Forte et fausse libertés asymptotiques de grandes matrices aléatoires

Camille Male

École Normale Supérieure de Lyon

Thèse de doctorat

Discipline : Mathématiques

présentée par

Camille Male

Forte et fausse libertés asymptotiques de grandes matrices aléatoires

---

dirigée par Alice Guionnet

après avis de M. Djalil Chafaï, M. Gilles Pisier et de M. Dan-Virgil Voiculescu,

dont la soutenance est prévue le 5 décembre 2011 devant le jury composé de :

M. Philippe Biane	examinateur
M. Djalil Chafaï	rapporteur
Mme Catherine Donati-Martin	examinatrice
M Damien Gaboriau	examinateur
Mlle Alice Guionnet	directrice
M. Michel Ledoux	examinateur

Unité de mathématiques pûres et appliquées
École Normale Supérieure de Lyon
46, allée d'Italie
69364 Lyon Cedex 07
France

École doctorale Informatique et Mathématiques
Université Lyon1 - Bât. Braconnier
43, Bvd. du 11 novembre 1918
69622 Villeurbanne Cedex
France

Le contenu de ce mémoire

Cette thèse s'inscrit dans la théorie des matrices aléatoires, à l'intersection avec la théorie des probabilités libres et des algèbres d'opérateurs. Elle s'insère dans une démarche générale qui a fait ses preuves ces dernières décennies : importer les techniques et les concepts de la théorie des probabilités non commutatives pour l'étude du spectre de grandes matrices aléatoires. La notion de liberté, qui est un analogue non commutatif de la notion d'indépendance statistique, joue le rôle central dans cette démarche.

On s'intéresse ici à des généralisations du théorème de liberté asymptotique de Voiculescu. Dans la Partie I, nous montrons un résultat de liberté asymptotique forte pour des matrices gaussiennes, unitaires aléatoires et déterministes. Dans la Partie II, nous introduisons la notion de fausse liberté asymptotique pour des matrices déterministes et certaines matrices hermitiennes à entrées sous diagonales indépendantes, interpolant les modèles de matrices de Wigner et de Lévy.

L'introduction de ce mémoire débute avec un court historique de la théorie des matrices aléatoires, en lien avec le contenu de cette thèse (Section 0.1). Dans un second temps, nous présentons les modèles étudiés dans une courte zoologie (Section 0.2). Dans la Section 0.3, nous rappelons des notions élémentaires de probabilités libres. Cette dernière partie est destinée principalement aux lecteurs probabilistes n'ayant pas de notions de probabilités libres (pour des ouvrages sur le sujet, voir [AGZ10, Gui09, NS06]). Les Sections 0.4 et 0.5 constituent une présentation des travaux de thèse.

Les Chapitres 1 et 2, formant la Partie I, sont sur le thème de la forte liberté asymptotique de grandes matrices aléatoires. Les Chapitre 3 et 4 constituent la Partie II sur le thème de la fausse liberté asymptotique.

Le Chapitre 1 est extrait d'une publication dans la revue *Probability theory and related fields*. Le Chapitre 2 est extrait d'un article en prépublication sur Arxiv.org en collaboration avec Benoît Collins. Les Chapitres 3 et 4 sont des comptes rendus de travaux en cours qui n'ont pas encore été prépubliés, le Chapitre 4 étant issu d'un travail en collaboration avec Florent Benaych-Georges et Alice Guionnet.

Mots-clefs : Théorie des matrices aléatoires, probabilités libres, liberté asymptotique, $\mathcal{C}^*$-algèbres, théorie spectrale des graphes

Keywords : Random matrix theory, free probability, asymptotic freeness, $\mathcal{C}^*$-algebra, spectral graph theory

Table des matières

Introduction

0.1 Un historique de la théorie des matrices aléatoires

Les matrices aléatoires apparaissent dans différents domaines des mathématiques et de la physique comme les théories des probabilités et des statistiques, la mécanique quantique, la théorie des opérateurs, la combinatoire, la théorie quantique de l'information, etc.

Elles ont d'abord été introduite dans les années 1930 par le statisticien Wishart [Wis28] dans le but d'étudier le spectre de matrices de covariance empiriques. Il a fallu attendre la fin des années cinquante pour que le sujet gagne en importance, lorsque le physicien Wigner introduisit le concept de distribution statistique des noyaux atomiques. Il remarqua [Wig58] que pour des noyaux lourds, la distribution d'énergie moyenne est très bien approximée par la distribution des valeurs propres de matrices aléatoires hermitiennes. Les fondations mathématiques de la théorie des matrices aléatoires furent établies par Dyson (voir le livre de Mehta [Meh04]).

Des progrès significatifs ont eu lieu ces dernières décennies dans le domaine des statistiques multivariées grâce à la théorie des matrices aléatoires, et cette intéraction est toujours très dynamique aujourd'hui. Les questions sous-jacentes sont naturellement liées aux rapides développements des technologies modernes. En effet, nous sommes aujourd'hui confrontés à l'analyse de données de très grandes dimensions, par exemple dans les domaines des télécommunications [TV04], de la finance [PBL05] ou de la génétique [VO05].

L'intérêt des probabilistes et statisticiens pour les matrices aléatoires a pris un nouveau souffle après 1967, lorsque Marchenko et Pastur [MP67] ont donné une forme simple pour la distribution spectrale asymptotique d'une matrice de Wishart. Leur résultat a été généralisé dans bien des directions. Des distributions spectrales limites de grandes matrices aléatoires ont été déterminées pour un grand nombre de modèles, voir les travaux de Bai, Yin et Krishnaiah [BYK86], Grenander et Silverstein [GS77], Jonsson [Jon82], Wachter [Wac78], Yin [Yin86], Yin et Krishnaiah [YK83].

Estimer la position des valeurs propres extrêmes d'une grande matrice aléatoire est un problème important en analyse en composantes principales. Le phénomène de convergence de ces valeurs propres vers le bord du support de la distribution spectrale limite pour une matrice de Wigner ou de Wishart a été démontré à partir des années 1980. Notons les contributions de Geman [Gem80], Juhász [Juh81], Füredi et Komlós [FK81], Jonsson [Jon85], Silverstein [Sil89, Sil85], Bai et Yin [BY88], Yin, Bai et Krishnaiah [YBK88], et Bai, Silverstein et Yin [BSY88].

Les phénomènes "pas de valeurs propres en dehors du spectre limite" et de "séparation exacte des valeurs propres" ont été démontrés par Bai et Silverstein [BS98, BS99]. Ces deux résultats sont d'une très grande importance puisqu' ils donnent une information très précise sur le comportement des valeurs propres d'une large classe de grandes matrices aléatoires (voir également [PS09]).

Plus récemment, des méthodes dites "d'équivalents déterministes" ont été développées pour une grande variété de modèles de matrices : celles-ci sont des combinaisons algébriques de matrices de type bruit et de matrices déterministes. Un équivalent déterministe donne une prédiction pour le spectre de grandes matrices aléatoires, sans hypothèses asymptotiques sur les matrices déterministes. Voir les travaux de Hachem, Loubaton, et Najim [HLN07], Couillet, Debbah, et Silverstein [CDS10] et Couillet, Hoydis, et Debbah [CHD10].

Paralèlement, les matrices aléatoires ont pris de l'importance dans le domaine des espaces d'opérateurs. Voiculescu [Voi85] a introduit la théorie des probabilités libres, qui est une théorie des probabilités dans un cadre non commutatif, pour étudier les algèbres de von Neumann des groupes libres. En particulier il a défini la notion d'entropie libre dans le but de répondre à la question de l'isomorphisme entre les facteurs libres. Bien que cette question ne soit pas encore résolue, l'approche des probabilités libres a permis de grands progrès dans la compréhension des algèbres de von Neumann [Voi96].

Dans les années 1990, Voiculescu [Voi95b] établit un lien entre les propriétés spectrales asymptotiques de grandes matrices aléatoires et les algèbres de von Neumann des groupes libres (voir [Voi95a]) en démontrant un premier théorème dit de " liberté asymptotique". Cette connexion a été renforcée dans diverses directions, par exemple avec les travaux de Ben Arous et Guionnet [BAG97] concernant un principe de grande déviations pour les matrices de Wigner, en lien avec l'entropie libre. Les travaux de Haagerup et Thorbjørnsen [HT05] concernent la convergence du rayon spectral de grandes matrices hermitiennes. De leur résultat principales, ils déduisent des propriétés de la $\mathcal{C}^*$-algèbre réduite du groupe libre et, par ailleurs, répondent à une question de la théorie des espaces d'opérateurs posée par Pisier dans son ouvrage [Pis03, Chapter 20].

Suivre la méthodologie de Voiculescu autour du théorème de liberté asymptotique défend un double intérêt :
- utiliser les matrices aléatoires pour décrire des espaces opérateurs,
- utiliser les outils et les concepts de la théorie des probabilités libres pour tirer des informations sur le spectre de grandes matrices aléatoires.

Les travaux présentés dans ce mémoire s'inscrivent dans la seconde dérmarche, avec pour but de répondre à des questions reliées aux problèmes de statistiques énoncés plus haut dans un grand degré de généralité. Cette approche n'est pas rare dans la littérature contemporaine, les probabilités libres étant aujourd'hui couramment utilisées pour répondre à des problèmes de télécommunication [SSH05, RD08, RFOBS08, WZCM09].

0.2 Une courte zoologie de matrices aléatoires

Une matrice aléatoire est une matrice dont les entrées sont des variables aléatoires. Dans le cas des matrices à coefficients complexes, il s'agit donc d'une application mesurable $\Omega \to \mathrm{M}_{N,N'}(\mathbb{C})$, où $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace de probabilité et $\mathrm{M}_{N,N'}(\mathbb{C})$ est l'ensemble des matrices N par N' à coefficients dans $\mathbb{C}$.

Nous présentons des exemples de matrices aléatoires étudiés. Nous partons de la matrice la plus simple à introduire, celle dont les entrées sont des variables aléatoires gaussiennes indépendantes, et construisons à partir de celle ci les trois ensembles de matrices les plus populaires. Ensuite, nous présentons d'autres modèles qui sont des généralisations de ces ensembles.

Cette zoologie aboutit au modèle générique que nous étudions et sur une brève présentation de nos résultats.

0.2.1 Ensemble gaussien unitaire, matrices de Haar sur le groupe unitaire et matrices de Wishart

Construction des ensembles

Considérons $M_{N,N'}$ la matrice aléatoire de taille N par N' dont les entrées sont indépendantes, identiquement distribuées selon la loi gaussienne complexe $\mathcal{N}_{\mathbb{C}}(0, \frac{1}{\sqrt{N'}})$. Dans le cas des matrices carrées, nous noterons $M_N = M_{N,N}$. Nous pouvons construire à partir de cette matrice d'autres matrices aléatoires en utilisant des manipulations simples d'algèbre linéaire.

Symétrisation de M_N : l'ensemble gaussien unitaire (GUE)

Posons $X_N = \frac{M_N + M_N^*}{2}$, où M_N^* est la transposée complexe de M_N. Cette matrice est hermitienne, ainsi ses valeurs propres constituent-elles un processus ponctuel sur la droite réelle.

Les entrées sous diagonales de X_N sont gaussiennes, indépendantes et centrées. Elles sont réelles et de variance $\frac{1}{\sqrt{N}}$ sur la diagonale, alors que les entrées strictement sous diagonales sont des variables complexes $\mathcal{N}_{\mathbb{C}}(0, \frac{1}{\sqrt{N}})$. Ainsi, la distribution de la matrice X_N est proportionnelle à la mesure

$$\exp\Big(-\frac{N}{2}\mathrm{Tr}(X^2)\Big) \prod_{i \leqslant j} \mathrm{d}\ \mathrm{Re}X_{i,j} \prod_{i<j} \mathrm{d}\ \mathrm{Im}X_{i,j}. \tag{1}$$

La distribution de X_N est donc la loi gaussienne standard dans l'espace hilbertien des matrices hermitiennes de taille N par N, muni du produit scalaire $(A, B) \mapsto N \times \mathrm{Tr}(AB)$. La distribution de X_N est donc invariante par conjugaison par une matrice unitaire. C'est pour cette raison que la loi de cette matrice est appellée ensemble gaussien unitaire.

Décomposition de X_N : matrices de Haar sur $\mathcal{U}_N$

D'après le théorème spectral, pour presque tout ω dans l'espace de probabilité sous-jacent pour M_N, il existe une matrice unitaire $U_N(\omega)$ et une matrice diagonale $\Delta_N(\omega)$ telle que l'on ait

$$X_N(\omega) = U_N(\omega)\Delta_N(\omega)U_N(\omega)^*. \tag{2}$$

Une paramètrisation adéquate permet de rendre les applications $\omega \to U_N(\omega)$ et $\omega \to \Delta_N(\omega)$ mesurables (U_N et Δ_N sont donc bien des matrices aléatoires) de sorte que la distribution de U_N est la mesure de Haar sur le groupe unitaire $\mathcal{U}_N$. Il s'agit de l'unique mesure de probabilité sur le groupe métrique compact $\mathcal{U}_N$ invariante par multiplication à gauche et à droite par une matrice unitaire. De plus, les matrices U_N et Δ_N s'avèrent être indépendantes.

Matrices de covariances empiriques : ensemble de Wishart

Soit Σ_N une matrice déterministe hermitienne, définie positive de taille N par N. Notons $\Sigma_N^{\frac{1}{2}}$ la matrice hermitienne positive racine carrée de Σ_N. Posons

$$W_{N,N'} = \Big(\Sigma_N^{\frac{1}{2}} \times M_{N,N'}\Big)\Big(\Sigma_N^{\frac{1}{2}} \times M_{N,N'}\Big)^*,$$

qui est une matrice hermitienne de taille N par N. La matrice $W_{N,N'}$ n'est autre que la matrice de covariance empirique d'un échantillon de N' vecteurs gaussiens de taille N, centrés et de matrice de covariance Σ_N.

La matrice $W_{N,N'}$ est appellée matrice de Wishart non blanche. Dans le cas où Σ_N est la matrice identité, on parle de matrice de Wishart blanche.

Spectre limite d'une grande matrice aléatoire

Etant donnée une matrice aléatoire H_N de taille N par N, nous nous intéressons à la distribution jointe de ses valeurs propres. Les modèles de références présentés plus haut ont la particularité d'avoir une distribution des valeurs propres explicite. Mais cette propriété est très rare parmi les ensembles de matrices aléatoires considérés usuellement.

C'est pour cette raison que nous étudions le plus souvent la distribution asymptotique d'une matrice aléatoire lorsque sa taille tend vers l'infini : implicitement, nous considérons une suite $(H_N)_{N\geqslant 1}$, où pour tout $N \geqslant 1$ la matrice H_N est de taille N par N. Afin de coder le spectre limite éventuel de H_N, nous introduisons la distribution empirique de ses valeurs propres

$$\mathcal{L}_{H_N} = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i^{(N)}}, \tag{3}$$

où $\lambda_1^{(N)}, \dots, \lambda_N^{(N)}$ désignent les valeurs propres de H_N et δ_λ est la masse de Dirac en λ.

Lorsque l'on parle de l'étude du spectre limite d'une matrice aléatoire H_N, nous voulons dire la convergence (en un sens qui dépend du contexte) de la mesure aléatoire $\mathcal{L}_{H_N}$ vers une mesure de probabilité μ.

Le résultat le plus populaire en théorie des matrices aléatoires est le théorème de Wigner [Wig58], qui affirme que la distribution empirique des valeurs propres d'une matrice X_N du GUE converge presque sûrement et en espérance en topologie faible-$*$ vers la mesure semicirculaire de rayon 2. En d'autres termes, pour toute fonction $f : \mathbb{R} \to \mathbb{C}$ continue bornée ou polynomiale, on a

$$\mathcal{L}_{H_N}(f) = \frac{1}{N} \sum_{i=1}^{N} f(\lambda_i^{(N)}) \underset{N\to\infty}{\longrightarrow} \int_{-2}^{2} f(t) \frac{1}{\sqrt{2\pi}} \sqrt{4-t^2} \mathrm{d}(t) \tag{4}$$

presque sûrement et en espérance. La distribution empirique des valeurs propres d'une matrice de Haar sur le groupe unitaire converge presque sûrement et en espérance vers la mesure uniforme sur le cercle de rayon 1 dans le plan complexe (voir [AGZ10]).

La question de la convergence de la distribution empirique des valeurs d'une matrice de covariance $W_{N,N'}$ est plus délicate. D'abord, il faut préciser un régime pour les croissances relatives des nombres N et N'. Faisons l'hypothèse que N' est fonction de N de sorte que le rapport $\frac{N}{N'}$ converge vers une constante $c > 0$.

Ensuite, il convient de préciser une asymptotique pour la matrice déterministe Σ_N. Supposons que la distribution empirique des valeurs propres de Σ_N converge vers une mesure de probabilité ν. Alors, celle de la matrice de Wishart $W_{N,N'}$ converge vers une mesure de probabilité qui ne dépend que de ν et de la constante c. Il s'agit du théorème de Marchenko-Pastur [MP67].

Matrices à entrées réelles et quaternioniques

Les matrices X_N, U_N et $W_{N,N'}$ ont été construites à partir de la matrice $M_{N,N'}$ dont les entrées sont des variables aléatoires gaussiennes complexes. Si nous remplaçons, dans la construction précédente, les variables complexes par des variables réelles (respectivement quaternioniques), nous obtenons

1. pour la distribution de X_N, l'ensemble gaussien orthogonal (respectivement symplectique),
2. pour la distribution de U_N, la mesure de Haar sur le groupe orthogonal (respectivement symplectique),
3. pour la distribution de $W_{N,N'}$, le modèle de Wishart réel (respectivement symplectique).

Ces modèles sont très proches dans leur structure des modèles à entrées complexes. En particuliers, les résultats de convergence des distributions empiriques de valeurs propres restent inchangés.

0.2.2 Matrices symétriques ou hermitiennes à entrées sous diagonales indépendantes

Matrices de Wigner

Le modèle de matrice de Wigner généralise le modèle du GUE : la distribution gaussienne dans la définition d'une matrice du GUE est remplacée par une distribution arbitraire (suffisamment régulière à l'infini). Plus précisément, une matrice N par N aléatoire $X_N = (X_{i,j}^{(N)})_{i,j=1,\dots,N}$ est dite de Wigner dès lors que

1. presque sûrement X_N est hermitienne,
2. les variables aléatoires $(\sqrt{N}X_{i,j})_{1\leqslant i\leqslant j\leqslant N}$ sont indépendantes, de variance finie,
3. les variables aléatoires $(\sqrt{N}X_{i,j})_{1\leqslant i<j\leqslant N}$ sont identiquement distribuées selon une loi qui ne dépend pas de N.
4. idem pour les variables aléatoires $(\sqrt{N}X_{i,i})_{i=1,\dots,N}$, avec éventuellement une loi commune différente de celle des entrées strictement sous diagonales.

La pertinence d'un tel modèle à ses raisons par la notion d'universalité : de nombreuses statistiques asymptotiques d'une grande matrice de Wigner ne dépendent pas de la loi de ses entrées, et donc sont les mêmes que pour une matrice du GUE. Par exemple, la distribution empirique des valeurs propres d'une matrice de Wigner converge presque sûrement et en espérance vers la loi semicirculaire.

Matrices de Lévy

Le modèle des matrices de Lévy est une variante de celui des matrices de Wigner où les entrées sous diagonales sont indépendantes, mais leur distribution est à queues lourdes. Plus précisément, une matrice aléatoire symétrique $X_N = (X_{i,j}^{(N)})_{i,j=1,\dots,N}$ est dite de Lévy de paramètre α lorsque pour tout $i, j = 1, \dots, N$,

$$X_N(i,j) = \frac{x_{i,j}}{\sigma_N},$$

où les variables aléatoires $(x_{i,j})_{1\leqslant i\leqslant j\leqslant N}$ sont indépendantes, identiquement distribuées selon une loi ne dépendant pas de N et appartenant au domaine d'attraction d'une loi α stable pour un nombre α dans $]0,2[$. En d'autres termes, il existe un fonction $L : \mathbb{R} \to \mathbb{R}$ à variations lentes, telle que

$$\mathbb{P}\Big(|x_{1,1}| \geqslant u\Big) = \frac{L(u)}{u^\alpha}, \forall u \in \mathbb{R}.$$

De plus, nous avons noté la constante normalisatrice

$$\sigma_N = \inf\left\{ u \in \mathbb{R}^+ \,\middle|\, \mathbb{P}\Big(|x_{1,1}| \geqslant u\Big) \leqslant \frac{1}{N} \right\}.$$

D'après un résultat de Ben Arous et Guionnet [BAG08], la distribution empirique des valeurs propres d'une matrice de Lévy de paramètre α converge en topologie faible-$*$ vers une mesure de probabilité μ_α. Cette mesure ne dépend que du nombre α et est à support non borné. En outre, nous n'avons pas d'expression explicite pour μ_α mais seulement une équation caractérisant sa transformée de Stieltjes.

Matrices de Wigner lourde

Une matrice de Lévy se distingue d'une matrice de Wigner car ses entrées sous diagonales n'ont pas leur moment d'ordre 2. Une matrice de Wigner lourde se distingue d'une matrice de Wigner par le fait que la loi commune des ses entrées sous diagonales peut dépendre de N, de sortes que ses moments peuvent être grands. Plus précisément, une matrice X_N est dite de Wigner lourde lorsque

1. pour tout $N \geqslant 1$, la matrice $A_N = \sqrt{N} X_N$ est N par N, réelle symétrique. Les entrées sous diagonales de A_N sont indépendantes et identiquement distribuées selon une mesure $p^{(N)}$ sur $\mathbb{R}$ qui possède tout ses moments,
2. pour tout $k \geqslant 1$, la suite des $2k$-ièmes moments satisfait
$$a_k := \lim_{N\to\infty} \frac{\int t^{2k} \mathrm{d}p^{(N)}(t)}{N^{k-1}} \text{ existe dans } \mathbb{R},$$
3. et on a $\sqrt{N} \int t \mathrm{d}p^{(N)}(t) = o(N^\beta)$ pour tout $\beta > 0$.

On a le même résultat de convergence du spectre pour les matrices de Wigner lourdes que pour les matrices de Lévy, avec une description combinatoire de la distribution asymptotique des valeurs propres [Zak06]. Ce modèle interpole ceux de Wigner et de Lévy.

0.2.3 Matrices aléatoires rencontrées en statistiques

Nous avons donné l'exemple des matrices de covariance empiriques d'échantillons de vecteurs gaussiens dans la Section 0.2.1. De nombreuses matrices aléatoires issues des statistiques sont des variantes de ce modèle. Notre objectif n'est pas d'en donner une liste exhaustive, mais de souligner leur mode de construction : ces matrices sont obtenues comme une combinaison algébrique de matrices aléatoires de type bruit (non nécessairement gaussien) et de matrices déterministes de type signal.

1. **Matrice de covariance séparable**
$$H_{N,N'} = A_N^{\frac{1}{2}} M_{N,N'} B_{N'} M_{N,N'}^* A_N^{\frac{1}{2}},$$
où
   - $\sqrt{N'} M_{N,N'}$ est une matrice N par N', à entrées indépendantes et identiquement distribuées,
   - $A_N^{\frac{1}{2}}$ est la racine carrée hermitienne positive d'une matrice déterministe A_N, hermitienne positive et de taille N par N,
   - B_N est une matrice déterministe, hermitienne positive et de taille N' par N'.
2. **Modèle information plus bruit**
$$H_{N,N'} = (M_{N,N'} + A_{N,N'})(M_{N,N'} + A_{N,N'})^*,$$
où $M_{N,N'}$ est comme précédemment et $A_{N,N'}$ est une matrice déterministe de taille N par N'.

3. **Perturbation d'une matrice hermitienne**

$$H_{N,N} = A_N + X_N,$$

où A_N est une matrice déterministe hermitienne et $\sqrt{N}X_N$ est une matrice hermitienne à entrées sous diagonales indépendantes et identiquement distribuées.

Pour chacun de ces exemples, lorsque les entrées des matrices aléatoires ont leurs moments finis et indépendants de N, il s'avère qu'on peut calculer la distribution asymptotique des valeurs propres d'une telle matrice connaissant :

1. la taille des matrices de type bruit ; par exemple le paramètre c dans le cas des matrices de covariance empiriques,
2. les spectres limites des matrices de type signal, ou éventuellement les distributions limites de leurs valeurs singulières lorsque les matrices ne sont pas carrées ; par exemple, la distribution limite des valeurs propres de Σ_N dans le cas des matrices de covariance empiriques.

La théorie des probabilités libres donne une vision unifiée qui permet pour chacun de ces modèles, par exemple, de calculer leur distribution de valeurs propres limite en fonction de la distribution asymptotique des valeurs propres ou singulières des matrices déterministes. Ces résultats témoignent du phénomène dit de liberté asymptotique des matrices aléatoires.

0.2.4 Le modèle générique étudié dans ce mémoire et survol des résultats de cette thèse

De manière informelle, suivant la méthodologie héritée du théorème de liberté asymptotique de Voiculescu, nous étudierons des modèles de matrices aléatoires hermitiennes de la forme

$$H_N = P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*), \tag{5}$$

où

1. $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ est une famille de type bruit multi-matriciel : les matrices de $\mathbf{X}_N$ sont aléatoires, indépendantes, hermitiennes et à entrées sous diagonales indépendantes.
2. $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ est une famille de type signal multi-matriciel : les matrices de $\mathbf{Y}_N$ sont possiblement aléatoires mais $\mathbf{Y}_N$ est indépendante de $\mathbf{X}_N$.
3. P est un polynôme en $p+2q$ indéterminées non commutatives à coefficients dans $\mathbb{C}$, indépendant de N et tel que H_N est une matrice hermitienne.

Il arrivera souvent que nous supposions que la famille des matrices déterministes $\mathbf{Y}_N$ satisfait des hypothèses asymptotiques qui dépendront du contexte (convergence, éventuellement forte, au sens des espaces de probabilités non commutatives aux Chapitres 1 et 2, ou convergence au sens des distribution de trafics aux Chapitres 3 et 4).

Toutefois, dans les éléments techniques de la preuve du résultat principal des Chapitres 1 et 2, nous ne ferons aucune hypothèse de nature asymptotique sur $\mathbf{Y}_N$. Ce fait a son intérêt dans le contexte des statistiques des matrices aléatoires où l'on parle de recherche d'un équivalent déterministe pour une telle approche.

Dans le Chapitre 1, la famille $\mathbf{X}_N$ est constituée de matrices du GUE indépendantes. Nous précisons des hypothèses sur les matrices de $\mathbf{Y}_N$ de sorte que toute matrice hermitienne de la forme $H_N = P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)$ satisfait la propriété suivante :

Presque sûrement, la distribution empirique des valeurs propres de H_N converge vers une mesure de probabilité μ sur $\mathbb{R}$, et pour N assez grand les valeurs propres de H_N sont contenues dans un voisinage du support de μ.

La convergence vers la mesure μ est connue par le théorème de liberté asymptotique de Voiculescu, le résultat nouveau étant la seconde partie de l'énoncé. L'hypothèse sur $\mathbf{Y}_N$ relève de la convergence au sens des $\mathcal{C}^*$-algèbres. Ce résultat s'étend pour des matrices non hermitiennes rectangulaires, ou composées de matrices de Haar sur le groupe unitaire (voir les applications du Chapitre 1 et le résultat principal du Chapitre 2).

Dans les Chapitre 3 et 4, la famille $\mathbf{X}_N$ est constituée de matrices de Wigner lourdes indépendantes. Nous précisons des hypothèse sur les matrices de $\mathbf{Y}_N$ de sorte que toute matrice hermitienne de la forme $H_N = P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)$ satisfait la propriété suivante :

La distribution empirique moyennisée des valeurs propres de H_N converge en moments vers une mesure de probabilité μ sur $\mathbb{R}$, et les statistiques linéaires normalisées par un facteur $\sqrt{N}$ en les valeurs propres de H_N satisfont un théorème central limite.

L'hypothèse sur $\mathbf{Y}_N$ est une généralisation commune des convergences au sens des probabilités non commutatives et au sens des graphes, que nous appelons convergence en distribution de trafics. La question de la convergence des supports n'a pas été abordée pour de telles matrices : en général, le spectre limite d'une matrice de Wigner lourde est non borné, et il est connu que les valeurs propres extrêmes d'une matrice de Lévy ont un comportement très différents de celles d'une matrice de Wigner (voir [ABAP09, Sos04]).

Remarque : Dans les exemples issus de la statistiques, nous avons considéré des modèles constitués de matrices non carrées. Ces modèles peuvent être étudiés via un choix judicieux de matrices H_N de la forme générique (5). Prenons l'exemple de la matrice de covariance séparée

$$H_{N,N'} = A_N^{\frac{1}{2}} M_{N,N'} B_{N'} M_{N,N'}^* A_N^{\frac{1}{2}}.$$

Soit $X_{N,N'}$ une matrice de taille $(N+N')$ par $(N+N')$, hermitienne, et dont les entrées sous diagonales sont indépendantes, telle que

$$X_{N,N'} = \begin{pmatrix} X_N^{(1)} & M_{N,N'} \\ M_{N,N'}^* & X_N^{(2)} \end{pmatrix}. \tag{6}$$

Posons les matrices par blocks de taille $(N+N')$ par $(N+N')$

$$e_1^{(N)} = \begin{pmatrix} \mathbf{1}_N & \mathbf{0}_{N,N'} \\ \mathbf{0}_{N',N} & \mathbf{0}_{N'} \end{pmatrix}, \qquad e_2^{(N)} = \begin{pmatrix} \mathbf{0}_N & \mathbf{0}_{N,N'} \\ \mathbf{0}_{N',N} & \mathbf{1}_{N'} \end{pmatrix},$$
$$\tilde{A}_{N,N'}^{\frac{1}{2}} = \begin{pmatrix} A_N^{\frac{1}{2}} & \mathbf{0}_{N,N'} \\ \mathbf{0}_{N',N} & \mathbf{0}_{N'} \end{pmatrix}, \qquad \tilde{B}_{N,N'}^{(N)} = \begin{pmatrix} \mathbf{0}_N & \mathbf{0}_{N,N'} \\ \mathbf{0}_{N',N} & B_{N'} \end{pmatrix}.$$

Dès lors, on peut étudier les propriétés spectrales de $H_{N,N'}$ via celles de la matrice

$$\begin{aligned} \tilde{H}_{N,N'} &= \tilde{A}_{N,N'}^{\frac{1}{2}} a_1^{(N)} X_{N,N'} e_2^{(N)} \tilde{B}_{N,N'} e_2^{(N)} X_{N,N'} e_1^{(N)} \tilde{A}_{N,N'}^{\frac{1}{2}} \\ &= \begin{pmatrix} H_{N,N'} & \mathbf{0}_{N,N'} \\ \mathbf{0}_{N',N} & \mathbf{0}_{N'} \end{pmatrix}, \end{aligned}$$

qui est bien de la forme générique (5) en posant $\mathbf{X}_N = (X_{N,N'})$ et $\mathbf{Y}_N = (e_1^{(N)}, e_2^{(N)}, \tilde{A}_{N,N'}^{\frac{1}{2}}, \tilde{B}_{N,N'})$.

0.3 La théorie des probabilités libres pour l'étude du spectre de grandes matrices aléatoires

La théorie des probabilités libres permet entre autre de décrire le spectre asymptotique des matrices H_N de la Section 0.2.4 lorsque le bruit multi-matriciel est constitué de matrices de Wigner indépendantes.

Heuristiquement, l'idée est de ne plus voir une matrice aléatoire comme une collection d'un grand nombre de variables aléatoires, mais comme une variable aléatoire à part entière et d'une nature non commutative. La théorie des probabilités libres introduite par Voiculescu donne un cadre formel à ce principe, ainsi qu'une intuition probabiliste pour manipuler ces objets : en effet, dans le cadre des *-espaces de probabilité, est définie la notion de liberté qui joue un rôle analogue à la notion d'indépendance dans le cadre classique des probabilités.

0.3.1 Probabilités non commutatives

Définitions et exemples

Un *-espace de probabilité est la donnée d'un triplet $(\mathcal{A}, .^*, \tau)$, où

1. $\mathcal{A}$ est une algèbre unifère sur $\mathbb{C}$,
2. $.^*$ est une involution anti-linéaire sur $\mathcal{A}$ vérifiant $(ab)^* = b^*a^*$ pour tout a, b dans $\mathcal{A}$,

3. τ est une forme linéaire sur $\mathcal{A}$, appelée état, telle que $\tau[\mathbf{1}_{\mathcal{A}}] = 1$ et $\tau[a^*a] \geqslant 0$ pour tout a dans $\mathcal{A}$.

L'état sera toujours supposé tracial, c'est à dire satisfaisant $\tau[ab] = \tau[ba]$ pour tout a, b dans $\mathcal{A}$. Très souvent, nous supposerons également que l'état est fidèle, c'est à dire satisfaisant $\tau[a^*a] = 0$ si et seulement si $a = 0$.

Voyons comme exemples de référence les deux espaces usuels suivants.

- **Espaces de probabilité classiques :**
  Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. L'ensemble $\bigcap_{p\geqslant 1} L^p(\Omega)$ des variables aléatoires complexes sur Ω possédant tout leurs moments est un *-espace de probabilité lorsqu'il est muni de la conjugaison complexe $\bar{\cdot}$ et de l'espérance $\mathbb{E}$ relative à $\mathbb{P}$.

- **Espaces de matrices :**
  L'ensemble $\mathrm{M}_N(\mathbb{C})$ des matrices de taille N par N à coefficients dans $\mathbb{C}$ est un *-espace de probabilité lorsqu'il est muni de la transposition complexe et de la trace normalisée $\tau_N = \frac{1}{N}\mathrm{Tr}$.

Le fait que la notion de *-espace de probabilité modélise bien un espace de probabilité vient du résultat suivant. Soit h un élément normal dans un *-espace de probabilité $(\mathcal{A}, .^*, \tau)$, c'est à dire vérifiant $hh^* = h^*h$. Alors, il existe une mesure de probabilité μ_h sur $\mathbb{C}$, telle que pour tout polynôme P à deux indéterminées, on a

$$\tau\Big[P(h, h^*)\Big] = \int P(z, \bar{z})\mathrm{d}\mu_h(z). \tag{7}$$

De plus, cette mesure est unique dès lors qu'elle est caractérisée par ses moments (par exemple si elle est à support compact).

Ce fait motive les définitions suivantes. Les éléments de $\mathcal{A}$ sont appelés des variables aléatoires non commutatives. La loi jointe d'une famille $\mathbf{a} = (a_1, \dots, a_p)$ d'éléments de $\mathcal{A}$ (appelée également loi non commutative s'il y a risque de confusion) est la forme linéaire

$$\begin{array}{rccc} \tau_{\mathbf{a}}: & \mathbb{C}\langle \mathbf{z}, \mathbf{z}^*\rangle & \to & \mathbb{C} \\ & P & \mapsto & \tau\Big[P(\mathbf{a}, \mathbf{a}^*)\Big], \end{array} \tag{8}$$

où $\mathbb{C}\langle \mathbf{z}, \mathbf{z}^*\rangle$ désigne l'ensemble des polynômes en $2p$ indéterminées non commutatives $z_1, \dots, z_p, z_1^*, \dots, z_p^*$ et $P(\mathbf{a}, \mathbf{a}^*)$ est une notation pour $P(a_1, \dots, a_p, a_1^*, \dots, a_p^*)$. Enfin, la convergence en distribution (dite aussi en loi non commutative) d'une suite de familles $(\mathbf{a}_N)_{N\geqslant 1}$ est la convergence simple de la suite de fonctions $(\tau_{\mathbf{a}_N})_{N\geqslant 1}$.

Retournons aux exemples de référence pour mieux appréhender les notions de loi jointe et de convergence en distribution.

- **Espaces de probabilité classiques :**
  Soient $A_1, \dots, A_p$ des variables aléatoires complexes admettant tout leurs

moments. Alors, leur loi non commutative n'est autre que la donnée des moments joints en ces variables et en leurs conjuguées, c'est à dire, de la collection des nombres complexes

$$\mathbb{E}[A_1^{n_1}\bar{A}_1^{m_1}\dots A_p^{n_p}\bar{A}_p^{m_p}] \tag{9}$$

pour tout entiers $n_1,\dots,n_p,m_1,\dots,m_p \geqslant 0$. Si la loi de probabilité de $(A_1,\dots,A_p)$ est caractérisée par ses moments (par exemple si les variables aléatoires sont bornées), alors celle-ci coïncide avec la loi non commutative de $(A_1,\dots,A_p)$. La convergence en distribution est alors la convergence en moments.

– **Espaces de matrices :**
Soit A_N une matrice normale de taille N par N. Notons $\lambda_1,\dots,\lambda_N$ ses valeurs propres. Alors, pour tout polynôme P en une variable, on a

$$\tau_N\Big[P(A_N)\Big] = \frac{1}{N}\sum_{i=1}^{N} P(\lambda_i). \tag{10}$$

Ainsi, la distribution de A_N n'est autre que la mesure empirique de ses valeurs propres. Pour une famille de matrice $\mathbf{A}_N = (A_1^{(N)},\dots,A_p^{(N)})$, la donnée des nombres complexes

$$\tau_N\Big[P(\mathbf{A}_N,\mathbf{A}_N^*)\Big] \tag{11}$$

pour tout polynôme non commutatif P est plus riche que la simple donnée des spectres des matrices $A_1^{(N)},\dots,A_p^{(N)}$. La distribution de $\mathbf{A}_N$ tient compte des positions relatives des sous espaces propres de $A_1^{(N)},\dots,A_p^{(N)}$.

Intérêt de la notion de convergence en distribution pour l'étude du spectre de grandes matrices aléatoires

Considérons une famille $\mathbf{X}_N = (X_1^{(N)},\dots,X_p^{(N)})$ de matrice de taille N par N et une famille $\mathbf{x} = (x_1,\dots,x_p)$ de variables aléatoires non commutative dans un espace $(\mathcal{A}, .^*, \tau)$. Alors, la convergence en distribution de $\mathbf{X}_N$ vers $\mathbf{x}$ signifie la convergence, pour tout polynôme P en $2p$ indéterminées non commutatives, de la suite de nombres

$$\tau_N\Big[P(\mathbf{X}_N,\mathbf{X}_N^*)\Big] \underset{N\to\infty}{\longrightarrow} \tau\Big[P(\mathbf{x},\mathbf{x}^*)\Big]. \tag{12}$$

Fixons un polynôme Q tel que la matrice $H_N = Q(\mathbf{X}_N,\mathbf{X}_N^*)$ et la variable aléatoire non commutative $h = Q(\mathbf{x},\mathbf{x}^*)$ sont normales. Soit μ_h une mesure de probabilité sur $\mathbb{C}$ satisfaisant la formule (7). Alors, en appliquant la convergence (12) avec $P = Q^k$ pour tout entier $k \geqslant 1$, nous obtenons que la distribution empirique des valeurs propres de H_N converge en moments vers la mesure μ_h.

Nous retiendrons donc le principe suivant :

La convergence en distribution d'une famille de matrices aléatoires $\mathbf{X}_N = (X_1^{(N)},\dots,X_p^{(N)})$ implique la convergence en moments de la mesure empirique des valeurs propres de toute matrice hermitienne de la forme $H_N = P(\mathbf{X}_N,\mathbf{X}_N^*)$, où P est indépendant de N.

0.3.2 La liberté asymptotiques de grandes matrices aléatoires

Motivation : le spectre de la somme de deux matrices hermitiennes

Illustrons la problématique de la liberté asymptotique à travers l'exemple suivant. Soient X_N et Y_N deux matrices hermitiennes de taille N par N, possiblement aléatoires mais indépendantes. Pour toute matrice hermitienne H_N, notons $\lambda_1^{(N)}(H_N) \leqslant \cdots \leqslant \lambda_N^{(N)}(H_N)$ ses valeurs propres triées, $\mathcal{L}_{H_N}$ leur distribution empirique et Δ_{H_N} la matrice diagonale diag $\big(\lambda_1^{(N)}(H_N), \ldots, \lambda_N^{(N)}(H_N)\big)$.

La donnée des valeurs propres de X_N et de Y_N ne permet pas de déterminer les valeurs propres de leur somme X_N+Y_N. En effet, il existe une matrice unitaire U_N telle que le spectre de X_N+Y_N est le spectre de la matrice $U_N\Delta_{X_N}U_N^*+\Delta_{Y_N}$.

Cas commutatifs : Dans les deux exemples qui suivent, les matrices X_N et Y_N commutent. Ainsi, par le procédé de diagonalisation simultané, nous sommes ramené à un problème de couplage de mesures de probabilité sur $\mathbb{R}$.

- **couplage monotone :** si U_N est la matrice identité, alors on a $\lambda_i^{(N)}(X_N+Y_N) = \lambda_i^{(N)}(X_N)+\lambda_i^{(N)}(Y_N)$ pour tout $i = 1, \ldots, N$. Ainsi $\mathcal{L}_{X_N+Y_N}$ est la loi de la somme de deux variables aléatoires x et y, où x et y sont distribuées selon les lois $\mathcal{L}_{X_N}$ et $\mathcal{L}_{Y_N}$ respectivement et suivent le couplage monotone standard des variables aléatoires réelles.
- **convolution des mesures :** considérons le cas où U_N est une matrice aléatoire, distribuée uniformément sur l'ensemble des matrices de permutation. Il s'avère alors que la distribution empirique des valeurs propres de $X_N + Y_N$ a la loi de la convolution des mesures $\mathcal{L}_{X_N}$ et $\mathcal{L}_{Y_N}$.

La théorie de probabilités libres permet de décrire le spectre de X_N+Y_N dans le cas suivant :

1. la matrice U_N est distribuée selon la mesure de Haar sur le groupe unitaire et est indépendante de $(\Delta_{X_N}, \Delta_{Y_N})$,
2. la taille N des matrices tend vers l'infini, les mesures $\mathcal{L}_{X_N}$ et $\mathcal{L}_{Y_N}$ ayant une limite en moments $\mathcal{L}_x$ et $\mathcal{L}_y$ respectivement.

Voiculescu a défini dans le contexte des *-espace de probabilité la notion de liberté, analogue de la notion d'indépendance des variables aléatoires classiques. Il s'avère alors que la mesure $\mathcal{L}_{X_N+Y_N}$ converge vers une mesure notée $\mathcal{L}_x \boxplus \mathcal{L}_y$ qui est décrite comme la somme de deux variables aléatoires non commutatives x et y "libres", distribuées selon les lois $\mathcal{L}_x$ et $\mathcal{L}_y$ respectivement. La mesure $\mathcal{L}_x \boxplus \mathcal{L}_y$ est appellée convolution libre des mesure $\mathcal{L}_x$ et $\mathcal{L}_y$.

A noter que si X_N est une matrice du GUE, alors d'après la Section 0.2.1, nous sommes dans ce cas d'application.

Définition de la liberté

Soit $(\mathcal{A}, .^*, \tau)$ un *-espace de probabilité. Soient $\mathcal{A}_1, \ldots, \mathcal{A}_k$ des $*$-sous algèbres unifères de $\mathcal{A}$. Ces algèbres sont dites libres dès lors que pour tout $n \geqslant 1$, tout

$a_i \in \mathcal{A}_{j_i}$ $(i = 1, \ldots, n,\ j_i \in \{1, \ldots, k\})$, on a

$$\tau\Big[\Big(a_1 - \tau[a_1]\Big) \cdots \Big(a_n - \tau[a_n]\Big)\Big] = 0$$

dès lors que $j_1 \neq j_2,\ j_2 \neq j_3, \ldots, j_{n-1} \neq j_n$. Des familles de variables aléatoires non commutatives sont dites libres lorsque les sous algèbres qu'elles engendrent le sont.

En pratique, lorsque les familles $\mathbf{a}_1, \ldots, \mathbf{a}_k$ sont libres, la distribution jointe de l'ensemble des familles $(\mathbf{a}_1, \ldots, \mathbf{a}_k)$ est complètement déterminée par les distributions jointes de chaque famille $\mathbf{a}_j$ pour $j = 1, \ldots, k$.

La liberté formalise une relation de haute non commutativité entre des variables non commutatives comme en témoigne l'exemple suivant. Soient a et b deux variables non commutatives dans un *-espace de probabilité $(\mathcal{A}, .^*, \tau)$ dont l'état τ est fidèle. Supposons que a et b sont libres et centrées ($\tau[a] = \tau[b] = 0$). Alors, par définition de la liberté, on a

$$\tau[aba^*b^*] = 0 \qquad (13)$$

$$\tau\Big[\Big(aa^* - \tau[aa^*]\Big)\Big(bb^* - \tau[bb^*]\Big)\Big] = 0. \qquad (14)$$

Supposons par ailleurs que les variables a^* et b commutent. Dès lors, on a $0 = \tau[aba^*b^*] = \tau[aa^*bb^*] = \tau[aa^*]\tau[bb^*]$. Ainsi, $\tau[aa^*] = 0$ ou $\tau[bb^*] = 0$. Mais l'état étant fidèle, nous obtenons alors $a = 0$ ou $b = 0$.

La liberté asymptotique de grandes matrices aléatoires

Le théorème de Voiculescu de liberté asymptotique affirme la chose suivante (voir [AGZ10] pour une démonstration). Pour tout entier $N \geqslant 1$, considérons

- $\mathbf{X}_N = (X_1^{(N)}, \ldots, X_p^{(N)})$ une famille de matrices de Wigner N par N indépendantes,
- $\mathbf{U}_N = (U_1^{(N)}, \ldots, U_q^{(N)})$ une famille de matrices de Haar sur le groupe unitaire N par N indépendantes,
- $\mathbf{Y}_N = (Y_1^{(N)}, \ldots, Y_r^{(N)})$ une famille de matrices N par N, possiblement aléatoires,
- les familles $\mathbf{X}_N, \mathbf{U}_N$ et $\mathbf{Y}_N$ étant supposées indépendantes.

Dans un *-espace de probabilité $(\mathcal{A}, .^*, \tau)$, considérons

- $\mathbf{x} = (x_1, \ldots, x_p)$ un système semicirculaire libre, i.e. les variables aléatoires non commutatives sont libres, auto-adjointes ($x_j^* = x_j$ pour tout $j = 1, \ldots, p$) et pour tout $j = 1, \ldots, p$ et tout $k \geqslant 1$, on a

$$\tau[x_j^k] = \int_{-2}^{2} t^k \frac{1}{\sqrt{2\pi}} \sqrt{4 - t^2} \mathrm{d}(t), \qquad (15)$$

- $\mathbf{u} = (u_1, \ldots, u_q)$ une famille d'unités de Haar libres, i.e les variables aléatoires non commutatives sont libres, normales ($u_j u_j^* = u_j^* u_j$ pour tout $j = 1, \ldots, p$) et pour tout $j = 1, \ldots, p$ et tout $k, l \geqslant 1$, on a

$$\tau\Big[u_j^k (u_j^*)^l\Big] = \delta_{k=l}, \qquad (16)$$

  où δ désigne le symbole de Kronecker,

– $\mathbf{y} = (y_1, \dots, y_r)$ une famille de variables aléatoires non commutative,
– les familles $\mathbf{x}, \mathbf{u}$ et $\mathbf{y}$ étant supposées libres.

On suppose que la famille des matrices $\mathbf{Y}_N$ satisfait :

1. Presque sûrement, la distribution de $\mathbf{Y}_N$ dans $\big(\mathrm{M}_N(\mathbb{C}), .^*, \tau_N\big)$ converge vers la distribution de $\mathbf{y}$ dans $(\mathcal{A}, .^*, \tau)$.
2. Presque sûrement, pour tout $j = 1, \dots, r$, la norme d'opérateur de $Y_j^{(N)}$ est bornée indépendamment de N.

Alors, la distribution de $(\mathbf{X}_N, \mathbf{U}_N, \mathbf{Y}_N)$ dans $\big(\mathrm{M}_N(\mathbb{C}), .^*, \tau_N\big)$ converge presque sûrement vers la distribution de $(\mathbf{x}, \mathbf{u}, \mathbf{y})$ dans $(\mathcal{A}, .^*, \tau)$, c'est à dire, presque sûrement, pour tout polynôme P en $p + 2q + 2r$ indéterminées, on a

$$\tau_N\Big[P(\mathbf{X}_N, \mathbf{U}_N, \mathbf{U}_N^*, \mathbf{Y}_N, \mathbf{Y}_N^*)\Big] \underset{N\to\infty}{\longrightarrow} \tau\Big[P(\mathbf{x}, \mathbf{u}, \mathbf{u}^*, \mathbf{y}, \mathbf{y}^*)\Big]. \qquad (17)$$

En vertue du principe énoncé à la fin de la Section 0.3.1, le théorème de liberté asymptotique permet de calculer le spectre limite d'une large classe de matrices aléatoires.

0.4 Présentation de la Partie I : la forte liberté asymptotique

Dans leur article [HT05], Haagerup et Thorbjørnsen ont établit un renforcement du théorème de liberté asymptotique pour des matrices du GUE indépendantes. Celui-ci s'exprime sur une structure plus riche que celle des *-espace de probabilité et est appelé convergence forte en distribution. La convergence forte en distribution pour une grande matrice hermitienne permet de comprendre plus en détails son spectre : elle implique que ses valeurs propres appartiennent à un petit voisinage du spectre limite lorsque la taille des matrices est assez grande.

Définitions et exemples

Un $\mathcal{C}^*$-espace de probabilité $(\mathcal{A}, .^*, \tau, \| \cdot \|)$ est la donnée d'un *-espace de probabilité $(\mathcal{A}, .^*, \tau)$ tel que $(\mathcal{A}, .^*, \| \cdot \|)$ est une $\mathcal{C}^*$-algèbre, c'est à dire que $\mathcal{A}$ est une algèbre de Banach et que la norme $\| \cdot \|$ satisfait $\|a^*a\| = \|a\|^2$ pour tout a dans $\mathcal{A}$. Par la construction de Gelfand-Naimark-Segal, on peut toujours réaliser une C^*-algèbre comme sous algèbre de l'algèbre des opérateurs bornés sur un espace de Hilbert. En outre, on peut utiliser le calcul fonctionnel sur ces espaces.

Cette structure s'applique sur nos exemples de référence.

– **Espaces de probabilité classiques :**
Etant donné $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, $(L^\infty(\Omega), \bar{\cdot}, \mathbb{E}, \| \cdot \|_\infty)$ est un $\mathcal{C}^*$-espace de probabilité, $\| \cdot \|_\infty$ désignant la norme infini essentielle.

- **Espaces de matrices :**
  L'ensemble $(\mathrm{M}_N(\mathbb{C}), .^*, \tau_N, \|\cdot\|)$ est un $\mathcal{C}^*$-espace de probabilité, $\|\cdot\|$ désignant la norme d'opérateur, i.e. $\|M\| = \sqrt{\rho(M^*M)}$, ρ étant le rayon spectral. Si M est hermitienne, alors il y a égalité entre rayon spectral et norme d'opérateur.

Une structure de $\mathcal{C}^*$-espace de probabilité $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ a surtout un intérêt lorsque l'état τ est fidèle. Dans ce cas, la norme s'exprime en fonction de l'état par la formule suivante : pour tout a dans $\mathcal{A}$, on a

$$\|a\| = \lim_{k\to\infty} \Big(\tau\Big[(a^*a)^k\Big]\Big)^{\frac{1}{2k}}. \tag{18}$$

Soit h un élément auto-adjoint dans un $\mathcal{C}^*$-espace de probabilité $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ dont l'état est fidèle. On a vu qu'il existe une mesure de probabilité μ_h sur $\mathbb{R}$ telle que pour tout polynôme P, on a

$$\tau\Big[P(h)\Big] = \int P \mathrm{d}\mu_h. \tag{19}$$

Alors, μ_h est nécessairement à support compact et la norme de h est donnée par

$$\|h\| = \max_{t\in \mathrm{Supp}\ (\mu_h)} |t|. \tag{20}$$

Pour tout n dans $\mathbb{N}\cup\{\infty\}$, soit $\mathbf{a}_N = (a_1^{(N)}, \dots, a_p^{(N)})$ une famille de variables aléatoires non commutatives dans un $\mathcal{C}^*$-espace de probabilité $(\mathcal{A}_N, .^*, \tau_N, \|\cdot\|_{\mathcal{A}_N})$. On dit que $\mathbf{a}_N$ converge fortement en distribution vers $\mathbf{a}_\infty$ lorsque pour tout polynôme P en $2p$ indéterminées non commutatives, on a

$$\begin{aligned} \tau_N\Big[P(\mathbf{a}_N, \mathbf{a}_N^*)\Big] &\underset{N\to\infty}{\longrightarrow} \tau_\infty\Big[P(\mathbf{a}_\infty, \mathbf{a}_\infty^*)\Big], \\ \Big\|P(\mathbf{a}_N, \mathbf{a}_N^*)\Big\|_{\mathcal{A}_N} &\underset{N\to\infty}{\longrightarrow} \Big\|P(\mathbf{a}_\infty, \mathbf{a}_\infty^*)\Big\|_{\mathcal{A}_\infty}. \end{aligned}$$

Intérêt de la notion de convergence forte en distribution pour l'étude du spectre de grandes matrices aléatoires

Soit $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ une famille de matrices dans $(\mathrm{M}_N(\mathbb{C}), .^*, \tau_N, \|\cdot\|)$ convergeant fortement en distribution vers une famille $\mathbf{x} = (x_1, \dots, x_p)$ dans un $\mathcal{C}^*$-espace de probabilité $(\mathcal{A}_N, .^*, \tau_N, \|\cdot\|)$ dont l'état est fidèle. Soit Q un polynôme tel que la matrice $H_N = Q(\mathbf{X}_N, \mathbf{X}_N^*)$ est hermitienne. Par fidélité de l'état, la variable aléatoire non commutative $h = Q(\mathbf{x}, \mathbf{x}^*)$ est toujours auto-adjointe.

La convergence forte en distribution de H_N vers h s'applique alors. Dès lors, il s'avère que pour toute fonction continue f (et non simplement polynomiale) on a

$$\Big\|f(H_N)\Big\| \underset{N\to\infty}{\longrightarrow} \Big\|f(h)\Big\|, \tag{21}$$

où $f(H_N)$ et $f(h)$ sont donnés par le calcul fonctionnel. Soit $\varepsilon > 0$ un nombre réel. Choisissons pour fonction continue une fonction f_ε, s'annulant sur Supp (h) +

$(-\frac{\varepsilon}{2}, \frac{\varepsilon}{2})$, et constante égale à 1 sur le complémentaire de Supp $(h) + (-\varepsilon, \varepsilon)$. On a donc

$$\Big\| f_\varepsilon(H_N) \Big\| \underset{N\to\infty}{\longrightarrow} \Big\| f_\varepsilon(h) \Big\| = 0, \tag{22}$$

et ainsi, pour N assez grand, toutes les valeurs propres de H_N sont contenues dans Supp $(h) + (-\varepsilon, \varepsilon)$.

Nous retiendrons donc le principe suivant :

La convergence forte en distribution d'une famille de matrices aléatoires $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ implique le phénomène "aucune valeur propre en dehors d'un voisinage du support limite" pour toute matrice hermitienne de la forme $H_N = P(\mathbf{X}_N, \mathbf{X}_N^*)$, où P est indépendant de N.

La liberté asymptotique forte de grandes matrices aléatoires

Dans les Chapitres 1 et 2, nous montrons le résultat suivant (voir les Théorèmes 1.1.6 et 2.1.4).

Théorème 0.4.1 (La liberté asymptotique forte de grandes matrices aléatoires)**.** Pour tout entier $N \geqslant 1$, considérons les familles de matrices $\mathbf{X}_N, \mathbf{U}_N$ et $\mathbf{Y}_N$ comme dans le théorème de liberté asymptotique de la Section 0.3.2. Dans un $\mathcal{C}^*$-espace de probabilité $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ dont l'état est fidèle, considérons les familles de variables aléatoires non commutatives $\mathbf{x}, \mathbf{u}$ et $\mathbf{y}$ comme dans la Section 0.3.2. On suppose que presque sûrement la distribution de $\mathbf{Y}_N$ converge fortement vers la distribution de $\mathbf{y}$. Alors, presque sûrement la distribution de $(\mathbf{X}_N, \mathbf{U}_N, \mathbf{Y}_N)$ dans $\Big(\mathrm{M}_N(\mathbb{C}), .^*, \tau_N, \|\cdot\|\Big)$ converge fortement vers la distribution de $(\mathbf{x}, \mathbf{u}, \mathbf{y})$ dans $(\mathcal{A}, .^*, \tau, \|\cdot\|)$, c'est à dire, presque sûrement, pour tout polynôme P en $p + 2q + 2r$ indéterminées, on a

$$\tau_N\Big[P(\mathbf{X}_N, \mathbf{U}_N, \mathbf{U}_N^*, \mathbf{Y}_N, \mathbf{Y}_N^*)\Big] \underset{N\to\infty}{\longrightarrow} \tau\Big[P(\mathbf{x}, \mathbf{u}, \mathbf{u}^*, \mathbf{y}, \mathbf{y}^*)\Big], \tag{23}$$

$$\Big\|P(\mathbf{X}_N, \mathbf{U}_N, \mathbf{U}_N^*, \mathbf{Y}_N, \mathbf{Y}_N^*)\Big\| \underset{N\to\infty}{\longrightarrow} \Big\|P(\mathbf{x}, \mathbf{u}, \mathbf{u}^*, \mathbf{y}, \mathbf{y}^*)\Big\|. \tag{24}$$

En vertu du principe énoncé à la fin de la Section 0.3.1, le théorème de liberté asymptotique permet d'obtenir le phénomène "aucune valeur propre en dehors d'un voisinage du support limite" de Bai et Silverstein [BS98] pour une large classe de matrices aléatoires.

En outre, en démontrant la liberté asymptotique forte pour des matrices de Haar sur le groupe unitaires indépendantes (sans matrices gaussiennes, ni matrices déterministes) nous répondons à une question naturelle de la théorie des espaces d'opérateurs. La convergence fortes pour des matrices de Haar sur le groupe orthogonal (respectivement symplectique) est également montrée dans le Chapitre 2.

Le résultat initial de Haagerup et Thorbjørnsen [HT05] est la liberté asymptotique forte des matrices $\mathbf{X}_N$ seulement. Dans le Chapitre 1, nous généralisons leur méthode afin d'établir la convergence forte pour $\mathbf{X}_N$ et $\mathbf{Y}_N$. L'idée nouvelle par rapport à [HT05] est d'utiliser une méthode de Bai et Silverstein [BS98] d'équivalent déterministe. Dans le Chapitre 2, nous montrons la liberté asymptotique forte lorsqu'on adjoint les matrices de $\mathbf{U}_N$. La preuve de ce résultat est basée sur un couplage entre une matrice du GUE et une matrice de Haar.

Un corollaire non direct de ce résultat traite de la convergence du support de la mesure empirique des valeurs propres de la somme de deux matrices. Soient X_N et Y_N deux matrices aléatoires indépendantes dont l'une est invariante en loi par conjugaison par une matrice unitaire. Supposons que, presque sûrement, chacune des matrices admet une distribution asymptotique des valeurs propres à support compact, et que pour N assez grand ses valeurs propres sont contenues dans un petit voisinage du spectre limite. Alors, presque sûrement, pour N assez grand les valeurs propres de $X_N + Y_N$ sont contenues dans un petit voisinage du support de la convolution libre des distributions des valeurs propres limites de X_N et de Y_N.

0.5 Présentation de la partie II : la fausse liberté asymptotique

Dans le Chapitre 3, nous introduisons un analogue de la théorie des probabilités libres qui permet de décrire le spectre des matrices H_N de la Section 0.2.4 lorsque le bruit multi-matriciel est constitué de matrices de Wigner lourdes indépendantes.

Illustration : retour sur le spectre de la somme de deux matrices hermitiennes

Soient X_N et Y_N deux matrices hermitiennes de taille N par N. Voyons X_N comme une matrice de "type bruit", aléatoire, et Y_N comme une matrice déterministe de "type signal" (soumise à des hypothèses asymptotiques). Par le théorème spectral, on peut écrire $X_N = U_N \Delta_N U_N^*$, où U_N est une matrice unitaire et Δ_N est une matrice diagonale.

Si X_N est une matrice de Wigner, alors par le théorème de Voiculescu, les matrices X_N et Y_N sont asymptotiquement libres. Ainsi, la distribution spectrale limite de $X_N + Y_N$ est la même que celle dans le cas où X_N est distribuée selon le GUE. Rappelons qu'alors, la matrice unitaire U_N est distribuée selon la mesure de Haar sur le groupe unitaire, et est indépendante de Δ_N.

Si maintenant X_N est une matrice de Wigner lourde, il s'avère que la distribution de U_N est très différente de la mesure de Haar. En particulier, le spectre limite possible pour $X_N + Y_N$ dépend de plus d'information sur Y_N que la simple connaissance de son spectre limite. Par exemple, nous montrons dans le Chapitre 3 que l'on a presque surement

- si Y_N est la réalisation d'une matrice du Z_N du GUE, alors X_N et Y_N sont asymptotiquement libres,
- ça n'est pas le cas si Y_N est la matrice diagonale Δ_{Z_N} des valeurs propres d'une réalisation de Z_N.

Présentation du Chapitre 3 : Distributions de trafics

Nous introduisons une notion de distribution qui, en particulier, est une façon de capter cette information supplémentaire sur Y_N nécessaire pour décrire le spectre limite de $X_N + Y_N$ lorsque X_N est une matrice de Wigner lourde. Nous appelons distribution de trafics cette donnée. Heuristiquement, l'idée est de ne plus voir une matrice aléatoire comme une collection d'un grand nombre de variables aléatoires, mais comme un grand graphe dont les arêtes sont étiquetées par des variables aléatoires. Ces graphes étiquetés sont traditionnellement appelés réseaux. La théorie des probabilités libres sert de support méthodologique pour introduire une notion de produit entre les distributions de trafics, appelé faux produit libre (à noter que cette notion n'est définie que dans un cas particulier dans ce mémoire et est développée dans un travail en préparation).

Voir une grande matrice aléatoire comme un grand réseau est une idée qui a fait ses preuves pour l'étude du spectre de grandes matrices de Lévy. En effet, Bordenave, Caputo et Chafaï [BCC10, BCC11] ont montré la convergence locale d'opérateur d'une matrice de Wigner lourde vers un réseau appelé l'arbre infini aux poids poissoniens. Recouper la démarche de ces auteurs avec celle de la distribution de trafics sera un problème intéressant. Cela permettrait de comprendre un analogue des probabilités libres permettant de décrire le spectre des matrices H_N de la Section 0.2.4 lorsque le bruit multi-matriciel est constitué de matrices de Lévy indépendantes.

Dans le contexte des distributions de trafics, le théorème central de ce Chapitre est le suivant (voir le Théorème 3.3.8 pour un énoncé précis).

Théorème 0.5.1 (La convergence en distribution de trafics de grandes matrices aléatoires)**.** Soit $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ une famille de matrices indépendantes de Wigner lourdes de taille N par N et $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ une famille de matrices déterministes de taille N par N. Alors, outre des hypothèses techniques sur $\mathbf{Y}_N$, si la famille $\mathbf{Y}_N$ admet une distribution de trafics limite, c'est aussi le cas pour la famille $(\mathbf{X}_N, \mathbf{Y}_N)$.

De ce résultat, nous déduisons que dans ces conditions la famille $(\mathbf{X}_N, \mathbf{Y}_N)$ dans $(\mathrm{M}_N(\mathbb{C}), .^*, \mathbb{E}[\tau_N])$ converge en distribution (au sens des *-espaces de probabilités) vers une famille $(\mathbf{x}, \mathbf{y})$ de variables aléatoires non libres en général. Ceci donne donc un moyen de calculer le spectre limite des matrices H_N de la Section 0.2.4 dans le cas où le bruit multi-matriciel est constitué de matrices de Wigner lourdes. Nous décrivons également des outils combinatoires et formels pour calculer des moments joints de $(\mathbf{x}, \mathbf{y})$, voir par exemple le Lemme 3.5.4 et le Théorème 3.6.2.

Par ailleurs, nous définissons la notion distribution de trafics pour un graphe aléatoire enraciné stationnaire (G, v). De manière informelle, cette distribution porte l'information du nombre moyen d'injections de petits graphes dans G. Il s'avère que la donnée de la distribution de trafics de (G, v) est équivalente à celle de sa loi de graphe aléatoire.

En outre, nous établissons une équivalence entre la convergence en distribution de trafics d'un graphe et sa convergence sous un autre mode, appelé "la convergence locale faible". Cette dernière notion a été introduite par Benjamini et Schramm [BS01], puis développée par Aldous et Steele [AS04] (voir aussi les travaux d'Aldous et Lyons [AL07]). Heuristiquement, elle consiste en l'étude asymptotique locale d'un graphe autour d'un sommet tiré uniformément dans le graphe. Nous montrons dans ce Chapitre le résultat suivant (voir le Théorème 3.4.6 pour un énoncé plus riche).

Théorème 0.5.2 (Convergence locale faible et convergence en distribution de trafics pour les graphes)**.** Soit G_N un graphe à N sommets. Alors, G_N a une distribution limite de trafics si et seulement si G_N a une limite faible locale et les deux limites, dans chacun des sens, sont en correspondance.

Présentation du Chapitre 4 : Un théorème central limite

Nous établissons un théorème central limite dans le cadre de la convergence en distribution de trafics pour une famille de matrices de Wigner lourdes indépendantes. De ce résultat, nous déduisons un théorème central limite pour les statistiques linéaires en les valeurs propres d'une matrice de Wigner lourde (voir le Théorème 4.2.2).

Théorème 0.5.3 (Un théorème central limite pour les statistiques linéaire spectrales de matrices de Wigner lourdes)**.** Soit X_N une matrice de Wigner lourde. Pour tout polynôme P, la variable aléatoire

$$\sqrt{N}\Big(\tau_N\big[P(X_N)\big] - \mathbb{E}\big[\tau_N\big[P(X_N)\big]\big]\Big) \tag{25}$$

est asymptotiquement gaussienne.

La normalisation par un facteur $\sqrt{N}$ n'est pas usuelle en théorie des matrices aléatoires. Pour une matrice de Wigner non lourde, il est connu [Jon82] qu'un théorème central limite a lieu avec un facteur normalisant N. Le facteur $\sqrt{N}$ est usuel pour les statistiques linéaires en des variables aléatoires indépendantes. Ainsi, les corrélations entre les valeurs propres d'une matrice de Wigner lourde sont d'une nature très différente de celles entre les valeurs propres d'une matrice de Wigner non lourde.

Le Théorème central de ce Chapitre renforce l'idée que la notion de distribution de trafics est adaptée à l'étude d'une famille $\mathbf{X}_N$ de matrices de Wigner lourdes indépendantes. En effet, la covariance dans le théorème central limite (Théorème 4.2.2) a une expression simple, et fait ne fait intervenir que la distribution de trafics limite de $\mathbf{X}_N$.

Première partie

Forte Liberté Asymptotique

Chapitre 1

The norm of polynomials in large random and deterministic matrices

With an appendix by Dimitri Shlyakhtenko.

ABSTRACT:

Let $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ *be a family of* $N \times N$ *independent, normalized random matrices from the Gaussian Unitary Ensemble. We state sufficient conditions on matrices* $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$*, possibly random but independent of* $\mathbf{X}_N$*, for which the operator norm of* $P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)$ *converges almost surely for all polynomials* P*. Limits are described by operator norms of objects from free probability theory. Taking advantage of the choice of the matrices* $\mathbf{Y}_N$ *and of the polynomials* P*, we get for a large class of matrices the "no eigenvalues outside a neighborhood of the limiting spectrum" phenomena. We give examples of diagonal matrices* $\mathbf{Y}_N$ *for which the convergence holds. Convergence of the operator norm is shown to hold for block matrices, even with rectangular Gaussian blocks, a situation including non-white Wishart matrices and some matrices encountered in MIMO systems.*

1.1 Introduction and statement of result

For a Hermitian $N \times N$ matrix H_N, let $\mathcal{L}_{H_N}$ denote its empirical eigenvalue distribution, namely

$$\mathcal{L}_{H_N} = \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i},$$

where δ_λ is the Dirac mass in λ and $\lambda_1, \dots, \lambda_N$ are the eigenvalues of H_N. The empirical eigenvalue distribution of large dimensional random matrices has been studied with much interest for a long time. One pioneering result is Wigner's theorem [Wig58], from 1958. Let W_N be an $N \times N$ Wigner matrix. Then the theorem states that, under appropriate assumptions, the n-th moment of $\mathcal{L}_{W_N}$ converges in expectation to the n-th moment of the semicircular law as N goes to infinity for any integer n. This result has been generalized in many directions, notably by Arnold [Arn67] for the almost sure convergence of the moments. The convergence of the empirical eigenvalue distribution for covariance matrices was first shown by Marčenko and Pastur [MP67] in 1967, and has been generalized in the late 1970's and the early 1980's by many people, including Grenander and Silverstein [GS77], Wachter [Wac78], Jonsson [Jon82], Yin and Krishnaiah [YK83], Bai, Yin and Krishnaiah [BYK86] and Yin [Yin86].

In 1991, Voiculescu [Voi91] discovered a connection between large random matrices and free probability theory. He showed the so-called asymptotic freeness theorem, which has been generalized for instance in [HP00, Tho00, Voi98], which implies the almost sure weak convergence of the empirical eigenvalue distribution for Hermitian matrices H_N of the form

$$H_N = P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*),$$

where
- P is a fixed polynomial in $2p+q$ non commutative indeterminates,
- $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ is a family of independent $N \times N$ matrices of the normalized Gaussian Unitary Ensemble (GUE),
- $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ are $N \times N$ matrices with appropriate assumptions (see Theorem 1.1.3 below).

The limiting empirical eigenvalue distribution of H_N can be computed by using the notion of freeness. Recall that an $N \times N$ random matrix $X^{(N)}$ is said to be a normalized GUE matrix if it is Hermitian with entries $(X_{n,m}^{(N)})_{1 \leqslant n,m \leqslant N}$, such that the set of random variables $(X_{n,n}^{(N)})_{1 \leqslant n \leqslant N}$, and $(\sqrt{2}\mathrm{Re}\,(X_{n,m}^{(N)}), \sqrt{2}\mathrm{Im}\,(X_{n,m}^{(N)})\,)_{1 \leqslant n < m \leqslant N}$ forms a centered Gaussian vector with covariance matrix $\frac{1}{N}\mathbf{1}_{N^2}$. Moreover, the result of Voiculescu holds even for independent Wigner or Wishart matrices instead of GUE matrices, as it has been proved by Dykema [Dyk93] and Capitaine and Casalis [CC04] respectively.

Currently, it is known for some random matrices, as for example Wigner and Wishart matrices, that, almost surely, the eigenvalues of the matrix belong to a small neighborhood of the limiting eigenvalue distribution for N large enough. More formally, if H_N is a Hermitian matrix whose empirical eigenvalue distribution converges weakly to a probability measure μ it is observed in many situations [BY88, YBK88, BSY88, BS98, PS09] that : for all $\varepsilon > 0$, almost surely there exists $N_0 \geqslant 1$ such that for all $N \geqslant N_0$ one has

$$\mathrm{Sp}\Big(\ H_N\ \Big) \subset \mathrm{Supp}\ \Big(\ \mu\ \Big) + (-\varepsilon, \varepsilon), \tag{1.1}$$

where " Sp " means the spectrum and " Supp " means the support.

The convergence of the extremal eigenvalues to the edges of the spectrum of a single Wigner or Wishart matrix has been shown in the early 1980's by Geman [Gem80], Juhász [Juh81], Füredi and Komlós [FK81], Jonsson [Jon85] and Silverstein [Sil89, Sil85]. In 1988, in the case of a real Wigner matrix, Bai and

Yin stated in [BY88] necessary and sufficient conditions for the convergence in terms of the first four moments of the entries of these matrices. In the case of a Wishart matrix, the similar result is due to Yin, Bai, and Krishnaiah [YBK88] and Bai, Silverstein, and Yin [BSY88]. The case of a complex matrix has been investigated later by Bai [Bai99]. The phenomenon "no eigenvalues outside (a small neighborhood of) the support of the limiting distribution" has been shown in 1998 by Bai and Silverstein [BS98] for large sample covariance matrices and in 2008 by Paul and Silverstein [PS09] for large separable covariance matrices.

In 2005, Haagerup and Thorbjørnsen [HT05] have shown (1.1) using operator algebra techniques for matrices $H_N = P(X_1^{(N)}, \dots, X_p^{(N)})$, where P is a polynomial in p non commutative indeterminates and $X_1^{(N)}, \dots, X_p^{(N)}$ are independent, normalized $N \times N$ GUE matrices. This constitutes a real breakthrough in the context of free probability. Their method has been used by Schultz [Sch05] to obtain the same result for Gaussian random matrices with real or symplectic entries, and by Capitaine and Donati-Martin [CDM07] for Wigner matrices with symmetric distribution of the entries satisfying a Poincaré inequality and for Wishart matrices.

A consequence of the main result of the present article is that the phenomenon (1.1) holds in the setting considered by Voiculescu, i.e. for certain Hermitian matrices H_N of the form $H_N = P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)$.

Theorem 1.1.1 (The spectrum of large Hermitian random matrices)**.** Let $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ be a family of independent, normalized GUE matrices and $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ be a family of $N \times N$ matrices, possibly random but independent of $\mathbf{X}_N$. Assume that for every Hermitian matrix H_N of the form

$$H_N = P(\mathbf{Y}_N, \mathbf{Y}_N^*),$$

where P is a polynomial in $2q$ non commutative indeterminates, we have with probability one that:

1. **Convergence of the empirical eigenvalue distribution:** there exists a compactly supported measure μ on the real line such that the empirical eigenvalue distribution of H_N converges weakly to μ as N goes to infinity.
2. **Convergence of the spectrum:** for any $\varepsilon > 0$, almost surely there exists N_0 such that for all $N \geqslant N_0$,

$$\mathrm{Sp}\Big(H_N \Big) \subset \mathrm{Supp}\ \Big(\mu \Big) + (-\varepsilon, \varepsilon). \tag{1.2}$$

Then almost surely the convergences of the empirical eigenvalue distribution and of the spectrum also hold for all Hermitian matrices $H_N = P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)$, where P is a polynomial in $p + 2q$ non commutative indeterminates.

Theorem 1.1.1 is a straightforward consequence of Theorem 1.1.6 below, where the language of free probability is used. Moreover, Theorem 1.1.6 specifies Theorem 1.1.1 by giving a description of the limit of the empirical eigenvalue distribution. For readers convenience, we recall some definitions (see [NS06] and [AGZ10] for details).

Definition 1.1.2. 1. A *-probability space $(\mathcal{A}, .^*, \tau)$ consists of a unital $\mathbb{C}$-algebra $\mathcal{A}$ endowed with an antilinear involution $.^*$ such that $(ab)^* = b^*a^*$ for all a, b in $\mathcal{A}$, and a state τ. A state τ is a linear functional $\tau : \mathcal{A} \mapsto \mathbb{C}$ satisfying

$$\tau[\mathbf{1}] = 1, \quad \tau[a^*a] \geqslant 0 \; \forall a \in \mathcal{A}. \tag{1.3}$$

The elements of $\mathcal{A}$ are called non commutative random variables. We will always assume that τ is a trace, i.e. that it satisfies $\tau[ab] = \tau[ba]$ for every $a, b \in \mathcal{A}$. The trace τ is said to be faithful when it satisfies $\tau[a^*a] = 0$ only if $a = 0$.

2. The non commutative law of a family $\mathbf{a} = (a_1, \dots, a_p)$ of non commutative random variables is defined as the linear functional $P \mapsto \tau\Big[P(\mathbf{a}, \mathbf{a}^*)\Big]$, defined on the set of polynomials in $2p$ non commutative indeterminates. The convergence in law is the pointwise convergence relative to this functional.

3. The families of non commutative random variables $\mathbf{a}_1, \dots, \mathbf{a}_n$ are said to be free if for all K in $\mathbb{N}$, for all non commutative polynomials $P_1, \dots, P_K$

$$\tau\Big[P_1(\mathbf{a}_{i_1}, \mathbf{a}_{i_1}^*) \dots P_K(\mathbf{a}_{i_K}, \mathbf{a}_{i_K}^*)\Big] = 0 \tag{1.4}$$

as soon as $i_1 \neq i_2$Ê $\neq \dots \neq i_K$ and $\tau\Big[P_k(\mathbf{a}_{i_k}, \mathbf{a}_{i_k}^*)\Big] = 0$ for $k = 1, \dots, K$.

4. A family of non commutative random variables $\mathbf{x} = (x_1, \dots, x_p)$ is called a free semicircular system when the non commutative random variables are free, selfadjoint ($x_i = x_i^*$, $i = 1, \dots, p$), and for all k in $\mathbb{N}$ and $i = 1, \dots, p$, one has

$$\tau[x_i^k] = \int t^k d\sigma(t), \tag{1.5}$$

with $d\sigma(t) = \frac{1}{2\pi}\sqrt{4 - t^2}\; \mathbf{1}_{|t| \leqslant 2}\; dt$ the semicircle distribution.

Recall first the statement of Voiculescu's asymptotic freeness theorem.

Theorem 1.1.3 ([HP00, Tho00, Voi95b, Voi98] The asymptotic freeness of $X_1^{(N)}, \dots, X_p^{(N)}$ and $\mathbf{Y}_N$). Let $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ be a family of independent, normalized GUE matrices and $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ be a family of $N \times N$ matrices, possibly random but independent of $\mathbf{X}_N$. Let $\mathbf{x} = (x_1, \dots, x_p)$ be a free semicircular system in a *-probability space $(\mathcal{A}, .^*, \tau)$ and $\mathbf{y} = (y_1, \dots, y_q)$ in $\mathcal{A}^q$ be a family of non commutative random variables free from $\mathbf{x}$. Assume the following.

1. **Convergence of $\mathbf{Y}_N$:** Almost surely, the non commutative law of $\mathbf{Y}_N$ in $(\mathrm{M}_N(\mathbb{C}), .^*, \tau_N)$ converges to the non commutative law of $\mathbf{y}$, which means that for all polynomial P in $2q$ non commutative indeterminates, one has

$$\tau_N\Big[P(\mathbf{Y}_N, \mathbf{Y}_N^*)\Big] \underset{N\to\infty}{\longrightarrow} \tau\Big[P(\mathbf{y}, \mathbf{y}^*)\Big], \tag{1.6}$$

where τ_N denotes the normalized trace of $N \times N$ matrices.

2. **Boundedness of the spectrum:** Almost surely, for $j = 1, \dots, q$ one has

$$\limsup_{N\to\infty} \|Y_j^{(N)}\| < \infty, \tag{1.7}$$

where $\|\cdot\|$ denotes the operator norm.

Then the non commutative law of $(\mathbf{X}_N, \mathbf{Y}_N)$ in $(\mathrm{M}_N(\mathbb{C}), .^*, \tau_N)$ converges to the non commutative law of $(\mathbf{x}, \mathbf{y})$, i.e. for all polynomial P in $p+2q$ non commutative indeterminates, one has

$$\tau_N\Big[P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)\Big] \underset{N\to\infty}{\longrightarrow} \tau\Big[P(\mathbf{x}, \mathbf{y}, \mathbf{y}^*)\Big]. \tag{1.8}$$

In [HT05] Haagerup and Thorbjørnsen strengthened the connection between random matrices and free probability. Limits of random matrices have now to be seen in more elaborated structure, called $\mathcal{C}^*$-probability space, which is endowed with a norm.

Definition 1.1.4. A $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ consists of a *-probability space $(\mathcal{A}, .^*, \tau)$ and a norm $\|\cdot\|$ such that $(\mathcal{A}, .^*, \|\cdot\|)$ is a $\mathcal{C}^*$-algebra.

By the Gelfand-Naimark-Segal construction, one can always realize $\mathcal{A}$ as a norm-closed $\mathcal{C}^*$-subalgebra of the algebra of bounded operators on a Hilbert space. Hence we can use functional calculus on $\mathcal{A}$. Moreover, if τ is a faithful trace, then the norm $\|\cdot\|$ is uniquely determined by the following formula (see [NS06, Proposition 3.17]):

$$\|a\| = \lim_{k\to\infty}\Big(\ \tau\big[\ (a^*a)^k\ \big]\ \Big)^{\frac{1}{2k}}, \forall a \in \mathcal{A}. \tag{1.9}$$

The main result of [HT05] is the following.

Theorem 1.1.5 ([HT05] The strong asymptotic freeness of independent GUE matrices)**.** Let $X_1^{(N)}, \dots, X_p^{(N)}$ be independent, normalized $N \times N$ GUE matrices and let $x_1, \dots, x_p$ be a free semicircular system in a $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \|\cdot\hat{\mathrm{E}}\|)$ with a faithful trace. Then almost surely, one has: for all polynomials P in p non commutative indeterminates, one has

$$\Big\|P(X_1^{(N)}, \dots, X_p^{(N)})\Big\| \underset{N\to\infty}{\longrightarrow} \|P(x_1, \dots, x_p)\|. \tag{1.10}$$

This article is mainly devoted to the following theorem which is a generalization of Theorem 1.1.5 in the setting of Theorem 1.1.3.

Theorem 1.1.6 (The strong asymptotic freeness of $X_1^{(N)}, \dots, X_p^{(N)}, \mathbf{Y}_N$)**.** Let $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ be a family of independent, normalized GUE matrices and $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ be a family of $N \times N$ matrices, possibly random but independent of $\mathbf{X}_N$. Let $\mathbf{x} = (x_1, \dots, x_p)$ and $\mathbf{y} = (y_1, \dots, y_q)$ be a family of non commutative random variables in a $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ with a faithful trace, such that $\mathbf{x}$ is a free semicircular system free from $\mathbf{y}$. Assume the following.

Strong convergence of $\mathbf{Y}_N$: Almost surely, for all polynomials P in $2q$ non commutative indeterminates, one has

$$\tau_N\Big[P(\mathbf{Y}_N, \mathbf{Y}_N^*)\Big] \underset{N\to\infty}{\longrightarrow} \tau[P(\mathbf{y}, \mathbf{y}^*)], \tag{1.11}$$

$$\Big\|P(\mathbf{Y}_N, \mathbf{Y}_N^*)\Big\| \underset{N\to\infty}{\longrightarrow} \|P(\mathbf{y}, \mathbf{y}^*)\|. \tag{1.12}$$

Then, almost surely, for all polynomials P in $p+2q$ non commutative indeterminates, one has

$$\tau_N\Big[P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)\Big] \underset{N\to\infty}{\longrightarrow} \tau[P(\mathbf{x}, \mathbf{y}, \mathbf{y}^*)], \tag{1.13}$$

$$\Big\|P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)\Big\| \underset{N\to\infty}{\longrightarrow} \|P(\mathbf{x}, \mathbf{y}, \mathbf{y}^*)\|. \tag{1.14}$$

The convergence of the normalized traces stated in (1.13) is the content of Voiculescu's asymptotic freeness theorem and is recalled in order to give a coherent and complete statement. Theorem 1.1.1 is easily deduced from Theorem 1.1.6 by applying Hamburger's theorem [Ham21] for the convergence of the measure and functional calculus for the convergence of the spectrum.

Organization of the paper: In Section 1.2 we give applications of Theorem 1.1.6 which are proved in Section 1.9. Sections 1.3 to 1.8 are dedicated to the proof of Theorem 1.1.6.

Acknowledgments: The author would like to thank Alice Guionnet for dedicating much time for many discussions to the subjects of this paper and, along with Manjunath Krishnapur and Ofer Zeitouni, for the communication of Lemma 1.8.2. He is very much obliged to Dimitri Shlyakhtenko for his contribution to this paper. He would like to thank Benoit Collins for pointing out an error in a previous version of Corollary 1.2.1 and giving the idea to fix it. He also likes to thank Mikael de la Salle for useful discussions.

Yin, Bai, and Krishnaiah [YBK88] and Bai, Silverstein, and Yin [BSY88]. The case of a complex matrix has been investigated later by Bai [Bai99]. In this series of papers, where the assumptions on the matrices were progressively relaxed up to the optimal ones, proofs were basically combinatorial, and based on the truncation of entries.

1.2 Applications

1.2.1 Diagonal matrices

The first and the simpler matrix model that may be investigated to play the role of matrices $\mathbf{Y}_N$ in Theorem 1.1.6 consists of deterministic diagonal matrices with real entries and prescribed asymptotic spectral measure.

Corollary 1.2.1 (diagonal matrices)**.** Let $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ be a family of independent, normalized GUE matrices and let $\mathbf{D}_N = (D_1^{(N)}, \dots, D_q^{(N)})$ be $N \times N$ deterministic real diagonal matrices, such that for any $j = 1, \dots, q$,

1. the empirical spectral distribution of $D_j^{(N)}$ converges weakly to a compactly supported probability measure μ_j,
2. the diagonal entries of $D_j^{(N)}$ are non decreasing:

$$D_j^{(N)} = \text{ diag } \Big(\lambda_1^{(N)}(j), \dots, \lambda_N^{(N)}(j) \Big), \text{ with } \lambda_1^{(N)}(j) \leqslant \dots \leqslant \lambda_N^{(N)}(j),$$

3. for all $\varepsilon > 0$, there exists N_0 such that for all $N \geqslant N_0$, for all $j = 1 \dots q$,

$$\mathrm{Sp}\Big(D_j^{(N)} \Big) \subset \mathrm{Supp}\Big(\mu_j \Big) + (-\varepsilon, \varepsilon).$$

Let $v = (v_1, \dots, v_q)$ in $[0,1]^q$. We set $\mathbf{D}_N^v = \Big(D_1^{(N)}(v_1), \dots, D_q^{(N)}(v_q)\Big)$, where for any $j = 1, \dots, q$,

$$D_j^{(N)}(v_j) = \text{ diag } \Big(\lambda_{1+\lfloor v_j N \rfloor}^{(N)}(j), \dots, \lambda_{N+\lfloor v_j N \rfloor}^{(N)}(j) \Big), \text{ with indices modulo } N.$$

Let $\mathbf{x} = (x_1, \dots, x_p)$ and $\mathbf{d}^v = \Big(d_1(v), \dots, d_q(v)\Big)$ be non commutative random variables in a $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \| \cdot \|)$ with a faithful trace, such that

1. $\mathbf{x}$ is a free semicircular system, free from $\mathbf{d}^v$,
2. The variables $d_1(v), \dots, d_q(v)$ commute, are selfadjoint and for all polynomials P in q indeterminates, one has

$$\tau[P(\mathbf{d}^v)\ ] = \int_0^1 P\Big(F_1^{-1}(u + v_1), \dots, F_q^{-1}(u + v_q) \Big) du. \tag{1.15}$$

   For any $j = 1 \dots q$, the application F_j^{-1} is the (periodized) generalized inverse of the cumulative distribution function $F_j : t \mapsto \mu_j\Big( ]-\infty, t]\Big)$ of μ_j defined by: F_j^{-1} is 1-periodic and for all u in $]0,1]$, $F_j^{-1}(u) = \inf\Big\{t \in \mathbb{R} \ \Big|\ F_j(t) \geqslant u\Big\}$.

Then, with probability one, for all polynomials P in $p+q$ non commutative indeterminates, one has

$$\tau_N\Big[P(\mathbf{X}_N, \mathbf{D}_N^v)\Big] \underset{N\to\infty}{\longrightarrow} \tau[P(\mathbf{x}, \mathbf{d}^v)] \tag{1.16}$$

$$\Big\|P(\mathbf{X}_N, \mathbf{D}_N^v)\Big\| \underset{N\to\infty}{\longrightarrow} \|P(\mathbf{x}, \mathbf{d}^v)\|, \tag{1.17}$$

for any v in $[0,1]^q$ except in a countable set.

Remark that the non commutative random variables $d_1, \dots, d_q$ can be realized as classical random variables, d_j being μ_j-distributed for $j = 1, \dots, q$. The dependence between the random variables is trivial since Formula (1.15) exhibits a deterministic coupling. The convergence of the normalized trace (1.16) actually holds for any v. In general, the convergence (1.17) of the norm can fail: the family of matrices $\mathbf{D}_N = (D_1^{(N)}, D_2^{(N)})$ where

$$D_1^{(N)} = \text{ diag } (\mathbf{0}_{\lfloor N/2 \rfloor}, \mathbf{1}_{N - \lfloor N/2 \rfloor}), \ D_1^{(N)} = \text{ diag } (\mathbf{0}_{\lfloor N/2 \rfloor + 1}, \mathbf{1}_{N - \lfloor N/2 \rfloor - 1})$$

gives a counterexample (consider their difference). Furthermore, let mention that it is clear that we always can take one of the v_i to be zero.

1.2.2 Non-white Wishart matrices

Theorem 1.1.6 may be used to deduce the same result for some Wishart matrices as for the GUE matrices. Let $r, s_1, \ldots, s_p \geqslant 1$ be integers. Let $\mathbf{Z}_N = (Z_1^{(N)}, \ldots, Z_p^{(N)})$ be a family of independent positive definite Hermitian random matrices such that for $j = 1, \ldots, p$ the matrix $Z_j^{(N)}$ is of size $s_jN \times s_jN$. Let $\mathbf{W}_N = \mathbf{W}_N(\mathbf{Z}) = (W_1^{(N)}, \ldots, W_p^{(N)})$ be the family of $rN \times rN$ matrices defined by: for each $j = 1, \ldots, p$, $W_j^{(N)} = M_j^{(N)}\ Z_j^{(N)}\ M_j^{(N)*}$, where $M_j^{(N)}$ is a $rN \times s_jN$ matrix whose entries are random variables,

$$M_j^{(N)} = (M_{n,m})_{\substack{1\leqslant n\leqslant rN \\ 1\leqslant m\leqslant s_jN}},$$

and the random variables $(\sqrt{2}\text{Re }(M_{n,m}), \sqrt{2}\text{Im }(M_{n,m})\ )_{1\leqslant n\leqslant rN, 1\leqslant m\leqslant s_jN}$ form a centered Gaussian vector with covariance matrix $\frac{1}{rN}\mathbf{1}_{2rs_jN^2}$. We assume that $M_1^{(N)}, \ldots, M_p^{(N)}, \mathbf{Z}_N$ are independent. The matrices $W_1^{(N)}, \ldots, W_p^{(N)}$ are called non-white Wishart matrices, the white case occurring when the matrices $Z_j^{(N)}$ are the identity matrices.

Corollary 1.2.2 (Wishart matrices)**.** Let $\mathbf{Y}_N = (Y_1^{(N)}, \ldots, Y_q^{(N)})$ be a family of $rN \times rN$ random matrices, independent of $\mathbf{Z}_N$ and $\mathbf{W}_N$. Assume that the families of matrices $(Z_1^{(N)}), \ldots, (Z_q^{(N)}), \mathbf{Y}_N$ satisfy separately the assumptions of Theorem 1.1.6. Then, almost surely, for all polynomials P in $p + 2q$ non commutative indeterminates, one has

$$\left\| P(\mathbf{W}_N, \mathbf{Y}_N, \mathbf{Y}_N^*) \right\| \underset{N\to\infty}{\longrightarrow} \|P(\mathbf{w}, \mathbf{y}, \mathbf{y}^*)\|, \tag{1.18}$$

where $\|\cdot\|$ is given by Formula (1.9) with τ a faithful trace for which the non commutative random variables $\mathbf{w} = (w_1, \ldots, w_p)$ and $\mathbf{y} = (y_1, \ldots, y_q)$ are free.

In [PS09], motivated by applications in statistics and wireless communications, the authors study the global limiting behavior of the spectrum of the following matrix, referred as separable covariance matrix:

$$C_n = \frac{1}{n} A_n^{1/2} X_n B_n X_n^* A_n^{1/2},$$

where X_n is a $n \times m$ random matrix, $A_n^{1/2}$ is a nonnegative definite square root of the nonnegative definite $n \times n$ Hermitian matrix A_n and B_n is a $m \times m$ diagonal matrix with nonnegative diagonal entries. It is shown in [PS09] that, for n large enough, almost surely the eigenvalues of C_n belong in a small neighborhood of the limiting distribution under the following assumptions:

1. $m = m(n)$ with $c_n := n/m \underset{n\to\infty}{\longrightarrow} c > 0$.
2. The entries of X_n are independent, identically distributed, standardized complex and with a finite fourth moment.
3. The empirical eigenvalue distribution $\mathcal{L}_{A_n}$ (respectively $\mathcal{L}_{B_n}$) of A_n (respectively B_n) converges weakly to a compactly supported probability measure ν_a (respectively ν_b) and the operator norms of A_n and B_n are uniformly bounded.

4. By assumptions 1,2 and 3, it is known that almost surely $\mathcal{L}_{C_n}$ converges weakly to a probability measure $\mu^{(c)}_{\nu_a,\nu_b}$. This define a map $\Phi : (x, \nu_1, \nu_2) \mapsto \mu^{(x)}_{\nu_1,\nu_2}$ (the input x is a positive real number, the inputs ν_1 and ν_2 are probability measures on $\mathbb{R}^+$). Assume that for every $\varepsilon > 0$, there exists $n_0 \geqslant 1$ such that, for all $n \geqslant n_0$, one has

$$\mathrm{Supp}\ \Big(\mu^{(c_n)}_{\mathcal{L}_{A_n},\mathcal{L}_{B_N}} \Big) \subset \mathrm{Supp}\ \Big(\mu^{(c)}_{\nu_a,\nu_b} \Big) \ + (-\varepsilon, \varepsilon).$$

Now consider the following situation, where Corollary 1.2.2 may be applied

1' $n = n(N) = rN$, $m = m(N) = sN$ for fixed positive integers r and s,

2' the entries of X_n are independent, identically distributed, standardized complex Gaussian,

3' the empirical eigenvalue distribution of A_n (respectively B_n) converges weakly to a compactly supported probability measure,

4' for N large enough, the eigenvalues of A_n (respectively B_n) belong in a small neighborhood of its limiting distribution.

Then we obtain by Corollary 1.2.2 that for N large enough, almost surely the eigenvalues of C_n belong in a small neighborhood of the limiting distribution. The advantage of our version is the replacement of assumption 4 by assumption 4'. Replacing assumptions 1' and 2' by assumptions 1 and 2 could be an interesting question.

1.2.3 Block matrices

It will be shown as a consequence of Theorem 1.1.6 that the convergence of norms (1.14) also holds for block matrices.

Corollary 1.2.3 (Block matrices)**.** Let $\mathbf{X}_N, \mathbf{Y}_N, \mathbf{x}, \mathbf{y}$ and τ be as in Theorem 1.1.6. Almost surely, for all positive integer ℓ and for all non commutative polynomials $(P_{u,v})_{1 \leqslant u,v \leqslant \ell}$, the operator norm of the $\ell N \times \ell N$ block matrix

$$\begin{pmatrix} P_{1,1}(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*) & \dots & P_{1,\ell}(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*) \\ \vdots & & \vdots \\ P_{\ell,1}(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*) & \dots & P_{\ell,\ell}(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*) \end{pmatrix} \tag{1.19}$$

converges to the norm $\| \cdot \|_{\tau_\ell \otimes \tau}$ of

$$\begin{pmatrix} P_{1,1}(\mathbf{x}, \mathbf{y}, \mathbf{y}^*) & \dots & P_{1,\ell}(\mathbf{x}, \mathbf{y}, \mathbf{y}^*) \\ \vdots & & \vdots \\ P_{\ell,1}(\mathbf{x}, \mathbf{y}, \mathbf{y}^*) & \dots & P_{\ell,\ell}(\mathbf{x}, \mathbf{y}, \mathbf{y}^*) \end{pmatrix}, \tag{1.20}$$

where $\| \cdot \|_{\tau_\ell \otimes \tau}$ is given by the faithful trace $\tau_\ell \otimes \tau$ defined by

$$(\tau_\ell \otimes \tau) \left[\begin{pmatrix} P_{1,1}(\mathbf{x}, \mathbf{y}, \mathbf{y}^*) & \dots & P_{1,\ell}(\mathbf{x}, \mathbf{y}, \mathbf{y}^*) \\ \vdots & & \vdots \\ P_{\ell,1}(\mathbf{x}, \mathbf{y}, \mathbf{y}^*) & \dots & P_{\ell,\ell}(\mathbf{x}, \mathbf{y}, \mathbf{y}^*) \end{pmatrix} \right] = \tau \Big[\frac{1}{\ell} \sum_{i=1}^{\ell} P_{i,i}(\mathbf{x}, \mathbf{y}, \mathbf{y}^*) \Big].$$

1.2.4 Channel matrices

We give a potential application of Theorem 1.1.6 in the context of communication, where rectangular block random matrices are sometimes investigated for the study of wireless Multiple-input Multiple-Output (MIMO) systems [LS03, TV04]. In the case of Intersymbol-Interference, the channel matrix H reflects the channel effect during a transmission and is of the form

$$H = \begin{pmatrix} A_1 & A_2 & \dots & A_L & \mathbf{0} & \dots & & \dots & \mathbf{0} \\ \mathbf{0} & A_1 & A_2 & \dots & A_L & \mathbf{0} & & & \vdots \\ \vdots & \mathbf{0} & A_1 & A_2 & \dots & A_L & \mathbf{0} & & \\ & & \ddots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & & \dots & \mathbf{0} & A_1 & A_2 & \dots & A_L \end{pmatrix}, \tag{1.21}$$

$(A_l)_{1\leqslant \ell \leqslant L}$ are $n_R \times n_T$ matrices that are very often modeled by random matrices e.g. $A_1, \dots, A_L$ are independent and for $\ell = 1, \dots, L$ the entries of the matrix A_ℓ are independent identically distributed with finite variance. The number of matrices L is the length of the impulse response of the channel, n_T is the number of transmitter antennas and n_R is the number of receiver antennas.
In order to calculate the capacity of such a channel, one must know the singular value distribution of H, which is predicted by free probability theory. Theorem 1.1.6 may be used to obtain the convergence of the singular spectrum for a large class of such matrices. For instance we investigate in Section 1.9.3 the following case:

Corollary 1.2.4 (Rectangular band matrices)**.** Let r and t be integers. Consider a matrix H of the form (1.21) such that for any $\ell = 1, \dots, L$ one has $A_\ell = C_\ell M_\ell D_\ell$ where

1. $\mathbf{M} = (M_1, \dots, M_L)$ is a family of independent $rN \times tN$ random matrices such that for $\ell = 1, \dots, L$ the entries of M_ℓ are independent, Gaussian and centered with variance σ_ℓ^2/N,
2. the family of $rN \times rN$ matrices $\mathbf{C} = (C_1, \dots, C_L)$ and the family of $tN \times tN$ matrices $\mathbf{D} = (D_1, \dots, D_L)$ satisfy separately the assumptions of Theorem 1.1.6,
3. the families of matrices $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{D}$ are independent.

Then, almost surely, the empirical eigenvalue distribution of HH^* converges weakly to a measure μ. Moreover, for any $\varepsilon > 0$, almost surely there exists N_0 such that the singular values of H belong to $\mathrm{Supp}(\mu) + (-\varepsilon, \varepsilon)$.

1.3 The strategy of proof

Let $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ and $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ be as in Theorem 1.1.6. We start with some remarks in order to simplify the proof.

1. We can suppose that the matrices of $\mathbf{Y}_N$ are Hermitian. Indeed for any $j = 1, \dots, q$, one has $Y_j^{(N)} = \text{Re } Y_j^{(N)} + i \text{ Im } Y_j^{(N)}$, where
$$\text{Re } Y_j^{(N)} := \frac{1}{2}\Big(Y_j^{(N)} + Y_j^{(N)*}\Big), \quad \text{Im } Y_j^{(N)} := \frac{1}{2i}\Big(Y_j^{(N)} - Y_j^{(N)*}\Big)$$
are Hermitian matrices. A polynomial in $\mathbf{Y}_N, \mathbf{Y}_N^*$ is obviously a polynomial in the matrices $\text{Re } Y_1^{(N)}, \dots, \text{Re } Y_q^{(N)}$, and $\text{Im } Y_1^{(N)}, \dots, \text{Im } Y_q^{(N)}$ and so the latter satisfies the assumptions of Theorem 1.1.6 as soon as $\mathbf{Y}_N$ does.
2. It is sufficient to prove the theorem for deterministic matrices $\mathbf{Y}_N$. Indeed, the matrices $\mathbf{X}_N$ and $\mathbf{Y}_N$ are independent. Then we can choose the underlying probability space to be of the form $\Omega = \Omega_1 \times \Omega_2$, with $\mathbf{X}_N$ (respectively $\mathbf{Y}_N$) a measurable function on Ω_1 (respectively Ω_2). The event "for all polynomials P the convergences (1.13) and (1.14) hold" is a measurable set $\tilde{\Omega} \subset \Omega$. Assume that the theorem holds for deterministic matrices. Then for almost all $\omega_2 \in \Omega_2$, there exists a set $\tilde{\Omega}_1(\omega_2)$ for which for all $\omega_1 \in \tilde{\Omega}_1$, (1.13) and (1.14) hold for $(\mathbf{X}_N(\omega_1), \mathbf{Y}_N(\omega_2))$. The set of such couples (ω_1, ω_2) is of outer measure one and is contained in $\tilde{\Omega}$, hence by Fubini's theorem $\tilde{\Omega}$ is of measure one.
3. It is sufficient to prove that for any polynomial the convergence of the norm in (1.14) holds almost surely (instead of almost surely the convergence holds for all polynomials). Indeed we can switch the words "for all polynomials with rational coefficients" and "almost surely" and both the left and the right hand side in (1.14) are continuous in P.

In the following, when we say that $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ is as in Section 1.3, we mean that $\mathbf{Y}_N$ is a family of deterministic Hermitian matrices satisfying (1.11) and (1.12).

Remark that by (1.12), almost surely the supremum over N of $\|Y_j^{(N)}\|$ is finite for all $j = 1, \dots, q$. Hence by Theorem 1.1.3, with probability one the non commutative law of $(\mathbf{X}_N, \mathbf{Y}_N)$ in $(\mathrm{M}_N(\mathbb{C}), .^*, \tau_N)$ converges to the law of non commutative random variables $(\mathbf{x}, \mathbf{y})$ in a *-probability space $(\mathcal{A}, .^*, \tau,)$: almost surely, for all polynomials P in $p + q$ non commutative indeterminates, one has
$$\tau_N\Big[P(\mathbf{X}_N, \mathbf{Y}_N)\ \Big] \underset{N\to\infty}{\longrightarrow} \tau[P(\mathbf{x}, \mathbf{y})], \tag{1.22}$$
where the trace τ is completely defined by:
- $\mathbf{x} = (x_1, \dots, x_p)$ is a free semicircular system,
- $\mathbf{y} = (y_1, \dots, y_q)$ is the limit in law of $\mathbf{Y}_N$,
- $\mathbf{x}, \mathbf{y}$ are free.

Since τ is faithful on the *-algebra spanned by $\mathbf{x}$ and $\mathbf{y}$, we can always assume that τ is a faithful trace on $\mathcal{A}$. Moreover, the matrices $\mathbf{Y}_N$ are uniformly bounded in operator norm. If we define $\|\cdot\|$ in $\mathcal{A}$ by Formula (1.9), then $\|y_j\|$ is finite for every $j = 1, \dots, q$. Hence, we can assume that $\mathcal{A}$ is a $\mathcal{C}^*$-probability space endowed with the norm $\|\cdot\|$.

Haagerup and Thorbjørnsen describe in [HT05] a method to show that for all non commutative polynomials P, almost surely one has

$$\left\| P(\mathbf{X}_N) \right\| \underset{N\to\infty}{\longrightarrow} \| P(\mathbf{x}) \|. \tag{1.23}$$

We present in this section this method with some modification to fit our situation. First, it is easy to see the following.

Proposition 1.3.1. For all non commutative polynomials P, almost surely one has

$$\liminf_{N\to\infty} \left\| P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*) \right\| \geqslant \| P(\mathbf{x}, \mathbf{y}, \mathbf{y}^*) \|. \tag{1.24}$$

Proof. In a $\mathcal{C}^*$-algebra $(\mathcal{A}, .^*, \|\cdot\|)$, one has $\forall a \in \mathcal{A}$, $\|a\|^2 = \|a^*a\|$. Hence, without loss of generality, we can suppose that $H_N := P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)$ is non negative Hermitian and $h := P(\mathbf{x}, \mathbf{y}, \mathbf{y}^*)$ is selfadjoint. Let $\mathcal{L}_N$ denote the empirical spectral distribution of H_N:

$$\mathcal{L}_N = \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i},$$

where $\lambda_1, \dots, \lambda_N$ denote the eigenvalues of H_N and δ_λ the Dirac measure in $\lambda \in \mathbb{R}$. By (1.22) and Hamburger's theorem [Ham21], almost surely $\mathcal{L}_N$ converges weakly to the compactly supported probability measure μ on $\mathbb{R}$ given by: for all polynomial P,

$$\int P \mathrm{d}\mu = \tau[P(h)].$$

Since τ is faithful, the extrema of the support of μ is $\|h\|$ ([NS06, proposition 3.15]). In particular, if $f : \mathbb{R} \to \mathbb{R}$ is a non negative continuous function whose support is the closure of a neighborhood of $\|h\|$ (f not indentically zero), then almost surely there exists a $N_0 \geqslant 0$ such that for all $N \geqslant N_0$ one has $\mathcal{L}_N(f) > 0$. Hence for $N \geqslant N_0$ some eigenvalues of H_N belong to the considered neighborhood of $\|h\|$ and so $\|H_N\| \geqslant \|h\|$. □

It remains to show that the limsup is smaller than the right hand side in (1.24). The method is carried out in many steps.

Step 1. A linearization trick: *With inequality* (1.24) *established, the question of almost sure convergence of the norm of any polynomial in the considered random matrices can be reduced to the question of the convergence of the spectrum of any matrix-valued selfadjoint degree one polynomials in these matrices. More precisely, in order to get* (1.23)*, it is sufficient to show that for all* $\varepsilon > 0$*,* k *positive integer,* L *selfadjoint degree one polynomial with coefficients in* $\mathrm{M}_k(\mathbb{C})$*, almost surely there exists* N_0 *such that for all* $N \geqslant N_0$*,*

$$\mathrm{Sp}\Big(\, L(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*) \,\Big) \subset \mathrm{Sp}\Big(\, L(\mathbf{x}, \mathbf{y}, \mathbf{y}^*) \,\Big) + (-\varepsilon, \varepsilon). \tag{1.25}$$

We refer the readers to [HT05, Parts 2 and 7] for the proof of this step, which is based on $\mathcal{C}^*$-algebra and operator space techniques. We only recall here the main ingredients. By an argument of ultraproduct it is sufficient to show the following: Let $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ be elements of a $\mathcal{C}^*$-algebra. Assume that for all selfadjoint degree one polynomials L with coefficients in $\mathrm{M}_k(\mathbb{C})$, one has

$$\mathrm{Sp}\Big(\ L(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*)\ \Big) \subset \mathrm{Sp}\Big(\ L(\mathbf{x}, \mathbf{y}, \mathbf{y}^*)\ \Big). \tag{1.26}$$

Then for all polynomials P one has $\|P(\mathbf{x}, \mathbf{y}, \mathbf{y}^*)\| \geqslant \|P(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*)\|$. The linearization trick used to prove that fact arises from matrix manipulations and Arveson's theorem: with a dilation argument, one deduces from (1.26) that there exists ϕ a unital $*$-homomorphism between the $\mathcal{C}^*$-algebra spanned by $(\mathbf{x}, \mathbf{y})$ and the one spanned by $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ such that one has $\phi(x_i) = \tilde{x}_i$ for $i = 1, \dots, p$, and $\phi(y_i) = \tilde{y}_i$ for $i = 1, \dots, q$. A $*$-homomorphism being always contractive, one gets the result.

We fix a selfadjoint degree one polynomial L with coefficients in $\mathrm{M}_k(\mathbb{C})$. To prove (1.25) we apply the method of Stieltjes transforms. We use an idea from Bai and Silverstein in [BS98]: we do not compare the Stieltjes transform of $L(\mathbf{X}_N, \mathbf{Y}_N)$ with the one of $L(\mathbf{x}, \mathbf{y})$, but with an intermediate quantity, where in some sense we have taken partially the limit N goes to infinity, only for the GUE matrices. To make it precise, we realize the non commutative random variables $\big(\mathbf{x}, \mathbf{y}, (\mathbf{Y}_N)_{N\geqslant 1}\big)$ in a same $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ with faithful trace, where

- the families $\mathbf{x}$, $\mathbf{y}$, $\mathbf{Y}_1$, $\mathbf{Y}_2, \dots, \mathbf{Y}_N, \dots$ are free,
- for any polynomials P in q non commutative indeterminates $\tau[P(\mathbf{Y}_N)] := \tau_N[P(\mathbf{Y}_N)]$.

The intermediate object $L(\mathbf{x}, \mathbf{Y}_N)$ is therefore well defined as an element of $\mathcal{A}$. We use a theorem about norm convergence, due to D. Shlyakhtenko and stated in Appendix 1.10, to relate the spectrum of $L(\mathbf{x}, \mathbf{Y}_N)$ with the spectrum of $L(\mathbf{x}, \mathbf{y})$.

Step 2. An intermediate inclusion of spectrum: *for all $\varepsilon > 0$ there exists N_0 such that for all $N \geqslant N_0$, one has*

$$\mathrm{Sp}\Big(\ L(\mathbf{x}, \mathbf{Y}_N)\ \Big) \subset \mathrm{Sp}\Big(\ L(\mathbf{x}, \mathbf{y})\ \Big) + (-\varepsilon, \varepsilon). \tag{1.27}$$

We define the Stieltjes transforms g_{L_N} and g_{ℓ_N} of $L_N = L(\mathbf{X}_N, \mathbf{Y}_N)$ and respectively $\ell_N = L(\mathbf{x}, \mathbf{Y}_N)$ by the formulas

$$g_{L_N}(\lambda) = \mathbb{E}\Big[(\tau_k \otimes \tau_N)\Big[\Big(\lambda \mathbf{1}_k \otimes \mathbf{1}_N - L(\mathbf{X}_N, \mathbf{Y}_N)\ \Big)^{-1}\ \Big]\Big], \tag{1.28}$$

$$g_{\ell_N}(\lambda) = (\tau_k \otimes \tau)\Big[\Big(\lambda \mathbf{1}_k \otimes \mathbf{1} - L(\mathbf{x}, \mathbf{Y}_N)\ \Big)^{-1}\ \Big], \tag{1.29}$$

for all complex numbers λ such that $\mathrm{Im}\ \lambda > 0$.

Step 3. From Stieltjes transform to spectra: *In order to show (1.26) with (1.27) granted,*
it is sufficient to show the following: for every $\varepsilon > 0$*, there exist* $N_0, \gamma, c, \alpha > 0$ *such that for all* $N \geqslant N_0$*, for all* λ *in* $\mathbb{C}$ *such that* $\varepsilon \leqslant (\mathrm{Im}\ \lambda)^{-1} \leqslant N^{\gamma}$*, one has*

$$|g_{L_N}(\lambda) - g_{\ell_N}(\lambda)| \leqslant \frac{c}{N^2}(\mathrm{Im}\ \lambda)^{-\alpha}. \tag{1.30}$$

The proof of Estimate (1.30) represents the main work of this paper. For this task we consider a generalization of the Stieltjes transform. We define the $\mathrm{M}_k(\mathbb{C})$-valued Stieltjes transforms G_{L_N} and G_{ℓ_N} of $L_N = L(\mathbf{X}_N, \mathbf{Y}_N)$ and respectively $\ell_N = L(\mathbf{x}, \mathbf{Y}_N)$ by the formulas

$$G_{L_N}(\Lambda) = \mathbb{E}\Big[(\mathrm{id}_k \otimes \tau_N)\Big[\Big(\Lambda \otimes \mathbf{1}_N - L(\mathbf{X}_N, \mathbf{Y}_N)\ \Big)^{-1}\ \Big]\Big], \tag{1.31}$$

$$G_{\ell_N}(\Lambda) = (\mathrm{id}_k \otimes \tau)\Big[\Big(\Lambda \otimes \mathbf{1} - L(\mathbf{x}, \mathbf{Y}_N)\ \Big)^{-1}\ \Big], \tag{1.32}$$

for all $k \times k$ matrices Λ such that the Hermitian matrix $\mathrm{Im}\ \Lambda := (\Lambda - \Lambda^*)/(2i)$ is positive definite. Since $g_{L_N}(\lambda) = \tau_k[G_{L_N}(\lambda \mathbf{1}_k)]$ and $g_{\ell_N}(\lambda) = \tau_k[G_{\ell_N}(\lambda \mathbf{1}_k)]$, a uniform control of $\|G_{L_N}(\Lambda) - G_{\ell_N}(\Lambda)\|$ will be sufficient to show (1.30). Here $\|\cdot\|$ denotes the operator norm.

Due to the block structure of the matrices under consideration, these quantities are more relevant than the classical Stieltjes transforms. The polynomial L is selfadjoint and of degree one, so we can write $L_N = a_0 \otimes \mathbf{1}_N + S_N + T_N$, $\ell_N = a_0 \otimes \mathbf{1} + s + T_N$, where

$$S_N = \sum_{j=1}^{p} a_j \otimes X_j^{(N)},\ \ s = \sum_{j=1}^{p} a_j \otimes x_j,\ \ T_N = \sum_{j=1}^{q} b_j \otimes Y_j^{(N)},$$

and $a_0, \ldots, a_p, b_1, \ldots, b_q$ are Hermitian matrices in $\mathrm{M}_k(\mathbb{C})$. We also need to introduce the $\mathrm{M}_k(\mathbb{C})$-valued Stieltjes transforms G_{T_N} of T_N:

$$G_{T_N}(\Lambda) = (\mathrm{id}_k \otimes \tau_N)\Big[\Big(\Lambda \otimes \mathbf{1} - T_N\ \Big)^{-1}\ \Big], \tag{1.33}$$

for all Λ in $\mathrm{M}_k(\mathbb{C})$ such that $\mathrm{Im}\ \Lambda$ is positive definite.

The families $\mathbf{x}$ and $\mathbf{Y}_N$ being free in $\mathcal{A}$ and $\mathbf{x}$ being a free semicircular system, the theory of matrix-valued non commutative random variables gives us the following equation relating G_{ℓ_N} and G_{T_N}. It encodes the fundamental property of $\mathcal{R}$-transforms, namely the linearity under free convolution.

Step 4. The subordination property for $\mathrm{M}_k(\mathbb{C})$-valued non commutative random variables: *For all Λ in $\mathrm{M}_k(\mathbb{C})$ such that* Im Λ *is positive definite, one has*

$$G_{\ell_N}(\Lambda) = G_{T_N}\Big(\Lambda - a_0 - \mathcal{R}_s\big(G_{\ell_N}(\Lambda)\ \big)\ \Big), \tag{1.34}$$

where

$$\mathcal{R}_s : M \mapsto \sum_{j=1}^{p} a_j M a_j.$$

We show that the fixed point equation implicitly given by (1.34) is, in a certain sense, stable under perturbations. On the other hand, by the asymptotic freeness of $\mathbf{X}_N$ and $\mathbf{Y}_N$, it is expected that Equation (1.34) is asymptotically satisfied when G_{ℓ_N} is replace by G_{L_N}. Since, in order to apply Step 3, we want an uniform control, we make this connection precise by showing the following:

Step 5. The asymptotic subordination property for random matrices: *For all Λ in $\mathrm{M}_k(\mathbb{C})$ such that* Im Λ *is positive definite, one has*

$$G_{L_N}(\Lambda) = G_{T_N}\Big(\Lambda - a_0 - \mathcal{R}_s\big(G_{L_N}(\Lambda)\ \big)\ \Big) + \Theta_N(\Lambda), \tag{1.35}$$

where $\Theta_N(\Lambda)$ satisfies

$$\|\Theta_N(\Lambda)\| \ \leqslant\ \frac{c}{N^2}\left\|(\mathrm{Im}\ \Lambda)^{-1}\right\|^5$$

for a constant c and with $\|\cdot\|$ denoting the operator norm.

Organization of the proof

We tackle the different points of the proof described above in the following order:

– **Proof of Step 4.** The precise statement of the subordination property for $\mathrm{M}_k(\mathbb{C})$-valued non commutative random variables is contained in Proposition 1.4.2 and Proposition 1.4.3. We highlight in this section the relevance of matrix-valued Stieltjes transforms in a quite general framework.
– **Proof of Step 5.** The asymptotic subordination property for random matrices is stated in Theorem 1.5.1 in a more general situation. The matrices $\mathbf{Y}_N$ can be random, independent of $\mathbf{X}_N$, satisfying a Poincaré inequality, without assumption on their asymptotic properties. This result is based on the Schwinger-Dyson equation and on the Poincaré inequality satisfied by the law of $\mathbf{X}_N$.
– **Proof of Estimate (1.30).** The estimate will follow easily from the two previous items.
– **Proof of Step 2.** This part is based on $\mathcal{C}^*$-algebra techniques. Step 2 is a consequence of a result due to D. Shlyakhtenko which is stated Theorem 1.10.1 of Appendix 1.10. In a previous version of this article, when we did not know this result, we used the subordination property with $L(\mathbf{x}, \mathbf{Y}_N)$ replaced by $L(\mathbf{x}, \mathbf{y})$ and T_N replaced by its limit in law $t = \sum_{j=1}^{q} b_j \otimes y_j$.

Hence we obtained Theorem 1.1.6 with additional assumptions on $\mathbf{Y}_N$, notably a uniform rate of convergence of G_{T_N} to the $\mathrm{M}_k(\mathbb{C})$-valued Stieltjes transform of t.

– **Proof of Step 3.** The method is quite standard once Steps 2 and 4 are established. We use a version due to [GKZ] which is based on the use of local concentration inequalities.

1.4 Proof of Step 4: the subordination property for matrix-valued non commutative random variables

In random matrix theory, a classical method lies in the study of empirical eigenvalue distribution by the analysis of its Stieltjes transform. In many situation, it is shown that this functional satisfies a fixed point equation and a lot of properties of the considered random matrices are deduced from this fact. The purpose of this section is to emphasize that this method can be generalized in the case where the matrices have a macroscopic block structure.

Let $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ be a $\mathcal{C}^*$-probability space with a faithful trace and $k \geqslant 1$ an integer. The algebra $\mathrm{M}_k(\mathbb{C}) \otimes \mathcal{A}$, formed by the $k \times k$ matrices with coefficients in $\mathcal{A}$, inherits the structure of $\mathcal{C}^*$-probability space with trace $(\tau_k \otimes \tau)$ and norm $\|\cdot\|_{\tau_k\otimes\tau}$ defined by (1.9) with $\tau_k \otimes \tau$ instead of τ. We also shall consider the linear functional $(\mathrm{id}_k \otimes \tau)$, called the partial trace.

For any matrix Λ in $\mathrm{M}_k(\mathbb{C})$ we denote $\mathrm{Im}\ \Lambda$ the Hermitian matrix $\frac{1}{2i}(\Lambda - \Lambda^*)$. We write $\mathrm{Im}\ \Lambda > 0$ whenever the matrix $\mathrm{Im}\ \Lambda$ is positive definite and we denote

$$\mathrm{M}_k(\mathbb{C})^+ = \Big\{\Lambda \in \mathrm{M}_k(\mathbb{C}) \ \Big|\ \mathrm{Im}\ \Lambda > 0\Big\}.$$

This lemma will be used throughout this paper. See [HT05, Lemma 3.1] for a proof.

Lemma 1.4.1. Let z in $\mathrm{M}_k(\mathbb{C}) \otimes \mathcal{A}$ be selfadjoint. Then for any $\Lambda \in \mathrm{M}_k(\mathbb{C})^+$, the element $(\Lambda \otimes \mathbf{1} - z)$ is invertible and

$$\Big\|(\Lambda \otimes \mathbf{1} - z)^{-1}\Big\|_{\tau_k\otimes\tau} \leqslant \|(\mathrm{Im}\ \Lambda)^{-1}\|. \tag{1.36}$$

On the right hand side, $\|\cdot\|$ denotes the operator norm in $\mathrm{M}_k(\mathbb{C})$.
For a selfadjoint non commutative random variable z in $\mathrm{M}_k(\mathbb{C}) \otimes \mathcal{A}$, its $\mathrm{M}_k(\mathbb{C})$-valued Stieltjes transform is defined by

$$\begin{array}{cccc} G_z : & \mathrm{M}_k(\mathbb{C})^+ & \to & \mathrm{M}_k(\mathbb{C}) \\ & \Lambda & \mapsto & (\mathrm{id}_k \otimes \tau)\Big[\Big(\Lambda \otimes \mathbf{1} - z\Big)^{-1}\Big]. \end{array}$$

The functional G_z is well defined by Lemma 1.4.1 and satifies

$$\forall \Lambda \in \mathrm{M}_k(\mathbb{C})^+,\ \|G_z(\Lambda)\| \leqslant \|(\mathrm{Im}\ \Lambda)^{-1}\|.$$

It maps $\mathrm{M}_k(\mathbb{C})^+$ to $\mathrm{M}_k(\mathbb{C})^- = \Big\{\Lambda \in \mathrm{M}_k(\mathbb{C}) \ \Big| \ -\Lambda \in \mathrm{M}_k(\mathbb{C})^+\Big\}$ and is analytic (in k^2 complex variables on the open set $\mathrm{M}_k(\mathbb{C})^+ \subset \mathbb{C}^{k^2}$). Moreover, it can be shown (see [Voi95b]) that G_z is univalent on a set of the form $U_\delta = \Big\{\Lambda \in \mathrm{M}_k(\mathbb{C})^+ \ \Big| \ \|\Lambda^{-1}\| < \delta \ \Big\}$ for some $\delta > 0$, and its inverse $G_z^{(-1)}$ in U_δ is analytic on a set of the form $V_\gamma = \Big\{\Lambda \in \mathrm{M}_k(\mathbb{C})^- \ \Big| \ \|\Lambda\| < \gamma\Big\}$ for some $\gamma > 0$.

The amalgamated $\mathcal{R}$-transform over $\mathrm{M}_k(\mathbb{C})$ of $z \in \mathrm{M}_k(\mathbb{C}) \otimes \mathcal{A}$ is the function $\mathcal{R}_z : G_z(U_\delta) \to \mathrm{M}_k(\mathbb{C})$ given by

$$\mathcal{R}_z(\Lambda) = G_z^{(-1)}(\Lambda) - \Lambda^{-1}, \quad \forall \Lambda \in G_z(U_\delta).$$

The following proposition states the fundamental property of the amalgamated $\mathcal{R}$-transform, namely the subordination property, which is the keystone of our proof of Theorem 1.1.6.

Proposition 1.4.2. Let $\mathbf{x} = (x_1, \dots, x_p)$ and $\mathbf{y} = (y_1, \dots, y_q)$ be selfadjoint elements of $\mathcal{A}$ and let $\mathbf{a} = (a_1, \dots, a_p)$ and $\mathbf{b} = (b_1, \dots, b_q)$ be $k \times k$ Hermitian matrices. Define the elements of $\mathrm{M}_k(\mathbb{C}) \otimes \mathcal{A}$

$$s = \sum_{j=1}^{p} a_j \otimes x_j, \quad t = \sum_{j=1}^{q} b_j \otimes y_j.$$

Suppose that the families $\mathbf{x}$ and $\mathbf{y}$ are free. Then one has

1. **Linearity property:** There is a γ such that, in the domain V_γ, one has
$$\mathcal{R}_{s+t} = \mathcal{R}_s + \mathcal{R}_t. \tag{1.37}$$

2. **Subordination property:** There is δ such that, for every Λ in U_δ, one has
$$G_{s+t}(\Lambda) = G_t\Big(\Lambda - \mathcal{R}_s\big(\ G_{s+t}(\Lambda)\ \big)\ \Big). \tag{1.38}$$

3. **Semicircular case:** If $(x_1, \dots, x_p)$ is a free semicircular system, then we get
$$\mathcal{R}_s : \Lambda \mapsto \sum_{j=1}^{p} a_j \Lambda a_j. \tag{1.39}$$

Proof. The linearity property has been shown by Voiculescu in [Voi95b] and the $\mathcal{R}$-transform of s has been computed by Lehner in [Leh99]. We deduce easily the subordination property since by Equation (1.37): there exists $\gamma > 0$ such that for all $\Lambda \in V_\gamma$,

$$G_t^{(-1)}(\Lambda) = G_{s+t}^{(-1)}(\Lambda) - \mathcal{R}_s(\Lambda).$$

Then there exists a $\delta > 0$ such that, with $G_{s+t}(\Lambda)$ instead of Λ in the previous equality,

$$G_t^{(-1)}\Big(G_{s+t}(\Lambda)\ \Big) = \Lambda - \mathcal{R}_s\Big(G_{s+t}(\Lambda)\ \Big).$$

We compose by $G_t^{(-1)}$ to obtain the result. $\square$

The subordination property plays a key role in our problem: it describes G_{s+t} as a fixed point of a simple function involving s and t separately. Such a fixed point is unique and stable under some perturbation, as it is stated in Proposition 1.4.3 below. Remark first that, for $\mathcal{R}_s$ given by (1.39), for any Λ in $\mathrm{M}_k(\mathbb{C})^+$ and M in $\mathrm{M}_k(\mathbb{C})^-$,

$$\mathrm{Im}\,\Big(\Lambda - \mathcal{R}_s(M)\Big) = \mathrm{Im}\,\Lambda - \sum_{j=1}^p a_j \,\mathrm{Im}\, M \, a_j > 0 \tag{1.40}$$

and

$$\Big\| \Big(\mathrm{Im}\,\Big(\Lambda - \mathcal{R}_s(M)\Big) \Big)^{-1} \Big\| \leqslant \| \, (\mathrm{Im}\,\Lambda)^{-1} \|. \tag{1.41}$$

In particular, by analytic continuation, the subordination property holds actually for any $\Lambda \in \mathrm{M}_k(\mathbb{C})^+$ when $\mathbf{x}$ is a free semicircular system.

Proposition 1.4.3. Let s and t be as in Proposition 1.4.2, with $\mathbf{x}$ a free semicircular system.

1. **Uniqueness of the fixed point:** For all $\Lambda \in \mathrm{M}_k(\mathbb{C})^+$ such that

$$\big\|(\mathrm{Im}\,\Lambda)^{-1}\big\| < \sqrt{\sum_{j=1}^p \|a_j\|^2},$$

   the following equation in $G_\Lambda \in \mathrm{M}_k(\mathbb{C})^-$,

$$G_\Lambda = G_t\Big(\Lambda - \mathcal{R}_s(\,G_\Lambda\,)\Big), \tag{1.42}$$

   admits a unique solution G_Λ in $M_k(\mathbb{C})^-$ given by $G_\Lambda = G_{s+t}(\Lambda)$.

2. **Stability under analytic perturbations:** Let $G : \Omega \to \mathrm{M}_k(\mathbb{C})^-$ be an analytic function on a simply connected open subset $\Omega \subset \mathrm{M}_k(\mathbb{C})^+$ containing matrices Λ such that $\|(\mathrm{Im}\,\Lambda)^{-1}\|$ is arbitrary small. Suppose that G satisfies: for all $\Lambda \in \Omega$,

$$G(\Lambda) = G_t\Big(\Lambda - \mathcal{R}_s\Big(G(\Lambda)\Big)\Big) + \Theta(\Lambda), \tag{1.43}$$

   where the function $\Theta : \Omega \to \mathrm{M}_k(\mathbb{C})$ is analytic and satisfies: there exists $\varepsilon > 0$ such that for all Λ in Ω,

$$\kappa(\Lambda) := \|\Theta(\Lambda)\| \; \|(\mathrm{Im}\,\Lambda)^{-1}\| \; \sum_{j=1}^p \|a_j\|^2 < 1 - \varepsilon.$$

   Then one has: $\forall \Lambda \in \Omega$

$$\|G(\Lambda) - G_{s+t}(\Lambda)\| \leqslant \Big(1 + c\, \|(\mathrm{Im}\,\Lambda)^{-1}\|^2\Big) \, \|\Theta(\Lambda)\|, \tag{1.44}$$

   where $c = \frac{1}{\varepsilon}\sum_{j=1}^p \|a_j\|^2$.

Proof. 1. Uniqueness of the fixed point:
Fix $\Lambda \in \mathrm{M}_k(\mathbb{C})^+$ such that

$$\left\|(\mathrm{Im}\ \Lambda)^{-1}\right\| < \sqrt{\sum_{j=1}^{p} \|a_j\|^2}. \tag{1.45}$$

Denote for any M in $\mathrm{M}_k(\mathbb{C})^-$ the matrix $\psi(M) = \Lambda - \mathcal{R}_s(M)$, which is in $\mathrm{M}_k(\mathbb{C})^+$ by (1.40). We show that the function

$$\Phi_\Lambda : M \to G_t\Big(\psi(M)\ \Big)$$

is a contraction on $\mathrm{M}_k(\mathbb{C})^-$. Remark that Φ_Λ maps $\mathrm{M}_k(\mathbb{C})^-$ into $\mathrm{M}_k(\mathbb{C})^-$. Moreover for all $M, \tilde{M}$ in $\mathrm{M}_k(\mathbb{C})^-$,

$$\begin{aligned}
&\|\Phi_\Lambda(M) - \Phi_\Lambda(\tilde{M})\| \\
&= \left\|(id_k \otimes \tau)\left[\Big(\psi(M) \otimes \mathbf{1} - t\Big)^{-1} - \Big(\psi(\tilde{M}) \otimes \mathbf{1} - t\Big)^{-1}\right]\right\| \\
&= \left\|(id_k \otimes \tau)\left[\Big(\psi(M) \otimes \mathbf{1} - t\Big)^{-1}\Big(\sum_{j=1}^{p} a_j(M - \tilde{M})a_j\Big) \otimes \mathbf{1}_N \right.\right. \\
&\qquad\qquad\qquad \left.\left. \times \Big(\psi(\tilde{M}) \otimes \mathbf{1} - t\Big)^{-1}\right]\right\| \\
&\leqslant \left\|\Big(\mathrm{Im}\ \big(\psi(M) \otimes \mathbf{1} - t\big)\Big)^{-1}\right\| \left\|\Big(\mathrm{Im}\ \big(\psi(\tilde{M}) \otimes \mathbf{1} - t\big)\Big)^{-1}\right\| \\
&\qquad\qquad\qquad \times \sum_{j=1}^{p} \|a_j\|^2 \left\|M - \tilde{M}\right\| \\
&\leqslant \left\|(\mathrm{Im}\ \Lambda)^{-1}\right\|^2 \sum_{j=1}^{p} \|a_j\|^2\ \|M - \tilde{M}\|.
\end{aligned}$$

Hence the function Φ_Λ is a contraction and by Picard's theorem the fixed point equation $M = \Phi_\Lambda(M)$ admits a unique solution M_Λ on the closed set of $k \times k$ matrices whose imaginary part is non positive semi-definite, which is necessarily G_{s+t} by the subordination property.

2. Stability under analytic perturbations:
We set $\tilde{G} : \Omega \to \mathrm{M}_k(\mathbb{C})^-$ given by: for all $\Lambda \in \Omega$,

$$\tilde{G}(\Lambda) = G(\Lambda) - \Theta(\Lambda) = G_t\Big(\Lambda - \mathcal{R}_s\Big(G(\Lambda)\ \Big)\ \Big).$$

We set $\tilde{\Lambda} : \Omega \to \mathrm{M}_k(\mathbb{C})$ given by: for all $\Lambda \in \Omega$

$$\tilde{\Lambda}(\Lambda) = \Lambda - \mathcal{R}_s(\Theta(\Lambda)) = \Lambda - \mathcal{R}_s\Big(G(\Lambda)\ \Big) + \mathcal{R}_s\Big(\tilde{G}(\Lambda)\ \Big).$$

In the following, we use $\tilde{\Lambda}$ as a shortcut for $\tilde{\Lambda}(\Lambda)$. One has $\tilde{\Lambda} - \mathcal{R}_s\Big(\tilde{G}(\Lambda)\ \Big) = \Lambda - \mathcal{R}_s\Big(G(\Lambda)\ \Big)$ which is in $\mathrm{M}_k(\mathbb{C})^+$ by (1.40). Hence we have: for all $\Lambda \in \Omega$,

$$\tilde{G}(\Lambda) \ = \ G_t\Big(\tilde{\Lambda} - \mathcal{R}_s\Big(\tilde{G}(\Lambda)\ \Big)\ \Big). \tag{1.46}$$

We want to estimate $\|(\text{Im } \tilde\Lambda)^{-1}\|$ in terms of $\|(\text{Im } \Lambda)^{-1}\|$. For all Λ in Ω, we use the definition of $\tilde\Lambda$ and we write:

$$\text{Im } \tilde\Lambda = \text{Im } \Lambda \Big(\mathbf{1}_k - (\text{Im } \Lambda)^{-1} \mathcal{R}_s\big(\Theta(\Lambda)\ \big) \Big).$$

Remark that

$$\|(\text{Im } \Lambda)^{-1}\mathcal{R}_s\big(\Theta(\Lambda)\ \big)\| \leqslant \kappa(\Lambda) = \|\Theta(\Lambda)\| \ \|(\text{Im } \Lambda)^{-1}\| \ \sum_{j=1}^p \|a_j\|^2 < 1 - \varepsilon$$

by assumption. Then $\text{Im } \tilde\Lambda$ is invertible and one has

$$(\text{Im } \tilde\Lambda)^{-1} = \sum_{\ell \geqslant 0} \Big(\ (\text{Im } \Lambda)^{-1}\mathcal{R}_s\big(\Theta(\Lambda)\ \big) \Big)^\ell \ (\text{Im } \Lambda)^{-1}.$$

We then obtain the following estimate

$$\begin{aligned} \|(\text{Im } \tilde\Lambda)^{-1}\| &\leqslant \Big\| \sum_{\ell\geqslant 0} \Big(\ (\text{Im } \Lambda)^{-1}\mathcal{R}_s\big(\Theta(\Lambda)\ \big) \Big)^\ell \ (\text{Im } \Lambda)^{-1} \Big\| \\ &\leqslant \frac{1}{1-\kappa(\Lambda)} \|(\text{Im } \Lambda)^{-1}\| < \frac{1}{\varepsilon}\|(\text{Im } \Lambda)^{-1}\|. \end{aligned}$$

By uniqueness of the fixed point and by (1.46), for all $\Lambda \in \Omega$ such that $\|(\text{Im } \Lambda)^{-1}\| < \varepsilon \sqrt{\sum_{j=1}^p \|a_j\|^2}$, one has $\tilde G(\Lambda) = G_{s+t}(\tilde\Lambda)$ (such matrices Λ exist by assumption on Ω). But the functions are analytic (in k^2 complex variables) so that the equality extends to Ω. Then for all $\Lambda \in \Omega$,

$$\|G(\Lambda) - G_{s+t}(\Lambda)\| \ \leqslant \ \|G(\Lambda) - \tilde G(\Lambda)\| + \|G_{s+t}(\tilde\Lambda) - G_{s+t}(\Lambda)\|.$$

For the first term we have by definition of $\tilde G$ that $\|G(\Lambda) - \tilde G(\Lambda)\| \leqslant \|\Theta(\Lambda)\|$. On the other hand, one has

$$\begin{aligned} &\|G_{s+t}(\Lambda) - G_{s+t}(\tilde\Lambda)\| \\ &\quad = \Big\| (\text{id}_k \otimes \tau)\big[(\Lambda \otimes \mathbf{1} - s - t)^{-1} - (\tilde\Lambda \otimes \mathbf{1} - s - t)^{-1} \ \big] \Big\| \\ &\quad = \Big\| (\text{id}_k \otimes \tau)\big[(\Lambda \otimes \mathbf{1} - s - t)^{-1} (\tilde\Lambda \otimes \mathbf{1} - \Lambda \otimes \mathbf{1}) (\tilde\Lambda \otimes \mathbf{1} - s - t)^{-1} \ \big] \Big\| \\ &\quad \leqslant \|(\Lambda \otimes \mathbf{1} - s - t)^{-1}\| \ \|\tilde\Lambda - \Lambda\| \ \|(\tilde\Lambda \otimes \mathbf{1} - s - t)^{-1}\| \\ &\quad \leqslant \frac{1}{\varepsilon} \ \Big\| \mathcal{R}_s\big(\tilde G(\Lambda)\ \big) - \mathcal{R}_s\big(G(\Lambda)\ \big) \ \Big\| \ \|(\text{Im } \Lambda)^{-1}\|^2 \hat{\text{E}} \\ &\quad \leqslant \frac{1}{\varepsilon} \ \sum_{j=1}^p \|a_j\|^2 \ \|(\text{Im } \Lambda)^{-1}\|^2 \ \|\Theta(\Lambda)\|. \end{aligned}$$

We then obtain as expected

$$\|G(\Lambda) - G_{s+t}(\Lambda)\| \leqslant \Big(1 + \frac{1}{\varepsilon} \ \sum_{j=1}^p \|a_j\|^2 \ \|(\text{Im } \Lambda)^{-1}\|^2 \Big) \ \|\Theta(\Lambda)\|.$$

□

1.5 Proof of Step 5: the asymptotic subordination property for random matrices

The purpose of this section is to prove Theorem 1.5.1 below, where it is stated that, for N fixed, the matrix-valued Stieltjes transforms of certain random matrices satisfy an asymptotic subordination property i.e. an equation as in (1.43). This result is independent with the previous part and does not involve the language of free probability.

Let $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ be a family of independent, normalized $N \times N$ matrices of the GUE and $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ be a family of $N \times N$ random Hermitian matrices, independent of $\mathbf{X}_N$. We fix an integer $k \geqslant 1$ and Hermitian matrices $a_0, \dots, a_p, b_1, \dots, b_q \in \mathrm{M}_k(\mathbb{C})$. We set S_N and T_N the $kN \times kN$ block matrices

$$S_N = \sum_{j=1}^{p} a_j \otimes X_j^{(N)}, \qquad T_N = \sum_{j=1}^{q} b_j \otimes Y_j^{(N)}.$$

Define the $\mathrm{M}_k(\mathbb{C})$-valued Stieltjes transforms of $S_N + T_N$ and T_N: for all $\Lambda \in \mathrm{M}_k(\mathbb{C})^+ = \Big\{\Lambda \in \mathrm{M}_k(\mathbb{C}) \ \Big| \ \mathrm{Im}\ \Lambda > 0\Big\}$,

$$\begin{aligned} G_{S_N+T_N}(\Lambda) &= \mathbb{E}\Big[(\mathrm{id}_k \otimes \tau_N)\Big[\Big(\Lambda \otimes \mathbf{1}_N - S_N - T_N\Big)^{-1}\Big]\ \Big], \\ G_{T_N}(\Lambda) &= \mathbb{E}\Big[(\mathrm{id}_k \otimes \tau_N)\Big[\Big(\Lambda \otimes \mathbf{1}_N - T_N\Big)^{-1}\Big]\ \Big]. \end{aligned}$$

We denote by $\mathcal{R}_s$ the functional

$$\begin{aligned} \mathcal{R}_s: \ \mathrm{M}_k(\mathbb{C}) &\to \mathrm{M}_k(\mathbb{C}) \\ M &\mapsto \sum_{j=1}^{p} a_j\ M\ a_j. \end{aligned}$$

Theorem 1.5.1 (Asymptotic subordination property)**.** Assume that there exists $\sigma \geqslant 1$ such that the joint law of the entries of the matrices $\mathbf{Y}_N$ satisfies a Poincaré inequality with constant σ/N, i.e. for any $f : \mathbb{R}^{2qN^2} \to \mathbb{C}$ function of the entries of q matrices, of class $\mathcal{C}^1$ and such that $\mathbb{E}\Big[\ |f(\mathbf{Y}_N)|^2\ \Big] < \infty$, one has

$$\mathbb{V}\mathrm{ar}\Big(f(\mathbf{Y}_N)\ \Big) \leqslant \frac{\sigma}{N}\ \mathbb{E}\Big[\|\nabla f(\mathbf{Y}_N)\|^2\Big], \tag{1.47}$$

where ∇f denotes the gradient of f, $\mathbb{V}\mathrm{ar}$ denotes the variance, $\mathbb{V}\mathrm{ar}(\ x\ ) = \mathbb{E}\Big[\ \Big|\ x - \mathbb{E}[\ x\ ]\ \Big|^2\Big]$.
Then for any $\Lambda \in \mathrm{M}_k(\mathbb{C})^+$, the Stieltjes transforms $G_{S_N+T_N}$ and G_{T_N} satisfy

$$G_{S_N+T_N}(\Lambda) = G_{T_N}\Big(\Lambda - \mathcal{R}_s\Big(G_{S_N+T_N}(\Lambda)\ \Big)\ \Big) + \Theta_N(\Lambda), \tag{1.48}$$

where Θ is analytic $\mathrm{M}_k(\mathbb{C})^+ \to \mathrm{M}_k(\mathbb{C})$ and satisfies

$$\|\Theta_N(\Lambda)\| \leqslant \frac{c}{N^2} \left\|(\mathrm{Im}\ \Lambda)^{-1}\right\|^5,$$

with $c = 2k^{9/2}\sigma \sum_{j=1}^p \|a_j\|^2 \Big(\sum_{j=1}^p \|a_j\| + \sum_{j=1}^q \|b_j\| \Big)^2$, $\|\cdot\|$ denoting the operator norm in $\mathrm{M}_k(\mathbb{C})$.

The proof of Theorem 1.5.1 is carried out in two steps.

- In Section 1.5.1 we state a mean Schwinger-Dyson equation for random Stieltjes transforms (Proposition 1.5.2).
- In Section 1.5.2 we deduce from Proposition 1.5.2 a Schwinger-Dyson equation for mean Stieltjes transforms (Proposition 1.5.3).

Theorem 1.5.1 is a direct consequence of Proposition 1.5.3 as it is shown in Section 1.5.3.

1.5.1 Mean Schwinger-Dyson equation for random Stieltjes transforms

For Λ, Γ in $\mathrm{M}_k(\mathbb{C})^+$, define the elements of $\mathrm{M}_k(\mathbb{C}) \otimes \mathrm{M}_N(\mathbb{C})$

$$\begin{aligned} h_{S_N+T_N}(\Lambda) &= (\Lambda \otimes \mathbf{1}_N - S_N - T_N)^{-1}, \\ h_{T_N}(\Gamma) &= (\Gamma \otimes \mathbf{1}_N - T_N)^{-1}, \end{aligned}$$

and $H_{S_N+T_N}(\Lambda) = (\mathrm{id}_k \otimes \tau_N)\Big[h_{S_N+T_N}(\Lambda)\Big]$, $H_{T_N}(\Lambda) = (\mathrm{id}_k \otimes \tau_N)\Big[h_{T_N}(\Lambda)\Big]$.

Proposition 1.5.2 (Mean Schwinger-Dyson equation for random Stieltjes transforms)**.** For all $\Lambda, \Gamma \in \mathrm{M}_k(\mathbb{C})^+$ we have

$$\begin{aligned} 0 = \mathbb{E}\Bigg[& H_{S_N+T_N}(\Lambda) - H_{T_N}(\Gamma) \qquad (1.49) \\ & -(\mathrm{id}_k \otimes \tau_N)\Big[h_{T_N}(\Gamma)\Big(\mathcal{R}_s\big(H_{S_N+T_N}(\Lambda)\ \big) - \Lambda + \Gamma\Big) \otimes \mathbf{1}_N\ h_{S_N+T_N}(\Lambda)\Big]\Bigg]. \end{aligned}$$

The result is a consequence of integration by parts for Gaussian densities and of the formula for the differentiation of the inverse of a matrix. If $(g_1, \dots, g_N)$ are independent identically distributed centered real Gaussian variables with variance σ^2 and $F : \mathbb{R}^N \to \mathbb{C}$ a differentiable map such that F and its partial derivatives are polynomially bounded, one has for $i = 1, \dots, N$

$$\mathbb{E}\Big[\ g_i\ F(g_1, \dots, g_N)\ \Big] = \sigma^2 \mathbb{E}\Big[\frac{\partial F}{\partial x_i}(g_1, \dots, g_N)\ \Big].$$

This induces an analogue formula for independent matrices of the GUE, called the Schwinger-Dyson equation, where the Hermitian symmetry of the matrices plays a key role. For instance, if P is a monomial in p non commutative indeterminates, one has for $i = 1, \dots, p$,

$$\mathbb{E}\Big[\tau_N\Big[\ X_i^{(N)}\ P(\mathbf{X}_N)\ \Big]\ \Big] = \sum_{P = L x_i R} \mathbb{E}\Big[\tau_N\Big[L(\mathbf{X}_N)\ \Big]\ \tau_N\Big[R(\mathbf{X}_N)\ \Big]\Big],$$

the sum over all decompositions $P = Lx_iR$ for L and R monomials being viewed as the partial derivative.

This formula has an analogue for analytical maps instead of polynomials. The case of the function $\mathbf{X}_N \mapsto (\Lambda \otimes \mathbf{1}_N - S_N)^{-1}$ is investigated in details in [HT05, Formula (3.9)], our proof is obtained by minor modifications.

Proof. Denote by $(\epsilon_{m,n})_{m,n=1,\dots,N}$ the canonical basis of $\mathrm{M}_N(\mathbb{C})$. By [HT05, Formula (3.9)] with minor modification, we get the following: for all Λ, Γ in $\mathrm{M}_k(\mathbb{C})^+$ and $j = 1, \dots, p$,

$$\begin{aligned}
&\mathbb{E}\Big[(\mathbf{1}_k \otimes X_j^{(N)})(\Lambda \otimes \mathbf{1}_N - S_N - T_N)^{-1} \,\Big|\, T_N\Big] \\
&\quad = \mathbb{E}\Big[\frac{1}{N}\sum_{m,n=1}^{N} (\mathbf{1}_k \otimes \epsilon_{m,n})(\Lambda \otimes \mathbf{1}_N - S_N - T_N)^{-1} \\
&\qquad \times (a_j \otimes \epsilon_{n,m})(\Lambda \otimes \mathbf{1}_N - S_N - T_N)^{-1} \,\Big|\, T_N\Big].
\end{aligned}$$

In these equations, $\mathbb{E}[\cdot|T_N]$ stands for the conditional expectation with respect to T_N. Furthermore, for any M in $\mathrm{M}_k(\mathbb{C}) \otimes \mathrm{M}_N(\mathbb{C})$, one has

$$\frac{1}{N}\sum_{m,n=1}^{N} (\mathbf{1}_k \otimes \epsilon_{m,n})\ M\ (\mathbf{1}_k \otimes \epsilon_{n,m}) = (\mathrm{id}_k \otimes \tau_N)[\ M\ ] \otimes \mathbf{1}_N.$$

Indeed the formula is clear if M is of the form $M = \tilde{M} \otimes \epsilon_{u,v}$ and extends by linearity. In particular, with $M = (\Lambda \otimes \mathbf{1}_N - S_N - T_N)^{-1}(a_j \otimes \mathbf{1}_N)$, we obtain that: for all Λ, Γ in $\mathrm{M}_k(\mathbb{C})^+$ and $j = 1, \dots, p$,

$$\begin{aligned}
&\mathbb{E}\Big[(a_j \otimes X_j^{(N)})(\Lambda \otimes \mathbf{1}_N - S_N - T_N)^{-1} \,\Big|\, T_N\Big] \\
&\quad = \mathbb{E}\Big[(a_j \otimes \mathbf{1}_N)\Big((\mathrm{id}_k \otimes \tau_N)\Big[(\Lambda \otimes \mathbf{1}_N - S_N - T_N)^{-1}\Big]a_j \otimes \mathbf{1}_N\Big) \\
&\qquad \times (\Lambda \otimes \mathbf{1}_N - S_N - T_N)^{-1} \,\Big|\, T_N\Big] \\
&\quad = \mathbb{E}\Big[\Big(a_j H_{S_N+T_N} a_j \otimes \mathbf{1}_N\Big)\ h_{S_N+T_N} \,\Big|\, T_N\Big].
\end{aligned}$$

Recall that $S_N = \sum_{j=1}^{p} a_j \otimes X_j^{(N)}$ and $\mathcal{R}_s : M \mapsto \sum_{j=1}^{p} a_j M a_j$, so that for all Λ, Γ

in $\mathrm{M}_k(\mathbb{C})^+$, one has

$$\begin{aligned}
&\mathbb{E}\Big[(\Gamma\otimes \mathbf{1}_N - T_N)^{-1}\ S_N\ (\Lambda\otimes \mathbf{1}_N - S_N - T_N)^{-1}\Big] \\
&\quad = \mathbb{E}\Big[\ (\Gamma\otimes \mathbf{1}_N - T_N)^{-1}\ \sum_{j=1}^p \mathbb{E}\Big[(a_j\otimes X_j^{(N)}) \qquad (1.50)\\
&\qquad \times(\Lambda\otimes \mathbf{1}_N - S_N - T_N)^{-1}\ \Big|\ T_N\Big]\ \Big] \\
&\quad = \mathbb{E}\Big[\ h_{T_N}(\Gamma)\ \mathbb{E}\Big[\Big(\sum_{j=1}^p a_j H_{S_N+T_N}(\Lambda) a_j\otimes \mathbf{1}_N\Big)\ h_{S_N+T_N}(\Lambda)\ \Big|\ T_N\Big] \\
&\quad = \mathbb{E}\Big[h_{T_N}(\Gamma)\ \Big(\mathcal{R}_s\Big(H_{S_N+T_N}(\Lambda)\ \Big)\otimes \mathbf{1}_N\Big)\ h_{S_N+T_N}(\Lambda)\Big]. \qquad (1.51)
\end{aligned}$$

We take the partial trace in Equation (1.51) to obtain:

$$\begin{aligned}
&\mathbb{E}\Big[(\mathrm{id}_k\otimes\tau_N)\Big[h_{T_N}(\Gamma)\ S_N\ h_{S_N+T_N}(\Lambda)\Big]\Big] \qquad (1.52)\\
&\quad = \ \mathbb{E}\Big[(\mathrm{id}_k\otimes\tau_N)\Big[h_{T_N}(\Gamma)\ \Big(\mathcal{R}_s\Big(H_{S_N+T_N}(\Lambda)\ \Big)\otimes \mathbf{1}_N\Big)\ h_{S_N+T_N}(\Lambda)\Big]\Big].
\end{aligned}$$

We now rewrite S_N as follow:

$$S_N = (\Lambda-\Gamma)\otimes \mathbf{1}_N + (\Gamma\otimes \mathbf{1}_N - T_N) - (\Lambda\otimes \mathbf{1}_N - S_N - T_N).$$

Re-injecting this expression in the left hand side of Equation (1.52), one gets Equation (1.49):

$$\begin{aligned}
&\mathbb{E}\Big[(\mathrm{id}_k\otimes\tau_N)\Big[h_{T_N}(\Gamma)\ \Big(\mathcal{R}_s\Big(H_{S_N+T_N}(\Lambda)\ \Big)\otimes \mathbf{1}_N\Big)\ h_{S_N+T_N}(\Lambda)\Big]\Big] \\
&\quad = \ \mathbb{E}\Big[(\mathrm{id}_k\otimes\tau_N)\Big[h_{T_N}(\Gamma)\ (\Lambda-\Gamma)\otimes \mathbf{1}_N h_{S_N+T_N}(\Lambda) \\
&\qquad + h_{S_N+T_N}(\Lambda) - h_{T_N}(\Gamma)\Big]\Big] \\
&\quad = \ \mathbb{E}\Big[(\mathrm{id}_k\otimes\tau_N)\Big[\ h_{T_N}(\Gamma)\ \Big((\Lambda-\Gamma)\otimes \mathbf{1}_N\Big)\ h_{S_N+T_N}(\Lambda)\Big] \\
&\qquad + H_{S_N+T_N}(\Lambda)\ -\ H_{T_N}(\Gamma)\Big].
\end{aligned}$$

□

1.5.2 Schwinger-Dyson equation for mean Stieltjes transforms

We use the concentration properties of the law of $(\mathbf{X}_N, \mathbf{Y}_N)$ to get from Equation (1.49) a relation between $G_{S_N+T_N}$ and G_{T_N}. We define the centered version of $H_{S_N+T_N}$ by: for all Λ in $\mathrm{M}_k(\mathbb{C})^+$,

$$K_{S_N+T_N}(\Lambda) \ = \ H_{S_N+T_N}(\Lambda) - G_{S_N+T_N}(\Lambda),\ \text{in } \mathrm{M}_k(\mathbb{C}). \qquad (1.53)$$

We introduce the random linear map

$$l_{N,\Lambda,\Gamma} : \begin{array}{ccc} \mathrm{M}_k(\mathbb{C}) \otimes \mathrm{M}_N(\mathbb{C}) & \to & \mathrm{M}_k(\mathbb{C}) \otimes \mathrm{M}_N(\mathbb{C}) \\ M & \mapsto & h_{T_N}(\Gamma) \ M \ h_{S_N+T_N}(\Lambda) \end{array} \tag{1.54}$$

and its mean

$$L_{N,\Lambda,\Gamma} : M \mapsto \mathbb{E}\Big[l_{N,\Lambda,\Gamma}(M) \ \Big]. \tag{1.55}$$

Remark that if M is a random matrix, then

$$L_{N,\Lambda,\Gamma}(M) = \mathbb{E}\Big[h_{\tilde{T}_N}(\Gamma) \ M \ h_{\tilde{S}_N+\tilde{T}_N}(\Lambda) \Big| M \Big],$$

where $(\tilde{S}_N + \tilde{T}_N)$ is an independent copy of $(S_N + T_N)$ independent of M.

Proposition 1.5.3 (Schwinger-Dyson equation for mean Stieltjes transforms)**.** For all Λ, Γ in $\mathrm{M}_k(\mathbb{C})^+$, one has

$$\begin{aligned} & G_{S_N+T_N}(\Lambda) - G_{T_N}(\Gamma) \\ & -(\mathrm{id}_k \otimes \tau_N)\Big[L_{N,\Lambda,\Gamma}\Big(\Big(\mathcal{R}_s \hat{\mathrm{E}}\Big(G_{S_N+T_N}(\Lambda) \ \Big) - \Lambda + \Gamma \Big) \otimes \mathbf{1}_N \Big) \ \Big] = \Theta_N(\Lambda, \Gamma), \end{aligned} \tag{1.56}$$

where

$$\Theta_N(\Lambda,\Gamma) = \mathbb{E}\Big[(\mathrm{id}_k \otimes \tau_N)\Big[\ (l_{N,\Lambda,\Gamma} - L_{N,\Lambda,\Gamma}) \Big(\mathcal{R}_s\Big(K_{S_N+T_N}(\Lambda) \ \Big) \otimes \mathbf{1}_N \Big) \ \Big] \Big] \tag{1.57}$$

is controlled in operator norm by the following estimate:

$$\|\Theta_N(\Lambda,\Gamma)\| \leqslant \frac{c}{N^2} \left\| (\mathrm{Im}\ \Gamma)^{-1} \right\| \left\| (\mathrm{Im}\ \Lambda)^{-1} \right\|^3 \ \Big(\|(\mathrm{Im}\ \Gamma)^{-1}\| + \|(\mathrm{Im}\ \Lambda)^{-1}\| \Big), \tag{1.58}$$

with $c = k^{9/2} \sigma \sum_{j=1}^p \|a_j\|^2 \Big(\sum_{j=1}^p \|a_j\| + \sum_{j=1}^q \|b_j\| \Big)^2$.

Proof of Proposition 1.5.3. We first expand $\Theta_N(\Lambda, \Gamma)$: for all Λ, Γ in $\mathrm{M}_k(\mathbb{C})^+$, we have

$$\begin{aligned} \Theta_N(\Lambda,\Gamma) \ := \ & \mathbb{E}\Big[(\mathrm{id}_k \otimes \tau_N)\Big[\ (l_{N,\Lambda,\Gamma} - L_{N,\Lambda,\Gamma}) \\ & \quad \times \Big(\mathcal{R}_s\Big(H_{S_N+T_N}(\Lambda) - G_{S_N+T_N}(\Lambda) \ \Big) \otimes \mathbf{1}_N \Big) \ \Big] \Big] \\ = \ & \mathbb{E}\Big[(\mathrm{id}_k \otimes \tau_N)\Big[l_{N,\Lambda,\Gamma}\Big(\mathcal{R}_s\Big(H_{S_N+T_N}(\Lambda) \ \Big) \otimes \mathbf{1}_N \Big) \ \Big] \Big] \\ & -(\mathrm{id}_k \otimes \tau_N)\Big[L_{N,\Lambda,\Gamma}\Big(\mathcal{R}_s\Big(G_{S_N+T_N}(\Lambda) \ \Big) \otimes \mathbf{1}_N \Big) \ \Big]. \end{aligned}$$

By Equation (1.49), we get the following:

$$\begin{aligned}
&\mathbb{E}\left[(\mathrm{id}_k\otimes\tau_N)\left[l_{N,\Lambda,\Gamma}\Big(\mathcal{R}_s\big(H_{S_N+T_N}(\Lambda)\ \big)\otimes\mathbf{1}_N\Big)\right]\right]\\
&= \ \mathbb{E}\left[(\mathrm{id}_k\otimes\tau_N)\left[l_{N,\Lambda,\Gamma}\Big(\ (\Lambda-\Gamma)\otimes\mathbf{1}_N\Big)\ \right]-H_{T_N}(\Gamma)+H_{S_N+T_N}(\Lambda)\right]\\
&= \ (\mathrm{id}_k\otimes\tau_N)\left[L_{N,\Lambda,\Gamma}\Big(\ (\Lambda-\Gamma)\otimes\mathbf{1}_N\Big)\ \right]-G_{T_N}(\Gamma)+G_{S_N+T_N}(\Lambda),
\end{aligned}$$

which gives Equation (1.56).

We use the Poincaré inequality to control the operator norm of Θ_N: if $(g_1,\dots,g_K)$ are independent identically distributed centered real Gaussian variables with variance v^2 and F is a differentiable map $\mathbb{R}^K\to\mathbb{C}$ such that F and its partial derivatives are polynomially bounded, then (see [Che82a, Theorem 2.1])

$$\mathbb{V}\mathrm{ar}\Big(F(g_1,\dots,g_K)\ \Big)\leqslant v^2\mathbb{E}\Big[\ \|\nabla F(g_1,\dots,g_K)\ \|^2\ \Big].$$

The Poincaré inequality is compatible with tensor product and then such a formula is still valid when F is a function of the matrices $\mathbf{X_N}$ and $\mathbf{Y_N}$ with $v^2=\frac{\sigma}{N}$.

We will often deal with matrices of size $k\times k$. Since the integer k is fixed, we can use intensively the equivalence of norms, the constants appearing will not modify the order of convergence. For any integer K, we denote the Euclidean norm of a $K\times K$ matrix $A=(a_{m,n})_{1\leqslant m,n\leqslant K}$ by

$$\|A\|_e=\sqrt{\sum_{m,n=1}^{K}|a_{m,n}|^2},$$

and its infinity norm by

$$\|A\|_\infty=\max_{m,n=1,\dots,K}|a_{m,n}|.$$

Recall that if A,B are $K\times K$ matrices we have the following inequalities

$$\|A\|\leqslant\|A\|_e\leqslant\sqrt{K}\|A\|, \tag{1.59}$$

$$\|A\|\leqslant\sqrt{K}\|A\|_\infty\leqslant\sqrt{K}\|A\|_e, \tag{1.60}$$

$$\|AB\|\leqslant\|A\|_e\ \|B\|. \tag{1.61}$$

When A is in $\mathrm{M}_k(\mathbb{C})\otimes\mathrm{M}_N(\mathbb{C})$, its Euclidean norm is defined by considering A as a $kN\times kN$ matrix. In the following we will write an element Z of $\mathrm{M}_k(\mathbb{C})\otimes\mathrm{M}_N(\mathbb{C})$

$$\begin{aligned}
Z &= \sum_{m,n=1}^{N}\sum_{u,v=1}^{k}Z_{u,v}^{m,n}\ \epsilon_{u,v}\otimes\epsilon_{m,n}=\sum_{m,n=1}^{N}Z^{(m,n)}\otimes\epsilon_{m,n} \qquad (1.62)\\
&= \sum_{u,v=1}^{k}\epsilon_{u,v}\otimes Z_{(u,v)},
\end{aligned}$$

where for $m, n = 1, \ldots, N$ and $u, v = 1, \ldots, k$, $Z_{u,v}^{m,n}$ is a complex number, $Z^{(m,n)}$ is a $k \times k$ matrix, and $Z_{(u,v)}$ is a $N \times N$ matrix; we use the same notation for the canonical bases of $\mathrm{M}_k(\mathbb{C})$ and $\mathrm{M}_N(\mathbb{C})$.
We fix Λ, Γ in $\mathrm{M}_k(\mathbb{C})^+$ until the end of this proof and we use for convenience the following notations:

$$\begin{aligned} M_N &= \mathcal{R}_s\Big(K_{S_N+T_N}(\Lambda)\Big) \\ h_N^{(1)} &= h_{S_N+T_N}(\Lambda) \\ h_N^{(2)} &= h_{T_N}(\Gamma) \\ l_N &= l_{N,\Lambda,\Gamma} \\ L_N &= L_{N,\Lambda,\Gamma}. \end{aligned}$$

We consider $(\tilde{h}_N^{(1)}, \tilde{h}_N^{(2)})$ an independent copy of $(h_N^{(1)}, h_N^{(2)})$, independent of $\mathbf{X}_N$ and $\mathbf{Y}_N$ (and hence of all the random variables considered). Recall that by definitions (1.54) and (1.55): for all Λ, Γ in $\mathrm{M}_k(\mathbb{C})^+$, we have

$$\begin{aligned} l_N &: A \in \mathrm{M}_k(\mathbb{C}) \mapsto h_N^{(2)}\ A\ h_N^{(1)} \in \mathrm{M}_k(\mathbb{C}), \\ L_N &: A \in \mathrm{M}_k(\mathbb{C}) \mapsto \mathbb{E}\Big[l_N(A)\Big] \in \mathrm{M}_k(\mathbb{C}). \end{aligned}$$

With the notations of (1.62) we have

$$\begin{aligned} &(\mathrm{id}_k \otimes \tau_N)\Big[(l_N - L_N)\ (M_N \otimes \mathbf{1}_N)\Big] \\ &\quad = (\mathrm{id}_k \otimes \tau_N)\Big[h_N^{(2)}\ (M_N \otimes \mathbf{1}_N)\ h_N^{(1)}\Big] \\ &\qquad -\mathbb{E}\Big[(\mathrm{id}_k \otimes \tau_N)\Big[\tilde{h}_N^{(2)}\ (M_N \otimes \mathbf{1}_N)\ \tilde{h}_N^{(1)}\Big]\ \Big|\ M_N\Big] \\ &\quad = \frac{1}{N}\sum_{m,n=1}^N \Bigg[\Big(h_N^{(2)}\Big)^{(m,n)} M_N\ \Big(h_N^{(1)}\Big)^{(n,m)} \\ &\qquad -\mathbb{E}\Big[\Big(\tilde{h}_N^{(2)}\Big)^{(m,n)} M_N\ \Big(\tilde{h}_N^{(1)}\Big)^{(n,m)}\ \Big|\ M_N\Big]\Bigg]. \end{aligned}$$

To estimate the operator norm of Θ_N we use the domination by the infinity norm

(1.60) in order to split the contributions due to M_N and due to $l_N - L_N$: we get

$$\begin{aligned}
\|\Theta_N(\Lambda,\Gamma)\| &= \Big\| \mathbb{E}\Big[(\mathrm{id}_k \otimes \tau_N)\,[(l_N - L_N)\ (M_N \otimes \mathbf{1}_N)] \Big] \Big\| \\
&\leqslant \sqrt{k} \Bigg\| \mathbb{E}\Bigg[\frac{1}{N} \sum_{m,n=1}^N \Big(h_N^{(2)}\Big)^{(m,n)} M_N \ \Big(h_N^{(1)}\Big)^{(n,m)} \\
&\qquad\qquad - \mathbb{E}\Big[\Big(\tilde h_N^{(2)}\Big)^{(m,n)} M_N \ \Big(\tilde h_N^{(1)}\Big)^{(n,m)} \ \Big| \ M_N \Big] \Bigg] \Bigg\|_\infty \\
&\leqslant k^{5/2} \max_{\substack{1\leqslant u,v \leqslant k \\ 1 \leqslant u',v' \leqslant k}} \Bigg| \mathbb{E}\Bigg[(M_N)_{u',v'} \times \frac{1}{N} \sum_{m,n=1}^N \Big(h_N^{(2)}\Big)^{m,n}_{u,u'} \Big(h_N^{(1)}\Big)^{n,m}_{v',v} \\
&\qquad\qquad - \mathbb{E}\Big[\Big(h_N^{(2)}\Big)^{m,n}_{u,u'} \Big(h_N^{(1)}\Big)^{n,m}_{v',v} \Big] \Bigg] \Bigg| \\
&\leqslant k^{5/2} \max_{u,v,u',v'} \mathbb{E}\Bigg[|(M_N)_{u',v'}| \times \Big| \tau_N \Big[\Big(h_N^{(1,2)}\Big)_{\substack{u,v\\u',v'}} \Big] - \mathbb{E}\Big[\tau_N \Big[\Big(h_N^{(1,2)}\Big)_{\substack{u,v\\u',v'}} \Big] \Big] \Big| \Bigg] \\
&\leqslant k^{5/2} \max_{u,v,u',v'} \mathbb{E}\Bigg[|(M_N)_{u',v'}| \times \Big| \tau_N \Big[\Big(k_N^{(1,2)}\Big)_{\substack{u,v\\u',v'}} \Big] \Big| \Bigg],
\end{aligned}$$

where we have denoted the $N \times N$ matrices

$$\begin{aligned}
\Big(h_N^{(1,2)}\Big)_{\substack{u,v\\u',v'}} &= \Big(h_N^{(2)}\Big)_{(u,u')} \Big(h_N^{(1)}\Big)_{(v',v)}, \\
\Big(k_N^{(1,2)}\Big)_{\substack{u,v\\u',v'}} &= \Big(h_N^{(1,2)}\Big)_{\substack{u,v\\u',v'}} - \mathbb{E}\Big[\Big(h_N^{(1,2)}\Big)_{\substack{u,v\\u',v'}} \Big].
\end{aligned}$$

Remark that by (1.61), for $u', v' = 1, \dots, k$,

$$\begin{aligned}
|(M_N)_{u',v'}| &= \Big| \Big(\sum_{j=1}^p a_j K_{S_N+T_N}(\Lambda) a_j \Big)_{u',v'} \Big| \\
&\leqslant \Big\| \sum_{j=1}^p a_j K_{S_N+T_N}(\Lambda) a_j \Big\|_e \\
&\leqslant \sum_{j=1}^p \|a_j\|^2 \ \|K_{S_N+T_N}(\Lambda)\|_e.
\end{aligned}$$

Then by Cauchy-Schwarz inequality we get:

$$\begin{aligned}
\|\Theta_N(\Lambda,\Gamma)\| &\leqslant k^{5/2} \sum_{j=1}^p \|a_j\|^2 \Big(\mathbb{E}\Big[\|K_{S_N+T_N}(\Lambda)\|_e^2 \Big] \\
&\qquad\qquad \times \max_{u,v,u',v'} \mathbb{E}\Big[\Big| \tau_N \Big[\Big(k_N^{(1,2)}\Big)_{\substack{u,v\\u',v'}} \Big] \Big|^2 \Big] \Big)^{1/2} \\
&\leqslant k^{5/2} \sum_{j=1}^p \|a_j\|^2 \Big(\sum_{u,v=1}^k \mathbb{V}\mathrm{ar}\Big(H_{S_N+T_N}(\Lambda) \Big)_{u,v} \qquad (1.63) \\
&\qquad\qquad \times \max_{u,v,u',v'} \mathbb{V}\mathrm{ar}\Big(\tau_N \Big[\Big(h_N^{(1,2)}\Big)_{\substack{u,v\\u',v'}} \Big] \Big) \Big)^{1/2}.
\end{aligned}$$

One is reduced to the study of variances of random variables. To use the Poincaré inequality, we write for $u, v, u', v' = 1, \ldots, k$,

$$\begin{aligned} \Big(H_{S_N+T_N}(\Lambda)\Big)_{u,v} &= F^{(1)}_{u,v}\Big(\mathbf{X}_N, \mathbf{Y}_N\Big), \\ \tau_N\Big[\Big(h_N^{(1,2)}\Big)_{\substack{u,v\\ u',v'}}\Big] &= F^{(2)}_{u,v,u',v'}\Big(\mathbf{X}_N, \mathbf{Y}_N\Big), \end{aligned}$$

where for all selfadjoint matrices $\mathbf{A} = (A_1, \ldots, A_p)$ in $\mathrm{M}_N(\mathbb{C})$, for all $\mathbf{B} = (B_1, \ldots, B_q)$ in $\mathrm{M}_N(\mathbb{C})$ and with $\tilde{S}_N = \sum_{j=1}^p a_j \otimes A_j$, $\tilde{T}_N = \sum_{j=1}^q b_j \otimes B_j$, we have set

$$\begin{aligned} F^{(1)}_{u,v}(\mathbf{A}, \mathbf{B}) &= \Big((\mathrm{id}_k \otimes \tau_N)\Big[(\Lambda \otimes \mathbf{1}_N - \tilde{S}_N - \tilde{T}_N)^{-1}\Big]\Big)_{u,v} \\ &= \frac{1}{N}(\mathrm{Tr}_k \otimes \mathrm{Tr}_N)\Big[(\epsilon_{v,u} \otimes \mathbf{1}_N)(\Lambda \otimes \mathbf{1}_N - \tilde{S}_N - \tilde{T}_N)^{-1}\Big], \end{aligned}$$

$$\begin{aligned} &F^{(2)}_{u,v,u',v'}(\mathbf{A}, \mathbf{B}) \\ &= \tau_N\Big[\Big((\Lambda \otimes \mathbf{1}_N - \tilde{S}_N - \tilde{T}_N)^{-1}\Big)_{(u,u')}\Big((\Gamma \otimes \mathbf{1}_N - \tilde{T}_N)^{-1}\Big)_{(v',v)}\Big] \\ &= \frac{1}{N}(\mathrm{Tr}_k \otimes \mathrm{Tr}_N)\Big[(\epsilon_{v,u} \otimes \mathbf{1}_N)(\Gamma \otimes \mathbf{1}_N - \tilde{T}_N)^{-1} \\ &\qquad \times(\epsilon_{u',v'} \otimes \mathbf{1}_N)\ (\Lambda \otimes \mathbf{1}_N - \tilde{S}_N - \tilde{T}_N)^{-1}\Big]. \end{aligned}$$

The functions and their partial derivatives are bounded (see [HT05, Lemma 4.6] with minor modifications), so that, since the law of $(\mathbf{X}_N, \mathbf{Y}_N)$ satisfies a Poincaré inequality with constant $\frac{\sigma}{N}$, one has

$$\begin{aligned} \mathbb{V}\mathrm{ar}\ \Big(H_{S_N+T_N}(\Lambda)\Big)_{u,v} &\leqslant \frac{\sigma}{N}\mathbb{E}\Big[\Big\|\nabla\ F^{(1)}_{u,v}(\mathbf{X}_N, \mathbf{Y}_N)\Big\|^2\Big], \\ \mathbb{V}\mathrm{ar}\Big(\tau_N\Big[\Big(h_N^{(1,2)}\Big)_{\substack{u,v\\ u',v'}}\Big]\Big) &\leqslant \frac{\sigma}{N}\mathbb{E}\Big[\Big\|\nabla\ F^{(2)}_{u,v,u',v'}(\mathbf{X}_N, \mathbf{Y}_N)\Big\|^2\Big]. \end{aligned}$$

We define the set $\mathcal{W}$ of families $(\mathbf{V}, \mathbf{W})$ of $N \times N$ Hermitian matrices, with $\mathbf{V} = (V_1, \ldots, V_p)$, $\mathbf{W} = (W_1, \ldots, W_q)$, of unit Euclidean norm in $\mathbf{R}^{(p+q)N^2}$. Then we have

$$\begin{aligned} &\mathbb{V}\mathrm{ar}\Big(H_{S_N+T_N}(\Lambda)\Big)_{u,v} \\ &\leqslant \frac{\sigma}{N}\mathbb{E}\Big[\max_{(\mathbf{V},\mathbf{W})\in\mathcal{W}}\Big|\frac{d}{dt}_{|t=0}\ F^{(1)}_{u,v}(\mathbf{X}_N + t\mathbf{V}, \mathbf{Y}_N + t\mathbf{W})\Big|^2\Big], \end{aligned}$$

$$\begin{aligned} &\mathbb{V}\mathrm{ar}\Big(\tau_N\Big[\Big(h_N^{(1,2)}\Big)_{\substack{u,v\\ u',v'}}\Big]\Big) \\ &\leqslant \frac{\sigma}{N}\mathbb{E}\Big[\max_{(\mathbf{V},\mathbf{W})\in\mathcal{W}}\Big|\frac{d}{dt}_{|t=0}\ F^{(2)}_{u,v,u',v'}(\mathbf{X}_N + t\mathbf{V}, \mathbf{Y}_N + t\mathbf{W})\Big|^2\Big]. \end{aligned}$$

For all $(\mathbf{V},\mathbf{W})$ in $\mathcal{W}$, for all selfadjoint $N\times N$ matrices $\mathbf{A}=(A_1,\dots,A_1)$, $\mathbf{B}=(B_1,\dots,B_1)$:

$$\begin{aligned}
&\left|\frac{d}{dt}_{|t=0}\, F^{(1)}_{u,v}(\mathbf{A}+t\mathbf{V},\mathbf{B}+t\mathbf{W})\right|^2\\
&= \left|\frac{d}{dt}_{|t=0}\, \frac{1}{N}(\mathrm{Tr}_k\otimes\mathrm{Tr}_N)\Big[(\epsilon_{v,u}\otimes\mathbf{1}_N)\right.\\
&\qquad\left.\times\Big(\Lambda\otimes\mathbf{1}_N-\sum_{j=1}^p a_j\otimes(A_j+tV_j)-\sum_{j=1}^q b_j\otimes(B_j+tW_j)\Big)^{-1}\Big]\right|^2\\
&= \left|\frac{1}{N}(\mathrm{Tr}_k\otimes\mathrm{Tr}_N)\Big[(\epsilon_{v,u}\otimes\mathbf{1}_N)(\Lambda\otimes\mathbf{1}_N-\tilde{S}_N-\tilde{T}_N)^{-1}\right.\\
&\qquad\left.\times\Big(\sum_{j=1}^p a_j\otimes V_j+\sum_{j=1}^q b_j\otimes W_j\Big)(\Lambda\otimes\mathbf{1}_N-\tilde{S}_N-\tilde{T}_N)^{-1}\Big]\right|^2.
\end{aligned}$$

The Cauchy-Schwarz inequality for $\mathrm{Tr}_k\otimes\mathrm{Tr}_N$ (i.e. for Tr_{kN}) gives

$$\begin{aligned}
&\left|\frac{d}{dt}_{|t=0}\, F^{(1)}_{u,v}(\mathbf{A}+t\mathbf{V},\mathbf{B}+t\mathbf{W})\right|^2\\
&\leqslant \frac{1}{N^2}\left\|(\epsilon_{v,u}\otimes\mathbf{1}_N)(\Lambda\otimes\mathbf{1}_N-\tilde{S}_N-\tilde{T}_N)^{-1}\right\|_e^2\\
&\qquad\times\left\|\Big(\sum_{j=1}^p a_j\otimes V_j+\sum_{j=1}^q b_j\otimes W_j\Big)(\Lambda\otimes\mathbf{1}_N-\tilde{S}_N-\tilde{T}_N)^{-1}\right\|_e^2.
\end{aligned}$$

Using (1.61) to split Euclidean norms into the product of an operator norm and an Euclidean norm, we get:

$$\begin{aligned}
&\left|\frac{d}{dt}_{|t=0}\, F^{(1)}_{u,v}(\mathbf{A}+t\mathbf{V},\mathbf{B}+t\mathbf{W})\right|^2\\
&\leqslant \frac{1}{N^2}\|\epsilon_{v,u}\otimes\mathbf{1}_N\|_e^2\ \|(\Lambda\otimes\mathbf{1}_N-\tilde{S}_N-\tilde{T}_N)^{-1}\|^2\\
&\qquad\times\left\|\sum_{j=1}^p a_j\otimes V_j+\sum_{j=1}^q b_j\otimes W_j\right\|_e^2\\
&\leqslant \frac{k}{N}\ \|(\mathrm{Im}\ \Lambda)^{-1}\|^4\ \left\|\sum_{j=1}^p a_j\otimes V_j+\sum_{j=1}^q b_j\otimes W_j\right\|_e^2.
\end{aligned}$$

Remark that, since $(\mathbf{V},\mathbf{W})\in\mathcal{W}$, the norm of the matrices V_j and W_j is bounded by one. Then we have the following:

$$\begin{aligned}
&\left\|\sum_{j=1}^p a_j\otimes V_j+\sum_{j=1}^q b_j\otimes W_j+b_j^*\otimes W_j^*\right\|_e\\
&\leqslant \sum_{j=1}^p\|a_j\|_e+2\sum_{j=1}^q\|b_j\|_e\leqslant\sqrt{k}\Big(\sum_{j=1}^p\|a_j\|+\sum_{j=1}^q\|b_j\|\Big).
\end{aligned}$$

Hence we finally obtain an estimate of $\mathbb{V}\mathrm{ar}(H_{S_N+T_N}(\Lambda)\ )_{u,v})$:

$$\mathbb{V}\mathrm{ar}\Big(H_{S_N+T_N}(\Lambda)\ \Big)_{u,v} \leqslant \frac{k^2\sigma}{N^2}\Big(\sum_{j=1}^{p}\|a_j\| + \sum_{j=1}^{q}\|b_j\|\Big)^2 \ \|(\mathrm{Im}\ \Lambda)^{-1}\|^4. \tag{1.64}$$

We obtain a similar estimate for $\mathbb{V}\mathrm{ar}\Big(\tau_N\Big[\Big(h_N^{(1,2)}\Big)_{\substack{u,v\\u',v'}}\Big]\ \Big)$. The partial derivative of $F^{(2)}_{u,v,u',v'}$ gives two terms: $\forall(\mathbf{V},\mathbf{W})\in\mathcal{W}$, $\forall(\mathbf{A},\mathbf{B})\in \mathrm{M}_N(\mathbb{C})^{p+q}$

$$\begin{aligned}
&\frac{d}{dt}_{|t=0}\ F^{(2)}_{u,v,u',v'}(\mathbf{A}+t\mathbf{V},\mathbf{B}+t\mathbf{W})\\
&= \frac{1}{N}(\mathrm{Tr}_k\otimes\mathrm{Tr}_N)\Big[(\epsilon_{v,u}\otimes\mathbf{1}_N)(\Gamma\otimes\mathbf{1}_N-\tilde{T}_N)^{-1}\Big(\sum_{j=1}^{q}b_j\otimes W_j\Big)\\
&\qquad\times(\Gamma\otimes\mathbf{1}_N-\tilde{T}_N)^{-1}(\epsilon_{u',v'}\otimes\mathbf{1}_N)(\Lambda\otimes\mathbf{1}_N-\tilde{S}_N-\tilde{T}_N)^{-1}\\
&\quad+(\epsilon_{v,u}\otimes\mathbf{1}_N)(\Gamma\otimes\mathbf{1}_N-\tilde{T}_N)^{-1}(\epsilon_{u',v'}\otimes\mathbf{1}_N)(\Lambda\otimes\mathbf{1}_N-\tilde{S}_N-\tilde{T}_N)^{-1}\\
&\qquad\times\Big(\sum_{j=1}^{p}a_j\otimes V_j^{(N)}+\sum_{j=1}^{q}b_j\otimes W_j^{(N)}\Big)(\Lambda\otimes\mathbf{1}_N-\tilde{S}_N-\tilde{T}_N)^{-1}\Big].
\end{aligned}$$

We then get the following:

$$\begin{aligned}
&\left|\frac{d}{dt}_{|t=0}\ F^{(2)}_{u,v,u',v'}(\mathbf{A}+t\mathbf{V},\mathbf{B}+t\mathbf{W})\right|^2\\
&\leqslant \frac{k^2}{N}\Big(\sum_{j=1}^{p}\|a_j\|+\sum_{j=1}^{q}\|b_j\|\Big)^2\ \|(\mathrm{Im}\ \Gamma)^{-1}\|^2\\
&\qquad\times\|(\mathrm{Im}\ \Lambda)^{-1}\|^2\Big(\|(\mathrm{Im}\ \Lambda)^{-1}\|+\|(\mathrm{Im}\ \Gamma)^{-1}\|\Big)^2.
\end{aligned}$$

Hence we have

$$\begin{aligned}
&\mathbb{V}\mathrm{ar}\Big(\tau_N\Big[\Big(h_N^{(1,2)}\Big)_{\substack{u,v\\u',v'}}\Big]\ \Big)\\
&\leqslant \frac{k^2\sigma}{N^2}\Big(\sum_{j=1}^{p}\|a_j\|+\sum_{j=1}^{q}\|b_j\|\Big)^2\|(\mathrm{Im}\ \Gamma)^{-1}\|^2\\
&\qquad\times\|(\mathrm{Im}\ \Lambda)^{-1}\|^2\ \Big(\|(\mathrm{Im}\ \Gamma)^{-1}\|+\|(\mathrm{Im}\ \Lambda)^{-1}\|\Big)^2.
\end{aligned}\tag{1.65}$$

We then obtain as desired, by (1.63), (1.64) and (1.65):

$$\begin{aligned}
\|\Theta_N(\Lambda,\Gamma)\| &\leqslant k^{5/2}\sum_{j=1}^{p}\|a_j\|^2\Big(\sum_{u,v=1}^{k}\mathbb{V}\mathrm{ar}\Big(H_{S_N+T_N}(\Lambda)\Big)_{u,v}\\
&\qquad\times\max_{u,v,u',v'}\mathbb{V}\mathrm{ar}\Big(\tau_N\Big[\Big(h_N^{(1,2)}\Big)_{\substack{u,v\\u',v'}}\Big]\ \Big)\ \Big)^{1/2}\\
&\leqslant \frac{c}{N^2}\left\|(\mathrm{Im}\ \Gamma)^{-1}\right\|\left\|(\mathrm{Im}\ \Lambda)^{-1}\right\|^3\\
&\qquad\times\Big(\|(\mathrm{Im}\ \Gamma)^{-1}\|+\|(\mathrm{Im}\ \Lambda)^{-1}\|\Big),
\end{aligned}$$

where $c=k^{9/2}\sigma\sum_{j=1}^{p}\|a_j\|^2\Big(\sum_{j=1}^{p}\|a_j\|+\sum_{j=1}^{q}\|b_j\|\Big)^2$.

□

1.5.3 Proof of Theorem 1.5.1

By (1.40), for all Λ in $\mathrm{M}_k(\mathbb{C})^+$, the matrix $\Lambda - \mathcal{R}_s\Big(G_{S_N+T_N}(\Lambda)\Big)$ is in $\mathrm{M}_k(\mathbb{C})^+$ and then it makes sense to choose $\Gamma = \Lambda - \mathcal{R}_s\Big(G_{S_N+T_N}(\Lambda)\Big)$ in Equation (1.56). We obtain for all Λ in $\mathrm{M}_k(\mathbb{C})^+$,

$$G_{S_N+T_N}(\Lambda) = G_{T_N}\bigg(\Lambda - \mathcal{R}_s\Big(G_{S_N+T_N}(\Lambda)\Big)\bigg) + \Theta_N(\Lambda),$$

where $\Theta_N(\Lambda) = \Theta_N\bigg(\Lambda, \Lambda - \mathcal{R}_s\Big(G_{S_N+T_N}(\Lambda)\Big)\bigg)$ is analytic in k^2 complex variables. Recall that by (1.41), we have $\bigg\|\Big(\Lambda - \mathcal{R}_s\Big(G_{S_N+T_N}(\Lambda)\Big)\Big)^{-1}\bigg\| \leqslant \|(\Lambda)^{-1}\|$, which gives (when replacing c in (1.58) by $c/2$) the expected estimate of $\Theta_N(\Lambda)$.

1.6 Proof of Estimate (1.30)

Let $(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{x}, \mathbf{y})$ be as in Section 1.3. We assume that $\big(\mathbf{x}, \mathbf{y}, (\mathbf{Y}_N)_{N\geqslant 1}\big)$ are realized in a same $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ with faithful trace, where

– the families $\mathbf{x}$, $\mathbf{y}$, $\mathbf{Y}_1$, $\mathbf{Y}_2, \dots, \mathbf{Y}_N, \dots$ are free,
– for any polynomials P in q non commutative indeterminates $\tau[P(\mathbf{Y}_N)] := \tau_N[P(\mathbf{Y}_N)]$.

Consider L a degree one selfadjoint polynomial with coefficients in $\mathrm{M}_k(\mathbb{C})$. Define the Stieltjes transform of $L_N = L(\mathbf{X}_N, \mathbf{Y}_N)$ and $\ell_N = L(\mathbf{x}, \mathbf{Y}_N)$: for all $\lambda \in \mathbb{C}^+ = \Big\{z \in \mathbb{C} \Big| \ \mathrm{Im}\ z > 0\Big\}$,

$$g_{L_N}(\lambda) = \mathbb{E}\bigg[(\tau_k \otimes \tau_N)\Big[\Big(\lambda \mathbf{1}_k \otimes \mathbf{1}_N - L_N\Big)^{-1}\Big]\bigg], \tag{1.66}$$

$$g_{\ell_N}(\lambda) = (\tau_k \otimes \tau)\Big[\Big(\lambda \mathbf{1}_k \otimes \mathbf{1} - \ell_N\Big)^{-1}\Big]. \tag{1.67}$$

One can always write $L_N = a_0 \otimes \mathbf{1}_N + S_N + T_N$, $\ell_N = a_0 \otimes \mathbf{1} + s + T_N$, where

$$S_N = \sum_{j=1}^{p} a_j \otimes X_j^{(N)}, \ s = \sum_{j=1}^{p} a_j \otimes x_j, \ T_N = \sum_{j=1}^{q} b_j \otimes Y_j^{(N)},$$

and $a_0, \dots, a_p, b_1, \dots, b_q$ are Hermitian matrices in $\mathrm{M}_k(\mathbb{C})$. Define the $\mathrm{M}_k(\mathbb{C})$-valued Stieltjes transforms of $S_N + T_N$ and $s + T_N$: for all $\Lambda \in \mathrm{M}_k(\mathbb{C})^+ = \Big\{\Lambda \in \mathrm{M}_k(\mathbb{C}) \ \Big| \ \mathrm{Im}\ \Lambda > 0\Big\}$,

$$G_{S_N+T_N}(\Lambda) = \mathbb{E}\bigg[(\mathrm{id}_k \otimes \tau_N)\Big[\Big(\Lambda \otimes \mathbf{1}_N - S_N - T_N\Big)^{-1}\Big]\bigg],$$

$$G_{s+T_N}(\Lambda) = (\mathrm{id}_k \otimes \tau)\Big[\Big(\Lambda \otimes \mathbf{1} - s - T_N\Big)^{-1}\Big].$$

Then one has: for all λ in $\mathbb{C}^+$

$$g_{L_N}(\lambda) = \tau_k\Big[\ G_{S_N+T_N}(\lambda \mathbf{1}_k - a_0)\ \Big], \ g_{\ell_N}(\lambda) = \tau_k\Big[\ G_{s+T_N}(\lambda \mathbf{1}_k - a_0)\ \Big].$$

By Proposition 1.4.2, for any $\Lambda \in \mathrm{M}_k(\mathbb{C})^+$, one has

$$G_{s+T_N}(\Lambda) = G_{T_N}\Big(\Lambda - \mathcal{R}_s\big(G_{s+T_N}(\Lambda)\ \big)\ \Big).$$

On the other hand, since the matrices of $\mathbf{Y}_N$ are deterministic, we can apply Theorem 1.5.1 with $\sigma = 1$

$$G_{S_N+T_N}(\Lambda) = G_{T_N}\Big(\Lambda - \mathcal{R}_s\big(G_{S_N+T_N}(\Lambda)\ \big)\ \Big) + \Theta_N(\Lambda),$$

where $\|\Theta_N(\Lambda)\| \leqslant \frac{c}{N^2}\,\|(\mathrm{Im}\ \Lambda)^{-1}\|^5$ for a constant $c > 0$. Define

$$\Omega_\eta^{(N)} = \Big\{\Lambda \in \mathrm{M}_k(\mathbb{C})^+\Big|\ \|(\mathrm{Im}\ \Lambda)^{-1}\| < N^\eta\Big\}.$$

Then for $\eta < 1/3$, there exists N_0 such that for all $N \geqslant N_0$ and for any Λ in $\Omega_\eta^{(N)}$, one has

$$\kappa(\Lambda) := \|\Theta_N(\Lambda)\|\ \|(\mathrm{Im}\ \Lambda)^{-1}\|\ \sum_{j=1}^p \|a_j\|^2 \leqslant \frac{c}{N^2}\|(\mathrm{Im}\ \Lambda)^{-1}\|^6 \leqslant cN^{6\eta-2} \leqslant \frac{1}{2}.$$

Then by Proposition 1.4.3 with $(t, G, \Theta, \Omega, \varepsilon) = (T_N, G_{S_N+T_N}, \Theta_N, \Omega_\eta^{(N)}, 1/2)$, one has

$$\begin{aligned}
&\|G_{s+T_N}(\Lambda) - G_{S_N+T_N}(\Lambda)\| \\
&\leqslant \ \Big(1 + 2\sum_{j=1}^p \|a_j\|^2\ \|(\mathrm{Im}\Lambda)^{-1}\|^2\Big)\ \|\Theta(\Lambda)\| \\
&\leqslant \ c\Big(1 + 2\sum_{j=1}^p \|a_j\|^2\ \|(\mathrm{Im}\ \Lambda)^{-1}\|^2\Big)\ \frac{\|(\mathrm{Im}\ \Lambda)^{-1}\|^5}{N^2}.
\end{aligned}$$

Hence for every $\varepsilon > 0$, there exist N_0 and γ such that for all $N \geqslant N_0$, for all λ in $\mathbb{C}$ such that $\varepsilon \leqslant (\mathrm{Im}\ \lambda)^{-1} \leqslant N^\gamma$, one has

$$\begin{aligned}
&|g_{L_N}(\lambda) - g_{\ell_N}(\lambda)| \\
&\quad \leqslant \|G_{s+T_N}(\lambda\mathbf{1}_k - a_0) - G_{S_N+T_N}(\lambda\mathbf{1}_k - a_0)\| \leqslant \frac{c}{N^2}(\mathrm{Im}\ \lambda)^{-7}, \quad (1.68)
\end{aligned}$$

where c denotes now the constant $c = k^{9/2}\sum_{j=1}^p \|a_j\|\Big(\sum_{j=1}^p \|a_j\| + \sum_{j=1}^q \|b_j\|\Big)^2\Big(\varepsilon^{-2} + 2\sum_{j=1}^p \|a_j\|^2\Big)$.

1.7 Proof of Step 2: An intermediate inclusion of spectrum

For a review on the theory of $\mathcal{C}^*$-algebras, we refer the readers to [Con00] and [BO08]. Notably, Appendix A of the second reference contains facts about ultrafilters and ultraproducts that are used in this section.

Let $\Big(\mathbf{x}, \mathbf{y}, (\mathbf{Y}_N)_{N\geqslant 1}\Big)$ be as in Section 1.3. We assume that these non commutative random variables are realized in the same $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ with faithful trace, where

– the families $\mathbf{x}$, $\mathbf{y}$, $\mathbf{Y}_1$, $\mathbf{Y}_2, \dots, \mathbf{Y}_N, \dots$ are free,
– for any polynomials P in q non commutative indeterminates $\tau[P(\mathbf{Y}_N)] := \tau_N[P(\mathbf{Y}_N)]$.

A consequence of Voiculescu's theorem and of Shlyakhtenko's Theorem 1.10.1 in Appendix 1.10 is that for all polynomials P in $p+q$ non commutative indeterminates,

$$\tau[P(\mathbf{x}, \mathbf{Y}_N)] \underset{N\to\infty}{\longrightarrow} \tau[P(\mathbf{x}, \mathbf{y})], \tag{1.69}$$

$$\|P(\mathbf{x}, \mathbf{Y}_N)\| \underset{N\to\infty}{\longrightarrow} \|P(\mathbf{x}, \mathbf{y})\|. \tag{1.70}$$

In order to prove Step 2, it remains to show that (1.70) still holds when the polynomials P are $\mathrm{M}_k(\mathbb{C})$-valued. This fact is a folklore result in $\mathcal{C}^*$-algebra theory, we give a proof for readers convenience. We need first the two following lemmas.

Lemma 1.7.1. Let $\mathcal{A}$ and $\mathcal{B}$ be unital $\mathcal{C}^*$-algebra. Let $\pi : \mathcal{A} \to \mathcal{B}$ be a morphism of unital *-algebra. Then π is contractive.

Proof. It is easy to see that for any a in $\mathcal{A}$, the spectrum of $\pi(a)$ is included in the spectrum of a (since $\lambda \mathbf{1}_{\mathcal{A}} - a$ invertible implies that $\lambda \mathbf{1}_{\mathcal{A}} - \pi(a)$ is also invertible). Hence we get that for all a in $\mathcal{A}$

$$\|\pi(a)\|^2 = \|\pi(a^*a)\| \leqslant \|a^*a\| = \|a\|^2.$$

□

Lemma 1.7.2. Let $\mathcal{A}$ be a unital $\mathcal{C}^*$-algebra. Then for any integer $k \geqslant 1$, there exists a unique $\mathcal{C}^*$-algebra structure on $\mathrm{M}_k(\mathbb{C}) \otimes \mathcal{A}$ compatible with the structure on $\mathcal{A}$. In particular, if $\mathcal{A}$ is a $\mathcal{C}^*$-probability space equipped with a faithful tracial state τ, then $\mathrm{M}_k(\mathbb{C}) \otimes \mathcal{A}$ is a $\mathcal{C}^*$-probability space with trace $(\tau_k \otimes \tau)$ and norm $\| \cdot \|_{\tau_k \otimes \tau}$, where τ_k is the normalized trace on $\mathrm{M}_k(\mathbb{C})$ and $\| \cdot \|_{\tau_k\otimes\tau}$ is given by Formula (1.9).

Sketch of the proof. For the existence we consider the norm given by the spectral radius. The uniqueness follows from Lemma 1.7.1. □

Proposition 1.7.3. Let $k \geqslant 1$ be an integer. For all $N \geqslant 1$, let $\mathbf{z}_N = (z_1^{(N)}, \dots, z_p^{(N)})$, respectively $\mathbf{z} = (z_1, \dots, z_p)$, be self-adjoint non commutative random variables in a $\mathcal{C}^*$- probability space $(\mathcal{A}_N, .^*, \tau_N, \| \cdot \|_{\tau_N})$, respectively $(\mathcal{A}, .^*, \tau, \| \cdot \|_\tau)$. Assume that the traces τ_N and τ are faithful (hence the notation for the norms) and that for any polynomial P in p non commutative indeterminates,

$$\tau_N[P(\mathbf{z}_N)] \underset{N\to\infty}{\longrightarrow} \tau[P(\mathbf{z})], \tag{1.71}$$

$$\|P(\mathbf{z}_N)\|_{\tau_N} \underset{N\to\infty}{\longrightarrow} \|P(\mathbf{z})\|_\tau. \tag{1.72}$$

Then for any polynomial P in p non commutative indeterminates with coefficients in $\mathrm{M}_k(\mathbb{C})$,

$$\|P(\mathbf{z}_N)\|_{\tau_k\otimes\tau_N} \underset{N\to\infty}{\longrightarrow} \|P(\mathbf{z})\|_{\tau_k\otimes\tau}. \tag{1.73}$$

We abuse notation and write with the same symbol the traces in $\mathrm{M}_k(\mathbb{C})$ and $\mathcal{A}_N$ when $N = k$. There is no danger of confusion.

Proof. For any positive integer k and any ultrafilter $\mathcal{U}$ on $\mathbb{N}$, we define the ultraproduct

$$\mathfrak{A}^{(k)} = \prod^{\mathcal{U}} \mathrm{M}_k(\mathbb{C}) \otimes \mathcal{A}_N,$$

which is the quotient of

$$\left\{ (a_N)_{N\geqslant 1} \;\middle|\; \forall N \geqslant 1,\ a_N \in \mathrm{M}_k(\mathbb{C}) \otimes \mathcal{A}_N \text{ and } \sup_{N\geqslant 1} \|a_N\| < \infty \right\},$$

by

$$\left\{ (a_N)_{N\geqslant 1} \;\middle|\; \forall N \geqslant 1,\ a_N \in \mathrm{M}_k(\mathbb{C}) \otimes \mathcal{A}_N \text{ and } \lim_{N\to\mathcal{U}} \|a_N\| = 0 \right\}.$$

The algebra $\mathfrak{A}^{(k)}$ is a $\mathcal{C}^*$-algebra whose norm $\|\cdot\|_{\mathfrak{A}^{(k)}}$ is given by: for all a in $\mathfrak{A}^{(k)}$, equivalence class of $(a_N)_{N\geqslant 1}$

$$\|a\|_{\mathfrak{A}^{(k)}} = \lim_{N\to\mathcal{U}} \|a_N\|_{\tau_k\otimes\tau_N}.$$

Furthermore $\mathfrak{A}^{(k)}$ is a $\mathcal{C}^*$-probability space which can be identified with $\mathrm{M}_k(\mathbb{C}) \otimes \mathfrak{A}^{(1)}$. The trace $\tilde{\tau}$ on $\mathfrak{A}^{(1)}$ is given by: for all a in $\mathfrak{A}^{(1)}$, equivalence class of $(A_N)_{N\geqslant 1}$, one has

$$\tilde{\tau}[a] = \lim_{N\to\mathcal{U}} \tau[A_N].$$

If the classical limit as N goes to infinity exists, then the trace of a does not depends on the ultrafilter $\mathcal{U}$ and is given by the limit. The trace on $\mathfrak{A}^{(k)}$ is $(\tau_k \otimes \tilde{\tau})$. Notice that $(\tau_k \otimes \tilde{\tau})$ on $\mathfrak{A}^{(k)}$ is not faithful in general, which implies that the norm $\|\cdot\|_{\mathfrak{A}^{(k)}}$ and the norm $\|\cdot\|_{\tau_k\otimes\tilde{\tau}}$ given by $(\tau_k \otimes \tilde{\tau})$ with Formula (1.9) are not equal on the whole $\mathcal{C}^*$-algebra.
At last, we can equip $\mathfrak{A}^{(k)}$ with a structure of operator-valued $\mathcal{C}^*$-probability space. Define the unital sub-algebra $\mathcal{B}$ of $\mathfrak{A}^{(k)}$ as the set

$$\left\{ b \otimes \mathbf{1}_{\mathfrak{A}^{(1)}} \;\middle|\; b \in \mathrm{M}_k(\mathbb{C}) \right\} \subset \mathfrak{A}^{(k)}.$$

The conditional expectation in $\mathfrak{A}^{(k)}$ is given by $(\mathrm{id}_k \otimes \tilde{\tau}) : \mathfrak{A}^{(k)} \to \mathcal{B}$.

For $j = 1, \ldots, p$, we denote by $\tilde{z}_j$ in $\mathfrak{A}^{(1)}$ the equivalence class of the sequence $(z_j^{(N)})_{N\geqslant 1}$. We have by definition of $\mathfrak{A}^{(k)}$: for all polynomial P in $p + 2q$ non commutative indeterminates with coefficients in $\mathrm{M}_k(\mathbb{C})$,

$$\|P(\mathbf{z}_N)\|_{\tau_N} \underset{N\to\mathcal{U}}{\longrightarrow} \|P(\tilde{\mathbf{z}})\|_{\mathfrak{A}^{(k)}}$$

Let $\mathcal{C}^*(\tilde{\mathbf{z}})$ be the sub-algebra spanned by $\tilde{\mathbf{z}} = (\tilde{z}_1, \ldots, \tilde{z}_p)$ in $\mathfrak{A}^{(1)}$ and let $\mathcal{C}^*(\mathbf{z})$ be the sub-algebra spanned by $\mathbf{z}$ in $\mathcal{A}$. Then by (1.72), the $\mathcal{C}^*$-algebras $\mathcal{C}^*(\tilde{\mathbf{z}})$ and $\mathcal{C}^*(\mathbf{z})$ are isomorphic. Hence we get an isomorphism of the *-algebras $\mathrm{M}_k(\mathbb{C}) \otimes \mathcal{C}^*(\tilde{\mathbf{z}})$ and $\mathrm{M}_k(\mathbb{C}) \otimes \mathcal{C}^*(\mathbf{z})$, and so an isomorphism of the $\mathcal{C}^*$-algebras by Lemma

1.7.1. Hence, for all polynomial P in $p+2q$ non commutative indeterminates with coefficients in $\mathrm{M}_k(\mathbb{C})$,

$$\|P(\tilde{\mathbf{z}})\|_{\mathfrak{A}^{(k)}} = \|P(\mathbf{z})\|_{\tau_k\otimes\tilde{\tau}}$$

Hence we get

$$\|P(\mathbf{z}_N)\|_{\tau_k\otimes\tau_N} \underset{N\to\mathcal{U}}{\longrightarrow} \|P(\mathbf{z})\|_{\tau_k\otimes\tilde{\tau}}$$

for all ultrafilter $\mathcal{U}$. Then the convergence holds when N goes to infinity. □

Proof of Step 2. Let L be a selfadjoint degree one polynomial in $p+q$ non commutative indeterminates with coefficients in $\mathrm{M}_k(\mathbb{C})$. Define $\ell_N = L(\mathbf{x}, \mathbf{Y}_N)$ and $\ell = L(\mathbf{x}, \mathbf{y})$. Then by Proposition 1.7.3, for all commutative polynomials P, one has

$$\|P(\ell_N)\|_{\tau_k\otimes\tau} \underset{N\to\infty}{\longrightarrow} \|P(\ell)\|_{\tau_k\otimes\tau}.$$

The convergence extends to continuous function on the real line and then, with an appropriate choice of test functions, Step 2 follows. □

1.8 Proof of Step 3: from Stieltjes transforms to spectra

Let $\mathbf{X}_N, \mathbf{Y}_N, \mathbf{x}$ and $\mathbf{y}$ be as in Section 1.3. As before $\mathbf{x}, \mathbf{y}$, and $\mathbf{Y}_N$ are assumed to be realized in a same $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ with faithful trace. Let L be a selfadjoint degree one polynomial with coefficients in $\mathrm{M}_k(\mathbb{C})$.

For any function $f:\mathbb{R}\to\mathbb{R}$ and any Hermitian matrix A with spectral decomposition $A = U\mathrm{diag}\,(\lambda_1,\dots,\lambda_K)U^*$, with U unitary, we set the Hermitian matrix $f(A) = U\mathrm{diag}\,(f(\lambda_1),\dots,f(\lambda_K))U^*$. For any function $f:\mathbb{R}\mapsto\mathbb{R}$, we set

$$D_N(f) = (\tau_k\otimes\tau_N)\Big[f(L(\mathbf{X}_N, \mathbf{Y}_N))\Big].$$

By Step 2, for all $\varepsilon>0$, there exists $N_0\geqslant 1$ such that for all $N\geqslant N_0$, one has

$$\mathrm{Sp}\Big(\ L(\mathbf{x}, \mathbf{Y}_N)\ \Big) \subset \mathrm{Sp}\Big(\ L(\mathbf{x}, \mathbf{y})\ \Big) + (-\varepsilon, \varepsilon).$$

Hence, for any function f vanishing on a neighborhood of the spectrum of $L(\mathbf{x}, \mathbf{y})$, there exists $N_0\geqslant 1$ such that for all $N\geqslant N_0$, the function f actually vanishes on a neighborhood of the spectrum of $L(\mathbf{x}, \mathbf{Y}_N)$. In particular, with μ_N (respectively ν_N) denoting the empirical eigenvalue distribution of $L_N = L(\mathbf{X}_N, \mathbf{Y}_N)$ (respectively $\ell_N = L(\mathbf{x}, \mathbf{Y}_N)$), one has

$$\mathbb{E}\Big[D_N(f)\Big] = \mathbb{E}\Big[\ \int f\ \mathrm{d}\mu_N\ \Big] = \mathbb{E}\Big[\ \int f\ \mathrm{d}\mu_N\ \Big] - \int f\ \mathrm{d}\nu_N. \tag{1.74}$$

Furthermore, by Estimate (1.30), with the Stieltjes transforms of L_N and of ℓ_N defined by: for all λ in $\mathbb{C}^+$

$$\begin{aligned} g_{L_N}(\lambda) &= \mathbb{E}\Big[(\tau_k \otimes \tau_N)\Big[\Big(\lambda \mathbf{1}_k \otimes \mathbf{1}_N - L_N \Big)^{-1} \Big] \Big] = \mathbb{E}\Big[\int \frac{1}{\lambda - t} \mathrm{d}\mu_N(t) \Big] \\ g_{\ell_N}(\lambda) &= (\tau_k \otimes \tau)\Big[\Big(\lambda \mathbf{1}_k \otimes \mathbf{1} - \ell_N \Big)^{-1} \Big] = \int \frac{1}{\lambda - t} \mathrm{d}\nu_N(t), \end{aligned}$$

we have shown that: for any $\varepsilon > 0$ and $A > 0$, there exist $N_0, c, \eta, \gamma, \alpha > 0$ such that for all $N \geqslant N_0$, for all λ in $\mathbb{C}$ such that $\varepsilon \leqslant (\text{Im } \lambda)^{-1} \leqslant N^\gamma$ and $|\text{Re } \lambda| \leqslant A$

$$|g_{L_N}(\lambda) - g_{\ell_N}(\lambda)| \leqslant \frac{c}{N^2} (\text{Im } \lambda)^{-\alpha}. \tag{1.75}$$

With (1.74) and (1.75) established, it is easy to show with minor modifications of [AGZ10, Lemma 5.5.5] the following result.

Lemma 1.8.1. For every smooth function $f : \mathbb{R} \to \mathbb{R}$ non negative, compactly supported and vanishing on a neighborhood of the spectrum of $L(\mathbf{x}, \mathbf{y})$, there exists a constant such that for all N large enough

$$\Big| \mathbb{E}\Big[D_N(f) \Big] \Big| \leqslant \frac{c}{N^2}. \tag{1.76}$$

To get an almost sure control of $D_N(f)$, we use the fact that the entries of the matrices $\mathbf{X}_N$ satisfy a concentration inequality.

Lemma 1.8.2. With f as in Lemma 1.8.1, there exists $\kappa > 0$ such that, almost surely

$$N^{1+\kappa} D_N(f) \underset{N \to \infty}{\longrightarrow} 0. \tag{1.77}$$

Proof. The law of the random matrices satisfying a Poincaré inequality with constant $\frac{1}{N}$ and L being a polynomial of degree one, for all Lipschitz function $\Psi : \mathrm{M}_{kN}(\mathbb{C}) \mapsto \mathbb{R}$, by [Gui09, Lemma 5.2] one has:

$$\mathbb{P}\Big(\Big| \Psi(L_N) - \mathbb{E}\Big[\Psi(L_N) \Big] \Big| \geqslant \delta \Big) \leqslant K_1 e^{-K_2 \frac{\sqrt{N}\delta}{|\Psi|_{\mathcal{L}}}}, \tag{1.78}$$

where K_1, K_2 are positive constants and $|\Psi|_{\mathcal{L}} = \sup_{A \neq B \in \mathrm{M}_{kN}(\mathbb{C})} \frac{|\Psi(A) - \Psi(B)|}{\|A - B\|_e}$. Recall that the Euclidean norm $\| \cdot \|_e$ of a matrix $A = (a_{i,j})_{i,j=1}^{kN}$ is given by

$$\|A\|_e = \sqrt{\sum_{i,j=1}^{kN} |a_{i,j}|^2}.$$

For any Hermitian matrices A in $\mathrm{M}_{kN}(\mathbb{C})$ and any function $f : \mathbb{R} \to \mathbb{R}$, we set

$$\Phi_N^{(f)}(A) = (\tau_k \otimes \tau_N)\Big[f(A) \Big]. \tag{1.79}$$

For all smooth function $f : \mathbb{R} \to \mathbb{R}$, $N \geqslant 1$ and $0 < \kappa < \frac{1}{2}$, we define

$$\mathcal{B}^{(f)}_{N,\kappa} = \left\{ A \in \mathrm{M}_{kN}(\mathbb{C}) \;\middle|\; A \text{ is Hermitian and } \left|\Phi^{(f'^2)}_N(A)\right| \leqslant \frac{1}{N^{4\kappa}} \right\}, \tag{1.80}$$

and denote $\rho^{(f)}_{N,\kappa} = |(\Phi^{(f)}_N)_{|\mathcal{B}_{N,\kappa}}|_{\mathcal{L}}$. Define $\Psi^{(f)}_N : \mathrm{M}_{kN}(\mathbb{C}) \mapsto \mathbb{R}$ by: $\forall A \in \mathrm{M}_N(\mathbb{C})$

$$\Psi^{(f)}_N(A) = \sup_{B \in \mathcal{B}^{(f)}_{N,\kappa}} \left\{ \Phi^{(f)}_N(B) - \rho^{(f)}_{N,\kappa} \;\|A - B\|_2 \;\right\}, \tag{1.81}$$

and denote $\tilde{D}_N(f) = \Psi^{(f)}_N(L_N)$. By [Gui09, Proof of Lemma 5.9], $\Psi^{(f)}_N$ coincides with $\Phi^{(f)}_N$ on $\mathcal{B}^{(f)}_{N,\kappa}$ and is Lipschitz with constant $|\Psi^{(f)}_N|_{\mathcal{L}} \leqslant \rho^{(f)}_{N,\kappa}$.

For all Hermitian matrices A in $\mathrm{M}_{kN}(\mathbb{C})$, M in $\mathrm{M}_{kN}(\mathbb{C})$ and $n \geqslant 1$, one has $\frac{d}{dt}_{|t=0}(A + tM)^n = \sum_{m=0}^n A^m M A^{n-m-1}$ and then $\frac{d}{dt}_{|t=0}(\tau_k \otimes \tau_N)[(A + tM)^n] = (\tau_k \otimes \tau_N)[nA^{n-1}M]$. So for all polynomials P, one has $\mathrm{D}_A\Phi^{(P)}_N(M) = (\tau_k \otimes \tau_N)[P'(A)M]$. Hence, by density of polynomials, for any smooth function $f : \mathbb{R} \to \mathbb{R}$ one has $\mathrm{D}_A\Phi^{(f)}_N(M) = (\tau_k \otimes \tau_N)[f'(A)M]$. By the Cauchy-Schwarz inequality, we get

$$\begin{aligned} \left|\mathrm{D}_A\Phi^{(f)}_N(M)\right|^2 &= |(\tau_k \otimes \tau_N)[f'(A)M]|^2 \\ &\leqslant (\tau_k \otimes \tau_N)[f'(A)^2] \times (\tau_k \otimes \tau_N)[M^*M] \\ &= \Phi^{(f'^2)}_N(A) \times \frac{\|M\|_e}{kN}. \end{aligned}$$

Then, for any smooth function f, one has

$$\rho^{(f)}_{N,\kappa} \;\leqslant \frac{1}{\sqrt{kN}} \| \;(\Phi^{(f'^2)}_N)_{|\mathcal{B}^{(f)}_{N,\kappa}} \;\|^{1/2}_\infty, \tag{1.82}$$

where $\|\cdot\|_\infty$ denotes the supremum of the considered function on the set of $kN \times kN$ Hermitian matrices. Hence we get that $|\Psi^{(f)}_N|_{\mathcal{L}} \leqslant \rho^{(f)}_{N,\kappa} \leqslant \frac{1}{\sqrt{k}} N^{-1/2-2\kappa}$.

We fix f a smooth function, non negative, compactly supported and vanishing on a neighborhood of the spectrum of $L(\mathbf{x}, \mathbf{y})$. By the Tchebychev inequality

$$\mathbb{P}(L_N \notin \mathcal{B}^{(f)}_{N,\kappa}) \;=\; \mathbb{P}\Big(D_N(f'^2) \geqslant \frac{1}{N^{4\kappa}}\Big) \leqslant N^{4\kappa}\mathbb{E}\Big[D_N(f'^2)\Big] \leqslant \frac{c}{N^{2-4\kappa}}, \tag{1.83}$$

where we have used Lemma 1.8.1 (f'^2 also vanishes in a neighborhood of the spectrum of $L(\mathbf{x}, \mathbf{y})$). Moreover, since $\Psi^{(f)}_N$ and $\Phi^{(f)}_N$ are equals in $\mathcal{B}^{(f)}_{N,\kappa}$ and $\|\Psi^{(f)}_N\|_\infty \leqslant \|\Phi^{(f)}_N\|_\infty$,

$$\left|\mathbb{E}\Big[\tilde{D}_N(f) - D_N(f)\Big]\;\right| \leqslant \|\Phi^{(f)}_N\|_\infty \mathbb{P}(L_N \notin \mathcal{B}^{(f)}_{N,\kappa}) \leqslant \|\Phi^{(f)}_N\|_\infty \frac{c}{N^{2-4\kappa}} \tag{1.84}$$

Now, by (1.78) applied to $\Psi_N^{(f)}$: for all $\delta > 0$

$$\begin{aligned}
&\mathbb{P}\bigg(\Big|D_N(f) - \mathbb{E}\big[D_N(f)\ \big]\ \Big| > \frac{\delta}{N^{1+\kappa}} \ \text{ and } \ L_N \in \mathcal{B}_{N,\kappa}^{(f)}\bigg) \\
&\leqslant \ P\bigg(\Big|\tilde{D}_N(f) - \mathbb{E}\big[\tilde{D}_N(f)\ \big]\ \Big| > \frac{\delta}{N^{1+\kappa}} \ - \ \Big|\mathbb{E}\big[\tilde{D}_N(f) - D_N(f)\big]\ \Big|\ \bigg) \\
&\leqslant \ K_1 \exp\bigg(- \sqrt{k} K_2 N^{\kappa}(\delta - \ \Big|\mathbb{E}\big[\tilde{D}_N(f) - D_N(f)\big]\ \Big|)\bigg)
\end{aligned}$$

By (1.83), (1.84), Lemma 1.8.1 and the Borel-Cantelli lemma, $D_N(f)$ is almost surely of order $N^{1+\kappa}$ at most. □

Proposition 1.8.3. For every $\varepsilon > 0$, there exists N_0 such that for $N \geqslant N_0$

$$\mathrm{Sp}\Big(\ L(\mathbf{X}_N, \mathbf{Y}_N)\ \Big) \ \subset \mathrm{Sp}\Big(\ L(\mathbf{x}, \mathbf{y})\ \Big) + (-\varepsilon, \varepsilon) \tag{1.85}$$

Proof. By (1.11) and [AGZ10, Exercise 2.1.27], almost surely there exists $N_0 \in \mathbb{N}$ and $D \geqslant 0$ such that the spectral radii of the matrices $(\mathbf{X}_N, \mathbf{Y}_N)$ is bounded by D for all $N \geqslant N_0$. Hence, there exists $M \geqslant 0$ such that almost surely one has

$$\mathrm{Sp}\Big(\ L(\mathbf{X}_N, \mathbf{Y}_N)\ \Big) \subset [-M, M].$$

Let $f : \mathbb{R} \mapsto \mathbb{R}$ non negative, compactly supported, vanishing on $\mathrm{Sp}(\ L(\mathbf{x}, \mathbf{y})\ ) + (-\varepsilon/2, \varepsilon/2)$ and equal to one on $[-M, M] \setminus \Big(\mathrm{Sp}(\ L(\mathbf{x}, \mathbf{y})) + (-\varepsilon, \varepsilon)\ \Big)$. Then almost surely for N large enough, no eigenvalue of $L(\mathbf{X}_N, \mathbf{Y}_N)$ belongs to the complementary of $\mathrm{Sp}(\ L(\mathbf{x}, \mathbf{y})\ ) + (-\varepsilon, \varepsilon)$, since otherwise

$$(\tau_k \otimes \tau_N)\Big[f\Big(L(\mathbf{X}_N, \mathbf{Y}_N)\ \Big)\Big] \geqslant N^{-1} \geqslant N^{-1-\kappa}$$

in contradiction with Lemma 1.8.2. □

1.9 Proof of Corollaries 1.2.1, 1.2.2 and 1.2.4

1.9.1 Proof of Corollary 1.2.1: diagonal matrices

Let $\mathbf{D}_N = (D_1^{(N)}, \dots, D_q^{(N)})$ be as in Corollary 1.2.1. For any $j = 1, \dots, p$, the number of jump of F_j^{-1} is countable. We show that the convergence of the norm (1.17) holds when we chose $v = (v_1, \dots, v_q)$ in $[0,1]^q$ such that for any $k \neq \ell$ in $\{1, \dots, q\}$, the sets of jump points of $u \mapsto F_k^{-1}(u + v_k)$ and $u \mapsto F_\ell^{-1}(u + v_\ell)$ are disjoint. We show that for such a v, the family $\mathbf{D}_N^v$ satisfies the assumptions of Theorem 1.1.6. In all this section, we always denote $\lambda_i(j)$ instead of $\lambda_i^{(N)}(j)$ for any $i = 1, \dots, N$ and any $j = 1, \dots, q$.

The convergence of traces, case $v = (0, \dots, 0)$: Since the matrices commute, we only consider commutative polynomials. We start by showing that for all polynomials P,

$$\tau_N\Big[P(\mathbf{D}_N) \Big] \underset{N\to\infty}{\longrightarrow} \int_0^1 P\Big(F_1^{-1}(u), \dots, F_q^{-1}(u) \Big) du. \tag{1.86}$$

Denote by μ the probability distribution of the random variable $\Big(F_1^{-1}(U), \dots, F_q^{-1}(U) \Big) \in \mathbb{R}^q$, where U is distributed according to the uniform distribution on $[0,1]$. In order to get (1.86), we show that the sequence of measure in $\mathbb{R}^q$

$$\Big(\frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i(1)}, \dots, \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i(q)} \Big)$$

converges weakly to μ. This sequence is tight, since there exists a $B > 0$ such that for all $j = 1 \dots q$, for all $i = 1 \dots N$, one has $\lambda_i(j) \in [-B, B]$. Hence it is sufficient to show the following: for all real numbers $a_1, \dots, a_q$, for all $\varepsilon > 0$, there exists $\eta > 0$ such that

$$\begin{aligned} \limsup_{N\to\infty} \Big| \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{]-\infty, a_1+\eta]}\Big(\lambda_i(1)\Big) \times \dots \times \mathbf{1}_{]-\infty, a_q+\eta]}\Big(\lambda_i(q)\Big) \\ -\mu\Big( ]-\infty, a_1] \times \dots \times]-\infty, a_q] \Big) \Big| \leqslant \varepsilon. \end{aligned} \tag{1.87}$$

Fix $(a_1, \dots, a_q)$ in $\mathbb{R}^q$ and $\varepsilon > 0$. Remark that one has

$$\mu\Big( ]-\infty, a_1] \times \dots \times]-\infty, a_q] \Big) = \min_{j=1\dots q} F_j(a_j).$$

Let j_0 be an integer such that $F_{j_0}(a_{j_0}) = \mu\Big( ]-\infty, a_1] \times \dots \times]-\infty, a_q] \Big)$. For any $j = 1, \dots, q$, the empirical spectral distribution of $D_j^{(N)}$ converges to μ_j. Then for all a in $\mathbb{R}$ point of continuity for F_j, one has

$$\frac{1}{N} \sum_{i=1}^N \mathbf{1}_{]-\infty, a]}\Big(\lambda_i(j)\hat{\mathrm{E}}\Big) \underset{N\to\infty}{\longrightarrow} \mu_j\Big( ]-\infty, a] \Big). \tag{1.88}$$

Let $\eta > 0$ such that

- $\mu_{j_0}\Big( ]a_{j_0}, a_{j_0} + \eta] \Big) < \varepsilon/2$.
- for all $j = 1, \dots, q$, the real numbers $a_j + \eta$ and $a_{j_0} + \eta$ are points of continuity for F_j.

By (1.88) with $a = a_j + \eta$, there exists $N_0 \geqslant 1$ such that for all $N \geqslant N_0$ and $j = 1, \dots, q$, one has

$$F_j(a_j + \eta) - \varepsilon \leqslant \frac{1}{N} \mathrm{Card} \Big\{ i = 1 \dots N \;\Big|\; \lambda_i(j) \leqslant a_j + \eta \Big\}.$$

But $F_j(a_j + \eta) \geqslant F_j(a_j) \geqslant F_{j_0}(a_{j_0})$. Then we have

$$N\Big(F_{j_0}(a_{j_0}) - \varepsilon \Big) \leqslant \mathrm{Card} \Big\{ i = 1 \dots N \;\Big|\; \lambda_i(j) \leqslant a_j + \eta \Big\}.$$

The $\lambda_i(j)$ are non decreasing, so we get

$$\forall j = 1 \dots q,\ \forall i \leqslant N\Big(F_{j_0}(a_{j_0}) - \varepsilon\Big),\ \lambda_i(j) \leqslant a_j + \eta. \tag{1.89}$$

On the other hand, by (1.88) with $j = j_0$ and $a = a_{j_0} + \eta$, there exists $N_0 \geqslant 1$ such that, for all $N \geqslant N_0$, one has

$$\frac{1}{N}\text{Card}\left\{i = 1 \dots N \,\middle|\, \lambda_i(j_0) \leqslant a_{j_0} + \eta \right\} \leqslant F_{j_0}(a_{j_0} + \eta) + \varepsilon/2.$$

But $F_{j_0}(a_{j_0} + \eta) \leqslant F_{j_0}(a_{j_0}) + \varepsilon/2$, so that

$$\text{Card}\left\{i = 1 \dots N \,\middle|\, \lambda_i(j_0) \leqslant a_{j_0} + \eta \right\} \leqslant N\Big(F_{j_0}(a_{j_0}) + \varepsilon\Big).$$

The $\lambda_i(j_0)$ are non decreasing, then we get

$$\forall i \geqslant N\Big(F_{j_0}(a_{j_0}) + \varepsilon\Big),\ \lambda_i(j_0) \geqslant a_{j_0} + \eta. \tag{1.90}$$

By (1.89) and (1.90) we obtain: for all $N \geqslant N_0$

$$\left|\frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{]-\infty, a_1+\eta]}\Big(\lambda_i(1)\Big) \times \cdots \times \mathbf{1}_{]-\infty, a_q+\eta]}\Big(\lambda_i(q)\Big) - F_{j_0}(a_{j_0} + \eta)\right| \leqslant \varepsilon,$$

and then (1.87) is satisfied. So the convergence (1.86) holds when v is zero.

The convergence of traces, case v in $[0,1]^q$: To deduce the general case we shall need the following lemmas.

Lemma 1.9.1 (Quantiles of real diagonal matrices with sorted entries)**.** Let $D_N =$ diag $(\lambda_1, \dots, \lambda_N)$ be an $N \times N$ real diagonal matrix with non decreasing entries along its diagonal. Assume that the empirical eigenvalue distribution of D_N converges weakly to a compactly supported probability measure μ. Let F denote the cumulative distribution function of μ and F^{-1} its generalized inverse. Let v in $(0,1)$ a point of continuity for F^{-1} and $(i_N)_{N\geqslant 1}$ a sequence of integers, with i_N in $\{1, \dots, N\}$, such that i_N/N tends to v. Then, one has

$$\lambda_{i_N} \underset{N\to\infty}{\longrightarrow} F^{-1}(v).$$

In particular, we have the convergence of the quantile of order v:

$$\lambda_{1+\lfloor vN \rfloor} \underset{N\to\infty}{\longrightarrow} F^{-1}(v).$$

Proof. Denote $w = F^{-1}(v)$. We can always find $\eta \geqslant 0$, arbitrary small, such that $w - \eta$ and $w + \eta$ and points of continuity for F. Then, one has

$$\frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{]-\infty, w-\eta]}\Big(\lambda_i\Big) \underset{N\to\infty}{\longrightarrow} \mu\Big(\ ]-\infty, w-\eta]\ \Big) = F(w - \eta).$$

Then, the λ_i being non decreasing, for any $\varepsilon > 0$ there exists N_0 such that for any $N \geqslant N_0$, one has

$$\forall i \geqslant \Big(F(w-\eta)+\varepsilon\Big)N, \quad \lambda_i \geqslant w-\eta. \tag{1.91}$$

Since v is a point of continuity for F^{-1}, we get that $F(w-\eta) < v$. We chose $\varepsilon < v - F(w-\eta)$. Then, we get $F(w-\eta)+\varepsilon < v$. Hence, there exists N_0 such that, for any $N \geqslant N_0$, one has $i_N \geqslant \Big(F(w-\eta)+\varepsilon\Big)N$ and so, by (1.91): for any $\eta > 0$, there exists N_0 such that for all $N \geqslant N_0$, one has $w - \eta \leqslant \lambda_{i_N}$. Hence, we get for all $\eta > 0$,

$$w - \eta \leqslant \liminf_{N\to\infty} \lambda_{i_N}.$$

With the same reasoning, we get that

$$\limsup_{N\to\infty} \lambda_{i_N} \geqslant w + \eta,$$

and hence, letting η go to zero, we obtain the expected result. □

Lemma 1.9.2 (Truncation of real diagonal matrices with sorted entries)**.** Let $D_N =$ diag $(\lambda_1, \dots, \lambda_N)$ an $N \times N$ real diagonal matrix with non decreasing entries along its diagonal. Assume that the empirical eigenvalue distribution of D_N converges weakly to a compactly supported probability measure μ. For any $v_1 < v_2$ in $[0, 1]$, we set

$$D_N^{(v_1,v_2)} = \text{diag } (\lambda_{1+\lfloor v_1 N\rfloor}, \dots, \lambda_{\lfloor v_2 N\rfloor}).$$

Let F denote the cumulative distribution function of μ and F^{-1} its generalized inverse. We set $w_1 = F^{-1}(v_1)$, $w_2 = F^{-1}(v_2)$, $a_1 = F(w_1) - v_1$ and $a_2 = v_2 - F(w_2^-)$. Then, the empirical eigenvalue distribution of $D_N^{(v_1,v_2)}$ converges weakly the probability measure proportional to

$$a_1\delta_{w_1} + \mu\Big(\ \cdot \cap\ ]w_1, w_2[\ \Big) + a_2\delta_{w_2}.$$

Proof. We only show the lemma for $v_2 = 0$, the general case can be deduce by adapting the reasoning. We then use, for conciseness, the symbols v, w and a instead of v_1, w_1 and a_1 respectively.

If F is not continuous in w (i.e. if $\mu(w) \neq 0$) and $v \neq F(w)$, then for any α in $]0, (F(w)-v)/2[$, the map F^{-1} is continuous at the points $v+\alpha$ and $F(w)-\alpha$. By Lemma 1.9.1, we get that

$$\lim_{N\to\infty} \lambda_{1+\lfloor(v+\alpha)N\rfloor} = \lim_{N\to\infty} \lambda_{1+\lfloor(F(w)-\alpha)N\rfloor} = w. \tag{1.92}$$

Hence, for any continuous function f, we get

$$\frac{1}{N} \sum_{i=1+\lfloor(v+\alpha)N\rfloor}^{1+\lfloor(F(w)-\alpha)N\rfloor} f(\lambda_i) \underset{N\to\infty}{\longrightarrow} (a-2\alpha)f(w). \tag{1.93}$$

If F is continuous in w, we take $\alpha = 0$ in the following.

We can always find $\beta > 0$, arbitrary small, such that $F(w) + \beta$ is a point of continuity for F^{-1}. Remark that we then have

$$w = F^{-1}\Big(F(w)\Big) < F^{-1}\Big(F(w) + \beta\Big).$$

By Lemma 1.9.1, we get

$$\lambda_{1+\lfloor (F(w)+\beta)N \rfloor} \underset{N\to\infty}{\longrightarrow} F^{-1}\Big(F(w) + \beta\Big). \tag{1.94}$$

Moreover, we can always find γ in $]0, F^{-1}\Big(F(w) + \beta\Big) - w[$, arbitrary small, such that $w + \gamma$ is a point of continuity for F and $F(w + \gamma) < F(w) + \beta$. Then, by (1.94), we get that, for N large enough

$$\mathrm{Card}\Big\{\ i \geqslant 1 + \lfloor (F(w) - \alpha)N \rfloor\ \Big|\ \lambda_i \leqslant w + \gamma\ \Big\} \leqslant \lfloor (F(w) + \beta)N \rfloor - \lfloor (F(w) - \alpha)N \rfloor.$$

Hence, for any continuous function f, we get that for N large enough

$$\begin{aligned}
&\left| \frac{1}{N} \sum_{i=1+\lfloor (F(w)-\alpha)N \rfloor}^{N} f(\lambda_i) - \int_{]\omega,+\infty]} f(x) \mathrm{d}\mu(x) \right| \\
&\leqslant \left| \frac{1}{N} \sum_{i=1}^{N} f(\lambda_i) \mathbf{1}_{]w+\gamma,+\infty]}(\lambda_i) - \int_{]\omega,+\infty]} f(x) \mathrm{d}\mu(x) \right| \\
&\quad + \|f\|_\infty \frac{\lfloor (F(w) + \beta)N \rfloor - \lfloor (F(w) - \alpha)N \rfloor}{N}.
\end{aligned} \tag{1.95}$$

By (1.93) and (1.95), we obtain

$$\begin{aligned}
&\limsup_{N\to\infty} \left| \frac{1}{N} \sum_{i=1+\lfloor vN \rfloor}^{N} f(\lambda_i) - a f(w) - \int_{]\omega,+\infty]} f(x) \mathrm{d}\mu(x) \right| \\
&\leqslant \|f\|_\infty \Big(4\alpha + \beta + \mu\Big(]w, w + \gamma]\Big) \Big).
\end{aligned}$$

Letting α, β, γ go to zero, we get the result.

□

Let v in $[0, 1]^q$. We now show that, for any polynomial P, one has

$$\tau_N \Big[P(\mathbf{D}_N^v) \Big] \underset{N\to\infty}{\longrightarrow} \int_0^1 P\Big(F_1^{-1}(u + v_1), \ldots, F_q^{-1}(u + v_q) \Big) du. \tag{1.96}$$

At the possible price of relabeling the matrices, we assume $v_1 \geqslant \cdots \geqslant v_q$ and set

$$\begin{aligned}
N_1 &= N - \lfloor v_1 N \rfloor, \\
N_j &= \lfloor v_{j-1} N \rfloor - \lfloor v_j N \rfloor,\ \forall j = 1, \ldots, q.
\end{aligned}$$

For any $j = 1, \ldots, q$, we decompose the matrices $D_j^{(N)}(v_j)$ into

$$D_j^{(N)}(v_j) = \ \mathrm{diag}\ (D_{j,1}^{(N)}, \ldots, D_{j,q}^{(N)}),$$

where for any $i = 1, \dots, q$, the matrix $D_{j,i}^{(N)}$ is $N_i \times N_i$. We set for any $i = 1, \dots, q$, the family $\mathbf{D}_N(i) = (D_{1,i}^{(N)}, \dots, D_{q,i}^{(N)})$. For any $i, j = 1, \dots, q$, we denote by $F_{i,j}$ the cumulative distribution function of the measure obtained in Lemma 1.9.2 with (D_N, μ, v_1, v_2) replaced by $(D_j^{(N)}, \mu_j, v_{i-1}, v_i)$. Then, for any polynomial P, one as

$$\tau_N[P(\mathbf{D}_N^v)] = \sum_{i=1}^q \frac{N_i}{N} \tau_{N_i}[P(\mathbf{D}_N(i))].$$

By Lemma 1.9.2 and by the case $v = (0, \dots, 0)$, we deduce that

$$\tau_{N_i}[P(\mathbf{D}_N(i))] \underset{N\to\infty}{\longrightarrow} \frac{1}{v_{q-1} - v_q} \int_{v_q}^{v_{q-1}} P\Big(F_{i,1}^{-1}(u + v_1), \dots, F_{i,q}^{-1}(u + v_q) \Big) du,$$

with the convention $v_0 = 1$. The merge of the different terms for $i = 1, \dots, q$ gives as expected

$$\tau_N\Big[P(\mathbf{D}_N^v) \Big] \underset{N\to\infty}{\longrightarrow} \int_0^1 P\Big(F_1^{-1}(u + v_1), \dots, F_q^{-1}(u + v_q) \Big) du. \tag{1.97}$$

The convergence of norms: Let $v = (v_1, \dots, v_q)$ in $[0, 1]^q$ such that for any $k \neq \ell$ in $\{1, \dots, q\}$, the sets of jump points of $u \mapsto F_k^{-1}(u + v_k)$ and $u \mapsto F_\ell^{-1}(u + v_\ell)$ are disjoint. We now show that, for all polynomials P, one has

$$\|P(\mathbf{D}_N^v)\| \underset{N\to\infty}{\longrightarrow} \underset{\text{Supp } \mu^v}{\text{Sup}} \Big| P \Big|,$$

where μ^v is the probability distribution of the random variable $\Big(F_1^{-1}(U + v_1), \dots, F_q^{-1}(U + v_q) \Big) \in \mathbb{R}^q$, where U is distributed according to the uniform distribution on $[0, 1]$. In view of the above, we have

$$\liminf \ \|P(\mathbf{D}_N^v)\| \geqslant \underset{\text{Supp } \mu^v}{\text{Sup}} \Big| P \Big|.$$

It is sufficient then to show that, for any $\eta > 0$, there exists $N_0 \geqslant N$ such that for all $i = 1, \dots, N$, one has

$$\Big(\lambda_{i + \lfloor v_1 N \rfloor}(1), \dots, \lambda_{i + \lfloor v_q N \rfloor}(q) \Big) \in \text{Supp } \mu^v + (-\eta, \eta)^q. \tag{1.98}$$

Indeed, by uniform continuity, for any polynomial P and $\varepsilon > 0$, there exists $\eta \geqslant 0$ such that, for all $(x_1, \dots, x_q)$ in Supp $\mu^v + [-1, 1]^q$ and $(y_1, \dots, y_q)$ in $\mathbb{R}^q$, one has

$$|y_j - x_j| < \eta \ \Rightarrow \ \Big| P(x_1, \dots, x_q) - P(y_1, \dots, y_q) \Big| < \varepsilon$$

and hence: for all $\varepsilon > 0$, there exist $\eta \geqslant 0$ and $N_0 \geqslant 1$ such that for all $N \geqslant N_0$, for all $i = 1, \dots, N$

$$\max_{i=1\dots N} \Big| P\Big(\lambda_{i + \lfloor v_1 N \rfloor}(1), \dots, \lambda_{i + \lfloor v_q N \rfloor}(q) \Big) \Big| \leqslant \max_{\text{Supp } \mu^v + (-\eta, \eta)^q} \Big| P \Big| \leqslant \max_{\text{Supp } \mu^v} |P| + \varepsilon.$$

Suppose that (1.98) is not true: there exist $\eta > 0$ and $(N_k)_{k \geqslant 1}$ an increasing sequence of positive integer such that for all $k \geqslant 1$, there exists i_k such that

$$\Big(\lambda_{i_k + \lfloor v_1 N_k \rfloor}^{(N_k)}(1), \dots, \lambda_{i_k + \lfloor v_q N_k \rfloor}^{(N_k)}(q) \Big) \notin \text{Supp } \mu^v + (-\eta, \eta)^q.$$

By compactness, one can always assume that i_k/N_k converges to u_0 in $[0,1]$. For all j in $\{1,\dots,q\}$ except a possible j_0, we have that u_0+v_j is a point of continuity for F_j^{-1} and so, by Lemma 1.9.1, $\lambda^{(N_k)}_{i_k+\lfloor v_j N_k\rfloor}(j)$ converges to $F_j^{-1}(u_0+v_j)$. Recall that

$$\text{Supp } \mu^v = \Big\{ \Big(F_1^{-1}(u+v_1), \dots, F_q^{-1}(u+v_q) \Big) \ \Big| \ u \in [0,1] \Big\}.$$

Then we have, for N large enough and for all u in $[0,1]$, that $\Big| \lambda^{(N_k)}_{i_k+\lfloor v_{j_0} N_k\rfloor}(j_0) - F_{j_0}^{-1}(u+v_{j_0}) \Big| > \eta$ i.e.

$$\text{dist } \Big(\lambda^{(N_k)}_{i_k+\lfloor v_{j_0} N_k\rfloor}(j_0), \text{Supp } \mu_{j_0} \Big) > \eta,$$

which is in contradiction with the fact that for N large enough the eigenvalues of $D_{j_0}^{(N)}$ belong to a small neighborhood of the support of μ_{j_0}.

1.9.2 Proof of Corollary 1.2.2: Wishart matrices

Let $r, s_1, \dots, s_p \geqslant 1$ and $(\mathbf{W}_N, \mathbf{Y}_N)$ be as in Corollary 1.2.2 and denote $s = s_1 + \dots + s_p$. We use matrix manipulations in order to see the norm of a polynomial in the $rN \times rN$ matrices $\mathbf{W}_N, \mathbf{Y}_N, \mathbf{Y}_N^*$ as the norm of a polynomial in $(r+s)N \times (r+s)N$ matrices $\tilde{\mathbf{X}}_N, \tilde{\mathbf{Y}}_N, \tilde{\mathbf{Y}}_N^*, \tilde{\mathbf{Z}}_N$ and some elementary matrices, where $\tilde{\mathbf{X}}_N$ is a family of independent GUE matrices and $\tilde{\mathbf{Y}}_N, \tilde{\mathbf{Z}}_N$ are modifications of $\mathbf{Y}_N, \mathbf{Z}_N$. We will obtain the result as a consequence of Theorem 1.1.6.

Define the $(r+s)N \times (r+s)N$ matrices $\mathbf{e}_N = (e_0^{(N)}, e_1^{(N)}, \dots, e_p^{(N)})$:

$$e_0^{(N)} = \begin{pmatrix} \mathbf{1}_{rN} & \mathbf{0}_{rN,sN} \\ \mathbf{0}_{sN,rN} & \mathbf{0}_{sN} \end{pmatrix}, \tag{1.99}$$

$$e_j^{(N)} = \begin{pmatrix} \mathbf{0}_{rN} & & & \\ & \mathbf{0}_{(s_1+\dots+s_{j-1})N} & & \\ & & \mathbf{1}_{s_j N} & \\ & & & \mathbf{0}_{(s_{j+1}+\dots+s_p)N} \end{pmatrix}. \tag{1.100}$$

for $j = 1, \dots, p$. Recall that by definition of the Wishart matrix model for $j = 1, \dots, p$

$$W_j^{(N)} = M_j^{(N)} Z_j^{(N)} M_j^{(N)*}, \tag{1.101}$$

where $M_j^{(N)}$ is an $rN \times s_j N$ complex Gaussian matrix with independent identically distributed entries, centered and of variance $1/rN$. Let $\tilde{\mathbf{X}}_N = (\tilde{X}_1^{(N)}, \dots, \tilde{X}_p^{(N)})$ be a family of p independent, normalized GUE matrices of size $(r+s)N \times (r+s)N$, independent of $\mathbf{Y}_N$ and $\mathbf{Z}_N$ and such that for $j = 1, \dots, p$, the $rN \times s_j N$ matrix $M_j^{(N)}$ appears as a sub-matrix of $\sqrt{\frac{r+s}{r}} \tilde{X}_j^{(N)}$ in the following way: if we denote $\tilde{M}_j^{(N)} = \sqrt{\frac{r+s}{r}} e_0^{(N)} \tilde{X}_j^{(N)} e_j^{(N)}$ then

$$\tilde{M}_j^{(N)} = \begin{pmatrix} \mathbf{0}_{rN} & & M_j^{(N)} & \\ & \mathbf{0}_{(s_1+\dots+s_{j-1})N} & & \\ & & \mathbf{0}_{s_j N} & \\ & & & \mathbf{0}_{(s_{j+1}+\dots+s_p)N} \end{pmatrix}. \tag{1.102}$$

Let $\tilde{\mathbf{Y}}_N = (\tilde{Y}_1^{(N)}, \dots, \tilde{Y}_q^{(N)})$ and $\tilde{\mathbf{Z}}_N = (\tilde{Z}_1^{(N)}, \dots, \tilde{Z}_p^{(N)})$ be the families of $(r+s)N \times (r+s)N$ matrices defined by:

$$\tilde{Y}_j^{(N)} = \begin{pmatrix} Y_j^{(N)} & \mathbf{0}_{rN,sN} \\ \mathbf{0}_{sN,rN} & \mathbf{0}_{sN} \end{pmatrix}, \quad j = 1, \dots, q, \tag{1.103}$$

$$\tilde{Z}_j^{(N)} = \begin{pmatrix} \mathbf{0}_{rN} & & & \\ & \mathbf{0}_{(s_1+\dots+s_{j-1})N} & & \\ & & Z_j^{(N)} & \\ & & & \mathbf{0}_{(s_{j+1}+\dots+s_p)N} \end{pmatrix}, \quad j = 1, \dots, p. \tag{1.104}$$

By assumption, with probability one the non commutative law of $\mathbf{Y}_N$ converges to the law of non commutative random variables $\mathbf{y} = (y_1, \dots, y_q)$ in a $\mathcal{C}^*$-probability space $(\mathcal{A}_0, .^*, \tau, \| \cdot \|)$ and for $j = 1 \dots p$ the non commutative law of Z_j converges to the law of a non commutative random variable z_j in a $\mathcal{C}^*$-probability space $(\mathcal{A}_j, .^*, \tau, \| \cdot \|)$ (we use the same notations for the functionals in the different spaces). All the traces under consideration are faithful. Let $\mathcal{B}$ denotes the product algebra $\mathcal{B}_0 \times \mathcal{B}_1 \times \dots \times \mathcal{B}_p$. We equip $\mathcal{B}$ with the involution $.^*$ and the trace $\tilde{\tau}$ defined by: for all $(b_0, \dots, b_p)$ in $\mathcal{B}$

$$(b_0, \dots, b_p)^* = (b_0^*, \dots, b_p^*),$$

$$\tilde{\tau}\Big[\ (b_0, \dots, b_p)\ \Big] = \frac{r}{r+s}\tau(b_0) + \frac{s_1}{r+s}\tau(b_1) + \dots + \frac{s_p}{r+s}\tau(b_p).$$

The trace $\tilde{\tau}$ is a faithful tracial state on $\mathcal{B}$. Equipped with $.^*$, $\tilde{\tau}$ and with the norm $\| \cdot \|$ defined by (1.9), the algebra $\mathcal{B}$ is a $\mathcal{C}^*$-probability space. Define $\tilde{\mathbf{y}} = (\tilde{y}_1, \dots, \tilde{y}_q)$, $\tilde{\mathbf{z}} = (\tilde{z}_1, \dots, \tilde{z}_q)$ and $\mathbf{e} = (e_0, \dots, e_p)$ by

$$\tilde{y}_j = (y_j, \mathbf{0}_{\mathcal{B}_1}, \dots, \mathbf{0}_{\mathcal{B}_p}),\ j = 1, \dots, q,$$

$$\tilde{z}_j = (\mathbf{0}_{\mathcal{B}_0}, \dots, \mathbf{0}_{\mathcal{B}_{j-1}}, z_j, \mathbf{0}_{\mathcal{B}_{j+1}}, \dots, \mathbf{0}_{\mathcal{B}_p}),\ j = 1, \dots, p,$$

$$e_j = (\mathbf{0}_{\mathcal{B}_0}, \dots, \mathbf{0}_{\mathcal{B}_{j-1}}, \mathbf{1}_{\mathcal{B}_j}, \mathbf{0}_{\mathcal{B}_{j+1}}, \dots, \mathbf{0}_{\mathcal{B}_p}),\ j = 0, \dots, q.$$

Lemma 1.9.3. With probability one, the non commutative law of $(\tilde{\mathbf{Y}}_N, \tilde{\mathbf{Z}}_N, \mathbf{e}_N)$ in $(\mathrm{M}_{(r+s)N}(\mathbb{C}), .^*, \tau_{(r+s)N})$ converges to the law of $(\tilde{\mathbf{y}}, \tilde{\mathbf{z}}, \mathbf{e})$ in $(\mathcal{B}, .^*, \tilde{\tau})$.

Proof. Let P be a polynomial in $2p + 2q + 1$ non commutative indeterminates:

$$\begin{aligned}
&\tau_{(r+s)N}\Big[P(\tilde{\mathbf{Y}}_N, \tilde{\mathbf{Y}}_N^*, \mathbf{Z}_N, \mathbf{e}_N)\ \Big] \\
&= \quad \frac{r}{r+s}\tau_{rN}\Big[P(\tilde{\mathbf{Y}}_N, \tilde{\mathbf{Y}}_N^*, \underbrace{\mathbf{0}_{rN}, \dots, \mathbf{0}_{rN}}_{p}, \mathbf{1}_{rN}, \underbrace{\mathbf{0}_{rN}, \dots, \mathbf{0}_{rN}}_{p})\ \Big] \\
&+ \quad \sum_{j=1}^p \frac{s_j}{s+r}\tau_{s_j}\Big[P(\underbrace{\mathbf{0}_{s_jN}, \dots, \mathbf{0}_{s_jN}}_{2q+j-1}, Z_j^{(N)}, \underbrace{\mathbf{0}_{s_jN}, \dots, \mathbf{0}_{s_jN}}_{p}, \mathbf{1}_{s_jN}, \underbrace{\mathbf{0}_{s_jN}, \dots, \mathbf{0}_{s_jN}}_{p-j})\ \Big] \\
&\underset{N\to\infty}{\longrightarrow} \quad \frac{r}{r+s}\tau\Big[P(\mathbf{y}, \mathbf{y}^*, \underbrace{\mathbf{0}, \dots, \mathbf{0}}_{p}, \mathbf{1}, \underbrace{\mathbf{0}, \dots, \mathbf{0}}_{p})\ \Big] \\
&\qquad + \quad \sum_{j=1}^p \frac{s_j}{s+r}\tau\Big[P(\underbrace{\mathbf{0}, \dots, \mathbf{0}}_{2q+j-1}, z_j, \hat{\mathrm{E}}\,\underbrace{\mathbf{0}, \dots, \mathbf{0}}_{p}, \mathbf{1}, \underbrace{\mathbf{0}, \dots, \mathbf{0}}_{p-1})\ \Big] \qquad (1.105) \\
&= \quad \tilde{\tau}[P(\tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*, \tilde{\mathbf{z}}, \mathbf{e})\ ], \qquad (1.106)
\end{aligned}$$

where the convergence holds almost surely since each term of the sum converges almost surely. □

Lemma 1.9.4. For all polynomials P in $2p+2q+1$ non commutative indeterminates, almost surely

$$\left\| P(\tilde{\mathbf{Y}}_N, \tilde{\mathbf{Y}}_N^*, \mathbf{Z}_N, \mathbf{e}_N) \right\| \underset{N\to\infty}{\longrightarrow} \| P(\tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*, \tilde{\mathbf{z}}, \mathbf{e}) \|.$$

Proof. Lemma 1.9.4 follows easily since for any polynomial P in $2p+2q+1$ non commutative indeterminates, $\left\| P(\tilde{\mathbf{Y}}_N, \tilde{\mathbf{Y}}_N^*, \mathbf{Z}_N, \mathbf{e}_N) \right\|$ is the maximum of the $p+1$ real numbers

- $\| P(\tilde{\mathbf{Y}}_N, \tilde{\mathbf{Y}}_N^*, \underbrace{\mathbf{0}_{rN}, \dots, \mathbf{0}_{rN}}_{p}, \mathbf{1}_{rN}, \underbrace{\mathbf{0}_{rN}, \dots, \mathbf{0}_{rN}}_{p}) \|$,
- $\| P(\underbrace{\mathbf{0}_{s_jN}, \dots, \mathbf{0}_{s_jN}}_{2q+j-1}, Z_j^{(N)}, \underbrace{\mathbf{0}_{s_jN}, \dots, \mathbf{0}_{s_jN}}_{p}, \mathbf{1}_{s_jN}, \underbrace{\mathbf{0}_{s_jN}, \dots, \mathbf{0}_{s_jN}}_{p-j}) \|$, $j = 1, \dots, p$,

and $\| P(\tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*, \tilde{\mathbf{z}}, \mathbf{e}) \|_{\tilde{\tau}}$ is the maximum of the $p+1$ real numbers

- $\| P(\mathbf{y}, \mathbf{y}^*, \underbrace{\mathbf{0}, \dots, \mathbf{0}}_{p}, \mathbf{1}, \underbrace{\mathbf{0}, \dots, \mathbf{0}}_{p}) \|$,
- $\| P(\underbrace{\mathbf{0}, \dots, \mathbf{0}}_{2q+j-1}, z_j, \underbrace{\mathbf{0}, \dots, \mathbf{0}}_{p}, \mathbf{1}, \underbrace{\mathbf{0}, \dots, \mathbf{0}}_{p-1}) \|$, $j = 1, \dots, p$.

□

Let $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_p)$ be a free semicircular system in $\mathcal{C}^*$-probability space. Let $\tilde{\mathcal{A}}$ be the reduced free product $\mathcal{C}^*$-algebra of $\mathcal{B}$ and the $\mathcal{C}^*$-algebra spanned by $\tilde{\mathbf{x}}$. We still denotes by $\tilde{\tau}$ the trace on $\tilde{\mathcal{A}}$ and the norm considered $\| \cdot \|$ is given by (1.9) since the trace is faithful. By Voiculescu's theorem and by the independence of $\tilde{\mathbf{X}}_N$ and $(\tilde{\mathbf{Y}}_N, \tilde{\mathbf{Z}}_N)$, with probability one the non commutative law of $(\tilde{\mathbf{X}}_N, \tilde{\mathbf{Y}}_N, \tilde{\mathbf{Z}}_N, \mathbf{e}_N)$ in $(\mathrm{M}_{(r+s)N}(\mathbb{C}), .^*, \tau_{(r+s)N})$ converges to the non commutative law of $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \tilde{\mathbf{z}}, \mathbf{e})$ in $(\tilde{\mathcal{A}}, .^*, \tilde{\tau})$. Define the non commutative random variables $\tilde{\mathbf{m}} = (\tilde{m}_1, \dots, \tilde{m}_q)$ and $\tilde{\mathbf{w}} = (\tilde{w}_1, \dots, \tilde{w}_q)$ in $\tilde{\mathcal{A}}$ by: for $j = 1, \dots, q$,

$$\tilde{m}_j = \sqrt{\frac{r+s}{r}} e_0 \tilde{x}_j e_j, \quad \tilde{w}_j = e_0(\tilde{m}_j \ \tilde{z}_j + \tilde{m}_j^*)^2. \tag{1.107}$$

Lemma 1.9.5. For any polynomial P in $p+2q$ non commutative indeterminates, there exists a polynomial $\tilde{P}$ in $3p+2q+1$ non commutative indeterminates, such that one has

$$\begin{pmatrix} P(\mathbf{W}_N, \mathbf{Y}_N, \mathbf{Y}_N^*) & \mathbf{0}_{rN,sN} \\ \mathbf{0}_{sN,rN} & \mathbf{0}_{sN} \end{pmatrix} = \tilde{P}(\tilde{\mathbf{X}}_N, \tilde{\mathbf{Y}}_N, \tilde{\mathbf{Y}}_N^*, \tilde{\mathbf{Z}}_N, \mathbf{e}_N), \tag{1.108}$$

$$e_0 P(\tilde{\mathbf{w}}, \tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*) = \tilde{P}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*, \tilde{\mathbf{z}}, \mathbf{e}).$$

Proof. We set $\tilde{\mathbf{W}}_N = (W_1^{(N)}, \dots, W_p^{(N)})$ given by: for $j = 1, \dots, p$,

$$\tilde{W}_j^{(N)} := e_0^{(N)} (\tilde{M}_j^{(N)} \tilde{Z}_j^{(N)} + \tilde{M}_j^{(N)*})^2 = \begin{pmatrix} W_j^{(N)} & \mathbf{0}_{rN,sN} \\ \mathbf{0}_{sN,rN} & \mathbf{0}_{sN} \end{pmatrix}. \tag{1.109}$$

Let P be a polynomial in $p+2q$ non commutative indeterminates. By the block decomposition of $\tilde{\mathbf{W}}_N$ and $\tilde{\mathbf{Y}}_N$, one has

$$\begin{pmatrix} P(\mathbf{W}_N, \mathbf{Y}_N, \mathbf{Y}_N^*) & \mathbf{0}_{rN,sN} \\ \mathbf{0}_{sN,rN} & \mathbf{0}_{sN} \end{pmatrix} = e_0^{(N)} \ P(\tilde{\mathbf{W}}_N, \tilde{\mathbf{Y}}_N, \tilde{\mathbf{Y}}_N^*).$$

Furthermore, By definitions of $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{W}}$: for $j=1,\dots,p$

$$\begin{aligned} \tilde{W}_j^{(N)} &= e_0^{(N)}(\tilde{M}_j^{(N)}\tilde{Z}_j^{(N)} + \tilde{M}_j^{(N)*})^2 \\ &= e_0^{(N)}\frac{r+s}{r}(e_0^{(N)}\tilde{X}_j^{(N)}e_j^{(N)}\tilde{Z}_j^{(N)} + e_j^{(N)}\tilde{X}_j^{(N)}e_0^{(N)})^2. \end{aligned}$$

Define for $j=1,\dots,p$ the non commutative polynomial P_j deduced by the formula

$$P_j(\tilde{x}_j, \tilde{z}_j, \mathbf{e}) = e_0\frac{r+s}{r}(e_0\tilde{x}_j e_j \tilde{z}_j + e_j\tilde{x}_j e_0)^2, \tag{1.110}$$

and define $\tilde{P}$ deduced by

$$\tilde{P}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*, \tilde{\mathbf{z}}, \mathbf{e}) = e_0 \ P\Big(P_1(\tilde{x}_1, \tilde{z}_1, \mathbf{e}), \dots, P_p(\tilde{x}_p, \tilde{z}_p, \mathbf{e}), \tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*\Big). \tag{1.111}$$

The polynomials are defined without ambiguity if $\tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*, \tilde{\mathbf{z}}, \mathbf{e}$ are seen as families of non commutative indeterminates (without any algebraic relation) instead of non commutative random variables. Remark that, by definition, for all $j = 1,\dots,p$ the non commutative random variable w_j equals $P_j(\tilde{x}_j, \tilde{z}_j, \mathbf{e})$. Hence it follows as expected that

$$\begin{aligned} \begin{pmatrix} P(\mathbf{W}_N, \mathbf{Y}_N, \mathbf{Y}_N^*) & \mathbf{0}_{rN,sN} \\ \mathbf{0}_{sN,rN} & \mathbf{0}_{sN} \end{pmatrix} &= \tilde{P}(\tilde{\mathbf{X}}_N, \tilde{\mathbf{Y}}_N, \tilde{\mathbf{Y}}_N^*, \tilde{\mathbf{Z}}_N, \mathbf{e}_N), \\ e_0 P(\tilde{\mathbf{w}}, \tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*) &= \tilde{P}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*, \tilde{\mathbf{z}}, \mathbf{e}). \end{aligned}$$

□

It is well known as a generalization of Voiculescu's theorem that, under Assumption 1 separately for $Z_1^{(N)}, , \dots, Z_p^{(N)}, \mathbf{Y}_N$ and by independence of the families, with probability one the non commutative law of $(\mathbf{W}_N, \mathbf{Y}_N)$ in $(\mathrm{M}_N(\mathbb{C}), .^*, \tau_N)$ converges to the non commutative law of $(\mathbf{w}, \mathbf{y})$ in a $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ with faithful trace, where

1. $\mathbf{w} = (w_1, \dots, w_p)$ are free selfadjoint non commutative random variables,
2. $\mathbf{y} = (y_1, \dots, y_q)$ is the limit in law of $\mathbf{Y}_N$,
3. $\mathbf{w}$ and $\mathbf{y}$ are free.

For any polynomial P in $p+2q$ non commutative indeterminates

$$\begin{aligned} \tau[P(\mathbf{w}, \mathbf{y}, \mathbf{y}^*)] &= \lim_{N\to\infty} \tau_{rN}\Big[P(\mathbf{W}_N, \mathbf{Y}_N, \mathbf{Y}_N^*) \ \Big] \\ &= \lim_{N\to\infty} \frac{r+s}{r}\tau_{(r+s)N}\left[\begin{pmatrix} P(\mathbf{W}_N, \mathbf{Y}_N, \mathbf{Y}_N^*) & \mathbf{0}_{rN,sN} \\ \mathbf{0}_{sN,rN} & \mathbf{0}_{sN} \end{pmatrix}\right] \\ &= \lim_{N\to\infty} \frac{r+s}{r}\tau_{(r+s)N}\Big[\tilde{P}(\tilde{\mathbf{X}}_N, \tilde{\mathbf{Y}}_N, \tilde{\mathbf{Y}}_N^*, \tilde{\mathbf{Z}}_N, \mathbf{e}_N) \ \Big] \\ &= \frac{r+s}{r}\tilde{\tau}\Big[\tilde{P}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*, \tilde{\mathbf{z}}, \mathbf{e}) \ \Big] \\ &= \frac{r+s}{r}\tilde{\tau}\Big[e_0 P(\tilde{\mathbf{w}}, \tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*) \ \Big], \end{aligned}$$

where the limits are almost sure. In particular we obtain that, for all polynomials P in $p+2q$ non commutative indeterminates, one has

$$\|e_0 P(\tilde{\mathbf{w}}, \tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*)\| = \|P(\mathbf{w}, \mathbf{y}, \mathbf{y}^*)\|. \tag{1.112}$$

By Lemmas 1.9.3 and 1.9.4, the family of $(r+s)N\times(r+s)N$ matrices $(\tilde{\mathbf{Y}}_N, \tilde{\mathbf{Z}}_N, \mathbf{e}_N)$ satisfies the assumptions of Theorem 1.1.6, hence for all polynomials P in $3p+2q+1$ non commutative indeterminates, with $\tilde{P}$ as in Lemma 1.9.5, almost surely one has

$$\|\tilde{P}(\tilde{\mathbf{X}}_N, \tilde{\mathbf{Y}}_N, \tilde{\mathbf{Y}}_N^*, \tilde{\mathbf{Z}}_N, \mathbf{e}_N)\| \underset{N\to\infty}{\longrightarrow} \|\tilde{P}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*, \tilde{\mathbf{z}}, \mathbf{e})\|. \tag{1.113}$$

Remark that

$$\begin{aligned} \|P(\mathbf{W}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)\| &= \left\| \begin{pmatrix} P(\mathbf{W}_N, \mathbf{Y}_N, \mathbf{Y}_N^*) & \mathbf{0}_{rN,sN} \\ \mathbf{0}_{sN,rN} & \mathbf{0}_{sN} \end{pmatrix} \right\| \\ &= \|\tilde{P}(\tilde{\mathbf{X}}_N, \tilde{\mathbf{Y}}_N, \tilde{\mathbf{Y}}_N^*, \tilde{\mathbf{Z}}_N, \mathbf{e}_N)\|, \end{aligned}$$

$$\|\tilde{P}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*, \tilde{\mathbf{z}}, \mathbf{e})\| = \|e_0 P(\tilde{\mathbf{w}}, \tilde{\mathbf{y}}, \tilde{\mathbf{y}}^*)\| = \|P(\mathbf{w}, \mathbf{y}, \mathbf{y}^*)\|.$$

Together with (1.113), this gives the expected result.

1.9.3 Proof of Corollary 1.2.4: Rectangular band matrices

We only give a sketch of the proof. Details are obtained by minor modification of the proofs of Corollaries 1.2.2 and 1.2.3. Let H be as in Corollary 1.2.4:

$$H = \begin{pmatrix} A_1 & A_2 & \dots & A_L & \mathbf{0} & \dots & & \dots & \mathbf{0} \\ \mathbf{0} & A_1 & A_1 & \dots & A_L & \mathbf{0} & & & \vdots \\ \vdots & \mathbf{0} & A_1 & A_2 & \dots & A_L & \mathbf{0} & & \\ & & \ddots & \ddots & \ddots & & \ddots & \vdots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & & \dots & \mathbf{0} & A_1 & A_2 & \dots & A_L \end{pmatrix}. \tag{1.114}$$

We start with the following observation: the operator norm of H is the square root of the operator norm of H^*H, which is a square block matrix. Its blocks consist of sums of $tN \times tN$ matrices of the form $A_l^* A_m$, $l, m = 1 \dots L$. By minor modifications of the proof of Corollary 1.2.2, we get the almost sure convergence of the normalized trace and of the norm for any polynomial in the matrices $\mathbf{A}_N = (A_l^* A_m)_{l,m=1..L}$ as N goes to the infinity. By Proposition 1.7.3, we get that the convergences hold for square block matrices and in particular for any polynomial in H^*H. Hence the result follows by functional calculus.

1.10 A theorem about norm convergence, by D. Shlyakhtenko [1]

Lemma Let (A,τ) be a C^*-algebra with a faithful trace τ, and consider B to be the universal C^*-algebra generated by A and elements $L^{(1)},\dots,L^{(n)}$ satisfying $L^{(i)*}xL^{(j)}=\delta_{i=j}\tau(x)$ for all $x\in A$. Moreover, consider the linear functional ψ determined on $*-\mathrm{Alg}(A,\{L^{(j)}\}_j)$ by:
$\psi|_A=\tau$,
$\psi(x_0L^{(i_1)}x_1\cdots x_{k-1}L^{(i_k)}x_ky_0L^{(j_1)*}y_1\cdots y_{l-1}L^{(j_l)*}y_l) = 0$ whenever $x_1,\dots,x_k,y_0,\dots,y_l\in A$ and at least one of k and l is nonzero.

Then ψ extends to a state on B having a faithful GNS representation. Moreover, $(B,\psi)\cong(A,\tau)*(\mathcal{E},\phi)$ where $(\mathcal{E},\phi)$ is the C^*-algebra generated by n free creation operators $\ell_1,\dots,\ell_n$ on the full Fock space $\mathcal{F}(\mathbb{C}^n)$ and ϕ is the vacuum expectation.

Sketch of proof. Consider the A,A-Hilbert bimodule $\mathcal{H}=L^2(A,\tau)\otimes A$ with the inner product

$$\langle\xi\otimes a,\xi'\otimes a'\rangle_A=\langle\xi,\xi'\rangle_{L^2(\tau)}a^*a'$$

and the left and right A actions given by

$$x\cdot(\xi\otimes a)\cdot y=x\xi\otimes ay.$$

Let B be the extended Cuntz-Pimsner algebra associated to $\mathcal{H}^{\oplus n}$ (see [Pim97]), i.e. the universal C^*-algebra generated by A and operators $L_h : h\in\mathcal{H}$ satisfying the relations

$$\begin{aligned} L_h^*L_g &= \langle h,g\rangle_A, & h,g\in\mathcal{H}^{\oplus n}\\ aL_hb &= L_{ahb}, & h\in\mathcal{H}^{\oplus n},\ a,b\in A. \end{aligned}$$

It follows from the results of [Shl98] that if we denote by $(\hat{B},\hat{\psi})$ the free product $(A,\tau)*(\mathcal{E},\phi)$, then:

$$\begin{gathered} \ell_i^*x\ell_j=\delta_{i=j}\tau(x), \qquad \forall x\in A,\\ \hat{\psi}(x_0\ell_{i_1}x_1\cdots x_{k-1}\ell_{i_k}x_ky_0\ell_{j_1}^*y_1\cdots y_{l-1}\ell_{j_l}^*y_l)=0,\\ \forall x_1,\dots,x_k,y_1,\dots,y_l\in A,\ k+l>0 \end{gathered}$$

If $h=(\sum_i\xi_i^{(k)}\otimes a_i^{(k)})_{k=1}^n\in(A\otimes A)^{\oplus n}\subset\mathcal{H}^{\oplus n}$ is a finite tensor, write

$$\ell_h=\sum_{k,i}\xi_i^{(k)}\ell_ka_i^{(k)}.$$

It then follows that

$$\begin{aligned} \ell_h^*\ell_g &= \langle h,g\rangle_A, & h,g\in\mathcal{H}^{\oplus n}\\ a\ell_hb &= \ell_{ahb}, & a,b\in A,\ h\in\mathcal{H}^{\oplus n} \end{aligned}$$

1. Research supported by NSF grant DMS-0900776

which in particular means that $\|\ell_h\|_2^2 = \|\ell_h^*\ell_h\| = \|h\|^2$ so that the mapping $h \mapsto \ell_h$ is an isometry. We then extend ℓ to a map from $\mathcal{H}^{\oplus n}$ into $\hat{B}$. Note that the extension of ℓ still satisfies $a\ell_h b = \ell_{ahb}$ whenever $a, b \in A$ and $h \in \mathcal{H}^{\oplus n}$.

From this we see that (by the universal property of B) there exists a $*$-homomorphism $\pi : B \to \hat{B}$, so that $\psi = \hat{\psi} \circ \pi$. Thus all we need to prove is that π is injective. But by [Pim97, Prop. 3.3], it follows that B is isomorphic to the Toeplitz algebra $\mathcal{T}$ (since in this case obviously $\langle \mathcal{H}^{\oplus n}, \mathcal{H}^{\oplus n}\rangle_A = A$) acting on the Fock space $\mathcal{F} = \bigoplus_{k\geqslant 0}(\mathcal{H}^{\oplus n})^{\otimes_A k}$. If we denote by E the canonical conditional expectation from $\mathcal{T}$ onto A and consider the state $\theta = \tau \circ E$, then the resulting Hilbert space is the closure of $\mathcal{F}$ in the (faithful) norm $\|\xi\| = \tau(\langle \xi, \xi\rangle_A)^{1/2}$; from this we see that the GNS representation of B associated to the state θ on B is faithful. Since $\hat{B}$ is exactly this GNS representation, it follows that π is injective. □

If A_N is a sequence of C^*-algebras and $\omega \in \beta\mathbb{N} \setminus \mathbb{N}$ is a free ultrafilter, we shall denote by

$$\mathfrak{A} = \prod^{\omega} A_N$$

the quotient

$$\prod^{\omega} A_N = \left(\prod_{N=1}^{\infty} A_N\right) / \left\{(a_j)_{N=1}^{\infty} : \lim_{N\to\omega} \|a_N\| = 0\right\}.$$

Then $\mathfrak{A}$ is a C^*-algebra.

Let now $X_N^{(j)}$, $j = 1, \ldots, n$, $N = 1, 2, \ldots$ be self-adjoint random variables and assume that $X^{(j)}$, $j = 1, \ldots, n$ are such that for any non-commutative polynomial P,

$$\begin{aligned} \tau_N(P(X_N^{(1)}, \ldots, X_N^{(n)})) &\to \tau(P(X^{(1)}, \ldots, X^{(n)})) \\ \|P(X_N^{(1)}, \ldots, X_N^{(n)})\| &\to \|P(X^{(1)}, \ldots, X^{(n)})\|. \end{aligned}$$

Let $L^{(j)}$, $j = 1, \ldots, n$ be a family of free creation operators, free from each other and from $\{X_N^{(j)}\}_{N,j} \cup \{X^{(j)}\}_j$. In other words, they satisfy:

$$L^{(j)*} x L^{(j)} = \tau(x), \qquad \forall x \in C^*(\{X_N^{(j)}\}_{N,j} \cup \{X^{(j)}\}_j)$$

We use the notations

$$\begin{aligned} A_N &= C^*(X_N^{(1)}, \ldots, X_N^{(n)}), \qquad B_N = C^*(X_N^{(1)}, \ldots, X_N^{(n)}, L^{(1)}, \ldots, L^{(n)}) \\ A &= C^*(X^{(1)}, \ldots, X^{(n)}), \qquad B = C^*(X^{(1)}, \ldots, X^{(n)}, L^{(1)}, \ldots, L^{(n)}) \end{aligned}$$

and we denote by τ_N and ψ_N the respective states on A_N and B_N ($\cong (A_N, \tau_N) * (\mathcal{E}, \phi)$). We denote by τ and ψ the respective states on A and B ($\cong (A, \tau) * (\mathcal{E}, \phi)$).

Consider now the ultrapowers

$$\mathfrak{A} = \prod^{\omega} A_N \subset \mathfrak{B} = \prod^{\omega} B_N.$$

The formula

$$\psi : (x_N)_{N=1}^{\infty} \mapsto \lim_{N\to\omega} \psi_N(x_N)$$

defines a state on $\mathfrak{B}$.

We shall denote by $\hat{X}^{(j)} \in \mathfrak{A}$ the sequence $(X_N^{(j)})_{j=1}^N$. Then by assumption, we have that the map α taking $X^{(j)}$ to $\hat{X}^{(j)}$ extends to a state-preserving isomorphism from (A, τ) into $\mathcal{B}$ with range $\hat{A} = C^*(\hat{X}^{(1)}, \ldots, \hat{X}^{(n)})$.

We shall also denote by $\hat{L}^{(j)}$ the constant sequence $(L^{(j)})_{N=1}^\infty \in \mathfrak{B}$. Then for any element of $\hat{A}$ represented by the sequence $x = (x_N)_{N=1}^\infty$ we have:

$$\hat{L}^{(j)*} x \hat{L}^{(i)} = \delta_{i=j} (\tau_N(x_N))_{N=1}^\infty$$

which (since the L^2 and operator norms coincide on multiples of identity) is equal to $\tau(x) 1 \delta_{i=j} \in \mathfrak{A}$. It follows from the universality property that

$$\hat{B} \stackrel{\text{def}}{=} C^*(\hat{X}^{(1)}, \ldots, \hat{X}^{(n)}, \hat{L}^{(1)}, \ldots, \hat{L}^{(n)})$$

is a quotient of $(A, \tau) * (\mathcal{E}, \phi)$, the quotient map β determined by the fact that it is α on A and takes ℓ_j to $\hat{L}^{(j)}$. On the other hand, if we consider the GNS-representation π of $\hat{B}$ with respect to the restriction of ψ, we easily get (by freeness from $\hat{A}$ and $\{\hat{L}^{(j)}\}_j$) that the image is isomorphic to $(A, \tau) * (\mathcal{E}, \phi)$. Thus $\pi \circ \beta = \text{id}$ so that actually

$$\beta : (A, \tau) * (\mathcal{E}, \phi) \to \hat{B} = C^*(\hat{X}^{(1)}, \ldots, \hat{X}^{(n)}, \hat{L}^{(1)}, \ldots, \hat{L}^{(n)})$$

is an isomorphism.

Consider now a non-commutative $*$-polynomial P. Then

$$\begin{aligned} & \|P(X^{(1)}, \ldots, X^{(n)}, \ell^{(1)}, \ldots \ell^{(n)})\|_{(A,\tau)*(\mathcal{E},\phi)} \\ = \ & \|P(\hat{X}^{(1)}, \ldots, \hat{X}^{(n)}, \hat{L}^{(1)}, \ldots, \hat{L}^{(n)})\|_{\mathfrak{B}} \\ = \ & \lim_{N \to \omega} \|P(X_N^{(1)}, \ldots, X_N^{(n)}, L^{(1)}, \ldots, L^{(n)})\|_{B_N}. \end{aligned}$$

Since the left hand side does not depend on ω, we have proved:

Theorem 1.10.1. Let $X_N^{(j)} \in (A_N, \tau_N)$, $j = 1, \ldots, n$, $N = 1, 2, \ldots$ be self-adjoint random variables and assume that $X^{(j)} \in (A, \tau)$, $j = 1, \ldots, n$ are such that for any non-commutative polynomial P,

$$\begin{aligned} \tau(P(X_N^{(1)}, \ldots, X_N^{(n)})) &\to \tau(P(X^{(1)}, \ldots, X^{(n)})) \\ \|P(X_N^{(1)}, \ldots, X_N^{(n)})\|_{A_N} &\to \|P(X^{(1)}, \ldots, X^{(n)})\|_A. \end{aligned}$$

Let $(\ell_1, \ldots, \ell_n) \in \mathcal{E}$ be free creation operators, and let $B_N = (\mathcal{E}, \phi) * (A_N, \tau_N)$, $B = (\mathcal{E}, \phi) * (A, \tau)$. Assume that the traces τ_j are faithful. Then for any non-commutative $*$-polynomial Q,

$$\|Q(X_N^{(1)}, \ldots, X_N^{(n)}, \ell_1, \ldots, \ell_n)\|_{B_N} \to \|Q(X^{(1)}, \ldots, X^{(n)}, \ell_1, \ldots, \ell_n)\|_B.$$

It should be noted that if $S_1, \ldots, S_n$ are free semicircular variables, free from $\{X_N^{(j)}\}_{N,j} \cup \{X^{(j)}\}_j$, then $C_N = C^*(X_N^{(1)}, \ldots, X_N^{(n)}, S_1, \ldots, S_n)$ is isometrically contained in B_N, while $C = C^*(X^{(1)}, \ldots, X^{(n)}, S_1, \ldots, S_n)$ is isometrically contained in B. Thus the analog of Theorem A with ℓ_j's replaced by a free semicircular family also holds.

Chapter 2

The strong asymptotic freeness of Haar and deterministic matrices

In collaboration with Benoit Collins

ABSTRACT:

In this paper, we are interested in sequences of q-tuple of $N \times N$ random matrices having a strong limiting distribution (i.e. given any non-commutative polynomial in the matrices and their conjugate transpose, its normalized trace and its norm converge). We start with such a sequence having this property, and we show that this property pertains if the q-tuple is enlarged with independent unitary Haar distributed random matrices. Besides, the limit of norms and traces in non-commutative polynomials in the enlarged family can be computed with reduced free product construction. This extends results of one author (C. M.) and of Haagerup and Thorbjørnsen. We also show that a p-tuple of independent orthogonal and symplectic Haar matrices have a strong limiting distribution, extending a recent result of Schultz.

2.1 Introduction and statement of the main results

Following random matrix notation, we call GUE the Gaussian Unitary Ensemble, i.e. any sequence $(X_N)_{N\geqslant 1}$ of random variables where X_N is an $N \times N$ selfadjoint random matrix whose distribution is proportional to the measure $\exp\Big(-\frac{N}{2}\mathrm{Tr}(A^2)\Big)\mathrm{d}A$, where $\mathrm{d}A$ denotes the Lebesgue measure on the set of $N \times N$ Hermitian matrices.

We recall for readers convenience the following definitions from free probability theory (see [AGZ10, NS06]).

Definition 2.1.1. 1. A $\mathcal{C}^*$**-probability space** $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ consists of a unital C^*-algebra $(\mathcal{A}, .^*, \|\cdot\|)$ endowed with a state τ, i.e. a linear map

$\tau \colon \mathcal{A} \to \mathbb{C}$ satisfying $\tau[\mathbf{1}_{\mathcal{A}}] = 1$ and $\tau[aa^*] \geqslant 0$ for all a in $\mathcal{A}$. In this paper, we always assume that τ is a trace, i.e. that it satisfies $\tau[ab] = \tau[ba]$ for every a, b in $\mathcal{A}$. A trace is said to be **faithful** if $\tau[aa^*] > 0$ whenever $a \neq 0$. An element of $\mathcal{A}$ is called a (non commutative) random variable.

2. Let $\mathcal{A}_1, \dots, \mathcal{A}_k$ be $*$-subalgebras of $\mathcal{A}$ having the same unit as $\mathcal{A}$. They are said to be **free** if for any integer $n \geqslant 1$, any $a_i \in \mathcal{A}_{j_i}$ ($i = 1, \dots, n$, $j_i \in \{1, \dots, k\}$) such that $\tau[a_i] = 0$, one has
$$\tau[a_1 \cdots a_n] = 0$$
as soon as $j_1 \neq j_2$, $j_2 \neq j_3, \dots, j_{n-1} \neq j_n$. Collections of random variables are said to be free if the unital subalgebras they generate are free.
3. Let $\mathbf{a} = (a_1, \dots, a_k)$ be a k-tuple of random variables. The *joint distribution* of the family $\mathbf{a}$ is the linear form $P \mapsto \tau\Big[P(\mathbf{a}, \mathbf{a}^*)\Big]$ on the set of polynomials in $2p$ non commutative indeterminates. By **convergence in distribution**, for a sequence of families of variables $(\mathbf{a}_N)_{N \geqslant 1} = (a_1^{(N)}, \dots, a_p^{(N)})_{N \geqslant 1}$, we mean the pointwise convergence of the map
$$P \mapsto \tau\Big[P(\mathbf{a}_N, \mathbf{a}_N^*)\Big],$$
and by **strong convergence in distribution**, we mean convergence in distribution, and pointwise convergence of the map
$$P \mapsto \Big\| P(\mathbf{a}_N, \mathbf{a}_N^*) \Big\|.$$
4. A family of non commutative random variables $\mathbf{x} = (x_1, \dots, x_p)$ is called a **free semicircular system** when the non commutative random variables are free, selfadjoint ($x_i = x_i^*$, $i = 1, \dots, p$), and for all k in $\mathbb{N}$ and $i = 1, \dots, p$, one has
$$\tau[x_i^k] = \int t^k d\sigma(t),$$
with $d\sigma(t) = \frac{1}{2\pi}\sqrt{4 - t^2}\ \mathbf{1}_{|t| \leqslant 2}\ dt$ the semicircle distribution.
5. A non commutative random variable u is called a **Haar unitary** when it is unitary ($uu^* = u^*u = \mathbf{1}$) and for all n in $\mathbb{N}$, one has
$$\tau[u^n] = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{otherwise.} \end{cases}$$

In their seminal paper [HT05], Haagerup and Thorbjørnsen proved the following result.

Theorem 2.1.2 ([HT05] The strong asymptotic freeness of independent GUE matrices)**.**
For any integer $N \geqslant 1$, let $X_1^{(N)}, \dots, X_p^{(N)}$ be $N \times N$ independent GUE matrices and let $(x_1, \dots, x_p)$ be a free semicircular system in a $\mathcal{C}^*$-probability space with faithful state. Then, almost surely, for all polynomials P in p non commutative indeterminates, one has
$$\Big\| P(X_1^{(N)}, \dots, X_p^{(N)}) \Big\| \underset{N \to \infty}{\longrightarrow} \Big\| P(x_1, \dots, x_p) \Big\|,$$

where $\|\cdot\|$ denotes the operator norm in the left hand side and the $\mathcal{C}^*$-algebra in the right hand side.

This theorem is a very deep result in random matrix theory, and had an important impact. Firstly, it had significant applications to C^*-algebra theory [HT05, Pis03], and more recently to quantum information theory [BCN, Che82b]. Secondly, it was generalized in many directions. Schultz [Sch05] has shown that Theorem 2.1.2 is true when the GUE matrices are replaced by matrices of the Gaussian Orthogonal Ensemble (GOE) or by matrices of the Gaussian Symplectic Ensemble (GSE). Capitaine and Donati-Martin [CDM07] and, very recently, Anderson [And] has shown the analogue for certain Wigner matrices.

An other significant extension of Haagerup and Thorbjørnsen's result was obtained by one author (C. M.) in [Mal11], where he managed to show that if in addition to independent GUE matrices, one also has an extra family of independent matrices with strong limiting distribution, the result still holds.

Theorem 2.1.3 ([Mal11] The strong asymptotic freeness of $X_1^{(N)}, \dots, X_p^{(N)}$ and $\mathbf{Y}_N$)**.**
For any integer $N \geqslant 1$, we consider
- a family $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ of $N \times N$ independent GUE matrices,
- a family $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ of $N \times N$ matrices, possibly random but independent of $\mathbf{X}_N$.

In a $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ with faithful trace, we consider
- a free semicircular system $\mathbf{x} = (x_1, \dots, x_p)$,
- a family $\mathbf{y} = (y_1, \dots, y_q)$ of non commutative random variables, free from $\mathbf{x}$.

Then, if $\mathbf{y}$ is the strong limit in distribution of $\mathbf{Y}_N$, we have that $(\mathbf{x}, \mathbf{y})$ is the strong limit in distribution of $(\mathbf{X}_N, \mathbf{Y}_N)$. In other words, if we assume that almost surely, for all polynomials P in $2q$ non commutative indeterminates, one has

$$\tau_N\Big[P(\mathbf{Y}_N, \mathbf{Y}_N^*)\Big] \underset{N\to\infty}{\longrightarrow} \tau\Big[P(\mathbf{y}, \mathbf{y}^*)\Big], \tag{2.1}$$

$$\Big\|P(\mathbf{Y}_N, \mathbf{Y}_N^*)\Big\| \underset{N\to\infty}{\longrightarrow} \Big\|P(\mathbf{y}, \mathbf{y}^*)\Big\|, \tag{2.2}$$

then, almost surely, for all polynomials P in $p + 2q$ non commutative indeterminates, one has

$$\tau_N\Big[P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)\Big] \underset{N\to\infty}{\longrightarrow} \tau\Big[P(\mathbf{x}, \mathbf{y}, \mathbf{y}^*)\Big], \tag{2.3}$$

$$\Big\|P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)\Big\| \underset{N\to\infty}{\longrightarrow} \Big\|P(\mathbf{x}, \mathbf{y}, \mathbf{y}^*)\Big\|. \tag{2.4}$$

It is natural to wonder whether the same property holds for unitary Haar matrices, instead of GUE matrices. The main result of this paper is the following theorem.

Theorem 2.1.4 (The strong asymptotic freeness of $U_1^{(N)}, \dots, U_p^{(N)}, \mathbf{Y}_N$)**.**
For any integer $N \geqslant 1$, we consider

- a family $\mathbf{U}_N = (U_1^{(N)}, \dots, U_p^{(N)})$ of $N \times N$ independent unitary Haar matrices,
- a family $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ of $N \times N$ matrices, possibly random but independent of $\mathbf{U}_N$.

In a $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \| \cdot \|)$ with faithful trace, we consider

- a family $\mathbf{u} = (u_1, \dots, u_p)$ of free Haar unitaries,
- a family $\mathbf{y} = (y_1, \dots, y_q)$ of non commutative random variables, free from $\mathbf{u}$.

Then, if $\mathbf{y}$ is the strong limit in distribution of $\mathbf{Y}_N$, we have that $(\mathbf{u}, \mathbf{y})$ is the strong limit in distribution of $(\mathbf{U}_N, \mathbf{Y}_N)$.

The convergence in distribution of $(\mathbf{U}_N, \mathbf{Y}_N)$ is the content of Voiculescu's asymptotic freeness theorem and is recalled in order to give a coherent and complete statement (see [AGZ10, Theorem 5.4.10] for a proof).

In order to solve this problem, it looks at first sight natural to attempt to mimic the proof of Haagerup and Thorbjørnsen [HT05] and write a Master equation in the case of unitary matrices. While this could be attempted via a Schwinger-Dyson type argument, the computation are much more difficult than for GUE matrices because of the non linearity of the $\mathcal{R}$-transform in the unitary case. In this paper, we take a completely different route to tackle this problem by building on Theorem 2.1.3 and using a series of folklore facts of classical probability and random matrix theory.

Our method applies with minor modifications to the cases of Haar matrices on the orthogonal and the symplectic groups by building on the result of Schultz [Sch05]. Since an analogue of Theorem 2.1.3 for GOE or GSE matrices does not exist yet, the result stated in this paper as Theorem 2.1.5 is less general than Theorem 2.1.4 is for unitary Haar matrices. We show the following.

Theorem 2.1.5 (The strong asymptotic freeness of independent Haar matrices)**.**
For any integer $N \geqslant 1$, let $U_1^{(N)}, \dots, U_p^{(N)}$ be a family of $N \times N$ independent orthogonal Haar matrices or $2N \times 2N$ independent symplectic Haar matrices and let $u_1, \dots, u_p$ be free unitaries in a $\mathcal{C}^*$-probability space with faithful state. Then, almost surely, for all polynomials P in $2p$ non commutative indeterminates, one has

$$\left\| P(U_1^{(N)}, \dots, U_p^{(N)}, U_1^{(N)*}, \dots, U_p^{(N)*}) \right\| \underset{N \to \infty}{\longrightarrow} \left\| P(u_1, \dots, u_p, u_1^*, \dots, u_p^*) \right\|,$$

where $\| \cdot \|$ denotes the operator norm in the left hand side and the $\mathcal{C}^*$-algebra in the right hand side.

Our paper is organized as follows. Section 2.2 provides the proofs of Theorem 2.1.4 and Theorem 2.1.5. Section 2.3 consists of further applications and concluding remarks.

2.2 Proof of Theorems 2.1.4 and 2.1.5

2.2.1 Idea of the proof

The keystone of the proof is the existence of an explicit coupling (U_N, X_N) of an $N \times N$ Haar matrix U_N and an $N \times N$ GUE matrix X_N, consisting of

- a trivial coupling of the eigenvalues of U_N and X_N (they are independent),
- a deterministic coupling of their eigenvectors (U_N and X_N are diagonalizable in a same basis),

such that the relative orders of the eigenvalues of X_N and of the arguments of the eigenvalues of U_N with respect to a numeration of their eigenvectors are consistent. Such a coupling is possible thanks to the unitary invariance of the GUE law and of the Haar measure. Moreover, we can construct a function $h_N : \mathbb{R} \to \mathbb{S}^1$, referred as the folding map, such that almost surely one has

$$U_N = h_N(X_N). \tag{2.5}$$

Formally, the function h_N depends measurably on the pair (U_N, X_N), but we will make a slight abuse of notation and denote it h_N (note that actually the dependence of h_N on (U_N, X_N) becomes negligible as $N \to \infty$ with probability one - this observation will be made rigorous in the proof). Recall that for a map $f : \mathbb{C} \to \mathbb{C}$ and a normal matrix $M = V\text{diag }(x_1, \dots, x_N)V^*$, with V unitary, the symbol $f(M)$ denotes the normal matrix $V\text{diag }\Big(f(x_1), \dots, f(x_K)\Big)V^*$. The map h_N is a not continuous and is random. It is obtained by combination of the empirical cumulative functions of the eigenvalues of X_N and of the arguments of the eigenvalues of U_N and X_N (see definition (2.9) below). The construction of h_N is quite a classical trick in probability on the real line, sometimes referred as the folding/unfolding of random variables, hence the name.

At the level of non commutative random variables, we have an analogue coupling

$$u = h(x), \tag{2.6}$$

between a Haar unitary u and a semicircular variable x in a $\mathcal{C}^*$-probability space. The map $h : \mathbb{R} \to \mathbb{S}^1$ is continuous. In particular, the symbol $h(x)$ is computed by functional calculus. If we consider $\tilde{U}_N = h(X_N)$, we can deduce from Theorem 2.1.3 that $(\tilde{U}_N, \mathbf{Y}_N)$ converges strongly to $(u, \mathbf{y})$ (i.e. we have the convergence of normalized trace and norm for any polynomial). This idea is used in [HT05, Part 8] to deduce results of $\mathcal{C}^*$-algebra theory from the convergence of random matrices.

Now, knowing the coupling (U_N, X_N) described above, it is actually possible to get directly the strong convergence for $(U_N, \mathbf{Y}_N)$. We only have to estimate $\|U_N - \tilde{U}_N\|$. This amounts to show the uniform convergence of the empirical cumulative function of the eigenvalues of X_N and of the general inverse of the empirical cumulative function of the arguments of the eigenvalues of U_N, which is obtained as a byproduct of Wigner's theorem and Dini's type theorems.

2.2.2 An almost sure coupling for random matrices

We first recall, in Proposition 2.2.1 below, the spectral theorem for unitary invariant random matrices, a well known result of random matrices theory.

Proposition 2.2.1 (Spectral theorem for unitary invariant random matrices)**.** Let M_N be an $N \times N$ Hermitian or unitary random matrix whose distribution is invariant under conjugacy by unitary matrices. Then, M_N can be written $M_N = V_N \Delta_N V_N^*$ almost surely, where

- V_N is distributed according to the Haar measure on the unitary group,
- Δ_N is the diagonal matrix of the eigenvalues of M_N, arranged in increasing order if M_N Hermitian, and in increasing order with respect to the set of arguments in $[-\pi, \pi[$ if M_N is unitary,
- V_N and Δ_N are independent.

We recall a proof for the convenience of the readers. We actually use the proposition only for unitary Haar and GUE matrices, which are two cases where almost surely the eigenvalues are distinct. This fact brings slight conceptual simplifications, which nevertheless do not change the proof. Hence, we prefer to state the proposition without any restriction on the multiplicity of the matrices.

Proof. By reasoning conditionally, one can always assume that the multiplicities of the eigenvalues of M_N is almost surely constant. We denote by $(N_1, \dots, N_K)$ the sequence of multiplicities when the eigenvalues are considered in the natural order in $\mathbb{R}$ or in increasing order with respect to their argument in $[-\pi, \pi[$.

Since almost surely M_N is normal, it can be written $M_N = \tilde{V}_N \Delta_N \tilde{V}_N$, where $\tilde{V}_N$ is a random unitary matrix and Δ_N is as announced. The choice of $\tilde{V}_N$ can be made in a measurable way, for instance by requiring that the first nonzero element of each column of V_N is a positive real number.

Let $(u_1, \dots, u_K)$ be a family of independent random matrices, independent of $(\Delta_N, \tilde{V}_N)$ and such that for any $k = 1, \dots, K$, the matrix u_k is distributed according to the Haar measure on $\mathcal{U}(N_k)$, the group of $N_k \times N_k$ unitary matrices. We set

$$V_N = \tilde{V}_N \text{ diag } (u_1, \dots, u_K),$$

and claim that the law of V_N depends only on the law of M_N, not in the choice of the random matrix $\tilde{V}_N$. Indeed, let $M_N = \bar{V}_N \Delta_N \bar{V}_N$ be an other decomposition, where $\bar{V}_N$ is a unitary random matrix, independent of $(u_1, \dots, u_K)$. The multiplicities of the eigenvalues being $N_1, \dots, N_K$, there exists $(v_1, \dots, v_K)$ in $\mathcal{U}(N_1) \times \dots \times \mathcal{U}(N_K)$, independent of $(u_1, \dots, u_K)$, such that $\bar{V}_N = \tilde{V}_N \text{ diag } (v_1, \dots, v_K)$. Hence, we get $\bar{V}_N \text{ diag } (u_1, \dots, u_K) = \tilde{V}_N \text{ diag } (v_1 u_1, \dots, v_K u_K)$, which is equal in law to V_N. This proves the claim.

Let W_N be an $N \times N$ unitary matrix. Then $W_N M_N W_N^* = (W_N \tilde{V}_N) \Delta_N (W_N \tilde{V}_N)^*$. By the above, since M_N and $W_N M_N W_N^*$ are equal in law, then V_N and $W_N V_N$ are also equal in law. Hence V_N is Haar distributed in $\mathcal{U}(N)$.

It remains to show the independence between V_N and Δ_N. Let $f : \mathcal{U}(N) \to \mathbb{C}$ and $g : \mathrm{M}_N(\mathbb{C}) \to \mathbb{C}$ two bounded measurable functions such that g depends only on the eigenvalues of its entries. Then one as $\mathbb{E}\Big[f(V_N)g(\Delta_N)\Big] = \mathbb{E}\Big[f(V_N)g(M_N)\Big]$. Let W_N be Haar distributed in $\mathcal{U}(N)$, independent of (V_N, Δ_N). Then by the invariance under unitary conjugacy of the law of M_N, one has

$$\begin{aligned}
\mathbb{E}\Big[f(V_N)g(\Delta_N)\Big] &= \mathbb{E}\Big[f(W_N V_N)g(W_N M_N W_N^*)\Big] \\
&= \mathbb{E}\Big[f(W_N V_N)g(\Delta_N)\Big] \\
&= \mathbb{E}\Big[\mathbb{E}\Big[f(W_N V_N)\Big|V_N, \Delta_N\Big]g(\Delta_N)\Big] \\
&= \mathbb{E}\Big[f(W_N)\Big]\mathbb{E}\Big[g(\Delta_N)\Big] = \mathbb{E}\Big[f(V_N)\Big]\mathbb{E}\Big[g(\Delta_N)\Big].
\end{aligned}$$

□

We are ready to construct the desired coupling. For the purposes of this paper, we start with a Haar unitary matrix, and then construct a GUE matrix.

Let U_N be an $N \times N$ unitary Haar matrix. By Proposition 2.2.1, we can write $U_N = V_N \Delta_N V_N^*$, where V_N is a Haar unitary matrix, independent of $\Delta_N = \mathrm{diag}\,(\mathrm{e}^{i\theta_1^{(N)}}, \ldots, \mathrm{e}^{i\theta_N^{(N)}})$, and

$$-\pi \leqslant \theta_1^{(N)} \leqslant \cdots \leqslant \theta_N^{(N)} < \pi.$$

We consider a random diagonal matrix $\tilde{\Delta}_N = \mathrm{diag}\,(\lambda_1^{(N)}, \ldots, \lambda_N^{(N)})$, independent of (V_N, Δ_N) and such that the random vector $(\lambda_1^{(N)}, \ldots, \lambda_N^{(N)})$ has the law of the eigenvalues of a GUE matrix, sorted in increasing order. We set

$$X_N := V_N \tilde{\Delta}_N V_N^*,$$

which is a GUE matrix by Proposition 2.2.1. Hence the announced coupling (U_N, X_N).

We now define the map h_N which gives $U_N = h_N(X_N)$. In the sequel, we will omit the superscript (N) and replace the notations $\lambda_1^{(N)}, \ldots, \lambda_N^{(N)}$ by $\lambda_1, \ldots, \lambda_N$ and $\theta_1^{(N)}, \ldots, \theta_N^{(N)}$ by $\theta_1, \ldots, \theta_N$. Let $F_{X_N} : \mathbb{R} \to [0, 1]$ be the empirical cumulative distribution function of $\{\lambda_1, \ldots, \lambda_N\}$, i.e. for all t in $\mathbb{R}$,

$$F_{X_N}(t) = N^{-1} \sum_{j=1}^{N} \mathbf{1}_{]-\infty, \lambda_j]}(t). \tag{2.7}$$

The eigenvalues of a GUE matrix are distinct with probability one, and $\lambda_1, \ldots, \lambda_N$ are arranged in increasing order. Then, almost surely and for any $j = 1, \ldots, N$, one has $F_{X_N}(\lambda_j) = j/N$. Remark that the push forward of the uniform measure on the spectrum of X_N is the uniform measure on $\{1/N, 2/N, \ldots, 1\}$, a phenomenon sometimes referred as the unfolding trick.

Let $F_{U_N} : [-\pi, \pi] \to [0,1]$ be the empirical cumulative distribution function of $\{\theta_1, \dots, \theta_N\}$ (defined as in (2.7) with the λ_j's replaced by the θ_j's). Let $F_{U_N}^{-1} : [0,1] \to [-\pi, \pi]$ be its generalized inverse i.e. for all s in $]0,1]$,

$$F_{U_N}^{-1}(s) = \inf\Big\{ t \in [-\pi, \pi] \ \Big| \ F_{U_N}(t) \geqslant s \Big\}. \tag{2.8}$$

By the arrangement of the eigenvalues of U_N, for any $j = 1, \dots, N$, one has $F_{U_N}^{-1}(j/N) = \theta_j$. Remark that the push forward of the uniform measure on $\{1/N, 2/N, \dots, 1\}$ is the uniform measure on the spectrum of U_N. This step is sometimes called the folding trick.

We set the random function

$$\begin{array}{rcll} h_N & : & \mathbb{R} & \to \quad \mathbb{S}^1 \\ & & t & \mapsto \quad \exp\Big(iF_{U_N}^{-1} \circ F_{X_N}(t)\Big). \end{array} \tag{2.9}$$

By construction, almost surely for any $j = 1, \dots, N$, one has $h_N(\lambda_j) = e^{i\theta_j}$, and hence, we get the expected relation between U_N and X_N: almost surely one has

$$h_N(X_N) = V_N \ \mathrm{diag}\ \Big(h_N(\lambda_1), \dots, h_N(\lambda_N)\Big)\ V_N^* = U_N. \tag{2.10}$$

In the following, we call h_N the folding map associated to the coupling (U_N, X_N).

2.2.3 A coupling for non commutative random variables

Let $F_x : \mathbb{R} \to [0,1]$ be the cumulative distribution function of the semicircular law with radius two, i.e. for all t in $\mathbb{R}$,

$$F_x(t) = \int_{-\infty}^{t} \frac{1}{2\pi} \sqrt{4 - y^2} \mathrm{d}y. \tag{2.11}$$

Let $F_u^{-1} : [0,1] \to [-\pi, \pi]$ be the inverse of the cumulative distribution function of the Lebesgue measure on $[-\pi, \pi]$, i.e. for all s in $[0,1]$,

$$F_u^{-1}(s) = 2\pi \Big(s - \frac{1}{2} \Big). \tag{2.12}$$

We define the continuous function

$$\begin{array}{rcll} h & : & \mathbb{R} & \to \quad \mathbb{S}^1 \\ & & t & \mapsto \quad \exp\Big(iF_u^{-1} \circ F_x(t)\Big). \end{array} \tag{2.13}$$

By construction, the push forward of the semicircular law with radius two is the uniform measure on the unit circle. Let u be a Haar unitary and x be a semicircular variable in a $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ (we do not care about the possible relation between u and x). Let $\mathbf{y}$ be a family of non commutative random variables in $\mathcal{A}$, free from u and x. Then, one has the equality in non commutative law

$$\Big(h(x), \mathbf{y}\Big) \overset{\mathcal{L}^{n.c.}}{=} \Big(u, \mathbf{y}\Big), \tag{2.14}$$

In other words, for any polynomial P in $2+q$ non commutative indeterminates, one has $\tau\Big[P(h(x), h(x)^*, \mathbf{y})\Big] = \tau\Big[P(u, u^*, \mathbf{y})\Big]$ and then $\|P(h(x), h(x)^*, \mathbf{y})\| = \|P(u, u^*, \mathbf{y})\|$ if τ is faithful. The symbol $h(x)$ is computed by functional calculus (see [NS06, Lecture 3]).

2.2.4 Proof of Theorem 2.1.4

Let $\mathbf{U}_N, \mathbf{Y}_N, \mathbf{u}, \mathbf{y}$ be as in Theorem 2.1.4. Without loss of generality, one can assume that the matrices $\mathbf{Y}_N$ are Hermitian, at the possible cost of replacing the collection of matrices by the collection their real and imaginary parts.

Let $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ be a family of independent $N \times N$ GUE matrices such that

- $(U_1^{(N)}, X_1^{(N)}), \dots, (U_p^{(N)}, X_p^{(N)}), \mathbf{Y}_N$ are independent,
- for any $j = 1, \dots, p$, $(U_j^{(N)}, X_j^{(N)})$ is a coupling constructed by the method of Section 2.2.2, whose folding map is denoted $h_j^{(N)}$.

Let h the function defined in Section 2.2.3 by formula (2.13). For any $j = 1, \dots, p$, we set the $N \times N$ unitary random matrix $\tilde{U}_j^{(N)} = h(X_j^{(N)})$. We denote $\tilde{\mathbf{U}}_N = (\tilde{U}_1^{(N)}, \dots, \tilde{U}_p^{(N)})$. Theses matrices are not Haar distributed: for instance, as it is noticed in [HT05, Remark 8.3], the matrix $\tilde{U}_1^{(N)}$ is the identity matrix with (small but) nonzero probability. Nevertheless, it is a known consequence of Theorem 2.1.3 that the family of matrices $\tilde{\mathbf{U}}_N$ converges strongly to the family $\mathbf{u}$ of free Haar unitaries (see [HT05, Section 8]). We only need here the norm convergence, and we recall a proof for the convenience of the readers.

Lemma 2.2.2. Almost surely, for every polynomial P in $2+q$ non commutative indeterminates, one has

$$\Big\| P(\tilde{\mathbf{U}}_N, \tilde{\mathbf{U}}_N^*, \mathbf{Y}_N) \Big\| \underset{N\to\infty}{\longrightarrow} \Big\| P(\mathbf{u}, \mathbf{u}^*, \mathbf{y}) \Big\|,$$

where $\tilde{\mathbf{U}}_N = \Big(h(X_1^{(N)}), \dots, h(X_p^{(N)}) \Big)$.

We shall need the following lemma.

Lemma 2.2.3. Let $\mathbf{a} = (a_1, \dots, a_n)$ and $\mathbf{b} = (b_1, \dots, b_n)$ be families of elements in a $\mathcal{C}^*$-algebra $(\mathcal{A}, \| \cdot \|)$. Denote D the supremum of $\|a_1\|, , \dots, \|a_n\|$ and $\|b_1\|, \dots, \|b_n\|, 1$. Then for every polynomial P in n non commutative indeterminates one has

$$\Big\| P(\mathbf{a}) - P(\mathbf{b}) \Big\| \leqslant \beta D^{\alpha-1} \sum_{i=1}^{n} \|a_i - b_i\|,$$

where the constant β depends only on P and α is the total degree of P.

Proof of Lemma 2.2.3. It is sufficient to show that there exist β such that, for any $a, b, c = (c_1, \dots, c_{n-1})$ in $\mathcal{A}$, with $D = \sup(\|a\|, \|b\|, \|c_1, \dots, c_{n-1}\|)$, one has

$$\Big\| P(a, \mathbf{c}) - P(b, \mathbf{c}) \Big\| \leqslant \beta D^{\alpha-1} \|a - b\|,$$

and then apply n times this fact. Moreover, it is sufficient to show this inequality when P is a monic monomial, of positive degree in the first indeterminate. For such a polynomial P, there exist two monic monomial L and R such that

$P(a, \mathbf{c}) = L(\mathbf{c})aR(a, \mathbf{c})$, $P(b, \mathbf{c}) = L(\mathbf{c})bR(b, \mathbf{c})$. Then, one has

$$\begin{aligned}
&\Big\| P(a, \mathbf{c}) - P(b, \mathbf{c}) \Big\| \\
&\quad \leqslant \Big\| L(\mathbf{c}) \Big\| \times \Big\| aR(a, \mathbf{c}) - bR(b, \mathbf{c}) \Big\| \\
&\quad \leqslant \Big\| L(\mathbf{c}) \Big\| \Big(\Big\| aR(a, \mathbf{c}) - bR(a, \mathbf{c}) \Big\| + \Big\| bR(a, \mathbf{c}) - bR(b, \mathbf{c}) \Big\| \Big) \\
&\quad \leqslant D^{\alpha - 1} \| a - b \| + \Big\| L(\mathbf{c}) \Big\| \times \| b \| \times \Big\| R(a, \mathbf{c}) - R(b, \mathbf{c}) \Big\|.
\end{aligned}$$

By induction on the degree of the monomials, we get the result. □

Proof of Lemma 2.2.2. In the following we use the notation $f(\mathbf{a})$ for $\big(f(a_1), \dots, f(a_k)\big)$ whenever $\mathbf{a} = (a_1, \dots, a_k)$ is a family of normal elements of a $\mathcal{C}^*$-algebra and $f : \mathbb{C} \to \mathbb{C}$ a continuous map. For any $\varepsilon > 0$, let h_ε be a polynomial such that $|h(x) - h_\varepsilon(x)| \leqslant \varepsilon$ for all x in $[-3, 3]$. For any polynomial P in $2p + q$ non commutative indeterminates, one has

$$\begin{aligned}
&\Big| \Big\| P(\tilde{\mathbf{U}}_N, \tilde{\mathbf{U}}_N^*, \mathbf{Y}_N) \Big\| - \Big\| P(\mathbf{u}, \mathbf{u}^*, \mathbf{y}) \Big\| \Big| \\
&\quad = \Big| \Big\| P\Big(h(\mathbf{X}_N), \overline{h}(\mathbf{X}_N), \mathbf{Y}_N\Big) \Big\| - \Big\| P\Big(h(\mathbf{x}), \overline{h}(\mathbf{x}), \mathbf{y}\Big) \Big\| \Big| \\
&\quad \leqslant \Big\| P\Big(h(\mathbf{X}_N), \overline{h}(\mathbf{X}_N), \mathbf{Y}_N\Big) - P\Big(h_\varepsilon(\mathbf{X}_N), \overline{h_\varepsilon}(\mathbf{X}_N), \mathbf{Y}_N\Big) \Big\| \\
&\qquad + \Big| \Big\| P\Big(h_\varepsilon(\mathbf{X}_N), \overline{h_\varepsilon}(\mathbf{X}_N), \mathbf{Y}_N\Big) \Big\| - \Big\| P\Big(h_\varepsilon(\mathbf{x}), \overline{h_\varepsilon}(\mathbf{x}), \mathbf{y}\Big) \Big\| \Big| \\
&\qquad + \Big\| P\Big(h(\mathbf{x}), \overline{h}(\mathbf{x}), \mathbf{y}\Big) - P\Big(h_\varepsilon(\mathbf{x}), \overline{h_\varepsilon}(\mathbf{x}), \mathbf{y}\Big) \Big\|
\end{aligned}$$

By Theorem 2.1.3, one has almost surely

$$\Big| \Big\| P\Big(h_\varepsilon(\mathbf{X}_N), \overline{h_\varepsilon}(\mathbf{X}_N), \mathbf{Y}_N\Big) \Big\| - \Big\| P\Big(h_\varepsilon(\mathbf{x}), \overline{h_\varepsilon}(\mathbf{x}), \mathbf{y}\Big) \Big\| \Big| \underset{N \to \infty}{\longrightarrow} 0.$$

On the other hand, by Lemma 2.2.3, we have almost surely

$$\begin{aligned}
&\Big\| P\Big(h(\mathbf{X}_N), \overline{h}(\mathbf{X}_N), \mathbf{Y}_N\Big) - P\Big(h_\varepsilon(\mathbf{X}_N), \overline{h_\varepsilon}(\mathbf{X}_N), \mathbf{Y}_N\Big) \Big\| \\
&\quad \leqslant C \sum_{j=1}^{p} \Big\| h(X_j^{(N)}) - h_\varepsilon(X_j^{(N)}) \Big\| \qquad (2.15)
\end{aligned}$$

$$\begin{aligned}
&\Big\| P\Big(h(\mathbf{x}), \overline{h}(\mathbf{x}), \mathbf{y}\Big) - P\Big(h_\varepsilon(\mathbf{x}), \overline{h_\varepsilon}(\mathbf{x}), \mathbf{y}\Big) \Big\| \\
&\quad \leqslant C \sum_{j=1}^{p} \Big\| h(x_j) - h_\varepsilon(x_j) \Big\|, \qquad (2.16)
\end{aligned}$$

where C is a constant that only depends on P and on a (random) bound D such that for any $j = 1, \dots, q$, one has $\|Y_j^{(N)}\| \leqslant D$. By Theorem 2.1.2, almost surely there exists N_0 such that for any $N \geqslant N_0$ and $j = 1, \dots, p$, one has $\|X_j^{(N)}\| \leqslant 3$. Moreover, the support of the semicircular distribution is $[-2, 2]$. Then, almost surely for N large enough, the two quantities (2.15) and (2.16) are bounded by $C\varepsilon$. Hence, we have shown that almost surely,

$$\left\|P(\tilde{\mathbf{U}}_N, \tilde{\mathbf{U}}_N^*, \mathbf{Y}_N)\right\| \underset{N\to\infty}{\longrightarrow} \left\|P(\mathbf{u}, \mathbf{u}^*, \mathbf{y})\right\|. \tag{2.17}$$

Since a countable intersection of probability one sets is again of probability one, we get that almost surely, (2.17) holds for all polynomials P with coefficients in $\mathbb{Q}$. Both sides in (2.17) are continuous in P, hence we obtain the expected result by density of polynomials with rational coefficients. □

Let P be a polynomial in $2p+q$ non commutative indeterminates. We want to show that: almost surely one has

$$\left\|P(\mathbf{U}_N, \mathbf{U}_N^*, \mathbf{Y}_N)\right\| \underset{N\to\infty}{\longrightarrow} \left\|P(\mathbf{u}, \mathbf{u}^*, \mathbf{y})\right\|,$$

which will be enough to show Theorem 2.1.4 by the same reasoning as in the end of the proof of Lemma 2.2.2. We set the random variable

$$\varepsilon_N = \left|\left\|P(\tilde{\mathbf{U}}_N, \tilde{\mathbf{U}}_N^*, \mathbf{Y}_N)\right\| - \left\|P(\mathbf{u}, \mathbf{u}^*, \mathbf{y})\right\|\right|,$$

which tends to zero almost surely by Lemma 2.2.2. Now, one has by Lemma 2.2.3

$$\begin{aligned} &\left|\left\|P(\mathbf{U}_N, \mathbf{U}_N^*, \mathbf{Y}_N)\right\| - \left\|P(\mathbf{u}, \mathbf{u}^*, \mathbf{y})\right\|\right| \\ &\quad\leqslant \left\|P(\mathbf{U}_N, \mathbf{U}_N^*, \mathbf{Y}_N) - P(\tilde{\mathbf{U}}_N, \tilde{\mathbf{U}}_N^*, \mathbf{Y}_N)\right\| + \varepsilon_N \qquad (2.18) \\ &\quad\leqslant C\sum_{j=1}^{p} \|U_j^{(N)} - \tilde{U}_j^{(N)}\| + \varepsilon_N, \qquad (2.19) \end{aligned}$$

where C is a constant that only depends on P and on a bound D such that for any $j = 1, \dots, q$, one has $\|Y_j^{(N)}\| \leqslant D$. It remains to show that, for any $j = 1, \dots, p$, almost surely $\|U_j^{(N)} - \tilde{U}_j^{(N)}\|$ tends to zero as N goes to infinity. For any $j = 1, \dots, p$, recall that almost surely

$$U_j^{(N)} = h_j^{(N)}(X_j^{(N)}), \quad \tilde{U}_j^{(N)} = h(X_j^{(N)}),$$

where $h_j^{(N)}$ is the folding map associated to the coupling $(U_j^{(N)}, X_j^{(N)})$ and h is given by formula (2.13). For any $j = 1, \dots, p$, we denote by $\lambda_1(j), \dots, \lambda_N(j)$ the eigenvalues of $X_j^{(N)}$. Hence, one has

$$\begin{aligned}
&\|U_j^{(N)} - \tilde{U}_j^{(N)}\| \\
&= \left\| h_j^{(N)}(X_j^{(N)}) - h(X_j^{(N)}) \right\| \\
&= \left\| \exp\left(iF^{-1}_{U_j^{(N)}} \circ F_{X_j^{(N)}}(X_j^{(N)}) \right) - \exp\left(iF_u^{-1} \circ F_x(X_j^{(N)}) \right) \right\| \\
&\leqslant \sup_{n=1,\dots,N} \left| \exp\left(iF^{-1}_{U_j^{(N)}} \circ F_{X_j^{(N)}}(\lambda_n(j)) \right) - \exp\left(iF_u^{-1} \circ F_x(\lambda_n(j)) \right) \right| \\
&\leqslant \sup_{n=1,\dots,N} \left| F^{-1}_{U_j^{(N)}} \circ F_{X_j^{(N)}}(\lambda_n(j)) - F_u^{-1} \circ F_x(\lambda_n(j)) \right| \\
&\leqslant \sup_{n=1,\dots,N} \left| F^{-1}_{U_j^{(N)}} \circ F_{X_j^{(N)}}(\lambda_n(j)) - F_u^{-1} \circ F_{X_j^{(N)}}(\lambda_n(j)) \right| \\
&\qquad\qquad + \sup_{n=1,\dots,N} \left| F_u^{-1} \circ F_{X_j^{(N)}}(\lambda_n(j)) - F_u^{-1} \circ F_x(\lambda_n(j)) \right| \\
&\leqslant \left\| F^{-1}_{U_j^{(N)}} - F_u^{-1} \right\|_{\mathrm{L}^\infty([0,1])} + 2\pi \left\| F_{X_j^{(N)}} - F_x \right\|_{\mathrm{L}^\infty([0,1])}. \qquad (2.20)
\end{aligned}$$

We shall need two lemmas in order to conclude the proof. The first one is famous in real analysis and is known as Dini's lemma.

Lemma 2.2.4. For any n in $\mathbb{N} \cup \{\infty\}$, let $f_n : \mathbb{R} \to [0,1]$ be a non decreasing function such that $\lim_{x\to-\infty} f_n(x) = 0$ and $\lim_{x\to+\infty} f_n(x) = 1$. Assume that f_∞ is continuous and that f_n converges pointwise to f_∞ on $\mathbb{R}$. Then f_n converges uniformly to f_∞ on $\mathbb{R}$.

Proof. Let $\varepsilon > 0$. We set K the ceiling of $2/\varepsilon$. For any $j = 1, \dots, K-1$, we set $x_j = f_\infty^{-1}(\frac{i}{K})$, where f_∞^{-1} denotes the generalized inverse of f_∞^{-1} defined as in (2.8). We also set $x_0 = -\infty$ and $x_K = +\infty$. In the following we use the convention $f_n(-\infty) = f_\infty(-\infty) = 0$ and $f_n(+\infty) = f_\infty(+\infty) = 1$. By the pointwise convergence of f_n to f_∞ at the points $x_1, \dots, x_{K-1}$: there exists n_0 such that for any $n \geqslant n_0$ and $j = 1, \dots, K-1$, one has

$$|f_n(x_j) - f_\infty(x_j)| \leqslant \frac{\varepsilon}{2}. \qquad (2.21)$$

Let $n \geqslant n_0$. For any x in $\mathbb{R}$, let j in $\{0, \dots, K\}$ such that $x_j \leqslant x < x_{j+1}$. Since the functions are non decreasing, one has $f_n(x_i) - f_\infty(x_{i+1}) \leqslant f_n(x) - f_\infty(x) \leqslant f_n(x_{i+1}) - f_\infty(x_i)$, and so, by (2.21), we get

$$-\frac{\varepsilon}{2} - f_\infty(x_i) + f_\infty(x_{i+1}) \leqslant f_n(x) - f_\infty(x) \leqslant \frac{\varepsilon}{2} + f_\infty(x_{i+1}) - f_\infty(x_i).$$

The continuity of f_∞ implies that $f_\infty(x_i) = i/K$. Hence we get $|f_n(x) - f_\infty(x)| \leqslant 1/K + \varepsilon/2 \leqslant \varepsilon$. □

Lemma 2.2.5. For any n in $\mathbb{N} \cup \{\infty\}$, let $f_n : [a,b] \to [0,1]$ be a non decreasing function. Assume that f_∞ is differentiable in $[a,b]$, its derivative is positive and f_n converges uniformly to f_∞ as n goes to infinity. Then f_n^{-1} converges uniformly to f_∞^{-1} as n goes to infinity, where f^{-1} stands for the generalized inverse of f_n, defined as in (2.8).

Proof. It is sufficient to prove the pointwise convergence of f_n^{-1} to f_∞^{-1}. Indeed, f_∞^{-1} is continuous on $[0,1]$. So, the pointwise convergence granted, we can extend for any n in $\mathbb{N}\cup\{\infty\}$ the map f_n^{-1} on $\mathbb{R}$ by $f_n(x) = a$ if $x < 0$ and $f_n(x) = b$ if $x > 1$, and then apply Lemma 2.2.4 to $(f_n^{-1} - a)/(b-a)$.

Let $\alpha > 0$ such that $f'_\infty(x) \geqslant \alpha$ for any x in $[a,b]$. By the mean value theorem, we get that for any $\varepsilon > 0$

$$\begin{aligned} U_\varepsilon \quad &:= \quad \Big\{(x,y) \in [a,b]\times[0,1] \ \Big| \ |y - f_\infty(x)| \leqslant \varepsilon\Big\} \\ &\subset \quad V_\varepsilon := \Big\{(x,y) \in [a,b]\times[0,1] \ \Big| \ |x - f_\infty^{-1}(y)| \leqslant \frac{\varepsilon}{\alpha}\Big\}. \end{aligned} \tag{2.22}$$

Let $\varepsilon > 0$. By the uniform convergence, there exists n_0 such that for any $n \geqslant n_0$, the graph of f_n is contained in $U_{\alpha\varepsilon}$. Let $n \geqslant n_0$ and t in [0,1]. If $f_n^{-1}(t)$ is a point of continuity for f_n, then $f_n \circ f_n^{-1}(t) = t$. So $(f_n^{-1}(t), t)$ is in the graph of f_n and it belongs to $U_{\alpha\varepsilon}$.

Otherwise, denote by t_1, respectively t_2, the left limit, respectively the right limit, of f_n in $f_n^{-1}(t)$. These limits exist since f_n is non decreasing. By definition of the generalized inverse, t belongs to the interval $[t_1, t_2]$. Moreover, the vertical sections of $U_{\alpha\varepsilon}$ are convex. Hence, if we show that $(f_n^{-1}(t), t_1)$ and $(f_n^{-1}(t), t_2)$ are in $U_{\alpha\varepsilon}$, we get that $(f_n^{-1}(t), t)$ also belongs to this set. Since f_∞ is continuous then $U_{\alpha\varepsilon}$ is closed in $\mathbb{R}^2$. On the other hand, we can find $\eta > 0$ arbitrary small such that $f_n^{-1}(t) - \eta$ is a point of continuity for f_n, and hence $\Big(f_n^{-1}(t) - \eta, f_n\big(f_n^{-1}(t) - \eta\big)\Big)$ belongs to $U_{\alpha\varepsilon}$. As η goes to zero, $\Big(f_n^{-1}(t) - \eta, f_n\big(f_n^{-1}(t) - \eta\big)\Big)$ converges to $(f_n^{-1}(t), t_1)$ and hence $(f_n^{-1}(t), t_1)$ belongs to $U_{\alpha\varepsilon}$. With the same reasoning with t_2, we get as expected that $(f_n^{-1}(t), t)$ is in $U_{\alpha\varepsilon}$. Hence by (2.22) we obtain that $(f_n^{-1}(t), t)$ belongs to V_ε, i.e. $|f_n^{-1}(t) - f_\infty^{-1}(t)| \leqslant \varepsilon$. □

By Wigner's theorem [Gui09, Theorem 1.13], almost surely the empirical eigenvalue distribution of $X_j^{(N)}$ converges to the semicircular law with radius two, and hence $F_{X_j^{(N)}}$ converges pointwise to F_x. By Lemma 2.2.4, we get that almost surely $\|F_{X_j^{(N)}} - F_x\|_{\mathrm{L}^\infty([0,1])}$ goes to zero as N goes to infinity.

Similarly, almost surely the empirical eigenvalue distribution of $U_j^{(N)}$ converges to the uniform measure on the unit circle [AGZ10, Theorem 5.4.10]. Hence we get that almost surely $\|F_{U_j^{(N)}} - F_u\|_{\mathrm{L}^\infty([0,1])}$ tends to zero and by Lemma 2.2.5 we have that almost surely $\|F^{-1}_{U_j^{(N)}} - F_u^{-1}\|_{\mathrm{L}^\infty([0,1])}$ goes to zero as N goes to infinity.

Hence, by (2.19) and (2.20) we obtain that: for any polynomial P, almost surely one has

$$\Big\|P(\mathbf{U}_N, \mathbf{U}_N^*, \mathbf{Y}_N)\Big\| \underset{N\to\infty}{\longrightarrow} \Big\|P(\mathbf{u}, \mathbf{u}^*, \mathbf{y})\Big\|, \tag{2.23}$$

which completes the proof.

2.2.5 Proof of Theorem 2.1.5

The proof of Theorem 2.1.5 is obtained by changing the words unitary, Hermitian and GUE into orthogonal, symmetric and GOE, respectively symplectic, self dual and GSE, by taking $\mathbf{Y}_N = \mathbf{0}$ and citing the main results of [Sch05] instead of Theorem 2.1.3. In the symplectic case, we also have to consider matrices of even size.

2.3 Applications

Our main result has the potential for many applications in random matrix theory.

2.3.1 The spectrum of the sum and the product of Hermitian random matrices

Corollary 2.3.1. Let A_N, B_N be two $N \times N$ independent Hermitian random matrices. Assume that:

1. the law of one of the matrices is invariant under unitary conjugacy,
2. almost surely, the empirical eigenvalue distribution of A_N (respectively B_N) converges to a compactly supported probability measure μ (respectively ν),
3. almost surely, for any neighborhood of the support of μ (respectively ν), for N large enough, the eigenvalues of A_N (respectively B_N) belong to the respective neighborhood.

Then, one has

- almost surely, for N large enough, the eigenvalues of $A_N + B_N$ belong to a small neighborhood of the support of $\mu \boxplus \nu$, where $\boxplus$ denotes the free additive convolution (see [NS06, Lecture 12]).
- if moreover B_N is nonnegative, then the eigenvalues of $(B_N)^{1/2} A_N (B_N)^{1/2}$ belong to a small neighborhood of the support of $\mu \boxtimes \nu$, where $\boxtimes$ denotes the free multiplicative convolution (see [NS06, Lecture 14]).

Corollary 2.3.1 can be applied in the following situation. Let A_N be an $N \times N$ Hermitian random matrix whose law is invariant under unitary conjugacy. Assume that, almost surely, the empirical eigenvalue distribution of A_N converges to a compactly supported probability measure μ and its eigenvalues belong to the support of μ for N large enough. Let Π_N be the matrix of the projection on first p_N coordinates, $\Pi_N = \text{diag}\ (\mathbf{1}_{p_N}, \mathbf{0}_{N-p_N})$, where $p_N \sim tN$, $t \in (0,1)$. We consider the empirical eigenvalue distribution μ_N of the Hermitian random matrix

$$\Pi_n A_n \Pi_n.$$

Then, it follows from a Theorem of Voiculescu [Voi98] (see also [Col03]) that almost surely μ_N converges weakly to the probability measure $\mu^{(t)} = \mu \boxtimes [(1-t)\delta_0 + t\delta_1]$. This distribution is important in free probability theory because of its close relationship to the free additive convolution semigroup (see [NS06, Exercise 14.21]). Besides, the eigenvalue counting measure μ_N was proved to be a determinantal point process obtained as the push forward of a uniform measure in a Gelfand-Cetlin cone [Def10]. Very recently, it was proved by Metcalfe [Met] that the eigenvalues satisfy universality property inside the bulk of the spectrum. Our result complement his, by showing that almost surely, for N large enough there is no eigenvalue outside of any neighborhood of the spectrum of $\mu^{(t)}$.

Proof of Corollary 2.3.1. Without loss of generality, assume that the law of A_N is invariant under unitary conjugacy. Let $D_1^{(N)} = \text{ diag } (\lambda_1^{(N)}, \dots, \lambda_N^{(N)})$ be the diagonal matrix whose entries are the eigenvalues of B_N, sorted in non decreasing order. For any ρ in $[0,1]$, we set

$$D_1^{(N)}(\rho) = \text{ diag } (\lambda_{1+\lfloor \rho N \rfloor}^{(N)}, \dots, \lambda_{N+\lfloor \rho N \rfloor}^{(N)}), \text{ with indices modulo } N.$$

By the spectral theorem, we can write $B_N = V_N(\rho) D_1^{(N)}(\rho) V_N(\rho)^*$, where $V_N(\rho)$ is unitary, $(V_N(\rho), D_1^{(N)}(\rho))$ being independent of A_N. The law of the Hermitian matrix $V_N(\rho)^* A_N V_N(\rho)$ is still invariant under unitary conjugacy. Then, by Proposition 2.2.1, we can write $V_N(\rho)^* A_N V_N(\rho) = U_N D_2^{(N)} U_N^*$, where U_N is a Haar unitary matrix, $D_2^{(N)}$ is a real diagonal matrix whose entries are non decreasing along the diagonal, $U_N, D_1^{(N)}, D_2^{(N)}$ are independent.

By [Mal11, Corollary 2.1], there exists ρ in $[0,1]$ such that, almost surely, the non commutative law of $(D_1^{(N)}(\rho), D_2^{(N)})$ converges strongly to the law of a couple of non commutative random variables (d_1, d_2) in a $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ with faithful trace. Let u be a Haar unitary in $\mathcal{A}$, free from (d_1, d_2). By Theorem 2.1.4, we get that almost surely $U_N D_1^{(N)}(\rho) U_N^* + D_2^{(N)}$ converges strongly to $u d_1 u^* + d_2$. The spectrum of $A_N + B_N$ being the spectra of $U_N D_1^{(N)}(\rho) U_N^* + D_2^{(N)}$, we get the first point of Corollary 2.3.1 since strong convergence of random matrices implies convergence of the support.

We get the second point of Corollary 2.3.1 with the same reasoning on $\big((D_1^{(N)}(\rho))^{1/2}, D_2^{(N)}\big)$. The application stated after Corollary 2.3.1 follows by taking $\Pi_N = B_N$, which satisfies the assumptions since $t \in (0,1)$, and remarking that $\Pi_N^{1/2} = \Pi_N$.

□

2.3.2 Questions from operator space theory

The following question was raised by Gilles Pisier to one author (B.C.) ten years ago: Let $U_1^{(N)}, \dots, U_p^{(N)}$ be $N \times N$ independent unitary Haar random matrices. Is it true that

$$\Big\| \sum_{i=1}^{p} U_i^{(N)} \Big\| \underset{N\to\infty}{\longrightarrow} 2\sqrt{p-1} \tag{2.24}$$

almost surely? This question is very natural from the operator space theory point of view, and although at least ten years old, it was still open before this paper.

Our main theorem implies immediately that the answer is positive since $2\sqrt{p-1}$ is the norm of the sum of p free Haar unitaries, a computation that goes back to a paper of Akemann and Ostrand [AO76]. We can give some generalizations of (2.24).

From [AO76], we can deduce more generally that for any complex numbers $a_1, \ldots, a_p$, almost surely one has

$$\Big\| \sum_{i=1}^p a_i U_i^{(N)} \Big\| \underset{N\to\infty}{\longrightarrow} \min_{t\geqslant 0} \Big\{ 2t + \sum_{i=1}^p \Big(\sqrt{t^2+|a_i|^2} - t \Big) \Big\}.$$

By a result of Kesten [Kes59], the norm of the sum of p free Haar unitaries and of their conjugate equals $2\sqrt{2k-1}$. Hence, we get from our result that almost surely one has

$$\Big\| \sum_{i=1}^p \Big(U_i^{(N)} + U_i^{(N)*} \Big) \Big\| \underset{N\to\infty}{\longrightarrow} 2\sqrt{2p-1}.$$

Furthermore, recall that from Theorem 2.1.4 we can deduce the following corollary (see [Mal11, Proposition 7.3] for a proof). We use the notations of Theorem 2.1.4.

Corollary 2.3.2. Let $k \geqslant 1$ be an integer. For any polynomial P with coefficients in $\mathrm{M}_k(\mathbb{C})$, almost surely one has

$$\|P(\mathbf{U}_N, \mathbf{U}_N^*, \mathbf{Y}_N, \mathbf{Y}_N^*)\| \underset{N\to\infty}{\longrightarrow} \|P(\mathbf{u}, \mathbf{u}^*, \mathbf{y}, \mathbf{y}^*)\|,$$

where $\|\cdot\|$ stands in the left hand side for the operator norm in $\mathrm{M}_{kN}(\mathbb{C})$ and in the right hand side for the $\mathcal{C}^*$-algebra norm in $\mathrm{M}_k(\mathcal{A})$.

By Corollary 2.3.2 and Fell's absorption principle [Pis03, Proposition 8.1], we can answer the question asked by Pisier in [Pis03, Chapter 20]: for any $k \times k$ unitary matrices $a_1, \ldots, a_p$, almost surely one has

$$\Big\| \sum_{i=1}^p a_i \otimes U_i^{(N)} \Big\| \underset{N\to\infty}{\longrightarrow} 2\sqrt{p-1}.$$

2.3.3 Haagerup's inequalities

Let $\mathbf{u} = (u_1, \ldots, u_p)$ be free Haar unitaries in a $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \|\cdot\|)$. For any integer $d \geqslant 1$, we denote by W_d the set of elements of $\mathcal{A}$ of length d in $(\mathbf{u}, \mathbf{u}^*)$, i.e.

$$W_d \;=\; \Big\{ u_{j_1}^{\varepsilon_1} \ldots u_{j_d}^{\varepsilon_d} \;\Big|\; j_1 \neq \cdots \neq j_d,\ \varepsilon_j \in \{1, *\}\ \forall j = 1, \ldots, d \Big\}.$$

In 1979, Haagerup [Haa79] has shown that one has

$$\Big\| \sum_{n\geqslant 1} \alpha_n x_n \Big\| \leqslant (d+1) \|\alpha\|_2, \tag{2.25}$$

for any sequence $(x_n)_{n\geqslant 1}$ of elements in W_d and sequence $\alpha = (\alpha_n)_{n\geqslant 1}$ of complex numbers whose ℓ^2-norm is denoted by

$$\|\alpha\|_2 = \sqrt{\sum_{n\geqslant 1} |\alpha|^2}.$$

This result, known as Haagerup's inequality, has many applications and has been generalized in many ways. For instance, Buchholz has generalized (2.25) in an estimate of $\sum_{n\geqslant 1} a_n \otimes x_n$, where the a_n are now $k \times k$ matrices. Let $\mathbf{U}_N$ be a family of p independent $N \times N$ unitary Haar matrices. As a byproduct of our main result, we then get from (2.25) an estimate of the norm of matrices of the form

$$\sum_{n\geqslant 1} \alpha_n X_n^{(N)},$$

where for any $n \geqslant 1$, the matrix $X_n^{(N)}$ is a word of fixed length in $(\mathbf{U}_N, \mathbf{U}_N^*)$.

Kemp and Speicher [KS07] have generalized Haagerup's inequality for $\mathcal{R}$-diagonal elements in the so-called holomorphic case. Theorem 2.1.4 established, the consequence for random matrices sounds relevant since it allows to consider combinations of Haar and deterministic matrices. The result of [KS07] we state below has been generalized by de la Salle [dlS09] in the case where the non commutative random variables have matrix coefficients. This situation could be interesting for practical applications, where block random matrices are sometimes considered (see [TV04] for applications of random matrices in telecommunication). Nevertheless, we only consider the scalar version for simplicity.

Recall that a non commutative random variable a is called an $\mathcal{R}$-diagonal element if it can be written $a = uy$, for u a Haar unitary free from y (see [NS06]). Let $\mathbf{a} = (a_1, \dots, a_p)$ be a family of free, identically distributed $\mathcal{R}$-diagonal elements in a $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \| \cdot \|)$. We denote by W_d^+ the set of elements of $\mathcal{A}$ of length d in $\mathbf{a}$ (and not its conjugate), i.e.

$$W_d^+ = \left\{ a_{j_1} \dots a_{j_d} \ \middle| \ j_1 \neq \dots \neq j_d \right\}.$$

Kemp and Speicher have shown the following, where the interesting fact is that the constant $(d+1)$ is replace by a constant of order $\sqrt{d+1}$: for any sequence $(x_n)_{n\geqslant 1}$ of elements of W_d^+ and any sequence $\alpha = (\alpha_n)_{n\geqslant 1}$, one has

$$\Big\| \sum_{n\geqslant 1} \alpha_n x_n \Big\| \leqslant e\sqrt{d+1} \Big\| \sum_{n\geqslant 1} \alpha_n x_n \Big\|_2, \tag{2.26}$$

where $\| \cdot \|_2$ denotes the L^2-norm in $\mathcal{A}$, given by $\|x\|_2 = \tau[x^*x]^{1/2}$ for any a in $\mathcal{A}$. In particular, if $\mathbf{a}$ is a family of free unitaries (i.e. $y = \mathbf{1}$) then we get $\| \sum_{n\geqslant 1} \alpha_n x_n \Big\|_2 = \|\alpha\|_2$, so that (2.26) is already an improvement of (2.25) without the generalization on $\mathcal{R}$-diagonal elements.

Now let $\mathbf{U}_N = (U_1^{(N)}, \dots, U_p^{(N)}), \mathbf{V}_N = (V_1^{(N)}, \dots, V_p^{(N)})$ be families of $N \times N$ independent unitary Haar matrices and $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_p^{(N)})$ be a family of $N \times N$ deterministic Hermitian matrices. Assume that for any $j = 1, \dots, p$, the empirical spectral distribution of $Y_j^{(N)}$ converges weakly to a measure μ (that does not depend on j) and that for N large enough, the eigenvalues of $Y_j^{(N)}$ belong to a small neighborhood of the support of μ. We set for any $j = 1, \dots, p$ the random matrix

$$A_j^{(N)} = U_j^{(N)} Y_j^{(N)} V_j^{(N)*}.$$

From Theorem 2.1.4 and [Mal11, Corollary 2.1], we can deduce that almost surely the family $(A_1, \dots, A_p)$ converges strongly in law to a family of free $\mathcal{R}$-diagonal elements $(a_1, \dots, a_p)$, identically distributed. Hence, inequality (2.26) gives an asymptotic bound for the norm of a random matrix of the form

$$\sum_{n \geqslant 1} \alpha_n X_n^{(N)},$$

where for any $n \geqslant 1$, the matrix $X_n^{(N)}$ is a word of fixed length in $A_j^{(N)}, \dots, A_j^{(N)*}$.

2.4 Acknowledgments

B.C. would like to thank Gilles Pisier for asking him Question (2.24) many years ago, as it was a source of inspiration for this paper. C.M. would like thank his Ph.D. supervisor Alice Guionnet for submitting him as a part of his Ph.D. project a question that lead to the main theorem of this paper and for her insightful guidance. He would also like to thank Ofer Zeitouni and Mikael de la Salle for useful discussions during the preparation of this article.

This paper was initiated and mostly completed during the Erwin Schrödinger International Institute for Mathematical Physics workshop “Bialgebras in Free Probability” in April 2011. Both authors gratefully acknowledge the ESI and the organizers of the meeting for providing an inspiring working environment.

Both authors acknowledge the financial support of the ESI. The research of B.C. was supported in part by a Discovery grant from the Natural Science and Engineering Research Council of Canada, an Early Research Award of the Government of Ontario, and the ANR GranMa.

Part II

Fausse Liberté Asymptotique

Chapter 3

Free probability on traffics: the limiting distribution of heavy Wigner and deterministic matrices.

Work in progress.

ABSTRACT:

We characterize the limiting eigenvalue distribution of N by N Hermitian matrices obtained as polynomial in certain random and deterministic matrices when their size goes to infinity. The random matrices, called heavy Wigner matrices, are independent, Hermitian and their sub-diagonal entries are independent, distributed according to a probability measure whose moments are large when N is large. The deterministic matrices are assumed to satisfy a new kind of convergence, called the convergence in distribution of traffics. This convergence carries much more information on the N by N matrices than the convergence in the sense of free probability. For an adjacency matrix of a graph, this convergence is equivalent to the weak local convergence of the graph.

3.1 Introduction

The ensemble of Wigner matrices has been introduced by Wigner [Wig58] in 1958. An N by N real matrix X_N is called a Wigner matrix whenever it is Hermitian and the sub-diagonal entries of $\sqrt{N}X_N$ are independent, identically distributed according to a probability measure whose moments are finite. This ensemble forms a large class of universality, in the sense that most of the statistical properties of the spectrum of a large Wigner matrix does not depend on the detail of the law of its entries.

The most famous result of universality is Wigner's semicircular law. Let X_N be a Wigner matrix such that the sub-diagonal entries of $\sqrt{N}X_N$ are of variance

1. Then, Wigner has proved [Wig58] that the mean eigenvalue distribution $\mathcal{L}_{X_N}$ of X_N converges in moments to the semicircular law with radius 2, i.e. for any polynomial P, one has

$$\mathcal{L}_{X_N}(P) = \mathbb{E}\Big[\tau_N\Big[P(X_N)\Big]\Big] \underset{N\to\infty}{\longrightarrow} \int_{-2}^{2} P(t)\frac{1}{2\pi}\sqrt{4-t^2}\mathrm{d}t,$$

where τ_N denotes the normalized trace of N by N matrices and $\mathbb{E}$ denotes the expectation relative to the entries of X_N. The matrix $P(X_N)$ is obtained by functional calculus.

This result has been generalized in many directions. In this article, we are interested in the situation where many matrices are involved. The pioneering works in this context are due to Voiculescu [Voi91] in 1991. They had a strong impact since Voiculescu has established in these papers a connection between random matrix theory and free probability. His main theorem in [Voi91] have been generalized by Dykema [Dyk93] who has shown in 1993 the following result. Let $X_1^{(N)}, \dots, X_p^{(N)}$ be independent N by N Wigner matrices (in [Voi91] the matrices are Gaussian). Let P be a polynomial in p non commutative indeterminates such that almost surely the matrix $H_N = P(X_1^{(N)}, \dots, X_p^{(N)})$ is Hermitian. Then, the mean eigenvalue distribution of H_N converges in moments to a probability measure on the real line which only depends on the polynomial P. Moreover, its limit can be described in the context of Voiculescu's free probability theory.

This article is motivated by the following question: how can we generalize Voiculescu's theorem for symmetric matrices with independent heavy tailed entries, and then understand an analogue of free probability theory for these matrices? In term of methodology, the main difficulty is the absence of reference model for heavy tailed matrices, as the Gaussian matrices are for Wigner matrices. To avoid this difficulty, we consider the following ensemble of random matrices.

Definition 3.1.1 (Heavy Wigner matrices)**.**
A sequence of random matrices $(X_N)_{N\geqslant 1}$ is called a sequence of heavy Wigner matrices whenever

1. for any $N \geqslant 1$, the matrix $A_N = \sqrt{N}X_N$ is N by N, real symmetric. The sub-diagonal entries of A_N are independent, identically distributed according to a measure $p^{(N)}$ on $\mathbb{R}$ which possesses all its moments,
2. for any $k \geqslant 1$, the sequence of $2k$-th moments satisfies

$$a_k := \lim_{N\to\infty} \frac{\int t^{2k}\mathrm{d}p^{(N)}(t)}{N^{k-1}} \text{ exists in } \mathbb{R},$$

3. one has $\sqrt{N}\int t\mathrm{d}p^{(N)}(t) = o(N^\beta)$ for any $\beta > 0$.

The sequence of non negative numbers $(a_k)_{k\geqslant 1}$ is called the parameter of $(X_N)_{N\geqslant 1}$. When we say that an N by N random matrix X_N is a heavy Wigner matrix, we implicitly mean that we have considered a sequence $(X_N)_{N\geqslant 1}$; by the parameter of X_N, we mean the parameter of this sequence. We say that the parameter $(a_k)_{k\geqslant 1}$ of a heavy Wigner matrix is trivial as soon as $a_k = 0$ for any $k \geqslant 2$.

A Wigner matrix is then a heavy Wigner matrix whose common law of entries does not depend on N. As explained at the end of this introduction, the ensemble of heavy Wigner matrices is an approximating model for matrices with heavy tailed entries.

Heavy Wigner matrices have been previously introduced and studied independently by two authors. In 2005, Zakharevich [Zak06] has studied the limiting mean eigenvalue distribution of a single heavy Wigner matrix.

Theorem 3.1.2 (The spectrum of a single heavy Wigner matrix)**.**
Let X_N be a heavy Wigner matrix with parameter $a = (a_k)_{k\geqslant 1}$. Then, its mean eigenvalue distribution $\mathcal{L}_{X_N}$ of X_N converges in moments to a symmetric probability measure μ_a on $\mathbb{R}$ depending only on a, i.e. for any polynomial P, one has

$$\mathcal{L}_{X_N}(P) := \mathbb{E}\Big[\tau_N\big[P(X_N)\big]\Big] \underset{N\to\infty}{\longrightarrow} \int P(t)\mathrm{d}\mu_a(t). \tag{3.1}$$

The measure μ_a is shown [Zak06] to be the semicircular distribution with radius $\sqrt{a_1}$ as soon as the parameter of X_N is trivial. Otherwise, little is known about μ_a. Zakharevich has shown a formula to compute the moments of μ_a based on the enumeration of certain colored rooted tress and she proved that μ_a has a unbounded support.

Ryan [Rya98] has established in 1997 a more general version of (3.1) for independent heavy Wigner matrices in the context of free probability.

Theorem 3.1.3 (The limiting distribution of independent heavy Wigner matrices)**.** Let $X_1^{(N)},\dots,X_p^{(N)}$ be a family of independent heavy Wigner matrices. Denote by $\mathbb{C}\langle x_1,\dots,x_p\rangle$ the set of non commutative polynomials in p non commutative indeterminates $x_1,\dots,x_p$. Then, for any polynomial P in $\mathbb{C}\langle x_1,\dots,x_p\rangle$,

$$\tau[P] := \underset{N\to\infty}{\mathbb{E}}\Big[\tau_N\big[P(X_1^{(N)},\dots,X_p^{(N)})\big]\Big] \text{ exists,} \tag{3.2}$$

and the linear form τ on $\mathbb{C}\langle x_1,\dots,x_p\rangle$ depends only on the parameters of the matrices.

Let $\mathbf{X}_N = (X_1^{(N)},\dots,X_p^{(N)})$ be a family of independent heavy Wigner matrices. Consider the Hermitian matrix

$$H_N = Q(X_1^{(N)},\dots,X_p^{(N)}),$$

where Q is a polynomial in p non commutative indeterminates (fixed and such that H_N is Hermitian). Then the convergence (3.2) applied to the polynomials Q^k for any $k\geqslant 1$ gives the convergence in moments of the mean eigenvalue distribution $\mathcal{L}_{H_N}$ of H_N.

Non commutative probability theory gives a conceptual framework to handle this kind of convergence. Recall the following definitions (see [AGZ10, Gui09, NS06]).

Definition 3.1.4 (Non commutative probability vocabulary)**.**

1. A *-probability space $(\mathcal{A}, .^*, \tau)$ consists of a unital $\mathbb{C}$-algebra $\mathcal{A}$ endowed with an antilinear involution $.^*$ such that $(ab)^* = b^*a^*$ for all a, b in $\mathcal{A}$, and a tracial state τ. A tracial state τ is a linear functional $\tau : \mathcal{A} \mapsto \mathbb{C}$ satisfying
$$\tau[\mathbf{1}] = 1, \quad \tau[ab] = \tau[ba], \quad \tau[a^*a] \geqslant 0 \ \forall a, b \in \mathcal{A}. \tag{3.3}$$
The elements of $\mathcal{A}$ are called non commutative random variables.
2. The joint distribution of a family $\mathbf{a} = (a_1, \dots, a_p)$ of non commutative random variables is the linear form
$$\begin{array}{rccc} \tau_{\mathbf{a}}: & \mathbb{C}\langle \mathbf{x}, \mathbf{x}^* \rangle & \to & \mathbb{C} \\ & P & \mapsto & \tau\Big[P(\mathbf{a}, \mathbf{a}^*)\Big], \end{array}$$
where $\mathbb{C}\langle \mathbf{x}, \mathbf{x}^* \rangle$ is the set of polynomials in $2p$ non commutative indeterminates $x_1, \dots, x_p, x_1^*, \dots, x_p^*$ and $P(\mathbf{a}, \mathbf{a}^*)$ is a shortcut for $P(a_1, \dots, a_p, a_1^*, \dots, a_p^*)$.
3. The convergence in distribution of a sequence of families $(\mathbf{a}_N)_{N\geqslant 1}$ is the pointwise convergence of sequence of functionals $(\tau_{\mathbf{a}_N})_{N\geqslant 1}$.

A family of independent heavy Wigner matrices $X_1^{(N)}, \dots, X_n^{(N)}$ is a n-tuple in the algebra $\cap_{p\geqslant 1} L^p\Big(\Omega, M_N(\mathbb{C})\Big)$ of N by N random matrices whose entries admitting all their moments. This algebra is equipped with $\mathbb{E}[\tau_N]$, the expectation of the normalized trace and $.^*$ the conjugate transpose.

Voiculescu [Voi95a] has introduced the notion of freeness for non commutative random variables. It describes the structure of the *-probability space where the limit in distribution of independent Wigner matrices lives. It is a non commutative analogue of the notion of independence for random variables which allows to compute the joint distribution of non commutative random variables from the knowledge of the marginal distributions only.

Definition 3.1.5 (Freeness)**.** Let $(\mathcal{A}, .^*, \tau)$ be a *-probability space. Let $\mathcal{A}_1, \dots, \mathcal{A}_k$ be $*$-subalgebras of $\mathcal{A}$ having the same unit as $\mathcal{A}$. They are said to be free if for any integer $n \geqslant 1$, any $a_i \in \mathcal{A}_{j_i}$ $(i = 1, \dots, n,\ j_i \in \{1, \dots, k\})$, one has
$$\tau\Big[\Big(a_1 - \tau[a_1]\Big) \cdots \Big(a_n - \tau[a_n]\Big)\Big] = 0$$
as soon as $j_1 \neq j_2, j_2 \neq j_3, \dots, j_{n-1} \neq j_n$. Collections of random variables are said to be free if the unital subalgebras they generate are free.

As a generalization of Voiculescu's theorem for independent Gaussian matrices, it is known since the works of Dykema [Dyk93] that the law of independent Wigner matrices $X_1^{(N)}, \dots, X_p^{(N)}$ in $\Big(\cap_{p\geqslant 1} L^p\Big(\Omega, M_N(\mathbb{C})\Big), .^*, \mathbb{E}[\tau_N]\Big)$ converges to the law of free semicircular variables $x_1, \dots, x_p$ in a *-probability space $(\mathcal{A}, .^*, \tau)$, i.e.

1. for any $m = 1, \dots, p$, there exists $a_m \geqslant 0$, such that for every $k \geqslant 1$, one has
$$\tau_{SC}[x_m^k] = \int t^k \mathrm{d}\sigma(t)$$
where $\mathrm{d}\sigma(t) = \frac{1}{2\pi}\sqrt{4 - \frac{t^2}{a_m}}\mathbf{1}_{|t| \leqslant 2\sqrt{a_m}}$ is the semicircular distribution of radius $2\sqrt{a_m}$,
2. the variables $x_1, \dots, x_p$ are free.

In contrast, Ryan has established in his Ph.D. thesis that the limiting distribution of independent heavy Wigner matrices is the distribution of free non commutative random variables if and only if at most one matrix has a non trivial parameter. The purpose of this paper is to generalize Ryan's result in the following way: we consider a family $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ of independent N by N heavy Wigner matrices and a family $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ of N by N deterministic matrices and we state sufficient assumptions on the family $\mathbf{Y}_N$ such that the joint distribution of $(\mathbf{X}_N, \mathbf{Y}_N)$ in $(\cap_{p \geqslant 1} L^p\big(\Omega, M_N(\mathbb{C})\big), .^*, \mathbb{E}[\tau_N])$ converges in distribution (Theorem 3.3.8).

When the matrices $\mathbf{X}_N$ are Wigner matrices, this result is known as Voiculescu's asymptotic freeness theorem for random matrices (see [AGZ10, Theorem 5.4.5]). Up to technical conditions, if the family $\mathbf{Y}_N$ has a limiting distribution, then the families $\mathbf{X}_N$ and $\mathbf{Y}_N$ are asymptotically free. In particular the limiting joint distribution of $(\mathbf{X}_N, \mathbf{Y}_N)$ depends only on the limiting distribution of $\mathbf{Y}_N$.

We show that this fact is not true when the matrices $\mathbf{X}_N$ are heavy Wigner matrices. Strictly more information on $\mathbf{Y}_N$ is needed to describe the possible limiting distributions of $(\mathbf{X}_N, \mathbf{Y}_N)$. This phenomenon reflects an other one that appears in the study of the spectrum of related random matrix models. Ben Arous and Guionnet [BAG08] have shown that the limiting Stieltjes transform of the empirical eigenvalue distribution of a single Lévy matrix (see the definition below) can be characterized by a closed equation which involves the limiting eigenvalue distribution of the uniform measure on the diagonal elements of the resolvant of the matrix (see also [BDG09]). Khorunzhy, Shcherbina and Vengerovsky [KSV04] have observed the same fact in the study of the adjacency matrix of large weighted graphs.

We introduce in Section 3.3.1 the notion of distribution of traffics, which encodes the information needed on the family $\mathbf{Y}_N$ of deterministic matrices to infer the limiting joint distribution of $(\mathbf{X}_N, \mathbf{Y}_N)$. Heuristically, the idea is to replace in the definition of distribution non commutative polynomials by finite, connected, directed graphs whose edges are labelled by indeterminates. We prove that, up to technical conditions, if $\mathbf{Y}_N$ has a limiting distribution of traffics then $(\mathbf{X}_N, \mathbf{Y}_N)$ also satisfies this property (see Theorem 3.3.8).

The distribution of traffics contains the information about the distribution in sense of *-probability spaces. For instance, let A_N, B_N be deterministic N by N matrices having a limiting distribution of traffics with A_N Hermitian. Let X_N

be an N by N heavy Wigner matrix. Under technical assumptions, we obtain from Theorem 3.3.8 the convergence of the mean eigenvalue distributions of the matrices $A_N + X_N$ and $B_N X_N B_N^*$ when N goes to infinity.

It turns out that the convergence in distribution of traffics generalizes the so-called weak local convergence of graphs introduced by Benjamini and Schramm [BS01] and developed by Aldous and Steele [AS04] (see Section 3.4). Let G_N be a graph with N vertices. Under technical assumptions, the convergence of a graph G_N is the weak local sense is equivalent to the convergence in distribution of traffics for one of its adjacency matrix.

At last, the language of distribution of traffics is useful to shed light on the non free relation between limits of heavy Wigner matrices. Let $\mathbf{X}_N$ be a family of independent heavy Wigner matrices and $\mathbf{Y}_N$ a family of deterministic matrices satisfying the assumptions of Theorem 3.3.8. Then, $(\mathbf{X}_N, \mathbf{Y}_N)$ converges in the sense of *-probability space to a family $(\mathbf{x}, \mathbf{y})$ of non commutative random variables. In general, the families $\mathbf{x}$ and $\mathbf{y}$ are not free. Nevertheless, in some heuristic sense the architecture of freeness rules the distribution of $(\mathbf{x}, \mathbf{y})$. Hence, we use the term false freeness for the relationship between the families $\mathbf{x}$ and $\mathbf{y}$.

The lack of freeness of $\mathbf{x}$ and $\mathbf{y}$ can be measured by using the following multi-linear forms. For any integer $K \geqslant 1$, we set

$$\begin{array}{cccc} \Phi_N^{(K)} & : & \mathrm{M}_N(\mathbb{C})^K & \rightarrow & \mathbb{C} \\ & & (A_1, \dots, A_K) & \mapsto & \tau_N[A_1 \circ A_2 \circ \cdots \circ A_K], \end{array}$$

where the symbol $\circ$ designates the entry-wise matrix multiplication, known as the Hadamard product. Remark that in particular one has $\Phi_N^{(1)} = \tau_N$. The convergence in distribution of traffics of $(\mathbf{X}_N, \mathbf{Y}_N)$ implies that for any polynomial $P_1, \dots, P_K$,

$$\Phi^{(K)}(P_1, \dots, P_K) := \lim_{N \to \infty} \Phi^{(K)}\Big(P_1(\mathbf{X}_N, \mathbf{Y}_N), \dots, P_K(\mathbf{X}_N, \mathbf{Y}_N)\Big) \text{ exists.}$$

These maps are useful to compute joint moments in $(\mathbf{x}, \mathbf{y})$ and to see when the freeness properties between $\mathbf{x}$ and $\mathbf{y}$ is broken. Moreover, when the deterministic matrices are diagonal, we show that the family of multi-linear forms $\big(\Phi^{(K)}\big)_{K \geqslant 1}$ satisfies a system of equations that generalizes the Schwinger-Dyson equation for free semicircular variables.

Before going further, we precise the connection between the model of heavy Wigner matrices studied in this paper and the model of symmetric matrices with heavy tailed entries, called Lévy matrices.

Definition 3.1.6 (Lévy matrices)**.**
An N by N symmetric random matrix $X_N = \big(X_N(i,j)\big)_{i,j=1,\dots,N}$ is called a Lévy matrix whenever for any $i, j = 1, \dots, N$,

$$X_N(i,j) = \frac{x_{i,j}}{\sigma_N},$$

where the random variables $(x_{i,j})_{1\leqslant i\leqslant j\leqslant N}$ are independent, identically distributed according to a law that belongs to the domain of attraction of an α stable law for an α in $]0,2[$. In other words, there exists a function $L:\mathbb{R}\to\mathbb{R}$ slowly varying such that

$$\mathbb{P}\Big(|x_{1,1}|\geqslant u\Big)=\frac{L(u)}{u^\alpha},\forall u\in\mathbb{R}.$$

Moreover, we have denoted the normalizing sequence

$$\sigma_N=\inf\left\{\ u\in\mathbb{R}^+\ \middle|\ \mathbb{P}\Big(|x_{1,1}|\geqslant u\Big)\leqslant\frac{1}{N}\ \right\}.$$

The number α is called the parameter of X_N.

Let X_N be a Lévy matrix of parameter α in $]0,2[$. With the notations of Definition 3.1.6, we consider for any $B>0$ the random matrix X_N^B whose entries are given by: for any $i,j=1,\dots,N$,

$$X_N^B(i,j)=\frac{x_{i,j}}{\sigma_N}\mathbf{1}_{|x_{i,j}|\leqslant Ba_N}.$$

Then [BAG08, Lemme 9.1], the matrix X_N^B is a heavy Wigner matrix whose parameter $(a_k^B)_{k\geqslant1}$ is: for any $k\geqslant1$

$$a_k^B\ =\ \frac{2-\alpha}{2k-\alpha}\Big(\frac{2-\alpha}{\alpha}B^\alpha\Big)^{k-1}.$$

Hence, the ensemble of Lévy matrices is at the frontier of the ensemble of heavy Wigner matrices.

The model of Lévy matrices has been introduced in 1994 by Bouchaud and Cizeau [BC94]. Pioneering works are due to Ben Arous and Guionnet [BAG08] in 2007, who have shown the convergence of the mean eigenvalue distribution of a single Lévy matrix. Belinschi, Dembo et Guionnet [BDG09] has studied in 2009 the perturbation of a Lévy matrix by a diagonal matrix and a band Lévy matrices. Moreover, Bordenave, Caputo and Chafaï [BCC11] has given in 2010 an other characterization of the limiting distribution of a Lévy matrix than the one of Ben Arous and Guionnet. It is based on the local operator convergence of a Lévy matrix to a certain graph whose entries are labelled by random variables, the Poissonian weighted infinite tree. This convergence is not far from being a convergence of traffics.

Organization of the paper:
The sections 3.2 to 3.6 are devoted to the presentation of the results. In Section 3.2, we give our approach to describe the limiting distribution of independent heavy Wigner matrices via cycles coloring a tree. We also state the so-called false freeness property which is useful for practical computations of moments. Section 3.3 is devoted to the definition of the convergence in distribution of traffics and to the statement of the main result of this paper. In Section 3.4, we remind the notion of weak local convergence for graphs and show that it is equivalent to the

convergence in distribution of traffics. Sections 3.5 and 3.6 are devoted to the applications of our main result. Sections 3.7 to 3.9 contains the proofs of our results. In a appendix at the end of the article, we give a short discussion on the model of heavy Wigner matrices.

Acknowledgment:
The author would like to gratefully thank Alice Guionnet for suggesting such an interesting problem. He would like to thank Florent Benaych-Georges, Charles Bordenave and Mikael de la Salle for useful discussions. He also acknowledges the financial support of the ANR GranMa.

This paper has known significant progress during some travels in 2011: a visit of Ashkan Nikeghbali and Kim Dang at the Institut für Mathematik of Zurich, the school "Vicious walkers and random matrices" in les Houches organized by the CNRS, and the summer school "Random matrix theory and its applications to high-dimensional statistics" in Changchun, founded jointly by the CNRS of France and the NSF of China. The author gratefully acknowledges the organizers of these events for providing an inspiring working environment.

3.2 The limiting distribution of heavy Wigner matrices via cycles coloring a tree

In this section, we give a new description of the limiting distribution of independent heavy Wigner matrices. A common way to compute joint moments of free semicircular variables consists of the enumeration of cycles coloring a tree. This point of view can be easily adapted to describe the asymptotic of heavy Wigner matrices. Moreover, it has the advantage of requiring less definitions than Ryan's formulation [Rya98] which is based on the so-called clickable partitions.

3.2.1 Reminder on free semicircular variables

Let τ_{SC} be the joint distribution of p free semicircular variables $x_1, \dots, x_p$ (as in the introduction). Assume that the variables are standard ($\tau_{SC}[x_m] = 0$ and $\tau_{SC}[x_m^2] = 1$ for any $m = 1, \dots, p$). Let $L \geqslant 1$ be an integer and $\ell = (\ell_1, \dots, \ell_L)$ in $\{1, \dots, p\}^L$ which will be referred to as a sequence of colors. We recall a formula for the joint moment $\tau_{SC}[x_{\ell_1} \dots x_{\ell_L}]$. By linearity, this formula characterizes τ_{SC} as a linear functional on non commutative polynomials.

Given a path of length L on a graph, the sequence of colors ℓ gives a coloration of the steps of c: for any $n = 1, \dots, L$, the n-th step of c is said to be of color ℓ_n. We say that such a path colors a graph via ℓ whenever it visits all the edges of the graph and any edge is visited by steps of the same color. We set $\mathcal{L}_{SC}^{(\ell)}$ the set of couples (T, c) where

- T is a rooted tree (one edge is specified), embedded in the plane with exactly $L/2$ edges.

- c is a cycle of length L coloring the tree T via ℓ, starting from the root and visiting T in the clockwise direction relative to the embedding of the tree.

Then, with these notations and definitions, one has

$$\tau_{SC}[x_{\ell_1} \dots x_{\ell_L}] = \text{Card}\left(\mathcal{L}_{SC}^{(\ell)}\right). \tag{3.4}$$

3.2.2 The limiting distribution of independent heavy Wigner matrices, heavy semicircular variables

Let τ be the limiting distribution of independent heavy Wigner matrices $X_1^{(N)}, \dots, X_p^{(N)}$, i.e. for any polynomial P in $\mathbb{C}\langle x_1, \dots, x_p \rangle$,

$$\mathbb{E}\Big[\tau_N\Big[P(X_1^{(N)}, \dots, X_p^{(N)})\Big]\Big] \underset{N\to\infty}{\longrightarrow} \tau[P]. \tag{3.5}$$

Let $L \geqslant 1$ be an integer and $\ell = (\ell_1, \dots, \ell_L)$ in $\{1, \dots, p\}^L$ be a sequence of colors. We give a formula for the joint moment $\tau[x_{\ell_1} \dots x_{\ell_L}]$ that generalizes (3.4). Our language is the following.

Definition 3.2.1 (Cycles coloring a tree)**.**

1. Given an integer $L \geqslant 1$ and a sequence of colors $\ell = (\ell_1, \dots, \ell_L)$ in $\{1, \dots, p\}^L$, we denote $\mathcal{L}^{(\ell)}$ the set of couples (T, c) where
   - T is a rooted tree, embedded in the plane with at most $L/2$ edges.
   - c is a cycle of length L coloring the tree T via ℓ, starting from the root and such that when c visits a new edge (necessarily moving away from the root), it visits the first unvisited edge relatively to the clockwise orientation.
2. Given such a couple (T, c), for any edge e of T we denote by $2n(e)$ the number of times c visits the edge e and by $\eta(e)$ the color of the steps of c corresponding to e.

The difference with the definition of $\mathcal{L}_{SC}^{(\ell)}$ is that in this situation a cycle is allowed to come back on edges it has already visited. In particular, the set $\mathcal{L}_{SC}^{(\ell)}$ is included in $\mathcal{L}^{(\ell)}$.

Theorem 3.2.2 (The limiting distribution of independent heavy Wigner matrices)**.** Let τ be the limiting distribution of independent heavy Wigner matrices $X_1^{(N)}, \dots, X_p^{(N)}$. For any $m = 1, \dots, p$, we set $(a_{m,k})_{k\geqslant 1}$ the parameter of the matrix $X_m^{(N)}$. Then, for any integer $L \geqslant 1$ and any $\ell = (\ell_1, \dots, \ell_L)$ in $\{1, \dots, p\}^L$, one has

$$\tau[x_{\ell_1} \dots x_{\ell_L}] = \sum_{(T,c)\in\mathcal{L}^{(\ell)}} \omega^{(1)}(c), \tag{3.6}$$

where

$$\omega^{(1)}(c) = \prod_{e \text{ edge of } T} a_{\eta(e),n(e)}.$$

This theorem is a special case of Theorem 3.3.8 which is proved in Section 3.7. Given p sequences of non negative integers $(a_{1,k})_{k\geqslant 1}, \dots, (a_{p,k})_{k\geqslant 1}$, we give for convenience a name to the distribution τ given by Formula (3.6). The choice in the terminology will become more meaningful in Section 3.5.

Definition 3.2.3 (Heavy semicircular variables)**.**
Let $x_1, \dots, x_p$ be non commutative random variables in a *-probability space $(\mathcal{A}, .^*, \tau)$. We say that $x_1, \dots, x_p$ are heavy semicircular variables whenever their distribution τ is given by Formula (3.6) for a certain sequences $(a_{1,k})_{k\geqslant 1}, \dots, (a_{p,k})_{k\geqslant 1}$. The sequence $\mathbf{a} = (a_{m,k})_{m=1,\dots,p}^{k\geqslant 1}$ forms the parameter of τ, and $(a_{m,k})_{k\geqslant 1}$ is called the parameter of x_m, $m = 1, \dots, p$.

The possible parameters of heavy Wigner matrices have strong restrictions: if a sequence $(a_k)_{k\geqslant 1}$ is a parameter, then it is the null sequence or it is the sequence of even moments of a Borel measure (see Appendix 3.10). For instance, if $(a_k)_{k\geqslant 1}$ is not a trivial parameter, then $a_k > 0$ for any $k \geqslant 1$. The role played by these numbers in the distribution τ looks very different from the role played by moments. That is why we have chosen to not taking into account this restriction in our definition of heavy semicircular variables.

Let τ be a distribution of heavy semicircular variables and denote by $\mathbf{a} = (a_{m,k})_{m=1,\dots,p}^{k\geqslant 1}$ its parameter. From Formula (3.6), we get easily the following facts.

- $\tau = \tau_{SC}$ as soon as the parameters of the matrices are trivial.
- $\tau[x_m] = 0$ and $\tau[x_m^2] = a_{m,1}$ for any $m = 1, \dots, p$.
- $\tau[x_{\ell_1} \dots x_{\ell_L}]$ vanishes as soon as the number of occurrence of one variable is odd.
- for any integers $n_1, \dots, n_L \geqslant 0$ and any distinct indices $m_1, \dots, m_L$ in $\{1, \dots, p\}$, one has $\tau[x_{m_1}^{n_1} \dots x_{m_L}^{n_L}] = \tau[x_{m_1}^{n_1}] \dots \tau[x_{m_L}^{n_L}]$.

A more subtle but direct consequence of this formula is that if all the parameters of $x_1, \dots, x_p$, except possibly one, are trivial, then τ is the distribution of free variables. Indeed, we get in this situation the classical Schwinger-Dyson equation for the semicircular variables (see Section 3.6 for a generalization of this fact). The reciprocal is true but is less easy to see, this is the purpose of the next section to make it clear.

3.2.3 The false freeness property of heavy semicircular variables.

The false freeness property is a simple observation. Let $L \geqslant 1$ be an integer and $\ell = (\ell_1, \dots, \ell_L)$ in $\{1, \dots, p\}^L$ be a sequence of colors. Then, we can get all the elements of $\mathcal{L}^{(\ell)}$ by "folding" the trees of $\mathcal{L}_{SC}^{(\ell)}$. This fact turns out to be particularly useful to compute joint moments of heavy semicircular variables and to understand when the freeness property is broken. We will deduce the following.

Proposition 3.2.4 (The lack of freeness of heavy semicircular variables)**.**
Let τ be a distribution of heavy semicircular variables. If the parameters of

at least two indeterminates are not trivial, then τ is not a distribution of free variables.

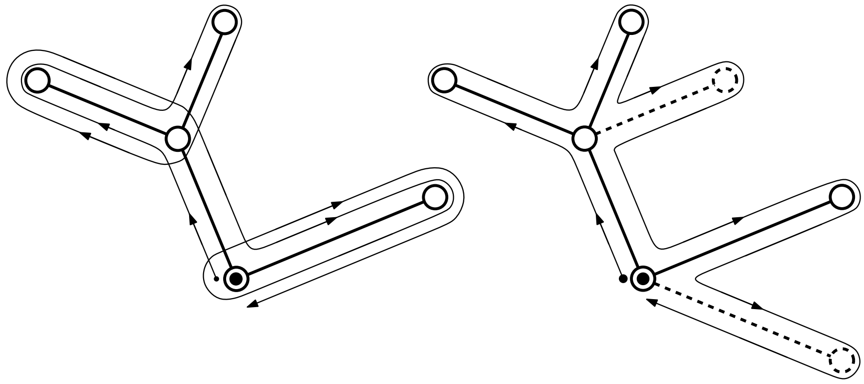

Figure 3.1: At the left, an element of $\mathcal{L}^{(\ell)}$, at the right an element of $\mathcal{L}^{(\ell)}_{SC}$, for $\ell = (1,\dots,1)$. They are related each other by the folding/unfolding trick.

We first describe the folding trick, that consists of obtaining an element of $\mathcal{L}^{(\ell)}_{SC}$ from an element of $\mathcal{L}^{(\ell)}$. Let (T,c) an element of $\mathcal{L}^{(\ell)} \setminus \mathcal{L}^{(\ell)}_{SC}$. After some steps, leaving a vertex s the cycle c comes back in an edge it has already visited. Then it induces a sub-cycle $\hat{c}$ on the tree of the descendent of s. We create a copy $\hat{T}$ of the sub-tree induces by $\hat{c}$, forget its original embedding and embed it in such a way $\hat{c}$ respects the rules concerning the order of visits of the edges of $\hat{T}$. Then we attach $\hat{T}$ endowed with this new orientation at the vertex s, between the edges it has already visited and the others. If some edges of the tree of the descendent of s where only visited by $\hat{c}$, then we erase them. We then keep an element of $\mathcal{L}^{(\ell)}$. Iterating this procedure a finite number of times, we then get an element of $\mathcal{L}^{(\ell)}_{SC}$.

Reciprocally, let (T,c) be an element of $\mathcal{L}^{(\ell)}$. Chose an edge e_1 of the tree. If possible, chose an other edge e_2, which shares the same vertex toward the root and which is of the same color as e_1. Then, merge these two edges, draw the tree of the descendant of e_1 at the right of the the tree of the descendant of e_2 and redirect the cycle c in this new tree. We then obtain an new element of $\mathcal{L}^{(\ell)}$. For any element (T_0,c_0) of $\mathcal{L}^{(\ell)}_{SC}$, we denote by $fold(T_0,c_0)$ the set of all elements of $\mathcal{L}^{(\ell)}$ we get by applying many times this trick. By the reverse construction above, we get the following.

Proposition 3.2.5 (The false freeness property)**.**
For any ℓ in $\{1,\dots,p\}^L$, one has

$$\sum_{(T,c)\in\mathcal{L}^{(\ell)}} \omega^{(1)}(c) = \sum_{(T_0,c_0)\in\mathcal{L}^{(\ell)}_{SC}} \ \sum_{(T,c)\in fold(T_0,c_0)} \omega^{(1)}(c). \tag{3.7}$$

In Figure 3.1, we have drawn two trees related by the folding/unfolding trick involved in the computation of $\tau\Big[(x_1)^{12}\Big]$. Using the false freeness property, we show the following result which implies Proposition 3.2.4.

Lemma 3.2.6 (Application of the false freeness property)**.**
Let τ be the distribution of two heavy semicircular variables x_1, x_2 with non trivial parameters. For $i = 1, 2$, denote by $(a_{i,k})_{k\geqslant 1}$ the parameter of x_i. We set $k_i = \min\{k \geqslant 2 | a_{i,k} \neq 0\} < \infty$, $i = 1, 2$. Then, for any integers $L \geqslant 2$ and any $n_1, \dots, n_L, m_1, \dots, m_L \geqslant 1$ such that $n_1 + \dots + n_L = k_1$ et $m_1 + \dots + m_L = k_2$, one has

$$\tau\Big[\Big(x_1^{2n_1} - \tau[x_1^{2n_1}]\Big)\Big(x_2^{2m_1} - \tau[x_2^{2m_1}]\Big)\dots\Big(x_1^{2n_L} - \tau[x_1^{2n_L}]\Big)\Big(x_2^{2m_L} - \tau[x_2^{2m_L}]\Big)\Big] = a_{1,k_1}a_{2,k_2}.$$

This lemma is proved in Section 3.9.1. Furthermore, the false freeness property gives a method to reasonably compute joint moments of heavy semicircular variables.

1. Enumerate the elements of $\mathcal{L}_{SC}^{(\ell)}$.
2. Fold the branches of these colored trees.
3. Then, read the contribution of all elements.

Example of computation: We apply this method to show that

$$\tau[x_1^2 x_2^2 x_1^2 x_2^2] = 3a_{1,1}^2 a_{2,1}^2 + a_{1,1}^2 a_{2,2} + a_{1,2}a_{2,1}^2 + a_{1,2}a_{2,2}$$

in Section 3.9.3

3.2.4 Motivations for the introduction of distribution of traffics: semicircular variables free from arbitrary variables

Let τ be the joint distribution of a family $\mathbf{x} = (x_1, \dots, x_p)$ of free semicircular variables, free from an arbitrary family $\mathbf{y} = (y_1, \dots, y_q)$. The freeness of the families $\mathbf{x}$ and $\mathbf{y}$ implies that the knowledge of the distribution of the family $\mathbf{y}$ determines completely τ. We recall a formula to compute the joint moments in $\mathbf{x}$ and $\mathbf{y}$ from the joint moments $\mathbf{y}$. It uses the language of cycles on trees. Then, based on the false freeness property, we guess what could become this formula when the free semicircular variables are replaced by heavy Wigner matrices.

Assume that the semicircular variables are standard. Let $L \geqslant 1$ be an integer, $\ell = (\ell_1, \dots, \ell_L)$ in $\{1, \dots, p\}^L$ be a sequence of colors and $Q_1, \dots, Q_L$ be monomials in $\mathbb{C}\langle \mathbf{y} \rangle$. The knowledge of $\tau[x_{\ell_1}Q_1 \dots x_{\ell_L}Q_L]$ for any such monomials completely determines τ by linearity and traciality.

Let (T, c) be in the set $\mathcal{L}_{SC}^{(\ell)}$ introduced in Section 3.2.4. The cycle c induces a partition π_c of $\{1, \dots, L\}$: two integers n and m are in the same block of π_c if and only if the n-th and the m-th steps of c reach the same vertex of T. We write $\pi = \{B_v\}_{v \text{ vertex of } T}$ and $B_v = \{j_{v,1}, \dots, j_{v,r_v}\}$ where $j_{v,1} < \dots < j_{v,r_v}$. Then, one has

$$\tau[x_{\ell_1}Q_1 \dots x_{\ell_L}Q_L] = \sum_{(T,c)\in\mathcal{L}_{SC}^{(\ell)}} \prod_{v \text{ vertex of } T} \tau[Q_{j_{v,1}} \dots Q_{j_{v,r_v}}]. \quad (3.8)$$

To infer an analogue of Equation (3.8) for heavy semicircular variables, we try to find a Formula which satisfies the false freeness property. Let (T, c) be in $\mathcal{L}_{SC}^{(\ell)}$. Recall that the n-th step of c has color ℓ_n, which means that it corresponds to the variable x_{ℓ_n} in the word $x_{\ell_1} Q_1 \dots x_{\ell_L} Q_L$. We magnify the tree T into a graph G and double the number of steps of c in order to include steps corresponding to the variables $Q_1, \dots, Q_L$: in Figure 3.2, we start with the tree at the left and we get the graph on the middle (we give a precise definition of this trick latter). Schematically, the vertices of T are transformed into cycles whose edges can be labelled by the monomials $Q_1, \dots, Q_L$ in the same way we have colored the steps of the original cycle. The cycles that replace the vertices of the original tree gives the traces in Equation (3.8).

Now, we apply the folding trick of the false freeness property to one edge of the graph which comes from the original branch of the tree: see the graph at the right in Figure 3.2. It turns out that the cycle at the source and the goal of this branch are folded into graphs which are no longer cycles. To show the convergence of the distribution of heavy Wigner and arbitrary matrices we give a sense to the trace in such graphs, labelled by monomials. This is the purpose of the next section.

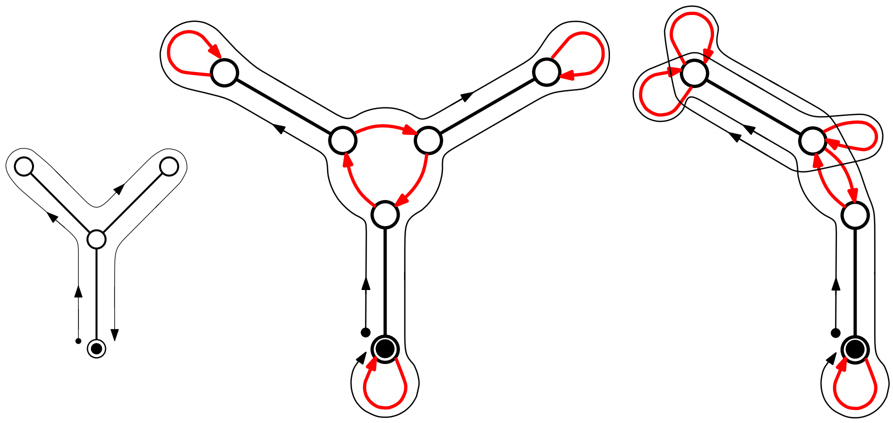

Figure 3.2: Left: a cycle on a tree (with only one color). Middle: each vertex has been replaced by cycles. Right: The upper rightmost edge is folded into the left one.

3.3 The convergence in distribution of traffics of heavy Wigner and deterministic matrices

3.3.1 Distribution of traffics

Definition and examples

Our setup is the following.

Definition 3.3.1 (Test graphs)**.**

1. A test graph is a finite, connected, directed graph whose edges are labelled by indeterminates. More formally, test graph in p variables (or indeterminates) is a triplet $T = (V, E, \gamma)$ where
   – (V, E) is a finite connected directed graph with possible multiple edges: V is its set of vertices, E is its set of edges, multi-set of couples of vertices.
   – γ is a map $E \to \{1, \dots, p\}$ which indicates the indeterminates corresponding to each edge.

   We sometimes denote a test graph $T = (G, \gamma)$ instead of $T = (V, E, \gamma)$ when $G = (V, E)$ is a finite directed graph.

2. A test graph $T = (G, \gamma)$ is said to be cyclic whenever we can cover G by a cycle that visits exactly one time each edge in the sense of its orientation.

3. The set of all test graphs in p variables is denote by $\mathcal{G}\langle x_1, \dots, x_p\rangle$, where the symbols $x_1, \dots, x_p$ refers to the indeterminates. The set of all cyclic test graphs in p variables is denote by $\mathcal{G}_{cyc}\langle x_1, \dots, x_p\rangle$.

We define an analogue of the normalized trace for test graphs as for polynomials in matrices.

Definition 3.3.2 (The distribution of traffics of matrices)**.**

1. Given a test graph $T = (V, E, \gamma)$ in $\mathcal{T}\langle x_1, \dots, x_p\rangle$ and a family $\mathbf{A}_N = (A_1^{(N)}, \dots, A_p^{(N)})$ of $N \times N$ matrices, we define the trace of T in $\mathbf{A}_N$ by

$$\tau_N\Big[T(\mathbf{A}_N)\Big] = \frac{1}{N} \sum_{\phi: V \to \{1,\dots,N\}} \prod_{e \in E} A^{(N)}_{\gamma(e)}\big(\phi(e)\big),$$

   where for any directed edge $e = (v_1, v_2)$, we have set $\phi(e) = (\phi(v_1), \phi(v_2))$ and for any matrix M and any integers n, m, the number $M(n, m)$ denotes the entry (n, m) of M.

2. The distribution of traffics of a family $\mathbf{A}_N = (A_1^{(N)}, \dots, A_p^{(N)})$ of $N \times N$ matrices is the map

$$\begin{array}{rccc} \tau_N: & \mathcal{G}_{cyc}\langle x_1, \dots, x_p, x_1^*, \dots, x_p^*\rangle & \to & \mathbb{C} \\ & T & \mapsto & \tau_N\Big[T(\mathbf{A}_N, \mathbf{A}_N^*)\Big]. \end{array}$$

3. Let $\mathbf{A}_N$ be a family of p matrices of size N by N. We say that $\mathbf{A}_N$ has a limiting distribution of traffics τ whenever, for any cyclic test graph T,

$$\tau[T] := \lim_{N \to \infty} \tau_N\Big[T(\mathbf{A}_N, \mathbf{A}_N^*)\Big] \text{ exists.}$$

The two following examples give clues about the amount of information which is contained in a limiting distribution of traffic τ.

Examples

1. Let $P = x_{\ell_1} \dots x_{\ell_L}$ be a monic monomial in p indeterminates $x_1, \dots, x_p$, where for any $n = 1, \dots, L$ one has ℓ_n in $\{1, \dots, p\}$. Let G be the graph whose vertices are $1, 2, \dots, L$ and edges are $(1,2), \dots, (L-1, L), (L, 1)$. We set $\gamma\Big((i, i+1)\Big) = \ell_i$ (with the convention $(p, p+1) = (p, 1)$) and we consider the test graph $T_P = (G, \gamma)$ in p variables. Then, for any $\mathbf{A}_N = (A_1^{(N)}, \dots, A_p^{(N)})$ in $\mathrm{M}_N(\mathbb{C})^p$, one has
$$\tau_N\Big[T_P(\mathbf{A}_N)\Big] = \tau_N\Big[P(\mathbf{A}_N)\Big].$$
In the left hand side, τ_N denotes the trace of test graphs in N by N matrices whereas in the right hand side it denotes the usual normalized trace of matrices. Hence, the distribution of traffics of a family $(\mathbf{A}_N, \mathbf{A}_N^*)$ of N by N matrices contains the information about the joint distribution of $\mathbf{A}_N$ in the sense of *-probability space.

2. Let A_N be an adjacency matrix of a directed graph G_N with N vertices and no multiple edges. Informally, for any test graph $T = (G, \gamma)$ in one variables, the number $N \times \tau_N\Big[T(A_N)\Big]$ is the number of times the graph G appears as a subgraph of G_N. This fact is made clear and exploited in Section 3.4 to show that the convergence in distribution of traffics generalizes also the weak local convergence of graphs.

Let $\mathbb{C}_0\langle x_1, \dots, x_p, x_1^*, \dots, x_p^*\rangle$ be the set of test graphs we can obtain with monic monomials as in the first example. Let τ be a map $\mathcal{G}_{cyc}\langle x_1, \dots, x_p\rangle \to \mathbb{C}$. Then, τ can be restricted on $\mathbb{C}_0\langle x_1, \dots, x_p, x_1^*, \dots, x_p^*\rangle$, and then extended by linearity on $\mathbb{C}\langle x_1, \dots, x_p, x_1^*, \dots, x_p^*\rangle$, with the convention $\tau[1] = 1$ (this maps is still denoted by τ). Then τ is always tracial and is called the trace induced.

Definition 3.3.3 (Distribution of traffics).
A distribution of traffic in p variables is a map $\tau : \mathcal{G}_{cyc}\langle x_1, \dots, x_p\rangle \to \mathbb{C}$ such that the trace induced in $\mathbb{C}\langle x_1, \dots, x_p, x_1^*, \dots, x_p^*\rangle$ by τ is a state, i.e. $\tau[PP^*] \geqslant 0$ for any polynomial P. By convergence in distribution of traffics we means the pointwise convergence of these maps.

The injective trace

The definition of the injective trace is natural both in the context of random matrices and for the analysis of random graphs. For matrices, its definition is the following.

Definition 3.3.4 (Injective trace for matrices).
Let $T = (V, E, \gamma)$ be a test graph in p variables and $\mathbf{A}_N = (A_1^{(N)}, \dots, A_p^{(N)})$ be a family of N by N matrices. We define the injective trace of T in $\mathbf{A}_N$ by
$$\tau_N^0\Big[T(\mathbf{A}_N)\Big] = \frac{1}{N} \sum_{\substack{\phi: V \to \{1, \dots, N\} \\ \text{injective}}} \prod_{e \in E} A_{\gamma(e)}^{(N)}\Big(\phi(e)\Big).$$

The knowledge of the distribution of traffics of a family $\mathbf{A}_N$ of N by N matrices is equivalent to the knowledge of the injective trace of cyclic test graphs in $\mathbf{A}_N$. To see this fact, we need the following definition.

Given a test graph $T = (V, E, \gamma)$ and a partition π of V, we define a new test graph $\pi(T) = \big(\pi(V), \pi(E), \pi(\gamma)\big)$, where we have identified the vertices that belong to a same block. The set of vertices $\pi(V)$ are the blocks of π. If V is the multiset $\big\{\!\!\big\{(v_1, v_2), \dots, (v_{2K-1}, v_{2K})\big\}\!\!\big\}$, then $\pi(V)$ is the multiset $\big\{\!\!\big\{(\pi(v_1), \pi(v_2)), \dots, (\pi(v_{2K-1}), \pi(v_{2K}))\big\}\!\!\big\}$, where for any v in V, $\pi(v)$ denotes the block of π containing v. For any $e = (\pi(v_{2k-1}, \pi(v_{2k}))$ in $\pi(V)$, we set $\pi(\gamma)(e) = \gamma(v_{2k-1}, v_{2k})$.

Lemma 3.3.5 (Injective trace vs. non-injective trace)**.**
Let $T = (V, E, \gamma)$ be a test graph in p variables. Then, for any p-tuple $\mathbf{A}_N$ of $N \times N$ matrices, one has

$$\tau_N\Big[T(\mathbf{A}_N)\Big] = \sum_{\sigma \in \mathcal{P}(V)} \tau_N^0\Big[\sigma(T)(\mathbf{A}_N)\Big], \tag{3.9}$$

where $P(V)$ is the set of partitions of V. Hence, one has

$$\tau_N^0\Big[T(\mathbf{A}_N)\Big] = \sum_{\sigma \in \mathcal{P}(V)} \tau_N\Big[\sigma(T)(\mathbf{A}_N)\Big] \times \mu_V(\sigma), \tag{3.10}$$

where μ_V is the Möbius function of the finite poset $\mathcal{P}(V)$ (see [NS06]).

This proposition motivates the following definition for general distributions of traffics.

Definition 3.3.6 (Injective trace)**.**
Let $\tau : \mathcal{G}_{cyc}\langle x_1, \dots, x_p\rangle \to \mathbb{C}$ be a distribution of traffics in p variables. The injective version of τ is the functional $\tau^0 : \mathcal{G}_{cyc}\langle x_1, \dots, x_p\rangle \to \mathbb{C}$ defined by: for any test graph T in p variables,

$$\tau^0\Big[T\Big] \ = \sum_{\sigma \in \mathcal{P}(V)} \tau\Big[\sigma(T)\Big] \times \mu_V(\sigma), \tag{3.11}$$

where μ_V is as in Proposition 3.3.5. Hence, we have

$$\tau\Big[T\Big] \ = \sum_{\sigma \in \mathcal{P}(V)} \tau^0\Big[\sigma(T)\Big]. \tag{3.12}$$

Evaluating graph test on monomials

Let $T = (V, E, \gamma)$ be a test graph in p indeterminates and $Q_1, \dots, Q_p$ be monic monomials in q non commutative indeterminates. We define the test graph $T(Q_1, \dots, Q_p)$ in q indeterminates by replacing each edge e of T by a chain corresponding to $Q_{\gamma(e)}$.

More precisely, for any directed edge e in E we apply the following trick. We write $Q_{\gamma(e)} = x_{\ell_1} \dots x_{\ell_L}$ as a product of indeterminates and denote $e = (v_0, v_L)$, where v_0 in V is the starting vertex of e and v_L in V is its end. If $L = 0$ (i.e. $Q_\gamma = 1$) then we simply merge the vertices v_0 and v_L and forget the edge e. Otherwise, we introduce new vertices $v_1, \dots, v_{L-1}$ and denote for any $i = 1, \dots, L$ by e_i the directed edge (v_{i-1}, v_i). Then, we replace e by the path $e_1 \circ \dots \circ e_L$. At last, we define $\gamma(e_i) = \ell_i$ for any $i = 1, \dots, L$. This defines the test graph $T(Q_1, \dots, Q_p)$ which is connected as soon as T is connected and cyclic as soon as T is cyclic.

3.3.2 The convergence of heavy Wigner and deterministic matrices

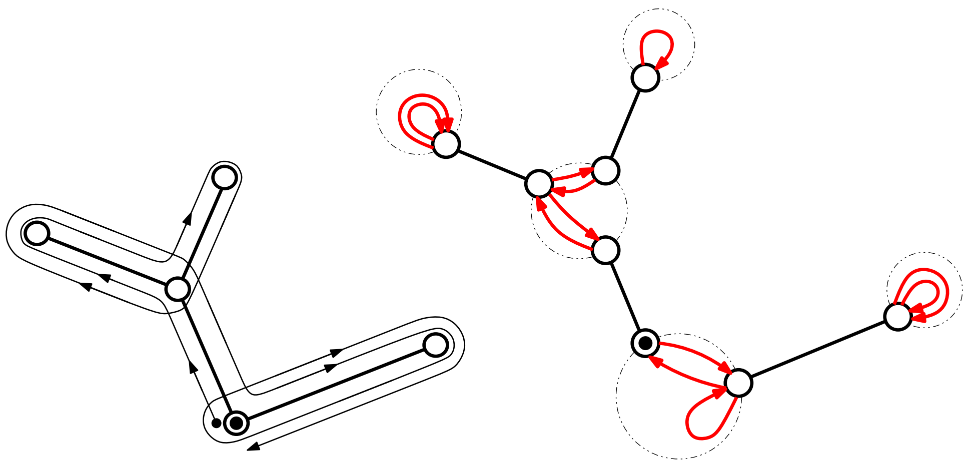

Figure 3.3: Left: a cycle on a tree (with only one color). Right: each vertex has been replaced by the graph of the associated test graph

Definition 3.3.7 (Test graphs associated to a cycle)**.** Let $L \geqslant 1$ be an integer and $\ell = (\ell_1, \dots, \ell_L)$ be a sequence of colors. Let (T, c) be in $\mathcal{L}^{(\ell)}$ and write $c = e_1 \circ \dots \circ e_L$, as a composition of directed edges of T. For any vertex v of T, we associate a test graph $T_{v,c} = (G_{v,c}, \gamma_{v,c})$. The vertices of $G_{v,c}$ are the incident edges of T in v. If the n-th step of c is incident at v, then we get an edge $\tilde{e}$ between the undirected edges corresponding to e_n and e_{n+1} (with the convention $e_{L+1} = e_1$). We set $\gamma_{v,c}(\tilde{e}) = n$.

In Figure 3.3, we have drawn a construction of test graphs from cycles on a tree. We can now state our main result.

Theorem 3.3.8 (The convergence of heavy Wigner and deterministic matrices)**.** Let $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ be a family of $N \times N$ independent heavy Wigner matrices. Let $\mathbf{Y}_N = (Y_1, \dots, Y_q)$ be $N \times N$ deterministic matrices. Let $\mathbf{x} = (x_1, \dots, x_p)$ and $\mathbf{y} = (y_1, \dots, y_q)$ be families of non commutative indeterminates. Assume that,

1 Convergence in distribution of traffics: For any cyclic test graph T in q variables, one has

$$\tau\Big[T\Big] := \lim_{N\to\infty} \tau_N\Big[T(\mathbf{Y}_N)\Big] \text{ exists.} \tag{3.13}$$

2 Control of traces of connected test graphs: For any connected test graph T in q variables and any $\beta > 0$, one has

$$\tau_N\Big[T(\mathbf{Y}_N)\Big] = o(N^\beta). \tag{3.14}$$

Then, $(\mathbf{X}_N, \mathbf{Y}_N)$ has a limiting distribution, i.e. for any polynomial P in $\mathbb{C}\langle \mathbf{x}, \mathbf{y}\rangle$,

$$\tau[P] := \lim_{N\to\infty} \mathbb{E}\Big[\tau_N\Big[P(\mathbf{X}_N, \mathbf{Y}_N)\Big]\Big] \text{ exists.}$$

More precisely, let $L \geqslant 1$ be an integer, $\ell = (\ell_1, \dots, \ell_L)$ in $\{1, \dots, p\}^L$ be a sequence of colors and $Q_1, \dots, Q_L$ be monic monomials in $\mathbb{C}\langle \mathbf{y}\rangle$. Then, one has

$$\tau[x_{\ell_1} Q_1 \dots x_{\ell_L} Q_L] = \sum_{(T,c)\in\mathcal{L}^{(\ell)}} \omega^{(1)}(c) \times \omega^{(2)}(c), \tag{3.15}$$

where we have denoted

$$\begin{aligned} \omega^{(1)}(c) &= \prod_{e \text{ edge of } T} a_{\eta(e),n(e)}, \\ \omega^{(2)}(c) &= \prod_{v \text{ vertex of } T} \tau\Big[T_{v,c}(Q_1 \dots Q_L)\Big]. \end{aligned}$$

More generally, $(\mathbf{X}_N, \mathbf{Y}_N)$ has a limiting distribution of traffics (see Theorem 3.7.1).

This theorem is proved in Section 3.7.

3.4 The distribution of traffics of a random graph

It is natural to wonder what means the convergence in distribution of traffics for matrices that are an adjacency matrix of a graph. It turns out that this convergence is equivalent to an other type of convergence for graphs, called the weak local convergence.

3.4.1 Distribution of traffics of finite graphs

Definition 3.4.1 (The distribution of traffics of a graph)**.**
Let G_N be a undirected graph (without multiple edges) with N vertices. We arbitrary label its vertices by the integers $\{1, \dots, N\}$. The adjacency matrix $A_N = \big(A_N(i,j)\big)_{i,j=1,\dots,N}$ of G_N associated to this labeling is the N by N symmetric matrix given by: for any $i, j = 1, \dots, N$, $A_N(i,j)$ is one if $\{i,j\}$ is an edge of G_N and is zero otherwise. Given a test graph T in one variable, we set

$\tau_N\Big[T(G_N)\Big] := \tau_N\Big[T(A_N)\Big]$, which does not depend on the choice of the labeling for the vertices of G. The distribution of traffics of G_N is the map

$$\begin{array}{cccc} \tau_N : & \mathcal{G}_{cyc}\langle x\rangle & \to & \mathbb{C} \\ & T & \mapsto & \tau_N\Big[T(\mathbf{G}_N)\Big] \end{array}$$

We only consider test graphs $T = (G, \gamma)$ in one variable in this section, a case where the map γ is trivial. With a slight abuse, we will use the symbol T to mean its graph G.

The language of traffics is relevant to describe the statistical geometry of a graph. Let G_N be a finite directed graph with N vertices and T be a test graph in one variable. Then, the number $N \times \tau_N^0\Big[T(G_N)\Big]$ is the number of way we can embed T into G_N.

3.4.2 Stationary random rooted graphs and their distribution of traffics

We denote by $\mathcal{G}$ the set of all undirected graphs (without multiple edges) whose degree of each vertex is finite (up to isomorphism of graphs). We denote by $\mathcal{G}_*$ the set of couples (G, v) where G is in $\mathcal{G}$ and v is a vertex of G, called its root. For any integer $p \geqslant 0$, we define $\mathcal{G}_{*,p}$ as the set of all connected, rooted graphs in $\mathcal{G}_*$ whose vertices are at distance at most p of the root. For any integer $p \geqslant 0$ and any (G, v) in $\mathcal{G}_*$, we denote by $(G, v)_p$ in $\mathcal{G}_{*,p}$ the connected sub-graph of G, rooted at v, constituted by the vertices of G that are at distance at most p of v and by the edges linking these vertices.

We do not give the exact definition of random elements in $\mathcal{G}_*$ (see [AS04]) and give only the definition of its distribution.

Definition 3.4.2 (random rooted graphs)**.**
The law of a random rooted graph (G, v) in $\mathcal{G}_*$ is the knowledge of $\mathbb{P}\Big((G, v)_p = (H, w)\Big)$ for any integer $p \geqslant 0$ and any connected rooted graph (H, w) in $\mathcal{G}_{*,p}$ (the equality of rooted graphs is up to isomorphism).

Let (G, v) and (T, r) be two rooted graphs in $\mathcal{G}_*$, where (T, r) is connected, finite and deterministic and (G, v) is random (think (T, r) as a test graph in one variable with an arbitrary chosen vertex). We denote by $\tau^0\Big[(T, r)(G, v)\Big]$ the expectation of the number of embeddings of T into G such that r is sent to v. The random rooted graph (G, v) is said to be stationary whenever for any T finite in $\mathcal{G}$ and any r vertex of T, the quantity $\tau^0\Big[(T, r)(G, v)\Big]$ does not depend on r.

Definition 3.4.3 (Distribution of traffics of stationary random rooted graphs)**.**
Let (G, v) be a stationary random rooted graph. The distribution of traffics of (G, v) is the map

$$\begin{array}{rccc} \tau: & \mathcal{G}_{cyc}\langle x\rangle & \to & \mathbb{C} \\ & T & \mapsto & \tau\Big[T(G,v)\Big], \end{array}$$

whose injective version τ^0 is given by: for any test graph T in one variable, $\tau^0\Big[T(G,v)\Big]$ is the common value of $\tau^0\Big[(T,r)(G,v)\Big]$ for any choice of vertex r of T.

We claim that the distribution of traffics characterizes the law of a stationary random rooted graph. This fact comes easily from the following proposition.

Proposition 3.4.4 (the relation between the law and the distribution of traffics of a stationary random rooted graph)**.**
Let (G,v) be a random rooted graph in $\mathcal{G}_*$. Then, for any integer $p \geqslant 1$ and any (T,r) in $G_{*,p}$, one has

$$\begin{aligned} &\tau^0\Big[(T,r)(G,v)\Big] \\ &\quad = \sum_{\substack{(H,w)\in\mathcal{G}_{*,p} \\ (H,w)\geqslant(T,r)}} \tau^0\Big[(T,r)(H,w)\Big] \times \mathbb{P}\Big((G,v)_p = (H,w)\Big). \end{aligned} \quad (3.16)$$

The symbol $(H,w) \geqslant (T,r)$ means that T is a subgraph of H up to an isomorphism which sends r to w. Hence, we get

$$\begin{aligned} &\tau^0\Big[(T,r)(T,r)\Big] \times \mathbb{P}\Big((G,v)_p = (T,r)\Big) \\ &\quad = \sum_{\substack{(H,w)\in\mathcal{G}_{*,p} \\ (H,w)\geqslant(T,r)}} \tau^0\Big[(H,w)(G,v)\Big] \times \mu_p\Big((H,w),(T,r)\Big), \end{aligned} \quad (3.17)$$

where μ_p is the Möbius map of the poset $\mathcal{G}_{*,p}$ (see [NS06]).

Remark: By Definition 3.4.3, a distribution of traffics τ of a stationary random rooted graph is only defined on cyclic test graphs. Nevertheless, the definition makes sense for T arbitrary. But for any test graph T in one variable, we can add edges to its multi-set of edges in order to obtain a cyclic test graph $\tilde{T}$ such that, for any distribution of traffics τ of a stationary random rooted graph, one has $\tau[T] = \tau[\tilde{T}]$. Hence, the restriction to cyclic test graphs in Definition 3.4.3 is not a real one.

3.4.3 The convergence in distribution of traffics and the weak local convergence

We have two notions of convergence for graphs whose number of vertices goes to the infinity. The first one is the convergence in distribution of traffics introduced above. The second one is the weak local convergence introduced by Benjamini and Schramm [BS01] and developed by Aldous and Steele [AS04].

Definition 3.4.5 (The weak local convergence of finite graphs)**.**
Let $(G_N)_{N\geqslant 1}$ be a sequence of finite graphs in $\mathcal{G}$ and (G,v) a rooted graph in

$\mathcal{G}_*$. We say that G_N converges weakly locally to (G, v) whenever for any integer $p \geqslant 1$ and any rooted graph (H, w) in $\mathcal{G}_{*,p}$, one has

$$\mathbb{P}\Big((G_N, v_N)_p = (H, w)\Big) \underset{N\to\infty}{\longrightarrow} \mathbb{P}\Big((G, v)_p = (H, w)\Big), \qquad (3.18)$$

where the root v_N is chosen uniformly on the vertices of G_N.

Theorem 3.4.6 (The equivalence between weak local convergence and convergence in distribution of traffics)**.**
Let G_N be a graph in $\mathcal{G}$ with N vertices. Then, G_N has a limiting distribution of traffics τ if and only if G_N weakly locally converges to a random rooted graph (G, v). In this case, (G, v) is stationary and τ is the distribution of traffics of (G, v).

Sketch of proof. Let G_N be a finite graph in $\mathcal{G}$ with N vertices and v_N be a random vertex of G_N chosen uniformly. It is easy to see that (G_N, v_N) is stationary and the distribution of traffics of G_N is the distribution of traffics of the random rooted graph (G_N, v_N). By an easy application of Proposition 3.4.4, the only non trivial thing we have to show is that if a sequence of finite graphs has a limiting distribution of traffics τ, then τ is the distribution of a random rooted graph.

Let τ be a limiting distribution of a sequence $(G_N)_{N\geqslant 1}$ of graphs in $\mathcal{G}$, where for any $N \geqslant 1$, G_N has N vertices. For any $p \geqslant 1$ and any (T, r) in $\mathcal{G}_{*,p}$, we set

$$\begin{aligned}
\eta_{p,N}(T, r) &= \frac{1}{\tau^0\Big[(T, r)(T, r)\Big]} \sum_{\substack{(H,w)\in\mathcal{G}_{*,p} \\ (H,w)\geqslant(T,r)}} \tau_N^0\Big[H(G_N)\Big]\mu_p\Big((H, w), (T, r)\Big), \\
\eta_p(T, r) &= \frac{1}{\tau^0\Big[(T, r)(T, r)\Big]} \sum_{\substack{(H,w)\in\mathcal{G}_{*,p} \\ (H,w)\geqslant(T,r)}} \tau^0\Big[H\Big]\mu_p\Big((H, w), (T, r)\Big),
\end{aligned}$$

where τ^0 is the injective version of τ. Then, since $\eta_p(T, r)$ is the limit of $\eta_{p,N}(T, r)$, it belongs to $[0, 1]$ and

$$\sum_{(T,r)\in\mathcal{G}_{*,p}} \eta_p(T, r) = 1.$$

Hence, the collection of numbers $\eta_p(T, r)$ well defines a random rooted graph (G, v) which is necessarily stationary. □

3.5 Distribution of traffics and free probability

3.5.1 A false free product construction

Theorem 3.3.8 motivates the following definition.

Definition 3.5.1 (Distribution of traffics in *-probability space)**.**
Let $(\mathcal{A}, \tau)$ be a *-probability space. We say that a family of non commutative random variable $\mathbf{a} = (a_1, \dots, a_q)$ has a distribution of traffics whenever we have specified a map

$\mathcal{G}_{cyc}\langle y_1, \dots, y_q \rangle \to \mathbb{C}$, still denoted by τ, such that for any monic monomial P one has $\tau\Big[T_P(\mathbf{a})\Big] = \tau\Big[P(\mathbf{a})\Big]$, where T_P is the test graph defined in the example 1 of Section 3.3.1.

In Definition 3.2.3, we have given the definition of heavy semicircular variables $x_1, \dots, x_p$ in a non commutative probability space $(\mathcal{A}, \tau)$. This definition is based on Theorem 3.2.2, where is computed the limiting distribution of heavy Wigner matrices. In Theorem 3.3.8, we have generalized Theorem 3.2.2 and state the convergence of the distribution of traffics of heavy Wigner matrices (the precise statement of this convergence is given in Theorem 3.7.1). We then refine the definition of heavy semicircular variables by specifying the distribution of traffics τ for $x_1, \dots, x_p$ given by Theorem 3.7.1.

Let $\mathbf{y} = (y_1, \dots, y_q)$ be a family of non commutative random variable having a distribution of traffics. Then, Formula (3.15) gives a canonical way to consider in a same non commutative probability space a family of heavy semicircular variable together with the family $\mathbf{y}$. The procedure of enlarging such a family $\mathbf{y}$ with heavy semicircular variables $\mathbf{x} = (x_1, \dots, x_p)$ by Formula (3.15) is referred as **the false free product construction** and we say that $\mathbf{x}$ and $\mathbf{y}$ are falsely free. In general, the two families are not free as we will see in Section 3.5.4.

This construction is actually a product of algebra. The false free product construction exhibits a product between an algebra $\mathcal{A}_\mathbf{x}$ spanned by heavy semicircular variables and an arbitrary algebra $\mathcal{A}_\mathbf{y}$ whose elements have a distribution of traffics. This fact suggests two interesting problems: finding a canonical construction for the false free product, as the Fock space construction for the usual free product and finding a general free product construction between two arbitrary algebras whose elements have a distribution of traffics. This second question is investigated in a work in preparation.

3.5.2 Diagonal non commutative random variables

Let $(\mathcal{A}, \tau)$ be a non commutative probability space and $\mathcal{A}_0 \subset \mathcal{A}$ a commutative unital subalgebra of $\mathcal{A}$. We can extend τ as a distribution of traffics for the elements of $\mathcal{A}_0$. Let $\mathbf{d} = (d_1, \dots, d_q)$ be a family in $\mathcal{A}_0$. For any cyclic test graph $T = (V, E, \gamma)$ in q variables, we set

$$\tau\Big[T(\mathbf{d})\Big] := \tau\Big[\prod_{e \in E} d_{\gamma(e)}\Big].$$

There is not ambiguity in this formula since the elements commute. When we specify this distribution of traffics for a family of commuting random variables, we will say that this family is diagonal. If $\mathbf{D}_N = (D_1, \dots, D_q)$ is a family of N by N diagonal matrices having a limiting distribution, then $\mathbf{D}_N$ has a limiting distribution of traffics which is diagonal.

When such a family $\mathbf{d}$ is extended with a family of heavy semicircular variables $\mathbf{x} = (x_1, \dots, x_p)$ by the false product construction, the families $\mathbf{x}$ and $\mathbf{d}$ are not free in general as we will see in Section 3.5.4. Moreover, in Section 3.6, we state a Schwinger-Dyson like system of equations for the joint family of $(\mathbf{x}, \mathbf{d})$.

3.5.3 The multilinear forms $\big(\Phi^{(K)}\big)_{K\geqslant 1}$

Let $(\mathcal{A}, \tau)$ be a non commutative probability space whose elements have a distribution of traffics. We introduce a family of multilinear forms on $\mathcal{A}$, that generalizes the usual trace. This family will be useful to shed light on the false freeness property and will play the main role in the Schwinger-Dyson equations stated in Section 3.6.

We first introduce these functionals in the special case of matrix spaces. For any integer $K \geqslant 1$, we set

$$\begin{array}{rccc} \Phi_N^{(K)} & : & \mathrm{M}_N(\mathbb{C})^K & \to & \mathbb{C} \\ & & (A_1, \dots, A_K) & \mapsto & \tau_N[A_1 \circ A_2 \circ \cdots \circ A_K], \end{array}$$

where the symbol $\circ$ designates the entry-wise matrix multiplication, known as the Hadamard product. Remark that in particular one has $\Phi_N^{(1)} = \tau_N$. These multilinear maps can be written as traces of certain test graphs. Let G be the graph with a single vertex and with K edges, $e_1, \dots, e_K$, linking the vertex to itself. We set $\gamma(e_i) = i$ for any $i = 1, \dots, K$ and consider the test graph in K variables $T^{(K)} = (G, \gamma)$. Then, for any $\mathbf{A}_N = (A_1^{(N)}, \dots, A_K^{(N)})$ in $\mathrm{M}_N(\mathbb{C})^p$, one has

$$\tau_N\Big[T^{(K)}(\mathbf{A}_N)\Big] = \Phi_N^{(K)}(A_1, \dots, A_K).$$

Now, let $\mathbf{a} = (a_1, \dots, a_p)$ be non commutative random variables having a distribution of traffics τ. For any integer $K \geqslant 1$, we define a K-linear form on $\mathbb{C}\langle \mathbf{x} \rangle$ by setting, for any monomials $P_1, \dots, P_K$ in $\mathbb{C}\langle \mathbf{x} \rangle$,

$$\Phi^{(K)}(P_1, \dots, P_K) := \tau\Big[T^{(K)}\Big(P_1(\mathbf{a}), \dots, P_K(\mathbf{a})\Big)\Big].$$

For heavy semicircular variables, we have a formula to compute $\Phi^{(K)}$ in terms of cycles visiting a tree which is very closed to the formula for the trace.

Definition 3.5.2 (Chain of cycles coloring a tree)**.** Let $K \geqslant 1$ be an integer, $\mathbf{L} = (L_1, \dots, L_K)$ be a family of non negative integers and ℓ in $\{1, \dots, p\}^L$ be a sequence of colors, where $L = L_1 + \cdots + L_K$. We denote by $\mathcal{L}_{\mathbf{L}}^{\ell}$ the set of couples (T, c) in $\mathcal{L}^{(\ell)}$ such that c is the composition of K cycles, $c = c_1 \circ \cdots \circ c_K$, where for any $k = 1, \dots, K$, the cycle c_k is of length L_k.

Theorem 3.5.3 (The multilinear forms $\big(\Phi^{(K)}\big)_{K\geqslant 1}$ in heavy semicircular variables)**.**
Let $\mathbf{x} = (x_1, \dots, x_p)$ be a family of heavy semicircular variables falsely free from variables $\mathbf{y} = (y_1, \dots, y_q)$. For any $m = 1, \dots, p$, we set $(a_{m,k})_{k\geqslant 1}$ the parameter

of the variable x_m. Let $K \geqslant 1$ be an integer, $\mathbf{L} = (L_1, \dots, L_K)$ a sequence of positive integers. For any $k = 1, \dots, K$, let $\ell^{(k)} = (\ell_{k,1}, \dots, \ell_{k,L_k})$ in $\{1, \dots, p\}^{L_k}$. We set the sequence of colors $\ell = (\ell_{1,1}, \dots, \ell_{1,L_1}, \dots, \ell_{K,1}, \dots, \ell_{K,L_K})$ in $\{1, \dots, p\}^L$, where $L = L_1 + \dots + L_K$. For any $k = 1, \dots, K$, let $Q_{k,0}, \dots, Q_{k,L_k}$ be monic monomials in the variables $y_1, \dots, y_q$. We set for any $k = 1, \dots, K$ the monic monomial in $\mathbb{C}\langle \mathbf{x}, \mathbf{y} \rangle$

$$P_j = Q_{k,0} x_{\ell_{k,1}} Q_{k,1} \dots x_{\ell_{k,L_k}} Q_{k,L_k}.$$

Then, one has

$$\Phi^{(K)}\Big(P_1, \dots, P_K\Big) = \sum_{(T,c) \in \mathcal{L}_{\mathbf{L}}^{(\ell)}} \omega^{(1)}(c) \omega^{(2)}(c), \tag{3.19}$$

where the weights are as in Theorem 3.3.8, i.e.

$$\begin{aligned} \omega^{(1)}(c) &= \prod_{e \text{ edge of } T} a_{\gamma(e), n(e)}, \\ \omega^{(2)}(c) &= \prod_{v \text{ vertex of } T} \tau\Big[T_{v,c}(Q_{j_{v,1}} \dots Q_{j_{v,r_v}}) \Big]. \end{aligned}$$

This theorem is proved as a corollary of Theorem 3.7.1 in Section 3.7.3. Informally, the examples below suggest that the multilinear forms $\Phi^{(K)}$ give a measurement of the "diagonality" of non commutative random variables.

Examples:

1. If $x_1, \dots, x_p$ are free semicircular variables (with their canonical distribution of traffics when seen as heavy semicircular variables), then one has: for any $K \geqslant 1$ and any polynomial $P_1, \dots, P_K$ in $\mathbb{C}\langle x_1, \dots, x_p \rangle$,

$$\Phi^{(K)}(P_1, \dots, P_K) = \tau[P_1] \dots \tau[P_K].$$

2. In contrast, if $y_1, \dots, y_q$ are diagonal non commutative random variables, then one has: for any $K \geqslant 1$ and any polynomial $P_1, \dots, P_K$ in $\mathbb{C}\langle \mathbf{y} \rangle$,

$$\Phi^{(K)}(P_1, \dots, P_K) = \tau[P_1 \dots P_K].$$

3.5.4 The false freeness property revisited

A false freeness property still holds for the joint distribution of a family of heavy semicircular variables falsely free from an arbitrary family. As in Proposition 3.2.5, it is simply based on the fact that the sum in Formula (3.15) of Theorem 3.3.8 is over elements of $\mathcal{L}^{(\ell)}$. Applying this idea, we get the following.

Lemma 3.5.4 (Application of the false freeness property revisited)**.**
Let x be a heavy semicircular variable falsely free from a family of variables $\mathbf{y}$. Denote by $(a_k)_{k \geqslant 1}$ the parameter of x. Assume it is not trivial and set

$k_0 = \min\{k \geqslant 2 | a_k \neq 0\}$. Then for any integers $L \geqslant 2$, any $n_1, \dots, n_L \geqslant 1$ such that $n_1 + \dots + n_L = k_0$, and any $m_1, \dots, m_L$ monomials in $\mathbf{y}$, one has

$$\tau\Big[\Big(x^{2n_1} - \tau[x^{2n_1}]\Big)\Big(m_1 - \tau[m_1]\Big)\dots\Big(x^{2n_L} - \tau[x^{2n_L}]\Big)\Big(m_L - \tau[m_L]\Big)\Big]$$
$$= a_{k_0}\Phi^{(L)}\Big(m_1 - \tau[m_1], \dots, m_L - \tau[m_L]\Big). \tag{3.20}$$

This lemma is proved in Section 3.9.2. Let x be a heavy semicircular variable with non trivial parameter, falsely free from a diagonal variable y. Then, $\Phi^{(2)}\Big(y - \tau[y], y - \tau[y]\Big) = \tau\Big[(y - \tau[y])^2\Big]$ and so x and y are not free as soon as y has a non trivial variance.

3.6 A Schwinger-Dyson system of equations for the distribution of heavy semicircular and diagonal variables

In the classical case of semicircular variables, the Schwinger Dyson equation is useful since it provides a bridge between the combinatorial and the analytical point of view. We first recall its statement in the following section and then give an analogue for heavy semicircular variables.

3.6.1 The Schwinger-Dyson equation for semicircular variables

Proposition 3.6.1 (Schwinger-Dyson equation for semicircular variables)**.**
Let $\mathbf{x} = (x_1, \dots, x_p)$ and $\mathbf{y} = (y_1, \dots, y_q)$ be two families of elements in a *-probability space $(\mathcal{A}, .^*, \tau)$. Assume that the variables $x_1, \dots, x_p$ are selfadjoint and standard, i.e. for any $m = 1, \dots, p$ one has $x_m = x_m^*$, $\tau[x_m] = 0, \tau[x_m^2] = 1$. Then, the two following statements are equivalent.

- $x_1, \dots, x_p$ are semicircular variables, and $(x_1), \dots, (x_p), \mathbf{y}$ are free
- For any monomial P in $\mathbb{C}\langle \mathbf{x}, \mathbf{y}, \mathbf{y}^* \rangle$ and any $m = 1, \dots, p$, one has

$$\tau\Big[x_m P\Big] = \sum_{P = L x_j R} \tau[L]\tau[R], \tag{3.21}$$

  where the last sum is over all decompositions of the monomial P as a product $L x_m R$.

3.6.2 The case of heavy Wigner matrices

It is natural to look for an analogue of the Schwinger-Dyson equation for heavy semicircular variables. The approach via the cycles coloring trees turns out to be appropriate. Indeed, an easy way to prove Proposition 3.6.1 is to classify the trees of $\mathcal{L}_{SC}^{(\ell)}$ according to the size of the sub-tree descendant of the first edge.

In our case, we classify the elements (T,c) of $\mathcal{L}^{(\ell)}$ according to the number of times c visits the first edge, and then according to the length of induced sub-cycles (see Figure 3.4).

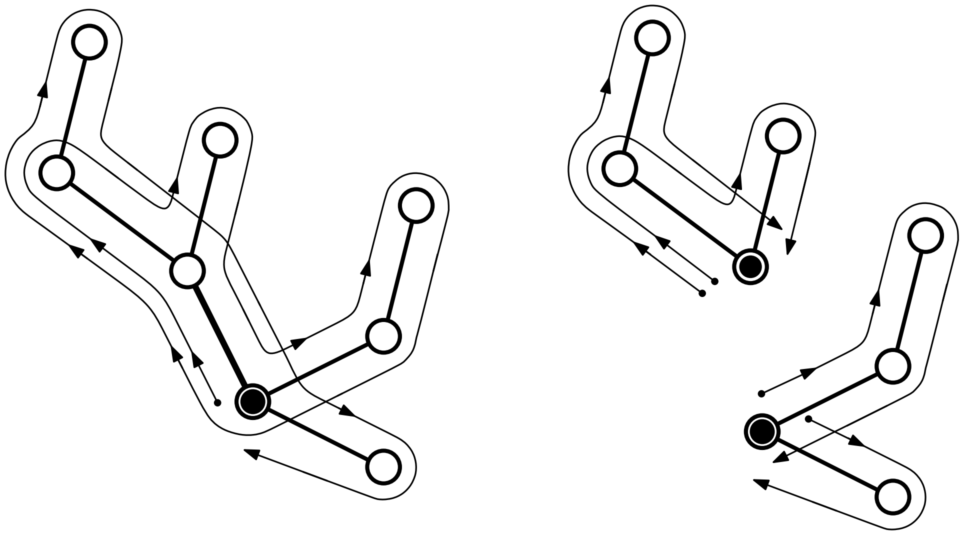

Figure 3.4: Left: the cycle visits 4 times the first vertex. Right: the two couples of cycles induces.

Theorem 3.6.2 (Schwinger-Dyson system of equations)**.**
Let $\mathbf{x} = (x_1, \dots, x_p)$ be a family of heavy semicircular variables falsely free from diagonal variables $\mathbf{y} = (y_1, \dots, y_q)$. For any $m = 1, \dots, p$, we set $(a_{m,k})_{k\geqslant 1}$ the parameter of the variable x_m. Then, the family of linear forms $\big(\Phi^{(K)}\big)_{K\geqslant 1}$ associated to the distribution of $(\mathbf{x}, \mathbf{y})$ satisfies the following equations. For any $m = 1, \dots, p$, denote by $(a_{m,k})_{k\geqslant 1}$ the parameter of x_m. For any monomial P in $\mathbb{C}\langle \mathbf{x}, \mathbf{y}\rangle$ and any $j = 1, \dots, p$, one has

$$\tau[x_j P] = \sum_{k\geqslant 1} a_{j,k} \sum_{x_j P = (x_j L_1 x_j) R_1 \dots (x_j L_k x_j) R_k} \Phi^{(k)}\big(L_1, \dots, L_k\big) \Phi^{(k)}\big(R_1, \dots, R_k\big). \tag{3.22}$$

More generally, for any integer $K \geqslant 1$, any monomials $P_1, \dots, P_K$ in $\mathbb{C}\langle \mathbf{x}, \mathbf{y}\rangle$ and any $j = 1, \dots, p$, one has

$$\begin{aligned} &\Phi^{(K)}(x_j P_1, P_2, \dots, P_K) \\ &\quad = \sum_{k\geqslant 1} a_{j,k} \sum_{\substack{s_1 + \dots + s_K = k \\ s_1 \geqslant 1,\ s_2, \dots, s_K \geqslant 0}} \sum_{\mathbf{L}, \mathbf{R}} \Phi^{(k)}\big(\mathbf{L}\big) \Phi^{(k+K-1)}\big(\mathbf{R}\big), \end{aligned} \tag{3.23}$$

where the last sum is over all the families of monomials

$$\begin{aligned} \mathbf{L} &= (L_1^{(1)}, \dots, L_{s_1}^{(1)}, \dots, L_1^{(K)}, \dots, L_{s_K}^{(K)}), \\ \mathbf{R} &= (R_1^{(1)}, \dots, R_{s_1}^{(1)}, R_0^{(2)}, \dots, R_{s_2}^{(2)}, \dots, R_0^{(K)}, \dots, R_{s_K}^{(K)}), \end{aligned}$$

such that

$$\begin{array}{rcl} x_j P_1 & = & (x_j L_1^{(1)} x_j) R_1^{(1)} \dots (x_j L_{s_1}^{(1)} x_j) R_{s_1}^{(1)} \\ P_2 & = & R_0^{(2)} (x_j L_1^{(2)} x_j) R_1^{(2)} \dots (x_j L_{s_2}^{(2)} x_j) R_{s_2}^{(2)}, \\ \vdots & & \\ P_K & = & R_0^{(K)} (x_j L_1^{(K)} x_j) R_1^{(K)} \dots (x_j L_{s_K}^{(K)} x_j) R_{s_K}^{(K)}. \end{array}$$

This theorem is proved in Section 3.8. Remark that this system of equations characterizes the family $\left(\Phi^{(K)}\right)_{K\geqslant 1}$ among all the families of multilinear forms $\left(\Psi^{(K)}\right)_{K\geqslant 1}$ such that, for any $K \geqslant 1$

- $\Psi^{(K)}$ is a symmetric K-linear form on the set of polynomials in p non commutative indeterminates,
- for any polynomials $P_1, \dots, P_K$, any polynomial Q in $(\mathbf{y}, \mathbf{y}^*)$ and any $i = 1, \dots, K$, one has

$$\begin{array}{rl} & \Psi^{(K)}(P_1, \dots, P_{i-1}, QP_i, P_{i+1}, \dots, P_K) \\ = & \Psi^{(K)}(P_1, \dots, P_{i-1}, P_iQ, P_{i+1}, \dots, P_K) \\ = & \Psi^{(K)}(QP_1, P_2, \dots, P_K). \end{array}$$

- for any polynomials $P_1, \dots, P_K$ in $(\mathbf{y}, \mathbf{y}^*)$, one has

$$\Psi^{(K)}(P_1, \dots, P_K) = \tau[P_1 \times \cdots \times P_K],$$

  where τ is the distribution of $(\mathbf{y}, \mathbf{y}^*)$.

Example of computation: We apply this method to show that

$$\tau[x_1^2 x_2^2 x_1^2 x_2^2] = 3a_{1,1}^2 a_{2,1}^2 + a_{1,1}^2 a_{2,2} + a_{1,2} a_{2,1}^2 + a_{1,2} a_{2,2}$$

in Section 3.9.3.

3.6.3 Application : a characterization of the law of a single heavy semicircular variable

Let x be a heavy semicircular variable of parameter $(a_k)_{k\geqslant 1}$. Let $\Phi^{(K)}$ be the family of multilinear forms on $\mathbb{C}\langle x\rangle = \mathbb{C}[x]$ associated to x. For any $K \geqslant 1$, we set the for formal power series in $\frac{1}{\lambda}$

$$\mu^\lambda(K) := \frac{1}{\lambda^K} \sum_{n\geqslant 0} \frac{1}{\lambda^n} \sum_{\substack{n_1+\dots n_K = n \\ n_1,\dots,n_K\geqslant 1}} \Phi^{(K)}(x^{n_1}, \dots, x^{n_K}).$$

This quantity is simply a formal analogue of

$$\Phi^{(K)}\Big((\lambda - x)^{-1}, \dots, (\lambda - x)^{-1}\Big).$$

Proposition 3.6.3 (A formal characterization of the law of a heavy semicircular variable)**.**
For any $K \geqslant 1$, we have the following equality between formal power series in $\frac{1}{\lambda}$:

$$\lambda \mu^\lambda(K) \quad = \quad \mu^\lambda(K-1) + \sum_{k \geqslant 1} a_k \binom{K+k-2}{K-1} \mu^\lambda(k)\mu^\lambda(k+K-1).$$

These equations characterize the sequence $\big(\mu^\lambda(K)\big)_{K\geqslant 1}$ among the set of formal power series $\big(\nu^\lambda(K)\big)_{K\geqslant 1}$ such that for any $K \geqslant 1$, the valence of $\nu^\lambda(K)$ is larger than K.

This proposition is proved in Section 3.9.4.

Remark : Given an N by N Hermitian matrix X_N and a complex number λ whose imaginary part is positive, then for any $K \geqslant 1$

$$\Phi_N^{(K)}\Big((\lambda - X_N)^{-1}, \dots, (\lambda - X_N)^{-1} \Big) = \frac{1}{N} \sum_{i=1}^N \Big((\lambda - X_N)^{-1} \Big)_{i,i}^K$$

is the moment of order K of the uniform probability measure on the diagonal elements of the resolvant of X_N. This measure is at the center of the analysis in [BAG08], [BDG09] and [KSV04] for other matrix models. Shedding light on this connection could be an interesting problem.

3.7 Proof of Theorem 3.3.8

Let $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ be a family of independent heavy Wigner matrices. For $m = 1, \dots, p$ we denote the $A_m^{(N)} = \sqrt{N} X_m^{(N)}$, which sub diagonal entries are independent, identically distributed according to a measure $p_m^{(N)}$. By assumption, one has for every $k \geqslant 0$

$$a_{m,k} := \lim_{N\to\infty} \frac{\int t^{2k} \mathrm{d}p_m^{(N)}(t)}{N^{k-1}} \text{ exists in } \mathbb{R}, \tag{3.24}$$

$$\sqrt{N} \int t \mathrm{d}p^{(N)}(t) = o(N^\beta), \ \forall \beta > 0. \tag{3.25}$$

Remark that by the Cauchy-Schwarz's inequality, we get that for any $k \geqslant 2$, one has (see Section 3.10.1)

$$\frac{\int t^k \mathrm{d}p_m^{(N)}(t)}{N^{\frac{k}{2}-1}} = O(1). \tag{3.26}$$

Let $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ be a family of deterministic matrices satisfying the assumptions of Theorem 3.3.8.

3.7.1 The injective trace of a cyclic test graph

We consider a cyclic test graph $T = (G, \gamma)$ in $\mathcal{G}_{cyc}\langle z_1, \dots, z_{p+q}\rangle$. By the definition of the injective trace, one has

$$\mathbb{E}\Big[\tau_N^0\big[T(\mathbf{X}_N, \mathbf{Y}_N)\big]\Big] = \mathbb{E}\Big[\frac{1}{N}\sum_{\substack{\phi: V\to\{1,\dots,N\}\\ \text{injective}}} \prod_{e\in E} Z^{(N)}_{\gamma(e)}\big(\phi(e)\big)\Big], \tag{3.27}$$

where

- V is the set of vertices of G, E is its multi-set of edges,
- for any directed edge $e = (v_1, v_2)$, we have set $\phi(e) = \big(\phi(v_1), \phi(v_2)\big)$,
- $Z_i^{(N)} = X_i^{(N)}$ for any $i = 1, \dots, p$ and $Z_{p+i}^{(N)} = Y_i^{(N)}$ for any $i = 1, \dots, q$,
- for any $m = 1, \dots, p+q$, $Z_m^{(N)}(i,j)$ is the (i,j) entry of $Z_m^{(N)}$.

Let $W \subset E$ be the multi-set of edges labelled by an integer in $\{1, \dots, p\}$. We denote, for any injective map $\phi : V \to \{1, \dots, N\}$,

$$P_N^{(1)}(\phi) = \prod_{e\in W} X^{(N)}_{\gamma(e)}\big(\phi(e)\big), \quad P_N^{(2)}(\phi) = \prod_{e\notin W} Y^{(N)}_{\gamma(e)-p}\big(\phi(e)\big),$$

so that one has

$$\mathbb{E}\Big[\tau_N^0\big[T(\mathbf{X}_N, \mathbf{Y}_N)\big]\Big] = \frac{1}{N}\sum_{\substack{\phi: V\to\{1,\dots,N\}\\ \text{injective}}} \mathbb{E}\big[P_N^{(1)}(\phi)\big]P_N^{(2)}(\phi).$$

The contribution of heavy Wigner matrices

We denote by $\bar G = (\bar V, \bar E)$ the undirected graph with no multiple edges obtained from G by forgetting the orientation of the edges and their multiplicity. The cycle c on G induces a cycle $\bar c$ on $\bar G$, written $\bar c = \bar e_1 \circ \cdots \circ \bar e_L$, where $\bar e_1, \dots, \bar e_L$ are directed edges of $\bar G$. For any $n = 1, \dots, L$, the n-th step of $\bar c$ is called a heavy step whenever $\gamma(e_n)$ is in $\{1, \dots, p\}$. In this case, $\gamma(e_n)$ is referred as the color of the n-th step of $\bar c$. In Figure 3.5 we have plotted an example of cyclic test graph $T = (G, \gamma)$ and the graph $\bar G$ induced, equipped with its cycle $\bar c$.

For any $m = 1, \dots, p$ and $k \geqslant 1$ we denote by $\eta_{m,k}$ the number of edges of $\bar G$ that are visited by $\bar c$ exactly k times by a heavy step of color m. Then, for any $\phi : V \to \{1, \dots, p\}$ injective, by the independence of the entries of heavy Wigner matrices one has

$$\mathbb{E}\big[P_N^{(1)}\big] = \prod_{m=1}^{p}\prod_{k\geqslant 1}\left(\frac{\int t^k \mathrm{d}p_m^{(N)}(t)}{N^{\frac{k}{2}}}\right)^{\eta_{m,k}}.$$

We set

$$\begin{aligned} B &:= \sum_{m=1}^{p}\sum_{k\geqslant 1}\eta_{m,k}, \\ \omega_N^{(1)} &:= N^B\,\mathbb{E}\big[P_N^{(1)}\big] = \prod_{m=1}^{p}\prod_{k\geqslant 1}\left(\frac{\int t^k \mathrm{d}p_m^{(N)}(t)}{N^{\frac{k}{2}-1}}\right)^{\eta_{m,k}}. \end{aligned}$$

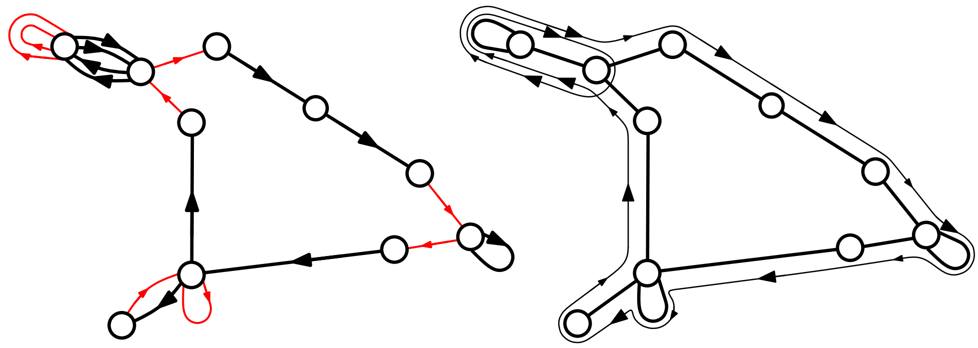

Figure 3.5: Left: a cyclic test graph $T = (G, \gamma)$ in two variables x and y. The first variable corresponds to a heavy Wigner matrix, the second one correspond to a deterministic matrix. The edges labelled by x are plotted in black with a large lines and arrows, the others are plotted in red with smaller lines and arrows. Right: The graph $\bar{G}$ and its cycle $\bar{c}$. The heavy steps of $\bar{c}$ are marked with a larger arrow than its light steps.

Then, one has

$$\mathbb{E}\Big[\tau_N^0\Big[T(\mathbf{X}_N, \mathbf{Y}_N)\Big]\Big] = \frac{\omega_N^{(1)}}{N^{B+1}} \sum_{\substack{\phi:\bar{V}\to\{1,\dots,N\}\\ \text{injective}}} P_N^{(2)}(\phi). \tag{3.28}$$

where

$$\omega_N^{(1)} \underset{N\to\infty}{\longrightarrow} \prod_{m=1}^{p} \prod_{k\geqslant 1} \Big(a_{m,\frac{k}{2}}\Big)^{\eta_{m,k}}$$

as soon as $\eta_{m,k} = 0$ for any $m = 1, \dots, p$ and any k odd, and $\omega_N^{(1)} = o(N^{\beta})$ for any $\beta > 0$ otherwise.

The contribution of deterministic matrices

Recall that for the traffic $T = (G, \gamma) = (V, E, \gamma)$, we have denoted by W the multi-set of edges labelled by an integer in $\{1, \dots, p\}$. We set $G_1, \dots, G_d$ the connected components of the graph $(V, E \setminus W)$. The map γ induces a labeling of the vertices of these components, and then we get test graphs $T_i = (G_i, \gamma_i)$ in $\mathcal{G}\langle y_1, \dots, y_q\rangle$, $i = 1, \dots, d$. In Figure 3.6, we have plotted the test graphs induced by the test graph of Figure 3.5.

We have

$$\frac{1}{N^d} \sum_{\substack{\phi:\bar{V}\to\{1,\dots,N\}\\ \text{injective}}} P_N^{(2)} = \frac{1}{N^d} \sum_{\phi_1,\dots,\phi_d} \prod_{i=1}^{d} \prod_{e\in E_i} Y_{\gamma_i(e)}^{(N)}\big(\phi_1(e)\big), \tag{3.29}$$

where for any $i = 1, \dots, d$ we have denoted $G_i = (V_i, E_i)$ and the first sum is over all injective maps $\phi_1 : V_1 \to \{1, \dots, N\}, \dots, \phi_d : V_d \to \{1, \dots, N\}$, such that the

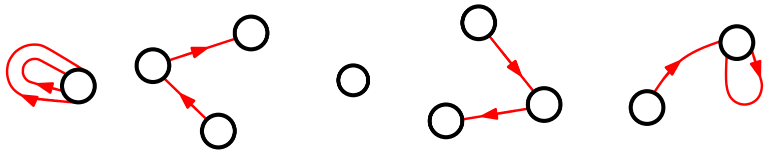

Figure 3.6: The five test graphs in the variable y induced by the test graph of Figure 3.5.

images of $\phi_1, \dots, \phi_d$ are disjoint. If we could drop this last condition, we would obtain $\tau_N^0\big[T_1(\mathbf{Y}_N)\big] \times \dots \times \tau_N^0\big[T_d(\mathbf{Y}_N)\big]$.

But, one has

$$N^d \tau_N^0\big[T_1(\mathbf{Y}_N)\big] \times \dots \times \tau_N^0\big[T_d(\mathbf{Y}_N)\big] = \sum_\pi \sum_{\phi_1,\dots,\phi_d} \prod_{i=1}^d \prod_{e \in E_i} Y_{\gamma_i(e)}^{(N)}\big(\phi_1(e)\big),$$

where

- the first sum is over all partitions π of V whose blocks contain at most one element of each V_i,
- the second sum is over all injective map $\phi_i : V_i \to \{1, \dots, p\}$, such that whenever $v \in V_i$ and $w \in V_j$ belong to a same block of π, then $\phi_i(v) = \phi_j(w)$.

Let π be such a partition which is not the finest one. Then, the term

$$\sum_{\phi_1,\dots,\phi_d} \prod_{i=1}^d \prod_{e \in E_i} Y_{\gamma_i(e)}^{(N)}\big(\phi_1(e)\big)$$

is the product of $\tilde d$ injective traces of graph tests in $\mathbf{Y}_N$ times $N^{\tilde d}$, where $\tilde d$ is strictly smaller than d. By assumption on the deterministic matrices, for any $\beta > 0$ one has

$$\frac{1}{N^d} \sum_{\substack{\phi:\bar V \to \{1,\dots,N\} \\ \text{injective}}} P_N^{(2)} = \omega_N^{(2)} + o\Big(\frac{1}{N^{1-\beta}}\Big), \tag{3.30}$$

where $\omega_N^{(2)} = \tau_N^0\big[T_1(\mathbf{Y}_N)\big] \times \dots \times \tau_N^0\big[T_d(\mathbf{Y}_N)\big]$. Moreover, one has

$$\omega_N^{(2)} \underset{N\to\infty}{\longrightarrow} \tau^0\big[T_1\big] \times \dots \times \tau^0\big[T_d\big]$$

as soon as all the test graphs are cyclic, and $\omega_N^{(2)} = o(N^\beta)$ for any $\beta > 0$ otherwise.

Conclusion

Recall that given the test graph $T = (G, \gamma) = (V, E, \gamma)$, we have defined in Section 3.7.1 the graph $\bar G = (\bar V, \bar E)$ obtained from G when the orientation of the edges and their multiplicity are forgotten. The graph $\bar G$ is equipped with a cycle

$\bar{c}$ which allows to recover the initial test graph. Its steps are called heavy steps when they carry heavy Wigner matrices and light steps otherwise.

We introduce the connected, non oriented graph G_{map}, which can have multiple edges: informally, G_{map} is the graph obtained from $\bar{G}$ when we merge the vertices linked by light steps of $\bar{c}$.

More precisely, we first denote by $B^h \subset \bar{E}$ the set of edges of $\bar{G}$ visited by $\bar{c}$ with a heavy step and never visited with a light step. We denote by $B^s \subset \bar{E}$ the set of edges of $\bar{G}$ visited by $\bar{c}$ with a heavy step, without taking into account the light steps. The elements of B^h are called hard bridges, the elements of B^s are called soft bridges. Let $\bar{G}_1, \dots, \bar{G}_d$ be the connected components of the graph $(\bar{V}, \bar{E} \setminus B^h)$. This number d is the same as in Section 3.7.1. The vertices of G_{map} are $\bar{G}_1, \dots, \bar{G}_d$ and for $i, j = 1, \dots, d$, two vertices $\bar{G}_i$ and $\bar{G}_i$ are connected by exactly n edges in G_{map} if there exists exactly n soft bridges between an element of $\bar{G}_i$ and an element of $\bar{G}_j$. Moreover, the heavy steps of $\bar{c}$ induces a cycle on G_{map}, denoted c_{map} whose steps are colored. In Figure 3.7 we have plotted the graph G_{map} equipped with its cycle c_{map} induced by the test graph of Figure 3.5.

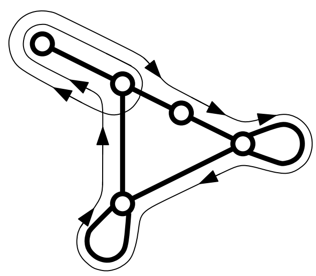

Figure 3.7: The graph G_{map} and its cycle c_{map} for the test graph of figure 3.5.

The number b of edges of G_{map} is the number of soft bridges. Its number of vertices is d, which is the number of test graphs we have considered in Section 3.7.1. Recall that in Section 3.7.1, we have set for any $m = 1, \dots, p$ and $k \geqslant 1$ the number $\eta_{m,k}$ of edges of $\bar{G}$ that are visited by $\bar{c}$ exactly k times by a heavy step of color m. Then, we have set

$$B = \sum_{m=1}^{p} \sum_{k \geqslant 1} \eta_{m,k}.$$

Then, B is actually the number of edges of G_{map} visited by c_{map}, this number being counted with multiplicity with respect to the colors of the steps: an edge visited by exactly steps of n different colors is counted n times. So we get that $b \leqslant B$. Moreover, by the relation between the number d of vertices and the number b of edges in the connected graph G_{map}, we know that $d \leqslant b + 1$. The equality occurs if and only if G_{map} is a tree (see [Gui09]).

If $d < B + 1$, then by (3.28) and (3.30), one has

$$\mathbb{E}\Big[\tau_N^0\Big[T(\mathbf{X}_N,\mathbf{Y}_N)\Big]\Big] = \frac{\omega_N^{(1)}\omega_N^{(2)}}{N^{B+1-d}} = o(N^{\beta-1})$$

for any $\beta > 0$, and so $\mathbb{E}\Big[\tau_N^0\Big[T(\mathbf{X}_N,\mathbf{Y}_N)\Big]\Big] \underset{N\to\infty}{\longrightarrow} 0$.

At the contrary, saying that $d = B+1$ is equivalent to say that (G_{map}, c_{map}) is a colored tree, i.e.

- The graph G_{map} is a tree.
- The colored cycle c_{map} visits each edge of G_{map} with steps of the same color.

In that case, we get that for any $m = 1,\dots,p$ the number $\eta_{m,k}$ vanishes as soon as k is an odd number and the test graphs $T_1,\dots,T_d$ are cyclic. We then have proved the following

Theorem 3.7.1 (The convergence in distribution of traffics of heavy Wigner and deterministic matrices)**.** For any cyclic traffic T in $p+q$ variables, one has $\mathbb{E}\Big[\tau_N^0\Big[T(\mathbf{X}_N,\mathbf{Y}_N)\Big]\Big] \underset{N\to\infty}{\longrightarrow} 0$ as soon as the map of T is not a colored tree. Otherwise, one has

$$\mathbb{E}\Big[\tau_N^0\Big[T(\mathbf{X}_N,\mathbf{Y}_N)\Big]\Big] \underset{N\to\infty}{\longrightarrow} \tau^0[T] := \omega^{(1)}\omega^{(2)}, \tag{3.31}$$

where

$$\begin{aligned}\omega^{(1)} &= \prod_{m=1}^{p}\prod_{k\geqslant 1}\Big(a_{m,\frac{k}{2}}\Big)^{\eta_{m,k}},\\ \omega^{(2)} &= \tau^0\Big[T_1\Big]\times\cdots\times\tau^0\Big[T_d\Big].\end{aligned}$$

3.7.2 The convergence of the trace of polynomial

We consider an integer $L \geqslant 1$, a sequence of colors $\ell = (\ell_1,\dots,\ell_L)$ in $\{1,\dots,p\}^L$ and monic monomials $Q_1,\dots,Q_L$ in $\mathbb{C}\langle\mathbf{y}\rangle$. Let $M_1^{(N)},\dots,M_L^{(N)}$ be the matrices given by $M_k^{(N)} = Q_k(\mathbf{Y}_N)$ for every $k = 1,\dots,L$. The entry (i,j) the the matrix $M_k^{(N)}$ is denoted by $M_k^{(N)}(i,j)$. We consider the matrix

$$H_N = X_{\ell_1}^{(N)}M_1^{(N)}\dots X_{\ell_L}^{(N)}M_L^{(N)}, \tag{3.32}$$

and the polynomial in $\mathbb{C}\langle\mathbf{x},\mathbf{y}\rangle$

$$h = x_{\ell_1}Q_1\dots x_{\ell_L}Q_L, \tag{3.33}$$

so that $H_N = h(\mathbf{X}_N,\mathbf{Y}_N)$. If we can compute the limit of $\mathbb{E}\Big[\tau_N[H_N]\Big]$, then by linearity and traciality we will get the joint limiting distribution of $(\mathbf{X}_N,\mathbf{Y}_N)$. We introduce the test graph $T_h = (V_h, E_h, \gamma_h)$ in $\mathcal{G}_{cyc}\langle x_1,\dots,x_p,y_1,\dots,y_q\rangle$ corresponding to h. First let $\tilde{T}_h = (G,\gamma)$ be the test graph in $\mathcal{G}_{cyc}\langle x_1,\dots,x_p,z_1,\dots,z_L\rangle$ such that: the vertices of G are $1,2,\dots,2L$

and its edges are $(1,2),\dots,(2L-1,2L),(2L,1)$, and we set $\gamma\big((2i+1,2i+2)\big)=\ell_i$ for $i=0,\dots,L-1$ and $\gamma\big((2i,2i+1)\big)=p+i$ for $i=1,\dots,L$ (with the convention $(p,p+1)=(p,1)$). Then, we set $T_h=T(x_1,\dots,x_p,Q_1,\dots,Q_L)$. Hence, one has

$$\tau_N\Big[T_h(\mathbf{X}_N,\mathbf{Y}_N)\Big]=\tau_N[H_N].$$

We expand $\mathbb{E}\Big[\tau_N[H_N]\Big]$ in term of a sum of injective traces as in Proposition 3.3.5

$$\mathbb{E}\Big[\tau_N[H_N]\Big]=\sum_{\sigma\in\mathcal{P}(V_h)}\mathbb{E}\Big[\tau_N^0\Big[\sigma(T)(\mathbf{X}_N,\mathbf{Y}_N)\Big]\Big]. \tag{3.34}$$

By Theorem 3.7.1, for any σ in $\mathcal{P}(V_h)$, one has

$$\mathbb{E}\Big[\tau_N^0\Big[\sigma(T)(\mathbf{X}_N,\mathbf{Y}_N)\Big]\Big]\underset{N\to\infty}{\longrightarrow}\tau^0\Big[\sigma(T)\Big]=\omega^{(1)}(\sigma)\times\omega^{(2)}(\sigma),$$

where $\omega^{(1)}(\sigma)$ and $\omega^{(2)}(\sigma)$ are as in the previous section for the test graph $\sigma(T)$. For any σ in $\mathcal{P}(V_h)$, we denote by $\Big(G_{map}(\sigma),c_{map}(\sigma)\Big)$ the map of $\sigma(T)$. Recall that $\mathcal{L}^{(\ell)}$ is the set of couples (T,c) where T is an embedded rooted tree with at most $\frac{L}{2}$ edges, c is a cycle coloring T and visiting the edges of T in the order relatively to the clockwise orientation (see Definition 3.2.1). Let σ in $\mathcal{P}(V_h)$ such that $G_{map}(\sigma)$ is a tree colored by $c_{map}(\sigma)$. The initial cycle on the test graph T_h is chosen to be the only one that starts at the edge corresponding to x_{ℓ_1}. The choice of this cycle induces a root for the tree G_{map} (the starting vertex of $c_{map}(\sigma)$ and there exists a unique embedding of G_{map} in the plane such that $\Big(G_{map}(\sigma),c_{map}(\sigma)\Big)$ is in $\mathcal{L}^{(\ell)}$. With this convention we get the following

$$\mathbb{E}\Big[\tau_N[H_N]\Big]\underset{N\to\infty}{\longrightarrow}\sum_{(G,c)\in\mathcal{L}^{(\ell)}}\sum_{\sigma\in\mathcal{P}(V_h)}\mathbf{1}_{(G_{map}(\sigma),c_{map}(\sigma))=(G,c)}\omega^{(1)}(\sigma)\times\omega^{(2)}(\sigma),$$

(Equality between graphs is up to isomorphism). But for any σ in $\mathcal{P}(V_h)$ and (G,c) in $\mathcal{L}^{(\ell)}$, if $(G_{map}(\sigma),c_{map}(\sigma))=(G,c)$ then

$$\begin{aligned}\omega^{(1)}(\sigma) &= \prod_{e\hat{E}\text{ edge of }G_{map}} a_{\eta(e),n(e)},\\ \sum_{\sigma\in\mathcal{P}(V_h)}\omega^{(2)}(\sigma) &= \prod_{v\hat{E}\text{ vertex of }G_{map}}\tau\Big[T_{v,c}\Big],\end{aligned}$$

where $2n(e)$ is the number of times c visits e, $\eta(e)$ is the color of e, the $T_{v,c}$ are the test graphs induced by c as in Definition 3.3.7 and τ is the limiting distribution of traffics of $\mathbf{Y}_N$. In particular, these number does not depend on σ and can be denoted $\omega^{(1)}(c)$ and $\omega^{(2)}(c)$ respectively. Hence, one has as expected

$$\mathbb{E}\Big[\tau_N[H_N]\Big]\underset{N\to\infty}{\longrightarrow}\sum_{(G,c)\in\mathcal{L}^{(\ell)}}\omega^{(1)}(c)\times\omega^{(2)}(c).$$

3.7.3 Proof of Theorem 3.5.3

Let $K \geqslant 1$ be an integer. We consider now K matrices $H_{1,N}, \dots, H_{K,N}$ of the following form: for any $k = 1, \dots, K$,

$$H_{k,N} = M_{k,0}^{(N)} X_{\ell_{k,1}}^{(N)} M_{k,1}^{(N)} \dots X_{\ell_{k,L_k}}^{(N)} M_{k,L_k}^{(N)},$$

where $L_k \geqslant 1$ is an integer, $\ell^{(k)} = (\ell_{k,1}, \dots, \ell_{k,L_k})$ in $\{1, \dots, p\}^{L_k}$ is a sequence of colors, and for any $j = 0, \dots, L_k$, one has $M_{k,j}^{(N)} = Q_{k,j}(\mathbf{Y}_N, \mathbf{Y}_N^*)$ where $Q_{k,0}, \dots, Q_{k,L_k}$ are monic monomials in $\mathbb{C}\langle \mathbf{y}, \mathbf{y}^* \rangle$. We set the sequence of colors $\ell = (\ell_{1,1}, \dots, \ell_{1,L_1}, \dots, \ell_{K,1}, \dots, \ell_{K,L_K})$, the integer $L = L_1 + \dots + L_K$ and the family of integers $\mathbf{L} = (L_1, \dots, L_K)$. We also consider the following polynomials: for any $k = 1, \dots, K$

$$h_k = Q_{k,0} x_{\ell_{k,1}} Q_{k,1} \dots x_{\ell_{k,L_k}} Q_{k,L_k}, \tag{3.35}$$

so that $H_{k,N} = h_k(\mathbf{X}_N, \mathbf{Y}_N)$. We introduce a test graph $T_\mathbf{h}$ such that $\tau_N\Big[T_\mathbf{h}(\mathbf{X}_N, \mathbf{Y}_N)\Big] = \Phi_N^{(K)}[H_{1,N}, \dots, H_{K,N}]$. First, let $T^{(K)} = (G^{(K)}, \gamma)$ in $\mathcal{G}_{cyc}\langle z_1, \dots, z_K \rangle$ be the test graph such that: $G^{(K)}$ has a single vertex and K edges, $e_1, \dots, e_K$, linking the vertex to itself. We set $\gamma(e_i) = i$ for any $i = 1, \dots, K$. Then we set $T_\mathbf{h} = T^{(K)}(h_1, \dots, h_K)$. With the same notations as in the previous section, one has

$$\mathbb{E}\Big[\tau_N^0\Big[\sigma(T)(\mathbf{X}_N, \mathbf{Y}_N)\Big]\Big] \underset{N \to \infty}{\longrightarrow} \tau^0\Big[\sigma(T)\Big] = \omega^{(1)}(\sigma) \times \omega^{(2)}(\sigma).$$

Recall that $\mathcal{L}_\mathbf{L}^{(\ell)}$ is the set of couples (T, c) in $\mathcal{L}_\mathbf{L}^{(\ell)}$ such that c is the composition of K cycles, $c = c_1 \circ \dots \circ c_K$, where for any $k = 1, \dots, K$ the cycle c_k is of length L_k. Let σ in $\mathcal{P}(V_h)$ such that $G_{map}(\sigma)$ is a tree colored by $c_{map}(\sigma)$. The initial cycle on $T_\mathbf{h}$ is chosen to be the one starting at the edge corresponding to $Q_{k,0}$, covering the loop corresponding to h_1, and visiting the loops corresponding to $h_2, \dots, h_K$ in this order. The choice of this cycle induces a root for the tree G_{map} (the starting vertex of $c_{map}(\sigma)$ and there exists a unique embedding of G_{map} in the plane such that $\Big(G_{map}(\sigma), c_{map}(\sigma)\Big)$ is in $\mathcal{L}^{(\ell)}$. Necessarily, the map is actually in $\mathcal{L}_\mathbf{L}^{(\ell)}$ and the end of the proof is as in the previous section with minor modifications.

3.8 Proof of the Schwinger-Dyson equations

Let $\mathbf{x} = (x_1, \dots, x_p)$ be a family of heavy semicircular variables falsely free from a family of diagonal non commutative random variables $\mathbf{y}$. For any $j = 1, \dots, p$, we denote by $(a_{j,k})_{k \geqslant 1}$ the parameter of x_j. We start by given the Schwinger-Dyson equation for the trace, and then we give the minor modifications necessary to get the equations for the functionals $\Phi^{(K)}$, $K \geqslant 1$.

3.8.1 The trace of monomials in $(\mathbf{x}, \mathbf{y})$

Let h be a monic monomial as in (3.33):

$$h = x_{\ell_1} Q_1 \dots x_{\ell_L} Q_L.$$

By the definition of heavy semicircular variables and false freeness, using the fact that the non commutative random variables of $\mathbf{y}$ are diagonal we get easily

$$\tau\Big[h(\mathbf{x}, \mathbf{y})\Big] = \sum_{(T,c)\in\mathcal{L}^{(\ell)}} \omega^{(1)}(c) \times \omega^{(2)}(c), \quad (3.36)$$

where

– the weight $\omega^{(1)}(c)$ is obtained by counting the visits of the edges of T

$$\omega^{(1)}(c) = \prod_{e\hat{\text{E}} \text{ edge of } T} a_{\eta(e),n(e)},$$

– the weight $\omega^{(2)}(c)$ is obtained by recording the order of visits of the vertices of T

$$\omega^{(2)}(c) = \prod_{v \text{ vertex of } T} \tau\Big[Q_{j_{v,1}} \dots Q_{j_{v,r_v}}\Big],$$

where π_c is the partition of $\{1, \dots, L\}$ induced by c: two integers i and j belong to the same block of π whenever the $2i-1$-th and the $2j-1$-th steps of c reach the same vertex. We have denote the partition induced by c by $\pi_c = \{B_v\}_{v \text{ vertex of } T}$ and $B_v = \{j_{v,1}, \dots, j_{v,r_v}\}$ in increasing order.

The Schwinger-Dyson equation for this quantity appears when we discuss on the number of times the cycles visits the first vertex.

3.8.2 Cycle visiting $2K$ times the first edge

In the rest of the proof, given a element (T, c), we enumerate the vertices of T in the following way. The starting vertex of T is labelled by the number 1, the second vertex visited by c is labelled by 2 and so on.

Let $(T_c, c) \in \mathcal{L}^{(\ell)}$. The first edge visited by c is the directed edge $a = (1, 2)$. Moreover, c visits the undirected edge $\{1, 2\}$ an even number of times.

Saying that $\{1, 2\}$ is visited exactly $2K$ times is equivalent to say that there exist cycles $d^{(1)}, \dots, d^{(K)}$ and $e^{(1)}, \dots, e^{(K)}$ such that

1. for any $k = 1, \dots, K$, the cycle $d^{(k)}$ starts at the vertex 2,
2. for any $k = 1, \dots, K$, the cycle $e^{(k)}$ starts at the vertex 1,
3. the cycles $d^{(1)}, \dots, d^{(K)}$ and $e^{(1)}, \dots, e^{(K)}$ do not visit the edge $\{1, 2\}$,
4. one has

$$c = a \circ d^{(1)} \circ a^* \circ e^{(1)} \circ a \circ d^{(2)} \circ a^* \circ e^{(2)} \circ \dots \circ a \circ d^{(K)} \circ a^* \circ e^{(K)}, \quad (3.37)$$

where a^* denotes the directed edge $(2,1)$. See Figure 3.4 for an example. Assume that c is of this form. Since the edge $\{1,2\}$ can only be visited by steps of color ℓ_1, we can write

$$\ell = (\ell_1, f^{(1)}, \ell_1, g^{(1)}, \ell_1, f^{(2)}, \ell_1, g^{(2)}, \dots, \ell_1, f^{(K)}, \ell_1, g^{(K)}), \tag{3.38}$$

where for any $k = 1, \dots, K$

– $f^{(k)} = (f_1^{(k)}, \dots, f_{L_k^{(f)}}^{(k)})$ is in $\{1,\dots,p\}^{L_k^{(f)}}$ and $g^{(k)} = (g_1^{(k)}, \dots, g_{L_k^{(g)}}^{(k)})$ is in $\{1,\dots,p\}^{L_k^{(g)}}$ for any $k = 1, \dots, K$,

– $L_k^{(f)}$ is the length of the cycle $d^{(k)}$ and $L_k^{(g)}$ is the length of the cycle $e^{(k)}$.

We define the two cycles

$$d = d^{(1)} \circ \cdots \circ d^{(K)}, \tag{3.39}$$

$$e = e^{(1)} \circ \cdots \circ e^{(K)}. \tag{3.40}$$

For any $k = 1, \dots, K$, We define the sequences of colors

$$f = (f_1^{(1)}, \dots, f_{L_1^{(f)}}^{(1)}, \dots, f_1^{(K)}, \dots, f_{L_K^{(f)}}^{(K)}),$$

$$g = (g_1^{(1)}, \dots, g_{L_1^{(g)}}^{(1)}, \dots, g_1^{(K)}, \dots, g_{L_K^{(g)}}^{(K)}).$$

Write $T_c = (V_c, E_c)$, which is the graph induces by c. Since T_c is tree, the subgraph T_d induced by d is also a tree. We can define a coloration of the cycle d with respect to f. On the other hand, the cycle c, which is colored by ℓ, induces a coloration of d. By the compatibility of the decomposition (3.37) of c and the decomposition (3.38) of ℓ, the two colorations are the same. Hence the cycle d colors the tree T_d. The same fact is true for the cycle e.

Denote by k the number of vertices of T_c, s the number of vertices visited by d and r the number of vertices visited by e. Since T_c is a tree, there is neither edge nor vertices visited both by d and c, so that $V_c = V_d \sqcup V_e$ and $E_c = E_d \sqcup (1,2) \sqcup E_e$. Then, there exist (unique) injective maps $\phi_1 : V_d \to \{1,\dots,s\}$ and $\phi_2 : V_e \to \{1,\dots,r\}$ such that the relabelings of the vertices of T_c by ϕ_1 and ϕ_2 define standard cycles $\bar{d}$ and $\bar{e}$ coloring the trees $\overline{T}_{\bar{d}}$ and $\overline{T}_{\bar{e}}$ respectively. Hence, one has that $(\overline{T}_{\bar{d}}, \bar{d})$ is in $\mathcal{L}^{(f)}$ and $(\overline{T}_{\bar{e}}, \bar{e})$ is in $\mathcal{L}^{(g)}$. If we denote $\mathbf{L}^{(f)} = (L_1^{(f)}, \dots, L_K^{(f)})$ and $\mathbf{L}^{(g)} = (L_1^{(g)}, \dots, L_K^{(g)})$, then by (3.39) and (3.40) we get that actually $(\overline{T}_{\bar{d}}, \bar{d})$ belongs to $\mathcal{L}^{(f)}_{\mathbf{L}^{(f)}}$, $(\overline{T}_{\bar{e}}, \bar{e})$ belongs to $\mathcal{L}^{(g)}_{\mathbf{L}^{(g)}}$.

3.8.3 Reciprocal construction

Let $K \geqslant 1$ be an integer and consider a decomposition of ℓ

$$\ell = (\ell_1, f^{(1)}, \ell_1, g^{(1)}, \ell_1, f^{(2)}, \ell_1, g^{(2)}, \dots, \ell_1, f^{(K)}, \ell_1, g^{(K)}), \tag{3.41}$$

where for any $k = 1, \dots, K$ one has $f^{(k)}$ is in $\{1,\dots,p\}^{L_k^{(f)}}$ and $g^{(k)}$ in $\{1,\dots,p\}^{L_K^{(g)}}$ for sequences of integers $\mathbf{L}^{(f)} = (L_1^{(f)}, \dots, L_K^{(f)})$ and $\mathbf{L}^{(g)} = (L_1^{(g)}, \dots, L_K^{(g)})$. Define

$$f = (f^{(1)}, \dots, f^{(K)}),$$

$$g = (g^{(1)}, \dots, g^{(K)}).$$

Let $(\overline{T}_{\bar d}, \bar d)$ in $\mathcal{L}^{(f)}_{\mathbf{L}^{(f)}}$ and $(\overline{T}_{\bar e}, \bar e)$ in $\mathcal{L}^{(g)}_{\mathbf{L}^{(g)}}$. We write

$$\begin{aligned} \bar d &= \bar d^{(1)} \circ \cdots \circ \bar d^{(K)}, \\ \bar e &= \bar e^{(1)} \circ \cdots \circ \bar e^{(K)}, \end{aligned}$$

where for $k = 1, \dots, K$, $\bar d^{(k)}$ and $\bar e^{(k)}$ are cycles of length $L^{(f)}_k$ and $L^{(g)}_k$ respectively, and so cycles starting from the vertex $\{1\}$. Denote by r the number of vertices visited by $\bar d$ and s the number of vertices visited by $\bar e$. We set $k = r + s$.

We define $\psi_1 : \{1, \dots, s\} \to \{1, \dots, k\}$ and $\psi_2 : \{1, \dots, r\} \to \{1, \dots, k\}$ as follow.

1. We set $\psi_1(1) = 2$ and $\psi_2(1) = 1$.
2. As $\bar d$ makes it first $L^{(f)}_1$-th steps, it visits the vertices $\{1, 2, \dots, w_1\}$. We set $\psi_1(i) = 1 + i$ for any $i = 2, \dots, w_1$.
3. As $\bar e$ makes it first $L^{(g)}_1$-th steps, it visits the vertices $\{1, 2, \dots, z_1\}$. We set $\psi_2(i) = 1 + w_1 + i$ for any $i = 2, \dots, z_1$.
4. As $\bar d$ makes it steps number $L^{(f)}_1 + 1, L^{(f)}_1 + 2, \dots, L^{(f)}_2$, it possibly visits new vertices $\{w_1 + 1, \dots, w_2\}$. We then set $\psi_1(i) = 1 + w_1 + z_1 + i$ for any $i = w_1 + 1, \dots, w_2$.
5. As $\bar e$ makes it steps number $L^{(g)}_1 + 1, L^{(g)}_1 + 2, \dots, L^{(g)}_2$, it possibly visits new vertices $\{z_1 + 1, \dots, z_2\}$. We then set $\psi_2(i) = 1 + w_1 + z_1 + w_2 + i$ for any $i = z_1 + 1, \dots, z_2$. And so on.

The maps ψ_1 and ψ_2 are injective and their ranks are disjoint. Moreover, ψ_1 sends $\bar d$ to a cycle $d = d^{(1)} \circ \cdots \circ d^{(K)}$, and ψ_2 sends $\bar e$ to a cycle $e = e^{(1)} \circ \cdots \circ e^{(K)}$, such that we can define

$$c(\bar d, \bar e) = a \circ d^{(1)} \circ a^* \circ e^{(1)} \circ a \circ d^{(2)} \circ a^* \circ e^{(2)} \circ \cdots \circ a \circ d^{(K)} \circ a^* \circ e^{(K)}$$

which is a cycle on $G_{\mathbb{N}}$. If we denote by $T_{c(\bar d, \bar e)}$ the tree it induces, then $(T_{c(\bar d, \bar e)}, c(\bar d, \bar e))$ belongs to $\mathcal{L}^{(\ell)}$ and visits $\{1, 2\}$ exactly $2K$ times. At last, $\bar d$ and $\bar e$ are the cycles we obtain from $c(\bar d, \bar e)$ with the construction above.

We have proved that

$$\tau\Big[h(\mathbf{x}, \mathbf{y})\Big] = \sum_{K \geqslant 1} \sum_{\substack{(f^{(1)}, \dots, f^{(K)}) \\ (g^{(1)}, \dots, g^{(K)}) \\ \text{as in (3.38)}}} \sum_{\substack{(\overline{T}_{\bar d}, \bar d) \in \mathcal{L}^{(f)}_{\mathbf{L}^{(f)}} \\ (\overline{T}_{\bar e}, \bar e) \in \mathcal{L}^{(g)}_{\mathbf{L}^{(g)}}}} \omega^{(1)}\Big(c(\bar d, \bar e)\Big) \times \omega^{(2)}\Big(c(\bar d, \bar e)\Big).$$

3.8.4 Computation of $\omega^{(1)}\big(c(\bar d, \bar e)\big)$

We consider $\eta(c)$ in $\mathcal{T}^{(\ell)}$, $\eta(\bar d)$ in $\mathcal{T}^{(f)}$ and $\eta(\bar e)$ in $\mathcal{T}^{(g)}$ the arrays obtained by the coloration of c by ℓ, of $\bar d$ by f and of $\bar e$ by g respectively. Since $E_c =$

$E_d \sqcup \Big\{\{1,2\}\Big\} \sqcup E_e$ one has: for all $m = 1, \ldots, p$ and $k \geqslant 1$

$$\left\{ \begin{array}{lcl} \eta_{m,k}(c) & = & \eta_{m,k}(\bar{d}) + \eta_{m,k}(\bar{e}) \quad \text{if } m \neq \ell_1 \text{ or } k \neq K, \\ \eta_{\ell_1,K}(c) & = & \eta_{\ell_1,K}(\bar{d}) + \eta_{\ell_1,K}(\bar{e}) + 1. \end{array} \right. \tag{3.42}$$

Therefore, we get that $\omega^{(1)}\Big(c(\bar{d},\bar{e})\Big) = a_{\ell_1,K} \times \omega^{(1)}(\bar{d}) \times \omega^{(1)}(\bar{e})$.

3.8.5 Computation of $\omega^{(2)}\Big(c(\bar{d},\bar{e})\Big)$

We denote V_d the set of vertices of T visited by d. Its complementary, V_e, consists on the vertices visited by e (in particular, 1 belongs to V_e and 2 belongs to V_d). Define

$$\pi_1 = \Big\{B_q \mid q \in V_e\Big\}, \quad \pi_2 = \Big\{B_q \mid q \in V_d\Big\},$$

Recall the decomposition (3.37) of c:

$$c = a \circ d^{(1)} \circ a^* \circ e^{(1)} \circ a \circ d^{(2)} \circ a^* \circ e^{(2)} \circ \cdots \circ a \circ d^{(K)} \circ a^* \circ e^{(K)}.$$

Let $1 = i_1 < j_1 < i_2 < j_2 < \ldots i_K < j_K \leqslant L$ be the integers such that the steps a in this decomposition are the i_1-th, i_2-th, $\ldots, i_K$-th steps of c, and the steps a^* are the j_1-th, j_2-th, $\ldots, j_K$-th. Then π_1 and π_2 are respectively partitions of the sets of integers N_1 and N_2 given by

$$N_1 = \{1\} \sqcup \{j_1+1, \ldots, i_2\} \sqcup \{j_2+1, \ldots, i_3\} \sqcup \ldots \{j_K+1, \ldots, L\}.$$

$$N_2 = \{2, \ldots, j_1\} \sqcup \{i_2+1, \ldots, j_2\} \sqcup \ldots \{i_K+1, \ldots, j_K\},$$

By relabeling the blocks of π_1 and π_2 in increasing order in $\{1, \ldots, |V_e|\}$ and $\{1, \ldots, |V_d|\}$ respectively, it is clear that

$$\begin{aligned} \omega^{(2)}\Big(c(\bar{d},\bar{e})\Big) &= \prod_{v \text{ vertex of } T} \tau\Big[Q_{j_{v,1}} \ldots Q_{j_{v,r_v}}\Big] \\ &= \prod_{v \in V_d} \tau\Big[Q_{j_{v,1}} \ldots Q_{j_{v,r_v}}\Big] \times \prod_{v \in V_e} \tau\Big[Q_{j_{v,1}} \ldots Q_{j_{v,r_v}}\Big] \\ &= \omega^{(2)}\Big(\bar{d}\Big) \times \omega^{(2)}\Big(\bar{e}\Big). \end{aligned}$$

3.8.6 Conclusion

We have obtained that

$$\tau\Big[h(\mathbf{x},\mathbf{y})\Big] = \sum_{K \geqslant 1} a_{\ell_1,K} \sum_{\substack{(f^{(1)},\ldots,f^{(K)}) \\ (g^{(1)},\ldots,g^{(K)}) \\ \text{as in (3.38)}}} \sum_{d \in \mathcal{L}^{(f)}_{\mathbf{L}^{(f)}}} \omega^{(1)}(d)\, \omega^{(2)}(d) \sum_{e \in \mathcal{L}^{(g)}_{\mathbf{L}^{(g)}}} \omega^{(1)}(e) \omega^{(2)}(e)$$

Only a finite numbers of the first sum are nonzero. On the other hand, we know that for any $(g^{(1)}, \ldots, g^{(K)})$ as in the sum, with the notations above, one has

$$\sum_{e \in \mathcal{L}^{(g)}_{\mathbf{L}^{(g)}}} \omega^{(1)}(e) \omega^{(2)}(e) \quad = \quad \Phi^{(K)}(R^{(1)}, \ldots, R^{(K)}),$$

where we have set

$$\begin{aligned} R^{(1)} &= Q_{j_1} x_{\ell_{j_1+1}} Q_{j_1+1} \dots x_{\ell_{i_2-1}} Q^{(N)}_{i_2-1} \\ &\vdots \\ R^{(K-1)} &= Q_{j_{K-1}} x_{\ell_{j_{K-1}+1}} Q_{j_{K-1}+1} \dots x_{\ell_{i_K-1}} Q_{i_K-1} \\ R^{(K)} &= Q_{j_K} x_{\ell_{j_K+1}} Q_{j_K+1} \dots x_{\ell_L} Q_{i_L}, \end{aligned}$$

and for $(f^{(1)}, \dots, f^{(K)})$ as in the sum, one has

$$\sum_{d \in \mathcal{L}^{(f)}_{\mathbf{L}^{(f)}}} \omega^{(1)}(d)\omega^{(2)}(d) = \Phi^{(K)}(L^{(1)}, \dots, L^{(K)}),$$

where

$$\begin{aligned} L^{(1)} &= Q_{i_1} x_{\ell_{i_1+1}} Q_{i_1+1} \dots x_{\ell_{j_1-1}} Q_{j_1-1} \\ &\vdots \\ L^{(K)} &= Q_{i_K} x_{\ell_{i_K+1}} Q_{i_K+1} \dots x_{\ell_{j_K-1}} Q_{j_{K-1}}. \end{aligned}$$

We define the polynomial $P = Q_1 x_{\ell_2} \dots x_{\ell_L} Q_L$ in $\mathbb{C}\langle \mathbf{x}, \mathbf{y}, \mathbf{y}^* \rangle$. We then have proved that

$$\begin{aligned} \tau[x_{\ell_1} P] = \sum_{k \geqslant 1} a_{j,k} \sum_{x_{\ell_1} P = (x_{\ell_1} L_1 x_{\ell_1}) R_1 \dots (x_{\ell_1} L_K x_{\ell_1}) R_K} & \Phi^{(K)}(L_1, \dots, L_K) \\ & \times \Phi^{(K)}(R_1, \dots, R_K). \end{aligned}$$

3.8.7 The Schwinger-Dyson equation for $\Phi^{(K)}$

Let $h_1, \dots, h_K$ be monic monomials as in (3.35), with $Q_{1,0} = \dots Q_{K,0} = 1$: for any $j = 1, \dots, K$,

$$h_k = x_{\ell_{k,1}} Q_{k,1} \dots x_{\ell_{k,L_k}} Q_{k,L_k},$$

By Theorem 3.5.3 and the definition of heavy semicircular variables and false freeness, we get

$$\Phi^{(K)}(h_1, \dots, h_K)\Big] = \sum_{c \in \mathcal{L}^{(\ell)}_{\mathbf{L}}} \omega^{(1)}(c) \times \omega^{(2)}(c),$$

where are as in (3.36). In this situation, the sum is over cycles c which can be written $c = c_1 \circ \dots c_K$, where for any $j = 1, \dots, K$, c_j is a cycle of length L_j. Assume $L_1 \geqslant 1$. Saying that c visits $\{1, 2\}$ exactly $2k$ times is equivalent to say there exists non negative integers $s_1, \dots, s_K$ such that

- $s_1 \geqslant 1$,
- $s_1 + \dots + s_K = k$,
- for any $j = 1, \dots, K$, the cycle c_j visits a exactly $2s_j$ times.

Assume that for any $j = 1, \dots, K$, the cycle c_j visits $\{1,2\}$ exactly $2s_j$ times. Then we get a decomposition

$$c_1 = a \circ d^{(1,1)} \circ a^* \circ e^{(1,1)} \circ a \circ d^{(1,2)} \circ a^* \circ e^{(1,2)} \circ \dots \circ a \circ d^{(1,s_1)} \circ a^* \circ e^{(1,s_1)},$$

and for any $j = 2, \dots, K$,

$$c_j = e^{(j,0)} \circ a \circ d^{(j,1)} \circ a^* \circ e^{(j,1)} \circ a \circ d^{(j,2)} \circ a^* \circ e^{(j,2)} \circ \dots \circ a \circ d^{(j,s_1)} \circ a^* \circ e^{(j,s_1)}.$$

The only difference is that the cycles $c_1, \dots, c_j$ are not constrained to visit $\{1,2\}$ during their first step. The remain of the proof can be written as we made for the proof of (3.22), without any new niceties. We the same reasoning as before, we obtain the expected result, i.e. Theorem 3.6.2.

3.9 Other proofs

3.9.1 Proof of Lemma 3.2.6

Let (x_1, x_2) be a family of heavy semicircular variables, with x_1 of parameter $(a_{1,k})_{k\geqslant 1}$ and x_2 of parameter $(a_{2,k})_{k\geqslant 1}$. Let (y_1, y_2) be free centered semicircular variables such that $\tau[y_i^2] = \tau[x_i^2]$ for $i = 1, 2$. Assume that the heavy Wigner matrices are non trivial and denote

$$k_i = \min\{k \geqslant 2 \mid a_{i,k} \neq 0\}.$$

The following lemma follows easily from the false freeness property.

Lemma 3.9.1. Let $K \geqslant 2$ and $p_1, \dots, p_K, q_1, \dots, q_K \geqslant 1$.

– if $p_1 + \dots + p_K < k_1$ and $q_1 + \dots + q_K < k_2$, then

$$\tau[x_1^{p_1} x_2^{q_1} \dots x_1^{p_K} x_2^{q_K}] = \tau[y_1^{p_1} y_2^{q_1} \dots y_1^{p_K} y_2^{q_K}].$$

– if $p_1 + \dots + p_K = k_1$ and $q_1 + \dots + q_K < k_2$, then

$$\tau[x_1^{p_1} x_2^{q_1} \dots x_1^{p_K} x_2^{q_K}] = \tau[y_1^{p_1} y_2^{q_1} \dots y_1^{p_K} y_2^{q_K}] + a_{1,k_1} \tau[y_2^{q_1}] \dots \tau[y_2^{q_K}].$$

– if $p_1 + \dots + p_K = k_1$ and $q_1 + \dots + q_K = k_2$, then

$$\begin{aligned} \tau[x_1^{p_1} x_2^{q_1} \dots x_1^{p_K} x_2^{q_K}] &= \tau[y_1^{p_1} y_2^{q_1} \dots y_1^{p_K} y_2^{q_K}] + a_{1,k_1} \tau[y_2^{q_1}] \dots \tau[y_2^{q_K}] \\ &\quad + a_{2,k_2} \tau[y_1^{p_1}] \dots \tau[y_1^{p_K}] + a_{1,k_1} q_{2,k_2}. \end{aligned}$$

Proof. We denote

$$\ell = (\underbrace{1, \dots, 1}_{p_1}, \underbrace{2, \dots, 2}_{q_1}, \dots, \underbrace{1, \dots, 1}_{p_K}, \underbrace{2, \dots, 2}_{q_K}),$$

so that one has

$$\tau[x_1^{p_1} x_2^{q_1} \dots x_1^{p_K} x_2^{q_K}] = \sum_{(T,c)\in\mathcal{L}^{(\ell)}} \prod_{m=1,\dots,p} \prod_{k\geqslant 1} a_{m,k}^{\eta_{m,k}(c)}.$$

In the first case, the only terms that contributes are those for which (T,c) is in $\mathcal{L}_{SC}^{(\ell)}$. In the second case, we have also the contribution of couples (T,c) obtained by folding a tree T_0 where all the edges of the color 1 are attached to the root, and where all these edges are fold into an unique edge. Folding the edges of color 2 does not give any contribution. Moreover, for any $k=1,\dots,K$, the edges corresponding to the terms $x_2^{q_k}$ must form a tree attached to the root of T_0. Hence the contribution $a_{1,k_1}\tau[y_2^{q_1}]\dots\tau[y_2^{q_K}]$. The third case is a combination of the second one and its analogue when we exchange the roles played by x_1 and x_2, in addition to the contribution of the tree where all the edges are attached to the root and all the edges of a same color are folded into an unique edge.

□

We can now prove Lemma 3.2.6. We consider integers $L\geqslant 1$, $n_1,\dots,n_L$ and $m_1,\dots,m_L$ such that $n_1+\dots+n_L=k_1$, $m_1+\dots+m_L=k_2$, and set

$$\Delta=\tau\Big[\Big(x_1^{2n_1}-\tau[x_1^{2n_1}]\Big)\Big(x_2^{2m_1}-\tau[x_2^{2m_1}]\Big)\dots\Big(x_1^{2n_L}-\tau[x_1^{2n_L}]\Big)\Big(x_2^{2m_L}-\tau[x_2^{2m_L}]\Big)\Big].$$

We first expend Δ in the following way.

$$\begin{aligned}\Delta \;=\; &\sum_{r\in\{0,1\}^L}\sum_{s\in\{0,1\}^L}(-1)^{r_1+\dots+r_L+s_1+\dots+s_L}\tau[x_1^{2n_1r_1}]\dots\tau[x_1^{2n_Lr_L}]\\ &\times\tau[x_2^{2m_1s_1}]\dots\tau[x_2^{2m_Ls_L}]\\ &\times\tau[x_1^{2n_1(1-r_1)}x_2^{2m_1(1-s_1)}\dots x_1^{2n_L(1-r_L)}x_2^{2m_L(1-s_L)}].\end{aligned}$$

We separate in the sum above the different cases outlined in Lemma 3.9.1 (and the case where we exchange the roles played by x_1 and x_2).

$$\begin{aligned}\Delta \;=\; &\tau[x_1^{2n_1}x_2^{2m_1}\dots x_1^{2n_L}x_2^{2m_L}]\\ &+\sum_{\substack{r\in\{0,1\}^L\\ r\neq\{0,\dots,0\}}}(-1)^{r_1+\dots+r_L}\tau[x_1^{2n_1r_1}]\dots\tau[x_1^{2n_Lr_L}]\\ &\quad\times\tau[x_1^{2n_1(1-r_1)}x_2^{2m_1}\dots x_1^{2n_L(1-r_L)}x_2^{2m_L}]\\ &+\sum_{\substack{s\in\{0,1\}^L\\ s\neq\{0,\dots,0\}}}(-1)^{s_1+\dots+s_L}\tau[x_2^{2m_1s_1}]\dots\tau[x_2^{2m_Ls_L}]\\ &\quad\times\tau[x_1^{2n_1}x_2^{2m_1(1-s_1)}\dots x_1^{2n_L}x_2^{2m_L(1-s_L)}]\\ &+\sum_{\substack{r\in\{0,1\}^L\\ r\neq\{0,\dots,0\}}}\sum_{\substack{s\in\{0,1\}^L\\ s\neq\{0,\dots,0\}}}(-1)^{r_1+\dots+r_L+s_1+\dots+s_L}\tau[x_1^{2n_1r_1}]\dots\tau[x_1^{2n_Lr_L}]\\ &\quad\times\tau[x_2^{2m_1s_1}]\dots\tau[x_2^{2m_Ls_L}]\\ &\quad\times\tau[x_1^{2n_1(1-r_1)}x_2^{2m_1(1-s_1)}\dots x_1^{2n_L(1-r_L)}x_2^{2m_L(1-s_L)}].\end{aligned}$$

When we apply Lemma 3.9.1, we get the analogue of Δ where we have replaced

(x_1, x_2) by (y_1, y_2) which is zero by freeness, plus the additional terms

$$\begin{aligned}\Delta &= a_{1,k_1}a_{2,k_2} + a_{2,k_2}\sum_{r\in\{0,1\}^L}(-1)^{r_1+\cdots+r_L}\tau[y_1^{2n_1r_1}]\dots\tau[y_1^{2n_Lr_L}]\\ &\quad\times\tau[y_1^{2n_1(1-r_1)}]\dots\tau[y_1^{2n_L(1-r_L)}]\\ &\quad+a_{1,k_1}\sum_{s\in\{0,1\}^L}(-1)^{s_1+\cdots+s_L}\tau[y_2^{2m_1s_1}]\dots\tau[y_2^{2m_Ls_L}]\\ &\quad\times\tau[y_2^{2m_1(1-s_1)}]\dots\tau[y_2^{2m_L(1-s_L)}].\end{aligned}$$

But the two sums are actually zero since,

$$\begin{aligned}&\sum_{r\in\{0,1\}^L}(-1)^{r_1+\cdots+r_L}\tau[y_1^{2n_1r_1}]\dots\tau[y_1^{2n_Lr_L}]\\ &\times\ \tau[y_1^{2n_1(1-r_1)}]\dots\tau[y_1^{2n_L(1-r_L)}]\\ &\qquad= \tau[y_1^{2n_L}]\tau[y_1^{2n_1}]\sum_{r_1\in\{0,1\}}(-1)^{r_1}\dots\sum_{r_L\in\{0,1\}}(-1)^{r_1}=0,\end{aligned}$$

and the same holds for the second sum.

3.9.2 Proof of Lemma 3.5.4

Let $(x, \mathbf{y})$ be as in Lemma 3.5.4 and let x_0 a semicircular variable, with the same variance as x and free from $\mathbf{y}$. Assume $k_0 = \min\{k \geqslant 2 | a_k \neq 0\} < \infty$. Then, with the same reasoning as in the proof of Lemma 3.9.1, we get: for any $K \geqslant 1$, any $p_1, \dots, p_K \geqslant 1$ and any monomials $Q_1, \dots, Q_K$, one has

– if $p_1 + \cdots + p_K < k_0$, then

$$\tau[x^{p_1}Q_1(\mathbf{y})\dots x^{p_K}Q_K(\mathbf{y})] = \tau[x_0^{p_1}Q_1(\mathbf{y})\dots x_0^{p_K}Q_K(\mathbf{y})].$$

– if $p_1 + \cdots + p_K = k_0$, then

$$\begin{aligned}\tau[x^{p_1}Q_1(\mathbf{y})\dots x^{p_K}Q_K(\mathbf{y})] &= \tau[x_0^{p_1}Q_1(\mathbf{y})\dots x_0^{p_K}Q_K(\mathbf{y})]\\ &\quad+a_{k_0}\Phi^{(K)}\Big(Q_1(\mathbf{y}),\dots,Q_K(\mathbf{y})\Big).\end{aligned}$$

Consider $L \geqslant 2$, $n_1, \dots, n_L \geqslant 1$ such that $n_1 + \cdots + n_L = k_0$, and $m_1, \dots, m_L$ monomials in $\mathbf{y}$. We set

$$\Delta = \tau\Bigg[\Big(x^{2n_1}-\tau[x^{2n_1}]\Big)\Big(m_1-\tau[m_1]\Big)\dots\Big(x^{2n_L}-\tau[x^{2n_L}]\Big)\Big(m_L-\tau[m_L]\Big)\Bigg].$$

With the same computation as in the previous section, one has

$$\begin{aligned}\Delta &= a_{k_0}\sum_{\mathbf{r}\in\{0,1\}^L\hat{\mathrm{E}}}(-1)^{r_1+\cdots+r_L}\Phi^{(L)}\Big(m_1^{1-r_1},\dots,m_L^{1-r_L}\Big)\\ &= a_{k_0}\Phi^{(L)}\Big(m_1-\tau[m_1],\dots,m_L-\tau[m_L]\Big).\end{aligned}$$

3.9.3 Examples of computations

Computation of $\tau[x_1^2x_2^2x_1^2x_2^2]$ by the false freeness property

First, we enumerate (Figure 3.8) the non crossing pair partitions associated to this word, and then deduce the trees of the corresponding set $\mathcal{L}_{SC}^{(\ell)}$ (see [AGZ10] for a correspondence between these two family of objects). The only tree that can be folded is the third one. Then, we enumerate (Figure 3.9) the cycles coloring a tree we deduce by folding this tree. By counting the contribution of each tree, we get

$$\tau[x_1^2x_2^2x_1^2x_2^2] = 3a_{1,1}^2a_{2,1}^2 + a_{1,1}^2a_{2,2} + a_{1,2}a_{2,1}^2 + a_{1,2}a_{2,2}.$$

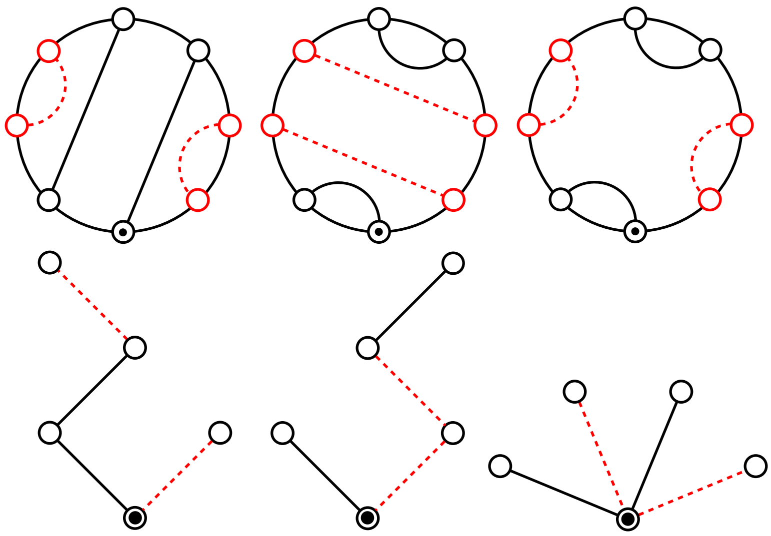

Figure 3.8: Enumeration of non crossing pair partition (top) in the computation of $\tau[x_1^2x_2^2x_1^2x_2^2]$ and the tree associated (bottom).

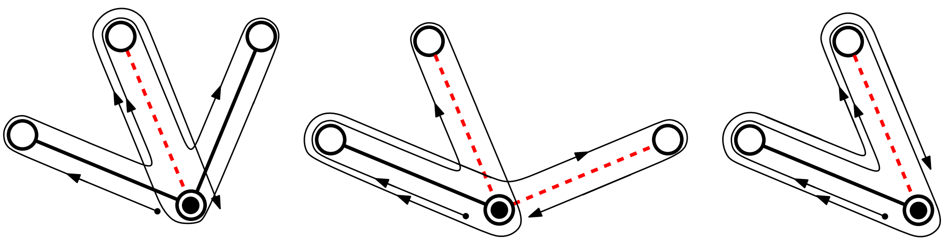

Figure 3.9: Tree cycles coloring a tree in the computation of $\tau[x_1^2x_2^2x_1^2x_2^2]$.

Computation of $\tau[x_1^2x_2^2x_1^2x_2^2]$ by the Schwinger-Dyson equations

First, we enumerate the decompositions

$$\begin{aligned} x_1^2x_2^2x_1^2x_2^2 &= (x_1\times 1\times x_1)x_2^2x_1^2x_2^2 \\ &= (x_1\times x_1x_2^2\times x_1)x_1x_2^2 \\ &= (x_1\times x_1x_2^2x_1\times x_1)x_2^2 \\ &= (x_1\times 1\times x_1)x_2^2(x_1\times 1\times x_1)x_2^2. \end{aligned}$$

Then, by Theorem 3.6.2 we get

$$\begin{aligned} \tau[x_1^2x_2^2x_1^2x_2^2] &= a_{1,1}\Big(\tau[1]\tau[x_2^2x_1^2x_2^2]+\tau[x_1x_2^2]\tau[x_1x_2^2]+\tau[x_1x_2^2x_1]\tau[x_2^2]\Big) \\ &\quad +a_{1,2}\Phi^{(2)}(1,1)\Phi^{(2)}(x_2^2,x_2^2) \\ &= a_{1,1}\Big(\tau[x_1^2]\tau[x_2^4]+0+\tau[x_1^2]\tau[x_2^2]^2\Big)+a_{1,2}\Phi^{(2)}(x_2^2,x_2^2) \\ &= a_{1,1}^2a_{2,1}^2+a_{1,1}^2\tau[x_2^4]+a_{1,2}\Phi^{(2)}(x_2^2,x_2^2), \end{aligned}$$

where we have used the facts that $\tau[x_1^nx_2^m]=\tau[x_1^n]\tau[x_2^m]$ for any $n,m\geqslant 1$ and that $\tau[x_i^2]=a_{i,1}$ for $i=1,2$. By Theorem 3.6.2, one has with a similar computation

$$\begin{aligned} \tau[x_2^4] &= a_{2,1}\Big(\tau[1]\tau[x_2^2]+\tau[x_2]\tau[x_2]+\tau[x_2]\tau[1]\Big)+a_{2,2}\Phi^{(2)}(1,1)\Phi^{(2)}(1,1) \\ &= 2a_{2,1}^2+a_{2,2}. \end{aligned}$$

To compute $\Phi^{(2)}(x_2^2,x_2^2)$ with Theorem 3.6.2, we enumerate the decompositions

$$(x_2^2,x_2^2)=\Big((x_2\times 1\times x_2)1,x_2^2\Big)=\Big((x_2\times 1\times x_2)1,1(x_2\times 1\times x_2)1\Big).$$

So we have

$$\begin{aligned} \Phi^{(2)}(x_2^2,x_2^2) &= a_{2,1}\tau[1]\Phi^{(2)}(1,x_2^2)+a_{2,2}\Phi^{(2)}(1,1)\Phi^{(3)}(1,1,1) \\ &= a_{2,1}^2+a_{2,2}. \end{aligned}$$

We then get as expected

$$\begin{aligned} \tau[x_1^2x_2^2x_1^2x_2^2] &= a_{1,1}^2a_{2,1}^2+a_{1,1}^2(2a_{2,1}^2+a_{2,2})+a_{1,2}(a_{2,1}^2+a_{2,2}) \\ &= 3a_{1,1}^2a_{2,1}^2+a_{1,1}^2a_{2,2}+a_{1,2}a_{2,1}^2+a_{1,2}a_{2,2}. \end{aligned}$$

3.9.4 Proof of Proposition 3.6.3

We manipulate truncated sums. Let $N\geqslant 1$ be an integer. Then, by Theorem 3.6.2

$$\begin{aligned} &\frac{1}{\lambda^K}\sum_{n=0}^{N}\sum_{n_1+\cdots+n_K=n}\frac{1}{\lambda^n}\Phi^{(K)}(x^{n_1+1},x^{n_2},\ldots,x^{n_K}) \\ &= \frac{1}{\lambda^K}\sum_{n=0}^{N}\sum_{n_1+\cdots+n_K=n}\frac{1}{\lambda^n}\sum_{1\leqslant k\leqslant\frac{n+1}{2}}a_k \\ &\quad\times \sum_{\substack{s_1+\cdots+s_K=k\\ 1\leqslant s_1\leqslant\frac{n_1+1}{2}\\ 0\leqslant s_2\leqslant\frac{n_2}{2},\ldots,0\leqslant s_K\leqslant\frac{n_K}{2}}}\sum_{(\mathbf{r},\mathbf{t})}\Phi^{(k)}(x^{\mathbf{r}})\Phi^{(k+K-1)}(x^{\mathbf{t}}). \end{aligned}$$

The last sum is over all families of non negative integers

$$\mathbf{r} = (r_1^{(1)}, \dots, r_{s_1}^{(1)}, \dots, r_1^{(K)}, \dots, r_{s_K}^{(K)}),$$

$$\mathbf{t} = (t_1^{(1)}, \dots, t_{s_1}^{(1)}, t_0^{(2)}, \dots, t_{s_2}^{(2)}, \dots, t_0^{(K)}, \dots, t_{s_K}^{(K)}),$$

such that

$$\begin{array}{rcl} r_1^{(1)} + \cdots + r_{s_1}^{(1)} + t_1^{(1)} + \cdots + t_{s_1}^{(1)} & = & n_1 + 1 - 2s_1, \\ r_1^{(2)} + \cdots + r_{s_2}^{(2)} + t_0^{(2)} + \cdots + t_{s_2}^{(2)} & = & n_2 - 2s_2, \\ & \vdots & \\ r_1^{(K)} + \cdots + r_{s_K}^{(K)} + t_0^{(K)} + \cdots + t_{s_K}^{(K)} & = & n_K - 2s_K. \end{array}$$

We have used (and we will use) the notation

$$\Phi^{(k)}(x^{\mathbf{r}}) = \Phi^{(k)}(x^{r_1^{(1)}}, \dots, x^{r_{s_1}^{(1)}}, \dots, x^{r_1^{(K)}}, \dots, x^{r_{s_K}^{(K)}}).$$

The restrictions on the second and third sums follow from consideration on the degree on the monomials we compute. Then one has

$$\begin{aligned} & \frac{1}{\lambda^K} \sum_{n=0}^{N} \sum_{n_1+\cdots+n_K=n} \frac{1}{\lambda^n} \Phi^{(K)}(x^{n_1+1}, x^{n_2}, \dots, x^{n_K}) \\ & = \frac{1}{\lambda^K} \sum_{1 \leqslant k \leqslant \frac{N+1}{2}} a_k \sum_{\substack{s_1+\cdots+s_K=k \\ s_1 \geqslant 1,\ s_2,\dots,s_K \geqslant 0}} \sum_{2k-1 \leqslant n \leqslant N} \frac{1}{\lambda^n} \\ & \quad \times \sum_{\substack{n_1+\cdots+n_K=n \\ n_1 \geqslant 2s_1-1 \\ n_2 \geqslant 2s_2,\dots,n_K \geqslant 2s_K}} \sum_{\mathbf{l}} \sum_{\mathbf{r}} \Phi^{(k)}(x^{\mathbf{r}}) \sum_{\mathbf{t}} \Phi^{(k+K-1)}(x^{\mathbf{t}}). \end{aligned}$$

By the sum over $\mathbf{l}$, we mean the sum over all families of non negative integers $\mathbf{l} = (l_1, \dots, l_K)$ such that

$$\begin{array}{rcl} 0 \leqslant & l_1 & \leqslant n_1 + 1 - 2s_1, \\ 0 \leqslant & l_2 & \leqslant n_2 - 2s_2, \\ \vdots & & \\ 0 \leqslant & l_K & \leqslant n_K - 2s_K. \end{array}$$

By the sum over $\mathbf{r}$, we mean the sum over all families of non negative integers

$$\mathbf{r} = (r_1^{(1)}, \dots, r_{s_1}^{(1)}, \dots, r_1^{(K)}, \dots, r_{s_K}^{(K)}),$$

such that

$$\begin{array}{rcl} r_1^{(1)} + \cdots + r_{s_1}^{(1)} & = & l_1, \\ r_1^{(2)} + \cdots + r_{s_2}^{(2)} & = & l_2, \\ & \vdots & \\ r_1^{(K)} + \cdots + r_{s_K}^{(K)} & = & l_K. \end{array}$$

At last, by the sum over $\mathbf{t}$, we mean the sum over all families of non negative integers

$$\mathbf{t} = (t_1^{(1)}, \dots, t_{s_1}^{(1)}, t_0^{(2)}, \dots, t_{s_2}^{(2)}, \dots, t_0^{(K)}, \dots, t_{s_K}^{(K)}),$$

such that

$$\begin{array}{rcl} t_1^{(1)} + \cdots + t_{s_1}^{(1)} & = & n_1 + 1 - 2s_1 - l_1, \\ t_0^{(2)} + \cdots + t_{s_2}^{(2)} & = & n_2 - 2s_2 - l_2, \\ & \vdots & \\ t_0^{(K)} + \cdots + t_{s_K}^{(K)} & = & n_k - 2s_K - l_K. \end{array}$$

Given $k, s_1, \dots, s_2$ as in the previous formula, we set the change of variable for $n, n_1, \dots, n_K$

$$\begin{array}{rcl} m & = & n + 1 - 2k, \\ m_1 & = & n_1 + 1 - 2s_1, \\ m_2 & = & n_2 - 2s_2, \\ \vdots & & \\ m_K & = & n_K - 2s_K. \end{array}$$

Remark first that

$$\frac{1}{\lambda^K} \times \frac{1}{\lambda^n} = \frac{1}{\lambda^m} \times \frac{1}{\lambda^{k+K-1}} \times \frac{1}{\lambda^k}.$$

Hence we get

$$\begin{aligned} &\frac{1}{\lambda^K} \sum_{n=0}^{N} \sum_{n_1+\cdots+n_K=n} \frac{1}{\lambda^n} \Phi^{(K)}(x^{n_1+1}, x^{n_2}, \dots, x^{n_K}) \\ &= \sum_{1\leqslant k \leqslant \frac{N+1}{2}} a_k \sum_{\substack{s_1+\cdots+s_K=k \\ s_1\geqslant 1,\ s_2,\dots,s_K\geqslant 0}} \sum_{m=0}^{N+1-2k} \frac{1}{\lambda^m} \\ &\quad \times \sum_{m_1+\cdots+m_K=m} \sum_{\substack{l_1=0\dots m_1 \\ \cdots \\ l_K=0\dots m_K}} \sum_{\mathbf{r}} \frac{1}{\lambda^k} \Phi^{(k)}(x^{\mathbf{r}}) \sum_{\mathbf{t}} \frac{1}{\lambda^{k+K-1}} \Phi^{(k+K-1)}(x^{\mathbf{t}}). \end{aligned}$$

The sum over $\mathbf{r}$ is the same as before, and now last sum is over all families of non negative integers

$$\mathbf{t} = (t_1^{(1)}, \dots, t_{s_1}^{(1)}, t_0^{(2)}, \dots, t_{s_2}^{(2)}, \dots, t_0^{(K)}, \dots, t_{s_K}^{(K)}),$$

such that

$$\begin{array}{rcl} t_1^{(1)} + \cdots + t_{s_1}^{(1)} & = & m_1 - l_1, \\ & \vdots & \\ t_0^{(K)} + \cdots + t_{s_K}^{(K)} & = & m_K - l_K. \end{array}$$

We replace the set variables $(m_1, \dots, m_K, l_1, \dots, l_K)$ by variables $p_1, \dots, p_K$ and $q_1, \dots, q_K$ where for any $i = 1, \dots, K$ we have set $p_i = m_i - l_i$ and $q_i = l_i$. Then we get

$$\begin{aligned}
&\frac{1}{\lambda^K} \sum_{n=0}^{N} \sum_{n_1+\dots+n_K=n} \frac{1}{\lambda^n} \Phi^{(K)}(x^{n_1+1}, x^{n_2}, \dots, x^{n_K}) \\
&= \sum_{1 \leqslant k \leqslant \frac{N+1}{2}} a_k \sum_{m=0}^{N+1-2k} \frac{1}{\lambda^m} \sum_{(\mathbf{p},\mathbf{q})} \sum_{\substack{s_1+\dots+s_K=k \\ s_1 \geqslant 1,\ s_2,\dots,s_K \geqslant 0}} \\
&\qquad \times \sum_{\mathbf{r}} \frac{1}{\lambda^k} \Phi^{(k)}(x^{\mathbf{r}}) \sum_{\mathbf{t}} \frac{1}{\lambda^{k+K-1}} \Phi^{(k+K-1)}(x^{\mathbf{t}}).
\end{aligned}$$

The sum over $(\mathbf{p}, \mathbf{q})$ is the sum over all families of non negative integers $\mathbf{p} = (p_1, \dots, p_K)$ and $\mathbf{q} = (q_1, \dots, q_K)$ such that

$$p_1 + \dots + p_K + q_1 + \dots q_K = m.$$

The sum over $\mathbf{r}$ is the sum over all families of non negative integers

$$\mathbf{r} = (r_1^{(1)}, \dots, r_{s_1}^{(1)}, \dots, r_1^{(K)}, \dots, r_{s_K}^{(K)}),$$

such that

$$\begin{aligned}
r_1^{(1)} + \dots + r_{s_1}^{(1)} &= q_1, \\
&\vdots \\
r_1^{(K)} + \dots + r_{s_K}^{(K)} &= q_K.
\end{aligned}$$

The sum over $\mathbf{t}$ is the sum over all families of non negative integers

$$\mathbf{t} = (t_1^{(1)}, \dots, t_{s_1}^{(1)}, t_0^{(2)}, \dots, t_{s_2}^{(2)}, \dots, t_0^{(K)}, \dots, t_{s_K}^{(K)}),$$

such that

$$\begin{aligned}
t_1^{(1)} + \dots + t_{s_1}^{(1)} &= p_1, \\
t_0^{(2)} + \dots + t_{s_2}^{(2)} &= p_2, \\
&\vdots \\
t_0^{(K)} + \dots + t_{s_K}^{(K)} &= p_K.
\end{aligned}$$

Let $K \geqslant 1$ and $k \geqslant 1$ be integers. Then there exist $\binom{K+k-2}{K-1}$ tuples of non negative integers $(s_1, \dots, s_K)$ such that $s_1 + \dots s_K = k$, $s_1 \geqslant 1$ and $s_2, \dots, s_K \geqslant 0$. Hence we get

$$\begin{aligned}
&\frac{1}{\lambda^K} \sum_{n=0}^{N} \sum_{n_1+\dots+n_K=n} \frac{1}{\lambda^n} \Phi^{(K)}(x^{n_1+1}, x^{n_2}, \dots, x^{n_K}) \\
&= \sum_{1 \leqslant k \leqslant \frac{N+1}{2}} a_k \binom{K+k-2}{K-1} \\
&\qquad \times \sum_{0 \leqslant p+q \leqslant N+1-2k} \frac{1}{\lambda^k} \sum_{r_1+\dots+r_k=q} \frac{1}{\lambda^q} \Phi^{(k)}(x^{r_1}, \dots, x^{r_q}) \\
&\qquad \times \frac{1}{\lambda^{k+K-1}} \sum_{t_1+\dots+t_{k+K-1}=p} \frac{1}{\lambda^p} \Phi^{(k+K-1)}(x^{t_1}, \dots, x^{t_{k+K-1}}).
\end{aligned}$$

This gives the expected result by identification of the coefficients. The uniqueness of the solution of the equations follows directly from the observation of the valence of the formal power series.

3.10 Appendix: A short discussion on the model of heavy Wigner matrices

3.10.1 On the assumptions

We use a slight different definition of the model of heavy Wigner matrices that the ones given by Zakharevich [Zak06] and Ryan [Rya98]. Let X_N be a heavy Wigner matrix and denote by $p^{(N)}$ the common law of its entries. In [Zak06], it is assumed that the even moments of the entries satisfy: for any integer $k \geqslant 0$

$$\lim_{N\to\infty} \frac{\int t^{2k+1} \mathrm{d}p^{(N)}(t)}{N^{k-\frac{1}{2}}} \text{ exists in } \mathbb{R}. \tag{3.43}$$

In [Rya98], the matrices considered has actually complex entries and are Hermitian whereas in this paper we only consider symmetric matrices. In [Rya98], when we only consider orthogonal matrices, it is assumed that

$$\frac{\int t^{2k+1} \mathrm{d}p^{(N)}(t)}{N^{k-\frac{1}{2}}} = o(N^{\beta}), \quad \forall \beta > 0. \tag{3.44}$$

The assumption (3.43) in [Zak06] could actually be replaced by the assumption (3.44) with minor modifications. Moreover, under the assumption we make in our definition of the model, which is for any integer $K \geqslant 1$

$$\lim_{N\to\infty} \frac{\int t^{2k} \mathrm{d}p^{(N)}(t)}{N^{k-1}} \text{ exists in } \mathbb{R}, \tag{3.45}$$

we get by the Cauchy-Schwarz's inequality that

$$\frac{\int t^{2k+1} \mathrm{d}p^{(N)}(t)}{N^{k-\frac{1}{2}}} \leqslant \sqrt{\frac{\int t^{4k} \mathrm{d}p^{(N)}(t)}{N^{2k-1}} \int t^2 \mathrm{d}p^{(N)}(t)} = O(1),$$

which implies (3.44).

3.10.2 The possible parameters of heavy Wigner matrices

It is natural to ask when a sequence of integers $(a_k)_{k\geqslant 1}$ can be a parameter of a heavy Wigner matrix. The answer is given by the Hamburger's moment problem, which characterizes sequence of numbers which are moments of measures.

Proposition 3.10.1. If a sequence $(a_k)_{k\geqslant 1}$ of real numbers is a parameter of a heavy Wigner matrix, then it is the null sequence or it is the sequence of even moments of a Borel measure m with finite moments, i.e. for any $k \geqslant 1$, $a_k = \int t^{2k-2} \mathrm{d}m(t)$. In particular, if the parameter $(a_k)_{k\geqslant 1}$ is non trivial then one has $a_k > 0$ for any $k \geqslant 1$.

Proof. By the Hamburger's theorem [Ham21], a sequence of real numbers $\big(\mu(k)\big)_{k\geqslant 1}$ is a sequence of moments if and only if, for any sequence $(x_k)_{k\geqslant 0}$ of complex numbers with finite support, one has

$$\sum_{j,k\geqslant 0} \mu(j+k)x_j\bar{x}_k \geqslant 0. \tag{3.46}$$

Let X_N be a heavy Wigner matrix of parameter $(a_k)_{k\geqslant 1}$ et let $p^{(N)}$ be the common law of its entries. One can always assume that $p^{(N)}$ is symmetric, since we get the same parameter for the heavy Wigner matrix whose common law of the entries is the symmetrization of $p^{(N)}$. Denote by $(\mu^{(N)}(k))_{k\geqslant 0}$ its sequence of moments. For any sequence $(y_k)_{k\geqslant 1}$ of complex numbers with finite support such that $y_0 = 0$, we apply (3.46) with $(x_k)_{k\geqslant 1} = (N^{\frac{k}{2}}y_k)_{k\geqslant 1}$: we get

$$\begin{aligned} \sum_{j,k\geqslant 0} \mu^{(N)}(j+k)x_j\bar{x}_k &= N\sum_{j,k\geqslant 0} \frac{\mu^{(N)}(j+k)}{N^{\frac{j+k}{2}-1}} y_j\bar{y}_k \\ &= N\sum_{j,k\geqslant 2} a_{\frac{j+k}{2}} y_j\bar{y}_k + o(1), \end{aligned}$$

where we have set $a_k = 0$ whenever k is odd. This gives the necessary condition.

Now, assume that $(a_k)_{k\geqslant 1}$ is non trivial. Consider the 3 by 3 matrix obtained from $A_m^{(N)}$ by keeping only the 3 last lines and column, $m \geqslant 3$. Since $A_m^{(N)}$ is positive definite, the determinant of this matrix is positive. But it converges to $a_{m-1}(a_{m-2}a_m - a_{m-1}^2)$. We then get by recurrence that if $a_{m-2} = 0$, then $a_k = 0$ for any $k \geqslant m-2$, and one the other hand that if $a_1, a_2 > 0$ then $a_k > 0$ for any $k \geqslant 1$.

□

Chapter 4

A central limit theorem for the injective trace of test graphs in independent heavy Wigner matrices

Work in progress, with Florent Benaych-Georges and Alice Guionnet

ABSTRACT:

We prove that, properly rescaled and centered, the injective traces of cyclic test graphs in a family of independent N by N heavy Wigner matrices converges to a multivariate gaussian processes as N goes to infinity. The covariance function of this process is written via the limiting distribution of traffics of the heavy Wigner matrices. In particular, we show a central limit theorem for linear statistics of the empirical eigenvalue distribution of a heavy Wigner matrix X_N.

4.1 Introduction

Given a polynomial P, we show that the random variable

$$\sqrt{N}\Big(\tau_N\big[P(X_N)\big] - \mathbb{E}\big[\tau_N\big[P(X_N)\big]\big]\Big) \tag{4.1}$$

converges to a Gaussian random variable, where τ_N denotes the normalized trace. Normalizing the centered trace by a factor $\sqrt{N}$ is unusual in random matrix theory. If X_N were a Wigner or a Wishart matrix, then we know that a central limit theorem holds with the normalizing factor N ([Jon82]). The fluctuations of linear statistic for heavy Wigner matrices are then at the same scale than the fluctuations of independent identically distributed random variables.

Our result is actually more general since we consider a family $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ of independent N by N matrices and work with the formalism of distribution of traffics introduced in Chapter 3. Given a cyclic test graph T in p indeterminates, we denote

$$Z_N(T) = \sqrt{N}\bigg(\tau_N^0\Big[T(\mathbf{X}_N)\Big] - \mathbb{E}\Big[\tau_N^0\Big[T(\mathbf{X}_N)\Big]\Big]\bigg),$$

where τ^0 is the injective normalized trace of N by N matrices, and show a multivariate central limit theorem for $\big(Z_N(T)\big)_{T\in\mathcal{G}_{cyc}\langle x_1,\dots,x_p\rangle}$.

The random variable in (4.1) can be written as a linear combination of $Z_N(T_1),\dots,Z_N(T_K)$, where $T_1,\dots,T_K$ are traffics in one variable. Hence, the multivariate central limit theorem for $\big(Z_N(T)\big)_{T\in\mathcal{G}_{cyc}\langle x_1,\dots,x_p\rangle}$ shown in this paper give a central limit theorem for the linear statistic of the empirical eigenvalue distribution.

Being Gaussian, the limiting process $\big(z(T)\big)_{T\in\mathcal{G}_{cyc}\langle x_1,\dots,x_p\rangle}$ of $\big(Z_N(T)\big)_{T\in\mathcal{G}_{cyc}\langle x_1,\dots,x_p\rangle}$ is completely characterized by its covariance map

$$\begin{array}{cccc} \rho: & \mathcal{G}_{cyc}\langle x_1,\dots,x_p\rangle^2 & \to & \mathbb{C} \\ & (T,T') & \mapsto & \mathbb{E}\Big[z(T)\times z(T')\Big]. \end{array}$$

We give a simple formula for $\delta(T,T')$ in terms of the limiting distribution of traffics of $\mathbf{X}_N$. Hence, the notion of distribution of traffics seems robust enough to have its "second order false freeness theory“, as in free probability with Mingo and Speicher's second order freeness theory [MS06].

Organization of the proof:
In Section 4.2, we give the precise statement of our result which is shown in Section 4.3.

Acknowledgment:
The authors gratefully aknowledge the financial support of the ANR GranMa.

4.2 Statement of results

By Theorem 3.3.8 of Chapter 3, we have the following description of the limiting distribution of traffics of a family of independent heavy Wigner matrices. Let $T=(G,\gamma)$ be a test graph in $\mathcal{G}\langle x_1,\dots,x_p\rangle$. We say that T is a colored tree whenever the graph $\bar G$ obtained from G by forgetting the orientation and the multiplicity of the edges, and the edges linking a same pair of vertices in $\bar G$ are the same. For such a colored tree, any edge e of $\bar G$ as a color $\eta(e)$ in $\{1,\dots,p\}$, which is the common color of the corresponding edges in G. Moreover, we denote by $n(e)$ the multiplicity of the edges corresponding to e in G.

Theorem 4.2.1. Let $\mathbf{X}_N=(X_1^{(N)},\dots,X_p^{(N)})$ be a family of N by N independent heavy Wigner matrices. For any $m=1,\dots,p$, we set $(a_{m,k})_{k\geqslant 1}$ the parameter of

$X_m^{(N)}$. Then, for any test graph $T = (G, \gamma)$ in $\mathcal{G}_{cyc}\langle x_1, \dots, x_p \rangle$, one has

$$\mathbb{E}\Big[\tau_N^0\big[T(\mathbf{X}_N)\big]\Big] \underset{N\to\infty}{\longrightarrow} \tau^0[T] := \begin{cases} \prod_{e\in\bar{G}} a_{\eta(e),n(e)} & \text{if } T \text{ is a colored tree} \\ 0 & \text{otherwise.} \end{cases} \tag{4.2}$$

To describe the fluctuations of

$$Z_N(T) = \sqrt{N}\bigg(\tau_N^0\big[T(\mathbf{X}_N)\big] - \mathbb{E}\Big[\tau_N^0\big[T(\mathbf{X}_N)\big]\Big]\bigg), \tag{4.3}$$

we need the following definitions.

Let T, T' be test graphs in $\mathcal{G}\langle x_1, \dots, x_p\rangle$. We define $\delta(T, T') \subset \mathcal{G}\langle x_1, \dots, x_p\rangle$, called the set of amalgamated product of T and T', as the set of all test graphs obtained by merging certain vertices of T and T'. More precisely, write $T = (V, E, \gamma)$ and $T' = (V', E', \gamma')$. Then $\delta(T, T')$ is the set of test graphs $T'' = (V'', E'', \gamma'')$ of the following form. There exist non empty sets of the same cardinal $W \subset V, W' \subset V'$ and a bijection $\psi : W \to W'$. This bijection is extended trivially to a bijection $\psi : W \sqcup (V' \setminus W') \to V'$. The set of vertices V'' of the test graph T'' is

$$V'' \quad = \quad V \sqcup (V' \setminus W'),$$

and a directed edge $e = (v_1, v_2)$ is in the multi-set E'' whenever
- v_1, v_2 are in V and e is an edge of T, or
- v_1, v_2 are in $W \sqcup (V' \setminus W')$ and $\big(\psi'(v_1), \psi(v_2)\big)$ is in E',

this enumeration taking account to the multiplicity of the edges in the multi-sets E and E' (with a certain abuse, we can think E'' as the set $E\hat{\text{E}} \sqcup E'$). The map γ'' is induced by the maps γ and γ'.

When T'' is obtained by such a construction, we denote

$$T'' = T \underset{\psi : W \to W'}{*} T'.$$

We also define $\delta_\sharp(T, T')$ as the set of all amalgamated products T'' of T and T' such that, with the notations above, there exists a pair of vertices in W'' linked both by an edge from E and an edge from E'. The main result of this note is the following.

Theorem 4.2.2. The process $\big(Z_N(T)\big)_{T\in\mathcal{G}_{cyc}\langle x_1,\dots,x_p\rangle}$ converges in mean moments to a centered Gaussian random process $\big(z(T)\big)_{T\in\mathcal{G}_{cyc}\langle x_1,\dots,x_p\rangle}$, i.e. for any integer $n \geqslant 1$, for any polynomial P in n indeterminates and any cyclic test graph $T_1, \dots, T_n$, one has

$$\mathbb{E}\Big[P\Big(Z_N(T_1), \dots, Z_N(T_n)\Big)\Big] \underset{N\to\infty}{\longrightarrow} \mathbb{E}\Big[P\Big(z(T_1), \dots, z(T_n)\Big)\Big]. \tag{4.4}$$

The covariance map of $\big(z(T)\big)_{T\in\mathcal{G}_{cyc}\langle x_1,\dots,x_p\rangle}$ is given by: for any cyclic test graphs T_1 and T_2,

$$\rho(T,T') \;:=\; \mathbb{E}\Big[z(T)\times z(T')\Big] = \sum_{T''\in\delta_\sharp(T,T')} \tau^0[T''], \tag{4.5}$$

where τ^0 is as in (4.2).

4.3 Proof of Theorem 4.2.2

Let $\mathbf{X}_N = (X_1^{(N)},\dots,X_p^{(N)})$ be a family of N by N independent heavy Wigner matrices. For any $m=1,\dots,p$, we set $(a_{m,k})_{k\geqslant 1}$ the parameter of $X_m^{(N)}$. By Wick's formula (need ref), to show Theorem 4.2.2 it is sufficient to show that, for any integer $n\geqslant 2$ and any $T_1,\dots,T_n$ in $\mathcal{G}_{cyc}\langle x_1,\dots,x_p\rangle$, one has

$$\mathbb{E}\Big[Z_N(T_1)\dots Z_N(T_n)\Big] \underset{N\to\infty}{\longrightarrow} \sum_{\pi\in\mathcal{PP}(n)} \prod_{\{i_1,i_2\}\in\pi} \rho(T_{i_1},T_{i_2}), \tag{4.6}$$

where $\mathcal{PP}(n)$ is the set of all pair partitions of $\{1,\dots,n\}$. We first show this fact for $n=2$ and then for general $n\geqslant 3$.

4.3.1 Convergence of the covariance

For any cyclic test graph $T=(V,E,\gamma), T'=(V',E',\gamma')$ in $\mathcal{G}_{cyc}\langle x_1,\dots,x_p\rangle$, we set

$$\rho_N(T,T') \;:=\; \mathbb{E}\Big[Z_N(T)Z_N(T')\Big]$$

Then, one has

$$\begin{aligned}\rho_N(T,T') \;=\; N\bigg(&\mathbb{E}\Big[\tau_N^0\big[T(\mathbf{X}_N)\big]\tau_N^0\big[T'(\mathbf{X}_N)\big]\Big]\\ &-\mathbb{E}\Big[\tau_N^0\big[T(\mathbf{X}_N)\big]\Big]\times\mathbb{E}\Big[\tau_N^0\big[T'(\mathbf{X}_N)\big]\Big]\bigg).\end{aligned}$$

By the definition of the injective trace, one has

$$\rho_N(T,T') = \frac{1}{N}\sum_{\substack{\phi:V\to\{1,\dots,N\}\\ \text{injective}}}\sum_{\substack{\phi':V'\to\{1,\dots,N\}\\ \text{injective}}} \mathbb{E}\Big[P_N(\phi)P'_N(\phi')\Big]-\mathbb{E}\Big[P_N(\phi)\Big]\mathbb{E}\Big[P'_N(\phi')\Big],$$

where for any ϕ,ϕ' as in the sums, we have denoted

$$P_N(\phi) := \prod_{e\in E} X_{\gamma(e)}^{(N)}\big(\phi(e)\big),$$
$$P'_N(\phi') := \prod_{e\in E'} X_{\gamma'(e)}^{(N)}\big(\phi'(e)\big).$$

We classify the terms in the sums above in the following way.

$$\rho_N(T,T') = \sum_{\substack{\psi:W\to W' \\ \text{bijective} \\ W\subset V, W'\subset V'}} \frac{1}{N}\sum_{\phi,\phi'} \mathbb{E}\Big[P_N(\phi)P'_N(\phi')\Big] - \mathbb{E}\Big[P_N(\phi)\Big]\mathbb{E}\Big[P'_N(\phi')\Big],$$

where the last sum is over all injective maps $\phi : V \to \{1,\dots,N\}$ and $\phi' : V' \to \{1,\dots,N\}$ such that, for any v in W one has $\phi(v) = \phi'\big(\psi(v)\big)$ and $\phi(V\setminus W)\cap \phi'(V'\setminus W') = \emptyset$. Let $\psi : W \to W'$ be as in the first sum and consider the amalgamated product of test graphs

$$T'' = (V'',E'',\gamma'') = T \underset{\psi:W\to W'}{*} T'.$$

By independence of the entries of the matrices, if T'' is not in $\delta_\sharp(T,T')$, then for any ϕ,ϕ' as above, one has $\mathbb{E}\Big[P_N(\phi)P'_N(\phi')\Big] = \mathbb{E}\Big[P_N(\phi)\Big]\mathbb{E}\Big[P'_N(\phi')\Big]$. Then, we get

$$\rho_N(T,T') = \sum_{T''=T \underset{\psi:W\to W'}{*} T'} \frac{1}{N} \sum_{\substack{\phi'':V''\to\{1,\dots,N\} \\ \text{injective}}} \mathbb{E}\Big[P''_N(\phi'')\Big] - \mathbb{E}\Big[P_N(\phi)\Big]\mathbb{E}\Big[P'_N(\phi')\Big], \tag{4.7}$$

where we have set

$$P''_N(\phi'') := \prod_{e\in E''} X^{(N)}_{\gamma''(e)}\Big(\phi''(e)\Big),$$

and $\phi = \phi''_{|V}$, $\phi' = \phi''_{|W\sqcup(V'\setminus W')}\circ\psi^{-1}$ (the map ψ is extended trivially to a bijection $W\sqcup(V'\setminus W')\to V$). Remark that for any T'' as in the first sum, one has

$$\frac{1}{N}\sum_{\substack{\phi'':V''\to\{1,\dots,N\} \\ \text{injective}}} \mathbb{E}\Big[P''_N(\phi'')\Big] = \mathbb{E}\Big[\tau^0_N\Big[T''(\mathbf{X}_N)\Big]\Big],$$

which tends to $\tau^0[T'']$ by Theorem 4.2.1 (T'' is well a cyclic test graph). It turns out that, for any bijection $\psi : W \to W'$ as in the first sum of (4.7), the term

$$\epsilon_N(\psi) := \frac{1}{N}\sum_{\substack{\phi'':V''\to\{1,\dots,N\} \\ \text{injective}}} \mathbb{E}\Big[P_N(\phi)\Big]\mathbb{E}\Big[P'_N(\phi')\Big]$$

is negligible. To show this fact, we use the same kind of analysis as in our preceding paper.

We write the cyclic test graph $T = (G,\gamma)$ and we consider the graph $\bar G = (\bar V,\bar E)$ obtained from G by forgetting the orientation and the multiplicity of its edges. Let c be a cycle on G that visits exactly one time each edge in the sense of their orientation. The cycle c induces a cycle $\bar c$ on $\bar G$ whose steps are colored by the labels in $\{1,\dots,p\}$ of the edges in T. For any $m = 1,\dots,p$ and $k \geqslant 1$, we denote by $\eta_{m,k}$ the number of edges of $\bar G$ visited by $\bar c$ exactly k times by a step of color m. We write $\bar G' = (\bar V', \bar E')$, $\eta'_{m,k}$ and $\bar G'' = (\bar V'', \bar E'')$, $\eta''_{m,k}$ for the same objects with T' and T'' respectively instead of T.

By the independence of the entries of the matrices, we get that

$$\begin{aligned} \epsilon_N(\psi) &= \frac{1}{N} \sum_{\substack{\phi'':V''\to\{1,\dots,N\}\\ \text{injective}}} \prod_{m=1}^p \prod_{k\geqslant 1} \left(\frac{\int t^k \mathrm{d}\mu_m^{(N)}(t)}{N^{\frac{k}{2}}} \right)^{\eta_{m,k}+\eta'_{m,k}} \\ &\underset{N\to\infty}{\sim} N^{|\bar V''|-1} \prod_{m=1}^p \prod_{k\geqslant 1} \left(\frac{\int t^k \mathrm{d}\mu_m^{(N)}(t)}{N^{\frac{k}{2}}} \right)^{\eta_{m,k}+\eta'_{m,k}} . \end{aligned}$$

Moreover, by the definition of heavy Wigner matrices, one has for any $m = 1,\dots,p$ and $k \geqslant 1$ that

$$\frac{\int t^k \mathrm{d}\mu_m^{(N)}(t)}{N^{\frac{k}{2}-1}} = o(N^\beta),\ \forall \beta > 0.$$

We set

$$B = \sum_{m=1}^p \sum_{k\geqslant 1} \eta_{m,k},\ B' = \sum_{m=1}^p \sum_{k\geqslant 1} \eta'_{m,k},$$

so that

$$\epsilon_N(\psi) = o\Big(N^{|\bar V''|-1-B-B'+\beta}\Big), \forall \beta > 0.$$

Remark that B is the number of edges of $\bar G$ counted with multiplicity with respect to the colors of the steps $\bar c$. Hence, one has $B \geqslant |\bar E|$. Similarly, one has $B' \geqslant |\bar E'|$. Moreover, the relation between the number of edges and the number of vertices in a connected graph tells us that $|V''| \leqslant |\bar E''| + 1$. At last, since the traffic T'' is in $\delta_\sharp(T,T')$, one has $|\bar E''| \leqslant |\bar E| + |\bar E'| - 1$. Hence we get

$$\epsilon_N(\psi) = o\Big(\frac{1}{N^{1-\beta}}\Big), \forall \beta > 0,$$

and so

$$\rho_N(T,T') \underset{N\to\infty}{\longrightarrow} \sum_{T''\in\delta_\sharp(T,T')} \tau^0[T''] = \rho(T,T')$$

as expected.

4.3.2 Proof of (4.6) for $n \geqslant 3$

Let $n \geqslant 3$ and $T_1 = (V_1, E_1, \gamma_1), \dots, T_n = (V_n, E_n, \gamma_n)$ be cyclic test graphs in $\mathcal{G}_{cyc}\langle x_1,\dots,x_p\rangle$. Then, one has

$$\begin{aligned} \mathbb{E}\Big[Z_N(T_1)\dots Z_N(T_n)\Big] &= N^{\frac{n}{2}} \mathbb{E}\Bigg[\prod_{i=1}^n \bigg(\tau_N^0\Big[T_i(\mathbf{X}_N)\Big] - \mathbb{E}\Big[\tau_N^0\Big[T_i(\mathbf{X}_N)\Big]\Big] \bigg)\Bigg] \\ &= \sum_{\pi\in\mathcal{P}(n)} \frac{1}{N^{\frac{n}{2}}} \sum_{\phi_1,\dots,\phi_n} \mathbb{E}\Bigg[\prod_{i=1}^n \bigg(P_N^{(i)}(\phi_i) - \mathbb{E}\Big[P_N^{(i)}(\phi_i)\Big]\bigg)\Bigg], \end{aligned}$$

where

- $\mathcal{P}(n)$ is the set of all partitions of $\{1,\dots,n\}$,
- the second sum is over all injective maps $\phi_i : V_i \to \{1,\dots,N\}$, $i = 1,\dots,n$, such that

$$i \sim_\pi j \Leftrightarrow \phi_i(V_i) \cap \phi_j(V_j) \neq \emptyset,$$

- for any $i = 1,\dots,n$, we have denoted

$$P_N^{(i)}(\phi_i) := \prod_{e\in E_i} X^{(N)}_{\gamma_i(e)}\Big(\phi_i(e)\Big).$$

Let π be in $\mathcal{P}(n)$ and $B = \{i_1,\dots,i_k\}$ one of its block. By the independence of the entries of the matrices, one has

$$\mathbb{E}\Big[Z_N(T_1)\dots Z_N(T_n)\Big] \tag{4.8}$$
$$= \sum_{\pi\in\mathcal{P}(n)} \frac{1}{N^{\frac{n}{2}-|\pi|}} \prod_{B=\{i_1,\dots,i_k\}\in\pi} \frac{1}{N} \sum_{\phi_1,\dots,\phi_k} \mathbb{E}\Bigg[\prod_{j=1}^k \Big(P_N^{(i_j)}(\phi_i) - \mathbb{E}\Big[P_N^{(i_j)}(\phi_i)\Big]\Big)\Bigg],$$

where the last sum is over all injective maps $\phi_j : V_{i_j} \to \{1,\dots,N\}$, such that for any $j_1, j_2 = 1,\dots,k$ one has $\phi_{j_1}(V_{i_{j_1}}) \cap \phi_{j_2}(V_{i_{j_2}}) \neq \emptyset$. By the same analysis as in the previous section, we get that for any π in $\mathcal{P}(n)$ and any $B = \{i_1,\dots,i_k\}$ in π, one has

$$\frac{1}{N}\sum_{\phi_1,\dots,\phi_k} \mathbb{E}\Bigg[\prod_{j=1}^k \Big(P_N^{(i_j)}(\phi_i) - \mathbb{E}\Big[P_N^{(i_j)}(\phi_i)\Big]\Big)\Bigg] = O(1).$$

Moreover, is B has only one element, by the centering this term vanishes. Hence, the only terms that contributes in the first sum of (4.8) are the partition π such that σ_π is a pair partition. Then, one has

$$\mathbb{E}\Big[Z_N(T_1)\dots Z_N(T_n)\Big]$$
$$= \sum_{\pi\in\mathcal{PP}(n)} \prod_{\{i_1,i_2\}\in\pi} \rho_N(T_{i_1},T_{i_2}) + O\Big(\frac{1}{N}\Big).$$

We then get as expected

$$\mathbb{E}\Big[Z_N(T_1)\dots Z_N(T_n)\Big] \underset{N\to\infty}{\longrightarrow} \sum_{\pi\in\mathcal{PP}(n)} \prod_{\{i_1,i_2\}\in\pi} \rho(T_{i_1},T_{i_2}). \tag{4.9}$$

Bibliography

[ABAP09] A. Auffinger, G. Ben Arous et S. Péché : Poisson convergence for the largest eigenvalues of heavy tailed random matrices. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(3):589–610, 2009.

[AGZ10] G. W. Anderson, A. Guionnet et O. Zeitouni : *An Introduction to Random Matrices*, volume 118 de *Cambridge studies in advanced mathematics*. Cambridge University Press, 2010.

[AL07] D. Aldous et R. Lyons : Processes on unimodular random networks. *Electron. J. Probab.*, 12:no. 54, 1454–1508, 2007.

[And] G. W. Anderson : Convergence of the largest singular value of a polynomial in independent wigner matrices. arXiv:1103.4825v1, preprint, http://arxiv.org/abs/1103.4825.

[AO76] C. A. Akemann et P. A. Ostrand : Computing norms in group C^*-algebras. *Amer. J. Math.*, 98(4):1015–1047, 1976.

[Arn67] L. Arnold : On the asymptotic distribution of the eigenvalues of random matrices. *J. Math. Anal. Appl.*, 20:262–268, 1967.

[AS04] D. Aldous et J. M. Steele : The objective method: probabilistic combinatorial optimization and local weak convergence. *In Probability on discrete structures*, volume 110 de *Encyclopaedia Math. Sci.*, pages 1–72. Springer, Berlin, 2004.

[BAG97] G. Ben Arous et A. Guionnet : Large deviations for Wigner's law and Voiculescu's non-commutative entropy. *Probab. Theory Related Fields*, 108(4):517–542, 1997.

[BAG08] G. Ben Arous et A. Guionnet : The spectrum of heavy tailed random matrices. *Comm. Math. Phys.*, 278(3):715–751, 2008.

[Bai99] Z. D. Bai : Methodologies in spectral analysis of large-dimensional random matrices, a review. *Statist. Sinica*, 9(3):611–677, 1999. With comments by G. J. Rodgers and Jack W. Silverstein; and a rejoinder by the author.

[BC94] J. Bouchaud et P Cizeau : Theory of lévy matrices. *Phys. Rev. E*, 50:1810Ð1822, 1994.

[BCC10] C. Bordenave, P. Caputo et D. Chafaï : Spectrum of large random reversible Markov chains: two examples. *ALEA Lat. Am. J. Probab. Math. Stat.*, 7:41–64, 2010.

[BCC11] C. Bordenave, P. Caputo et D. Chafaï : Spectrum of large random reversible markov chains: heavy tailed weights on the complete graph. *Ann. Probab.*, 39:1544–1590, 2011.

[BCN] S. Belinschi, B. Collins et I. Nechita : Laws of large numbers for eigenvectors and eigenvalues associated to random subspaces in a tensor product. arXiv:1008.3099v2, preprint, http://arxiv.org/abs/1008.3099v2.

[BDG09] S. Belinschi, A. Dembo et A. Guionnet : Spectral measure of heavy tailed band and covariance random matrices. *Comm. Math. Phys.*, 289(3):1023–1055, 2009.

[BO08] N. P. Brown et N. Ozawa : *C^*-algebras and finite-dimensional approximations*, volume 88 de *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2008.

[BS98] Z. D. Bai et J. W. Silverstein : No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.*, 26(1):316–345, 1998.

[BS99] Z. D. Bai et J. W. Silverstein : Exact separation of eigenvalues of large-dimensional sample covariance matrices. *Ann. Probab.*, 27(3): 1536–1555, 1999.

[BS01] I. Benjamini et O. Schramm : Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 13 pp. (electronic), 2001.

[BSY88] Z. D. Bai, J. W. Silverstein et Y. Q. Yin : A note on the largest eigenvalue of a large-dimensional sample covariance matrix. *J. Multivariate Anal.*, 26(2):166–168, 1988.

[BY88] Z. D. Bai et Y. Q. Yin : Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix. *Ann. Probab.*, 16(4):1729–1741, 1988.

[BYK86] Z. D. Bai, Y. Q. Yin et P. R. Krishnaiah : On limiting spectral distribution of product of two random matrices when the underlying distribution is isotropic. *J. Multivariate Anal.*, 19(1):189–200, 1986.

[CC04] M. Capitaine et M. Casalis : Asymptotic freeness by generalized moments for Gaussian and Wishart matrices. Application to beta random matrices. *Indiana Univ. Math. J.*, 53(2):397–431, 2004.

[CDM07] M. Capitaine et C. Donati-Martin : Strong asymptotic freeness for Wigner and Wishart matrices. *Indiana Univ. Math. J.*, 56(2): 767–803, 2007.

[CDS10] R. Couillet, M. Debbah et J. Silverstein : A deterministic equivalent for the analysis of mimo multiple access channels, 2010.

[CHD10] R. Couillet, J. Hoydis et M. Debbah : A deterministic equivalent approach to the performance analysis of isometric random precoded systems, 2010.

[Che82a] L. H. Y. Chen : An inequality for the multivariate normal distribution. *J. Multivariate Anal.*, 12(2):306–315, 1982.

[Che82b] L. H. Y. Chen : An inequality for the multivariate normal distribution. *J. Multivariate Anal.*, 12(2):306–315, 1982.

[Col03] B. Collins : Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability. *Int. Math. Res. Not.*, (17):953–982, 2003.

[Con00] J. B. Conway : *A course in operator theory*, volume 21 de *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2000.

[Def10] M. Defosseux : Orbit measures, random matrix theory and interlaced determinantal processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(1):209–249, 2010.

[dlS09] M. de la Salle : Strong Haagerup inequalities with operator coefficients. *J. Funct. Anal.*, 257(12):3968–4002, 2009.

[Dyk93] K. Dykema : On certain free product factors via an extended matrix model. *J. Funct. Anal.*, 112(1):31–60, 1993.

[FK81] Z. Füredi et J. Komlós : The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3):233–241, 1981.

[Gem80] S. Geman : A limit theorem for the norm of random matrices. *Ann. Probab.*, 8(2):252–261, 1980.

[GKZ] A. Guionnet, M. Krishnapur et O. Zeitouni : The single-ring theorem. arXiv:0909.2214v1, preprint, http://arxiv4.library.cornell.edu/abs/0909.2214.

[GS77] U. Grenander et J. W. Silverstein : Spectral analysis of networks with random topologies. *SIAM J. Appl. Math.*, 32(2):499–519, 1977.

[Gui09] A. Guionnet : *Large random matrices: lectures on macroscopic asymptotics*, volume 1957 de *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009. Lectures from the 36th Probability Summer School held in Saint-Flour, 2006.

[Haa79] U. Haagerup : An example of a nonnuclear C^*-algebra, which has the metric approximation property. *Invent. Math.*, 50(3):279–293, 1978/79.

[Ham21] H. Hamburger : Über eine Erweiterung des Stieltjesschen Momentenproblems. *Math. Ann.*, 82(3-4):168–187, 1921.

[HLN07] W. Hachem, P. Loubaton et J. Najim : Deterministic equivalents for certain functionals of large random matrices. *Ann. Appl. Probab.*, 17(3):875–930, 2007.

[HP00] F. Hiai et D. Petz : Asymptotic freeness almost everywhere for random matrices. *Acta Sci. Math. (Szeged)*, 66(3-4):809–834, 2000.

[HT05] U. Haagerup et S. Thorbjørnsen : A new application of random matrices: $\mathrm{Ext}(C^*_{\mathrm{red}}(\mathbb{F}_2))$ is not a group. *Ann. of Math. (2)*, 162(2): 711–775, 2005.

[Jon82] D. Jonsson : Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.*, 12(1):1–38, 1982.

[Jon85] D. Jonsson : On the largest eigenvalue of a sample covariance matrix. *In Multivariate analysis VI (Pittsburgh, Pa., 1983)*, pages 327–333. North-Holland, Amsterdam, 1985.

[Juh81] F. Juhász : On the spectrum of a random graph. *In Algebraic methods in graph theory, Vol. I, II (Szeged, 1978)*, volume 25 de *Colloq. Math. Soc. János Bolyai*, pages 313–316. North-Holland, Amsterdam, 1981.

[Kes59] H. Kesten : Symmetric random walks on groups. *Trans. Amer. Math. Soc.*, 92:336–354, 1959.

[KS07] T. Kemp et R. Speicher : Strong Haagerup inequalities for free $\mathcal{R}$-diagonal elements. *J. Funct. Anal.*, 251(1):141–173, 2007.

[KSV04] O. Khorunzhy, M. Shcherbina et V. Vengerovsky : Eigenvalue distribution of large weighted random graphs. *J. Math. Phys.*, 45(4):1648–1672, 2004.

[Leh99] F. Lehner : Computing norms of free operators with matrix coefficients. *Amer. J. Math.*, 121(3):453–486, 1999.

[LS03] E. G. Larsson et P. Stoica : *Space-time block coding for wireless communications*. Cambridge University Press, Cambridge, 2003.

[Mal11] C. Male : The norm of polynomials in large random and deterministic matrices. *Probab. Theory Related Fields*, 2011.

[Meh04] M. L. Mehta : *Random matrices*, volume 142 de *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, third édition, 2004.

[Met] A. Metcalfe : Universality properties of gelfand-tsetlin patterns. arXiv:1105.1272v1, preprint, http://arxiv.org/abs/1105.1272.

[MP67] V. A. Marčenko et L. A. Pastur : Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, 72 (114):507–536, 1967.

[MS06] J. A. Mingo et R. Speicher : Second order freeness and fluctuations of random matrices. I. Gaussian and Wishart matrices and cyclic Fock spaces. *J. Funct. Anal.*, 235(1):226–270, 2006.

[NS06] A. Nica et R. Speicher : *Lectures on the combinatorics of free probability*, volume 335 de *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2006.

[PBL05] M. Potters, J. P. Bouchaud et L. Laloux : Financial applications of random matrix theory: old laces and new pieces. *Acta Phys. Polon. B*, 36(9):2767–2784, 2005.

[Pim97] M. V. PIMSNER : A class of C^*-algebras generalizing both Cuntz-Krieger algebras and crossed products by **Z**. *In Free probability theory (Waterloo, ON, 1995)*, volume 12 de *Fields Inst. Commun.*, pages 189–212. Amer. Math. Soc., Providence, RI, 1997.

[Pis03] G. PISIER : *Introduction to operator space theory*, volume 294 de *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2003.

[PS09] D. PAUL et J. W. SILVERSTEIN : No eigenvalues outside the support of the limiting empirical spectral distribution of a separable covariance matrix. *J. Multivariate Anal.*, 100(1):37–57, 2009.

[RD08] Ø. RYAN et M. DEBBAH : Channel capacity estimation using free-probability theory. *IEEE Trans. Signal Process.*, 56(11):5654–5667, 2008.

[RFOBS08] R. RASHIDI FAR, T. ORABY, W. BRYC et R. SPEICHER : On slow-fading MIMO systems with nonseparable correlation. *IEEE Trans. Inform. Theory*, 54(2):544–553, 2008.

[Rya98] Ø. RYAN : On the limit distributions of random matrices with independent or free entries. *Comm. Math. Phys.*, 193(3):595–626, 1998.

[Sch05] H. SCHULTZ : Non-commutative polynomials of independent Gaussian random matrices. The real and symplectic cases. *Probab. Theory Related Fields*, 131(2):261–309, 2005.

[Shl98] D. SHLYAKHTENKO : Some applications of freeness with amalgamation. *J. Reine Angew. Math.*, 500:191–212, 1998.

[Sil85] J. W. SILVERSTEIN : The smallest eigenvalue of a large-dimensional Wishart matrix. *Ann. Probab.*, 13(4):1364–1368, 1985.

[Sil89] J. W. SILVERSTEIN : On the weak limit of the largest eigenvalue of a large-dimensional sample covariance matrix. *J. Multivariate Anal.*, 30(2):307–311, 1989.

[Sos04] Alexander SOSHNIKOV : Poisson statistics for the largest eigenvalues of Wigner random matrices with heavy tails. *Electron. Comm. Probab.*, 9:82–91 (electronic), 2004.

[SSH05] A. SKUPCH, D. SEETHALER et F. HLAWATSCH : Free probability based capacity calculation for MIMO channels with transmit or receive correlation. *In Proceedings of the 2005 International Conference on Wireless Networks Communications and Mobile Computing*, pages 1041–1046, Maui, Hawaii, USA, June 2005.

[Tho00] S. THORBJØRNSEN : Mixed moments of Voiculescu's Gaussian random matrices. *J. Funct. Anal.*, 176(2):213–246, 2000.

[TV04] A. M. TULINO et S. VERDÚ : *Random matrices and wireless communications*, volume vol. 1 de *Fundations and Trends in Communications and Information Theory*. Now Publishers Inc., 2004.

[VO05] G. VERNIZZI et H. ORLAND : Large-N random matrices for RNA folding. *Acta Phys. Polon. B*, 36(9):2821–2827, 2005.

[Voi85] D. VOICULESCU : Symmetries of some reduced free product C^*-algebras. *In Operator algebras and their connections with topology and ergodic theory (Buşteni, 1983)*, volume 1132 de *Lecture Notes in Math.*, pages 556–588. Springer, Berlin, 1985.

[Voi91] D. VOICULESCU : Limit laws for random matrices and free products. *Invent. Math.*, 104(1):201–220, 1991.

[Voi95a] D. VOICULESCU : Free probability theory: random matrices and von Neumann algebras. *In Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)*, pages 227–241, Basel, 1995. Birkhäuser.

[Voi95b] D. VOICULESCU : Operations on certain non-commutative operator-valued random variables. *Astérisque*, .(232):243–275, 1995. Recent advances in operator algebras (Orléans, 1992).

[Voi96] D. VOICULESCU : The analogues of entropy and of Fisher's information measure in free probability theory. III. The absence of Cartan subalgebras. *Geom. Funct. Anal.*, 6(1):172–199, 1996.

[Voi98] D. VOICULESCU : A strengthened asymptotic freeness result for random matrices with applications to free entropy. *Internat. Math. Res. Notices*, .(1):41–63, 1998.

[Wac78] K. W. WACHTER : The strong limits of random matrix spectra for sample matrices of independent elements. *Ann. Probability*, 6(1):1–18, 1978.

[Wig58] E. P. WIGNER : On the distribution of the roots of certain symmetric matrices. *Ann. of Math. (2)*, 67:325–327, 1958.

[Wis28] J. WISHART : The generalized product moment distribution in samples from a normal multivariate population. *Biometrika*, 20 A: 32–56, 1928.

[WZCM09] L. WANG, B. ZHENG, J. CUI et Q. MENG : Cooperative mimo spectrum sensing using free probability theory. *In Proceedings of the 5th International Conference on Wireless communications, networking and mobile computing*, WiCOM'09, pages 1455–1458, Piscataway, NJ, USA, 2009. IEEE Press.

[YBK88] Y. Q. YIN, Z. D. BAI et P. R. KRISHNAIAH : On the limit of the largest eigenvalue of the large-dimensional sample covariance matrix. *Probab. Theory Related Fields*, 78(4):509–521, 1988.

[Yin86] Y. Q. YIN : Limiting spectral distribution for a class of random matrices. *J. Multivariate Anal.*, 20(1):50–68, 1986.

[YK83] Y. Q. YIN et P. R. KRISHNAIAH : A limit theorem for the eigenvalues of product of two random matrices. *J. Multivariate Anal.*, 13(4):489–507, 1983.

[Zak06] I. ZAKHAREVICH : A generalization of Wigner's law. *Comm. Math. Phys.*, 268(2):403–414, 2006.