure 1.2. En revanche, lorsque $\rho = 1$, $a_\rho = 0$ et on observe un bord « dur » : aucune valeur propre ne peut être inférieure à $a_\rho$. La densité associée explose au voisinage de 0, voir figure 1.2. On s'attend donc à des comportements locaux des valeurs propres différents selon la valeur du paramètre, notamment en ce qui concerne les plus petites valeurs propres.

Par simplicité, selon le contexte, $\rho_{sc}$ et $\mu_{MP}(\rho)$ désignent soit les lois en elles-mêmes, soit les densités de ces lois.

### 1.1.2 Fonction de comptage

Dans ce paragraphe, nous nous intéressons aux fluctuations de la fonction de comptage des valeurs propres de matrices gaussiennes de Wigner et de covariance. Cette fonction de comptage est définie de la manière suivante.

**Définition 1.5.** *La fonction de comptage des valeurs propres* $\mathcal{N}_x(A)$ *d'une matrice* $A$ *est le nombre de valeurs propres de* $A$ *qui sont dans l'intervalle* $(-\infty, x]$.

La fonction de comptage est une quantité globale, au sens où elle concerne l'intégralité du spectre. Cependant, elle a des liens très forts avec des propriétés locales des valeurs propres, celles qui ne concernent qu'un petit nombre de valeurs propres. On a en effet l'équivalence suivante

$$\mathcal{N}_x(A) \geqslant k \Leftrightarrow \lambda_k(A) \leqslant x. \tag{1.11}$$

Cette « formule magique » permet de transférer des propriétés connues sur $\mathcal{N}_x(A)$ aux valeurs propres de $A$ et inversement. Par ailleurs, le nombre de valeurs propres dans un petit intervalle autour de $x \in (-2, 2)$ est proche de la partie imaginaire de la transformée de Stieltjes de la matrice, qui est un outil important pour l'étude locale du spectre. Le chapitre 2 donne plus de détails à ce sujet. La fonction de comptage est donc une quantité clé pour étudier les comportements locaux et globaux des valeurs propres.

Cette fonction de comptage normalisée par $\frac{1}{N}$, $N$ étant la taille de la matrice $A$, est également la fonction de répartition empirique des valeurs propres. Les théorèmes de Wigner et de Marchenko-Pastur donnent une loi des grands nombres pour cette fonction de comptage.

$$\frac{1}{N}\mathcal{N}_x(W_N) \underset{N\to\infty}{\to} \rho_{sc}\Big((-\infty, x]\Big) \text{ et } \frac{1}{n}\mathcal{N}_x(S_{m,n}) \underset{n\to\infty}{\to} \mu_{MP}(\rho)\Big((-\infty, x]\Big) \text{ p.s.} \tag{1.12}$$

La question des fluctuations se pose alors naturellement. Dans le cas de l'ensemble gaussien unitaire et de l'ensemble unitaire de Laguerre, la loi jointe des valeurs propres est connue. Ces valeurs propres forment un processus déterminantal et la fonction de comptage a de ce fait la même loi qu'une somme de variables aléatoires de Bernoulli indépendantes. Plus de précisions à ce sujet sont apportées au chapitre 2. Costin et Lebowitz ont montré dans [21] un théorème central limite pour la fonction de comptage de ces processus déterminantaux. Gustavsson, dans son article [49], en a déduit le théorème central limite suivant pour la fonction de comptage des valeurs propres d'une matrice du GUE.

$$\frac{\mathcal{N}_x(W_N) - \mathbb{E}\Big[\mathcal{N}_x(W_N)\Big]}{\sqrt{\operatorname{Var}\Big(\mathcal{N}_x(W_N)\Big)}} \underset{N\to\infty}{\to} \mathcal{N}(0, 1), \tag{1.13}$$

pour $x \in (-2, 2)$, dépendant éventuellement de $N$, vérifiant $N(2-x)^{3/2} \to +\infty$ ou $N(x+2)^{3/2} \to +\infty$ quand $N \to \infty$. Gustavsson poursuit son article en calculant les asymptotiques de l'espérance et de la variance, de telle sorte qu'il obtient un autre théorème central limite, dit « numérique », que nous écrivons uniquement dans le cas où $x$ est à l'intérieur du spectre limite.

**Théorème 1.6** (Gustavsson [49])**.** *Soit $\delta > 0$. Soit $x \in [-2+\delta, 2-\delta]$ dépendant éventuellement de $N$. On a alors*

$$\mathbb{E}\Big[\mathcal{N}_x(W_N)\Big] = N\rho_{sc}\Big((-\infty, x]\Big) + O\Big(\tfrac{\log N}{N}\Big), \tag{1.14}$$

$$\mathrm{Var}\,\Big(\mathcal{N}_x(W_N)\Big) = \frac{1}{2\pi^2}\log N\Big(1+o(1)\Big). \tag{1.15}$$

*Par conséquent,*

$$\frac{\mathcal{N}_x(W_N) - N\rho_{sc}\Big((-\infty, x]\Big)}{\sqrt{\frac{1}{2\pi^2}\log N}} \underset{N\to\infty}{\to} \mathcal{N}(0,1). \tag{1.16}$$

Pour $x$ au bord du spectre vérifiant $N(2-x)^{3/2} \to +\infty$ ou $N(x+2)^{3/2} \to +\infty$, les asymptotiques calculées par Gustavsson sont différentes et un théorème central limite est également dérivé dans ce cas.

Procédant de la même manière, Su [83] montre un théorème analogue pour les matrices du LUE, pour lequel nous introduisons la loi de Marchenko-Pastur « approchée » définie par :

$$d\mu_{m,n}(x) = \frac{1}{2\pi x}\sqrt{(b_{m,n}-x)(x-a_{m,n})}\mathbb{1}_{[a_{m,n}, b_{m,n}]}(x)\,dx, \tag{1.17}$$

où $a_{m,n} = (1-\sqrt{\frac{m}{n}})^2$ et $b_{m,n} = (1+\sqrt{\frac{m}{n}})^2$. De même que pour $\rho_{sc}$ et $\mu_{MP}(\rho)$, $\mu_{m,n}$ peut désigner selon le contexte la loi ou sa densité.

**Théorème 1.7** (Su [83])**.** *Soit $\delta > 0$. Soit $x \in [a_{m,n}+\delta, b_{m,n}-\delta]$ dépendant éventuellement de $m$ et $n$. On a alors*

$$\mathbb{E}\Big[\mathcal{N}_x(S_{m,n})\Big] = n\mu_{m,n}\Big((-\infty, x]\Big) + O(1), \tag{1.18}$$

$$\mathrm{Var}\,\Big(\mathcal{N}_x(S_{m,n})\Big) = \frac{1}{2\pi^2}\log n\Big(1+o(1)\Big). \tag{1.19}$$

*Par conséquent,*

$$\frac{\mathcal{N}_x(S_{m,n}) - n\mu_{m,n}\Big((-\infty, x]\Big)}{\sqrt{\frac{1}{2\pi^2}\log n}} \underset{n\to\infty}{\to} \mathcal{N}(0,1). \tag{1.20}$$

De même, Su établit également un théorème similaire dans le cas où $x$ est près du bord droit du spectre, tout en vérifiant $n(b_{m,n}-x)^{3/2} \to +\infty$. Dans le cas des matrices réelles, des théorèmes analogues peuvent être démontrés. Les valeurs propres des matrices du LOE et du GOE ont en effet une loi jointe qui peut être explicitement calculée. Cependant, elles ne forment plus un processus déterminantal. La loi jointe des valeurs propres fait en effet intervenir un pfaffien à la place d'un déterminant et on ne peut pas appliquer directement les techniques précédentes. O'Rourke utilise les propriétés d'entrelacement des valeurs propres du GUE et du GOE, montrées par Forrester et Rains dans [38], pour étendre les résultats de Gustavsson au cas du GOE (voir [69]). On peut de la même manière étendre ceux de Su au cas du LOE. Les résultats dans le cas réel diffèrent du cas complexe essentiellement d'un facteur 2.

### 1.1.3 Statistiques linéaires

La fonction de comptage est un cas particulier de statistiques linéaires des valeurs propres, quantités que nous considérons dans ce paragraphe et dont de nombreuses propriétés sont décrites dans le livre [72]. Nous nous intéressons particulièrement aux fluctuations de ces statistiques linéaires, définies de la manière suivante.

**Définition 1.8.** *Soit $W_N$ une matrice de Wigner normalisée. Soit $\varphi$ une fonction de $\mathbb{R}$ dans $\mathbb{R}$. On appelle statistique linéaire des valeurs propres de $W_N$ le long de $\varphi$ la variable aléatoire*

$$\mathcal{N}_N[\varphi] = \sum_{j=1}^{N} \varphi(\lambda_j). \tag{1.21}$$

*On définit de manière similaire les statistiques linéaires des valeurs propres d'une matrice de covariance $\mathcal{N}_{m,n}[\varphi]$.*

Notons que la fonction de comptage $\mathcal{N}_x$ est une statistique linéaire avec pour fonction de test la fonction indicatrice $\varphi = \mathbb{1}_{(-\infty,x]}$.

Les théorèmes de Wigner et Marchenko-Pastur fournissent une loi des grands nombres pour des fonctions de test continues bornées :

$$\frac{1}{N}\mathcal{N}_N[\varphi] \underset{N\to\infty}{\to} \int \varphi \, d\rho_{sc} \quad \text{et} \quad \frac{1}{n}\mathcal{N}_{m,n}[\varphi] \underset{n\to\infty}{\to} \int \varphi \, d\mu_{MP}(\rho) \quad \text{p.s.} \tag{1.22}$$

L'étude des fluctuations donne lieu à deux types de comportements différents selon la régularité de la fonction test $\varphi$. D'un côté, nous avons le comportement de la fonction de comptage, qui vérifie un théorème central limite avec une normalisation qui dépend de $N$ (respectivement $n$) comme l'indiquent (1.16) et (1.20). De l'autre côté, lorsque la fonction $\varphi$ est régulière, la variance limite ne dépend pas de $N$ (respectivement $n$). De nombreux auteurs ont travaillé sur les fluctuations des statistiques linéaires dans le cas où $\varphi$ est régulière. Citons par exemple Jonsson [55], Johansson [51], Sinai et Soshnikov [79], Cabanal-Duvillard [19], Bai et Silverstein [7], Anderson et Zeitouni [2], Pastur [71] et Lytova et Pastur [63]. Le livre de Pastur et Shcherbina [72] contient beaucoup d'information à ce sujet. Le meilleur résultat connu actuellement pour les matrices gaussiennes est issu de l'article de Lytova et Pastur [63].

**Théorème 1.9** (Lytova-Pastur [63])**.** *Soit $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction continue bornée, dérivable, dont la dérivée est bornée. Soit $W_N$ une matrice du GUE ($\beta = 2$) ou du GOE ($\beta = 1$) normalisée. Alors*

$$\mathcal{N}_N[\varphi] - \mathbb{E}\Big[\mathcal{N}_N[\varphi]\Big] \underset{N\to\infty}{\to} \mathcal{N}\Big(0, V^{\beta}_{Wig}[\varphi]\Big), \tag{1.23}$$

*avec*

$$V^{\beta}_{Wig}[\varphi] = \frac{1}{2\beta\pi^2}\int_{-2}^{2}\int_{-2}^{2}\Big(\frac{\varphi(x)-\varphi(y)}{x-y}\Big)^2\frac{4-xy}{\sqrt{4-x^2}\sqrt{4-y^2}}\,dxdy. \tag{1.24}$$

Ils montrent ce théorème en utilisant la transformée de Fourier de la statistique linéaire. Un des points clé de la démonstration consiste à établir des bornes sur plusieurs quantités grâce à la majoration de la variance de $\mathcal{N}_N[\varphi]$. Dans ce même article [63], et en utilisant les mêmes techniques, les auteurs montrent également un théorème analogue pour les matrices du LUE et du LOE.

**Théorème 1.10** (Lytova-Pastur [63])**.** *Soit $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction continue bornée, dérivable, dont la dérivée est bornée. Soit $S_{m,n}$ une matrice du LUE ($\beta = 2$) ou du LOE ($\beta = 1$) normalisée. On suppose que $\frac{m}{n} \to \rho \in [1,\infty)$. Alors*

$$\mathcal{N}_{m,n}[\varphi] - \mathbb{E}\Big[\mathcal{N}_{m,n}[\varphi]\Big] \underset{n\to\infty}{\to} \mathcal{N}\Big(0, V^{\beta}_{Cov}[\varphi]\Big), \tag{1.25}$$

*avec*

$$V^{\beta}_{Cov}[\varphi] = \frac{1}{2\beta\pi^2}\int_{a_\rho}^{b_\rho}\int_{a_\rho}^{b_\rho}\Big(\frac{\varphi(x)-\varphi(y)}{x-y}\Big)^2\frac{4\rho-(x-c_\rho)(y-c_\rho)}{\sqrt{4\rho-(x-c_\rho)^2}\sqrt{4\rho-(y-c_\rho)^2}}\,dxdy, \tag{1.26}$$

*où $c_\rho = \frac{1}{2}(a_\rho + b_\rho)$.*

Dans ces deux théorèmes, la convergence a lieu sans facteur normalisant dépendant de $N$ (respectivement $n$). En revanche, la variance limite dépend de la fonction test $\varphi$. Peu de choses sont connues pour les fonctions de test moins régulières que $\mathcal{C}^1$ mais plus régulière qu'une fonction indicatrice et, notamment, rien ne permet encore de faire le transition entre ces deux comportements si différents de la variance des statistiques linéaires. Récemment, Bao, Pan et Zhou se sont intéressés aux statistiques linéaires partielles, c'est-à-dire pour lesquelles on ne somme qu'une partie des termes $\varphi(\lambda_j)$ (voir [14]). Une des manières possibles est de fixer un seuil $u \in \mathbb{R}$ et de ne sommer que les termes pour lesquels $\lambda_j \leqslant u$. Ceci revient à considérer une statistique linéaire « classique » pour la fonction de test $\varphi\mathbb{1}_{(-\infty,u]}$, qu'on notera $\varphi_u$. Ils établissent alors le théorème suivant.

**Théorème 1.11** (Bao-Pan-Zhou [14])**.** *Soit $W_N$ une matrice du GUE normalisée. Soit $\delta > 0$ fixé. Soit $u \in [-2+\delta, 2-\delta]$. Si $\varphi \in \mathcal{C}^1([-2-\delta, 2+\delta])$ et si $\varphi(u) \neq 0$,*

$$\frac{\mathcal{N}_N[\varphi_u] - N\int_{-2}^{u}\varphi(x)d\rho_{sc}(x)}{\sqrt{\frac{\varphi(u)^2}{2\pi^2}\log N}} \underset{N\to\infty}{\to} \mathcal{N}(0,1). \tag{1.27}$$

Dans ce cas, la fonction test est lisse sur $[-2-\delta, 2+\delta]$ sauf en $u$, où elle présente une discontinuité. Le facteur normalisant dépend alors à la fois de $N$ et de la fonction de test. Leur raisonnement est probablement adaptable aux matrices de covariance.

## 1.2 Étude locale du spectre des matrices gaussiennes

Nous abordons maintenant l'étude des propriétés locales du spectre des matrices gaussiennes. Après avoir établi la convergence globale du spectre, Wigner s'est intéressé aux espacements entre deux valeurs propres consécutives. Différents travaux, notamment ceux de Mehta, ont établi le comportement de ces espacements. Nous présentons un aperçu volontairement orienté de ces résultats dans le premier paragraphe. Les livres de Mehta [67] et de Anderson, Guionnet et Zeitouni [1] contiennent de nombreuses informations sur ces espacements. Le comportement des valeurs propres individuelles a également suscité de nombreuses études. Nous nous intéressons particulièrement à leurs fluctuations dans les deux paragraphes suivants, le premier étant consacré aux valeurs propres extrêmes et le second aux valeurs propres à l'intérieur du spectre. Ces résultats et de nombreux compléments peuvent être trouvés dans les ouvrages de référence [1] et [8].

### 1.2.1 Espacements entre deux valeurs propres consécutives

L'expression de la loi jointe des valeurs propres (1.2), (1.7) fait apparaître le terme $\prod_{i<j}|\lambda_j - \lambda_i|^\beta$, qui est petit lorsque les valeurs propres sont proches. Les valeurs propres vont donc avoir tendance à se repousser les unes des autres. Toutefois, les théorèmes de Wigner 1.3 et Marchenko-Pastur 1.4, alliés au fait que les valeurs propres extrêmes convergent vers les bords des supports des lois limites (sujet qui sera discuté au paragraphe 1.2.2), montrent que les valeurs propres ont tendance à rester dans un intervalle proche du support de la limite, $[-2, 2]$ dans le cas de matrices du GUE ou du GOE, $[a_\rho, b_\rho]$ pour le LUE et le LOE. Afin qu'elles soient le plus éloignées les unes des autres, on s'attend à ce que les espacements, au moins à l'intérieur du spectre, soient de taille environ $\frac{1}{N}$, soit $\frac{1}{\sqrt{N}}$ pour les valeurs propres de $M_N$ (respectivement $\frac{m}{n}$ pour $S_{m,n}$). Par ailleurs, le théorème de Wigner

fournit par exemple l'égalité $\mathcal{N}_I(W_N) = N \int_I \rho_{sc}(x)\,dx + o(N)$ pour un intervalle $I$ de $\mathbb{R}$. Si on omet le terme $o(N)$, le nombre de valeurs propres de $W_N$ dans l'intervalle $I$ sera nul dès que $\int_I \rho_{sc}(x)\,dx$ sera plus petit que $\frac{1}{N}$, ce qui correspond pour un intervalle $I$ à l'intérieur du spectre limite à un intervalle de taille plus petite que $\frac{1}{N}$. L'espacement entre deux valeurs propres consécutives à l'intérieur du spectre semble donc bien être d'ordre $\frac{1}{N}$. Le raisonnement analogue peut être mené pour les matrices de covariance.

Nous abordons la question des espacements pour une matrice du GUE $M_N$ à travers la probabilité dite de « trou », c'est-à-dire la probabilité de ne trouver aucune valeur propre dans un ensemble donné. Le survol de Guionnet [46] est une référence générale possible à ce sujet. En s'appuyant sur la forme déterminantale de la loi jointe des valeurs propres (1.4), Mehta et Gaudin ont montré dans [66] et [68] les résultats suivants.

**Lemme 1.12** (Mehta-Gaudin [66, 68])**.** *Soit $A$ un ensemble mesurable de $\mathbb{R}$. Soit $M_N$ une matrice du GUE.*

$$\mathbb{P}\Big(\bigcap_{j=1}^{N} \{\lambda_j(M_N) \in A\}\Big) = 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_{A^c} \cdots \int_{A^c} \det\Big(K_N(x_i,x_j)\Big)_{i,j=1}^{k} \prod_{j=1}^{k} dx_j. \tag{1.28}$$

La probabilité qu'il n'y ait aucune valeur propre dans l'ensemble $A$ s'exprime donc en fonction d'un déterminant de Fredholm du noyau $K_N$, qu'on notera $\Delta(A^c, K_N)$. La formule de Christoffel-Darboux appliquée à ce noyau fournit

$$K_N(x,y) = \sqrt{N}\frac{\psi_N(x)\psi_{N-1}(y) - \psi_{N-1}(x)\psi_N(y)}{x-y}. \tag{1.29}$$

Les asymptotiques des polynômes d'Hermite sont connues, le chapitre 2, en particulier la section 2.1, contient des références à ce sujet. Il est donc possible d'en déduire le comportement des fonctions $\psi$, puis du noyau $K_N$. On obtient alors, pour $x \in (-2,2)$ à l'intérieur du spectre,

$$\lim_{N\to\infty} \frac{1}{\rho_{sc}(x)\sqrt{N}} K_N\Big(N^{1/2}x + \frac{y_1}{\rho_{sc}(x)\sqrt{N}}, N^{1/2}x + \frac{y_2}{\rho_{sc}(x)\sqrt{N}}\Big) = K_{\sin}(y_1,y_2), \tag{1.30}$$

avec $K_{\sin}(y_1,y_2) = \frac{\sin\pi(y_1-y_2)}{\pi(y_1-y_2)}$. Le déterminant de Fredholm étant continu sur l'ensemble des noyaux muni de la topologie de la convergence uniforme, il est possible de déduire le résultat suivant sur la probabilité de « trou », en posant $A = (N^{1/2}x + N^{-1/2}B)^c$ où $B$ est un ensemble compact de $\mathbb{R}$.

$$\lim_{N\to\infty} \mathbb{P}\Big(\lambda_j \notin N^{1/2}x + N^{-1/2}\rho_{sc}(x)B, j = 1\dots N\Big) = \Delta(B, K_{\sin}).$$

Le théorème suivant reprend le résultat précédent.

**Théorème 1.13** (Mehta-Gaudin [66, 68]). *Soit $W_N$ une matrice du GUE renormalisée. Soit $B$ un compact de $\mathbb{R}$.*

$$\lim_{N\to\infty} \mathbb{P}\Big(\lambda_j(W_N) \notin x + \frac{\rho_{sc}(x)}{N}B, j = 1\dots N\Big) = \Delta(B, K_{\sin}). \tag{1.31}$$

La probabilité de trouver un ensemble vide autour du point $x$ converge donc vers une quantité non triviale quand cet ensemble est de taille $\frac{C}{N}\rho_{sc}(x)$, où $C$ est une constante positive. Le facteur normalisant $\frac{1}{N}$ confirme l'intuition initiale de la taille de l'espacement entre deux valeurs propres consécutives. On observe l'apparition d'un autre facteur normalisant, qui dépend de la position du point $x$ dans $(-2, 2)$, $\rho_{sc}(x)$. Par ailleurs, l'étude des fluctuations de la $j^{\text{ème}}$ valeur propre $\lambda_j$ située à l'intérieur du spectre, effectuée par Gustavsson [49] et dont nous parlerons plus en détail dans le paragraphe 1.2.3, montre que cette valeur propre fluctue autour de $\gamma_j$, sa position théorique définie par

$$\frac{j}{N} = \int_{-2}^{\gamma_j} d\rho_{sc}(x).$$

On s'attend donc à ce que les valeurs propres soient fortement concentrées autour de leurs positions théoriques. C'est le cas pour les matrices du GUE, comme pour les matrices de Wigner non gaussiennes, ainsi que cela sera présenté à la section 1.3 et détaillé dans le chapitre 2. Cette propriété est appelée « rigidité des valeurs propres ».

Très récemment, Tao a étudié les espacements individuels $\lambda_{i+1} - \lambda_i$ pour un indice $i$ vérifiant $\eta N \leqslant i \leqslant (1-\eta)N$ (voir [84]). Ces espacements individuels avaient été étudiés en moyenne par Mehta et Gaudin [68]. Les auteurs déduisent des théorèmes précédents que les espacements moyens sont régis à la limite par la distribution de Gaudin, $P_{\text{Gaudin}}$, dont la densité $p$ est donnée par la relation

$$\det\Big(1 - \mathbb{1}_{[0,y]} K_{\sin} \mathbb{1}_{[0,y]}\Big) = \int_y^{\infty} p(y)(x-y)\,dx.$$

Dans [84], Tao « localise » ce résultat et montre qu'un espacement entre deux valeurs propres consécutives à l'intérieur du spectre correctement normalisé est régi à la limite par cette même distribution de Gaudin. Ce résultat est basé sur les propriétés des processus déterminantaux et sur le fait que $\mathcal{N}_x$ est approximativement indépendant de l'évènement qu'il n'y ait aucune valeur propre dans $[x, x+s]$ lorsque la matrice $M_N$ est normalisée correctement.

**Théorème 1.14** (Tao [84]). *Soit $M_N$ une matrice du GUE. Soit $\eta > 0$. On considère $j$ tel que $\eta N \leqslant j \leqslant (1-\eta)N$. Alors*

$$\sqrt{N}\rho_{sc}(\gamma_j)(\lambda_{j+1} - \lambda_j) \to P_{Gaudin}. \tag{1.32}$$

Reprenons l'égalité $\mathcal{N}_I(W_N) = N \int_I \rho_{sc}(x)\,dx + o(N)$ pour un intervalle $I$ inclus dans $[-2, 2]$ mais proche d'un des bords. Si on omet le terme $o(N)$, le nombre de valeurs propres de $W_N$ dans l'intervalle $I$ sera nul dès que $\int_I \rho_{sc}(x)\,dx$ sera plus petit que $\frac{1}{N}$. Au bord du spectre, la densité de la loi du demi-cercle n'est plus minorée par un réel strictement positif. Il est alors possible de minorer la densité par $\frac{1}{\sqrt{2\pi}}\sqrt{2-x}$ lorsque l'intervalle est proche du bord droit du spectre, ce qui conduit à un intervalle $I$ au bord du spectre limite dont la taille est majorée par $\frac{1}{N^{2/3}}$ (il en est de même sur le bord gauche). L'espacement entre deux valeurs propres consécutives au bord du spectre semble donc être plus grand qu'à l'intérieur, d'ordre $\frac{1}{N^{2/3}}$. En utilisant les mêmes techniques d'asymptotiques de polynômes orthogonaux que pour l'intérieur, Tracy et Widom ont étudié la probabilité de « trou » pour les intervalles situés au bord du spectre (voir par exemple [46], [31] ou [1] pour des références générales à ce sujet). Ils ont établi le théorème suivant.

**Théorème 1.15** (Tracy-Widom [91])**.** *Soit $W_N$ une matrice du GUE renormalisée.*

$$\lim_{N\to\infty} \mathbb{P}\Big(\lambda_j \notin 2 + N^{-2/3}[t, t'], j = 1 \dots N\Big) = \Delta([t, t'], K_{Ai}), \tag{1.33}$$

*où $K_{Ai}(x, y) = \frac{Ai(x)Ai'(y) - Ai'(x)Ai(y)}{x-y}$ est le noyau d'Airy et Ai est la fonction spéciale d'Airy.*

Le comportement sur le bord gauche est identique. On observe alors un espacement typique de l'ordre de $\frac{C}{N^{2/3}}$. L'étude des espacements au bord permet de déduire des informations sur le comportement des valeurs propres extrêmes, qui sera détaillé à la section suivante 1.2.2. En effet, comme les valeurs propres sont ordonnées, l'évènement qu'il n'y en ait aucune dans un intervalle du type $[t, \infty)$ correspond à l'évènement que la plus grande valeur propre $\lambda_N$ soit inférieure à $t$.

L'étude des espacements pour les matrices du GOE mène aux mêmes conclusions. Elle est cependant plus compliquée car les valeurs propres ne forment plus un processus déterminantal. Toutefois, la loi jointe des valeurs propres fait intervenir un pfaffien et l'étude reste possible. Elle a été menée par Tracy et Widom [93] et est détaillée dans l'ouvrage [1].

Pour les matrices de covariance, les mêmes procédés permettent d'obtenir les comportements locaux. Les matrices du LUE ont en effet les mêmes propriétés déterminantales que celles du GUE, les polynômes de Laguerre remplaçant ceux d'Hermite. A l'intérieur du spectre, on retrouve un comportement de type noyau sinus. Sur le bord droit du spectre, c'est-à-dire autour de la plus grande valeur propre, on retrouve un comportement de type noyau d'Airy. Sur le bord gauche du spectre, les espacements et le comportement de la plus petite valeur propre sont différents selon la valeur de $\rho$. Si $\rho > 1$, c'est-à-dire lorsque le bord gauche est un bord « mou », le comportement est de type noyau d'Airy. Dans le cas du bord

« dur » (pour $\rho > 1$), c'est le noyau de Bessel qui intervient sur le bord gauche du spectre. Il est défini par

$$K_J(x,y) = \frac{J_\alpha(\sqrt{x})\sqrt{y}J'_\alpha(\sqrt{y}) - \sqrt{x}J'_\alpha(\sqrt{x})J_\alpha(\sqrt{y})}{2(x-y)}, \tag{1.34}$$

où $\alpha = m - n$ est supposé indépendant de $n$. Tracy et Widom ont obtenu ces résultats dans [92]. Ils peuvent être trouvés dans les références générales [1], [46], [67], [72].

Récemment, Ben Arous et Bourgade se sont intéressés à la loi limite des espacements extrêmes pour des matrices du GUE [4]. Ils obtiennent un espacement minimal d'ordre $N^{-4/3}$ et un espacement maximal d'ordre $\frac{\sqrt{\log N}}{N}$. L'idée est d'utiliser les asymptotiques connues sur la densité des espacements $p_2(s)$ lorsque l'espacement tend vers 0 et vers $+\infty$, tout en considérant que les espacements sont indépendants. La difficulté consiste ensuite à montrer que cette approximation est asymptotiquement correcte.

### 1.2.2 Fluctuations des valeurs propres extrêmes

Pour certaines applications des matrices aléatoires, comme le traitement du signal, une question importante est de savoir s'il y a des valeurs propres à l'extérieur du support du spectre limite et à quel point elles en sont éloignées (voir par exemple [8]). De nombreux chercheurs se sont donc intéressés au comportement des valeurs propres extrêmes. Le livre de Bai et Silverstein [8] constitue une référence générale à ce sujet, ainsi que celui de Anderson, Guionnet et Zeitouni [1]. Le théorème de Wigner établit la convergence globale du spectre mais ne donne pas d'information complète sur les valeurs propres individuelles. En particulier, on pourrait tout à fait imaginer que la plus grande valeur propre « saute » du spectre limite et converge vers une valeur strictement plus grande que 2. Le théorème de Wigner fournit pour seule information sur la plus grande valeur propre que $\liminf \lambda_N \geqslant 2$, ainsi qu'une information similaire pour la plus petite valeur propre $\lambda_1$. En 1988, Bai et Yin établissent des conditions nécessaires et suffisantes pour que la plus grande et la plus petite valeurs propres d'une matrice de Wigner convergent effectivement vers les bords du support [11].

**Théorème 1.16** (Bai-Yin [11]). *Sous des hypothèses sur les quatrièmes moments des coefficients de la matrice,*

$$\lambda_1 \to -2 \quad et \quad \lambda_N \to 2 \quad p.s. \tag{1.35}$$

Ce théorème se démontre en utilisant la méthode des moments, qui permet de traiter directement le cas de matrices de Wigner assez générales, satisfaisant une condition sur les quatrièmes moments des coefficients. Comme $-M_N$ est aussi une matrice de Wigner satisfaisant les mêmes conditions, l'étude de la plus petite valeur propre se ramène à celle de la plus grande. Leurs comportements sont en effet symétriques. Dans le cas des matrices de covariance, sous des conditions similaires, la plus petite et la plus grande valeurs propres convergent vers les bords du support de la loi de Marchenko-Pastur. Mais contrairement aux matrices de Wigner, ces valeurs propres extrémales n'ont plus des comportements symétriques et il faut les traiter séparément. L'étude de la plus grande $\lambda_n$ est moins difficile et des conditions nécessaires et suffisantes pour sa convergence sont établies en combinant les deux articles [6] et [9]. Ces conditions nécessaires et suffisantes sont similaires à celles pour les matrices de Wigner et imposent des conditions sur les quatrièmes moments des coefficients de la matrice. Le cas de la plus petite valeur propre est plus difficile en raison des comportements différents de la loi de Marchenko-Pastur, notamment lorsqu'on s'approche du bord dur, c'est-à-dire lorsque $\rho$ est proche de 1. La convergence de $\lambda_1$ vers $a_\rho$ est établie pour tout $\rho$ par Bai et Yin dans [12], toujours sous une hypothèse de finitude des quatrièmes moments des coefficients de la matrice.

La question des fluctuations est fortement liée à celle des espacements au bord du spectre. En effet, pour une matrice du GUE renormalisée $W_N$, le théorème 1.15 appliqué à l'intervalle $[t, +\infty)$ pour $t \in \mathbb{R}$ permet de connaître la limite de $\mathbb{P}\Big(\lambda_j \notin 2 + N^{-2/3}[t, +\infty)\Big)$, soit de $\mathbb{P}\Big(N^{2/3}(\lambda_N - 2) \leqslant t\Big)$, d'après (1.33). La limite est donnée en termes de déterminant de Fredholm du noyau d'Airy. Tracy et Widom ont alors montré que cette limite était bien une fonction de répartition et étudié la distribution correspondante [91]. Le cas des matrices du GOE est traité dans un autre de leurs articles [93]. Leurs résultats sont présentés dans le théorème suivant.

**Théorème 1.17** (Tracy-Widom [91, 93])**.** *Soit $M_N$ une matrice du GUE. On pose $W_N = \frac{1}{\sqrt{N}} M_N$. Alors*

$$N^{2/3}(2 - \lambda_N) \underset{N \to \infty}{\to} F_2 \quad \textit{et} \quad N^{2/3}(\lambda_1 + 2) \underset{N \to \infty}{\to} F_2, \tag{1.36}$$

*où $F_2$ est la distribution de Tracy-Widom, dont on définit la fonction de répartition de la manière suivante. Soit $q : \mathbb{R} \mapsto \mathbb{R}$ l'unique solution de l'équation différentielle*

$$q''(x) = xq(x) + 2q(x)^3$$

*vérifiant $q(x) \underset{x \to +\infty}{\sim} A_i(x)$. On pose alors*

$$F_2(s) = \exp\Big( - \int_s^{+\infty} (x - s) q(x)^2 \, dx \Big).$$

*Dans le cas où la matrice est issue du GOE, les mêmes quantités convergent vers* $F_1$ *avec*

$$F_1(s) = \sqrt{F_2(s)}\exp\Big( - \frac{1}{2}\int_0^{+\infty} q(x)\,dx\Big).$$

Dans le cas des matrices de covariance gaussiennes, le comportement des valeurs propres extrêmes dépend de la situation dans laquelle on se trouve : bord « dur » ou bord « mou ». Le comportement de la plus grande valeur propre est similaire à celui de la plus de grande valeur propre de matrices du GUE. C'est Johansson, en 2000, qui a montré ce résultat dans le cas du LUE et Johnstone, en 2001, l'a montré pour les matrices du LOE.

**Théorème 1.18** (Johansson [52], Johnstone [54])**.** *Soit* $S_{m,n}$ *une matrice du LUE ou du LOE. On suppose que* $\frac{m}{n} \to \rho \in [1, +\infty)$*. Alors*

$$n^{2/3}\frac{\lambda_n - b_{m,n}}{b_{m,n}^{2/3}\left(\frac{m}{n}\right)^{-1/6}} \to F_\beta. \tag{1.37}$$

En 2003, Borodin et Forrester établissent le comportement de la plus petite valeur propre dans le cas du bord « mou » [16]. On retrouve un comportement similaire à celui de la plus petite valeur propre d'une matrice du GUE : la normalisation est en $n^{2/3}$ et la loi limite est la loi de Tracy-Widom. Pour $S_{m,n}$ une matrice du LUE ou du LOE, avec $\frac{m}{n} \to (1, +\infty)$,

$$n^{2/3}\frac{a_{m,n} - \lambda_1}{a_{m,n}^{2/3}\left(\frac{m}{n}\right)^{-1/6}} \to F_\beta. \tag{1.38}$$

Dans le cas du bord dur, Edelman a calculé en 1988 la loi exacte de la plus petite valeur propre du LUE et du LOE lorsque $m = n$. Il se base pour cela sur l'expression exacte de la loi jointe des valeurs propres et procède par intégration. Il en déduit les fluctuations suivantes.

**Théorème 1.19** (Edelman [30])**.** *Si* $S_{n,n}$ *est une matrice du LUE,*

$$n^2\lambda_1 \underset{n\to\infty}{\to} \mathcal{E}(1), \tag{1.39}$$

*où* $\mathcal{E}(1)$ *est la loi exponentielle de paramètre* $1$*. Si* $S_{n,n}$ *est une matrice du LOE, la même quantité converge vers une variable aléatoire dont la densité de probabilité est donnée par* $\frac{1+\sqrt{x}}{2\sqrt{x}}e^{-(x/2+\sqrt{x})}\mathbb{1}_{\mathbb{R}_+}(x)\,dx$*.*

Notons qu'il est possible de montrer que les $k$ plus grandes et les $k$ plus petites valeurs propres (pour $k > 0$ fixé) de matrices gaussiennes convergent également vers une loi de Tracy-Widom. Ceci est vrai pour des matrices du GUE ou du GOE, ainsi que du LUE ou du LOE, le résultat sur les plus petites valeurs propres étant vrai uniquement dans le cas du bord mou. Tous ces résultats peuvent être trouvés dans les ouvrages de référence [8, 1, 72].

Ces distributions de Tracy-Widom sont apparues depuis dans de nombreux autres problèmes plus ou moins en lien avec les valeurs propres extrêmes de matrices aléatoires, comme par exemple les temps de dernier passage en percolation [52] et la longueur de la plus grande sous-suite croissante dans une permutation aléatoire [13]. Le survol de Tracy et Widom [94] recense d'autres occurrences de ces distributions.

### 1.2.3 Fluctuations des valeurs propres à l'intérieur du spectre

Le comportement des valeurs propres extrêmes se déduit du comportement des espacements au bord du spectre. À l'intérieur du spectre, le lien entre les espacements et les valeurs propres individuelles est moins direct. Il est toutefois possible d'étudier leur comportement individuel, c'est-à-dire le comportement de chaque $\lambda_j$, avec $\frac{j}{N} \to a \in (0,1)$. Ainsi, de même que pour les valeurs propres extrêmes, une sorte de loi des grands nombres peut être établie pour ces valeurs propres. Leurs fluctuations peuvent également être étudiées.

Rappelons qu'on définit la position théorique de la $j^{\text{ème}}$ valeur propre $\gamma_j$ par $\frac{j}{N} = \int_{-2}^{\gamma_j} d\rho_{sc}(x)$ pour les matrices de Wigner et $\gamma_j^{m,n}$ par $\frac{j}{n} = \int_{a_{m,n}}^{\gamma_j^{m,n}} d\mu_{m,n}(x)$ pour les matrices de covariance. En utilisant le théorème de Wigner et le fait que les valeurs propres extrêmes sont avec grande probabilité à une distance inférieure à $\varepsilon$ des bornes du spectre limite, on peut montrer que, pour $\eta N \leqslant j \leqslant (1-\eta)N$ (avec $\eta > 0$ fixé),

$$\lambda_j - \gamma_j \underset{N\to\infty}{\to} 0, \tag{1.40}$$

presque sûrement, uniformément en $j$. Un énoncé similaire est disponible pour les matrices de covariance. Ces résultats proviennent de l'utilisation du théorème de Glivenko-Cantelli, qui permet d'affirmer que la convergence de la fonction de répartition empirique des valeurs propres vers celle de la loi du demi-cercle ou de la loi de Marchenko-Pastur est une convergence uniforme.

Dans [49], Gustavsson s'intéresse à la fonction de comptage uniquement dans le but d'établir un théorème central limite pour les valeurs propres à l'intérieur du spectre. Il utilise (1.11) pour transférer le théorème central limite sur la fonction de comptage aux valeurs propres. Il obtient alors le théorème suivant.

**Théorème 1.20** (Gustavsson [49])**.** *Soit $W_N$ une matrice du GUE renormalisée.*

*Pour tout $j$ tel que $\frac{j}{N} \to x \in (0,1)$,*

$$\frac{\lambda_j - \gamma_j}{\sqrt{\frac{2\log N}{(4-\gamma_j^2)N^2}}} \underset{N\to\infty}{\to} \mathcal{N}(0,1). \tag{1.41}$$

Il montre un théorème similaire pour les valeurs propres dites « intermédiaires », c'est-à-dire les valeurs propres vérifiant $\frac{j}{N} \to 1$ et $N - j \to \infty$ pour le bord droit, le comportement de celles sur le bord gauche pouvant être déduit par symétrie.

$$\frac{\lambda_j - \left(2 - \left(\frac{3\pi(N-j)}{N}\right)^{2/3}\right)}{\sqrt{\left(\frac{1}{2\pi}\right)^{2/3} \frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}}} \underset{N\to\infty}{\to} \mathcal{N}(0,1). \tag{1.42}$$

En se basant sur le théorème central limite pour la fonction de comptage d'une matrice du GOE, dont la démonstration repose sur des résultats d'entrelacement des valeurs propres, O'Rourke montre que le comportement des valeurs propres de cette matrice est le même que celui des valeurs propres du GUE, à un facteur 2 près (voir [69]).

Dans [83], Su montre des résultats similaires pour les matrices de covariance. Les techniques utilisées sont identiques : il se base sur le théorème central limite pour la fonction de comptage et sur le calcul des asymptotiques de l'espérance et de la variance de cette fonction de comptage. Il obtient alors le théorème suivant.

**Théorème 1.21** (Su [83])**.** *Soit $S_{m,n}$ une matrice du LUE. On suppose que $m \geqslant n$. Pour tout $j$ tel que $\frac{j}{n} \to x \in (0,1)$,*

$$\mu_{m,n}(\gamma_j^{m,n}) \frac{\lambda_j - \gamma_j^{m,n}}{\sqrt{\frac{1}{2\pi^2}\frac{\log n}{n^2}}} \underset{n\to\infty}{\to} \mathcal{N}(0,1). \tag{1.43}$$

Les matrices de covariance ont un spectre limite dont la densité et le support ne sont pas symétriques sur les deux bords. Le comportement des valeurs propres « intermédiaires » peut donc être différent selon le côté du spectre où l'on se place. Ainsi, Su montre un théorème central limite pour les valeurs propres « intermédiaires » sur le bord droit du spectre.

$$\frac{\left(3\pi(b_{m,n} - a_{m,n})\right)^{1/3}}{\sqrt{2}b_{m,n}^{2/3}} \frac{\lambda_j - \left(b_{m,n} - \left(\frac{3\pi b_{m,n}(n-j)}{\sqrt{b_{m,n}-a_{m,n}}n}\right)^{2/3}\right)}{\sqrt{\frac{\log(n-j)}{n^{4/3}(n-j)^{2/3}}}} \underset{n\to\infty}{\to} \mathcal{N}(0,1). \tag{1.44}$$

Il n'effectue pas les calculs nécessaires au traitement du bord gauche mais il est probable qu'un résultat similaire existe pour ces valeurs propres « intermédiaires » lorsque le bord gauche est un bord mou. En revanche, lorsque c'est un bord dur, le comportement est probablement très différent et, à ma connaissance, rien n'est établi pour les valeurs propres intermédiaires dans ce cas.

## 1.3 Vers des théorèmes d'universalité

Les propriétés locales et globales du spectre de matrices de Wigner et de covariance ont été conjecturées comme étant universelles mais cette universalité pour des classes de matrices assez générales n'a été atteinte que récemment pour nombre des propriétés spectrales présentées dans les sections précédentes. Nous présentons dans cette section certains des résultats d'universalité correspondants. Les outils nécessaires à l'établissement de ces résultats sont détaillés au chapitre 2.

Les premiers théorèmes d'universalité sont ceux de Wigner 1.3 et de Marchenko-Pastur 1.4 concernant la convergence de la mesure spectrale empirique. Ils ont été obtenus directement ou améliorés rapidement pour des matrices très générales et sont basés sur la méthode des moments ou la transformée de Stieltjes. Par conséquent, les lois des grands nombres pour la fonction de comptage (1.12) et pour les statistiques linéaires (1.22) sont valables directement pour ces mêmes matrices de Wigner et de covariance. Les propriétés locales de type lois des grands nombres ont également été montrées rapidement pour des matrices assez générales. Ainsi, la loi des grands nombres pour les valeurs propres extrêmes (1.35) a été obtenue à la fin des années 1980 par Bai, Yin, Silverstein *et al* pour des matrices de Wigner et de covariance satisfaisant une condition sur le moment d'ordre 4 des coefficients, comme cela a été décrit dans le paragraphe 1.2.2. La convergence presque sûre des valeurs propres à l'intérieur du spectre (1.40) découle de ce résultat sur les valeurs propres extrêmes et sur les théorèmes de Wigner et de Marchenko-Pastur. Cette propriété est donc établie comme étant universelle depuis la fin des années 1980.

En revanche, l'universalité des autres propriétés présentées dans les sections précédentes, c'est-à-dire celles qui concernent les fluctuations des quantités considérées, a été bien plus difficile à établir. En effet, les techniques de démonstration utilisées pour les matrices du GUE, du GOE, du LUE et du LOE sont fortement liées à la structure gaussienne. Elles sont donc inutilisables pour des matrices non gaussiennes et de nouvelles méthodes doivent être mises en œuvre.

L'universalité des fluctuations des valeurs propres extrêmes, en lien avec celle des espacements au bord du spectre, a été étudiée par différents auteurs. Soshnikov, en utilisant la méthode des grandes traces, c'est-à-dire l'estimation par des techniques de combinatoire de $\mathbb{E}[Tr(M_N^{2k_N})]$ avec $k_N$ qui tend vers l'infini avec $N$, a établi l'universalité des fluctuations des valeurs propres extrêmes de matrices de Wigner sous l'hypothèse que la loi des coefficients de la matrice est symétrique et a des moments sous-gaussiens [80].

**Théorème 1.22** (Soshnikov [80])**.** *La conclusion du théorème 1.17 reste vraie pour des matrices de Wigner sous certaines hypothèses.*

Les travaux de Péché-Soshnikov [74], Ruzmakina [78] et Khorunzhiy [56] ont réduit l'hypothèse de moments sous-gaussiens à une hypothèse sur le douzième moment mais n'ont pas permis d'enlever l'hypothèse de symétrie de la loi des coefficients. Tous ces articles utilisent des techniques de combinatoire. Les matrices de covariance ont été étudiées avec les mêmes techniques. Soshnikov [82] puis Péché [73] ont établi l'universalité des fluctuations de la plus grande valeur propre de matrices de covariance sous une hypothèse de symétrie de la loi des coefficients. Feldheim et Sodin ont étudié les fluctuations de la plus petite valeur propre, toujours sous une hypothèse de symétrie de la loi des coefficients [36]. Ils utilisent également des techniques de combinatoire. L'universalité pour des matrices de Wigner et de covariance sans hypothèse de symétrie sur la loi des coefficients a été atteinte grâce aux travaux récents et simultanés de Erdös, Schlein, Yau *et al* d'une part, et de Tao et Vu d'autre part. Des conditions nécessaires et suffisantes pour les fluctuations des valeurs propres extrêmes de matrices de Wigner viennent d'être établies par Lee et Yin dans [60]. Ils obtiennent le théorème suivant.

**Théorème 1.23** (Lee-Yin [60])**.** *Soit $M_N$ une matrice de Wigner complexe ou réelle. Alors les fluctuations des valeurs propres extrêmes sont régies par la loi de Tracy-Widom* (1.36) *si et seulement si*

$$\lim_{s\to+\infty} s^4 \mathbb{P}(|(M_N)_{12}| \geqslant s) = 0. \tag{1.45}$$

L'étude des espacements à l'intérieur du spectre connait une grande avancée lorsque Johansson introduit un modèle dont les coefficients ont une petite composante gaussienne [53]. Le modèle exact est décrit plus précisément dans le chapitre suivant. Il fait alors la remarque fondamentale suivante : ce modèle possède encore des propriétés déterminantales. Il montre ainsi l'universalité de la loi des espacements à l'intérieur du spectre pour des matrices dont les coefficients ont une petite composante gaussienne complexe et des moments d'ordre six finis. Les mêmes travaux récents de Erdös, Schlein, Yau *et al* et de Tao et Vu permettent d'étendre ces résultats à des classes très générales de matrices de Wigner et de covariance.

**Théorème 1.24** (Erdös-Ramirez-Schlein-Tao-Vu-Yau [34])**.** *Soit $M_N$ une matrice de Wigner complexe dont les coefficients satisfont une hypothèse de décroissance exponentielle. On note $W_N$ la matrice renormalisée par $\frac{1}{\sqrt{N}}$. Soit $B$ un compact de $\mathbb{R}$.*

$$\lim_{N\to\infty} \mathbb{P}\Big(\lambda_j(W_N) \notin x + \frac{\rho_{sc}(x)}{N} B, j = 1\dots N\Big) = \Delta(B, K_{\sin}), \tag{1.46}$$

*où $\Delta(B, K_{\sin})$ est le déterminant de Fredholm du noyau sinus. Pour les matrices de Wigner réelles, l'universalité est obtenue pour des matrices de Wigner dont les coefficients satisfont une hypothèse de décroissance exponentielle et dont les quatre premiers moments coïncident avec ceux d'une matrice du GOE.*

L'universalité pour les matrices de covariance est obtenue sous des hypothèses similaires. Ces travaux reposent sur des techniques probabilistes qui seront décrites plus précisément au chapitre 2. Ils réduisent les hypothèses sur la loi des coefficients, notamment en enlevant l'hypothèse de symétrie de ces lois. Leurs résultats sont puissants et permettent d'établir l'universalité du comportement de nombreuses quantités locales, parfois même globales, pour des matrices complexes comme pour des matrices réelles. En effet, ils permettent notamment de montrer que les fluctuations des valeurs propres à l'intérieur du spectre et des valeurs propres dites « intermédiaires » sont universelles. Les résultats de Gustavsson et O'Rourke (1.41), (1.42), ainsi que ceux de Su (1.43), (1.44) sont donc établis pour de grandes familles de matrices de Wigner et de covariance. Par (1.11), les propriétés des valeurs propres individuelles se transmettent à la fonction de comptage. Il est donc possible d'étendre le théorème central limite « numérique » vérifié par cette fonction de comptage à des classes générales de matrices de Wigner et de covariance [24]. Les estimées quantitatives de $\mathbb{E}[\mathcal{N}_t]$ et de $\mathrm{Var}(\mathcal{N}_t)$ peuvent elles aussi être étendues. Ce sujet sera discuté plus en détail au chapitre 3, consacré aux résultats non asymptotiques.

L'universalité du comportement des statistiques linéaires globales a fait l'objet de nombreux travaux. Beaucoup concernent des fonctions de test polynomiales ou analytiques, les premières étant traitées par le biais de la méthode des moments, les secondes par transformée de Stieltjes. Le cas d'une fonction de test lipschitzienne peut être traité par des techniques de concentration de la mesure. Citons en particulier Sinai et Soshnikov [79], Cabanal-Duvillard [19], Bai et Silverstein [7], Anderson et Zeitouni [2]. Utilisant la transformée de Fourier et une interpolation avec les matrices gaussiennes, Lytova et Pastur [63] établissent l'universalité des fluctuations des statistiques linéaires pour des fonctions de test essentiellement $\mathcal{C}^4$ et des conditions sur les quatre premiers moments des coefficients de la matrice.

**Théorème 1.25** (Lytova-Pastur [63])**.** *Soit $W_N$ une matrice de Wigner renormalisée satisfaisant une hypothèse sur les quatre premiers moments des coefficients. Soit $\varphi$ une fonction de test dont la transformée de Fourier $\hat{\varphi}$ vérifie*

$$\int_{\mathbb{R}} (1 + |x|^4)|\hat{\varphi}(x)| \, dx < +\infty.$$

*Alors $\mathcal{N}_N[\varphi]$ satisfait* (1.24)*.*

Ce théorème est également vrai pour les matrices de covariance. Bao, Pan et Zhou abordent l'universalité de statistiques linéaires partielles à l'aide de techniques de transformée de Fourier et d'un théorème adapté des travaux de Tao et Vu [14].

# Chapitre 2

# Outils

Ce chapitre contient un descriptif de certains outils développés pour montrer les résultats asymptotiques présentés au chapitre 1. Toutefois, il n'a pas vocation à constituer un descriptif exhaustif des techniques mises en œuvre pour l'étude asymptotique des matrices aléatoires. Nous insistons presque exclusivement sur les outils qui nous seront utiles par la suite, notamment dans le cadre de l'étude des propriétés non asymptotiques. La première section décrit les principales méthodes utilisées dans le cas des matrices gaussiennes. Des références générales possibles pour cette section sont [1] et [72]. La deuxième section présente ensuite les méthodes qui ont été utilisées par différents auteurs pour montrer les théorèmes d'universalité du chapitre 1. Nous nous intéresserons particulièrement aux outils développés par Erdös, Schlein, Yau *et al*, ainsi que par Tao et Vu depuis environ cinq ans. L'objectif de ces travaux était en particulier l'universalité des espacements à l'intérieur du spectre. Ils fournissent des résultats puissants, qui ont révolutionné le monde des matrices aléatoires. Des références générales possibles sont [46], [32] et [90].

## 2.1 Méthodes exactes : cas gaussien

Comme annoncé, nous commençons par les méthodes développées pour les matrices des ensembles unitaires et orthogonaux gaussiens et de Laguerre. Ce sont des méthodes exactes, s'appuyant sur le fait que la structure gaussienne fournit des formules explicites. Si elles permettent d'obtenir le comportement asymptotique du spectre, elles sont également utiles pour fournir des résultats non asymptotiques, aspect qui sera développé dans le chapitre 3. Des références possibles pour ces outils sont [1] et [31].

Le point de départ de l'étude des matrices gaussiennes est le calcul de la loi jointe des valeurs propres. Ce calcul est réalisable du fait de l'invariance de la loi

(1.1) ou (1.6) par transformation unitaire. Il est en effet possible de déduire de cette propriété d'invariance que les valeurs propres d'une matrice de Wigner ou de covariance gaussienne sont indépendantes de ses vecteurs propres. Considérant l'application qui à une matrice associe ses valeurs propres et ce résultat d'indépendance, un changement de variables et le calcul de son jacobien permettent d'obtenir l'expression explicite de la densité de la loi jointe des valeurs propres. Notons toutefois que de nombreuses difficultés techniques apparaissent lors de ce changement de variables, ne serait-ce que parce que l'application considérée n'est pas bijective. Il est notamment nécessaire de restreindre les ensembles considérés, en enlevant certaines parties qui sont de mesure nulle. On peut par exemple trouver ce résultat dans l'ouvrage de Forrester [37].

Dans le cas des matrices gaussiennes complexes, c'est-à-dire du GUE et du LUE, ces lois ont une structure particulière. Les valeurs propres forment en effet un processus déterminantal dont les noyaux $K_N$ et $K_{m,n}$ s'expriment en fonction des polynômes orthogonaux de Hermite (pour le GUE, voir (1.4)) et de Laguerre (pour le LUE, voir (1.8)). Ces propriétés déterminantales sont cruciales pour l'étude des espacements entre valeurs propres consécutives, comme cela a été détaillé au paragraphe 1.2.1. Elles sont également très importantes pour établir les fluctuations de la fonction de comptage $\mathcal{N}_t$. En effet, l'article [50] établit le fait que

$$\mathcal{N}_t \stackrel{(d)}{=} \sum_{k \in \mathbb{N}} \xi_k, \tag{2.1}$$

où les $\xi_k$ sont des variables aléatoires de Bernoulli indépendantes de paramètres $\nu_k$, les $\nu_k$ étant les valeurs propres associées à l'opérateur $K_N$ restreint à l'intervalle $(-\infty, t]$,

$$\begin{aligned} K_N^t : & f \in L^2\Big((-\infty,t]\Big) \mapsto K_N^t f \in L^2\Big((-\infty,t]\Big) \\ \text{avec} \quad & K_N^t f : x \in (-\infty,t] \mapsto \int_{-\infty}^t K_N(x,y) f(y)\, dy. \end{aligned}$$

Ainsi, la fonction de comptage $\mathcal{N}_t$ a même loi qu'une somme de variables aléatoires de Bernoulli indépendantes. De plus, $\mathcal{N}_t$ étant bornée par $N$, cette somme ne comporte pas plus de $N$ variables de Bernoulli dont le paramètre n'est pas nul. Cette écriture sera reprise dans le chapitre 3 et utilisée dans les chapitres 5 et 6. Par ailleurs, Costin et Lebowitz ont établi dans [21] un théorème central limite pour la fonction de comptage associée à un processus déterminantal. C'est ce théorème, qu'il est possible de démontrer également en se basant sur l'écriture (2.1), que Gustavsson [49] et Su [83] utilisent dans leurs articles pour établir des théorèmes centraux limites pour les fonctions de comptage de matrices du GUE (1.16) et du LUE (1.20). Une autre propriété des processus déterminantaux dont ces articles se servent est l'expression de l'espérance et la variance de la fonction de comptage en fonction du noyau. Ainsi

$$\mathbb{E}[\mathcal{N}_t] = \int_{-\infty}^{t} K_N(x,x)\,dx \quad \text{et} \quad \mathrm{Var}(\mathcal{N}_t) = \int_{-\infty}^{t}\int_{t}^{+\infty} K_N^2(x,y)\,dx\,dy.$$

Grâce à la formule de Christoffel-Darboux (1.29), il est possible de remplacer les termes $K_N(x,x)$ et $K_N(x,y)$ par des expressions ne faisant intervenir que les fonctions harmoniques $\psi_j$ associées aux polynômes orthogonaux de Hermite. Or les asymptotiques de ces fonctions ont été calculées dans [25]. Gustavsson s'appuie sur les résultats de cet article pour calculer les intégrales précédentes et en déduire les asymptotiques de l'espérance et de la variance de la fonction de comptage (1.14) et (1.15). Su traite le cas du LUE, prenant pour base les expressions précédentes dans lesquelles $K_N$ est remplacé par $K_{m,n}$, ainsi que les asymptotiques des polynômes orthogonaux de Laguerre, calculées dans [96]. Il obtient des asymptotiques analogues pour l'espérance et la variance de la fonction de comptage (1.18).

Les valeurs propres des matrices gaussiennes réelles, c'est-à-dire du GOE et du LOE, ne forment plus un processus déterminantal. La loi jointe des valeurs propres ne peut plus s'écrire comme le déterminant d'un noyau. Il existe toutefois une écriture faisant intervenir un pfaffien, mais l'étude précédente n'a pas été reproduite directement. Une manière de les traiter passe par les entrelacements. Les ensembles gaussiens unitaires et orthogonaux satisfont en effet des relations d'entrelacement établies par Forrester et Rains dans [38].

**Théorème 2.1** (Forrester-Rains [38]). *Les ensembles gaussiens unitaire et orthogonal vérifient la relation suivante :*

$$\mathrm{GUE}_N = \mathrm{pair}(\mathrm{GOE}_N \cup \mathrm{GOE}_{N+1}). \tag{2.2}$$

*Dans le cas des ensembles unitaire et orthogonal de Laguerre,*

$$\mathrm{LUE}_{m,n} = \mathrm{pair}(\mathrm{LOE}_{m,n} \cup \mathrm{LOE}_{m+1,n+1}) \tag{2.3}$$

Détaillons la signification de la première relation. Considérons deux matrices indépendantes du GOE, l'une de taille $N$ et l'autre de taille $N+1$. Elles ont respectivement $N$ et $N+1$ valeurs propres. On ordonne ces $2N+1$ valeurs propres et on considère uniquement les $N$ éléments pairs de ce $(2N+1)$-uplet. Ces $N$ éléments ont alors la même loi que les $N$ valeurs propres d'une matrice du GUE de taille $N$. La deuxième relation correspond au résultat analogue pour les matrices de covariance gaussiennes. Ces relations peuvent être mises à profit pour déduire du cas complexe des informations sur le cas réel. Pour l'étude de la fonction de comptage, O'Rourke [69] a utilisé cette première relation pour déduire les résultats dans le cas de matrices du GOE à partir des résultats obtenus pour les matrices du GUE. Grâce au théorème d'entrelacement de Cauchy, qui permet de faire le lien entre les valeurs propres d'une matrice de taille $N+1$ et celles d'une de ses matrices extraites de taille $N$, il obtient la relation suivante sur la fonction de comptage.

$$\mathcal{N}_t(M_N^{\mathbb{C}}) \overset{(d)}{=} \frac{1}{2}\Big(\mathcal{N}_t(M_N^{\mathbb{R}}) + \mathcal{N}_t(\tilde{M}_N^{\mathbb{R}})\Big) + \zeta_N(t),$$

où $M_N^{\mathbb{R}}$ et $\tilde{M}_N^{\mathbb{R}}$ sont deux matrices indépendantes du GOE et $\zeta_N(t)$ prend ses valeurs dans $\left\{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\right\}$. Des résultats analogues peuvent être obtenus pour les matrices de covariance gaussiennes, grâce à la relation (2.3).

## 2.2 Étude des matrices non gaussiennes

Dans le cas de matrices non gaussiennes, les méthodes exactes présentées précédemment ne fonctionnent plus. Elles sont trop liées à la structure gaussienne pour pouvoir être adaptées dans le cas général. L'objet de cette section est de présenter différentes techniques développées pour traiter le cas non gaussien. Le premier paragraphe passe en revue ces techniques, tandis que les trois suivants sont consacrés aux outils de Erdös, Schlein, Yau *et al* et de Tao et Vu. Des références générales possibles pour ces outils sont les survols récents de Erdös [32], de Tao et Vu [90] et de Guionnet [46].

### 2.2.1 Des démonstrations souvent basées sur une comparaison avec le cas gaussien

Comme nous l'avons dit, les outils décrits précédemment, permettant de faire des calculs exacts, ne sont utilisables que dans le cas gaussien. Tous les résultats d'universalité reposent à ce jour sur des techniques d'approximation. Beaucoup d'entre eux, notamment tous ceux concernant les statistiques locales des valeurs propres, s'appuient sur le cas gaussien. Le schéma classique consiste à montrer que la statistique dans le cas général est « proche » de la même statistique dans le cas gaussien, ce qui permet d'étendre le résultat gaussien à des matrices plus générales. Il existe toutefois quelques démonstrations directes, par exemple celles des théorèmes de Wigner 1.3 et Marchenko-Pastur 1.4. Pour ces deux théorèmes, on utilise la méthode des moments, employée initialement par Wigner [102], ou la transformée de Stieltjes, méthode développée par Pastur [70]. La méthode des moments consiste à montrer que les moments de la quantité qu'on considère (ici de la loi spectrale empirique) convergent vers les moments d'une certaine quantité limite (ici la loi du demi-cercle ou la loi de Marchenko-Pastur), et à vérifier que cela suffit pour montrer la convergence désirée. Cette convergence de moments est établie par des techniques combinatoires. La transformée de Stieltjes d'une mesure de probabilité $\nu$ sur $\mathbb{R}$ est définie de la manière suivante. Pour tout $z \in \mathbb{C}$ tel que $\Im(z) > 0$,

$$s_\nu(z) = \int_{\mathbb{R}} \frac{1}{x-z}\, d\nu(x). \tag{2.4}$$

La méthode par transformée de Stieltjes repose sur le fait que la convergence de la transformée de Stieltjes de la loi spectrale empirique vers celle de la loi du demi-cercle ou de Marchenko-Pastur implique la convergence des mesures elles-mêmes. Dans le cas de la mesure spectrale empirique, la transformée de Stieltjes s'exprime à l'aide de la résolvante de la matrice $(W_N - zI_N)^{-1}$. Pour tout $z \in \mathbb{C}$ tel que $\Im(z) > 0$,

$$s_{L_N}(z) = \frac{1}{N}\sum_{j=1}^{N} \frac{1}{\lambda_j - z} = \frac{1}{N}\operatorname{Tr}\Big((W_N - zI_N)^{-1}\Big). \tag{2.5}$$

L'étude de cette quantité est alors facilitée par les formules d'algèbre linéaire. La transformée de Stieltjes de la loi du demi-cercle peut être aisément calculée. Pour tout $z \in \mathbb{C}$ tel que $\Im(z) > 0$,

$$s_{sc}(z) = \frac{1}{2}(-z + \sqrt{z^2-4}),$$

où on choisit la racine de $z^2 - 4$ dont la partie imaginaire est strictement positive. Elle vérifie la relation suivante : pour tout $z \in \mathbb{C}$ tel que $\Im(z) > 0$,

$$s_{sc}(z) = -\frac{1}{s_{sc}(z) + z}. \tag{2.6}$$

L'enjeu de la démonstration est de montrer que la transformée de Stieltjes de la mesure spectrale empirique $s_{L_N}$ vérifie quasiment cette même relation, en contrôlant l'erreur. Notons que $s_{L_N}$ est concentrée autour de son espérance. La démonstration du théorème de Marchenko-Pastur par transformée de Stieltjes est similaire, la transformée de Stieltjes de la loi de Marchenko-Pastur vérifiant une relation similaire à (2.6).

Tous les autres résultats d'universalité utilisent à ce jour les résultats du cas gaussien. Par exemple, les premiers pas vers l'universalité des espacements au bord du spectre, et de fait les fluctuations des valeurs propres extrêmes (1.36), ont été effectués par Soshnikov, voir [80] et [82]. Il utilise la méthode des moments, en calculant l'espérance de traces de grandes puissances de $W_N$ (ou $S_{m,n}$), puissances d'ordre $k = tN^{2/3}$, permettant essentiellement de décrire la loi jointe des plus grandes valeurs propres centrées par leur limite et renormalisées. Soshnikov montre alors que ces grandes traces sont suffisamment proches de celles dans le cas de matrices gaussiennes pour que l'universalité des fluctuations au bord du spectre soit vraie. Ses travaux nécessitent des techniques combinatoires très fines pour majorer ces grandes traces. Ces résultats d'universalité ne sont donc montrés que sous des hypothèses assez restrictives sur les lois des coefficients, notamment le fait que ces lois sont symétriques (ce qui permet d'annuler un grand nombre de moments). Ses travaux ont été améliorés et complétés par Péché [73] et Feldheim-Sodin [36], mais sans que l'hypothèse de symétrie de la loi des coefficients ne puisse être enlevée.

D'autres auteurs ont également eu recours à une approximation par des modèles gaussiens, en faisant une interpolation entre une matrice gaussienne et une matrice non gaussienne. Plus précisément, Johansson a étendu les résultats concernant les espacements à l'intérieur du spectre aux matrices de la forme $\sqrt{\varepsilon}M'_N + \sqrt{1-\varepsilon}M_N$, avec $\varepsilon > 0$ fixé, où $M'_N$ est une matrice du GUE et $M_N$ est une matrice de Wigner indépendante de $M'_N$, voir [53]. La remarque fondamentale est que la loi jointe des valeurs propres possède encore des propriétés déterminantales. De manière plus précise, elle s'exprime comme une moyenne de déterminants et il est alors possible de reproduire les calculs faits dans le cas gaussien, bien que l'opération soit beaucoup plus ardue. Ce même schéma d'interpolation est utilisé par Lytova et Pastur pour étendre le comportement des statistiques linéaires globales aux matrices non gaussiennes, voir [63].

Très récemment, de formidables avancées ont été obtenues par Erdös, Schlein, Yau *et al* d'un côté et Tao et Vu *et al* de l'autre. Ces deux équipes ont mis au point des outils puissants pour montrer l'universalité des statistiques locales, dans le but notamment d'établir l'universalité des espacements à l'intérieur du spectre (théorème 1.24). Il est important de noter que ces résultats permettent de transférer les résultats existants dans le cas gaussien ou dans le cas de matrices un peu plus générales comme celles étudiées par Johansson par exemple. Ils ne fournissent pas de démonstration directe du comportement des statistiques locales. Comme précédemment, ces outils constituent des méthodes d'interpolation avec le cas gaussien. Dans les trois paragraphes qui suivent, nous nous intéressons de plus près aux outils et méthodes déployés par ces deux groupes. Nous commençons par les travaux de Erdös, Schlein, Yau *et al*, qui établissent une loi du demi-cercle ou de Marchenko-Pastur « locale » ainsi que la délocalisation des vecteurs propres. Nous traitons ensuite les résultats de Tao et Vu, basés sur la stratégie de Lindeberg. Ces deux paragraphes contiennent les outils sur lesquels sont basés les résultats de cette thèse, résultats qui seront présentés au chapitre 3. Nous terminons par un paragraphe expliquant brièvement comment l'universalité des statistiques locales est obtenue pour des matrices de Wigner et de covariance très générales à partir de chacune des approches précédentes.

### 2.2.2 Lois locales, propriétés de localisation et délocalisation des éléments propres

Dans tout ce paragraphe, nous notons $\mathrm{llog} N = \log\log N$, $N$ étant supposé assez grand pour que cette quantité soit bien définie. Nous supposons de plus que les matrices de Wigner considérées satisfont la condition $(C0)$ définie de la manière suivante.

**Définition 2.2.** *Une matrice de Wigner $M_N$ satisfait la condition $(C0)$ si la loi de ses coefficients $(\xi_{i,j})$ a une décroissance exponentielle, c'est-à-dire s'il existe des constantes $C$ et $C'$ telles que, pour tout $t \geqslant C'$,*

$$\mathbb{P}(|\xi_{i,j}| > t^C) \leqslant e^{-t}.$$

Cette condition est une hypothèse générale souple d'utilisation mais susceptible d'améliorations.

Nous présentons donc dans ce paragraphe certains des résultats de Erdös, Schlein, Yau *et al.* Des références possibles sont le survol récent de Erdös [32] et celui de Guionnet [46]. Le fondement de leurs travaux est l'établissement d'une loi du demi-cercle locale. En effet, le théorème de Wigner donne pour un intervalle $I = [a, b]$ la relation suivante (ici, $a$ et $b$ sont des réels fixés).

$$\frac{1}{N}\mathcal{N}_I \to \int_a^b \rho_{sc}(x)dx.$$

Le nombre de valeurs propres dans l'intervalle $I$ est donc supposé être de l'ordre de $N(b-a)$, quand $b$ et $a$ sont à l'intérieur du support de la loi du demi-cercle $[-2, 2]$. Le théorème de Wigner est une loi des grands nombres, le résultat devrait donc être valable dès que le nombre d'objets considérés tend vers l'infini, c'est-à-dire ici dès que $\mathcal{N}_I$ tend vers l'infini. Si l'approximation précédente est vraie pour $a$ et $b$ quelconques, en particulier dépendant de $N$, le théorème de Wigner devrait être vrai dès que l'intervalle $I$ a une taille supérieure à $\frac{1}{N}$. Ce que Erdös, Schlein, Yau *et al* appellent la loi du demi-cercle locale est justement de montrer que, pour des intervalles de taille quasiment $\frac{1}{N}$ autour d'un point $x$ dans $(-2, 2)$, le nombre de valeurs propres dans $I$ est de l'ordre de $\rho_{sc}(x)|I|$ avec grande probabilité.

Pour contrôler le nombre de valeurs propres dans de tels intervalles, les auteurs font appel à la transformée de Stieltjes. Sur de petits intervalles de la forme $[E-\eta, E+\eta]$, le nombre de valeurs propres peut en effet être contrôlé précisément à l'aide de la partie imaginaire de $s_{L_N}(E+i\eta)$. En reprenant la démonstration du théorème de Wigner par transformée de Stieltjes, ils se sont aperçus que l'une des majorations pouvait être affinée, permettant ainsi de quantifier (et d'améliorer) la précision sur la vitesse de convergence. En effet, un terme d'erreur était classiquement contrôlé par $\frac{1}{\Im(z)}$ et peut être contrôlé plus précisément, à condition d'avoir une majoration a priori sur le nombre de valeurs propres autour de $\Re(z)$. Ces arguments ont été affinés dans une série d'articles de Erdös, Schlein, Yau *et al*, permettant d'atteindre la loi du demi-cercle locale jusqu'à des échelles d'ordre quasiment $\frac{1}{N}$. Dans l'article [35], la loi du demi-cercle locale en termes de transformée de Stieltjes s'exprime de la manière suivante.

$$\mathbb{P}\Big(\sup_{z\in S}|s_{L_N}(z)-s_{sc}(z)|>\frac{(\log N)^{C\mathrm{llog}N}}{N\eta}\Big)\leqslant C'\exp\Big(-c(\log N)^{c'\mathrm{llog}N}\Big), \quad (2.7)$$

avec $S=\{z=E+i\eta;|E|\leqslant 5,\frac{1}{N}(\log N)^{C''\mathrm{llog}N}<\eta<10\}$. Il est alors possible d'en déduire une majoration fine de la différence entre la fonction de comptage normalisée et la loi du demi-cercle.

$$\mathbb{P}\Big(\sup_{|t|\leqslant 5}\Big|\frac{1}{N}\mathcal{N}_t-\rho_{sc}\Big((-\infty,t]\Big)\Big|>\frac{(\log N)^{c\mathrm{llog}N}}{N}\Big)\leqslant C\exp\Big(-c'(\log N)^{C'\mathrm{llog}N}\Big). \quad (2.8)$$

Cette loi locale vient confirmer ce que la formule limite des espacements à l'intérieur du spectre pour les matrices du GUE (1.31) semblait indiquer, à savoir que le comportement des valeurs propres semble être bien décrit par les prévisions fournies par la loi du demi-cercle. Par conséquent, il semble naturel de penser que les valeurs propres vont être localisées autour de leurs positions théoriques fournies par la loi du demi-cercle, les $\gamma_j$, comme c'est le cas pour le GUE. Dans cette même série d'articles, Erdös, Schlein, Yau *et al* obtiennent des résultats de localisation de plus en plus précis pour ces valeurs propres. Les meilleurs résultats à ce jour sont contenus dans [35], dans lequel Erdös, Yau et Yin déduisent effectivement de la loi du demi-cercle locale une propriété de localisation des valeurs propres, appelée théorème de rigidité des valeurs propres.

**Théorème 2.3** (Erdös-Yau-Yin [35])**.** *Soit $W_N$ une matrice de Wigner normalisée vérifiant la condition $(C0)$. Alors, pour tout $j\in[\![1,N]\!]$,*

$$\mathbb{P}\Big(|\lambda_j-\gamma_j|\geqslant\frac{(\log N)^{cllogN}}{\min(j,N+1-j)^{1/3}N^{2/3}}\Big)\leqslant C'\exp\Big(-c'(\log N)^{llogN}\Big). \quad (2.9)$$

Ce théorème fournit des renseignements précis et quantifiés sur la position des valeurs propres. Combiné au théorème des quatre moments de Tao et Vu (voir section 2.2.3), nous l'utiliserons à plusieurs reprises dans les chapitres 4, 5, 6.

Par ailleurs, Erdös, Schlein et Yau montrent que les vecteurs propres des matrices de Wigner sont délocalisés avec grande probabilité, c'est-à-dire que les coordonnées d'un vecteur propre $v$ de norme 2 égale à 1 vont être à peu près tous du même ordre de grandeur $\frac{1}{\sqrt{N}}$. Ce résultat provient également de la loi du demi-cercle locale.

Des résultats analogues sont disponibles pour les matrices de covariance et ont été montrés par Pillai et Yin dans [75]. Le raisonnement suit exactement le même chemin. Les auteurs établissent d'abord une loi de Marchenko-Pastur locale, en déduisent les propriétés de localisation des valeurs propres et de délocalisation des vecteurs propres, ce qui permettra par la suite de montrer l'universalité des statistiques locales des valeurs propres, comme décrit dans le paragraphe 2.2.4.

### 2.2.3 Stratégie de Lindeberg et théorème des quatre moments

Tao et Vu, de leur côté, développent la méthode de Lindeberg (rappelée plus bas) pour comparer les statistiques locales de deux matrices aléatoires en remplaçant un à un les coefficients de la première matrice par ceux de la deuxième. Pour contrôler précisément l'erreur effectuée à chaque pas, ils ont besoin des résultats de Erdös, Schlein, Yau *et al* sur la loi du demi-cercle locale et la délocalisation des vecteurs propres. Des références possibles pour ce paragraphe sont les survols récents [46] et [90]. Avant de citer le théorème, dit des quatre moments, qu'ils obtiennent, nous introduisons deux définitions.

**Définition 2.4.** *Une matrice de Wigner $M_N$ satisfait la condition $(C1)$ de constante $C$ si la loi de ses coefficients $(\xi)_{i,j}$ admet un moment d'ordre $C$ borné, c'est-à-dire s'il existe une constante $C' > 0$ indépendante de $N$ telle que*

$$\mathbb{E}[|\xi_{i,j}|^C] \leqslant C'.$$

Notons que la condition $(C0)$ décrite à la définition 2.2 implique que la condition $(C1)$ est vérifiée pour toute constante $C$.

**Définition 2.5.** *Deux variables aléatoires $\xi$ et $\xi'$ ont des moments qui coïncident jusqu'à l'ordre $k \in \mathbb{N}$ si*

$$\mathbb{E}[\Re(\xi)^i \Im(\xi)^j] = \mathbb{E}[\Re(\xi')^i \Im(\xi')^j]$$

*pour tous entiers $i$ et $j$ tels que $i + j \leqslant k$.*

Nous pouvons maintenant présenter le théorème des quatre moments de Tao et Vu.

**Théorème 2.6** (Tao-Vu [87, 86]). *Il existe une constante $c_0 > 0$ telle que ce qui suit est vrai. Soient $M_N$ et $M'_N$ deux matrices de Wigner satisfaisant la condition $(C1)$. Supposons que les quatre premiers moments des coefficients non diagonaux* *de ces deux matrices coïncident, ainsi que les deux premiers moments des coefficients diagonaux. Soit $k$ un entier naturel fixé (ou dépendant de $N$ dans la limite $k \leqslant N^{c_0}$). Soit $G : \mathbb{R}^k \to \mathbb{R}$ une fonction lisse qui vérifie l'hypothèse suivante.*

$$\forall j \in [\![0, 5]\!], |\nabla G^j(x)| \leqslant N^{c_0}. \tag{2.10}$$

*Alors, pour tout choix d'entiers $1 \leqslant i_1 < \cdots < i_k \leqslant N$,*

$$\left|\mathbb{E}\Big[G(\lambda_{i_1}(A_N), \ldots, \lambda_{i_k}(A_N))\Big] - \mathbb{E}\Big[G(\lambda_{i_1}(A'_N), \ldots, \lambda_{i_k}(A'_N))\Big]\right| \leqslant N^{-c_0}, \tag{2.11}$$

*avec $A_N = \sqrt{N} M_N$ et $A'_N = \sqrt{N} M'_N$.*

Voici un exemple d'application de ce théorème, qui sera utilisé plusieurs fois dans les chapitres 4, 5 et 6. Considérons des matrices de Wigner $M_N$ et $M'_N$ satisfaisant les conditions d'application du théorème 2.6. On note $I = [a, b]$, $I^+ = [a - N^{-c_0/10}, b + N^{-c_0/10}]$ et $I^- = [a + N^{-c_0/10}, b - N^{-c_0/10}]$, pour $a < b$ deux réels. Prenons pour fonction test $G$ une fonction plateau lisse telle que $G(x) = 1$ si $x \in I$ et $G(x) = 0$ si $x \notin I^+$ vérifiant les conditions (2.10) (cette fonction peut être obtenue par convolution). Alors le théorème 2.6 s'applique. Pour tout $j \in \{1, \ldots, N\}$ (éventuellement dépendant de $N$),

$$\left|\mathbb{E}\Big[G(\lambda_j(A_N))\Big] - \mathbb{E}\Big[G(\lambda_j(A'_N))\Big]\right| \leqslant N^{-c_0}.$$

De plus, $\mathbb{P}(\lambda_j(A_N) \in I) \leqslant \mathbb{E}[G(\lambda_j(A_N))]$ et $\mathbb{P}(\lambda_j(A'_N) \in I^+) \geqslant \mathbb{E}[G(\lambda_j(A'_N))]$. D'où, d'après l'inégalité triangulaire,

$$\mathbb{P}(\lambda_j(A_N) \in I) \leqslant \mathbb{P}(\lambda_j(A'_N) \in I^+) + N^{-c_0}.$$

Effectuant le même raisonnement à l'aide d'une autre fonction plateau lisse $H$ qui vérifie $H(x) = 1$ si $x \in I^-$ et $H(x) = 0$ si $x \notin I$, on obtient

$$\mathbb{P}(\lambda_j(A'_N) \in I^-) - N^{-c_0} \leqslant \mathbb{P}(\lambda_j(A_N) \in I).$$

La combinaison des inégalités précédentes fournit l'encadrement suivant.

$$\mathbb{P}(\lambda_j(A'_N) \in I^-) - N^{-c_0} \leqslant \mathbb{P}(\lambda_j(A_N) \in I) \leqslant \mathbb{P}(\lambda_j(A'_N) \in I^+) + N^{-c_0}. \tag{2.12}$$

À l'aide d'encadrements de ce type, il est possible de transférer des informations asymptotiques connues sur $M'_N$ à la matrice $M_N$. Ainsi, l'universalité des fluctuations des valeurs propres (1.16), (1.20), (1.36) peut être obtenue de cette manière.

Pour montrer ce théorème, Tao et Vu utilisent la stratégie de Lindeberg, faisant ainsi référence à la démonstration du théorème central limite proposée par Lindeberg [61] (voir également [22]), dont nous reproduisons ici une partie. La stratégie est en effet bien plus facile à comprendre dans la démonstration du théorème central limite que dans celle du théorème des quatre moments où de nombreuses difficultés techniques apparaissent.

L'idée de la démonstration dans le cas du théorème central limite est la suivante. Considérons des variables aléatoires indépendantes et identiquement distribuées à valeurs réelles $(X_n)_{n\in\mathbb{N}}$ et $(Y_n)_{n\in\mathbb{N}}$. Les $X_n$ et les $Y_n$ sont de plus supposées indépendantes, de moyenne nulle et de variance 1. Notons toutefois que cette hypothèse peut être remplacée par le fait que $X_n$ et $Y_n$ ont mêmes premier et deuxième moments (et que ces moments sont finis). Soit $F$ une fonction lisse à support compact sur $\mathbb{R}$. Afin de comparer $\mathbb{E}\Big[F(\frac{1}{\sqrt{n}}\sum_{j=1}^n X_j)\Big]$ et $\mathbb{E}\Big[F(\frac{1}{\sqrt{n}}\sum_{j=1}^n Y_j)\Big]$, on va remplacer une à une les variables $X_j$ par les variables $Y_j$. La première étape consiste à remplacer $X_n$ par $Y_n$. Il faut alors comparer $\mathbb{E}\Big[F(S_{n-1}+\frac{1}{\sqrt{n}}X_n)\Big]$ et $\mathbb{E}\Big[F(S_{n-1}+\frac{1}{\sqrt{n}}Y_n)\Big]$, où $S_{n-1}=\sum_{j=1}^{n-1}X_j$. D'après les formules de Taylor, on sait que

$$F\Big(S_{n-1}+\frac{1}{\sqrt{n}}X_n\Big)=F(S_{n-1})+F'(S_{n-1})\frac{1}{\sqrt{n}}X_n+F''(S_{n-1})\frac{1}{2n}X_n^2+O\Big(\frac{1}{n^{3/2}}|X_n|^3\Big)$$

et

$$F\Big(S_{n-1}+\frac{1}{\sqrt{n}}Y_n\Big)=F(S_{n-1})+F'(S_{n-1})\frac{1}{\sqrt{n}}Y_n+F''(S_{n-1})\frac{1}{2n}Y_n^2+O\Big(\frac{1}{n^{3/2}}|Y_n|^3\Big).$$

Comme $S_{n-1}$, $X_n$ et $Y_n$ sont indépendants, on obtient

$$\begin{aligned}\mathbb{E}\Big[F\Big(S_{n-1}+\frac{1}{\sqrt{n}}X_n\Big)\Big]=\mathbb{E}[F(S_{n-1})]&+\mathbb{E}[F'(S_{n-1})]\frac{1}{\sqrt{n}}\mathbb{E}[X_n]\\&+\mathbb{E}[F''(S_{n-1})]\frac{1}{2n}\mathbb{E}[X_n^2]+O\Big(\frac{1}{n^{3/2}}\mathbb{E}[|X_n|^3]\Big)\end{aligned}$$

et une expression similaire pour la deuxième équation. Mais $X_n$ et $Y_n$ ont les mêmes deux premiers moments. Si on suppose de plus que leurs troisièmes moments sont bornés indépendamment de $n$, on a

$$\mathbb{E}\Big[F(S_{n-1}+\frac{1}{\sqrt{n}}X_n)\Big]=\mathbb{E}\Big[F(S_{n-1}+\frac{1}{\sqrt{n}}Y_n)\Big]+O(n^{-3/2}).$$

En effectuant $n$ fois cette opération et en sommant, on obtient

$$\mathbb{E}\Big[F(S_n(X))\Big]=\mathbb{E}\Big[F(S_n(Y))\Big]+O(n^{-1/2}).$$

Il est alors possible d'en déduire la convergence en loi recherchée, avec de plus une estimation de la vitesse de convergence en $n^{-1/2}$ (théorème de Berry-Esseen, voir par exemple [28]). Par ailleurs, il est possible d'enlever l'hypothèse de troisième moment borné mais l'estimation de la vitesse de convergence n'est plus disponible dans ce cas, voir [22]. Il est intéressant de noter que si les variables aléatoires considérées ont des moments d'ordre supérieur qui coïncident, la vitesse de convergence est améliorée. À chaque coïncidence supplémentaire, la vitesse est augmentée d'un facteur $n^{-1/2}$.

Tao et Vu exploitent cette idée de remplacer un par un les coefficients de la matrice $A_N$ par ceux de la matrice $A'_N$. Plus précisément, à chaque étape, ils remplacent le coefficient $(i,j)$ et le coefficient $(j,i)$ de la matrice initiale $A$ par ceux de la matrice qu'on cherche à atteindre $A'$. Nous décrivons cette procédure dans le cas où la matrice est symétrique réelle. Dans le cas où les coefficients sont complexes, la partie réelle et la partie imaginaire sont remplacées séparément, doublant ainsi le nombre d'étapes nécessaires. On parcourt la partie triangulaire supérieure de la matrice ligne par ligne, suivant l'ordre lexicographique des numéros des coefficients. À l'étape $(i,j)$, on dispose d'une matrice $A_{i,j}$ dont les coefficients placés avant $(i,j)$ dans l'ordre lexicographique sont ceux de $A'_N$. À partir de $(i,j)$, les coefficients sont ceux de $A_N$. Les symétriques sont remplacés simultanément, $A_{i,j}$ est donc une matrice symétrique réelle (c'est une matrice de Wigner réelle, les coefficients des deux matrices $A_N$ et $A'_N$ étant supposés indépendants). Au cours de cette étape, on va remplacer le coefficient $(i,j)$ et son symétrique par celui de $A'_N$. Pour contrôler l'erreur associée à cette étape, on notera $A_{i,j}(x)$ la matrice $A_{i,j}$ dans laquelle on a remplacé le coefficient $(i,j)$ et son symétrique par $x$. Suivant la méthode de Lindeberg, Tao et Vu utilisent un développement de Taylor de la fonction $F_{i,j}(x) = G(\lambda_{i_1},\ldots,\lambda_{i_k})\Big(A_{i,j}(x)\Big)$. On cherche à obtenir un terme d'erreur en $N^{-c_0}$. Comme il y a environ $N^2$ étapes pour les coefficients non diagonaux et $N$ étapes pour les coefficients diagonaux, l'étape de remplacement d'un coefficient non diagonal ne doit pas induire une erreur d'ordre supérieur à $N^{-2-c_0}$, tandis qu'elle ne doit pas excéder l'ordre $N^{-1-c_0}$ pour les éléments diagonaux. Pour obtenir l'ordre $N^{-2-c_0}$, Tao et Vu développent la fonction $F_{i,j}$ à l'ordre 4, atteignant ainsi une erreur d'ordre $N^{-5/2+O(c_0)}$. Pour que la méthode fonctionne, quatre moments doivent coïncider. Sur la diagonale, on atteint une erreur d'ordre $N^{-3/2+O(c_0)}$, nécessitant donc uniquement la coïncidence de deux moments. Ce développement de Taylor est calculé à l'aide des formules de variation d'Hadamard, dont la première et la deuxième sont données ci-dessous pour des valeurs propres $\lambda_k$ simples.

$$\frac{d\lambda_k(A_{i,j}(x))}{dx} = \Big\langle u_k\Big(A_{i,j}(x)\Big), \partial_x A_{i,j}(x) u_k\Big(A_{i,j}(x)\Big)\Big\rangle \tag{2.13}$$

$$\frac{d^2\lambda_k(A_{i,j}(x))}{dx^2} = -\sum_{p\neq k} \frac{\Big\langle u_k\Big(A_{i,j}(x)\Big), \partial_x A_{i,j}(x) u_k\Big(A_{i,j}(x)\Big)\Big\rangle^2}{\lambda_p(A_{i,j}(x)) - \lambda_k(A_{i,j}(x))}. \tag{2.14}$$

Les $u_j$ apparaissant dans les expressions ci-dessus sont des vecteurs propres normés associés aux valeurs propres $\lambda_j$. Ils sont uniquement déterminés à des nombres complexes de module 1 près, qui n'interviennent pas dans les formules ci-dessus. Afin de contrôler les erreurs effectuées, il est nécessaire d'utiliser la propriété de délocalisation des vecteurs propres obtenues par Erdös, Schlein et Yau. Il est également crucial de se placer sur l'évènement que les valeurs propres ne sont pas trop proches les unes des autres, afin que les termes $\frac{1}{\lambda_p - \lambda_k}$ n'explosent pas. Il faut donc ensuite estimer la probabilité de cet évènement, ce qui est fait via la loi du demi-cercle locale.

Ce théorème a ensuite été adapté pour d'autres modèles de matrices. Citons en particulier le cas des matrices de covariance, traité dans les articles [88] de Tao et Vu et [101] de Wang.

### 2.2.4 Universalité des statistiques locales

Les deux approches décrites précédemment ont été développées dans le but d'établir l'universalité des statistiques locales, en particulier des espacements. Certains des résultats obtenus ont été abordés dans la section 1.3. Ce paragraphe présente la manière dont ces théorèmes d'universalité sont obtenus à partir de la loi du demi-cercle locale (2.7) et du théorème des quatre moments 2.6. Des références possibles sont [32], [46] et [90]. Nous supposons dans ce paragraphe que toutes les matrices de Wigner considérées satisfont la condition $(C0)$.

S'appuyant sur la loi du demi-cercle locale et ses conséquences directes, Erdös, Péché, Ramirez, Schlein et Yau approfondissent les techniques développées par Johansson dans l'article [33]. Utilisant des formules explicites impliquant des déterminants, ils montrent que les espacements à l'intérieur du spectre sont bien régis par le noyau sinus $K_{\sin}$ pour des matrices du type $e^{-t/2}M_N + \sqrt{1-e^{-t}}M'_N$, avec $M_N$ une matrice du GUE et $M'_N$ une matrice de Wigner indépendante de $M_N$, et pour lesquelles $t > 0$ est autorisé à dépendre de $N$. Ils atteignent ainsi des $t$ aussi petits que $N^{-1+\delta}$ pour tout $\delta > 0$. L'universalité est donc montrée pour des matrices de Wigner complexes dont la loi des coefficients est divisible par une gaussienne, la composante gaussienne étant de taille très faible. Ajoutant à ceci un argument d'approximation d'une distribution lisse par une loi divisible par une gaussienne, l'universalité est atteinte pour des matrices dont la loi des coefficients a une densité de la forme $h(x) = e^{-g(x)}$, avec $g$ essentiellement $\mathcal{C}^6(\mathbb{R})$. Cette condition exclut malheureusement toutes les lois discrètes, et notamment les matrices de Bernoulli.

Quasiment simultanément, Tao et Vu montrent l'universalité des espacements à l'intérieur du spectre à l'aide de leur théorème des quatre moments en se basant sur les résultats de Johansson [87, 86] pour les matrices complexes et sur ceux du GOE pour les matrices réelles. Cette méthode permet alors d'atteindre des matrices de Wigner complexes dont la loi des coefficients a au moins trois points dans son support et les matrices de Wigner réelles dont les coefficients ont les mêmes quatre premiers moments que ceux d'une matrice du GOE, ce qui exclut également les matrices de Bernoulli.

Combinant les deux approches dans un article commun [34], les deux équipes établissent l'universalité pour les matrices de Wigner complexes sous l'hypothèse seule que la matrice satisfait la condition ($C0$). La remarque fondamentale de cet article est que le théorème des quatre moments de Tao et Vu reste vrai si les moments d'ordre 3 et 4 ne coïncident plus, mais sont proches. Ceci, allié aux résultats obtenus dans [33] par Erdös, Péché, Ramirez, Schlein et Yau, donne le résultat (voir par exemple le théorème 1.24).

Depuis, Erdös, Schlein, Yau *et al* ont introduit une approche dynamique permettant d'étendre les résultats de Johansson pour des $t > 0$ petits sans utiliser les formules explicites déterminantales. La méthode devient donc applicable à de nombreux modèles de matrices, notamment aux matrices de Wigner réelles et aux matrices de covariance complexes ou réelles. S'inspirant du théorème des quatre moments de Tao et Vu et de la stratégie de Lindeberg, ils mettent au point un théorème des quatre moments permettant de comparer les transformées de Stieltjes des mesures spectrales empiriques, appelé théorème de comparaison des fonctions de Green [35]. Grâce aux nombreuses formules disponibles pour les transformées de Stieltjes et les résolvantes, la démonstration est beaucoup plus simple que celle du théorème des quatre moments de Tao et Vu. La forme atteinte est également bien adaptée au travail ultérieur sur les fonctions de corrélation. Le théorème de Tao et Vu est en revanche beaucoup plus facile à utiliser lorsqu'on veut des informations sur une ou plusieurs valeur(s) propre(s) donnée(s) fixe(s) dans le spectre, comme par exemple, la valeur propre numérotée $\frac{n}{2}$. Ces informations sont atteignables par le biais du théorème des quatre moments de Erdös, Yau et Yin mais de manière moins directe.

# Chapitre 3

# Quelques aspects non-asymptotiques

Ce chapitre est consacré à certains aspects non asymptotiques de l'étude des matrices aléatoires. L'objectif est d'en expliquer l'intérêt et de présenter des résultats quantitatifs concernant les matrices de Wigner et les matrices de covariance. Les travaux effectués au cours de cette thèse se situent dans le cadre de cette étude non asymptotique. Ils concernent plus particulièrement les fluctuations de la fonction de comptage et des bornes optimales sur les valeurs propres de matrices de Wigner et de covariance. Les textes des articles publiés ou soumis suite à ces travaux figurent dans les chapitres 4, 5 et 6. Ce chapitre 3 est l'occasion de mettre ces travaux en perspective avec les résultats asymptotiques et non asymptotiques précédemment établis, ainsi que de donner des esquisses de démonstration.

## 3.1 Les enjeux de l'étude non asymptotique

Comme annoncé, nous commençons par présenter l'intérêt de l'étude quantitative. Nous avons évoqué à plusieurs reprises dans les deux chapitres précédents l'existence d'une théorie non asymptotique des matrices aléatoires, par opposition aux résultats asymptotiques. À l'origine, l'introduction des matrices aléatoires par Wishart dans le domaine des statistiques multivariées les destinait plutôt à une utilisation non asymptotique. Le domaine des statistiques est en effet intéressé par des résultats concernant des échantillons de taille grande mais fixée, et non infinie. Les matrices aléatoires apparaissent également dans les années 40 lors de la naissance de l'informatique et de l'analyse numérique matricielle dans les travaux de Goldstine et von Neumann, voir par exemple [39], [40]. Mais la théorie des matrices aléatoires ne prend vraiment son essor que lorsque Wigner les introduit en physique statistique et conjecture l'universalité du comportement spectral. Les résultats asymptotiques sont particulièrement adaptés à ce domaine de la physique statistique. En effet, les matrices aléatoires sont dans ce cadre un modèle pour des opérateurs de dimension infinie. Les propriétés asymptotiques correspondent donc aux propriétés des objets qu'on cherche à étudier. Cette conjecture d'universalité popularise les matrices aléatoires et elles sont utilisées dans des domaines divers, comme évoqué au chapitre 1. Parmi les applications possibles des matrices aléatoires, nombreuses sont celles nécessitant des résultats non asymptotiques. Ces applications mettent en effet en jeu des objets de dimension grande mais finie, comme par exemple en statistiques multivariées, pour l'étude de grands échantillons (de taille finie), en analyse fonctionnelle géométrique dans laquelle sont étudiés des opérateurs aléatoires sur des espaces de dimension finie, ou en traitement du signal, domaine dans lequel les signaux sont échantillonnés aléatoirement en un nombre grand mais fixé de points. Des références générales possibles à ce sujet sont les survols de Ledoux [58], de Rudelson et Vershynin [77] et de Vershynin [99].

L'exemple du théorème central limite, l'un des tout premiers résultats d'universalité, est symptomatique des similitudes et des intérêts divergents des propriétés asymptotiques et non asymptotiques. Posons $S_N = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} X_j$, les $X_j$ étant des variables aléatoires indépendantes, identiquement distribuées, centrées et de variance 1. Le théorème central limite affirme que $S_N$ converge en loi vers une variable aléatoire $G$ de loi gaussienne centrée réduite, fournissant ainsi une propriété asymptotique. Un échantillon $(X_1, \ldots, X_N)$ de taille fixée $N$ étant donné, il est tentant de considérer que $S_N$ suit une loi gaussienne centrée réduite. Il est alors important de se demander à quel point cette approximation est vraie, c'est-à-dire à quel point la loi de $S_N$ est proche de celle de $G$, quantitativement parlant. Le théorème de Berry-Esseen (voir par exemple [28]) fournit une version non asymptotique du théorème central limite, en contrôlant quantitativement, c'est-à-dire à $N$ fixé, l'écart entre les fonctions de répartition de ces deux lois. Sous l'hypothèse que les $X_j$ ont un troisième moment fini,

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}(S_N \leqslant x) - \mathbb{P}(G \leqslant x) \right| \leqslant \frac{C\mathbb{E}\left[ |X_1|^3 \right]}{\sqrt{N}},$$

où $C$ est une constante universelle strictement positive. Cette information est moins précise que le théorème central limite du fait de la présence d'une constante universelle qui est possiblement grande. Mais elle est vraie pour tout $N$ et peut donc être appliquée rigoureusement sur un échantillon de taille finie. Cela peut permettre par exemple de déterminer la taille d'échantillon nécessaire pour assurer un contrôle donné sur l'erreur effectuée en considérant que $S_N$ suit effectivement une loi gaussienne centrée réduite.

Une autre manière d'obtenir une version quantitative du théorème central limite est de considérer des inégalités exponentielles sur la queue de la distribution.

Ainsi, pour une variable aléatoire gaussienne centrée réduite $G$, $\mathbb{P}(|G| > x) \leqslant e^{-x^2/2}$ pour $x > 0$. Il est tout à fait naturel de se demander si $S_N$ vérifie une inégalité similaire et donc de chercher des constantes universelles $C$, $C'$ et $c$ strictement positives telles que $\mathbb{P}(|S_N| > x) \leqslant Ce^{-cx^2}$ pour $x > C'$. Par exemple, l'inégalité de Bernstein s'écrit de la manière suivante, pour des variables aléatoires $X_j$ indépendantes et bornées par $M$ (voir par exemple [95]),

$$\forall u > 0, \quad \mathbb{P}\Big(S_N - \mathbb{E}[S_N] \geqslant u\Big) \leqslant \exp\Big(-\frac{u^2}{2\,\mathrm{Var}(S_N) + Mu}\Big).$$

Ce type d'inégalité est extrêmement utile et permet par exemple d'obtenir des bornes sur les moments de $S_N$ ou de mettre au point des tests statistiques.

Le lien entre propriétés asymptotiques et non asymptotiques pour les matrices aléatoires est similaire. Considérant les théorèmes asymptotiques qui ont été montrés, notamment ceux du type théorème central limite qui concernent les fluctuations d'une quantité, on cherche à obtenir des informations quantitatives en lien avec ces théorèmes. Ainsi, les fluctuations des valeurs propres et valeurs singulières ont été étudiées de manière non asymptotique en cherchant des inégalités de déviation précises, reflétant la distribution qui contrôle les fluctuations. La fonction de comptage et les statistiques linéaires peuvent également être étudiées de cette manière. Les théorèmes asymptotiques de type théorème central limite fournissent également une moyenne et une variance asymptotiques. Un autre axe d'étude possible est donc le contrôle précis de ces moyennes et variances à taille de matrice fixée. D'autres auteurs ont cherché à quantifier la convergence du spectre global dans les théorèmes de Wigner et de Marchenko-Pastur, par le biais de l'étude de la distance entre la mesure spectrale empirique et la mesure limite. Différentes distances ont été considérées pour ce problème.

Cette étude non asymptotique s'est révélée assez ardue, notamment pour les matrices non gaussiennes. Dans le cas des modèles gaussiens, les calculs exacts possibles grâce aux propriétés évoquées dans le chapitre précédent fournissent un moyen direct d'obtenir des propriétés non asymptotiques pour ces matrices. Les calculs se révèlent toutefois compliqués en général. En revanche, les matrices non gaussiennes ne permettent plus de faire ces calculs exacts et de nouvelles techniques doivent être mises en œuvre. La méthode des moments est basée sur un contrôle de l'écart entre les moments de la quantité considérée dans les cas gaussien et non gaussien. Elle devrait donc théoriquement permettre d'étendre les inégalités obtenues dans le cas gaussien à des matrices non gaussiennes. Toutefois, cela nécessite un contrôle très précis de l'erreur, cette erreur devant être plus faible que la borne qu'on cherche à étendre. Étendre des inégalités exponentielles avec ce genre d'arguments est donc compliqué, voire impossible, les erreurs obtenues étant en général polynomiales. D'autres techniques ont donc été développées. Certains des résultats reposent sur une analyse géométrique du problème, initiée par l'article [62]. Les valeurs singulières extrêmes sont particulièrement adaptées à ces arguments géométriques. Elles s'écrivent en effet de la manière suivante. $s_N(A) = \max_{\|x\|_2=1} \|Ax\|_2$ et $s_1(A) = \min_{\|x\|_2=1} \|Ax\|_2$. Des arguments de réseau et de chaînage sont souvent utilisés pour contrôler ce minimum et ce maximum. Rudelson et Vershynin ont développé particulièrement ces techniques géométriques pour étudier la plus petite valeur singulière de matrices de covariance et de Wigner, voir par exemple [76] et [97]. D'une toute autre nature, les outils de Erdös, Schlein, Yau *et al* et ceux de Tao et Vu, développés dans le cadre de l'étude asymptotique, sont des outils non asymptotiques. Initialement, les auteurs ne destinaient pas ces outils à une telle utilisation mais leurs démonstrations fournissent des estimations quantitatives relativement précises sur les erreurs effectuées lors de la comparaison entre modèles gaussiens et non gaussiens. Depuis leur publication, ces outils ont été adaptés et utilisés de multiples fois, dans le cadre de l'étude asymptotique comme dans celui de l'étude non asymptotique des propriétés spectrales. Citons par exemple l'article de Döring et Eichelsbacher [29] sur les déviations modérées de la fonction de comptage, ainsi que celui de Bao, Pan et Zhou [14] sur les statistiques linéaires partielles.

Le champ des applications possibles est riche et divers. En statistique par exemple, l'étude non asymptotique des matrices de covariance permet de donner des bornes quantitatives sur la taille d'échantillon nécessaire à une estimation précise de la matrice de covariance, voir par exemple les travaux de Vershynin à ce sujet [98]. Le compressed sensing est un autre exemple de domaine dans lequel les propriétés non asymptotiques se sont révélées utiles. Se basant sur le fait que le signal envoyé est parcimonieux (ce qui est le cas de nombreux signaux réels), le compressed sensing permet de réduire significativement le nombre de mesures à effectuer sur ce signal pour le reconstruire de manière exacte. Imaginons qu'on envoie le signal $x_0 \in \mathbb{R}^n$ dont le support (inconnu) est supposé de taille au plus $s$, avec $s$ très petit devant $n$. Le destinataire reçoit le signal $y = Ax_0 \in \mathbb{R}^N$, avec $N$ également petit devant $n$ mais plus grand que $s$, sans quoi la reconstruction du signal est impossible. Le problème est mal conditionné mais l'hypothèse de parcimonie du signal permet de passer outre cette difficulté. Si la matrice de mesures $A$ est quasiment une isométrie pour les vecteurs $s$ parcimonieux, il est possible de reconstruire le signal. Le fait qu'elle soit une quasi isométrie se traduit par le biais de la propriété d'isométrie restreinte de constante $\delta$, que $A$ vérifie si, pour tout vecteur $x \in \mathbb{R}^n$ $s$ parcimonieux,

$$(1-\delta)\|x\|_2^2 \leqslant \|Ax\|_2^2 \leqslant (1+\delta)\|x\|_2^2.$$

Candès et Tao ont révolutionné le domaine du traitement du signal en montrant que, si la matrice $A$ satisfait la propriété d'isométrie restreinte avec une constante $\delta < \sqrt{2}-1$, alors minimiser la norme $\ell_1$ de $x$ sous la contrainte $Ax = y$ permet de reconstruire de manière exacte le signal $s$ parcimonieux envoyé. Toutefois, il est difficile à l'heure actuelle de décider en temps raisonnable si une matrice $A$ donnée vérifie la propriété d'isométrie restreinte. En revanche, il est possible de montrer que certaines matrices aléatoires vérifient cette propriété avec grande probabilité (quantifiée). Par exemple, une matrice du LOE possède cette propriété pour $s \leqslant Cm \log(\frac{m}{n})$. Ceci fournit une nouvelle manière d'échantillonner les signaux produisant des échantillons de taille beaucoup plus faible que ce qui était fait auparavant et permettant de reconstruire de manière exacte le signal initial. L'article [20] est un survol récent sur ce sujet.

Dans la suite de ce chapitre, nous présentons des versions non asymptotiques de certains résultats asymptotiques abordés au chapitre 1, parmi lesquels figurent les travaux effectués durant cette thèse. Nous commençons par des résultats quantitatifs connus. Les premiers concernent certaines statistiques globales des valeurs propres. Le paragraphe 3.2.1 présente les résultats concernant la vitesse de convergence du spectre. Cette vitesse de convergence du spectre a été explorée au travers de diverses distances entre la mesure spectrale empirique et la loi du demi-cercle ou de Marchenko-Pastur, nous abordons ici les résultats obtenus sur la distance de Kolmogorov et la distance de Wasserstein d'ordre 1. Nous présentons ensuite au paragraphe 3.2.2 des bornes sur l'espérance et la variance de la fonction de comptage $\mathcal{N}_t$ et présentons des inégalités de déviation satisfaites par $\mathcal{N}_t$ dans le cas de matrices gaussiennes. Les résultats suivants concernent les statistiques locales des valeurs propres, plus particulièrement des valeurs propres extrêmes. Le paragraphe 3.2.3 s'intéresse aux inégalités de déviation précises pour les valeurs propres extrêmes, reflétant la convergence vers la loi de Tracy-Widom, de la forme

$$\mathbb{P}(\lambda_N \geqslant 2 + \varepsilon) \leqslant Ce^{-cN\varepsilon^{3/2}}.$$

De nombreuses inégalités sont disponibles pour les matrices gaussiennes et certaines sont partiellement étendues à des matrices plus générales. Enfin, le paragraphe 3.2.4 présente le problème d'inversibilité des matrices aléatoires, c'est-à-dire l'étude de la taille typique de la plus petite valeur singulière. Nous présentons ensuite dans une troisième section les résultats obtenus durant la thèse et nous donnons la trame des démonstrations associées, dont le détail est contenu dans les chapitre suivants. Ces résultats concernent la vitesse de convergence du spectre en termes de distance de Wasserstein d'ordre 2, des bornes sur la variance et l'espérance de la fonction de comptage, des inégalités de déviation pour les valeurs propres à l'intérieur du spectre ainsi que des bornes sur la variance des valeurs propres. Ils sont valables pour des familles assez larges de matrices de Wigner et de covariance, hormis en ce qui concerne les inégalités de déviation pour les valeurs propres à l'intérieur du spectre, qui ne sont établies que pour les modèles gaussiens.

## 3.2 Quelques résultats quantitatifs connus

Comme annoncé, nous présentons dans les paragraphes suivants des résultats non asymptotiques correspondants à certains théorèmes asymptotiques décrits dans le chapitre 1. Nous considérons en particulier les théorèmes de Wigner 1.3 et de Marchenko-Pastur 1.4, les fluctuations de la fonction de comptage (théorème 1.6) et celles des valeurs propres extrêmes (théorème 1.17).

### 3.2.1 Vitesse de convergence du spectre

Les théorèmes de Wigner 1.3 et de Marchenko-Pastur 1.4 établissent la convergence du spectre vers la loi du demi-cercle dans le cas des matrices de Wigner et vers la loi de Marchenko-Pastur pour les matrices de covariance. Différents auteurs se sont attachés à estimer la vitesse de convergence dans ces théorèmes, au travers de plusieurs distances entre la mesure spectrale empirique et la loi limite.

La distance de Kolmogorov, définie de la manière suivante, a été étudiée de façon approfondie,

$$d_K(L_N, \rho_{sc}) = \sup_{x \in \mathbb{R}} \left| \tfrac{1}{N} \mathcal{N}_x - G_{sc}(x) \right|,$$

$G_{sc}$ étant la fonction de répartition de la loi du demi-cercle. Götze et Tikhomirov ont récemment établi des bornes précises sur cette distance [45]. Plus précisément, ils ont démontré qu'il existe des constantes universelles $C$, $C'$, $c$ et $c'$ positives telles que pour toute matrice de Wigner satisfaisant la condition $(C0)$,

$$\mathbb{P}\Big( d_K(L_N, \rho_{sc}) > \frac{(\log N)^C}{N} \Big) \leqslant C' e^{-c' (\mathrm{llog} N)^c}.$$

Rappelons la notation $\mathrm{llog} N = \log\log N$. Ceci implique que

$$\mathbb{E}\big[ d_K(L_N, \rho_{sc}) \big] \leqslant \frac{(\log N)^C}{N},$$

pour $C$ une constante universelle. Notons que ces constantes dépendent de celles intervenant dans la définition de la condition $(C0)$ 2.2. Dans une autre contribution récente [44], les mêmes auteurs établissent le résultat analogue sur la distance $d_K(L_{m,n}, \mu_{MP}(\rho))$ pour les matrices de covariance. Dans ces deux articles, les auteurs utilisent les propriétés de la transformée de Stieltjes pour montrer les résultats précédemment cités.

D'autres auteurs se sont intéressés aux distances de Wasserstein et notamment à la distance de Wasserstein d'ordre 1. Pour des mesures de probabilité $\mu$ et $\nu$ sur $\mathbb{R}$, ces distances sont définies par

$$W_p(\mu, \nu) = \inf \left( \int_{\mathbb{R}^2} |x - y|^p \, d\pi(x, y) \right)^{1/p},$$

où la borne inférieure est prise sur les mesures de probabilité $\pi$ sur $\mathbb{R}^2$ dont la première marginale est $\mu$ et la seconde est $\nu$. Ces distances sont aisément ordonnées : si $p \leqslant q$, $W_p(\mu, \nu) \leqslant W_q(\mu, \nu)$ pour toutes mesures de probabilité $\mu$ et $\nu$. Par ailleurs, il est possible de comparer $W_1$ et la distance de Kolmogorov. En effet, si $\nu$ a une densité bornée par rapport à la mesure de Lebesgue, $d_K(\mu, \nu) \leqslant c\sqrt{W_1(\mu, \nu)}$, où $c > 0$ dépend de la borne sur cette densité. Cette relation est due à une écriture de $W_1$ en termes de fonctions lipschitziennes.

$$W_1(\mu, \nu) = \sup_{f\,1\text{-Lipschitz}} \left( \int_{\mathbb{R}} f d\mu - \int_{\mathbb{R}} f d\nu \right) \tag{3.1}$$

(voir par exemple [100]). Récemment, Meckes et Meckes [65] ont utilisé l'écriture (3.1) pour établir des inégalités de concentration sur $W_1(L_N, \mathbb{E}[L_N])$. Ils en déduisent la borne suivante :

$$\mathbb{E}\Big[W_1(L_N, \mathbb{E}[L_N])\Big] \leqslant CN^{-2/3}.$$

La distance de Wasserstein d'ordre 2 entre la mesure spectrale empirique et la loi limite a été étudiée durant cette thèse, pour des matrices de Wigner et de covariance assez générales. Le paragraphe 3.3.1 présente les résultats obtenus et les idées de la démonstration, dont le détail est contenu dans les sections 5.4 et 6.4.

### 3.2.2 Variance et déviations de la fonction de comptage

Nous nous penchons maintenant sur l'étude non asymptotique de la fonction de comptage des valeurs propres $\mathcal{N}_t$. Rappelons que $\mathcal{N}_t$ est le nombre de valeurs propres dans l'intervalle $(-\infty, t]$. Nous avons vu au chapitre 1 que $\mathcal{N}_t$ satisfait un théorème central limite (1.13), que nous rappelons ici dans le cas où $t$ est fixé, à l'intérieur du spectre limite $[-2, 2]$

$$\frac{\mathcal{N}_t(W_N) - \mathbb{E}\big[\mathcal{N}_t(W_N)\big]}{\sqrt{\mathrm{Var}\left(\mathcal{N}_t(W_N)\right)}} \underset{N\to\infty}{\to} \mathcal{N}(0, 1),$$

pour $t \in (-2, 2)$. Gustavsson montre ce résultat dans le cas de matrices du GUE [49]. Il calcule ensuite des asymptotiques pour l'espérance et la variance de $\mathcal{N}_t$ (1.14) et (1.15) pour en déduire un théorème central limite « numérique » (1.16). De la démonstration de l'asymptotique de la variance, il est possible de déduire une borne non asymptotique pour cette variance. Pour tout $\delta > 0$, il existe $c_\delta > 0$ dépendant uniquement de $\delta$ telle que pour tout $t \in (-2 + \delta, 2 - \delta)$,

$$\mathrm{Var}\left(\mathcal{N}_t(W_N)\right) \leqslant c_\delta \log N. \tag{3.2}$$

Su effectue les calculs analogues dans le cas de matrices du LUE (1.18), voir [83]. Il est possible de déduire de ses calculs la même relation sur $\mathrm{Var}\Big(\mathcal{N}_t(S_{m,n})\Big)$ pour $t \in (a_{m,n}+\delta, b_{m,n}-\delta)$. En ce qui concerne les espérances, les calculs de Gustavsson et Su fournissent des majorations mais Götze et Tikhomirov ont montré pour les matrices du GUE le résultat plus précis suivant [43]. Il existe une constante universelle $C > 0$ telle que

$$\sup_{t\in\mathbb{R}} \Big|\mathbb{E}[\mathcal{N}_t] - N\int_{-\infty}^{t} \rho_{sc}(x)\,dx\Big| \leqslant C. \tag{3.3}$$

Dans ce même article, ils montrent un résultat similaire pour les matrices de covariance gaussiennes complexes. Sous l'hypothèse qu'il existe $A_1$ et $A_2$ des constantes telles que $1 < A_1 \leqslant \frac{m}{n} \leqslant A_2$, (3.3) est établi avec une constante $C$ qui dépend uniquement de $A_1$ et $A_2$. Dans le cas où $m = n$, le résultat est également vrai, la constante $C$ étant universelle. Durant cette thèse, nous avons étendu les bornes sur la variance de la fonction de comptage (3.2) au cas des matrices non gaussiennes. Les résultats et les idées de démonstration sont présentés au paragraphe 3.3.2, le détail est disponible au chapitre 4.

Un autre angle d'approche de l'étude non asymptotique de la fonction de comptage est l'établissement d'inégalités de déviation. Dans le cas des matrices du GUE ou du LUE, les propriétés déterminantales fournissent immédiatement ces inégalités. Rappelons en effet que la fonction de comptage a la même loi qu'une somme de variables aléatoires indépendantes de Bernoulli (voir 2.1). On peut alors par exemple utiliser l'inégalité de Bernstein. Combinant ceci avec (3.3), on obtient pour tout $u \geqslant 0$

$$\mathbb{P}\Big(\Big|\mathcal{N}_t - N\int_{-\infty}^{t} \rho_{sc}(x)\,dx\Big| \geqslant u + C\Big) \leqslant 2\exp\Big(-\frac{u^2}{2\,\mathrm{Var}(\mathcal{N}_t)+u}\Big). \tag{3.4}$$

Un résultat analogue existe pour les matrices du LUE, ainsi que pour celles du GOE et du LOE. Les inégalités de déviation dans le cas réel sont établies à l'aide des relations d'entrelacement décrites au chapitre précédent.

Le cas des matrices non gaussiennes est plus difficile à appréhender. Dans leurs travaux sur la loi du demi-cercle locale et sur la rigidité des valeurs propres, Erdös, Yau et Yin montrent une inégalité de déviation pour des matrices de Wigner assez générales (2.8), que nous rappelons ci-dessous :

$$\mathbb{P}\Big(\sup_{|t|\leqslant 5}\Big|\mathcal{N}_t - N\int_{-\infty}^{t} \rho_{sc}(x)\,dx\Big| > (\log N)^{\mathrm{cllog}N}\Big) \leqslant C\exp\Big(-c'(\log N)^{C'\mathrm{llog}N}\Big).$$

Pillai et Yin montrent par la suite un résultat similaire sur les matrices de covariance, voir [75]. Tao et Vu, dans leur récente contribution [89], affinent cette inégalité de déviation et montrent pour $u \geqslant 0$

$$\mathbb{P}\Big(\Big|\mathcal{N}_t - N\int_{-\infty}^{t} \rho_{sc}(x)\,dx\Big| > u\Big) \leqslant CN^c e^{-C'u^{c'}},$$

les constantes $C$, $C'$, $c$ et $c'$ dépendant uniquement des constantes dans la condition $(C1)$ (définition 2.4). Cette inégalité n'est exploitable que lorsque $u$ est assez grand, de l'ordre de $(\log N)^a$, $a$ étant une constante positive. La question d'une inégalité de déviation précise pour des $u$ petits reste ouverte dans le cas général.

Il est légitime de se poser ces mêmes questions non asymptotiques pour les statistiques linéaires des valeurs propres $\mathcal{N}_N[\varphi]$ et $\mathcal{N}_{m,n}[\varphi]$, notamment la question des inégalités de déviation. Lorsque la fonction $\varphi$ est suffisamment régulière, c'est-à-dire lorsqu'elle est lipschitzienne, les outils de concentration de la mesure permettent d'obtenir de telles inégalités dans le cas de matrices gaussiennes. En effet, pour une fonction $F : \mathbb{R}^N \to \mathbb{R}$ lipschitzienne, $\mathbb{R}^N$ étant muni de la norme 2, pour tout $r > 0$,

$$\mathbb{P}\Big(\Big|F(\lambda_1,\dots,\lambda_N) - \mathbb{E}[F(\lambda_1,\dots,\lambda_N)]\Big| \geqslant r\Big) \leqslant 2e^{-r^2/2\|F\|_{Lip}},$$

les $\lambda_j$ étant les valeurs propres d'une matrice $M_N$ du GUE. Si on considère la fonction $F$ définie par $F(x_1,\dots,x_N) = \frac{1}{N}\sum_{j=1}^N \varphi(x_j)$, avec $\varphi$ une fonction lipschitzienne de $\mathbb{R}$ dans $\mathbb{R}$, on obtient, pour tout $u > 0$,

$$\mathbb{P}\Big(\frac{1}{N}\Big|\mathcal{N}_N[\varphi] - \mathbb{E}[\mathcal{N}_N[\varphi]]\Big| \geqslant u\Big) \leqslant 2e^{-N^2u^2/2\|\varphi\|_{Lip}}.$$

Les valeurs propres considérées sont ici celles de la matrice $W_N = \frac{1}{\sqrt{N}}M_N$. Une référence générale possible pour ces questions de concentration de la mesure est l'ouvrage de Ledoux [57]. Ces outils ont été utilisés également pour obtenir des inégalités de déviation pour les statistiques linéaires dans le cas de matrices non gaussiennes, les coefficients de la matrice devant vérifier certaines conditions, voir par exemple [47] et [48]. Dès que la fonction test devient moins régulière, les inégalités ci-dessus ne fournissent plus aucun renseignement. Rien ne semble connu à ce jour permettant de faire un lien entre ces inégalités obtenues pour une fonction de test régulière et celles obtenues pour la fonction de comptage.

Nous nous intéressons maintenant aux valeurs propres extrêmes. Ces valeurs propres sont étudiées asymptotiquement et non asymptotiquement depuis longtemps. Une des raisons de cette étude approfondie est leur lien avec le conditionnement $\kappa$, défini par $\kappa(A) = \frac{s_{\max}(A)}{s_{\min}(A)}$ où $s_{\max}$ et $s_{\min}$ sont respectivement la plus grande et la plus petite valeurs singulières de la matrice $A$. Ce conditionnement influe sur la précision et parfois le temps de calcul de certains algorithmes. Il est donc utile de connaitre la taille typique de $\kappa(A)$ et pour cela d'étudier non asymptotiquement les valeurs propres et valeurs singulières extrêmes.

Dans les deux prochains paragraphes, nous présentons des versions non asymptotiques des propriétés locales décrites dans le paragraphe 1.2.2, c'est-à-dire les fluctuations des valeurs propres extrêmes (théorèmes 1.17, 1.18, 1.19). Ces résultats peuvent parfois être établis en suivant la démonstration de la propriété asymptotique correspondante mais il est souvent nécessaire d'utiliser des techniques plus fines.

### 3.2.3 Inégalités de déviation pour les valeurs propres extrêmes

Les valeurs propres extrêmes des matrices de Wigner et de covariance ont des fluctuations régies par la loi de Tracy-Widom (voir théorème 1.17), hormis la plus petite valeur propre des matrices de covariance dans le cas où $\frac{m}{n} \to 1$. De plus, les fluctuations sont d'ordre $N^{-2/3}$ (ou $n^{-2/3}$ dans le cas des matrices de covariance). Nous sommes intéressés par des inégalités de déviation vérifiées par ces valeurs propres, qui reflètent la bonne vitesse de convergence et le fait que c'est la loi de Tracy-Widom qui régit les fluctuations. La fonction de répartition de la loi de Tracy-Widom présente des comportements différents en $\pm\infty$. Ainsi, elle est de l'ordre de $e^{ct^3}$ pour $t \to -\infty$ et de l'ordre de $e^{-c't^{3/2}}$ pour $t \to +\infty$. Les inégalités recherchées sont donc de la forme $\mathbb{P}(\lambda_N \geqslant 2+\varepsilon) \leqslant Ce^{-N\varepsilon^{3/2}/C}$ et $\mathbb{P}(\lambda_N \leqslant 2-\varepsilon) \leqslant Ce^{-N^2\varepsilon^3/C}$, pour $\varepsilon > 0$. Ces inégalités seront précises dans le cas où $\varepsilon$ est petit. Le théorème asymptotique donne en effet la convergence de $\mathbb{P}(\lambda_N \geqslant 2+\varepsilon)$ vers la fonction de répartition de la loi de Tracy-Widom quand $\varepsilon = tN^{-2/3}$. Lorsque $\varepsilon$ est grand, on s'attend à une majoration de ces probabilités par $e^{-CN\varepsilon^2}$.

Même dans le cas des matrices gaussiennes, cette étude se révèle compliquée. Dans [5], Aubrun suit pas à pas la démonstration de Tracy et Widom et contrôle les déterminants de Fredholm mis en jeu pour obtenir effectivement l'inégalité attendue sur la plus grande valeur propre de matrices du GUE. Pour tout $\varepsilon \geqslant 0$,

$$\mathbb{P}(\lambda_N \geqslant 2+\varepsilon) \leqslant Ce^{-N\varepsilon^{3/2}/C}, \tag{3.5}$$

la constante $C > 0$ étant universelle. Il est également possible d'obtenir cette inégalité à partir des travaux de Johansson [52], comme annoncé dans [58]. L'inégalité obtenue est plus précise car elle reflète à la fois le comportement pour $\varepsilon$ petit et pour $\varepsilon$ grand :

$$\mathbb{P}(\lambda_N \geqslant 2+\varepsilon) \leqslant Ce^{-N\max(\varepsilon^2,\ \varepsilon^{3/2})/C}.$$

Dans [36], Feldheim et Sodin mettent au point une variante de la méthode des moments qui leur permet d'établir en même temps l'universalité des propriétés asymptotiques et non asymptotiques des valeurs propres extrêmes des matrices de Wigner et de covariance. L'inégalité (3.5) est alors étendue à des matrices de Wigner plus générales, dont la loi des coefficients est symétrique et sous-gaussienne. L'article récent de Tao et Vu [89] fournit une inégalité de déviation pour les valeurs propres extrêmes de matrices de Wigner beaucoup plus générales reflétant bien le taux de convergence en $N^{-2/3}$.

$$\forall u \geqslant C \log N, \quad \mathbb{P}\Big( |\lambda_N - 2| \geqslant \frac{u}{N^{2/3}} \Big) \leqslant e^{-cu^{c'}}.$$

Notons qu'elle concerne autant la déviation à droite que celle à gauche. Toutefois, la majoration n'est pas optimale dans le sens où elle ne fournit pas d'information pour des $u$ petits et où elle ne reflète pas le comportement de la queue de distribution de la loi de Tracy-Widom.

Les déviations à gauche sont plus difficiles à étudier. Pour les matrices gaussiennes, Ledoux et Rider proposent dans [59] des démonstrations unifiées, à partir de la représentation tridiagonale (1.5), pour les inégalités de déviation à droite et à gauche du bon ordre pour tous les $\beta$ ensembles gaussiens, ce qui recouvre le GUE et le GOE. On obtient pour la déviation à gauche,

$$\mathbb{P}\Big( \lambda_N \leqslant 2 - \varepsilon \Big) \leqslant C \exp\Big( - N^2 \varepsilon^3 / C \Big). \tag{3.6}$$

Cet article fournit les mêmes inégalités de déviation optimales à droite et à gauche pour la plus grande valeur propre des matrices du LUE et du LOE. Il traite également le cas plus difficile de la plus petite valeur propre des matrices du LUE et du LOE, uniquement pour les déviations à gauche.

$$\mathbb{P}(\lambda_1 \leqslant a_{m,n} - \varepsilon) \leqslant C e^{-n\varepsilon^{3/2}/C}, \tag{3.7}$$

pour $\varepsilon > 0$ dans le cas où $m \geqslant A_1 n$ avec $A_1 > 1$. Feldheim et Sodin, dans l'article cité précédemment, établissent l'universalité de ces inégalités de déviation pour la plus grande et pour la plus petite valeurs propres de matrices de covariance sous l'hypothèse que les coefficients ont une loi symétrique et sous-gaussienne.

### 3.2.4 Problème d'inversibilité de matrices aléatoires

Nous évoquions précédemment le problème de l'estimation du conditionnement des matrices aléatoires. Pour cela, il est nécessaire de connaître la taille typique des valeurs singulières extrêmes de ces matrices. Plaçons-nous dans le cas des matrices de covariance $S_{m,n} = \frac{1}{m} X^* X$. La plus grande valeur singulière $s_n(\frac{1}{\sqrt{m}} X)$ est alors la racine carrée de $\lambda_n$. Les inégalités de déviation décrites au paragraphe précédent donnent donc les informations nécessaires sur $s_n(X)$. De plus, lorsque $m \geqslant A_1 n$ avec $A_1 > 1$, la plus petite valeur singulière $s_1(X)$, qui s'exprime aisément en fonction de $\lambda_1$, a une taille typique de $\sqrt{m} - \sqrt{n}$. La question la plus difficile est celle de la taille typique de $s_1$ lorsque $\frac{m}{n} \to 1$, l'estimation précédente devenant totalement inutile. Dans le cas des matrices gaussiennes, Edelman a calculé à partir de la loi jointe des valeurs propres la loi exacte de la plus petite valeur singulière pour $m = n$ [30]. Il est aisé d'en déduire que

$$\mathbb{P}\Big(s_1(X) \leqslant \varepsilon n^{-1/2}\Big) \leqslant C\varepsilon.$$

Par le biais d'arguments géométriques et d'une analyse fine de la structure des colonnes de la matrice aléatoire $X$, Rudelson et Vershynin montrent dans [76] l'estimation suivante lorsque $m = n$, sous une hypothèse de décroissance sous-gaussienne de la loi des coefficients.

$$\mathbb{P}\Big(s_1(X) \leqslant \varepsilon n^{-1/2}\Big) \leqslant C\varepsilon + c^n,$$

les constantes $C > 0$ et $c \in (0, 1)$ dépendant uniquement du moment sous-gaussien de la loi des coefficients. Lorsque $\varepsilon = 0$, on obtient la probabilité que $X$ ne soit pas inversible, qui est non nulle, contrairement au cas gaussien. Unifiant le résultat précédent avec celui sur les matrices rectangulaires, ils obtiennent pour tous $m \geqslant n$, $\varepsilon \geqslant 0$,

$$\mathbb{P}\Big(s_1(X) \leqslant \varepsilon(\sqrt{m} - \sqrt{n})\Big) \leqslant (C\varepsilon)^{m-n+1} + c^m,$$

sous les mêmes hypothèses très peu restrictives sur la loi des coefficients. Le terme $c^n$ traduit le fait que la matrice $X$ n'est en général pas inversible mais qu'elle tend à être inversible quand ses dimensions deviennent grandes. Notons par exemple le cas des matrices $X$ carrées de taille $n$ dont les coefficients sont des variables aléatoires de Bernoulli à valeurs $\pm 1$. Ces matrices ont une probabilité non nulle d'être singulières. En effet, deux colonnes sont égales avec probabilité $\frac{1}{2^n}$. La probabilité de non inversibilité dans ce cas est conjecturée d'ordre $c^n$ avec $c = \frac{1}{2} + o(1)$ mais la meilleure constante connue actuellement vaut $\frac{1}{\sqrt{2}} + o(1)$ [18]. Par ailleurs, Vershynin a récemment établi un résultat similaire pour la plus petite valeur singulière des matrices de Wigner, qui est la plus petite valeur propre *en valeur absolue* de la matrice et non pas la plus petite valeur propre $\lambda_1$ [97].

## 3.3 Résultats non asymptotiques

Le but de cette section est de présenter les travaux effectués au cours de cette thèse dans le cadre de l'étude non asymptotique des matrices aléatoires. Ces travaux se basent sur deux éléments principaux : les propriétés de la fonction de comptage dans le cas gaussien, avec notamment l'écriture sous forme de somme de variables aléatoires de Bernoulli indépendantes (2.1), et la combinaison du théorème des quatre moments de Tao et Vu 2.6 avec le théorème de localisation de Erdös, Yau et Yin 2.3.

Cette section est divisée en quatre paragraphes. Le premier concerne la vitesse de convergence du spectre, qui est étudiée ici en termes de distance de Wasserstein d'ordre 2. Le deuxième traite de l'extension au cas de matrices non gaussiennes des fluctuations et des bornes sur la variance de la fonction de comptage. Les deux derniers s'intéressent aux valeurs propres individuelles. Nous établissons tout d'abord des inégalités de déviation pour les valeurs propres à l'intérieur du spectre pour des matrices gaussiennes. Nous nous intéressons ensuite établir des bornes optimales sur la variance des valeurs propres. Chacun des paragraphes contient les énoncés des résultats et les idées principales des démonstrations. Nous renvoyons aux chapitres suivants pour les détails.

### 3.3.1 Vitesse de convergence du spectre

Dans l'article [23] dont le texte est contenu dans le chapitre 5, nous établissons des bornes sur la variance des valeurs propres et en déduisons une majoration sur la distance de Wasserstein d'ordre 2 entre $L_N$ et $\rho_{sc}$. Pour cela, nous utilisons une autre écriture de $W_2$, à savoir

$$W_2(\mu,\nu)^2 = \int_0^1 \Big(F_\mu^{-1}(x) - F_\nu^{-1}(x)\Big)^2 \, dx,$$

où $F_\mu^{-1}$ est l'inverse généralisé de la fonction de répartition $F_\mu$ associée à $\mu$ (et de même pour $F_\nu^{-1}$). Appliquant cette formule pour $\mu = L_N$ et $\nu = \rho_{sc}$, les propriétés de ces deux mesures permettent de majorer $\mathbb{E}\Big[W_2(L_N,\rho_{sc})^2\Big]$ par $\frac{1}{N}\sum_{j=1}^N \mathrm{Var}(\lambda_j) + \frac{C}{N^2}$. Les bornes obtenues dans ce même article et présentées dans 3.3.4 sur les variances des valeurs propres donnent la majoration suivante pour des matrices de Wigner satisfaisant la condition $(C0)$ et ayant les mêmes quatre premiers moments qu'une matrice du GUE

$$\mathbb{E}\Big[W_2^2(L_N,\rho_{sc})\Big] \leqslant C\frac{\log N}{N^2}, \tag{3.8}$$

$C$ étant une constante universelle. Le résultat est également vrai pour les matrices de Wigner réelles. Le cas des matrices de covariance est très similaire. Toutefois, l'établissement des bornes sur la variance des valeurs propres repose sur des calculs de variance pour la fonction de comptage réalisés par Su [83]. Or il ne traite pas le bord gauche, ce qui ne permet donc pas de conclure pour les valeurs propres proches du bord gauche du spectre. Cependant, il est fortement probable que ces calculs au niveau du bord gauche soient une simple adaptation de ceux au niveau du bord droit, dans le cas où le rapport entre le nombre de lignes et celui de colonnes est borné inférieurement par $A_1 > 1$. Dans ce cas, on obtiendrait la même majoration pour la distance de Wasserstein d'ordre 2 entre $L_{m,n}$ et $\mu_{m,n}$. Le détail des calculs, très similaires à ceux effectués dans le cas des matrices de Wigner, est présenté au chapitre 6. Lorsque $\frac{m}{n} \to 1$, c'est-à-dire dans le cas d'un bord dur, le théorème central limité établi par Edelman 1.19 montre que la variance de la plus petite valeur propre est asymptotiquement de l'ordre de $n^{-4}$, un ordre radicalement différent de celui en $n^{-4/3}$ « habituel » au bord du spectre. Les variances des valeurs propres sur le bord gauche sont donc très certainement d'ordre très différent de celui dans le cas du bord mou. Peu de choses sont connues et on ne peut conclure directement quant à la majoration de la distance. Il est cependant probable que ces termes ne soient pas prépondérants dans la majoration et qu'on puisse toujours conclure quant à la borne (3.8)

### 3.3.2 Variance de la fonction de comptage

Dans un article écrit conjointement avec V. Vu [24] et dont le texte est présenté au chapitre 4, nous montrons l'universalité du théorème central limite et des asymptotiques de Gustavsson.

**Théorème 3.1** (Théorème 4.2)**.** *Soit $M_N$ une matrice hermitienne de Wigner satisfaisant la condition $(C0)$ et dont les coefficients ont les mêmes quatre premiers moments que ceux d'une matrice du GUE. On pose $W_N = \frac{1}{\sqrt{N}} M_N$. Alors, pour tout $x \in (-2, 2)$,*

$$\mathbb{E}[\mathcal{N}_x] = N\rho_{sc}\Big((-\infty, x]\Big) + o(1) \quad et \quad \mathrm{Var}(\mathcal{N}_x) = \Big(\frac{1}{2\pi^2} + o(1)\Big) \log N.$$

*Par conséquent, la fonction de comptage $\mathcal{N}_x$ satisfait le théorème central limite suivant :*

$$\frac{\mathcal{N}_x - \mathbb{E}[\mathcal{N}_x]}{\sqrt{\mathrm{Var}(\mathcal{N}_x)}} \underset{N\to\infty}{\to} \mathcal{N}(0,1).$$

Pour cela, nous nous basons sur le théorème central limite numérique vérifié par le GUE (1.16) et nous l'étendons au cas général à l'aide du théorème des quatre moments de Tao et Vu 2.6, comme décrit au chapitre 2. Il est plus difficile de montrer l'universalité de (1.14) et (1.15). Nous expliquons la démarche sur le calcul de l'espérance. Le raisonnement est identique pour la variance mais un peu plus technique. L'idée est de majorer l'écart entre $\mathbb{E}[\mathcal{N}_t(W_N)]$ et $\mathbb{E}[\mathcal{N}_t(W'_N)]$, pour $W_N$ une matrice de Wigner générale renormalisée et $W'_N$ une matrice du GUE renormalisée. Nous écrivons

$$\mathcal{N}_t = \mathcal{N}_t(W_N) = \sum_{j=1}^{N} A_j \quad \text{avec} \quad A_j = \mathbb{1}_{\lambda_j \leqslant t},$$

et $\mathcal{N}'_t = \mathcal{N}_t(W'_N) = \sum_{j=1}^N A'_j$ de manière similaire. Il nous faut contrôler $|\mathbb{E}[A_j] - \mathbb{E}[A'_j]|$. Le théorème de localisation de Erdös, Yau et Yin 2.3 permet de montrer que, dès que la position théorique $\gamma_j$ de la valeur propre est assez loin de $t$, $\mathbb{E}[A_j]$ et $\mathbb{E}[A'_j]$ vont être tous les deux très petits ou tous les deux très proches de 1. La différence est donc très petite, suffisamment petite pour que la somme de tous ces termes (au plus $N$) tende vers 0, à une vitesse contrôlable. Par ailleurs, il est possible de montrer qu'il y a peu de termes pour lesquels $\gamma_j$ est proche de $t$ (la notion de proximité nécessaire ayant été définie à l'étape précédente), ce nombre de termes étant d'ordre une grande puissance de $\log N$. Pour ces termes, l'utilisation du théorème des quatre moments de Tao et Vu 2.6 fournit une majoration directe : chacun d'entre eux est inférieur à $N^{-c_0}$, ce qui permet de conclure. Nous avons conservé dans cet article l'écriture des résultats sur l'espérance et la variance sous forme asymptotique. Il est toutefois possible d'en extraire un résultat non asymptotique sur la variance similaire à celui de Gustavsson (3.2). Tant que $t$ est dans $[-2+\delta, 2-\delta]$, la différence $|\operatorname{Var}(\mathcal{N}_t) - \operatorname{Var}(\mathcal{N}'_t)|$ majorée comme indiqué précédemment est bornée uniformément en $t$. On en déduit alors qu'il existe une constante $c'_\delta$ dépendant uniquement de $\delta$ telle que

$$\operatorname{Var}(\mathcal{N}_t) \leqslant c'_\delta \log N,$$

pour toute matrice de Wigner satisfaisant les conditions d'application des théorèmes 2.3 et 2.6. Ce raisonnement ne permet toutefois pas d'étendre le résultat de Götze et Tikhomirov (3.3) à des matrices de Wigner plus générales. En effet, la procédure décrite précédemment ne fonctionne plus si $t$ est trop proche d'un des bords du spectre. Le contrôle du nombre de termes $|\mathbb{E}[A_j] - \mathbb{E}[A'_j]|$ pour lesquels on applique le théorème des quatre moments n'est pas assez précis pour montrer que la somme de ces termes tend vers 0. Les résultats analogues sont disponibles pour les matrices de covariance, les démonstrations sont identiques.

### 3.3.3 Déviations des valeurs propres individuelles à l'intérieur du spectre

Dans l'article [23], nous nous intéressons aux valeurs propres individuelles à l'intérieur du spectre. Ces valeurs propres à l'intérieur du spectre satisfont un théorème central limite, la distribution limite étant gaussienne (théorèmes 1.20 et 1.21). Dans les chapitres 5 et 6, nous établissons des inégalités de déviation pour chacune des valeurs propres à l'intérieur du spectre lorsque les matrices sont gaussiennes, le chapitre 5 étant consacré aux matrices de Wigner, le 6 aux matrices de covariance. Précisément, nous montrons pour une matrice du GUE la proposition suivante.

**Proposition 3.2** (Proposition 5.6)**.** *Soit* $0 < \eta \leqslant \frac{1}{2}$ *et soit* $\eta N \leqslant j \leqslant (1-\eta)N$. *Il existe* $C > 0$, $c > 0$, $c' > 0$ *et* $\delta \in (0,2)$ *(dépendant uniquement de* $\eta$*) telles que, pour tout* $c \leqslant u \leqslant c'N$,

$$\mathbb{P}\Big(|\lambda_j - \gamma_j| \geqslant \frac{u}{N}\Big) \leqslant 4\exp\Big(-\frac{C^2u^2}{2c_\delta \log N + Cu}\Big). \tag{3.9}$$

La démonstration est basée sur l'inégalité de déviation établie pour la fonction de comptage au début de ce chapitre (3.4) et sur la borne non asymptotique (3.2) sur la variance de $\mathcal{N}_t$ issue des travaux de Gustavsson. Le schéma de démonstration est le suivant. Nous commençons par majorer $\mathbb{P}(\lambda_j \geqslant \gamma_j + \frac{u}{N})$ en utilisant le lien entre fonction de comptage et $j^{\text{ème}}$ valeur propre (1.11). On doit alors majorer

$$\mathbb{P}\Big(|\mathcal{N}_{\gamma_j+\frac{u}{N}} - N\rho_{\gamma_j+\frac{u}{N}}| > N(\rho_{\gamma_j+\frac{u}{N}} - \rho_{\gamma_j})\Big)$$

où $\rho_t = \rho_{sc}((-\infty,t])$. Pour cela, nous minorons la différence $\rho_{\gamma_j+\frac{u}{N}} - \rho_{\gamma_j}$ grâce à des bornes classiques sur la densité de la loi du demi-cercle. Utilisant le fait que $\eta N \leqslant j \leqslant (1-\eta)N$, nous obtenons que cette quantité est minorée par $C\frac{u}{N}$ et donc que

$$\mathbb{P}\Big(|\mathcal{N}_{\gamma_j+\frac{u}{N}} - N\rho_{\gamma_j+\frac{u}{N}}| > N(\rho_{\gamma_j+\frac{u}{N}} - \rho_{\gamma_j})\Big) \leqslant \mathbb{P}(|\mathcal{N}_{\gamma_j+\frac{u}{N}} - N\rho_{\gamma_j+\frac{u}{N}}| > Cu).$$

De plus, le fait que $\eta N \leqslant j \leqslant (1-\eta)N$ permet de s'assurer que $\gamma_j + \frac{u}{N}$ reste dans un intervalle $[-2+\delta, 2-\delta]$, tant que $u \leqslant c'N$. Ceci permet d'utiliser (3.4) pour conclure.

Dans le chapitre 6, nous montrons une proposition analogue pour les matrices du LUE. La démonstration est complètement similaire et se base sur les mêmes éléments : l'inégalité de déviation établie pour la fonction de comptage grâce aux propriétés déterminantales, la version non asymptotique des calculs de Su [83] sur la variance de la fonction de comptage et le lien entre $\mathcal{N}_t$ et la $j^{\text{ème}}$ valeur propre (1.11). Le cas des matrices du GOE et du LOE est traité via les relations d'entrelacement dont nous avons parlé au paragraphe 2.1.

Suivant le même raisonnement, il est également possible de montrer des inégalités de déviation similaires pour les valeurs propres intermédiaires, c'est-à-dire les valeurs propres $\lambda_j$ avec $N - j \to +\infty$ avec $\frac{j}{N} \to 1$ pour le bord droit du spectre des matrices du GUE. On définit de manière analogue ces valeurs propres intermédiaires sur le bord gauche et dans le cas de matrices de covariance.

**Proposition 3.3** (Proposition 5.10)**.** *Il existe des constantes positives universelles* $C$ *et* $c$ *telles que ce qui suit est vrai. Soient* $K > 20\sqrt{2}$ *et* $0 < \eta \leqslant \frac{1}{2}$. *Prenons* $(1-\eta)N \leqslant j \leqslant N - K\log N$. *Alors il existe* $C' > 0$ *et* $c' > 0$ *dépendant uniquement de* $K$ *et* $\eta$ *telles que, pour tout* $c \leqslant u \leqslant c'(N-j)$,

$$\mathbb{P}\Big(|\lambda_j - \gamma_j| \geqslant \frac{u}{N^{2/3}(N-j)^{1/3}}\Big) \leqslant 4\exp\Big(-\frac{C^2u^2}{C'\log(N-j) + Cu}\Big). \tag{3.10}$$

La démonstration est très similaire. Pour les matrices du LUE, cette proposition n'est établie que sur le bord droit. Su n'a en effet pas effectué les calculs de la variance nécessaires au traitement du bord gauche. Toutefois, il est hautement probable que la borne

$$\mathrm{Var}(\mathcal{N}_t) \leqslant c_{\tilde{\delta},\tilde{K}} \log n(b_{m,n} - t)^{3/2},$$

valable pour tout $t$ tel que $0 < b_{m,n} - t < \tilde{\delta}$ et $n(b_{m,n} - t)^{3/2} \geqslant \tilde{K} \log n$, soit transposable dans le cas du bord gauche en $\mathrm{Var}(\mathcal{N}_t) \leqslant c_{\tilde{\delta},\tilde{K}} \log n(t - a_{m,n})^{3/2}$ quand le rapport entre les nombres de lignes et de colonnes de la matrice $X$, $\frac{m}{n}$, est minoré par $A_1 > 1$. Dans ce cas, le raisonnement précédent fonctionne et l'inégalité de déviation en découle immédiatement.

Ces inégalités de déviation reposent sur un raisonnement applicable uniquement aux matrices gaussiennes. Dans le cas des matrices de Wigner et de covariance plus générales, il n'existe pas d'inégalités aussi précises pour les petites déviations. Les inégalités disponibles actuellement ne concernent que les grandes déviations, c'est-à-dire pour $u$ grand. En particulier, le théorème de rigidité des valeurs propres de Erdös, Yau et Yin (théorème 2.3) fournit l'inégalité suivante pour $\eta N \leqslant j \leqslant (1 - \eta)N$.

$$\mathbb{P}\Big(|\lambda_j - \gamma_j| \geqslant \frac{(\log N)^{C\mathrm{llog} N}}{N}\Big) \leqslant e^{-c(\log N)^{c'\mathrm{llog} N}}.$$

Ce théorème donne donc une information uniquement pour $u \geqslant (\log N)^{C\mathrm{llog} N}$, dans le cas où la loi des coefficients a une décroissance exponentielle. Des informations similaires sont disponibles pour les valeurs propres intermédiaires, ainsi que pour les matrices de covariance, grâce au théorème de rigidité des valeurs propres établi par Pillai et Yin [75]. Tao et Vu [89] ont récemment amélioré ces inégalités de déviation sous une hypothèse supplémentaire de troisième moment qui s'annule, obtenant des informations pour $u \geqslant C(\log N)^c$. Ces inégalités récentes ne traitant pas des petites déviations, il est impossible d'en déduire une information optimale sur les variances de ces valeurs propres. Nous discutons de ces variances dans le paragraphe suivant.

### 3.3.4 Bornes sur la variance des valeurs propres

Dans l'article [23], nous établissons des bornes sur la variance des valeurs propres de matrices de Wigner et de covariance. Les chapitres 5 et 6 contiennent les détails.

Les bornes sur la variance des valeurs propres sont d'abord obtenues dans le cas gaussien. Il est en effet aisé de déduire ces bornes des inégalités de déviation (3.9), (3.10), (3.5), (3.6). Il suffit d'intégrer les inégalités précédentes, en utilisant le fait que $\mathbb{E}[Z^2] = \int_0^{+\infty} \mathbb{P}(|Z| > s) 2s\, ds$ pour une variable aléatoire $Z$. On obtient alors une borne sur $\mathbb{E}[|\lambda_j - \gamma_j|^2]$, ce qui fournit une borne sur la variance de $\lambda_j$. Dans le cas des matrices gaussiennes renormalisées, on obtient alors les majorations suivantes :

$$\mathbb{E}\Big[(\lambda_N - 2)^2\Big] \leqslant CN^{-4/3} \quad \text{et} \quad \mathbb{E}[(\lambda_1 + 2)^2] \leqslant CN^{-4/3}, \tag{3.11}$$

$$\mathbb{E}\Big[(\lambda_j - \gamma_j)^2\Big] \leqslant C(\eta) \frac{\log N}{N^2} \quad \text{pour } \eta N \leqslant j \leqslant (1-\eta)N, \tag{3.12}$$

$$\mathbb{E}\Big[(\lambda_j - \gamma_j)^2\Big] \leqslant C(K, \eta) \frac{\log(\min(j, N-j))}{N^{4/3} \min(j, N-j)^{2/3}}, \tag{3.13}$$

pour $K \log N \leqslant \min(j, N-j) \leqslant \eta N$.

Les inégalités de déviation connues actuellement sur les matrices de Wigner générales ne permettent pas d'obtenir ces bornes optimales sur les variances du fait de l'absence d'information pour les petites déviations. En revanche, il est possible de combiner le théorème des quatre moments de Tao et Vu 2.6 et le théorème de rigidité des valeurs propres de Erdös, Yau et Yin 2.3 pour transférer les bornes précédentes à des matrices de Wigner plus générales.

**Théorème 3.4** (Théorèmes 5.1, 5.2 et 5.3)**.** *Soit $M_N$ une matrice de Wigner satisfaisant la condition* $(C0)$ *et dont les coefficients ont les mêmes quatre premiers moments qu'une matrice du GUE (ou du GOE dans le cas de matrices symétriques réelles). Alors les valeurs propres $\lambda_j$ de $W_N = \frac{1}{\sqrt{N}} M_N$ vérifient* (3.11)*,* (3.12) *et* (3.13)*.*

L'idée de la démonstration est la suivante. Si on pouvait utiliser le théorème des quatre moments 2.6 avec la fonction $H_j : \mathbb{R} \mapsto \mathbb{R}$ définie par $H_j(x) = (x - N\gamma_j)^2$, on obtiendrait

$$\Big|\mathbb{E}\Big[(\lambda_j(A_N) - N\gamma_j)^2\Big] - \mathbb{E}\Big[(\lambda_j(A'_N) - N\gamma_j)^2\Big]\Big| \leqslant N^{-c_0}.$$

D'où

$$\mathbb{E}\Big[(\lambda_j(W_N) - \gamma_j)^2\Big] \leqslant \mathbb{E}\Big[(\lambda_j(W'_N) - \gamma_j)^2\Big] + N^{-c_0-2}.$$

Mais d'après les majorations sur $\mathbb{E}[(\lambda_j(W'_N) - \gamma_j)^2]$ (3.11), (3.12) et (3.13), le terme d'erreur $N^{-c_0-2}$ est plus petit que la borne gaussienne et on obtient bien les majorations voulues. Ce raisonnement est toutefois impossible à mener, la fonction $H_j$ ne vérifiant pas les conditions d'applications du théorème des quatre moments, à savoir le fait que pour tout $x \in \mathbb{R}$, $|H_j^{(k)}(x)| \leqslant N^{c_0}$ pour tout $k \leqslant 5$. Nous utilisons donc le théorème de rigidité des valeurs propres 2.3 qui nous fournit l'information que la valeur propre $\lambda_j$ est fortement localisée autour de sa valeur théorique $\gamma_j$. Ceci nous permet de remplacer la fonction par une troncature lisse de $H_j$ dans le raisonnement précédent. Le théorème de rigidité fournit alors un contrôle quantitatif sur la probabilité que cette troncature induise une erreur. Plus précisément, nous prenons $\psi$ une fonction lisse à support dans $[-2,2]$ et à valeurs dans $[0,1]$ telle que $\psi(x) = \frac{1}{10}x^2$ pour tout $x \in [-1,1]$ (cette fonction peut être obtenue par convolution). On pose alors $G_j : x \in \mathbb{R} \mapsto \psi\Big(\frac{x-N\gamma_j}{R_N^{(j)}}\Big)$, où $R_N^{(j)} = (\log N)^{C\mathrm{llog}N} N^{1/3} \min(j, N+1-j)^{-1/3}$ (le terme apparaissant dans l'inégalité de déviation fournie par le théorème de rigidité des valeurs propres). Comme $\psi$ est lisse à support compact, la fonction $G_j$ va vérifier les hypothèses du théorème des quatre moments. Cette démarche nous permet de comparer $\mathbb{E}[(\lambda_j - \gamma_j)^2]$ pour une matrice de Wigner renormalisée $W_N$ à $\mathbb{E}[(\lambda'_j - \gamma_j)^2]$ pour une matrice du GUE ou du GOE renormalisée $W'_N$. Le terme d'erreur obtenu est d'ordre $O\Big((R_N^{(j)})^2 N^{-c_0-2}\Big)$, ce qui s'avère plus petit que les majorations qu'on cherche à transférer (quel que soit l'indice $j$). Le résultat en découle immédiatement. Ce résultat et sa démonstration sont contenus dans le chapitre 5.

Il est possible de mener exactement la même étude pour les matrices de covariance. La démonstration est identique et repose sur le théorème des quatre moments de Tao, Vu [88] et Wang [101], ainsi que sur le théorème de rigidité des valeurs propres établi par Pillai et Yin pour ces matrices [75]. Cette démonstration est présentée dans le chapitre 6.

# Chapitre 4

# Théorème central limite pour la fonction de comptage

[*Ce chapitre contient le texte de l'article « A note on the Central Limit Theorem for the Eigenvalue Counting Function of Wigner Matrices », écrit avec Van H. Vu, paru dans Electronic Journal of Probability, 16 (2011), p. 314-322.*]

**Abstract.** *The purpose of this note is to establish a Central Limit Theorem for the number of eigenvalues of a Wigner matrix in an interval. The proof relies on the correct asymptotics of the variance of the eigenvalue counting function of GUE matrices due to Gustavsson, and its extension to large families of Wigner matrices by means of the Tao and Vu Four Moment Theorem and recent localization results by Erdös, Yau and Yin.*

## 4.1 Introduction

This note is concerned with the asymptotic behavior of the eigenvalue counting function, that is the number $N_I$ of eigenvalues falling in an interval $I$, of families of Wigner matrices, when the size of the matrix goes to infinity. Wigner matrices are random Hermitian matrices $M_n$ of size $n$ such that, for $i < j$, the real and imaginary parts of $(M_n)_{ij}$ are iid, with mean 0 and variance $\frac{1}{2}$, $(M_n)_{ii}$ are iid with mean 0 and variance 1. An important example of Wigner matrices is the case where the entries are Gaussian, giving rise to the so-called Gaussian Unitary Ensemble (GUE). GUE matrices will be denoted by $M'_n$. In this case, the joint law of the eigenvalues is known, allowing for complete descriptions of their limiting behavior both in the global and local regimes (cf. for example [1]).

Denote by $\lambda_1, \ldots, \lambda_n$ the real eigenvalues of the normalized Wigner matrix $W_n = \frac{1}{\sqrt{n}} M_n$. The classical Wigner theorem states that the empirical distribution

$\frac{1}{n}\sum_{j=1}^{n}\delta_{\lambda_j}$ on the eigenvalues of $W_n$ converges weakly almost surely as $n \to \infty$ to the semicircle law $d\rho_{sc}(x) = \frac{1}{2\pi}\sqrt{4-x^2}\mathbf{1}_{[-2,2]}(x)dx$. Consequently, for any interval $I \subset \mathbb{R}$,

$$\frac{1}{n}N_I(W_n) = \frac{1}{n}\sum_{j=1}^{n}\mathbf{1}_{\{\lambda_j \in I\}} \underset{n\to\infty}{\to} \rho_{sc}(I) \quad \text{almost surely.}$$

At the fluctuation level, it is known, due to the particular determinantal structure of the GUE, that

**Theorem 4.1** (Costin-Lebowitz [21], Soshnikov [81] (see [1]))**.** *Let $M_n'$ be a GUE matrix. Set $W_n' = \frac{1}{\sqrt{n}}M_n'$. Let $I_n$ be an interval in $\mathbb{R}$. If $\mathrm{Var}(N_{I_n}(W_n')) \underset{n\to\infty}{\to} \infty$, then*

$$\frac{N_{I_n}(W_n') - \mathbb{E}[N_{I_n}(W_n')]}{\sqrt{\mathrm{Var}(N_{I_n}(W_n'))}} \underset{n\to\infty}{\to} \mathcal{N}(0,1) \tag{4.1}$$

*in distribution.*

In 2005, Gustavsson [49] was able to fully describe for GUE matrices the asymptotic behavior of the variance of the counting function $N_I(W_n')$ for intervals $I = [y, +\infty)$ with $y \in (-2, 2)$ strictly in the bulk of the semicircle law. He established that

$$\mathbb{E}[N_I(W_n')] = n\rho_{sc}(I) + O\Big(\frac{\log n}{n}\Big) \quad \text{and} \quad \mathrm{Var}(N_I(W_n')) = \Big(\frac{1}{2\pi^2} + o(1)\Big)\log n. \tag{4.2}$$

In particular therefore, if $I = [y, +\infty)$ with $y \in (-2, 2)$,

$$\frac{N_I(W_n') - \mathbb{E}[N_I(W_n')]}{\sqrt{\mathrm{Var}(N_I(W_n'))}} \underset{n\to\infty}{\to} \mathcal{N}(0,1). \tag{4.3}$$

as well as

$$\frac{N_I(W_n') - n\rho_{sc}(I)}{\sqrt{\frac{1}{2\pi^2}\log n}} \underset{n\to\infty}{\to} \mathcal{N}(0,1) \tag{4.4}$$

(which we call below the CLT with numerics).

The purpose of this note is to extend these conclusions to non-Gaussian Wigner matrices. Tao and Vu's Four Moment Theorem leads to (4.4) for non-Gaussian Wigner matrices, as will be developped in Section 2. Unfortunately, this theorem does not give (4.2) and (4.3) for these matrices. Indeed, the Four Moment Theorem deals with a finite number of eigenvalues, whereas the computation of $\mathbb{E}[N_I(W_n)]$ and $\mathrm{Var}\, N_I(W_n)$ involves all the eigenvalues of the matrix. To achieve (4.2), we make use of recent results by Erdös, Yau and Yin [35] providing suitable localization properties of the eigenvalues in the bulk. This result, combined with the Four Moment Theorem, will give (4.2) and consequently (4.3). Section 3 is devoted to this main step.

The class of Wigner matrices covered by our results is described by the following condition. Say that $M_n$ satisfies condition $(C0)$ if the real part $\xi$ and the imaginary part $\tilde{\xi}$ of $(M_n)_{ij}$ are independent and have an exponential decay: there are two constants $C$ and $C'$ such that

$$\mathbb{P}(|\xi| \geq t^C) \leq e^{-t} \quad \text{and} \quad \mathbb{P}(|\tilde{\xi}| \geq t^C) \leq e^{-t},$$

for all $t \geq C'$.

Say that two complex random variables $\xi$ and $\xi'$ match to order $k$ if

$$\mathbb{E}\left[\Re(\xi)^m \Im(\xi)^l\right] = \mathbb{E}\left[\Re(\xi')^m \Im(\xi')^l\right]$$

for all $m, l \geq 0$ such that $m + l \leq k$.

The following theorem is the main result of this note.

**Theorem 4.2.** *Let $M_n$ be a random Hermitian matrix whose entries satisfy condition $(C0)$ and match the corresponding entries of GUE up to order 4. Set $W_n = \frac{1}{\sqrt{n}} M_n$. Then, for any $y \in (-2, 2)$ and $I(y) = [y, +\infty)$, setting $Y_n := N_{I(y)}(W_n)$, we have*

$$\mathbb{E}[Y_n] = n\rho_{sc}(I(y)) + o(1) \quad \textit{and} \quad \mathrm{Var}(Y_n) = \left(\frac{1}{2\pi^2} + o(1)\right) \log n.$$

*As a consequence, the sequence $(Y_n)$ satisfies the CLT in the form*

$$\frac{Y_n - \mathbb{E}[Y_n]}{\sqrt{\mathrm{Var}(Y_n)}} \underset{n \to \infty}{\to} \mathcal{N}(0, 1).$$

The theorem is established in the next two sections. As announced, in a first step, relying on Gustavsson's results and its extension to Wigner matrices by Tao and Vu [87], we establish that $(Y_n)$ satisfies the CLT with numerics (4.4). In a second step, we use recent results of Erdös, Yau and Yin [35] on the localization of eigenvalues in order to prove that $\mathbb{E}[Y_n]$ and $\mathrm{Var}(Y_n)$ are close to those of $M_n'$ (GUE) and therefore satisfy (4.2).

## 4.2 CLT with numerics and eigenvalues in the bulk

In this section, we prove (4.4) for non-Gaussian Wigner matrices.

On the basis of the CLT with numerics, Gustavsson [49] described the Gaussian behavior of eigenvalues in the bulk of the semicircle law in the form of

$$\sqrt{\frac{4-t(i/n)^2}{2}}\,\frac{\lambda_i(W_n')-t(i/n)}{\frac{\sqrt{\log n}}{n}} \underset{n\to\infty}{\to} \mathcal{N}(0,1) \tag{4.5}$$

in distribution, where $t(x) \in [-2,2]$ is defined for $x \in [0,1]$ by

$$x = \int_{-2}^{t(x)} d\rho_{sc}(t) = \frac{1}{2\pi}\int_{-2}^{t(x)} \sqrt{4-x^2}dx.$$

More informally, $\lambda_i(W_n') \approx t(i/n) + \mathcal{N}(0, \frac{2\log n}{(4-t(i/n)^2)n^2})$. This is achieved by the tight relation between eigenvalues and the counting function expressed by the elementary equivalence, for $I(y) = [y, +\infty)$, $y \in \mathbb{R}$,

$$N_{I(y)}(W_n) \le n-i \quad \text{if and only if} \quad \lambda_i \le y. \tag{4.6}$$

The result (4.5) was extended in [87] to large families of Wigner matrices satisfying condition ($C0$) by means of the Four Moment Theorem (see [87] and [86]). Now using the reverse strategy based on (4.6), (4.5) may be shown to imply back the CLT with numerics (4.4) for Wigner matrices whose entries match those of the GUE up to order 4. We provide some details in this regard relying on the following simple consequence of the Four Moment Theorem.

**Proposition 4.3.** *Let $M_n$ and $M_n''$ be two random matrices satisfying condition ($C0$) such that their entries match up to order 4. There exists $c > 0$ such that, if $n$ is large enough, for any $y \in (-2,2)$ and any $(i,j) \in \{1,\dots,n\}^2$, if $I(y) = [y, +\infty)$,*

$$\mathbb{P}(\lambda_i'' \in I_-(y)) - n^{-c} \le \mathbb{P}(\lambda_i \in I(y)) \le \mathbb{P}(\lambda_i'' \in I_+(y)) + n^{-c},$$

*and*

$$\mathbb{P}(\lambda_i'' \in I_-(y) \wedge \lambda_j'' \in I_-(y)) - n^{-c} \le \mathbb{P}(\lambda_i \in I(y) \wedge \lambda_j \in I(y)),$$

*and*

$$\mathbb{P}(\lambda_i \in I(y) \wedge \lambda_j \in I(y)) \le \mathbb{P}(\lambda_i'' \in I_+(y) \wedge \lambda_j'' \in I_+(y)) + n^{-c},$$

*where $I_+(y) = [y + n^{-c-1}, +\infty)$ et $I_-(y) = [y - n^{-c-1}, +\infty)$.*

As announced, we would like to show that the behavior of eigenvalues in the bulk (4.5) extended to Wigner matrices leads to the CLT with numerics for such matrices, namely,

$$\frac{N_{I(y)}(W_n) - n\rho_{sc}(I(y))}{\sqrt{\frac{1}{2\pi^2}\log n}} \underset{n\to\infty}{\to} \mathcal{N}(0,1), \tag{4.7}$$

in distribution for Wigner matrices $W_n$ satisfying $(C0)$. To prove this, observe that for every $x \in \mathbb{R}$.

$$\mathbb{P}\Big(\frac{N_{I(y)}(W_n) - n\rho_{sc}(I(y))}{\sqrt{\frac{1}{2\pi^2}\log n}} \leq x\Big) = \mathbb{P}(N_{I(y)}(W_n) \leq n - i_n),$$

where $i_n = n\rho_{sc}((-\infty, y]) - x\sqrt{\frac{1}{2\pi^2}\log n}$. Then, by (4.6),

$$\begin{aligned}\mathbb{P}\Big(\frac{N_{I(y)}(W_n) - n\rho_{sc}(I(y))}{\sqrt{\frac{1}{2\pi^2}\log n}} \leq x\Big) &= \mathbb{P}(\lambda_{i_n}(W_n) \leq y) \\ &= \mathbb{P}\Big(\sqrt{\frac{4 - t(i_n/n)^2}{2}}\frac{\lambda_{i_n}(W_n) - t(i_n/n)}{\frac{\sqrt{\log n}}{n}} \leq x_n\Big),\end{aligned}$$

where $x_n = \sqrt{\frac{4-t(i_n/n)^2}{2}}\frac{y-t(i_n/n)}{\frac{\sqrt{\log n}}{n}}$. Now $\frac{i_n}{n} \to \rho_{sc}((-\infty, y]) \in (0,1)$. Furthermore, $x_n \to x$ since

$$\begin{aligned} t(i_n/n) &= t\Big(\rho_{sc}((-\infty, y]) - \frac{x}{n}\sqrt{\frac{1}{2\pi^2}\log n}\Big) \\ &= t\Big(\rho_{sc}((-\infty, y])\Big) - t'\Big(\rho_{sc}((-\infty, y])\Big)\frac{x}{n}\sqrt{\frac{1}{2\pi^2}\log n} + o\Big(\frac{\sqrt{\log n}}{n}\Big) \\ &= y - x\sqrt{\frac{2}{4-y^2}}\frac{\sqrt{\log n}}{n} + o\Big(\frac{\sqrt{\log n}}{n}\Big). \end{aligned}$$

Hence $\frac{y-t(i_n/n)}{\frac{\sqrt{\log n}}{n}} = x\sqrt{\frac{2}{4-y^2}} + o(1)$, from which $x_n \to x$.

Applying (4.5) (extended to Wigner matrices), we obtain that

$$\mathbb{P}\Big(\sqrt{\frac{4 - t(i_n/n)^2}{2}}\frac{\lambda_{i_n}(W_n) - t(i_n/n)}{\frac{\sqrt{\log n}}{n}} \leq x_n\Big) \underset{n\to\infty}{\to} \mathbb{P}(X \leq x),$$

where $X \sim \mathcal{N}(0,1)$, implying (4.7).

## 4.3 Estimating the mean and the variance of $Y_n$

In this section, we prove that the mean and the variance of the eigenvalue counting function are of the same order for Gaussian and non-Gaussian Wigner matrices.

To reach the CLT of Theorem 4.2 from the CLT with numerics (4.7), it is necessary to suitably control the expectation and variance $\mathbb{E}[Y_n]$ and $\mathrm{Var}(Y_n)$ of the eigenvalue counting function, and to show that their behaviors are identical to the ones for GUE matrices. The direct use of the Four Moment Theorem is unfortunately not enough to this purpose since it only provides proximity on a small number of eigenvalues. But recent results of Erdös, Yau and Yin [35], presented in the following statement, describe strong localization of the eigenvalues of Wigner matrices which provides the additional step necessary to complete the argument.

**Theorem 4.4.** *Let $M_n$ be a random Hermitian matrix whose entries satisfy condition $(C0)$. There is a constant $C > 0$ such that, for any $i \in \{1, \dots, n\}$,*

$$\mathbb{P}(|\lambda_i - t(i/n)| \geq (\log n)^{C \log\log n} \min(i, n-i+1)^{-1/3} n^{-2/3}) \leq n^{-3}.$$

Note that if $n\varepsilon \leq i \leq (1-\varepsilon)n$ for some small $\varepsilon > 0$, then $\min(i, n-i+1) \geq n\varepsilon$ so that

$$\mathbb{P}\Big(|\lambda_i - t(i/n)| \geq n^{-1}\varepsilon^{-1/3}(\log n)^{C\log\log n}\Big) \leq n^{-3}. \tag{4.8}$$

The next lemma presents the main conclusion on the expectation and variance of the eigenvalue counting function, extending Gustavsson's conclusion (4.2) for the GUE to Wigner matrices of the class $(C0)$. Once this lemma is established, Theorem 4.2 will follow.

**Lemma 4.5.** *Set $W_n = \frac{1}{\sqrt{n}} M_n$, $I(y) = [y, +\infty)$ where $y \in (-2, 2)$, and $Y_n = N_{I(y)}(W_n)$. Then*

$$\mathbb{E}[Y_n] = n\rho_{sc}(I(y)) + o(1) \quad \textit{and} \quad \mathrm{Var}(Y_n) = \Big(\frac{1}{2\pi^2} + o(1)\Big)\log n.$$

*Proof.* By Gustavsson's results (4.2) therefore, if $Y_n'$ denotes $N_{I(y)}(W_n')$ in the case $M_n'$ is GUE,

$$\mathbb{E}[Y_n'] = n\rho_{sc}(I(y)) + O\Big(\frac{\log n}{n}\Big) \quad \text{and} \quad \mathrm{Var}(Y_n') = \Big(\frac{1}{2\pi^2} + o(1)\Big)\log n.$$

Hence, to establish Lemma 4.5, it suffices to show that $\mathbb{E}[Y_n] = E[Y_n'] + o(1)$ and $\mathrm{Var}(Y_n) = \mathrm{Var}(Y_n') + o(1)$. Below, we only deal with the variance, the argument for the expectation being similar and actually simpler.

Set $A_i = \mathbf{1}_{\{\lambda_i \in I\}}$, for $i \in \{1, \dots, n\}$. Notice that

$$|\operatorname{Var}(Y_n) - \operatorname{Var}(Y_n')| \leq \sum_{1 \leq i,j \leq n} |(\mathbb{E}[A_iA_j] - \mathbb{E}[A_i]\mathbb{E}[A_j]) - (\mathbb{E}[A_i'A_j'] - \mathbb{E}[A_i']\mathbb{E}[A_j'])|.$$

Call an index $i$ *first class* if $\mathbb{E}[A_i] \geq 1 - n^{-3}$ or $\leq n^{-3}$ and *second class* otherwise.

Note that if $j$ is first class, then, for all $i \in \{1, \dots, n\}$,

$$|\mathbb{E}[A_i A_j] - \mathbb{E}[A_i]\mathbb{E}[A_j]| = O(n^{-3}).$$

Indeed, if $\mathbb{E}[A_j] \le n^{-3}$, then both terms between the absolute value signs are $O(n^{-3})$, so that $|\mathbb{E}[A_i A_j] - \mathbb{E}[A_i]\mathbb{E}[A_j]| = O(n^{-3})$. The other case can be brought back to this case by the identity

$$|\mathbb{E}[A_i A_j] - \mathbb{E}[A_i]\mathbb{E}[A_j]| = |\mathbb{E}[B_i B_j] - \mathbb{E}[B_i]\mathbb{E}[B_j]|,$$

where $B_i$ is the complement of $A_i$. Consequently,

$$\sum_{\substack{i \text{ or } j \\ 1^{\text{st}} \text{ class}}} \left|\Big(\mathbb{E}[A_i A_j] - \mathbb{E}[A_i]\mathbb{E}[A_j]\Big) - \Big(\mathbb{E}[A_i' A_j'] - \mathbb{E}[A_i']\mathbb{E}[A_j']\Big)\right| = O(n^{-1}). \tag{4.9}$$

Theorem 4.4 shows that there are only $O((\log n)^{C \log\log n})$ second class indices. Indeed, set $\eta_n = n^{-1}\varepsilon^{-1/3}(\log n)^{C\log\log n}$ and suppose first that $i \in \{1, \dots, n\}$ is such that $t(i/n) < y - \eta_n$:

– if $t(i/n) > t(\varepsilon)$, (4.8) is true for $W_n$. Then

$$\mathbb{P}(\lambda_i \in I_n) \le \mathbb{P}\Big(|\lambda_i - t(i/n)| \ge \eta_n\Big) \le n^{-3}.$$

– if $t(i/n) < t(\varepsilon)$, choose $j$ such that $t(\varepsilon) < t(j/n) < y - \eta_n$ (take $\varepsilon$ small enough and $n$ large enough such that there is such a $j$). Then $\lambda_i \le \lambda_j$ and $\mathbb{P}(\lambda_i \in I_n) = \mathbb{P}(\lambda_i \ge y) \le \mathbb{P}(\lambda_j \ge y) = \mathbb{P}(\lambda_j \in I_n) \le n^{-3}$.

Similarly one can show that if $i \in \{1, \dots, n\}$ is such that $t(i/n) > y + \eta_n$, then $i$ is first class.

As a consequence of this discussion, $i \in \{1, \dots, n\}$ can only be second class if $y - \eta_n < t(i/n) < y + \eta_n$. We need to count these possible $i$'s. By definition of $t(i/n)$, $i = \frac{n}{2\pi}\int_{-2}^{t(i/n)} \sqrt{4 - x^2}dx$. Thus,

$$\frac{n}{2\pi}\int_{-2}^{y-\eta_n} \sqrt{4 - x^2}dx \le i \le \frac{n}{2\pi}\int_{-2}^{y+\eta_n} \sqrt{4 - x^2}dx.$$

In this case $i$ belongs to an interval of length

$$\frac{n}{2\pi}\int_{y-\eta_n}^{y+\eta_n} \sqrt{4 - x^2}dx \le \frac{2}{\pi\varepsilon^{1/3}}(\log n)^{C\log\log n}.$$

Therefore, there are at most $\frac{2}{\pi\varepsilon^{1/3}}(\log n)^{C\log\log n} + 1 = O((\log n)^{C\log\log n})$ second class $i$'s.

Next, by Proposition 4.3, it is easily seen that if both $i, j$ are second class, then

$$|(\mathbb{E}[A_i A_j] - \mathbb{E}[A_i]\mathbb{E}[A_j]) - (\mathbb{E}[A_i' A_j'] - \mathbb{E}[A_i']\mathbb{E}[A_j'])| = O(n^{-c})$$

for some positive constant $c$. Since the number of such pairs is $O((\log n)^{2C \log\log n})$, we have

$$\sum_{\substack{i \text{ and } j \\ 2^{\text{nd}} \text{ class}}} |\Big(\mathbb{E}[A_iA_j]-\mathbb{E}[A_i]\mathbb{E}[A_j]\Big)-\Big(\mathbb{E}[A'_iA'_j]-\mathbb{E}[A'_i]\mathbb{E}[A'_j]\Big)| = O(n^{-c}(\log n)^{2C\log\log n}). \tag{4.10}$$

To conclude,

$$\begin{aligned} |\operatorname{Var}(Y_n) - \operatorname{Var}(Y'_n)| \leq & \sum_{\substack{i \text{ or } j \\ 1^{\text{st}} \text{ class}}} |\Big(\mathbb{E}[A_iA_j] - \mathbb{E}[A_i]\mathbb{E}[A_j]\Big) - \Big(\mathbb{E}[A'_iA'_j] - \mathbb{E}[A'_i]\mathbb{E}[A'_j]\Big)| \\ & + \sum_{\substack{i \text{ and } j \\ 2^{\text{nd}} \text{ class}}} |\Big(\mathbb{E}[A_iA_j] - \mathbb{E}[A_i]\mathbb{E}[A_j]\Big) - \Big(\mathbb{E}[A'_iA'_j] - \mathbb{E}[A'_i]\mathbb{E}[A'_j]\Big)|, \end{aligned}$$

so that (4.9) and (4.10) lead to

$$|\operatorname{Var}(Y_n) - \operatorname{Var}(Y'_n)| \leq O(n^{-1}) + O(n^{-c}(\log n)^{2C\log\log n}) = o(1),$$

as claimed. This shows that $\operatorname{Var}(Y_n) = \left(\frac{1}{2\pi^2} + o(1)\right)\log n$. As mentioned earlier, it may be shown similarly that $\mathbb{E}[Y_n] = n\rho_{sc}(I(y)) + o(1)$ and the proof of Lemma 4.5 is thus complete. □

## 4.4 About real Wigner matrices

In this section, we briefly indicate how the preceding results for Hermitian random matrices may be stated similarly for real Wigner symmetric matrices. To this task, we follow the same scheme of proof, relying in particular on the corollary of Tao and Vu Four Moment Theorem (Proposition 4.3) which also holds in the real case (cf. [69]).

Real Wigner matrices are random symmetric matrices $M_n$ of size $n$ such that, for $i < j$, $(M_n)_{ij}$ are iid, with mean 0 and variance 1, $(M_n)_{ii}$ are iid with mean 0 and variance 2. As in the complex case, an important example of real Wigner matrices is the case where the entries are Gaussian, giving rise to the so-called Gaussian Orthogonal Ensemble (GOE).

The main issue is actually to establish first the conclusions for the GOE. This has been suitably developed by O'Rourke in [69] by means of interlacing formulas (cf. [38]).

**Theorem 4.6** (Forrester-Rains)**.** *The following relation holds between matrix ensembles:*

$$\mathrm{GUE}_n = \mathrm{even}(\mathrm{GOE}_n \cup \mathrm{GOE}_{n+1}).$$

This statement can be interpreted in the following way. Take two independent matrices from the GOE, one of size $n$ and the other of size $n+1$. If we superimpose the $2n+1$ eigenvalues on the real line and then take the $n$ even ones, they have the same distribution as the eigenvalues of a $n \times n$ matrix from the GUE.

Let $I$ be an interval in $\mathbb{R}$. Let $M_n^{\mathbb{R}}$ be a GOE matrix and $M_n^{\mathbb{C}}$ be a GUE matrix. $W_n^{\mathbb{R}}$ and $W_n^{\mathbb{C}}$ are the corresponding normalized matrices. The preceding interlacing formula leads to

– $\mathbb{E}[N_I(W_n^{\mathbb{R}})] = \mathbb{E}[N_I(W_n^{\mathbb{C}})] + O(1)$
– $\mathrm{Var}(N_I(W_n^{\mathbb{R}})) = 2\,\mathrm{Var}(N_I(W_n^{\mathbb{C}})) + O(1)$, if $\mathrm{Var}(N_I(W_n^{\mathbb{C}})) \underset{n\to\infty}{\to} \infty$.

Relying on this result and on the GUE case, O'Rourke proved the following theorem:

**Theorem 4.7.** *Let $M_n^{\mathbb{R}}$ be a GOE matrix. Set $W_n^{\mathbb{R}} = \frac{1}{\sqrt{n}} M_n^{\mathbb{R}}$. Then, for any $y \in (-2,2)$ and $I(y) = [y, +\infty)$, setting $Y_n^{\mathbb{R}} := N_{I(y)}(W_n^{\mathbb{R}})$, we have*

$$\mathbb{E}[Y_n^{\mathbb{R}}] = n\rho_{sc}(I(y)) + O(1) \quad \textit{and} \quad \mathrm{Var}(Y_n^{\mathbb{R}}) = \Big(\frac{1}{\pi^2} + o(1)\Big)\log n.$$

*As a consequence, the sequence $(Y_n^{\mathbb{R}})$ satisfies the CLT in the form*

$$\frac{Y_n^{\mathbb{R}} - \mathbb{E}[Y_n^{\mathbb{R}}]}{\sqrt{\mathrm{Var}(Y_n^{\mathbb{R}})}} \underset{n\to\infty}{\to} \mathcal{N}(0,1).$$

Following exactly the same scheme as for complex Wigner matrices leads to the same conclusion: Theorem 4.7 is true for Wigner symmetric matrices, provided their entries match the corresponding entries of GOE up to order 4 and satisfy condition $(C0)$.

The CLT for the eigenvalue counting function has been investigated as well for families of covariance matrices. The main conclusion of this work holds similarly in this case conditioned however on the validity of the Erdös-Yau-Yin rigidity theorem for covariance matrices. There is strong indication that the current approach by Erdös, Yin and Yau for Wigner matrices will indeed yield such a result. All the other ingredients of the proof are besides available. Indeed, Su (cf. [83]) carried out computations for Gaussian covariance matrices and proved both the CLT and the correct asymptotics for mean and variance. Tao and Vu in [88] extended their Four Moment Theorem to such matrices. If a localization result is proved for these matrices, arguing as for Wigner matrices, we could reach in the same way the asymptotics for the mean and the variance of the eigenvalue counting function and, consequently, a CLT in the same form for suitable families of covariance matrices.

# Chapitre 5

# Bornes sur la variance des valeurs propres de matrices de Wigner

[*Ce chapitre contient le texte de l'article « Eigenvalue variance bounds of Wigner and covariance random matrices », paru dans la revue Random Matrices : Theory and Applications, 1 (2012), et disponible sur Hal et ArXiv.*]

**Abstract.** *This work is concerned with finite range bounds on the variance of individual eigenvalues of Wigner random matrices, in the bulk and at the edge of the spectrum, as well as for some intermediate eigenvalues. Relying on the GUE example, which needs to be investigated first, the main bounds are extended to families of Hermitian Wigner matrices by means of the Tao and Vu Four Moment Theorem and recent localization results by Erdös, Yau and Yin. The case of real Wigner matrices is obtained from interlacing formulas. As an application, bounds on the expected* 2*-Wasserstein distance between the empirical spectral measure and the semicircle law are derived. Similar results are available for random covariance matrices.*

Two different models of random Hermitian matrices were introduced by Wishart in the twenties and by Wigner in the fifties. Wishart was interested in modeling tables of random data in multivariate analysis and worked on random covariance matrices. In this paper, the results for covariance matrices are very close to those for Wigner matrices. Therefore, it deals mainly with Wigner matrices. Definitions and results regarding covariance matrices are available in the last section.

Random Wigner matrices were first introduced by Wigner to study eigenvalues of infinite-dimensional operators in statistical physics (see [67]) and then propagated to various fields of mathematics involved in the study of spectra of random matrices. Under suitable symmetry assumptions, the asymptotic properties of the eigenvalues of a random matrix were soon conjectured to be universal, in the sense they do not depend on the individual distribution of the matrix entries. This opened the way to numerous developments on the asymptotics of various statistics of the eigenvalues of random matrices, such as for example the global behavior of the spectrum, the spacings between the eigenvalues in the bulk of the spectrum or the behavior of the extreme eigenvalues. Two main models have been considered, invariant matrices and Wigner matrices. In the invariant matrix models, the matrix law is unitary invariant and the eigenvalue joint distribution can be written explicitly in terms of a given potential. In the Wigner models, the matrix entries are independent (up to symmetry conditions). The case where the entries are Gaussian is the only model belonging to both types. In the latter case, the joint distribution of the eigenvalues is thus explicitly known and the previous statistics have been completely studied. One main focus of random matrix theory in the past decades was to prove that these asymptotic behaviors were the same for non-Gaussian matrices (see for instance [1], [8] and [72]).

However, in several fields such as computer science or statistics for example, asymptotic statements are often not enough, and more quantitative finite range results are required. Several recent developments have thus been concerned with non-asymptotic random matrix theory towards quantitative bounds (for instance on the probability for a certain event to occur) which are valid for all $N$, where $N$ is the size of the given matrix. See for example [99] for an introduction to some problems considered in non-asymptotic random matrix theory. In this paper, we investigate in this respect variance bounds on the eigenvalues of families of Wigner random matrices.

Wigner matrices are Hermitian or real symmetric matrices $M_N$ such that, if $M_N$ is complex, for $i < j$, the real and imaginary parts of $(M_N)_{ij}$ are independent and identically distributed (iid) with mean 0 and variance $\frac{1}{2}$, $(M_N)_{ii}$ are iid, with mean 0 and variance 1. In the real case, $(M_n)_{ij}$ are iid, with mean 0 and variance 1 and $(M_N)_{ii}$ are iid, with mean 0 and variance 2. In both cases, set $W_N = \frac{1}{\sqrt{N}} M_N$. An important example of Wigner matrices is the case where the entries are Gaussian. If $M_N$ is complex, then it belongs to the so-called Gaussian Unitary Ensemble (GUE). If it is real, it belongs to the Gaussian Orthogonal Ensemble (GOE). The matrix $W_N$ has $N$ real eigenvalues $\lambda_1 \leqslant \cdots \leqslant \lambda_N$. In the Gaussian case, the joint law of the eigenvalues is known, allowing for complete descriptions of their limiting behavior both in the global and local regimes (see for example [1], [8] and [72]).

Among universality results, at the global level, the classical Wigner's Theorem states that the empirical distribution $L_N = \frac{1}{N} \sum_{j=1}^{N} \delta_{\lambda_j}$ on the eigenvalues of $W_N$ converges weakly almost surely to the semicircle law defined by $d\rho_{sc}(x) = \frac{1}{2\pi}\sqrt{4 - x^2}\mathbb{1}_{[-2,2]}(x)dx$ (see for example [8] for a proof in the more general setting).

This gives the global asymptotic behavior of the spectrum. However, this theorem is not enough to deduce some information on individual eigenvalues. Define, for all $1 \leqslant j \leqslant N$, the theoretical location of the $j^{\text{th}}$ eigenvalue $\gamma_j$ by $\int_{-2}^{\gamma_j} d\rho_{sc}(x) = \frac{j}{N}$. Bai and Yin proved in [11] with assumptions on higher moments that almost surely the smallest and the largest eigenvalues converge to their theoretical locations, which means that $\lambda_N \to 2$ and $\lambda_1 \to -2$ almost surely (see also [8] for a proof of this theorem). From this almost sure convergence of the extreme eigenvalues and Wigner's Theorem, it is possible to deduce information on individual eigenvalues in the bulk of the spectrum. Indeed, according to the Glivenko-Cantelli Theorem (see for example [26]), the normalized eigenvalue function $\frac{1}{N}\mathcal{N}_x$, where $\mathcal{N}_x$ is the number of eigenvalues which are in $(-\infty, x]$, converges uniformly on $\mathbb{R}$ almost surely to the distribution function of the semicircle law $G$ (with no more assumptions on the matrix entries). Then, using that $\frac{1}{N}\mathcal{N}_{\lambda_j} = \frac{j}{N} = G(\gamma_j)$ together with crude bounds on the semicircle density function and the fact that $\lambda_j$ is almost surely between $-2-\varepsilon$ and $2+\varepsilon$ shows that almost surely $\lambda_j - \gamma_j \to 0$ uniformly for $\eta N \leqslant j \leqslant (1-\eta)N$ (for any fixed $\eta > 0$).

At the fluctuation level, eigenvalues inside the bulk and at the edge of the spectrum do not have the same behavior. Tracy and Widom showed in [91] that the largest eigenvalue fluctuates around 2 according to the so-called Tracy-Widom law $F_2$. Namely,

$$N^{2/3}(\lambda_N - 2) \to F_2$$

in distribution as $N$ goes to infinity. They proved this result for Gaussian matrices of the GUE, later extended to families of non-Gaussian Wigner matrices by Soshnikov (see [80]). Recent results by Tao and Vu (see [86]) and by Erdös, Yau and Yin (see [35]) provide alternate proofs of this fact for larger families of Wigner matrices. According to this asymptotic property, the variance of the largest eigenvalue $\lambda_N$ is thus of the order of $N^{-4/3}$. In the bulk, Gustavsson proved in [49] again for the GUE, that, for any fixed $\eta > 0$ and all $\eta N \leqslant j \leqslant (1-\eta)N$,

$$\frac{\lambda_j - \gamma_j}{\sqrt{\frac{2\log N}{(4-\gamma_j^2)N^2}}} \to \mathcal{N}(0,1) \tag{5.1}$$

in distribution as $N$ goes to infinity. This result was extended by Tao and Vu in [87] to large families of non-Gaussian Wigner matrices. The variance of an eigenvalue $\lambda_j$ in the bulk is thus of the order of $\frac{\log N}{N^2}$. Right-side intermediate eigenvalues consist in the $\lambda_j$'s with $\frac{N}{2} \leqslant j \leqslant N$ such that $\frac{j}{N} \to 1$ but $N - j \to \infty$ as $N$ goes to infinity (the left-side can be deduced by symmetry). Gustavsson proved a Central Limit Theorem for these eigenvalues (see [49]) from which their variance is guessed to be of the order of $\frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}$. This result was again extended to large classes of Wigner matrices by Tao and Vu in [86].

The previous results are however asymptotic. As announced, the purpose of this work is to provide quantitative bounds on the variance of the eigenvalues of the correct order, in the bulk and at the edge of the spectrum, as well as for some intermediate eigenvalues. In the statements below, $M_N$ is a complex (respectively real) Wigner matrix satisfying condition $(C0)$. This condition, which will be detailed in Section 5.2.1, provides an exponential decay of the matrix entries. Assume furthermore that the entries of $M_N$ have the same first four moments as the entries of a GUE matrix (respectively GOE). Set $W_N = \frac{1}{\sqrt{N}} M_N$.

**Theorem 5.1** (in the bulk)**.** *For all $0 < \eta \leqslant \frac{1}{2}$, there exists a constant $C(\eta) > 0$ such that, for all $N \geqslant 2$ and for all $\eta N \leqslant j \leqslant (1-\eta)N$,*

$$\mathrm{Var}(\lambda_j) \leqslant C(\eta) \frac{\log N}{N^2}. \tag{5.2}$$

**Theorem 5.2** (at the edge)**.** *There exists a universal constant $C > 0$ such that, for all $N \geqslant 1$,*

$$\mathrm{Var}(\lambda_1) \leqslant CN^{-4/3} \quad \textit{and} \quad \mathrm{Var}(\lambda_N) \leqslant CN^{-4/3}. \tag{5.3}$$

**Theorem 5.3** (between the bulk and the edge)**.** *For all $0 < \eta \leqslant \frac{1}{2}$ and for all $K > 20\sqrt{2}$, there exists a constant $C(\eta, K) > 0$ such that, for all $N \geqslant 2$, for all $K \log N \leqslant j \leqslant \eta N$ and $(1-\eta)N \leqslant j \leqslant N - K \log N$,*

$$\mathrm{Var}(\lambda_j) \leqslant C(\eta, K) \frac{\log\Big(\min(j, N-j)\Big)}{N^{4/3}\Big(\min(j, N-j)\Big)^{2/3}}. \tag{5.4}$$

It should be mentioned that Theorem 5.1 does not seem to be known even for Gaussian matrices. The first step is thus to prove it for the GUE. This will be achieved via the analysis of the eigenvalue counting function, which due to the particular determinantal structure in this case, has the same distribution as a sum of independent Bernoulli variables. Sharp standard deviation inequalities are thus available in this case. These may then be transferred to the eigenvalues in the bulk together with Gustavsson's bounds on the variance of the eigenvalue counting function. As a result, we actually establish that

$$E\Big[|\lambda_j - \gamma_j|^2\Big] \leqslant C(\eta) \frac{\log N}{N^2}$$

leading thus to Theorem 5.1 in this case. Similarly, Theorem 5.3 does not seem to be known for Gaussian matrices. The proof follows exactly the same scheme and we establish that

$$E\Big[|\lambda_j - \gamma_j|^2\Big] \leqslant C(\eta, K) \frac{\log\Big(\min(j, N-j)\Big)}{N^{4/3}\Big(\min(j, N-j)\Big)^{2/3}}.$$

On the other hand, Theorem 5.2 for the GUE and GOE has been known for some time (see [59]). On the basis of these results for the GUE (and GOE), Theorems 5.1, 5.2 and 5.3 are then extended to families of Wigner matrices by a suitable combination of Tao and Vu's Four Moment Theorem (see [87], [86]) and Erdös, Yau and Yin's Localization Theorem (see [35]). The basic idea is that while the localization properties almost yield the correct order, the Four Moment Theorem may be used to reach the optimal bounds by comparison with the Gaussian models. Theorems 5.1, 5.2 and 5.3 are established first in the complex case. The real case is deduced by means of interlacing formulas. Furthermore, analogous results are established for covariance matrices, for which non-asymptotic quantitative results are needed, as they are useful in several fields, such as compressed sensing (see [99]), wireless communication and quantitative finance (see [8]).

The method developed here do not seem powerful enough to strengthen the variance bounds into exponential tail inequalities. In the recent contribution [89], Tao and Vu proved such exponential tail inequalities by a further refinement of the replacement method leading to the Four Moment Theorem. While much more powerful than variance bounds, they do not seem to yield at this moment the correct order of the variance bounds of Theorems 5.1, 5.2 and 5.3.

As a corollary of the latter results, a bound on the rate of convergence of the empirical spectral measure $L_N$ can be achieved. This bound is expressed in terms of the so-called 2-Wasserstein distance $W_2$ between $L_N$ and the semicircle law $\rho_{sc}$. For $p \in [1, \infty)$, the $p$-Wasserstein distance $W_p(\mu, \nu)$ between two probability measures $\mu$ and $\nu$ on $\mathbb{R}$ is defined by

$$W_p(\mu, \nu) = \inf \left( \int_{\mathbb{R}^2} |x - y|^p \, d\pi(x, y) \right)^{1/p}$$

where the infimum is taken over all probability measure $\pi$ on $\mathbb{R}^2$ such that its first marginal is $\mu$ and its second marginal is $\nu$. Note for further purposes that the rate of convergence of this empirical distribution has been investigated in various forms. For example, the Kolmogorov distance between $L_N$ and $\rho_{sc}$ has been considered in this respect in several papers (see for example [41], [15] and [45]). It is given by

$$d_K(L_N, \rho_{sc}) = \sup_{x \in \mathbb{R}} \left| \tfrac{1}{N} \mathcal{N}_x - G(x) \right|,$$

where $\mathcal{N}_x$ is the eigenvalue counting function and $G$ is the distribution function of the semicircle law. More and more precise bounds were established. For example, under the hypothesis of an exponential decay of the matrix entries, Götze and Tikhomirov recently showed that, with high probability,

$$d_K(L_N, \rho_{sc}) \leqslant \frac{(\log N)^c}{N}$$

for some universal constant $c > 0$ (see [45]). The rate of convergence in terms of $W_1$, also called the Kantorovich-Rubinstein distance, was studied by Guionnet and Zeitouni in [47] and recently by Meckes and Meckes in [65]. In [65], the authors proved that $\mathbb{E}[W_1(L_N, \mathbb{E}[L_N])] \leqslant CN^{-2/3}$. The following is concerned with the distance between $L_N$ and $\rho_{sc}$ and strengthens the preceding conclusion on $W_1$.

**Corollary 5.4.** *There exists a numerical constant $C > 0$ such that, for all $N \geqslant 2$,*

$$\mathbb{E}\Big[W_2^2(L_N, \rho_{sc})\Big] \leqslant C\frac{\log N}{N^2}. \tag{5.5}$$

The proof of this corollary relies on the fact that $\mathbb{E}\Big[W_2^2(L_N, \rho_{sc})\Big]$ is bounded above, up to a constant, by the sum of the expectations $\mathbb{E}\Big[(\lambda_j - \gamma_j)^2\Big]$. The previously established bounds then easily yield the result.

Turning to the organization of the paper, Section 5.1 describes Theorems 5.1, 5.2 and 5.3 in the GUE case. Section 5.2 emphasizes to start with the Four Moment Theorem of Tao and Vu (see [87] and [86]) and the Localization Theorem of Erdös, Yau and Yin (see [35]). On the basis of these results and the GUE case, the main Theorems 5.1, 5.2 and 5.3 are then established for families of Wigner matrices. Section 5.3 is devoted to the corresponding statements for real matrices, while Section 5.5 describes the analogous results for covariance matrices. Section 5.4 deals with Corollary 5.4 and the rate of convergence of $L_N$ towards $\rho_{sc}$ in terms of 2-Wasserstein distance.

Throughout this paper, $C(\cdot)$ and $c(\cdot)$ will denote positive constants which depend on the specified quantities and may differ from one line to another.

## 5.1 Deviation inequalities and variance bounds in the GUE case

The results and proofs developped in this section for the GUE model heavily rely on its determinantal structure which allows for a complete description of the joint law of the eigenvalues (see [1]). In particular the eigenvalue counting function is known to have the same distribution as a sum of independent Bernoulli variables (see [50], [1]), whose mean and variance were computed by Gustavsson (see [49]). Deviation inequalities for individual eigenvalues can thus be established, leading to the announced bounds on the variance.

### 5.1.1 Eigenvalues in the bulk of the spectrum

The aim of this section is to establish the following result.

**Theorem 5.5.** *Let $M_N$ be a GUE matrix. Set $W_N = \frac{1}{\sqrt{N}} M_N$. For any $0 < \eta \leqslant \frac{1}{2}$, there exists a constant $C(\eta) > 0$ such that for all $N \geqslant 2$ and all $\eta N \leqslant j \leqslant (1-\eta)N$,*

$$\mathbb{E}\Big[|\lambda_j - \gamma_j|^2\Big] \leqslant C(\eta) \frac{\log N}{N^2}. \tag{5.6}$$

*In particular,*

$$\mathrm{Var}(\lambda_j) \leqslant C(\eta) \frac{\log N}{N^2}. \tag{5.7}$$

As announced, the proof is based on the connection between the distribution of eigenvalues and the eigenvalue counting function. For every $t \in \mathbb{R}$, let $\mathcal{N}_t = \sum_{i=1}^N \mathbb{1}_{\lambda_i \leqslant t}$ be the eigenvalue counting function. Due to the determinantal structure of the GUE, it is known (see [50]) that $\mathcal{N}_t$ has the same distribution as a sum of independent Bernoulli random variables. Bernstein's inequality for example (although other, even sharper, inequalities may be used) may then be applied to get that for every $u \geqslant 0$,

$$\mathbb{P}\Big(\Big|\mathcal{N}_t - \mathbb{E}[\mathcal{N}_t]\Big| \geqslant u\Big) \leqslant 2 \exp\Big(-\frac{u^2}{2\sigma_t^2 + u}\Big), \tag{5.8}$$

where $\sigma_t^2$ is the variance of $\mathcal{N}_t$ (see for example [95]). Note that the upper-bound is non-increasing in $u$ while non-decreasing in the variance. Set for simplicity $\rho_t = \rho_{sc}\Big((-\infty, t]\Big)$, $t \in \mathbb{R}$. It has been shown in [43] that for some numerical constant $C_1 > 0$,

$$\sup_{t \in \mathbb{R}} \Big|\mathbb{E}[\mathcal{N}_t] - N\rho_t\Big| \leqslant C_1.$$

In particular thus, together with (5.8), for every $u \geqslant 0$,

$$\mathbb{P}\Big(\Big|\mathcal{N}_t - N\rho_t\Big| \geqslant u + C_1\Big) \leqslant 2 \exp\Big(-\frac{u^2}{2\sigma_t^2 + u}\Big). \tag{5.9}$$

As a main conclusion of the work by Gustavsson [49], for every $\delta \in (0,2)$, there exists $c_\delta > 0$ such that

$$\sup_{t \in I_\delta} \sigma_t^2 \leqslant c_\delta \log N, \tag{5.10}$$

where $I_\delta = [-2+\delta, 2-\delta]$. On the basis of inequalities (5.9) and (5.10), it is then possible to derive a deviation inequality for eigenvalues $\lambda_j$ in the bulk from their theoretical locations $\gamma_j \in [-2,2]$, $1 \leqslant j \leqslant N$, defined by $\rho_{\gamma_j} = \frac{j}{N}$.

**Proposition 5.6** (Deviation inequality for $\lambda_j$)**.** *Let $\eta \in (0, \frac{1}{2}]$ and $\eta N \leqslant j \leqslant (1-\eta)N$. There exist $C > 0$, $c > 0$, $c' > 0$ and $\delta \in (0,2)$ (all depending on $\eta$) such that, for all $c \leqslant u \leqslant c'N$,*

$$\mathbb{P}\Big(|\lambda_j - \gamma_j| \geqslant \frac{u}{N}\Big) \leqslant 4 \exp\Big(-\frac{C^2 u^2}{2c_\delta \log N + Cu}\Big). \tag{5.11}$$

*Proof.* Let $\eta \in (0, \frac{1}{2}]$. To start with, evaluate, for $\frac{N}{2} \leqslant j \leqslant (1-\eta)N$ and $u \geqslant 0$, the probability $\mathbb{P}(|\lambda_j - \gamma_j| > \frac{u}{N})$. We have

$$\begin{aligned}
\mathbb{P}\Big(\lambda_j > \gamma_j + \frac{u}{N}\Big) &= \mathbb{P}\Big(\sum_{i=1}^{N} \mathbb{1}_{\lambda_i \leqslant \gamma_j + \frac{u}{N}} < j\Big) \\
&= \mathbb{P}\Big(\mathcal{N}_{\gamma_j + \frac{u}{N}} < j\Big) \\
&= \mathbb{P}\Big(N\rho_{\gamma_j + \frac{u}{N}} - \mathcal{N}_{\gamma_j + \frac{u}{N}} > N\rho_{\gamma_j + \frac{u}{N}} - j\Big) \\
&= \mathbb{P}\Big(N\rho_{\gamma_j + \frac{u}{N}} - \mathcal{N}_{\gamma_j + \frac{u}{N}} > N(\rho_{\gamma_j + \frac{u}{N}} - \rho_{\gamma_j})\Big)
\end{aligned}$$

where it has been used that $\rho_{\gamma_j} = \frac{j}{N}$. Then

$$\mathbb{P}\Big(\lambda_j > \gamma_j + \frac{u}{N}\Big) \leqslant \mathbb{P}\Big(|\mathcal{N}_{\gamma_j + \frac{u}{N}} - N\rho_{\gamma_j + \frac{u}{N}}| > N(\rho_{\gamma_j + \frac{u}{N}} - \rho_{\gamma_j})\Big).$$

But

$$\begin{aligned}
\rho_{\gamma_j + \frac{u}{N}} - \rho_{\gamma_j} &= \int_{\gamma_j}^{\gamma_j + \frac{u}{N}} \frac{1}{2\pi}\sqrt{4 - x^2}\, dx \\
&\geqslant \frac{1}{\sqrt{2}\pi} \int_{\gamma_j}^{\gamma_j + \frac{u}{N}} \sqrt{2 - x}\, dx \\
&\geqslant \frac{\sqrt{2}}{3\pi}(2 - \gamma_j)^{3/2}\left(1 - \left(1 - \frac{\frac{u}{N}}{2 - \gamma_j}\right)^{3/2}\right) \\
&\geqslant \frac{\sqrt{2}}{3\pi}(2 - \gamma_j)^{1/2}\frac{u}{N},
\end{aligned}$$

if $u \leqslant (2 - \gamma_j)N$. By definition, $1 - \frac{j}{N} = \int_{\gamma_j}^{2} \frac{1}{2\pi}\sqrt{4 - x^2}\, dx$. Then

$$\frac{1}{\sqrt{2}\pi} \int_{\gamma_j}^{2} \sqrt{2 - x}\, dx \leqslant 1 - \frac{j}{N} \leqslant \frac{1}{\pi} \int_{\gamma_j}^{2} \sqrt{2 - x}\, dx.$$

Computing the preceding integrals yields

$$\left(\frac{3\pi}{2}\frac{N - j}{N}\right)^{2/3} \leqslant 2 - \gamma_j \leqslant \left(\frac{3\pi}{\sqrt{2}}\frac{N - j}{N}\right)^{2/3}. \tag{5.12}$$

Therefore, $u \leqslant (2 - \gamma_j)N$ if $u \leqslant c'N$ with $c' \leqslant (\frac{3\pi}{2})^{2/3}\eta^{2/3}$. In this case, (5.12) yields

$$\rho_{\gamma_j + \frac{u}{N}} - \rho_{\gamma_j} \geqslant 2C\frac{u}{N}$$

where $C > 0$. Therefore

$$\begin{aligned}
\mathbb{P}\Big(\lambda_j > \gamma_j + \frac{u}{N}\Big) &\leqslant \mathbb{P}\Big(|\mathcal{N}_{\gamma_j + \frac{u}{N}} - N\rho_{\gamma_j + \frac{u}{N}}| > 2Cu\Big) \\
&\leqslant \mathbb{P}\Big(|\mathcal{N}_{\gamma_j + \frac{u}{N}} - N\rho_{\gamma_j + \frac{u}{N}}| > Cu + C_1\Big)
\end{aligned}$$

when $u \geqslant \frac{C_1}{C} = c$. Then, applying (5.9) leads to

$$\mathbb{P}\Big(\lambda_j > \gamma_j + \frac{u}{N}\Big) \leqslant 2\exp\Big(-\frac{C^2u^2}{2\sigma^2_{\gamma_j+\frac{u}{N}} + Cu}\Big).$$

As $\gamma_j$ is in the bulk, there exist $\delta$ and $c' < (\frac{3\pi}{2})^{2/3}\eta^{2/3}$ such that $\gamma_j + \frac{u}{N} \in I_\delta$, for all $\eta N \leqslant j \leqslant (1-\eta)N$ and for all $c \leqslant u \leqslant c'N$, for all $N \geqslant 1$ (both $\delta$ and $c'$ depend on $\eta$). Then

$$\mathbb{P}\Big(\lambda_j > \gamma_j + \frac{u}{N}\Big) \leqslant 2\exp\Big(-\frac{C^2u^2}{2c_\delta \log N + Cu}\Big).$$

Repeating the argument leads to the same bound on $\mathbb{P}\Big(\lambda_j < \gamma_j - \frac{u}{N}\Big)$. Therefore,

$$\mathbb{P}\Big(|\lambda_j - \gamma_j| \geqslant \frac{u}{N}\Big) \leqslant 4\exp\Big(-\frac{C^2u^2}{2c_\delta \log N + Cu}\Big).$$

The proposition is thus established.

□

*Proof of Theorem 5.5.* Note first that, for every $j$,

$$\mathbb{E}\Big[\lambda_j^4\Big] \leqslant \sum_{i=1}^{N} \mathbb{E}\Big[\lambda_i^4\Big] = \mathbb{E}\Big[\operatorname{Tr}(W_N^4)\Big].$$

The mean of this trace can be easily computed and is equal to $2N + \frac{1}{N}$. Consequently, for all $N \geqslant 1$,

$$\mathbb{E}\Big[\lambda_j^4\Big] \leqslant 3N. \tag{5.13}$$

Choose next $M = M(\eta) > 0$ large enough such that $\frac{C^2M^2}{2c_\delta + CM} > 5$. Setting $Z = N|\lambda_j - \gamma_j|$,

$$\begin{aligned}
\mathbb{E}[Z^2] &= \int_0^\infty \mathbb{P}(Z \geqslant v)2v\,dv \\
&= \int_0^c \mathbb{P}(Z \geqslant v)2v\,dv + \int_c^{M\log N} \mathbb{P}(Z \geqslant v)2v\,dv + \int_{M\log N}^\infty \mathbb{P}(Z \geqslant v)2v\,dv \\
&\leqslant c^2 + I_1 + I_2.
\end{aligned}$$

The two latter integrals are handled in different ways. The first one $I_1$ is bounded using (5.11) while $I_2$ is controlled using the Cauchy-Schwarz inequality and (5.13).

Starting thus with $I_2$,

$$\begin{aligned}
I_2 &= \int_{M\log N}^{+\infty} \mathbb{P}(Z \geqslant v) 2v\, dv \\
&\leqslant \mathbb{E}\Big[Z^2 \mathbb{1}_{Z\geqslant M\log N}\Big] \\
&\leqslant \sqrt{\mathbb{E}\big[Z^4\big]}\sqrt{\mathbb{P}\big(Z \geqslant M\log N\big)} \\
&\leqslant 2A \exp\left(\frac{1}{2}\Big(5 - \frac{C^2M^2}{2c_\delta + CM}\Big)\log N\right),
\end{aligned}$$

where $A > 0$ is a numerical constant. As $\exp\left(\frac{1}{2}\Big(5 - \frac{C^2M^2}{2c_\delta+CM}\Big)\log N\right) \underset{N\to\infty}{\to} 0$, there exists a constant $C(\eta) > 0$ such that

$$I_2 \leqslant C(\eta).$$

Turning to $I_1$, recall that Proposition 5.6 gives, for $c \leqslant v \leqslant c'N$,

$$P(Z \geqslant v) = \mathbb{P}\Big(|\lambda_j - \gamma_j| \geqslant \frac{v}{N}\Big) \leqslant 4\exp\Big(-\frac{C^2v^2}{2c_\delta \log N + Cv}\Big).$$

Hence in the range $v \leqslant M\log N$,

$$P(Z \geqslant v) \leqslant 4\exp\Big(-\frac{B}{\log N}v^2\Big),$$

where $B = B(\eta) = \frac{C^2}{2c_\delta + CM}$. There exists thus a constant $C(\eta) > 0$ such that

$$I_1 \leqslant C(\eta)\log N.$$

Summarizing the previous steps, $\mathbb{E}\big[Z^2\big] \leqslant C(\eta)\log N$. Therefore

$$\mathbb{E}\Big[|\lambda_j - \gamma_j|^2\Big] \leqslant C(\eta)\frac{\log N}{N^2},$$

which is the claim. The proof of Theorem 5.5 is complete. $\square$

It may be shown similarly that, under the assumptions of Theorem 5.5,

$$\mathbb{E}\Big[|\lambda_j - \gamma_j|^p\Big] \leqslant C(p,\eta)\frac{(\log N)^{p/2}}{N^p}.$$

### 5.1.2 Eigenvalues at the edge of the spectrum

In [59], Ledoux and Rider gave unified proofs of precise small deviation inequalities for the extreme eigenvalues of $\beta$-ensembles. The results hold in particular for GUE matrices ($\beta = 2$) and for GOE matrices ($\beta = 1$). The following theorem summarizes some of the relevant inequalities for the GUE.

**Theorem 5.7.** *There exists a universal constant $C > 0$ such that the following holds. Let $M_N$ be a GUE matrix. Set $W_N = \frac{1}{\sqrt{N}} M_N$ and denote by $\lambda_N$ the maximal eigenvalue of $W_N$. Then, for all $N \in \mathbb{N}$ and all $0 < \varepsilon \leqslant 1$,*

$$\mathbb{P}\Big(\lambda_N \geqslant 2(1+\varepsilon)\Big) \leqslant C \exp\Big(-\frac{2N}{C}\varepsilon^{3/2}\Big), \tag{5.14}$$

*and*

$$\mathbb{P}\Big(\lambda_N \leqslant 2(1-\varepsilon)\Big) \leqslant C^2 \exp\Big(-\frac{2N^2}{C}\varepsilon^{3}\Big). \tag{5.15}$$

There exists also a right-tail large deviation inequality for $\varepsilon = O(1)$ of the form

$$\mathbb{P}\Big(\lambda_N \geqslant 2(1+\varepsilon)\Big) \leqslant C \exp\Big(-\frac{2N}{C}\varepsilon^{2}\Big), \tag{5.16}$$

where $C > 0$ is a universal constant. Similar inequalities hold for the smallest eigenvalue $\lambda_1$. As stated in [59], bounds on the variance straightly follow from these deviation inequalities.

**Corollary 5.8.** *Let $M_N$ be a GUE matrix. Set $W_N = \frac{1}{\sqrt{N}} M_N$. Then there exists a universal constant $C > 0$ such that for all $N \geqslant 1$,*

$$\mathrm{Var}(\lambda_N) \leqslant \mathbb{E}\Big[(\lambda_N - 2)^2\Big] \leqslant C N^{-4/3}.$$

Similar results are probably true for the $k^{\text{th}}$ smallest or largest eigenvalue (for $k \in \mathbb{N}$ fixed).

### 5.1.3 Eigenvalues between the bulk and the edge of the spectrum

The aim of this section is to establish a bound on the variance for some intermediate eigenvalues. The proof is very similar to what was done for eigenvalues in the bulk. It relies on the fact that the eigenvalue counting function has the same distribution as a sum of independent Bernoulli variables due to the determinantal properties of the eigenvalues. A deviation inequality for individual eigenvalues can then be derived and the bound on the variance straightly follows. Parts of the proof which are identical to the proof for eigenvalues in the bulk will be omitted. In what follows we only consider the right-side of the spectrum. Results and proofs for the left-side can be deduced by replacing $N-j$ by $j$. The precise statement is the following.

**Theorem 5.9.** *Let $M_N$ be a GUE matrix. Set $W_N = \frac{1}{\sqrt{N}} M_N$. For all $K \geqslant 20\sqrt{2}$ and for all $\eta \in (0, \frac{1}{2}]$, there exists a constant $C(\eta, K) > 0$ such that for all $N \geqslant 2$ and all $(1-\eta)N \leqslant j \leqslant N - K \log N$,*

$$\mathbb{E}\Big[|\lambda_j - \gamma_j|^2\Big] \leqslant C(\eta, K) \frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}. \tag{5.17}$$

*In particular,*

$$\operatorname{Var}(\lambda_j) \leqslant C(\eta, K) \frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}. \tag{5.18}$$

The preceding theorem does not concern all intermediate eigenvalues since $N-j$ has to be at least of the order of $20\sqrt{2}\log N$ for the method used here to yield the correct order on the variance. This restriction seems however to be technical since Gustavsson [49] proved that a Central Limit Theorem holds for all eigenvalues $\lambda_j$ such that $N - j \to \infty$ but $\frac{N-j}{N} \to 0$. From this CLT, the variance of such an individual eigenvalue $\lambda_j$ is guessed to be similarly of the order of $\frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}$ for this range.

As for eigenvalues in the bulk, the proof relies on the deviation inequality for the eigenvalue counting function (5.9). From this and a bound on the variance of this counting function (5.10), it was then possible to derive a deviation inequality for eigenvalues in the bulk. The work of Gustavsson [49] suggests that for all $0 < \tilde{\eta} < 4$ and for all $\tilde{K} > 0$, there exists a constant $c_{\tilde{\eta},\tilde{K}} > 0$ such that the following holds. For every sequence $(t_N)_{N\in\mathbb{N}}$ such that, for all $N$, $0 < 2 - t_N \leqslant \tilde{\eta}$ and $N(2-t_N)^{3/2} \geqslant \tilde{K} \log N$,

$$\sigma_{t_N}^2 \leqslant c_{\tilde{\eta},\tilde{K}} \log\Big(N(2-t_N)^{3/2}\Big). \tag{5.19}$$

Similarly to the bulk case, the following proposition can be established.

**Proposition 5.10** (Deviation inequality for intermediate $\lambda_j$)**.** *There exist universal positive constants $C$ and $c$ such that the following holds. Let $K > 20\sqrt{2}$ and $\eta \in (0, \frac{1}{2}]$. Set $(1-\eta)N \leqslant j \leqslant N - K \log N$. There exists $C' > 0$ and $c' > 0$ depending on $K$ and $\eta$ such that for all $c \leqslant u \leqslant c'(N-j)$,*

$$\mathbb{P}\Big(|\lambda_j - \gamma_j| \geqslant \frac{u}{N^{2/3}(N-j)^{1/3}}\Big) \leqslant 4\exp\Big(-\frac{C^2u^2}{C'\log(N-j) + Cu}\Big). \tag{5.20}$$

*Proof.* Set $C = 2^{-5/6}(3\pi)^{-2/3}$. Let $K > 20\sqrt{2}$ and $\eta \in (0, \frac{1}{2}]$. Take $\alpha \in \Big(\frac{20\sqrt{2}}{K}, 1\Big)$ and set $c' = \alpha\Big(\frac{3\pi}{2}\Big)^{2/3}$. For $(1-\eta)N \leqslant j \leqslant N - K\log N$ and $u \geqslant 0$, set $u_{N,j} = \frac{u}{N^{2/3}(N-j)^{1/3}}$. As in the proof of Proposition 5.6, evaluating the probability $\mathbb{P}(|\lambda_j - \gamma_j| > u_{N,j})$ yields

$$\mathbb{P}\Big(\lambda_j > \gamma_j + u_{N,j}\Big) \leqslant \mathbb{P}\Big(|\mathcal{N}_{\gamma_j+u_{N,j}} - N\rho_{\gamma_j+u_{N,j}}| > N(\rho_{\gamma_j+u_{N,j}} - \rho_{\gamma_j})\Big).$$

But, if $u \leqslant N^{2/3}(N-j)^{1/3}(2-\gamma_j)$,

$$\rho_{\gamma_j+u_{N,j}} - \rho_{\gamma_j} \geqslant \frac{\sqrt{2}}{3\pi}(2-\gamma_j)^{1/2}u_{N,j}.$$

Similarly to the proof of Proposition 5.6, this condition holds if $u \leqslant c'(N-j)$. In this case,

$$\mathbb{P}\Big(\lambda_j > \gamma_j + u_{N,j}\Big) \leqslant \mathbb{P}\Big(|\mathcal{N}_{\gamma_j+u_{N,j}} - N\rho_{\gamma_j+u_{N,j}}| > Cu + C_1\Big),$$

when $u \geqslant c = \frac{C_1}{C}$. Then, applying (5.9) leads to

$$\mathbb{P}\Big(\lambda_j > \gamma_j + u_{N,j}\Big) \leqslant 2\exp\Big(-\frac{C^2u^2}{2\sigma^2_{\gamma_j+u_{N,j}} + Cu}\Big)$$

for all $c \leqslant u \leqslant c'(N-j)$. Gustavsson's result (5.19) gives a bound on $\sigma^2_{\gamma_j+u_{N,j}}$. Set $t_N = \gamma_j + u_{N,j}$. As $j \geqslant (1-\eta)N$ and $u \geqslant 0$, $0 \leqslant 2 - t_N \leqslant 2 - \gamma_j \leqslant \Big(\frac{3\pi}{2}\eta\Big)^{2/3} = \tilde{\eta}$ for all $N$. Moreover,

$$\begin{aligned}
N(2-t_N)^{3/2} &= N(2-\gamma_j-u_{N,j})^{3/2} \\
&= N(2-\gamma_j)^{3/2}\Big(1 - \frac{u_{N,j}}{2-\gamma_j}\Big)^{3/2} \\
&\geqslant \frac{3\pi}{2}(N-j)\Big(1 - \frac{c'(N-j)^{2/3}}{N^{2/3}(2-\gamma_j)}\Big)^{3/2} \\
&\geqslant \frac{3\pi}{2}(1-\alpha)^{3/2}K\log N \\
&\geqslant \tilde{K}\log N,
\end{aligned}$$

where $\tilde{K} = \frac{3\pi}{2}(1-\alpha)^{3/2}K > 0$. From (5.19), for all $c \leqslant u \leqslant c'(N-j)$,

$$\mathrm{Var}(\mathcal{N}_{\gamma_j+u_{N,j}}) \leqslant c_{\tilde{\eta},\tilde{K}}\log\Big(N(2-t_N)^{3/2}\Big).$$

But

$$
\begin{aligned}
N(2-t_N)^{3/2} &= N(2-\gamma_j-u_{N,j})^{3/2} \\
&\leqslant N(2-\gamma_j)^{3/2} \\
&\leqslant \frac{3\pi}{\sqrt{2}}(N-j).
\end{aligned}
$$

Hence $\log\Big(N(2-t_N)^{3/2}\Big) \leqslant \log(N-j)+\log(\frac{3\pi}{\sqrt{2}})$. As $K > 20\sqrt{2}$ and $N \geqslant 2$, $N-j \geqslant K\log N \geqslant \frac{3\pi}{\sqrt{2}}$ and $\mathrm{Var}(\mathcal{N}_{\gamma_j+u_{N,j}}) \leqslant 2c_{\tilde{\eta},\tilde{K}}\log(N-j)$. Therefore

$$
\mathbb{P}\Big(\lambda_j > \gamma_j+u_{N,j}\Big) \leqslant 2\exp\Big(-\frac{C^2u^2}{4c_{\tilde{\eta},\tilde{K}}\log(N-j)+Cu}\Big).
$$

The proof is concluded similarly to Proposition 5.6. □

On the basis of Proposition 5.10, we may then conclude the proof of Theorem 5.9.

*Proof of Theorem 5.9.* Setting $Z = N^{2/3}(N-j)^{1/3}|\lambda_j-\gamma_j|$,

$$
\begin{aligned}
\mathbb{E}[Z^2] &= \int_0^\infty \mathbb{P}(Z \geqslant v)2v\,dv \\
&= \int_0^c \mathbb{P}(Z \geqslant v)2v\,dv + \int_c^{\frac{C'}{C}\log(N-j)} \mathbb{P}(Z \geqslant v)2v\,dv \\
&+ \int_{\frac{C'}{C}\log(N-j)}^{c'(N-j)} \mathbb{P}(Z \geqslant v)2v\,dv + \int_{c'(N-j)}^{\infty} \mathbb{P}(Z \geqslant v)2v\,dv \\
&\leqslant c^2 + J_1 + J_2 + J_3.
\end{aligned}
$$

Repeating the computations carried out with $I_2$ in the proof of Theorem 5.5 yields

$$
J_3 \leqslant 2AN^{11/6}(N-j)^{2/3}\exp\Big(-\frac{1}{2}\frac{C^2c'^2(N-j)^2}{C'\log(N-j)+Cc'(N-j)}\Big),
$$

where $A > 0$ is a numerical constant. For $N$ large enough (depending on $\eta$ and $K$), $C'\log(N-j) \leqslant Cc'(N-j)$ and

$$
J_3 \leqslant 2AN^{5/2}\exp\Big(-\frac{Cc'}{4}(N-j)\Big).
$$

Then, as $N-j \geqslant K\log N$,

$$
J_3 \leqslant 2AN^{5/2}\exp\Big(-\frac{KCc'}{4}\log N\Big).
$$

But $\frac{KCc'}{4} = \frac{K\alpha}{8\sqrt{2}} > \frac{5}{2}$. The right-hand side goes thus to 0 when $N$ goes to infinity. As a consequence, there exists a constant $C(\eta, K) > 0$ such that

$$J_3 \leqslant C(\eta, K).$$

The integral $J_1$ is handled as $I_1$, using that, in the range $v \leqslant \frac{C'}{C}\log(N-j)$,

$$P(Z \geqslant v) \leqslant 4\exp\Big(-\frac{B}{\log(N-j)}v^2\Big),$$

where $B = B(\eta, K) = \frac{C^2}{2C'}$ (this is due to Proposition 5.10). Hence, there exists a constant $C(\eta, K)$ such that

$$J_1 \leqslant C(\eta, K)\log(N-j).$$

Finally, $J_2$ is handled similarly. From Proposition 5.10, in the range $\frac{C'}{C}\log(N-j) \leqslant v \leqslant c'(N-j)$,

$$P(Z \geqslant v) \leqslant 4\exp\Big(-\frac{C}{2}v\Big).$$

Thus

$$J_2 \leqslant 4\int_{\frac{C'}{C}\log(N-j)}^{c'(N-j)} \exp\Big(-\frac{C}{2}v\Big)2v\,dv \leqslant 4\int_0^\infty \exp\Big(-\frac{C}{2}v\Big)2v\,dv.$$

Then $J_2$ is bounded by a constant, which is independent of $\eta$ and $K$. There exists thus a constant $C > 0$ such that

$$J_2 \leqslant C.$$

Summarizing the previous steps, $\mathbb{E}\Big[Z^2\Big] \leqslant C(\eta, K)\log(N-j)$. Therefore

$$\mathbb{E}\Big[|\lambda_j - \gamma_j|^2\Big] \leqslant C(\eta, K)\frac{\log(N-j)}{N^{2/3}(N-j)^{1/3}},$$

which is the claim. □

## 5.2 Variance bounds for Wigner Hermitian matrices

As announced, the goal of this section is to prove Theorems 5.1, 5.2 and 5.3 for Wigner Hermitian matrices. The eigenvalues of a Wigner Hermitian matrix do not form a determinantal process. Therefore it does not seem easy to provide deviation inequalities for the counting function and for individual eigenvalues. However the sharp non-asymptotic bounds established in the Gaussian case can still be reached by a comparison procedure.

### 5.2.1 Localization of the eigenvalues and the Four Moment Theorem

Two main recent theorems will be used in order to carry out this comparison procedure. First, Erdös, Yau and Yin proved in [35] a Localization Theorem which gives a high probability non-asymptotic bound on the distance between an eigenvalue $\lambda_j$ and its theoretical value $\gamma_j$. Secondly, Tao and Vu's Four Moment Theorem (see [87] and [86]) provides a very useful non-asymptotic bound on the error made by approximating a statistics of the eigenvalues of a Wigner matrix by the same statistics but with the eigenvalues of a GUE matrix.

Let $M_N$ be a Wigner Hermitian matrix. Say that $M_N$ satisfies condition $(C0)$ if the real part $\xi$ and the imaginary part $\tilde{\xi}$ of $(M_N)_{ij}$ are independent and have an exponential decay: there are two positive constants $B_1$ and $B_2$ such that

$$\mathbb{P}\Big(|\xi| \geqslant t^{B_1}\Big) \leqslant e^{-t} \quad \text{and} \quad \mathbb{P}\Big(|\tilde{\xi}| \geqslant t^{B_1}\Big) \leqslant e^{-t}$$

for all $t \geqslant B_2$.

**Theorem 5.11** (Localization [35]). *Let $M_N$ be a random Hermitian matrix whose entries satisfy condition $(C0)$. There are positive universal constants $c$ and $C$ such that, for any $1 \leqslant j \leqslant N$,*

$$\mathbb{P}\Big(|\lambda_j - \gamma_j| \geqslant (\log N)^{C \log\log N} N^{-2/3} \min(j, N+1-j)^{-1/3}\Big) \leqslant C e^{-(\log N)^{c \log\log N}}. \tag{5.21}$$

This a strong localization result and it almost yields the correct order on the bound on the variance. Indeed, by means of the Cauchy-Schwarz inequality and (5.13), it can be shown that, for $\eta N \leqslant j \leqslant (1-\eta)N$,

$$\mathrm{Var}(\lambda_j) \leqslant \mathbb{E}\Big[(\lambda_j - \gamma_j)^2\Big] \leqslant \frac{C(\eta)}{N^2} (\log N)^{C \log\log N}.$$

In Tao and Vu's recent paper [89] on deviation inequalities, the authors proved a more precise localization result. They indeed established a bound similar to (5.21) but with $(\log N)^A$ instead of $(\log N)^{C \log\log N}$ (where $A > 0$ is fixed). However this more precise bound is of no help below, as the final bounds on the variances remain unchanged.

We turn now to Tao and Vu's Four Moment Theorem, in order to compare $W_N$ with a GUE matrix $W_N'$. Say that two complex random variables $\xi$ and $\xi'$ match to order $k$ if

$$\mathbb{E}\Big[\Re(\xi)^m \Im(\xi)^l\Big] = \mathbb{E}\Big[\Re(\xi')^m \Im(\xi')^l\Big]$$

for all $m, l \geqslant 0$ such that $m + l \leqslant k$.

**Theorem 5.12** (Four Moment Theorem [87], [86]). *There exists a small positive constant $c_0$ such that the following holds. Let $M_N = (\xi_{ij})_{1\leqslant i,j\leqslant N}$ and $M'_N = (\xi'_{ij})_{1\leqslant i,j\leqslant N}$ be two random Wigner Hermitian matrices satisfying condition $(C0)$. Assume that, for $1 \leqslant i < j \leqslant N$, $\xi_{ij}$ and $\xi'_{ij}$ match to order 4 and that, for $1 \leqslant i \leqslant N$, $\xi_{ii}$ and $\xi'_{ii}$ match to order 2. Set $A_N = \sqrt{N}M_N$ and $A'_N = \sqrt{N}M'_N$. Let $G : \mathbb{R} \to \mathbb{R}$ be a smooth fonction such that:*

$$\forall\ 0 \leqslant k \leqslant 5, \quad \forall x \in \mathbb{R}, \quad \left|G^{(k)}(x)\right| \leqslant N^{c_0}. \tag{5.22}$$

*Then, for all $1 \leqslant i \leqslant N$ and for $N$ large enough (depending on constants $B_1$ and $B_2$ in condition $(C0)$),*

$$\left|\mathbb{E}\Big[G(\lambda_i(A_N))\Big] - \mathbb{E}\Big[G(\lambda_i(A'_N))\Big]\right| \leqslant N^{-c_0}. \tag{5.23}$$

Actually Tao and Vu proved this theorem in a more general form, involving a finite number of eigenvalues. In this work, it will only be used with one given eigenvalue. See [87] and [86] for more details.

It should be mentionned that Tao and Vu extended in [87] Gustavsson's result (see equation (5.1)) via this theorem. By means of a smooth bump function $G$, they compared the probability for $\lambda_j$ to be in a given interval for a non-Gaussian matrix with almost the same probability but for a GUE matrix. Applying this technique to $\mathbb{P}\Big(|\lambda_j - \gamma_j| > \frac{u}{N}\Big)$ in order to extend directly the deviation inequality leads to the following in the general Wigner case: for all $\eta N \leqslant j \leqslant (1-\eta)N$ and for all $C' \leqslant u \leqslant c'N$,

$$\mathbb{P}\Big(|\lambda_j - \gamma_j| > \frac{u}{N}\Big) \leqslant C \exp\Big(-\frac{u^2}{c\log N + u}\Big) + O(N^{-c_0}),$$

where $C$, $C'$, $c$ and $c'$ are positive constants depending only on $\eta$. The bound is not exponential anymore and is not enough to conclude towards sharp bounds on the variance or higher moments.

### 5.2.2 Comparison with Gaussian matrices

Let $M_N$ be a Hermitian Wigner matrix and $M'_N$ be a GUE matrix such that they satisfy the hypotheses of Theorem 5.12. As the function $G : x \in \mathbb{R} \mapsto x^2$ does not satisfy (5.22), Theorem 5.12 will be applied to a truncation of $G$. Theorem 5.11 will provide a small area around the theoretical location $\gamma_j$ where the eigenvalue $\lambda_j$ is very likely to be in, so that the error due to the truncation will be well controlled. Note that this procedure is valid for eigenvalues in the bulk and at the edge of the spectrum, as well as for intermediate eigenvalues.

Let $1 \leqslant j \leqslant N$. Set $R_N^{(j)} = (\log N)^{C \log\log N} N^{1/3} \min(j, N+1-j)^{-1/3}$ and $\varepsilon_N = Ce^{-(\log N)^{c \log\log N}}$. Then Theorem 5.11 leads to:

$$\mathbb{P}\Big(|\lambda_j - \gamma_j| \geqslant \frac{R_N^{(j)}}{N}\Big) \leqslant \varepsilon_N. \tag{5.24}$$

Let $\psi$ be a smooth function with support $[-2, 2]$ and values in $[0, 1]$ such that $\psi(x) = \frac{1}{10}x^2$ for all $x \in [-1; 1]$. Set $G_j : x \in \mathbb{R} \mapsto \psi\Big(\frac{x - N\gamma_j}{R_N^{(j)}}\Big)$. We want to apply Tao and Vu's Four Moment Theorem 5.12 to $G_j$. As $\psi$ is smooth and has compact support, its first five derivatives are bounded by $M > 0$. Then, for all $0 \leqslant k \leqslant 5$, for all $x \in \mathbb{R}$,

$$\Big|G_j^{(k)}(x)\Big| \leqslant \frac{M}{(R_N^{(j)})^k} \leqslant N^{c_0},$$

where the last inequality holds for $N$ large enough (depending only on $M$ and $c_0$). Then, the Four Moment Theorem 5.12 yields:

$$\Big|\mathbb{E}\Big[G_j(\lambda_j(A_N))\Big] - \mathbb{E}\Big[G_j(\lambda_j(A'_N))\Big]\Big| \leqslant N^{-c_0} \tag{5.25}$$

for large enough $N$. But

$$\begin{aligned}
&\mathbb{E}\Big[G_j(\lambda_j(A_N))\Big] \\
&\quad = \tfrac{1}{10}\mathbb{E}\Big[\Big(\tfrac{\lambda_j(A_N) - N\gamma_j}{R_N^{(j)}}\Big)^2 \mathbb{1}_{\frac{|\lambda_j(A_N) - N\gamma_j|}{R_N^{(j)}} \leqslant 1}\Big] + \mathbb{E}\Big[G_j(\lambda_j(A_N))\mathbb{1}_{\frac{|\lambda_j(A_N) - N\gamma_j|}{R_N^{(j)}} > 1}\Big] \\
&\quad = \frac{N^2}{10\Big(R_N^{(j)}\Big)^2}\mathbb{E}\Big[(\lambda_j - \gamma_j)^2 \mathbb{1}_{|\lambda_j - \gamma_j| \leqslant \frac{R_N^{(j)}}{N}}\Big] + \mathbb{E}\Big[G_j(\lambda_j(A_N))\mathbb{1}_{\frac{|\lambda_j(A_N) - N\gamma_j|}{R_N^{(j)}} > 1}\Big].
\end{aligned}$$

On the one hand,

$$\mathbb{E}\Big[G_j(\lambda_j(A_N))\mathbb{1}_{\frac{|\lambda_j - N\gamma_j|}{R_N^{(j)}} > 1}\Big] \leqslant \mathbb{P}\Big(|\lambda_j(W_N) - \gamma_j| > \tfrac{R_N^{(j)}}{N}\Big) \leqslant \varepsilon_N.$$

On the other hand,

$$\mathbb{E}\Big[(\lambda_j - \gamma_j)^2 \mathbb{1}_{|\lambda_j - \gamma_j| \leqslant \frac{R_N^{(j)}}{N}}\Big] = \mathbb{E}\Big[(\lambda_j - \gamma_j)^2\Big] - \mathbb{E}\Big[(\lambda_j - \gamma_j)^2 \mathbb{1}_{|\lambda_j - \gamma_j| > \frac{R_N^{(j)}}{N}}\Big].$$

From the Cauchy-Schwarz inequality and (5.13),

$$\begin{aligned}
\mathbb{E}\Big[(\lambda_j - \gamma_j)^2 \mathbb{1}_{|\lambda_j - \gamma_j| > \frac{R_N^{(j)}}{N}}\Big] &\leqslant \sqrt{\mathbb{E}\Big[(\lambda_j - \gamma_j)^4\Big]\mathbb{P}\Big(|\lambda_j - \gamma_j| > \tfrac{R_N^{(j)}}{N}\Big)} \\
&\leqslant A\sqrt{N\varepsilon_N}
\end{aligned}$$

where $A > 0$ is a numerical constant. Then

$$\begin{aligned}\mathbb{E}\Big[G_j(\lambda_j(A_N))\Big] &= \frac{N^2}{10\Big(R_N^{(j)}\Big)^2}\Big(\mathbb{E}\Big[(\lambda_j-\gamma_j)^2\Big] + O\Big(N^{1/2}\varepsilon_N^{1/2}\Big)\Big) + O(\varepsilon_N) \\ &= \frac{N^2}{10\Big(R_N^{(j)}\Big)^2}\mathbb{E}\Big[(\lambda_j-\gamma_j)^2\Big] + O\Big(N^{5/2}\varepsilon_N^{1/2}\Big(R_N^{(j)}\Big)^{-2}\Big) + O(\varepsilon_N).\end{aligned}$$

Repeating the same computations gives similarly

$$\mathbb{E}\Big[G_j(\lambda_j(A'_N))\Big] = \frac{N^2}{10\Big(R_N^{(j)}\Big)^2}\mathbb{E}\Big[(\lambda'_j-\gamma_j)^2\Big] + O\Big(N^{5/2}\varepsilon_N^{1/2}\Big(R_N^{(j)}\Big)^{-2}\Big) + O(\varepsilon_N).$$

Then (5.25) leads to

$$\mathbb{E}\Big[(\lambda_j-\gamma_j)^2\Big] = \mathbb{E}\Big[(\lambda'_j-\gamma_j)^2\Big] + O\Big(N^{1/2}\varepsilon_N^{1/2} + N^{-2}\Big(R_N^{(j)}\Big)^2\varepsilon_N + \Big(R_N^{(j)}\Big)^2 N^{-c_0-2}\Big).$$

As the first two error terms are smaller than the third one, the preceding equation becomes

$$\mathbb{E}\Big[(\lambda_j-\gamma_j)^2\Big] = \mathbb{E}\Big[(\lambda'_j-\gamma_j)^2\Big] + O\Big(\Big(R_N^{(j)}\Big)^2 N^{-c_0-2}\Big). \tag{5.26}$$

### 5.2.3 Combining the results

We distinguish between the bulk, the edge and the intermediate cases. Note that the constants $C(\eta)$ and $C$ depend on the constants $B_1$ and $B_2$ in condition $(C0)$.

#### Eigenvalues in the bulk of the spectrum

Let $0 < \eta \leqslant \frac{1}{2}$ and $\eta N \leqslant j \leqslant (1-\eta)N$. From Theorem 5.5, $\mathbb{E}\Big[(\lambda'_j-\gamma_j)^2\Big] \leqslant C(\eta)\frac{\log N}{N^2}$. Thus, from (5.26), it remains to show that the error term is smaller than $\frac{\log N}{N^2}$. But

$$R_N^{(j)} = (\log N)^{C\log\log N} N^{1/3}\min(j, N+1-j)^{-1/3} \leqslant \eta^{-1/3}(\log N)^{C\log\log N}.$$

Then $\Big(R_N^{(j)}\Big)^2 N^{-c_0-2} = o_\eta\Big(\frac{\log N}{N^2}\Big)$. As a consequence,

$$\mathbb{E}\Big[(\lambda_j-\gamma_j)^2\Big] = \mathbb{E}\Big[(\lambda'_j-\gamma_j)^2\Big] + o_\eta\Big(\frac{\log N}{N^2}\Big)$$

and we get the desired result

$$\mathbb{E}\Big[(\lambda_j-\gamma_j)^2\Big] \leqslant C(\eta)\frac{\log N}{N^2}.$$

**Eigenvalues at the edge of the spectrum**

From Corollary 5.8, $\mathbb{E}\Big[(\lambda'_N-\gamma_N)^2\Big]=\mathbb{E}\Big[(\lambda'_N-2)^2\Big]\leqslant CN^{-4/3}$. By means of (5.26), it remains to prove that the error term is smaller than $N^{-4/3}$. We have

$$R_N^{(N)}=(\log N)^{C\log\log N}N^{1/3}.$$

Consequently $\Big(R_N^{(N)}\Big)^2N^{-c_0-2}=o\Big(N^{-4/3}\Big)$. Then

$$\mathbb{E}\Big[(\lambda_N-2)^2\Big]=\mathbb{E}\Big[(\lambda'_N-2)^2\Big]+o\Big(N^{-4/3}\Big)$$

and

$$\mathbb{E}\Big[(\lambda_N-2)^2\Big]\leqslant CN^{-4/3}.$$

As for Gaussian matrices, the same result is available for the smallest eigenvalue $\lambda_1$.

**Eigenvalues between the bulk and the edge of the spectrum**

Let $0<\eta\leqslant\frac{1}{2}$, $K>20\sqrt{2}$ and $(1-\eta)N\leqslant j\leqslant N-K\log N$. From Theorem 5.9, $\mathbb{E}\Big[(\lambda'_j-\gamma_j)^2\Big]\leqslant C(\eta,K)\frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}$. Thus, from (5.26), it remains to show that the error term is smaller than $\frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}$. But

$$R_N^{(j)}=(\log N)^{C\log\log N}N^{1/3}(N+1-j)^{-1/3}.$$

Then $\Big(R_N^{(j)}\Big)^2N^{-c_0-2}=o\Big(\frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}\Big)$. As a consequence,

$$\mathbb{E}\Big[(\lambda_j-\gamma_j)^2\Big]=\mathbb{E}\Big[(\lambda'_j-\gamma_j)^2\Big]+o\Big(\frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}\Big)$$

and we get the desired result

$$\mathbb{E}\Big[(\lambda_j-\gamma_j)^2\Big]\leqslant C(\eta,K)\frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}.$$

A similar result holds for the left-side of the spectrum.

## 5.3 Real matrices

The goal of this section is to prove Theorems 5.1, 5.2 and 5.3 for real Wigner matrices. Tao and Vu's Four Moment Theorem (Theorem 5.12) as well as Erdös, Yau and Yin's Localization Theorem (Theorem 5.11) still hold for real Wigner matrices. Section 5.2 is therefore valid for real matrices. The point is then to establish the results in the GOE case.

As announced in Section 5.1.2, the variance of eigenvalues at the edge of the spectrum is known to be bounded by $N^{-4/3}$ for GOE matrices (see [59]). The conclusion for the smallest and largest eigenvalues is then established for large families of real symmetric Wigner matrices.

$$\mathrm{Var}(\lambda_N) \leqslant \frac{\tilde{C}}{N^{4/3}} \quad \text{and} \quad \mathrm{Var}(\lambda_1) \leqslant \frac{\tilde{C}}{N^{4/3}}.$$

For eigenvalues in the bulk of the spectrum, O'Rourke proved in [69] a Central Limit Theorem which is very similar to the one established by Gustavsson in [49]. In particular, the normalisation is still of the order of $\left(\frac{\log N}{N^2}\right)^{1/2}$ and differs from the complex case only by a constant. It is therefore natural to expect the same bound on the variance for GOE matrices. The situation is completely similar for intermediate eigenvalues. But GOE matrices do not have the same determinantal properties as GUE matrices, and it is therefore not clear that a deviation inequality (similar to (5.9)) holds for the eigenvalue counting function. However, as explained by O'Rourke in [69], GOE and GUE matrices are linked by interlacing formulas established by Forrester and Rains (see [38]). These formulas lead to the following relation between the eigenvalue counting functions in the complex and in the real cases: for all $t \in \mathbb{R}$,

$$\mathcal{N}_t(W_N^{\mathbb{C}}) = \frac{1}{2}\Big(\mathcal{N}_t(W_N^{\mathbb{R}}) + \mathcal{N}_t(\tilde{W}_N^{\mathbb{R}})\Big) + \zeta_N(t), \tag{5.27}$$

where $W_N^{\mathbb{C}} = \frac{1}{\sqrt{N}} M_N^{\mathbb{C}}$ and $M_N^{\mathbb{C}}$ is from the GUE, $W_N^{\mathbb{R}} = \frac{1}{\sqrt{N}} M_N^{\mathbb{R}}$, $\tilde{W}_N^{\mathbb{R}} = \frac{1}{\sqrt{N}} \tilde{M}_N^{\mathbb{R}}$ and $M_N^{\mathbb{R}}$ and $\tilde{M}_N^{\mathbb{R}}$ are independent matrices from the GOE and $\zeta_N(t)$ takes values in $\left\{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\right\}$. See [69] for more details.

The aim is now to establish a deviation inequality for the eigenvalue counting function similar to (5.9). From (5.9), we know that for all $u \geqslant 0$,

$$\mathbb{P}\Big(|\mathcal{N}_t(W_N^{\mathbb{C}}) - N\rho_t| \geqslant u + C_1\Big) \leqslant 2\exp\Big(-\frac{u^2}{2\sigma_t^2 + u}\Big).$$

Set $C_1' = C_1 + 1$ and let $u \geqslant 0$. We can then write

$$\begin{aligned}
\mathbb{P}\Big(\mathcal{N}_t(W_N^{\mathbb{R}}) &- N\rho_t \geqslant u + C_1'\Big)^2 \\
&= \mathbb{P}\Big(\mathcal{N}_t(W_N^{\mathbb{R}}) - N\rho_t \geqslant u + C_1',\ \mathcal{N}_t(\tilde{W}_N^{\mathbb{R}}) - N\rho_t \geqslant u + C_1'\Big) \\
&\leqslant \mathbb{P}\Big(\tfrac{1}{2}\Big(\mathcal{N}_t(W_N^{\mathbb{R}}) + \mathcal{N}_t(\tilde{W}_N^{\mathbb{R}})\Big) - N\rho_t \geqslant u + C_1'\Big).
\end{aligned}$$

Then

$$\mathbb{P}\Big(\mathcal{N}_t(W_N^{\mathbb{R}}) - N\rho_t \geqslant u + C_1'\Big)^2 \leqslant \mathbb{P}\Big(\mathcal{N}_t(W_N^{\mathbb{C}}) - N\rho_t \geqslant u + C_1' - 1\Big)$$
$$\leqslant 2\exp\Big(-\frac{u^2}{2\sigma_t^2 + u}\Big).$$

Repeating the computations for $\mathbb{P}\Big(\mathcal{N}_t(W_N^{\mathbb{R}}) - N\rho_t \leqslant -u - C_1'\Big)$ and combining with the preceding yield

$$\mathbb{P}\Big(|\mathcal{N}_t(W_N^{\mathbb{R}}) - N\rho_t| \geqslant u + C_1'\Big) \leqslant 2\sqrt{2}\exp\Big(-\frac{u^2}{4\sigma_t^2 + 2u}\Big). \qquad (5.28)$$

Note that $\sigma_t^2$ is still the variance of $\mathcal{N}_t(W_N^{\mathbb{C}})$ in the preceding formula.

What remains then to be proved is very similar to the complex case. From (5.28) and Gustavsson's bounds on the variance $\sigma_t^2$ (see (5.10) for the bulk case and (5.19) for the intermediate case), deviation inequalities for individual eigenvalues can be deduced, as was done to prove Propositions 5.6 and 5.10. It is then straightforward to derive the announced bounds on the variances for GOE matrices. The argument developed in Section 5.2 in order to extend the GUE results to large families of Hermitian Wigner matrices can be reproduced to reach the desired bounds on the variances of eigenvalues in the bulk and between the bulk and the edge of the spectrum for families of real Wigner matrices. Then there exists a constant $C(\eta) > 0$ such that for all $\eta N \leqslant j \leqslant (1-\eta)N$,

$$\mathrm{Var}(\lambda_j) \leqslant \mathbb{E}\Big[(\lambda_j - \gamma_j)^2\Big] \leqslant C(\eta)\frac{\log N}{N^2},$$

and there exists a constant $C(\eta, K) > 0$ such that for all $(1-\eta)N \leqslant j \leqslant N - K\log N$ (and similarly for the left-side of the spectrum),

$$\mathrm{Var}(\lambda_j) \leqslant \mathbb{E}\Big[(\lambda_j - \gamma_j)^2\Big] \leqslant C(\eta, K)\frac{\log(N-j)}{N^{4/3}(N-j)^{2/3}}.$$

## 5.4 A corollary on the 2-Wasserstein distance

The bounds on the variances, more exactly on $\mathbb{E}[(\lambda_j - \gamma_j)^2]$, developed in the preceding sections lead to a bound on the rate of convergence of the empirical spectral measure $L_N$ towards the semicircle law $\rho_{sc}$ in terms of 2-Wasserstein distance. Recall that $W_2(L_N, \rho_{sc})$ is a random variable defined by

$$W_2(L_N, \rho_{sc}) = \inf\Big(\int_{\mathbb{R}^2} |x-y|^2\, d\pi(x,y)\Big)^{1/2},$$

where the infimum is taken over all probability measures $\pi$ on $\mathbb{R}^2$ with respective marginals $L_N$ and $\rho_{sc}$. To achieve the expected bound, we rely on another expression of $W_2$ in terms of distribution functions, namely

$$W_2^2(L_N, \rho_{sc}) = \int_0^1 \Big(F_N^{-1}(x) - G^{-1}(x)\Big)^2 \, dx, \tag{5.29}$$

where $F_N^{-1}$ (respectively $G^{-1}$) is the generalized inverse of the distribution function $F_N$ (respectively $G$) of $L_N$ (respectively $\rho_{sc}$) (see for example [100]). On the basis of this representation, the following statement may be derived.

**Proposition 5.13.** *There exists a universal constant $C > 0$ such that for all $N \geqslant 1$,*

$$W_2^2(L_N, \rho_{sc}) \leqslant \frac{2}{N} \sum_{j=1}^{N} (\lambda_j - \gamma_j)^2 + \frac{C}{N^2}. \tag{5.30}$$

*Proof.* From (5.29),

$$W_2^2(L_N, \rho_{sc}) = \int_0^1 \Big(F_N^{-1}(x) - G^{-1}(x)\Big)^2 \, dx.$$

Then,

$$\begin{aligned} W_2^2(L_N, \rho_{sc}) &= \sum_{j=1}^{N} \int_{\frac{j-1}{N}}^{\frac{j}{N}} \Big(\lambda_j - G^{-1}(x)\Big)^2 \, dx \\ &\leqslant \frac{2}{N} \sum_{j=1}^{N} (\lambda_j - \gamma_j)^2 + 2 \sum_{j=1}^{N} \int_{\frac{j-1}{N}}^{\frac{j}{N}} \Big(\gamma_j - G^{-1}(x)\Big)^2 \, dx. \end{aligned}$$

But $\gamma_j = G^{-1}\Big(\frac{j}{N}\Big)$ and $G^{-1}$ is non-decreasing. Therefore, $\Big|\gamma_j - G^{-1}(x)\Big| \leqslant \gamma_j - \gamma_{j-1}$ for all $x \in \Big[\frac{j-1}{N}, \frac{j}{N}\Big]$. Consequently,

$$W_2^2(L_N, \rho_{sc}) \leqslant \frac{2}{N} \sum_{j=1}^{N} (\lambda_j - \gamma_j)^2 + \frac{2}{N} \sum_{j=1}^{N} (\gamma_j - \gamma_{j-1})^2. \tag{5.31}$$

But if $j - 1 \geqslant \frac{N}{2}$ (and therefore $\gamma_{j-1} \geqslant 0$),

$$\begin{aligned} \frac{1}{N} &= \int_{\gamma_{j-1}}^{\gamma_j} \frac{1}{2\pi} \sqrt{4 - x^2} \, dx \\ &\geqslant \frac{1}{\sqrt{2}\pi} \int_{\gamma_{j-1}}^{\gamma_j} \sqrt{2 - x} \, dx \\ &\geqslant \frac{\sqrt{2}}{3\pi} (2 - \gamma_{j-1})^{3/2} \left(1 - \left(1 - \frac{\gamma_j - \gamma_{j-1}}{2 - \gamma_{j-1}}\right)^{3/2}\right) \\ &\geqslant \frac{\sqrt{2}}{3\pi} (2 - \gamma_{j-1})^{1/2} (\gamma_j - \gamma_{j-1}). \end{aligned}$$

Then

$$\frac{1}{N} \geqslant \frac{\sqrt{2}}{3\pi}\Big(\frac{3\pi}{2}\frac{N-j+1}{N}\Big)^{1/3}(\gamma_j - \gamma_{j-1}),$$

from (5.12). Then

$$\gamma_j - \gamma_{j-1} \leqslant \frac{(3\pi)^{2/3}2^{-1/6}}{N^{2/3}(N-j+1)^{2/3}}.$$

It may be shown that a similar bound holds if $j-1 \leqslant \frac{N}{2}$. As a summary, there exists a universal constant $c>0$ such that, for all $j \geqslant 2$,

$$\gamma_j - \gamma_{j-1} \leqslant \frac{c}{N^{2/3}\min(j, N+1-j)^{1/3}}. \tag{5.32}$$

This yields

$$\sum_{j=1}^{N}(\gamma_j - \gamma_{j-1})^2 \leqslant \frac{c^2}{N^{4/3}}\sum_{j=1}^{N}\frac{1}{\min(j, N+1-j)^{2/3}} \leqslant \frac{C}{N},$$

where $C>0$ is a universal constant. Then (5.31) becomes

$$W_2^2(L_N, \rho_{sc}) \leqslant \frac{2}{N}\sum_{j=1}^{N}(\lambda_j - \gamma_j)^2 + \frac{C}{N^2},$$

where $C>0$ is a universal constant, which is the claim. □

*Proof of Corollary 5.4.* Let $N \geqslant 2$. Due to Proposition 5.13,

$$\mathbb{E}\Big[W_2^2(L_N, \rho_{sc})\Big] \leqslant \frac{2}{N}\sum_{j=1}^{N}\mathbb{E}\Big[(\lambda_j - \gamma_j)^2\Big] + \frac{C}{N^2}.$$

We then make use of the bounds on $\mathbb{E}\Big[(\lambda_j - \gamma_j)^2\Big]$ produced in the previous sections. Set $\eta \in (0, \frac{1}{2}]$ and $K > 20\sqrt{2}$ so that $K\log N \leqslant \eta N$. We first decompose

$$\begin{aligned}
\sum_{j=1}^{N}\mathbb{E}\Big[(\lambda_j - \gamma_j)^2\Big] &= \sum_{j=1}^{K\log N}\mathbb{E}\Big[(\lambda_j - \gamma_j)^2\Big] + \sum_{j=K\log N+1}^{\eta N}\mathbb{E}\Big[(\lambda_j - \gamma_j)^2\Big] \\
&\quad + \sum_{j=\eta N+1}^{(1-\eta)N-1}\mathbb{E}\Big[(\lambda_j - \gamma_j)^2\Big] + \sum_{j=(1-\eta)N}^{N-K\log N-1}\mathbb{E}\Big[(\lambda_j - \gamma_j)^2\Big] \\
&\quad + \sum_{j=N-K\log N}^{N}\mathbb{E}\Big[(\lambda_j - \gamma_j)^2\Big] \\
&= \Sigma_1 + \Sigma_2 + \Sigma_3 + \Sigma_4 + \Sigma_5.
\end{aligned}$$

The sum $\Sigma_3$ will be bounded using the bulk case (Theorem 5.1), while Theorem 5.2 will be used to handle $\Sigma_2$ and $\Sigma_4$. A crude version of Theorem 5.11 will be enough to bound $\Sigma_1$ and $\Sigma_5$. To start with thus, from Theorem 5.1,

$$\Sigma_3 \leqslant \sum_{j=\eta N+1}^{(1-\eta)N-1} C(\eta)\frac{\log N}{N^2} \leqslant C(\eta)\frac{\log N}{N}.$$

Secondly, from Theorem 5.3,

$$\Sigma_2 + \Sigma_4 \ \leqslant\ \frac{C(\eta, K)}{N^{4/3}} \sum_{j=K\log N+1}^{\eta N} \frac{\log j}{j^{2/3}} \ \leqslant\ C(\eta, K)\frac{\log N}{N}.$$

Next $\Sigma_1$ and $\Sigma_5$ have only $K\log N$ terms. If each term is bounded by $\frac{C}{N}$ where $C$ is a positive universal constant, we get that $\Sigma_1 + \Sigma_5 \leqslant \frac{2KC\log N}{N}$, which is enough to prove the desired result on $\sum_{j=1}^{N}\mathbb{E}\Big[(\lambda_j-\gamma_j)^2\Big]$. For $N$ large enough depending only on constant $C$ in Theorem 5.11, $\frac{1}{\sqrt{N}} \geqslant \frac{(\log N)^{C\log\log N}}{N^{2/3}\min(j,N+1-j)^{1/3}}$ and Theorem 5.11 yields

$$\mathbb{P}\Big(|\lambda_j-\gamma_j| \geqslant \frac{1}{\sqrt{N}}\Big) \leqslant Ce^{-(\log N)^{c\log\log N}}.$$

Then, by the Cauchy-Schwarz inequality,

$$\begin{aligned}
\mathbb{E}\Big[(\lambda_j-\gamma_j)^2\Big] &\leqslant \mathbb{E}\Big[(\lambda_j-\gamma_j)^2\mathbb{1}_{|\lambda_j-\gamma_j|\leqslant\frac{1}{\sqrt{N}}}\Big] + \mathbb{E}\Big[(\lambda_j-\gamma_j)^2\mathbb{1}_{|\lambda_j-\gamma_j|>\frac{1}{\sqrt{N}}}\Big] \\
&\leqslant \frac{1}{N} + \sqrt{\mathbb{E}\Big[|\lambda_j-\gamma_j|^4\Big]}\sqrt{\mathbb{P}\Big(|\lambda_j-\gamma_j| > \frac{1}{\sqrt{N}}\Big)} \\
&\leqslant \frac{1}{N} + \sqrt{3}CN^{1/2}e^{-(\log N)^{c\log\log N}}.
\end{aligned}$$

As $\sqrt{3}CN^{1/2}e^{-(\log N)^{c\log\log N}} = o(\frac{1}{N})$, there exists a constant $C>0$ such that $\mathbb{E}\Big[(\lambda_j-\gamma_j)^2\Big] \leqslant \frac{C}{N}$. Then

$$\Sigma_1 + \Sigma_5 \leqslant 2KC\frac{\log N}{N}.$$

As a consequence,

$$\sum_{j=1}^{N}\mathbb{E}\Big[(\lambda_j-\gamma_j)^2\Big] \leqslant C\frac{\log N}{N}.$$

Therefore

$$\mathbb{E}\Big[W_2(L_N,\rho_{sc})^2\Big] \leqslant C\frac{\log N}{N^2},$$

where $C>0$ is a universal constant, which is the claim. The corollary is thus established. □

The preceding Corollary 5.4 implies that

$$\mathbb{E}\Big[W_1(L_N,\rho_{sc})\Big] \leqslant C\frac{\sqrt{\log N}}{N}.$$

This is an improvement of Meckes and Meckes' rate of convergence obtained in [65]. Note however that the distance studied in [65] is the expected 1-Wasserstein distance between $L_N$ and its mean instead of $\rho_{sc}$. The rate of convergence in 1-Wasserstein distance can be furthermore compared to the rate of convergence in Kolmogorov distance. Indeed, if $\mu$ and $\nu$ are two probability measures on $\mathbb{R}$ such that $\nu$ has a bounded density function with respect to Lebesgue measure, $d_K(\mu,\nu) \leqslant c\sqrt{W_1(\mu,\nu)}$, where $c>0$ depends on the bound on the density function. This is due to the fact that

$$W_1(\mu,\nu) = \sup_{f\,1\text{-Lipschitz}} \left(\int_{\mathbb{R}} f\,d\mu - \int_{\mathbb{R}} f\,d\nu\right)$$

(see for example [100]) and $d_K(\mu,\nu) = \sup_{x\in\mathbb{R}} \Big|\mu\Big((-\infty,x]\Big) - \nu\Big((-\infty,x]\Big)\Big|$. Approximating $\mathbb{1}_{(-\infty,x]}$ from above and from below by $\frac{1}{\varepsilon}$-Lipschitz functions and optimizing on $\varepsilon$ gives the result. Therefore, the preceding implies that

$$\mathbb{E}\Big[d_K(L_N,\rho_{sc})\Big] \leqslant c\frac{(\log N)^{1/4}}{N^{1/2}}$$

which is however far from Götze and Tikhomirov's recent bound [45]

$$\mathbb{E}\Big[d_K(L_N,\rho_{sc})\Big] \leqslant \frac{(\log N)^c}{N}.$$

## 5.5 Eigenvalue variance bounds for covariance matrices

This section provides the analogous non-asymptotic bounds on the variance of eigenvalues for covariance matrices. The proofs will be detailed in another redaction. Therefore, this section contains only the background and the results.

Random covariance matrices are defined by the following. Let $X$ be a $m\times n$ (real or complex) matrix, with $m \geqslant n$, such that its entries are independent, centered and have variance 1. Then $S_{m,n} = \frac{1}{m}X^*X$ is a random covariance matrix. An important example is the case when the entries of $X$ are Gaussian. Then $S_{m,n}$ belongs to the so-called Laguerre Unitary Ensemble (LUE) if the entries of $X$ are complex and to the Laguerre Orthogonal Ensemble (LOE) if they are real. $S_{m,n}$ is Hermitian (or real symmetric) and therefore has $n$ real eigenvalues. As $m \geqslant n$, none of these eigenvalues is trivial. Furthermore, these eigenvalues are nonnegative and will be denoted by $0 \leqslant \lambda_1 \leqslant \cdots \leqslant \lambda_n$.

Similarly to Wigner's Theorem, the classical Marchenko-Pastur theorem states that, if $\frac{m}{n} \to \rho \geqslant 1$ when $n$ goes to infinity, the empirical spectral measure $L_{m,n} = \frac{1}{n}\sum_{j=1}^n \delta_{\lambda_j}$ converges almost surely to a deterministic measure, called the Marchenko-Pastur distribution of parameter $\rho$. This measure is compactly supported and is absolutely continuous with respect to Lebesgue measure, with density

$$d\mu_{MP(\rho)}(x) = \frac{1}{2\pi x}\sqrt{\Big(b_\rho - x\Big)\Big(x - a_\rho\Big)}\mathbb{1}_{[a_\rho, b_\rho]}(x)dx,$$

where $a_\rho = (1-\sqrt{\rho})^2$ and $b_\rho = (1+\sqrt{\rho})^2$ (see for example [8]). Two different behaviors arise according to the value of $\rho$. If $\rho > 1$, $a_\rho > 0$ is a soft edge, which means that an eigenvalue $\lambda_j$ can be larger or smaller than $a_\rho$. On the contrary, if $\rho = 1$, $a_\rho = 0$ is called a hard edge: no eigenvalue can be less than $a_\rho$. Furthermore, the Marchenko-Pastur density function explodes at 0. It is the case in particular when $m = n$. We will denote by $\mu_{m,n}$ the approximate Marchenko-Pastur distribution whose density is defined by

$$\mu_{m,n}(x) = \frac{1}{2\pi x}\sqrt{(x - a_{m,n})(b_{m,n} - x)}\mathbb{1}_{[a_{m,n}\,;\,b_{m,n}]}(x),$$

with $a_{m,n} = (1-\sqrt{\frac{m}{n}})^2$ and $b_{m,n} = (1+\sqrt{\frac{m}{n}})^2$.

The asymptotic behaviors of individual eigenvalues for LUE matrices have been known for some time and extended to more general covariance matrices in the last decade. For an eigenvalue in the bulk of the spectrum, i.e. $\lambda_j$ such that $\eta n \leqslant j \leqslant (1-\eta)n$, a Central Limit Theorem was proved by Su (see [83]) and Tao-Vu (see [88]). From this theorem, the variance of such eigenvalues is guessed to be of the order of $\frac{\log n}{n^2}$. For right-side intermediate eigenvalues, i.e. $\lambda_j$ such that $\frac{j}{n} \to 1$ and $n - j \to \infty$, Su, and later Tao-Vu and Wang, proved a CLT (see [83], [88] and [101]). The variance appears to be of the order of $\frac{\log(n-j)}{n^{4/3}(n-j)^{2/3}}$. Similar results probably hold for the left-side of the spectrum, when $\rho > 1$. Finally, for the smallest and the largest eigenvalues $\lambda_1$ and $\lambda_n$, CLTs were proved for Gaussian matrices by Borodin-Forrester (see [16]). Several authors extended these results to more general covariance matrices. The latest results are due to Tao-Vu and Wang (see [88] and [101]). Their variances are then guessed to be of the order of $n^{-4/3}$. It should be mentionned that the result for the smallest eigenvalue only holds when $\rho > 1$. When $\rho = 1$, in the case of a squared matrix, Edelman proved a CLT for the smallest eigenvalue $\lambda_1$, extended by Tao and Vu in [85], from which the variance is guessed to be of the order of $n^{-4}$.

The following statement summarizes a number of quantitative bounds on the eigenvalues of covariance matrices which are proved by methods similar to the ones developed in the preceding sections. For simplicity, we basically assume that $\rho > 1$. More precisely, we assume that $1 < A_1 \leqslant \frac{m}{n} \leqslant A_2$ (where $A_1$ and $A_2$ are fixed constants) and that $S_{m,n}$ is a covariance matrix whose entries have an exponential decay (condition $(C0)$) and have the same first four moments as those of a LUE matrix.

**Theorem 5.14.** *1. **In the bulk of the spectrum.***
*Let $\eta \in (0, \frac{1}{2}]$. There exist a constant $C > 0$ (depending on $\eta$, $A_1$ and $A_2$) such that for all covariance matrix $S_{m,n}$, for all $\eta n \leqslant j \leqslant (1-\eta)n$,*

$$\mathrm{Var}(\lambda_j) \leqslant C \frac{\log n}{n^2}.$$

*2. **Between the bulk and the edge of the spectrum.***
*There exists a constant $\kappa > 0$ (depending on $A_1$ and $A_2$) such that the following holds. For all $K > \kappa$, for all $\eta \in (0, \frac{1}{2}]$, there exists a constant $C > 0$ (depending on $K$, $\eta$, $A_1$ and $A_2$) such that for all covariance matrix $S_{m,n}$, for all $(1-\eta)n \leqslant j \leqslant n - K \log n$,*

$$\mathrm{Var}(\lambda_j) \leqslant C \frac{\log(n-j)}{n^{4/3}(n-j)^{2/3}}.$$

*3. **At the edge of the spectrum.***
*There exists a constant $C > 0$ (depending on $A_1$ and $A_2$) such that, for all covariance matrix $S_{m,n}$,*

$$\mathrm{Var}(\lambda_n) \leqslant C n^{-4/3}.$$

As for Wigner matrices, the first two results of this theorem are first proved in the Gaussian case, using the fact that the eigenvalues of a LUE matrix form a determinantal process. It is then possible to derive a deviation inequality for the eigenvalue counting function and then for individual eigenvalues. Integrating leads to the results for LUE matrices. The result for the largest eigenvalue in the Gaussian case is already known, see [59]. These bounds are then extended to non-Gaussian matrices, relying on three recent papers. First, Pillai and Yin proved in [75] localization properties for individual eigenvalues of covariance matrices, very similar to the localization properties for Wigner matrices established by Erdös, Yau and Yin in [35]. Secondly, Tao and Vu (see [88]) and later Wang (see [101]) proved a Four Moment Theorem for these matrices. Combining these theorems as for Wigner matrices yield the theorem.

## Acknowledgements

I would like to thank my advisor, Michel Ledoux, for bringing up this problem to me, for the several discussions we had about this work, and Emmanuel Boissard for very useful conversations about Wasserstein distances.

# Chapitre 6

# Bornes sur la variance des valeurs propres de matrices de covariance

[*Ce chapitre contient le texte d'une rédaction en préparation, qui sera déposée sur Hal et ArXiv prochainement. Il est rédigé de manière à pouvoir être lu séparément et contient de ce fait des pans entiers très similaires au chapitre précédent.*]

Random covariance matrices, or Wishart matrices, were introduced by the statistician Wishart in 1928 to model tables of random data in multivariate satistics. The spectral properties of these matrices are indeed useful for elaborating statistical tests and for principal component analysis. Similarly to Wigner matrices, which were introduced by the physicist Wigner in the fifties in order to study infinite-dimensional operators in statistical physics, the asymptotic spectral properties were soon conjectured to be universal in the sense they do not depend on the distribution of the entries. Eigenvalues were studied asymptotically both at the global and local regimes, considering for instance the global behavior of the spectrum, the behavior of extreme eigenvalues or the spacings between eigenvalues in the bulk of the spectrum. In the Gaussian case, the eigenvalue joint distribution is explicitly known, allowing for a complete study of the asymptotic spectral properties (see for example [1], [8], [72]). One of the main goals of random matrix theory over the past decades was to extend these results to non-Gaussian covariance matrices.

However, in multivariate statistics, quantitative finite-range results are needed, rather than asymptotic properties. Furthermore, random covariance matrices have become useful in several other fields, such as compressed sensing (see [99]), wireless communication and quantitative finance (see [8]). In these fields too, quantitative results are needed. Several recent developments have thus been concerned with non-asymptotic random matrix theory, see for example some recent surveys and papers on this topic [77], [99] and [98]. In this paper, we investigate in this respect variance bounds on the eigenvalues of families of covariance matrices. In a preceding paper [23], we established similar bounds for Wigner matrices and the results for covariance matrices were stated but not proved. In the present paper, we provide the corresponding proofs. For the sake of completeness and in order to make the present paper readable separatly, we reproduce here some parts of the previous one [23].

Random covariance matrices are defined by the following. Let $X$ be a $m \times n$ (real or complex) matrix, with $m \geqslant n$, such that its entries are independent, centered and have variance 1. Then $S_{m,n} = \frac{1}{m} X^* X$ is a covariance matrix. An important example is the case when the entries of $X$ are Gaussian. Then $S_{m,n}$ belongs to the so-called Laguerre Unitary Ensemble (LUE) if the entries of $X$ are complex and to the Laguerre Orthogonal Ensemble (LOE) if they are real. $S_{m,n}$ is Hermitian (or real symmetric) and therefore has $n$ real eigenvalues. As $m \geqslant n$, none of these eigenvalues is trivial. Furthermore, these eigenvalues are nonnegative and will be denoted by $0 \leqslant \lambda_1 \leqslant \cdots \leqslant \lambda_n$.

Among universality results, the classical Marchenko-Pastur theorem states that, if $\frac{m}{n} \to \rho \geqslant 1$ when $n$ goes to infinity, the empirical spectral measure $\nu_{m,n} = \frac{1}{n} \sum_{j=1}^{n} \delta_{\lambda_j}$ converges almost surely to a deterministic measure, called the Marchenko-Pastur distribution of parameter $\rho$. This measure is compactly supported and is absolutely continuous with respect to Lebesgue measure, with density

$$d\mu_{MP(\rho)}(x) = \frac{1}{2\pi x} \sqrt{(b_\rho - x)(x - a_\rho)} \mathbb{1}_{[a_\rho, b_\rho]}(x) dx,$$

where $a_\rho = (1 - \sqrt{\rho})^2$ and $b_\rho = (1 + \sqrt{\rho})^2$ (see for example [8]). We denote by $\mu_{m,n}$ the approximate Marchenko-Pastur density

$$\mu_{m,n}(x) = \frac{1}{2\pi x} \sqrt{(x - a_{m,n})(b_{m,n} - x)} \mathbb{1}_{[a_{m,n}, b_{m,n}]}(x),$$

with $a_{m,n} = (1 - \sqrt{\frac{m}{n}})^2$ and $b_{m,n} = (1 + \sqrt{\frac{m}{n}})^2$. The behavior of individual eigenvalues was more difficult to achieve. At the edge of the spectrum, it was proved under a condition on the fourth moments of the entries that, almost surely,

$$a_{m,n} - \lambda_1 \underset{n \to \infty}{\to} 0 \quad \text{and} \quad \lambda_n - b_{m,n} \underset{n \to \infty}{\to} 0. \tag{6.1}$$

Once the behavior of eigenvalues at the edge of the spectrum is known, some local information on eigenvalues in the bulk can be deduced from the Marchenko-Pastur theorem. Indeed the Glivenko-Cantelli Theorem gives that almost surely

$$\sup_{x \in \mathbb{R}} |F_{m,n}(x) - G_{m,n}(x)| \to 0,$$

where $F_{m,n}$ is the distribution function of the empirical spectral distribution $L_{m,n}$ and $G_{m,n}$ is the distribution function of the approximate Marchenko-Pastur law.

Combining this with crude bounds on the Marchenko-Pastur density and with (6.1) leads to the following law of large numbers. For all $\eta > 0$, for all $\eta n \leqslant j \leqslant (1-\eta)n$, i.e. for eigenvalues in the bulk of the spectrum,

$$\lambda_j - \gamma_j^{m,n} \underset{n\to\infty}{\to} 0,$$

almost surely, where the theoretical location $\gamma_j^{m,n} \in [a_{m,n}, b_{m,n}]$ of the $j$-th eigenvalue $\lambda_j$ is defined by

$$\frac{j}{n} = \int_{a_{m,n}}^{\gamma_j^{m,n}} \mu_{m,n}(x)\, dx.$$

At the fluctuation level, the behavior of invidual eigenvalues depends heavily on their location in the spectrum and on the value of the parameter $\rho$, at least for the smallest eigenvalues. Indeed, when $\rho > 1$, the left-side of the limiting support $a_\rho$ is positive. As a consequence, eigenvalues, and in particular smallest eigenvalues, can be less than $a_\rho$, which is therefore called a soft edge. On the contrary, when $\rho = 1$, $a_\rho = 0$ and no eigenvalue can be less than $a_\rho$. In this case, the left-side is called a hard edge. Even the behavior of the Marchenko-Pastur density is different at the lower edge in these two cases. Therefore, the behavior of the smallest eigenvalue is expected to be different according to $\rho$. Indeed, on the one hand, when $m = n$ (which implies $\rho = 1$), Edelman proved that, for LUE matrices,

$$n^2 \lambda_1 \underset{n\to\infty}{\to} \mathcal{E}(1),$$

where $\mathcal{E}(1)$ is an exponential random variable with parameter 1. A similar result is available for LOE matrices, see [30] for more details. This theorem was later extended to more general covariance matrices by Tao and Vu in [85]. On the other hand, when $\rho > 1$, Borodin and Forrester proved that, for LUE matrices,

$$n^{2/3} \frac{a_{m,n} - \lambda_1}{a_{m,n}^{2/3} \left(\frac{m}{n}\right)^{-1/6}} \to F_2,$$

where $F_2$ is the so-called Tracy-Widom law (see [16]). A similar result holds for LOE matrices. These theorems were later extended to some non-Gaussian covariance matrices by Feldheim-Sodin in [36] and then to large families of covariance matrices by Wang (see [101]). On the contrary, the behavior of the largest eigenvalue relies much less on the value of the parameter $\rho$. Johansson (see [52]) proved that, for LUE matrices,

$$n^{2/3} \frac{\lambda_n - b_{m,n}}{b_{m,n}^{2/3} \left(\frac{m}{n}\right)^{-1/6}} \to F_2.$$

Johnstone proved a similar result for LOE matrices (see [54]). Soshnikov and Péché extended these theorems to more general covariance matrices in [82] and [73]. They were then extended to large families of non-Gaussian covariance matrices by Wang in [101]. From these central limit theorems, the variances of the smallest (when $\rho > 1$) and largest eigenvalues are guessed to be of the order of $n^{-4/3}$.

In the bulk of the spectrum, i.e. for all eigenvalues $\lambda_j$ such that $\eta n \leqslant j \leqslant (1-\eta)n$ for a fixed $\eta > 0$, Su proved in [83] that

$$\mu_{m,n}(\gamma_j^{m,n}) \frac{\lambda_j - \gamma_j^{m,n}}{\sqrt{\frac{1}{2\pi^2}\frac{\log n}{n^2}}} \underset{n\to\infty}{\to} \mathcal{N}(0,1),$$

in distribution. As for the largest eigenvalue, the value of parameter $\rho$ does not change significantly the behavior of eigenvalues in the bulk. This Central Limit Theorem was extended to families of non-Gaussian matrices by Tao and Vu in [88]. The variances of eigenvalues in the bulk are then guessed to be of the order of $\frac{\log n}{n^2}$. Su proved in [83] a similar Central Limit Theorem for right-side intermediate eigenvalues, which means eigenvalues $\lambda_j$ with $\frac{j}{n} \to 1$ and $n-j \to \infty$ when $n$ goes to infinity. From this theorem, the variance of such eigenvalues is guessed to be of the order of $\frac{\log(n-j)}{n^{4/3}(n-j)^{2/3}}$. This theorem was later extended to non-Gaussian covariance matrices by Wang in [101]. It seems probable that a similar result holds for left-side intermediate eigenvalues when $\rho > 1$ but Su did not carry out the computations in this case.

The aim of this paper is to provide sharp non-asymptotic bounds for the variance of individual eigenvalues of covariance matrices. For simplicity, we basically assume that $\rho > 1$. More precisely, we assume that $1 < A_1 \leqslant \frac{m}{n} \leqslant A_2$ (where $A_1$ and $A_2$ are fixed constants). When $m = n$ (therefore $\rho = 1$), it is possible to show that the following results in the bulk and on the right-side of the spectrum are true. It will be precised in the corresponding sections. Assume furthermore that $S_{m,n}$ is a complex covariance matrix (respectively real) whose entries have an exponential decay and have the same first four moments as those of a LUE (respectively LOE) matrix. This condition is called condition $(C0)$ and will be detailed in Section 6.2. The main results of this paper are the following theorems.

**Theorem 6.1** (in the bulk of the spectrum)**.** *For all $0 < \eta \leqslant \frac{1}{2}$, there exists a constant $C > 0$ depending only on $\eta$, $A_1$ and $A_2$ such that, for all $\eta n \leqslant j \leqslant (1-\eta)n$,*

$$\mathrm{Var}(\lambda_j) \leqslant C \frac{\log n}{n^2}. \tag{6.2}$$

**Theorem 6.2** (between the bulk and the edge of the spectrum)**.** *There exists a constant $\kappa > 0$ (depending on $A_1$ and $A_2$) such that the following holds. For all $K > \kappa$, for all $0 < \eta \leqslant \frac{1}{2}$, there exists a constant $C > 0$ (depending on $K$, $\eta$, $A_1$ and $A_2$) such that for all covariance matrix $S_{m,n}$, for all $(1-\eta)n \leqslant j \leqslant n - K\log n$,*

$$\mathrm{Var}(\lambda_j) \leqslant C \frac{\log(n-j)}{n^{4/3}(n-j)^{2/3}}. \tag{6.3}$$

**Theorem 6.3** (at the edge of the spectrum). *There exists a constant $C > 0$ depending only on $A_1$ and $A_2$ such that,*

$$\mathrm{Var}(\lambda_n) \leqslant Cn^{-4/3}. \tag{6.4}$$

It should be mentionned that Theorem 6.2 (respectively Theorem 6.3) probably holds for left-hand side intermediate eigenvalues (respectively the smallest eigenvalue $\lambda_1$), when $\rho > 1$. We refer to Section 6.1.2 for more details on that topic. On the contrary, when $\rho = 1$, the behavior of eigenvalues on the left-side of the spectrum is probably very different and much more difficult to study.

The first two theorems do not seem to be known even for LUE matrices. The first step is then to establish these results for such matrices. The proof relies on the fact that the eigenvalues of a LUE matrix form a determinantal process and therefore that the eigenvalue counting function has the same distribution as a sum of independent Bernoulli variables. Using a standard concentration inequality for Bernoulli variables, it is then possible to establish a deviation inequality for individual eigenvalues. A simple integration leads to the desired bounds on the variances. On the contrary, Theorem 6.3 on the largest eigenvalue $\lambda_n$ of LUE matrices has been known for some time, at least for the largest eigenvalue $\lambda_n$ (see [59]). From these results for the LUE, Theorems 6.1, 6.2 and 6.3 are then extended to large families of non-Gaussian covariance matrices by means of localization properties by Pillai and Yin (see [75]) and the Four Moment Theorem by Tao-Vu and Wang (see [88] and [101]). While the localization properties almost yield the correct order on the variance, the Four Moment Theorem is used to reach the optimal bound via a comparison with LUE matrices. Theorems 6.1, 6.2 and 6.3 are established first in the complex case. The real case is then achieved by means of interlacing formulas.

As a corollary of the preceding results and provided Theorem 6.1.2 holds also for left-hand side intermediate eigenvalues, a bound on the rate of convergence of the empirical spectral distribution $L_{m,n}$ towards the Marchenko-Pastur distribution can be achieved. It can be written in terms of the 2-Wasserstein distance between the approximate Marchenko-Pastur distribution $\mu_{m,n}$ and $L_{m,n}$, defined by the following.

$$W_2(\mu, \nu) = \inf \left( \int_{\mathbb{R}^2} |x-y|^2 \, d\pi(x,y) \right)^{1/2},$$

for $\mu$ and $\nu$ two probability measures on $\mathbb{R}$, where the infimum is taken over all probability measure $\pi$ on $\mathbb{R}^2$ such that its first marginal is $\mu$ and its second marginal is $\nu$. Note that the rate of convergence of this empirical distribution has also been investigated in terms of the Kolmogorov distance between $L_{m,n}$ and $\mu_{m,n}$ (see for example [42] and [44]). This distance is defined by

$$d_K(L_{m,n}, \mu_{m,n}) = \sup_{x \in \mathbb{R}} \left| \tfrac{1}{n}\mathcal{N}_x - G_{m,n}(x) \right|,$$

where $\mathcal{N}_x$ is the eigenvalue counting function and $G_{m,n}$ is the distribution function of the approximate Marchenko-Pastur distribution. Götze and Tikhomirov recently showed that, with high probability,

$$d_K(L_{m,n}, \mu_{m,n}) \leqslant \frac{(\log n)^c}{n}$$

for some universal constant $c > 0$ (see [44]). The rate of convergence in terms of the 1-Wasserstein distance $W_1$, also called the Kantorovich-Rubinstein distance, was studied by Guionnet and Zeitouni in [47], who proved that $\mathbb{E}[W_1(L_{m,n}, \mathbb{E}[L_{m,n}])]$ is bounded by $Cn^{-2/5}$. The following statement is concerned with the expectation of $W_2(L_{m,n}, \mu_{m,n})$.

**Corollary 6.4.** *Let $1 < A_1 < A_2$. Then there exists a constant $C > 0$ depending only on $A_1$ and $A_2$ such that, for all $m$ and $n$ such that $1 < A_1 \leqslant \frac{m}{n} \leqslant A_2$,*

$$\mathbb{E}\Big[W_2^2(L_{m,n}, \mu_{m,n})\Big] \leqslant C\frac{\log n}{n^2}. \tag{6.5}$$

The proof of this corollary relies on the fact that $\mathbb{E}[W_2^2(L_{m,n}, \mu_{m,n})]$ is bounded above, up to a constant, by the sum of the expectations $\mathbb{E}[(\lambda_j - \gamma_j^{m,n})^2]$. The previously established bounds then easily yield the result, provided Theorem 6.2 holds for left-hand side intermediate eigenvalues.

Turning now to the content of this paper, Section 6.1 describes Theorems 6.1, 6.2 and 6.3 in the LUE case. Section 6.2 starts with the Localization Theorem of Pillai and Yin (see [75]) and the Four Moment Theorem of Tao-Vu and Wang (see [88] and [101]). Theorems 6.1, 6.2 and 6.3 are then established for families of covariance matrices. Section 6.3 is devoted to real matrices. Section 6.4 deals with Corollary 6.4 and the rate of convergence of $L_{m,n}$ towards $\mu_{MP}$ in terms of 2-Wasserstein distance.

Throughout this paper, $C$, $c$ will denote positive constants, which depend on the indicated parameters and whose values may change from one line to another.

## 6.1 Deviation inequalities and variance bounds for LUE matrices

This section is concerned with Gaussian covariance matrices. The results and techniques used here heavily rely on the Gaussian structure, in particular on the determinantal properties of the eigenvalues. As a consequence of this determinantal structure, the eigenvalue counting function is known to have the same distribution as a sum of independent Bernoulli variables (see [50], [1]). Its mean and variance were computed by Su (see [83]). Deviation inequalities can therefore be established for individual eigenvalues, leading to the announced bounds on the variance. All the proofs are written in the case when $1 < A_1 \leqslant \frac{m}{n} \leqslant A_2$. Assuming $m = n$, if the results still hold, then the proofs are very similar and therefore are not reproduced.

### 6.1.1 Inside the bulk of the spectrum

The aim of this section is to prove the following theorem for eigenvalues in the bulk, i.e. for $\lambda_j$ with $\eta n \leqslant j \leqslant (1-\eta)n$.

**Theorem 6.5.** *Let $1 < A_1 < A_2$. Let $S_{m,n}$ be a LUE matrix. For any $0 < \eta \leqslant \frac{1}{2}$, there exists a constant $C > 0$ depending only on $\eta$, $A_1$ and $A_2$ such that for all $A_1 \leqslant \frac{m}{n} \leqslant A_2$ and all $\eta n \leqslant j \leqslant (1-\eta)n$,*

$$\mathbb{E}\Big[|\lambda_j - \gamma_j^{m,n}|^2\Big] \leqslant C\frac{\log n}{n^2}. \tag{6.6}$$

*In particular,*

$$\mathrm{Var}(\lambda_j) \leqslant C\frac{\log n}{n^2}. \tag{6.7}$$

The proof of this theorem relies on the properties of the eigenvalue counting function, denoted by $\mathcal{N}_t = \sum_{i=1}^n \mathbb{1}_{\lambda_i \leqslant t}$ for every $t \in \mathbb{R}$. As announced, $\mathcal{N}_t$ has the same distribution as a sum of independent Bernoulli variables. Consequently, sharp deviation inequalities are available for $\mathcal{N}_t$. Applying Bersntein's inequality leads to

$$\mathbb{P}\Big(\Big|\mathcal{N}_t - \mathbb{E}[\mathcal{N}_t]\Big| \geqslant u\Big) \leqslant 2\exp\Big(-\frac{u^2}{2\sigma_t^2 + u}\Big), \tag{6.8}$$

where $\sigma_t^2$ is the variance of $\mathcal{N}_t$ (see for example [95]). Götze and Tikhomirov proved in [42] that, as soon as $1 < A_1 \leqslant \frac{m}{n} \leqslant A_2$, there exists a positive constant $C_1$ depending only on $A_1$ and $A_2$ such that

$$\sup_{t\in\mathbb{R}} \Big|\mathbb{E}[\mathcal{N}_t] - n\int_{-\infty}^t \mu_{m,n}(x)\,dx\Big| \leqslant C_1, \tag{6.9}$$

for every LUE matrix $S_{m,n}$. For simplicity, we denote $\int_{-\infty}^t \mu_{m,n}(x)\,dx$ by $\mu_t$. Together with (6.8), for every $u \geqslant 0$,

$$\mathbb{P}\Big(|\mathcal{N}_t - n\mu_t| \geqslant u + C_1\Big) \leqslant 2\exp\Big(-\frac{u^2}{2\sigma_t^2 + u}\Big). \tag{6.10}$$

Among Su's results (see [83]), for every $\delta > 0$, there exists $c_\delta > 0$ such that

$$\sup_{t\in I_\delta} \sigma_t^2 \leqslant c_\delta \log n, \tag{6.11}$$

where $I_\delta = [a_{m,n} + \delta, b_{m,n} - \delta]$. Combining (6.10) and (6.11), deviation inequalities for individual eigenvalues in the bulk are then available, as stated in the following proposition.

**Proposition 6.6.** *Assume that* $1 < A_1 \leqslant \frac{m}{n} \leqslant A_2$. *Let* $\eta > 0$. *Then there exist positive constants* $C$, $c$, $c'$ *and* $\delta$ *such that the following holds. For any LUE matrix* $S_{m,n}$, *for all* $\eta n \leqslant j \leqslant (1-\eta)n$, *for all* $c \leqslant u \leqslant c'n$,

$$\mathbb{P}\Big(|\lambda_j - \gamma_j^{m,n}| \geqslant \frac{u}{n}\Big) \leqslant 4\exp\Big(-\frac{C^2u^2}{2c_\delta \log n + Cu}\Big). \tag{6.12}$$

*The constants* $C$, $c'$ *and* $\delta$ *depend only on* $\eta$ *and* $A_2$, *whereas the constant* $c$ *depends only on* $\eta$, $A_1$ *and* $A_2$.

Note that this proposition still holds if $m = n$, as Götze and Tikhomirov proved in [42] that (6.9) holds in that case.

*Proof.* Let $\eta > 0$ and $u \geqslant 0$. Assume first that $\frac{n}{2} \leqslant j \leqslant (1-\eta)n$. Start with estimating the probability that $\lambda_j$ is geater than $\gamma_j^{m,n} + \frac{u}{n}$.

$$\begin{aligned}
\mathbb{P}\Big(\lambda_j > \gamma_j^{m,n} + \frac{u}{n}\Big) &= \mathbb{P}\Big(\sum_{i=1}^{n} \mathbb{1}_{\lambda_i \leqslant \gamma_j^{m,n} + \frac{u}{n}} < j\Big) \\
&= \mathbb{P}\Big(\mathcal{N}_{\gamma_j^{m,n} + \frac{u}{n}} < j\Big) \\
&= \mathbb{P}\Big(n\mu_{\gamma_j^{m,n} + \frac{u}{n}} - \mathcal{N}_{\gamma_j^{m,n} + \frac{u}{n}} > n\mu_{\gamma_j^{m,n} + \frac{u}{n}} - j\Big) \\
&= \mathbb{P}\Big(n\mu_{\gamma_j^{m,n} + \frac{u}{n}} - \mathcal{N}_{\gamma_j^{m,n} + \frac{u}{n}} > n(\mu_{\gamma_j^{m,n} + \frac{u}{n}} - \mu_{\gamma_j^{m,n}})\Big) \\
&\leqslant \mathbb{P}\Big(|\mathcal{N}_{\gamma_j^{m,n} + \frac{u}{n}} - n\mu_{\gamma_j^{m,n} + \frac{u}{n}}| > n(\mu_{\gamma_j^{m,n} + \frac{u}{n}} - \mu_{\gamma_j^{m,n}})\Big),
\end{aligned}$$

where it has been used that $\mu_{\gamma_j^{m,n}} = \frac{j}{n}$. In order to use (6.10), a lower bound on $n(\mu_{\gamma_j^{m,n} + \frac{u}{n}} - \mu_{\gamma_j^{m,n}})$ is needed.

$$\begin{aligned}
\mu_{\gamma_j^{m,n} + \frac{u}{n}} - \mu_{\gamma_j^{m,n}} &= \int_{\gamma_j^{m,n}}^{\gamma_j^{m,n} + \frac{u}{n}} \frac{1}{2\pi x}\sqrt{(b_{m,n} - x)(x - a_{m,n})}\, dx \\
&\geqslant \frac{\sqrt{\gamma_j^{m,n} - a_{m,n}}}{2\pi b_{m,n}} \int_{\gamma_j^{m,n}}^{\gamma_j^{m,n} + \frac{u}{n}} \sqrt{b_{m,n} - x}\, dx \\
&\geqslant \frac{\sqrt{\gamma_j^{m,n} - a_{m,n}}}{3\pi b_{m,n}} (b_{m,n} - \gamma_j^{m,n})^{3/2}\Big(1 - \Big(1 - \frac{u/n}{b_{m,n} - \gamma_j^{m,n}}\Big)\Big) \\
&\geqslant \frac{\sqrt{\gamma_j^{m,n} - a_{m,n}}}{3\pi b_{m,n}} (b_{m,n} - \gamma_j^{m,n})^{1/2} \frac{u}{n},
\end{aligned}$$

if $\gamma_j^{m,n} + \frac{u}{n} \leqslant b_{m,n}$. Furthermore, by definition of $\gamma_j^{m,n}$,

$$
\begin{aligned}
1 - \frac{j}{n} &= \int_{\gamma_j^{m,n}}^{b_{m,n}} \frac{1}{2\pi x} \sqrt{(x - a_{m,n})(b_{m,n} - x)}\, dx \\
&\leqslant \frac{\sqrt{b_{m,n} - a_{m,n}}}{2\pi \gamma_j^{m,n}} \int_{\gamma_j^{m,n}}^{b_{m,n}} \sqrt{b_{m,n} - x}\, dx \\
&\leqslant \frac{\sqrt{b_{m,n} - a_{m,n}}}{3\pi \gamma_j^{m,n}} \left(b_{m,n} - \gamma_j^{m,n}\right)^{3/2}.
\end{aligned}
$$

Then

$$
b_{m,n} - \gamma_j^{m,n} \geqslant \left( \frac{3\pi}{\sqrt{b_{m,n} - a_{m,n}}} \gamma_j^{m,n} \left(1 - \frac{j}{n}\right)\right)^{2/3}. \tag{6.13}
$$

Moreover

$$
\begin{aligned}
\frac{j}{n} &= \int_{a_{m,n}}^{\gamma_j^{m,n}} \frac{1}{2\pi x} \sqrt{(x - a_{m,n})(b_{m,n} - x)}\, dx \\
&\leqslant \frac{\sqrt{b_{m,n} - a_{m,n}}}{2\pi} \int_{a_{m,n}}^{\gamma_j^{m,n}} \frac{1}{x} \sqrt{x - a_{m,n}}\, dx \\
&\leqslant \frac{\sqrt{b_{m,n} - a_{m,n}}}{2\pi} \int_0^{\sqrt{\gamma_j^{m,n} - a_{m,n}}} \frac{2v^2}{v^2 + a_{m,n}}\, dv,
\end{aligned}
$$

with change of variables $x = v^2 + a_{m,n}$. Therefore,

$$
\frac{j}{n} \leqslant \frac{\sqrt{b_{m,n} - a_{m,n}}}{\pi} \sqrt{\gamma_j^{m,n} - a_{m,n}}.
$$

Then

$$
\sqrt{\gamma_j^{m,n}} \geqslant \sqrt{\gamma_j^{m,n} - a_{m,n}} \geqslant \frac{\pi}{\sqrt{b_{m,n} - a_{m,n}}} \frac{j}{n} \geqslant \frac{\pi}{2\sqrt{b_{m,n} - a_{m,n}}}, \tag{6.14}
$$

as $j \geqslant \frac{n}{2}$. Therefore, as $1 - \frac{j}{n} \geqslant \eta$,

$$
b_{m,n} - \gamma_j^{m,n} \geqslant \left(\tfrac{3}{4}\right)^{2/3} \frac{\pi^2}{b_{m,n} - a_{m,n}} \eta^{2/3}. \tag{6.15}
$$

As a consequence, a lower bound on $\mu_{\gamma_j^{m,n} + \frac{u}{n}} - \mu_{\gamma_j^{m,n}}$ is achieved.

$$
\mu_{\gamma_j^{m,n} + \frac{u}{n}} - \mu_{\gamma_j^{m,n}} \geqslant \frac{C}{b_{m,n}(b_{m,n} - a_{m,n})} \eta^{1/3} \frac{u}{n},
$$

where $C > 0$ is a universal constant. As $\frac{m}{n} \leqslant A_2$,

$$\mu_{\gamma_j^{m,n}+\frac{u}{n}} - \mu_{\gamma_j^{m,n}} \geqslant \frac{C}{4\sqrt{A_2}(1+\sqrt{A_2})^2}\eta^{1/3}\frac{u}{n} = C(A_2,\eta)\frac{u}{n}.$$

Then

$$\mathbb{P}\Big(\lambda_j > \gamma_j^{m,n} + \frac{u}{n}\Big) \leqslant \mathbb{P}\Big(|\mathcal{N}_{\gamma_j^{m,n}+\frac{u}{n}} - n\mu_{\gamma_j^{m,n}+\frac{u}{n}}| > C(A_2,\eta)u\Big).$$

This is true for all $u \leqslant n(b_{m,n} - \gamma_j^{m,n})$. From (6.15), this will be true when $u \leqslant c'n$ where $c' > 0$ depends only on $A_2$ and $\eta$. If $u \geqslant c = \frac{2C_1}{C(A_2,\eta)}$, then

$$\mathbb{P}\Big(\lambda_j > \gamma_j^{m,n} + \frac{u}{n}\Big) \leqslant \mathbb{P}\Big(|\mathcal{N}_{\gamma_j^{m,n}+\frac{u}{n}} - n\mu_{\gamma_j^{m,n}+\frac{u}{n}}| > \tfrac{1}{2}C(A_2,\eta)u + C_1\Big).$$

Consequently, from (6.10), we get

$$\mathbb{P}\Big(\lambda_j > \gamma_j^{m,n} + \frac{u}{n}\Big) \leqslant 2\exp\Big(-\frac{u^2}{2\sigma^2_{\gamma_j^{m,n}+\frac{u}{n}} + u}\Big).$$

For $u \leqslant c'n$ (with a maybe smaller $c' > 0$ depending only on $\eta$ and $A_2$), there exists $\delta > 0$ depending on $\eta$ and $A_2$ such that $\gamma_j^{m,n} + \frac{u}{n} \in I_\delta$. Consequently, from (6.11), $\sigma^2_{\gamma_j^{m,n}+\frac{u}{n}} \leqslant c_\delta \log n$. Then, for $c \leqslant u \leqslant c'n$,

$$\mathbb{P}\Big(\lambda_j > \gamma_j^{m,n} + \frac{u}{n}\Big) \leqslant 2\exp\Big(-\frac{u^2}{2c_\delta \log n + u}\Big).$$

Repeating the argument leads to the same bound on $\mathbb{P}\Big(\lambda_j < \gamma_j^{m,n} - \frac{u}{n}\Big)$. Therefore,

$$\mathbb{P}\Big(|\lambda_j - \gamma_j^{m,n}| \geqslant \frac{u}{n}\Big) \leqslant 4\exp\Big(-\frac{C^2u^2}{2c_\delta \log n + Cu}\Big).$$

The case when $j \leqslant \frac{n}{2}$ is treated similarly. The proposition is thus established. $\square$

We turn now to the proof of Theorem 6.5.

*Proof of Theorem 6.5.* Note first that, for every $i$,

$$\mathbb{E}[\lambda_i^4] \leqslant \sum_{j=1}^{n} \mathbb{E}[\lambda_j^4] = \mathbb{E}[\mathrm{Tr}(S_{m,n}^4)].$$

From Hölder inequality,

$$\mathbb{E}[\mathrm{Tr}(S_{m,n}^4)] \leqslant \frac{1}{n^4} \sum_{\substack{j_1,\dots,j_4 \in [\![1,m]\!] \\ i_1,\dots,i_4 [\![1,n]\!]}} \Big(\mathbb{E}[|X_{i_1,j_1}|^8]\Big)^{1/8} \dots \Big(\mathbb{E}[|X_{i_4,j_1}|^8]\Big)^{1/8}. \tag{6.16}$$

As $S_{m,n}$ is from the LUE, the $8^{\text{th}}$ moment of its entries is $\mathbb{E}[|X_{i,j}|^8] = 105$. Then

$$\mathbb{E}[\mathrm{Tr}(S_{m,n}^4)] \leqslant 105m^4 \leqslant 105A_2n^4.$$

Consequently, for all $n \geqslant 1$, for all $1 \leqslant i \leqslant n$,

$$\mathbb{E}[\lambda_i^4] \leqslant 105A_2n^4. \tag{6.17}$$

Choose next $M > 0$ large enough such that $\frac{C^2M^2}{2c_\delta + CM} > 8$, where $C$ and $c_\delta$ are given by Proposition 6.6. M depends only on $\eta$ and $A_2$. Setting $Z = n|\lambda_j - \gamma_j^{m,n}|$,

$$\begin{aligned}
\mathbb{E}[Z^2] &= \int_0^\infty \mathbb{P}(Z \geqslant v)2v\,dv \\
&= \int_0^c \mathbb{P}(Z \geqslant v)2v\,dv + \int_c^{M\log n} \mathbb{P}(Z \geqslant v)2v\,dv + \int_{M\log n}^\infty \mathbb{P}(Z \geqslant v)2v\,dv \\
&\leqslant c^2 + I_1 + I_2.
\end{aligned}$$

The two latter integrals are handled in different ways. The first one $I_1$ is bounded using (6.12) while $I_2$ is controlled using the Cauchy-Schwarz inequality and (6.17). Starting thus with $I_2$,

$$\begin{aligned}
I_2 &= \int_{M\log n}^{+\infty} \mathbb{P}(Z \geqslant v)2v\,dv \\
&\leqslant \mathbb{E}\Big[Z^2 \mathbb{1}_{Z \geqslant M\log n}\Big] \\
&\leqslant \sqrt{\mathbb{E}\big[Z^4\big]}\sqrt{\mathbb{P}\big(Z \geqslant M\log n\big)} \\
&\leqslant An^4\sqrt{\mathbb{P}\Big(|\lambda_j - \gamma_j^{m,n}| \geqslant \frac{M\log n}{n}\Big)} \\
&\leqslant 2An^4 \exp\Big(-\frac{1}{2}\frac{C^2M^2}{2c_\delta + CM}\log n\Big) \\
&\leqslant 2A\exp\left(\frac{1}{2}\Big(8 - \frac{C^2M^2}{2c_\delta + CM}\Big)\log n\right),
\end{aligned}$$

where $A > 0$ is a numerical constant. As $\exp\left(\frac{1}{2}\Big(8 - \frac{C^2M^2}{2c_\delta+CM}\Big)\log n\right) \underset{n\to\infty}{\to} 0$, there exists a constant $C > 0$ (depending only on $\eta$ and $A_2$) such that

$$I_2 \leqslant C.$$

Turning to $I_1$, recall that Proposition 6.6 gives, for $c \leqslant v \leqslant c'n$,

$$P(Z \geqslant v) = \mathbb{P}\Big(|\lambda_j - \gamma_j^{m,n}| \geqslant \frac{v}{n}\Big) \leqslant 4\exp\Big(-\frac{C^2v^2}{2c_\kappa\log n + Cv}\Big).$$

Hence in the range $v \leqslant M \log n$,

$$P(Z \geqslant v) \leqslant 4 \exp\Big(-\frac{B}{\log n} v^2\Big),$$

where $B = B(A_2, \eta) = \frac{C^2}{2c_\delta + CM}$. As a consequence,

$$I_1 \leqslant 4 \int_c^{M \log n} \exp\Big(-\frac{B}{\log n} v^2\Big) 2v \, dv \leqslant \frac{4 \log n}{B} \int_0^\infty e^{-v^2} 2v \, dv.$$

There exists thus a constant $C > 0$ (depending only on $\eta$ and $A_2$) such that

$$I_1 \leqslant C \log n.$$

Summarizing the previous steps, $\mathbb{E}\Big[Z^2\Big] \leqslant C \log n$. Therefore

$$\mathbb{E}\Big[|\lambda_j - \gamma_j^{m,n}|^2\Big] \leqslant C \frac{\log n}{n^2},$$

$C$ depending only on $A_1$, $A_2$ and $\eta$, which is the claim. The proof of Theorem 6.5 is complete. $\square$

### 6.1.2 Between the bulk and the edge of the spectrum

The aim of this section is to prove an analogous theorem for some eigenvalues between the bulk and the right edge of the spectrum, i.e. for $\lambda_j$ such that $(1-\eta)n \leqslant j \leqslant n - K \log n$. The precise statement is the following.

**Theorem 6.7.** *There exists a constant $\kappa > 0$ (depending on $A_1$ and $A_2$) such that the following holds. For all $K > \kappa$, for all $0 < \eta \leqslant \frac{1}{2}$, there exists a constant $C > 0$ (depending on $K$, $\eta$, $A_1$ and $A_2$) such that for all covariance matrix $S_{m,n}$, for all $(1 - \eta)n \leqslant j \leqslant n - K \log n$,*

$$\mathbb{E}[(\lambda_j - \gamma_j^{m,n})^2] \leqslant C \frac{\log(n-j)}{n^{4/3}(n-j)^{2/3}}. \tag{6.18}$$

*In particular,*

$$\mathrm{Var}(\lambda_j) \leqslant C \frac{\log(n-j)}{n^{4/3}(n-j)^{2/3}}. \tag{6.19}$$

As for eigenvalues in the bulk, the proof relies on the determinantal structure of LUE matrices. Recall that this structure together with a bound on the mean counting function (6.9) leads to the following deviation inequality for the counting function $\mathcal{N}_t$.

$$\mathbb{P}\Big(|\mathcal{N}_t - n\mu_t| \geqslant u + C_1\Big) \leqslant 2 \exp\Big(-\frac{u^2}{2\sigma_t^2 + u}\Big).$$

Among Su's results (see [83]), for every $\tilde{\delta} > 0$, for every $\tilde{K} > 0$, there exists $c_{\tilde{\delta},\tilde{K}} > 0$ such that for all $t$ satisfying $0 < b_{m,n} - t < \tilde{\delta}$ and $n(b_{m,n} - t)^{3/2} \geqslant \tilde{K} \log n$,

$$\sigma_t^2 \leqslant c_{\tilde{\delta},\tilde{K}} \log n (b_{m,n} - t)^{3/2}. \tag{6.20}$$

Combining (6.10) and (6.20), deviation inequalities for individual eigenvalues in the bulk are then available.

**Proposition 6.8.** *Assume that $1 < A_1 \leqslant \frac{m}{n} \leqslant A_2$. There exists $\kappa > 0$ depending only on $A_1$ and $A_2$ such that the following holds. Let $K > \kappa$ and $0 < \eta \leqslant \frac{1}{2}$. Then there exist positive constants $C$, $c$, $C'$ and $c'$ such that the following holds. For any LUE matrix $S_{m,n}$, for all $(1-\eta)n \leqslant j \leqslant n - K \log n$, for all $c \leqslant u \leqslant c'n$,*

$$\mathbb{P}\Big(|\lambda_j - \gamma_j^{m,n}| \geqslant \frac{u}{n^{2/3}(n-j)^{1/3}}\Big) \leqslant 4 \exp\Big(-\frac{C^2u^2}{C' \log(n-j) + Cu}\Big). \tag{6.21}$$

*The constants $C$, $C'$ and $c'$ depend only on $K$, $\eta$ and $A_2$, whereas the constant $c$ depends only on $K$, $\eta$, $A_1$ and $A_2$.*

Note that this proposition still holds when $m = n$, for eigenvalues on the right-side of the spectrum. The proof is very similar to what was done for eigenvalues in the bulk. Therefore some details are not reproduced.

*Proof.* Let $\eta > 0$, $K > 0$ and $u \geqslant 0$. Assume that $(1-\eta)n \leqslant j \leqslant n - K \log n$. Set $u_{n,j} = \frac{u}{n^{2/3}(n-j)^{1/3}}$. As for the bulk case, we start with estimating the probability that $\lambda_j$ is greater than $\gamma_j^{m,n} + u_{n,j}$. We get

$$\mathbb{P}\Big(\lambda_j > \gamma_j^{m,n} + u_{n,j}\Big) \leqslant \mathbb{P}\Big(|\mathcal{N}_{\gamma_j^{m,n}+u_{n,j}} - n\mu_{\gamma_j^{m,n}+u_{n,j}}| > n(\mu_{\gamma_j^{m,n}+u_{n,j}} - \mu_{\gamma_j^{m,n}})\Big).$$

Furthermore,

$$\mu_{\gamma_j^{m,n}+u_{n,j}} - \mu_{\gamma_j^{m,n}} \geqslant \frac{\sqrt{\gamma_j^{m,n} - a_{m,n}}}{3\pi b_{m,n}} (b_{m,n} - \gamma_j^{m,n})^{1/2} u_{n,j},$$

if $\gamma_j^{m,n} + u_{n,j} \leqslant b_{m,n}$. From (6.13),

$$b_{m,n} - \gamma_j^{m,n} \geqslant \Big(\frac{3\pi}{\sqrt{b_{m,n} - a_{m,n}}} \gamma_j^{m,n} \Big(\frac{n-j}{n}\Big)\Big)^{2/3}.$$

Moreover, as $\eta \leqslant \frac{1}{2}$, from (6.14),

$$\sqrt{\gamma_j^{m,n}} \geqslant \sqrt{\gamma_j^{m,n} - a_{m,n}} \geqslant \frac{\pi}{2\sqrt{b_{m,n} - a_{m,n}}}.$$

Therefore

$$b_{m,n} - \gamma_j^{m,n} \geqslant \left(\tfrac{3}{4}\right)^{2/3} \frac{\pi^2}{b_{m,n} - a_{m,n}} \Big(\frac{n-j}{n}\Big)^{2/3}, \tag{6.22}$$

and

$$\mu_{\gamma_j^{m,n}+u_{n,j}} - \mu_{\gamma_j^{m,n}} \geqslant \frac{C}{b_{m,n}(b_{m,n} - a_{m,n})} \frac{u}{n},$$

where $C > 0$ is a universal constant. As $\frac{m}{n} \leqslant A_2$,

$$\mu_{\gamma_j^{m,n}+u_{n,j}} - \mu_{\gamma_j^{m,n}} \geqslant \frac{C}{4\sqrt{A_2}(1+\sqrt{A_2})^2} \frac{u}{n} = C(A_2)\frac{u}{n}.$$

Similarly to the bulk case, we get

$$\mathbb{P}(\lambda_j > \gamma_j^{m,n} + u_{n,j}) \leqslant 2\exp\Big(-\frac{u^2}{2\sigma^2_{\gamma_j^{m,n}+u_{n,j}} + u}\Big).$$

This relation holds if $c \leqslant u \leqslant n^{2/3}(n-j)^{1/3}(b_{m,n} - \gamma_j^{m,n})$, with $c$ depending only on $A_1$ and $A_2$. Let $\alpha \in (0,1)$. Set $c' = \alpha(\frac{3}{4})^{2/3}\frac{\pi^2}{4\sqrt{A_2}}$, depending only on $\alpha$ and $A_2$. If $u \leqslant c'(n-j)$, then, due to (6.13), the preceding relation holds. The bound (6.20) on $\sigma^2_{t_n}$ obtained by Su holds when $0 < b_{m,n} - t_n \leqslant \tilde{\delta}$ and $n(b_{m,n} - t_n)^{3/2} \geqslant \tilde{K}\log n$. Set $t_n = \gamma_j + u_{n,j}$. As $u \geqslant 0$, $0 < b_{m,n} - t_n \leqslant b_{m,n} - \gamma_j^{m,n}$. Therefore, as $j \geqslant (1-\eta)n$,

$$b_{m,n} - t_n \leqslant \Big(\frac{3b_{m,n}\sqrt{b_{m,n} - a_{m,n}}}{1-\eta}\eta\Big)^{2/3} \leqslant \Big(\frac{6(A_2)^{1/4}(1+\sqrt{A_2})^2}{1-\eta}\eta\Big)^{2/3} = \tilde{\delta}$$

for all $n$. Moreover,

$$\begin{aligned} n(b_{m,n} - t_n)^{3/2} &= n(b_{m,n} - \gamma_j^{m,n})^{3/2}\Big(1 - \frac{u_{u,j}}{b_{m,n} - \gamma_j^{m,n}}\Big)^{3/2} \\ &\geqslant \frac{3\pi^3}{4(b_{m,n} - a_{m,n})^{3/2}}(n-j)\Big(1 - \frac{c'(n-j)^{2/3}}{n^{2/3}(b_{m,n} - \gamma_j^{m,n})}\Big)^{3/2} \\ &\geqslant \frac{3\pi^3}{4(4\sqrt{A_2})^{3/2}}(1-\alpha)^{3/2}K\log n \\ &\geqslant \tilde{K}\log n, \end{aligned}$$

where $\tilde{K} = \frac{3\pi^3}{4(4\sqrt{A_2})^{3/2}}(1-\alpha)^{3/2}K > 0$. From (6.20), for all $c \leqslant u \leqslant c'(n-j)$,

$$\mathrm{Var}(\mathcal{N}_{\gamma_j^{m,n}+u_{n,j}}) \leqslant c_{\tilde{\eta},\tilde{K}} \log\Big(n(b_{m,n} - t_n)^{3/2}\Big).$$

But

$$n(b_{m,n}-t_n)^{3/2} \leqslant n\Big(b_{m,n}-\gamma_j^{m,n}\Big)^{3/2}.$$

Using the same techniques as for (6.13), it is possible to show that

$$(b_{m,n}-\gamma_j^{m,n})^{3/2} \leqslant 6b_{m,n}\sqrt{b_{m,n}-a_{m,n}}\frac{n-j}{n} \leqslant 12(A_2)^{1/4}\Big(1+\sqrt{A_2}\Big)^2\frac{n-j}{n}.$$

Hence $\log\Big(n(b_{m,n}-t_n)^{3/2}\Big) \leqslant \log(n-j)+\log(12(A_2)^{1/4}(1+\sqrt{A_2})^2)$. For $K>\kappa$ with $\kappa$ large enough depending only on $A_2$ and for $n\geqslant 2$, $n-j\geqslant K\log n \geqslant 12(A_2)^{1/4}(1+\sqrt{A_2})^2$ and $\mathrm{Var}(\mathcal{N}_{\gamma_j^{m,n}+u_{n,j}}) \leqslant 2c_{\tilde{\delta},\tilde{K}}\log(n-j)$. Therefore

$$\mathbb{P}\Big(\lambda_j>\gamma_j^{m,n}+u_{n,j}\Big) \leqslant 2\exp\Big(-\frac{C^2u^2}{4c_{\tilde{\delta},\tilde{K}}\log(n-j)+Cu}\Big).$$

The proof is concluded similarly to Proposition 6.6.

$\square$

We turn now to the proof of Theorem 6.7, in which some details are skipped, due to the similarity with the proof of Theorem 6.5.

*Proof of Theorem 6.7.* Setting $Z=n^{2/3}(n-j)^{1/3}|\lambda_j-\gamma_j^{m,n}|$,

$$\begin{aligned}
\mathbb{E}[Z^2] &= \int_0^\infty \mathbb{P}(Z\geqslant v)2v\,dv \\
&= \int_0^c \mathbb{P}(Z\geqslant v)2v\,dv + \int_c^{\frac{C'}{C}\log(n-j)} \mathbb{P}(Z\geqslant v)2v\,dv \\
&+ \int_{\frac{C'}{C}\log(n-j)}^{c'(n-j)} \mathbb{P}(Z\geqslant v)2v\,dv + \int_{c'(n-j)}^\infty \mathbb{P}(Z\geqslant v)2v\,dv \\
&\leqslant c^2+J_1+J_2+J_3.
\end{aligned}$$

Repeating the computations carried out with $I_2$ in the proof of Theorem 6.5 yields

$$\begin{aligned}
J_3 &\leqslant 2n^{4/3}(n-j)^{2/3}\sqrt{\mathbb{E}[(\lambda_j-\gamma_j^{m,n})^4]}\exp\Big(-\frac{1}{2}\frac{C^2C'^2(n-j)^2}{C'\log(n-j)+Cc'(n-j)}\Big) \\
&\leqslant 2An^4\exp\Big(-\frac{1}{2}\frac{C^2c'^2(n-j)^2}{C'\log(n-j)+Cc'(n-j)}\Big),
\end{aligned}$$

where $A>0$ is a numerical constant. The last inequality is due to (6.17). For $n$ large enough (depending on $\eta$, $A_2$ and $K$), $C'\log(n-j)\leqslant Cc'(n-j)$ and

$$J_3 \leqslant 2A\exp\Big(4\log n-\frac{Cc'}{4}(n-j)\Big).$$

Then, as $n - j \geqslant K \log n$,

$$J_3 \leqslant 2A \exp\Big(\Big(4 - \frac{KCc'}{4}\Big) \log n\Big).$$

Recall from the proof of Proposition 6.8 that $c' = \alpha c'(A_2)$ where $\alpha \in (0,1)$ is a universal constant and $c'(A_2)$ depends only on $A_2$. Furthermore, the constant $C$ depends only on $A_2$. Therefore, if we choose $\kappa > 0$ such that $\kappa > \frac{16}{Cc'(A_2)}$, then $\frac{KCc'}{4} > 4$. The right-hand side goes thus to 0 when $n$ goes to infinity. As a consequence, there exists a constant $C > 0$ depending only on $A_2$, $\eta$ and $K$ such that

$$J_3 \leqslant C.$$

The integral $J_1$ is handled as $I_1$, using that, in the range $v \leqslant \frac{C'}{C} \log(n-j)$,

$$P(Z \geqslant v) \leqslant 4 \exp\Big(-\frac{B}{\log(n-j)} v^2\Big),$$

where $B$ depends only on $K$, $\eta$ and $A_2$ (this is due to Proposition 6.8). Hence, there exists a constant $C$ depending only on $A_2$, $\eta$ and $K$ such that

$$J_1 \leqslant C \log(n-j).$$

Finally, $J_2$ is handled similarly. In the range $\frac{C'}{C} \log(n-j) \leqslant v \leqslant c'(n-j)$, from Proposition 6.8,

$$P(Z \geqslant v) \leqslant 4 \exp\Big(-\frac{C}{2} v\Big).$$

Thus

$$J_2 \leqslant 4 \int_{\frac{C'}{C}\log(n-j)}^{c'(n-j)} \exp\Big(-\frac{C}{2} v\Big) 2v \, dv \leqslant 4 \int_0^{\infty} \exp\Big(-\frac{C}{2} v\Big) 2v \, dv.$$

Then $J_2$ is bounded by a constant, which depends only on $A_2$. There exists thus a constant $C > 0$ such that

$$J_2 \leqslant C.$$

Summarizing the previous steps, $\mathbb{E}[Z^2] \leqslant C \log(n-j)$, where $C$ depends only on $A_1$, $A_2$, $\eta$ and $K$. Therefore

$$\mathbb{E}\Big[|\lambda_j - \gamma_j|^2\Big] \leqslant C \frac{\log(n-j)}{n^{2/3}(n-j)^{1/3}},$$

which is the claim. $\square$

### 6.1.3 At the edge of the spectrum

In [59], Ledoux and Rider gave unified proofs of precise small deviation inequalities for the largest eigenvalues of $\beta$-ensembles. The results hold in particular for LUE matrices ($\beta = 2$) and for LOE matrices ($\beta = 1$). The following theorem summarizes some of the relevant inequalities for the LUE.

**Theorem 6.9.** *Let $A_1 > 1$. There exists a constant $C > 0$ depending only on $A_1$ such that the following holds. Let $S_{m,n}$ be a LUE matrix. Denote by $\lambda_n$ the maximal eigenvalue of $S_{m,n}$. Then, for all $n \in \mathbb{N}$, for all $m \in \mathbb{N}$ such that $m > A_1 n$, and for all $0 < \varepsilon \leqslant 1$,*

$$\mathbb{P}\Big(\lambda_n \leqslant b_{m,n}(1-\varepsilon)\Big) \leqslant C^2 \exp\Big(-\frac{2}{C} n^2 \varepsilon^3\Big), \tag{6.23}$$

*and*

$$\mathbb{P}\Big(\lambda_n \geqslant b_{m,n}(1+\varepsilon)\Big) \leqslant C \exp\Big(-\frac{2}{C} n \varepsilon^{3/2}\Big). \tag{6.24}$$

The large deviation tails are also known. The following corollary can be deduced by integrating these inequalities.

**Corollaire 6.10.** *Let $S_{m,n}$ be a LUE matrix. Then there exists a universal constant $C > 0$ such that for all $n \geqslant 1$, for all $m \in \mathbb{N}$ such that $m > A_1 n$,*

$$\mathrm{Var}(\lambda_n) \leqslant \mathbb{E}\Big[(\lambda_n - b_{m,n})^2\Big] \leqslant C n^{-4/3}.$$

Similar results are probably true for the $k^{\text{th}}$ largest eigenvalue (for $k \in \mathbb{N}$ fixed). The authors established also a left-side deviation inequality for the smallest eigenvalue in the case when $m > A_1 n$.

$$\mathbb{P}\Big(\lambda_1 \leqslant a_{m,n}(1-\varepsilon)\Big) \leqslant C \exp\Big(-\frac{2}{C} n \varepsilon^{3/2}\Big), \tag{6.25}$$

for all $0 < \varepsilon \leqslant 1$. But no right-side deviation inequality seems to be known for the smallest eigenvalue $\lambda_1$ and therefore we cannot deduce a precise bound on the variance of the smallest eigenvalue.

## 6.2 Variance bounds for families of covariance matrices

The previously achieved bounds on the variance of eigenvalues for LUE matrices are then extended to families of more general covariance matrices. It is due to the combination of two very recent results, some localization properties established by Pillai and Yin [75] and to the Four Moment Theorem proved by Tao and Vu [88] and Wang [101].

### 6.2.1 Localization properties and the Four Moment Theorem

This subsection is devoted to the statement of the previously mentioned results which will be used in order to extend variance bounds to large families of non Gaussian covariance matrices. Matrices which are considered in this section are covariance matrices $S_{m,n}$ satisfying condition $(C0)$, defined by the following. Say that $S_{m,n}$ satisfies condition $(C0)$ if the entries $X_{ij}$ are independent and have an exponential decay: there are positive constants $B_1$ and $B_2$ such that

$$\forall\, i \in \{1,\ldots,n\}, \forall\, j \in \{1,\ldots,m\},\ \mathbb{P}\Big(|X_{ij}| \geqslant t^{B_1}\Big) \leqslant e^{-t}$$

for all $t \geqslant B_2$.

Pillai and Yin proved in [75] a Localization Theorem similar to the one proved by Erdös, Yau and Yin in [35]. This theorem establishes that the eigenvalues are highly localized around their theoretical locations $\gamma_j^{m,n}$.

**Theorem 6.11** (Localization [75])**.** *Let $S_{m,n}$ be a random covariance matrix whose entries satisfy condition $(C0)$. Suppose that $1 < A_1 \leqslant \frac{m}{n} \leqslant A_2 < +\infty$. There are positive universal constants $c$ and $C$ such that, for any $1 \leqslant j \leqslant n$,*

$$\mathbb{P}\bigg(|\lambda_j - \gamma_j^{m,n}| \geqslant (\log n)^{C \log\log n} n^{-2/3} \min(j, n+1-j)^{-1/3}\bigg) \leqslant C e^{-(\log n)^{c \log\log n}}. \tag{6.26}$$

This deviation inequality (6.26) can be used to reach an almost optimal bound on the variance. Indeed, due to (6.26) and the Cauchy-Schwartz inequality, $\mathrm{Var}(\lambda_j)$ may be bounded by $\frac{(\log n)^{2C \log\log n}}{n^2}$ in the bulk of the spectrum, which is almost the right order for the variance. In order to remove the $\log\log n$ term, we turn now to the Four Moment Theorem. This theorem was proved for the bulk of the spectrum by Tao and Vu [88] and extended to the edge by Wang [101]. From now, we consider covariance matrices $S_{m,n}$ which satisfy condition $(C0)$ and whose entries match the entries of a LUE matrix up to order 4. Say that two complex random variables $\xi$ and $\xi'$ match to order $k$ if

$$\mathbb{E}\Big[\Re(\xi)^m \Im(\xi)^l\Big] = \mathbb{E}\Big[\Re(\xi')^m \Im(\xi')^l\Big]$$

for all $m, l \geqslant 0$ such that $m + l \leqslant k$.

**Theorem 6.12** (Four Moment Theorem [88, 101])**.** *There exists a small positive constant $c_0$ such that the following holds. Let $S_{m,n} = \frac{1}{n} X^* X$ and $S'_{m,n} = \frac{1}{n} X'^* X'$ be two random covariance matrices satisfying condition $(C0)$. Assume that, for $1 \leqslant i \leqslant n$, $X_{ij}$ and $X'_{ij}$ match to order 4. Let $G : \mathbb{R} \to \mathbb{R}$ be a smooth fonction such that:*

$$\forall\ 0 \leqslant k \leqslant 5, \quad \forall x \in \mathbb{R}, \quad \Big|G^{(k)}(x)\Big| \leqslant n^{c_0}. \tag{6.27}$$

*Then, for all $1 \leqslant i \leqslant n$ and for $n$ large enough (depending on constants $B_1$ and $B_2$ in condition $(C0)$),*

$$\Big|\mathbb{E}[G(n\lambda_i)] - \mathbb{E}[G(n\lambda_i')]\Big| \leqslant n^{-c_0}. \tag{6.28}$$

Suppose Theorem 6.12 apply with $G_j : x \in \mathbb{R} \mapsto (x - n\gamma_j^{m,n})^2$. Then (6.12) writes

$$\Big|n^2\mathbb{E}[(\lambda_j - \gamma_j^{m,n})^2] - n^2\mathbb{E}[(\lambda_j' - \gamma_j^{m,n})^2]\Big| \leqslant n^{-c_0}.$$

As $\mathbb{E}[(\lambda_j' - \gamma_j^{m,n})^2]$ is bounded by $\frac{\log n}{n^2}$, $\frac{\log(n-j)}{n^{4/3}(n-j)^{2/3}}$ or $n^{-4/3}$, which are bigger than $n^{-2-c_0}$, the bounds could be extended. Unfortunately, $G_j$ does not satisfy (6.27). To get round this difficulty, Theorem 6.12 is applied to a smooth truncation of $G_j$. Theorem 6.11 provides a small area around $\gamma_j^{m,n}$ where $\lambda_j$ is very likely to be inn so that the error due to the truncation is well controlled. Details are contained in the following subsection.

### 6.2.2 Comparison with LUE matrices

Let $S_{m,n}$ be a covariance matrix and $S'_{m,n}$ be a LUE matrix such that they satisfy the hypotheses of Theorem 6.12. Note that the following procedure is valid for eigenvalues in the bulk and at the edge of the spectrum, as well as for intermediate eigenvalues.

Let $1 \leqslant j \leqslant n$. Set $R_n^{(j)} = (\log n)^{C \log\log n} n^{1/3} \min(j, n+1-j)^{-1/3}$ and $\varepsilon_n = Ce^{-(\log n)^{c \log\log n}}$. Then Theorem 6.11 leads to:

$$\mathbb{P}\Big(|\lambda_j - \gamma_j^{m,n}| \geqslant \frac{R_n^{(j)}}{n}\Big) \leqslant \varepsilon_n. \tag{6.29}$$

Let $\psi$ be a smooth function with support $[-2, 2]$ and values in $[0, 1]$ such that $\psi(x) = \frac{1}{10}x^2$ for all $x \in [-1; 1]$. Set $G_j : x \in \mathbb{R} \mapsto \psi\Big(\frac{x - n\gamma_j}{R_n^{(j)}}\Big)$. We want to apply Tao and Vu's Four Moment Theorem 6.12 to $G_j$. As $\psi$ is smooth and has compact support, its first five derivatives are bounded by $M > 0$. Then, for all $0 \leqslant k \leqslant 5$, for all $x \in \mathbb{R}$,

$$\Big|G_j^{(k)}(x)\Big| \leqslant \frac{M}{(R_n^{(j)})^k} \leqslant n^{c_0},$$

where the last inequality holds for $n$ large enough (depending only on $M$ and $c_0$). Then, the Four Moment Theorem 6.12 yields:

$$\Big|\mathbb{E}[G_j(n\lambda_j)] - \mathbb{E}[G_j(n\lambda_j')]\Big| \leqslant n^{-c_0} \tag{6.30}$$

for large enough $n$. But

$$\begin{aligned}
&\mathbb{E}[G_j(n\lambda_j)] \\
&\quad = \tfrac{1}{10}\mathbb{E}\Big[\Big(\tfrac{n\lambda_j - n\gamma_j^{m,n}}{R_n^{(j)}}\Big)^2 \mathbb{1}_{\frac{|n\lambda_j - n\gamma_j^{m,n}|}{R_n^{(j)}} \leqslant 1}\Big] + \mathbb{E}\Big[G_j(n\lambda_j)\mathbb{1}_{\frac{|n\lambda_j - n\gamma_j^{m,n}|}{R_n^{(j)}} > 1}\Big] \\
&\quad = \tfrac{n^2}{10(R_n^{(j)})^2}\mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2 \mathbb{1}_{|\lambda_j - \gamma_j^{m,n}| \leqslant \frac{R_n^{(j)}}{n}}\Big] + \mathbb{E}\Big[G_j(n\lambda_j)\mathbb{1}_{\frac{|n\lambda_j - n\gamma_j^{m,n}|}{R_n^{(j)}} > 1}\Big].
\end{aligned}$$

On the one hand,

$$\begin{aligned}
\mathbb{E}\Big[G_j(n\lambda_j)\mathbb{1}_{\frac{|n\lambda_j - n\gamma_j^{m,n}|}{R_n^{(j)}} > 1}\Big] &\leqslant \mathbb{P}\Big(|n\lambda_j - n\gamma_j^{m,n}| > R_n^{(j)}\Big) \\
&\leqslant \mathbb{P}\Big(|\lambda_j - \gamma_j^{m,n}| > \frac{R_n^{(j)}}{n}\Big) \\
&\leqslant \varepsilon_n.
\end{aligned}$$

On the other hand,

$$\mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2 \mathbb{1}_{|\lambda_j - \gamma_j^{m,n}| \leqslant \frac{R_n^{(j)}}{n}}\Big] = \mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big] - \mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2 \mathbb{1}_{|\lambda_j - \gamma_j^{m,n}| > \frac{R_n^{(j)}}{n}}\Big].$$

As condition $(C0)$ is satisfied, the 8-th moment of the entries is uniformly bounded by a constant which depends only on constants $B_1$ and $B_2$ in condition $(C0)$. Then, from the Cauchy-Schwarz inequality and (6.16),

$$\begin{aligned}
\mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2 \mathbb{1}_{|\lambda_j - \gamma_j| > \frac{R_n^{(j)}}{n}}\Big] &\leqslant \sqrt{\mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^4\Big]\mathbb{P}\Big(|\lambda_j - \gamma_j| > \tfrac{R_n^{(j)}}{n}\Big)} \\
&\leqslant An^2\sqrt{\varepsilon_n}
\end{aligned}$$

where $A > 0$ is a numerical constant. Then

$$\begin{aligned}
\mathbb{E}\Big[G_j(n\lambda_j)\Big] &= \frac{n^2}{10\Big(R_n^{(j)}\Big)^2}\Big(\mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big] + O\Big(n^2\varepsilon_n^{1/2}\Big)\Big) + O(\varepsilon_n) \\
&= \frac{n^2}{10\Big(R_n^{(j)}\Big)^2}\mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big] + O\Big(n^4\varepsilon_n^{1/2}\Big(R_n^{(j)}\Big)^{-2}\Big) + O(\varepsilon_n).
\end{aligned}$$

Repeating the same computations gives similarly

$$\mathbb{E}\Big[G_j(n\lambda_j')\Big] = \frac{n^2}{10\Big(R_n^{(j)}\Big)^2}\mathbb{E}\Big[(\lambda_j' - \gamma_j^{m,n})^2\Big] + O\Big(n^4\varepsilon_n^{1/2}\Big(R_n^{(j)}\Big)^{-2}\Big) + O(\varepsilon_n).$$

Then (6.30) yields

$$\mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big] = \mathbb{E}\Big[(\lambda'_j - \gamma_j^{m,n})^2\Big] + O\Big(n^2 \varepsilon_n^{1/2} + n^{-2}(R_n^{(j)})^2 \varepsilon_n + (R_n^{(j)})^2 n^{-c_0-2}\Big).$$

As the first two error terms are smaller than the third one, the preceding equation becomes

$$\mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big] = \mathbb{E}\Big[(\lambda'_j - \gamma_j^{m,n})^2\Big] + O\Big((R_n^{(j)})^2 n^{-c_0-2}\Big). \qquad (6.31)$$

### 6.2.3 Variance bounds

(6.31) is true for all eigenvalue $\lambda_j$. We estimate the error term $O\Big((R_n^{(j)})^2 n^{-c_0-2}\Big)$ differently according to the location of the eigenvalue in the spectrum, in order to get the announced bounds.

#### Inside the bulk of the spectrum

Let $0 < \eta \leqslant \frac{1}{2}$ and $\eta n \leqslant j \leqslant (1-\eta)n$. From Theorem 6.5, $\mathbb{E}\Big[(\lambda'_j - \gamma_j^{m,n})^2\Big] \leqslant C\frac{\log n}{n^2}$. Thus, from (6.31), it remains to show that the error term is smaller than $\frac{\log n}{n^2}$. But

$$R_n^{(j)} = (\log n)^{C \log\log n} n^{1/3} \min(j, n+1-j)^{-1/3} \leqslant \eta^{-1/3} (\log n)^{C \log\log n}.$$

Then $(R_n^{(j)})^2 n^{-c_0-2} = o_\eta\Big(\frac{\log n}{n^2}\Big)$. As a consequence,

$$\mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big] = \mathbb{E}\Big[(\lambda'_j - \gamma_j^{m,n})^2\Big] + o_\eta\Big(\frac{\log n}{n^2}\Big)$$

and we get the desired result

$$\mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big] \leqslant C\frac{\log n}{n^2},$$

$C$ depending only on $\eta$, $A_1$ and $A_2$.

#### Between the bulk and the edge of the spectrum

Let $0 < \eta \leqslant \frac{1}{2}$, $K > \kappa$ and $(1-\eta)n \leqslant j \leqslant n - K\log n$. From Theorem 6.7, $\mathbb{E}\Big[(\lambda'_j - \gamma_j^{m,n})^2\Big] \leqslant C\frac{\log(n-j)}{n^{4/3}(n-j)^{2/3}}$. Thus, from (6.31), it remains to show that the error term is smaller than $\frac{\log(n-j)}{n^{4/3}(n-j)^{2/3}}$. But

$$R_n^{(j)} = (\log n)^{C \log\log n} n^{1/3} (n+1-j)^{-1/3}.$$

Then $(R_n^{(j)})^2 n^{-c_0-2} = o\Big(\frac{\log(n-j)}{n^{4/3}(n-j)^{2/3}}\Big)$. As a consequence,

$$\mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big] = \mathbb{E}\Big[(\lambda'_j - \gamma_j^{m,n})^2\Big] + o\Big(\frac{\log(n-j)}{n^{4/3}(n-j)^{2/3}}\Big)$$

and we get the desired result

$$\mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big] \leqslant C\frac{\log(n-j)}{n^{4/3}(n-j)^{2/3}},$$

$C$ depending only on $\eta$, $A_1$, $A_2$ and $K$. A similar result probably holds for the left-side of the spectrum, when $\rho > 1$.

**At the edge of the spectrum**

From Corollary 6.10, $\mathbb{E}\Big[(\lambda'_n - \gamma_n^{m,n})^2\Big] = \mathbb{E}\Big[(\lambda'_n - b_{m,n})^2\Big] \leqslant Cn^{-4/3}$. By means of (6.31), it remains to prove that the error term is smaller than $n^{-4/3}$. We have

$$R_n^{(n)} = (\log n)^{C\log\log n} n^{1/3}.$$

Consequently $(R_n^{(n)})^2 n^{-c_0-2} = o\Big(n^{-4/3}\Big)$. Then

$$\mathbb{E}\Big[(\lambda_n - b_{m,n})^2\Big] = \mathbb{E}\Big[(\lambda'_n - 2)^2\Big] + o\Big(n^{-4/3}\Big)$$

and

$$\mathbb{E}\Big[(\lambda_n - 2)^2\Big] \leqslant Cn^{-4/3}.$$

If this bound holds for the smallest eigenvalue $\lambda_1$ of LUE matrices, the same result is available for non Gaussian covariance matrices.

## 6.3 Real Wishart matrices

The aim of this section is to prove Theorems 6.1, 6.2 and 6.3 for real covariance matrices. The Four Moment Theorem (Theorem 6.12) by Tao, Vu and Wang as well as Pillai and Yin's Localization Theorem (Theorem 6.11) still hold for real covariance matrices. Section 6.2 is therefore valid for real matrices. The point is then to establish the results in the LOE case.

As announced in Section 6.1.3, the variance of eigenvalues at the right edge of the spectrum is known to be bounded by $n^{-4/3}$ for LOE matrices (see [59]). The conclusion for the largest eigenvalue is then established for large families of real covariance matrices.

$$\mathrm{Var}(\lambda_n) \leqslant \frac{\tilde{C}}{n^{4/3}}.$$

For eigenvalues in the bulk of the spectrum, following O'Rourke's approach (see [69]), a Central Limit Theorem similar to the one established by Su in [83] may be proved. In particular, the normalisation is still of the order of $(\frac{\log n}{n^2})^{1/2}$ and differs from the complex case only by a constant. It is therefore natural to expect the same bound on the variance for LOE matrices. The situation is completely similar for intermediate eigenvalues. But LOE matrices do not have the same determinantal properties as LUE matrices, and it is therefore not clear that a deviation inequality (similar to (6.10)) holds for the eigenvalue counting function. However, LOE and LUE matrices are linked by interlacing formulas established by Forrester and Rains (see [38]). These formulas lead to the following relation between the eigenvalue counting functions in the complex and in the real cases: for all $t \in \mathbb{R}$,

$$\mathcal{N}_t(S_{m,n}^{\mathbb{C}}) = \frac{1}{2}\Big(\mathcal{N}_t(S_{m,n}^{\mathbb{R}}) + \mathcal{N}_t(\tilde{S}_{m,n}^{\mathbb{R}})\Big) + \zeta_N(t), \tag{6.32}$$

where $S_{m,n}^{\mathbb{C}}$ is from the LUE, $S_{m,n}^{\mathbb{R}}$, $\tilde{S}_{m,n}^{\mathbb{R}}$ are independent matrices from the LOE and $\zeta_N(t)$ takes values in $\Big\{-\frac{3}{2}, -1, -\frac{1}{2}, 0, \frac{1}{2}, 1, \frac{3}{2}\Big\}$. See [69] for more details.

The aim is now to establish a deviation inequality for the eigenvalue counting function similar to (6.10). From (6.10), we know that for all $u \geqslant 0$,

$$\mathbb{P}\Big(|\mathcal{N}_t(S_{m,n}^{\mathbb{C}}) - n\mu_t| \geqslant u + C_1\Big) \leqslant 2\exp\Big(-\frac{u^2}{2\sigma_t^2 + u}\Big).$$

Set $C_1' = C_1 + \frac{3}{2}$ and let $u \geqslant 0$. We can then write

$$\begin{aligned}
\mathbb{P}\Big(&\mathcal{N}_t(S_{m,n}^{\mathbb{R}}) - n\mu_t \geqslant u + C_1'\Big)^2 \\
&= \mathbb{P}\Big(\mathcal{N}_t(S_{m,n}^{\mathbb{R}}) - n\mu_t \geqslant u + C_1',\ \mathcal{N}_t(\tilde{S}_{m,n}^{\mathbb{R}}) - n\mu_t \geqslant u + C_1'\Big) \\
&\leqslant \mathbb{P}\Big(\tfrac{1}{2}\Big(\mathcal{N}_t(S_{m,n}^{\mathbb{R}}) + \mathcal{N}_t(\tilde{S}_{m,n}^{\mathbb{R}})\Big) - n\mu_t \geqslant u + C_1'\Big) \\
&\leqslant \mathbb{P}\Big(\mathcal{N}_t(S_{m,n}^{\mathbb{C}}) - n\mu_t \geqslant u + C_1' - \frac{3}{2}\Big) \\
&\leqslant 2\exp\Big(-\frac{u^2}{2\sigma_t^2 + u}\Big).
\end{aligned}$$

Repeating the computations for $\mathbb{P}\Big(\mathcal{N}_t(S_{m,n}^{\mathbb{R}}) - n\mu_t \leqslant -u - C_1'\Big)$ and combining with the preceding yield

$$\mathbb{P}\Big(|\mathcal{N}_t(S_{m,n}^{\mathbb{R}}) - n\mu_t| \geqslant u + C_1'\Big) \leqslant 2\sqrt{2}\exp\Big(-\frac{u^2}{4\sigma_t^2 + 2u}\Big). \tag{6.33}$$

Note that $\sigma_t^2$ is still the variance of $\mathcal{N}_t(S_{m,n}^{\mathbb{C}})$ in the preceding formula.

What remains then to be proved is very similar to the complex case. From (6.33) and Su's bounds on the variance $\sigma_t^2$ (see (6.11) for the bulk case and (6.20) for the intermediate case), deviation inequalities for individual eigenvalues can be deduced, as was done to prove Propositions 6.6 and 6.8. It is then straightforward to derive the announced bounds on the variances for LOE matrices. The argument developed in Section 6.2 in order to extend the LUE results to large families of covariance matrices can be reproduced to reach the desired bounds on the variances of eigenvalues in the bulk and between the bulk and the edge of the spectrum for families of real covariance matrices.

## 6.4 Rate of convergence towards the Marchenko-Pastur distribution

In this whole section, we suppose that Theorem 6.2 holds for left-side intermediate eigenvalues. The bounds on $\mathbb{E}[(\lambda_j - \gamma_j^{m,n})^2]$ developed in the preceding sections lead to a bound on the rate of convergence of the empirical spectral measure $L_{m,n}$ towards the Marchenko-Pastur distribution in terms of 2-Wasserstein distance. Recall that $W_2(L_{m,n}, \mu_{m,n})$ is a random variable defined by

$$W_2(L_{m,n}, \mu_{m,n}) = \inf \left( \int_{\mathbb{R}^2} |x-y|^2 \, d\pi(x,y) \right)^{1/2},$$

where the infimum is taken over all probability measures $\pi$ on $\mathbb{R}^2$ with respective marginals $L_{m,n}$ and $\mu_{m,n}$. To achieve the expected bound, we rely on another expression of $W_2$ in terms of distribution functions, namely

$$W_2^2(L_{m,n}, \mu_{m,n}) = \int_0^1 \left( F_{m,n}^{-1}(x) - G_{m,n}^{-1}(x) \right)^2 dx, \tag{6.34}$$

where $F_{m,n}^{-1}$ (respectively $G_{m,n}^{-1}$) is the generalized inverse of the distribution function $F_{m,n}$ (respectively $G_{m,n}$) of $L_{m,n}$ (respectively $\mu_{m,n}$) (see for example [100]). On the basis of this representation, the following statement may be derived.

**Proposition 6.13.** *There exists a constant $C > 0$ depending only on $A_2$ such that for all $1 \leqslant \frac{m}{n} \leqslant A_2$,*

$$W_2^2(L_{m,n}, \mu_{m,n}) \leqslant \frac{2}{n} \sum_{j=1}^{n} (\lambda_j - \gamma_j^{m,n})^2 + \frac{C}{n^2}. \tag{6.35}$$

*Proof.* From (6.34),

$$W_2^2(L_{m,n}, \mu_{m,n}) = \int_0^1 \left( F_{m,n}^{-1}(x) - G_{m,n}^{-1}(x) \right)^2 dx.$$

Then,

$$
\begin{aligned}
W_2^2(L_{m,n},\mu_{m,n}) &= \sum_{j=1}^{n}\int_{\frac{j-1}{n}}^{\frac{j}{n}}\left(\lambda_j - G_{m,n}^{-1}(x)\right)^2 dx\\
&\leqslant \frac{2}{n}\sum_{j=1}^{n}(\lambda_j-\gamma_j^{m,n})^2 + 2\sum_{j=1}^{n}\int_{\frac{j-1}{n}}^{\frac{j}{n}}\left(\gamma_j^{m,n} - G_{m,n}^{-1}(x)\right)^2 dx.
\end{aligned}
$$

But $\gamma_j^{m,n} = G_{m,n}^{-1}\left(\frac{j}{n}\right)$ and $G_{m,n}^{-1}$ is non-decreasing. Therefore, $\left|\gamma_j^{m,n} - G_{m,n}^{-1}(x)\right| \leqslant \gamma_j^{m,n} - \gamma_{j-1}^{m,n}$ for all $x \in \left[\frac{j-1}{n},\frac{j}{n}\right]$. Consequently,

$$
W_2^2(L_{m,n},\mu_{m,n}) \leqslant \frac{2}{n}\sum_{j=1}^{n}(\lambda_j-\gamma_j^{m,n})^2 + \frac{2}{n}\sum_{j=1}^{n}(\gamma_j^{m,n}-\gamma_{j-1}^{m,n})^2. \tag{6.36}
$$

But if $j-1 \geqslant \frac{n}{2}$ (and therefore $\gamma_{j-1}^{m,n} \geqslant 0$),

$$
\begin{aligned}
\frac{1}{n} &= \int_{\gamma_{j-1}^{m,n}}^{\gamma_j^{m,n}} \frac{1}{2\pi x}\sqrt{(b_{m,n}-x)(x-a_{m,n})}\,dx\\
&\geqslant \frac{\sqrt{\gamma_{j-1}^{m,n}-a_{m,n}}}{\sqrt{2}\pi b_{m,n}}\int_{\gamma_{j-1}^{m,n}}^{\gamma_j^{m,n}}\sqrt{b_{m,n}-x}\,dx\\
&\geqslant \frac{\sqrt{\gamma_{j-1}^{m,n}-a_{m,n}}}{3\pi b_{m,n}}(b_{m,n}-\gamma_{j-1}^{m,n})^{3/2}\left(1-\left(1-\frac{\gamma_j^{m,n}-\gamma_{j-1}^{m,n}}{b_{m,n}-\gamma_{j-1}^{m,n}}\right)^{3/2}\right)\\
&\geqslant \frac{1}{3b_{m,n}\sqrt{b_{m,n}-a_{m,n}}}(b_{m,n}-\gamma_{j-1}^{m,n})^{1/2}(\gamma_j^{m,n}-\gamma_{j-1}^{m,n})\\
&\geqslant C(A_2)\left(\frac{n-j+1}{n}\right)^{1/3}(\gamma_j^{m,n}-\gamma_{j-1}^{m,n}),
\end{aligned}
$$

from (6.13). Then

$$
\gamma_j^{m,n}-\gamma_{j-1}^{m,n} \leqslant \frac{C(A_2)}{n^{2/3}(n-j+1)^{2/3}}.
$$

It may be shown that a similar bound holds if $j-1 \leqslant \frac{n}{2}$. As a summary, there exists a constant $c>0$ depending only on $A_2$ such that, for all $j \geqslant 2$,

$$
\gamma_j^{m,n}-\gamma_{j-1}^{m,n} \leqslant \frac{c}{n^{2/3}\min(j,n+1-j)^{1/3}}. \tag{6.37}
$$

This yields

$$
\sum_{j=1}^{n}(\gamma_j^{m,n}-\gamma_{j-1}^{m,n})^2 \leqslant \frac{c^2}{n^{4/3}}\sum_{j=1}^{n}\frac{1}{\min(j,n+1-j)^{2/3}} \leqslant \frac{C}{n},
$$

where $C > 0$ depends only on $A_2$. Then (6.36) becomes

$$W_2^2(L_{m,n}, \mu_{m,n}) \leqslant \frac{2}{n} \sum_{j=1}^{n} (\lambda_j - \gamma_j^{m,n})^2 + \frac{C}{n^2},$$

where $C > 0$ depends only on $A_2$, which is the claim. □

*Proof of Corollary 6.4.* Let $n \geqslant 2$. Due to Proposition 6.13,

$$\mathbb{E}\Big[W_2^2(L_{m,n}, \mu_{m,n})\Big] \leqslant \frac{2}{n} \sum_{j=1}^{n} \mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big] + \frac{C}{n^2}.$$

We then make use of the bounds on $\mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big]$ produced in the previous sections. Set $0 < \eta \leqslant \frac{1}{2}$ and $K > \kappa$ so that $K \log n \leqslant \eta n$. We first decompose

$$\begin{aligned}
\sum_{j=1}^{n} \mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big] &= \sum_{j=1}^{K \log n} \mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big] + \sum_{j=K\log n+1}^{\eta n} \mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big] \\
&\quad + \sum_{j=\eta n+1}^{(1-\eta)n-1} \mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big] + \sum_{j=(1-\eta)n}^{n-K\log n-1} \mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big] \\
&\quad + \sum_{j=n-K\log n}^{n} \mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big] \\
&= \Sigma_1 + \Sigma_2 + \Sigma_3 + \Sigma_4 + \Sigma_5.
\end{aligned}$$

The sum $\Sigma_3$ will be bounded using the bulk case (Theorem 6.1), while Theorem 6.2 will be used to handle $\Sigma_2$ and $\Sigma_4$. A crude version of Theorem 6.11 will be enough to bound $\Sigma_1$ and $\Sigma_5$. To start with thus, from Theorem 6.1,

$$\Sigma_3 \leqslant \sum_{j=\eta n+1}^{(1-\eta)n-1} C \frac{\log n}{n^2} \leqslant C \frac{\log n}{n}.$$

Secondly, from Theorem 6.2,

$$\Sigma_2 + \Sigma_4 \ \leqslant \ \frac{C}{n^{4/3}} \sum_{j=K\log n+1}^{\eta n} \frac{\log j}{j^{2/3}} \ \leqslant \ C \frac{\log n}{n}.$$

Next $\Sigma_1$ and $\Sigma_5$ have only $K \log n$ terms. If each term is bounded by $\frac{C}{n}$ where $C$ is a positive universal constant, we get that $\Sigma_1 + \Sigma_5 \leqslant \frac{2KC \log n}{n}$, which is enough to prove the desired result on $\sum_{j=1}^{n} \mathbb{E}\Big[(\lambda_j - \gamma_j^{m,n})^2\Big]$. For $n$ large enough depending only on constant $C$ in Theorem 6.11, $\frac{1}{\sqrt{n}} \geqslant \frac{(\log n)^{C \log \log n}}{n^{2/3} \min(j, n+1-j)^{1/3}}$ and Theorem 6.11 yields

$$\mathbb{P}\Big(|\lambda_j - \gamma_j^{m,n}| \geqslant \frac{1}{\sqrt{n}}\Big) \leqslant C e^{-(\log n)^{c \log\log n}}.$$

Then, by the Cauchy-Schwarz inequality,

$$\begin{aligned}\mathbb{E}\Big[(\lambda_j-\gamma_j^{m,n})^2\Big] &\leqslant \mathbb{E}\Big[(\lambda_j-\gamma_j^{m,n})^2\mathbb{1}_{|\lambda_j-\gamma_j^{m,n}|\leqslant\frac{1}{\sqrt{n}}}\Big]+\mathbb{E}\Big[(\lambda_j-\gamma_j^{m,n})^2\mathbb{1}_{|\lambda_j-\gamma_j^{m,n}|>\frac{1}{\sqrt{n}}}\Big]\\ &\leqslant \frac{1}{n}+\sqrt{\mathbb{E}\Big[|\lambda_j-\gamma_j^{m,n}|^4\Big]}\sqrt{\mathbb{P}\Big(|\lambda_j-\gamma_j^{m,n}|>\frac{1}{\sqrt{n}}\Big)}\\ &\leqslant \frac{1}{n}+\sqrt{3}Cn^2e^{-(\log n)^{c\log\log n}}.\end{aligned}$$

As $\sqrt{3}Cn^2e^{-(\log n)^{c\log\log n}}=o(\frac{1}{n})$, there exists a constant $C>0$ such that $\mathbb{E}\Big[(\lambda_j-\gamma_j^{m,n})^2\Big]\leqslant\frac{C}{n}$. Then

$$\Sigma_1+\Sigma_5\leqslant 2KC\frac{\log n}{n}.$$

As a consequence,

$$\sum_{j=1}^{n}\mathbb{E}\Big[(\lambda_j-\gamma_j^{m,n})^2\Big]\leqslant C\frac{\log n}{n}.$$

Therefore

$$\mathbb{E}\Big[W_2(L_{m,n},\mu_{m,n})^2\Big]\leqslant C\frac{\log n}{n^2},$$

where $C>0$ depends only on $A_1$ and $A_2$, which is the claim. The corollary is thus established. $\square$

# Bibliographie


[1] G. Anderson, A. Guionnet et O. Zeitouni : *An introduction to random matrices*, volume 118 de *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[2] G. Anderson et O. Zeitouni : A CLT for a band matrix model. *Probab. Theory Related Fields*, 134(2):283–338, 2006.

[3] L. Arnold : On Wigner's semicircle law for the eigenvalues of random matrices. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 19:191–198, 1971.

[4] G. Ben Arous et P. Bourgade : Extreme gaps between eigenvalues of random matrices. *arXiv :1010.1294*, 2010.

[5] G. Aubrun : A sharp small deviation inequality for the largest eigenvalue of a random matrix. *In Séminaire de Probabilités XXXVIII*, volume 1857 de *Lecture Notes in Math.*, pages 320–337. Springer, Berlin, 2005.

[6] Z. Bai, P. Krishnaiah et Y. Yin : On the limit of the largest eigenvalue of the large-dimensional sample covariance matrix. *Probab. Theory Related Fields*, 78(4):509–521, 1988.

[7] Z. Bai et J. Silverstein : CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.*, 32(1A):553–605, 2004.

[8] Z. Bai et J. Silverstein : *Spectral analysis of large dimensional random matrices.* Springer Series in Statistics. Springer, New York, second édition, 2010.

[9] Z. Bai, J. Silverstein et Y. Yin : A note on the largest eigenvalue of a large-dimensional sample covariance matrix. *J. Multivariate Anal.*, 26(2):166–168, 1988.

[10] Z. Bai et Y. Yin : Convergence to the semicircle law. *Ann. Probab.*, 16(2):863–875, 1988.

[11] Z. Bai et Y. Yin : Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix. *Ann. Probab.*, 16(4):1729–1741, 1988.

[12] Z. Bai et Y. Yin : Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix. *Ann. Probab.*, 21(3):1275–1294, 1993.

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

[14] Z. Bao, G. Pan et W. Zhou : Central limit theorem for partial linear eigenvalue statistics of wigner matrices. *arXiv :1206.0508*, 2012.

[15] S. Bobkov, F. Götze et A. Tikhomirov : On concentration of empirical measures and convergence to the semi-circle law. *J. Theoret. Probab.*, 23(3): 792–823, 2010.

[16] A. Borodin et P. Forrester : Increasing subsequences and the hard-to-soft edge transition in matrix ensembles. *J. Phys. A*, 36(12):2963–2981, 2003. Random matrix theory.

[17] P. Bourgade, L. Erdös et H-T. Yau : Universality of general $\beta$-ensembles. *arXiv :1104.2272*, 2011.

[18] J. Bourgain, V. Vu et P. Wood : On the singularity probability of discrete random matrices. *J. Funct. Anal.*, 258(2):559–603, 2010.

[19] T. Cabanal-Duvillard : Fluctuations de la loi empirique de grandes matrices aléatoires. *Ann. Inst. H. Poincaré Probab. Statist.*, 37(3):373–402, 2001.

[20] E. Candès : Compressive sampling. *In International Congress of Mathematicians. Vol. III*, pages 1433–1452. Eur. Math. Soc., Zürich, 2006.

[21] O. Costin et J. Lebowitz : Gaussian fluctuations in random matrices. *Phys. Rev. Lett.*, 75, 1995.

[22] R. C. Dalang : Une démonstration élémentaire du théorème central limite. *Elem. Math.*, 61(2):65–73, 2006.

[23] S. Dallaporta : Eigenvalue variance bounds for wigner and covariance random matrices. *RMTA*, 1(3), 2012.

[24] S. Dallaporta et V. Vu : A note on the central limit theorem for the eigenvalue counting function of Wigner matrices. *Electron. Commun. Probab.*, 16:314–322, 2011.

[25] P. Deift, T. Kriecherbauer, K. McLaughlin, S. Venakides et X. Zhou : Strong asymptotics of orthogonal polynomials with respect to exponential weights. *Comm. Pure Appl. Math.*, 52(12):1491–1552, 1999.

[26] R. M. Dudley : *Real analysis and probability*, volume 74 de *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original.

[27] I. Dumitriu et A. Edelman : Matrix models for beta ensembles. *J. Math. Phys.*, 43(11):5830–5847, 2002.

[28] R. Durrett : *Probability : theory and examples.* Duxbury Press, Belmont, CA, second édition, 1996.

[29] H. Döring et P. Eichelsbacher : Moderate deviations for the eigenvalue counting function of wigner matrices. *arXiv :1104.0221*, 2011.

[30] A. Edelman : Eigenvalues and condition numbers of random matrices. *SIAM J. Matrix Anal. Appl.*, 9(4):543–560, 1988.

[31] A. Edelman et N. Rao : Random matrix theory. *Acta Numer.*, 14:233–297, 2005.

[32] L. Erdös : Universality of Wigner random matrices : a survey of recent results. *Uspekhi Mat. Nauk*, 66(3(399)):67–198, 2011.

[33] L. Erdős, S. Péché, J. Ramírez, B. Schlein et H-T. Yau : Bulk universality for Wigner matrices. *Comm. Pure Appl. Math.*, 63(7):895–925, 2010.

[34] L. Erdős, J. Ramírez, B. Schlein, T. Tao, V. Vu et H-T. Yau : Bulk universality for Wigner Hermitian matrices with subexponential decay. *Math. Res. Lett.*, 17(4):667–674, 2010.

[35] L. Erdös, H.-T. Yau et J. Yin : Rigidity of eigenvalues of generalized wigner matrices. *arXiv :1007.4652*, 2010.

[36] O. Feldheim et S. Sodin : A universality result for the smallest eigenvalues of certain sample covariance matrices. *Geom. Funct. Anal.*, 20(1):88–123, 2010.

[37] P. Forrester : *Log-gases and random matrices*, volume 34 de *London Mathematical Society Monographs Series.* Princeton University Press, Princeton, NJ, 2010.

[38] P. Forrester et E. Rains : Interrelationships between orthogonal, unitary and symplectic matrix ensembles. *In Random matrix models and their applications*, volume 40 de *Math. Sci. Res. Inst. Publ.*, pages 171–207. Cambridge Univ. Press, Cambridge, 2001.

[39] H. Goldstine et J. von Neumann : Numerical inverting of matrices of high order. *Bull. Amer. Math. Soc.*, 53:1021–1099, 1947.

[40] H. Goldstine et J. von Neumann : Numerical inverting of matrices of high order. II. *Proc. Amer. Math. Soc.*, 2:188–202, 1951.

[41] F. Götze et A. Tikhomirov : Rate of convergence to the semi-circular law. *Probab. Theory Related Fields*, 127(2):228–276, 2003.

[42] F. Götze et A. Tikhomirov : Rate of convergence in probability to the Marchenko-Pastur law. *Bernoulli*, 10(3):503–548, 2004.

[43] F. Götze et A. Tikhomirov : The rate of convergence for spectra of GUE and LUE matrix ensembles. *Cent. Eur. J. Math.*, 3(4):666–704, 2005.

[44] F. Götze et A. Tikhomirov : On the rate of convergence to the Marchenko-Pastur distribution. *arXiv :1110.1284*, 2011.

[45] F. Götze et A. Tikhomirov : On the rate of convergence to the semicircular law. *arXiv :1109.0611*, 2011.

[46] A. Guionnet : Grandes matrices aléatoires et théorèmes d'universalité (d'après Erdős, Schlein, Tao, Vu et Yau). *Séminaire Bourbaki. Vol. 2009/2010. Exposés 1012–1026*, (339):Exp. No. 1019, viii, 203–237, 2011.

[47] A. Guionnet et O. Zeitouni : Concentration of the spectral measure for large matrices. *Electron. Comm. Probab.*, 5:119–136, 2000.

[48] A. Guntuboyina et H. Leeb : Concentration of the spectral measure of large Wishart matrices with dependent entries. *Electron. Commun. Probab.*, 14:334–342, 2009.

[49] J. Gustavsson : Gaussian fluctuations of eigenvalues in the GUE. *Ann. Inst. H. Poincaré Probab. Statist.*, 41(2):151–178, 2005.

[50] J. Ben Hough, M. Krishnapur, Y. Peres et B. Virág : Determinantal processes and independence. *Probab. Surv.*, 3:206–229, 2006.

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

[52] K. Johansson : Shape fluctuations and random matrices. *Comm. Math. Phys.*, 209(2):437–476, 2000.

[53] K. Johansson : Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. *Comm. Math. Phys.*, 215(3):683–705, 2001.

[54] I. Johnstone : On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.*, 29(2):295–327, 2001.

[55] D. Jonsson : Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.*, 12(1):1–38, 1982.

[56] 0. Khorunzhiy : High moments of large Wigner random matrices and asymptotic properties of the spectral norm. *Random Oper. Stoch. Equ.*, 20(1):25–68, 2012.

[57] M. Ledoux : *The concentration of measure phenomenon*, volume 89 de *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 2001.

[58] M. Ledoux : Deviation inequalities on largest eigenvalues. *In Geometric aspects of functional analysis*, volume 1910 de *Lecture Notes in Math.*, pages 167–219. Springer, Berlin, 2007.

[59] M. Ledoux et B. Rider : Small deviations for beta ensembles. *Electron. J. Probab.*, 15(41):1319–1343, 2010.

[60] J.O. Lee et J. Yin : A necessary and sufficient condition for edge universality of wigner matrices. *arXiv :1206.2251*, 2012.

[61] J. W. Lindeberg : Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Math. Z.*, 15(1):211–225, 1922.

[62] A. Litvak, A. Pajor, M. Rudelson et N. Tomczak-Jaegermann : Smallest singular value of random matrices and geometry of random polytopes. *Adv. Math.*, 195(2):491–523, 2005.

[63] A. Lytova et L. Pastur : Central limit theorem for linear eigenvalue statistics of the Wigner and the sample covariance random matrices. *Metrika*, 69(2-3):153–172, 2009.

[64] V. A. Marčenko et L. A. Pastur : Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, 72 (114), 1967.

[65] E. Meckes et M. Meckes : Concentration and convergence rates for spectral measure of random matrices. *to appear in Probab. Theory Relat. Fields,arXiv :1109.5997*, 2011.

[66] M. Mehta : On the statistical properties of the level-spacings in nuclear spectra. *Nuclear Phys.*, 18:395–419, 1960.

[67] M. Mehta : *Random matrices*, volume 142 de *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, third édition, 2004.

[68] M. Mehta et M. Gaudin : On the density of eigenvalues of a random matrix. *Nuclear Phys.*, 18:420–427, 1960.

[69] S. O'Rourke : Gaussian fluctuations of eigenvalues in Wigner random matrices. *J. Stat. Phys.*, 138(6):1045–1066, 2010.

[70] L. Pastur : The spectrum of random matrices. *Teoret. Mat. Fiz.*, 10(1):102–112, 1972.

[71] L. Pastur : Limiting laws of linear eigenvalue statistics for Hermitian matrix models. *J. Math. Phys.*, 47(10):103303, 22, 2006.

[72] L. Pastur et M. Shcherbina : *Eigenvalue distribution of large random matrices*, volume 171 de *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2011.

[73] S. Péché : Universality results for the largest eigenvalues of some sample covariance matrix ensembles. *Probab. Theory Related Fields*, 143(3-4):481–516, 2009.

[74] S. Péché et A. Soshnikov : Wigner random matrices with non-symmetrically distributed entries. *J. Stat. Phys.*, 129(5-6):857–884, 2007.

[75] N. Pillai et J. Yin : Universality of covariance matrices. *arXiv :1110.2501*, 2011.

[76] M. Rudelson et R. Vershynin : The Littlewood-Offord problem and invertibility of random matrices. *Adv. Math.*, 218(2):600–633, 2008.

[77] M. Rudelson et R. Vershynin : Non-asymptotic theory of random matrices : extreme singular values. *In Proceedings of the International Congress of Mathematicians. Volume III*, pages 1576–1602, New Delhi, 2010. Hindustan Book Agency.

[78] A. Ruzmaikina : Universality of the edge distribution of eigenvalues of Wigner random matrices with polynomially decaying distributions of entries. *Comm. Math. Phys.*, 261(2):277–296, 2006.

[79] Y. Sinai et A. Soshnikov : Central limit theorem for traces of large random symmetric matrices with independent matrix elements. *Bol. Soc. Brasil. Mat. (N.S.)*, 29(1):1–24, 1998.

[80] A. Soshnikov : Universality at the edge of the spectrum in Wigner random matrices. *Comm. Math. Phys.*, 207(3):697–733, 1999.

[81] A. Soshnikov : Gaussian fluctuation for the number of particles in Airy, Bessel, sine, and other determinantal random point fields. *J. Statist. Phys.*, 100(3-4):491–522, 2000.

[82] A. Soshnikov : A note on universality of the distribution of the largest eigenvalues in certain sample covariance matrices. *J. Statist. Phys.*, 108(5-6):1033–1056, 2002.

[83] Z. Su : Gaussian fluctuations in complex sample covariance matrices. *Electron. J. Probab.*, 11(48):1284–1320, 2006.

[84] T. Tao : The asymptotic distribution of a single eigenvalue gap of a wigner matrix. *arXiv :1203.1605*, 2012.

[85] T. Tao et V. Vu : Random matrices : the distribution of the smallest singular values. *Geom. Funct. Anal.*, 20(1):206–297, 2010.

[86] T. Tao et V. Vu : Random matrices : universality of local eigenvalue statistics up to the edge. *Comm. Math. Phys.*, 298(2):549–572, 2010.

[87] T. Tao et V. Vu : Random matrices : universality of local eigenvalues statistics. *Acta Math.*, 206(1):127–204, 2011.

[88] T. Tao et V. Vu : Random covariance matrices : universality of local statistics of eigenvalues. *Annals of Probability*, 40(3):1285–1315, 2012.

[89] T. Tao et V. Vu : Random matrices : sharp concentration of eigenvalues. *arXiv :1201.4789*, 2012.

[90] T. Tao et V. Vu : Random matrices : The universality phenomenon for Wigner ensembles. *ArXiv :1202.0068*, 2012.

[91] C. Tracy et H. Widom : Level-spacing distributions and the Airy kernel. *Comm. Math. Phys.*, 159(1):151–174, 1994.

[92] C. Tracy et H. Widom : Level spacing distributions and the Bessel kernel. *Comm. Math. Phys.*, 161(2):289–309, 1994.

[93] C. Tracy et H. Widom : On orthogonal and symplectic matrix ensembles. *Comm. Math. Phys.*, 177(3):727–754, 1996.

[94] C. Tracy et H. Widom : Distribution functions for largest eigenvalues and their applications. *In Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002)*, pages 587–596, Beijing, 2002. Higher Ed. Press.

[95] A. van der Vaart : *Asymptotic statistics*, volume 3 de *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, 1998.

[96] M. Vanlessen : Strong asymptotics of Laguerre-type orthogonal polynomials and applications in random matrix theory. *Constr. Approx.*, 25(2):125–175, 2007.

[97] R. Vershynin : Invertibility of symmetric random matrices. *arXiv :1102.0300*, 2011.

[98] R. Vershynin : How close is the sample covariance matrix to the actual covariance matrix ? *J. Theoret. Probab.*, 25(3):655–686, 2012.

[99] R. Vershynin : *Introduction to the non-asymptotic analysis of random matrices*, chapitre 5 of Compressed sensing, theory and applications. Y. Eldar and G. Kutyniok. Cambridge University Press, 2012.

[100] C. Villani : *Topics in optimal transportation*, volume 58 de *Graduate Studies in Mathematics*. American Mathematical Society, 2003.

[101] K. Wang : Random covariance matrices : universality of local statistics of eigenvalues up to the edge. *arXiv :1104.4832*, 2011.

[102] E. Wigner : On the distribution of the roots of certain symmetric matrices. *Ann. of Math. (2)*, 67:325–327, 1958.

Cette thèse est consacrée à l'étude quantitative de la fonction de comptage et des valeurs propres de matrices aléatoires. Initialement introduites dans le cadre de la physique statistique, ces matrices servent de modèles pour des opérateurs de dimension infinie. Leurs propriétés asymptotiques ont donc été particulièrement étudiées. Devenues populaires grâce au phénomène d'universalité, i.e. le fait que ces propriétés asymptotiques ne dépendent pas de la loi des coefficients de la matrice, leur étude a intéressé de nombreux domaines pour lesquels les propriétés à taille de matrice fixée sont plus exploitables que les propriétés asymptotiques. Nous nous sommes intéressés à ce pan de l'étude des matrices aléatoires par le biais de la fonction de comptage des valeurs propres, c'est-à-dire du nombre de valeurs propres d'une matrice présentes dans un intervalle $I$.

Après avoir introduit les modèles de matrices aléatoires que nous étudions dans cette thèse, les matrices de Wigner et de covariance, nous présentons les principaux résultats asymptotiques en lien avec la fonction de comptage et plus globalement avec les valeurs propres de ces matrices. Les outils permettant d'établir ces résultats d'universalité sont ensuite détaillés. L'accent est mis sur ceux qui peuvent être utilisés dans le cadre de l'étude quantitative, notamment les résultats récents de Erdös *et al.* d'une part et de Tao et Vu d'autre part, qui ont permis une avancée spectaculaire des études asymptotiques et non asymptotiques. Nous discutons ensuite les enjeux de l'étude non asymptotique et présentons les travaux effectués durant cette thèse.

Dans une première étude, nous établissons un théorème central limite pour la fonction de comptage des valeurs propres de matrices de Wigner et de covariance et nous obtenons des estimées quantitatives sur l'espérance et la variance de cette fonction de comptage. Ces résultats se basent sur les résultats précédemment établis par Gustavsson et Su dans le cas de matrices gaussiennes et sont étendus à des familles plus générales de matrices de Wigner et de covariance par le biais de travaux récents de Erdös, Yau et Yin, Pillai et Yin et de Tao et Vu.

Dans une seconde étude, nous établissons des bornes quantitatives sur la variance des valeurs propres de matrices de Wigner. En s'appuyant sur les propriétés de la fonction de comptage, nous montrons d'abord une inégalité de déviation pour les valeurs propres individuelles à l'intérieur du spectre dans le cas de matrices gaussiennes et nous en déduisons les bornes dans ce cas. Afin d'étendre ces bornes au cas des matrices de Wigner plus générales, nous combinons de nouveau les outils de Erdös, Yau et Yin et de Tao et Vu. Nous établissons également des résultats analogues pour les matrices de covariance, en utilisant la même démarche.

---

This thesis is concerned by the quantitative study of the counting function and the eigenvalues of random matrices. These matrices were first introduced for use in statistical physics and are models for infinite dimensional operators. Therefore a thorough study of their asymptotic properties was initiated. They became popular due to the universality phenomenon, i.e. the fact that these asymptotic properties do not depend on the distribution of the entries. Lots of fields became interested in random matrices. Several of these fields need quantitative results rather than asymptotic ones. In this thesis, we are interested in the non asymptotic study of random matrices, through the study of the eigenvalue counting function, which is the number of eigenvalues which are in an interval $I$.

We first introduce models of random matrices which are considered in this thesis, Wigner matrices and covariance matrices. The main asymptotic results regarding the counting function and more generally regarding the eigenvalues of these matrices are presented. Tools to reach universality results are then described. We insist on those which are useful to obtain non asymptotic results as well, such as recent results by Erdös et al. on the one hand and by Tao and Vu on the other hand. We then discuss non asymptotic issues and present the work we did during this thesis.

In a first part, a Central Limit Theorem for the counting function is established for Wigner and covariance matrices. Quantitative asymptotics for the mean and the variance of the counting function are obtained. These results are based on previous results by Gustavsson and Su for Gaussian matrices and are extended to large families of Wigner and covariance matrices by recent results of Erdös, Yau and Yin, Pillai and Yin and of Tao and Vu.

In a second part, quantitative bounds on the variance of individual eigenvalues of Wigner matrices are established. Relying on properties of the counting function, a deviation inequality for eigenvalues in the bulk is proved for Gaussian matrices. The bounds on the variance in this case are then easily derived. These bounds are again extended to families of Wigner matrices by a combination of tools of Erdös, Yau and Yin and of Tao and Vu. Following the same scheme, analogous results are then established for covariance matrices.