Isolated eigenvalues of non Hermitian random matrices

Jean Rochet

UNIVERSITÉ PARIS DESCARTES
Laboratoire MAP5 UMR CNRS 8145
École doctorale 386 : Sciences Mathématiques de Paris Centre

THÈSE
Pour obtenir le grade de
DOCTEUR de l'UNIVERSITÉ PARIS DESCARTES

Spécialité : Mathématiques appliquées

Présentée par

Jean ROCHET

Isolated eigenvalues of non Hermitian random matrices

Soutenue le 15 juin 2016 devant le jury composé de

Florent BENAYCH-GEORGES	Université Paris Descartes	Directeur de thèse
Catherine DONATI-MARTIN	Université Versailles Saint-Quentin-en-Yvelines	Rapporteuse
Sean O'ROURKE	University of Colorado Boulder	Rapporteur
Djalil CHAFAÏ	Université Paris Dauphine	Examinateur
Jamal NAJIM	Université Paris-Est Marne la Vallée	Examinateur
Sandrine PÉCHÉ	Université Paris Diderot	Examinatrice

À Nana

RÉSUMÉ

Dans cette thèse, il est question de *spiked models* pour des matrices aléatoires non-hermitiennes. Plus précisément, on considère des matrices de type $\mathbf{A}+\mathbf{P}$, tel que le rang de $\mathbf{P}$ reste borné indépendamment de la taille de la matrice qui tend vers l'infini, et tel que $\mathbf{A}$ est une matrice aléatoire non-hermitienne. Tout d'abord, on prouve que dans le cas où la matrice $\mathbf{P}$ possède des valeurs propres hors du *bulk*, quelques valeurs propres de $\mathbf{A}+\mathbf{P}$ (appelées *outliers*) apparaissent loin de celui-ci. Ensuite, on regarde les fluctuations des *outliers* de $\mathbf{A}$ autour de leurs limites et on montre que celles-ci ont la même distribution que les valeurs propres d'une certaine matrice aléatoire de taille finie. Ce genre de phénomène avait déjà été observé pour des modèles hermitiens. De manière inattendue, on montre que les vitesses de convergence des *outliers* varient (en fonction de la Réduite de Jordan de $\mathbf{P}$) et que des corrélations peuvent apparaître entre des *outliers* situés à une distance macroscopique l'un de l'autre. Le premier modèle de matrices non-hermitiennes que l'on considère provient du théorème du *Single Ring* que l'on doit à Guionnet, Krishnapur et Zeitouni. Un autre modèle étudié est celui des matrices dites "presque" hermitiennes : c'est-à-dire lorsque $\mathbf{A}$ est hermitienne mais $\mathbf{P}$ ne l'est pas. Enfin, on regarde aussi les *outliers* pour des matrices Elliptiques Gaussiennes. Cette thèse traite aussi de la convergence en loi de variables aléatoires du type $\mathrm{Tr}(f(\mathbf{A})\mathbf{M})$ où $\mathbf{A}$ est une matrice du théorème du *Single Ring* et f est une fonction holomorphe sur un voisinage du *bulk* et la norme de Frobenius de $\mathbf{M}$ est de l'ordre de $\sqrt{N}$. En corollaire de ce résultat, on obtient des théorèmes type "Centrale Limite" pour les statistiques linéaires de $\mathbf{A}$ (pour des fonctions tests holomorphes) et des projections de rang finies de la matrice $\mathbf{A}$ (comme par exemple les entrées de la matrice).

ABSTRACT

This thesis is about spiked models of non Hermitian random matrices. More specifically, we consider matrices of the type $\mathbf{A}+\mathbf{P}$, where the rank of $\mathbf{P}$ stays bounded as the dimension goes to infinity and where the matrix $\mathbf{A}$ is a non Hermitian random matrix. We first prove that if $\mathbf{P}$ has some eigenvalues outside the bulk, then $\mathbf{A}+\mathbf{P}$ has some eigenvalues (called outliers) away from the bulk. Then, we study the fluctuations of the outliers of $\mathbf{A}$ around their limit and prove that they are distributed as the eigenvalues of some finite dimensional random matrices. Such facts had already been noticed for Hermitian models. More surprising facts are that outliers can here have very various rates of convergence to their limits (depending on the Jordan Canonical Form of $\mathbf{P}$) and that some correlations can appear between outliers at a macroscopic distance from each other. The first non Hermitian model studied comes from the Single Ring Theorem due to Guionnet, Krishnapur and Zeitouni. Then we investigated spiked models for nearly Hermitian random matrices : where $\mathbf{A}$ is Hermitian but $\mathbf{P}$ isn't. At last, we studied the outliers of Gaussian Elliptic random matrices. This thesis also investigates the convergence in distribution of random variables of the type $\mathrm{Tr}(f(\mathbf{A})\mathbf{M})$ where $\mathbf{A}$ is a matrix from the Single Ring Theorem and f is analytic on a neighborhood of the bulk and the Frobenius norm of $\mathbf{M}$ has order $\sqrt{N}$. As corollaries, we obtain central limit theorems for linear spectral statistics of $\mathbf{A}$ (for analytic test functions) and for finite rank projections of $f(\mathbf{A})$ (like matrix entries).

Remerciements [1]

Tout d'abord, un très grand merci à Florent Benaych-Georges pour m'avoir encadré pendant cette thèse. Travailler avec toi a été un véritable plaisir et je n'aurais sincèrement pas pu imaginer meilleur directeur de thèse. J'avoue que cela a fortement contribué à rendre plus difficile ma décision de choisir entre l'enseignement et la recherche.

J'adresse aussi mes plus sincères remerciements à Catherine Donati-Martin et Sean O'Rourke qui ont bien voulu accepté le long travail que constitue le fait de rapporter cette thèse. Je remercie aussi Djalil Chafaï, Jamal Najim et Sandrine Péché de me faire l'honneur de compléter ce jury.

Je tiens à dire que j'ai beaucoup apprécié ces quatre années passées au laboratoire du MAP5. Je remercie les deux directrices successives du laboratoire, Annie Raoult et Fabienne Comte, ainsi que la directrice de l'UFR Math-Info, Christine Graffigne. Je remercie aussi les équipes administratives et informatiques pour leurs aides diverses et variées : merci notamment à Christophe, Marie-Hélène, Azedine, Marie, Arnaud, Thierry, Isabelle.

Un grand merci aux thésards (et affiliés) du MAP5 qui ont largement contribué à rendre ces quatre ans très agréables (tout en restant studieuses, bien entendu). Je n'oublierai pas tous ces débats endiablés sur tout et n'importe quoi (et peut-être plus souvent sur la Grèce antique), qui rendaient les pauses café du midi plus animées et difficiles à terminer. J'ai une pensée particulière pour mes ex-co-organisatrices du célèbre GTTJD avec qui j'ai vécu tant d'aventures ! Merci à Charlotte D, Christèle et Gwen, sans vous José n'aurait jamais été José ! Merci aussi à Docteur Leclair qui a été mon mentor et mon modèle et surtout qui m'a initié à l'art du Karaoké. Merci aussi à mon petit frère Alkéos, avec qui, parler de matrices (ou autre) a toujours été enrichissant. Un grand merci aussi à tous les autres thésards (anciens et actuels) : Rémy, Rebecca, Laurent, José, Andréa, Claire, Thomas B, Gaëlle, Ronan, Christophe & Anne-Claire, Julie, Mélina, Maud, Kévin, Charlotte L, Loïc, Anne-Sophie, Pierre-Alexandre, Alasdair, Thomas P, Léo, Angélina, Samuel, Pierre, Sonia, Fabien.

Un grand merci aussi à l'équipe de proba du MAP5 et aux autres équipes : Merci notamment à Flora, Avner, Etienne, Thierry, Mireille, Maya, Manon, Julie, Sylvain, Anne, Jean-Claude, Joan, Bérénice, Raphaël, Eric, Camille, Sébastien, Lionel, Edoardo et Ellen.

[1] Pour ne pas créer une quelconque impression de hiérarchisation des personnes, toute liste a été classée par ordre alphabétique des noms de famille.

Je tiens aussi à remercier tous les membres du groupe de travail *Méthodes pour Matrices Aléatoires* et à celles et ceux qui ont bien voulu nous faire l'honneur de nous faire un mini-cours. J'ai toujours beaucoup apprécié nos sessions très instructives qui ont toujours l'occasion de passer des bons moments entre jeunes matriciens. Pourvu que ça dure ! Merci notamment à Marwa, Raphaël (à qui l'ont doit cette bonne idée!), Sandrine, Maxime, Adrien, Benjamin et Edouard. Je (re)-remercie d'ailleurs au passage Camille Male pour l'organisation du MEGA.

Enfin, je remercie chaleureusement ma famille et mes amis pour leur soutien, même dans les moments où je n'en avais pas besoin. Je remercie en particulier Léon et Suzanne qui m'ont toujours soutenu et incité à repousser mes propres limites. Bien entendu, un énorme merci à Elise qui a toujours été ma muse des mathématiques.

Dans le cas où vous estimez que vous deviez être remercié(e) mais que vous n'apparaissez pas dans les présents remerciement, ou bien que vous souhaitez être remercié(e) d'avantage, veuillez envoyer un mail à

remerciements.these.jrochet@gmail.com

Contents

Notation

Here are several notations used in this thesis.

- Bold letters $\mathbf{a}, \mathbf{A}$ always designate a vector or a matrix, whereas regular letters a, A stand for a real or complex number.
- $\mathbf{A}^*$ stands for the adjoint of matrix $\mathbf{A}$.
- $\mathscr{M}_N(\mathbb{R})$ (resp. $\mathscr{M}_N(\mathbb{C})$) is the set of square matrices of size $N \times N$ with real (resp. complex) entries.
- $\mathscr{M}_{p,q}(\mathbb{R})$ (resp. $\mathscr{M}_{p,q}(\mathbb{C})$) is the set of rectangular matrices with p rows and q columns with real (resp. complex) entries.
- For any matrix $\mathbf{A} \in \mathscr{M}_N(\mathbb{C})$ with eigenvalues $\lambda_1, \ldots, \lambda_N$, we shall denote by $\mu_{\mathbf{A}} := \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i}$, the *empirical spectral measure* of $\mathbf{A}$.
- For any real measure μ, we shall denote by $G_\mu(z) := \int \frac{1}{z-x}\mu(dx)$, the *Cauchy Transform* of μ.

Introduction

The first model ever studied in the random matrix theory was sample covariance matrices. Given a sample $X_1,\dots,X_N \in \mathbb{R}^p$ drawn from a centered law, the $p\times p$ matrix $S=[X_1\cdots X_N][X_1\cdots X_N]^*$ plays a key role in statistical analysis. We know, from the law of large numbers that $\frac{1}{N}S$ is a good approximation of $\Sigma=\mathbb{E}[X_1X_1^*]$, the population covariance matrix, as long as p is fixed and N goes to infinity. However, many problems involve high dimensional data which means that p may be on the same order than N so that the covariance matrix Σ is harder to catch. The study of the *largest* eigenvalues, as for example in the statistical method of principal components analysis, may give interesting information about Σ, especially in the almost null case, which means that Σ is not so far from the identity matrix. Johnstone, in 2001 (see [57]), proposed, what he called, a *spiked model* where the covariance matrix has all its eigenvalues equal to one, except a finite number of them (the *spikes*). Then, Baik, Ben Arous and Péché, in 2005, pointed out an interesting phenomenon of phase transition (called the *BBP phase transition*) where the asymptotic behavior of the largest eigenvalues depends on the spikes. Afterwards, this phenomenon has been widely studied for Hermitian models in [77, 43, 28, 29, 19, 20, 16, 17, 30, 61, 62]. Also, Tao, Benaych-Georges and myself, O'Rourke, Renfrew, Bordenave and Capitaine studied non-Hermitian models: in [92, 21, 73, 24, 85, 23] were considered spiked i.i.d., isotropic or elliptic random matrices, in the following sense: the authors added a perturbing matrix (with a finite number of non-zero eigenvalues which are the spikes) and proved that for large enough spikes, some *outliers* (eigenvalues at macroscopic distance of the *bulk*) appear at precise positions.

This thesis investigates mostly the asymptotic behavior of such isolated eigenvalues (outliers) of additive deformations for non Hermitian models (see Chapters I,III and IV). The challenge in the study of such models is that, due to the non Hermitian structure of the perturbation matrix, severals problems occur. First, the perturbation matrix is no longer diagonalizable, which means that we have to consider the Jordan Canonical Form of the matrix to truly understand the fluctuations : surprisingly, we proved that the converge rate of the outliers is not necessarily $\sqrt{N}$ as in the Hermitian case, but depends on the size of the Jordan block associated to the eigenvalue. Also, the outliers tend to locate around their limit at the vertices of a regular polygon. At last, unlike in the Hermitian case, the eigenvectors of the perturbation matrix cannot be supposed orthogonal : it leads to potential correlations between fluctuations around outliers even when are at a macroscopic distance from each other.

The first model of random matrices we study (see Chapter I), that we shall call the *Single Ring Theorem model*, is a non Hermitian model which is isotropic in the sense that, seen as a linear map, its distribution does not depend on the orthonormal bases chosen to represent its transformation matrix. The "Single Ring Theorem", firstly mentioned and partially proved by the Physicists Feinberg and Zee in [42] and completely proved by Guionnet, Krishnapur and Zeitouni in [49], was named after the fact that the eigenvalues of such matrices tend to concentrate on an unique ring.

The second model studied (see Chapters III), called sometimes *nearly Hermitian random matrices*, consists in Hermitian random matrices deformed by an additive non-Hermitian matrix with possibly complex eigenvalues. This model appears in the physics literature, especially in the case where the perturbation matrix is skew-Hermitian, i.e. $\mathbf{P}^* = -\mathbf{P}$ (see [44, 45, 46, 47]).

The last model we considered is the Gaussian Elliptic Ensemble (see Chapter IV) which is some sort of intermediate model between the *Gaussian Unitary Ensemble* and the *Ginibre Ensemble*. The generic location of the outliers has already been studied for elliptic random matrices (see [73]), but the authors did not consider the fluctuations.

This thesis also studies the asymptotic behavior of spectral linear statistics for random matrices from the Single Ring Theorem (see Chapter II). At last, matrix entries of random matrices whose distribution is invariant by unitary conjugation are also studied in a quite general framework, with application to fluctuations of outliers of elliptic matrices (see section 4.2.5).

0.1 The "Single Ring Theorem" Model

0.1.1 Singular values of a matrix

We remind here the definition of the singular values of matrix

Definition 0.1.1 (Singular values)**.** Let $\mathbf{A}$ be $N \times N$ matrix. The *singular values* $s_1, \ldots, s_N$ of $\mathbf{A}$ the eigenvalues of the positive-semidefinite Hermitian matrix $\sqrt{\mathbf{A}\mathbf{A}^*}$.

The positive-semidefinite Hermitian matrix $\sqrt{\mathbf{A}\mathbf{A}^*}$ is also the "Hermitian" part of the polar decomposition of $\mathbf{A}$, i.e. $\mathbf{A}$ can be written as a unique product of an unitary matrix $\mathbf{U}$ and a positive-semidefinite Hermitian matrix $\mathbf{H}$

$$\mathbf{A} = \mathbf{U}\mathbf{H}.$$

Hence, a geometric interpretation of the singular values of $\mathbf{A}$ is to see them as the values of the semiaxes of the ellipsoid

$$\mathscr{E}_{\mathbf{A}} := \{\mathbf{A}\mathbf{v}, \ \ \|\mathbf{v}\| = 1\}.$$

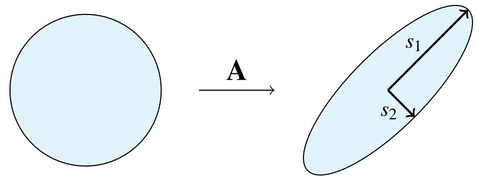


Figure 1: Singular values s_1 and s_2 of a matrix $\mathbf{A} \in \mathbb{R}^{2\times 2}$.

Moreover, unlike the eigenvalues, the singular values are a good criterion for the distance to the null matrix. In fact, some of the usual matrix norms can be defined thanks to the singular values.

$$\|\mathbf{A}\|_2 \ := \ \sqrt{\operatorname{Tr}\mathbf{A}\mathbf{A}^*} \ = \ \sqrt{\sum_{i=1}^{N} s_i^2}, \qquad \|\mathbf{A}\|_{\mathrm{op}} \ := \ \max_{\|\mathbf{v}\|_2=1} \|\mathbf{A}\mathbf{v}\|_2 \ = \ \max_{i=1}^{N} s_i.$$

0.1.2 Definition of the model

Let $\mathbf{A}$ be a $N \times N$ random matrix such that

$$\mathbf{A} \ := \ \mathbf{U} \begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_N \end{pmatrix} \mathbf{V} \tag{1}$$

where

- $\mathbf{U}$, $\mathbf{V}$ and the s_i's are independent,
- $\mathbf{U}$ and $\mathbf{V}$ are Haar-distributed,
- $(s_1, \ldots, s_N)$ is a random N-tuple of non negative real numbers.

Note that $\mathbf{A}$ is *isotropic* in the sense that the distribution of $\mathbf{A}$ is invariant by left and right multiplication by any unitary matrix (independent from $\mathbf{A}$) due to the properties of the Haar measure. Conversely, any random matrix whose distribution satisfies such invariant condition is of the form (1). Actually, one can see the distribution of $\mathbf{A}$ as : $\mathbf{A}$ is uniformly distributed among the matrices whose singular values are $(s_1, \ldots, s_N)$.

0.1.3 Examples

We can already give some examples.

- A Haar-distributed matrix fulfills the previous condition with $(s_1, \ldots, s_N) \sim \delta_1^{\otimes N}$.

- Another important example is the *Ginibre ensemble*, which is $N \times N$ random matrices $\mathbf{G}$ with i.i.d. entries whose distribution are centered Gaussian variables with variance $1/N$. It is easy to show that for any unitary matrix $\mathbf{U}$ (independent from $\mathbf{G}$), one has

$$\mathbf{G} \overset{\text{(d)}}{=} \mathbf{U}\mathbf{G} \overset{\text{(d)}}{=} \mathbf{G}\mathbf{U}.$$

- More generally, random matrices $\mathbf{A}_N$ distributed according to the law

$$\frac{1}{Z_N} \exp\left(-N \operatorname{Tr} V(\mathbf{X}\mathbf{X}^*)\right) d\mathbf{X},$$

where $d\mathbf{X}$ is the Lebesgue measure of the $N \times N$ complex matrices set, V is a polynomial with positive leading coefficient and Z_N is a normalization constant. One can notice that $V(x) = \frac{x}{2}$ gives the renormalized *Ginibre matrices*.

0.1.4 The Single Ring Theorem

One distinctive feature of such matrices is that their eigenvalues tend to spread over a single annulus centered in the origin (see Figures 3.11(a), 3.11(b), 2(c) and 2(d)). This result, partially proved by Feinberg and Zee in [42], is due to Guionnet, Krishnapur and Zeitouni ([49]) where they made the following assumptions. For all $N \geq 1$, let $\mathbf{A}_N$ be a $N \times N$ random matrix such that

$$\mathbf{A}_N \;:=\; \mathbf{U}_N \mathbf{T}_N \mathbf{V}_N,$$

where $\mathbf{T}_N = \operatorname{diag}(s_1, \ldots, s_N)$ is a non-negative diagonal matrix and $\mathbf{U}_N$ and $\mathbf{V}_N$ are both Haar-distributed, independent from each other and independent from the s_i's. Then we make additional assumptions.

Hypothesis 0.1.2. The *empirical spectral measure* of $\mathbf{T}_N$

$$\mu_{\mathbf{T}_N} \;:=\; \frac{1}{N} \sum_{i=1}^{N} \delta_{s_i}$$

converges in probability to a compactly supported measure ν.

Hypothesis 0.1.3. There exists $M > 0$, such that $\mathbb{P}(\|\mathbf{T}_N\|_{\text{op}} > M) \longrightarrow 0$,

Hypothesis 0.1.4.

$$\operatorname{Im}(z) \;>\; N^{-\kappa} \;\Longrightarrow\; \left|\operatorname{Im}\left(G_{\mu_{\mathbf{T}_N}}(z)\right)\right| \;\leq\; \frac{1}{\kappa},$$

where, for any measure Θ, G_Θ designates its *Cauchy Transform*

$$G_\Theta(z) \;:=\; \int \frac{1}{z-x} \Theta(dx).$$

Remark 0.1.5. There is actually another assumption in the *Single Ring Theorem* [49], but Rudelson and Vershynin showed in [86] that it was unnecessary. In [10], Basak and Dembo also weakened the hypotheses of the Single Ring Theorem (roughly allowing Hypothesis 0.1.4 not to hold on a small enough set, so that ν is allowed to have some atoms).

Theorem 0.1.6 (Single Ring Theorem, [49])**.** *If $\mathbf{A}_N$ is a sequence of random matrices satisfying the previous hypothesis, then*

1. *The* empirical spectral measure *of $\mathbf{A}_N$ converges in probability to a deterministic probability measure μ.*
2. *The measure μ possesses a radially-symmetric density $\rho(z)$ with respect to the Lebesgue measure on $\mathbb{C}$ depending only on ν and supported on a single ring $\{z \in \mathbb{C},\ a \leq z \leq b\}$ where*

$$a := \frac{1}{\sqrt{\int x^{-2}\nu(dx)}} \quad ; \quad b := \sqrt{\int x^2 \nu(dx)}.$$

Remark 0.1.7. With the convention $1/\infty = 0$, one can have $a = 0$ so that the support is a disk (see for example the Ginibre case figure 3.11(b)). Also, if ν is a Dirac measure, we have $a = b$ so that the support is a circle (see for example the Haar-distributed case figure 3.11(a)).

Remark 0.1.8. One should notice that the support of μ is a *single ring* and cannot consist in two concentric annuli for example, which means that the support is connected even if it is not the cas for the one of ν (see figure 2(d)).

The previous theorem gives us the global behavior of the spectrum of $\mathbf{A}_N$ but allows, for example, a finite number of eigenvalues to stay away from the *bulk* (i.e. the support of limit distribution). Guionnet and Zeitouni proved in 2012 ([50]) that the *Single Ring* model doesn't have these particular kind of eigenvalue, usually called *outliers*. More precisely, they showed the following result.

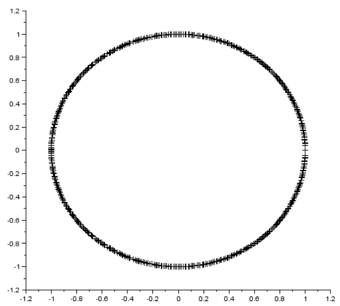
Theorem 0.1.9 (Convergence of the support in the Single Ring Theorem, [50])**.** *As N goes to infinity,*

$$\max_{\lambda \in \mathrm{Sp}(\mathbf{A}_N)} |\lambda| \xrightarrow{(\mathbb{P})} b, \qquad \min_{\lambda \in \mathrm{Sp}(\mathbf{A}_N)} |\lambda| \xrightarrow{(\mathbb{P})} a.$$

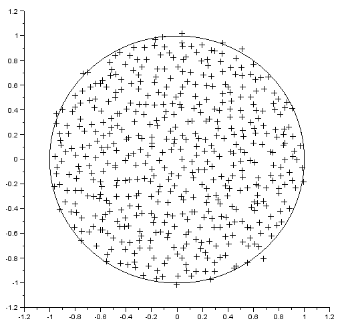
0.2 The Gaussian Elliptic Ensemble

Matrices from the *Gaussian elliptic ensemble*, first introduced in [89], can be defined as follows.

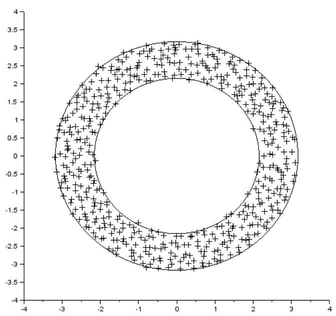
Definition 0.2.1. A *Gaussian elliptic matrix of parameter* $\rho \in [-1,1]$ is a random matrix $\mathbf{Y} = [y_{ij}]_{i,j=1}^N$ such that



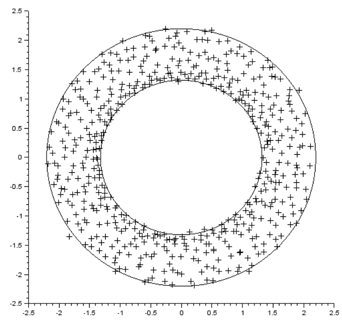
(a) Spectrum of a 500×500 Haar-distributed matrix.



(b) Spectrum of a 500×500 Ginibre matrix.



(c) Spectrum of a 500×500 random matrix of the form $\sim \mathbf{U}\,\mathrm{diag}(s_1, \ldots, s_N)\mathbf{V}$ where $(s_1, \ldots, s_N) \sim \frac{1}{4}\mathbb{1}_{[1,5]}(dx)^{\otimes N}$.



(d) Spectrum of a 500×500 random matrix of the form $\sim \mathbf{U}\,\mathrm{diag}(s_1, \ldots, s_N)\mathbf{V}$ where $(s_1, \ldots, s_N) \sim \frac{1}{2}(\delta_1 + \delta_3)^{\otimes N}$.

Figure 2: Spectrum of isotropic matrices.

- $\{(y_{ij}, y_{ji}), 1 \le i < j \le N\} \cup \{y_{ii}, 1 \le i \le N\}$ is a family of independent random vectors,
- $\{(y_{ij}, y_{ji}), 1 \le i < j \le N\}$ are i.i.d. Gaussian such that

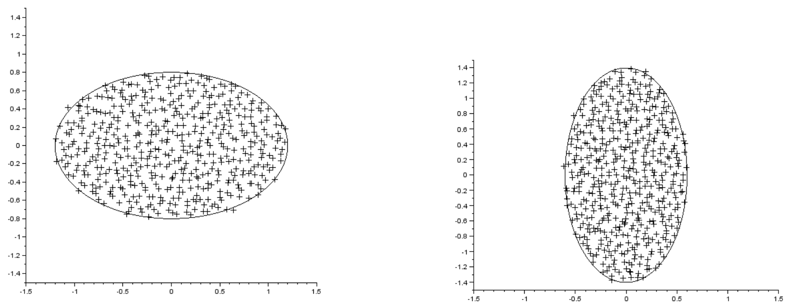
$$\mathbb{E}\,y_{ij}^2 = \mathbb{E}\,y_{ji}^2 = \mathbb{E}\,y_{ij}\overline{y_{ji}} = 0, \quad \mathbb{E}\,|y_{ij}|^2 = \mathbb{E}\,|y_{ji}|^2 = 1 \qquad \text{and} \qquad \mathbb{E}\,y_{ij}y_{ji} = \rho$$

- $\{y_{ii}, 1 \le i \le N\}$ are i.i.d. Gaussian such that

$$\mathbb{E}\,y_{ii}^2 = \rho \qquad \text{and} \qquad \mathbb{E}\,|y_{ii}|^2 = 1.$$

Remark 0.2.2. This ensemble is called *elliptic* due to the fact the eigenvalues of such matrices tend to spread over an ellipse. Indeed, it is known (see [89]) that when one renormalizes by $\frac{1}{\sqrt{N}}$ a matrix from the Gaussian elliptic ensemble, its limiting eigenvalue distribution, as N goes to infinity, is the uniform measure μ_ρ on the ellipse

$$\mathscr{E}_\rho := \left\{ z \in \mathbb{C}\,;\; \frac{\mathrm{Re}(z)^2}{(1+\rho)^2} + \frac{\mathrm{Im}(z)^2}{(1-\rho)^2} \leq 1 \right\}. \tag{2}$$



(a) Spectrum of a 500×500 Gaussian elliptic matrix where $\rho = 0.2$.

(b) Spectrum of a 500×500 Gaussian elliptic matrix where $\rho = -0.4$.

Figure 3: Spectrum of Gaussian elliptic matrices.

Remark 0.2.3. Gaussian elliptic matrices can be seen as an generalization of *GUE matrices* and the *Ginibre matrices*. Indeed, a Gaussian elliptic matrix $\mathbf{Y}$ of parameter ρ can be realized as

$$\mathbf{Y} = \sqrt{\frac{1+\rho}{2}}\mathbf{H}_1 + \mathrm{i}\sqrt{\frac{1-\rho}{2}}\mathbf{H}_2,$$

where $\mathbf{H}_1$ and $\mathbf{H}_2$ are two independent GUE matrices from the *GUE*. Hence GUE matrices (resp. Ginibre matrices) are Gaussian elliptic matrices of parameter 1 (resp. 0).

One can also define more general *elliptic random matrices* (see [70, 71, 73, 74] for more details). One feature of the Gaussian elliptic matrices is that their distribution is invariant by unitary conjugation (see for example Remark 0.2.3).

0.3 Spiked models

In this thesis, we study an additive deformation of random matrices. For any $N \geq 1$, let $\mathbf{A}_N$ be a $N \times N$ random matrix with no "natural outliers" : i.e. with a probability tending to one, there is no eigenvalue of $\mathbf{A}_N$ away from the bulk. Let $\mathbf{P}_N$ be a $N \times N$ matrix of perturbation which shall be "small" in the sense

- the rank of $\mathbf{P}_N$ is bounded : it means that there exists a fixed integer k, independent from N, such that $\operatorname{rank}(\mathbf{P}_N) \leq k$,
- the operator norm of $\mathbf{P}_N$ is also bounded, uniformly in N.

We also assume that $\mathbf{P}_N$ is independent from $\mathbf{A}_N$. Then, we consider

$$\widetilde{\mathbf{A}}_N \quad := \quad \mathbf{A}_N + \mathbf{P}_N. \tag{3}$$

As $\mathbf{P}_N$ has finite rank, it barely affects the global behavior of the spectrum, nevertheless, it can generate outliers. Since we assume that $\mathbf{A}_N$ has no natural outliers, the questions are :

- what are the conditions on $\mathbf{P}_N$ to observe outliers for $\widetilde{\mathbf{A}}_N$?
- is there a *phase transition* as in [9]?
- what are the behavior of such outliers (location, distribution of the fluctuations, rate of convergence, correlation ...)?

0.4 Results

In the first paper *Outliers in the Single Ring Theorem* [21], we did observe a *phase transition*: we proved that all eigenvalues of $\mathbf{P}_N$ with a modulus higher than b (the upper radius of the ring support) give birth to an outlier equal to this eigenvalue. Conversely, any eigenvalue with modulus lower than b generate no outlier, even inside the inner circle of the ring (i.e. the z with modulus lower than a) : see Figure 1.5.

The true novelty of this paper was the result on the fluctuations : due to the non Hermitian form of the additive perturbation, we had to consider the Jordan Canonical Form (JCF) of the matrix $\mathbf{P}$ and we showed that asymptotic behavior of the outliers highly depends on the JCF of $\mathbf{P}_N$: more precisely, unlike the Hermitian case, the convergence rate is no longer necessarily $\frac{1}{\sqrt{N}}$ but is equal to $\frac{1}{(\sqrt{N})^{1/p}}$ where p is the size of the Jordan block associated to the eigenvalue of $\mathbf{P}_N$. Moreover, the outliers tend to locate around their limit at the vertices of a p-sided regular polygon (see Figure 1.6). At last, since the perturbation is not Hermitian, its eigenvectors cannot necessarily be chosen orthogonal : we showed that it can make appear some correlations between outliers at macroscopique distance from each other. This last phenomenon has been already noticed by Knowles and Yin in [62] in the Hermitian case, but the correlation they observed was due to the non Gaussian structure of the model, whereas in our case we can observe such correlated outliers for random matrices from the Ginibre ensemble. Our proofs, as in most other results, rely on heavily the so-called *Weingarten calculus*, an integration method for the Haar measure on the unitary group developed by Collins and Śniady in [36, 38].

In a second time, using the same tools than the previous paper, we obtained results on the fluctuations of linear spectral statistics of a random matrix $\mathbf{A}$ form the Single Ring Theorem. More precisely, for any analytic function f on the ring $\{a \leq |z| \leq b\}$ and any deterministic matrix $\mathbf{N}$ (whose Frobenius norm is a $O\big(N^{-1/2}\big)$), we proved that

$$\mathrm{Tr}\left((f(\mathbf{A}) - f(0)\mathbf{I})\mathbf{N}\right)$$

weakly converges to a Gaussian distribution. One of the main difficulties of this second paper was since $\{a \leq |z| \leq b\}$ is not simply connected, we had to study Laurent series of matrix $\mathbf{A}$ and understand the joint distribution of $(\mathbf{A}, \mathbf{A}^{-1})$. As an application of the result, we were able to study a multiplicative perturbation which yields spikes inside the inner circle of the ring (see Figure 2.8).

Then, we wanted to use the good understanding of the non Hermitian structure of the additive perturbation on well-known models. Hence, in the third paper, we investigated nearly Hermitian random matrices which are Hermitian random matrices with a non Hermitian additive deformation. We found back the same kind of result from the first paper (phase transition, various rates of convergence due to the JCF of the perturbation, correlation between far away outliers,...) but we also proved that in some cases, the spikes can outnumber the rank of the perturbation (see Figure 3.9(b)), phenomenon already observed in [19] in the case where the support of the limit spectral distribution has a disconnected support (see also [12]). In our case, the phenomenon occurs even for connected support.

At last, we studied more generic models which are invariant in distribution by unitary conjugation. We obtained the limit distribution of the fluctuations of matrix entries and as an application of this result, we studied the outliers of low-ranked perturbation of matrices from the Gaussian elliptic ensemble. We proved, up to some changes, that main theorem of the first paper still apply for such matrices (see Figure 4.12).

I

Outliers in the Single Ring Theorem

Florent Benaych-Georges and Jean Rochet

Abstract :

This text is about spiked models of non-Hermitian random matrices. More specifically, we consider matrices of the type $\mathbf{A}+\mathbf{P}$, where the rank of $\mathbf{P}$ stays bounded as the dimension goes to infinity and where the matrix $\mathbf{A}$ is a non-Hermitian random matrix, satisfying an isotropy hypothesis: its distribution is invariant under the left and right actions of the unitary group. The macroscopic eigenvalue distribution of such matrices is governed by the so called *Single Ring Theorem*, due to Guionnet, Krishnapur and Zeitouni. We first prove that if $\mathbf{P}$ has some eigenvalues out of the maximal circle of the single ring, then $\mathbf{A}+\mathbf{P}$ has some eigenvalues (called *outliers*) in the neighborhood of those of $\mathbf{P}$, which is not the case for the eigenvalues of $\mathbf{P}$ in the inner cycle of the single ring. Then, we study the fluctuations of the outliers of $\mathbf{A}$ around the eigenvalues of $\mathbf{P}$ and prove that they are distributed as the eigenvalues of some finite dimensional random matrices. Such kind of fluctuations had already been shown for Hermitian models. More surprising facts are that outliers can here have very various rates of convergence to their limits (depending on the Jordan Canonical Form of $\mathbf{P}$) and that some correlations can appear between outliers at a macroscopic distance from each other (a fact already noticed by Knowles and Yin in [62] in the Hermitian case, but only for non Gaussian models, whereas spiked Gaussian matrices belong to our model and can have such correlated outliers). Our first result generalizes a result by Tao proved specifically for matrices with i.i.d. entries, whereas the second one (about the fluctuations) is new.

1.1 Introduction

We know that, most times, if one adds to a large random matrix, a finite rank perturbation, it barely modifies its spectrum. However, we observe that the extreme eigenvalues may be altered and deviated away from the bulk. This phenomenon has already been well understood in the Hermitian case. It was shown under several hypotheses in [77, 43, 28, 29, 19, 20, 16, 17, 30, 61, 62] that for a large random Hermitian matrix, if the strength of the added perturbation is above a certain threshold, then the extreme eigenvalues of the perturbed matrix deviate at a macroscopic distance from the bulk (such eigenvalues are usually called *outliers*) and have well understood fluctuations, otherwise they stick to the bulk and fluctuate as those of the non-perturbated matrix (this phenomenon is called the *BBP phase transition*, named after the authors of [9], who first brought it to light for empirical covariance matrices). Also, Tao, O'Rourke, Renfrew, Bordenave and Capitaine studied a non-Hermitian case: in [92, 73, 24] they considered spiked i.i.d. or elliptic random matrices and proved that for large enough spikes, some outliers also appear at precise positions. In this paper, we study finite rank perturbations for another natural model of non-Hermitian random matrices, namely the *isotropic* random matrices, i.e. the random matrices invariant, in law, under the left and right actions of the unitary group. Such matrices can be written

$$\mathbf{A} = \mathbf{U}\begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_N \end{pmatrix}\mathbf{V}, \tag{1.1}$$

with $\mathbf{U}$ and $\mathbf{V}$ independent Haar-distributed random matrices and the s_i's some positive numbers which are independent from $\mathbf{U}$ and $\mathbf{V}$. We suppose that the empirical distribution of the s_i's tends to a probability measure ν which is compactly supported on $\mathbb{R}^+$. We know that the singular values of a random matrix with i.i.d. entries satisfy this last condition (where ν is the Marčenko-Pastur quarter circular law with density $\pi^{-1}\sqrt{4-x^2}\mathbf{1}_{[0,2]}(x)dx$, see for example [2, 6, 91, 25]), so one can see this model as a generalization of the Ginibre matrices (i.e. matrices with i.i.d. standard complex Gaussian entries). In [49], Guionnet, Krishnapur and Zeitouni showed that the eigenvalues of $\mathbf{A}$ tend to spread over a single annulus centered in the origin as the dimension tends to infinity. Furthermore in [50], Guionnet and Zeitouni proved the convergence in probability of the support of its ESD (Empirical Spectral Distribution) which shows the lack of natural outliers for this kind of matrices (see Figure 1.4). This result has been recently improved in [15] with exponential bounds for the rate of convergence.

In this paper, we prove that, for a finite rank perturbation $\mathbf{P}$ with bounded operator norm, outliers of $\mathbf{A}+\mathbf{P}$ show up close to the eigenvalues of $\mathbf{P}$ which are outside the annulus whereas no outlier appears inside the inner circle of the ring. Then we show (and this is the main difficulty of the paper) that the outliers have fluctuations which are not necessarily Gaussian and whose convergence rates depend on the shape of the perturbation, more precisely on its Jordan Canonical Form[2]. Let us denote by $a < b$ the radiuses of the circles bounding the support of the limit spectral law of $\mathbf{A}$. We prove that for any eigenvalue θ of $\mathbf{P}$ such that $|\theta| > b$, if one denotes by

$$\underbrace{p_1,\ldots,p_1}_{\beta_1 \text{ times}} > \underbrace{p_2,\ldots,p_2}_{\beta_2 \text{ times}} > \cdots > \underbrace{p_\alpha,\ldots,p_\alpha}_{\beta_\alpha \text{ times}}$$

the sizes of the blocks of type $\mathbf{R}_p(\theta)$ (notation introduced in Footnote 2) in the Jordan Canonical Form of $\mathbf{P}$, then there are exactly $\beta_1 p_1 + \cdots + \beta_\alpha p_\alpha$ outliers of

[2]Recall that any matrix $\mathbf{M}$ in the set $\mathscr{M}_N(\mathbb{C})$ of $N \times N$ complex matrices is similar to a square block diagonal matrix

$$\begin{pmatrix} \mathbf{R}_{p_1}(\theta_1) & & & (0) \\ & \mathbf{R}_{p_2}(\theta_2) & & \\ & & \ddots & \\ (0) & & & \mathbf{R}_{p_r}(\theta_r) \end{pmatrix} \quad \text{where} \quad \mathbf{R}_p(\theta) = \begin{pmatrix} \theta & 1 & & (0) \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ (0) & & & \theta \end{pmatrix} \in \mathscr{M}_p(\mathbb{C}),$$

which is called the *Jordan Canonical Form* of $\mathbf{M}$, unique up to the order of the diagonal blocks [53, Chapter 3].

$\mathbf{A}+\mathbf{P}$ tending to θ and among them, $\beta_1 p_1$ go to θ at rate $N^{-1/(2p_1)}$, $\beta_2 p_2$ go to θ at rate $N^{-1/(2p_2)}$, etc... (see Figure 1.4). Moreover, we give the precise limit distribution of the fluctuations of these outliers around their limits. This limit distribution is not always Gaussian but corresponds to the law of the eigenvalues of some Gaussian matrices (possibly with correlated entries, depending on the eigenvectors of $\mathbf{P}$ and $\mathbf{P}^*$). A surprising fact is that some correlations can appear between the fluctuations of outliers with different limits. In [62], for spiked Wigner matrices, Knowles and Yin had already brought to light some correlations between outliers at a macroscopic distance from each other but it was for non Gaussian models, whereas spiked Ginibre matrices belong to our model and can have such correlated outliers.

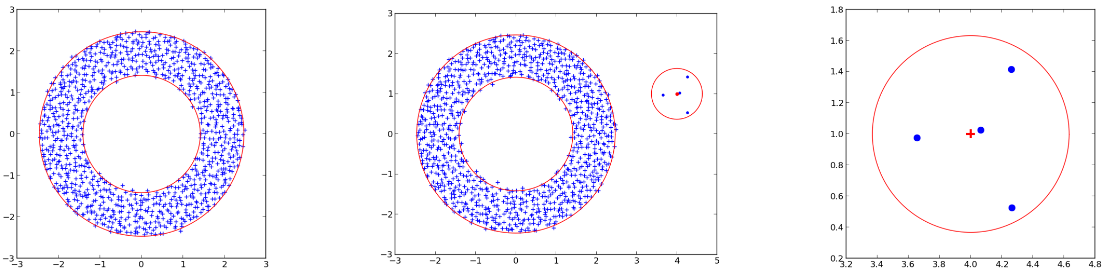


Figure 1.4: Spectrums of $\mathbf{A}$ (left), of $\mathbf{A}+\mathbf{P}$ (center) and zoom on a part of the spectrum of $\mathbf{A}+\mathbf{P}$ (right), for the same matrix $\mathbf{A}$ (chosen as in (1.1) for s_i's uniformly distributed on $[0.5,4]$ with $N=10^3$) and $\mathbf{P}$ with rank 4 having one block $\mathbf{R}_3(\theta)$ and one block $\mathbf{R}_1(\theta)$ in its Jordan Canonical Form ($\theta=4+i$). We see, on the right, four outliers around θ (θ is the red cross): three of them are at distance $\approx N^{-1/6}$ and one of them, much closer, is at distance $\approx N^{-1/2}$. One can notice that the three ones draw an approximately equilateral triangle. This phenomenon will be explained by Theorem 1.2.10.

The motivations behind the study of outliers in non-Hermitian models comes mostly from the general effort toward the understanding of the effect of a perturbation with small rank on the spectrum of a large-dimensional operator. The Hermitian case is now quite well understood, and this text provides a review of the question as far as outliers of isotropic non-Hermitian models are concerned. Besides, isotropic non-Hermitian matrix models also appear in wireless networks (see e.g. the recent preprint [94]).

1.2 Results

1.2.1 Setup and assumptions

Let, for each $N \geq 1$, $\mathbf{A}_N$ be a random matrix which admits the decomposition $\mathbf{A}_N = \mathbf{U}_N \mathbf{T}_N \mathbf{V}_n$ with $\mathbf{T}_N = \operatorname{diag}(s_1, \ldots, s_N)$ where the s_i's are non negative numbers (implicitly depending on N) and where $\mathbf{U}_N$ and $\mathbf{V}_N$ are two independent random unitary matrices which are Haar-distributed and independent from the matrix $\mathbf{T}_N$. We make (part of) the assumptions of the *Single Ring Theorem* [49] :

– **Hypothesis 1**: There is a deterministic number $b \geq 0$ such that as $N \to \infty$, we have the convergence in probability

$$\frac{1}{N}\operatorname{Tr}(\mathbf{T}_N^2) \longrightarrow b^2,$$

– **Hypothesis 2**: There exists $M > 0$, such that $\mathbb{P}(\|\mathbf{T}_N\|_{\mathrm{op}} > M) \longrightarrow 0$,

– **Hypothesis 3**: There exist a constant $\kappa > 0$ such that

$$\operatorname{Im}(z) \ > \ N^{-\kappa} \implies \left|\operatorname{Im}\left(G_{\mu_{\mathbf{T}_N}}(z)\right)\right| \ \leq \ \frac{1}{\kappa},$$

where for $\mathbf{M}$ a matrix, $\mu_{\mathbf{M}}$ denotes the empirical spectral distribution (ESD) of $\mathbf{M}$ and for μ a probability measure, G_μ denotes the *Stieltjes transform* of μ, that is $G_\mu(z) = \displaystyle\int \frac{\mu(dx)}{z-x}$.

Example 1.2.1. Thanks to [49], we know that our hypotheses are satisfied for example in the model of random complex matrices $\mathbf{A}_N$ distributed according to the law

$$\frac{1}{Z_N}\exp\left(-N\operatorname{Tr}V(\mathbf{X}\mathbf{X}^*)\right)d\mathbf{X},$$

where $d\mathbf{X}$ is the Lebesgue measure of the $N \times N$ complex matrices set, V is a polynomial with positive leading coefficient and Z_N is a normalization constant. It is quite a natural unitarily invariant model. One can notice that $V(x) = \frac{x}{2}$ gives the renormalized *Ginibre matrices*.

Remark 1.2.2. If one strengthens Hypothesis **1** into the convergence in probability of the ESD $\mu_{\mathbf{T}_N}$ of $\mathbf{T}_N$ to a limit probability measure ν, then by the Single Ring Theorem [49, 86], we know that the ESD $\mu_{\mathbf{A}_N}$ of $\mathbf{A}_N$ converges, in probability, weakly to a deterministic probability measure whose support is $\{z \in \mathbb{C},\ a \leq |z| \leq b\}$ where

$$\begin{aligned} a &= \left(\textstyle\int x^{-2}\nu(dx)\right)^{-1/2}, \\ b &= \left(\textstyle\int x^{2}\nu(dx)\right)^{1/2}. \end{aligned}$$

Remark 1.2.3. According to [50], with a bit more work (this works consists in extracting subsequences within which the ESD of $\mathbf{T}_N$ converges, so that we are in the conditions of the previous remark), we know that there is no natural outlier outside the circle centered at zero with radius b as long as $\|\mathbf{T}_N\|_{\mathrm{op}}$ is bounded, even if $\mathbf{T}_N$ has his own outliers. In Theorem 1.2.6, to make also sure there is no natural outlier inside the inner circle (when $a > 0$), we may suppose in addition that $\sup_{N\geq 1} \|\mathbf{T}_N^{-1}\|_{\mathrm{op}} < \infty$.

Remark 1.2.4. In the case where the matrix $\mathbf{A}$ is a *real* isotropic matrix (i.e. where $\mathbf{U}$ and $\mathbf{V}$ are Haar-distributed on the *orthogonal* group), despite the facts that the Single Ring Theorem still holds, as proved in [49], and that the Weingarten calculus works quite similarly, our proof does not work anymore: the reason is that we use in a crucial way the bound of Lemma 1.5.10, proved in [15] thanks to an explicit formula for the Weingarten function of the unitary group, which has no analogue for the orthogonal group. However, numerical simulations tend to show that similar behaviors occur, with the difference that the radial invariance of certain limit distributions is replaced by the invariance under the action of some discrete groups, reflecting the transition from the unitary group to the orthogonal one.

1.2.2 Main results

Let us now consider a sequence of matrices $\mathbf{P}_N$ (possibly random, but independent of $\mathbf{U}_N, \mathbf{T}_N$ and $\mathbf{V}_N$) with rank lower than a fixed integer r such that $\|\mathbf{P}_N\|_{\mathrm{op}}$ is also bounded. Then, we have the following theorem (note that in its statement, r_b, as the $\lambda_i(\mathbf{P}_N)$'s, can possibly depend on N and be random):

Theorem 1.2.5 (Outliers for finite rank perturbation)**.** *Suppose Hypothesis* **1** *to hold. Let $\varepsilon > 0$ be fixed and suppose that $\mathbf{P}_N$ hasn't any eigenvalues in the band $\{z \in \mathbb{C},\ b+\varepsilon < |z| < b+3\varepsilon\}$ for all sufficiently large N, and has r_b eigenvalues counted with multiplicity*[3] *$\lambda_1(\mathbf{P}_N), \ldots, \lambda_{r_b}(\mathbf{P}_N)$ with modulus higher than $b+3\varepsilon$.*

Then, with a probability tending to one, $\mathbf{A}_N + \mathbf{P}_N$ has exactly r_b eigenvalues with modulus higher than $b+2\varepsilon$. Furthermore, after labeling properly,

$$\forall i \in \{1, \ldots, r_b\},\ \ \lambda_i(\mathbf{A}_N + \mathbf{P}_N) - \lambda_i(\mathbf{P}_N) \xrightarrow{(\mathbb{P})} 0.$$

This first result is a generalization of Theorem 1.4 of Tao's paper [92], and so is its proof. However, things are different inside the annulus. Indeed, the following result establishes the lack of small outliers:

Theorem 1.2.6 (No outlier inside the bulk)**.** *Suppose that there exists $M' > 0$ such that*

$$\mathbb{P}(\|\mathbf{T}_N^{-1}\|_{\mathrm{op}} > M') \longrightarrow 0$$

and that there is $a > 0$ deterministic such that we have the convergence in probability

$$\frac{1}{N}\sum_{i=1}^{N} s_i^{-2} \longrightarrow \frac{1}{a^2}.$$

[3] To sort out misunderstandings: we call the *multiplicity* of an eigenvalue its order as a root of the characteristic polynomial, which is greater than or equal to the dimension of the associated eigenspace.

Then for all $\delta \in]0,a[$, with a probability tending to one,

$$\mu_{\mathbf{A}_N+\mathbf{P}_N}\left(\{z\in\mathbb{C}, |z|<a-\delta\}\right) \;=\; 0,$$

where $\mu_{\mathbf{A}_N+\mathbf{P}_N}$ is the Empirical Spectral Distribution of $\mathbf{A}_N+\mathbf{P}_N$.

Theorems 1.2.5 and 1.2.6 are illustrated in Figure 1.5 (see also Figure 1.4). We drew circles around each eigenvalues of $\mathbf{P}_N$ and we do observe the lack of outliers inside the annulus.

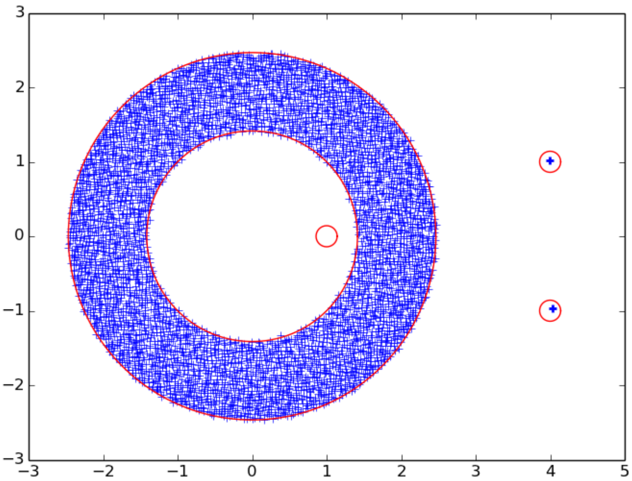


Figure 1.5: Eigenvalues of $\mathbf{A}_N+\mathbf{P}_N$ for $N=5.10^3$, ν the uniform law on $[0.5,4]$ and $\mathbf{P}_N = \operatorname{diag}(1,4+i,4-i,0,\ldots,0)$. The small circles are centered at $1,4+i$ and $4-i$, respectively, and each have a radius $\frac{10}{\sqrt{N}}$ (we will see later that in this particular case, the rate of convergence of $\lambda_i(\mathbf{A}_N+\mathbf{P}_N)$ to $\lambda_i(\mathbf{P}_N)$ is $\frac{1}{\sqrt{N}}$).

Let us now consider the fluctuations of the outliers. We need to be more precise about the perturbation matrix $\mathbf{P}_N$. Unlike Hermitian matrices, non-Hermitian matrices are not determined, up to a conjugation by a unitary matrix, only by their spectrums. A key parameter here will be the Jordan Canonical Form (JCF) of $\mathbf{P}_N$. From now on, we consider a deterministic perturbation $\mathbf{P}_N$ of rank $\leq r/2$ with r an integer independent of N (denoting the upper bound on the rank of $\mathbf{P}_N$ by $r/2$ instead of r will lighten the notations in the sequel).

As $\dim(\operatorname{Im}\mathbf{P}_N + (\ker\mathbf{P}_N)^\perp) \leq r$, one can find a unitary matrix $\mathbf{W}_N$ and an $r\times r$ matrix $\mathbf{Po}$ such that

$$\mathbf{P}_n \;=\; \mathbf{W}_N\begin{pmatrix}\mathbf{Po} & 0\\ 0 & 0\end{pmatrix}\mathbf{W}_N^*. \tag{1.2}$$

To simplify the problem, we shall suppose that $\mathbf{Po}$ does not depend on N (even though most of what follows can be extended to the case where **Po** depends on N but converges to a fixed $r\times r$ matrix as $N\to\infty$).

Let us now introduce the *Jordan Canonical Form* (JCF) of **Po** : we know that up to a basis change, one can write **Po** as a direct sum of *Jordan blocks*, i.e. blocks of the type

$$\mathbf{R}_p(\theta) \;=\; \begin{pmatrix} \theta & 1 & & (0) \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ (0) & & & \theta \end{pmatrix} \in \mathscr{M}_p(\mathbb{C}) \qquad (\theta\in\mathbb{C},\, p\geq 1). \tag{1.3}$$

Let us denote by $\theta_1,\ldots,\theta_q$ the distinct eigenvalues of **Po** which are in $\{|z|>b+3\varepsilon\}$ (for b as in Hypothesis **1** and ε as in the hypothesis of Theorem 1.2.5) and for each $i=1,\ldots,q$, introduce a positive integer α_i, some positive integers $p_{i,1}>\cdots>p_{i,\alpha_i}$ corresponding to the distinct sizes of the blocks relative to the eigenvalue θ_i and $\beta_{i,1},\ldots,\beta_{i,\alpha_i}$ such that for all j, $\mathbf{R}_{p_{i,j}}(\theta_i)$ appears $\beta_{i,j}$ times, so that, for a certain $\mathbf{Q}\in\mathrm{GL}_r(\mathbb{C})$, we have:

$$\mathbf{J}=\mathbf{Q}^{-1}\mathbf{PoQ}=(\text{Matrix with spec.}\subset\{|z|\leq b+3\varepsilon\})\bigoplus\bigoplus_{i=1}^{q}\bigoplus_{j=1}^{\alpha_i}\underbrace{\begin{pmatrix}\mathbf{R}_{p_{i,j}}(\theta_i) & & \\ & \ddots & \\ & & \mathbf{R}_{p_{i,j}}(\theta_i)\end{pmatrix}}_{\beta_{i,j}\text{ blocks}} \tag{1.4}$$

where $\oplus$ is defined, for square block matrices, by $\mathbf{M}\oplus\mathbf{N}:=\begin{pmatrix}\mathbf{M} & 0\\ 0 & \mathbf{N}\end{pmatrix}$.

The asymptotic orders of the fluctuations of the eigenvalues of $\widetilde{\mathbf{A}}_N:=\mathbf{A}_N+\mathbf{P}_N$ depend on the sizes $p_{i,j}$ of the blocks. Actually, for each θ_i, we know, by Theorem 1.2.5, there are $\sum_{j=1}^{\alpha_i}p_{ij}\times\beta_{i,j}$ eigenvalues of $\widetilde{\mathbf{A}}_N$ which tend to θ_i : we shall write them with a tilda and a θ_i on the top left corner: ${}^{\theta_i}\widetilde{\lambda}$. Theorem 1.2.10 below will state that for each block with size $p_{i,j}$ corresponding to θ_i of the JCF of **Po**, there are $p_{i,j}$ eigenvalues (we shall write them with $p_{i,j}$ on the bottom left corner : ${}^{\theta_i}_{p_{i,j}}\widetilde{\lambda}$) whose convergence rate will be $N^{-1/(2p_{i,j})}$. As there are $\beta_{i,j}$ blocks of size $p_{i,j}$, there are actually $p_{i,j}\times\beta_{i,j}$ eigenvalues tending to θ_i with convergence rate $N^{-1/(2p_{i,j})}$ (we shall write them ${}^{\theta_i}_{p_{i,j}}\widetilde{\lambda}_{s,t}$ with $s\in\{1,\ldots,p_{i,j}\}$ and $t\in\{1,\ldots,\beta_{i,j}\}$). It would be convenient to denote by $\Lambda_{i,j}$ the vector with size $p_{i,j}\times\beta_{i,j}$ defined by

$$\Lambda_{i,j} \;:=\; \left(N^{1/(2p_{i,j})}\cdot\left({}^{\theta_i}_{p_{i,j}}\widetilde{\lambda}_{s,t}-\theta_i\right)\right)_{\substack{1\leq s\leq p_{i,j}\\ 1\leq t\leq\beta_{i,j}}}. \tag{1.5}$$

Let us now define the family of random matrices that we shall use to characterize the limit distribution of the $\Lambda_{i,j}$'s. For each $i = 1, \ldots, q$, let $I(\theta_i)$ (resp. $J(\theta_i)$) denote the set, with cardinality $\sum_{j=1}^{\alpha_i} \beta_{i,j}$, of indices in $\{1, \ldots, r\}$ corresponding to the first (resp. last) columns of the blocks $\mathbf{R}_{p_{i,j}}(\theta_i)$ ($1 \le j \le \alpha_i$) in (1.4).

Remark 1.2.7. Note that the columns of $\mathbf{Q}$ (resp. of $(\mathbf{Q}^{-1})^*$) whose index belongs to $I(\theta_i)$ (resp. $J(\theta_i)$) are eigenvectors of $\mathbf{Po}$ (resp. of $\mathbf{Po}^*$) associated to θ_i (resp. $\overline{\theta_i}$). Indeed, if $k \in I(\theta_i)$ and $\mathbf{e}_k$ denotes the k-th vector of the canonical basis, then $\mathbf{J}\mathbf{e}_k = \theta_i \mathbf{e}_k$, so that $\mathbf{Po}(\mathbf{Q}\mathbf{e}_k) = \theta_i \mathbf{Q}\mathbf{e}_k$.

Now, let

$$\left(m_{k,\ell}^{\theta_i}\right)_{\substack{i=1,\ldots,q, \\ (k,\ell)\in J(\theta_i)\times I(\theta_i)}} \tag{1.6}$$

be the random centered complex Gaussian vector with covariance

$$\mathbb{E}\left(m_{k,\ell}^{\theta_i}\, m_{k',\ell'}^{\theta_{i'}}\right) = 0, \qquad \mathbb{E}\left(m_{k,\ell}^{\theta_i}\, \overline{m_{k',\ell'}^{\theta_{i'}}}\right) = \frac{b^2}{\theta_i \overline{\theta_{i'}} - b^2}\, \mathbf{e}_k^* \mathbf{Q}^{-1} (\mathbf{Q}^{-1})^* \mathbf{e}_{k'}\, \mathbf{e}_{\ell'}^* \mathbf{Q}^* \mathbf{Q}\, \mathbf{e}_\ell, \tag{1.7}$$

where $\mathbf{e}_1, \ldots, \mathbf{e}_r$ are the column vectors of the canonical basis of $\mathbb{C}^r$. Note that each entry of this vector has a rotationally invariant Gaussian distribution on the complex plane.

For each i, j, let $K(i,j)$ (resp. $K(i,j)^-$) be the set, with cardinality $\beta_{i,j}$ (resp. $\sum_{j'=1}^{j-1} \beta_{i,j'}$), of indices in $J(\theta_i)$ corresponding to a block of the type $\mathbf{R}_{p_{i,j}}(\theta_i)$ (resp. to a block of the type $\mathbf{R}_{p_{i,j'}}(\theta_i)$ for $j' < j$). In the same way, let $L(i,j)$ (resp. $L(i,j)^-$) be the set, with the same cardinality as $K(i,j)$ (resp. as $K(i,j)^-$), of indices in $I(\theta_i)$ corresponding to a block of the type $\mathbf{R}_{p_{i,j}}(\theta_i)$ (resp. to a block of the type $\mathbf{R}_{p_{i,j'}}(\theta_i)$ for $j' < j$). Note that $K(i,j)^-$ and $L(i,j)^-$ are empty if $j = 1$. Let us define the random matrices

$$\begin{aligned}
\mathrm{M}_j^{\theta_i,\mathrm{I}} &:= [m_{k,\ell}^{\theta_i}]_{\substack{k\in K(i,j)^- \\ \ell\in L(i,j)^-}} & \qquad \mathrm{M}_j^{\theta_i,\mathrm{II}} &:= [m_{k,\ell}^{\theta_i}]_{\substack{k\in K(i,j)^- \\ \ell\in L(i,j)}} \\
\mathrm{M}_j^{\theta_i,\mathrm{III}} &:= [m_{k,\ell}^{\theta_i}]_{\substack{k\in K(i,j) \\ \ell\in L(i,j)^-}} & \qquad \mathrm{M}_j^{\theta_i,\mathrm{IV}} &:= [m_{k,\ell}^{\theta_i}]_{\substack{k\in K(i,j) \\ \ell\in L(i,j)}}
\end{aligned} \tag{1.8}$$

and then let us define the matrix $\mathbf{M}_j^{\theta_i}$ as

$$\mathbf{M}_j^{\theta_i} := \theta_i \left(\mathrm{M}_j^{\theta_i,\mathrm{IV}} - \mathrm{M}_j^{\theta_i,\mathrm{III}} \left(\mathrm{M}_j^{\theta_i,\mathrm{I}}\right)^{-1} \mathrm{M}_j^{\theta_i,\mathrm{II}} \right) \tag{1.9}$$

Remark 1.2.8. It follows from the fact that the matrix $\mathbf{Q}$ is invertible, that $\mathrm{M}_j^{\theta_i,\mathrm{I}}$ is a.s. invertible and so is $\mathbf{M}_j^{\theta_i}$.

Remark 1.2.9. From the Remark 1.2.7 and (1.7), we see that each matrix $\mathbf{M}_j^{\theta_i}$ essentially depends on the eigenvectors of $\mathbf{P}_N$ and of $\mathbf{P}_N^*$ associated to blocks $\mathbf{R}_{p_{i,j}}(\theta_i)$ in (1.4) and the correlations between several $\mathbf{M}_j^{\theta_i}$'s depend essentially on the scalar products of such vectors.

Now, we can formulate our main result.

Theorem 1.2.10. *1. As N goes to infinity, the random vector*

$$(\Lambda_{i,j})_{\substack{1\le i\le q\\ 1\le j\le \alpha_i}}$$

defined at (1.5) converges jointly to the distribution of a random vector

$$(\Lambda^\infty_{i,j})_{\substack{1\le i\le q\\ 1\le j\le \alpha_i}}$$

with joint distribution defined by the fact that for each $1\le i\le q$ and $1\le j\le \alpha_i$, $\Lambda^\infty_{i,j}$ is the collection of the ${p_{i,j}}^{th}$ roots of the eigenvalues of $\mathbf{M}_j^{\theta_i}$ defined at (1.9).

2. *The distributions of the random matrices $\mathbf{M}_j^{\theta_i}$ are absolutely continuous with respect to the Lebesgue measure and none of the coordinates of the random vector $(\Lambda^\infty_{i,j})_{\substack{1\le i\le q\\ 1\le j\le \alpha_i}}$ has distribution supported by a single point.*

Remark 1.2.11. Each non zero complex number has exactly $p_{i,j}$ ${p_{i,j}}^{\text{th}}$ roots, drawing a regular $p_{i,j}$-sided polygon. Moreover, by the second part of the theorem, the spectrums of the $\mathbf{M}_j^{\theta_i}$'s almost surely do not contain 0, so each $\Lambda^\infty_{i,j}$ is actually a complex random vector with $p_{i,j}\times\beta_{i,j}$ coordinates, which draw $\beta_{i,j}$ regular $p_{i,j}$-sided polygons.

Example 1.2.12. For example, suppose that $\mathbf{P}_N$ has only one eigenvalue θ with modulus $> b+2\varepsilon$ (i.e. $q=1$), with multiplicity 4 (i.e. $r_b=4$). Then five cases can occur (illustrated by simulations in Figure 1.6, see also Figure 1.4, corresponding to the case (b)):

(a) The JCF of $\mathbf{P}_N$ for θ has one block with size 4 (so that $\alpha_1=1$, $(p_{1,1},\beta_{1,1})=(4,1)$) : then the 4 outliers of $\widetilde{\mathbf{A}}_N$ are the vertices of a square with center $\approx\theta$ and size $\approx N^{-1/8}$ (their limit distribution is the one of the four fourth roots of the complex Gaussian variable $\theta m^\theta_{1,1}$ with covariance given by (1.7)).

(b) The JCF of $\mathbf{P}_N$ for θ has one block with size 3 and one block with size 1 (so that $\alpha_1=2$, $(p_{1,1},\beta_{1,1})=(3,1)$, $(p_{1,2},\beta_{1,2})=(1,1))$: then the 4 outliers of $\widetilde{\mathbf{A}}_N$ are the vertices of an equilateral triangle with center $\approx\theta$ and size $\approx N^{-1/6}$ plus a point at distance $\approx N^{-1/2}$ from θ (the three first ones behave like the three third roots of the variable $\theta m^\theta_{1,1}$ and the last one behaves like $\theta(m^\theta_{4,4}-m^\theta_{1,4}m^\theta_{4,1}/m^\theta_{1,1})$ where $m^\theta_{1,1},m^\theta_{1,4},m^\theta_{4,1},m^\theta_{4,4}$ are Gaussian variables with correlations given by (1.7)).

(c) The JCF of $\mathbf{P}_N$ for θ has two blocks with size 2 (so that $\alpha_1 = 1$, $(p_{1,1}, \beta_{1,1}) = (2,2)$) : then the 4 outliers of $\widetilde{\mathbf{A}}_N$ are the extremities of two crossing segments with centers $\approx \theta$ and size $\approx N^{-1/4}$ (their limit distribution is the one of the square roots of the eigenvalues of the matrix

$$M_1^\theta \;=\; \theta \begin{pmatrix} m^\theta_{1,1} & m^\theta_{1,3} \\ m^\theta_{3,1} & m^\theta_{3,3} \end{pmatrix}$$

where $m^\theta_{1,1}, m^\theta_{1,3}, m^\theta_{3,1}, m^\theta_{3,3}$ are Gaussian variables with correlations given by (1.7)).

(d) The JCF of $\mathbf{P}_N$ for θ has one block with size 2 and two blocks with size 1 (so that $\alpha_1 = 2$, $(p_{1,1}, \beta_{1,1}) = (2,1)$, $(p_{1,2}, \beta_{1,2}) = (1,2)$) : then the 4 outliers of $\widetilde{\mathbf{A}}_N$ are the extremities of a segment with center $\approx \theta$ and size $\approx N^{-1/4}$ plus two points at distance $\approx N^{-1/2}$ from θ (the two first ones behave like the square roots of $\theta m^\theta_{1,1}$ and the two last ones behave like the eigenvalues of the matrix

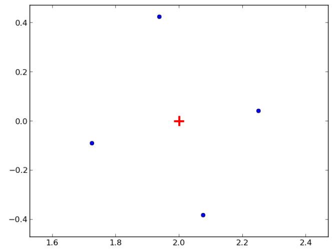
$$M_2^\theta \;=\; \theta \begin{pmatrix} m^\theta_{3,3} & m^\theta_{3,4} \\ m^\theta_{4,3} & m^\theta_{4,4} \end{pmatrix} - \frac{\theta}{m^\theta_{1,1}} \begin{pmatrix} m^\theta_{3,1} \\ m^\theta_{4,1} \end{pmatrix} \begin{pmatrix} m^\theta_{1,3} & m^\theta_{1,4} \end{pmatrix}$$

where the $m^\theta_{i,j}$'s are Gaussian variables with correlations given by (1.7)).

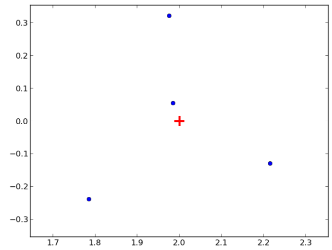
(e) The JCF of $\mathbf{P}_N$ for θ has four blocks with size 1 (so that $\alpha_1 = 1$, $(p_{1,1}, \beta_{1,1}) = (1,4)$) : then the 4 outliers of $\widetilde{\mathbf{A}}_N$ are four points at distance $\approx N^{-1/2}$ from θ (their limit distribution is the one of the eigenvalues of the matrix

$$M_1^\theta \;=\; \theta \begin{pmatrix} m^\theta_{1,1} & m^\theta_{1,2} & m^\theta_{1,3} & m^\theta_{1,4} \\ m^\theta_{2,1} & m^\theta_{2,2} & m^\theta_{2,3} & m^\theta_{2,4} \\ m^\theta_{3,1} & m^\theta_{3,2} & m^\theta_{3,3} & m^\theta_{3,4} \\ m^\theta_{4,1} & m^\theta_{4,2} & m^\theta_{4,3} & m^\theta_{4,4} \end{pmatrix}$$

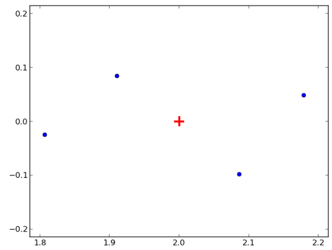
where the $m^\theta_{i,j}$'s are Gaussian variables with correlations given by (1.7)).



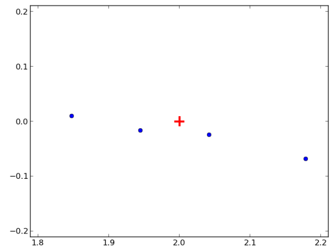
(a) The blue dots draw a **square** with center $\approx \theta$ at distance $\approx N^{-1/8}$ from θ



(b) The blue dots draw an **equilateral triangle** with center $\approx \theta$ at distance $\approx N^{-1/6}$ from θ plus a point at distance $\approx N^{-1/2}$ from θ



(c) The blue dots draw two crossing **segments** with centers $\approx \theta$ and lengths $\approx N^{-1/4}$



(d) The blue dots draw a **segment** with center $\approx \theta$ and length $\approx N^{-1/4}$ plus two points at distance $\approx N^{-1/2}$ from θ

Figure 1.6: The four first cases of Example 1.2.12 (the fifth one, less visual, does not appear here): the red cross is θ and the blue circular dots are the outliers of $\mathbf{A}_N + \mathbf{P}_N$ tending to θ_1. Each figure is made with the simulation of $\mathbf{A}_N$ a renormalized Ginibre matrix with size 2.10^3 plus $\mathbf{P}_N$ (whose choice depends of course of the case) with $\theta = 2$.

1.2.3 Examples

Uncorrelated case

Let us suppose that

$$\forall i,i' = 1,\ldots,q,\ \forall (k,\ell,k',\ell') \in J(\theta_i)\times I(\theta_i)\times J(\theta_{i'})\times I(\theta_{i'}), \tag{1.10}$$

$$\mathbf{e}_k^*\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^*\mathbf{e}_{k'} \cdot \mathbf{e}_{\ell'}^*\mathbf{Q}^*\mathbf{Q}\,\mathbf{e}_\ell \;=\; \mathbb{1}_{k=k',\ell=\ell'}$$

Note that it is the case when in (1.7), $\mathbf{Q}$ is unitary, i.e. when $\mathbf{P}$ is unitarily conjugated to $\begin{pmatrix}\mathbf{J} & 0\\ 0 & 0\end{pmatrix}$, with $\mathbf{J}$ as in (1.4).

By (1.7), Hypothesis (1.10) implies that the entries $m_{k,\ell}^{\theta_i}$ of the random vector of (1.6) are independent and that each $m_{k,\ell}^{\theta_i}$ has a distribution which depends only on θ_i. Let us introduce some notation. For β a positive integer, we define[4]

$$\mathrm{Ginibre}(\beta) := \beta\times\beta \text{ random matrix with i.i.d. } \mathscr{N}(0,1) \text{ entries}, \quad (1.11)$$

$$\mathrm{Ginibre}(\beta,\beta') := \beta\times\beta' \text{ random matrix with i.i.d. } \mathscr{N}(0,1) \text{ entries}, \quad (1.12)$$

and we get the following corollary:

Corollary 1.2.13. *If Hypothesis (1.10) holds, then :*

1. *the collection of random vectors $(\Lambda_{i,1},\Lambda_{i,2},\dots,\Lambda_{i,\alpha_i})$, indexed by $i=1,\dots,q$, i.e. by the distinct limit outliers θ_i, is asymptotically independent,*

2. *for each $i=1,\dots,q$ and each $j=1,\dots,\alpha_i$, the matrix $\mathbf{M}_j^{\theta_i}$ is distributed as:*
   - *if $j=1$, then*
   $$\mathbf{M}_j^{\theta_i}\sim\frac{\theta_i b}{\sqrt{|\theta_i|^2-b^2}}\mathrm{Ginibre}(\beta_{i,j}),$$
   - *if $j>1$, then*
   $$\mathbf{M}_j^{\theta_i}\sim\frac{\theta_i b}{\sqrt{|\theta_i|^2-b^2}}\left(\mathrm{Ginibre}(\beta_{i,j})-\mathrm{Ginibre}(\beta_{i,j},\rho_{i,j})\times\mathrm{Ginibre}(\rho_{i,j})^{-1}\times\mathrm{Ginibre}(\rho_{i,j},\beta_{i,j})\right),$$

   where the four Ginibre matrices involved if $j>1$ are independent and where $\rho_{i,j}=\sum_{j'=1}^{j-1}\beta_{i,j'}$.

Remark 1.2.14.

- The first part of this corollary means that under Hypothesis (1.10), the fluctuations of outliers of $\widetilde{\mathbf{A}}_N$ with different limits are independent. We will see below that it is not always true anymore if Hypothesis (1.10) does not hold.

- In the second part of this corollary, $j=1$ means that $p_{i,j}=\max_{j'}p_{i,j'}$, i.e. that we consider the outliers of $\widetilde{\mathbf{A}}_n$ at the largest possible distance ($\approx N^{-1/(2p_{i,1})}$) from θ_i.

[4] For any $\sigma>0$, $\mathscr{N}(0,\sigma^2)$ denotes the centered Gaussian law on $\mathbb{C}$ with covariance $\frac{1}{2}\begin{pmatrix}\sigma^2 & 0\\ 0 & \sigma^2\end{pmatrix}$.

- In the second part of the corollary, for $j>1$, the four matrices involved are independent, but the $\mathbf{M}_j^{\theta_i}$'s are not independent as j varies (the reason is that the matrix $M_j^{\theta_i,\mathrm{I}}$ of (1.9) contains $M_{j'}^{\theta_i,\mathrm{IV}}$ as a submatrix as soon as $j'<j$).

- If one weakens Hypothesis (1.10) by supposing it to hold only for $i=i'$ (resp. $i\neq i'$), then only the second (resp. first) part of the corollary stays true.

The $i=i'$ case of the last point of the previous remark implies the following corollary.

Corollary 1.2.15. *If, for a certain i, $\alpha_i=\beta_{i,1}=1$ (i.e. if θ_i is an eigenvalue of $\mathbf{P}$ with multiplicity[5] $p_{i,1}$ but with associated eigenspace having dimension one), then the random vector*

$$\left(N^{1/(2p_{i,1})}\cdot\left(\frac{\theta_i}{p_{i,1}}\widetilde{\lambda}_{s,1}-\theta_i\right)\right)_{1\leq s\leq p_{i,1}}$$

converges in distribution to the vector of the ${p_{i,1}}^{th}$ roots of a $\mathcal{N}(0,\frac{b^2}{|\theta_i|^2(|\theta_i|^2-b^2)})$ random variable.

Correlated case

If Hypothesis (1.10) does not hold anymore, then the individual and joint distributions of the random matrices $\mathbf{M}_j^{\theta_i}$ are not anymore related to Ginibre matrices as in Corollary 1.2.13: the entries of the matrices $M_j^{\theta_i,\mathrm{I,II,III,IV}}$ of (1.9) can have non uniform variances, even be correlated, and one can also have correlations between the entries of two matrices $\mathbf{M}_j^{\theta_i}$, $\mathbf{M}_{j'}^{\theta_{i'}}$ for $\theta_i\neq\theta_{i'}$. This last case has the surprising consequence that outliers of $\widetilde{\mathbf{A}}_N$ with different limits can be asymptotically correlated. Such a situation had so far only been brought to light, by Knowles and Yin in [62], for deformation of non Gaussian Wigner matrices. Note that in our model no restriction on the distributions of the deformed matrix $\mathbf{A}_N$ is made ($\mathbf{A}_N$ can for example be a renormalized Ginibre matrix). The following corollary gives an example of a simple situation where such correlations occur. This simple situation corresponds to the following case : we suppose that for some $i\neq i'$ in $\{1,\ldots,q\}$, we have $\beta_{i,1}=\beta_{i',1}=1$. We let ℓ and ℓ' (resp. k and k') denote the indices in $\{1,\ldots,r\}$ corresponding to the last (resp. first) columns of the block $\mathbf{R}_{p_{i,1}}(\theta_i)$ and of the block $\mathbf{R}_{p_{i',1}}(\theta_{i'})$ and set

$$K := \mathbf{e}_k^*\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^*\mathbf{e}_{k'}\cdot\mathbf{e}_{\ell'}^*\mathbf{Q}^*\mathbf{Q}\mathbf{e}_\ell. \tag{1.13}$$

We will see in the next corollary that as soon as $K\neq 0$, the fluctuations of outliers at macroscopic distance from each other (i.e. with distinct limits) are not indepen-

[5] Let us recall that what is here called the *multiplicity* of an eigenvalue its order as a root of the characteristic polynomial, which is not smaller than the dimension of the associated eigenspace.

dent. Set

$$\sigma^2 := \frac{|\theta_i|^2 b^2}{|\theta_i|^2 - b^2} \mathbf{e}_k^* \mathbf{Q}^{-1} (\mathbf{Q}^{-1})^* \mathbf{e}_k \cdot \mathbf{e}_\ell^* \mathbf{Q}^* \mathbf{Q} \mathbf{e}_\ell \,, \qquad \sigma'^2 := \frac{|\theta_{i'}|^2 b^2}{|\theta_{i'}|^2 - b^2} \mathbf{e}_{k'}^* \mathbf{Q}^{-1} (\mathbf{Q}^{-1})^* \mathbf{e}_{k'} \cdot \mathbf{e}_{\ell'}^* \mathbf{Q}^* \mathbf{Q} \mathbf{e}_{\ell'} . \tag{1.14}$$

Corollary 1.2.16. *Under this hypothesis, for any $1 \le s \le p_{i,1}$ and any $1 \le s' \le p_{i',1}$, as $n \to \infty$, the random vector*

$$(Z_N, Z'_N) := \left(\sqrt{N} \left(\underset{p_{i,1}}{\overset{\theta_i}{\widetilde{\lambda}}}_{s,1} - \theta_i \right)^{p_{i,1}} , \sqrt{N} \left(\underset{p_{i',1}}{\overset{\theta_{i'}}{\widetilde{\lambda}}}_{s',1} - \theta_{i'} \right)^{p_{i',1}} \right) \tag{1.15}$$

converges in distribution to a complex centered Gaussian vector (Z, Z') defined by

$$Z \sim \mathcal{N}(0, \sigma^2), \quad Z' \sim \mathcal{N}(0, \sigma'^2), \quad \mathbb{E}[ZZ'] = 0, \quad \mathbb{E}[Z\overline{Z'}] = \frac{\theta_i \overline{\theta_{i'}} \, b^2 K}{\theta_i \overline{\theta_{i'}} - b^2}. \tag{1.16}$$

Example 1.2.17. Let us illustrate this corollary (which is already an example) by a still more particular example. Suppose that $\mathbf{A}_n$ is a renormalized Ginibre matrix and that for $\theta = 1.5 + i$, $\theta' = 3 + i$ and for $\kappa \in \mathbb{R} \backslash \{-1, 1\}$, $\mathbf{Po}$ is given by

$$\mathbf{Po} = \mathbf{Q} \begin{pmatrix} \theta & 0 \\ 0 & \theta' \end{pmatrix} \mathbf{Q}^{-1} , \qquad \mathbf{Q} = \begin{pmatrix} 1 & \kappa \\ \kappa & 1 \end{pmatrix} .$$

In this case, $q = 2$, $\alpha_1 = \alpha_2 = p_{1,1} = p_{2,1} = \beta_{1,1} = \beta_{2,1} = 1$ and $\ell = k = 1$, $\ell' = k' = 2$. Thus $\mathbf{A}_N + \mathbf{P}_N$ has two outliers $\widetilde{\lambda}_N := \underset{p_{1,1}}{\overset{\theta}{\widetilde{\lambda}}}_{1,1}$ and $\widetilde{\lambda}'_N := \underset{p_{2,1}}{\overset{\theta'}{\widetilde{\lambda}}}_{1,1}$ and one can compute the numbers K, σ, σ' of (1.13), (1.14) and get

$$\sigma^2 = \frac{(1+\kappa^2)^2}{(1 - |\theta|^{-2})(1-\kappa^2)^2} \quad \sigma'^2 = \frac{(1+\kappa^2)^2}{(1 - |\theta'|^{-2})(1-\kappa^2)^2} \quad \mathbb{E}[Z\overline{Z'}] = \frac{-4\kappa^2}{(1 - (\theta\overline{\theta'})^{-1})(1-\kappa^2)^2}. \tag{1.17}$$

We see that for $\kappa = 0$, $Z_N = \sqrt{N}(\widetilde{\lambda} - \theta)$ and $Z'_N = \sqrt{N}(\widetilde{\lambda}' - \theta')$ are asymptotically independent, but that for $\kappa \neq 0$, Z_N and Z'_N are not asymptotically independent anymore. This phenomenon and the accuracy of the approximation $(Z_N, Z'_N) \approx (Z, Z')$ for $N \gg 1$ are illustrated by Table 1.1 and Figure 1.7, where 10^3 samples of (Z_N, Z'_N) have been simulated for $N = 10^3$.

1.2.4 Preliminaries to the proofs

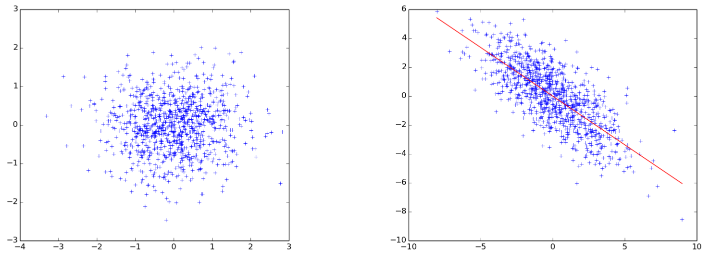
First, for notational brevity, from now on, N will be an implicit parameter ($\mathbf{A} := \mathbf{A}_N$, $\mathbf{P} := \mathbf{P}_N, \ldots$), except in case of ambiguity.

Secondly, from now on, we shall suppose that $\mathbf{T}$ is deterministic. Indeed, once the results established with $\mathbf{T}$ deterministic, as $\mathbf{T}$ is independent from the others random variables and the only relevant parameter b is deterministic, we can condition on $\mathbf{T}$ and apply the deterministic result. So we suppose that $\mathbf{T}$ is deterministic and that there is a constant M independent of N such that for all N,

$$\|\mathbf{T}\|_{\text{op}} \le M.$$

	$\mathbb{E}[\lvert Z\rvert^2]$		$\mathbb{E}[\lvert Z'\rvert^2]$		$\mathbb{E}[Z\overline{Z'}]$	
	$\kappa = 0$	$\kappa = 2^{-1/2}$	$\kappa = 0$	$\kappa = 2^{-1/2}$	$\kappa = 0$	$\kappa = 2^{-1/2}$
Theorical	1.444	13.0	1.111	10.0	0.0	$-8.755 - 1.358i$
Empirical	1.492	12.72	1.107	10.04	$0.00616 - 0.00235i$	$-8.917 - 1.317i$

Table 1.1: Comparison between theoretical asymptotic formulas (1.16) and (1.17) and a Monte-Carlo numerical computation made out of 10^3 matrices with size $N = 10^3$.



(a) $\kappa = 0$: uncorrelated case.

(b) $\kappa = 2^{-1/2}$: correlated case. The straight line is the theoretical optimal regression line (i.e. the line with equation $y = ax$ where a minimizes the variance of $Y - aX$, computed thanks to the asymptotic formulas (1.16) and (1.17)): one can notice that it fits well with the empirical datas.

Figure 1.7: **Lack of correlation/correlation** between outliers with **different limits** : abscissas (resp. ordinates) of the dots are $X := \Im(Z_N)$ (resp. $Y := \Im(Z'_N)$) for 10^3 independent copies of (Z_N, Z'_N) (computed thanks to matrices with size $N = 10^3$ as for Table 1.1).

Thirdly, as the set of probability measures supported by $[0,M]$ is compact, up to an extraction, one can suppose that there is a probability measure Θ on $[0,M]$ such that the ESD of $\mathbf{T}$ converges to Θ as $N \to \infty$. We will work within this subsequence. This could seem to give a partial convergence result, but in fact, what is proved is that from any subsequence, one can extract a subsequence which converges to the limit given by the theorem. This is of course enough for the proof. Note that by Hypothesis **1**, we have $b^2 = \displaystyle\int x^2 \Theta(dx)$. Having supposed that the ESD of $\mathbf{T}$ converges to Θ insures that $\mathbf{A}$ satisfies the hypotheses[6] of the Single Ring Theorem

[6]There is actually another assumption in the *Single Ring Theorem* [49], but Rudelson and Vershynin recently showed in [86] that it was unnecessary. In [10], Basak Dembo also weakened the hypotheses (roughly allowing Hypothesis **3** not to hold on a small enough set, so that ν is allowed

of [49] and of the paper [50]. We will use it once, in the proof of Lemma 1.6.1, where we need one of the preliminary results of [50].

At last, notice that $\mathbf{A}+\mathbf{P}$ and $\mathbf{V}(\mathbf{A}+\mathbf{P})\mathbf{V}^*$ have the same spectrum, that

$$\mathbf{V}(\mathbf{A}+\mathbf{P})\mathbf{V}^* \;=\; \mathbf{VUT}+\mathbf{VPV}^*, \tag{1.18}$$

and that as $\mathbf{U}$ and $\mathbf{V}$ are independent Haar-distributed matrices, $\mathbf{VU}$ and $\mathbf{V}$ are also Haar-distributed and independent. It follows that we shall, instead of the hypotheses made above the statement of Hypotheses **1**, **2**, and **3**, suppose that:

$$\mathbf{A}=\mathbf{UT} \qquad \text{with} \qquad \mathbf{T} \text{ deterministic and } \mathbf{U} \text{ Haar-distributed} \tag{1.19}$$

and $\qquad$ $\mathbf{P}$ is independent of $\mathbf{A}$ and invariant, in law, by conjugation by any unitary matrix.

In the sequel $\mathbb{E}_{\mathbf{U}}$ will denote the expectation with respect to the randomness of $\mathbf{U}$ and not to the one of $\mathbf{P}$. In the same way, $\mathbb{E}_{\mathbf{P}}$ will denote the expectation with respect to the randomness of $\mathbf{P}$.

1.2.5 Sketch of the proofs

We start with the following trick, now quite standard in spiked models. Let $\mathbf{B} \in \mathscr{M}_{N\times r}(\mathbb{C})$ and $\mathbf{C} \in \mathscr{M}_{r\times N}(\mathbb{C})$ such that $\mathbf{P}=\mathbf{BC}$ (where $\mathscr{M}_{p\times q}(\mathbb{C})$ denotes the rectangular complex matrices of size $p\times q$). Then

$$\begin{aligned}
\det(z\mathbf{I}-\widetilde{\mathbf{A}}) &= \det(z\mathbf{I}-(\mathbf{A}+\mathbf{P})) \\
&= \det(z\mathbf{I}-\mathbf{A})\det\left(\mathbf{I}-(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{P}\right) \\
&= \det(z\mathbf{I}-\mathbf{A})\det\left(\mathbf{I}-(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{BC}\right) \\
&= \det(z\mathbf{I}-\mathbf{A})\det\left(\mathbf{I}-\mathbf{C}(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}\right). \qquad (1.20)
\end{aligned}$$

For the last step, we used the fact that for all $\mathbf{M} \in \mathscr{M}_{r\times N}$ and $\mathbf{N} \in \mathscr{M}_{N\times r}(\mathbb{C})$, $\det(\mathbf{I}_r+\mathbf{MN}) = \det(\mathbf{I}_N+\mathbf{NM})$. Therefore, the eigenvalues z of $\widetilde{\mathbf{A}}$ which are not eigenvalues of $\mathbf{A}$ are characterized by

$$\det\left(\mathbf{I}-\mathbf{C}(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}\right) \;=\; 0. \tag{1.21}$$

In view of (1.20), as previously done by Tao in [92], we introduce the meromorphic functions (implicitly depending on N)

$$f(z) \;:=\; \det\left(\mathbf{I}-\mathbf{C}(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}\right) \;=\; \frac{\det(z\mathbf{I}-\widetilde{\mathbf{A}})}{\det(z\mathbf{I}-\mathbf{A})}, \tag{1.22}$$

$$g(z) \;:=\; \det\left(\mathbf{I}-\mathbf{C}(z\mathbf{I})^{-1}\mathbf{B}\right) \;=\; \frac{\det(z\mathbf{I}-\mathbf{P})}{\det(z\mathbf{I})} \tag{1.23}$$

---

to have some atoms). As it follows from the recent preprint [15] that the convergence of the extreme eigenvalues first established in [50] also works in this case, we could harmlessly weaken our hypotheses down to the ones of [10].

and aim to study the zeros of f.

• The proof of Theorem 1.2.5 (eigenvalues outside the outer circle) relies on the fact that on the domain $\{|z| > b+2\varepsilon\}$, $f(z) \approx g(z)$. This follows from the fact that for $|z| > b+2\varepsilon$, the $N \times N$ matrix $(z\mathbf{I}-\mathbf{A})^{-1} - z^{-1}\mathbf{I}$ has small entries, and even satisfies

$$\mathbf{x}^*((z\mathbf{I}-\mathbf{A})^{-1} - z^{-1}\mathbf{I})\mathbf{y} \ll 1 \tag{1.24}$$

for deterministic unitary column vectors $\mathbf{x}, \mathbf{y}$.

• The proof of Theorem 1.2.6 (lack of eigenvalues inside the inner circle) relies on the fact that for $|z| < a-\delta$, $\left\|\mathbf{C}(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}\right\|_{\mathrm{op}} < 1$. We will see that it follows from estimates as the one of (1.24) for $\mathbf{A}$ replaced by $\mathbf{A}^{-1}$.

• The most difficult part of the article is the proof of Theorem 1.2.10 about the fluctuations of the outliers around their limits θ_i $(1 \le i \le q)$. As the outliers are the zeros of f, we shall expand f around any fixed θ_i. Specifically, for each block size $p_{i,j}$ $(1 \le j \le \alpha_i)$, we prove at Lemma 1.5.1 that for $\pi_{i,j} := \sum_{l>j} \beta_{i,l} p_{i,l}$ and $\mathbf{M}_j^{\theta_i}$ the matrix with size[7] $\beta_{i,j}$ defined above, we have

$$f\left(\theta_i + \frac{z}{N^{1/(2p_{i,j})}}\right) \approx z^{\pi_{i,j}} \cdot \det\left(z^{p_{i,j}} - \mathbf{M}_j^{\theta_i}\right). \tag{1.25}$$

This proves that $\mathbf{A}+\mathbf{P}$ has $\pi_{i,j}$ outliers tending to θ_i at rate $\ll N^{-1/(2p_{i,j})}$, has $p_{i,j} \times \beta_{i,j}$ outliers tending to θ_i at rate $n^{-1/(2p_{i,j})}$ and that these $p_{i,j} \times \beta_{i,j}$ outliers are distributed as the $p_{i,j}{}^{\text{th}}$ roots of the eigenvalues of $\mathbf{M}_j^{\theta_i}$. We see that the key result in this proof is the estimate (1.25). To prove it, we first specify the choice of the already introduced matrices $\mathbf{B} \in \mathscr{M}_{N\times r}(\mathbb{C})$ and $\mathbf{C} \in \mathscr{M}_{r\times N}(\mathbb{C})$ such that $\mathbf{P} = \mathbf{BC}$ by imposing moreover that $\mathbf{CB} = \mathbf{J}$ (recall that $\mathbf{J}$ is the $r \times r$ Jordan Canonical Form of $\mathbf{P}$ of (1.4)). Then, for

$$\tilde{z} := \theta_i + \frac{z}{N^{1/(2p_{i,j})}}, \qquad \mathbf{X}_N^{\tilde{z}} := \sqrt{N}\mathbf{C}((\tilde{z}\mathbf{I}-\mathbf{A})^{-1} - \tilde{z}^{-1}\mathbf{I})\mathbf{B},$$

we write

$$\begin{aligned} f(\tilde{z}) &= \det\left(\mathbf{I} - \frac{1}{\tilde{z}}\mathbf{J} - \frac{1}{\sqrt{N}}\mathbf{X}_N^{\tilde{z}}\right) \\ &= \det\left(\mathbf{I} - \theta_i^{-1}\mathbf{J} + \theta_i^{-1}\left(1 - \frac{1}{1+N^{-1/(2p_{i,j})}z\theta_i^{-1}}\right)\mathbf{J} - \frac{1}{\sqrt{N}}\mathbf{X}_N^{\tilde{z}}\right) \\ &\approx \det\left(\mathbf{I} - \theta_i^{-1}\mathbf{J} + \frac{z\theta_i^{-2}}{N^{1/(2p_{i,j})}}\mathbf{J} - \frac{1}{\sqrt{N}}\mathbf{X}_N^{\tilde{z}}\right) \end{aligned} \tag{1.26}$$

[7] Recall the $\beta_{i,j}$ is the *number* of blocks $\mathbf{R}_{p_{i,j}}(\theta_i)$ in the JCF of $\mathbf{P}$.

At this point, one has to note that (obviously) $\det\left(\mathbf{I}-\theta_i^{-1}\mathbf{J}\right)=0$ and that (really not obviously) the $r\times r$ random array $\mathbf{X}_N^{\tilde{z}}$ converges in distribution to a Gaussian array as $N\to\infty$ (this is proved thanks to the Weingarten calculus). Then the result will follow from a Taylor expansion of (1.26) and a careful look at the main contributions to the determinant.

1.3 Eigenvalues outside the outer circle : proof of Theorem 1.2.5

We start with Equations (1.20) and (1.21), established in the previous Section, and the functions f and g, introduced at (1.22) and (1.23).

Lemma 1.3.1. *As N goes to infinity, we have*

$$\sup_{|z|\geq b+2\varepsilon}|f(z)-g(z)|\xrightarrow{(\mathbb{P})}0.$$

Before proving the lemma, let us explain how it allows to conclude the proof of Theorem 1.2.5. The poles of f and g are respectively eigenvalues of the $\mathbf{A}$ and of the null matrix, hence for N large enough, they have no pole in the region $\{z\in\mathbb{C}\,;\,|z|>b+2\varepsilon\}$, whereas their zeros in this region are precisely the eigenvalues of respectively $\widetilde{\mathbf{A}}$ and $\mathbf{P}$ that are in this region. But $|g|$ admits the following lower bound on the circle with radius $b+\varepsilon$: as we assumed that any eigenvalue of $\mathbf{P}$ is at least at distance at least ε from $\{z\in\mathbb{C}\,;\ |z|=b+2\varepsilon\}$, one has

$$\inf_{|z|=b+2\varepsilon}|g(z)|\;=\;\inf_{|z|=b+2\varepsilon}\frac{\prod_{i=1}^N|z-\lambda_i(\mathbf{P})|}{|z|^N}\;\geq\;\left(\frac{\varepsilon}{b+2\varepsilon}\right)^r,$$

so that by the previous lemma, with probability tending to one,

$$\forall z\in\mathbb{C},|z|=b+2\varepsilon\implies|f(z)-g(z)|<|g(z)|,$$

and so, by Rouché's Theorem [11, p. 131], we know that inside the region $\{z\in\mathbb{C},|z|\leq b+2\varepsilon\}$, f and g have the same number of zeros (since they both have n poles). Therefore, as their total number of zeros is n, f and g have the same number of zeros outside this region.

Also, Lemma 1.3.1 allows to conclude that, after a proper labeling

$$\forall i\in\{1,\ldots,r_b\},\ \ \lambda_i(\widetilde{\mathbf{A}})-\lambda_i(\mathbf{P})\xrightarrow{(\mathbb{P})}0.$$

Indeed, for each fixed $i\in\{1,\ldots,r_b\}$,

$$\begin{aligned}\prod_{j=1}^r\left|1-\frac{\lambda_j(\mathbf{P})}{\lambda_i(\widetilde{\mathbf{A}})}\right|\;=\;\left|g(\lambda_i(\widetilde{\mathbf{A}}))\right|\;&=\;\left|f(\lambda_i(\widetilde{\mathbf{A}}))-g(\lambda_i(\widetilde{\mathbf{A}}))\right|\\&\leq\;\sup_{|z|\geq b+2\varepsilon}|f(z)-g(z)|\xrightarrow{(\mathbb{P})}0.\end{aligned}$$

Let us now explain how to prove Lemma 1.3.1. One can notice at first that it suffices to prove that

$$\sup_{|z|\geq b+2\varepsilon} \left\|\mathbf{C}(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}-\mathbf{C}(z\mathbf{I})^{-1}\mathbf{B}\right\|_{\mathrm{op}} \xrightarrow{(\mathbb{P})} 0, \tag{1.27}$$

simply because the function $\det : \mathscr{M}_r(\mathbb{C}) \to \mathbb{C}$ is Lipschitz over every bounded set of $\mathscr{M}_r(\mathbb{C})$. Then, the proof of Lemma 1.3.1 is based on both following lemmas (whose proofs are postponed to Section 1.6).

Lemma 1.3.2. *There exists a constant $C_1 > 0$ such that the event*

$$\mathscr{E}_N := \{\forall k \geq 1,\ \ \|\mathbf{A}^k\|_{\mathrm{op}} \ \leq\ C_1 \cdot (b+\varepsilon)^k\}$$

has probability tending to one as N tends to infinity.

Lemma 1.3.3. *For all $k \geq 0$, as N goes to infinity, we have*

$$\|\mathbf{C}\mathbf{A}^k\mathbf{B}\|_{\mathrm{op}} \xrightarrow{(\mathbb{P})} 0.$$

On the event $\mathscr{E}_N$ defined at Lemma 1.3.2 above, we write, for $|z| \geq b+2\varepsilon$,

$$\mathbf{C}(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}-\mathbf{C}(z\mathbf{I})^{-1}\mathbf{B} \quad = \quad \mathbf{C}\sum_{k=1}^{+\infty}\frac{\mathbf{A}^k}{z^{k+1}}\mathbf{B}.$$

and it suffices to write that for any $\delta > 0$,

$$\begin{aligned}\mathbb{P}\left(\sup_{|z|\geq b+2\varepsilon} \left\|\mathbf{C}(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}-\mathbf{C}(z\mathbf{I})^{-1}\mathbf{B}\right\|_{\mathrm{op}} > \delta\right) \quad &\leq \quad \mathbb{P}(\mathscr{E}_N^c)+\mathbb{P}\left(\sum_{k=1}^{k_0}\frac{\left\|\mathbf{C}\mathbf{A}^k\mathbf{B}\right\|_{\mathrm{op}}}{(b+2\varepsilon)^{k+1}} > \frac{\delta}{2}\right) \\ &\quad +\mathbb{P}\left(\mathscr{E}_N \text{ and } \left\|\mathbf{C}\sum_{k=k_0+1}^{+\infty}\frac{\mathbf{A}^k}{(b+2\varepsilon)^{k+1}}\mathbf{B}\right\|_{\mathrm{op}} > \frac{\delta}{2}\right).\end{aligned}$$

By to Lemma 1.3.2 and the fact that $\mathbf{C}$ and $\mathbf{B}$ are uniformly bounded (see Remark 1.5.2), we can find k_0 so that the last event has a vanishing probability. Then, by Lemma 1.3.3, the probability of the last-but-one event goes to zero as N tends to infinity. This gives (1.27) and then Lemma 1.3.1.

1.4 Lack of eigenvalues inside the inner circle : proof of Theorem 1.2.6

Our goal here is to show that for all $\delta \in]0,a[$, with probability tending to one, the function f defined at (1.22) has no zero in the region $\{z \in \mathbb{C}, |z| < a-\delta\}$. Recall that

$$f(z) \quad = \quad \det\left(\mathbf{I}-\mathbf{C}(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}\right),$$

so that a simple sufficient condition would be $\left\|\mathbf{C}(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}\right\|_{\mathrm{op}} < 1$ for all $|z| < a-\delta$. Thus, it suffices to prove that with probability tending to one as n tends to infinity,

$$\sup_{|z|<a-\delta} \left\|\mathbf{C}(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}\right\|_{\mathrm{op}} < 1.$$

By Remark 1.2.3, we know that $\mathbf{A}$ is invertible. As in Section 1.3, we write, for all $|z| < a-\delta$,

$$\begin{aligned} \mathbf{C}\,(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{B} &= -\mathbf{C}\mathbf{A}^{-1}\left(\mathbf{I}-z\mathbf{A}^{-1}\right)^{-1}\mathbf{B} \\ &= -\mathbf{C}\sum_{k=1}^{\infty} z^{k-1}\mathbf{A}^{-k}\mathbf{B}. \end{aligned}$$

The idea is to see $\mathbf{A}^{-1}$ as an isotropic random matrix such as $\mathbf{A}$, since $\mathbf{A}^{-1} = \mathbf{V}^*\operatorname{diag}(\frac{1}{s_1},\ldots,\frac{1}{s_N})\mathbf{U}^*$, and satisfies the same kind of hypothesis. Indeed, Hypotheses **1** and **2** are automatiquelly satisfied because $a > 0$ (see Remark 1.2.3), and the following lemma, proved in Section 1.6.2, insures us that Hypotheses **3** is also satisfied.

Lemma 1.4.1. *There exist a constant $\widetilde{\kappa} > 0$ such that*

$$\operatorname{Im}(z) \;>\; N^{-\widetilde{\kappa}} \quad\Rightarrow\quad \left|\operatorname{Im}\left(G_{\mu_{\mathbf{T}^{-1}}}(z)\right)\right| \;\leq\; \frac{1}{\widetilde{\kappa}}.$$

Thus, according to [50], the support of $\mu_{\mathbf{A}^{-1}}$ converges in probability to the annulus
$\left\{z\in\mathbb{C},\ b^{-1}\leq |z|\leq a^{-1}\right\}$ as $N\to\infty$, and so, according to (1.27),

$$\sup_{|\xi|>a^{-1}+\varepsilon} \mathbf{C}\sum_{k=1}^{\infty}\frac{\mathbf{A}^{-k}}{\xi^{k+1}}\mathbf{B} \xrightarrow{(\mathbb{P})} 0.$$

Therefore

$$\mathbb{P}\left(\sup_{|z|<a-\delta}\left\|\mathbf{C}(zI-\mathbf{A})^{-1}\mathbf{B}\right\|_{\mathrm{op}} < 1\right) \geq 1-\mathbb{P}\left(\sup_{|\xi|>a^{-1}+\varepsilon}\left\|\mathbf{C}\sum_{k=1}^{\infty}\frac{\mathbf{A}^{-k}}{\xi^{k-1}}\mathbf{B}\right\|_{\mathrm{op}} > 1\right) \longrightarrow 1\ ,$$

with a proper choice for ε. □

1.5 Proof of Theorem 1.2.10

1.5.1 Lemma 1.5.1 granted proof of Theorem 1.2.10

Recall that we write $\mathbf{P} = \mathbf{BC}$ and we know that

$$\sup_{|z|>b+2\varepsilon}\left\|\mathbf{C}\,(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}-z^{-1}\mathbf{CB}\right\|_{\mathrm{op}} \xrightarrow{(\mathbb{P})} 0, \tag{1.28}$$

(again, for notational brevity, N will be an implicit parameter, except in case of ambiguity).

Following the ideas of [16], we shall need to differentiate the function f defined at (1.22) to understand the fluctuations of $\widetilde{\lambda}-\theta$, and to do so, we shall need to be more accurate in the convergence in (1.28).

Let us first state our key lemma, whose proof is postponed in Section 1.5.3. Recall from (1.4) that we supposed the JCF of $\mathbf{P}$ to have, for the eigenvalue θ_i, $\beta_{i,1}$ blocks with size $p_{i,1}$, ..., β_{i,α_i} blocks with size p_{i,α_i}. Recall also that

$$f(z) \quad = \quad \det\left(\mathbf{I}-\mathbf{C}(z\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}\right).$$

Lemma 1.5.1. *For all $j\in\{1,\ldots,\alpha_i\}$, let $F_j^{\theta_i}(z)$ be the rational function defined by*

$$F_j^{\theta_i}(z) \;:=\; f\left(\theta_i+\frac{z}{N^{1/(2p_{i,j})}}\right). \tag{1.29}$$

Then, there exists a collection of positive constants $(\gamma_{i,j})_{\substack{1\le i\le q\\1\le j\le\alpha_i}}$ and a collection of non vanishing random variables $(C_{i,j})_{\substack{1\le i\le q\\1\le j\le\alpha_i}}$ independent of z, such that we have the convergence in distribution (for the topology of the uniform convergence over any compact set)

$$\left(N^{\gamma_{i,j}}F_j^{\theta_i}(\cdot)\right)_{\substack{1\le i\le q\\1\le j\le\alpha_i}} \underset{N\to\infty}{\longrightarrow} \left(z\in\mathbb{C}\;\mapsto\; z^{\pi_{i,j}}\cdot C_{i,j}\cdot\det\left(z^{p_{i,j}}-\mathbf{M}_j^{\theta_i}\right)\right)_{\substack{1\le i\le q\\1\le j\le\alpha_i}}$$

where $\mathbf{M}_j^{\theta_i}$ is the random matrix introduced at (1.7) and $\pi_{i,j} \;:=\; \sum_{l>j}\beta_{i,l}p_{i,l}$.

To end the proof of Theorem 1.2.10, we make sure that we have the right number of eigenvalues of $\widetilde{\mathbf{A}}$ thanks to complex analysis considerations (Cauchy formula) :

- Eigenvalues tending to θ_i with the highest convergence rate :
  - Lemma 1.5.1 tells us that on any compact set, $F_j^{\theta_i}$ and $z^{\pi_{i,j}}\det(z^{p_{i,j}}-\mathbf{M}_j^{\theta_i})$ have the exact same number of roots (for any large enough N, the poles of $F_j^{\theta_i}$ leave any compact set), so, for the smallest block size p_{i,α_i}, we know that $F_{\alpha_i}^{\theta_i}$ has exactly $\beta_{i,\alpha_i}\times p_{i,\alpha_i}$ roots which do not eventually leave any compact set as N goes to infinity.
  - Moreover, we know that the only roots of $F_{\alpha_j}^{\theta_i}$ are the $N^{1/(2p_{i,\alpha_i})}(\widetilde{\lambda}-\theta_i)$'s where $\widetilde{\lambda}$ are the eigenvalues of $\widetilde{\mathbf{A}}$.
  - We conclude that there are exactly $\beta_{i,\alpha_i}\times p_{i,\alpha_i}$ eigenvalues $\left({}^{\theta_i}_{p_{i,\alpha_i}}\widetilde{\lambda}_{s,t}\right)_{\substack{1\le s\le p_{i,\alpha_i}\\1\le t\le\beta_{i,\alpha_i}}}$ of $\widetilde{\mathbf{A}}$ such that

$$N^{1/(2p_{i,\alpha_i})}\left({}^{\theta_i}_{p_{i,\alpha_i}}\widetilde{\lambda}_{s,t}-\theta_i\right) \;=\; O(1),$$

and thanks to Lemma 1.5.1, we know that the $N^{1/2p_{i,\alpha_i}}\left(\overset{\theta_i}{\underset{p_{i,\alpha_i}}{}}\widetilde{\lambda}_{s,t}-\theta_i\right)$'s satisfy the equation

$$\det(z^{p_{i,\alpha_i}}-\mathbf{M}^{\theta_i}_{\alpha_i})+o(1) \;=\; 0$$

and so are tighted and converge jointly in distribution to the $p_{i,j}{}^{\text{th}}$ roots of the eigenvalues of $\mathbf{M}^{\theta_i}_{\alpha_i}$. As $\mathbf{M}^{\theta_i}_{\alpha_i}$ is a. s. invertible (recall Remark 1.2.8), none of the $N^{1/2p_{i,\alpha_i}}\left(\overset{\theta_i}{\underset{p_{i,\alpha_i}}{}}\widetilde{\lambda}_{s,t}-\theta_i\right)$'s converge to 0.

- Then, we take the second smallest size p_{i,α_i-1} and work likewise: we know there are exactly

$$\pi_{i,\alpha_i-1}+\beta_{i,\alpha_i-1}\times p_{i,\alpha_i-1}=\beta_{i,\alpha_i}\times p_{i,\alpha_i}+\beta_{i,\alpha_i-1}\times p_{i,\alpha_i-1}$$

eigenvalues of $\widetilde{\mathbf{A}}$ such that

$$N^{1/2p_{i,\alpha_i-1}}\left(\widetilde{\lambda}-\theta_i\right) \;=\; O(1).$$

We know that the eigenvalues $\overset{\theta_i}{\underset{p_{i,\alpha_i}}{}}\widetilde{\lambda}_{s,t}$ $(1\leq s\leq p_{i,\alpha_i},\, 1\leq t\leq \beta_{i,\alpha_i})$ are among them (because $p_{i,\alpha_i-1}>p_{i,\alpha_i}$) so there are $\beta_{i,\alpha_i-1}\times p_{i,\alpha_i-1}$ other eigenvalues $\left(\overset{\theta_i}{\underset{p_{i,\alpha_i-1}}{}}\widetilde{\lambda}_{s,t}\right)_{\substack{1\leq s\leq p_{i,\alpha_i-1}\\ 1\leq t\leq \beta_{i,\alpha_i-1}}}$ of $\widetilde{\mathbf{A}}$ such that

$$N^{1/2p_{i,\alpha_i-1}}\left(\overset{\theta_i}{\underset{p_{i,\alpha_i-1}}{}}\widetilde{\lambda}_{s,t}-\theta_i\right) \;=\; O(1)\,.$$

It follows that $\left(\overset{\theta_i}{\underset{p_{i,\alpha_i-1}}{}}\widetilde{\lambda}_{s,t}\right)_{\substack{1\leq s\leq p_{i,\alpha_i-1}\\ 1\leq t\leq \beta_{i,\alpha_i-1}}}$ converges jointly in distribution to the $p^{\text{th}}_{i,\alpha_{i-1}}$ roots of the eigenvalues of $\mathbf{M}^{\theta_i}_{\alpha_{i-1}}$ (which are almost surely non zero).

- At each step, $\pi_{p_{i,j}}$ corresponds to the number of eigenvalues we have already "discovered" and which go to θ_i faster than $N^{-1/(2p_{i,j})}$ (because $p_{i,\alpha_i}<\cdots<p_{i,1}$), and so it explains the presence of the factor $z^{\pi_{i,j}}$ before $\det(z^{p_{i,j}}-\mathbf{M}^{\theta_i}_j)$ the previous lemma. So one can continue this induction and conclude. that way, we get the exact number of eigenvalues of $\widetilde{\mathbf{A}}$.

It remains now to prove Lemma 1.5.1. We begin with the convergence of $z\mapsto\mathbf{X}^z_N$.

1.5.2 Convergence of $z \mapsto \mathbf{X}_N^z$.

Recall that in order to simplify, we wrote, at (1.2),

$$\mathbf{P} = \mathbf{W}\begin{pmatrix}\mathbf{Po} & 0\\ 0 & 0\end{pmatrix}\mathbf{W}^* = \mathbf{W}\begin{pmatrix}\mathbf{QJQ}^{-1} & 0\\ 0 & 0\end{pmatrix}\mathbf{W}^*,$$

where $\mathbf{J}$ is a Jordan Canonical Form and $\mathbf{W}$ is supposed to be Haar-distributed from (1.19). We also wrote $\mathbf{P} = \mathbf{BC}$ without specifying any choice. For now on, we shall set down

$$\mathbf{B} := \mathbf{W}\begin{pmatrix}\mathbf{QJ}\\ 0\end{pmatrix} \in \mathscr{M}_{N\times r}(\mathbb{C}) \quad \text{and} \quad \mathbf{C} := \begin{pmatrix}\mathbf{Q}^{-1} & 0\end{pmatrix}\mathbf{W}^* \in \mathscr{M}_{r\times N}(\mathbb{C}). \tag{1.30}$$

One can easily notice that

$$\mathbf{CB} = \mathbf{J} \quad ; \quad \mathbf{B}^*\mathbf{B} = \mathbf{J}^*\mathbf{Q}^*\mathbf{QJ} \quad ; \quad \mathbf{CC}^* = \mathbf{Q}^{-1}(\mathbf{Q}^{-1})^*, \tag{1.31}$$

so that all these matrix products do not depend on N.

Remark 1.5.2. With this specific choice, the norm of the matrix $\mathbf{B}$ (resp. $\mathbf{C}$) is uniformly bounded by $\|\mathbf{QJ}\|_{\mathrm{op}}$ (resp. $\|\mathbf{Q}^{-1}\|$) which doesn't depend on N.

For $|z| > b + 2\varepsilon$, we define the $\mathscr{M}_r(\mathbb{C})$-valued random variable

$$\mathbf{X}_n^z := \sqrt{n}\mathbf{C}\left((z\mathbf{I}-\mathbf{A})^{-1} - z^{-1}\right)\mathbf{B}. \tag{1.32}$$

Lemma 1.5.3. *As N goes to infinity, the finite dimensional marginals of $(\mathbf{X}_N^z)_{|z|>b+2\varepsilon}$ converge to the ones of a centered complex Gaussian process $(\mathbf{X}^z = [x_{i,j}^z]_{1\le i,j\le r})_{|z|>b+2\varepsilon}$ such that for all θ, θ' in $\{|z| > b+2\varepsilon\}$,*

- $x_{i,j}^\theta \sim \mathscr{N}\left(0, \frac{b^2}{|\theta|^2}\frac{1}{|\theta|^2-b^2}\cdot \mathbf{e}_i^*\mathbf{CC}^*\mathbf{e}_i \cdot \mathbf{e}_j^*\mathbf{B}^*\mathbf{B}\mathbf{e}_j\right)$,
- $\mathbb{E}\left(x_{i,j}^\theta x_{k,l}^{\theta'}\right) = 0, \ \mathbb{E}\left(x_{i,j}^\theta \overline{x_{k,l}^{\theta'}}\right) = \frac{b^2}{\theta\overline{\theta'}}\frac{1}{\theta\overline{\theta'}-b^2}\cdot \mathbf{e}_i^*\mathbf{CC}^*\mathbf{e}_k \cdot \mathbf{e}_l^*\mathbf{B}^*\mathbf{B}\mathbf{e}_j$.

Recall now that the event $\mathscr{E}_N$ has been defined at Lemma 1.3.2 and has probability tending to one.

Lemma 1.5.4. *There is C finite such that for N large enough, on $\{|z| > b+2\varepsilon\}$,*

$$\mathbb{E}\left(\mathbb{1}_{\mathscr{E}_N}\left\|\frac{\partial}{\partial z}\mathbf{X}_N^z\right\|^4\right) \le C,$$

where $\|\cdot\|$ denotes a norm on $\mathscr{M}_r(\mathbb{C})$.

We deduce, by e.g. [59, Cor. 14.9] (slightly modified because of the presence of $\mathbb{1}_{\mathscr{E}_N}$), that as $N\to\infty$, the random process $(\mathbf{X}_N^z)_{|z|>b+2\varepsilon}$ converges weakly, for the topology of uniform convergence on compact subsets, to the random process $(\mathbf{X}^z)_{|z|>b+2\varepsilon}$

Proof of lemma 1.5.3

Let us fix an integer p, some complex numbers $z_1,\dots,z_p$ from $\{|z|>b+2\varepsilon\}$, some complex numbers $\nu_1,\dots,\nu_p$ and some integers $i_1,j_1,\dots,i_p,j_p$ in $\{1,\dots,r\}$ and define

$$G_N \;:=\; \sum_{t=1}^{p} \nu_t \mathbf{e}_{i_t}^* \mathbf{X}_N^{z_t} \mathbf{e}_{j_t}.$$

At first, we notice that on the event $\mathscr{E}_N$ of Lemma 1.3.2, we can rewrite G_N this way

$$G_N \;=\; \sqrt{N}\sum_{t=1}^{p} \nu_t \mathbf{e}_{i_t}^* \mathbf{C} \sum_{k\geq 1} \frac{\mathbf{A}^k}{z_t^{k+1}} \mathbf{B}\mathbf{e}_{j_t} \;=\; \sqrt{N}\sum_{t=1}^{p} \nu_t \mathbf{c}_t^* \sum_{k\geq 1} \frac{\mathbf{A}^k}{z_t^{k+1}} \mathbf{b}_t.$$

where $\mathbf{b}_t$ designates the j_t-th column of $\mathbf{B}$ and $\mathbf{c}_t$ the i_t-th column of $\mathbf{C}^*$. As $\mathbb{P}(\mathscr{E}_N) \longrightarrow 1$, $\mathscr{E}_N^c$ is irrelevant to weak convergence (see details below at (1.36)), here is what we shall do :

• **Step one :** We set

$$\sigma^2 \;:=\; \sum_{i,i'} \nu_i \overline{\nu}_{i'} \frac{b^2}{z_i \overline{z}_{i'}} \frac{\mathbf{b}_{i'}^* \mathbf{b}_i \mathbf{c}_i^* \mathbf{c}_{i'}}{z_i \overline{z}_{i'} - b^2} > 0, \tag{1.33}$$

and prove that for all fixed integer k_0, there is η_{k_0} such that

$$G_{N,k_0} := \sqrt{N}\sum_{t=1}^{p} \nu_t \sum_{k=1}^{k_0} \frac{\mathbf{c}_t^* \mathbf{A}^k \mathbf{b}_t}{z_t^{k+1}} \;\xrightarrow{(\mathrm{d})}\; Z_{k_0} \overset{(\mathrm{d})}{:=} \mathscr{N}\left(0, \sigma^2 - \eta_{k_0}\right), \tag{1.34}$$

ant that $\eta_{k_0} \to 0$ when $k_0 \to \infty$. Note that σ^2 doesn't depend on N thanks to (1.31).

• **Step two :** We show that the rest shall be neglected for large enough k_0. More precisely, for all $\delta > 0$, we prove that there exists a large enough integer k_0 such that

$$\limsup_{N\to\infty} \mathbb{E}\left(\mathbb{1}_{\mathscr{E}_N} \times \left| \sqrt{N}\sum_{t=1}^{p} \nu_t \sum_{k>k_0} \frac{\mathbf{c}_t^* \mathbf{A}^k \mathbf{b}_t}{z_t^{k+1}} \right| \right) \;\leq\; \delta. \tag{1.35}$$

(for $\mathscr{E}_N$ the event of Lemma 1.3.2 above). After that, we shall easily conclude. Indeed, to prove that G_N converges in distribution to $\mathscr{N}\left(0,\sigma^2\right)$ it suffices to prove that, for any Lipstichtz bounded test function F with Lipschitz constant $\mathscr{L}_F$,

$$\mathbb{E}\left[F(G_N)\right] \;\longrightarrow\; \mathbb{E}\left[F(Z)\right],$$

where Z is a random variable such that $Z \overset{(\mathrm{d})}{=} \mathscr{N}\left(0,\sigma^2\right)$. So, we write

$$\begin{aligned} \left|\mathbb{E}\left[F(G_N) - F(Z)\right]\right| \;&\leq\; \left|\mathbb{E}\left[F(G_N) - F(G_{N,k_0})\right]\right| + \left|\mathbb{E}\left[F(G_{N,k_0}) - F(Z_{k_0})\right]\right| + \qquad (1.36)\\ &\qquad \left|\mathbb{E}\left[F(Z_{k_0}) - F(Z)\right]\right| \\ &\leq\; 2\|F\|_\infty \mathbb{P}(\mathscr{E}_N^c) + \mathscr{L}_F\, \mathbb{E}\left(\mathbb{1}_{\mathscr{E}_N} \times \left| \sqrt{N}\sum_{t=1}^{p} \nu_t \sum_{k>k_0} \frac{\mathbf{c}_t^* \mathbf{A}^k \mathbf{b}_t}{z_t^{k+1}} \right| \right) + \\ &\qquad \left|\mathbb{E}\left[F(G_{N,k_0})\right] - \mathbb{E}\left[F(Z_{k_0})\right]\right| + \mathscr{L}_F\, \mathbb{E}\left|Z_{k_0} - Z\right| \end{aligned}$$

which can be made as small as needed by (1.34) and (1.35) if Z and Z_{k_0} are coupled in the right way.

• **Proof of step one :** *Convergence of the finite sum.*
Let us fix a positive integer k_0. Our goal here is to determine the limits of all the moments of the r.v. G_{N,k_0} defined at (1.34) to conclude it is indeed asymptotically Gaussian. More precisely, we have

Lemma 1.5.5. *There exists $\sigma > 0$ and η_{k_0} such that $\lim_{k_0\to\infty} \eta_{k_0} = 0$ and such that for all large enough k_0 and all non negative distinct integers q, s,*

$$\mathbb{E}\left[|G_{N,k_0}|^{2q}\right] = q!\cdot(\sigma^2 - \eta_{k_0})^q + o(1) \qquad \text{and} \qquad \mathbb{E}\left[G^q_{N,k_0}\overline{G^s_{N,k_0}}\right] = o(1).$$

To prove Lemma 1.5.5, we need to recall a main result about integration with respect to the Haar measure on unitary group, (see [38, Cor. 2.4 and Cor. 2.7]),

Proposition 1.5.6. *Let k be a positive integer and $U = (u_{i,j})$ a Haar-distributed matrix. Let $(i_1,\dots,i_k)$, $(i'_1,\dots,i'_k)$, $(j_1,\dots,j_k)$ and $(j'_1,\dots,j'_k)$ be four k-tuple of $\{1,\dots,N\}$. Then*

$$\mathbb{E}\left[u_{i_1,j_1}\cdots u_{i_k,j_k}\overline{u_{i'_1,j'_1}}\cdots\overline{u_{i'_k,j'_k}}\right] = \sum_{\sigma,\tau\in S_k} \delta_{i_1,i'_{\sigma(1)}}\dots\delta_{i_k,i'_{\sigma(k)}}\delta_{j_1,j'_{\tau(1)}}\dots\delta_{j_k,j'_{\tau(k)}}\mathrm{Wg}(\tau\sigma^{-1}), \tag{1.37}$$

where Wg *is a function called the* Weingarten function. *Moreover, for $\sigma \in S_k$, the asymptotical behavior of* $\mathrm{Wg}(\sigma)$ *is given by*

$$N^{k+|\sigma|}\mathrm{Wg}(\sigma) = \mathrm{Moeb}(\sigma) + O\left(\frac{1}{N^2}\right), \tag{1.38}$$

where $|\sigma|$ denotes the minimal number of factors necessary to write σ as a product of transpositions, and Moeb *denotes a function called the* Möbius function.

Remark 1.5.7. a) The permutation σ for which $\mathrm{Wg}(\sigma)$ will have the largest order is the only one satisfying $|\sigma| = 0$, i.e. $\sigma = id$. As a consequence, the only thing we have to know here about the Möbius function is that $\mathrm{Moeb}(id) = 1$ (see [38]).

b) Notice that if for all $p \neq q$, $i_p \neq i_q$ and $j_p \neq j_q$, then there is at most one non zero term in the RHT of (2.22).

Lemma 1.5.5 follows from the following technical lemma (we use the index m in $\{\cdot\}_m$ to denote a *multiset*, i.e. $\{x_1,\dots,x_k\}_m$ is the class of the k-tuple $(x_1,\dots,x_k)$ under the action of the symmetric group S_k).

Lemma 1.5.8. *Let $k_1,\dots,k_q$ and $l_1,\dots,l_s$ be some positive integers, let $i_1,\dots,i_q,i'_1,\dots,i'_s$ be some integers of $\{1,\dots,r\}$. Then :*

1. *If $\{k_1,\dots,k_q\}_m \neq \{l_1,\dots,l_s\}_m$, we have*

$$\mathbb{E}\left[\sqrt{N}\mathbf{c}^*_{i_1}\mathbf{A}^{k_1}\mathbf{b}_{i_1}\cdots\sqrt{N}\mathbf{c}^*_{i_q}\mathbf{A}^{k_q}\mathbf{b}_{i_q}\overline{\sqrt{N}\mathbf{c}^*_{i'_1}\mathbf{A}^{l_1}\mathbf{b}_{i'_1}}\cdots\overline{\sqrt{N}\mathbf{c}^*_{i'_s}\mathbf{A}^{l_s}\mathbf{b}_{i'_s}}\right] = o(1)$$

2. *In the other case, $s = q$ and one can suppose that $l_1 = k_1, \dots, l_q = k_q$. Under such an assumption, we have*

$$\mathbb{E}\left[\sqrt{N}\mathbf{c}_{i_1}^*\mathbf{A}^{k_1}\mathbf{b}_{j_1}\cdots\sqrt{N}\mathbf{c}_{i_q}^*\mathbf{A}^{k_q}\mathbf{b}_{j_q}\overline{\sqrt{N}\mathbf{c}_{i'_1}^*\mathbf{A}^{l_1}\mathbf{b}_{i'_1}}\cdots\overline{\sqrt{N}\mathbf{c}_{i'_s}^*\mathbf{A}^{l_s}\mathbf{b}_{i'_s}}\right]$$
$$= b^{2(k_1+\cdots+k_q)}\sum_{\sigma\in S_{k_1,\dots,k_q}}\prod_{t=1}^{q}\mathbf{b}_{i'_{\sigma(t)}}^*\mathbf{b}_{i_t}\mathbf{c}_{i_t}^*\mathbf{c}_{i'_{\sigma(t)}}+o(1)$$

where $S_{k_1,\dots,k_q}$ is the set of permutations of $\{1,\dots,q\}$ such that for each $t = 1,\dots,q$, $k_t = k_{\sigma(t)}$.

3. *Moreover,*

$$\sum_{\substack{1\le k_1,\dots,k_q=k_0\\ 1\le k'_1,\dots,k'_q\le k_0}}\mathbb{E}\left[\sqrt{N}\mathbf{c}_{i_1}^*\frac{\mathbf{A}^{k_1}}{z_{i_1}^{k_1+1}}\mathbf{b}_{i_1}\overline{\sqrt{N}\mathbf{c}_{i'_1}^*\frac{\mathbf{A}^{k'_1}}{z_{i'_1}^{k'_1+1}}\mathbf{b}_{i'_1}}\cdots\sqrt{N}\mathbf{c}_{i_q}^*\frac{\mathbf{A}^{k_q}}{z_{i_q}^{k_q+1}}\mathbf{b}_{i_q}\overline{\sqrt{N}\mathbf{c}_{i'_q}^*\frac{\mathbf{A}^{k'_q}}{z_{i'_q}^{k'_q+1}}\mathbf{b}_{i'_q}}\right]$$
$$= \sum_{\sigma\in S_q}\prod_{t=1}^{q}\frac{b^2}{z_{i_t}\overline{z_{i'_{\sigma(t)}}}}\frac{1-\left(\frac{b^2}{z_{i_t}\overline{z_{i'_{\sigma(t)}}}}\right)^{k_0}}{z_{i_t}\overline{z_{i'_{\sigma(t)}}}-b^2}\mathbf{b}_{i'_{\sigma(t)}}^*\mathbf{b}_{i_t}\mathbf{c}_{i_t}^*\mathbf{c}_{i'_{\sigma(t)}}+o(1).$$

Let us briefly explain the main ideas of the proof of this lemma (detailed proof is given in Section 1.6). First, let us recall that $\mathbf{A} = \mathbf{UT}$, so that these expectations expand as sums of terms as

$$\mathbb{E}\left[u_{i_{0,1},i_{1,1}}\cdots u_{i_{k_1-1,1},i_{k_1,1}}u_{i_{0,2},i_{1,2}}\cdots u_{i_{k_r-1,r},i_{k_r,r}}\overline{u_{j_{0,1},j_{1,1}}\cdots u_{j_{l_1-1,1},j_{l_1,1}}u_{j_{0,2},j_{1,2}}\cdots u_{j_{l_s-1,s},j_{l_s,s}}}\right].$$

If the $u_{i,j}$'s were independent and distributed as $\mathcal{N}\left(0,\frac{1}{N}\right)$, the result would be easily proved because most of these expectations would be equal to zero. In our case, the difficulty is that, according to Proposition 2.5.1, lots of expectations do not vanish and they are expressed with the Weingarten function (which is a very complicated function). However, we notice that when these expectations do not vanish as in the Gaussian case, $\mathrm{Wg}(id)$ never occurs in (2.22), so that they are negligible thanks to (2.23).

At last, it is easy to conclude the proof of Lemma 1.5.5 thanks to Lemma 1.5.8. Indeed, for any integers $q \neq s$, we have from (1) of Lemma 1.5.8 that

$\mathbb{E}\left[G_{N,k_0}^q \overline{G_{N,k_0}^s}\right] = o(1)$. Moreover, we have

$$\begin{aligned}
\mathbb{E}\left[\left|G_{N,k_0}\right|^{2q}\right] &= N^q \mathbb{E}\left[\left(\sum_i v_i \mathbf{c}_i^* \sum_{k=1}^{k_0} \frac{\mathbf{A}^k}{z_i^{k+1}} \mathbf{b}_i\right)^q \overline{\left(\sum_{i'} v_{i'} \mathbf{c}_{i'}^* \sum_{k'=1}^{k_0} \frac{\mathbf{A}^{k'}}{\theta_{i'}^{k'+1}} \mathbf{b}_{i'}\right)}^q\right] \\
&= \sum_{\substack{i_1,\ldots,i_q \\ i'_1,\ldots,i'_q}} \prod_{t=1}^q v_{i_t} \overline{v}_{i'_t} \sum_{\substack{k_1,\ldots,k_q \\ k'_1,\ldots,k'_q}} \mathbb{E}\left[\sqrt{N}\mathbf{c}_{i_1}^* \mathbf{A}^{k_1} \mathbf{b}_{i_1} \overline{\sqrt{N}\mathbf{c}_{i'_1}^* \mathbf{A}^{k'_1} \mathbf{b}_{i'_1}} \cdots \sqrt{N}\mathbf{c}_{i_q}^* \mathbf{A}^{k_q} \mathbf{b}_{i_q} \overline{\sqrt{N}\mathbf{c}_{i'_q}^* \mathbf{A}^{k'_q} \mathbf{b}_{i'_q}}\right] \\
&= \sum_{\substack{i_1,\ldots,i_q \\ i'_1,\ldots,i'_q}} \prod_{t=1}^q v_{i_t} \overline{v}_{i'_t} \sum_{\sigma \in S_q} \prod_{t=1}^q \frac{b^2}{z_{i_t}\overline{z}_{i'_{\sigma(t)}}} \frac{1-\left(\frac{b^2}{z_{i_t}\overline{z}_{i'_{\sigma(t)}}}\right)^{k_0}}{z_{i_t}\overline{z}_{i'_{\sigma(t)}} - b^2} \mathbf{b}_{i'_{\sigma(t)}}^* \mathbf{b}_{i_t} \mathbf{c}_{i_t}^* \mathbf{c}_{i'_{\sigma(t)}} + o(1) \\
&= q! \times \left(\sigma^2 - \eta_{k_0}\right)^q + o(1)
\end{aligned}$$

where for σ is given by (1.33) and $|\eta_{k_0}| < \sigma^2 \cdot (\frac{b}{b+2\varepsilon})^{2k_0}$.

• **Proof of step two :** *Vanishing of the tail of the sum.*
Our goal here is to prove that the rest can be neglected, i.e. that for all $\delta > 0$, there exists a large enough integer k_0 such that for any $t \in \{1,\ldots,p\}$ and for $\mathscr{E}_N$ the event of Lemma 1.3.2 above,

$$\limsup_{N\to\infty} \mathbb{E}\left(\mathbb{1}_{\mathscr{E}_N} \times \left|\sqrt{N}\sum_{k>k_0} \frac{\mathbf{c}_t^* \mathbf{A}^k \mathbf{b}_t}{z_t^{k+1}}\right|\right) \leq \delta. \tag{1.39}$$

First, using the fact that

$$\mathbb{E}\left[\mathbb{1}_{\mathscr{E}_N} \times \left|z_t^{-k-1} \mathbf{c}_t^* \mathbf{A}^k \mathbf{b}_t\right|\right] \leq \|\mathbf{B}\|_{\mathrm{op}} \|\mathbf{C}\|_{\mathrm{op}} C_1 \frac{(b+\varepsilon)^k}{(b+2\varepsilon)^{k+1}},$$

it is easy to show that for a large enough positive constant C (depending only on ε), we have

$$\sqrt{N} \sum_{k > C\log N} \mathbb{E}\left[\mathbb{1}_{\mathscr{E}_N} \times \left|z_t^{-k-1} \mathbf{c}_t^* \mathbf{A}^k \mathbf{b}_t\right|\right] = o(1).$$

Now, we only need to prove that

$$\forall \delta > 0,\ \exists k_0,\ \text{ for all } N \text{ large enough}, \quad \sum_{k_0 < k < C\log N} \frac{\sqrt{N}}{|z_t|^{k+1}} \sqrt{\mathbb{E}\left[\mathbb{1}_{\mathscr{E}_N} \times \left|\mathbf{c}_t^* \mathbf{A}^k \mathbf{b}_t\right|^2\right]} \leq \delta.$$

At first, we notice that

$$\mathbb{E}\left(\mathbb{1}_{\mathscr{E}_N} \times \left|\sqrt{N}\sum_{k=k_0}^{C\log N} \frac{\mathbf{c}_t^* \mathbf{A}^k \mathbf{b}_t}{z_t^{k+1}}\right|\right) \leq \sum_{k=k_0}^{C\log N} \frac{\sqrt{N}}{|z_t|^{k+1}} \mathbb{E}\left(\mathbb{1}_{\mathscr{E}_N} \times \left|\mathbf{c}_t^* \mathbf{A}^k \mathbf{b}_t\right|\right) \leq \sum_{k=k_0}^{C\log N} \frac{\sqrt{N}}{|z_t|^{k+1}} \sqrt{\mathbb{E}\left[\mathbb{1}_{\mathscr{E}_N} \times \left|\mathbf{c}_t^* \mathbf{A}^k \mathbf{b}_t\right|^2\right]}$$

Then we condition with respect to the σ-algebra of $\mathbf{U}$, i.e. write

$$\mathbb{E}\left[\mathbb{1}_{\mathscr{E}_N}\times\left|\mathbf{c}_t^*\mathbf{A}^k\mathbf{b}_t\right|^2\right]=\mathbb{E}\left[\mathbb{1}_{\mathscr{E}_N}\times\mathbb{E}_{\mathbf{P}}\left(\left|\mathbf{c}_t^*\mathbf{A}^k\mathbf{b}_t\right|^2\right)\right]=\mathbb{E}\left[\mathbb{1}_{\mathscr{E}_N}\times\mathbb{E}_{\mathbf{P}}\left(\mathbf{c}_t^*\mathbf{A}^k\mathbf{b}_t\mathbf{b}_t^*(\mathbf{A}^*)^k\mathbf{c}_t\right)\right].$$

Let us now remember that we have supposed, at (1.19), that $\mathbf{P}=\mathbf{BC}$ is invariant, in law, by conjugation by any unitary matrix. Hence one can introduce a Haar-distributed unitary matrix $\mathbf{V}$, independent of all other random variables, and write $\mathbf{P}\overset{(\mathrm{d})}{=}\mathbf{VPV}^*$, so that

$$\begin{aligned}\mathbb{E}_{\mathbf{P}}\left(\mathbf{c}_t^*\mathbf{A}^k\mathbf{b}_t\mathbf{b}_t^*(\mathbf{A}^*)^k\mathbf{c}_t\right) &= \mathbb{E}_{\mathbf{P}}\left(\operatorname{Tr}\mathbf{A}^k\mathbf{b}_t\mathbf{b}_t^*(\mathbf{A}^*)^k\mathbf{c}_t\mathbf{c}_t^*\right) \\ &= \mathbb{E}_{\mathbf{P}}\left(\mathbb{E}_{\mathbf{V}}\left(\operatorname{Tr}\mathbf{A}^k\mathbf{V}\mathbf{b}_t\mathbf{b}_t^*\mathbf{V}^*(\mathbf{A}^*)^k\mathbf{V}\mathbf{c}_t\mathbf{c}_t^*\mathbf{V}^*\right)\right),\end{aligned}\tag{1.40}$$

where $\mathbb{E}_{\mathbf{V}}$ denotes the expectation with respect to the randomness of $\mathbf{V}$.

Then, we shall use the following lemma, whose proof is postponed to Section 1.6.4.

Lemma 1.5.9. *Let $\mathbf{V}$ be an $N\times N$ Haar-distributed unitary matrix and let $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ be some deterministic $N\times N$ matrices. Then*

$$\begin{aligned}\mathbb{E}\operatorname{Tr}\mathbf{AVBV}^*\mathbf{CVDV}^* &= \frac{1}{N^2-1}\left\{\operatorname{Tr}\mathbf{AC}\operatorname{Tr}\mathbf{B}\operatorname{Tr}\mathbf{D}+\operatorname{Tr}\mathbf{A}\operatorname{Tr}\mathbf{C}\operatorname{Tr}\mathbf{BD}\right\} \\ &\quad -\frac{1}{N(N^2-1)}\left\{\operatorname{Tr}\mathbf{AC}\operatorname{Tr}\mathbf{BD}+\operatorname{Tr}\mathbf{A}\operatorname{Tr}\mathbf{C}\operatorname{Tr}\mathbf{B}\operatorname{Tr}\mathbf{D}\right\}.\end{aligned}\tag{1.41}$$

By this lemma, one easily gets

$$|\mathbb{E}_{\mathbf{V}}\left(\operatorname{Tr}\mathbf{A}^k\mathbf{V}\mathbf{b}_t\mathbf{b}_t^*\mathbf{V}^*(\mathbf{A}^*)^k\mathbf{V}\mathbf{c}_t\mathbf{c}_t^*\mathbf{V}^*\right)|\le\frac{2}{N-1}\left(\|\mathbf{A}^k\|_{\mathrm{op}}^2+\left|\operatorname{Tr}\left(\mathbf{A}^k\right)\right|^2\right)\|\mathbf{B}\|_{\mathrm{op}}^2\|\mathbf{C}\|_{\mathrm{op}}^2$$

hence as $\mathbf{B}$ and $\mathbf{C}$ are supposed to be bounded, there is a constant C such that

$$\mathbb{E}_{\mathbf{P}}\left(\mathbf{c}_t^*\mathbf{A}^k\mathbf{b}_t\mathbf{b}_t^*(\mathbf{A}^*)^k\mathbf{c}_t\right)\le\frac{C}{N}\left(\|\mathbf{A}^k\|_{\mathrm{op}}^2+\left|\operatorname{Tr}\left(\mathbf{A}^k\right)\right|^2\right).$$

Then, we use the following lemma, a weaker version of [15, Theorem 1].

Lemma 1.5.10. *There exists a positive constant K such that for all $k\le C\log N$, for all large enough N,*

$$\mathbb{E}\left[\left|\operatorname{Tr}\left(\mathbf{A}^k\right)\right|^2\right]\;\le\;K\,(b+\varepsilon)^{2k}.$$

By (1.40) and Lemma 1.5.10, for all $k\le C\log N$, there exists some positive constant C' such that

$$\sqrt{\mathbb{E}\left[\mathbb{1}_{\mathscr{E}_N}\times\left|\mathbf{c}_t^*\mathbf{A}^k\mathbf{b}_t\right|^2\right]}\le\frac{C'(b+\varepsilon)^k}{\sqrt{N}}.$$

Hence as $|z_t|\ge b+2\varepsilon$ for N large enough, (1.39) is proved.

Proof of Lemma 1.5.4

The proof relies on the same tricks of the proof of Lemma 1.5.3, using the already noticed fact that for $|z| > b + 2\varepsilon$,

$$\mathbb{1}_{\mathscr{E}_N} \mathbf{X}_N^z = \sqrt{N} \mathbb{1}_{\mathscr{E}_N} \sum_{k \geq 1} \mathbf{C} \frac{\mathbf{A}^k}{z^{k+1}} \mathbf{B},$$

so that

$$\frac{\partial}{\partial z} \mathbb{1}_{\mathscr{E}_N} \mathbf{X}_N^z = -\sqrt{N} \mathbb{1}_{\mathscr{E}_N} \sum_{k \geq 1} (k+1) \mathbf{C} \frac{\mathbf{A}^k}{z^{k+2}} \mathbf{B}.$$

1.5.3 Proof of Lemma 1.5.1

To prove Lemma 1.5.1, we shall need to do a Taylor expansion of $F_j^{\theta_i}(z)$. From now on, we fix a compact set K and consider $z \in K$. Recall that $F_j^{\theta_i}(z)$ and $\mathbf{X}_N^z$ have been defined respectively at (1.29) and (1.32) as

$$F_j^{\theta_i}(z) = \det\left(\mathbf{I} - \mathbf{C}\left(\left(\theta_i + N^{-1/(2p_{i,j})} z\right) \mathbf{I} - \mathbf{A}\right)^{-1} \mathbf{B}\right) \qquad \mathbf{X}_N^z = \sqrt{N} \mathbf{C}\left((z\mathbf{I} - \mathbf{A})^{-1} - z^{-1}\right) \mathbf{B},$$

hence, using Lemma 1.5.4 and the convergence of $\mathbf{X}_N^z$ to $\mathbf{X}^z$ established at Section 1.5.2,

$$\begin{aligned} F_j^{\theta_i}(z) &= \det\left(\mathbf{I} - \frac{1}{\theta_i + N^{-1/(2p_{i,j})} z} \mathbf{J} - \frac{1}{\sqrt{N}} \mathbf{X}_N^{\theta_i + N^{-1/(2p_{i,j})} z}\right) \\ &= \det\left(\mathbf{I} - \theta_i^{-1} \mathbf{J} + \theta_i^{-1}\left(1 - \frac{1}{1 + N^{-1/(2p_{i,j})} z \theta_i^{-1}}\right) \mathbf{J} - \frac{1}{\sqrt{N}} \mathbf{X}^{\theta_i} + o\left(\frac{1}{\sqrt{N}}\right)\right) \\ &= \det\left(\mathbf{I} - \theta_i^{-1} \mathbf{J} + z \delta_N \theta_i^{-2} \mathbf{J} + \frac{1}{\sqrt{N}} \mathbf{G}\right) \end{aligned}$$

where we define

$$\delta_N := \frac{\theta_i}{z}\left(1 - \frac{1}{1 + N^{-1/(2p_{i,j})} z \theta_i^{-1}}\right) = N^{-1/(2p_{i,j})}(1 + o(1)) \qquad \text{and} \qquad \mathbf{G} = -\mathbf{X}^{\theta_i} + o(1).$$

Let us write $\mathbf{J}$ by blocks

$$\mathbf{J} = \left(\begin{array}{c|c|c} * & (0) & (0) \\ \hline (0) & \mathbf{J}(\theta_i) & (0) \\ \hline (0) & (0) & * \end{array}\right)$$

where $\mathbf{J}(\theta_i)$ is the part with the blocks associated to θ_i. And so, we write

$$\mathbf{I} - \theta_i^{-1} \mathbf{J} = \left(\begin{array}{c|c|c} \mathbf{N}' & (0) & (0) \\ \hline (0) & \mathbf{N} & (0) \\ \hline (0) & (0) & \mathbf{N}'' \end{array}\right)$$

where $\mathbf{N}'$ and $\mathbf{N}''$ are invertible matrices and $\mathbf{N}$ is the diagonal by blocks matrix

$$\begin{aligned}\mathbf{N} &= \mathbf{I}-\theta_i^{-1}\operatorname{diag}(\underbrace{\mathbf{R}_{p_{i,1}}(\theta_i),\ldots,\mathbf{R}_{p_{i,1}}(\theta_i)}_{\beta_{i,1}\text{ blocks}},\ldots\ldots,\underbrace{\mathbf{R}_{p_{i,\alpha_i}}(\theta_i),\ldots,\mathbf{R}_{p_{i,\alpha_i}}(\theta_i)}_{\beta_{i,\alpha_i}\text{ blocks}}) \\ &= -\theta_i^{-1}\operatorname{diag}(\underbrace{\mathbf{R}_{p_{i,1}}(0),\ldots,\mathbf{R}_{p_{i,1}}(0)}_{\beta_{i,1}\text{ blocks}},\ldots\ldots,\underbrace{\mathbf{R}_{p_{i,\alpha_i}}(0),\ldots,\mathbf{R}_{p_{i,\alpha_i}}(0)}_{\beta_{i,\alpha_i}\text{ blocks}}) \qquad (1.42)\end{aligned}$$

with $\mathbf{R}_p(\theta)$ as defined at (1.3) for p an integer and $\theta\in\mathbb{C}$.

Let us now expand the determinant $\det\left(\mathbf{I}-\theta_i^{-1}\mathbf{J}+z\delta_N\theta_i^{-2}\mathbf{J}+N^{-1/2}\mathbf{G}\right)$ using the *columns replacement approach* of following formula, where the M_k's and the H_k's are the columns of two $r\times r$ matrices $\mathbf{M}$ and $\mathbf{H}$ (that one will think of as an error term, even though the formula below is exact)

$$\begin{aligned}\det(\mathbf{M}+\mathbf{H}) &= \det(\mathbf{M})+\sum_{k=1}^{r}\det\left(M_1\middle|M_2\middle|\ldots\middle|H_k\middle|\ldots\middle|M_r\right)+\sum_{1\le k_1<k_2\le r}\det\left(M_1\middle|\ldots\middle|H_{k_1}\middle|\ldots\middle|H_{k_2}\middle|\ldots\middle|M_r\right) \\ &\quad+\ \ldots\ +\sum_{1\le k_1<k_2<\cdots<k_s\le r}\det\left(M_1\middle|\ldots\middle|H_{k_1}\middle|\ldots\middle|H_{k_s}\middle|\ldots\middle|M_r\right) \\ &\quad+\ \ldots\ +\sum_{k=1}^{r}\det\left(H_1\middle|H_2\middle|\ldots\middle|M_k\middle|\ldots\middle|H_r\right)+\det(\mathbf{H})\end{aligned}$$

We shall use this formula with $\mathbf{M}=\mathbf{I}-\theta_i^{-1}\mathbf{J}$ and $\mathbf{H}=z\delta_N\theta_i^{-2}\mathbf{J}+N^{-1/2}\mathbf{G}$, and we shall keep only higher terms. It means that the determinant is a summation of determinants of $\mathbf{M}$ where some of the columns are replaced by the corresponding column of $z\delta_N\theta_i^{-2}\mathbf{J}$ or of $N^{-1/2}\mathbf{G}$. Recall that $\mathbf{M}$ has several columns of zeros (the ones corresponding to null columns of $\mathbf{N}$), so we know that we have to replace at least these columns to get a non-zero determinant. Moreover, we won't replace the columns of $\mathbf{N}'$ or $\mathbf{N}''$ because this would necessarily make appear negligible terms (recall that $\mathbf{N}'$ and $\mathbf{N}''$ are invertible), so all the non-negligible determinants will be factorizable by $\det(\mathbf{N}')\det(\mathbf{N}'')$. So now, let us understand what are the non-negligible terms in the summation.

To make things clear, let us start with an example. We choose $p_{i,j}=3$ and the matrix $\mathbf{N}$ given, via (1.42), by

$$\mathbf{N} = -\theta_i^{-1}\left(\begin{array}{cccc|ccc|c} 0&1&0&0& & & & \\ 0&0&1&0& &(0)& & \\ 0&0&0&1& & & & \\ 0&0&0&0& & & & \\ \hline & & & &0&1&0& \\ & & & &0&0&1& \\ &(0)& & &0&0&0& \\ \hline & & & & & & &0\end{array}\right),$$

we know we have to replace at least 3 columns (the first, the fifth and the last ones) which correspond to the first column of each diagonal blocks, and we shall deal with one block at the time. Let us deal with the first one. If we replace this column by the corresponding column of $z\delta_N\theta_i^{-2}\mathbf{J}$, we get

$$z\delta_N \begin{vmatrix} \frac{1}{\theta_i} & \frac{1}{\theta_i} & 0 & 0 \\ 0 & 0 & \frac{1}{\theta_i} & 0 \\ & & 0 & \frac{1}{\theta_i} \\ & (0) & & \\ & & 0 & 0 \end{vmatrix}$$

We see that in this case, some non linearly independent columns appear. It follows that once one has replaced a null column by a column from $z\delta_N\theta_i^{-2}\mathbf{J}$, the whole block needs to be replaced to get a non zero determinant :

$$z\delta_N \begin{vmatrix} \frac{1}{\theta_i} & \frac{1}{\theta_i} & 0 & 0 \\ 0 & 0 & \frac{1}{\theta_i} & 0 \\ & & 0 & \frac{1}{\theta_i} \\ & (0) & & \\ & & 0 & 0 \end{vmatrix} \longrightarrow (z\delta_N)^2 \begin{vmatrix} \frac{1}{\theta_i} & \frac{1}{\theta_i^2} & 0 & 0 \\ 0 & \frac{1}{\theta_i} & \frac{1}{\theta_i} & 0 \\ & & 0 & \frac{1}{\theta_i} \\ & (0) & & \\ & & 0 & 0 \end{vmatrix} \longrightarrow (z\delta_N)^3 \begin{vmatrix} \frac{1}{\theta_i} & \frac{1}{\theta_i^2} & 0 & 0 \\ 0 & \frac{1}{\theta_i} & \frac{1}{\theta_i^2} & 0 \\ & & \frac{1}{\theta_i} & \frac{1}{\theta_i} \\ & (0) & & \\ & & 0 & 0 \end{vmatrix}$$

$$\longrightarrow (z\delta_N)^4 \begin{vmatrix} \frac{1}{\theta_i} & \frac{1}{\theta_i^2} & 0 & 0 \\ 0 & \frac{1}{\theta_i} & \frac{1}{\theta_i^2} & 0 \\ & & \frac{1}{\theta_i} & \frac{1}{\theta_i^2} \\ & (0) & & \\ & & 0 & \frac{1}{\theta_i} \end{vmatrix},$$

Another possibility to fill a null column would be to replace it by the corresponding one in $N^{-1/2}\mathbf{G}$:

$$\begin{vmatrix} g_{1,1} & \frac{1}{\theta_i} & 0 & 0 \\ g_{2,1} & 0 & \frac{1}{\theta_i} & 0 \\ g_{3,1} & 0 & 0 & \frac{1}{\theta_i} \\ g_{4,1} & 0 & 0 & 0 \end{vmatrix}$$

We obtain an invertible block directly (i.e. without having to replace the whole block as above). However, in this example, $\delta_N \gg N^{-1/2}$ (because $p_{i,j} = 3$), this term might be negligible. If $\delta_N^4 \gg N^{-1/2}$, then first choice is relevant (the other would be negligible), or else, if $\delta_N^4 \ll N^{-1/2}$, we would make the second choice.

Our strategy is to choose $\mathbf{J}$ on the blocks of size $p < p_{i,j}$ (because $\delta_N^p \gg N^{-1/2}$) and $\mathbf{G}$ on the blocks of size $p > p_{i,j}$ (because $\delta_N^p \ll \frac{1}{\sqrt{N}}$). For the blocks of size $p = p_{i,j}$, we can choose both (because $\delta_N^{p_{i,j}} \approx \frac{1}{\sqrt{N}}$). So in our example, the non negligible terms are

$$\det\left(\frac{G_1}{\sqrt{N}}\Big|N_2\Big|N_3\Big|N_4\Big|\frac{z\delta_n}{\theta_i^2}J_5\Big|\frac{z\delta_n}{\theta_i^2}J_6\Big|\frac{z\delta_N}{\theta_i^2}J_7\Big|\frac{z\delta_N}{\theta_i^2}J_8\right) = -z^4\cdot\frac{\delta_N^4}{\sqrt{N}}\begin{vmatrix}
g_{1,1} & \frac{1}{\theta_i} & 0 & 0 & & & & \\
g_{2,1} & 0 & \frac{1}{\theta_i} & 0 & & (0) & & \\
g_{3,1} & 0 & 0 & \frac{1}{\theta_i} & & & & \\
g_{4,1} & 0 & 0 & 0 & & & & \\
g_{5,1} & & & & \frac{1}{\theta_i} & \frac{1}{\theta_i^2} & 0 & 0 \\
g_{6,1} & & (0) & & 0 & \frac{1}{\theta_i} & \frac{1}{\theta_i^2} & 0 \\
g_{7,1} & & & & 0 & 0 & \frac{1}{\theta_i} & 0 \\
g_{8,1} & & & & 0 & 0 & 0 & \frac{1}{\theta_i}
\end{vmatrix}$$

and

$$\det\left(\frac{G_1}{\sqrt{N}}\Big|N_2\Big|N_3\Big|N_4\Big|\frac{G_5}{\sqrt{N}}\Big|N_6\Big|N_7\Big|\frac{z\delta_N}{\theta_i^2}J_8\right) = -z\cdot\frac{\delta_N}{N}\cdot\begin{vmatrix}
g_{1,1} & \frac{1}{\theta_i} & 0 & 0 & g_{1,5} & & & \\
g_{2,1} & 0 & \frac{1}{\theta_i} & 0 & g_{2,5} & & (0) & \\
g_{3,1} & 0 & 0 & \frac{1}{\theta_i} & g_{3,5} & & & \\
g_{4,1} & 0 & 0 & 0 & g_{4,5} & & & \\
g_{5,1} & & & & g_{5,5} & \frac{1}{\theta_i} & 0 & 0 \\
g_{6,1} & & (0) & & g_{6,5} & 0 & \frac{1}{\theta_i} & 0 \\
g_{7,1} & & & & g_{7,5} & 0 & 0 & 0 \\
g_{8,1} & & & & g_{8,5} & 0 & 0 & \frac{1}{\theta_i}
\end{vmatrix}$$

and one can easily notice that the sum of the non negligible terms is

$$\begin{aligned}
&\det\left(\frac{G_1}{\sqrt{N}}\Big|N_2\Big|N_3\Big|N_4\Big|\frac{z\delta_N}{\theta_i^2}J_5\Big|\frac{z\delta_N}{\theta_i^2}J_6\Big|\frac{z\delta_N}{\theta_i^2}J_7\Big|\frac{z\delta_N}{\theta_i^2}J_8\right)+\det\left(\frac{G_1}{\sqrt{N}}\Big|N_2\Big|N_3\Big|N_4\Big|\frac{G_5}{\sqrt{N}}\Big|N_6\Big|N_7\Big|\frac{z\delta_N}{\theta_i^2}J_8\right)\\
&= \frac{z}{\theta_i^6}\frac{1}{N^{1+\frac{1}{2p_{i,j}}}}\begin{vmatrix} g_{4,1} & g_{4,5} \\ g_{7,1} & g_{7,5}-\frac{z^3}{\theta_i}\end{vmatrix} + o\left(\frac{1}{N^{1+\frac{1}{2p_{i,j}}}}\right).
\end{aligned}$$

Now that this example is well understood, let us treat the general case :

– We know that there are $\beta_{i,1}+\cdots+\beta_{i,j-1}$ blocks of size larger than $p_{i,j}$ so we will replace the first column of each of these blocks by the corresponding column of $N^{-1/2}\mathbf{G}$.

– For all the blocks of lower size, we replace all the columns by the corresponding column of $z\delta_N\theta_i^{-2}\mathbf{J}$. The number of such columns is $\pi_{i,j} := \beta_{i,j+1}\times p_{i,j+1}+\cdots\beta_{i,\alpha_i}\times p_{i,\alpha_i}$.

– We also know that there are $\beta_{i,j}$ blocks of size $p_{i,j}$ and for each block, we have two choices so that represents $2^{\beta_{i,j}}$ non negligible terms.

And so, we conclude that :

- The statement holds for $\gamma_{i,j} = \frac{1}{2}\sum_{l=1}^{j-1}\beta_{i,l} + \frac{\pi_{i,j}}{2p_{i,j}} + \frac{1}{2}\beta_{i,j}$.

- All the non negligible terms are factorizable by $z^{\pi_{i,j}}$.

- Using notations from (1.8), we define the matrices

$$\mathrm{M}_j^{\theta_i,\mathrm{I}} := [g_{k,\ell}^{\theta_i}]_{\substack{k\in K(i,j)^- \\ \ell\in L(i,j)^-}} \qquad \mathrm{M}_j^{\theta_i,\mathrm{II}} := [g_{k,\ell}^{\theta_i}]_{\substack{k\in K(i,j)^- \\ \ell\in L(i,j)}}$$

$$\mathrm{M}_j^{\theta_i,\mathrm{III}} := [g_{k,\ell}^{\theta_i}]_{\substack{k\in K(i,j) \\ \ell\in L(i,j)^-}} \qquad \mathrm{M}_j^{\theta_i,\mathrm{IV}} := [g_{k,\ell}^{\theta_i}]_{\substack{k\in K(i,j) \\ \ell\in L(i,j)}}$$

and with a simple calculation, one can sum up all the non-negligible terms by

$$C\cdot\frac{z^{\pi_{i,j}}}{N^{\gamma_{i,j}}}\begin{vmatrix} \mathrm{M}_j^{\theta_i,\mathrm{I}} & \mathrm{M}_j^{\theta_i,\mathrm{II}} \\ \mathrm{M}_j^{\theta_i,\mathrm{III}} & \mathrm{M}_j^{\theta_i,\mathrm{IV}} - \frac{z^{p_{i,j}}}{\theta_i}\mathbf{I}_{\beta_{i,j}} \end{vmatrix} + o\left(\frac{1}{N^{\gamma_{i,j}}}\right)$$

where C is a deterministic constant equal to $\pm$ a power of θ_i^{-1}. Then, using a well-know formula (see for example Eq. (A1) of [2] p. 414), we have

$$\begin{vmatrix} \mathrm{M}_j^{\theta_i,\mathrm{I}} & \mathrm{M}_j^{\theta_i,\mathrm{II}} \\ \mathrm{M}_j^{\theta_i,\mathrm{III}} & \mathrm{M}_j^{\theta_i,\mathrm{IV}} - \frac{z^{p_{i,j}}}{\theta_i}\mathbf{I}_{\beta_{i,j}} \end{vmatrix} = \theta_i^{-\beta_{i,j}}\det\left(\mathrm{M}_j^{\theta_i,\mathrm{I}}\right)\det\left(\theta_i(\mathrm{M}_j^{\theta_i,\mathrm{IV}} - \mathrm{M}_j^{\theta_i,\mathrm{III}}(\mathrm{M}_j^{\theta_i,\mathrm{I}})^{-1}\mathrm{M}_j^{\theta_i,\mathrm{II}}) - z^{p_{i,j}}\mathbf{I}_{\beta_{i,j}}\right).$$

- Thanks to Lemma 1.5.3, we know that

$$\mathbb{E}\left(m_{k,\ell}^{\theta_i} m_{k',\ell'}^{\theta_{i'}}\right) = 0, \qquad \mathbb{E}\left(m_{k,\ell}^{\theta_i}\overline{m_{k',\ell'}^{\theta_{i'}}}\right) = \frac{b^2}{\theta_i\overline{\theta_{i'}}}\frac{1}{\theta_i\overline{\theta_{i'}} - b^2}\mathbf{e}_k^*\mathbf{C}\mathbf{C}^*\mathbf{e}_{k'}\,\mathbf{e}_{\ell'}^*\mathbf{B}^*\mathbf{B}\,\mathbf{e}_\ell,$$

and from (1.31), we write

$$\mathbf{e}_k^*\mathbf{C}\mathbf{C}^*\mathbf{e}_{k'}\,\mathbf{e}_{\ell'}^*\mathbf{B}^*\mathbf{B}\mathbf{e}_\ell = \mathbf{e}_k^*\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^*\mathbf{e}_{k'}\,\mathbf{e}_{\ell'}^*\mathbf{J}^*\mathbf{Q}^*\mathbf{Q}\mathbf{J}\mathbf{e}_\ell$$

then, from the definition of the set $L(i,j)$, we know that if $\ell\in L(i,j)$ then $\mathbf{J}\mathbf{e}_\ell = \theta_i\mathbf{e}_\ell$, so, finally,

$$\mathbb{E}\left(m_{k,\ell}^{\theta_i} m_{k',\ell'}^{\theta_{i'}}\right) = 0, \qquad \mathbb{E}\left(m_{k,\ell}^{\theta_i}\overline{m_{k',\ell'}^{\theta_{i'}}}\right) = \frac{b^2}{\theta_i\overline{\theta_{i'}} - b^2}\mathbf{e}_k^*\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^*\mathbf{e}_{k'}\,\mathbf{e}_{\ell'}^*\mathbf{Q}^*\mathbf{Q}\,\mathbf{e}_\ell.$$

1.6 Proofs of the technical results

1.6.1 Proofs of Lemmas 1.3.2 and 1.3.3

Lemma 1.6.1. *There exists a constant C_1, independent of N, such that with probability tending to one,*

$$\sup_{|z|=b+\varepsilon} \left\|(z\mathbf{I}-\mathbf{A})^{-1}\right\|_{\mathrm{op}} \quad \leq \quad C_1.$$

Proof. Note that for any $\eta > 0$,

$$\|(z\mathbf{I}-\mathbf{A})^{-1}\|_{\mathrm{op}} \leq \frac{1}{\eta} \iff \nu^z([-\eta,\eta]) = 0,$$

where $\nu^z := \frac{1}{2N}\sum_{i=1}^N(\delta_{-s_i^z}+\delta_{s_i^z})$ and the s_i^z's are the singular values of $z\mathbf{I}-\mathbf{A}$.

By Corollary 10 of [50], for any z such that $|z| > b$, there is β_z such that with probability tending to one,

$$\nu^z([-\beta_z,\beta_z]) = 0.$$

It follows from standard perturbation inequalities that with probability tending to one, for any z' such that $|z'-z| < \frac{\beta_z}{2}$,

$$\nu^{z'}([-\frac{\beta_z}{2},\frac{\beta_z}{2}]) = 0.$$

Then with a compacity argument, one concludes easily. □

Proof of Lemma 1.3.2

Note first that thanks to the Cauchy formula, for all $x \in \mathbb{C}$,

$$|x| < b+\varepsilon \quad \Longrightarrow \quad \forall k \geq 0, \quad x^k = \frac{1}{2i\pi}\int_{|z|=b+\varepsilon}\frac{z^k}{z-x}dz.$$

Moreover, by [50, Th. 2], the spectral radius of $\mathbf{A}$ converges in probability to b, so that with probability tending to one, by application of the holomorphic functional calculus (which is working even for non-Hermitian matrices) to $\mathbf{A}$,

$$\forall k \geq 0, \qquad \mathbf{A}^k \quad = \quad \frac{1}{2i\pi}\int_{|z|=b+\varepsilon} z^k\,(z-\mathbf{A})^{-1}\,dz.$$

Thus with probability tending to one,

$$\forall k \geq 0, \qquad \|\mathbf{A}^k\|_{\mathrm{op}} \quad \leq \quad \frac{1}{2\pi}\sup_{|z|=b+\varepsilon}\|(z\mathbf{I}-\mathbf{A})^{-1}\|_{\mathrm{op}} \times \int_{|z|=b+\varepsilon}|z|^k dz.$$

Then one concludes using the previous lemma.

Proof of Lemma 1.3.3

Since $\mathbf{CA}^k\mathbf{B}$ is a square $r \times r$ matrix, it suffices to prove that each entry tends, in probability, to 0. And since $\mathbf{B}$ and $\mathbf{C}$ are uniformly bounded (see Remark 1.5.2) , one just has to show that for all unit vectors $\mathbf{b}$ and $\mathbf{c}$,

$$\mathbf{c}^*\mathbf{A}^k\mathbf{b} \quad \xrightarrow{(\mathbb{P})} \quad 0. \tag{1.43}$$

Recall that $\mathbf{A} = \mathbf{UT}$ and

$$\mathbf{c}^*\mathbf{A}^k\mathbf{b} \quad = \quad \sum_{i_0,i_1,\ldots,i_k} \overline{c_{i_0}} b_{i_k} u_{i_0,i_1} s_{i_1} u_{i_1,i_2} s_{i_2} \cdots u_{i_{k-1},i_k} s_{i_k},$$

and so we have

$$\mathbb{E}_{\mathbf{U}} \left|\mathbf{c}^*\mathbf{A}^k\mathbf{b}\right|^2 \quad = \quad \sum_{\substack{i_0,\ldots,i_k \\ j_0,\ldots,j_k}} \overline{c_{i_0}} c_{j_0} b_{i_k} \overline{b_{j_k}} s_{i_1} s_{j_1} \ldots s_{i_k} s_{j_k} \, \mathbb{E}\left[u_{i_0,i_1} u_{i_1,i_2} \cdots u_{i_{k-1},i_k} \overline{u}_{j_0,j_1} \overline{u}_{j_1,j_2} \ldots \overline{u}_{j_{k-1},j_k}.\right]$$

Let $(i_1,\ldots,i_k),(i'_1,\ldots,i'_k),(j_1,\ldots,j_k)$ and $(j'_1,\ldots,j'_k)$ be k-tuples of intergers lower than n. By Proposition 2.5.1, we know that

$$\mathbb{E}\left[u_{i_1,j_1} \cdots u_{i_k,j_k} \overline{u}_{i'_1,j'_1} \cdots \overline{u}_{i'_k,j'_k}\right] \neq 0$$

if and only if there are two permutations σ and τ so that for all $p \in \{1,\ldots,k\}$, $i_{\sigma(p)} = i'_p$ and $j_{\tau(p)} = j'_p$. In our case, we know that for a $(i_0,\ldots,i_k)$ fixed, there will be no more than $(k+1)!$ tuples $(j_0,\ldots,j_k)$ leading to a non-zero expectation. By Proposition 2.5.1 again, we know that all these expectations are $O\left(N^{-k}\right)$. So, one concludes with the following computation

$$\begin{aligned}
\mathbb{E}_{\mathbf{U}} |\mathbf{c}^*\mathbf{A}^k\mathbf{b}|^2 \quad &\leq \quad \sum_{\mu \in S_{k+1}} \sum_{i_0,\ldots,i_k} \left|\overline{c}_{i_0} c_{i_{\mu(0)}} b_{i_k} \overline{b}_{i_{\mu(k)}}\right| s_{i_1}^2 \ldots s_{i_k}^2 \times O\left(\frac{1}{N^k}\right) \\
&\leq \quad \sum_{\mu \in S_{k+1}} \sum_{i_0,\ldots,i_k} \frac{1}{2}\left[|c_{i_0}|^2 |b_{i_k}|^2 + |c_{i_{\mu(0)}}|^2 |b_{i_{\mu(k)}}|^2\right] s_{i_1}^2 \ldots s_{i_k}^2 \times O\left(\frac{1}{N^k}\right) \\
&\leq \quad (k+1)! \sum_{i_0,\ldots,i_k} |c_{i_0}|^2 |b_{i_k}|^2 s_{i_1}^2 \cdots s_{i_k}^2 O\left(N^{-k}\right) \\
&\leq \quad \left(\frac{1}{N} \sum_{i=1}^N s_i^2\right)^{k-1} \left(\sum_{j=1}^N |b_j|^2 s_j^2\right) \times O\left(\frac{1}{N}\right) \; = \; O\left(\frac{1}{N}\right).
\end{aligned}$$

1.6.2 Proof of Lemma 1.4.1

Lemma 1.4.1 is a direct consequence of the following lemma.

Lemma 1.6.2. *Let μ be a probability measure whose support is contained in an interval $[m,M] \subset]0,+\infty[$. Let μ^{-1} denote the push-forward of μ by the map $t \longmapsto 1/t$. Then for all $x \in \mathbb{R}$, $y > 0$,*

$$|\operatorname{Im} G_{\mu^{-1}}(x+iy)| \leq \begin{cases} M & \text{if } x \notin [1/(2M), 2/m], \\ \frac{8M^4}{m^2} \left|\operatorname{Im} G_\mu\left(\frac{1}{x} + i\frac{m^2 y}{2}\right)\right| & \text{otherwise.} \end{cases} \tag{1.44}$$

Proof. Note that

$$|\operatorname{Im} G_{\mu^{-1}}(x+iy)| = \int \frac{y}{(x-1/t)^2 + y^2} \mu(dt).$$

If $x \notin [1/(2M), 2/m]$, then for all $t \in [m,M]$, $|x - 1/t| \geq 1/(2M)$, and (1.44) follows from the fact that for all $y > 0$, we have

$$\frac{y}{y^2 + (1/(2M))^2} \leq M.$$

If $x \in [1/(2M), 2/m]$, then for all $t \in [m,M]$,

$$\frac{1}{2M^2} \leq \frac{x}{t} \leq \frac{2}{m^2}$$

hence

$$\left(x - \frac{1}{t}\right)^2 + y^2 = \frac{x^2}{t^2}\left(\left(\frac{1}{x} - t\right)^2 + \left(\frac{yt}{x}\right)^2\right) \geq \frac{1}{4M^4}\left(\left(\frac{1}{x} - t\right)^2 + \left(\frac{m^2 y}{2}\right)^2\right)$$

and (1.44) follows directly. □

1.6.3 Proof of Lemma 1.5.8

First of all, as $\|\mathbf{B}\|_{\text{op}}$ and $\|\mathbf{C}\|_{\text{op}}$ are bounded (see Remark 1.5.2) and for any unitary matrix $\mathbf{V}$, $\mathbf{VB} \overset{(d)}{=} \mathbf{B}$ and $\mathbf{CV}^* \overset{(d)}{=} \mathbf{C}$, we know, by for example Theorem 2 of [54], that there is constant C such that with a probability tending to one,

$$\forall N \geq 1, \quad \max_{\substack{1 \leq i \leq N \\ 1 \leq j \leq r}} |b_{i,j}| \leq C\sqrt{\frac{\log N}{N}} \quad \text{and} \quad \max_{\substack{1 \leq i \leq N \\ 1 \leq j \leq r}} |c_{j,i}| \leq C\sqrt{\frac{\log N}{N}}. \tag{1.45}$$

Outline of the proof

If we expand the following expectation

$$\mathbb{E}\left[\mathbf{c}_{i_1}^* \mathbf{A}^{k_1} \mathbf{b}_{i_1} \cdots \mathbf{c}_{i_q}^* \mathbf{A}^{k_q} \mathbf{b}_{i_q} \overline{\mathbf{c}_{i'_1}^* \mathbf{A}^{l_1} \mathbf{b}_{i'_1}} \cdots \overline{\mathbf{c}_{i'_s}^* \mathbf{A}^{l_s} \mathbf{b}_{i'_s}}\right]$$

(where the expectation is with respect to the randomness of $\mathbf{U}$), we get a summation of terms such as

$$\mathbb{E}\left[\underbrace{u_{t_{1,1}t_{1,2}}\cdots u_{t_{1,k_1}t_{1,k_1+1}}}_{k_1\text{ factors}}\cdots\underbrace{u_{t_{q,1}t_{q,2}}\cdots u_{t_{q,k_q}t_{q,k_q+1}}}_{k_q\text{ factors}}\underbrace{\overline{u}_{t'_{1,1}t'_{1,2}}\cdots\overline{u}_{t'_{1,l_1}t'_{1,l_1+1}}}_{l_1\text{ factors}}\cdots\underbrace{\overline{u}_{t'_{s,1}t'_{s,2}}\cdots\overline{u}_{t'_{s,l_s}t'_{s,l_s+1}}}_{l_s\text{ factors}}\right]. \tag{1.46}$$

Our goal is to find out which of these terms will be negligible before the others. First, we know from Proposition 2.5.1 that the expectation vanishes unless the set of the first indices (resp. second) of the u_{ij}'s is the same as the set of the first indices (resp. second) of the $\overline{u}_{ij}$'s. Secondly, each expectation is computed thanks to the following formula (see Proposition 2.5.1):

$$\mathbb{E}\left[u_{i_1,j_1}\cdots u_{i_k,j_k}\overline{u_{i'_1,j'_1}}\cdots\overline{u_{i'_k,j'_k}}\right] = \sum_{\sigma,\tau\in S_k}\delta_{i_1,i'_{\sigma(1)}}\ldots\delta_{i_k,i'_{\sigma(k)}}\delta_{j_1,j'_{\tau(1)}}\ldots\delta_{j_k,j'_{\tau(k)}}\mathrm{Wg}(\tau\sigma^{-1}). \tag{1.47}$$

Then, by (2.23) of Proposition 2.5.1, we know that the prevailing terms are the ones involving $\mathrm{Wg}(id)$, i.e. those allowing to match the i's in the u_{ij}'s with the i''s in the $\overline{u}_{i'j'}$'s thanks to a permutation which also matches j's in the u_{ij}'s with the j''s in the $\overline{u}_{i'j'}$'s. To prove the second part of Lemma 1.5.8, we shall characterize such terms among those of the type of (1.46) and prove that the other ones are negligible. To prove the first part of the lemma, we shall prove that only negligible terms occur, i.e. that if $\{k_1,\ldots,k_q\}_m \neq \{l_1,\ldots,l_s\}_m$, then $\mathrm{Wg}(id)$ can never occur in (1.46). Then, the third part of the lemma is only a straightforward summation following from the first and second parts.

Proof of (2) of Lemma 1.5.8:

Now, we reformulate the (2) from Lemma 1.5.8 this way : let $k_1 > k_2 > \cdots > k_q$ be distinct positive integers and $m_1,\ldots,m_q$ positive integers, and let $\left(i_{\alpha,\beta}\right)_{\substack{1\le\beta\le q\\1\le\alpha\le m_\beta}}$ and $\left(i'_{\alpha,\beta}\right)_{\substack{1\le\beta\le q\\1\le\alpha\le m_\beta}}$ be some integers of $\{1,\ldots,r\}$. Our goal is to prove that

$$\begin{aligned}&\mathbb{E}\Big[\sqrt{N}\mathbf{c}^*_{i_{1,1}}\mathbf{A}^{k_1}\mathbf{b}_{i_{1,1}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{1,1}}\mathbf{A}^{k_1}\mathbf{b}_{i'_{1,1}}}\cdots\sqrt{N}\mathbf{c}^*_{i_{m_1,1}}\mathbf{A}^{k_1}\mathbf{b}_{i_{m_1,1}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{m_1,1}}\mathbf{A}^{k_1}\mathbf{b}_{i'_{m_1,1}}}\times\\ &\qquad\sqrt{N}\mathbf{c}^*_{i_{1,2}}\mathbf{A}^{k_2}\mathbf{b}_{i_{1,2}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{1,2}}\mathbf{A}^{k_2}\mathbf{b}_{i'_{1,2}}}\cdots\cdots\sqrt{N}\mathbf{c}^*_{i_{m_q,q}}\mathbf{A}^{k_q}\mathbf{b}_{i_{m_q,q}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{m_q,q}}\mathbf{A}^{k_q}\mathbf{b}_{i'_{m_q,q}}}\Big]\\ &= b^{2\sum k_i m_i}\times\prod_{t=1}^{q}\left[\sum_{\mu_t\in S_{m_t}}\prod_{s=1}^{m_t}\left(\mathbf{b}^*_{i'_{\mu_t(s),t}}\mathbf{b}_{i_{s,t}}\cdot\mathbf{c}^*_{i_{s,t}}\mathbf{c}_{i'_{\mu_t(s),t}}\right)\right]+o(1).\end{aligned}$$

We will denote the coordinate of $\mathbf{b}_{i_{\alpha,\beta}}$: $\left(b_t^{i_{\alpha,\beta}}\right)_{1\le t\le N}$. We write

$$\mathbf{c}^*_{i_{\alpha,\beta}}\mathbf{A}^{k_1}\mathbf{b}_{i_{\alpha,\beta}} = \sum_{1\le t_0,\ldots,t_{k_1}\le N}\overline{c_{t_0}^{i_{\alpha,\beta}}}u_{t_0,t_1}s_{t_1}u_{t_1,t_2}s_{t_2}\cdots u_{t_{k_1-1},t_{k_1}}s_{t_{k_1}}b_{t_{k_1}}^{i_{\alpha,\beta}}. \tag{1.48}$$

In order to simply notation, we shall use bold letters to designate tuples of consecutive indices. For example, we set $\mathbf{t}_{i,j} := (t_{0,i,j}, t_{1,i,j}, \ldots, t_{k_j,i,j})$ and write

$$u_{\mathbf{t}_{i,j}} := u_{t_{0,i,j},t_{1,i,j}} \cdots u_{t_{k_j-1,i,j},t_{k_j,i,j}} \quad ; \quad s_{\mathbf{t}_{i,j}} := s_{t_{1,i,j}} \cdots s_{t_{k_j,i,j}}, \tag{1.49}$$

so if we expand the whole expectation in (1.48) with respect to the randomness of $\mathbf{U}$, we get terms as

$$\begin{aligned} & \mathbb{E}\, u_{\mathbf{t}_{1,1}} \cdots u_{\mathbf{t}_{m_1,1}} \overline{u}_{\mathbf{t}'_{1,1}} \cdots \overline{u}_{\mathbf{t}'_{m_1,1}} \cdots\cdots u_{\mathbf{t}_{1,q}} \cdots u_{\mathbf{t}_{m_q,q}} \overline{u}_{\mathbf{t}'_{1,q}} \cdots \overline{u}_{\mathbf{t}'_{m_q,q}} \\ = \; & \mathbb{E} \prod_{1\le c\le q} \; \prod_{1\le b\le m_c} \; \prod_{0\le a\le k_c-1} u_{t_{a,b,c},t_{a+1,b,c}} \overline{u}_{t'_{a,b,c},t'_{a+1,b,c}}, \end{aligned}$$

and by Proposition 2.5.1, for this expectation to be non-zero, the set of the first indices (resp. second) of the $u_{i,j}$'s must be the same than the set of the first indices (resp. second) of the $\overline{u}_{i,j}$'s, which can be expressed by the following equalities of multisets[8]

$$\begin{aligned} &\left\{t_{a,b,c}\,;\; 1\le c\le q,\, 1\le b\le m_c,\, 0\le a\le k_c-1\right\}_m = \\ &\qquad\qquad \left\{t'_{a,b,c}\,;\; 1\le c\le q,\, 1\le b\le m_c,\, 0\le a\le k_c-1\right\}_m \end{aligned} \tag{1.50}$$

$$\begin{aligned} &\left\{t_{a,b,c}\,;\; 1\le c\le q,\, 1\le b\le m_c,\, 1\le a\le k_c\right\}_m = \\ &\qquad\qquad \left\{t'_{a,b,c}\,;\; 1\le c\le q,\, 1\le b\le m_c,\, 1\le a\le k_c\right\}_m \end{aligned} \tag{1.51}$$

For now on, we denote by $K = \sum m_i k_i$ and by $\rho = \sum m_i$.

To make a better use of these multisets equalities, we shall need to reason on the $t_{a,b,c}$'s which are pairwise distincts and so, in a first place, we prove that the summation over the non pairwise indices is negligible.

To do so, we deduce first from (1.50) and (1.51) that for any fixed collection of indices $(t_{a,b,c})$, there is only a $O(1)$ choices of collection of indices $(t'_{a,b,c})$ leading to a non vanishing expectation. Then, noticing that

$$\operatorname{Card}\{t_{a,b,c}\,;\, 1\le c\le q, 1\le b\le m_c, 0\le a\le k_c\} \;=\; O\left(N^{K+\rho-1}\right),$$

(where $K = \sum m_i k_i$ and $\rho = \sum m_i$), we know that the summation contains a $O\big(N^{K+\rho-1}\big)$ of terms. At last, we use the fact that the expectation over the $u_{i,j}$'s and the $\overline{u}_{i,j}$'s is at most $O\big(N^{-K}\big)$, that $\sup_i s_i < M$ and that $|b_i|$ and $|c_i| = O\left(\sqrt{\frac{\log N}{N}}\right)$ (recall

[8]Recall that the notion of *multiset* has been defined before Lemma 1.5.8: a multiset is roughly a collection of elements with possible repetitions (as in tuples) but where the order of appearance is insignificant (contrarily to tuples). For example,

$$\{1,2,2,3\}_m = \{3,2,1,2\}_m \neq \{1,2,3\}_m.$$

(1.45)), to claim that each term of the sum is at most a $O\big((\log N)^{2\rho}N^{-K-\rho}\big)$ (one should not forget that each $\mathbf{c}\mathbf{A}^k\mathbf{b}$ is multiply by $\sqrt{N}$). We conclude that the summation over the non pairwise distinct indices is a $O\Big(\frac{(\log N)^{2\rho}}{N}\Big)$.

For now on, we consider only the pairwise distinct indices so that (1.50) and (1.51) can be seen as set equalities (instead of multiset). Also, if one sees the sets as K-tuple, the equalities (1.50) and (1.51) means that there exists two permutations σ_1 and σ_2 in S_K so that for all $1 \le c \le q$, $1 \le b \le m_q$, $0 \le a \le k_q - 1$ (resp. $1 \le a \le k_q$), we have $t'_{a,b,c} = t_{\sigma_1(a,b,c)}$ (resp. $t'_{a,b,c} = t_{\sigma_2(a,b,c)}$).

Remark 1.6.3. The notation $t_{\sigma_1(a,b,c)}$ is a little improper: the set

$$\{(a,b,c)\ ;\ 1 \le c \le q,\ 1 \le b \le m_c,\ 0 \le a \le k_c - 1\}$$

is identified with $\{1,\dots,K\}$ thanks to the colexicographical order (where $K = \sum m_i k_i$).

Thanks to the Proposition 2.5.1 and the Remark 2.5.3, we know that the expectation of the $u_{t_{a,b,c}}$'s and the $\overline{u}_{t'_{a,b,c}}$'s is equal to $\mathrm{Wg}(\sigma_1 \circ \sigma_2^{-1})$ and so, we know that we can neglect all of these with $\sigma_1 \neq \sigma_2$. For now on, we suppose $\sigma_1 = \sigma_2$.

One needs to understand that the sets $\{t_{a,b,c}\ ;\ 1 \le c \le q,\ 1 \le b \le m_c,\ 0 \le a < k_c\}$ and
$\{t_{a,b,c}\ ;\ 1 \le c \le q,\ 1 \le b \le m_c,\ 1 \le a \le k_c\}$ are very similar except for the shift for the first index. Due to this likeness and the fact that they are both mapped onto the $t'_{a,b,c}$'s in the same way (i.e. $\sigma_1 = \sigma_2$), we prove that the choice of σ_1 is very specific :

– First, using the distinctness of the indices, it easy to see that the equalities (1.50) and (1.51) lead us to these new equalities of sets

$$\{t_{0,b,c}\ ;\ 1 \le c \le q,\ 1 \le b \le m_c\} = \big\{t'_{0,b,c}\ ;\ 1 \le c \le q,\ 1 \le b \le m_c\big\}, \quad (1.52)$$

and

$$\{t_{k_c,b,c}\ ;\ 1 \le c \le q,\ 1 \le b \le m_c\} = \big\{t'_{k_c,b,c}\ ;\ 1 \le c \le q,\ 1 \le b \le m_c\big\}, \quad (1.53)$$

– According to the equality (1.52), we know that $\{t_{0,b,c},\ 1 \le c \le q,\ 1 \le b \le m_c\}$ is an invariant set of σ_1. Indeed, we know that

$$\big\{t_{\sigma_1(0,b,c)}\ ;\ 1 \le c \le q,\ 1 \le b \le m_c\big\} \subset \big\{t'_{a,b,c}\ ;\ 1 \le c \le q,\ 1 \le b \le m_c,\ 0 \le a \le k_c - 1\big\},$$

and with the condition (1.52), to avoid non pairwise distinct indices, we must have

$$\big\{t_{\sigma_1(0,b,c)}\ ;\ 1 \le c \le q,\ 1 \le b \le m_c\big\} = \big\{t'_{0,b,c}\ ;\ 1 \le c \le q,\ 1 \le b \le m_c\big\},$$

and so, we deduce that

$$\{t_{0,b,c}\ ;\ 1 \le c \le q,\ 1 \le b \le m_c\} = \big\{t_{\sigma_1(0,b,c)}\ ;\ 1 \le c \le q,\ 1 \le b \le m_c\big\}.$$

– As $\sigma_1 = \sigma_2$, σ_2 permutes $\{t_{1,b,c}\ ;\ \ 1 \le c \le q,\ 1 \le b \le m_c\}$ in the same way (actually, the sets $\{(a,b,c)\ ;\ \ 1 \le c \le q,\ 1 \le b \le m_c,\ 0 \le a \le k_c - 1\}$ and $\{(a,b,c)\ ;\ \ 1 \le c \le p,\ 1 \le b \le m_c,\ 1 \le a \le k_c\}$ are indentified to the same set (with cardinality K) thanks to the colexicographical order, and so, the action of σ_1 and σ_2 must be seen on this common set).

– As each element of $\{t_{1,b,c}\ ;\ \ 1 \le c \le q,\ 1 \le b \le m_c\}$ has only one corresponding $t'_{d,e,f}$ (indeed by (1.50) and (1.51) and as the t's and the t''s are pairwise distinct, to each t corresponds a unique t'), we deduce that σ_1 permutes $\{t_{1,b,c}\ ;\ \ 1 \le c \le q,\ 1 \le b \le m_c\}$ in the same way (indeed, it allows to claim that

$$\left(t_{\sigma_1(1,b,c)}\ ;\ \ 1 \le c \le q,\ 1 \le b \le m_c\right) \ = \ \left(t_{\sigma_2(1,b,c)}\ ;\ \ 1 \le c \le q,\ 1 \le b \le m_c\right)$$

as K-tuples.

– As $\sigma_1 = \sigma_2$, we know that σ_2 permutes $\{t_{2,b,c}\ ;\ \ 1 \le c \le q,\ 1 \le b \le m_c\}$ in the same way, and so on until one shows that σ_2 permutes $\left\{t_{k_q,b,c}\ ;\ \ 1 \le c \le q,\ 1 \le b \le m_c\right\}$ in the same way.

– However, according to (1.53), we know that $\{t_{k_c,b,c}\ ;\ \ 1 \le c \le q,\ 1 \le b \le m_c\}$ is an invariant set of σ_2.

Therefore, as $\left\{t_{k_q,b,c}\ ;\ \ 1 \le c \le q,\ 1 \le b \le m_c\right\}$ and $\{t_{k_c,b,c}\ ;\ \ 1 \le c \le q,\ 1 \le b \le m_c\}$ are invariant sets by σ_2, we know that

$$\left\{t_{k_q,b,c}\ ;\ 1 \le c \le q,\ 1 \le b \le m_c\right\} \cap \left\{t_{k_c,b,c}\ ;\ 1 \le c \le q,\ 1 \le b \le m_c\right\} \ = \ \left\{t_{k_q,b,q}\ ;\ \ 1 \le b \le m_q\right\}$$

is also an invariant set of σ_2 and we deduce that σ_1 permutes in the same way every set of the form $\left\{t_{l,b,q},\ 1 \le b \le m_q\right\}$ for $l \in \left\{0, k_q - 1\right\}$. And so, we rewrite the equalities (1.52) and (1.53)

$$\{t_{0,b,c}\ ;\ \ 1 \le c \le q-1,\ 1 \le b \le m_c\} \ = \ \{t'_{0,b,c}\ ;\ \ 1 \le c \le q-1,\ 1 \le b \le m_c\} \tag{1.54}$$

and

$$\{t_{k_c,b,c}\ ;\ \ 1 \le c \le q-1,\ 1 \le b \le m_c\} \ = \ \{t'_{k_c,b,c}\ ;\ \ 1 \le c \le q-1,\ 1 \le b \le m_c\} \tag{1.55}$$

and one can make an induction on q to show that there exist $\mu_1 \in S_{m_1}$, $\mu_2 \in S_{m_2}, \ldots, \mu_q \in S_{m_q}$ such that for all $1 \le c \le q$, $1 \le b \le m_c$, $1 \le a \le k_c$, we have

$$t'_{a,b,c} \ = \ t_{\sigma_1(a,b,c)} \ = \ t_{a,\mu_c(b),c}$$

and so, to sum up, we deduce that the non negligible terms that we get when we expand the whole expectation are terms such as

$$\mathbb{E} \prod_{1 \le c \le q} \ \prod_{1 \le b \le m_c} \ \prod_{0 \le a \le k_c - 1} u_{t_{a,b,c}, t_{a+1,b,c}} \overline{u}_{t_{a,\mu_c(b),c}, t_{a+1,\mu_c(b),c}}$$

where for all c, μ_c belongs to S_{m_c} and so, one can easily deduce that

$$\begin{aligned}
&\mathbb{E}\Big[\sqrt{N}\mathbf{c}^*_{i_{1,1}}\mathbf{A}^{k_1}\mathbf{b}_{i_{1,1}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{1,1}}\mathbf{A}^{k_1}\mathbf{b}_{i'_{1,1}}}\cdots\sqrt{N}\mathbf{c}^*_{i_{m_1,1}}\mathbf{A}^{k_1}\mathbf{b}_{i_{m_1,1}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{m_1,1}}\mathbf{A}^{k_1}\mathbf{b}_{i'_{m_1,1}}}\times \\
&\qquad \sqrt{N}\mathbf{c}^*_{i_{1,2}}\mathbf{A}^{k_2}\mathbf{b}_{i_{1,2}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{1,2}}\mathbf{A}^{k_2}\mathbf{b}_{i'_{1,2}}}\cdots\cdots\sqrt{N}\mathbf{c}^*_{i_{m_q,q}}\mathbf{A}^{k_q}\mathbf{b}_{i_{m_q,q}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{m_q,q}}\mathbf{A}^{k_q}\mathbf{b}_{i'_{m_q,q}}}\Big] \\
&= \left(\frac{1}{N}\sum_{i=1}^{N}s_i^2\right)^{K-\rho}\cdot\prod_{u=1}^{q}\left[\sum_{\mu_u\in S_{m_u}}\prod_{s=1}^{m_u}\left(\sum_{\substack{1\le t_{0,s,u}\le N\\ 1\le t_{k_u,s,u}\le N}}\overline{b_{t_{k_u,s,u}}^{i'_{\mu_u(s),u}}}\,b_{t_{k_u,s,u}}^{i_{s,u}}s^2_{t_{k_u,s,u}}\overline{c_{t_{0,s,u}}^{i_{s,u}}}\,c_{t_{0,s,u}}^{i'_{\mu_u(s),u}}\right)\right]+o(1) \\
&= b^{2K}\times\prod_{u=1}^{q}\left[\sum_{\mu_u\in S_{m_u}}\prod_{s=1}^{m_u}\left(\mathbf{b}^*_{i'_{\mu_u(s),u}}\mathbf{b}_{i_{s,u}}\cdot\mathbf{c}^*_{i_{s,u}}\mathbf{c}_{i'_{\mu_u(s),u}}\right)\right]+o(1),
\end{aligned}$$

and we can conclude.

Remark 1.6.4. We used the fact that

$$\frac{1}{N}\sum_{i=1}^{N}s_i^2 = b^2+o(1) \tag{1.56}$$

$$\sum_{j=1}^{N}\overline{b_j^{i_{\alpha,\beta}}}\,b_j^{i_{\gamma,\delta}}s_j^2 = b^2\mathbf{b}^*_{i_{\alpha,\beta}}\mathbf{b}_{i_{\gamma,\delta}}+o(1). \tag{1.57}$$

The relation (1.56) is obvious and the (1.57) can be proved using the fact $\mathbf{P}$ is invariant, in law, by conjugation by any unitary matrix (we explained at Section 1.2.4 that we can add this hypothesis).

Proof of (1) of Lemma 1.5.8:

The proof of (1) goes along the same lines as the previous proof. Our goal is to show that

$$\mathbb{E}\left[\sqrt{N}\mathbf{c}^*_{i_1}\mathbf{A}^{k_1}\mathbf{b}_{i_1}\cdots\sqrt{N}\mathbf{c}^*_{i_q}\mathbf{A}^{k_q}\mathbf{b}_{i_q}\overline{\sqrt{N}\mathbf{c}^*_{i'_1}\mathbf{A}^{l_1}\mathbf{b}_{i'_1}}\cdots\overline{\sqrt{N}\mathbf{c}^*_{i'_s}\mathbf{A}^{l_s}\mathbf{b}_{i'_s}}\right] = o(1).$$

At first, one can notice that if $\sum k_i \neq \sum l_j$, the expectation is equal to zero. We assume now that $\sum k_i = \sum l_j$, and let K denote the common value. Then, we distinguish two cases.

• First case : $q = s$

Then we can also focus on the "pairwise distinct indices" summation, by similar argument as in the previous proof. We suppose that there exists j such that $k_j \neq l_j$ (otherwise, one should read the previous proof). Our goal is to show that there is no expectation equal to $\mathrm{Wg}(id)$ (which means that we cannot have $\sigma_1 = \sigma_2$) in that case and so we shall conclude that

$$\mathbb{E}\left[\sqrt{N}\mathbf{c}^*_{i_1}\mathbf{A}^{k_1}\mathbf{b}_{i_1}\cdots\sqrt{N}\mathbf{c}^*_{i_q}\mathbf{A}^{k_q}\mathbf{b}_{i_q}\overline{\sqrt{N}\mathbf{c}^*_{i'_1}\mathbf{A}^{l_1}\mathbf{b}_{i'_1}}\cdots\overline{\sqrt{N}\mathbf{c}^*_{i'_q}\mathbf{A}^{l_q}\mathbf{b}_{i'_q}}\right] = O\left(\frac{1}{N}\right).$$

Let us gather the k_i's which are equal and in order to simply the expressions, we shall use notations in the same spirit than (1.49)

$$\begin{aligned} \sqrt{N}(\mathbf{c}^*\mathbf{A}^{k_\alpha}\mathbf{b})_{\mathbf{i}_\alpha} &:= \sqrt{N}\mathbf{c}^*_{i_{1,\alpha}}\mathbf{A}^{k_\alpha}\mathbf{b}_{i_{1,\alpha}}\cdots\sqrt{N}\mathbf{c}^*_{i_{m_\alpha,\alpha}}\mathbf{A}^{k_\alpha}\mathbf{b}_{i_{m_\alpha,\alpha}}, \\ \sqrt{N}(\mathbf{c}^*\mathbf{A}^{\ell_\beta}\mathbf{b})_{\mathbf{i}'_\beta} &:= \sqrt{N}\mathbf{c}^*_{i'_{1,\beta}}\mathbf{A}^{\ell_\beta}\mathbf{b}_{i'_{1,\beta}}\cdots\sqrt{N}\mathbf{c}^*_{i'_{n_\beta,\beta}}\mathbf{A}^{\ell_\beta}\mathbf{b}_{i'_{n_\beta,\beta}}, \end{aligned} \tag{1.58}$$

so that we rewrite our expectation

$$\mathbb{E}\left[\sqrt{N}(\mathbf{c}^*\mathbf{A}^{k_1}\mathbf{b})_{\mathbf{i}_1}\cdots\sqrt{N}(\mathbf{c}^*\mathbf{A}^{k_r}\mathbf{b})_{\mathbf{i}_r}\overline{\sqrt{N}(\mathbf{c}^*\mathbf{A}^{\ell_1}\mathbf{b})_{\mathbf{i}'_1}}\cdots\overline{\sqrt{N}(\mathbf{c}^*\mathbf{A}^{\ell_s}\mathbf{b})_{\mathbf{i}'_s}}\right]$$

with $\sum_{i=1}^r m_i k_i = \sum_{j=1}^s n_j l_j$ and $k_1 > \cdots > k_r$ and $l_1 > \cdots > l_s$. Without loss of generality, we shall assume that $(k_r, m_r) \neq (l_s, n_s)$ (indeed, otherwise, we can start the induction from the previous proof until we find an integer x such that $(k_{r-x}, m_{r-x}) \neq (l_{s-x}, n_{s-x})$ to show that the expectation is equal to

$$\begin{aligned} &\mathbb{E}\left[\sqrt{N}(\mathbf{c}^*\mathbf{A}^{k_1}\mathbf{b})_{\mathbf{i}_1}\cdots\sqrt{N}(\mathbf{c}^*\mathbf{A}^{k_{r-x}}\mathbf{b})_{\mathbf{i}_{r-x}}\overline{\sqrt{N}(\mathbf{c}^*\mathbf{A}^{\ell_1}\mathbf{b})_{\mathbf{i}'_1}}\cdots\overline{\sqrt{N}(\mathbf{c}^*\mathbf{A}^{\ell_{s-x}}\mathbf{b})_{\mathbf{i}'_{s-x}}}\right] \\ \times &\prod_{t=r-x+1}^{r}\sum_{\mu_t\in S_{m_t}}\prod_{s=1}^{m_t}\mathbb{E}\left[\sqrt{N}\mathbf{c}^*_{i_{s,t}}\mathbf{A}^{k_t}\mathbf{b}_{i_{s,t}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{\mu_t(s),t}}\mathbf{A}^{k_t}\mathbf{b}_{i'_{\mu_t(s),t}}}\right]+o(1), \end{aligned}$$

and the following of the proof is the same). We shall also assume that $k_r \leq l_s$.

According to Proposition 2.5.1, we have the following equalities

$$\begin{aligned} &\left\{t_{a,b,c}\,;\ 1\leq c\leq r,\ 1\leq b\leq m_c,\ 0\leq a\leq k_c-1\right\} \\ &\qquad = \left\{t'_{a,b,c}\,;\ 1\leq c\leq s,\ 1\leq b\leq n_c,\ 0\leq a\leq l_c-1\right\}, \end{aligned} \tag{1.59}$$

and

$$\begin{aligned} &\left\{t_{a,b,c}\,;\ 1\leq c\leq r,\ 1\leq b\leq m_c,\ 1\leq a\leq k_c\right\} \\ &\qquad = \left\{t'_{a,b,c}\,;\ 1\leq c\leq s,\ 1\leq b\leq n_c,\ 1\leq a\leq l_c\right\}, \end{aligned} \tag{1.60}$$

and let σ_1 and σ_2 the two permutations describing these equalities. Let us prove by contradiction that $\sigma_1 \neq \sigma_2$ and so let us suppose that $\sigma_1 = \sigma_2$. As we consider only pairwise distinct indices, we have also

$$\left\{t_{0,b,c}\,;\ 1\leq c\leq r,\ 1\leq b\leq m_c\right\} = \left\{t'_{0,b,c}\,;\ 1\leq c\leq s,\ 1\leq b\leq n_c\right\}, \tag{1.61}$$

and

$$\left\{t_{k_c,b,c}\,;\ 1\leq c\leq r,\ 1\leq b\leq m_c\right\} = \left\{t'_{l_c,b,c}\,;\ 1\leq c\leq s,\ 1\leq b\leq n_c\right\}, \tag{1.62}$$

According to the fact that $\sigma_1 = \sigma_2$ and (1.61), we can deduce that

$$\left\{t_{k_r,b,c}\,;\ 1\leq c\leq r,\ 1\leq b\leq m_c\right\} = \left\{t'_{k_r,b,c}\,;\ 1\leq c\leq s,\ 1\leq b\leq n_c\right\}, \tag{1.63}$$

and here comes the contradiction. Indeed, if $k_r < l_s$, then

$$\left\{t'_{l_c,b,c}\ ;\ 1\leq c\leq s,\ 1\leq b\leq n_c\right\}\cap\left\{t'_{k_r,b,c}\ ;\ 1\leq c\leq s,\ 1\leq b\leq n_c\right\}=\emptyset,$$

otherwise, $k_r = l_s$ (which means $m_r \neq n_s$), let us suppose $m_r < n_s$, so that

$$\text{Card}\left\{t'_{l_c,b,c}\ ;\ 1\leq c\leq s,\ 1\leq b\leq n_c\right\}\cap\left\{t'_{k_r,b,c}\ ;\ 1\leq c\leq s,\ 1\leq b\leq n_c\right\} = n_s$$

however,

$$\text{Card}\left\{t_{k_c,b,c}\ ;\ 1\leq c\leq r,\ 1\leq b\leq m_c\right\}\cap\left\{t_{k_r,b,c}\ ;\ 1\leq c\leq r,\ 1\leq b\leq m_r\right\} = m_r,$$

which is, according to (1.62) and (1.63), impossible.

• Second case : $q \neq s$

Without loss of generality, we suppose that $q > s$. We cannot consider here the pairwise distinct indices simply because the cardinal of the $t_{i,j}$'s is different than the one of the $t'_{i,j}$'s.

Expanding the product

$$\mathbb{E}\left[\sqrt{N}\mathbf{c}^*_{i_1}\mathbf{A}^{k_1}\mathbf{b}_{i_1}\cdots\sqrt{N}\mathbf{c}^*_{i_q}\mathbf{A}^{k_q}\mathbf{b}_{i_q}\overline{\sqrt{N}\mathbf{c}^*_{i'_1}\mathbf{A}^{l_1}\mathbf{b}_{i'_1}}\cdots\overline{\sqrt{N}\mathbf{c}^*_{i'_s}\mathbf{A}^{l_s}\mathbf{b}_{i'_s}}\right],$$

we get terms such as

$$\mathbb{E}_{\mathbf{U}}\left[u_{t_{0,1},t_{1,1}}\cdots u_{t_{k_1-1,1},t_{k_1,1}}u_{t_{0,2},t_{1,2}}\cdots u_{t_{k_q-1,q},t_{k_q,q}}\overline{u_{t'_{0,1},t'_{1,1}}\cdots u_{t'_{l_1-1,1},t'_{l_1,1}}u_{t'_{0,2},t'_{1,2}}\cdots u_{t'_{l_s-1,s},t'_{l_s,s}}}\right]$$

According to Proposition 1.3.1, for the expectation to be non zero, one needs to have the equality of sets

$$\begin{aligned}
\{t_{0,1},t_{1,1},\ldots,t_{k_1-1,1},t_{0,2},\ldots,t_{k_2-1,2},\ldots,t_{k_q-1,q}\} &= \{t'_{0,1},t'_{1,1},\ldots,t'_{l_1-1,1},t'_{0,2},\ldots,t'_{l_2-1,2},\ldots,t'_{l_s-1,s}\},\\
\{t_{1,1},t_{2,1},\ldots,t_{k_1,1},t_{1,2},\ldots,t_{k_2,2},t_{1,3},\ldots,t_{k_q,q}\} &= \{t'_{1,1},t'_{2,1},\ldots,t'_{l_1,1},t'_{1,2},\ldots,t'_{l_2,2},t'_{1,3},\ldots,t'_{l_s,s}\}.
\end{aligned}$$

Set $\mathscr{A} := \{t_{a,b},\ 1\leq b\leq r,\ 0\leq a\leq k_b\}$ and $\mathscr{B} = \left\{t'_{a,b},\ 1\leq b\leq s,\ 0\leq a\leq l_b\right\}$. According to the previous inequalities, to (1.45) and to the fact that all the expectations are $O\left(N^{-K}\right)$, we write

$$\begin{aligned}
&\left|\mathbb{E}\left[\sqrt{N}\mathbf{c}^*_{i_1}\mathbf{A}^{k_1}\mathbf{b}_{i_1}\cdots\sqrt{N}\mathbf{c}^*_{i_q}\mathbf{A}^{k_q}\mathbf{b}_{i_q}\overline{\sqrt{N}\mathbf{c}^*_{i'_1}\mathbf{A}^{l_1}\mathbf{b}_{i'_1}}\cdots\overline{\sqrt{N}\mathbf{c}^*_{i'_s}\mathbf{A}^{l_s}\mathbf{b}_{i'_s}}\right]\right|\\
\leq\ & C^{4q}\frac{\log^{2q}(N)}{N^{2q}}\cdot N^{\frac{q+s}{2}}\cdot\sum_{\substack{t_{0,1},\ldots,t_{k_1,1}\\ \vdots\\ t_{0,q},\ldots,t_{k_q,q}}}\ \sum_{\mu\in\mathscr{B}^{\mathscr{A}}}\prod_{a=1}^{q}s_{t_{1,a}}s_{\mu(t_{1,a})}\cdots s_{\mu(t_{k_a,a})}\times O\left(N^{-K}\right)\\
\leq\ & O\left(\log^{2q}(N)\cdot N^{\frac{q+s}{2}-K-2q}\right)\sum_{\substack{\mu\in\mathscr{B}^{\mathscr{A}}\\ t_{0,1},\ldots,t_{k_1,1}\\ \vdots\\ t_{0,q},\ldots,t_{k_q,q}}}\frac{1}{2}\left[\prod_{a=1}^{q}s^2_{t_{1,a}}\cdots s^2_{t_{k_a,a}}+\prod_{a=1}^{q}s^2_{\mu(t_{1,a})}\cdots s^2_{\mu(t_{k_a,a})}\right]
\end{aligned}$$

On the one hand,

$$\sum_{\substack{\mu\in\mathscr{B}^{\mathscr{A}}\\ t_{0,1},\ldots,t_{k_1,1}\\ \vdots \\ t_{0,q},\ldots,t_{k_q,q}}} \prod_{a=1}^{q} s_{t_{1,a}}^2\cdots s_{t_{k_a,a}}^2 \quad\leq\quad \mathrm{Card}(\mathscr{B}^{\mathscr{A}})\times \mathrm{Card}\left(\{1,\ldots,N\}^{K+q}\right)\times M^{2K} \;=\; O\left(N^{K+q}\right).$$

On the other,

$$\sum_{\substack{\mu\in\mathscr{B}^{\mathscr{A}}\\ t_{0,1},\ldots,t_{k_1,1}\\ \vdots \\ t_{0,q},\ldots,t_{k_q,q}}} \prod_{a=1}^{q} s_{\mu(t_{1,a})}^2\cdots s_{\mu(t_{k_a,a})}^2 \quad=\quad O(1)\times \sum_{\substack{t'_{0,1},\ldots,t'_{l_1,1}\\ \vdots \\ t'_{0,s},\ldots,t'_{l_s,s}}} \prod_{a=1}^{s} s_{t'_{1,a}}^2\cdots s_{t'_{l_a,a}}^2 \;=\; O\left(N^{K+s}\right).$$

Indeed, for any fixed $J=\{t'_{0,1},t'_{1,1},\ldots,t'_{l_1,1},t'_{0,2},\ldots,t'_{l_2,2},\ldots,t'_{l_s,s}\}$, there are $O(1)$ of μ's in $\mathscr{S}_{\mathscr{I}\to\mathscr{J}}$ and $I=\{t_{0,1},t_{1,1},\ldots,t_{k_1,1},t_{0,2},\ldots,t_{k_2,2},\ldots,t_{k_q,q}\}$ such that $\mu(I)=J$.

Therefore,

$$\left|\mathbb{E}_{\mathbf{U}}\left[\sqrt{N}\mathbf{c}_{i_1}^*\mathbf{A}^{k_1}\mathbf{b}_{i_1}\cdots\sqrt{N}\mathbf{c}_{i_q}^*\mathbf{A}^{k_q}\mathbf{b}_{i_q}\overline{\sqrt{N}\mathbf{c}_{i'_1}^*\mathbf{A}^{l_1}\mathbf{b}_{i'_1}}\cdots\overline{\sqrt{N}\mathbf{c}_{i'_s}^*\mathbf{A}^{l_s}\mathbf{b}_{i'_s}}\right]\right| \quad=\quad O\left(\log^{2q}(N)\cdot N^{-\frac{q-s}{2}}\right),$$

and, since $q>s$, it is at least a $O\left(\frac{\log^{2q}(N)}{\sqrt{N}}\right)$.

Proof of (3) of Lemma 1.5.8:

If $\left\{k_1,\ldots,k_q\right\}_m\neq\left\{k'_1,\ldots,k'_q\right\}_m$, we know that it contributes to the $o(1)$. So we rewrite

$$\begin{aligned}
&\sum_{\substack{k_1,\ldots,k_q=1\\ k'_1,\ldots,k'_q=1}}^{k_0}\mathbb{E}\left[\prod_{\alpha=1}^{q}\sqrt{N}\mathbf{c}_{i_\alpha}^*\frac{\mathbf{A}^{k_\alpha}}{z_{i_1}^{k_\alpha+1}}\mathbf{b}_{i_\alpha}\overline{\sqrt{N}\mathbf{c}_{i'_\alpha}^*\frac{\mathbf{A}^{k'_\alpha}}{z_{i'_\alpha}^{k'_\alpha+1}}\mathbf{b}_{i'_\alpha}}\right]\\
=\;&\sum_{k_1,\ldots,k_q=1}^{k_0}\;\sum_{\substack{k'_1,\ldots,k'_q=1\\ \{k_1,\ldots,k_q\}_m=\{k'_1,\ldots,k'_q\}_m}}^{k_0}\mathbb{E}\left[\prod_{\alpha=1}^{q}\sqrt{N}\mathbf{c}_{i_\alpha}^*\frac{\mathbf{A}^{k_\alpha}}{z_{i_1}^{k_\alpha+1}}\mathbf{b}_{i_\alpha}\overline{\sqrt{N}\mathbf{c}_{i'_\alpha}^*\frac{\mathbf{A}^{k'_\alpha}}{z_{i'_\alpha}^{k'_\alpha+1}}\mathbf{b}_{i'_\alpha}}\right]+o(1).
\end{aligned}$$

Then, we fixed $(k_1,\ldots,k_q)$, and let us calculate

$$\sum_{\substack{k'_1,\ldots,k'_q=1\\ \{k_1,\ldots,k_q\}_m=\{k'_1,\ldots,k'_q\}_m}}^{k_0}\mathbb{E}\left[\prod_{\alpha=1}^{q}\sqrt{N}\mathbf{c}_{i_\alpha}^*\frac{\mathbf{A}^{k_\alpha}}{z_{i_1}^{k_\alpha+1}}\mathbf{b}_{i_\alpha}\overline{\sqrt{N}\mathbf{c}_{i'_\alpha}^*\frac{\mathbf{A}^{k'_\alpha}}{z_{i'_\alpha}^{k'_\alpha+1}}\mathbf{b}_{i'_\alpha}}\right], \qquad (1.64)$$

to do so, we will use the previous notations and write

$$\Big(\underbrace{k_1,\ldots,k_1}_{m_1},\underbrace{k_2,\ldots,k_2}_{m_2},\ldots,\underbrace{k_s,\ldots,k_s}_{m_s}\Big)$$

and we shall show that (1.64) doesn't depend on the m_i's but depends only on $q=\sum m_i k_i$. We rewrite the summation

$$\begin{aligned}\sum_{\{k'_1,\ldots,k'_q\}_m=\{k_1,\ldots,k_s\}_m} \mathbb{E}\Bigg[& \sqrt{N}\mathbf{c}^*_{i_{1,1}}\frac{\mathbf{A}^{k_1}}{z_{i_{1,1}}^{k_1+1}}\mathbf{b}_{i_{1,1}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{1,1}}\frac{\mathbf{A}^{k'_1}}{z_{i'_{1,1}}^{k'_1+1}}\mathbf{b}_{i'_{1,1}}}\cdots\sqrt{N}\mathbf{c}^*_{i_{m_1,1}}\frac{\mathbf{A}^{k_1}}{z_{i_{m_1,1}}^{k_1+1}}\mathbf{b}_{i_{m_1,1}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{m_1,1}}\frac{\mathbf{A}^{k'_{m_1}}}{z_{i'_{m_1,1}}^{k'_{m_1}+1}}\mathbf{b}_{i'_{m_1,1}}}\\ & \sqrt{N}\mathbf{c}^*_{i_{1,2}}\frac{\mathbf{A}^{k_2}}{z_{i_{1,2}}^{k_2+1}}\mathbf{b}_{i_{1,2}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{1,2}}\frac{\mathbf{A}^{k'_{m_1+1}}}{z_{i'_{1,2}}^{k'_{m_1+1}+1}}\mathbf{b}_{i'_{1,2}}}\cdots\sqrt{N}\mathbf{c}^*_{i_{m_2,2}}\frac{\mathbf{A}^{k_2}}{z_{i_{m_2,2}}^{k_2+1}}\mathbf{b}_{i_{m_2,2}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{m_2,2}}\frac{\mathbf{A}^{k'_{m_1+m_2}}}{z_{i'_{m_2,2}}^{k'_{m_1+m_2}+1}}\mathbf{b}_{i'_{m_2,2}}}\\ & \cdots\sqrt{N}\mathbf{c}^*_{i_{m_s,s}}\frac{\mathbf{A}^{k_s}}{z_{i_{m_s,s}}^{k_s+1}}\mathbf{b}_{i_{m_s,s}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{m_s,s}}\frac{\mathbf{A}^{k'_q}}{z_{i'_{m_s,s}}^{k'_q+1}}\mathbf{b}_{i'_{m_s,s}}}\Bigg].\end{aligned} \tag{1.65}$$

We gather the k_i's which are equal, so we rewrite the summation thanks to permutations of the set
$\mathscr{I}=\{(\alpha,\beta),\ 1\le\beta\le s,\ 1\le\alpha\le m_s\}$:

$$\begin{aligned}\sum_{\mu\in S_{\mathscr{I}}}\mathbb{E}\Bigg[& \sqrt{N}\mathbf{c}^*_{i_{1,1}}\frac{\mathbf{A}^{k_1}}{z_{i_{1,1}}^{k_1+1}}\mathbf{b}_{i_{1,1}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{\mu(1,1)}}\frac{\mathbf{A}^{k_1}}{z_{i'_{\mu(1,1)}}^{k_1+1}}\mathbf{b}_{i'_{\mu(1,1)}}}\cdots\sqrt{N}\mathbf{c}^*_{i_{m_1,1}}\frac{\mathbf{A}^{k_1}}{z_{i_{m_1,1}}^{k_1+1}}\mathbf{b}_{i_{m_1,1}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{\mu(m_1,1)}}\frac{\mathbf{A}^{k_1}}{z_{i'_{\mu(m_1,1)}}^{k_1+1}}\mathbf{b}_{i'_{\mu(m_1,1)}}}\\ & \sqrt{N}\mathbf{c}^*_{i_{1,2}}\frac{\mathbf{A}^{k_2}}{z_{i_{1,2}}^{k_2+1}}\mathbf{b}_{i_{1,2}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{\mu(1,2)}}\frac{\mathbf{A}^{k_2}}{z_{i'_{\mu(1,2)}}^{k_2+1}}\mathbf{b}_{i'_{\mu(1,2)}}}\cdots\sqrt{N}\mathbf{c}^*_{i_{m_s,s}}\frac{\mathbf{A}^{k_s}}{z_{i_{m_s,s}}^{k_s+1}}\mathbf{b}_{i_{m_s,s}}\overline{\sqrt{N}\mathbf{c}^*_{i'_{\mu(m_s,s)}}\frac{\mathbf{A}^{k_s}}{z_{i'_{\mu(m_s,s)}}^{k_s+1}}\mathbf{b}_{i'_{\mu(m_s,s)}}}\Bigg]\end{aligned}$$

except that we count several times each terms. Indeed, for example, if one wants to rearrange

$$\mathbb{E}\left[\mathbf{c}^*_{i_{1,1}}\frac{\mathbf{A}^{k_1}}{z_{i_{1,1}}^{k_1}}\mathbf{b}_{i_{1,1}}\mathbf{c}^*_{i_{2,1}}\frac{\mathbf{A}^{k_1}}{z_{i_{2,1}}^{k_1}}\mathbf{b}_{i_{2,1}}\mathbf{c}^*_{i_{1,2}}\frac{\mathbf{A}^{k_2}}{z_{i_{1,2}}^{k_2}}\mathbf{b}_{i_{1,2}}\overline{\mathbf{c}^*_{i'_{1,1}}\frac{\mathbf{A}^{k_1}}{z_{i'_{1,1}}^{k_1}}\mathbf{b}_{i'_{1,1}}\mathbf{c}^*_{i'_{2,1}}\frac{\mathbf{A}^{k_2}}{z_{i'_{2,1}}^{k_2}}\mathbf{b}_{i'_{2,1}}\mathbf{c}^*_{i'_{1,2}}\frac{\mathbf{A}^{k_1}}{z_{i'_{1,2}}^{k_1}}\mathbf{b}_{i'_{1,2}}}\right], \tag{1.66}$$

there are two ways to do it :

$$\mathbb{E}\left[\mathbf{c}^*_{i_{1,1}}\frac{\mathbf{A}^{k_1}}{z_{i_{1,1}}^{k_1}}\mathbf{b}_{i_{1,1}}\overline{\mathbf{c}^*_{i'_{1,1}}\frac{\mathbf{A}^{k_1}}{z_{i'_{1,1}}^{k_1}}\mathbf{b}_{i'_{1,1}}}\mathbf{c}^*_{i_{2,1}}\frac{\mathbf{A}^{k_1}}{z_{i_{2,1}}^{k_1}}\mathbf{b}_{i_{2,1}}\overline{\mathbf{c}^*_{i'_{1,2}}\frac{\mathbf{A}^{k_1}}{z_{i'_{1,2}}^{k_1}}\mathbf{b}_{i'_{1,2}}}\mathbf{c}^*_{i_{1,2}}\frac{\mathbf{A}^{k_2}}{z_{i_{1,2}}^{k_2}}\mathbf{b}_{i_{1,2}}\overline{\mathbf{c}^*_{i'_{2,1}}\frac{\mathbf{A}^{k_2}}{z_{i'_{2,1}}^{k_2}}\mathbf{b}_{i'_{2,1}}}\right],$$

or

$$\mathbb{E}\left[\mathbf{c}^*_{i_{1,1}}\frac{\mathbf{A}^{k_1}}{z_{i_{1,1}}^{k_1}}\mathbf{b}_{i_{1,1}}\overline{\mathbf{c}^*_{i'_{1,2}}\frac{\mathbf{A}^{k_1}}{z_{i'_{1,2}}^{k_1}}\mathbf{b}_{i'_{1,2}}}\mathbf{c}^*_{i_{2,1}}\frac{\mathbf{A}^{k_1}}{z_{i_{2,1}}^{k_1}}\mathbf{b}_{i_{2,1}}\overline{\mathbf{c}^*_{i'_{1,1}}\frac{\mathbf{A}^{k_1}}{z_{i'_{1,1}}^{k_1}}\mathbf{b}_{i'_{1,1}}}\mathbf{c}^*_{i_{1,2}}\frac{\mathbf{A}^{k_2}}{z_{i_{1,2}}^{k_2}}\mathbf{b}_{i_{1,2}}\overline{\mathbf{c}^*_{i'_{2,1}}\frac{\mathbf{A}^{k_2}}{z_{i'_{2,1}}^{k_2}}\mathbf{b}_{i'_{2,1}}}\right],$$

and so (1.66) would be counted twice. Actually, it is easy to see that μ_1 and μ_2 give us the same terms if and only if $\sigma = \mu_1 \circ \mu_2^{-1}$ is a permutation such that for all $(i,j) \in \mathscr{I}$, $\sigma(i,j) = (i',j)$ (it means that σ doesn't change the second index). Let us denote by $S_{k_1,\ldots,k_q}$ the set of such σ's in $S_{\mathscr{I}}$. Then the expression of (1.65) rewrites

$$
\begin{aligned}
& \frac{1}{\operatorname{Card} S_{k_1,\ldots,k_q}} \sum_{\mu \in S_{\mathscr{I}}} \mathbb{E}\left[\prod_{\alpha=1}^{s}\prod_{\beta=1}^{m_\alpha} \sqrt{N}\mathbf{c}_{i_{\beta,\alpha}}^* \frac{\mathbf{A}^{k_\alpha}}{z_{i_{\beta,\alpha}}^{k_\alpha+1}}\mathbf{b}_{i_{\beta,\alpha}} \overline{\sqrt{N}\mathbf{c}_{i'_{\mu(\beta,\alpha)}}^* \frac{\mathbf{A}^{k_\alpha}}{z_{i'_{\mu(\beta,\alpha)}}^{k_\alpha+1}}\mathbf{b}_{i'_{\mu(\beta,\alpha)}}}\right] \\
= & \frac{1}{\operatorname{Card} S_{k_1,\ldots,k_q}} \sum_{\mu \in S_{\mathscr{I}}} \sum_{\sigma \in S_{k_1,\ldots,k_q}} \prod_{t=1}^{s}\prod_{u=1}^{m_t} \frac{1}{z_{i_{u,t}}\overline{z}_{i'_{\mu(u,t)}}}\left(\frac{b^2}{z_{i_{u,t}}\overline{z}_{i'_{\mu(u,t)}}}\right)^{k_t} \mathbf{b}_{i'_{\sigma\circ\mu(u,t)}}^* \mathbf{b}_{i_{u,t}} \mathbf{c}_{i_{u,t}}^* \mathbf{c}_{i'_{\sigma\circ\mu(u,t)}} + o(1) \\
= & \frac{1}{\operatorname{Card} S_{k_1,\ldots,k_q}} \sum_{\sigma \in S_{k_1,\ldots,k_q}} \sum_{\mu \in S_{\mathscr{I}}} \prod_{t=1}^{s}\prod_{u=1}^{m_t} \frac{1}{z_{i_{u,t}}\overline{z}_{i'_{\mu(u,t)}}}\left(\frac{b^2}{z_{i_{u,t}}\overline{z}_{i'_{\mu(u,t)}}}\right)^{k_t} \mathbf{b}_{i'_{\mu(u,t)}}^* \mathbf{b}_{i_{u,t}} \mathbf{c}_{i_{u,t}}^* \mathbf{c}_{i'_{\mu(u,t)}} + o(1) \\
= & \sum_{\mu \in S_{\mathscr{I}}} \prod_{t=1}^{s}\prod_{u=1}^{m_t} \frac{1}{z_{i_{u,t}}\overline{z}_{i'_{\mu(u,t)}}}\left(\frac{b^2}{z_{i_{u,t}}\overline{z}_{i'_{\mu(u,t)}}}\right)^{k_t} \mathbf{b}_{i'_{\mu(u,t)}}^* \mathbf{b}_{i_{u,t}} \mathbf{c}_{i_{u,t}}^* \mathbf{c}_{i'_{\mu(u,t)}} + o(1)
\end{aligned}
$$

and if we go back to the notation $\{1,\ldots,k_q\}$, we have

$$
\sum_{\mu \in S_{\mathscr{I}}} \prod_{t=1}^{s}\prod_{u=1}^{m_t} \frac{1}{z_{i_{u,t}}\overline{z}_{i'_{\mu(u,t)}}}\left(\frac{b^2}{z_{i_{u,t}}\overline{z}_{i'_{\mu(u,t)}}}\right)^{k_t} \mathbf{b}_{i'_{\mu(u,t)}}^* \mathbf{b}_{i_{u,t}} \mathbf{c}_{i_{u,t}}^* \mathbf{c}_{i'_{\mu(u,t)}} = \sum_{\sigma \in S_q} \prod_{t=1}^{q} \frac{1}{z_{i_t}\overline{z}_{i'_{\sigma(t)}}}\left(\frac{b^2}{z_{i_t}\overline{z}_{i'_{\sigma(t)}}}\right)^{k_t} \mathbf{b}_{i'_{\sigma(t)}}^* \mathbf{b}_{i_t} \cdot \mathbf{c}_{i_t}^* \mathbf{c}_{i'_{\sigma(t)}}
$$

and so

$$
\sum_{\substack{k_1,\ldots,k_q=1\\ k'_1,\ldots,k'_q=1}}^{k_0} \mathbb{E}\left[\sqrt{N}\mathbf{c}_{i_1}^* \frac{\mathbf{A}^{k_1}}{z_{i_1}^{k_1+1}}\mathbf{b}_{i_1} \overline{\sqrt{N}\mathbf{c}_{i'_1}^* \frac{\mathbf{A}^{k'_1}}{z_{i'_1}^{k'_1+1}}\mathbf{b}_{i'_1}} \cdots \sqrt{N}\mathbf{c}_{i_q}^* \frac{\mathbf{A}^{k_q}}{z_{i_q}^{k_q+1}}\mathbf{b}_{i_{1,q}} \overline{\sqrt{N}\mathbf{c}_{i'_q}^* \frac{\mathbf{A}^{k'_q}}{z_{i'_q}^{k'_q+1}}\mathbf{b}_{i'_q}}\right]
$$

$$
\begin{aligned}
= & \sum_{k_1,\ldots,k_q=1}^{k_0} \sum_{\substack{k'_1,\ldots,k'_q=1\\ \{k_1,\ldots,k_q\}_m=\{k'_1,\ldots,k'_q\}_m}}^{k_0} \mathbb{E}\left[\sqrt{N}\mathbf{c}_{i_1}^* \frac{\mathbf{A}^{k_1}}{z_{i_1}^{k_1+1}}\mathbf{b}_{i_1} \overline{\sqrt{N}\mathbf{c}_{i'_1}^* \frac{\mathbf{A}^{k'_1}}{z_{i'_1}^{k'_1+1}}\mathbf{b}_{i'_1}} \cdots \sqrt{N}\mathbf{c}_{i_q}^* \frac{\mathbf{A}^{k_q}}{z_{i_q}^{k_q+1}}\mathbf{b}_{i_{1,q}} \overline{\sqrt{N}\mathbf{c}_{i'_q}^* \frac{\mathbf{A}^{k'_q}}{z_{i'_q}^{k'_q+1}}\mathbf{b}_{i'_q}}\right] + o(1) \\
= & \sum_{k_1,\ldots,k_q=1}^{k_0} \sum_{\sigma \in S_q} \prod_{t=1}^{q} \frac{1}{z_{i_t}\overline{z}_{i'_{\sigma(t)}}}\left(\frac{b^2}{z_{i_t}\overline{z}_{i'_{\sigma(t)}}}\right)^{k_t} \mathbf{b}_{i'_{\sigma(t)}}^* \mathbf{b}_{i_t} \cdot \mathbf{c}_{i_t}^* \mathbf{c}_{i'_{\sigma(t)}} + o(1) \\
= & \sum_{\sigma \in S_q} \prod_{t=1}^{q} \frac{b^2}{z_{i_t}\overline{z}_{i'_{\sigma(t)}}} \frac{1-\left(\frac{b^2}{z_{i_t}\overline{z}_{i'_{\sigma(t)}}}\right)^{k_0}}{z_{i_t}\overline{z}_{i'_{\sigma(t)}} - b^2} \mathbf{b}_{i'_{\sigma(t)}}^* \mathbf{b}_{i_t} \mathbf{c}_{i_t}^* \mathbf{c}_{i'_{\sigma(t)}} + o(1).
\end{aligned}
$$

This allows to conclude directly.

1.6.4 Proof of Lemma 1.5.9

We want to compute

$$E := \mathbb{E}\,\mathrm{Tr}\,\mathbf{AVBV}^*\mathbf{CVDV}^*.$$

Let us denote the entries of $\mathbf{V}$ by v_{ij}, the entries of $\mathbf{A}$ by a_{ij}, the entries of $\mathbf{B}$ by b_{ij}... Then, expanding the trace, we have

$$E = \sum_{1\le\alpha,\beta,i,j,\gamma,\tau,k,l\le n} \underbrace{\mathbb{E}[a_{\alpha\beta}v_{\beta i}b_{ij}\overline{v}_{\gamma j}c_{\gamma\tau}v_{\tau k}d_{kl}\overline{v}_{\alpha l}]}_{:=\mathbb{E}_{\alpha,\beta,i,j,\gamma,\tau,k,l}}$$

By the left and right invariance of the Haar measure on the unitary group (see Proposition 2.5.1), for the expectation of a product of entries of $\mathbf{V}$ and $\overline{\mathbf{V}}$ to be non zero, we need each row to appear as much times in $\mathbf{V}$ as in $\overline{\mathbf{V}}$ and each column to appear as much times in $\mathbf{V}$ as in $\overline{\mathbf{V}}$. It follows that for $\mathbb{E}_{\alpha,\beta,i,j,\gamma,\tau,k,l}$ to be non zero, we need to have the equalities of multisets:

$$\{\beta,\tau\}_m = \{\alpha,\gamma\}_m, \qquad \{i,k\}_m = \{j,l\}_m$$

The first condition is equivalent to one of the three conditions

$$\alpha=\beta=\gamma=\tau \qquad \text{or} \qquad \alpha=\beta\neq\gamma=\tau \qquad \text{or} \qquad \alpha=\tau\neq\beta=\gamma$$

and the second condition is equivalent to one of the three conditions

$$i=j=k=l \qquad \text{or} \qquad i=j\neq k=l \qquad \text{or} \qquad i=l\neq j=k.$$

Hence we have 9 cases to consider below. In each one, the involved moments of the v_{ij}'s are computed thanks to e.g. Proposition 4.2.3 of [52]: for any a,b,c,d, we have

- •• $$\mathbb{E}[|v_{ab}|^4] = \frac{2}{N(N+1)},$$
- •• $$b\neq d \implies \mathbb{E}[|v_{ab}|^2|v_{cd}|^2] = \begin{cases} \frac{1}{N(N+1)} & \text{if } a=c \\ \frac{1}{N^2-1} & \text{if } a\neq c \end{cases}$$
- •• $$a\neq b \text{ and } c\neq d \implies \mathbb{E}[v_{ac}\overline{v_{ad}}v_{bd}\overline{v_{bc}}] = -\frac{1}{N(N^2-1)}$$

So let us treat the 9 cases:

• Under condition $\alpha=\beta=\gamma=\tau$ and $i=j=k=l$, we have

$$\sum_{\alpha,\beta,i,j,\gamma,\tau,k,l} \mathbb{E}_{\alpha,\beta,i,j,\gamma,\tau,k,l} = \sum_{\alpha,i} \mathbb{E}[a_{\alpha\alpha}v_{\alpha i}b_{ii}\overline{v}_{\alpha i}c_{\alpha\alpha}v_{\alpha i}d_{ii}\overline{v}_{\alpha i}] = \frac{2}{N(N+1)}\sum_{\alpha}a_{\alpha\alpha}c_{\alpha\alpha}\sum_i b_{ii}d_{ii}$$

• Under condition $\alpha=\beta=\gamma=\tau$ and $i=j\neq k=l$, we get $\frac{1}{N(N+1)}\sum_{\alpha}a_{\alpha\alpha}c_{\alpha\alpha}\sum_{i\neq k}b_{ii}d_{kk}$

- Under condition $\alpha=\beta=\gamma=\tau$ and $i=l\neq j=k$, we get $\dfrac{1}{N(N+1)}\sum_{\alpha}a_{\alpha\alpha}c_{\alpha\alpha}\sum_{i\neq j}b_{ij}d_{ji}$
- Under condition $\alpha=\beta\neq\gamma=\tau$ and $i=j=k=l$, we get $\dfrac{1}{N(N+1)}\sum_{\alpha\neq\gamma}a_{\alpha\alpha}c_{\gamma\gamma}\sum_{i}b_{ii}d_{ii}$
- Under condition $\alpha=\beta\neq\gamma=\tau$ and $i=j\neq k=l$, we get $\dfrac{-1}{N(N^2-1)}\sum_{\alpha\neq\gamma}a_{\alpha\alpha}c_{\gamma\gamma}\sum_{i\neq k}b_{ii}d_{kk}$
- Under condition $\alpha=\beta\neq\gamma=\tau$ and $i=l\neq j=k$, we get $\dfrac{1}{N^2-1}\sum_{\alpha\neq\gamma}a_{\alpha\alpha}c_{\gamma\gamma}\sum_{i\neq j}b_{ij}d_{ji}$
- Under condition $\alpha=\tau\neq\beta=\gamma$ and $i=j=k=l$, we get $\dfrac{1}{N(N+1)}\sum_{\alpha\neq\beta}a_{\alpha\beta}c_{\beta\alpha}\sum_{i}b_{ii}d_{ii}$
- Under condition $\alpha=\tau\neq\beta=\gamma$ and $i=j\neq k=l$, we get $\dfrac{1}{N^2-1}\sum_{\alpha\neq\beta}a_{\alpha\beta}c_{\beta\alpha}\sum_{i\neq k}b_{ii}d_{kk}$
- Under condition $\alpha=\tau\neq\beta=\gamma$ and $i=l\neq j=k$, we get $\dfrac{-1}{N(N^2-1)}\sum_{\alpha\neq\beta}a_{\alpha\beta}c_{\beta\alpha}\sum_{i\neq j}b_{ij}d_{ji}$

Summing up the nine previous sums, we easily get the desired result:

$$\begin{aligned}
N(N+1)E &= 2\sum_{\alpha}a_{\alpha\alpha}c_{\alpha\alpha}\sum_{i}b_{ii}d_{ii}+\sum_{\alpha}a_{\alpha\alpha}c_{\alpha\alpha}\sum_{i\neq k}b_{ii}d_{kk}+\sum_{\alpha}a_{\alpha\alpha}c_{\alpha\alpha}\sum_{i\neq j}b_{ij}d_{ji}\\
&\quad+\sum_{\alpha\neq\gamma}a_{\alpha\alpha}c_{\gamma\gamma}\sum_{i}b_{ii}d_{ii}-\frac{1}{N-1}\sum_{\alpha\neq\gamma}a_{\alpha\alpha}c_{\gamma\gamma}\sum_{i\neq k}b_{ii}d_{kk}+\frac{N}{N-1}\sum_{\alpha\neq\gamma}a_{\alpha\alpha}c_{\gamma\gamma}\sum_{i\neq j}b_{ij}d_{ji}\\
&\quad+\sum_{\alpha\neq\beta}a_{\alpha\beta}c_{\beta\alpha}\sum_{i}b_{ii}d_{ii}+\frac{N}{N-1}\sum_{\alpha\neq\beta}a_{\alpha\beta}c_{\beta\alpha}\sum_{i\neq k}b_{ii}d_{kk}-\frac{1}{N-1}\sum_{\alpha\neq\beta}a_{\alpha\beta}c_{\beta\alpha}\sum_{i\neq j}b_{ij}d_{ji}\\
&= \sum_{\alpha}a_{\alpha\alpha}c_{\alpha\alpha}\sum_{i,k}b_{ii}d_{kk}+\sum_{\alpha}a_{\alpha\alpha}c_{\alpha\alpha}\sum_{i,j}b_{ij}d_{ji}\\
&\quad+\frac{N}{N-1}\sum_{\alpha\neq\gamma}a_{\alpha\alpha}c_{\gamma\gamma}\sum_{i,j}b_{ij}d_{ji}-\frac{1}{N-1}\sum_{\alpha\neq\gamma}a_{\alpha\alpha}c_{\gamma\gamma}\sum_{i,k}b_{ii}d_{kk}\\
&\quad+\frac{N}{N-1}\sum_{\alpha\neq\beta}a_{\alpha\beta}c_{\beta\alpha}\sum_{i,k}b_{ii}d_{kk}-\frac{1}{N-1}\sum_{\alpha\neq\beta}a_{\alpha\beta}c_{\beta\alpha}\sum_{i,j}b_{ij}d_{ji}\\
&= \mathrm{Tr}\,\mathbf{B}\,\mathrm{Tr}\,\mathbf{D}\left\{\sum_{\alpha}a_{\alpha\alpha}c_{\alpha\alpha}-\frac{1}{N-1}\sum_{\alpha\neq\gamma}a_{\alpha\alpha}c_{\gamma\gamma}+\frac{N}{N-1}\sum_{\alpha\neq\beta}a_{\alpha\beta}c_{\beta\alpha}\right\}\\
&\quad+\mathrm{Tr}\,\mathbf{BD}\left\{\sum_{\alpha}a_{\alpha\alpha}c_{\alpha\alpha}+\frac{N}{N-1}\sum_{\alpha\neq\gamma}a_{\alpha\alpha}c_{\gamma\gamma}-\frac{1}{N-1}\sum_{\alpha\neq\beta}a_{\alpha\beta}c_{\beta\alpha}\right\}\\
&= \frac{N}{N-1}\left\{\mathrm{Tr}\,\mathbf{AC}\,\mathrm{Tr}\,\mathbf{B}\,\mathrm{Tr}\,\mathbf{D}+\mathrm{Tr}\,\mathbf{A}\,\mathrm{Tr}\,\mathbf{C}\,\mathrm{Tr}\,\mathbf{BD}\right\}\\
&\quad-\frac{1}{N-1}\left\{\mathrm{Tr}\,\mathbf{AC}\,\mathrm{Tr}\,\mathbf{BD}+\mathrm{Tr}\,\mathbf{A}\,\mathrm{Tr}\,\mathbf{C}\,\mathrm{Tr}\,\mathbf{B}\,\mathrm{Tr}\,\mathbf{D}\right\}.
\end{aligned}$$

Acknowledgments: We would like to thank J. Novak for discussions on Weingarten calculus.

II

Fluctuations in the Single Ring Theorem

Florent Benaych-Georges and Jean Rochet

Abstract :

We consider a non-Hermitian random matrix $\mathbf{A}$ whose distribution is invariant under the left and right actions of the unitary group. The so-called *Single Ring Theorem*, proved by Guionnet, Krishnapur and Zeitouni [49], states that the empirical eigenvalue distribution of $\mathbf{A}$ converges to a limit measure supported by a ring S. In this text, we establish the convergence in distribution of random variables of the type $\mathrm{Tr}(f(\mathbf{A})\mathbf{M})$ where f is analytic on S and the Frobenius norm of $\mathbf{M}$ has order $\sqrt{N}$. As corollaries, we obtain central limit theorems for linear spectral statistics of $\mathbf{A}$ (for analytic test functions) and for finite rank projections of $f(\mathbf{A})$ (like matrix entries). As an application, we locate outliers in multiplicative perturbations of $\mathbf{A}$.

2.1 Introduction

The Single Ring Theorem, by Guionnet, Krishnapur and Zeitouni [49], describes the empirical distribution of the eigenvalues of a large generic matrix with prescribed singular values, *i.e.* an $N \times N$ matrix of the form $\mathbf{A} = \mathbf{UTV}$, with $\mathbf{U}, \mathbf{V}$ some independent Haar-distributed unitary matrices and $\mathbf{T}$ a deterministic matrix whose singular values are the ones prescribed. More precisely, under some technical hypotheses, as the dimension N tends to infinity, if the empirical distribution of the singular values of $\mathbf{A}$ converges to a compactly supported limit measure Θ on the real line, then the empirical eigenvalues distribution of $\mathbf{A}$ converges to a limit measure μ on the complex plane which depends only on Θ. The limit measure μ (see Figure 2.8(a)) is rotationally invariant in $\mathbb{C}$ and its support is the annulus $S := \{z \in \mathbb{C} \,;\, a \leq |z| \leq b\}$, with $a, b \geq 0$ such that

$$a^{-2} = \int x^{-2} \mathrm{d}\Theta(x) \qquad \text{and} \qquad b^2 = \int x^2 \mathrm{d}\Theta(x).$$

In this text, we consider such a matrix $\mathbf{A}$ and we study (Theorem 2.2.4) the joint weak convergence, as $N \to \infty$, of random variables of the type

$$\mathrm{Tr}(f(\mathbf{A})\mathbf{M}),$$

for f an analytic function on the ring S whose Laurent series expansion has null constant term and $\mathbf{M}$ a deterministic $N \times N$ matrix satisfying some limit conditions. These limit conditions (see (2.1)) allow to consider both:

– fluctuations, around their limits as predicted by the Single Ring Theorem, of linear spectral statistics of $\mathbf{A}$ (for $\mathbf{M} = \mathbf{I}$):

$$\mathrm{Tr}\, f(\mathbf{A}) = \sum_{i=1}^{N} f(\lambda_i),$$

where $\lambda_1, \ldots, \lambda_N$ denote the eigenvalues of $\mathbf{A}$,

– finite rank projections of $f(\mathbf{A})$ (for $\mathbf{M} = \sqrt{N} \times$(a matrix with bounded rank)), like the matrix entries of $f(\mathbf{A})$.

Let us present both of these directions with more details.

2.1.1 Linear spectral statistics of A

As far as Hermitian random matrices are concerned, linear spectral statistics fluctuations usually come right after the macroscopic behavior, with the microscopic one, in the natural questions that arise (see e.g., among the wide literature on the subject, [58, 60, 88, 56, 8, 7, 64, 87, 1, 64, 5, 32, 18]). For unitary or orthogonal matrices, also, many results have been proved (see e.g. the results of Diaconis *et al* in [39, 40], the ones of Soshnikov in [90] or the ones of Lévy and Maïda in [63]). For non-Hermitian matrices, established results are way less numerous: the first one was [83], by Rider and Silverstein, for analytic test functions of matrices with i.i.d. entries, then came the paper [84] by Rider and Virág, who managed, thanks to the explicit determinantal structure of the correlation functions of the Ginibre ensemble, to study the fluctuations of linear spectral statistics of such matrices for $\mathscr{C}^1$ test functions. Recently, in [74], O'Rourke and Renfrew studied the fluctuations of linear spectral statistics of elliptic matrices for analytic test functions. The reason why, except for the breakthrough of Rider and Virág in [84], many results are limited to analytic test functions is that when non-normal matrices are concerned, functional calculus makes sense only for analytic functions: if one denotes by $\lambda_1,\ldots,\lambda_N$ the eigenvalues of a non-Hermitian matrix $\mathbf{A}$, one can estimate $\sum_{i=1}^N f(\lambda_i)$ out of the numbers $\operatorname{Tr}\mathbf{A}^k$ or $\operatorname{Tr}((z-\mathbf{A})^{-1})$ only when f is analytic. For a $\mathscr{C}^2$ test function f, one relies on the explicit joint distribution of the λ_i's or on Girko's so-called *Hermitization technique*, which expresses the empirical spectral measure of $\mathbf{A}$ as the Laplacian of the function $z \longmapsto \log|\det(z-\mathbf{A})|$ (see e.g. [48, 25]). This is a way more difficult problem, which we consider in a forthcoming project.

In this text, as a corollary of our main theorem, we prove that for $\mathbf{A} = \mathbf{UTV}$ an $N\times N$ matrix of the type introduced above and $f(z) = \sum_{n\in\mathbb{Z}} a_n z^n$ an analytic function on a neighborhood of the limit support S of the empirical eigenvalue distribution of $\mathbf{A}$, the random variable

$$\operatorname{Tr} f(\mathbf{A}) - Na_0$$

converges in distribution, as $N\to\infty$, to a centered complex Gaussian random variable with a given covariance matrix (see Corollary 2.2.7). This is a first step in the study of the noise in the Single Ring Theorem. We notice a quite common fact in random matrix theory: the random variable $\operatorname{Tr} f(\mathbf{A}) - Na_0 = \sum_{i=1}^N f(\lambda_i) - \mathbb{E} f(\lambda_i)$ does not need to be renormalized to have a limit in distribution, which reflects the eigenvalue repulsion phenomenon (indeed, would the λ_i's have been i.i.d., this random variable would have had order $\sqrt{N}$).

Next, two corollaries are given (Corollaries 2.2.11 and 2.2.13), one about the Bergman kernel and the resolvant and one about the log-correlated limit distribution of the characteristic polynomial out of the support.

2.1.2 Finite rank projections and matrix entries

A century ago, in 1906, Émile Borel proved in [26] that, for a uniformly distributed point $(X_1,\dots,X_N)$ on the unit Euclidian sphere $\mathbb{S}^{N-1}$, the scaled first coordinate $\sqrt{N}X_1$ converges weakly to the Gaussian distribution as the dimension N tends to infinity. As explained in the introduction of the paper [3] of Diaconis *et al.*, this means that the features of the “microcanonical" ensemble in a certain model for statistical mechanics (uniform measure on the sphere) are captured by the “canonical" ensemble (Gaussian measure). Since then, a long list of further-reaching results about the asymptotic normality of entries of random orthogonal or unitary matrices have been obtained (see e.g. [3, 65, 33, 37, 55, 14]).

In this text, as a corollary of our main theorem, we prove that for $\mathbf{A}=\mathbf{UTV}$ an $N\times N$ matrix of the type introduced above, $f(z)=\sum_{n\in\mathbb{Z}} a_n z^n$ an analytic function on a neighborhood of the limit support S of the empirical eigenvalue distribution of $\mathbf{A}$ and $\mathbf{a},\mathbf{b}$ some unit column vectors, the random variables of the type

$$\sqrt{N}(\mathbf{b}^* f(\mathbf{A})\mathbf{a}-a_0\mathbf{b}^*\mathbf{a})$$

converge in joint distribution, as $N\to\infty$, to centered complex Gaussian random variables with a given covariance matrix (see Corollary 2.2.15). This allows for example to consider matrix entries of $f(\mathbf{A})$, in the vein of the works of Soshnikov *et al.* for Wigner matrices in [76, 80] (see Corollary 2.2.18 and Remark 2.2.19). It also applies to the study of finite rank perturbations of $\mathbf{A}$ of *multiplicative type*: the BBP phase transition (named after the authors of the seminal paper [9]) is well understood for additive or multiplicative perturbations ($\widetilde{\mathbf{A}}=\mathbf{A}+\mathbf{P}$ or $\widetilde{\mathbf{A}}=\mathbf{A}(\mathbf{I}+\mathbf{P})$) of general Hermitian models (see [77, 30, 19] or [9, 20]), for additive perturbations of various non-Hermitian models (see [92, 21, 73, 24]), but multiplicative perturbations of non-Hermitian models were so far unexplored. In Remark 2.2.16 and Figure 2.8, we explain briefly how our results allow to enlighten a BBP transition for such perturbations.

Organisation of the paper: In the next section, we state our main theorem (Theorem 2.2.4) and its corollaries, all of which go without proof, except Corollary 2.2.13. The rest of the paper is devoted to the proof of Theorem 2.2.4, to the proof of Corollary 2.2.13 and to the appendix, where we state several technical results needed here.

2.2 Main result

Let $\mathbf{A}$ be a random $N\times N$ matrix implicitly depending on N such that $\mathbf{A}=\mathbf{UTV}$, with $\mathbf{U},\mathbf{V},\mathbf{T}$ independent and $\mathbf{U},\mathbf{V}$ Haar-distributed on the unitary group. We make the following hypotheses on $\mathbf{T}$:

Assumption 2.2.1. As $N \to \infty$, the sequence $\left(N^{-1}\operatorname{Tr}\mathbf{T}\mathbf{T}^*\right)^{1/2}$ converges in probability to a deterministic limit $b > 0$ and there is $M < \infty$ such that with probability tending to one, $\|\mathbf{T}\|_{\mathrm{op}} \leq M$.

Assumption 2.2.2. With the convention $1/\infty = 0$ and $1/0 = \infty$, the sequence

$$\left(N^{-1}\operatorname{Tr}((\mathbf{T}\mathbf{T}^*)^{-1})\right)^{-1/2}$$

converges in probability to a deterministic limit $a \geq 0$. If $a > 0$, we also suppose that there is $M' < \infty$ such that with probability tending to one, $\|\mathbf{T}^{-1}\|_{\mathrm{op}} \leq M'$.

The following assumption, which could possibly be relaxed following Basak and Dembo's approach of [10], is made to control tails of Laurent series but can be removed if the f_j's have finite Laurent expansion, like in Corollary 2.2.8 or in Remark 2.2.19.

Assumption 2.2.3. There exist a constant $\kappa > 0$ such that

$$\operatorname{Im}(z) \;>\; n^{-\kappa} \implies N^{-1}\left|\operatorname{Im}\operatorname{Tr}((z-\sqrt{\mathbf{T}\mathbf{T}^*})^{-1})\right| \;\leq\; \frac{1}{\kappa}.$$

For f an analytic function on a neighborhood of the ring

$$S := \{z \in \mathbb{C}\,;\, a \leq |z| \leq b\},$$

the matrix $f(\mathbf{A})$ is well defined with probability tending to one as $N \to \infty$, as it was proved in [50, 15] that the spectrum of $\mathbf{A}$ is contained in any neighborhood of S with probability tending to one. We denote the Laurent series expansion, on S, of any such function f by

$$f(z) = \sum_{n\in\mathbb{Z}} a_n(f) z^n.$$

Theorem 2.2.4. *For each $N \geq 1$, let $\mathbf{M}_1, \dots, \mathbf{M}_k$ be $N \times N$ deterministic matrices such that for all i, j, as $N \to \infty$,*

$$\frac{1}{N}\operatorname{Tr}\mathbf{M}_i \longrightarrow \tau_i, \quad \frac{1}{N}\operatorname{Tr}\mathbf{M}_i\mathbf{M}_j^* \longrightarrow \alpha_{ij}, \quad \frac{1}{N}\operatorname{Tr}\mathbf{M}_i\mathbf{M}_j \longrightarrow \beta_{ij} \tag{2.1}$$

Let $f_1, \dots, f_k$ be analytic on a neighborhood of S. Then, as $N \to \infty$, the random vector

$$\Big(\operatorname{Tr} f_j(\mathbf{A})\mathbf{M}_j - a_0(f_j)\operatorname{Tr}\mathbf{M}_j\Big)_{j=1}^k$$

converges to a centered complex Gaussian vector $(\mathscr{G}(f_1), \dots, \mathscr{G}(f_k))$ whose distribution is defined by

$$\begin{aligned}
\mathbb{E}\mathscr{G}(f_i)\mathscr{G}(f_j) &= \sum_{n\geq 1}\big((n-1)\tau_i\tau_j + \beta_{ij}\big)\big(a_n(f_i)a_{-n}(f_j) + a_{-n}(f_i)a_n(f_j)\big)\\
\mathbb{E}\mathscr{G}(f_i)\overline{\mathscr{G}(f_j)} &= \sum_{n\geq 1}\big((n-1)\tau_i\overline{\tau_j} + \alpha_{ij}\big)\big(a_n(f_i)\overline{a_n(f_j)}b^{2n} + a_{-n}(f_i)\overline{a_{-n}(f_j)}a^{-2n}\big).
\end{aligned}$$

Remark 2.2.5. Note that if $a = 0$, as the f_j's are analytic on S, we have $a_{-n}(f_j) = 0$ for each $n \geq 1$ and each j, so that the above expression still makes sense. Besides, it seems reasonable to verify that the two series above converge:

$$\begin{aligned}
\sum_{n\geq 1} n|a_n(f_i)||a_n(f_j)|b^{2n} &\leq \big(\max_{n\geq 1}|a_n(f_j)|b^n\big)\sum_{n\geq 1} n|a_n(f_i)|b^n < \infty\\
\sum_{n\geq 1} n|a_n(f_i)||a_{-n}(f_j)| &\leq \big(\max_{n\geq 1}|a_n(f_i)|b^n\big)\sum_{n\geq 1} n|a_{-n}(f_j)|a^{-n} < \infty.
\end{aligned}$$

Remark 2.2.6 (Relation to second order freeness)**.** A theory has been developed recently about Gaussian fluctuations (called *second order limits*) of traces of large random matrices, the most emblematic articles in this theory being [66, 68, 69, 34]. Theorem 2.2.4 can be compared to some of these results. However, our hypotheses on the matrices we consider are of a different nature than the ones of the previously cited papers, since the convergence of the non commutative distributions is not required here: our hypotheses are satisfied for example by matrices like $\mathbf{M}_j = \sqrt{N}\times$(an elementary $N \times N$ matrix), which have no bounded moments of order higher than two.

Our two main applications are the case where the $\mathbf{M}_j$'s are all $\mathbf{I}$ (Corollaries 2.2.7 and 2.2.8) and the cases where the $\mathbf{M}_j$'s are $\sqrt{N}$ times matrices with bounded rank and norm, like elementary matrices (Corollaries 2.2.15 and 2.2.18). In the case $\mathbf{M} = \mathbf{I}$, we immediately obtain the following corollary about linear spectral statistics of $\mathbf{A}$.

Corollary 2.2.7. *Let $f_1, \dots, f_k$ be analytic on a neighborhood of S. Then, as $N \to \infty$, the random vector*

$$\Big(\operatorname{Tr} f_j(\mathbf{A}) - Na_0(f_j)\Big)_{j=1}^k$$

converges to a centered complex Gaussian vector $(\mathscr{G}(f_1),\ldots,\mathscr{G}(f_k))$ *such that*

$$\begin{aligned}\mathbb{E}\mathscr{G}(f_i)\mathscr{G}(f_j) &= \sum_{n\in\mathbb{Z}}|n|a_n(f_i)a_{-n}(f_j)\\ \mathbb{E}\mathscr{G}(f_i)\overline{\mathscr{G}(f_j)} &= \sum_{n\geq 1}n\big(a_n(f_i)\overline{a_n(f_j)}b^{2n}+a_{-n}(f_i)\overline{a_{-n}(f_j)}a^{-2n}\big).\end{aligned}$$

For $n\geq 1$, let us define the functions

$$\varphi_n^{\pm}(z) := \left(\frac{z}{b}\right)^n \pm \left(\frac{a}{z}\right)^n.$$

These functions (plus the constant one) define a basis of the space of analytic functions on a neighborhood of S and we have the change of basis formula

$$\sum_{n\in\mathbb{Z}}a_nz^n = a_0+\sum_{n\geq 1}c_n^+\varphi_n^+(z)+c_n^-\varphi_n^-(z) \iff \forall n\geq 1,\ \begin{pmatrix}a_n\\ a_{-n}\end{pmatrix} = \begin{pmatrix}b^{-n} & b^{-n}\\ a^n & -a^n\end{pmatrix}\begin{pmatrix}c_n^+\\ c_n^-\end{pmatrix},$$

implying that

$$\sum_{n\geq 1}|a_n(f)|^2b^{2n}+|a_{-n}(f)|^2a^{-2n} = 2\sum_{n\geq 1}|c_n^+(f)|^2+|c_n^-(f)|^2.$$

Besides, these functions allow to identify the underlying white noise in Theorem 2.2.4 (we only state it here in the case $\mathbf{M}_j=\mathbf{I}$, but this of course extends to the case of general $\mathbf{M}_j$'s, allowing for example to state analogous results for the matrix entries).

Corollary 2.2.8 (Underlying white noise)**.** *The finite dimensional marginal distributions of*

$$(\operatorname{Tr}\varphi_n^+(\mathbf{A}))_{n\geq 1}\cup(\operatorname{Tr}\varphi_n^-(\mathbf{A}))_{n\geq 1}$$

converge to the ones of a collection $(\mathscr{G}_n^+)_{n\geq 1}\cup(\mathscr{G}_n^-)_{n\geq 1}$ *of independent centered complex Gaussian random variables satisfying*

$$\mathbb{E}\,|\mathscr{G}_n^{\pm}|^2 = 2 \quad ; \quad \mathbb{E}(\mathscr{G}_n^{\pm})^2 = \pm 2n(a/b)^n.$$

Remark 2.2.9 (Ginibre matrices)**.** In the particular case where $\mathbf{A}$ is a Ginibre matrix (*i.e.* with i.i.d. entries with law $\mathscr{N}_{\mathbb{C}}(0,N^{-1})$), we find back the result of Rider and Silverstein [83], noticing that in this case $a=0$ and $b=1$, so that $a_n(f)=0$

when $n<0$ and $\mathbb{E}\mathscr{G}(f_i)\mathscr{G}(f_j)=0$, and, for $\mathrm{d}m(z)$ the Lebesgue measure on $\mathbb{C}$,

$$\begin{aligned}
&\frac{1}{\pi}\int_{|z|<1}\frac{\partial}{\partial z}f_i(z)\overline{\frac{\partial}{\partial z}f_j(z)}\mathrm{d}m(z)\\
=\;&\frac{1}{\pi}\int_{|z|<1}-\frac{1}{4\pi^2}\oint_{\mathrm{Circle}(1+\varepsilon)}\oint_{\mathrm{Circle}(1+\varepsilon)}\frac{f_i(\xi_1)}{(\xi_1-z)^2}\frac{\overline{f_j(\xi_2)}}{(\overline{\xi_2}-\overline{z})^2}\mathrm{d}\xi_1\mathrm{d}\xi_2\mathrm{d}m(z)\\
=\;&-\frac{1}{4\pi^2}\oint_{\mathrm{Circle}(1+\varepsilon)}\oint_{\mathrm{Circle}(1+\varepsilon)}\frac{f_i(\xi_1)\overline{f_j(\xi_2)}}{\xi_1^2\overline{\xi_2}^2}\frac{1}{\pi}\int_{|z|<1}\sum_{n,n'\geq 1}nn'\left(\frac{z}{\xi_1}\right)^{n-1}\left(\frac{\overline{z}}{\overline{\xi}_2}\right)^{n'-1}\mathrm{d}m(z)\mathrm{d}\xi_1\mathrm{d}\xi_2\\
=\;&-\frac{1}{4\pi^2}\oint_{\mathrm{Circle}(1+\varepsilon)}\oint_{\mathrm{Circle}(1+\varepsilon)}f_i(\xi_1)\overline{f_j(\xi_2)}\sum_{n\geq 1}n(\xi_1\overline{\xi_2})^{-n-1}\mathrm{d}\xi_1\mathrm{d}\xi_2\\
=\;&\sum_{n\geq 1}na_n(f_i)\overline{a_n(f_j)}
\end{aligned}$$

Remark 2.2.10. If $\mathbf{T}=\mathbf{I}$, and the f_k's are polynomial, we find back a result of Diaconis and Shahshahani [39, Theorem 1] on the limit joint distribution of

$$\left(\mathrm{Tr}(\mathbf{U}^k)\right)_{k=1}^{n},$$

where $\mathbf{U}$ is Haar-distributed. Actually, the Corollary 2.2.7 is slightly stronger, since the result holds for $\mathbf{A}=\mathbf{UT}$ or $\mathbf{A}=\mathbf{UTV}$ as long as $\mathbf{T}$ satisfies

$$\lim_{N\to\infty}\frac{1}{N}\mathrm{Tr}(\mathbf{T}^2)\;=\lim_{N\to\infty}\frac{1}{N}\mathrm{Tr}(\mathbf{T}^{-2})=1, \tag{2.2}$$

in which case $\mathbf{A}$ may be seen as a multiplicative perturbation of $\mathbf{U}$. $\mathbf{T}$ satisfies the condition (2.2) if for example all its diagonal coefficients are equal to 1 except $o(N)$ of them (which stay away from 0 and ∞).

Corollary 2.2.11 (Bergman kernel and resolvant)**.** *The random process*

$$\left(\mathrm{Tr}(z-\mathbf{A})^{-1}\right)_{|z|<a}\cup\left(\mathrm{Tr}(z-\mathbf{A})^{-1}\right)_{|z|>b}$$

converges, for the finite-dimensional distributions, to a centered complex Gaussian process

$$(\mathscr{G}_z)_{|z|<a}\cup(\mathscr{H}_z)_{|z|>b}$$

with covariance defined by

$$\mathbb{E}\,\mathscr{G}_z\overline{\mathscr{G}_{z'}}=\frac{b^2}{(b^2-z\overline{z'})^2},\qquad \mathbb{E}\,\mathscr{H}_z\overline{\mathscr{H}_{z'}}=\frac{a^2}{(a^2-z\overline{z'})^2},\qquad \mathbb{E}\,\mathscr{G}_z\mathscr{H}_{z'}=-\frac{1}{(z'-z)^2}.$$

and by the fact that

$$\forall\theta\in\mathbb{R},\qquad \left(\mathrm{e}^{-\mathrm{i}\theta}\mathscr{G}_z\right)_{|z|<a}\cup\left(\mathrm{e}^{\mathrm{i}\theta}\mathscr{H}_z\right)_{|z|>b}\overset{law}{=}(\mathscr{G}_z)_{|z|<a}\cup(\mathscr{H}_z)_{|z|>b}.$$

Remark 2.2.12. The kernel of the limit Gaussian analytic function, in the previous corollary, is, up to a constant factor, the *Bergman kernel* (see [13, 78]).

Corollary 2.2.13 (Characteristic polynomial out of the support)**.** *The random process*

$$(\log|\det(z-\mathbf{A})| - \operatorname{Tr}\log\mathbf{T})_{|z|<a} \cup (\log|\det(z-\mathbf{A})| - N\log|z|)_{|z|>b}$$

converges, for the finite-dimensional distributions, to a centered real Gaussian process

$$(\mathscr{G}_z)_{|z|<a} \cup (\mathscr{H}_z)_{|z|>b}$$

with covariance defined by

$$2\,\mathbb{E}\,\mathscr{G}_z\mathscr{G}_{z'} = -\log|1-\frac{z\overline{z'}}{a^2}|, \qquad 2\,\mathbb{E}\,\mathscr{H}_z\mathscr{H}_{z'} = -\log|1-\frac{b^2}{z\overline{z'}}|, \qquad 2\,\mathbb{E}\,\mathscr{G}_z\mathscr{H}_{z'} = -\log|1-\frac{z}{z'}|.$$

Remark 2.2.14. As $z \neq z'$ both tend to the same point of the boundary of S, the above covariances are equivalent to $-\log|z-z'|$. In the light of the log-correlation approach to the Gaussian Free Field (see [41]), it supports the idea that on the limit support S, the characteristic polynomial of $\mathbf{A}$ should tend to an object related to the Gaussian Free Field, as for Ginibre matrices (see Corollary 2 of [84]). It would be interesting to see in what extent such a convergence depends on the hypotheses made on the precise distribution of the singular values of $\mathbf{T}$.

In the case $\mathbf{M}_j = \sqrt{N}\mathbf{a}_j\mathbf{b}_j^*$, we immediately obtain the following corollary:

Corollary 2.2.15. *For each $N \geq 1$, let $\mathbf{a}_1, \mathbf{b}_1, \ldots, \mathbf{a}_k, \mathbf{b}_k$ be deterministic column vectors with size N such that for all i, j, as $N \to \infty$,*

$$\mathbf{a}_i^*\mathbf{a}_j \longrightarrow \kappa_{ij}^{\mathrm{a,a}} \in \mathbb{C} \qquad ; \qquad \mathbf{b}_i^*\mathbf{a}_j \longrightarrow \kappa_{ij}^{\mathrm{b,a}} \in \mathbb{C} \qquad ; \qquad \mathbf{b}_i^*\mathbf{b}_j \longrightarrow \kappa_{ij}^{\mathrm{b,b}} \in \mathbb{C}$$

Let $f_1, \ldots, f_k$ be analytic on a neighborhood of $S = \mathrm{Ring}(a,b)$. Then, as $N \to \infty$, the random vector

$$\sqrt{N}\Big(\mathbf{b}_j^* f_j(\mathbf{A})\mathbf{a}_j - \mathbf{b}_j^*\mathbf{a}_j a_0(f_j)\Big)_{j=1}^{k}$$

converges to a centered complex Gaussian vector $(\mathscr{G}(f_1), \ldots, \mathscr{G}(f_k))$ such that

$$\begin{aligned} \mathbb{E}\,\mathscr{G}(f_i)\mathscr{G}(f_j) &= \sum_{n\geq 1} \kappa_{ji}^{\mathrm{a,a}}\kappa_{ij}^{\mathrm{b,b}}\big(a_n(f_i)a_{-n}(f_j) + a_{-n}(f_i)a_n(f_j)\big) \\ \mathbb{E}\,\mathscr{G}(f_i)\overline{\mathscr{G}(f_j)} &= \sum_{n\geq 1} \kappa_{ji}^{\mathrm{b,a}}\kappa_{ij}^{\mathrm{b,a}}\big(a_n(f_i)\overline{a_n(f_j)}b^{2n} + a_{-n}(f_i)\overline{a_{-n}(f_j)}a^{-2n}\big). \end{aligned}$$

Remark 2.2.16 (Application to multiplicative finite rank perturbations of $\mathbf{A}$)**.** The previous corollary has several applications to the study of the outliers of spiked models related to the Single Ring Theorem. It allows for example to understand easily, using the technics developed in [21], the impact of *multiplicative* finite rank perturbations on the spectrum of $\mathbf{A}$ (whereas only *additive* perturbations had been studied so far). For example, one can deduce from this corollary that for $\mathbf{P}$ a deterministic matrix with bounded operator norm and rank one, if one defines $\widetilde{\mathbf{A}} := \mathbf{A}(\mathbf{I}+\mathbf{P})$ and $\hat{\mathbf{A}} := \mathbf{A}(\mathbf{I}+\mathbf{AP})$, then

- the matrix $\widetilde{\mathbf{A}}$ has no outlier (*i.e.* the support of its spectrum still converges to S),
- the matrix $\hat{\mathbf{A}}$ has no outlier with modulus $> b$, but each eigenvalue λ of $\mathbf{P}$ such that $|\lambda| > a^{-1}$ gives rise to an outlier of $\hat{\mathbf{A}}$ located approximately at $-\lambda^{-1}$ (besides, when the multiplicity of λ as an eigenvalue of $\mathbf{P}$ is 1, the fluctuations of the outlier around $-\lambda^{-1}$ are Gaussian and with order $N^{-1/2}$).

This phenomenon is illustrated by Figure 2.8.

To state the next corollary, let us first give the definition of Gaussian elliptic matrices.

Definition 2.2.17 (Gaussian elliptic matrices)**.** Let $\rho \in \mathbb{C}$ such that $|\rho| \leq 1$. A *Gaussian elliptic matrix with parameter* ρ is an $N \times N$ Gaussian centered complex random matrix $\mathbf{X} = (x_{ij})$ satisfying :

(1) the random vectors $\left(x_{ij}, x_{ji}\right)_{i \leq j}$ are independent,

(2) $\forall i$, $\mathbb{E}|x_{ii}|^2 = 1$ and $\mathbb{E}x_{ii}^2 = \rho$,

(3) $\forall i \neq j$, $\mathbb{E}|x_{ij}|^2 = 1$, $\mathbb{E}x_{ij}^2 = 0$, $\mathbb{E}x_{ij}x_{ji} = \rho$ and $\mathbb{E}x_{ij}\overline{x_{ji}} = 0$.

This matrix ensemble was introduced by Girko in [48], and its name is due to the fact that its empirical eigenvalue distribution is the uniform distribution inside an ellipse. In the case where $\rho = 0$ (resp. $\rho = 1$), we get a Ginibre (resp. GUE) matrix.

Corollary 2.2.18. *Let f be analytic on a neighborhood of $S = \mathrm{Ring}(a,b)$ such that*

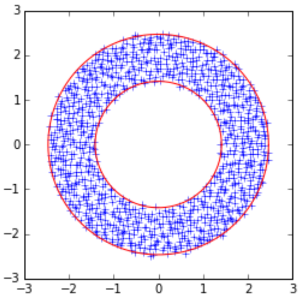
$$\sum_{n \geq 1} |a_n(f)|^2 b^{2n} + |a_{-n}(f)|^2 a^{-2n} \;=\; 1.$$

Let k be a fixed positive integer and let $I = I(N)$ be a (possibly depending on N) subset of $\{1, \ldots, N\}$ with cardinality k. Let us define the random $k \times k$ matrix

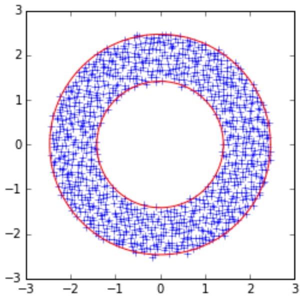
$$\mathbf{X}_N := \sqrt{N}\left[f(\mathbf{A})_{ij} - a_0(f)\delta_{ij}\right]_{(i,j) \in I \times I}.$$

Then, as $N \to \infty$, the matrix $\mathbf{X}_N$ converges in distribution to a $k \times k$ Gaussian elliptic matrix $\mathbf{X}$ with parameter ρ,

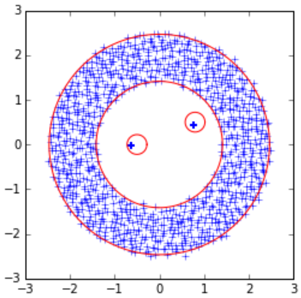
$$\rho \;:=\; 2\sum_{n \geq 1} a_n(f) a_{-n}(f).$$



(a) Spectrum of $\mathbf{A}$



(b) Spectrum of $\widetilde{\mathbf{A}} := \mathbf{A}(\mathbf{I}+\mathbf{P})$



(c) Spectrum of $\hat{\mathbf{A}} := \mathbf{A}(\mathbf{I}+\mathbf{AP})$ (small circles are centered at the theoretical limit locations of the outliers)

Figure 2.8: **Outliers/lack of outliers for multiplicative perturbations:** simulation realized with a single $10^3 \times 10^3$ matrix $\mathbf{A} = \mathbf{UTV}$ when the singular values of $\mathbf{T}$ are uniformly distributed on $[0.5, 4]$ and $\mathbf{P} = \operatorname{diag}(-2, (0.8+0.5\mathrm{i})^{-1}, 1/3, 0, \dots, 0)$. As predicted, none of these matrices has any outlier outside the outer circle, nor do the two first ones inside the inner circle, but $\hat{\mathbf{A}}$ has two outliers inside the inner circle, close to the predicted locations.

Remark 2.2.19. a) In the case where $f(z) = z$, we find back the well-known result that any principal submatrix with fixed size of $\sqrt{N}\mathbf{U}$ matrix converges to a Ginibre matrix (see e.g. [3, 55]).

b) By Corollary 2.2.18, the statement of the first part of this remark happens to stay true, up to a constant multiplicative factor, if $\mathbf{U}$ is replaced by $\mathbf{A} = \mathbf{UTV}$ or even by $\mathbf{A}^n$ or by $f(\mathbf{A})$ if f is analytic in a neighborhood of the disc $\overline{B}(0,b)$.

c) It also follows from what precedes that for any $n \geq 1$, any sequence of principal submatrices with fixed size of $\sqrt{N/2}(\mathbf{U}^n + \mathbf{U}^{-n})$ and $\sqrt{N/2}(\mathbf{U}^n - \mathbf{U}^{-n})$ converge in distribution to a GUE matrix and i times a GUE matrix, both being independent.

2.3 Proof of Theorem 2.2.4

To avoid having to treat the cases $a > 0$ and $a = 0$ separately all along the proof, we shall suppose that $a > 0$ (the case $a = 0$ is more simple, as sums run only on $n \geq 0$). Besides, note that by invariance of the Haar measure, the distribution of the random matrix $\mathbf{A}$ depends on $\mathbf{T}$ only through its singular values, so we shall suppose that $\mathbf{T} = \mathrm{diag}(s_1, \ldots, s_N)$, with $s_i \geq 0$. At last, as the limit distributions, in Theorem 2.2.4, only depend on $\mathbf{T}$ only through the deterministic parameters a, b, up to a conditioning, one can suppose that $\mathbf{T}$ is deterministic (and that both $\|\mathbf{T}\|_{\mathrm{op}}$ and $\|\mathbf{T}^{-1}\|_{\mathrm{op}}$ are uniformly bounded, by Asssumptions 2.2.1 and 2.2.2).

2.3.1 Randomization of the $\mathbf{M}_j$'s

Let us define $\mathbf{W} := \mathbf{VU}$. The random matrix $\mathbf{W}$ is also Haar-distributed and independent from $\mathbf{V}$. Besides, for each j, as $\mathbf{A} = \mathbf{UTV} = \mathbf{V}^*\mathbf{WTV}$,

$$\mathrm{Tr}\, f_j(\mathbf{A})\mathbf{M}_j = \mathrm{Tr}\, \mathbf{V}^* f_j(\mathbf{WT})\mathbf{VM}_j = \mathrm{Tr}\, f_j(\mathbf{WT})\mathbf{VM}_j\mathbf{V}^*$$

As a consequence, we shall suppose that $\mathbf{A} = \mathbf{UT}$ (instead of $\mathbf{A} = \mathbf{UTV}$) and that there is a Haar-distributed random unitary matrix $\mathbf{V}$, independent of $\mathbf{U}$, such that for each j, $\mathbf{M}_j = \mathbf{V}\widetilde{\mathbf{M}}_j\mathbf{V}^*$, with $\widetilde{\mathbf{M}}_1, \ldots, \widetilde{\mathbf{M}}_k$ a collection of deterministic matrices also satisfying (2.1).

2.3.2 Tails of the series

Let us first prove that Theorem 2.2.4 can be deduced from the particular case where there is n_0 such that for all n, we have

$$|n| > n_0 \implies \forall j = 1, \ldots, k,\ a_n(f_j) = 0.$$

Let $\varepsilon \in (0, a/2)$ such that the domain of each f_j contains $\mathrm{Ring}(a - 2\varepsilon, b + 2\varepsilon)$.

Lemma 2.3.1. *There is a constant C independent of N such that for any n such that $n^6 \leq N$ and any $j = 1, \ldots, k$, we have*

$$\mathbb{E}\,|\mathrm{Tr}\,\mathbf{A}^n\mathbf{M}_j|^2 \leq Cn^2\left(\mathbb{1}_{n\geq 0}(b+\varepsilon)^{2n} + \mathbb{1}_{n\leq 0}(a-\varepsilon)^{2n}\right)$$

Proof. With the notation of Section 2.3.1, let $\mathbb{E}_{\mathbf{V}}$ denote the expectation with respect to the randomness of $\mathbf{V}$. For each $n \in \mathbb{Z}$ and each j, by Lemma 2.5.6, we have

$$\begin{aligned}
\mathbb{E}_{\mathbf{V}}\,|\mathrm{Tr}\,\mathbf{A}^n\mathbf{M}_j|^2 &= \mathbb{E}_{\mathbf{V}}\,\mathrm{Tr}\,\mathbf{A}^n\mathbf{V}\widetilde{\mathbf{M}}_j\mathbf{V}^*\,\mathrm{Tr}(\mathbf{A}^*)^n\mathbf{V}\widetilde{\mathbf{M}}_j^*\mathbf{V}^* \\
&= \frac{1}{N^2-1}\left(\mathrm{Tr}\,\mathbf{A}^n\,\mathrm{Tr}(\mathbf{A}^*)^n\,\mathrm{Tr}\,\widetilde{\mathbf{M}}_j\,\mathrm{Tr}\,\widetilde{\mathbf{M}}_j^* + \mathrm{Tr}\,\mathbf{A}^n(\mathbf{A}^*)^n\,\mathrm{Tr}\,\widetilde{\mathbf{M}}_j\widetilde{\mathbf{M}}_j^*\right) \\
&\quad -\frac{1}{N(N^2-1)}\left(\mathrm{Tr}\,\mathbf{A}^n\,\mathrm{Tr}(\mathbf{A}^*)^n\,\mathrm{Tr}\,\widetilde{\mathbf{M}}_j\widetilde{\mathbf{M}}_j^* + \mathrm{Tr}\,\mathbf{A}^*(\mathbf{A}^*)^n\,\mathrm{Tr}\,\widetilde{\mathbf{M}}_j\,\mathrm{Tr}\,\widetilde{\mathbf{M}}_j^*\right) \\
&\leq \frac{1}{N^2-1}\left(|\mathrm{Tr}\,\mathbf{A}^n|^2\,\mathrm{Tr}\,\widetilde{\mathbf{M}}_j\,\mathrm{Tr}\,\widetilde{\mathbf{M}}_j^* + \mathrm{Tr}\,\mathbf{A}^n(\mathbf{A}^*)^n\,\mathrm{Tr}\,\widetilde{\mathbf{M}}_j\widetilde{\mathbf{M}}_j^*\right) \\
&\leq C\left(|\mathrm{Tr}\,\mathbf{A}^n|^2 + N^{-1}\,\mathrm{Tr}\,\mathbf{A}^n(\mathbf{A}^*)^n\right),
\end{aligned}$$

where C is a constant independent of N. Then the conclusion follows from Lemma 2.5.5. $\square$

Lemma 2.3.2. *There are some constants $C>0$ and $c\in(0,1)$ and a sequence $\mathscr{E}=\mathscr{E}_N$ of events such that*

$$\mathbb{P}(\mathscr{E}) \underset{N\to\infty}{\longrightarrow} 1$$

and such that for all N, all $n_1\geq 0$ and all $j=1,\dots,k$, we have

$$\mathbb{E}\left|\mathbb{1}_{\mathscr{E}}\sum_{|n|>n_1} a_n(f_j)\operatorname{Tr}(\mathbf{A}^n\mathbf{M}_j)\right| \leq CN(1-c)^{n_1}.$$

Proof. By [21, Lem. 3.2], we known that there is a constant C_1 such that the event

$$\mathscr{E}=\mathscr{E}_N := \{\forall n\geq 0,\ \|\mathbf{A}^n\|_{\mathrm{op}}\leq C_1(b+\varepsilon)^n\}\cap\{\forall n\leq 0,\ \|\mathbf{A}^n\|_{\mathrm{op}}\leq C_1(a-\varepsilon)^n\}$$

has probability tending to one.

Then one concludes easily, noting first that by non-commutative Hölder inequalities (see [2, Eq. (A.13)]), we have

$$\mathbb{1}_{\mathscr{E}}|\operatorname{Tr}(\mathbf{A}^n\mathbf{M}_j)|\leq\begin{cases}C_1(b+\varepsilon)^n N\sqrt{N^{-1}\operatorname{Tr}\mathbf{M}_j\mathbf{M}_j^*} & \text{if } n\geq 0\\ C_1(a-\varepsilon)^n N\sqrt{N^{-1}\operatorname{Tr}\mathbf{M}_j\mathbf{M}_j^*} & \text{if } n\leq 0\end{cases}$$

and secondly that there is $c\in(0,1)$ such that for each j, the sequences

$$\left(a_n(f_j)\frac{(b+\varepsilon)^n}{(1-c)^n}\right)_{n\geq 0}\qquad;\qquad\left(a_n(f_j)\frac{(a-\varepsilon)^n}{(1-c)^{-n}}\right)_{n\leq 0}$$

are bounded. $\square$

As a consequence of Lemmas 2.3.1 and 2.3.2, for any $0<n_0<n_1\leq N^{1/6}$ and any $j=1,\dots,k$,

$$\begin{aligned}\mathbb{E}\left|\mathbb{1}_{\mathscr{E}}\sum_{|n|>n_0} a_n(f_j)\operatorname{Tr}(\mathbf{A}^n\mathbf{M}_j)\right| &\leq \sum_{n_0<|n|\leq n_1}|a_n(f_j)|\sqrt{\mathbb{P}(\mathscr{E})}\sqrt{\mathbb{E}\,|\operatorname{Tr}(\mathbf{A}^n\mathbf{M}_j)|^2}\\ &\quad+\mathbb{E}\left|\mathbb{1}_{\mathscr{E}}\sum_{|n|>n_1}a_n(f_j)\operatorname{Tr}(\mathbf{A}^n\mathbf{M}_j)\right|\\ &\leq \sum_{n_0<|n|\leq n_1}C|a_n(f_j)|n^2\left(\mathbb{1}_{n\geq 0}(b+\varepsilon)^{2n}+\mathbb{1}_{n\leq 0}(a-\varepsilon)^{2n}\right)\\ &\quad+CN(1-c)^{n_1}\end{aligned}$$

Choosing first $n_1=\lfloor A\log N\rfloor$ for A a large enough constant and then using the fact that for any $j=1,\dots,k$,

$$\sum_{n\in\mathbb{Z}}|a_n(f_j)|n^2\left(\mathbb{1}_{n\geq 0}(b+\varepsilon)^{2n}+\mathbb{1}_{n\leq 0}(a-\varepsilon)^{2n}\right)<\infty,$$

we deduce that for any $\delta > 0$, there is $n_0 > 0$ fixed such that for all N large enough,

$$\sum_{j=1}^{k} \mathbb{E}\left| \mathbb{1}_{\mathscr{E}} \sum_{|n|>n_0} a_n(f_j) \operatorname{Tr}(\mathbf{A}^n \mathbf{M}_j) \right| \leq \delta,$$

for $\mathscr{E} = \mathscr{E}_N$ as in Lemma 2.3.2.

Let us now suppose Theorem 2.2.4 to be proved in the particular case where there is n_0 such that for all n, we have

$$|n| > n_0 \implies \forall j = 1, \ldots, k,\ a_n(f_j) = 0$$

and let us prove it in the general case. Let X_N denote the random vector of (2.2). We want to prove that as $N \to \infty$, the distribution of X_N tends to the one of $\mathscr{G} := (\mathscr{G}(f_1), \ldots, \mathscr{G}(f_k))$, *i.e.* that for any function $F : \mathbb{C}^k \to \mathbb{C}$ which is 1 Lipschitz and bounded by 1, we have

$$\mathbb{E} F(X_N) \underset{N\to\infty}{\longrightarrow} \mathbb{E} F(\mathscr{G}).$$

To do so, we first set

$$X_{N,n_0} \ := \ \left(\sum_{|n|<n_0} a_n(f_j) \operatorname{Tr}\left(\mathbf{A}^n \mathbf{M}_j\right) - \operatorname{Tr}(\mathbf{M}_j) a_0(f_j) \right)_{j=1}^{k}$$

By hypothesis, for any fixed n_0, X_{N,n_0} converges in distribution to a centered complex Gaussian vector $\mathscr{G}_{n_0} := (\mathscr{G}_{n_0}(f_1), \ldots, \mathscr{G}_{n_0}(f_k))$ such that

$$\begin{aligned}
\mathbb{E}\left[\mathscr{G}_{n_0}(f_i)\mathscr{G}_{n_0}(f_j)\right] &= \mathbb{E}\left[\mathscr{G}(f_i)\mathscr{G}(f_j)\right] + \eta_{ij}^{n_0} \\
\mathbb{E}\left[\mathscr{G}_{n_0}(f_i)\overline{\mathscr{G}_{n_0}(f_j)}\right] &= \mathbb{E}\left[\mathscr{G}(f_i)\overline{\mathscr{G}(f_j)}\right] + \delta_{ij}^{n_0},
\end{aligned}$$

where $\lim_{n_0\to\infty} \sum_{1\leq i,j\leq k} |\eta_{ij}^{n_0}| + |\delta_{ij}^{n_0}| = 0$. Therefore,

$$\begin{aligned}
& |\mathbb{E}\left[F(X_N) - F(\mathscr{G})\right]| \\
\leq\ & |\mathbb{E}\left[F(X_N) - F(X_{N,n_0})\right]| + |\mathbb{E}\left[F(X_{N,n_0}) - F(\mathscr{G}_{n_0})\right]| + |\mathbb{E}\left[F(\mathscr{G}_{n_0}) - F(\mathscr{G})\right]| \\
\leq\ & 2\mathbb{P}(\mathscr{E}^c) + \mathbb{E}\left[\mathbb{1}_{\mathscr{E}} \|X_N - X_{N,n_0}\|_2^2\right] + |\mathbb{E}\left[F(X_{N,n_0}) - F(\mathscr{G}_{n_0})\right]| + \mathbb{E}\left[\|\mathscr{G}_{n_0} - \mathscr{G}\|_2^2\right]
\end{aligned}$$

which can be as small as we want by (2.3) and the fact that $X_{N,n_0} \xrightarrow{(\mathrm{d})} \mathscr{G}_{n_0}$ if $\mathscr{G}_{n_0}$ and $\mathscr{G}$ are coupled the right way.

2.3.3 Proof of Theorem 2.2.4 when the f_j's are polynomial in z and z^{-1}

We suppose here that there is $n_0 > 0$ such that for all $n > n_0$ and all $1 \leq j \leq k$, $a_n(f_j) = 0$. Without any loss of generality, we also assume that for all j, $a_0(f_j) = 0$. In this case, any linear combination of the $\operatorname{Tr} f_j(\mathbf{A})\mathbf{M}_j$'s can be written

$$G_N \ := \ \sum_{j=1}^{k} \nu_j \operatorname{Tr} f_j(\mathbf{A}) \mathbf{M}_j \ = \ \sum_{|n|\leq n_0} \operatorname{Tr} \mathbf{A}^n \mathbf{N}_n$$

where $\mathbf{N}_n := \sum_{j=1}^{k} v_j a_n(f_j)\mathbf{M}_j$. Written this way, we notice that to prove that the limit distribution of G_N is Gaussian, we simply have to prove that the random vector

$$(\operatorname{Tr} \mathbf{A}^n \mathbf{N}_n)_{-n_0 \le n \le n_0}$$

converges in distribution to a Gaussian vector. We will prove it by computing the limit of the joint moments.

Before going any further, recall that by the preliminary randomization of the $\mathbf{N}_j$'s from section 2.3.1, we suppose that $\mathbf{A} = \mathbf{UT}$ (instead of $\mathbf{A} = \mathbf{UTV}$) and that there is a Haar-distributed random unitary matrix $\mathbf{V}$, independent of $\mathbf{U}$, such that for each j, $\mathbf{N}_j = \mathbf{V}\widetilde{\mathbf{N}}_j\mathbf{V}^*$, with $\widetilde{\mathbf{N}}_j$ a deterministic matrix.

We shall proceed in three steps:

a) First, we prove the asymptotic independence of the random vectors

$$\left(\operatorname{Tr} \mathbf{A}^n \mathbf{N}_n, \operatorname{Tr} \mathbf{A}^{-n} \mathbf{N}_{-n}\right)_{n \ge 1}$$

by the factorization of the joint moments. More precisely, we prove, thanks to Corollary 2.5.4, that for any $(p_n)_{n=1}^{n_0}$, $(q_n)_{n=1}^{n_0}$, $(r_n)_{n=1}^{n_0}$, $(s_n)_{n=1}^{n_0}$,

$$\begin{aligned}
&\mathbb{E}\Big[\prod_{1 \le n \le n_0} (\operatorname{Tr} \mathbf{A}^n \mathbf{N}_n)^{p_n} \overline{(\operatorname{Tr} \mathbf{A}^n \mathbf{N}_n)^{q_n}} (\operatorname{Tr} \mathbf{A}^{-n} \mathbf{N}_{-n})^{r_n} \overline{(\operatorname{Tr} \mathbf{A}^{-n} \mathbf{N}_{-n})^{s_n}}\Big] \\
= &\prod_{1 \le n \le n_0} \mathbb{E}\Big[(\operatorname{Tr} \mathbf{A}^n \mathbf{N}_n)^{p_n} \overline{(\operatorname{Tr} \mathbf{A}^n \mathbf{N}_n)^{q_n}} (\operatorname{Tr} \mathbf{A}^{-n} \mathbf{N}_{-n})^{r_n} \overline{(\operatorname{Tr} \mathbf{A}^{-n} \mathbf{N}_{-n})^{s_n}}\Big] + o(1)
\end{aligned}$$

b) Then, we prove for any fixed n, the random complex vector

$$(\operatorname{Tr} \mathbf{A}^n \mathbf{N}_n, \operatorname{Tr} \mathbf{A}^{-n} \mathbf{N}_{-n})$$

converges in distribution to a centered complex Gaussian vector thanks to the criterion provided by the Lemma 2.5.7. This criterion consists in proving that the joint moments, at the large N limit, satisfy the same induction relation as the moments of a Gaussian distribution.

c) It will follow from **a)** and **b)** that when all f_j's are polynomials in z and z^{-1}, the random vector of (2.2) converges in distribution to a centered Gaussian vector. To conclude the proof, the last step will be to prove that the limit covariance is the one given in Theorem 2.2.4.

In the proofs of **a)** and **b)**, we shall need to compute expectations with respect to the randomness of the Haar-distributed matrix $\mathbf{U}$. More precisely, we shall need to compute sums of expectations with respect of $\mathbf{U}$ resulting from the expansion of products of traces involving powers of $\mathbf{A}$ (such as $\operatorname{Tr} \mathbf{A}^n \mathbf{N}_n$). To do so, we will use the Weingarten calculus (see Proposition 2.5.1) and shall always proceed in the following way: first, we use (2.22) to state that all the terms of the sum are null except those for which the left (resp. right) indices involved in u are obtained by permuting the left (resp. right) ones involved in $\overline{u}$. Then, we claim, by Remark 2.5.2, that among the remaining terms, we can neglect all those whose indices are not pairwise distinct. At last, once all the remaining terms are, up to multiplicative constant, equal to $\mathrm{Wg}(\sigma)$ for some permutation σ, we neglect all those for which $\sigma \neq id$ (see Remark 2.5.3) and the summation finally gets easy to compute. We introduce here a notation that we shall use several times :

$$\mathrm{I}_n^{\neq} := \left\{(i_1,\dots,i_n) \in \{1,\dots,N\}^n \; ; \; i_1,\dots,i_n \text{ are pairwise distinct}\right\} \tag{2.3}$$

(this set implicitly depends on N).

2.3.4 Proof of b)

In this part, as n is fixed, we shall denote $\mathbf{N}_n$ (resp. $\mathbf{N}_{-n}$) by $\mathbf{M} = [M_{ij}]$ (resp. $\mathbf{K} = [K_{ij}]$). For any non-negative integers p,q,r,s, wet set

$$m_{p,q,r,s} \quad := \quad \mathbb{E}\left(\mathrm{Tr}\,\mathbf{A}^n\mathbf{M}\right)^p \overline{\left(\mathrm{Tr}\,\mathbf{A}^n\mathbf{M}\right)}^q \left(\mathrm{Tr}\,\mathbf{A}^{-n}\mathbf{K}\right)^r \overline{\left(\mathrm{Tr}\,\mathbf{A}^{-n}\mathbf{K}\right)}^s$$

and our goal is to show that, as N goes to infinity, the numbers $m_{p,q,r,s}$ have limits satisfying conditions (2.26), (2.27), (2.29), (2.30) and (2.31) of the Lemma 2.5.7. Note that (2.26) and (2.29) follow from the fact the the Haar measure on the unitary group is invariant by multiplication by any $\mathrm{e}^{\mathrm{i}\theta}$, $\theta \in \mathbb{R}$. We shall use the following notations

$$\lim_{N\to\infty} \frac{1}{N}\mathrm{Tr}\left(\mathbf{M}\mathbf{M}^*\right) =: \alpha_{\mathbf{M}} \quad ; \quad \lim_{N\to\infty} \frac{1}{N}\mathrm{Tr}\left(\mathbf{K}\mathbf{K}^*\right) =: \alpha_{\mathbf{K}} \quad ; \quad \lim_{N\to\infty} \frac{1}{N}\mathrm{Tr}\left(\mathbf{M}\mathbf{K}\right) =: \beta_{\mathbf{M},\mathbf{K}}$$
$$\lim_{N\to\infty} \frac{1}{N}\mathrm{Tr}\,\mathbf{M} =: \tau_{\mathbf{M}} \quad ; \quad \lim_{N\to\infty} \frac{1}{N}\mathrm{Tr}\,\mathbf{K} =: \tau_{\mathbf{K}}.$$

$\mathrm{Tr}\,\mathbf{A}^n\mathbf{M}$ and $\mathrm{Tr}\,\mathbf{A}^{-n}\mathbf{K}$ are asymptotically two circular Gaussian complex variables satisfying conditions (2.27) and (2.31)

We simply have to show that for any integer $p \geq 1$

$$\mathbb{E}\left|\mathrm{Tr}\,\mathbf{A}^n\mathbf{M}\right|^{2p} = p!\left(b^{2n}((n-1)|\tau_{\mathbf{M}}|^2 + \alpha_{\mathbf{M}})\right)^p + o(1)\,, \tag{2.4}$$
$$\mathbb{E}\left|\mathrm{Tr}\,\mathbf{A}^{-n}\mathbf{K}\right|^{2p} = p!\left(a^{-2n}((n-1)|\tau_{\mathbf{K}}|^2 + \alpha_{\mathbf{K}})\right)^p + o(1)\,, \tag{2.5}$$
$$\mathbb{E}\left(\mathrm{Tr}\,\mathbf{A}^n\mathbf{M}\;\mathrm{Tr}\,\mathbf{A}^{-n}\mathbf{K}\right)^p = p!\left((n-1)\tau_{\mathbf{M}}\tau_{\mathbf{K}} + \beta_{\mathbf{M},\mathbf{K}}\right)^p + o(1)\,. \tag{2.6}$$

We shall prove it by induction on p. So first, we show the previous relation for $p = 1$. Recall that we assume that $\mathbf{M} = \mathbf{V}\widetilde{\mathbf{M}}\mathbf{V}^*$ and $\mathbf{K} = \mathbf{V}\widetilde{\mathbf{K}}\mathbf{V}^*$, where $\widetilde{\mathbf{M}}$ and $\widetilde{\mathbf{K}}$ are deterministic, so that, using the Lemma 2.5.6, we have (denoting again by $\mathbb{E}_{\mathbf{V}}$

the expectation with respect to the randomness of $\mathbf{V}$),

$$
\begin{aligned}
\mathbb{E}_{\mathbf{V}}\left|\operatorname{Tr}\mathbf{A}^n\mathbf{V}\widetilde{\mathbf{M}}\mathbf{V}^*\right|^2 &= \frac{1}{N}\operatorname{Tr}\mathbf{A}^n(\mathbf{A}^*)^n\left(\frac{1}{N}\operatorname{Tr}\widetilde{\mathbf{M}}\widetilde{\mathbf{M}}^* - \left|\frac{1}{N}\operatorname{Tr}\widetilde{\mathbf{M}}\right|^2\right) \\
&+ \left|\operatorname{Tr}\mathbf{A}^n\right|^2\left|\frac{1}{N}\operatorname{Tr}\widetilde{\mathbf{M}}\right|^2 + O\left(\frac{1}{N}\right) \\
\mathbb{E}_{\mathbf{V}}\left|\operatorname{Tr}\mathbf{A}^{-n}\mathbf{V}\widetilde{\mathbf{K}}\mathbf{V}^*\right|^2 &= \frac{1}{N}\operatorname{Tr}\mathbf{A}^{-n}(\mathbf{A}^*)^{-n}\left(\frac{1}{N}\operatorname{Tr}\left(\widetilde{\mathbf{K}}\widetilde{\mathbf{K}}^*\right) - \left|\frac{1}{N}\operatorname{Tr}\widetilde{\mathbf{K}}\right|^2\right) \\
&+ \left|\operatorname{Tr}\mathbf{A}^{-n}\right|^2\left|\frac{1}{N}\operatorname{Tr}\widetilde{\mathbf{K}}\right|^2 + O\left(\frac{1}{N}\right) \\
\mathbb{E}_{\mathbf{V}}\operatorname{Tr}\mathbf{A}^n\mathbf{V}\widetilde{\mathbf{M}}\mathbf{V}\operatorname{Tr}\mathbf{A}^{-n}\mathbf{V}\widetilde{\mathbf{K}}\mathbf{V}^* &= \frac{1}{N}\operatorname{Tr}\widetilde{\mathbf{M}}\widetilde{\mathbf{K}} - \frac{1}{N}\operatorname{Tr}\widetilde{\mathbf{M}}\frac{1}{N}\operatorname{Tr}\widetilde{\mathbf{K}} \\
&+ \operatorname{Tr}\mathbf{A}^n\operatorname{Tr}\mathbf{A}^{-n}\frac{1}{N}\operatorname{Tr}\widetilde{\mathbf{M}}\frac{1}{N}\operatorname{Tr}\widetilde{\mathbf{K}} + O\left(\frac{1}{N}\right).
\end{aligned}
$$

This is asymptotically determined by the limits of $\mathbb{E}\left|\operatorname{Tr}\mathbf{A}^n\right|^2$, $\mathbb{E}\left|\operatorname{Tr}\mathbf{A}^{-n}\right|^2$, $\mathbb{E}\operatorname{Tr}\mathbf{A}^n\operatorname{Tr}\mathbf{A}^{-n}$, $N^{-1}\mathbb{E}\operatorname{Tr}\mathbf{A}^n(\mathbf{A}^*)^n$, $N^{-1}\mathbb{E}\operatorname{Tr}\mathbf{A}^{-n}(\mathbf{A}^*)^{-n}$. First, we compute $\mathbb{E}\left|\operatorname{Tr}\mathbf{A}^n\right|^2$ for $n \geq 1$. We write

$$
\mathbb{E}\left|\operatorname{Tr}\mathbf{A}^n\right|^2 = \sum_{\substack{1\leq i_1,\dots,i_n\leq N \\ 1\leq j_1,\dots,j_n\leq N}} \mathbb{E}\left[u_{i_1i_2}\cdots u_{i_ni_1}\overline{u_{j_1j_2}}\cdots\overline{u_{j_nj_1}}\right]s_{i_1}s_{j_1}\cdots s_{i_n}s_{j_n}. \tag{2.7}
$$

From (2.22), we have a condition on the i_k's and the j_k's for a non-vanishing expectation, which is the multiset[9] equality

$$
\{i_1,\dots,i_n\}_m = \{j_1,\dots,j_n\}_m, \tag{2.8}
$$

The first consequence of (2.8) is that the sum is in fact over $O(N^n)$ terms which all are at most $O(N^{-n})$, which means that any sub-summation over $o(N^n)$ terms might be neglected. So from now on, we shall only sum over the n-tuples $(i_1,\dots,i_n) \in \mathrm{I}_n^{\neq}$ (recall notation (2.3)). Then (2.7) becomes

$$
\mathbb{E}\left|\operatorname{Tr}\mathbf{A}^n\right|^2 = \sum_{(i_1,\dots,i_n)\in\mathrm{I}_n^{\neq}} s_{i_1}^2\cdots s_{i_n}^2 \sum_{\sigma\in S_n}\mathbb{E}\left[u_{i_1i_2}\cdots u_{i_ni_1}\overline{u_{i_{\sigma(1)}i_{\sigma(2)}}}\cdots\overline{u_{i_{\sigma(n)}i_{\sigma(1)}}}\right] + o(1)
$$

Let $c \in S_n$ be the cycle $(1\ 2\ \cdots n)$. From (2.22) (see Remark 2.5.2), as long as the i_k's are pairwise distinct, one can write

$$
\mathbb{E}\left[u_{i_1i_2}\cdots u_{i_ni_1}\overline{u_{i_{\sigma(1)}i_{\sigma(2)}}}\cdots\overline{u_{i_{\sigma(n)}i_{\sigma(1)}}}\right] = \operatorname{Wg}\left(\sigma c^{-1}\sigma^{-1}c\right)
$$

and from (2.23) and Remark 2.5.3, we know that the non-negligible terms are the ones such that $\sigma c^{-1}\sigma^{-1}c = id$, *i.e.* $\sigma c = c\sigma$, which means that σ must be a

[9] We use the index m in $\{\cdot\}_m$ to denote a *multiset*, which means that $\{x_1,\dots,x_n\}_m$ is the class of the n-tuple $(x_1,\dots,x_n)$ under the action of the symmetric group S_n.

power of c and so, only n permutations σ contribute to the non negligible terms. At last, as $\mathrm{Moeb}(id)=1$, we have

$$\begin{aligned}\mathbb{E}\big|\operatorname{Tr}\mathbf{A}^n\big|^2 &= \sum_{(i_1,\ldots,i_n)\in \mathrm{I}_n^{\neq}} s_{i_1}^2\cdots s_{i_n}^2 \times nN^{-n}\big(1+o(1)\big)+o(1)\\ &= n\left(\frac{1}{N}\sum_{i=1}^{N}s_i^2\right)^n+o(1) \;=\; nb^{2n}+o(1)\,.\end{aligned}$$

In the same way, one can get

$$\begin{aligned}\mathbb{E}\big|\operatorname{Tr}\mathbf{A}^{-n}\big|^2 &= na^{-2n}+o(1)\,,\\ \mathbb{E}\operatorname{Tr}\mathbf{A}^n\operatorname{Tr}\mathbf{A}^{-n} &= n+o(1)\,.\end{aligned}$$

Let us now consider $N^{-1}\,\mathbb{E}\operatorname{Tr}\mathbf{A}^n(\mathbf{A}^*)^n$ for $n\geq 1$. We have

$$\frac{1}{N}\mathbb{E}\operatorname{Tr}\mathbf{A}^n(\mathbf{A}^*)^n = N^{-1}\sum_{\substack{1\leq i_0,i_1,\ldots,i_n\leq N\\ 1\leq j_0,j_1,\ldots,j_n\leq N\\ i_0=j_0,\ i_n=j_n}} \mathbb{E}\big[u_{i_0i_1}\cdots u_{i_{n-1}i_n}\overline{u_{j_0j_1}}\cdots\overline{u_{j_{n-1}j_n}}\big]\,s_{i_1}s_{j_1}\cdots s_{i_n}s_{j_n} \tag{2.9}$$

As previously, we know that by (2.22), that for the expectation to be non zero, we must have the multiset equality

$$\{i_0,\ldots,i_n\}_m = \{j_0,\ldots,j_n\}_m\,, \tag{2.10}$$

The first consequence of (2.10) is that the sum is in fact over $O\big(N^{n+1}\big)$ terms which are all at most $O\big(N^{-n-1}\big)$, so that any sub-summation over $o\big(N^{n+1}\big)$ terms might be neglected. As previously, we shall sum over the pairwise distinct indices $\mathrm{I}_{n+1}^{\neq}$ (see notation (2.3)). Hence (2.9) becomes

$$N^{-1}\,\mathbb{E}\operatorname{Tr}\mathbf{A}^n(\mathbf{A}^*)^n = N^{-1}\sum_{(i_0,i_1,\ldots,i_n)\in\mathrm{I}_{n+1}^{\neq}} s_{i_1}^2\cdots s_{i_n}^2\sum_{\substack{\sigma\in S_{n+1}\\ \sigma(0)=0\\ \sigma(n)=n}}\mathbb{E}\Big[u_{i_0i_1}\cdots u_{i_{n-1}i_n}\overline{u_{i_{\sigma(0)}i_{\sigma(1)}}}\cdots\overline{u_{i_{\sigma(n-1)}i_{\sigma(n)}}}\Big]$$

Let $c\in S_{n+1}$ be the cycle $(0\;1\;2\;\cdots n)$. From (2.22) (see Remark 2.5.2) one can write

$$\mathbb{E}\Big[u_{i_0i_1}\cdots u_{i_{n-1}i_n}\overline{u_{i_{\sigma(0)}i_{\sigma(1)}}}\cdots\overline{u_{i_{\sigma(n-1)}i_{\sigma(n)}}}\Big] = \mathrm{Wg}\big(\sigma c^{-1}\sigma^{-1}c\big).$$

As previously, σ must be a power of c for the asymptotic contribution to be non-negligible. However, this time, we impose $\sigma(0)=0$ and $\sigma(n)=n$, so that the only possible choice is $\sigma=id$ which means that only one term contributes this time. At last,

$$\frac{1}{N}\mathbb{E}\operatorname{Tr}\mathbf{A}^n(\mathbf{A}^*)^n = b^{2n}+o(1)\,,$$

The same way, one can get

$$\frac{1}{N}\mathbb{E}\operatorname{Tr}\mathbf{A}^{-n}(\mathbf{A}^*)^{-n} = a^{-2n}+o(1)\,.$$

This concludes the first step of the induction.

In the second step, we have to prove the following induction relation: for any $p \geq 2$,

$$\mathbb{E}\left|\operatorname{Tr}\mathbf{A}^n\mathbf{M}\right|^{2p} = p\,\mathbb{E}\left|\operatorname{Tr}\mathbf{A}^n\mathbf{M}\right|^2 \mathbb{E}\left|\operatorname{Tr}\mathbf{A}^n\mathbf{M}\right|^{2(p-1)} + o(1) \tag{2.11}$$

$$\mathbb{E}\left|\operatorname{Tr}\mathbf{A}^{-n}\mathbf{K}\right|^{2p} = p\,\mathbb{E}\left|\operatorname{Tr}\mathbf{A}^{-n}\mathbf{K}\right|^2 \mathbb{E}\left|\operatorname{Tr}\mathbf{A}^{-n}\mathbf{K}\right|^{2(p-1)} + o(1) \tag{2.12}$$

$$\mathbb{E}(\operatorname{Tr}\mathbf{A}^n\mathbf{M}\operatorname{Tr}\mathbf{A}^{-n}\mathbf{K})^p = p\,\mathbb{E}\left[\operatorname{Tr}\mathbf{A}^n\mathbf{M}\ \operatorname{Tr}\mathbf{A}^{-n}\mathbf{K}\right]\mathbb{E}\left[\left(\operatorname{Tr}\mathbf{A}^n\mathbf{M}\ \operatorname{Tr}\mathbf{A}^{-n}\mathbf{K}\right)^{p-1}\right] + o(1) \tag{2.13}$$

Let us first consider $\mathbb{E}\left|\operatorname{Tr}\mathbf{A}^n\mathbf{M}\right|^{2p}$. We shall use the following notation

$$\operatorname{Tr}\mathbf{A}^n\mathbf{M} = \sum_{i_0,i_1,\dots,i_n} u_{i_0i_1}s_{i_1}\cdots u_{i_{n-1}i_n}s_{i_n}M_{i_ni_0} =: \sum_{\mathbf{i}} u_{\mathbf{i}}s_{\mathbf{i}}M_{i_ni_0},$$

where the bold letter $\mathbf{i}$ denotes the $(n+1)$-tuple $(i_0,\dots,i_n)$ and where we set

$$u_{\mathbf{i}} := u_{i_0i_1}\cdots u_{i_{n-1}i_n} \qquad ; \qquad s_{\mathbf{i}} := s_{i_1}\cdots s_{i_n}.$$

Hence,

$$\mathbb{E}\left|\operatorname{Tr}\mathbf{A}^n\mathbf{M}\right|^{2p} = \sum_{\substack{\mathbf{i}^1,\dots,\mathbf{i}^p\\ \mathbf{j}^1,\dots,\mathbf{j}^p}} \mathbb{E}\left[u_{\mathbf{i}^1}\cdots u_{\mathbf{i}^p}\overline{u_{\mathbf{j}^1}}\cdots\overline{u_{\mathbf{j}^p}}\right] s_{\mathbf{i}^1}M_{i_n^1i_0^1}s_{\mathbf{j}^1}\overline{M_{j_n^1j_0^1}}\cdots s_{\mathbf{i}^p}M_{i_n^pi_0^p}s_{\mathbf{j}^p}\overline{M_{j_n^pj_0^p}} \tag{2.14}$$

As usual, we know we can sum over the $\mathbf{i}^k$'s satisfying that $(\mathbf{i}^1,\dots,\mathbf{i}^p)$ (the $p(n+1)$-tuple obtained by concatenation of the $\mathbf{i}$'s) has pairwise distinct entries and such that we have the set equality:

$$\left\{i_\mu^\lambda,\ 1\leq\lambda\leq p,\ 0\leq\mu\leq n\right\} = \left\{j_\mu^\lambda,\ 1\leq\lambda\leq p,\ 0\leq\mu\leq n\right\}. \tag{2.15}$$

Then, in order to have $\mathrm{Wg}(id)$, we must match each of the $(n+1)$-tuples $\mathbf{i}^1,\dots,\mathbf{i}^p$ with one of the $(n+1)$-tuples $\mathbf{j}^1,\dots,\mathbf{j}^p$, *i.e.* that for all $1\leq\lambda\leq p$, there is a $1\leq\lambda'\leq p$ such that we have the set equality

$$\{\mathbf{i}^\lambda\} =: \left\{i_0^\lambda, i_1^\lambda,\dots,i_n^\lambda\right\} = \left\{j_0^{\lambda'}, j_1^{\lambda'},\dots,j_n^{\lambda'}\right\} := \{\mathbf{j}^{\lambda'}\}.$$

We rewrite (2.14) by summing according the possible choice to match $\{\mathbf{i}^1\} = \{i_0^1, i_1^1,\dots,i_n^1\}$

$$\mathbb{E}\left|\operatorname{Tr}\mathbf{A}^n\mathbf{M}\right|^{2p} = \sum_{\lambda=1}^{p}\sum_{\substack{(\mathbf{i}^1,\dots,\mathbf{i}^p)\in I^{\neq}_{p(n+1)}\\ (\mathbf{j}^1,\dots,\mathbf{j}^p)\in I^{\neq}_{p(n+1)}\\ \mathbf{i}^1\leftrightarrow\mathbf{j}^\lambda}} \mathbb{E}\left[u_{\mathbf{i}^1}\cdots u_{\mathbf{i}^p}\overline{u_{\mathbf{j}^1}}\cdots\overline{u_{\mathbf{j}^p}}\right] s_{\mathbf{i}^1}M_{i_n^1i_0^1}s_{\mathbf{j}^1}\overline{M_{j_n^1j_0^1}}\cdots s_{\mathbf{i}^p}M_{i_n^pi_0^p}s_{\mathbf{j}^p}\overline{M_{j_n^pj_0^p}} + o(1),$$

where $\mathbf{i}^1 \leftrightarrow \mathbf{j}^\lambda$ stands for the set equality $\{i_0^1, i_1^1, \ldots, i_n^1\} = \left\{j_0^\lambda, j_1^\lambda, \ldots, j_n^\lambda\right\}$. Then, we know that the set of indices $\{i_0^1, i_1^1, \ldots, i_n^1\}$ is disjoint from the others, so that by Corollary 2.5.4,

$$\mathbb{E}\left[u_{\mathbf{i}^1}\cdots u_{\mathbf{i}^p}\overline{u_{\mathbf{j}^1}}\cdots\overline{u_{\mathbf{j}^p}}\right] = \mathbb{E}\left[u_{\mathbf{i}^1}\overline{u_{\mathbf{j}^\lambda}}\right]\mathbb{E}\left[u_{\mathbf{i}^2}\cdots u_{\mathbf{i}^p}\overline{u_{\mathbf{j}^1}}\cdots\overline{u_{\mathbf{j}^{\lambda-1}}u_{\mathbf{j}^{\lambda+1}}}\cdots\overline{u_{\mathbf{j}^p}}\right]$$

and up to a proper relabeling of the indices, all the choices lead to the same value of the expectation, so that

$$\begin{aligned}
&\mathbb{E}\left|\operatorname{Tr}\mathbf{A}^n\mathbf{M}\right|^{2p} \\
= \; & p\sum_{\substack{\mathbf{i}^1\in\mathrm{I}^{\neq}_{n+1}\\ \mathbf{j}^1\in\mathrm{I}^{\neq}_{n+1}}}\mathbb{E}\left[u_{\mathbf{i}^1}\overline{u_{\mathbf{j}^1}}\right]s_{\mathbf{i}^1}M_{i_n^1,i_0^1}s_{\mathbf{j}^1}\overline{M_{j_n^1,j_0^1}}\sum_{\substack{(\mathbf{i}^2,\ldots,\mathbf{i}^p)\in\mathrm{I}^{\neq}_{(p-1)(n+1)}\\ (\mathbf{j}^2,\ldots,\mathbf{j}^p)\in\mathrm{I}^{\neq}_{(p-1)(n+1)}}}\mathbb{E}\left[u_{\mathbf{i}^2}\cdots\overline{u_{\mathbf{j}^p}}\right]s_{\mathbf{i}^2}M_{i_n^2i_0^2}s_{\mathbf{j}^2}\overline{M_{j_n^2j_0^2}}\cdots s_{\mathbf{i}^p}M_{i_n^pi_0^p}s_{\mathbf{j}^p}\overline{M_{j_n^pj_0^p}}+o(1) \\
= \; & p\,\mathbb{E}\left[\left|\operatorname{Tr}(\mathbf{A}^n\mathbf{M})\right|^2\right]\mathbb{E}\left[\left|\operatorname{Tr}(\mathbf{A}^n\mathbf{M})\right|^{2(p-1)}\right]+o(1)\,.
\end{aligned}$$

This proves (2.11). In the same way, we prove (2.12) and (2.13), and thus conclude the proof of the induction.

Remark 2.3.3. In the last computation, we split the expectation and so we separated the summation implying that

$$\mathrm{I}^{\neq}_{p(n+1)} = \mathrm{I}^{\neq}_{n+1}\times\mathrm{I}^{\neq}_{(p-1)(n+1)}$$

which is obviously inaccurate. Nevertheless, we easily have that

$$\operatorname{Card}\mathrm{I}^{\neq}_{p(n+1)} = \operatorname{Card}\left(\mathrm{I}^{\neq}_{n+1}\times\mathrm{I}^{\neq}_{(p-1)(n+1)}\right)\left(1+o(1)\right),$$

which means that this inaccuracy is actually contained in the $o(1)$.

To conclude the proof of **b)**, we have to prove that $\operatorname{Tr}\mathbf{A}^n\mathbf{M}$ and $\operatorname{Tr}\mathbf{A}^{-n}\mathbf{K}$ satisfy Condition (2.30) at the large N limit.

$\operatorname{Tr}\mathbf{A}^n\mathbf{M}$ and $\operatorname{Tr}\mathbf{A}^{-n}\mathbf{K}$ satisfy Condition (2.30) at the large N limit

We apply the same idea as previously, but for a slightly more complicated expectation. Let p,q,r,s be positive integers and such that $p-q=r-s$. We denote joint moments by $m_{p,q,r,s}$:

$$m_{p,q,r,s} := \mathbb{E}\left(\operatorname{Tr}\mathbf{A}^n\mathbf{M}\right)^p\overline{\left(\operatorname{Tr}\mathbf{A}^n\mathbf{M}\right)}^q\left(\operatorname{Tr}\mathbf{A}^{-n}\mathbf{K}\right)^r\overline{\left(\operatorname{Tr}\mathbf{A}^{-n}\mathbf{K}\right)}^s, \tag{2.16}$$

and as

$$\begin{aligned}
\operatorname{Tr}\mathbf{A}^n\mathbf{M} &= \sum_{i_0,i_1,\ldots,i_n}u_{i_0i_1}s_{i_1}\cdots u_{i_{n-1}i_n}s_{i_n}M_{i_ni_0} = \sum_{\mathbf{i}}u_{\mathbf{i}}s_{\mathbf{i}}M_{i_ni_0} \\
\operatorname{Tr}\mathbf{A}^{-n}\mathbf{K} &= \sum_{i_n,i_{n-1},\ldots,i_0}\overline{u}_{i_{n-1}i_n}s_{i_n}^{-1}\cdots\overline{u}_{i_0i_1}s_{i_1}K_{i_0i_n} = \sum_{\mathbf{i}}\overline{u}_{\mathbf{i}}s_{\mathbf{i}}^{-1}K_{i_0i_n},
\end{aligned}$$

we rewrite (2.16) as follow

$$\mathbb{E}\sum_{\substack{\mathbf{i}^1,\ldots,\mathbf{i}^p\\ \mathbf{j}^1,\ldots,\mathbf{j}^q\\ \mathbf{k}^1,\ldots,\mathbf{k}^r\\ \mathbf{l}^1,\ldots,\mathbf{l}^s}}\prod_{\substack{1\le\lambda\le p\\ 1\le\mu\le q\\ 1\le\nu\le r\\ 1\le\theta\le s}}\frac{s_{\mathbf{i}^\lambda}s_{\mathbf{j}^\mu}}{s_{\mathbf{k}^\nu}s_{\mathbf{l}^\theta}}M_{i_n^\lambda,i_0^\lambda}\overline{M}_{j_n^\mu,j_0^\mu}K_{k_0^\nu,k_n^\mu}\overline{K}_{\ell_0^\theta,\ell_n^\theta}u_{\mathbf{i}^\lambda}u_{\mathbf{l}^\theta}\overline{u_{\mathbf{j}^\mu}u_{\mathbf{k}^\nu}} \tag{2.17}$$

(recall that the $s_{\mathbf{i}} = s_{i_1}\cdots s_{i_n}$ for $\mathbf{i} = (i_0,\ldots,i_n)$). As previously, we deduce from Proposition 2.5.1 that for the non vanishing expectations, we must have the following multiset equality

$$\begin{aligned} &\left\{i_\mu^\lambda,\ 1\le\lambda\le p, 0\le\mu\le n\right\}_m\cup\left\{\ell_\mu^\lambda,\ 1\le\lambda\le s, 0\le\mu\le n\right\}_m\\ = &\left\{j_\mu^\lambda,\ 1\le\lambda\le q, 0\le\mu\le n\right\}_m\cup\left\{k_\mu^\lambda,\ 1\le\lambda\le r, 0\le\mu\le n\right\}_m, \end{aligned}\tag{2.18}$$

from which we deduce that we can restrict the summation to the tuples such that

$$(\mathbf{i}^1,\ldots,\mathbf{i}^p,\mathbf{l}^1,\ldots\mathbf{l}^s)\in\mathrm{I}^{\neq}_{(p+s)(n+1)}$$

and that, for the non negligible terms (*i.e.* those which lead to $\mathrm{Wg}(id)$), we must match each of the $(n+1)$-tuples involved in u (the $\mathbf{i}$'s and the $\mathbf{l}$'s) with one of those involved in $\overline{u}$ (the $\mathbf{j}$'s and the $\mathbf{k}$'s). For example, we sum according to the choice the "partner" of $\mathbf{i}^1$.

$$\begin{aligned} m_{p,q,r,s} &= \sum_{\alpha=1}^{q}\mathbb{E}\sum_{\substack{(\mathbf{i},\mathbf{l})\in\mathrm{I}^{\neq}\\ (\mathbf{j},\mathbf{k})\in\mathrm{I}^{\neq}\\ \mathbf{i}^1\leftrightarrow\mathbf{j}^\alpha}}\prod_{\substack{1\le\lambda\le p\\ 1\le\mu\le q\\ 1\le\nu\le r\\ 1\le\theta\le s}}\frac{s_{\mathbf{i}^\lambda}s_{\mathbf{j}^\mu}}{s_{\mathbf{k}^\nu}s_{\mathbf{l}^\theta}}M_{i_n^\lambda,i_0^\lambda}\overline{M}_{j_n^\mu,j_0^\mu}K_{k_0^\nu,k_n^\mu}\overline{K}_{\ell_0^\theta,\ell_n^\theta}u_{\mathbf{i}^\lambda}u_{\mathbf{l}^\theta}\overline{u_{\mathbf{j}^\mu}u_{\mathbf{k}^\nu}}\\ &+\sum_{\beta=1}^{r}\mathbb{E}\sum_{\substack{(\mathbf{i},\mathbf{l})\in\mathrm{I}^{\neq}\\ (\mathbf{j},\mathbf{k})\in\mathrm{I}^{\neq}\\ \mathbf{i}^1\leftrightarrow\mathbf{k}^\beta}}\prod_{\substack{1\le\lambda\le p\\ 1\le\mu\le q\\ 1\le\nu\le r\\ 1\le\theta\le s}}\frac{s_{\mathbf{i}^\lambda}s_{\mathbf{j}^\mu}}{s_{\mathbf{k}^\nu}s_{\mathbf{l}^\theta}}M_{i_n^\lambda,i_0^\lambda}\overline{M}_{j_n^\mu,j_0^\mu}K_{k_0^\nu,k_n^\mu}\overline{K}_{\ell_0^\theta,\ell_n^\theta}u_{\mathbf{i}^\lambda}u_{\mathbf{l}^\theta}\overline{u_{\mathbf{j}^\mu}u_{\mathbf{k}^\nu}}+o(1)\,. \end{aligned}$$

where to simplify, $(\mathbf{i},\mathbf{l})$ stands for the $(p+s)(n+1)$-tuple obtained by the concatenation of the $\mathbf{i}^\lambda$'s and the $\mathbf{l}^\mu$'s, and $\mathrm{I}^{\neq}$ implicitly means $\mathrm{I}^{\neq}_{(p+s)(n+1)}$. As previously, we use the Corollary 2.5.4 to split the expectations. Hence, one easily gets

$$m_{p,q,r,s} = q\mathbb{E}\left[\left|\operatorname{Tr}\mathbf{A}^n\mathbf{M}\right|^2\right]m_{p-1,q-1,r,s}+r\mathbb{E}\left[\operatorname{Tr}\mathbf{A}^n\mathbf{M}\operatorname{Tr}\mathbf{A}^{-n}\mathbf{K}\right]m_{p-1,q,r-1,s}+o(1)\,.$$

To get the other relations, we just sum according to the choice of the partner of $\mathbf{j}^1$ (resp. $\mathbf{k}^1$ and $\mathbf{l}^1$).

2.3.5 Proof of a): asymptotic factorisation of joint moments

The proof relies mostly on Corollary 2.5.4. We first expand the expectation

$$\mathbb{E}\Big[\prod_{1\le n\le n_0}\big(\operatorname{Tr}\mathbf{A}^n\mathbf{N}_n\big)^{p_n}\overline{\big(\operatorname{Tr}\mathbf{A}^n\mathbf{N}_n\big)^{q_n}}\big(\operatorname{Tr}\mathbf{A}^{-n}\mathbf{N}_{-n}\big)^{r_n}\overline{\big(\operatorname{Tr}\mathbf{A}^{-n}\mathbf{N}_{-n}\big)^{s_n}}\Big]. \tag{2.19}$$

Let $M_{ij}^{(n)}$ denote the (i,j)-th entry of $\mathbf{N}_n$ and recall that for $\mathbf{i}=(i_0,\dots,i_n)$, we set

$$u_{\mathbf{i}} := u_{i_0i_1}\cdots u_{i_{n-1}i_n} \qquad ; \qquad s_{\mathbf{i}} := s_{i_1}\cdots s_{i_n}\,.$$

We get

$$\begin{aligned}
\operatorname{Tr}\mathbf{A}^n\mathbf{N}_n &= \sum_{i_0,i_1,\dots,i_n} u_{i_0i_1}s_{i_1}\cdots u_{i_{n-1}i_n}s_{i_n}M^{(n)}_{i_ni_0} = \sum_{\mathbf{i}} u_{\mathbf{i}}s_{\mathbf{i}}M^{(n)}_{i_ni_0}\\
\operatorname{Tr}\mathbf{A}^{-n}\mathbf{N}_{-n} &= \sum_{i_n,i_{n-1},\dots,i_0} \overline{u}_{i_{n_1}i_n}s_{i_n}^{-1}\cdots\overline{u}_{i_0i_1}s_{i_1}M^{(-n)}_{i_0i_n} = \sum_{\mathbf{i}} \overline{u}_{\mathbf{i}}s_{\mathbf{i}}^{-1}M^{(-n)}_{i_0i_n},
\end{aligned}$$

so that

$$\mathbb{E}\prod_{1\le n\le n_0}\sum_{\substack{\mathbf{i}^{n,1},\dots,\mathbf{i}^{n,p_n}\\ \mathbf{j}^{n,1},\dots,\mathbf{j}^{n,q_n}\\ \mathbf{k}^{n,1},\dots,\mathbf{k}^{n,r_n}\\ \mathbf{l}^{n,1},\dots,\mathbf{l}^{n,s_n}}}\prod_{\substack{1\le\lambda\le p_n\\ 1\le\mu\le q_n\\ 1\le\nu\le r_n\\ 1\le\theta\le s_n}}\frac{s_{\mathbf{i}^{n,\lambda}}s_{\mathbf{j}^{n,\mu}}}{s_{\mathbf{k}^{n,\nu}}s_{\mathbf{l}^{n,\theta}}}M^{(n)}_{i_n^{n,\lambda}i_0^{n,\lambda}}\overline{M}^{(n)}_{j_n^{n,\mu}j_0^{n,\mu}}M^{(-n)}_{k_0^{n,\nu}k_n^{n,\mu}}\overline{M}^{(-n)}_{\ell_0^{n,\theta}\ell_n^{n,\theta}}u_{\mathbf{i}^{n,\lambda}}u_{\mathbf{l}^{n,\theta}}\overline{u_{\mathbf{j}^{n,\mu}}u_{\mathbf{k}^{n,\nu}}} \tag{2.20}$$

where we use bold letters such as $\mathbf{i}^{n,\lambda}$ to denote $(n+1)$-tuples $(i_0^{n,\lambda},i_1^{n,\lambda},\dots,i_n^{n,\lambda})$. We can use the same ideas as in [21, Lemma 5.8] to state that the non-negligible terms of the sum must satisfy that for all n, there are as much $(n+1)$-tuples involved in u as in $\overline{u}$, which means that

$$p_n+s_n \;=\; q_n+r_n,$$

and that we must have the multiset equalities

$$\begin{aligned}
&\bigcup_{n=1}^{n_0}\Big\{i_\mu^{n,\lambda},\ 1\le\lambda\le p_n, 0\le\mu\le n\Big\}_m \cup\Big\{\ell_\mu^{n,\lambda},\ 1\le\lambda\le s_n, 0\le\mu\le n\Big\}_m\\
=\;&\bigcup_{n=1}^{n_0}\Big\{j_\mu^{n,\lambda},\ 1\le\lambda\le q_n, 0\le\mu\le n\Big\}_m \cup\Big\{k_\mu^{n,\lambda},\ 1\le\lambda\le r_n, 0\le\mu\le n\Big\}_m
\end{aligned} \tag{2.21}$$

We deduce that there are a $O\big(N^{\sum_n n(p_n+s_n)}\big)$ non zero terms in (2.20) and we can easily show that any subsum over a $o\big(N^{\sum_n n(p_n+s_n)}\big)$ is negligible so that for now on we shall sum over the non pairwise indices. Then, we know that we can neglect any expectation $\mathbb{E}_{\mathbf{U}}$ which won't lead to $\mathrm{Wg}(id)$ (see (2.23)) so that (2.21) becomes

$$\begin{aligned}
\forall 1\le n\le n_0,\qquad &\Big\{i_\mu^{n,\lambda},\ 1\le\lambda\le p_n, 0\le\mu\le n\Big\}\cup\Big\{\ell_\mu^{n,\lambda},\ 1\le\lambda\le s_n, 0\le\mu\le n\Big\}\\
=\;&\Big\{j_\mu^{n,\lambda},\ 1\le\lambda\le q_n, 0\le\mu\le n\Big\}\cup\Big\{k_\mu^{n,\lambda},\ 1\le\lambda\le r_n, 0\le\mu\le n\Big\}.
\end{aligned}$$

It follows that the set of indices involved in the expansion of the $\operatorname{Tr}\mathbf{A}^n\mathbf{N}_n$, $\operatorname{Tr}\mathbf{A}^{-n}\mathbf{N}_{-n}$, $\overline{\operatorname{Tr}\mathbf{A}^n\mathbf{N}_n}$, $\overline{\operatorname{Tr}\mathbf{A}^{-n}\mathbf{N}_{-n}}$, is disjoint from the set of indices involved in the expansion of the $\operatorname{Tr}\mathbf{A}^m\mathbf{N}_m$, $\operatorname{Tr}\mathbf{A}^{-m}\mathbf{N}_{-m}$, $\overline{\operatorname{Tr}\mathbf{A}^m\mathbf{N}_m}$, $\overline{\operatorname{Tr}\mathbf{A}^{-m}\mathbf{N}_{-m}}$, as long as $n \neq m$. Therefore, Corollary 2.5.4 allows to conclude the proof of **a)**.

2.3.6 Proof of c): computation of the limit covariance

Let f, g be polynomials in z and z^{-1} and let $\mathbf{M}, \mathbf{N}$ be $N \times N$ deterministic matrices such that, as $N \to \infty$,

$$\frac{1}{N}\operatorname{Tr}\mathbf{M} \longrightarrow \tau \quad ; \quad \frac{1}{N}\operatorname{Tr}\mathbf{N} \longrightarrow \tau' \quad ; \quad \frac{1}{N}\operatorname{Tr}\mathbf{M}\mathbf{N}^* \longrightarrow \alpha \quad ; \quad \frac{1}{N}\operatorname{Tr}\mathbf{M}\mathbf{N} \longrightarrow \beta.$$

We need to check that the limits of both sequences

$$\mathbb{E}(\operatorname{Tr} f(\mathbf{A})\mathbf{M} - a_0(f)\operatorname{Tr}\mathbf{M})(\overline{\operatorname{Tr} g(\mathbf{A})\mathbf{N} - a_0(g)\operatorname{Tr}\mathbf{N}})$$

and

$$\mathbb{E}(\operatorname{Tr} f(\mathbf{A})\mathbf{M} - a_0(f)\operatorname{Tr}\mathbf{M})(\operatorname{Tr} g(\mathbf{A})\mathbf{N} - a_0(g)\operatorname{Tr}\mathbf{N})$$

are the ones given in the statement of Theorem 2.2.4. Note that it suffices to compute the limits for $f = g$ and $\mathbf{M} = \mathbf{N}$. Indeed, using the classical polarization identities for $\mathbf{M}$ and $\mathbf{N}$, first for general polynomials f, g, we reduce the problem to the case $\mathbf{M} = \mathbf{N}$. Then, we use polarization identities again to reduce the problem to $f = g$.

Also, recall that since $\mathbf{A} \overset{(\mathrm{d})}{=} e^{i\theta}\mathbf{A}$ for any deterministic θ, we know that for any positive distinct integers p, q, we have

$$\mathbb{E}\operatorname{Tr}\mathbf{A}^p\mathbf{M}\operatorname{Tr}\mathbf{A}^{-q}\mathbf{M} \;=\; \mathbb{E}\operatorname{Tr}\mathbf{A}^p\mathbf{M}\overline{\operatorname{Tr}\mathbf{A}^q\mathbf{M}} \;=\; 0.$$

It follows, using (2.4), (2.5) and (2.6), that

$$\begin{aligned}
\mathbb{E}\,|\operatorname{Tr} f(\mathbf{A})\mathbf{M} - a_0(f)\operatorname{Tr}\mathbf{M}|^2 \quad &= \quad \sum_{\substack{m,n\in\mathbb{Z}\\ \neq 0}} a_m(f)\overline{a_n(f)}\,\mathbb{E}\operatorname{Tr}\mathbf{A}^m\mathbf{M}\overline{\operatorname{Tr}\mathbf{A}^n\mathbf{M}} \\
&= \quad \sum_{n\geq 1}\Big(|a_n(f)|^2\,\mathbb{E}\,\big|\operatorname{Tr}\mathbf{A}^n\mathbf{M}\big|^2 + |a_{-n}(f)|^2\,\mathbb{E}\,\big|\operatorname{Tr}\mathbf{A}^{-n}\mathbf{M}\big|^2\Big) \\
&\longrightarrow \quad \sum_{n\geq 1}\big(|a_n(f)|^2 b^{2n} + |a_{-n}(f)|^2 a^{-2n}\big)\big((n-1)|\tau|^2 + \alpha\big), \\
\mathbb{E}\,(\operatorname{Tr} f(\mathbf{A})\mathbf{M} - a_0(f)\operatorname{Tr}\mathbf{M})^2 \quad &= \quad \sum_{\substack{m,n\in\mathbb{Z}\\ \neq 0}} a_m(f)a_n(f)\,\mathbb{E}\operatorname{Tr}\mathbf{A}^m\mathbf{M}\operatorname{Tr}\mathbf{A}^n\mathbf{M} \\
&= \quad \sum_{n\geq 1} 2a_n(f)a_{-n}(f)\,\mathbb{E}\operatorname{Tr}\mathbf{A}^n\mathbf{M}\operatorname{Tr}\mathbf{A}^{-n}\mathbf{M} \\
&\longrightarrow \quad 2\sum_{n\geq 1} a_n(f)a_{-n}(f)\big((n-1)\tau^2 + \beta\big),
\end{aligned}$$

which concludes the proof.

2.4 Proof of Corollary 2.2.13

It is easy to see that for any $z \notin S$, we have

$$\log|\det(z-\mathbf{A})| = \begin{cases} \operatorname{Tr}\log\mathbf{T} + \operatorname{Re}\mathscr{A}_z^N, & \text{with } \mathscr{A}_z^N := \operatorname{Tr}\sum_{n\leq -1}\frac{\mathbf{A}^n}{nz^n} \text{ if } |z|<a. \\ N\log|z| + \operatorname{Re}\mathscr{B}_z^N, & \text{with } \mathscr{B}_z^N := -\operatorname{Tr}\sum_{n\geq 1}\frac{\mathbf{A}^n}{nz^n} \text{ if } |z|>b, \end{cases}$$

(in the first case, we used the fact that $|\det\mathbf{A}| = \det\mathbf{T}$). Then, by Theorem 2.2.4,

$$\left(\mathscr{A}_z^N\right)_{|z|<a} \cup \left(\mathscr{B}_z^N\right)_{|z|>b}$$

converges, for the finite-dimensional distributions, to a centered complex Gaussian process

$$(\mathscr{A}_z)_{|z|<a} \cup (\mathscr{B}_z)_{|z|>b}$$

with covariance defined by

$$\mathbb{E}\,\mathscr{A}_z\mathscr{A}_{z'} = 0, \quad \mathbb{E}\,\mathscr{A}_z\overline{\mathscr{A}_{z'}} = -\log(1-\frac{z\overline{z'}}{a^2}),$$

$$\mathbb{E}\,\mathscr{B}_z\mathscr{B}_{z'} = 0, \quad \mathbb{E}\,\mathscr{B}_z\overline{\mathscr{B}_{z'}} = -\log(1-\frac{b^2}{z\overline{z'}}),$$

$$\mathbb{E}\,\mathscr{A}_z\mathscr{B}_{z'} = -\log(1-\frac{z'}{z}), \quad \mathbb{E}\,\mathscr{A}_z\overline{\mathscr{B}_{z'}} = 0,$$

where log denotes the canonical complex log on $B(1,1)$. Then, one concludes by noting that for $A,B\in\mathbb{C}$, $2\operatorname{Re}A\operatorname{Re}B = \operatorname{Re}(AB+A\overline{B})$.

2.5 Appendix

2.5.1 Weingarten calculus on the unitary group

Here comes a main result about integration with respect to the Haar measure on unitary group, (see [38, Cor. 2.4 and Cor. 2.7]). To simplify, for any k-tuples $\mathbf{i} = (i_1,\dots,i_k)$ and $\mathbf{j} = (j_1,\dots,j_k)$, we set

$$u_{\mathbf{i},\mathbf{j}} := u_{i_1j_1}u_{i_2j_2}\cdots u_{i_kj_k}$$

Proposition 2.5.1. *Let k be a positive integer and $\mathbf{U} = (u_{ij})$ a $N\times N$ Haar-distributed matrix. Let $\mathbf{i} = (i_1,\dots,i_k)$, $\mathbf{i}' = (i'_1,\dots,i'_k)$, $\mathbf{j} = (j_1,\dots,j_k)$ and $\mathbf{j}' = (j'_1,\dots,j'_k)$ be four k-tuples of $\{1,\dots,N\}$. Then*

$$\mathbb{E}\left[u_{\mathbf{i},\mathbf{j}}\overline{u_{\mathbf{i}',\mathbf{j}'}}\right] = \sum_{\sigma,\tau\in S_k}\delta_{i_1,i'_{\sigma(1)}}\dots\delta_{i_k,i'_{\sigma(k)}}\delta_{j_1,j'_{\tau(1)}}\dots\delta_{j_k,j'_{\tau(k)}}\operatorname{Wg}(\tau\sigma^{-1}), \tag{2.22}$$

where Wg *is a function called the* Weingarten function. *Moreover, for $\sigma\in S_k$, the asymptotical behavior of* $\operatorname{Wg}(\sigma)$ *is given by*

$$N^{k+|\sigma|}\operatorname{Wg}(\sigma) = \operatorname{Moeb}(\sigma) + O\left(\frac{1}{N^2}\right), \tag{2.23}$$

where $|\sigma|$ *denotes the minimal number of factors necessary to write* σ *as a product of transpositions, and* Moeb *denotes a function called the* Möbius function.

Remark 2.5.2. One should notice that if all k-tuples $(i_1,\ldots,i_k)$, $(j_1,\ldots,j_k)$, $(i'_1,\ldots,i'_k)$, and $(j'_1,\ldots,j'_k)$ have pairwise distinct entries, then (2.22) becomes simpler because in this case there is at most one non-zero term in the sum.

Remark 2.5.3. The permutation σ for which $\mathrm{Wg}(\sigma)$ will have the largest order is the only one satisfying $|\sigma|=0$, i.e. $\sigma = id$. Also, $\mathrm{Moeb}(id)=1$ (see [38]).

Here comes an useful corollary which permits to simplify many computations.

Corollary 2.5.4. *Let* $\mathbf{i}=(i_1,\ldots,i_p)$, $\mathbf{j}=(j_1,\ldots,j_p)$, $\mathbf{k}=(k_1,\ldots,k_q)$, $\mathbf{l}=(\ell_1,\ldots,\ell_q)$, $\mathbf{i}'=(i'_1,\ldots,i'_p)$, $\mathbf{j}'=(j'_1,\ldots,j'_p)$, $\mathbf{k}'=(k'_1,\ldots,k'_q)$, $\mathbf{l}'=(\ell'_1,\ldots,\ell'_q)$ *be tuples such the multisets defined by* $\mathbf{i}$ *and* $\mathbf{i}'$ *(resp. by* $\mathbf{j}$ *and* $\mathbf{j}'$*, by* $\mathbf{k}$ *and* $\mathbf{k}'$*, by* $\mathbf{l}$ *and* $\mathbf{l}'$*) are equal and such that*

$$\{i_1,\ldots,i_p\}\cap\{k_1,\ldots,k_q\} \;=\; \{j_1,\ldots,j_p\}\cap\{l_1,\ldots,l_q\} \;=\; \emptyset.$$

Then

$$\mathbb{E}\left[u_{\mathbf{i},\mathbf{j}}u_{\mathbf{k},\mathbf{l}}\overline{u_{\mathbf{i}',\mathbf{j}'}u_{\mathbf{k}',\mathbf{l}'}}\right] \;=\; \mathbb{E}\left[u_{\mathbf{i},\mathbf{j}}\overline{u_{\mathbf{i}',\mathbf{j}'}}\right]\mathbb{E}\left[u_{\mathbf{k},\mathbf{l}}\overline{u_{\mathbf{k}',\mathbf{l}'}}\right]\left(1+O\left(\frac{1}{N^2}\right)\right).$$

Proof. To prove this result, we first recall the exact expression of the Möbius function : for any permutation σ with cycle decomposition $C_1C_2\cdots C_r$,

$$\mathrm{Moeb}(\sigma) \;=\; \prod_{i=1}^{r}(-1)^{|C_i|-1}\,\mathrm{Cat}_{|C_i|-1},$$

where Cat_k is the k-th Catalan number, $\frac{1}{k+1}\binom{2k}{k}$. Then, obviously, if σ and τ are two permutations with disjoint supports, then

$$\mathrm{Moeb}(\sigma\circ\tau) \;=\; \mathrm{Moeb}(\sigma)\,\mathrm{Moeb}(\tau) \qquad \text{and} \qquad \left|\sigma\circ\tau\right| \;=\; |\sigma|+|\tau|.$$

Thus

$$\begin{aligned} N^{p+q+|\sigma\circ\tau|}\,\mathrm{Wg}(\sigma\circ\tau) \;&=\; \mathrm{Moeb}(\sigma\circ\tau)+O\left(\frac{1}{N^2}\right) \\ &=\; \mathrm{Moeb}(\sigma)\,\mathrm{Moeb}(\tau)+O\left(\frac{1}{N^2}\right) \\ &=\; n^{p+|\sigma|}\,\mathrm{Wg}(\sigma)N^{q+|\tau|}\,\mathrm{Wg}(\tau)\left(1+O\left(\frac{1}{N^2}\right)\right) \end{aligned}$$

So that

$$\mathrm{Wg}(\sigma\circ\tau) \;=\; \mathrm{Wg}(\sigma)\,\mathrm{Wg}(\tau)\left(1+O\left(\frac{1}{N^2}\right)\right).$$

One can easily conclude. □

We also need the following lemmas in the paper.

Lemma 2.5.5. *Let $\mathbf{A} = \mathbf{U}\mathbf{T}$ with $\mathbf{U}$ Haar-distributed on the unitary group and $\mathbf{T}$ deterministic. Let $\varepsilon > 0$. There is a finite constant C depending only on ε (in particular, independent of N and of $\mathbf{T}$) such that for all positive integer n such that $n^6 < (2-\varepsilon)N$, we have*

$$\mathbb{E}\operatorname{Tr}\mathbf{A}^n(\mathbf{A}^n)^*) \;\le\; CNn^2\left(m_2 + \frac{nm_\infty^2}{N}\right)^n$$

and

$$\mathbb{E}[|\operatorname{Tr}\mathbf{A}^n|^2] \;\le\; C\left(m_2 + \frac{nm_\infty^2}{N}\right)^n,$$

where $m_2 := N^{-1}\operatorname{Tr}\mathbf{T}\mathbf{T}^*$ and $m_\infty := \|\mathbf{T}\|_{\mathrm{op}}$.

Proof. See [15, Th. 1]. □

Lemma 2.5.6. *Let $\mathbf{V}$ be Haar-distributed and let $\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}$ be deterministic $N\times N$ matrices. Then we have*

$$\begin{aligned}\mathbb{E}\operatorname{Tr}\mathbf{AVBV}^*\operatorname{Tr}\mathbf{CVDV}^* \;&=\; \frac{1}{N^2-1}\left(\operatorname{Tr}\mathbf{A}\operatorname{Tr}\mathbf{C}\operatorname{Tr}\mathbf{B}\operatorname{Tr}\mathbf{D} + \operatorname{Tr}\mathbf{AC}\operatorname{Tr}\mathbf{BD}\right)\\ &\quad -\frac{1}{N(N^2-1)}\left(\operatorname{Tr}\mathbf{A}\operatorname{Tr}\mathbf{C}\operatorname{Tr}\mathbf{BD} + \operatorname{Tr}\mathbf{AC}\operatorname{Tr}\mathbf{B}\operatorname{Tr}\mathbf{D}\right)\end{aligned}$$

Proof. Let $\mathscr{M}_N(\mathbb{C})$ denote the set of $N\times N$ complex matrices. It has already been proved, in [21, Lem. 5.9], that for any matrices $\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}\in\mathscr{M}_N(\mathbb{C})$, we have

$$\begin{aligned}\mathbb{E}\operatorname{Tr}\mathbf{AVBV}^*\mathbf{CVDV}^* \;&=\; \frac{1}{N^2-1}\left\{\operatorname{Tr}\mathbf{AC}\operatorname{Tr}\mathbf{B}\operatorname{Tr}\mathbf{D} + \operatorname{Tr}\mathbf{A}\operatorname{Tr}\mathbf{C}\operatorname{Tr}\mathbf{BD}\right\}\\ &\quad -\frac{1}{N(N^2-1)}\left\{\operatorname{Tr}\mathbf{AC}\operatorname{Tr}\mathbf{BD} + \operatorname{Tr}\mathbf{A}\operatorname{Tr}\mathbf{C}\operatorname{Tr}\mathbf{B}\operatorname{Tr}\mathbf{D}\right\}.\end{aligned} \tag{2.24}$$

We deduce that

$$\begin{aligned}\mathbb{E}\,\mathbf{V}\otimes\mathbf{V}^*\otimes\mathbf{V}\otimes\mathbf{V}^* &= \frac{1}{N^2-1}\left\{\sum_{i,j,k,\ell} E_{j,k}\otimes E_{k,j}\otimes E_{i,\ell}\otimes E_{\ell,i} + E_{i,j}\otimes E_{k,\ell}\otimes E_{\ell,k}\otimes E_{j,i}\right\}\\ &\quad -\frac{1}{N(N^2-1)}\left\{\sum_{i,j,k,\ell} E_{j,k}\otimes E_{\ell,j}\otimes E_{i,\ell}\otimes E_{k,i} + E_{i,j}\otimes E_{j,k}\otimes E_{k,\ell}\otimes E_{\ell,i}\right\},\end{aligned} \tag{2.25}$$

where the $E_{r,s}$ denote the elementary matrices. Indeed, the linear morphism Ψ from $\mathscr{M}_N(\mathbb{C})^{\otimes 4}$ to the space of 4-linear forms on $\mathscr{M}_N(\mathbb{C})$ defined by

$$\Psi(\mathbf{M}\otimes\mathbf{N}\otimes\mathbf{P}\otimes\mathbf{Q})(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}) := \operatorname{Tr}\mathbf{AMBNCPDQ}$$

is an isomorphism, and (2.24) proves that the left and right hand terms of (2.25) have the same image by Ψ. Then, applying

$$\mathbf{M}\otimes\mathbf{N}\otimes\mathbf{P}\otimes\mathbf{Q} \longmapsto \operatorname{Tr}\left(\mathbf{AMBN}\right)\operatorname{Tr}\left(\mathbf{CPDQ}\right)$$

on both sides of (2.25), we deduce the lemma. □

2.5.2 Moments of a Gaussian vector with values in $\mathbb{C}^2$

The following lemma allows to prove that a complex random vector $(\tilde{X},\tilde{Y})$ is Gaussian without having to compute all its joint moments, by only proving an induction relation.

Lemma 2.5.7. *Let (X,Y) be a Gaussian random vector with values in $\mathbb{C}^2$ whose distribution is characterized by*

$$\mathbb{E}X \;=\; \mathbb{E}X^2 \;=\; \mathbb{E}Y \;=\; \mathbb{E}Y^2 \;=\; 0, \tag{2.26}$$

and

$$\mathbb{E}|X|^2 = \sigma_X \;;\; \mathbb{E}|Y|^2 = \sigma_Y \;;\; \mathbb{E}XY = \sigma_{XY} \;;\; \mathbb{E}X\overline{Y} = 0. \tag{2.27}$$

Then, the moments

$$m_{p,q,r,s} \;\;:=\;\; \mathbb{E}X^p\overline{X}^qY^r\overline{Y}^s \tag{2.28}$$

satisfy

$$p-q \neq r-s \quad\Longrightarrow\quad m_{p,q,r,s} \;=\; 0, \tag{2.29}$$

$$m_{p,q,r,s} \;\;=\;\; \begin{cases} q\sigma_X m_{p-1,q-1,r,s} + r\sigma_{XY} m_{p-1,q,r-1,s} \\ p\sigma_X m_{p-1,q-1,r,s} + s\overline{\sigma_{XY}} m_{p,q-1,r,s-1} \\ s\sigma_Y m_{p,q,r-1,s-1} + p\sigma_{XY} m_{p-1,q,r-1,s} \\ r\sigma_Y m_{p,q,r-1,s-1} + q\overline{\sigma_{XY}} m_{p,q-1,r,s-1} \end{cases} \tag{2.30}$$

and

$$m_{p,0,p,0} \;=\; \mathbb{E}(XY)^p \;\;=\;\; p!\sigma_{XY}^p. \tag{2.31}$$

Conversely, if $(\tilde{X},\tilde{Y})$ is a random vector with values in $\mathbb{C}^2$ such that both $\tilde{X}$ and $\tilde{Y}$ are Gaussian and have joint moments $\tilde{m}_{p,q,r,s}$ satisfying (2.26), (2.27), (2.29), (2.30) and (2.31), then $(\tilde{X},\tilde{Y})$ is Gaussian.

Proof. First, one easily obtains (2.29) by noticing that for any $\theta \in \mathbb{R}$, $(\mathrm{e}^{\mathrm{i}\theta}X, \mathrm{e}^{-\mathrm{i}\theta}Y) \overset{\text{law}}{=} (X,Y)$.

To prove the remaining, we consider (X,Y) as a real 4-tuple $(\Re(X),\Im(X),\Re(Y),\Im(Y)) =: (x_1,x_2,x_3,x_4)$ with covariance matrix

$$\Gamma \;:=\; \frac{1}{2}\begin{pmatrix} \sigma_X & 0 & \Re(\sigma_{XY}) & \Im(\sigma_{XY}) \\ 0 & \sigma_X & \Im(\sigma_{XY}) & -\Re(\sigma_{XY}) \\ \Re(\sigma_{XY}) & \Im(\sigma_{XY}) & \sigma_Y & 0 \\ \Im(\sigma_{XY}) & -\Re(\sigma_{XY}) & 0 & \sigma_Y \end{pmatrix}.$$

Its Fourier transform is given, for $\mathbf{t}=(t_1,t_2,t_3,t_4)$, by

$$\begin{aligned}\Phi(\mathbf{t}) &:= \mathbb{E}\exp\big(\mathrm{i}(t_1x_1+t_2x_2+t_3x_3+t_4x_4)\big)\\ &= \exp\Big\{-\frac{1}{4}\big(\sigma_X(t_1^2+t_2^2)+\sigma_Y(t_3^2+t_4^2)\big)-\frac{1}{2}\Re(\sigma_{XY})\big(t_1t_3-t_2t_4\big)-\frac{1}{2}\Im(\sigma_{XY})\big(t_2t_3+t_1t_4\big)\Big\}\end{aligned}$$

We define the differential operators

$$\begin{aligned}\partial_X &= \partial_1+\mathrm{i}\partial_2 \qquad & \partial_{\overline{X}} &= \partial_1-\mathrm{i}\partial_2\\ \partial_Y &= \partial_3+\mathrm{i}\partial_4 \qquad & \partial_{\overline{Y}} &= \partial_3-\mathrm{i}\partial_4\end{aligned}$$

so that

$$\mathbb{E}X^p\overline{X}^qY^r\overline{Y}^s = (-\mathrm{i})^{p+q+r+s}\partial_X^p\partial_{\overline{X}}^q\partial_Y^r\partial_{\overline{Y}}^s\Phi(\mathbf{t})\big|_{\mathbf{t}=0} \tag{2.32}$$

and

$$\begin{aligned}\partial_X\Phi(\mathbf{t}) &= -\frac{1}{2}\big(\sigma_X(t_1+\mathrm{i}t_2)+\sigma_{XY}(t_3-\mathrm{i}t_4)\big)\Phi(\mathbf{t})\\ \partial_{\overline{X}}\Phi(\mathbf{t}) &= -\frac{1}{2}\big(\sigma_X(t_1-\mathrm{i}t_2)+\overline{\sigma_{XY}}(t_3+\mathrm{i}t_4)\big)\Phi(\mathbf{t})\\ \partial_Y\Phi(\mathbf{t}) &= -\frac{1}{2}\big(\sigma_Y(t_3+\mathrm{i}t_4)+\sigma_{XY}(t_1-\mathrm{i}t_2)\big)\Phi(\mathbf{t})\\ \partial_{\overline{Y}}\Phi(\mathbf{t}) &= -\frac{1}{2}\big(\sigma_Y(t_3-\mathrm{i}t_4)+\overline{\sigma_{XY}}(t_1+\mathrm{i}t_2)\big)\Phi(\mathbf{t})\end{aligned}$$

We easily deduce that

$$\partial_X^p\partial_{\overline{X}}^q\partial_Y^r\partial_{\overline{Y}}^s\Phi(\mathbf{t})\big|_{\mathbf{t}=0} = -q\sigma_X\partial_X^{p-1}\partial_{\overline{X}}^{q-1}\partial_Y^r\partial_{\overline{Y}}^s\Phi(\mathbf{t})\big|_{\mathbf{t}=0}-r\sigma_{XY}\partial_X^{p-1}\partial_{\overline{X}}^q\partial_Y^{r-1}\partial_{\overline{Y}}^s\Phi(\mathbf{t})\big|_{\mathbf{t}=0}$$

so that

$$\mathbb{E}X^p\overline{X}^qY^r\overline{Y}^s = q\sigma_X\,\mathbb{E}X^{p-1}\overline{X}^{q-1}Y^r\overline{Y}^s+r\sigma_{XY}\,\mathbb{E}X^{p-1}\overline{X}^qY^{r-1}\overline{Y}^s.$$

This proves (2.30). To prove the last point, we simply write

$$\mathbb{E}(XY)^p=(-1)^p\partial_X^p\partial_Y^p\Phi(\mathbf{t})\big|_{\mathbf{t}=0} \quad ; \quad \partial_X^p\Phi(\mathbf{t})=\Big(-\frac{\sigma_X(t_1+\mathrm{i}t_2)+\sigma_{XY}(t_3-\mathrm{i}t_4)}{2}\Big)^p\Phi(\mathbf{t}).$$

Then, using Leibniz formula and noticing that for all $k\le p$,

$$\partial_Y^k\Big(-\frac{\sigma_X(t_1+\mathrm{i}t_2)+\sigma_{XY}(t_3-\mathrm{i}t_4)}{2}\Big)^p\Big|_{\mathbf{t}=0}=\begin{cases}0 & \text{if } k<p\\ p!(\sigma_{XY})^p & \text{if } k=p\end{cases},$$

one can easily conclude.

Conversely, let $(\tilde X,\tilde Y)$ be a random vector with values in $\mathbb{C}^2$ such that both $\tilde X$ and $\tilde Y$ are Gaussian and have joint moments $\tilde m_{p,q,r,s}$ satisfying (2.26), (2.27), (2.29), (2.30) and (2.31). Let $\mathbb{N}$ denote the set of non-negative integers and let us define the sets

$$\mathscr{K} := \left\{(p,q,r,s) \in \mathbb{N}^4\,;\;\; p-q \neq r-s\right\},$$

$$\mathscr{H}_0 := \left\{(p,q,r,s) \in \mathbb{N}^4\,;\, (r+s+|p-q|)\,(p+q+|r-s|)\,(p+r+|q-s|)\,(q+s+|p-r|) = 0\right\}$$

and, for $k \geq 1$,

$$\begin{aligned}\mathscr{H}_k \;\; := \;\; &\{(p,q,r,s) \in \mathbb{N}^4\,;\;\; (p-1,q-1,r,s) \text{ or } (p-1,q,r-1,s) \text{ or} \\ &\qquad\qquad (p,q-1,r,s-1) \text{ or } (p,q,r-1,s-1) \in \mathscr{H}_{k-1}\}\end{aligned}$$

Then by hypothesis, the joint moments function $\tilde{m}_{p,q,r,s}$ coincides with $m_{p,q,r,s}$ on $\mathscr{K} \cup \mathscr{H}_0$. Besides, by (2.30), if $\tilde{m}_{p,q,r,s}$ coincides with $m_{p,q,r,s}$ on $\mathscr{H}_{k-1}$, then $\tilde{m}_{p,q,r,s}$ coincides with $m_{p,q,r,s}$ on $\mathscr{H}_k$. As

$$\mathbb{N}^4 \;=\; \mathscr{K} \cup \bigcup_{k \geq 0} \mathscr{H}_k,$$

we deduce that $\tilde{m}_{p,q,r,s}$ coincides with $m_{p,q,r,s}$ on $\mathbb{N}^4$, wich implies that $(\tilde{X},\tilde{Y}) \overset{\text{law}}{=} (X,Y)$. $\square$

III

Complex outliers of Hermitian random matrices

Jean Rochet

Abstract :

In this paper, we study the asymptotic behavior of the outliers of the sum a Hermitian random matrix and a finite rank matrix which is not necessarily Hermitian. We observe several possible convergence rates and outliers locating around their limits at the vertices of regular polygons as in [21], as well as possible correlations between outliers at macroscopic distance as in [62] and [21]. We also observe that a single spike can generate several outliers in the spectrum of the deformed model, as already noticed in [19] and [12]. In the particular case where the perturbation matrix is Hermitian, our results complete the work of [16], as we consider fluctuations of outliers lying in “holes" of the limit support, which happen to exhibit surprising correlations.

3.1 Introduction

It is known that adding a finite rank perturbation to a large matrix barely changes the global behavior of its spectrum. Nevertheless, some of the eigenvalues, called *outliers*, can deviate away from the bulk of the spectrum, depending on the strength of the perturbation. This phenomenon, well known as the *BBP transition*, was first brought to light for empirical covariance matrices by Johnstone in [57], by Baik, Ben Arous and Péché in [9], and then shown under several hypothesis in the Hermitian case in [77, 43, 28, 29, 79, 19, 20, 16, 17, 30, 61, 62]. Non-Hermitian models have been also studied: i.i.d. matrices in [92, 24, 82], elliptic matrices in [73] and matrices from the Single Ring Theorem in [21]. In [21], and lately in [82], the authors have also studied the fluctuations of the outliers and, due to non-Hermitian structure, obtained unusual results: the distribution of the fluctuations highly depends on the shape of the Jordan Canonical Form of the perturbation, in particular, the convergence rate depends on the size of the Jordan blocks. Also, the outliers tend to locate around their limit at the vertices of a regular polygon. At last, they observe correlations between the fluctuations of outliers at a macroscopic distance with each other.
In this paper, we show that the same kind of phenomenon occurs when we perturb an Hermitian matrix $\mathbf{H}$ with a non-Hermitian one $\mathbf{A}$. More precisely, we study finite rank perturbations for Hermitian random matrices $\mathbf{H}$ whose spectral measure tends to a compactly supported measure μ and the perturbation $\mathbf{A}$ is just a complex matrix with a finite rank. With further assumptions, we prove that outliers of $\mathbf{H}+\mathbf{A}$ may appear at a macroscopic distance from the bulk and, following the ideas of [21], we show that they fluctuate with convergence rates which depend on the matrix $\mathbf{A}$ through its Jordan Canonical Form. Remind that any complex matrix is

similar to a block diagonal matrix with diagonal blocks of the type

$$\mathbf{R}_p(\theta) \;:=\; \begin{pmatrix} \theta & 1 & & (0) \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ (0) & & & \theta \end{pmatrix},$$

so that $\mathbf{A} \sim \operatorname{diag}\left(\mathbf{R}_{p_1}(\theta_1), \ldots, \mathbf{R}_{p_q}(\theta_q)\right)$, this last matrix being called the *Jordan Canonical Form* of $\mathbf{A}$ [53, Chapter 3]. We show, up to some hypothesis, that for any eigenvalue θ of $\mathbf{A}$, if we denote by

$$\underbrace{p_1, \ldots, p_1}_{\beta_1 \text{ times}} > \underbrace{p_2, \ldots, p_2}_{\beta_2 \text{ times}} > \cdots > \underbrace{p_\alpha, \ldots, p_\alpha}_{\beta_\alpha \text{ times}}$$

the sizes of the blocks associated to θ in the Jordan Canonical Form of $\mathbf{A}$ and introduce the (possibly empty) set

$$\mathscr{S}_\theta \;:=\; \left\{ \xi \in \mathbb{C},\; G_\mu(\xi) = \frac{1}{\theta} \right\}$$

where $G_\mu(z) := \displaystyle\int \frac{1}{z-x}\mu(dx)$ is the Cauchy transform of the measure μ, then there are exactly $\beta_1 p_1 + \cdots + \beta_\alpha p_\alpha$ outliers of $\mathbf{H}+\mathbf{A}$ tending to each element of $\mathscr{S}_\theta$. We also prove that for each element ξ in $\mathscr{S}_\theta$, there are exactly $\beta_1 p_1$ outliers tending to ξ at rate $N^{-1/(2p_1)}$, $\beta_2 p_2$ outliers tending to ξ at rate $N^{-1/(2p_2)}$, etc... (see Figure 3.10). Furthermore, the limit joint distribution of the fluctuations is explicit, not necessarily Gaussian, and might show correlations even between outliers at a macroscopic distance with each other. This phenomenon of correlations between the fluctuations of two outliers with distinct limits has already been proved for non-Gaussian Wigner matrices when $\mathbf{A}$ is Hermitian (see [62]), while in our case, Gaussian Wigner matrices can have such correlated outliers: indeed, the correlations that we bring to light here are due to the fact that the eigenspaces of $\mathbf{A}$ are not necessarily orthogonal or that one single spike generates several outliers. Indeed, we observe that the outliers may outnumber the rank of $\mathbf{A}$. This had already been noticed in [19, Remark 2.11] when the support of the limit spectral measure of $\mathbf{H}$ has some “holes” or in the different model of [12], where the authors study the case where $\mathbf{A}$ is Hermitian but with full rank and is invariant in distribution by unitary conjugation. Here, the phenomenon can be proved to occur even when the support of the limit spectral measure of $\mathbf{H}$ is connected. At last, if we apply our results in the particular case where $\mathbf{A}$ is Hermitian, we also see that two outliers at a macroscopic distance with each other are correlated if they both are generated by the same spike (which can occur only if the limit support is disconnected) and are independent otherwise (see Figure 3.11). From this point of view, this completes the work of [16], where fluctuations of outliers lying in “holes" of the limit support

had not been studied.
The fact to consider a non-Hermitian deformation on a Hermitian random matrix has already been studied in theoretical physics (see [44, 45, 46, 47]) in the particular case where $\mathbf{H}$ is a GOE/GUE matrix and $\mathbf{A}$ is a non negative Hermitian matrix times i (the square root of -1). They proved a weaker version of Theorem 3.2.3 in this specific case but didn't study the fluctuations.
The proofs of this paper rely essentially on the ideas of the paper [21] about outliers in the Single Ring Theorem and on the results proved in [79, 80, 16]. More precisely, the study of the fluctuations reproduce the outlines of the proofs of [21] as long as the model fulfills some conditions. Thanks to [79, 80], we show that these conditions are satisfied for Wigner matrices. At last, using [16] and the Weingarten calculus, we show the same for Hermitian matrices invariant in distribution by unitary conjugation. In the appendix, as a tool for the outliers study, we prove a result on the fluctuations of the entries of such matrices.

3.2 General Results

At first, we formulate the results in general settings and we shall give, in the next section, examples of random matrices on which these results apply.

3.2.1 Convergence of the outliers

Set up and assumptions

For all $N \geq 1$, let $\mathbf{H}_N$ be an Hermitian random $N \times N$ matrix whose *empirical spectral measure*, as N goes to infinity, converges weakly in probability to a compactly supported measure μ

$$\mu_N \quad := \quad \frac{1}{N}\sum_{i=1}^{N}\delta_{\lambda_i(\mathbf{H})} \longrightarrow \mu. \tag{3.1}$$

We shall suppose that μ is non trivial in the sense that μ is not a single Dirac measure. Also, we suppose that $\mathbf{H}_N$ does not possess any *natural outliers*, i.e.

Assumption 3.2.1. As N goes to infinity, with probability tending to one,

$$\sup_{\lambda \in \mathrm{Spec}(\mathbf{H}_N)} \mathrm{dist}(\lambda, \mathrm{supp}(\mu)) \quad \longrightarrow \quad 0.$$

For all $N \geq 1$, let $\mathbf{A}_N$ be an $N \times N$ random matrix independent from $\mathbf{H}_N$ (which does not satisfies necessarily $\mathbf{A}_N^* = \mathbf{A}_N$) whose rank is bounded by an integer r (independent from N). We know that we can write

$$\mathbf{A}_N \quad := \quad \mathbf{U}\begin{pmatrix}\mathbf{A}_0 & 0\\ 0 & 0\end{pmatrix}\mathbf{U}^* \tag{3.2}$$

where $\mathbf{U}$ is an $N\times N$ unitary matrix and $\mathbf{A}_0$ is $2r\times 2r$ matrix. We notice that $\mathbf{A}_N$ only depends on the $2r$-first columns of $\mathbf{U}$ so that, we shall write

$$\mathbf{A}_N \;:=\; \mathbf{U}_{2r}\mathbf{A}_0\mathbf{U}_{2r}^*,$$

where the $N\times 2r$ matrix $\mathbf{U}_{2r}$ designates the $2r$-first columns of $\mathbf{U}$. We shall assume that $\mathbf{A}_0$ is deterministic and independent from N. We shall denote by $\theta_1,\ldots,\theta_j$ the distinct non-zero eigenvalues of $\mathbf{A}_0$ and $k_1,\ldots,k_j$ their respective multiplicity[10] (note that $\sum_{i=1}^{j} k_i \leq r$). We consider the additive perturbation

$$\widetilde{\mathbf{H}}_N \;:=\; \mathbf{H}_N+\mathbf{A}_N, \tag{3.3}$$

We set

$$G_\mu(z) \;:=\; \int \frac{1}{z-x}\mu(dx). \tag{3.4}$$

the Cauchy transform of the measure μ. We introduce, for all $i\in\{1,\ldots,j\}$, the finite, possibly empty, set

$$\mathscr{S}_{\theta_i} \;:=\; \left\{\xi\in\mathbb{C}\setminus\operatorname{supp}(\mu),\; G_\mu(\xi)=\frac{1}{\theta_i}\right\}, \quad\text{and}\quad m_i \;:=\; \operatorname{Card}\mathscr{S}_{\theta_i} \tag{3.5}$$

We make the following assumption

Assumption 3.2.2. For any $\delta>0$, as N goes to infinity, we have

$$\sup_{\operatorname{dist}(z,\operatorname{supp}(\mu))>\delta}\left\|\mathbf{U}_{2r}^*\left(z\mathbf{I}-\mathbf{H}_N\right)^{-1}\mathbf{U}_{2r}-G_\mu(z)\mathbf{I}\right\|_{\mathrm{op}} \xrightarrow{(\mathbb{P})} 0.$$

Result

Theorem 3.2.3 (Convergence of the outliers)**.** *For $\theta_1,\ldots,\theta_j$, $k_1,\ldots,k_j$, $\mathscr{S}_{\theta_1},\ldots,\mathscr{S}_{\theta_j}$ and $m_1,\ldots,m_j$ as defined above, with probability tending to one, $\widetilde{\mathbf{H}}_N:=\mathbf{H}_N+\mathbf{A}_N$ possesses exactly $\displaystyle\sum_{i=1}^{j}k_im_i$ eigenvalues at a macroscopic distance of* $\operatorname{supp}\mu$ *(outliers). More precisely, for all small enough $\delta>0$, for all large enough N, for all $i\in\{1,\ldots,j\}$, if we set*

$$\mathscr{S}_{\theta_i} \;=\; \{\xi_{i,1},\ldots,\xi_{i,m_i}\},$$

there are m_i eigenvalues $\widetilde{\lambda}_{i,1},\ldots,\widetilde{\lambda}_{i,m_i}$ of $\widetilde{\mathbf{H}}_N$ in $\{z,\ \operatorname{dist}(z,\operatorname{supp}(\mu))>\delta\}$ satisfying

$$\widetilde{\lambda}_{i,n} \;=\; \xi_{i,n}+o(1), \quad \textit{for all } n\in\{1,\ldots,m_i\},$$

after a proper labeling.

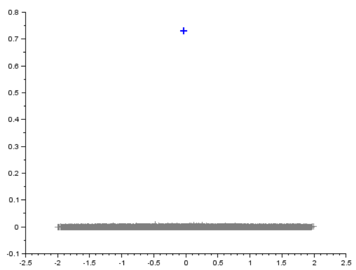
[10]The *multiplicity* of an eigenvalue is defined as its order as a root of the characteristic polynomial, which is greater than or equal to the dimension of the associated eigenspace.

Remark 3.2.4. If all the $\mathscr{S}_{\theta_i}$'s are empty, there is possibly no outlier at all. This condition is the analogous of the *phase transition* condition in [19, Theorem 2.1] in the case where the θ_i's are real, which is if

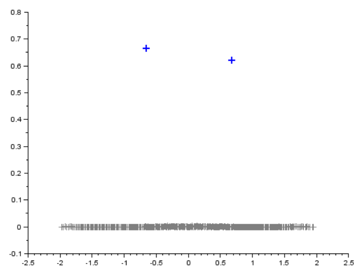
$$\frac{1}{\theta_i} \notin \left] \lim_{x \to a^-} G_\mu(x), \lim_{x \to b^+} G_\mu(x) \right[$$

where a (resp. b) designates the infimum (resp. the supremum) of the support of μ, then, θ_i does not generate any outlier. In our case, if $|\theta_i|$ is large enough, $\mathscr{S}_{\theta_i}$ is necessarily non-empty[11], which means that a strong enough perturbation always creates outliers.

Remark 3.2.5. We notice that the outliers can outnumber the rank of $\mathbf{A}$. This phenomenon was already observed in [19] in the case where the support of the limit spectral distribution has a disconnected support (see also [12]). In our case, the phenomenon occurs even for connected support (see Figure 3.9(b)).



(a) Spectrum of an Hermitian matrix of size $N = 2000$, whose spectral measure tends to the *semi-circle law* $\mu_{sc}(dx) := \frac{1}{2\pi}\sqrt{4-x^2}\mathbb{1}[-2,2](dx)$ (such as Wigner matrix), with perturbation matrix $\mathbf{A} = \mathrm{diag}(i\sqrt{2}, 0\ldots, 0)$.



(b) Spectrum of an Hermitian matrix of size $N = 2000$, whose spectral measure tends to $\frac{2}{5}\delta_{-1}(dx) + \frac{2}{5}\delta_1(dx) + \frac{1}{5}\mu_{sc}(dx)$ and with perturbation matrix $\mathbf{A} = \mathrm{diag}(i\sqrt{2}, 0\ldots, 0)$.

Figure 3.9: Spectrums of two Hermitian matrices with the same limit bulk but different limit spectral densities on this bulk, perturbed by the same matrix: both do not have the same number of outliers (the blue crosses "+").

3.2.2 Fluctuations of the outliers

To study the fluctuations, one needs to understand the limit distribution of

$$\sqrt{N} \left\| \mathbf{U}_{2r}^* (z\mathbf{I} - \mathbf{H}_N)^{-1} \mathbf{U}_{2r} - G_\mu(z)\mathbf{I} \right\|_{\mathrm{op}}. \tag{3.6}$$

[11] due to the fact that the Cauchy transform of a compactly supported measure can always be inverted in a neighborhood of infinity.

In the particular case where $\mathbf{H}_N$ is a Wigner matrix, we know from [79] that this quantity is tight but does not necessarily converge. Hence, we shall need additional assumptions.

Set up and assumptions

As $\mathbf{A}_N$ is not Hermitian, we need to introduce the Jordan Canonical Form (JCF) to describe the fluctuations. More precisely, we shall consider the JCF of $\mathbf{A}_0$ which does not depend on N. We know that, in a proper basis, $\mathbf{A}_0$ is a direct sum of *Jordan blocks*, i.e. blocks of the form

$$\mathbf{R}_p(\theta) \quad = \quad \begin{pmatrix} \theta & 1 & & (0) \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ (0) & & & \theta \end{pmatrix}, \; p\times p \text{ matrix}, \qquad \theta\in\mathbb{C}, p\geq 1 \quad (3.7)$$

Let us denote by $\theta_1,\dots,\theta_q$ the distinct eigenvalues of $\mathbf{A}_0$ such that $\mathscr{S}_\theta \neq \emptyset$ (see (3.5) for the definition of $\mathscr{S}_\theta$), and for each $i=1,\dots,q$, we introduce a positive integer α_i, some positive integers $p_{i,1}>\dots>p_{i,\alpha_i}$ corresponding to the distinct sizes of the blocks relative to the eigenvalue θ_i and $\beta_{i,1},\dots,\beta_{i,\alpha_i}$ such that for all j, $\mathbf{R}_{p_{i,j}}(\theta_i)$ appears $\beta_{i,j}$ times, so that, for a certain $2r\times 2r$ non singular matrix $\mathbf{Q}$, we have:

$$\mathbf{J} \quad = \quad \mathbf{Q}^{-1}\mathbf{A}_0\mathbf{Q} \; = \; \hat{\mathbf{A}} \bigoplus \bigoplus_{i=1}^{q} \bigoplus_{j=1}^{\alpha_i} \underbrace{\begin{pmatrix} \mathbf{R}_{p_{i,j}}(\theta_i) & & \\ & \ddots & \\ & & \mathbf{R}_{p_{i,j}}(\theta_i) \end{pmatrix}}_{\beta_{i,j}\text{ blocks}} \qquad (3.8)$$

where $\oplus$ is defined, for square block matrices, by $\mathbf{M}\oplus\mathbf{N} := \begin{pmatrix} \mathbf{M} & 0 \\ 0 & \mathbf{N} \end{pmatrix}$ and $\hat{\mathbf{A}}$ is a matrix such that its eigenvalues θ are such that $\mathscr{S}_\theta = \emptyset$ or null.

The asymptotic orders of the fluctuations of the eigenvalues of $\mathbf{H}_N+\mathbf{A}_N$ depend on the sizes $p_{i,j}$ of the blocks. Actually, for each θ_i and each $\xi_{i,n}\in\mathscr{S}_{\theta_i} = \{\xi_{i,1},\dots,\xi_{i,m_i}\}$, we know, by Theorem 3.2.3, there are $\sum_{j=1}^{\alpha_i} p_{ij}\times\beta_{i,j}$ eigenvalues $\widetilde{\lambda}$ of $\mathbf{H}_N+\mathbf{A}_N$ which tend to $\xi_{i,n}$: we shall write them with a $\xi_{i,n}$ on the top left corner, as follows

$${}^{\xi_{i,n}}\widetilde{\lambda}.$$

Theorem 3.2.10 below will state that for each block with size $p_{i,j}$ corresponding to θ_i in the JCF of $\mathbf{A}_0$, there are $p_{i,j}$ eigenvalues (we shall write them with $p_{i,j}$ on the bottom left corner : ${}^{\xi_{i,n}}_{p_{i,j}}\widetilde{\lambda}$) whose convergence rate will be $N^{-1/(2p_{i,j})}$. As there are $\beta_{i,j}$ blocks of size $p_{i,j}$, there are actually $p_{i,j}\times\beta_{i,j}$ eigenvalues tending to $\xi_{i,n}$ with convergence rate $N^{-1/(2p_{i,j})}$ (we shall write them ${}^{\xi_{i,n}}_{p_{i,j}}\widetilde{\lambda}_{s,t}$ with $s\in\{1,\dots,p_{i,j}\}$ and $t\in\{1,\dots,\beta_{i,j}\}$). It would be convenient to denote by $\Lambda_{i,j,n}$ the vector with size $p_{i,j}\times\beta_{i,j}$ defined by

$$\Lambda_{i,j,n} \;:=\; \left(N^{1/(2p_{i,j})}\cdot\left({}^{\xi_{i,n}}_{p_{i,j}}\widetilde{\lambda}_{s,t}-\xi_{i,n}\right)\right)_{\substack{1\le s\le p_{i,j}\\ 1\le t\le \beta_{i,j}}}. \tag{3.9}$$

In addition, we make an assumption on the convergence of (3.6).

Assumption 3.2.6.

(1) The vector $\left(\sqrt{N}\mathbf{U}_{2r}^*\left((\xi_{i,n}-\mathbf{H}_N)^{-1}-\frac{1}{\theta_i}\right)\mathbf{U}_{2r}\right)_{\substack{1\le i\le q\\ 1\le n\le m_i}}$ converges in distribution and none of its entries tends to zero.

(2) For all $k\ge 1$, all $i\in\{1,\dots,q\}$ and all $n\in\{1,\dots,m_i\}$,

$$\sqrt{N}\mathbf{U}_{2r}^*\left(\left(\xi_{i,n}-\mathbf{H}_N\right)^{-(k+1)}-\int\frac{\mu(dx)}{(\xi_{i,n}-x)^{k+1}}\right)\mathbf{U}_{2r}$$

is tight.

or

(0') For all $i\in\{1,\dots,q\}$ and all $n\in\{1,\dots,m_i\}$, as N goes to infinity,

$$\sqrt{N}\left(\frac{1}{N}\operatorname{Tr}\left(\xi_{i,n}-\mathbf{H}_N\right)^{-1}-\frac{1}{\theta_i})\right)\longrightarrow 0.$$

(1') The vector $\left(\sqrt{N}\mathbf{U}_{2r}^*\left((\xi_{i,n}-\mathbf{H}_N)^{-1}-\frac{1}{N}\operatorname{Tr}\left(\xi_{i,n}-\mathbf{H}_N\right)^{-1}\right)\mathbf{U}_{2r}\right)_{\substack{1\le i\le q\\ 1\le n\le m_i}}$ converges in distribution and none of its entries tends to zero.

(2') For all $k\ge 1$ and for all $i\in\{1,\dots,q\}$,

$$\sqrt{N}\mathbf{U}_{2r}^*\left(\left(\xi_{i,n}-\mathbf{H}_N\right)^{-(k+1)}-\frac{1}{N}\operatorname{Tr}\left(\xi_{i,n}-\mathbf{H}_N\right)^{-(k+1)}\right)\mathbf{U}_{2r}$$

is tight.

As in [21], we define now the family of random matrices that we shall use to characterize the limit distribution of the $\Lambda_{i,j,n}$'s. For each $i=1,\dots,q$, let $I(\theta_i)$ (resp. $J(\theta_i)$) denote the set, with cardinality $\sum_{j=1}^{\alpha_i}\beta_{i,j}$, of indices in $\{1,\dots,r\}$ corresponding to the first (resp. last) columns of the blocks $\mathbf{R}_{p_{i,j}}(\theta_i)$ $(1\le j\le\alpha_i)$ in (3.8).

Remark 3.2.7. Note that the columns of $\mathbf{Q}$ (resp. of $(\mathbf{Q}^{-1})^*$) whose index belongs to $I(\theta_i)$ (resp. $J(\theta_i)$) are eigenvectors of $\mathbf{A}_0$ (resp. of $\mathbf{A}_0^*$) associated to θ_i (resp. $\overline{\theta_i}$). See [21, Remark 2.7].

Now, let

$$\left(m_{k,\ell}^{\theta_i,n}\right)_{\substack{1\leq i\leq q\\ 1\leq n\leq m_i\\ (k,\ell)\in J(\theta_i)\times I(\theta_i)}} \tag{3.10}$$

be the multivariate random variable defined as the limit joint-distribution of

$$\left(\sqrt{N}\mathbf{e}_k^*\mathbf{Q}^{-1}\mathbf{U}_{2r}^*\left(\left(\xi_{i,n}-\mathbf{H}_N\right)^{-1}-\frac{1}{\theta_i}\right)\mathbf{U}_{2r}\mathbf{Q}\mathbf{e}_\ell\right)_{\substack{1\leq i\leq q\\ 1\leq n\leq m_i\\ (k,\ell)\in J(\theta_i)\times I(\theta_i)}} \xrightarrow[\text{jointly}]{\text{(d)}} \left(m_{k,\ell}^{\theta_i,n}\right)_{\substack{1\leq i\leq q\\ 1\leq n\leq m_i\\ (k,\ell)\in J(\theta_i)\times I(\theta_i)}} \tag{3.11}$$

(which does exist by Assumption 3.2.6) and where $\mathbf{e}_1,\ldots,\mathbf{e}_r$ are the column vectors of the canonical basis of $\mathbb{C}^r$).

For each i,j, let $K(i,j)$ (resp. $K(i,j)^-$) be the set, with cardinality $\beta_{i,j}$ (resp. $\sum_{j'=1}^{j-1}\beta_{i,j'}$), of indices in $J(\theta_i)$ corresponding to a block of the type $\mathbf{R}_{p_{i,j}}(\theta_i)$ (resp. to a block of the type $\mathbf{R}_{p_{i,j'}}(\theta_i)$ for $j'<j$). In the same way, let $L(i,j)$ (resp. $L(i,j)^-$) be the set, with the same cardinality as $K(i,j)$ (resp. as $K(i,j)^-$), of indices in $I(\theta_i)$ corresponding to a block of the type $\mathbf{R}_{p_{i,j}}(\theta_i)$ (resp. to a block of the type $\mathbf{R}_{p_{i,j'}}(\theta_i)$ for $j'<j$). Note that $K(i,j)^-$ and $L(i,j)^-$ are empty if $j=1$. Let us define the random matrices for each $n\in\{1,\ldots,m_i\}$

$$\begin{aligned} \mathrm{M}_{j,n}^{\theta_i,\mathrm{I}} &:= [m_{k,\ell}^{\theta_i,n}]_{\substack{k\in K(i,j)^-\\ \ell\in L(i,j)^-}} & \mathrm{M}_{j,n}^{\theta_i,\mathrm{II}} &:= [m_{k,\ell}^{\theta_i,n}]_{\substack{k\in K(i,j)^-\\ \ell\in L(i,j)}} \\ \mathrm{M}_{j,n}^{\theta_i,\mathrm{III}} &:= [m_{k,\ell}^{\theta_i,n}]_{\substack{k\in K(i,j)\\ \ell\in L(i,j)^-}} & \mathrm{M}_{j,n}^{\theta_i,\mathrm{IV}} &:= [m_{k,\ell}^{\theta_i,n}]_{\substack{k\in K(i,j)\\ \ell\in L(i,j)}} \end{aligned} \tag{3.12}$$

and then let us define the $\beta_{i,j}\times\beta_{i,j}$ matrix $\mathbf{M}_{j,n}^{\theta_i}$ as

$$\mathbf{M}_{j,n}^{\theta_i} := \theta_i\left(\mathrm{M}_{j,n}^{\theta_i,\mathrm{IV}}-\mathrm{M}_{j,n}^{\theta_i,\mathrm{III}}\left(\mathrm{M}_{j,n}^{\theta_i,\mathrm{I}}\right)^{-1}\mathrm{M}_{j,n}^{\theta_i,\mathrm{II}}\right) \tag{3.13}$$

Remark 3.2.8. It follows from the fact that the matrix $\mathbf{Q}$ is invertible, that $\mathrm{M}_{j,n}^{\theta_i,\mathrm{I}}$ is a.s. invertible and so is $\mathbf{M}_{j,n}^{\theta_i}$.

Remark 3.2.9. In the particular case where $\mathbf{A}_0$ is Hermitian (which means that $\mathbf{Q}^{-1}=\mathbf{Q}^*$ and the θ_i's are real), then the matrces $\mathbf{M}_{j,n}^{\theta_i}$ are also Hermitian.

Now, we can formulate the result on the fluctuations.

Result

Theorem 3.2.10. *1. As N goes to infinity, the random vector*

$$\left(\Lambda_{i,j,n}\right)_{\substack{1\leq i\leq q\\ 1\leq j\leq\alpha_i\\ 1\leq n\leq m_i}}$$

defined at (3.9) *converges to the distribution of a random vector*

$$\left(\Lambda^{\infty}_{i,j,n}\right)_{\substack{1\le i\le q\\ 1\le j\le \alpha_i\\ 1\le n\le m_i}}$$

with joint distribution defined by the fact that for each $1\le i\le q$, $1\le j\le \alpha_i$ *and* $1\le n\le m_i$, $\Lambda^{\infty}_{i,j,n}$ *is the collection of the* ${p_{i,j}}^{th}$ *roots of the eigenvalues of some random matrix* $\mathbf{M}^{\theta_i}_{j,n}$.

2. *The distributions of the random matrices* $\mathbf{M}^{\theta_i}_{j,n}$ *are absolutely continuous with respect to the Lebesgue measure and the random vector* $\left(\Lambda^{\infty}_{i,j,n}\right)_{\substack{1\le i\le q\\ 1\le j\le \alpha_i}}$ *has no deterministic coordinate.*

Theorem 3.2.10 is illustrated in Figure 3.10 with an example. We clearly see appearing regular polygons.

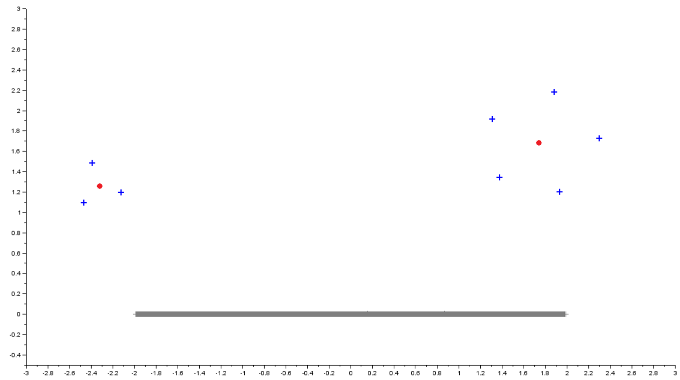


Figure 3.10: Spectrum of a Wigner matrix of size $N = 5000$ with perturbation matrix $\mathbf{A} = \mathrm{diag}\left(\mathbf{R}_5(1.5+2\mathrm{i}), \mathbf{R}_3(-2+1.5\mathrm{i}), 0, \ldots, 0\right)$. We see the blue crosses "+" (outliers) forming respectively a regular pentagon and an equilateral triangle around the red dots "•" (their limit). We also see a significant difference between the two rates of convergence, $N^{-1/10}$ and $N^{-1/6}$.

3.3 Applications

In this section, we give examples of random matrices which satisfy the assumptions of Theorem 3.2.3 and Theorem 3.2.10.

3.3.1 Wigner matrices

Let $\mathbf{H}_N = \frac{1}{\sqrt{N}}\mathbf{W}_N$ be a symmetric/Hermitian Wigner matrix with independent entries up to the symmetry. More precisely, we assume that

Assumption 3.3.1.
Real symmetric case :

- $\left(\mathbf{W}_N\right)_{i,j}, 1 \leq i \leq j \leq N,$ are independent,
- The $(\mathbf{W}_N)_{i,j}$'s for $i \neq j$ (resp. $i = j$), are identically distributed,
- $\mathbb{E}(\mathbf{W}_N)_{1,1} = \mathbb{E}(\mathbf{W}_N)_{1,2} = 0,\ \mathbb{E}(\mathbf{W}_N)_{1,1}^2 = 2\sigma^2,\ \mathbb{E}(\mathbf{W}_N)_{1,2}^2 = \sigma^2,$
- $c_3 := \mathbb{E}\left|(\mathbf{W}_N)_{1,1}\right|^3 < \infty,\ m_5 := \mathbb{E}\left|(\mathbf{W}_N)_{1,2}\right|^5 < \infty.$

Hermitian case :

- $\left(\operatorname{Re}\mathbf{W}_N\right)_{i,j}, \left(\operatorname{Im}\mathbf{W}_N\right)_{i,j}, 1 \leq i < j \leq N, \left(\mathbf{W}_N\right)_{i,i}, 1 \leq i \leq N,$ are independent.
- The $(\operatorname{Re}\mathbf{W}_N)_{i,j}$'s, $(\operatorname{Im}\mathbf{W}_N)_{i,j}$'s for $i \neq j$ (resp. $(\mathbf{W}_N)_{i,i}$'s), are identically distributed,
- $\mathbb{E}(\mathbf{W}_N)_{1,1} = \mathbb{E}(\mathbf{W}_N)_{1,2} = 0,\ \mathbb{E}(\mathbf{W}_N)_{1,1}^2 = \sigma^2,\ \mathbb{E}(\operatorname{Re}\mathbf{W}_N)_{1,2}^2 = \dfrac{\sigma^2}{2},$
- $c_3 := \mathbb{E}\left|(\mathbf{W}_N)_{1,1}\right|^3 < \infty,\ m_5 := \mathbb{E}\left|(\mathbf{W}_N)_{1,2}\right|^5 < \infty.$

In this case, we have the following version of Theorem 3.2.3

Theorem 3.3.2 (Convergence of the outliers for Wigner matrices)**.**
Let $\theta_1, \ldots, \theta_j$ be the eigenvalues of $\mathbf{A}_N$ such that $|\theta_i| > \sigma$. Then, with probability tending to one, for all large enough N, there are exactly j eigenvalues $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_j$ of $\widetilde{\mathbf{H}}_N := \frac{1}{\sqrt{N}}\mathbf{W}_N + \mathbf{A}_N$ at a macroscopic distance of $[-2\sigma, 2\sigma]$ (outliers). More precisely, for all small enough $\delta > 0$, for all large enough N, for all $i \in \{1, \ldots, j\}$,

$$\widetilde{\lambda}_i \quad = \quad \theta_i + \frac{\sigma^2}{\theta_i} + o(1),$$

after a proper labeling.

Proof. We just need to check that Assumptions 3.2.1 and 3.2.2 are satisfied.

- As long as the entries of $\mathbf{W}_N$ have a finite fourth moment, we know (see [6, Theorem 5.2]) that Assumption 3.2.1 is satisfied.

- Now, we need to show that for any $\delta > 0$, as N goes to infinity,

$$\sup_{\operatorname{dist}(z,\operatorname{supp}(\mu))>\delta} \left\|\mathbf{U}_{2r}^*\left(z\mathbf{I} - \mathbf{H}_N\right)^{-1}\mathbf{U}_{2r} - G_{\mu_{sc}}(z)\mathbf{I}\right\|_{\mathrm{op}} \xrightarrow{(\mathbb{P})} 0.$$

  Since we are dealing with $2r \times 2r$ sized matrices, it suffices to prove that for any unite vectors $\mathbf{u}$,$\mathbf{v}$ of $\mathbb{C}^N$, for any $\delta > 0$ and any $\eta > 0$, as N goes to infinity,

$$\mathbb{P}\Big(\sup_{\operatorname{dist}(z,\operatorname{supp}(\mu))>\delta} \left|\mathbf{u}^*\big(\left(z\mathbf{I} - \mathbf{H}_N\right)^{-1} - G_{\mu_{sc}}(z)\mathbf{I}\big)\mathbf{v}\right| > \eta\Big) \longrightarrow 0.$$

Moreover, as both $G_{\mu_{sc}}(z)$ and $\left\|(z\mathbf{I}-\mathbf{H}_N)^{-1}\right\|_{\mathrm{op}}$ goes to 0 when $|z|$ goes to infinity, we know there is a large enough constant M such that we just need to prove that

$$\mathbb{P}\Big(\sup_{\substack{\mathrm{dist}(z,\mathrm{supp}(\mu))>\delta\\ |z|\le M}}\left|\mathbf{u}^*\big((z\mathbf{I}-\mathbf{H}_N)^{-1}-G_{\mu_{sc}}(z)\mathbf{I}\big)\mathbf{v}\right|>\eta\Big)\longrightarrow 0.$$

Then, for any $\eta'>0$, the compact set $K=\{z,\ \mathrm{dist}(z,\mathrm{supp}(\mu))>\delta$ and $|z|\le M\}$ admits a η'-net, which is a finite set $\{z_1,\dots,z_p\}$ of K such that

$$\forall z\in K,\ \exists i\in\{1,\dots,p\},\quad |z-z_i|<\eta',$$

so that, using the uniform boundedness of the derivative of $G_{\mu_{sc}}(z)$ and $\mathbf{u}^*(z-\mathbf{H}_N)^{-1}\mathbf{v}$ on K, for a small enough η', we just need to prove that

$$\mathbb{P}\Big(\max_{i=1}^{p}\left|\mathbf{u}^*\big((z_i\mathbf{I}-\mathbf{H}_N)^{-1}-G_{\mu_{sc}}(z_i)\mathbf{I}\big)\mathbf{v}\right|>\eta/2\Big)\longrightarrow 0.$$

Then, we properly decompose each function $x\mapsto\frac{1}{z_i-x}$ as a sum of a smooth compactly supported function and one that vanishes on a neighborhood of $[-2\sigma,2\sigma]$ and conclude using [79, (ii) Theorem 1.6]

Moreover, in the Wigner case, we have

$$G_{\mu_{sc}}(z) = \frac{z-\sqrt{z^2-4\sigma^2}}{2\sigma^2},$$

where $\sqrt{z^2-4\sigma^2}$ is the branch of the square root with branch cut $[-2\sigma,2\sigma]$ so that for any z outside $[-2\sigma,2\sigma]$, the equation $G_{\mu_{sc}}(z)=\frac{1}{\theta}$ possesses one solution if and only if $|\theta|>\sigma$ and the unique solution is

$$\theta+\frac{\sigma^2}{\theta},$$

which means that in the Wigner case, the outliers cannot outnumber the rank of the perturbation, and the *phase transition* condition is simply : $|\theta|>\sigma$. Actually in [12] (see Remark 3.2), the authors explain that if μ is $\boxplus$-infinitely divisible, then the sets $\mathscr{S}_{\theta_i}$'s have at most one element, which means that for Wigner matrices, it is not possible to observe the phenomenon of "outliers outnumber the rank of $\mathbf{A}$".

Remark 3.3.3. One can find an other proof of Theorem 3.3.2 in [**?**] as a particular case of the Theorem 2.4 (see [**?**, Remark 2.5]) due to the fact that a Wigner matrix can be seen as a particular Elliptic matrix. Nevertheless, the authors of [**?**] don't deal with the matter of the fluctuations.

□

To study the fluctuations of the outliers in the Wigner case, we must make an additional assumption on the perturbation $\mathbf{A}_N$.

Assumption 3.3.4. The matrix $\mathbf{A}_N$ has only a finite number (independent of N) of entries which are non-zero.

Remark 3.3.5. Assumption 3.3.4 is equivalent to suppose that $\mathbf{U}_{2r}$ (the $2r$-first columns of $\mathbf{U}$), possesses only a finite number K (independent of N) of non-zero rows. Actually, this assumption is the analogous "the eigenvectors of $\mathbf{A}$ don't spread out" hypothesis corresponding to the "case a)" in [**?**].

Remark 3.3.6. If $\mathbf{U}$ is Haar-distributed and independent from $\mathbf{W}$, we can avoid making Assumption 3.3.4 (see section 3.3.2). One can also slightly weaken Assumption 3.3.4 by assuming that the $2r$-first rows of $\mathbf{U}$ correspond to the N first coordinates of a collection of non-random vectors $\mathbf{u}_1,\ldots,\mathbf{u}_{2r}$ in $\ell^2(\mathbb{N})$ (see [79, Theorem 1.7]).

Theorem 3.3.7 (Fluctuations for Wigner matrices)**.**
With Assumtions 3.3.1 and 3.3.4, Theorem 3.2.10 holds. Moreover, the distribution of the random vector

$$\left(m_{k,\ell}^{\theta_i}\right)_{\substack{1\le i\le q\\(k,\ell)\in J(\theta_i)\times I(\theta_i)}},$$

defined by (3.10)*, is*

$$\left(\mathbf{e}_k^*\mathbf{Q}^{-1}\mathbf{U}_{K,2r}^*\Upsilon(\xi_i)\mathbf{U}_{K,2r}\mathbf{Q}\mathbf{e}_\ell\right)_{\substack{1\le i\le q\\(k,\ell)\in J(\theta_i)\times I(\theta_i)}},$$

where $\xi_i:=\theta_i+\dfrac{\sigma^2}{\theta_i}$ *and where* $\Upsilon(z)$ *is a* $K\times K$ *random field defined by*

$$\Upsilon(z) \quad := \quad (G_{\mu_{sc}}(z))^2\left(\mathbf{W}^{(K)}+\mathbf{Y}(z)\right) \tag{3.14}$$

where $\mathbf{W}^{(K)}$ *is the* $K\times K$ *upper-left corner submatrix of a matrix* $\widetilde{\mathbf{W}}_N$ *such that* $\widetilde{\mathbf{W}}_N\overset{(\mathrm{d})}{=}\mathbf{W}_N$ *and* $\mathbf{Y}(z)$ *is a* $K\times K$ *Gaussian random field defined by [80, (2.7),(2.8),(2.9),(2.10),(2.11),(2.12)] in the real case and [80, (2.42),(2.43),(2.44),(2.45),(2.46),(2.47)] in the complex case.*

Remark 3.3.8. This provides an example of non universal fluctuations, in the sense that the $\left(m_{k,\ell}^{\theta_i}\right)$'s are not necessarily Gaussian. However, when $\mathbf{H}_N$ is a GOE or GUE matrix, the $\left(m_{k,\ell}^{\theta_i}\right)$'s are centered Gaussian variables such that

$$\begin{aligned}
\mathbb{E}\left(m_{k,\ell}^{\theta_i}\,m_{k',\ell'}^{\theta_{i'}}\right) &= \psi_{sc}(\xi_i,\xi_{i'})\left(\mathbf{e}_k^*\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^*\mathbf{e}_{k'}\;\mathbf{e}_\ell^*\mathbf{Q}^*\mathbf{Q}\mathbf{e}_{\ell'}\;+\;\delta_{k,\ell'}\delta_{k',\ell}\right), \\
\mathbb{E}\left(m_{k,\ell}^{\theta_i}\,\overline{m_{k',\ell'}^{\theta_{i'}}}\right) &= \psi_{sc}(\xi_i,\overline{\xi_{i'}})\left(\mathbf{e}_k^*\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^*\mathbf{e}_{k'}\;\mathbf{e}_\ell^*\mathbf{Q}^*\mathbf{Q}\mathbf{e}_{k'}\;+\;\delta_{k,\ell'}\delta_{k',\ell}\right),
\end{aligned} \tag{3.15}$$

for the GOE, and

$$\begin{aligned}
\mathbb{E}\left(m_{k,\ell}^{\theta_i}\,m_{k',\ell'}^{\theta_{i'}}\right) &= \psi_{sc}(\xi_i,\xi_{i'})\,\delta_{k,\ell'}\delta_{k',\ell}, \\
\mathbb{E}\left(m_{k,\ell}^{\theta_i}\,\overline{m_{k',\ell'}^{\theta_{i'}}}\right) &= \psi_{sc}(\xi_i,\overline{\xi_{i'}})\,\mathbf{e}_k^*\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^*\mathbf{e}_{k'}\;\mathbf{e}_{\ell'}^*\mathbf{Q}^*\mathbf{Q}\,\mathbf{e}_\ell,
\end{aligned} \tag{3.16}$$

for the GUE, where

$$\psi_{sc}(z,w) \;:=\; G^2_{\mu_{sc}}(z)G^2_{\mu_{sc}}(w)\big(\sigma^2+\sigma^4\varphi_{sc}(z,w)\big),$$
$$\varphi_{sc}(z,w) \;:=\; \int \frac{1}{z-x}\frac{1}{w-x}\mu_{sc}(dx).$$

We notice that, if $\mathbf{Q}^{-1} \neq \mathbf{Q}^*$, then we might observe correlations between the fluctuations of outliers at a macroscopic distance with each other. This phenomenon has already been observed in [62] for non-Gaussian Wigner matrices whereas, here, the phenomenon may still occur for GUE matrices. Actually, (3.15) and (3.16) can be simplified due to the fact

$$\sigma^2 G^2_{\mu_{sc}}(z) - zG_{\mu_{sc}}(z) + 1 \;=\; 0,$$

so that $\varphi_{sc}(z,w) = -\dfrac{G_{\mu_{sc}}(z)-G_{\mu_{sc}}(w)}{z-w}$ satisfies

$$\sigma^2 G_{\mu_{sc}}(z)G_{\mu_{sc}}(w)\varphi_{sc}(z,w) \;=\; \varphi_{sc}(z,w) - G_{\mu_{sc}}(z)G_{\mu_{sc}}(w). \tag{3.17}$$

Hence,

$$\begin{aligned}
& G^2_{\mu_{sc}}(\xi_i)G^2_{\mu_{sc}}(\xi_{i'})\big(\sigma^2+\sigma^4\varphi_{sc}(\xi_i,\xi_{i'})\big)\\
=\;& \sigma^2 G_{\mu_{sc}}(\xi_i)G_{\mu_{sc}}(\xi_{i'})\Big[G_{\mu_{sc}}(\xi_i)G_{\mu_{sc}}(\xi_{i'}) + \sigma^2 G_{\mu_{sc}}(\xi_i)G_{\mu_{sc}}(\xi_{i'})\varphi_{sc}(\xi_i,\xi_{i'})\Big]\\
=\;& \sigma^2 G_{\mu_{sc}}(\xi_i)G_{\mu_{sc}}(\xi_{i'})\varphi_{sc}(\xi_i,\xi_{i'})\\
=\;& \varphi_{sc}(\xi_i,\xi_{i'}) - G_{\mu_{sc}}(\xi_i)G_{\mu_{sc}}(\xi_{i'}) \;=\; \Phi_{sc}(\xi_i,\xi_{i'}).
\end{aligned}$$

and we fall back on the expression of the variance for the UCI model (see section 3.3.2), which is expected since the GUE belongs to the UCI model.

Proof. We show that the assumptions 3.3.1 and 3.3.4 imply Assumption 3.2.6, more precisely (1) and (2). For (1), we simply use [80, Theorem 2.1/2.5] to show that

$$\sqrt{N}\mathbf{U}^*_{2r}\left((z-\mathbf{H}_N)^{-1} - G_{\mu_{sc}}(z)\mathbf{I}\right)\mathbf{U}_{2r},$$

converges weakly (as it is done in [79]). The limit distribution is also given by [80, Theorem 2.1/2.5].
Then for (2), we know by [80, (i) of Theorem 2.3/2.7] (respectively [80, (iii) of Proposition 2.1]) that, for all $k \geq 1$, the diagonal entries (respectively the off-diagonal entries) of the matrix

$$\sqrt{N}\big((z-\mathbf{H}_N)^{-k-1} - \int (z-x)^{-k-1}\mu_{sc}(dx)\mathbf{I}\big)$$

converge in distribution so that

$$\sqrt{N}\mathbf{U}^*_{2r}\left((z-\mathbf{H}_N)^{-k-1} - \int (z-x)^{-k-1}\mu_{sc}(dx)\mathbf{I}\right)\mathbf{U}_{2r}$$

is tight. □

3.3.2 Hermitian matrices whose distribution is invariant by unitary conjugation

Let $\mathbf{H}_N$ be an Hermitian matrix such that for any unitary $N \times N$ matrix $\mathbf{U}_N$, we have

$$\mathbf{U}_N \mathbf{H}_N \mathbf{U}_N^* \overset{(\mathrm{d})}{=} \mathbf{H}_N. \tag{3.18}$$

$\mathbf{H}_N$ can be written $\mathbf{H}_N = \mathbf{U}_N \mathbf{D}_N \mathbf{U}_N^*$ where $\mathbf{D}_N$ is diagonal, $\mathbf{U}_N$ is Haar-distributed and $\mathbf{U}_N$ and $\mathbf{D}_N$ are independent. We also assume that $\mathbf{H}_N$ satisfies (3.1) and Assumption 3.2.1. We shall call such matrices *UCI matrices* (for Unitary Conjugation Invariance). In this case, as we can we can write

$$\widetilde{\mathbf{H}}_N = \mathbf{H}_N + \mathbf{A}_N = \mathbf{U}_N (\mathbf{D}_N + \mathbf{U}_N^* \mathbf{A}_N \mathbf{U}_N) \mathbf{U}_N^*,$$

so that, without any loss of generality, we can simply assume that $\mathbf{H}_N$ is a diagonal matrix and $\mathbf{A}_N$ is a matrix of the form

$$\mathbf{A}_N = \mathbf{U}_{2r} \mathbf{A}_0 \mathbf{U}_{2r}^*$$

where $\mathbf{U}_{2r}$ is the $2r$-first columns of an Haar-distributed matrix independent from $\mathbf{H}_N$.

Theorem 3.3.9 (Convergence of the outliers for UCI matrices)**.**
If $\mathbf{H}_N$ is an UCI matrix, then Theorem 3.2.3 holds.

Remark 3.3.10. Unlike the Wigner case, Theorem 3.2.3 does not need to be reformulated. In this case, we do observe the phenomenon of the outliers outnumbering the rank of $\mathbf{A}_N$.

Proof. We just need to check that Assumption 3.2.2 is satisfied. To do so, one can apply a slightly modified version of [16, Lemma 2.2], where we replace all the "$\mathrm{dist}(z,[a,b]) > \delta$" by "$\mathrm{dist}(z,\mathrm{supp}(\mu)) > \delta$", which does not change the ideas of the proof. □

For the fluctuations, we need to assume that for all $i \in \{1,\ldots,q\}$ and all $n \in \{1,\ldots,m_i\}$, as N goes to infinity,

$$\sqrt{N}\left(\frac{1}{N}\mathrm{Tr}\,(\xi_{i,n} - \mathbf{H}_N)^{-1} - \frac{1}{\theta_i})\right) \longrightarrow 0. \tag{3.19}$$

Remark 3.3.11. Actually, in [16], the authors make the same assumption ([16, Hypothesis 3.1]).

Theorem 3.3.12 (Fluctuations for UCI matrices)**.**
If $\mathbf{H}_N$ is an UCI matrix, then if satisfies Theorem 3.2.10. More precisely, the

$$\left(m_{k,\ell}^{\theta_i,n}\right)_{\substack{1\le i\le q \\ 1\le n\le m_i \\ (k,\ell)\in J(\theta_i)\times I(\theta_i)}},$$

defined by (3.10) *are centered Gaussian variables such that*

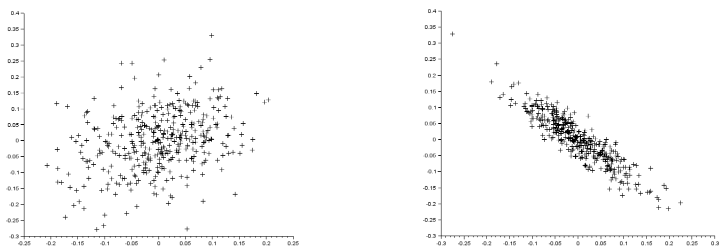
$$\begin{aligned}
\mathbb{E}\left(m_{k,\ell}^{\theta_i,n}\, m_{k',\ell'}^{\theta_{i'},n'}\right) &= \Phi(\xi_{i,n},\xi_{i',n'})\;\delta_{k,\ell'}\delta_{k',\ell},\\
\mathbb{E}\left(m_{k,\ell}^{\theta_i,n'}\, \overline{m_{k',\ell'}^{\theta_{i'},n'}}\right) &= \Phi(\xi_{i,n},\overline{\xi_{i',n'}})\;\mathbf{e}_k^*\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^*\mathbf{e}_{k'}\;\mathbf{e}_{\ell'}^*\mathbf{Q}^*\mathbf{Q}\,\mathbf{e}_\ell,
\end{aligned}$$

where

$$\begin{aligned}
\Phi(z,w) &:= \int \frac{1}{z-x}\frac{1}{w-x}\mu(dx) - \int\frac{1}{z-x}\mu(dx)\int\frac{1}{w-x}\mu(dx)\\
&= \begin{cases} -\frac{G_\mu(z)-G_\mu(w)}{z-w} - G_\mu(z)G_\mu(w) & : \textit{if } z\neq w,\\ -G'_\mu(z)-(G_\mu(z))^2 & : \textit{otherwise.}\end{cases}
\end{aligned}$$

Remark 3.3.13. Remind that we supposed that μ is not a single Dirac measure, so that Φ is not equal to zero.

Remark 3.3.14. If $\mathbf{A}_N$ is Hermitian, the size of all the Jordan blocks are equal to 1 and the fluctuations are real random variables (see Remark 3.2.9). We find back that, in the Hermitian case, fluctuations between outliers at a macroscopic distance are independent (see [16]) except if the two outliers come from the same eigenvalue of $\mathbf{A}$ (i.e. they both belong to the same set $\mathscr{S}_\theta$). In this case, the fluctuations of outliers belonging to the same set $\mathscr{S}_\theta$ are all correlated. This phenomenon is illustrated by Figures 3.11(a) and 3.11(b).



(a) Uncorrelated case : $G_\mu(\xi_1)\neq G_\mu(\xi_2)$, which means that ξ_1 and ξ_2 do not belong to the same set $\mathscr{S}_\theta$.

(b) Correlated case : $G_\mu(\xi_1) = G_\mu(\xi_2)$, which means that ξ_1 and ξ_2 belong to the same set $\mathscr{S}_\theta$.

Figure 3.11: Correlation between the fluctuations of two outliers $\xi_1\neq\xi_2$ for a sample of 400 matrices of size 500, in the case where both $\mathbf{H}$ and $\mathbf{A}$ are Hermitian and where the support of μ is disconnected. Here $\mu(dx)$ is taken equal to $\frac{1}{2}(\delta_{-1}(dx)+\mathbb{1}_{[1,2]}(dx))$.

Proof. We just need to check that $\mathbf{H}_N$ satisfies $(1'),(2')$ of Assumption 3.2.6 (since $(0')$ is assumed below). Actually, for any $k\geq 1$ and any $i\in\{1,\dots,q\}$, the diagonal matrix

$$\left(\xi_{i,n}-\mathbf{H}_N\right)^{-(k+1)}-\frac{1}{N}\operatorname{Tr}\left(\xi_{i,n}-\mathbf{H}_N\right)^{-(k+1)},$$

fulfill the assumptions of Theorem 3.4.6, so that $(2')$ is true. Then, $(1')$ is true thanks to Theorem 3.4.8. This theorem also gives us the covariance.

□

3.4 Proofs

3.4.1 Convergence of the outliers : proof of Theorem 3.2.3

In [12], the authors give an interpretation of why the limit is necessarily a solution of $G_\mu(z)=\frac{1}{\theta}$ with the subordinate functions of the free additive convolution of measures in the particular case where one of the measure is δ_0 (see [12, Example 4.1]). Actually, our definition of the sets $\mathscr{S}_{\theta_i}$'s corresponds to the one of the set O_θ in [12, Definition 4.1]. A quick (but inaccurate) way to see why the limit is $G_\mu^{-1}(\frac{1}{\theta})$ and to understand the approach of the proof, is to write

$$\det\left(z-(\mathbf{H}_N+\mathbf{A}_N)\right) \;=\; \det\left(z-\mathbf{H}_N\right)\det\left(\mathbf{I}-\left(z-\mathbf{H}_N\right)^{-1}\mathbf{A}_N\right),$$

then if $\left(z-\mathbf{H}_N\right)^{-1}\sim G_\mu(z)\mathbf{I}$, we can write

$$\det\left(z-(\mathbf{H}_N+\mathbf{A}_N)\right)\sim\det\left(z-\mathbf{H}_N\right)G_\mu(z)\det\left(\frac{1}{G_\mu(z)}\mathbf{I}-\mathbf{A}_N\right)$$

so that if z is an outlier of $\mathbf{H}_N+\mathbf{A}_N$, $\frac{1}{G_\mu(z)}$ must be an eigenvalue of $\mathbf{A}_N$.

To do it properly, we introduce the following function[12],

$$f(z) \;=\; \det\left(\mathbf{I}-\mathbf{U}_{2r}^*\left(z\mathbf{I}-\mathbf{H}\right)^{-1}\mathbf{U}_{2r}\mathbf{A}_0\right) \;=\; \frac{\det\left(z-(\mathbf{H}_N+\mathbf{A}_N)\right)}{\det\left(z-\mathbf{H}_N\right)}, \tag{3.20}$$

we know that the zeros of f are eigenvalues of $\widetilde{\mathbf{H}}_N$ which are not eigenvalues of $\mathbf{H}_N$. Then, we introduce the function

$$f_0(z) \;:=\; \det\left(\mathbf{I}-G_\mu(z)\mathbf{A}_0\right), \tag{3.21}$$

and the proof of Theorem 3.2.3 relies on the two following lemmas.

Lemma 3.4.1. *As N goes to infinity, we have*

$$\sup_{\operatorname{dist}(z,\operatorname{supp}(\mu))>\delta}|f(z)-f_0(z)|\xrightarrow{(\mathbb{P})}0.$$

[12] we used a classical trick of finite rank perturbation models which $\det(\mathbf{I}_m+\mathbf{AB})=\det(\mathbf{I}_n+\mathbf{BA})$ for any $m\times n$ matrix $\mathbf{A}$ and $n\times m$ matrix $\mathbf{B}$

Lemma 3.4.2. *Let K be a compact set and let $\varepsilon > 0$ such that*

- $\operatorname{dist}(K, \bigcup_{i=1}^{j} \mathscr{S}_{\theta_i}) \geq \varepsilon$,
- $\operatorname{dist}(K, \operatorname{supp}(\mu)) \geq \varepsilon$.

Then, with a probability tending to one,

$$\inf_{z \in K} \left| \det\left(\mathbf{I} - (z - \mathbf{H}_N)^{-1} \mathbf{A}_N \right) \right| \; > \; 0.$$

If these lemmas are true, the end of the proof goes as follow. We know that, with a probability tending to one, there is $\varepsilon > 0$, such that

- there is a constant $M > 0$ such that $\mathbf{H}_N + \mathbf{A}_N$ has no eigenvalues in the area $\{z, \ |z| > M\}$,
- $\operatorname{Spec}(\mathbf{H}_N) \subset \{z, \ \operatorname{dist}(z, \operatorname{supp}(\mu)) < \varepsilon\}$,

We set

$$\mathscr{S} \; := \; \bigcup_{i=1}^{j} \mathscr{S}_{\theta_i}$$

and we define

$$\mathscr{S}^{\varepsilon} \; := \; \bigcup_{i=1}^{j} \bigcup_{\xi \in \mathscr{S}_{\theta_i}} \{z, \ |z - \xi| < \varepsilon\} \tag{3.22}$$

with the convention that $\mathscr{S}^{\varepsilon} = \emptyset$ if $\mathscr{S} = \emptyset$. Up to a smaller choice of ε, we can suppose that none of the disk centered in the element of the $\mathscr{S}_{\theta_i}$'s and of radius ε intersects each other nor intersect $\{z, \ \operatorname{dist}(z, \operatorname{supp}(\mu)) < \varepsilon\}$. Then, using Lemma 3.4.2, with

$$K \; := \; \{z, \ |z| \leq M\} \setminus (\mathscr{S}^{\varepsilon} \cup \{z, \ \operatorname{dist}(z, \operatorname{supp}(\mu)) < \varepsilon\}),$$

we deduce all the eigenvalues of $\widetilde{\mathbf{H}}_N$ are contained in $\mathscr{S}^{\varepsilon} \cup \{z, \ \operatorname{dist}(z, \operatorname{supp}(\mu)) < \varepsilon\}$. Indeed, if z is an eigenvalue of $\widetilde{\mathbf{H}}_N$ such that $\operatorname{dist}(z, \operatorname{supp}(\mu)) > \varepsilon$, z must be a zero of f.
Moreover, for each $i \in \{1, \ldots, j\}$ and each $\xi \in \mathscr{S}_{\theta_i}$, we know that from Lemma 3.4.1

$$\sup_{z, |z - \xi| = \varepsilon} |f(z) - f_0(z)| \; \longrightarrow \; 0, \qquad \text{and} \qquad \inf_{z, |z - \xi| = \varepsilon} |f_0(z)| \; > \; 0.$$

we deduce by Rouché Theorem (see [11, p. 131]) that f and f_0, for all large enough N, have the same number of zeros inside the domain $\{z, \ |z - \xi| < \varepsilon\}$, for each ξ in the $\mathscr{S}_{\theta_i}$'s.
Now, we just need to prove the two previous lemmas.

Proof. [of Lemma 3.4.1] We know that, for some positive constant C,

$$\sup_{\mathrm{dist}(z,\mathrm{supp}(\mu))>\delta} |f(z)-f_0(z)| \;\leq\; C \sup_{\mathrm{dist}(z,\mathrm{supp}(\mu))>\delta} \left\| \mathbf{U}_{2r}^* \left((z-\mathbf{H}_N)^{-1} - G_\mu(z)\mathbf{I} \right) \mathbf{U}_{2r} \right\|_{\mathrm{op}}$$

and we conclude with Assumption 3.2.2. □

Proof. [of Lemma 3.4.2] We write, thanks to Assumption 3.2.2,

$$\begin{aligned} \det\left(\mathbf{I} - (z-\mathbf{H}_N)^{-1}\mathbf{A}_N\right) &= \det\left(\mathbf{I} - \mathbf{U}_{2r}^*(z-\mathbf{H}_N)^{-1}\mathbf{U}_{2r}\mathbf{A}_0\right) \\ &= \det\left(\mathbf{I} - G_\mu(z)\mathbf{A}_0 + o(1)\right) \\ &= \prod_{i=1}^k \left(1 - G_\mu(z)\theta_i\right) + o(1)\,. \end{aligned}$$

Then, since $z \in K$, it easy to show that for each i, $|1-G_\mu(z)\theta_i| > 0$. □

3.4.2 Fluctuations

The proof of Theorem 3.2.10 is the same than [21, Theorem 2.10] and all we need to do here is to prove this analogous version of [21, Lemma 5.1].

Lemma 3.4.3. *For all $j \in \{1,\dots,\alpha_i\}$ and all $n \in \{1,\dots,m_i\}$, let $F_{j,n}^{\theta_i}(z)$ be the rationnal function defined by*

$$F_{j,n}^{\theta_i}(z) \;:=\; f\left(\xi_{i,n} + \frac{z}{N^{1/(2p_{i,j})}}\right). \tag{3.23}$$

Then, there exists a collection of positive constants $(\gamma_{i,j})_{\substack{1\leq i\leq q\\ 1\leq j\leq \alpha_i}}$ and a collection of non vanishing random variables $(C_{i,j,n})_{\substack{1\leq i\leq q\\ 1\leq j\leq\alpha_i\\ 1\leq n\leq m_i}}$ independent of z, such that we have the convergence in distribution (for the topology of the uniform convergence over any compact set)

$$\left(N^{\gamma_{i,j}} F_{j,n}^{\theta_i}(\cdot)\right)_{\substack{1\leq i\leq q\\ 1\leq j\leq\alpha_i\\ 1\leq n\leq m_i}} \underset{N\to\infty}{\longrightarrow} \left(z\in\mathbb{C} \;\mapsto\; z^{\pi_{i,j}}\cdot C_{i,j,n}\cdot\det\left(z^{p_{i,j}} - \mathbf{M}_{j,n}^{\theta_i}\right)\right)_{\substack{1\leq i\leq q\\ 1\leq j\leq\alpha_i\\ 1\leq n\leq m_i}}$$

where $\mathbf{M}_{j,n}^{\theta_i}$ is the random matrix introduced at (3.11) and $\pi_{i,j} := \sum_{l>j}\beta_{i,l}p_{i,l}$.

Once this lemma proven, the Theorem 3.2.10 follows (see section 5.1 of [21] for more details). To prove Lemma 3.4.3, we shall proceed as it is done in [21] to prove Lemma 5.1. First, we write, for a fixed $\theta_i(=\theta)$, a fixed $n\in\{1,\dots,m_i\}$ and a fixed $j\in\{1,\dots,\alpha_i\}$ (which shall be implicit) and fixed $p_{i,j}(=p)$, recall that

$\mathbf{A}_0 = \mathbf{QJQ}^{-1}$,

$$\begin{aligned}
F^{\theta}_{j,n}(z) &= \det\left(\mathbf{I} - \left(\xi_n + \frac{z}{N^{1/(2p)}} - \mathbf{H}_N\right)^{-1} \mathbf{U}_{2r}\mathbf{QJQ}^{-1}\mathbf{U}^*_{2r}\right) \\
&= \det\left(\mathbf{I} - G_\mu\left(\xi_n + \frac{z}{N^{1/(2p)}}\right)\mathbf{J} - \frac{1}{\sqrt{N}}\mathbf{Z}_N\left(\xi_n + \frac{z}{N^{1/(2p)}}\right)\right) \\
&= \det\left(\mathbf{I} - \frac{\mathbf{J}}{\theta} - G'_\mu(\xi_n)\frac{z}{N^{1/(2p)}}(1+o(1))\mathbf{J} - \frac{1}{\sqrt{N}}\mathbf{Z}_N\left(\xi_n + \frac{z}{N^{1/(2p)}}\right)\right)
\end{aligned}$$

where

$$Z_N(z) := \sqrt{N}\mathbf{Q}^{-1}\mathbf{U}^*_{2r}\left((z-\mathbf{H}_N)^{-1} - G_\mu(z)\mathbf{I}\right)\mathbf{U}_{2r}\mathbf{QJ}.$$

Remind that by definition, $G_\mu(\xi_n) = \theta^{-1}$. From here, the reasoning to end the proof is the exact same than the one from [21, Lemma 5.1]. Nevertheless, we still have to prove that, for all θ and for all n, for all compact set K and for all $z \in K$,

$$\mathbf{Z}_N(\xi_n + \frac{z}{N^{1/(2p)}}) = \mathbf{Z}_N(\xi_n) + o(1), \qquad \text{and} \quad \mathbf{Z}_N(\xi_n) \text{ converges weakly} \tag{3.24}$$

To do so, we write (thanks to 3.4.4),

$$\begin{aligned}
\mathbf{Z}_N\left(\xi_n + \frac{z}{N^{1/(2p)}}\right) &= \sqrt{N}\mathbf{Q}^{-1}\mathbf{U}^*_{2r}\left((\xi_n - \mathbf{H}_N)^{-1} - \frac{1}{\theta}\right)\mathbf{U}_{2r}\mathbf{QJ} \\
&\quad + \sum_{k=1}^{p}\left(\frac{-z}{N^{1/(2p)}}\right)^k \sqrt{N}\mathbf{Q}^{-1}\mathbf{U}^*_{2r}\left((\xi_n - \mathbf{H}_N)^{-(k+1)} - \int\frac{\mu(dx)}{(\xi_n - x)^{k+1}}\right)\mathbf{U}_{2r}\mathbf{QJ} \\
&\quad + \frac{1}{N^{1/(2p)}}\mathbf{QU}^*_{2r}\left(\xi_n - \mathbf{H}_N\right)^{-(k+1)}\left(\xi_n + \frac{z}{N^{1/(2p)}} - \mathbf{H}_N\right)^{-1}\mathbf{U}_{2r}\mathbf{QJ} + o(1).
\end{aligned}$$

The last term is a $o(1)$ since $\operatorname{dist}(\xi_n, \operatorname{Spec}(\mathbf{H}_N)) > \varepsilon$ and one can conclude if $(1), (2)$ are satisfied in Assumption 3.2.6. Otherwise, if it's $(0'), (1'), (2')$, we write

$$\begin{aligned}
\mathbf{Z}_N\left(\xi_n + \frac{z}{N^{1/(2p)}}\right) &= \sqrt{N}\mathbf{Q}^{-1}\mathbf{U}^*_{2r}\left(\frac{1}{N}\operatorname{Tr}(\xi_n - \mathbf{H}_N)^{-1} - \frac{1}{\theta}\right)\mathbf{U}_{2r}\mathbf{QJ} \\
&\quad + \sqrt{N}\mathbf{Q}^{-1}\mathbf{U}^*_{2r}\left((\xi_n - \mathbf{H}_N)^{-1} - \frac{1}{N}\operatorname{Tr}(\xi_n - \mathbf{H}_N)^{-1}\right)\mathbf{U}_{2r}\mathbf{QJ} \\
&\quad + \sum_{k=1}^{p}\left(\frac{-z}{N^{1/(2p)}}\right)^k \sqrt{N}\mathbf{Q}^{-1}\mathbf{U}^*_{2r}\left((\xi_n - \mathbf{H}_N)^{-(k+1)} - \frac{1}{N}\operatorname{Tr}(\xi_n - \mathbf{H}_N)^{-1}\right)\mathbf{U}_{2r}\mathbf{QJ} \\
&\quad + \frac{1}{N^{1/(2p)}}\mathbf{QU}^*_{2r}\left(\xi_n - \mathbf{H}_N\right)^{-(p+1)}\left(\xi_n + \frac{z}{N^{1/(2p)}} - \mathbf{H}_N\right)^{-1}\mathbf{U}_{2r}\mathbf{QJ} + o(1).
\end{aligned}$$

Appendix

3.4.3 Linear algebra lemmas

Lemma 3.4.4. *Let* $\mathbf{A}$ *be a matrix and* $\lambda \in \mathbb{C}$ *be such that both* $\mathbf{A}$ *and* $\mathbf{A}+\lambda\mathbf{I}$ *are non singular. Then, for all* $p \geq 1$,

$$(\mathbf{A}+\lambda\mathbf{I})^{-1} \;=\; \sum_{k=1}^{p}(-\lambda)^{k-1}\mathbf{A}^{-k}+(-\lambda)^{p}\mathbf{A}^{-p}(\mathbf{A}+\lambda\mathbf{I})^{-1}$$

Lemma 3.4.5 (Schur's complement [53])**.** *For any* $\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}$*, one has, when it makes sense*

$$\begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{pmatrix}^{-1} = \begin{pmatrix}\left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1} & -\left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1}\mathbf{B}\mathbf{D}^{-1}\\ -\mathbf{D}^{-1}\mathbf{C}\left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1} & \mathbf{D}^{-1}+\mathbf{D}^{-1}\mathbf{C}\left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1}\mathbf{B}\mathbf{D}^{-1}\end{pmatrix}$$

3.4.4 Fluctuations of the entries of UCI random matrices

We give here some results on the fluctuations of the entries of UCI matrices, which means, matrices of the form $\mathbf{H} := \mathbf{U}\mathbf{D}\mathbf{U}^*$ where $\mathbf{U}$ is Haar-distributed and $\mathbf{D}$ is a complex diagonal matrix.

Theorem 3.4.6 (Fluctuations of the entries of UCI random matrices)**.** *Let* $\mathbf{T}$ *be an* $N \times N$ *diagonal matrix such that*

$$\operatorname{Tr}\mathbf{T} \;=\; 0, \qquad \frac{1}{N}\operatorname{Tr}\mathbf{T}\mathbf{T}^* \;\to\; \sigma^2, \qquad \frac{1}{N}\operatorname{Tr}\mathbf{T}^2 \;\to\; \tau^2, \qquad \forall k \geq 1,\ \frac{1}{N}\operatorname{Tr}(\mathbf{T}\mathbf{T}^*)^k \;=\; O(1) \tag{3.25}$$

Let $\mathbf{u}_{t_1},\ldots,\mathbf{u}_{t_p}$ *be* p *distinct columns of a Haar-distributed unitary matrix. Then*

$$\left(\sqrt{N}\langle \mathbf{u}_{t_i},\, \mathbf{T}\mathbf{u}_{t_j}\rangle\right)_{i,j=1}^{p},$$

converges in distribution to a centered complex Gaussian vector $\left(\mathscr{G}_{i,j}\right)_{i,j=1}^{p}$ *with covariance*

$$\mathbb{E}\left[\mathscr{G}_{i,j}\mathscr{G}_{k,\ell}\right] \;=\; \delta_{i,\ell}\delta_{j,k}\tau^2 \quad ; \quad \mathbb{E}\left[\mathscr{G}_{i,j}\overline{\mathscr{G}}_{k,\ell}\right] \;=\; \delta_{i,k}\delta_{j,\ell}\sigma^2$$

Remark 3.4.7. If $\mathbf{H} := \mathbf{U}\mathbf{D}\mathbf{U}^*$ satisfies (3.1) and Assumption 3.2.1, then $\mathbf{T} := \mathbf{D} - \frac{1}{N}\operatorname{Tr}\mathbf{D}$ satisfies (3.25).

Here comes a version of Theorem 3.4.6, with several matrices diagonal $\mathbf{T}$. Due to the complex values of the diagonal matrices, the following theorem is not a simple consequence of Theorem 3.4.6 and Cramér–Wold theorem.

Theorem 3.4.8. *Let* $\mathbf{T}_1,\ldots,\mathbf{T}_q$ *be* $N \times N$ *diagonal matrices such that for all* $m,n \in \{1,\ldots,q\}$

$$\operatorname{Tr}\mathbf{T}_m \;=\; 0, \qquad \frac{1}{N}\operatorname{Tr}\mathbf{T}_m\mathbf{T}_n^* \;\to\; \sigma_{m,n}^2, \qquad \frac{1}{N}\operatorname{Tr}\mathbf{T}_m\mathbf{T}_n \;\to\; \tau_{m,n}^2, \qquad \forall k \geq 1,\ \frac{1}{N}\operatorname{Tr}(\mathbf{T}_m\mathbf{T}_m^*)^k \;=\; O(1)\,.$$

Let $\mathbf{u}_{t_1},\dots,\mathbf{u}_{t_p}$ be p distinct columns of an Haar-distributed matrix. Then

$$\left(\sqrt{N}\langle \mathbf{u}_{t_i},\ \mathbf{T}_m\mathbf{u}_{t_j}\rangle\right)_{\substack{1\le i\le p\\ 1\le j\le p\\ 1\le m\le q}},$$

converges in distribution to a centered complex Gaussian vector $\left(\mathscr{G}_{i,j,m}\right)_{\substack{1\le i\le p\\ 1\le j\le p\\ 1\le m\le q}}$ with covariance

$$\mathbb{E}\left[\mathscr{G}_{i,j,m}\mathscr{G}_{k,\ell,n}\right] \;=\; \delta_{i,\ell}\delta_{j,k}\tau^2_{m,n} \quad ; \quad \mathbb{E}\left[\mathscr{G}_{i,j,m}\overline{\mathscr{G}}_{k,\ell,n}\right] \;=\; \delta_{i,k}\delta_{j,\ell}\sigma^2_{m,n}$$

Proof of Theorem 3.4.6. Without any loss of generality, due to the invariance by conjugation by a matrix of permutation, we can suppose that $t_1=1,t_2=2,\dots,t_p=p$. Then, we just need to show that

$$X \;=\; \sqrt{N}\,\mathrm{Tr}\left(\mathbf{U}^*\mathbf{T}\mathbf{U}\mathbf{A}\right)$$

where $\mathbf{A}$ is a $N\times N$ deterministic matrix of the form

$$\mathbf{A} \;=\; \begin{pmatrix}\mathbf{A}_p & 0\\ 0 & 0\end{pmatrix}, \qquad \mathbf{A}_p=\left(a_{i,j}\right)_{i,j=1}^p\in\mathscr{M}_p(\mathbb{C}),$$

is a asymptotically Gaussian. Before starting, we remind some definition. Let $(\mathbf{M}_1,\dots,\mathbf{M}_q)$ be q matrices. For any permutation $\sigma\in S_q$, with cycle decomposition

$$\sigma \;=\; (i_{1,1}\cdots i_{1,k_1})(i_{2,1}\cdots i_{2,k_2})\cdots(i_{r,1}\cdots i_{r,k_r})$$

we denote by

$$\mathrm{Tr}_\sigma\left(\mathbf{M}_t\right)_{t=1}^q \;:=\; \prod_{j=1}^r \mathrm{Tr}\left(\mathbf{M}_{t_{i_{j,1}}}\cdots\mathbf{M}_{t_{i_{j,r_j}}}\right). \tag{3.26}$$

For example, if $\sigma=(13)(256)\in S_6$, then

$$\mathrm{Tr}_\sigma\left(\mathbf{M}_t\right)_{t=1}^6 \;=\; \mathrm{Tr}(\mathbf{M}_1\mathbf{M}_3)\,\mathrm{Tr}(\mathbf{M}_2\mathbf{M}_5\mathbf{M}_6)\,\mathrm{Tr}(\mathbf{M}_4).$$

Let $M(2n)$ be the set of all *perfect matching* on $\{1,\dots,2n\}$ which is a subset of S_{2n} of the permutation which are the product of n transpositions with disjoint support. For example

$$M(4) \;=\; \{(12)(34),(13)(24),(14)(23)\}.$$

Then, if the following lemma is true, one can conclude the proof.

Lemma 3.4.9. *Let $\mathbf{T}_1,\dots,\mathbf{T}_q$ be q diagonal matrix such for all $i,j\in\{1,\dots,q\}$,*

$$\mathrm{Tr}\,\mathbf{T}_i \;=\; 0 \quad ; \quad \frac{1}{N}\mathrm{Tr}\,\mathbf{T}_i\mathbf{T}_j \;\longrightarrow\; \tau_{i,j} \quad ; \quad \forall k\ge 1,\ \frac{1}{N}\mathrm{Tr}\left(\mathbf{T}_i\mathbf{T}_i^*\right)^k \;=\; O(1). \tag{3.27}$$

Let $\mathbf{A}_1, \dots, \mathbf{A}_q$ be q matrices of the form

$$\mathbf{A}_i = \begin{pmatrix} \mathbf{A}_{0,i} & 0 \\ 0 & 0 \end{pmatrix}$$

where the $\mathbf{A}_{0,i}$'s are $K \times K$ matrices independent from N where K is a fixed integer. Let $\mathbf{U}$ be a Haar-distributed matrix. Then, as N goes to infinity,

$$\mathbb{E}\left[\prod_{t=1}^{q} \sqrt{N}\,\mathrm{Tr}\left(\mathbf{U}^*\mathbf{T}_t\mathbf{U}\mathbf{A}_t\right)\right] \longrightarrow \begin{cases} \displaystyle\sum_{\sigma\in M(q)} \mathrm{Tr}_\sigma\left(\mathbf{A}_i\right)_{i=1}^{q} \prod_{t=1}^{q/2} \tau_{\sigma(2t-1),\sigma(2t)} & \text{if } q \text{ is even.} \\ 0 & \text{if } q \text{ is odd.} \end{cases}$$

Indeed, once we suppose Lemma 3.4.9 satisfied, we need to compute for all p, q

$$\mathbb{E}\left[\left[\sqrt{N}\,\mathrm{Tr}\,\mathbf{U}^*\mathbf{T}\mathbf{U}\mathbf{A}\right]^p \left[\overline{\sqrt{N}\,\mathrm{Tr}\,\mathbf{U}^*\mathbf{T}\mathbf{U}\mathbf{A}}\right]^q\right] = \mathbb{E}\left[\left[\sqrt{N}\,\mathrm{Tr}\,\mathbf{U}^*\mathbf{T}\mathbf{U}\mathbf{A}\right]^p \left[\sqrt{N}\,\mathrm{Tr}\,\mathbf{U}^*\mathbf{T}^*\mathbf{U}\mathbf{A}^*\right]^q\right],$$

in order to apply Lemma 3.4.10. According to Lemma 3.4.9, for $\mathbf{T}_t \equiv \mathbf{T}$ and $\mathbf{A}_t \equiv \mathbf{A}$, we have

$$\mathbb{E}\left[\sqrt{N}\,\mathrm{Tr}\,\mathbf{U}^*\mathbf{T}\mathbf{U}\mathbf{A}\right]^q \longrightarrow \begin{cases} \mathrm{Card}\,M(q)\,\mathrm{Tr}(\mathbf{A}^2)^{q/2}\tau^{q/2} & \text{if q is even,} \\ 0 & \text{if q is odd.} \end{cases}$$

(remind that $\mathrm{Card}\,M(q) = (q-1)(q-3)\cdots 3$) which means that the limit distribution of X already satisfies (3.30) and (3.31). Let $p \geq 1$ and $q \geq 2$ be two fixed integers such that $p+q$ is even, then, using notations from (3.26), we know thanks to Lemma 3.4.9 that

$$\mathbb{E}\left[\left[\sqrt{N}\,\mathrm{Tr}\,\mathbf{U}^*\mathbf{T}\mathbf{U}\mathbf{A}\right]^p \left[\sqrt{N}\,\mathrm{Tr}\,\mathbf{U}^*\mathbf{T}^*\mathbf{U}\mathbf{A}^*\right]^q\right] = \frac{1}{N^{\frac{p+q}{2}}} \sum_{\sigma\in M(p+q)} \mathrm{Tr}_\sigma\left(\mathbf{A}_t\right)_{t=1}^{p+q} \mathrm{Tr}_\sigma\left(\mathbf{T}_t\right)_{t=1}^{p+q} \tag{3.28}$$

where

$$(\mathbf{T}_1, \dots, \mathbf{T}_{p+q}) = (\underbrace{\mathbf{T}, \dots, \mathbf{T}}_{p}, \underbrace{\mathbf{T}^*, \dots, \mathbf{T}^*}_{q}) \quad \text{and} \quad (\mathbf{A}_1, \dots, \mathbf{A}_{p+q}) = (\underbrace{\mathbf{A}, \dots, \mathbf{A}}_{p}, \underbrace{\mathbf{A}^*, \dots, \mathbf{A}^*}_{q}).$$

We rewrite the right side of (3.28) summing according to the value of $\sigma(1)$.

$$\begin{aligned}
\sum_{\sigma\in M(p+q)} \operatorname{Tr}_\sigma\left(\mathbf{A}_t\right)_{t=1}^{p+q} \operatorname{Tr}_\sigma\left(\mathbf{T}_t\right)_{t=1}^{p+q} &= \sum_{a=2}^{p+q} \sum_{\substack{\sigma\in M(p+q)\\ \sigma(1)=a}} \operatorname{Tr}_\sigma\left(\mathbf{A}_t\right)_{t=1}^{p+q} \operatorname{Tr}_\sigma\left(\mathbf{T}_t\right)_{t=1}^{p+q} \\
&= \sum_{a=2}^{p+q} \operatorname{Tr}(\mathbf{A}_1\mathbf{A}_a)\operatorname{Tr}(\mathbf{T}_1\mathbf{T}_a) \sum_{\substack{\sigma\in M(p+q)\\ \sigma(1)=a}} \operatorname{Tr}_{\sigma\circ(1a)}\left(\mathbf{A}_t\right)_{t=1}^{p+q} \operatorname{Tr}_{\sigma\circ(1a)}\left(\mathbf{T}_t\right)_{t=1}^{p+q} \\
&= \sum_{a=2}^{p} \operatorname{Tr}\mathbf{A}^2 \operatorname{Tr}\mathbf{T}^2 \sum_{\sigma\in M(p+q-2)} \operatorname{Tr}_\sigma\left(\widehat{\mathbf{A}}_t\right)_{t=1}^{p+q-2} \operatorname{Tr}_\sigma\left(\widehat{\mathbf{T}}_t\right)_{t=1}^{p+q-2} \\
&+ \sum_{a=p+1}^{p+q} \operatorname{Tr}\mathbf{A}\mathbf{A}^* \operatorname{Tr}\mathbf{T}\mathbf{T}^* \sum_{\sigma\in M(p+q-2)} \operatorname{Tr}_\sigma\left(\widetilde{\mathbf{A}}_t\right)_{t=1}^{p+q-2} \operatorname{Tr}_\sigma\left(\widetilde{\mathbf{T}}_t\right)_{t=1}^{p+q-2},
\end{aligned}$$

where

$$(\widehat{\mathbf{A}}_1,\ldots,\widehat{\mathbf{A}}_{p+q-2}) = (\underbrace{\mathbf{A},\ldots,\mathbf{A}}_{p-2},\underbrace{\mathbf{A}^*,\ldots,\mathbf{A}^*}_{q}) \quad\text{and}\quad (\widetilde{\mathbf{A}}_1,\ldots,\widetilde{\mathbf{A}}_{p+q-2}) = (\underbrace{\mathbf{A},\ldots,\mathbf{A}}_{p-1},\underbrace{\mathbf{A}^*,\ldots,\mathbf{A}^*}_{q-1}).$$

At last, one easily deduces that

$$\begin{aligned}
\mathbb{E}\left[\left[\sqrt{N}\operatorname{Tr}\mathbf{U}^*\mathbf{T}\mathbf{U}\mathbf{A}\right]^p\left[\sqrt{N}\operatorname{Tr}\mathbf{U}^*\mathbf{T}^*\mathbf{U}\mathbf{A}^*\right]^q\right] &= \frac{1}{N}\operatorname{Tr}\mathbf{T}^2\operatorname{Tr}\mathbf{A}^2(p-1)\,\mathbb{E}\left[\left[\sqrt{N}\operatorname{Tr}\mathbf{U}^*\mathbf{T}\mathbf{U}\mathbf{A}\right]^{p-2}\left[\sqrt{N}\operatorname{Tr}\mathbf{U}^*\mathbf{T}^*\mathbf{U}\mathbf{A}^*\right]^q\right] \\
&+ \frac{1}{N}\operatorname{Tr}\mathbf{T}\mathbf{T}^*\operatorname{Tr}\mathbf{A}\mathbf{A}^* q\,\mathbb{E}\left[\left[\sqrt{N}\operatorname{Tr}\mathbf{U}^*\mathbf{T}\mathbf{U}\mathbf{A}\right]^{p-1}\left[\sqrt{N}\operatorname{Tr}\mathbf{U}^*\mathbf{T}^*\mathbf{U}\mathbf{A}^*\right]^{q-1}\right] + o(1)
\end{aligned}$$

and so $\sqrt{N}\operatorname{Tr}(\mathbf{U}^*\mathbf{T}\mathbf{A}\mathbf{U})$ satisfies (3.32) which means according to Lemma 3.4.10 that its limit distribution is Gaussian.
At last, to compute to covariance of the $(\mathscr{G}_{i,j})$'s, one can simply use [22, Lemma A.6]. $\square$

Proof of Theorem 3.4.8. This time, we shall use Lemma 3.4.11 to show that for any $\mathbf{A}_1,\ldots,\mathbf{A}_r$, $N\times N$ deterministic matrix of the form

$$\mathbf{A}_m = \begin{pmatrix} \mathbf{A}_{m,p} & 0 \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{A}_{m,p} = (a^m_{i,j})_{i,j=1}^p \in \mathscr{M}_p(\mathbb{C}),$$

the vector

$$\left(\frac{1}{N}\operatorname{Tr}\left(\mathbf{U}^*\mathbf{T}_1\mathbf{U}\mathbf{A}_1\right),\ldots,\frac{1}{N}\operatorname{Tr}\left(\mathbf{U}^*\mathbf{T}_r\mathbf{U}\mathbf{A}_r\right)\right)$$

converges weakly to a Gaussian multivariate. Thanks to Theorem 3.4.6, we know that for each m,

$$\frac{1}{N}\operatorname{Tr}\left(\mathbf{U}^*\mathbf{T}_m\mathbf{U}\mathbf{A}_m\right)$$

is asymptotically Gaussian. Then, we show that

$$\mathbb{E}\left[\left(\frac{1}{N}\operatorname{Tr}\left(\mathbf{U}^*\mathbf{T}_1\mathbf{U}\mathbf{A}_1\right)\right)^{p_1}\left(\frac{1}{N}\operatorname{Tr}\left(\mathbf{U}^*\mathbf{T}_1^*\mathbf{U}\mathbf{A}_1^*\right)\right)^{q_1}\cdots\left(\frac{1}{N}\operatorname{Tr}\left(\mathbf{U}^*\mathbf{T}_r\mathbf{U}\mathbf{A}_r\right)\right)^{p_r}\left(\frac{1}{N}\operatorname{Tr}\left(\mathbf{U}^*\mathbf{T}_r^*\mathbf{U}\mathbf{A}_r^*\right)\right)^{q_r}\right]$$

satisfies (3.35) and (3.36) using Lemma 3.4.9. □

Proof of Lemma 3.4.9. We know from [**?**, Proposition 3.4]

$$\mathbb{E}\left[\prod_{t=1}^{q}\sqrt{N}\operatorname{Tr}\left(\mathbf{U}^*\mathbf{T}_t\mathbf{U}\mathbf{A}_t\right)\right] = N^{q/2}\sum_{\sigma,\tau\in S_q}\operatorname{Wg}(\tau\circ\sigma^{-1})\operatorname{Tr}_\tau\left(\mathbf{A}_i\right)_{i=1}^{q}\operatorname{Tr}_\sigma\left(\mathbf{T}_i\right)_{i=1}^{q}$$

where Wg is a function called the *Weingarten function*. Moreover, for $\sigma\in S_q$, the asymptotical behavior of $\operatorname{Wg}(\sigma)$ is at most given by

$$\operatorname{Wg}(\sigma) = O\left(N^{-q}\right). \tag{3.29}$$

First, one should notice that if σ has one invariant point (which means a cycle of size one in its cycle decomposition), then

$$\operatorname{Tr}_\sigma\left(\mathbf{T}_i\right)_{i=1}^{q} = 0,$$

also, if σ has r cycles in its cycle decomposition, then, by the Holder inequality,

$$\operatorname{Tr}_\sigma\left(\mathbf{T}_i\right)_{i=1}^{q} = O(N^r).$$

Actually, the maximum of cycles in its decomposition that can have σ without any 1-sized cycle is $\left\lfloor\frac{q}{2}\right\rfloor$ so that, using (3.29)

$$N^{q/2}\operatorname{Wg}(\sigma\circ\tau^{-1})\operatorname{Tr}_\tau\left(\mathbf{A}_i\right)_{i=1}^{q}\operatorname{Tr}_\sigma\left(\mathbf{T}_i\right)_{i=1}^{q} = O\left(N^{\lfloor\frac{q}{2}\rfloor-\frac{q}{2}}\right),$$

so that first, if q is odd

$$\mathbb{E}\left[\prod_{t=1}^{q}\sqrt{N}\operatorname{Tr}\left(\mathbf{U}^*\mathbf{T}_t\mathbf{U}\mathbf{A}_t\right)\right] = o(1).$$

Moreover, if $q=2r$, then the only way to have

$$N^{q/2}\operatorname{Wg}(\sigma\circ\tau^{-1})\operatorname{Tr}_\tau\left(\mathbf{A}_i\right)_{i=1}^{q}\operatorname{Tr}_\sigma\left(\mathbf{T}_i\right)_{i=1}^{q} \neq o(1)$$

is to have

- $\tau=\sigma$,
- σ is a product of $\frac{q}{2}=r$ transpositions with disjoint support.

One easily conclude. □

3.4.5 Moments of a complex Gaussian variable.

The following lemma allows to prove that a random variable is Gaussian if and only if its moments satisfy an induction relation.

Lemma 3.4.10. *Let Z be a complex Gaussian variable such that*

$$\mathbb{E}[Z] = 0, \quad \mathbb{E}\left[Z^2\right] = \tau^2, \quad \mathbb{E}\left[|Z|^2\right] = \sigma^2. \tag{3.30}$$

Then, for all $p \geq 1$

$$\mathbb{E}\left[Z^{2p}\right] = p!!\tau^{2p} \quad \text{and} \quad \mathbb{E}\left[Z^{2p+1}\right] = 0, \qquad \left(\text{where } p!! := \frac{(2p)!}{2^p p!}\right) \tag{3.31}$$

also, for all $p,q \geq 0$,

$$\begin{aligned} \mathbb{E}\left[Z^{p+2}\overline{Z}^{q+2}\right] &= \sigma^2(q+2)\mathbb{E}\left[Z^{p+1}\overline{Z}^{q+1}\right] + \tau^2(p+1)\mathbb{E}\left[Z^p\overline{Z}^{q+2}\right] \\ &= \sigma^2(p+2)\mathbb{E}\left[Z^{p+1}\overline{Z}^{q+1}\right] + \overline{\tau}^2(q+1)\mathbb{E}\left[Z^{p+2}\overline{Z}^q\right]. \end{aligned} \tag{3.32}$$

Conversely, any complex random variable Z satisfying (3.30),(3.31) and (3.32) is a complex Gaussian variable.

Proof. First, recall that if $Z = X_1 + \mathrm{i}X_2$ is a complex random Gaussian such that

$$\mathbb{E}[Z] = 0, \quad \mathbb{E}\left[Z^2\right] = \tau^2, \quad \mathbb{E}\left[|Z|^2\right] = \sigma^2,$$

then, its Fourier transform is given, for $t = t_1 + \mathrm{i}t_2 \in \mathbb{C}$, by

$$\begin{aligned} \Phi_Z(t) &:= \mathbb{E}\exp\left(i(X_1t_1 + X_2t_2)\right) \\ &= \exp\left(-\frac{1}{4}\left((t_1^2+t_2^2)\sigma^2 + (t_1^2-t_2^2)\operatorname{Re}(\tau^2) + 2t_1t_2\operatorname{Im}(\tau^2)\right)\right) \end{aligned}$$

We define the differential operators

$$\partial_t := \partial_1 + \mathrm{i}\partial_2 \quad ; \quad \partial_{\bar{t}} := \partial_1 - \mathrm{i}\partial_2 \tag{3.33}$$

so that

$$\mathbb{E}\left[Z^p\overline{Z}^q\right] = (-\mathrm{i})^{p+q}\partial_t^p\partial_{\bar{t}}^q\Phi(t)\big|_{t=0}. \tag{3.34}$$

One can easily compute

$$\partial_t\Phi(t) = -\frac{1}{2}\left(t\sigma^2 + \bar{t}\tau^2\right)\Phi(t) \quad ; \quad \partial_{\bar{t}}\Phi(t) = -\frac{1}{2}\left(\bar{t}\sigma^2 + t\overline{\tau}^2\right)\Phi(t)$$

therefore, for any $p \geq 0$, $q \geq 0$,

$$\begin{aligned}
\partial_t^{p+2}\Phi(t) &= \partial_t^{p+1}\left(-\frac{1}{2}\left(t\sigma^2+\bar{t}\tau^2\right)\Phi(t)\right) \\
&= -\frac{1}{2}t\sigma^2\partial_t^{p+1}\Phi(t) - \frac{1}{2}\tau^2\partial_t^{p+1}\left(\bar{t}\Phi(t)\right) \\
&= -\frac{1}{2}t\sigma^2\partial_t^{p+1}\Phi(t) - \frac{1}{2}\tau^2\bar{t}\partial_t^{p+1}\Phi(t) - \tau^2(p+1)\partial_t^p\Phi(t)
\end{aligned}$$

and

$$\begin{aligned}
\partial_t^{p+2}\partial_{\bar{t}}^{q+2}\Phi(t) &= \partial_t^{p+1}\partial_{\bar{t}}^{q+2}\left(-\frac{1}{2}\left(t\sigma^2+\bar{t}\tau^2\right)\Phi(t)\right) \\
&= -\frac{1}{2}\sigma^2\partial_{\bar{t}}^{q+2}\left(t\partial_t^{p+1}\Phi(t)\right) - \frac{1}{2}\tau^2\partial_t^{p+1}\left(\bar{t}\partial_{\bar{t}}^{q+2}\Phi(t)\right) \\
&= -\sigma^2\left(\frac{t}{2}\partial_t^{p+1}\partial_{\bar{t}}^{q+2}\Phi(t) + (q+2)\partial_t^{p+1}\partial_{\bar{t}}^{q+1}\Phi(t)\right) \\
&\quad - \tau^2\left(\frac{\bar{t}}{2}\partial_t^{p+1}\partial_{\bar{t}}^{q+2}\Phi(t) + (p+1)\partial_t^{p}\partial_{\bar{t}}^{q+2}\Phi(t)\right),
\end{aligned}$$

hence,

$$\begin{aligned}
\mathbb{E}\left[Z^{p+2}\right] &= \tau^2(p+1)\mathbb{E}\left[Z^p\right], \\
\mathbb{E}\left[Z^{p+2}\overline{Z}^{q+2}\right] &= \sigma^2(q+2)\mathbb{E}\left[Z^{p+1}\overline{Z}^{q+1}\right] + \tau^2(p+1)\mathbb{E}\left[Z^{p}\overline{Z}^{q+2}\right]
\end{aligned}$$

and the same way,

$$\begin{aligned}
\mathbb{E}\left[\overline{Z}^{p+2}\right] &= \overline{\tau}^2(p+1)\mathbb{E}\left[\overline{Z}^p\right], \\
\mathbb{E}\left[Z^{p+2}\overline{Z}^{q+2}\right] &= \sigma^2(p+2)\mathbb{E}\left[Z^{p+1}\overline{Z}^{q+1}\right] + \overline{\tau}^2(q+1)\mathbb{E}\left[Z^{p+2}\overline{Z}^{q}\right].
\end{aligned}$$

Conversely, one can easily prove by induction that any complex random variable Z satisfying (3.30),(3.31) and (3.32) has all its moments uniquely determined and since the complex Gaussian variable also satisfies (3.30),(3.31) and (3.32), one can conclude.

□

More generally, one can show the following lemma

Lemma 3.4.11. *Let $(X_1, \dots, X_r)$ be a centered complex Gaussian vector. Then, for all non negative integers $p_1, q_1, \dots, p_r, q_r$, for all $i \in \{1, \dots, r\}$*

$$\begin{aligned}
\text{if } p_i \geq 1, \quad \mathbb{E}\left[X_1^{p_1}\overline{X_1}^{q_1}\cdots X_r^{p_r}\overline{X_r}^{q_r}\right] &= (p_i-1)\mathbb{E}\left[X_i^2\right]\mathbb{E}\left[X_1^{p_1}\overline{X_1}^{q_1}\cdots X_i^{p_i-2}\overline{X_i}^{q_i}\cdots X_r^{p_r}\overline{X_r}^{q_r}\right] \\
&+ \sum_{\substack{j=1 \\ j\neq i}}^{r} p_j\,\mathbb{E}\left[X_iX_j\right]\mathbb{E}\left[X_1^{p_1}\overline{X_1}^{q_1}\cdots X_i^{p_i-1}\cdots X_j^{p_j-1}\cdots X_r^{p_r}\overline{X_r}^{q_r}\right] \quad (3.35) \\
&+ \sum_{j=1}^{r} q_j\,\mathbb{E}\left[X_i\overline{X_j}\right]\mathbb{E}\left[X_1^{p_1}\overline{X_1}^{q_1}\cdots X_i^{p_i-1}\cdots\overline{X_j}^{q_j-1}\cdots X_r^{p_r}\overline{X_r}^{q_r}\right]
\end{aligned}$$

$$\text{if } q_i \geq 1, \qquad \mathbb{E}\left[X_1^{p_1}\overline{X_1}^{q_1}\cdots X_r^{p_r}\overline{X_r}^{q_r}\right] = (q_i-1)\,\mathbb{E}\left[\overline{X_i^2}\right]\mathbb{E}\left[X_1^{p_1}\overline{X_1}^{q_1}\cdots X_i^{p_i}\overline{X_i}^{q_i-2}\cdots X_r^{p_r}\overline{X_r}^{q_r}\right]$$
$$+ \sum_{j=1}^{r} p_j\,\mathbb{E}\left[\overline{X_i}X_j\right]\mathbb{E}\left[X_1^{p_1}\overline{X_1}^{q_1}\cdots \overline{X_i^{q_i-1}}\cdots X_j^{p_j-1}\cdots X_r^{p_r}\overline{X_r}^{q_r}\right] \tag{3.36}$$
$$+ \sum_{\substack{j=1\\ j\neq i}}^{r} q_j\,\mathbb{E}\left[\overline{X_iX_j}\right]\mathbb{E}\left[X_1^{p_1}\overline{X_1}^{q_1}\cdots \overline{X_i^{q_i-1}}\cdots \overline{X_j^{q_j-1}}\cdots X_r^{p_r}\overline{X_r}^{q_r}\right]$$

with the convention that $X^{-1}=0$.
Conversely, if $X_1,\ldots,X_r$ *are* r *centered Gaussian variables satisfying* (3.35) *and* (3.36)*, then* $(X_1,\ldots,X_r)$ *is a centered complex Gaussian vector.*

Proof. In the same spirit as the proof of Lemma 3.4.10, we obtain (3.35) and (3.36) by derivating the Fourier transform. The converse is proved by induction. $\square$

IV

Fluctuation of matrix entries and application to outliers of elliptic matrices

Florent Benaych-Georges, Guillaume Cébron and Jean Rochet

Abstract :

For any family of $N \times N$ random matrices $(\mathbf{A}_k)_{k \in K}$ which is invariant, in law, under unitary conjugation, we give general conditions for central limit theorems of random variables of the type $\mathrm{Tr}(\mathbf{A}_k \mathbf{M})$, where the Euclidean norm of $\mathbf{M}$ has order $\sqrt{N}$ (such random variables include for example the normalized matrix entries $\sqrt{N}\mathbf{A}_k(i,j)$). A consequence is the asymptotic independence of the projection of the matrices $\mathbf{A}_k$ onto the subspace of null trace matrices from their projections onto the orthogonal of this subspace. This result is used to study the asymptotic behaviour of the outliers of a spiked elliptic random matrix. More precisely, we show that their fluctuations around their limits can have various rates of convergence, depending on the Jordan Canonical Form of the additive perturbation. Also, some correlations can arise between outliers at a macroscopic distance from each other. These phenomena have already been observed in [21] with random matrices from the Single Ring Theorem.

4.1 Introduction

This paper is firstly concerned with the fluctuations of sums of entries of unitarily invariant random matrices when the dimension tends to infinity and, secondly, with its application to the fluctuations of the outliers of spiked elliptic matrices.

The first problem is to find out conditions under which, for given collections $(\mathbf{A}_k)_{k \in K}$ of random matrices and $(\mathbf{M}_\ell)_{\ell \in L}$ of non-random matrices, the finite marginals of

$$\left(\mathrm{Tr}\left(\mathbf{A}_k \mathbf{M}_\ell\right) - \mathbb{E}\,\mathrm{Tr}\left(\mathbf{A}_k \mathbf{M}_\ell\right)\right)_{k \in K, \ell \in L} \tag{4.1}$$

converge as the dimension N tends to infinity. We shall always suppose that the $\mathbf{M}_\ell$'s have Euclidean norms of order $\sqrt{N}$, *i.e.* that the random variables

$$\frac{1}{N} \mathrm{Tr}\, \mathbf{M}_\ell \mathbf{M}_\ell^*$$

are bounded in probability. These hypotheses are satisfied by matrices

$$\mathbf{M}_\ell = \sqrt{N} \times (\text{an elementary matrix}) \tag{4.2}$$

but also by more complex matrices, like a matrices whose entries all have order $N^{-1/2}$, as Wigner matrices. In this framework, the main hypothesis we need for the random vector of (4.1) to be asymptotically Gaussian is the global invariance, in law, of $(\mathbf{A}_k)_{k \in K}$ under unitary conjugation, *i.e.* that for any unitary matrix $\mathbf{U}$,

$$\mathbf{A} \stackrel{\text{law}}{=} \left(\mathbf{U}\mathbf{A}_k \mathbf{U}^*\right)_{k \in K}.$$

It then appears that the question decomposes into two independent problems: one associated to the projections of the $\mathbf{A}_k$'s onto the space of null trace matrices (see Theorem 4.2.2) and one associated to the convergence of the centered traces of the

$\mathbf{A}_k$'s; and that both give rise to independent asymptotic fluctuations (see Theorem 4.2.5 and Corollary 4.2.6). These results extend an already proved partial result in this direction, Theorem 6.4 of [16] (see also Theorem 1.2 of [81] in the particular case of real symmetric matrices $\mathbf{A}_k$). The main advantages of Theorems 4.2.2 and 4.2.5 over the results of [81] and [16] is that they do not require the matrices $\mathbf{M}_\ell$ to be well approximated by matrices with finitely many non zero entries and that they give the asymptotic independence mentioned above. Besides, the technical hypotheses needed here are weaker than in the existing literature. Our proofs are based on the so-called *Weingarten calculus*, an integration method for the Haar measure on the unitary group developed by Collins and Śniady in [36, 38].

All these results belong to a long list of results begun in 1906 with the theorem by Borel [26] stating that any coordinate of a uniformly distributed random vector of the sphere of $\mathbb{R}^N$ with radius $\sqrt{N}$ is asymptotically standard Gaussian as $N \to \infty$, and continued e.g. with the papers [81, 3, 55, 65, 33, 37, 14, 31, 22] on central limit theorems on large symmetric spaces for the orthogonal and the unitary group. Some of the results of the previously cited papers can be deduced from Theorems 4.2.2 and 4.2.5. More generally, we notice in Corollary 4.2.7 that for collections of random matrices which are invariant, in law, under unitary conjugation, the convergence in second-order distribution implies that the matrix entries (or, more generally, random variables of the type of (4.1)) are asymptotically Gaussian. Second order freeness is a theory that has been developed these last ten years about Gaussian fluctuations (called *second order limits*) of traces of large random matrices, the most emblematic articles in this theory being [66, 68, 67, 34]. Our results cannot be deduced directly from this theory, because the "test matrices" we consider (*i.e.* the matrices $\mathbf{M}_\ell$) are not supposed to have second order limit distributions (and actually do not have any in many cases, as the one of (4.2)).

These general results about asymptotic fluctuations of matrix entries mentioned above are then applied to the fluctuations of the outliers of Gaussian elliptic matrices. It is easy to see that the global behavior of the spectrum of a large random matrix is not altered, from the macroscopic point of view, by a low rank additive perturbation. However, some of the eigenvalues, called *outliers*, can deviate away from the bulk, depending on the strength of the perturbation. Firstly brought to light for empirical covariance matrices by Johnstone in [57], this phenomenon, known as the *BBP transition*, was proved by Baik, Ben Arous and Péché in [9], and then shown under several hypothesis on the Hermitian case in [77, 43, 28, 29, 79, 19, 20, 16, 17, 30, 61, 62]. Non-Hermitian models have been also studied: i.i.d. matrices in [92, 24, 82], real elliptic matrices in [73], matrices from the Single Ring Theorem in [21] and also nearly Hermitian matrices [85, 75]. As an application of our main result, we investigate the fluctuations of the outliers and due to the non-Hermitian structure, we prove, as in [21, 24, 82, 85], that the distribution of the fluctuations highly depends on the shape of the Jordan Canonical Form of the perturbation, in particular, the convergence rate depends on the size of the Jordan blocks. Also, the outliers tend to locate around their limit at the vertices of a regular polygon (see Figure 4.12). At last, as in [21], we observe potential correlations between the fluctuations of outliers at a macroscopic distance with each other (see Remark 4.2.21).

The paper is organized as follows. In Section 4.2, we state our main results (Theorems 4.2.2, 4.2.5, 4.2.13 and 4.2.19) and their corollaries. These theorems are then proved in the following sections and an appendix is devoted to a technical result needed here.

4.2 Main results

4.2.1 General results

Let $\mathbf{A} = \left(\mathbf{A}_k\right)_{k\in K}$ be a collection of $N \times N$ random matrices and let $\left(\mathbf{M}_\ell\right)_{\ell\in L}$ be a collection $N \times N$ non random matrices, both implicitly depending on N.

Assumption 4.2.1.

(a) $\mathbf{A}$ is invariant in distribution under unitary conjugation: for any unitary matrix $\mathbf{U}$,

$$\mathbf{A} \overset{\text{law}}{=} \left(\mathbf{U}\mathbf{A}_k\mathbf{U}^*\right)_{k\in K};$$

(b) for each $k \in K$, and each $p, q \geq 1$, $\frac{1}{N}\operatorname{Tr}|\mathbf{A}_k|^{2p}$ is bounded in L^q independently of N;

(c) for each $k, k' \in K$, we have the following convergences, in L^2, to non random variables:

$$\lim_{N\to\infty} \frac{1}{N}\operatorname{Tr}\mathbf{A}_k\mathbf{A}_{k'} - \frac{1}{N}\operatorname{Tr}\mathbf{A}_k \cdot \frac{1}{N}\operatorname{Tr}\mathbf{A}_{k'} = \tau(k,k')$$

and

$$\lim_{N\to\infty} \frac{1}{N}\operatorname{Tr}\mathbf{A}_k\mathbf{A}_{k'}^* - \frac{1}{N}\operatorname{Tr}\mathbf{A}_k \cdot \frac{1}{N}\operatorname{Tr}\mathbf{A}_{k'}^* = \tau(k,\overline{k'});$$

(d) for each $\ell, \ell' \in L$, we have the following convergences:

$$\lim_{N\to\infty} \frac{1}{N}\operatorname{Tr}\mathbf{M}_\ell\mathbf{M}_{\ell'} - \frac{1}{N}\operatorname{Tr}\mathbf{M}_\ell \cdot \frac{1}{N}\operatorname{Tr}\mathbf{M}_{\ell'} = \eta_{\ell\ell'} \tag{4.3}$$

and

$$\lim_{N\to\infty} \frac{1}{N}\operatorname{Tr}\mathbf{M}_\ell\mathbf{M}_{\ell'}^* - \frac{1}{N}\operatorname{Tr}\mathbf{M}_\ell \cdot \frac{1}{N}\operatorname{Tr}\mathbf{M}_{\ell'}^* = \beta_{\ell\ell'}. \tag{4.4}$$

Under this sole hypothesis, we first have the following result, focused on the case where the $\mathbf{M}_\ell$'s all have null trace, *i.e.* focused on the projections of the $\mathbf{A}_k$'s onto the space of such matrices.

Theorem 4.2.2. *Under Hypothesis 4.2.1, if, for each ℓ, $\operatorname{Tr}(\mathbf{M}_\ell) = 0$, the finite marginals of*

$$\left(\operatorname{Tr}\left(\mathbf{A}_k\mathbf{M}_\ell\right)\right)_{k\in K,\ell\in L} \tag{4.5}$$

converge to the ones of a complex centered Gaussian vector $\left(\mathscr{G}_{k,\ell}\right)_{k\in K,\ell\in L}$ *with covariance*

$$\mathbb{E}\left[\mathscr{G}_{k,\ell}\mathscr{G}_{k',\ell'}\right]=\eta_{\ell\ell'}\tau(k,k'),\qquad \textit{and}\qquad \mathbb{E}\left[\mathscr{G}_{k,\ell}\overline{\mathscr{G}_{k',\ell'}}\right]=\beta_{\ell\ell'}\tau(k,\overline{k'}).$$

Remark 4.2.3. Note that by invariance of the distribution of $\mathbf{A}$ under unitary conjugation, we have

$$\mathbb{E}\operatorname{Tr}\left(\mathbf{A}_k\mathbf{M}_\ell\right)=\mathbb{E}\left(\frac{1}{N}\operatorname{Tr}\mathbf{A}_k\right)\operatorname{Tr}\mathbf{M}_\ell,$$

hence the random variables of (4.5) are centered and the ones of (4.7) below rewrite

$$\operatorname{Tr}\left(\mathbf{A}_k\mathbf{M}_\ell\right)-\mathbb{E}\left(\frac{1}{N}\operatorname{Tr}\mathbf{A}_k\right)\operatorname{Tr}\mathbf{M}_\ell.$$

The following theorem gives the joint fluctuations of the projections of the $\mathbf{A}_k$'s on null trace matrices and of their traces.

Assumption 4.2.4. The finite marginals of the process $\left(\operatorname{Tr}\mathbf{A}_k-\mathbb{E}\operatorname{Tr}\mathbf{A}_k\right)_{k\in K}$ converge to those of a random centered vector $(\mathscr{T}_k)_{k\in K}$ and for each $\ell\in L$, there is $\alpha_\ell\in\mathbb{C}$ such that

$$\lim_{N\to\infty}\frac{1}{N}\operatorname{Tr}\mathbf{M}_\ell=\alpha_\ell. \tag{4.6}$$

Theorem 4.2.5. *Under Hypotheses 4.2.1 and 4.2.4, the finite marginals of*

$$\left(\operatorname{Tr}\left(\mathbf{A}_k\mathbf{M}_\ell\right)-\mathbb{E}\operatorname{Tr}\left(\mathbf{A}_k\mathbf{M}_\ell\right)\right)_{k\in K,\ell\in L} \tag{4.7}$$

converge to the ones of $\left(\mathscr{G}_{k,\ell}+\alpha_\ell\mathscr{T}_k\right)_{k\in K,\ell\in L}$ *where* $\left(\mathscr{G}_{k,\ell}\right)_{k\in K,\ell\in L}$ *is a complex centered Gaussian vector independent from* $\left(\mathscr{T}_k\right)_{k\in K}$ *and with covariance*

$$\mathbb{E}\left[\mathscr{G}_{k,\ell}\mathscr{G}_{k',\ell'}\right]=\eta_{\ell\ell'}\tau(k,k'),\qquad \textit{and}\qquad \mathbb{E}\left[\mathscr{G}_{k,\ell}\overline{\mathscr{G}_{k',\ell'}}\right]=\beta_{\ell\ell'}\tau(k,\overline{k'}).$$

A consequence of this theorem is the asymptotic independence of the projection of the matrices $\mathbf{A}_k$ onto the subspace of null trace matrices from their projections onto the orthogonal of this subspace:

Corollary 4.2.6. *Under Hypotheses 4.2.1 and 4.2.4, suppose that for any* $\ell\in L$, *we have* $\operatorname{Tr}(\mathbf{M}_\ell)=0$. *Then the processes*

$$\left(\operatorname{Tr}\left(\mathbf{A}_k\mathbf{M}_\ell\right)\right)_{k\in K,\ell\in L}\qquad \textit{and}\qquad \left(\operatorname{Tr}\mathbf{A}_k-\mathbb{E}\operatorname{Tr}\mathbf{A}_k\right)_{k\in K}$$

are asymptotically independent.

4.2.2 Second-order freeness implies fluctuations of matrix elements

The following corollary of Theorem 4.2.5 is obvious. Let $\mathbb{C}\langle x_k, x_k^*, k \in K\rangle$ denote the space of polynomials in the non commutative variables x_k, x_k^*, indexed by $k \in K$.

Corollary 4.2.7. *Let $\left(\mathbf{A}_k\right)_{k\in K}$ be a collection of $N\times N$ random matrices which is invariant by unitary conjugation and which converges in second order $*$-distribution to some family $a = (a_k)_{k\in K}$ in $(\mathscr{A}, \tau_1, \tau_2)$ as $N \to \infty$. Let $\left(\mathbf{M}_\ell\right)_{\ell\in L}$ be a collection non random matrices satisfying (4.3), (4.4) and (4.6).*

Then the finite marginals of

$$\left(\mathrm{Tr}\left(P(\mathbf{A})\mathbf{M}_\ell\right) - \mathbb{E}\,\mathrm{Tr}\left(P(\mathbf{A})\mathbf{M}_\ell\right)\right)_{P\in\mathbb{C}\langle x_k, x_k^*, k\in K\rangle, \ell\in L} \tag{4.8}$$

converge to the ones of a complex centered Gaussian vector

$$\left(\mathscr{H}_{P,\ell}\right)_{P\in\mathbb{C}\langle x_k, x_k^*, k\in K\rangle, \ell\in L}$$

such that, for all $P, Q \in \mathbb{C}\langle x_k, x_k^*, k \in K\rangle$ and $\ell, \ell' \in L$,

$$\begin{aligned}
\mathbb{E}\,\mathscr{H}_{P,\ell}\mathscr{H}_{Q,\ell'} &= \alpha_\ell\alpha_{\ell'}\tau_2(P(a), Q(a)) + \eta_{\ell\ell'}\left(\tau_1(P(a)Q(a)) - \tau_1(P(a))\tau_1(Q(a))\right), \\
\mathbb{E}\,\mathscr{H}_{P,\ell}\overline{\mathscr{H}_{Q,\ell'}} &= \alpha_\ell\overline{\alpha_{\ell'}}\tau_2(P(a), Q(a)^*) + \beta_{\ell\ell'}\left(\tau_1(P(a)Q(a)^*) - \tau_1(P(a))\tau_1(Q(a)^*)\right).
\end{aligned}$$

Remark 4.2.8. The following matrices have been shown to converge in second order $*$-distribution:

- GUE matrices or so-called matrix models where the entries interact via a potential [56],
- Ginibre matrices [83],
- matrices arising from the Haar measure on the unitary group $U(N)$ [39],
- matrices arising from the heat kernel measure on $U(N)$ [63] and on $GL_N(\mathbb{C})$ [31],
- Wishart matrices, matrices of the type $\mathbf{UAV}$ or $\mathbf{UAU}^*$, with $\mathbf{U}, \mathbf{V}$ independent and Haar distributed on $U(N)$ and $\mathbf{A}$ deterministic with a limit spectral distribution [66, 68, 67, 34].

A consequence of Corollary 4.2.7 is that any non commutative polynomial in independent random matrices taken from the list above has asymptotically Gaussian entries, which are independent modulo a possible symmetry.

4.2.3 Left and right unitary invariant matrices

Here is another corollary on random matrices invariant by left and right unitary multiplication.

Corollary 4.2.9. *Let $\mathbf{A} = \left(\mathbf{A}_k\right)_{k\in K}$ be a collection of $N\times N$ random matrices such that:*

- *(a') $\mathbf{A}$ is invariant by left and right multiplication by unitary matrix: for any unitary matrix $\mathbf{U}$, $\mathbf{A} \overset{law}{=} \left(\mathbf{U}\mathbf{A}_k\right)_{k\in K} \overset{law}{=} \left(\mathbf{A}_k\mathbf{U}\right)_{k\in K}$;*
- *(b') for each k and each p,q, $\frac{1}{N}\operatorname{Tr}|\mathbf{A}_k|^{2p}$ is bounded in L^q independently of N;*
- *(c') for each k,k', the sequence $\frac{1}{N}\operatorname{Tr}\mathbf{A}_k\mathbf{A}_{k'}^*$ converges in L^2 to some non random limits denoted $\tau(k,\overline{k'})$.*

Let $\left(\mathbf{M}_\ell\right)_{\ell\in L}$ be a collection non random matrices satisfying (4.3), (4.4) and (4.6). Then the finite marginals of

$$(\operatorname{Tr}(\mathbf{A}_k\mathbf{M}_\ell))_{k\in K,\ell\in L}$$

converge to the ones of a complex centered Gaussian vector $\left(\mathscr{G}_{k,\ell}\right)_{k\in K,\ell\in L}$ with covariance

$$\mathbb{E}\mathscr{G}_{k,\ell}\mathscr{G}_{k',\ell'} = 0, \quad \textit{and} \quad \mathbb{E}\mathscr{G}_{k,\ell}\overline{\mathscr{G}_{k',\ell'}} = \beta_{\ell,\ell'}\tau(k,\overline{k'}).$$

The proof of this corollary is postponed to Section 4.3.4: we show that the hypotheses of the corollary imply Hypotheses 4.2.1 and 4.2.4.

4.2.4 Permutation matrix entries under randomized basis

In [93], the individual entries of a uniform random $N\times N$ permutation matrix S conjugated by a uniform random orthogonal matrix are studied. The limit distribution of the number of d-cycles, indexed by $d\geq 1$, is known to be asymptotically, as $N\to\infty$, a Poisson process $(\mathscr{Z}_d)_{d\geq 1}$ on the set of positive integers with intensity $1/d$ by [4], which means in particular that each trace $\operatorname{Tr}(S^k)$, where $k\geq 1$, converges in distribution to $\sum_{d|k} d\mathscr{Z}_d$. Thanks to Theorem 4.2.5 and Remark 4.2.3, we deduce from this limiting Poisson distribution the following result about the matrix entries of a uniform permutation matrix S conjugated by a uniform unitary matrix.

Corollary 4.2.10. *Let S be a $N\times N$ random permutation matrix which is uniformly distributed, U be a $N\times N$ random unitary matrix which is Haar distributed , and $\left(\mathbf{M}_\ell\right)_{\ell\in L}$ a collection of non random matrices satisfying (4.3), (4.4) and (4.6). Then the finite marginals of*

$$\left(\operatorname{Tr}\left(US^kU^*\mathbf{M}_\ell\right)\right)_{k\geq 1,\ell\in L}$$

converge to the ones of $\left(\mathscr{G}_{k,\ell}+\alpha_\ell\sum_{d|k} d\mathscr{Z}_d\right)_{k\geq 1,\ell\in L}$*, where* $\left(\mathscr{G}_{k,\ell}\right)_{k\geq 1,\ell\in L}$ *is a complex centered Gaussian vector with covariance*

$$\mathbb{E}\mathscr{G}_{k,\ell}\mathscr{G}_{k',\ell'} \;=\; 0, \quad \textit{and} \quad \mathbb{E}\mathscr{G}_{k,\ell}\overline{\mathscr{G}_{k',\ell'}} \;=\; \mathbb{1}_{k=k'}\beta_{\ell,\ell'},$$

and $(\mathscr{Z}_d)_{d\geq 1}$ *is a Poisson process on the set of positive integers with intensity* $1/d$ *which is independent from* $\left(\mathscr{G}_{k,\ell}\right)_{k\in\mathbb{N},\ell\in L}$*.*

4.2.5 Low rank perturbation for Gaussian elliptic matrices

Matrices from the *Gaussian elliptic ensemble*, first introduced in [89], can be defined as follows.

Definition 4.2.11. A *Gaussian elliptic matrix of parameter* $\rho\in[-1,1]$ is a random matrix $\mathbf{Y}=[y_{ij}]_{i,j=1}^N$ such that

- $\{(y_{ij},y_{ji}),1\leq i<j\leq N\}\cup\{y_{ii},1\leq i\leq N\}$ is a family of independent random vectors,
- $\{(y_{ij},y_{ji}),1\leq i<j\leq N\}$ are i.i.d. Gaussian such that

$$\mathbb{E}y_{ij}^2=\mathbb{E}y_{ji}^2=\mathbb{E}y_{ij}\overline{y_{ji}}=0, \quad \mathbb{E}|y_{ij}|^2=\mathbb{E}|y_{ji}|^2=1 \quad \text{and} \quad \mathbb{E}y_{ij}y_{ji}=\rho$$

- $\{y_{ii},1\leq i\leq N\}$ are i.i.d. Gaussian such that

$$\mathbb{E}y_{ii}^2=\rho \quad \text{and} \quad \mathbb{E}|y_{ii}|^2=1.$$

Remark 4.2.12. Gaussian elliptic matrices can be seen as an generalization of *GUE matrices* and the *Ginibre matrices*. Indeed, a Gaussian elliptic matrix $\mathbf{Y}$ of parameter ρ can be realized as

$$\mathbf{Y} \;=\; \sqrt{\frac{1+\rho}{2}}\mathbf{H}_1+\mathrm{i}\sqrt{\frac{1-\rho}{2}}\mathbf{H}_2,$$

where $\mathbf{H}_1$ and $\mathbf{H}_2$ are two independent GUE matrices from the *GUE*. Hence GUE matrices (resp. Ginibre matrices) are Gaussian elliptic matrices of parameter 1 (resp. 0).

One can also define more general *elliptic random matrices* (see [70, 71, 73, 74] for more details). In our case, it is easy to see that the Gaussian elliptic ensemble is invariant in distribution by unitary conjugation (see for example Remark 4.2.12) which allows us to use our Theorem 4.2.5 for this model. In this section, we are interested in the outliers in the spectrum of these matrices. Indeed, it is known (see [89]) that when you renormalize by $\frac{1}{\sqrt{N}}$ the matrix $\mathbf{Y}$, its limiting eigenvalue distribution is the uniform measure μ_ρ on the ellipse

$$\mathscr{E}_\rho \;:=\; \left\{z\in\mathbb{C}\,;\;\frac{\operatorname{Re}(z)^2}{(1+\rho)^2}+\frac{\operatorname{Im}(z)^2}{(1-\rho)^2}\;\leq\;1\right\}. \tag{4.9}$$

Also, we know that adding a finite rank matrix $\mathbf{P}$ to such a matrix $\mathbf{Y}$ barely alters its spectrum (see [71, Theorem 1.8]). In other words, most of the eigenvalues remain distributed according to μ_ρ but the perturbation $\mathbf{P}$ may give rise to outliers. The generic location of the outliers has already been studied for elliptic random matrices (see [73]), but the authors did not consider the fluctuations.

For all $N \geq 1$, let $\mathbf{X}_N := \frac{1}{\sqrt{N}}\mathbf{Y}_N$ where $\mathbf{Y}_N$ is an $N\times N$ Gaussian elliptic matrix of parameter ρ and let $\mathbf{P}_N$ be a $N\times N$ random matrix independent from $\mathbf{X}_N$ whose rank is bounded by an integer r (independent from N). We consider the additive pertubation

$$\widetilde{\mathbf{X}}_N \;:=\; \frac{1}{\sqrt{N}}\mathbf{Y}_N + \mathbf{P}_N \;=\; \mathbf{X}_N + \mathbf{P}_N.$$

Since, for any unitary matrix $\mathbf{U}$ which is independent from $\mathbf{X}_N$, we have $\mathbf{X}_N \overset{(d)}{=} \mathbf{U}\mathbf{X}_N\mathbf{U}^*$, we can assume that $\mathbf{P}_N$ has the following block structure

$$\mathbf{P}_N \;=\; \begin{pmatrix} \mathbf{P} & 0 \\ 0 & 0 \end{pmatrix}, \qquad \text{where } \mathbf{P} \text{ is a } 2r\times 2r \text{ matrix}$$

(indeed, any complex matrix is unitarily similar to a upper triangular matrix and since the rank of $\mathbf{P}_N$ is lower than r, we have $\dim(\operatorname{Im}\mathbf{P}_N + (\operatorname{Ker}\mathbf{P}_N)^\perp) \leq 2r$).

Theorem 4.2.13 (Outliers for finite rank perturbations of a Gaussian elliptic matrix)**.** *Let $\varepsilon > 0$. Suppose that $\mathbf{P}_N$ does not have any eigenvalue λ such that*

$$|\lambda| \;>\; 1 \quad \text{and} \quad 1+|\rho|+\varepsilon \;<\; \left|\lambda + \frac{\rho}{\lambda}\right| \;<\; 1+|\rho|+3\varepsilon, \tag{4.10}$$

and has exactly $j \leq r$ eigenvalues $\lambda_1(\mathbf{P}_N),\dots,\lambda_j(\mathbf{P}_N)$ (counted with multiplicity) such that, for each $i = 1,\dots,j$,

$$|\lambda_i(\mathbf{P}_N)| \;>\; 1 \quad \text{and} \quad \left|\lambda_i(\mathbf{P}_N) + \frac{\rho}{\lambda_i(\mathbf{P}_N)}\right| \;>\; 1+|\rho|+3\varepsilon. \tag{4.11}$$

Then, with probability tending to one, $\widetilde{\mathbf{X}}_N := \mathbf{X}_N + \mathbf{P}_N$ possesses exactly j eigenvalues $\widetilde{\lambda}_1,\dots,\widetilde{\lambda}_j$ in $\{z\in\mathbb{C}\,;\,|z| > 1+|\rho|+2\varepsilon\}$ and after a proper labeling

$$\widetilde{\lambda}_i \;=\; \lambda_i(\mathbf{P}_N) + \frac{\rho}{\lambda_i(\mathbf{P}_N)} + o(1), \tag{4.12}$$

for each $1\leq i\leq j$.

Remark 4.2.14. In [73], the authors prove this result for *real elliptic random matrices* and have a more precise statement. Indeed, they replace in our conditions (4.10) and (4.11) the area $\{z\in\mathbb{C}\,;\,1+|\rho|+\varepsilon < |z| < 1+|\rho|+3\varepsilon\}$ (resp. $\{z\in\mathbb{C},\ |z| > 1+|\rho|+3\varepsilon\}$) by $\mathscr{E}_{\rho,3\varepsilon}\backslash\mathscr{E}_{\rho,\varepsilon}$ (resp. by $\mathscr{E}^c_{\rho,3\varepsilon}$) where $\mathscr{E}_{\rho,\varepsilon}$ is a ε-neighborhood of the ellipse $\mathscr{E}_\rho$ (see (4.9)). But our proof relies on the identity

$$\operatorname{Tr}(z-\mathbf{X})^{-1} \;=\; \sum_{k\geq 0} z^{-k-1}\operatorname{Tr}\mathbf{X}^k,$$

this is why (4.10) and (4.11) are circular conditions, instead of elliptic ones.

To study the fluctuations of the outliers $\widetilde{\lambda}_i$ around their generic locations as given by (4.12), we need to specify the shape of the matrix $\mathbf{P}$ as it is done in [21]. Indeed, since $\mathbf{P}$ is not Hermitian, we need to introduce its Jordan Canonical Form (JCF) which does not depend on N. We know that, in a proper basis, $\mathbf{P}$ is a direct sum of *Jordan blocks*, i.e. blocks of the form

$$\mathbf{R}_p(\theta) = \begin{pmatrix} \theta & 1 & & (0) \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ (0) & & & \theta \end{pmatrix}, \; p\times p \text{ matrix}, \qquad \theta\in\mathbb{C}, p\geq 1 \tag{4.13}$$

Let us denote by $\theta_1,\dots,\theta_q$ the distinct eigenvalues of $\mathbf{P}$ satisfying condition (4.11). For convenience, we shall write from now on

$$\hat{\theta}_i := \theta_i + \frac{\rho}{\theta_i}. \tag{4.14}$$

We introduce a positive integer α_i, some positive integers $p_{i,1} > \cdots > p_{i,\alpha_i}$ corresponding to the distinct sizes of the blocks relative to the eigenvalue θ_i and $\beta_{i,1},\dots,\beta_{i,\alpha_i}$ such that for all j, $\mathbf{R}_{p_{i,j}}(\theta_i)$ appears $\beta_{i,j}$ times, so that, for a certain $2r\times 2r$ non singular matrix $\mathbf{Q}$, we have:

$$\mathbf{J} = \mathbf{Q}^{-1}\mathbf{P}\mathbf{Q} = \hat{\mathbf{P}} \bigoplus \bigoplus_{i=1}^{q} \bigoplus_{j=1}^{\alpha_i} \underbrace{\begin{pmatrix} \mathbf{R}_{p_{i,j}}(\theta_i) & & \\ & \ddots & \\ & & \mathbf{R}_{p_{i,j}}(\theta_i) \end{pmatrix}}_{\beta_{i,j} \text{ blocks}} \tag{4.15}$$

where $\oplus$ is defined, for square block matrices, by $\mathbf{M}\oplus\mathbf{N} := \begin{pmatrix} \mathbf{M} & 0 \\ 0 & \mathbf{N} \end{pmatrix}$ and $\hat{\mathbf{P}}$ is a matrix such that its eigenvalues θ are such that $|\theta| < 1$ or $|\hat{\theta}| = |\theta+\rho\theta^{-1}| < 1+\rho+\varepsilon$.

The asymptotic orders of the fluctuations of the eigenvalues of $\mathbf{X}_N+\mathbf{P}_N$ depend on the sizes $p_{i,j}$ of the blocks. We know, by Theorem 4.2.13, there are $\sum_{j=1}^{\alpha_i} p_{ij}\times\beta_{i,j}$ eigenvalues $\widetilde{\lambda}$ of $\mathbf{X}_N+\mathbf{P}_N$ which tend to $\hat{\theta}_i = \theta_i+\rho\theta_i^{-1}$: we shall write them with a $\hat{\theta}_i$ on the top left corner, as follows

$${}^{\hat{\theta}_i}\widetilde{\lambda}.$$

Theorem 4.2.19 below will state that, for each block with size $p_{i,j}$ corresponding to θ_i in the JCF of $\mathbf{P}$, there are $p_{i,j}$ eigenvalues (we shall write them with $p_{i,j}$ on the bottom left corner: ${}^{\hat{\theta}_i}_{p_{i,j}}\widetilde{\lambda}$) whose convergence rate will be $N^{-1/(2p_{i,j})}$. As there are $\beta_{i,j}$ blocks of size $p_{i,j}$, there are actually $p_{i,j} \times \beta_{i,j}$ eigenvalues tending to $\hat{\theta}_i$ with convergence rate $N^{-1/(2p_{i,j})}$ (we shall write them ${}^{\hat{\theta}_i}_{p_{i,j}}\widetilde{\lambda}_{s,t}$ with $s \in \{1,\ldots,p_{i,j}\}$ and $t \in \{1,\ldots,\beta_{i,j}\}$). It would be convenient to denote by $\Lambda_{i,j}$ the vector with size $p_{i,j} \times \beta_{i,j}$ defined by

$$\Lambda_{i,j} := \left(N^{1/(2p_{i,j})} \cdot \left({}^{\hat{\theta}_i}_{p_{i,j}}\widetilde{\lambda}_{s,t} - \hat{\theta}_i\right)\right)_{\substack{1 \le s \le p_{i,j} \\ 1 \le t \le \beta_{i,j}}}. \tag{4.16}$$

As in [21], we define now the family of random matrices that we shall use to characterize the limit distribution of the $\Lambda_{i,j}$'s. For each $i = 1,\ldots,q$, let $I(\theta_i)$ (resp. $J(\theta_i)$) denote the set, with cardinality $\sum_{j=1}^{\alpha_i} \beta_{i,j}$, of indices in $\{1,\ldots,2r\}$ corresponding to the first (resp. last) columns of the blocks $\mathbf{R}_{p_{i,j}}(\theta_i)$ $(1 \le j \le \alpha_i)$ in (4.15).

Remark 4.2.15. Note that the columns of $\mathbf{Q}$ (resp. of $(\mathbf{Q}^{-1})^*$) whose index belongs to $I(\theta_i)$ (resp. $J(\theta_i)$) are eigenvectors of $\mathbf{P}$ (resp. of $\mathbf{P}^*$) associated to θ_i (resp. $\overline{\theta_i}$). See [21, Remark 2.8].

Now, let

$$\left(m^{\theta_i}_{k,\ell}\right)_{\substack{1 \le i \le q \\ (k,\ell) \in J(\theta_i) \times I(\theta_i)}} \tag{4.17}$$

be the random centered complex Gaussian vector with covariance

$$\begin{aligned} \mathbb{E}\left[m^{\theta_i}_{k,\ell} m^{\theta_{i'}}_{k',\ell'}\right] &= \left(\frac{1}{\theta_i\theta_{i'} - \rho} - \frac{1}{\theta_i\theta_{i'}}\right)\delta_{k,\ell'}\delta_{k',\ell} \\ \mathbb{E}\left[m^{\theta_i}_{k,\ell} \overline{m^{\theta_{i'}}_{k',\ell'}}\right] &= \Phi_\rho(\hat{\theta}_i, \hat{\theta}_{i'})\mathbf{e}_k \mathbf{Q}^{-1}(\mathbf{Q}^{-1})^* \mathbf{e}_{k'} \cdot \mathbf{e}_{\ell'} \mathbf{Q}^* \mathbf{Q} \mathbf{e}_\ell, \end{aligned} \tag{4.18}$$

where $\mathbf{e}_1,\ldots,\mathbf{e}_{2r}$ are the column vectors of the canonical basis of $\mathbb{C}^{2r}$ and

$$\Phi_\rho(z,z') = \int \frac{1}{z-w}\frac{1}{\overline{z'}-\overline{w}}\mu_\rho(dw) - \int \frac{1}{z-w}\mu_\rho(dw)\int \frac{1}{\overline{z'}-\overline{w}}\mu_\rho(dw).$$

Remark 4.2.16. We will see in section 4.3.5 that (4.17) is in fact the limit distribution of

$$\left(\sqrt{N}\mathbf{e}_k^*\mathbf{Q}^{-1}\left(\left(\hat{\theta}_i - \mathbf{X}_N\right)^{-1} - \frac{1}{\theta_i}\right)\mathbf{Q}\mathbf{e}_\ell\right)_{\substack{1 \le i \le q \\ (k,\ell) \in J(\theta_i) \times I(\theta_i)}} \xrightarrow{(\mathrm{d})} \left(m^{\theta_i}_{k,\ell}\right)_{\substack{1 \le i \le q \\ (k,\ell) \in J(\theta_i) \times I(\theta_i)}}$$

This convergence is a consequence of Theorem 4.2.5.

Remark 4.2.17. When $\rho = 0$, one has

$$\begin{aligned}\Phi_0(z,z') &= \frac{1}{\pi}\int_{|w|\le 1}\frac{1}{z-w}\frac{1}{\overline{z'}-\overline{w}}dw - \frac{1}{\pi}\int_{|w|\le 1}\frac{1}{z-w}dw\frac{1}{\pi}\int_{|w|\le 1}\frac{1}{\overline{z'}-\overline{w}}dw \\ &= \frac{1}{z\overline{z'}-1} - \frac{1}{z\overline{z'}} = \frac{1}{z\overline{z'}(z\overline{z'}-1)}.\end{aligned}$$

We recover the expression of the covariance in the Ginibre case (see [21]). Also, the expression of Φ_1 corresponds to the covariance in the GUE case (see [85]).

For each i,j, let $K(i,j)$ (resp. $K(i,j)^-$) be the set, with cardinality $\beta_{i,j}$ (resp. $\sum_{j'=1}^{j-1}\beta_{i,j'}$), of indices in $J(\theta_i)$ corresponding to a block of the type $\mathbf{R}_{p_{i,j}}(\theta_i)$ (resp. to a block of the type $\mathbf{R}_{p_{i,j'}}(\theta_i)$ for $j' < j$). In the same way, let $L(i,j)$ (resp. $L(i,j)^-$) be the set, with the same cardinality as $K(i,j)$ (resp. as $K(i,j)^-$), of indices in $I(\theta_i)$ corresponding to a block of the type $\mathbf{R}_{p_{i,j}}(\theta_i)$ (resp. to a block of the type $\mathbf{R}_{p_{i,j'}}(\theta_i)$ for $j' < j$). Note that $K(i,j)^-$ and $L(i,j)^-$ are empty if $j = 1$. Let us define the random matrices

$$\begin{aligned}\mathrm{M}_j^{\theta_i,\mathrm{I}} &:= [m_{k,\ell}^{\theta_i,n}]_{\substack{k\in K(i,j)^- \\ \ell\in L(i,j)^-}} & \mathrm{M}_j^{\theta_i,\mathrm{II}} &:= [m_{k,\ell}^{\theta_i}]_{\substack{k\in K(i,j)^- \\ \ell\in L(i,j)}} \\ \mathrm{M}_{j,n}^{\theta_i,\mathrm{III}} &:= [m_{k,\ell}^{\theta_i}]_{\substack{k\in K(i,j) \\ \ell\in L(i,j)^-}} & \mathrm{M}_j^{\theta_i,\mathrm{IV}} &:= [m_{k,\ell}^{\theta_i,n}]_{\substack{k\in K(i,j) \\ \ell\in L(i,j)}}\end{aligned} \tag{4.19}$$

and then let us define the matrix $\mathbf{M}_j^{\theta_i}$ as

$$\mathbf{M}_{j,n}^{\theta_i} := \theta_i\left(\mathrm{M}_j^{\theta_i,\mathrm{IV}} - \mathrm{M}_j^{\theta_i,\mathrm{III}}\left(\mathrm{M}_j^{\theta_i,\mathrm{I}}\right)^{-1}\mathrm{M}_j^{\theta_i,\mathrm{II}}\right) \tag{4.20}$$

Remark 4.2.18. It follows from the fact that the matrix $\mathbf{Q}$ is invertible, that $\mathrm{M}_j^{\theta_i,\mathrm{I}}$ is a.s. invertible and so is $\mathbf{M}_j^{\theta_i}$.

Now, we can state the result on the fluctuations.

Theorem 4.2.19. *1. As N goes to infinity, the random vector*

$$\left(\Lambda_{i,j}\right)_{\substack{1\le i\le q \\ 1\le j\le \alpha_i}}$$

defined at (4.16) converges to the distribution of a random vector

$$\left(\Lambda_{i,j}^\infty\right)_{\substack{1\le i\le q \\ 1\le j\le \alpha_i}}$$

with joint distribution defined by the fact that, for each $1 \le i \le q$ and $1 \le j \le \alpha_i$, $\Lambda_{i,j}^\infty$ is the collection of the ${p_{i,j}}^{th}$ roots of the eigenvalues of some random matrix $\mathbf{M}_j^{\theta_i}$.

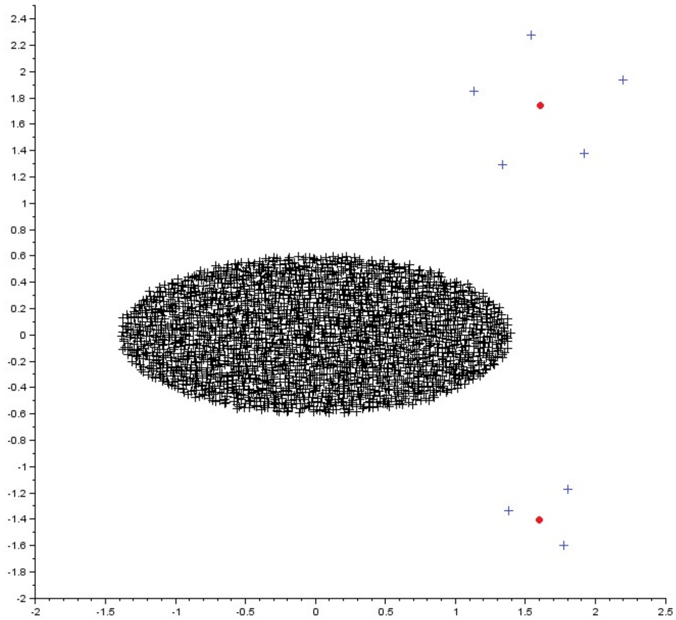


Figure 4.12: Spectrum of a Gaussian elliptic matrix of size $N = 2500$ with perturbation matrix $\mathbf{P} = \operatorname{diag}(\mathbf{R}_5(1.5+2.625\mathrm{i}), \mathbf{R}_3(1.5-1.5\mathrm{i}), 0, \ldots, 0)$. We see the blue crosses "+" (outliers) forming respectively a regular pentagon and an equilateral triangle around the red dots "•" (their limit). We also see a significant difference between the two rates of convergence, $N^{-1/10}$ and $N^{-1/6}$.

2. *The distributions of the random matrices $\mathbf{M}_j^{\theta_i}$ are absolutely continuous with respect to the Lebesgue measure and the random vector $\left(\Lambda_{i,j}^{\infty}\right)_{\substack{1\le i\le q\\ 1\le j\le \alpha_i}}$ has no deterministic coordinate.*

Remark 4.2.20. Each non zero complex number has exactly $p_{i,j}$ $p_{i,j}{}^{\text{th}}$ roots, drawing a regular $p_{i,j}$-sided polygon. Moreover, by the second part of the theorem, the spectrums of the $\mathbf{M}_j^{\theta_i}$'s almost surely do not contain 0, so each $\Lambda_{i,j}^{\infty}$ is actually a complex random vector with $p_{i,j} \times \beta_{i,j}$ coordinates, which draw $\beta_{i,j}$ regular $p_{i,j}$-sided polygons.

Remark 4.2.21. We notice that in the particular case where the matrix $\mathbf{Q}$ is unitary, the covariance of the Gaussian variables $\left(m_{k,\ell}^{\theta_i}\right)_{\substack{1\leq i\leq q\\(k,\ell)\in J(\theta_i)\times I(\theta_i)}}$ can be rewritten

$$\begin{aligned}\mathbb{E}\left[m_{k,\ell}^{\theta_i}m_{k',\ell'}^{\theta_{i'}}\right] &= \left(\frac{1}{\theta_i\theta_{i'}-\rho}-\frac{1}{\theta_i\theta_{i'}}\right)\delta_{k,\ell'}\delta_{k',\ell},\\ \mathbb{E}\left[m_{k,\ell}^{\theta_i}\overline{m_{k',\ell'}^{\theta_{i'}}}\right] &= \Phi(\hat\theta_i,\hat\theta_{i'})\delta_{k,k'}\delta_{\ell,\ell'}.\end{aligned}$$

Which means that for any i,i' such that $i\neq i'$, the family $\left(m_{k,\ell}^{\theta_i}\right)_{(k,\ell)\in J(\theta_i)\times I(\theta_i)}$ is independent from $\left(m_{k,\ell}^{\theta_{i'}}\right)_{(k,\ell)\in J(\theta_{i'})\times I(\theta_{i'})}$. Indeed, since the Jordan blocks associated to θ_i are distinct with those associated to $\theta_{i'}$, the sets $I(\theta_i)$ and $J(\theta_i)$ don't share any common index with $I(\theta_{i'})$ and $J(\theta_{i'})$. We can deduce that in this particular case, all the fluctuations around θ_i are independent from those around $\theta_{i'}$ (see [21, section 2.3.1.] for more details).
However, in the general case, there is no particular reason to have independance between the fluctuations around two spikes at macroscopique distance. To illustrate this phenomenon, we can take the same particular example than [21, Example 2.17] since a Ginibre matrix is also an Gaussian elliptic matrix. In this example, the authors of [21] took a matrix $\mathbf{P}$ of the form

$$\mathbf{P} = \mathbf{Q}\begin{pmatrix}\theta & 0\\ 0 & \theta'\end{pmatrix}\mathbf{Q}^{-1}, \qquad \mathbf{Q} = \begin{pmatrix}1 & \kappa\\ \kappa & 1\end{pmatrix}, \quad \kappa\neq\pm 1,$$

and they empirically confirmed that, in the case $\kappa\neq 0$, the fluctuations of the outliers around θ are correlated with these around θ'.

4.3 Proofs of Theorem 4.2.2 and Theorem 4.2.5

4.3.1 Preliminary result

Let $\left(\mathbf{B}_k\right)_{k\in K}$ be a collection of (implicitly depending on N) $N\times N$ random matrices such that

(i) for each $k\in K$, almost surely, $\operatorname{Tr}\mathbf{B}_k=0$;

(ii) for each $k\in K$, and each $p,q\geq 1$, $\frac{1}{N}\operatorname{Tr}|\mathbf{B}_k|^{2p}$ is bounded in L^q independently of N;

(iii) for each $k,k'\in K$, we have the following convergences to nonrandom variables in L^2

$$\lim_{N\to\infty}\frac{1}{N}\operatorname{Tr}\mathbf{B}_k\mathbf{B}_{k'}=\tau(k,k') \text{ and } \lim_{N\to\infty}\frac{1}{N}\operatorname{Tr}\mathbf{B}_k\mathbf{B}_{k'}^*=\tau(k,\overline{k'}).$$

Let also $\left(\mathbf{M}_\ell\right)_{\ell\in L}$ be a collection non random matrices such that

(iv) for each $\ell, \ell' \in L$, we have the following convergences

$$\lim_{N\to\infty} \frac{1}{N} \operatorname{Tr} \mathbf{M}_\ell \mathbf{M}_{\ell'} - \frac{1}{N} \operatorname{Tr} \mathbf{M}_\ell \cdot \frac{1}{N} \operatorname{Tr} \mathbf{M}_{\ell'} = \eta_{\ell\ell'}$$

and

$$\lim_{N\to\infty} \frac{1}{N} \operatorname{Tr} \mathbf{M}_\ell \mathbf{M}_{\ell'}^* - \frac{1}{N} \operatorname{Tr} \mathbf{M}_\ell \cdot \frac{1}{N} \operatorname{Tr} \mathbf{M}_{\ell'}^* = \beta_{\ell\ell'}.$$

At last, let $\mathbf{U} = \mathbf{U}^{(N)}$ be an $N \times N$ Haar-distributed unitary random matrix independent of $\left(\mathbf{B}_k\right)_{k\in K}$.

Proposition 4.3.1. *Let us fix $p \geq 1$, $(k_1, \ldots, k_p) \in K^p$ and $(\ell_1, \ldots, \ell_p) \in L^p$. If (i), (ii), (iii) and (iv) hold, then the centered vector*

$$\left(\operatorname{Tr}\left(\mathbf{U}\mathbf{B}_{k_i}\mathbf{U}^*\mathbf{M}_{\ell_i}\right)\right)_{1\leq i\leq p} \tag{4.21}$$

converges in distribution, as $N \to \infty$, to a complex centered Gaussian vector $\left(\mathscr{G}_i\right)_{1\leq i\leq p}$ such that, for all i, i',

$$\mathbb{E}\mathscr{G}_i\mathscr{G}_{i'} = \eta_{\ell_i\ell_{i'}}\tau(k_i, k_{i'}), \ \ and \ \ \mathbb{E}\mathscr{G}_i\overline{\mathscr{G}_{i'}} = \beta_{\ell_i\ell_{i'}}\tau(k_i, \overline{k_{i'}}).$$

Besides, for any sequence (Y_N) of bounded random variables such that Y_N is independent of $\mathbf{U}^{(N)}$ and $\mathbb{E}Y_N$ has a limit L and any polynomial f in p complex variables and their conjugates, we have

$$\lim_{N\to\infty} \mathbb{E}\left[Y_N f\left(\operatorname{Tr}\left(\mathbf{U}\mathbf{B}_{k_i}\mathbf{U}^*\mathbf{M}_{\ell_i}\right), 1\leq i\leq p\right)\right] \ = \ L\mathbb{E}\left[f\left(\mathscr{G}_i, 1\leq i\leq p\right)\right].$$

Proof. First, we can suppose the $\mathbf{B}_k$'s and the $\mathbf{M}_\ell$'s are all Hermitian (which makes the entries of the vector of (4.21) real), up to changing

$$\left(\mathbf{B}_k\right)_{k\in K} \quad \longrightarrow \quad \left(\mathbf{B}_{(k,1)} := \frac{1}{2}\left(\mathbf{B}_k + \mathbf{B}_k^*\right), \ \mathbf{B}_{(k,2)} := \frac{1}{2\mathrm{i}}\left(\mathbf{B}_k - \mathbf{B}_k^*\right)\right)_{(k,\varepsilon)\in K\times\{1,2\}},$$

$$\left(\mathbf{M}_\ell\right)_{\ell\in L} \quad \longrightarrow \quad \left(\mathbf{M}_{(\ell,1)} := \frac{1}{2}\left(\mathbf{M}_\ell + \mathbf{M}_\ell^*\right), \ \mathbf{M}_{(\ell,2)} := \frac{1}{2\mathrm{i}}\left(\mathbf{M}_\ell - \mathbf{M}_\ell^*\right)\right)_{(\ell,\varepsilon)\in L\times\{1,2\}}.$$

Second, as all $\mathbf{B}_k$'s have null trace, up to changing $\mathbf{M}_\ell \to \mathbf{M}_\ell - \frac{1}{N}\operatorname{Tr}\mathbf{M}_\ell$, one can suppose that all $\mathbf{M}_\ell$'s have null trace.

To prove the full proposition, it suffices to prove the convergence

$$\lim_{N\to\infty} \mathbb{E}\left[Y_N \prod_{i=1}^{n} \operatorname{Tr}\left(\mathbf{U}\mathbf{B}_{k_i}\mathbf{U}^*\mathbf{M}_{\ell_i}\right)\right] \ = \ L_Y\,\mathbb{E}\left[\prod_{i=1}^{n}\mathscr{H}_i\right].$$

for any $n \geq 1$, $(k_1, \ldots, k_n) \in K^n$, $(\ell_1, \ldots, \ell_n) \in L^n$ and any sequence (Y_N) of bounded random variables independent from $\mathbf{U}^{(N)}$ such that $\lim_{N\to\infty}\mathbb{E}Y_N = L_Y$. Indeed, we can take each k as many times as we want in $(k_1, \ldots, k_n)$ (and the same for ℓ), which implies the convergence of the expectation of any polynomials as wanted and consequently the convergence in distribution of finite dimensional marginals.

Let $n \geq 1$, and $\mathfrak{S}_n$ be the n-th symmetric group, and $\mathfrak{S}_{n,2}$ be the subset of permutations in $\mathfrak{S}_n$ with only cycles of length 2. We denote be $\#\sigma$ the number of cycles of $\sigma \in \mathfrak{S}_n$ and by $\mathrm{Fix}(\sigma)$ the number of fixed points of σ. The neutral element of $\mathfrak{S}_n$ is denoted by id_n. For any $\sigma \in \mathfrak{S}_n$, we set

$$\mathrm{Tr}_\sigma \left(\mathbf{N}_i\right)_{i=1}^n \;=\; \prod_{\substack{\text{cycle } (t_1 t_2 \cdots t_m) \\ \text{de } \sigma}} \mathrm{Tr}\left(\mathbf{N}_{t_1}\mathbf{N}_{t_2}\cdots\mathbf{N}_{t_m}\right)$$

For example, for $\sigma \in \mathfrak{S}_6$, $\sigma := (1,2,3,4,5,6) \;\mapsto\; (3,2,4,1,6,5)$

$$\mathrm{Tr}_\sigma \left(\mathbf{N}_i\right)_{i=1}^6 \;=\; \mathrm{Tr}\left(\mathbf{N}_1\mathbf{N}_3\mathbf{N}_4\right)\mathrm{Tr}\left(\mathbf{N}_2\right)\mathrm{Tr}\left(\mathbf{N}_5\mathbf{N}_6\right).$$

Lemma 4.3.2. *Let $n \geq 1$, $(k_1,\ldots,k_n) \in K^n$, $(\ell_1,\ldots,\ell_n) \in L^n$, and (Y_N) be any sequence of bounded random variables such that $\lim_{N\to\infty} \mathbb{E}\, Y_N = L$. With the above assumptions on $\mathbf{M}$ and $\mathbf{B}$, we have, for all γ and σ in $\mathfrak{S}_n$,*

$$\mathrm{Tr}_\gamma \left(\mathbf{M}_{\ell_i}\right)_{i=1}^n \;=\; \mathbb{1}_{\mathrm{Fix}(\gamma)=0}\, O\left(N^{n/2}\right)$$

and

$$\mathbb{E}\left[Y_N \mathrm{Tr}_\sigma \left(\mathbf{B}_{k_i}\right)_{i=1}^n\right] \;=\; \mathbb{1}_{\sigma\in\mathfrak{S}_{n,2}} N^{n/2} L_Y \prod_{\substack{\text{cycle } (i,j) \\ \text{de } \sigma}} \tau(k_i,k_j) + o(N^{n/2}).$$

Proof. Because $\mathbf{B}_k$'s and $\mathbf{M}_\ell$'s have null traces, the formulas are true in presence of fixed points. Thus, we can assume that σ and γ have no fixed point.

The first result comes from Lemma 4.3.6 and from the fact that, for each ℓ, $\mathrm{Tr}\,\mathbf{M}_\ell^2 \;=\; O(N)$.

The second result can be proved in two steps. First, if $\sigma \notin \mathfrak{S}_{n,2}$, the non-commutative Hölder's inequality (see [2, Appendix A.3]) and Hypothesis (ii) say us that

$$\left|\mathbb{E}\left[Y_N \mathrm{Tr}_\sigma \left(\mathbf{B}_{k_{i_j}}\right)_{j=1}^n\right]\right| = O(N^{\#\sigma}) = o(N^{n/2}).$$

If $\sigma \in \mathfrak{S}_{n,2}$ (and $n > 0$ is even), we decompose σ in 2-cycles $\sigma \;=\; (i_1\; j_1)\cdots(i_{n/2}\; j_{n/2})$. By classical Hölder's inequality, the absolute difference between

$$N^{-n/2}\,\mathbb{E}\left[Y_N \mathrm{Tr}_\sigma \left(\mathbf{B}_{k_j}\right)_{j=1}^n\right] = \mathbb{E}\left[Y_N \prod_{t=1}^{n/2} \frac{1}{N}\mathrm{Tr}\,\mathbf{B}_{k_{i_t}}\mathbf{B}_{k_{j_t}}\right] \text{ and } \mathbb{E}\left[Y_N \prod_{t=1}^{n/2-1} \frac{1}{N}\mathrm{Tr}\,\mathbf{B}_{k_{i_t}}\mathbf{B}_{k_{j_t}}\right]\mathbb{E}\left[\frac{1}{N}\mathrm{Tr}\,\mathbf{B}_{k_{i_1}}\mathbf{B}_{k_{j_1}}\right]$$

is less than

$$\mathbb{E}[Y_N^{2n}]^{1/2n} \prod_{t=1}^{n/2-1} \mathbb{E}\left[\left(\frac{1}{N}\mathrm{Tr}\,\mathbf{B}_{k_{i_t}}\mathbf{B}_{k_{j_t}}\right)^{2n}\right]^{1/2n} \mathbb{V}\mathrm{ar}\left(\frac{1}{N}\mathrm{Tr}\,\mathbf{B}_{k_{i_1}}\mathbf{B}_{k_{j_1}}\right)$$

and consequently converges to 0 - using again the non-commutative Hölder's inequality (see [2, Appendix A.3]) and Hypothesis (ii) to control $\mathbb{E}(\frac{1}{N}\operatorname{Tr}\mathbf{B}_{k_{i_t}}\mathbf{B}_{k_{j_t}})^{2n}$. By a direct induction on $n/2$, it means that the expectation of product

$$N^{-n/2}\,\mathbb{E}\left[Y_N \operatorname{Tr}_\sigma\left(\mathbf{B}_{k_j}\right)_{j=1}^n\right] = \mathbb{E}\left[Y_N \prod_{t=1}^{n/2}\frac{1}{N}\operatorname{Tr}\mathbf{B}_{k_{i_t}}\mathbf{B}_{k_{j_t}}\right]$$

has the same limit as the product of expectation

$$\mathbb{E}\left[Y_N\right]\prod_{t=1}^{n/2}\mathbb{E}\left[\frac{1}{N}\operatorname{Tr}\mathbf{B}_{k_{i_t}}\mathbf{B}_{k_{j_t}}\right],$$

and the result follows. □

Let $n \geq 1$, $(k_1,\dots,k_n)\in K^n$, $(\ell_1,\dots,\ell_n)\in L^n$, and (Y_N) be any sequence of bounded random variables such that $\lim_{N\to\infty}\mathbb{E}Y_N = L$. Using [67, Proposition 3.4] (and, first, an integration with respect to the randomness of $\mathbf{U}$, and then a "full expectation"), we have

$$\mathbb{E}\left[Y_N\prod_{i=1}^n \operatorname{Tr}\left(\mathbf{U}\mathbf{B}_{k_i}\mathbf{U}^*\mathbf{M}_{\ell_i}\right)\right] = \sum_{\sigma,\gamma\in\mathfrak{S}_n}\operatorname{Wg}(\sigma\gamma^{-1})\,\mathbb{E}[Y_N\operatorname{Tr}_\sigma\left(\mathbf{B}_i\right)_{i=1}^n]\operatorname{Tr}_\gamma\left(\mathbf{M}_i\right)_{i=1}^n \tag{4.22}$$

where Wg is the Weingarten function. We know from [38, Coro. 2.7] and [72, Propo 23.11] that, for any $\tau\in\mathfrak{S}_n$,

$$\operatorname{Wg}(\tau) = O\big(N^{\#\tau-2n}\big) \text{ and } \operatorname{Wg}(1_n) = N^{-n}+O(N^{-n-2}).$$

It implies, by Lemma 4.3.2, that for $\sigma,\gamma\in\mathfrak{S}_n$,

$$\operatorname{Wg}(\sigma\gamma^{-1})\,\mathbb{E}[Y_N\operatorname{Tr}_\sigma\left(\mathbf{B}_i\right)_{i=1}^n]\operatorname{Tr}_\gamma\left(\mathbf{M}_i\right)_{i=1}^n = \mathbb{1}_{\sigma\in\mathfrak{S}_{n,2}}\mathbb{1}_{\operatorname{Fix}(\gamma)=0}\,O(N^{(\#(\sigma\gamma^{-1})-n)}) = \mathbb{1}_{\gamma=\sigma\in\mathfrak{S}_{n,2}}O(1),$$

and more precisely, using the exact asymptotic for $\gamma=\sigma\in\mathfrak{S}_{n,2}$, that

$$\operatorname{Wg}(\sigma\gamma^{-1})\,\mathbb{E}[Y_N\operatorname{Tr}_\sigma\left(\mathbf{B}_i\right)_{i=1}^n]\operatorname{Tr}_\gamma\left(\mathbf{M}_i\right)_{i=1}^n = \mathbb{1}_{\gamma=\sigma\in\mathfrak{S}_{n,2}}\,L_Y\prod_{\substack{\text{cycle }(i,j)\\ \text{de }\sigma}}\tau(k_i,k_j)\eta_{\ell_i\ell_j}+o(1).$$

As a consequence, we can rewrite (4.22) as

$$\mathbb{E}\left[Y_N\prod_{i=1}^n \operatorname{Tr}\left(\mathbf{U}\mathbf{B}_{k_i}\mathbf{U}^*\mathbf{M}_{\ell_i}\right)\right] = L_Y\sum_{\sigma\in\mathfrak{S}_{n,2}}\prod_{\substack{\text{cycle }(i,j)\\ \text{de }\sigma}}\tau(k_i,k_j)\eta_{\ell_i\ell_j}+o(1),$$

which is the wanted convergence in order to prove the proposition, since

$$\mathbb{E}\left[\prod_{i=1}^n\mathscr{G}_i\right] = \sum_{\sigma\in\mathfrak{S}_{n,2}}\prod_{\substack{\text{cycle }(i,j)\\ \text{de }\sigma}}\tau(k_i,k_j)\eta_{\ell_i\ell_j}.$$

□

4.3.2 Proof of Theorem 4.2.5

First, note that, for each $k \in K$, $\mathbb{E}\mathbf{A}_k = \frac{1}{N}\mathbb{E}[\mathrm{Tr}(\mathbf{A}_k)]I$, hence for $\mathbf{B}_k := \mathbf{A}_k - \frac{1}{N}\mathrm{Tr}(\mathbf{A}_k)I$ and $T_k := \frac{1}{N}\mathrm{Tr}(\mathbf{A}_k) - \mathbb{E}\frac{1}{N}\mathrm{Tr}(\mathbf{A}_k)$, one can write

$$\left(\mathbf{A}_k - \mathbb{E}\mathbf{A}_k\right)_{b\in K} = \left(\mathbf{B}_k + T_k I\right)_{b\in K}.$$

Let us now introduce an (implicitly depending on N) Haar-distributed unitary matrix $\mathbf{U}$ independent of the collection $\mathbf{A}$. By unitary invariance, we get

$$\left(\mathbf{A}_k - \mathbb{E}\mathbf{A}_k\right)_{b\in K} = \left(\mathbf{B}_k + T_k I\right)_{b\in K} \overset{\text{law}}{=} \left(\mathbf{U}\mathbf{B}_k\mathbf{U}^* + T_k I\right)_{b\in K}.$$

Then, by Proposition 4.3.1, we know that, for any $n \geq 1$, any $k_1, \ldots, k_n \in K$ and any $\ell_1, \ldots, \ell_n \in L$, the random vector

$$\left(\mathrm{Tr}(\mathbf{U}\mathbf{B}_{k_i}\mathbf{U}^*\mathbf{M}_{\ell_i})\right)_{1\leq i\leq n}$$

converges in distribution to a complex centered Gaussian vector $\left(\mathscr{H}_i\right)_{1\leq i\leq n}$ such that, for all i, i',

$$\begin{aligned}
\mathbb{E}\,\mathscr{H}_i\mathscr{H}_{i'} &= \left(\tau(k_i,k_{i'}) - \tau(k_i)\tau(k_{i'})\right)\left(\eta_{\ell_i\ell_{i'}} - \alpha_{\ell_i}\alpha_{\ell_{i'}}\right), \\
\mathbb{E}\,\mathscr{H}_i\overline{\mathscr{H}_{i'}} &= \left(\tau(k_i,\overline{k_{i'}}) - \tau(k_i)\overline{\tau(k_{i'})}\right)\left(\eta_{\ell_i\ell_{i'}} - \alpha_{\ell_i}\overline{\alpha_{\ell_{i'}}}\right).
\end{aligned}$$

Besides, Proposition 4.3.1 also says that

$$\left(\mathrm{Tr}(\mathbf{U}\mathbf{B}_{k_i}\mathbf{U}^*\mathbf{M}_{\ell_i})\right)_{1\leq i\leq n}$$

is asymptotically independent from $\left(T_{k_i}\mathrm{Tr}(\mathbf{M}_{\ell_i})\right)_{1\leq i\leq n}$, which converges in distribution, by Hypothesis (d), to $\left(\alpha_{\ell_i}\mathscr{T}_{k_i}\right)_{1\leq i\leq n}$. As it is clear, from the covariance of $\left(\mathscr{G}_i\right)_{1\leq i\leq n}$, that for $\left(\mathscr{H}_i\right)_{1\leq i\leq n}$ independent from $\left(\alpha_{\ell_i}\mathscr{T}_{k_i}\alpha_{\ell_i}\right)_{1\leq i\leq n}$, we have

$$\left(\mathscr{G}_i\right)_{1\leq i\leq n} \overset{\text{law}}{=} \left(\mathscr{H}_i\right)_{1\leq i\leq n} + \left(\alpha_{\ell_i}\mathscr{T}_{k_i}\right)_{1\leq i\leq n},$$

the theorem is proved. □

4.3.3 Proof of Theorem 4.2.2

It is a direct application of Proposition 4.3.1 since if $\mathrm{Tr}\,\mathbf{M} = 0$, then

$$\mathrm{Tr}(\mathbf{A}\mathbf{M}) \;=\; \mathrm{Tr}\left[\left(\mathbf{A} - \frac{1}{N}(\mathrm{Tr}\,\mathbf{A})\mathbf{I}\right)\mathbf{M}\right],$$

so that one can assume that $\mathrm{Tr}\,\mathbf{A} = 0$. □

4.3.4 Proof of Corollary 4.2.9

We just need to show that the hypotheses of the corollary imply Hypotheses 4.2.1 and 4.2.4.

The proof of Hypothesis 4.2.1 comes down to the following computations, where we introduce a Haar-distributed unitary matrix $\mathbf{U}$ independent of $(\mathbf{A}_k)_{k\in K}$ and use Equation (33) of [22]. We have

$$\begin{aligned}\mathbb{E}\left|\frac{1}{N}\operatorname{Tr}\mathbf{A}_k\right|^2 &= \frac{1}{N^2}\mathbb{E}\left[\mathbb{E}_{\mathbf{U}}\left[\operatorname{Tr}(\mathbf{U}\mathbf{A}_k)\operatorname{Tr}(\mathbf{A}_k^*\mathbf{U}^*)\right]\right]\\ &= \frac{1}{N^3}(1+o(1))\,\mathbb{E}\left[\operatorname{Tr}(\mathbf{A}_k\mathbf{A}_k^*)\right] \;=\; O\left(\frac{1}{N^2}\right),\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}\left|\frac{1}{N}\operatorname{Tr}(\mathbf{A}_k\mathbf{A}_{k'})\right|^2 &= \frac{1}{N^2}\mathbb{E}\left[\mathbb{E}_{\mathbf{U}}\left[\operatorname{Tr}(\mathbf{A}_k\mathbf{U}\mathbf{A}_{k'}\mathbf{U})\operatorname{Tr}(\mathbf{A}_{k'}\mathbf{U}^*\mathbf{A}_k\mathbf{U}^*)\right]\right]\\ &= \frac{1}{N^4}(1+o(1))\,\mathbb{E}\left(\operatorname{Tr}(\mathbf{A}_k\mathbf{A}_k^*)\operatorname{Tr}(\mathbf{A}_{k'}\mathbf{A}_{k'}^*)+\operatorname{Tr}(\mathbf{A}_k\mathbf{A}_{k'}^*)\operatorname{Tr}(\mathbf{A}_{k'}\mathbf{A}_k^*)\right)\\ &= O\left(\frac{1}{N^2}\right).\end{aligned}$$

Now, in order to show Hypothesis 4.2.4, we want to prove that, for any fixed r, $\left(\operatorname{Tr}(\mathbf{A}_{k_1}),\dots,\operatorname{Tr}(\mathbf{A}_{k_r})\right)_{i=1}^r$ is asymptotically Gaussian. Let $n\geq 1$ and $i_1,j_1,\dots,i_n,j_n\in\{1,\dots,r\}$, using [67, Proposition 3.4], we have

$$\begin{aligned}\mathbb{E}\left[\prod_{\ell=1}^n\operatorname{Tr}(\mathbf{A}_{k_{i_\ell}})\operatorname{Tr}(\mathbf{A}_{k_{j_\ell}}^*)\right] &= \mathbb{E}\left[\mathbb{E}_{\mathbf{U}}\prod_{\ell=1}^n\operatorname{Tr}(\mathbf{U}\mathbf{A}_{k_{i_\ell}})\operatorname{Tr}(\mathbf{A}_{k_{j_\ell}}^*\mathbf{U}^*)\right]\\ &= \frac{1}{N^n}\sum_{\sigma\in\mathfrak{S}_n}\mathbb{E}\left[\prod_{\ell=1}^n\operatorname{Tr}\left(\mathbf{A}_{k_{i_\ell}}\mathbf{A}_{k_{j_{\sigma(\ell)}}}^*\right)\right]+o(1).\end{aligned}$$

Then, one can prove that

$$\frac{1}{N^n}\sum_{\sigma\in\mathfrak{S}_n}\mathbb{E}\left[\prod_{\ell=1}^n\operatorname{Tr}\left(\mathbf{A}_{k_{i_\ell}}\mathbf{A}_{k_{j_{\sigma(\ell)}}}^*\right)\right] = \frac{1}{N^n}\sum_{\sigma\in\mathfrak{S}_n}\prod_{\ell=1}^n\mathbb{E}\left[\operatorname{Tr}\left(\mathbf{A}_{k_{i_\ell}}\mathbf{A}_{k_{j_{\sigma(\ell)}}}^*\right)\right]+o(1)\tag{4.23}$$

Indeed, similarly from above, we use classical Hölder's inequality to state that the difference between

$$N^{-n}\,\mathbb{E}\left[\prod_{\ell=1}^n\operatorname{Tr}\left(\mathbf{A}_{k_{i_\ell}}\mathbf{A}_{k_{j_{\sigma(\ell)}}}^*\right)\right]$$

and

$$N^{-(n-1)}\,\mathbb{E}\left[\prod_{i=1}^{n-1}\operatorname{Tr}(\mathbf{A}_{k_{i_\ell}}\mathbf{A}_{k_{j_{\sigma(\ell)}}}^*)\right]\mathbb{E}\left[\frac{1}{N}\operatorname{Tr}(\mathbf{A}_{k_{i_n}}\mathbf{A}_{k_{j_{\sigma(n)}}}^*)\right]$$

is lower than

$$\prod_{i=1}^{n-1}\mathbb{E}\left[\left(\frac{1}{N}\operatorname{Tr}\left(\mathbf{A}_{k_{i_\ell}}\mathbf{A}^*_{k_{j_{\sigma(\ell)}}}\right)\right)^{2(n-1)}\right]^{\frac{1}{2(n-1)}}\mathbb{V}\mathrm{ar}\left(\frac{1}{N}\operatorname{Tr}(\mathbf{A}_{k_{i_n}}\mathbf{A}^*_{k_{j_{\sigma(n)}}})\right),$$

which tends to 0 thanks to the non-commutative Hölder's inequality and the fact that $\frac{1}{N}\operatorname{Tr}(\mathbf{A}_n\mathbf{A}^*_{\sigma(n)})$ converges in probability to a constant. We conclude the proof (4.23) with a simple induction. Once we have (4.23), we can conclude using the Wick Formula. □

4.3.5 Proofs of Theorem 4.2.13 and Theorem 4.2.19.

In this section, we will directly apply [85, Theorem 2.3 and Theorem 2.10] in order to prove both Theorems 4.2.13 and 4.2.19. To do so, we only need to prove that the Gaussian elliptic ensemble satisfies the assumptions of [85, Theorem 2.3 and Theorem 2.10]. This is the purpose of the following proposition.

Proposition 4.3.3. *Let $\mathbf{X}_N := \frac{1}{\sqrt{N}}\mathbf{Y}_N$ where $\mathbf{Y}_N$ is an $N\times N$ Gaussian elliptic matrix of parameter ρ. Then, as $N\to\infty$, we have:*

(i) $\|\mathbf{X}_N\|_{\mathrm{op}}$ *converges in probability to* $1+|\rho|$*;*

(ii) for any $\delta>0$, as N goes to infinity, we have the convergence in probability

$$\sup_{|z|>1+|\rho|+\delta}\ \max_{1\le i,j\le 2r}\left|\mathbf{e}_i^*\left(z\mathbf{I}-\mathbf{X}_N\right)^{-1}\mathbf{e}_j-\delta_{ij}m(z)\right| \longrightarrow 0,$$

where $m(z) := \displaystyle\int \frac{1}{z-w}\mu_\rho(dw)$*;*

(iii) for any z such that $|z|>1+|\rho|+\varepsilon$, we have the convergence in probability

$$\sqrt{N}\left(\frac{1}{N}\operatorname{Tr}\left(z-\mathbf{X}\right)^{-1}-m(z)\right) \longrightarrow 0;$$

(iv) the finite marginals of random process

$$\left(\sqrt{N}\left(\mathbf{e}_i^*\left(z-\mathbf{X}\right)^{-1}\mathbf{e}_j-\delta_{ij}\frac{1}{N}\operatorname{Tr}\left(z-\mathbf{X}\right)^{-1}\right)\right)_{\substack{|z|>1+|\rho|+\varepsilon\\ 1\le i,j\le 2r}}$$

converge to the ones of the complex centered Gaussian process

$$\left(\mathscr{G}_{i,j,z}\right)_{\substack{|z|>1+|\rho|+\varepsilon\\ 1\le i,j\le 2r}}$$

satisfying

$$\begin{aligned}
\mathbb{E}\left[\mathscr{G}_{i,j,z}\mathscr{G}_{i',j',z'}\right] &= \delta_{ij'}\delta_{i'j}\left(\int\frac{1}{(z-w)(z'-w)}\mu_\rho(dw)-\int\frac{1}{z-w}\mu_\rho(dw)\int\frac{1}{z'-w}\mu_\rho(dw)\right),\\
\mathbb{E}\left[\mathscr{G}_{i,j,z}\overline{\mathscr{G}_{i',j',z'}}\right] &= \delta_{ii'}\delta_{jj'}\left(\int\frac{1}{(z-w)(\overline{z'}-\overline{w})}\mu_\rho(dw)-\int\frac{1}{z-w}\mu_\rho(dw)\overline{\int\frac{1}{z'-w}\mu_\rho(dw)}\right);
\end{aligned}$$

(v) *for any $p \geq 1$, any $1 \leq i, j \leq 2r$ and any $|z| > 1 + |\rho| + \varepsilon$, the sequence*

$$\sqrt{N}\big(\mathbf{e}_i^*\left(z-\mathbf{X}\right)^{-p}\mathbf{e}_j - \delta_{ij}\frac{1}{N}\operatorname{Tr}\left(z-\mathbf{X}\right)^{-p}\big)$$

is tight.

Remark 4.3.4. One should be careful about the fact that our $m(z)$ is not the same from [73, Lemma 4.3] but the opposite. Moreover, for any $|\theta| > 1$, we still have (see [73, (5.2) and (5.3)])

$$m(z) \;=\; \frac{1}{\theta} \quad\Longleftrightarrow\quad z \;=\; \theta + \frac{\rho}{\theta},$$

so that it is easy to compute $\mathbb{E}\left[\mathscr{G}_{i,j,z}\mathscr{G}_{i',j',z'}\right]$ in (iv) for $z = \theta + \frac{\rho}{\theta}$ and $z' = \theta' + \frac{\rho}{\theta'}$, indeed

$$\begin{aligned}
&\int \frac{1}{(z-w)(z'-w)}\mu_\rho(dw) - \int \frac{1}{z-w}\mu_\rho(dw)\int \frac{1}{z'-w}\mu_\rho(dw)\\
=\;& -\frac{m(z)-m(z')}{z-z'} - m(z)m(z') \;=\; \frac{1}{\theta\theta'-\rho} - \frac{1}{\theta\theta'}.
\end{aligned}$$

Also, for any $|z| > 2\sqrt{|\rho|}$, it might be useful to write

$$m(z) \;=\; \sum_{k\geq 0}\rho^k \operatorname{Cat}(k) z^{-2k-1},$$

where $\operatorname{Cat}(k) = \frac{1}{k+1}\binom{2k}{k}$ is the k-th Catalan number.

Proof of Proposition 4.3.3. First, (i) is a direct consequence of [71, Theorem 1.9]. It implies that, with a probability tending one, for any fixed $|z| > 1 + |\rho|$, one can write

$$(z-\mathbf{X}_N)^{-1} \;=\; \sum_{k\geq 0} z^{-k-1}\mathbf{X}_N^k.$$

Moreover, if we apply the Theorem 4.2.5 with

$$(\mathbf{M}_\ell) = \big(\sqrt{N}\mathbf{E}_{ji}\big)_{1\leq i,j\leq 2r} \quad\text{and}\quad (\mathbf{A}_k) = \Big((z-\mathbf{X}_N)^{-1} - \frac{1}{N}\operatorname{Tr}(z-\mathbf{X}_N)^{-1}\mathbf{I}\Big)_{|z|>1+|\rho|+\varepsilon}$$

(since $\mathbf{X}_N$ is invariant in distribution by unitary conjugation, so is $(z-\mathbf{X}_N)^{-p}$ for any $p \geq 1$), we easily obtain (iv). The same for (v) by changing the exponent -1 into $-p$. At last, we just need to prove (ii) and (iii) and to do so, we prove the following lemma.

Lemma 4.3.5. *Let $\mathbf{X}_N$ be a $N\times N$ Gaussian elliptic matrix of parameter $-1 \leq \rho \leq 1$. Then, for any z such that $|z| > 1 + |\rho| + \varepsilon$,*

$$\frac{1}{N}\operatorname{Tr}(z-\mathbf{X}_N)^{-1} - m(z) \;=\; o\left(\frac{1}{\sqrt{N}}\right).$$

Indeed, once the previous Lemma is shown, we clearly have (iii), and here comes the proof of (ii).

Proof of (ii). Let $\eta > 0$ and let i, j be two integers lower than $2r$. Since $\|\mathbf{X}_N\|_{\text{op}}$ is bounded, we know that $\left\|(z-\mathbf{X}_N)^{-1}\right\|_{\text{op}}$ goes to 0 when $|z| \to \infty$, as the function $m(z)$, so that, we know there is a positive constant M such that

$$\mathbb{P}\big(\sup_{|z|>1+\rho+\varepsilon}\left|\mathbf{e}_i^*(z-\mathbf{X}_N)^{-1}\mathbf{e}_j-\delta_{ij}m(z)\right|>\eta\big)=\mathbb{P}\big(\sup_{1+\rho+\varepsilon<|z|<M}\left|\mathbf{e}_i^*(z-\mathbf{X}_N)^{-1}\mathbf{e}_j-\delta_{ij}m(z)\right|>\eta\big)+o(1).$$

Then, for any $\eta' > 0$, the compact set $K = \{1+\rho+\varepsilon \le |z| \le M\}$ admits a η'-net that we denote by $S_{\eta'}$, which is a finite set of K such that

$$\forall z \in K,\ \exists z' \in S_{\eta'},\quad |z-z'| < \eta',$$

so that, using the uniform boundedness of the derivative of $m(z)$ and $\mathbf{e}_i^*(z-\mathbf{X}_N)^{-1}\mathbf{e}_j$ on K, we have for a small enough η'

$$\mathbb{P}\big(\sup_{z\in K}\left|\mathbf{e}_i^*(z-\mathbf{X})^{-1}\mathbf{e}_j-\delta_{ij}m(z)\right|>\eta\big) \quad = \quad \mathbb{P}\big(\bigcup_{z\in S_{\eta'}}\left\{\left|\mathbf{e}_i^*(z-\mathbf{X})^{-1}\mathbf{e}_j-\delta_{ij}m(z)\right|>\eta/2\right\}\big)$$

At last, we write, for any $z \in S_{\eta'}$,

$$\begin{aligned}
&\mathbb{P}\big(\left|\mathbf{e}_i^*(z-\mathbf{X})^{-1}\mathbf{e}_j-\delta_{ij}m(z)\right|>\eta/2\big)\\
\le\ &\mathbb{P}\Big(\left|\mathbf{e}_i^*(z-\mathbf{X})^{-1}\mathbf{e}_j-\delta_{ij}\frac{1}{N}\operatorname{Tr}(z-\mathbf{X})^{-1}\right|>\eta/4\Big)+\mathbb{P}\Big(\delta_{ij}\left|\frac{1}{N}\operatorname{Tr}(z-\mathbf{X})^{-1}-m(z)\right|>\eta/4\Big).
\end{aligned}$$

The first term vanishes thanks to Theorem 4.2.5 with $\mathbf{M} = \sqrt{N}\mathbf{E}_{ji}$ and the second one vanishes by Lemma 4.3.5. □

Proof of Lemma 4.3.5. First of all, let us write, for any $\eta > 0$,

$$\begin{aligned}
&\mathbb{P}\left(\left|\frac{1}{N}\operatorname{Tr}(z-\mathbf{X}_N)^{-1}-m(z)\right|>\frac{\eta}{\sqrt{N}}\right)\\
\le\ &\frac{4N}{\eta^2}\left(\mathbb{E}\left|\frac{1}{N}\operatorname{Tr}(z-\mathbf{X}_N)^{-1}-\mathbb{E}\frac{1}{N}\operatorname{Tr}(z-\mathbf{X}_N)^{-1}\right|^2+\left|\mathbb{E}\frac{1}{N}\operatorname{Tr}(z-\mathbf{X}_N)^{-1}-m(z)\right|^2\right),
\end{aligned}$$

which means that we just have to prove that

$$\left|\mathbb{E}\frac{1}{N}\operatorname{Tr}(z-\mathbf{X}_N)^{-1}-m(z)\right|^2 \quad = \quad o\left(\frac{1}{N}\right), \tag{4.24}$$

and

$$\mathbb{E}\left|\frac{1}{N}\operatorname{Tr}(z-\mathbf{X}_N)^{-1}-\mathbb{E}\frac{1}{N}\operatorname{Tr}(z-\mathbf{X}_N)^{-1}\right|^2 \quad = \quad o\left(\frac{1}{N}\right). \tag{4.25}$$

Proof of (4.24): We know from [9, Theorem 1.1] that (4.24) is true for a Wigner matrix. Here, the idea of the proof is to use the fact that the Stieltjes transform of the semicircular law of variance $\sigma^2 = \rho$ is equal to $m(z)$ outside the ellipse $\mathscr{E}_\rho$ when $\rho > 0$. First, we shall suppose that $\rho \geq 0$ up to changing $\mathbf{X}_N$ into $\mathrm{i}\mathbf{X}_N$. For any $i \neq j$, we have

$$\mathbb{E}x_{ii}^2 \;=\; \mathbb{E}x_{ij}x_{ji} \;=\; \frac{\rho}{N}, \quad \text{and} \quad \mathbb{E}x_{ij}^2 \;=\; 0. \tag{4.26}$$

One can notice that if $\mathbf{W}_N$ is a real symmetric Gaussian matrix of variance ρ with iid entries such that, for any $i \neq j$,

$$\mathbb{E}w_{ii}^2 \;=\; \mathbb{E}w_{ij}w_{ji} \;=\; \mathbb{E}w_{ij}^2 \;=\; \frac{\rho}{N}, \tag{4.27}$$

then we have, by the Wick formula applied to the expansion of the traces,

- $\mathbb{E}\left[\operatorname{Tr}\mathbf{W}_N^k\right] \geq 0$ and $\mathbb{E}\left[\operatorname{Tr}\mathbf{X}_N^k\right] \geq 0$,
- $\mathbb{E}\left[\operatorname{Tr}\mathbf{W}_N^k\right] \geq \mathbb{E}\left[\operatorname{Tr}\mathbf{X}_N^k\right]$ since there are more non-zero terms for $\mathbf{W}_N$ than for $\mathbf{X}_N$.

Also, we know that, for any z such that $|z| > 1+\rho+\varepsilon$,

$$\mathbb{E}\frac{1}{N}\operatorname{Tr}(z-\mathbf{W}_N)^{-1} \;=\; \sum_{k\geq 0} z^{-k-1}\,\mathbb{E}\left[\frac{1}{N}\operatorname{Tr}\mathbf{W}_N^k\right] \;\text{ and }\; \mathbb{E}\frac{1}{N}\operatorname{Tr}(z-\mathbf{X}_N)^{-1} \;=\; \sum_{k\geq 0} z^{-k-1}\,\mathbb{E}\left[\frac{1}{N}\operatorname{Tr}\mathbf{X}_N^k\right]$$

converge to the same limit

$$m(z) \;=\; \sum_{k\geq 0}\operatorname{Cat}(k)\rho^k z^{-2k-1}, \quad \text{where } \operatorname{Cat}(k) \text{ is the } k\text{-th Catalan number.}$$

Moreover, by the Wick formula again, for $\mathscr{P}_2(2k)$ (resp. $NC_2(2k)$) the set of pairings (resp. non crossing pairings) of $\{1,\dots,2k\}$,

$$\begin{aligned}
\mathbb{E}\left[\operatorname{Tr}\mathbf{X}_N^{2k}\right] \;&=\; \sum_{1\leq i_1,\dots,i_{2k}\leq N}\;\sum_{\pi\in\mathscr{P}_2(2k)}\;\prod_{\{s,t\}\in\pi}\mathbb{E}\left[x_{i_s i_{s+1}}x_{i_t i_{t+1}}\right] \\
&=\; \sum_{\pi\in\mathscr{P}_2(2k)}\;\sum_{1\leq i_1,\dots,i_{2k}\leq N}\;\prod_{\{s,t\}\in\pi}\mathbb{E}\left[x_{i_s i_{s+1}}x_{i_t i_{t+1}}\right]
\end{aligned}$$

Note that using the Dyck path interpretation of $NC_2(2k)$ (see e.g. [72]), one can easily see that i the previous sum, the term associated to each $\pi \in NC_2(2k)$ is precisely ρ^k. Hence as the cardinality of $NC_2(2k)$ is $\operatorname{Cat}(k)$ (see [72] again) and each $\mathbb{E}\left[x_{i_s i_{s+1}}x_{i_t i_{t+1}}\right]$ is non negative, we have

$$\mathbb{E}\left[\operatorname{Tr}\mathbf{X}_N^{2k}\right] \;\geq\; \operatorname{Cat}(k)\rho^k.$$

At last, we know from e.g. [9, Theorem 1.1] that, for any z such that $|z| > 1+\rho+\varepsilon$, we have

$$\mathbb{E}\frac{1}{N}\operatorname{Tr}(z-\mathbf{W}_N)^{-1} - m(z) \;=\; o\left(\frac{1}{\sqrt{N}}\right),$$

so that, to conclude, it suffices to write

$$\begin{aligned}\left|\mathbb{E}\frac{1}{N}\operatorname{Tr}(z-\mathbf{X}_N)^{-1}-m(z)\right| &\leq \sum_{k\geq 0}\left(\mathbb{E}\frac{1}{N}\operatorname{Tr}\mathbf{X}_N^{2k}-\operatorname{Cat}(k)\rho^k\right)|z|^{-2k-1}\\ &\leq \sum_{k\geq 0}\left(\mathbb{E}\frac{1}{N}\operatorname{Tr}\mathbf{W}_N^{2k}-\operatorname{Cat}(k)\rho^k\right)|z|^{-2k-1}\\ &= \mathbb{E}\frac{1}{N}\operatorname{Tr}(|z|-\mathbf{W}_N)^{-1}-m(|z|) \;=\; o\left(\frac{1}{\sqrt{N}}\right).\end{aligned}$$

Proof of (4.25): We apply the same idea but this time $\mathbf{W}_N$ is a real symmetric Gaussian matrix of variance ρ with iid entries, such that, for any $i\neq j$,

$$\mathbb{E}w_{ii}^2 = \frac{1}{N} \quad ; \quad \mathbb{E}w_{ij}w_{ji} = \mathbb{E}w_{ij}^2 = \frac{\rho}{N}. \tag{4.28}$$

From e.g. [9, Theorem 1.1], we know that, for all $|z|>1+|\rho|+\varepsilon$,

$$\mathbb{E}\left|\frac{1}{N}\operatorname{Tr}(z-\mathbf{W}_N)^{-1}-\mathbb{E}\frac{1}{N}\operatorname{Tr}(z-\mathbf{W}_N)^{-1}\right|^2 = o\left(\frac{1}{N}\right).$$

Moreover, we can write

$$\begin{aligned}&\mathbb{E}\left|\frac{1}{N}\operatorname{Tr}(z-\mathbf{X}_N)^{-1}-\mathbb{E}\frac{1}{N}\operatorname{Tr}(z-\mathbf{X}_N)^{-1}\right|^2\\ =\;&\mathbb{E}\left|\frac{1}{N}\operatorname{Tr}(z-\mathbf{X}_N)^{-1}\right|^2-\left|\mathbb{E}\frac{1}{N}\operatorname{Tr}(z-\mathbf{X}_N)^{-1}\right|^2\\ =\;&\frac{1}{N^2}\sum_{k,\ell\geq 0}z^{-k-1}\overline{z}^{-\ell-1}\left(\mathbb{E}\left[\operatorname{Tr}\mathbf{X}_N^k\overline{\operatorname{Tr}\mathbf{X}_N^\ell}\right]-\mathbb{E}\operatorname{Tr}\mathbf{X}_N^k\,\mathbb{E}\overline{\operatorname{Tr}\mathbf{X}_N^\ell}\right).\end{aligned}$$

By the Wick formula, we see that, for all k,ℓ,

$$0 \leq \mathbb{E}\left[\operatorname{Tr}\mathbf{X}_N^k\overline{\operatorname{Tr}\mathbf{X}_N^\ell}\right]-\mathbb{E}\operatorname{Tr}\mathbf{X}_N^k\,\mathbb{E}\overline{\operatorname{Tr}\mathbf{X}_N^\ell} \leq \mathbb{E}\left[\operatorname{Tr}\mathbf{W}_N^k\overline{\operatorname{Tr}\mathbf{W}_N^\ell}\right]-\mathbb{E}\operatorname{Tr}\mathbf{W}_N^k\,\mathbb{E}\overline{\operatorname{Tr}\mathbf{W}_N^\ell}.$$

Indeed,

$$\begin{aligned}\mathbb{E}\left[\operatorname{Tr}\mathbf{X}_N^k\overline{\operatorname{Tr}\mathbf{X}_N^\ell}\right] &= \sum_{\substack{1\leq i_1,\dots,i_k\leq N\\ 1\leq i'_{k+1},\dots,i'_{k+\ell}\leq N}}\mathbb{E}\left[x_{i_1i_2}\cdots x_{i_ki_1}\overline{x_{i'_{k+1}i'_{k+2}}}\cdots\overline{x_{i'_{k+\ell}i'_{k+1}}}\right]\\ &= \sum_{\substack{1\leq i_1,\dots,i_k\leq N\\ 1\leq i'_{k+1},\dots,i'_{k+\ell}\leq N}}\;\sum_{\pi\in\mathscr{P}_2(k+\ell)}\;\prod_{\{a,b\}\in\pi}\mathbb{E}[x_ax_b]\end{aligned} \tag{4.29}$$

where $x_a = \begin{cases} x_{i_ai_{a+1}} & \text{if } 1\leq a\leq k-1\\ x_{i_ki_1} & \text{if } a=k\\ \overline{x_{i'_ai'_{a+1}}} & \text{if } k+1\leq a\leq k+\ell-1\\ \overline{x_{i'_{k+\ell}i'_{k+1}}} & \text{if } a=k+\ell.\end{cases}$

We have also

$$\mathbb{E}\,\mathrm{Tr}\,\mathbf{X}_N^k\,\mathbb{E}\,\overline{\mathrm{Tr}\,\mathbf{X}_N^\ell} \;=\; \sum_{\substack{1\le i_1,\dots,i_k\le N\\ 1\le i'_{k+1},\dots,i'_{k+\ell}\le N}}\;\sum_{\substack{\pi\in\mathscr{P}_2(k)\\ \mu\in\mathscr{P}_2(\ell)}}\;\prod_{\{a,b\}\in\pi}\mathbb{E}\,[x_a x_b]\prod_{\{c,d\}\in\mu}\mathbb{E}\,[x_c x_d]\,,$$

which is a subsum of (4.29). Hence, as all the $\mathbb{E}\,[x_a x_b]$'s are non negative (see (4.26)), we conclude that

$$\mathbb{E}\left[\mathrm{Tr}\,\mathbf{X}_N^k\,\overline{\mathrm{Tr}\,\mathbf{X}_N^\ell}\right] \;\ge\; \mathbb{E}\,\mathrm{Tr}\,\mathbf{X}_N^k\,\mathbb{E}\,\overline{\mathrm{Tr}\,\mathbf{X}_N^\ell}.$$

Since, for all a and b, $\mathbb{E}\,[x_a x_b] \le \mathbb{E}\,[w_a w_b]$ (see (4.28)), we deduce that

$$\mathbb{E}\left[\mathrm{Tr}\,\mathbf{X}_N^k\,\overline{\mathrm{Tr}\,\mathbf{X}_N^\ell}\right]-\mathbb{E}\,\mathrm{Tr}\,\mathbf{X}_N^k\,\mathbb{E}\,\overline{\mathrm{Tr}\,\mathbf{X}_N^\ell} \;\le\; \mathbb{E}\left[\mathrm{Tr}\,\mathbf{W}_N^k\,\overline{\mathrm{Tr}\,\mathbf{W}_N^\ell}\right]-\mathbb{E}\,\mathrm{Tr}\,\mathbf{W}_N^k\,\mathbb{E}\,\overline{\mathrm{Tr}\,\mathbf{W}_N^\ell}$$

At last, we can write

$$\begin{aligned}
&\mathbb{E}\left|\frac{1}{N}\mathrm{Tr}\,(z-\mathbf{X}_N)^{-1}-\mathbb{E}\,\frac{1}{N}\mathrm{Tr}\,(z-\mathbf{X}_N)^{-1}\right|^2\\
\le\;& \frac{1}{N^2}\sum_{k,\ell\ge 0}|z|^{-k-1}|z|^{-\ell-1}\left(\mathbb{E}\left[\mathrm{Tr}\,\mathbf{X}_N^k\,\overline{\mathrm{Tr}\,\mathbf{X}_N^\ell}\right]-\mathbb{E}\,\mathrm{Tr}\,\mathbf{X}_N^k\,\mathbb{E}\,\overline{\mathrm{Tr}\,\mathbf{X}_N^\ell}\right)\\
\le\;& \frac{1}{N^2}\sum_{k,\ell\ge 0}|z|^{-k-1}|z|^{-\ell-1}\left(\mathbb{E}\left[\mathrm{Tr}\,\mathbf{W}_N^k\,\overline{\mathrm{Tr}\,\mathbf{W}_N^\ell}\right]-\mathbb{E}\,\mathrm{Tr}\,\mathbf{W}_N^k\,\mathbb{E}\,\overline{\mathrm{Tr}\,\mathbf{W}_N^\ell}\right)\\
=\;& \mathbb{E}\left|\frac{1}{N}\mathrm{Tr}\,(|z|-\mathbf{W}_N)^{-1}-\mathbb{E}\,\frac{1}{N}\mathrm{Tr}\,(|z|-\mathbf{W}_N)^{-1}\right|^2 \;=\; o\left(\frac{1}{N}\right).
\end{aligned}$$

□

Appendix: a matrix inequality

Lemma 4.3.6. *For any $k\ge 2$ and any Hermitian matrix $\mathbf{H}$,*

$$\left|\mathrm{Tr}\,\mathbf{H}^k\right| \;\le\; \left(\mathrm{Tr}\,\mathbf{H}^2\right)^{k/2}$$

more generally, for any family of Hermitian matrices $\mathbf{H}_1,\dots,\mathbf{H}_k$,

$$\left|\mathrm{Tr}\left(\mathbf{H}_1\cdots\mathbf{H}_k\right)\right| \;\le\; \prod_{i=1}^k\sqrt{\mathrm{Tr}\,\mathbf{H}_i^2}$$

Proof. We know that, for any non negative Hermitian matrices $\mathbf{A}$ and $\mathbf{B}$, one has

$$\mathrm{Tr}\,\mathbf{A}\mathbf{B} \le \mathrm{Tr}\,\mathbf{A}\,\mathrm{Tr}\,\mathbf{B}$$

so that, for any $p\ge 1$,

$$\mathrm{Tr}\,\mathbf{H}^{2p} \le \left(\mathrm{Tr}\,\mathbf{H}^2\right)^p$$

also,

$$\operatorname{Tr}\mathbf{H}^{2p+1} \leq \sqrt{\operatorname{Tr}\mathbf{H}^2}\sqrt{\operatorname{Tr}\mathbf{H}^{4p}} \leq \sqrt{\operatorname{Tr}\mathbf{H}^2}\sqrt{\left(\operatorname{Tr}\mathbf{H}^2\right)^{2p}} \;=\; \left(\operatorname{Tr}\mathbf{H}^2\right)^{(2p+1)/2}.$$

Then, using the non-commutative Hölder's inequality (see [2, A.3]), we deduce that

$$\left|\operatorname{Tr}\left(\mathbf{H}_1\cdots\mathbf{H}_k\right)\right| \leq \prod_{i=1}^{k}\left(\operatorname{Tr}|\mathbf{H}_i|^k\right)^{1/k} \leq \prod_{i=1}^{k}\left(\left(\operatorname{Tr}\mathbf{H}_i^2\right)^{k/2}\right)^{1/k}.$$

□

Bibliography

[1] G. Anderson, O. Zeitouni *A CLT for a band matrix model*, Probab. Theory Rel. Fields, 2005, **134**, 283–338

[2] G. Anderson, A. Guionnet, O. Zeitouni *An Introduction to Random Matrices*. Cambridge studies in advanced mathematics, 118 (2009).

[3] A. D'Aristotile, P. Diaconis, C. Newman *Brownian motion and the classical groups*. With Probability, Statisitica and their applications: Papers in Honor of Rabii Bhattacharaya. Edited by K. Athreya *et al.* 97–116. Beechwood, OH: Institute of Mathematical Statistics, 2003.

[4] R. Arratia, S. Tavare *The Cycle Structure of Random Permutations*. Ann. of Probab., 20(3): 1567 1591 (1992).

[5] Z.D. Bai, J. Silverstein *CLT for linear spectral statistics of large-dimensional sample covariance matrices.* Ann. Probab. **32**, 2004, 533–605.

[6] Z. D. Bai, J. W. Silverstein *Spectral analysis of large dimensional random matrices*, Second Edition, Springer, New York, 2009.

[7] Z.D. Bai, X. Wang, W. Zhou *CLT for linear spectral statistics of Wigner matrices.* Electron. J. Probab. **14** (2009), no. 83, 2391–2417.

[8] Z.D. Bai, J. Yao *On the convergence of the spectral empirical process of Wigner matrices.* Bernoulli **11** (2005) 1059–1092.

[9] J. Baik, G. Ben Arous, S. Péché *Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices*. Ann. Probab., 33(5):1643–1697, 2005.

[10] A. Basak, A. Dembo *Limiting spectral distribution of sums of unitary and orthogonal matrices*. Electron. Commun. Probab. 18 (2013), no. 69, 19 pp.

[11] A. Beardon *Complex Analysis: the Winding Number principle in analysis and topology*. John Wiley and Sons (1979).

[12] S. T. Belinschi, H. Bercovici, M. Capitaine, M. Février, *Outliers in the spectrum of large deformed unitarily invariant models* arXiv:1207.5443v1, 2012

[13] S.R. Bell *The Cauchy transform, potential theory, and conformal mapping*. (1992) Studies in advanced Mathematics. CRC Press, Boca Raton, FL.

[14] F. Benaych-Georges *Central limit theorems for the Brownian motion on large unitary groups*, Bull. Soc. Math. France,Vol. 139, no. 4 (2011) 593–610.

[15] F. Benaych-Georges *Exponential bounds for the support convergence in the Single Ring Theorem*, J. Funct. Anal. 268 (2015), pp. 3492–3507.

[16] F. Benaych-Georges, A. Guionnet, M. Maida *Fluctuations of the extreme eigenvalues of finite rank deformations of random matrices*, Electron. J. Prob. Vol. 16 (2011), Paper no. 60, 1621–1662.

[17] F. Benaych-Georges, A. Guionnet, M. Maida *Large deviations of the extreme eigenvalues of random deformations of matrices*, Probab. Theory Related Fields Vol. 154, no. 3 (2012), 703–751.

[18] F. Benaych-Georges, A. Guionnet, C. Male *Central limit theorems for linear statistics of heavy tailed random matrices*. Comm. Math. Phys. Vol. 329 (2014), no. 2, 641–686.

[19] F. Benaych-Georges, R.N. Rao *The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices*, Adv. Math. (2011), Vol. 227, no. 1, 494–521.

[20] F. Benaych-Georges, R.N. Rao *The singular values and vectors of low rank perturbations of large rectangular random matrices*, J. Multivariate Anal., Vol. 111 (2012), 120–135.

[21] F. Benaych-Georges, J. Rochet *Outliers in the Single Ring Theorem*, arXiv:1308.3064.

[22] F. Benaych-Georges, J. Rochet *Fluctuations for analytic test functions in the Single Ring Theorem*, arXiv:arXiv:1504.05106.

[23] F. Benaych-Georges, G. Cébron, J. Rochet *Fluctuation of matrix entries and application to outliers of elliptic matrices*, arxiv.org/abs/1602.02929.

[24] C. Bordenave, M. Capitaine *Outlier eigenvalues for deformed i.i.d. random matrices*. arXiv:1403.6001.

[25] C. Bordenave, D. Chafaï *Around the circular law*, Probab. Surv. 9 (2012), 1–89.

[26] É. Borel *Sur les principes de la théorie cinétique des gaz.* Annales de l'École Normale Supérieure 23 (1906), 9–32.

[27] Y. Cao, L. Cai, C. Qiu, J. Gu, X. He, Q. Ai, Z.Jin *A random matrix theoretical approach to early event detection using experimental data*, arXiv:1503.08445

[28] M. Capitaine, C. Donati-Martin, D. Féral *The largest eigenvalues of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations* Ann. Probab. 37 (2009)1–47.

[29] M. Capitaine, C. Donati-Martin, D. Féral *Central limit theorems for eigenvalues of deformations of Wigner matrices*. Ann. Inst. Henri Poincaré Probab. Stat., 48, 1, 107–133, 2012.

[30] M. Capitaine, C. Donati-Martin, D. Féral, M. Février *Free convolution with a semi-circular distribution and eigenvalues of spiked deformations of Wigner matrices*, Electron. J. Prob. Vol. 16 (2011), 1750–1792.

[31] G. Cébron, T. Kemp *Fluctuations of Brownian Motions on GL_N*, to appear in Ann. Inst. Henri Poincaré Probab. Stat., arXiv:1409.5624.

[32] S. Chatterjee *Fluctuations of eigenvalues and second order Poincaré inequalities*, Probab. Theory Related Fields, **143**, 2009,1–40.

[33] S. Chatterjee, E. Meckes *Multivariate normal approximation using exchangeable pairs* ALEA 4 (2008).

[34] B. Collins, J.A. Mingo, P. Śniady, R. Speicher *Second order freeness and fluctuations of random matrices. III. Higher order freeness and free cumulants*. Doc. Math. 12 (2007), 1–70.

[35] B. Collins, P. Sniady *Integration with respect to the Haar measure on unitary, orthogonal and symplectic group*. Comm. Math. Phys., 264(3):773–795, 2006.

[36] B. Collins *Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability*. Int. Math. Res. Not. 2003, no. 17, 953–982.

[37] B. Collins, M. Stolz *Borel theorems for random matrices from the classical compact symmetric spaces.* Ann. Probab. Vol. 36, no. 3 (2008), 876–895.

[38] B. Collins, P. Śniady *Integration with respect to the Haar measure on unitary, orthogonal and symplectic group.* Comm. Math. Phys., 264(3):773–795, 2006.

[39] P. Diaconis, M. Shahshahani *On the eigenvalues of random matrices*, Studies in applied probability, J. Appl. Probab. 31A (1994), 49–62.

[40] P. Diaconis, S. Evans *Linear functionals of eigenvalues of random matrices*, Trans. Amer. Math. Soc., **353**, 2001, 2615–2633.

[41] B. Duplantier, R. Rhodes, S. Sheffield, V. Vargas *Log-correlated Gaussian fields: an overview*, arXiv.

[42] J. Feinberg, A. Zee, , *Non-Gaussian non-Hermitian random matrix theory: phase transition and addition formalism*, Nuclear Phys. B 501, 643–669, (1997).

[43] D. Féral, S. Péché *The largest eigenvalue of rank one deformation of large Wigner matrices* Comm. Math. Phys. 272 (2007)185–228.

[44] Y.V. Fyodorov and H.-J. Sommers, *Statistics of S-matrix poles in few-channel chaotic scattering: Crossover from isolated to overlapping resonances* JETP Lett. 63 (1996) 1026-1030.

[45] Y.V. Fyodorov and H.-J. Sommers *Statistics of resonance poles, phase shifts and time delays in quantum chaotic scattering: Random matrix approach for systems with broken time reversal invariance* J. Math. Phys. 38 (1997), 1918-1981.

[46] Y.V. Fyodorov and B.A. Khoruzhenko *Systematic Analytical Approach to Correlation Functions of Resonances in Quantum Chaotic Scattering* Phys. Rev. Lett. 83 (1999), 65-68.

[47] Y.V. Fyodorov and H.-J. Sommers *Random matrices close to Hermitian or unitary: overview of methods and results* J. Phys. A: Math. Gen. 36 (2003) 3303-3347.

[48] V.L. Girko *The elliptic law*, Teoriya Veroyatnostei i ee Primeneniya, 30(4):640–651, 1985.

[49] A. Guionnet, M. Krishnapur, O. Zeitouni *The Single Ring Theorem*. Ann. of Math. (2) 174 (2011), no. 2, 1189–1217.

[50] A. Guionnet, O. Zeitouni *Support convergence in the Single Ring Theorem*. Probab. Theory Related Fields 154 (2012), no. 3-4, 661–675.

[51] X. He, Q. Ai, C. Qiu, W. Huang, L. Piao, H. Liu *A Big Data architecture design for smart grids based on random matrix theory* arXiv:1501.07329

[52] F. Hiai, D. Petz *The semicircle law, free random variables, and entropy*. Amer. Math. Soc., Mathematical Surveys and Monographs Volume 77, 2000.

[53] R.A. Horn, C.R. Johnson, *Matrix Analysis*, Cambridge University Press, ISBN 978-0-521-38632-6 (1985).

[54] T. Jiang *Maxima of entries of Haar distributed matrices*. Probab. Theory Related Fields 131 (2005), no. 1, 121–144.

[55] T. Jiang *How many entries of a typical orthogonal matrix can be approximated by independent normals?* Ann. Probab. 34(4), 1497–1529. 2006.

[56] K. Johansson *On the fluctuations of eigenvalues of random Hermitian matrices.* Duke Math. J. **91** 1998, 151–204.

[57] I.M. Johnstone, *On the distribution of the largest eigenvalue in principal components analysis*, Ann. Statist. 29: 295327. 2001.

[58] D. Jonsson *Some limit theorems for the eigenvalues of a sample covariance matrix.* J. Mult. Anal. **12**, 1982, 1–38.

[59] O. Kallenberg *Foundations of Modern Probability*, Springer, 1997.

[60] A. M. Khorunzhy, B. A. Khoruzhenko, L. A. Pastur *Asymptotic properties of large random matrices with independent entries*, J. Math. Phys. 37 (1996) 5033–5060.

[61] A. Knowles, J. Yin *The isotropic semicircle law and deformation of Wigner matrices*, Comm. Pure Appl. Math. 66 (2013), no. 11, 1663–1750.

[62] A. Knowles, J. Yin *The outliers of a deformed Wigner matrix*, Ann. Probab. 42 (2014), no. 5, 1980–2031.

[63] T. Lévy, M. Maïda *Central limit theorem for the heat kernel measure on the unitary group*. J. Funct. Anal. 259 (2010), no. 12, 3163–3204.

[64] A. Lytova, L. Pastur *Central limit theorem for linear eigenvalue statistics of random matrices with independent entries*, Ann. Probab., **37**, 2009, 1778–1840.

[65] E. Meckes *Linear functions on the classical matrix groups*. Trans. Amer. Math. Soc. 360 (2008), no. 10, 5355–5366.

[66] J.A. Mingo, A. Nica *Annular noncrossing permutations and partitions, and second-order asymptotics for random matrices*. Int. Math. Res. Not. 2004, no. 28, 1413–1460.

[67] J.A. Mingo, P. Śniady, R. Speicher, *Second order freeness and fluctuations of random matrices. II. Unitary random matrices.* Adv. Math. 209 (2007), no. 1, 212-240.

[68] J.A. Mingo, R. Speicher *Second order freeness and fluctuations of random matrices. I. Gaussian and Wishart matrices and cyclic Fock spaces*. J. Funct. Anal. 235 (2006), no. 1, 226–270.

[69] J.A. Mingo, P. Śniady, R. Speicher, *Second order freeness and fluctuations of random matrices. II. Unitary random matrices*. Adv. Math. 209 (2007), no. 1, 212–240.

[70] A. Naumov, *Elliptic law for real random matrices*, arXiv:1201.1639.

[71] H. Nguyen, S. O'Rourke, *The elliptic law*, Int. Math. Res. Not. IMRN 2015, no. 17, 7620–7689.

[72] A. Nica, R. Speicher *Lectures on the combinatorics of free probability*. London Mathematical Society Lecture Note Series, 335. Cambridge University Press, Cambridge, 2006.

[73] S. O'Rourke, D. Renfrew *Low rank perturbations of large elliptic random matrices*, Electron. J. Probab. 19 (2014), no. 43, 65 pp.

[74] S. O'Rourke, D. Renfrew *Central limit theorem for linear eigenvalue statistics of elliptic random matrices*, arXiv:1410.4586v1.

[75] S. O'Rourke, P. Matchett Wood *Spectra of nearly Hermitian random matrices*, arXiv:1510.00039v1.

[76] S. O'Rourke, D. Renfrew, A. Soshnikov *On fluctuations of matrix entries of regular functions of Wigner matrices with non-identically distributed entries*. J. Theoret. Probab. 26 (2013), no. 3, 750–780.

[77] S. Péché *The largest eigenvalue of small rank perturbations of Hermitian random matrices*, Prob. Theory Relat. Fields, 134 127–173, 2006.

[78] Y. Peres, B. Virág *Zeros of the i.i.d. Gaussian power series: a confor- mally invariant determinantal process*. Acta. Math. 194. (2005) 1 – 35.

[79] A. Pizzo, D. Renfrew and A. Soshnikov, *On finite rank deformations of Wigner matrices*, arXiv preprint arXiv:1103.3731v4 2011

[80] A. Pizzo, D. Renfrew, A. Soshnikov *Fluctuations of matrix entries of regular functions of Wigner matrices*. J. Stat. Phys. 146 (2012), no. 3, 550–591.

[81] E. Rains *Normal limit theorems for symmetric random matrices* Probab. Theory Relat. Fields 112, 411–423 (1998).

[82] A.B. Rajagopalan *Outlier eigenvalue fluctuations of perturbed iid matrices*, arXiv:1507.01441.

[83] B. Rider, J. W. Silverstein *Gaussian fluctuations for non-Hermitian random matrix ensembles*. Ann. Probab. 34 (2006), no. 6, 2118–2143.

[84] B. Rider, B. Virág *The noise in the circular law and the Gaussian free field*. Int. Math. Res. Not. IMRN 2007, no. 2.

[85] J. Rochet *Complex outliers of Hermitian random matrices*, arXiv:1507.00455v2.

[86] M. Rudelson, R. Vershynin *Invertibility of random matrices: unitary and orthogonal perturbations*, J. Amer. Math. Soc. 27 (2014), 293–338.

[87] M. Shcherbina *Central Limit Theorem for Linear Eigenvalue Statistics of the Wigner and Sample Covariance Random Matrices*, Journal of Mathematical Physics, Analysis, Geometry, 7(2), (2011), 176–192.

[88] Y. Sinai, A. Soshnikov *Central limit theorem for traces of large random symmetric matrices with independent matrix elements*, Bol. Soc. Brasil. Mat. (N.S.), **29**, 1998, 1–24.

[89] H.-J. Sommers, A. Crisanti, H. Sompolinsky, Y. Stein *Spectrum of large random asymmetric matrices*. Phys. Rev. Lett. 60, 1895 1899 (1988).

[90] A. Soshnikov *The central limit theorem for local linear statistics in classical compact groups and related combinatorial identities*, Ann. Probab., 28 (2000), 1353–1370.

[91] T. Tao *Topics in random matrix theory*, Graduate Studies in Mathematics, AMS, 2012.

[92] T. Tao *Outliers in the spectrum of i.i.d. matrices with bounded rank perturbations*. Probab. Theory Related Fields 155 (2013), no. 1-2, 231–263.

[93] B. Tsou *The Distribution of Permutation Matrix Entries Under Randomized Basis*, arxiv:1509.06505.

[94] C. Zhang, R.C. Qiu *Data Modeling with Large Random Matrices in a Cognitive Radio Network Testbed: Initial Experimental Demonstrations with 70 Nodes*, arXiv:1404.3788.