



**HAL**  
open science

# Central limit theorems for linear statistics of heavy tailed random matrices

Florent Benaych-Georges, Alice Guionnet, Camille Male

► **To cite this version:**

Florent Benaych-Georges, Alice Guionnet, Camille Male. Central limit theorems for linear statistics of heavy tailed random matrices. 2013. hal-00769741v1

**HAL Id: hal-00769741**

**<https://hal.science/hal-00769741v1>**

Preprint submitted on 3 Jan 2013 (v1), last revised 25 Dec 2013 (v5)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# CENTRAL LIMIT THEOREMS FOR LINEAR STATISTICS OF HEAVY TAILED RANDOM MATRICES

FLORENT BENAYCH-GEORGES, ALICE GUIONNET, CAMILLE MALE

ABSTRACT. We show central limit theorems (CLT) for the Stieltjes transforms or more general analytic functions of symmetric matrices with independent heavy tailed entries, including entries in the domain of attraction of  $\alpha$ -stable laws and entries with moments exploding with the dimension, as in the adjacency matrices of Erdős-Rényi graphs. For the second model, we also prove a central limit theorem of the moments of its empirical eigenvalues distribution. The limit laws are Gaussian, but unlike to the case of standard Wigner matrices, the normalization is the one of the classical CLT for independent random variables.

## 1. INTRODUCTION AND STATEMENT OF RESULTS

Recall that a Wigner matrix is a symmetric random matrix  $A = (a_{i,j})_{i,j=1,\dots,N}$  such that

1. the sub-diagonal entries of  $A$  are independent and identically distributed (i.i.d.),
2. the random variables  $\sqrt{N}a_{i,j}$  are distributed according to a measure  $\mu$  that does not depend on  $N$  and have all moments finite.

This model was introduced in 1956 by Wigner [37] who proved the convergence of the moments

$$(1) \quad \lim_{N \rightarrow \infty} \mathbb{E} \left[ \frac{1}{N} \text{Tr}(A^p) \right] = \int x^p \frac{\sqrt{4-x^2}}{2\pi} dx$$

when  $\mu$  is centered with unit variance. Moments can be easily replaced by bounded continuous functions in the above convergence and this convergence holds almost surely. Assumption 2 can also be weakened to assume only that the second moment is finite. The

---

2000 *Mathematics Subject Classification.* 15A52;60F05.

*Key words and phrases.* Random matrices, heavy tailed random variables, central limit theorem.

FBG: florent.benaych-georges@parisdescartes.fr, MAP 5, UMR CNRS 8145 - Université Paris Descartes, 45 rue des Saints-Pères 75270 Paris Cedex 6, France and CMAP, École Polytechnique, route de Saclay, 91128 Palaiseau Cedex, France.

AG: aguionne@ens-lyon.fr, CNRS & École Normale Supérieure de Lyon, Unité de mathématiques pures et appliquées, 46 allée d'Italie, 69364 Lyon Cedex 07, France and MIT, Mathematics department, 77 Massachusetts Av, Cambridge MA 02139-4307, USA.

CM: male@math.univ-paris-diderot.fr, Laboratoire de probabilités et modèles aléatoires, Université Paris Diderot, 175 rue Chevaleret, 75013, France.

Research supported by ANR-08-BLAN-0311-01.

fluctuations around this limit or around the expectation were first studied by Jonsson [28] in the (slightly different) Wishart model, then by Sinai and Soshnikov [35] with  $p \ll N^{1/2}$  possibly going to infinity with  $N$ . Since then, a long list of further-reaching results have been obtained: the central limit theorem was extended to so-called matrix models where the entries interact via a potential in [27], the set of test functions was extended and the assumptions on the entries of the Wigner matrices weakened in [7, 6, 29], a more general model of band matrices was considered in [2] (see also [29, 5] for general covariance matrices), unitary matrices were considered in [19], and Chatterjee developed a general approach to these questions in [17], under the condition that the law  $\mu$  can be written as a transport of the Gaussian law. Finally, but this is not really our concern here, the fluctuations of the trace of words in several random matrices were studied in [16, 25, 32, 26]. It turns out that in these cases

$$\mathrm{Tr}(A^p) - \mathbb{E}[\mathrm{Tr}(A^p)]$$

converges towards a Gaussian variable whose covariance depends on the first four moments of  $\mu$ . Moments can also be replaced by regular enough functions and Assumption 2 can be weakened to assume that the fourth moment only is finite. The latter condition is however necessary as the covariance for the limiting Gaussian depends on it. The absence of normalization by  $\sqrt{N}$  shows that the eigenvalues of  $A$  fluctuate very little, as precisely studied by Erdős, Schlein, Yau, Tao, Vu and their co-authors, who analyzed their rigidity in e.g. [21, 22, 36].

In this article, we extend these results for two variations of the Wigner matrix model where Assumption 2 is removed or weakened: we consider cases where  $\mu$  does not have any second moment or cases where  $\mu$  depends on  $N$ , with moments growing with  $N$ . In these cases, Wigner's convergence theorem (1) does not hold, even when moments are replaced by smooth bounded functions. The analogue of the convergence (1) was studied when the common law  $\mu$  of the entries of  $\sqrt{N}A$  belongs to the domain of attraction of an  $\alpha$ -stable law (in which case the normalization in  $\sqrt{N}$  has to be adapted) or  $\mu$  depends on  $N$  and has moments blowing up with  $N$ . We describe these models below more precisely.

**Definition 1.1** (Models of symmetric heavy tailed matrices with i.i.d. sub-diagonal entries).

Let  $A = (a_{i,j})_{i,j=1,\dots,N}$  be a random symmetric matrix with i.i.d. sub-diagonal entries.

1. We say that  $A$  is a **Lévy matrix** of parameter  $\alpha$  in  $]0, 2[$  when  $A = X/a_N$  where the absolute values of the entries  $X_{ij}$  of  $X$  are in the domain of attraction of  $\alpha$ -stable distribution, more precisely

$$(2) \quad P(|X_{ij}| \geq u) = \frac{L(u)}{u^\alpha}$$

with a slowly varying function  $L$ , and

$$a_N = \inf\left\{u : P(|X_{ij}| \geq u) \leq \frac{1}{N}\right\}$$

( $a_N = \tilde{L}(N)N^{1/\alpha}$ , with  $\tilde{L}(\cdot)$  a slowly varying function).

2. We say that  $A$  is a **Wigner matrix with exploding moments** with parameter  $(C_k)_{k \geq 1}$  whenever the entries of  $A$  are centered, for any  $k \geq 1$

$$(3) \quad N\mathbb{E}[(a_{ij})^{2k}] \xrightarrow{N \rightarrow \infty} C_k,$$

and  $\sum_{k \geq 2} C_k > 0$ .

Notice that if there exists a  $C_k > 0$  for a  $k \geq 2$ , then actually  $C_k > 0$  for any  $k \geq 2$  [31, Appendix A]. A Lévy matrix whose entries are truncated in an appropriate way is a Wigner matrix with exploding moments [10, 31, 38]. The adjacency matrix of an Erdős-Rényi graph, *i.e.* a matrix  $A$  such that

$$(4) \quad A_{ij} = 1 \text{ with probability } p/N \text{ and } 0 \text{ with probability } 1 - p/N,$$

is also an exploding moments Wigner matrix. In this case the fluctuations were already studied in [34].

The weak convergence of the empirical eigenvalues distribution of a Lévy matrix has been established in [10] (see also [15, 8, 13]) where it was shown that for any bounded continuous function  $f$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \text{Tr}(f(A)) = \int f(x) d\mu_\alpha(x) \quad a.s.$$

where  $\mu_\alpha$  is a heavy tailed probability measure which depends only on  $\alpha$ . Moreover,  $\mu_\alpha$  converges towards the semicircle law as  $\alpha$  goes to 2.

The convergence in moments, in expectation and in probability, of the empirical eigenvalues distribution of a Wigner matrix with exploding moments has been established by Zakharevich in [38]. If we assume that the sequence  $(C_k)_{k \geq 1}$  has at most exponential growth, *i.e.* that for a constant  $C$ , for all  $k$ ,

$$(5) \quad C_k \leq C^k,$$

then for any continuous bounded function  $f$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \text{Tr}(f(A)) = \int f(x) d\mu_{\mathbf{C}}(x) \quad a.s.$$

where  $\mu_{\mathbf{C}}$  is a probability measure which depends only on the sequence  $\mathbf{C} := (C_k)_{k \geq 1}$ .

We shall first establish the fluctuations of moments of a Wigner matrix with exploding moments around their limit, namely prove the following theorem.

**Theorem 1.2** (The CLT for moments of Wigner matrices with exploding moments).

Let  $A = (a_{i,j})_{i,j=1,\dots,N}$  be a Wigner matrix with exploding moments with parameter  $(C_k)_{k \geq 1}$ . Then the process

$$(6) \quad \left( \frac{1}{\sqrt{N}} \text{Tr} A^K - \mathbb{E} \left[ \frac{1}{\sqrt{N}} \text{Tr} A^K \right] \right)_{K \geq 1}$$

converges in distribution to a centered Gaussian process.

The expression of the covariance of the process is of combinatorial nature, given it later in Section 2, Formula (30) and Theorem 2.2.

Note that the speed of the central limit theorem is  $N^{-1/2}$  as for independent square integrable random variables, differently from what happens for standard Wigner's matrices. This phenomenon has already been observed for adjacency matrix of random graphs [9, 20] and we will see below that it also holds for Lévy matrices. It suggests that the repulsive interactions exhibited by the eigenvalues of most models of random matrices with not too large entries (see the first paragraph of this introduction) are absent or much less strong in the context of heavy tailed matrices.

For Lévy matrices, moments do not make sense and one should consider smooth bounded test functions. We start as is common in random matrix theory, with the study of the normalized trace of the resolvent of  $A$ , given for  $z \in \mathbb{C} \setminus \mathbb{R}$  by

$$G(z) := \frac{1}{z - A}.$$

By the previous results, for both Lévy and exploding moments Wigner matrices satisfying (5),  $N^{-1} \text{Tr} G(z)$  converges in probability to a deterministic limit as the parameter  $N$  tends to infinity. We study the associated fluctuations. In fact, even in the case of Wigner matrices with exploding moments, the CLT for moments does not imply a priori the CLT for Stieltjes functions even though concentration inequalities hold on the right scale, see [14], as one cannot approximate for the total variation norm smooth bounded functions by polynomials unless on compact subset, a point which is not clear in this heavy tail setting. However, with additional arguments based on martingale technology, we shall prove the following result, valid for both Lévy matrices and Wigner matrices with exploding moments.

**Theorem 1.3** (CLT for Stieltjes transforms). *Let  $A$  be an  $N$  by  $N$  Hermitian matrix with i.i.d. real entries below the diagonal defined in Definition 1.1, satisfying (5) in the case of Wigner matrices with exploding moments. Let*

$$(7) \quad \phi_N(\lambda) := \mathbb{E}[\exp(-i\lambda|a_{11}|^2)],$$

*defined for any  $\lambda$  in  $\overline{\mathbb{C}^-}$ , with  $\mathbb{C}^- := \{\lambda \in \mathbb{C}; \Im \lambda < 0\}$ . Assume that there is a function  $\Phi$  on  $\overline{\mathbb{C}^-}$  such that as  $N \rightarrow \infty$ , we have the convergence*

$$(8) \quad N(\phi_N(\lambda) - 1) \longrightarrow \Phi(\lambda).$$

*uniformly on compact subsets of  $\overline{\mathbb{C}^-}$ .*

*Then, the process*

$$(9) \quad \left( \frac{1}{\sqrt{N}} \text{Tr} G(z) - \frac{1}{\sqrt{N}} \mathbb{E}[\text{Tr} G(z)] \right)_{z \in \mathbb{C} \setminus \mathbb{R}}$$

*converges in distribution (in the sense of finite marginals) to a centered complex Gaussian process  $Z(z)$ ,  $z \in \mathbb{C} \setminus \mathbb{R}$ , whose covariance is defined by the formulas:*

$$(10) \quad Z(\bar{z}) = \overline{Z(z)} \quad \text{and} \quad C(z, z') := \mathbb{E}[Z(z)Z(z')] = L(z, z') - L(z)L(z'),$$

where the analytic functions  $L(\cdot)$  and  $L(\cdot, \cdot)$  are given respectively in Formulas (13) and (14).

Note that the uniform convergence on compact subsets, in (8), implies that  $\Phi$  is analytic on  $\mathbb{C}^-$  and continuous on  $\overline{\mathbb{C}^-}$ .

As a consequence, we can extend the central limit theorem to a larger class of functions.

**Corollary 1.4.** *Let  $\mathcal{A}$  be the set of analytic functions  $f$  on the strip  $\{z = a + ib, a \in \mathbb{R}, |b| \leq \epsilon\}$  for some  $\epsilon > 0$  so that there exists  $\gamma > 1$  so that for a large enough*

$$(11) \quad |f(a \pm i\epsilon)| \leq a^{-\gamma}.$$

*Then under the hypotheses of Theorem 1.3, the process*

$$(12) \quad \left( Z_N(f) := \frac{1}{\sqrt{N}} \text{Tr} f(A) - \frac{1}{\sqrt{N}} \mathbb{E}[\text{Tr} f(A)]; f \in \mathcal{A} \right)$$

*converges in law towards a centered Gaussian process  $(Z(f); f \in \mathcal{A})$  with covariance defined by the facts that  $Z(\overline{f^c}) = \overline{Z(f)}$ , with  $\overline{f^c}$  the function of  $\mathcal{A}$  defined by the formula  $\overline{f^c}(z) := \overline{f(\overline{z})}$ , and*

$$\text{Cov}(f, g) := \frac{1}{(2i\pi)^2} \int_{\Gamma} dz \int_{\Gamma} dz' f(z)g(z')C(z, z'),$$

*where  $\Gamma = ]-\infty + i\epsilon, +\infty + i\epsilon[ \cup ]+\infty - i\epsilon, -\infty - i\epsilon[$  and  $C(z, z')$  is the function of (10). The result extends to the closure of  $\mathcal{A}$  for the total variation norm.*

Before giving the explicit form of the maps involved in the formula of the covariance, let us first discuss the hypothesis made at Equation (8).

- For a Lévy matrix, if  $A = X/a_N$  with  $|X_{ij}|$  in the domain of attraction of an  $\alpha$ -stable law,  $X_{ij}^2$  is in the domain of attraction of an  $\alpha/2$ -stable law, and we have (see [23]) that

$$\Phi(\lambda) = -\sigma(i\lambda)^{\alpha/2}$$

for some constant  $\sigma \geq 0$ .

- For a Wigner matrix with exploding moments, at least on the formal level, we have

$$\phi_N(\lambda) = 1 + \frac{1}{N} \sum_{k \geq 1} \frac{(-i\lambda)^k}{k!} N\mathbb{E}[|a_{11}|^{2k}].$$

Therefore, if (3) holds together with the reinforcement of (5) into

$$\sup_{N \geq 1} N\mathbb{E}[(a_{ij})^{2k}] \leq C^k,$$

then (8) is satisfied. In particular, the case where  $A$  is the adjacency matrix of an Erdős-Rényi graph defined at (4) satisfies the assumptions with

$$\Phi(\lambda) = p(e^{-i\lambda} - 1).$$

The covariance in Theorem 1.3 is defined in terms of the functions  $L(\cdot)$  and  $L(\cdot, \cdot)$  that have the following form: for any  $u$  in  $]0, 1[$  and any  $z, z'$  in  $\mathbb{C} \setminus \mathbb{R}$ ,

$$(13) \quad L(z) = - \int_0^{\operatorname{sgn} \Im(z)\infty} \frac{1}{t} \partial_z e^{itz + \rho_z(t)} dt$$

$$(14) \quad L(z, z') = \int_0^1 \Psi^u(z, z') du, \quad \text{where}$$

$$(15) \quad \Psi^u(z, z') = \int_0^{\operatorname{sgn} \Im(z)\infty} \int_0^{\operatorname{sgn} \Im(z')\infty} \frac{1}{tt'} \partial_{z, z'}^2 e^{itz + it'z' + \rho_{z, z'}^u(t, t')} dt dt'$$

and where the maps  $\rho_z(t)$  and  $\rho_{z, z'}^u(t, t')$  are given as the non random limits

$$(16) \quad \rho_z(t) = \lim_{N \rightarrow \infty} \frac{1}{N-1} \sum_{j=1}^{N-1} \Phi(tG_k(z)_{jj}), \quad a.s$$

$$(17) \quad \rho_{z, z'}^u(t, t') = \lim_{\substack{N \rightarrow \infty \\ k/N \rightarrow u}} \frac{1}{N-1} \sum_{j=1}^{N-1} \Phi(tG_k(z)_{jj} + t'G'_k(z')_{jj}), \quad a.s$$

where  $G_k(z) = (z - A_k)^{-1}$  for  $A_k$  the matrix obtained from  $A$  by deleting the  $k$ -th row and column, and  $G'_k(z') = (z' - A'_k)^{-1}$  for  $A'_k$  a copy of  $A_k$  where the entries  $(i, j)$  for  $i, j \geq k$  are independent of  $A_k$  and the other are those of  $A_k$ . Also, we assume  $t\Im z \geq 0$  and  $t'\Im z' \geq 0$  in (16) and (17).

The existence of the limits in (16) and (17) is a consequence of a generalized convergence in moments, namely the convergence in distribution of traffics, of  $(A_k, A'_k)$  stated in [31], see Lemma 3.2. However, under stronger assumptions, independent proofs of these convergences and more intrinsic characterizations of  $\rho_z$  and  $\rho_{z, z'}^u$  are provided in Theorems 1.5 and 1.6 below.

Let us mention that the map  $(t, z) \mapsto \rho_z(t)$  characterizes the limiting eigenvalues distribution of  $A$ , since under the assumptions of Theorem 1.3, for all  $z \in \mathbb{C}^+$ , we have the almost sure convergence

$$\lim_{N \rightarrow \infty} \frac{1}{N} \operatorname{Tr} G_N(z) = i \int_0^\infty e^{itz + \rho_z(t)} dt.$$

This fact was known in the Lévy case [10] and is proved in Corollary 3.3 for Wigner matrices with exploding moments.

**Theorem 1.5** (A fixed point equation for  $\rho_z(t)$ ). *We assume the hypotheses of Theorem 1.3 and, moreover, we assume that we have the following decomposition for all  $z \in \mathbb{C}^-$ :*

$$(18) \quad \Phi(z) = \int_0^\infty g(y) e^{i\frac{y}{z}} dy$$

where  $g(y)$  is a function such that for some constants  $K, \gamma > -1, \kappa \geq 0$ , we have

$$(19) \quad |g(y)| \leq K1_{y \leq 1} y^\gamma + K1_{y \geq 1} y^\kappa, \quad \forall y > 0.$$

Then the function  $(z \in \mathbb{C}^+, \lambda \in \mathbb{R}^+) \mapsto \rho_z(\lambda)$  is analytic in its first argument and continuous in the second, with non positive real part, and characterized among the set of such functions by the formula

$$(20) \quad \rho_z(\lambda) = \lambda \int_0^\infty g(\lambda y) e^{iyz + \rho_z(y)} dy.$$

The assumption (18) is satisfied for our main examples.

- In the heavy tailed case, we have  $\Phi(x) = -\sigma(ix)^{\frac{\alpha}{2}}$ , hence the Hypothesis made at (18) is true, with  $g(y) = Cy^{\frac{\alpha}{2}-1}$ . Indeed, for any  $z \in \mathbb{C}^+$ , an application of residues formula gives

$$(21) \quad \Gamma(\alpha/2) = -i \int_{t=0}^{+\infty} (-izt)^{\frac{\alpha}{2}-1} e^{itz} z dt,$$

(here,  $x^\gamma$  has cut on  $\mathbb{R}^-$ ).

- In the case of the Erdős-Rényi adjacency matrix,  $\Phi(x) = p(e^{-ix} - 1)$  so that

$$(22) \quad \Phi(x) = -p \int_0^\infty J_1(s) e^{\frac{ix-1}{4}s^2} ds = -p \int_0^\infty J_1(2\sqrt{s}) e^{ix^{-1}s} \frac{1}{\sqrt{s}} ds$$

where  $J_1(s) = \frac{s}{2} \sum_{k \geq 0} \frac{(-s^2/4)^k}{k!(k+1)!}$  is a Bessel function of the first kind.  $J_1$  is bounded on  $\mathbb{R}_+$ .

- More generally, if  $\Phi$  is so that  $\Phi(x^{-1})$  is in  $L^1$ , by Fourier integral formula we can write it

$$\Phi\left(\frac{1}{x}\right) = \int_{-\infty}^{\infty} \hat{\Phi}_{-1}(y) e^{iyx} dy$$

with  $\hat{\Phi}_{-1}$  the Fourier transform of  $\Phi(1/x)$ . Hence, if  $\hat{\Phi}_{-1}$  is supported on  $\mathbb{R}^+$  we have

$$\Phi(x) = \int_0^\infty \hat{\Phi}_{-1}(y) e^{-iy/x} dy,$$

so that our hypothesis is satisfied if  $\hat{\Phi}_{-1}$  satisfies (19).

Note that the asymptotics of Wigner matrices with bound moments is also described by (20). In this case  $\Phi(\lambda) = -i\lambda$ , so  $g(y) = -i$  and  $\rho_z(t) = -it \lim_{N \rightarrow \infty} \frac{1}{N} \text{Tr } G(z)$ , which leads to the classical quadratic equation

$$(23) \quad s(z) := \lim_{N \rightarrow \infty} \frac{1}{N} \text{Tr } G(z) = \frac{1}{z - s(z)}.$$

Let us now give a fixed point characterization for the function  $\rho_{z,z'}(t, t')$  of (17).

**Theorem 1.6** (A fixed point system of equations for  $\rho_{z,z'}^u(t, t')$ ). *We assume the hypotheses of Theorem 1.3 and, moreover, we assume that  $\Phi$  satisfies the formula*

$$(24) \quad \Phi(x + y) = \int_0^\infty \int_0^\infty e^{i\frac{v}{x} + i\frac{v'}{y}} d\tau(v, v') + \int e^{i\frac{v}{x}} d\mu(v) + \int e^{i\frac{v'}{y}} d\mu(v')$$



for some complex measures  $\tau, \mu$  on respectively  $\mathbb{R}^+ \times \mathbb{R}^+$  and  $\mathbb{R}^+$  such that for all  $b > 0$ ,  $\int e^{-bv} d|\mu|(v)$  is finite for some constants  $K > 0$ ,  $-1 < \gamma, \gamma' \leq 0$  and  $\kappa, \kappa' \geq 0$ , we have

$$\frac{d|\tau|(v, v')}{dv dv'} \leq K(v^\gamma 1_{v \in ]0,1]} + v^\kappa 1_{v \in ]1,\infty[})(v'^{\gamma'} 1_{v' \in ]0,1]} + v'^{\kappa'} 1_{v' \in ]1,\infty[})$$

Then,  $\rho_{z,z'}^u(t, s) = u\rho_{z,z'}^{u,1}(t, s) + (1-u)\rho_{z,z'}^{u,2}(t, s)$  where  $\rho^{u,1}, \rho^{u,2}$  are solution of the fixed point system of equations

$$\begin{aligned} \rho_{z,z'}^{u,1}(t, s) &= \int_0^{\operatorname{sgn}_z \infty} \int_0^{\operatorname{sgn}_{z'} \infty} e^{i\frac{v}{t}z + i\frac{v'}{s}z'} e^{\rho_{z,z'}^u(\frac{v}{t}, \frac{v'}{s})} d\tau(v, v') \\ &\quad + \int_0^{\operatorname{sgn}_z \infty} e^{i\frac{v}{t}} e^{\rho_z(\frac{v}{t})} d\mu(v) + \int_0^{\operatorname{sgn}_{z'} \infty} e^{i\frac{v'}{s}} e^{\rho_{z'}(\frac{v'}{s})} d\mu(v) \\ \rho_{z,z'}^{u,2}(t, s) &= \int_0^{\operatorname{sgn}_z \infty} \int_0^{\operatorname{sgn}_{z'} \infty} e^{i\frac{v}{t}z + i\frac{v'}{s}z'} e^{u\rho_{z,z'}^{u,1}(\frac{v}{t}, \frac{v'}{s}) + (1-u)[\rho_z(\frac{v}{t}) + \rho_{z'}(\frac{v'}{s})]} d\tau(v, v') \\ (25) \quad &\quad + \int_0^{\operatorname{sgn}_z \infty} e^{i\frac{v}{t}} e^{\rho_z(\frac{v}{t})} d\mu(v) + \int_0^{\operatorname{sgn}_{z'} \infty} e^{i\frac{v'}{s}} e^{\rho_{z'}(\frac{v'}{s})} d\mu(v) \end{aligned}$$

There is a unique analytic solution  $(\rho^{u,1}, \rho^{u,2})$  from  $\Lambda = \{t/\Im z > 0, s/\Im z' > 0\}$  into  $\{z; \Re(z) \leq 0\}^2$ .

In our main examples:

- (24) is satisfied in the case of the Erdős-Rényi graph adjacency matrices. Indeed, in this case, by (22),

$$\begin{aligned} \Phi(x+y) &= p(e^{x+y} - 1) \\ &= p(e^x - 1)(e^y - 1) + p(e^x - 1) + p(e^y - 1) \\ &= p \int_0^\infty J_1(v) dv \int_0^\infty J_1(v') dv' e^{i\frac{v^2}{4x} + i\frac{v'^2}{4y}} \\ &\quad - p \int_0^\infty J_1(v) dv e^{i\frac{v^2}{4x}} - p \int_0^\infty J_1(v') dv' e^{i\frac{v'^2}{4y}} \end{aligned}$$

so that if  $\nu$  denotes the measure on  $\mathbb{R}^+$  defined by

$$d\nu(v) = J_1(2\sqrt{v}) \frac{1}{\sqrt{v}} dv$$

then we can take  $\tau = p\nu^{\otimes 2}$  and  $\mu = -p\nu$ , which satisfies our assumption with  $\gamma' = \gamma = -1/2$ ,  $\kappa' = \kappa = 0$ .

- For Lévy matrices, where  $\Phi(x) = C_\alpha x^{\alpha/2}$ , (24) holds as well. In fact, in this case we have for all  $x, y \in \mathbb{C}^+$ ,

$$\begin{aligned}\Phi(x^{-1} + y^{-1}) &= C_\alpha \left(\frac{1}{x} + \frac{1}{y}\right)^{\alpha/2} \\ &= C_\alpha \frac{1}{(x)^{\alpha/2}} \frac{1}{(y)^{\alpha/2}} (x + y)^{\alpha/2} \\ &= C_\alpha \int_0^\infty dw \int_0^\infty dw' \int_0^\infty dv w^{\alpha/2-1} (w')^{\alpha/2-1} v^{-\alpha/2-1} e^{iwx+iw'y} (e^{iv(x+y)} - 1)\end{aligned}$$

so that our hypothesis holds with  $\mu = 0$  and

$$\frac{d\tau(v, v')}{dv dv'} = C_\alpha \int_{0 \leq u \leq v \wedge v'} du u^{-\alpha/2-1} [(v-u)^{\alpha/2-1} (v'-u)^{\alpha/2-1} - v^{\alpha/2-1} (v')^{\alpha/2-1}]$$

It is not hard to verify that  $\tau$  satisfies our hypotheses with  $\gamma = \gamma' = 1 - \alpha/4$  and  $\kappa' = \kappa = 0$ .

Let us conclude this introduction with three remarks.

- (1) Let  $A = X/a_N$ , with  $X = (X_{ij})$  be a symmetric real  $N \times N$  random matrix with i.i.d. entries independent of  $N$  such that

$$P(|X_{ij}| \geq u) = \frac{L(u)}{u^2},$$

with  $L(\cdot)$  a slowly varying function at  $+\infty$  and

$$a_N = \inf\{u : P(|X_{ij}| \geq u) \leq \frac{1}{N}\}$$

(in other words,  $A$  is an  $\alpha = 2$  version of the Lévy matrices defined at Definition 1.1). Then using Example c) p. 44. of [23] instead of the hypothesis made at Equation (7), one can prove that as  $N \rightarrow \infty$ , the empirical spectral law of  $A$  converges almost surely to the semi-circle law with support  $[-2, 2]$  (see (23)). This result somehow “fills the gap” between heavy-tailed matrices and finite second moment Wigner matrices. It allows for example to state that if  $P(|X_{ij}| \geq u) \sim cu^{-2}$ , with  $c > 0$ , even though the entries of  $X$  do not have any second moment, we have that the empirical spectral law of  $\frac{X}{\sqrt{cn \log(n)}}$  converges almost surely to the semi-circle law with support  $[-2, 2]$ .

- (2) About applications of our results to **standard Wigner matrices**, *i.e.* real symmetric matrices  $A$  with i.i.d. centered entries (up to the symmetry) with variance  $1/N$  : one can see from the proofs that Theorems 1.3 and 1.5 and Corollary 1.4 also apply for such matrices and that the function  $\Phi$  of (8) is then linear, which implies that  $L(z, z') = L(z)L(z')$  for all  $z, z'$ , so that the covariance is null. In

this case, (20) is the self-consistent equation satisfied by the Stieltjes transform of the semi-circle law, namely (23), and Corollary 1.4 only means that for functions  $f \in \mathcal{A}$ , we have, for the convergence in probability,

$$(26) \quad \text{Tr } f(A) - \mathbb{E}[\text{Tr } f(A)] = o(\sqrt{N}).$$

When the entries of the Wigner matrix  $A$  have a fourth moment, (26) does not bring any new information, because it is already well-known that its LHT converges in distribution to a Gaussian law. However, for Wigner matrices whose entries have a second but not a fourth moment, (26) brings new information. Indeed, for such matrices, which could be called “*semi heavy-tailed random matrices*”, the convergence to the semi circle law holds (see [4] or the remark right above that one) but the largest eigenvalues do not tend to the upper-bound of the support of the semi-circle law, are asymptotically in the scale  $N^{\frac{4-\alpha}{2\alpha}}$  (with  $\alpha \in (2, 4)$  as in Equation (2) when such an exponent exists) and distributed according to a Poisson process (see [3]), and it is not clear what the rate of convergence to the semi-circle law will be. Equation (26) says that this rate is  $\ll N^{-1/2}$ .

- (3) About recentering with respect to the limit instead of the expectation, everything depends only on the rate of convergence in (3) or in (7). For instance, if  $N\mathbb{E}(a_{11}^{2k}) - C_k = o(\sqrt{N}^{-1})$  for any  $k \geq 1$ , then

$$\sqrt{N} \left( \mathbb{E} \left[ \frac{1}{N} \text{Tr } A^k \right] - \lim_{N \rightarrow \infty} \mathbb{E} \left[ \frac{1}{N} \text{Tr } A^k \right] \right) \xrightarrow[N \rightarrow \infty]{} 0,$$

but otherwise a non trivial recentering should occur. See the end of Section 2.

**Notation:** In this article, the power functions are defined on  $\mathbb{C} \setminus \mathbb{R}_-$  via the standard determination of the argument on this set taking values in  $(-\pi, \pi)$ . The set  $\mathbb{C}^+$  (resp.  $\mathbb{C}^-$ ) denotes the open upper (resp. lower) half plane.

## 2. CLT FOR THE MOMENTS OF WIGNER MATRICES WITH EXPLODING MOMENTS

The goal of this section is to prove Theorem 1.2. In order to prove the CLT for the moments of the empirical eigenvalues distribution of  $A$ , we use a modification of the method of moments inspired by [30] which consists in studying more general functionals of the entries of the matrix (the so-called traffic cumulants) than only its moments. We describe this approach below. Let  $A$  be a Wigner matrix with exploding moments. Let  $K \geq 1$  be an integer. The trace of the  $K$ -th power of  $A$  can be expanded in the following way.

$$\begin{aligned} \frac{1}{N} \text{Tr } A^K &= \frac{1}{N} \sum_{i_1, \dots, i_K=1}^N A(i_1, i_2) \dots A(i_{K-1}, i_K) A(i_K, i_1) \\ &= \sum_{\pi \in \mathcal{P}(K)} \underbrace{\frac{1}{N} \sum_{\mathbf{i} \in S_\pi} A(i_1, i_2) \dots A(i_{K-1}, i_K) A(i_K, i_1)}_{\tau_N^0[\pi]}, \end{aligned}$$

where  $\mathcal{P}(K)$  is the set of partitions of  $\{1, \dots, K\}$  and  $S_\pi$  is the set of multi-indices  $\mathbf{i} = (i_1, \dots, i_K)$  in  $\{1, \dots, N\}^K$  such that  $n \sim_\pi m \Leftrightarrow i_n = i_m$ .

We interpret  $\tau_N^0$  as a functional on graphs instead of partitions. Let  $\pi$  be a partition of  $\{1, \dots, K\}$ . Let  $T^\pi = (V, E)$  be the undirected graph (with possibly multiple edges and loops) whose set of vertices  $V$  is  $\pi$  and with multi-set of edges  $E$  given by: there is one edge between two blocks  $V_i$  and  $V_j$  of  $\pi$  for each  $n$  in  $\{1, \dots, K\}$  such that  $n \in V_i$  and  $n+1 \in V_j$  (with notation modulo  $K$ ). Then, one has

$$(27) \quad \tau_N^0(\pi) = \tau_N^0[T^\pi]$$

if for graph  $T = (V, E)$ , we have denoted

$$(28) \quad \tau_N^0(T) = \frac{1}{N} \sum_{\substack{\phi: V \rightarrow [N] \\ \text{injective}}} \prod_{e \in E} A(\phi(e)),$$

where  $[N] = \{1, \dots, N\}$  and for any edge  $e = \{i, j\}$  we have denoted  $A(\phi(e)) = A(\phi(i), \phi(j))$ . There is no ambiguity in the previous definition since the matrix  $A$  is symmetric.

In order to prove the convergence of

$$(Z_N(K))_{K \geq 1} = \left( \frac{1}{\sqrt{N}} \text{Tr} A^K - \mathbb{E} \left[ \frac{1}{\sqrt{N}} \text{Tr} A^K \right] \right)_{K \geq 1}$$

to a Gaussian process, it is sufficient to prove the convergence of

$$(29) \quad (Z_N(T^\pi))_{\pi \in \cup_K \mathcal{P}(K)} := \left( \sqrt{N} \left( \tau_N^0[T^\pi] - \mathbb{E}[\tau_N^0[T^\pi]] \right) \right)_{\pi \in \cup_K \mathcal{P}(K)}$$

to a Gaussian process, since

$$(30) \quad (Z_N(K))_{K \geq 1} = \left( \sum_{\pi \in \mathcal{P}(K)} Z_N(T^\pi) \right)_{K \geq 1}.$$

Before giving the proof of this fact, we recall a result from [31], namely the convergence of  $\tau_N^0[T^\pi]$  for any partition  $\pi$ . These limits are involved in our computation of the covariance of the limiting process of  $(Z_N(T^\pi))_{\pi \in \cup_K \mathcal{P}(K)}$ , and this convergence will be useful in the proof of the CLT for Stieltjes transforms latter.

**Proposition 2.1** (Convergence of generalized moments). *Let  $A$  be a Wigner matrix with exploding moments with parameter  $(C_k)_{k \geq 1}$ . For any partition  $\pi$  in  $\cup_K \mathcal{P}(K)$ , with  $\tau_N^0[T^\pi]$  defined in (27),*

$$\mathbb{E}[\tau_N^0[T^\pi]] \xrightarrow{N \rightarrow \infty} \tau^0[T^\pi] := \begin{cases} \prod_{k \geq 1} C_k^{q_k} & \text{if } T^\pi \text{ is a fat tree,} \\ 0 & \text{otherwise,} \end{cases}$$

where a fat tree is a graph that becomes a tree when the multiplicity of the edges is forgotten, and for  $T$  such a graph we have denoted  $q_k$  the number of edges of  $T$  with multiplicity  $2k$ .

**Theorem 2.2** (Fluctuations of generalized moments). *Let  $A$  be a Wigner matrix with exploding moments. Then, the process  $(Z_N(T^\pi))_{\pi \in \cup_K \mathcal{P}(K)}$  defined by (29) converges to a centered Gaussian process  $(z(T^\pi))_{\pi \in \cup_K \mathcal{P}(K)}$  whose covariance is given by: for any  $T^{\pi_1}, T^{\pi_2}$ ,*

$$\mathbb{E}[z(T^{\pi_1})z(T^{\pi_2})] = \sum_{T \in \mathcal{P}_{\#}(T^{\pi_1}, T^{\pi_2})} \tau^0[T],$$

where  $\tau^0[T]$  is given by Proposition 2.1 and  $\mathcal{P}_{\#}(T^{\pi_1}, T^{\pi_2})$  is the set of graphs obtained by considering disjoint copies of the graphs  $T^{\pi_1}$  and  $T^{\pi_2}$  and gluing them by requiring that they have at least an edge (and therefore two “adjacent” vertices) in common.

*Proof.* We show the convergence of joint moments of  $(Z_N(T^\pi))$ . Gaussian distribution being characterized by its moments, this will prove the theorem. Let  $T_1 = (V_1, E_1), \dots, T_p = (V_n, E_n)$  be finite undirected graphs, each of them being of the form  $T^\pi$  for a partition  $\pi$ . We first write

$$\begin{aligned} \mathbb{E}[Z_N(T_1) \dots Z_N(T_n)] &= \frac{1}{N^{\frac{n}{2}}} \sum_{\substack{\phi_1, \dots, \phi_n \\ \phi_j: V_j \rightarrow [N] \text{ inj.}}} \mathbb{E} \left[ \underbrace{\prod_{j=1}^n \left( \prod_{e \in E_j} A(\phi_j(e)) - \mathbb{E} \left[ \prod_{e \in E_j} A(\phi_j(e)) \right] \right)}_{\omega_N(\phi_1, \dots, \phi_n)} \right] \\ &= \sum_{\sigma \in \mathcal{P}(V_1, \dots, V_n)} \frac{1}{N^{\frac{n}{2}}} \sum_{(\phi_1, \dots, \phi_n) \in S_\sigma} \omega_N(\phi_1, \dots, \phi_n), \end{aligned}$$

where

- $\mathcal{P}(V_1, \dots, V_n)$  is the set of partitions of the disjoint union of  $V_1, \dots, V_n$  whose blocks contain at most one element of each  $V_j$ ,
- $S_\sigma$  is the set of families of injective maps,  $\phi_j : V_j \rightarrow [N], j = 1, \dots, n$ , such that for any  $v \in V_j, v' \in V_{j'}$ , one has  $\phi_j(v) = \phi_{j'}(v') \Leftrightarrow v \sim_\sigma v'$ .

First, it should be noticed that by invariance in law of  $A$  by conjugacy by permutation matrices, for any  $\sigma$  in  $\mathcal{P}(V_1, \dots, V_n)$  and  $(\phi_1, \dots, \phi_n)$  in  $S_\sigma$ , the quantity  $\omega_N(\phi_1, \dots, \phi_n)$  depends only on  $\sigma$ . We then denote  $\omega_N(\sigma) = \omega_N(\phi_1, \dots, \phi_n)$ . Moreover, choosing a partition in  $\mathcal{P}(V_1, \dots, V_n)$  is equivalent to merge certain vertices of different graphs among  $T_1, \dots, T_n$ . We equip  $\mathcal{P}(V_1, \dots, V_n)$  with the edges of  $T_1, \dots, T_n$  and say that two vertices are adjacent if there is an edge between them. We denote by  $\mathcal{P}_{\#}(V_1, \dots, V_n)$  the subset of  $\mathcal{P}(V_1, \dots, V_n)$  such that any graph has two adjacent vertices that are merged to two adjacent vertices of another graph. By the independence of the entries of  $X$  and the centering of the components in  $\omega_N$ , for any  $\sigma$  in  $\mathcal{P}(V_1, \dots, V_n) \setminus \mathcal{P}_{\#}(V_1, \dots, V_n)$  one has

$\omega_N(\sigma) = 0$ . Hence, since the cardinal of  $S_\sigma$  is  $\frac{N!}{(N-|\sigma|)!}$ , we get

$$\begin{aligned}
\mathbb{E}\left[Z_N(T_1) \dots Z_N(T_n)\right] &= \sum_{\sigma \in \mathcal{P}_{\sharp}(V_1, \dots, V_n)} N^{-\frac{n}{2}} \frac{N!}{(N-|\sigma|)!} \omega_N(\sigma) \\
(31) \qquad \qquad \qquad &= \sum_{\sigma \in \mathcal{P}_{\sharp}(V_1, \dots, V_n)} N^{-\frac{n}{2}+|\sigma|} \omega_N(\sigma) (1 + O(N^{-1})).
\end{aligned}$$

Let  $\sigma$  in  $\mathcal{P}_{\sharp}(V_1, \dots, V_n)$ . We now analyze the term  $\omega_N(\sigma)$ . We first expand its product.

$$\omega_N(\sigma) = \sum_{B \subset \{1, \dots, n\}} (-1)^{n-|B|} \mathbb{E}\left[\prod_{j \in B} \prod_{e \in E_j} A(\phi_j(e))\right] \times \prod_{j \notin B} \mathbb{E}\left[\prod_{e \in E_j} A(\phi_j(e))\right].$$

Let  $B \subset \{1, \dots, n\}$ . Denote by  $T_B$  the graph obtained by merging the vertices of  $T_j$ ,  $j \in B$  that belong to a same block of  $\sigma$ . For any  $k \geq 1$  denote by  $p_k$  the number of vertices of  $T_B$  where  $k$  loops are attached. For any  $k \geq \ell \geq 0$ , denote by  $q_{k,\ell}$  the number of pair of vertices that are linked by  $k$  edges in one direction and  $\ell$  edges in the other. Denote by  $\mu_N$  the common law of the entries of  $\sqrt{N}A$ . By independence of the entries of  $A$ , for any  $(\phi_1, \dots, \phi_n)$  in  $S_\sigma$ , one has

$$\begin{aligned}
\mathbb{E}\left[\prod_{j \in B} \prod_{e \in E_j} A(\phi_j(e))\right] &= \prod_{k \geq 1} \left(\frac{\int t^k d\mu_N(t)}{N^{\frac{k}{2}}}\right)^{p_k} \prod_{k, \ell \geq 0} \left(\frac{\int t^{k+\ell} d\mu_N(t)}{N^{\frac{k+\ell}{2}}}\right)^{q_{k,\ell}} \\
(32) \qquad \qquad \qquad &= N^{-|\bar{E}_B|} \underbrace{\prod_{k \geq 1} \left(\frac{\int t^k d\mu_N(t)}{N^{\frac{k}{2}-1}}\right)^{p_k} \prod_{k, \ell \geq 0} \left(\frac{\int t^{k+\ell} d\mu_N(t)}{N^{\frac{k+\ell}{2}-1}}\right)^{q_{k,\ell}}}_{\delta_N(B)},
\end{aligned}$$

where  $|\bar{E}_B|$  is the number of edges of  $T_B$  once the multiplicity and the orientation of edges are forgotten. Recall assumption (3): for any  $k \geq 1$ ,

$$N \mathbb{E}[(a_{i,j})^{2k}] = \mathbb{E}\left[\frac{(\sqrt{N}a_{i,j})^{2k}}{N^{k-1}}\right] = \frac{\int t^{2k} d\mu_N(t)}{N^{k-1}} \xrightarrow{N \rightarrow \infty} C_k.$$

By the Cauchy-Schwarz inequality, for any  $k \geq 1$ ,

$$\frac{\int |t|^{k+1} d\mu_N(t)}{N^{\frac{k+1}{2}-1}} \leq \sqrt{\frac{\int t^{2k} d\mu_N(t)}{N^{k-1}}} \times \sqrt{\int t^2 d\mu_N(t)} = O(1).$$

Hence, since the measure  $\mu_N$  is centered, the quantity  $\delta_N(B)$  is bounded. Denote  $T_\sigma$  the graph obtained by merging the vertices of  $T_1, \dots, T_n$  that belong to a same block of  $\sigma$ ,  $|\bar{E}_\sigma|$  its number of edges when orientation and multiplicity is forgotten, and by  $c_\sigma$  its number

of components. We obtain from (31) and (32)

$$\begin{aligned}
& \mathbb{E} \left[ Z_N(T_1) \dots Z_N(T_n) \right] \\
&= \sum_{\sigma \in \mathcal{P}_{\sharp}(V_1, \dots, V_n)} \sum_{B \subset \{1, \dots, n\}} N^{-\frac{n}{2} + |\sigma| - |\bar{E}_B| - \sum_{j \notin B} |\bar{E}_{\{j\}}|} (-1)^{n-|B|} \delta_N(B) \prod_{j \notin B} \delta_N(\{j\}) (1 + O(N^{-1})) \\
&= \sum_{\sigma \in \mathcal{P}_{\sharp}(V_1, \dots, V_n)} \sum_{B \subset \{1, \dots, n\}} N^{c_\sigma - \frac{n}{2}} \times N^{|\bar{E}_\sigma| - |\bar{E}_B| - \sum_{j \notin B} |\bar{E}_{\{j\}}|} \times N^{|\sigma| - c_\sigma - |\bar{E}_\sigma|} \\
&\quad \times (-1)^{n-|B|} \delta_N(B) \prod_{j \notin B} \delta_N(\{j\}) \times (1 + O(N^{-1}))
\end{aligned}$$

A partition  $\sigma \in \mathcal{P}_{\sharp}(V_1, \dots, V_n)$  induces a partition  $\bar{\sigma}$  of  $\{1, \dots, n\}$ :  $i \sim_{\bar{\sigma}} j$  if and only if  $T_i$  and  $T_j$  belong to a same connected component of  $T_\sigma$ . Denote by  $\mathcal{P}_2(n)$  the set of pair partitions of  $\{1, \dots, n\}$ . One has

$$(33) \quad N^{c_\sigma - \frac{n}{2}} = 1_{\bar{\sigma} \in \mathcal{P}_2(n)} + O(N^{-1}).$$

Secondly, one has  $|\bar{E}_\sigma| - |\bar{E}_B| - \sum_{j \notin B} |\bar{E}_{\{j\}}| \leq 0$  with equality if and only if  $B = \{1, \dots, n\}$ , so that

$$(34) \quad N^{|\bar{E}_\sigma| - |\bar{E}_B| - \sum_{j \notin B} |\bar{E}_{\{j\}}|} = 1_{B = \{1, \dots, n\}} + O(N^{-1}).$$

Moreover, by the relation [24, Lemma 1.1] between the number of vertices, edges and components of a graph,  $|\sigma| - c_\sigma - |\bar{E}_\sigma|$  is the number of cycles of  $\bar{T}_\sigma$ , the graph obtained from  $T_\sigma$  by forgetting the multiplicity and the orientation of its edges. Hence, we have

$$(35) \quad N^{|\sigma| - c_\sigma - |\bar{E}_\sigma|} = 1_{\bar{T}_\sigma \text{ is a forest}} + O(N^{-1}).$$

By (33), (34) and (35), if we denote  $\delta_N(\sigma) = \delta_N(\{1, \dots, n\})$  we get

$$\begin{aligned}
& \mathbb{E} \left[ Z_N(T_1) \dots Z_N(T_n) \right] \\
&= \sum_{\pi \in \mathcal{P}_2(n)} \sum_{\substack{\sigma \in \mathcal{P}_{\sharp}(V_1, \dots, V_n) \\ \text{s.t. } \bar{\sigma} = \pi}} 1_{\bar{T}_\sigma \text{ is a forest of } \frac{n}{2} \text{ trees}} \delta_N(\sigma) + O(N^{-1}) \\
(36) \quad &= \sum_{\pi \in \mathcal{P}_2(n)} \prod_{\{i, j\} \in \pi} \sum_{\sigma \in \mathcal{P}_{\sharp}(V_i, V_j)} 1_{\bar{T}_\sigma \text{ is a tree}} \delta_N(\sigma) + O(N^{-1}),
\end{aligned}$$

where we have used the independence of the entries of  $A$  to split  $\delta_N$ . The case  $n = 2$  gives

$$\mathbb{E} \left[ Z_N(T_1) Z_N(T_2) \right] = \underbrace{\sum_{\sigma \in \mathcal{P}_{\sharp}(V_1, V_2)} 1_{\bar{T}_\sigma \text{ is a tree}} \delta(\sigma)}_{M^{(2)}(T_1, T_2)} + o(1),$$

where  $\delta(\sigma) = \lim_{N \rightarrow \infty} \delta_N(\sigma)$ , which exists since  $\bar{T}_\sigma$  is a tree. Indeed, in the definition (32) of  $\delta(\sigma)$ , we have  $p_k = q_{k, \ell} = 0$  for any  $k \neq \ell$ . Moreover, we obtain that  $\delta(\sigma) = \tau^0[T_\sigma]$  defined

in Proposition 2.1, and the sum over  $\sigma$  on  $\mathcal{P}_{\sharp}(V_1, V_2)$  can be replaced by a sum over graphs  $T$  obtained by identifying certain adjacent vertices of  $T_1$  with adjacent edges  $T_2$ , since  $\tau^0[T]$  is zero if  $T$  is not a fat tree. We then obtain as expected the limiting covariance

$$M^{(2)}(T_1, T_2) = \sum_{T \in \mathcal{P}_{\sharp}(T_1, T_2)} \tau^0[T].$$

The general case  $n \geq 3$  in (36) gives the Wick formula

$$\mathbb{E}\left[Z_N(T_1) \dots Z_N(T_n)\right] = \sum_{\pi \in \mathcal{P}_2(n)} \prod_{\{i,j\} \in \pi} M^{(2)}(T_i, T_j) + o(1),$$

which characterizes the Gaussian distribution.  $\square$

Remark about the recentering with respect to the limit instead of the expectation: in our computations, up to (36) the errors terms are of order  $O(N^{-1})$ , and so if  $N\mathbb{E}[a_{11}^{2k}] = C_k + o(\sqrt{N})$  for any  $k \geq 1$ , then  $\delta_N(\sigma) = \delta(\sigma) + o(\sqrt{N})$  and hence one can replace the  $o(1)$  by  $o(\sqrt{N})$  at the end of the proof.

### 3. CLT FOR STIELTJES TRANSFORM AND THE METHOD OF MARTINGALES

Let  $A = (a_{i,j})_{i,j=1,\dots,N}$  be either a Lévy matrix of parameter  $\alpha \in (0, 2)$ , or a Wigner matrix with exploding moments of parameter  $(C_k)_{k \geq 1}$ . For any  $z$  in  $\mathbb{C} \setminus \mathbb{R}$ , recall that  $G(z) = (z - A)^{-1}$ . To prove Theorem 1.3, we show that any linear combination of the random variables

$$Z_N(z) := \frac{1}{\sqrt{N}} \operatorname{Tr} G(z) - \mathbb{E}\left[\frac{1}{\sqrt{N}} \operatorname{Tr} G(z)\right], \quad z \in \mathbb{C} \setminus \mathbb{R},$$

and their complex transposes converges in distribution to a complex Gaussian variable. Since  $\overline{G(z)} = G(\bar{z})$ , it is enough to fix a linear combination

$$(37) \quad M(N) := \sum_{i=1}^p \lambda_i Z_N(z_i),$$

for some fixed  $p \geq 1$  and  $\lambda_1, \dots, \lambda_p \in \mathbb{C}$ ,  $z_1, \dots, z_p \in \mathbb{C} \setminus \mathbb{R}$ , and prove that  $M(N)$  is asymptotically Gaussian with the adequate covariance.

**3.1. The method of martingales differences and reduction to the case  $p = 1$ .** For all  $N$ , we have

$$M(N) - \mathbb{E}[M(N)] = \sum_{k=1}^N X_k \quad \text{with} \quad X_k := (\mathbb{E}_k - \mathbb{E}_{k-1})[M(N)],$$

where  $\mathbb{E}_k$  denotes the conditional expectation with respect to the  $\sigma$ -algebra generated by the  $k$  first columns of  $A$ . In view of Theorem 6.7 of the appendix about fluctuations of



martingales, it is enough to prove that we have

$$(38) \quad \sum_{k=1}^N \mathbb{E}_{k-1}[X_k^2] \xrightarrow{N \rightarrow \infty} \sum_{i,j=1}^p \lambda_i \lambda_j (L(z_i, z_j) - L(z_i)L(z_j)),$$

$$(39) \quad \sum_{k=1}^N \mathbb{E}_{k-1}[|X_k|^2] \xrightarrow{N \rightarrow \infty} \sum_{i,j=1}^p \lambda_i \bar{\lambda}_j (L(z_i, \bar{z}_j) - L(z_i)L(\bar{z}_j))$$

and that for each  $\varepsilon > 0$ ,

$$(40) \quad \sum_{k=1}^N \mathbb{E}[|X_k|^2 1_{|X_k| \geq \varepsilon}] \xrightarrow{N \rightarrow \infty} 0.$$

Notice first that (38) implies (39). Let us now prove (40). The proof of (38) will then be the main difficulty of the proof of Theorem 1.3.

*Proof of (40).* Let  $A_k$  be the symmetric matrix with size  $N-1$  obtained by removing the  $k$ -th row and the  $k$ -th column of  $A$  and set  $G_k(z) := \frac{1}{z-A_k}$ . Note that  $\mathbb{E}_k G_k(z) = \mathbb{E}_{k-1} G_k(z)$ , so that we can write

$$X_k = \sum_{i=1}^p \lambda_i \times (\mathbb{E}_k - \mathbb{E}_{k-1}) \left[ \frac{1}{\sqrt{N}} (\text{Tr } G(z_i) - \text{Tr } G_k(z_i)) \right].$$

Hence by (82) of Lemma 6.5 in the appendix, there is  $C$  such that for all  $N$  and all  $k$ ,

$$|X_k| \leq \frac{C \max_i |\lambda_i| |\Im z_i|^{-1}}{\sqrt{N}}.$$

Thus for  $N$  large enough, we have that for all  $k$ ,  $|X_k|^2 1_{|X_k| \geq \varepsilon} = 0$  and (40) is proved.  $\square$

We now pass to the main part of the proof of Theorem 1.3, namely the proofs of (38) and (39). It is divided into several steps.

Let us first show that we can get rid of the linear combination in (37) and assume  $p = 1$ . We have

$$X_k = (\mathbb{E}_k - \mathbb{E}_{k-1}) [M(N)] = \sum_{i=1}^p \lambda_i (\mathbb{E}_k - \mathbb{E}_{k-1}) \left[ \frac{1}{\sqrt{N}} \text{Tr } G(z_i) \right],$$

hence, as  $\overline{\text{Tr } G(z_i)} = \text{Tr } G(\bar{z}_i)$ , both  $X_k^2$  and  $|X_k|^2$  are linear combinations of terms of the form

$$(\mathbb{E}_k - \mathbb{E}_{k-1}) \left[ \frac{1}{\sqrt{N}} \text{Tr } G(z) \right] \times (\mathbb{E}_k - \mathbb{E}_{k-1}) \left[ \frac{1}{\sqrt{N}} \text{Tr } G(z') \right],$$

with  $z, z' \in \mathbb{C} \setminus \mathbb{R}$ . As a consequence, we shall only fix  $z, z' \in \mathbb{C} \setminus \mathbb{R}$  and prove that for

$$Y_k := (\mathbb{E}_k - \mathbb{E}_{k-1}) \left[ \frac{1}{\sqrt{N}} \operatorname{Tr} G(z) \right] \quad \text{and} \quad Y'_k := (\mathbb{E}_k - \mathbb{E}_{k-1}) \left[ \frac{1}{\sqrt{N}} \operatorname{Tr} G(z') \right],$$

we have the convergence in probability

$$(41) \quad C_N := \sum_{k=1}^N \mathbb{E}_{k-1} [Y_k Y'_k] \xrightarrow[N \rightarrow \infty]{} L(z, z') - L(z)L(z').$$

First, for  $G_k$  as introduced in the proof of (40) above, as  $\mathbb{E}_k G_k(z) = \mathbb{E}_{k-1} G_k(z)$  again, we have  $Y_k = (\mathbb{E}_k - \mathbb{E}_{k-1}) \left[ \frac{1}{\sqrt{N}} (\operatorname{Tr} G(z) - \operatorname{Tr} G_k(z)) \right]$ . Hence by Lemma 6.5, we can write

$$(42) \quad Y_k = \frac{1}{\sqrt{N}} (\mathbb{E}_k - \mathbb{E}_{k-1}) \frac{1 + \mathbf{a}_k^* G_k(z)^2 \mathbf{a}_k}{z - a_{kk} - \mathbf{a}_k^* G_k(z) \mathbf{a}_k},$$

where  $\mathbf{a}_k$  is the  $k$ -th column of  $A$  where the  $k$ -th entry has been removed. To prove (41), we shall first show that we can get rid of the off diagonal terms  $\sum_{j \neq \ell} \mathbf{a}_k(j) \mathbf{a}_k(\ell) G_k(z)_{j\ell}$  and  $\sum_{j \neq \ell} \mathbf{a}_k(j) \mathbf{a}_k(\ell) (G_k(z)^2)_{j\ell}$  in the above expression.

**3.2. Removing the off-diagonal terms.** In this section, we prove that we can replace  $Y_k$  in (41) by

$$(43) \quad \tilde{Y}_k := \frac{1}{\sqrt{N}} (\mathbb{E}_k - \mathbb{E}_{k-1}) \frac{1 + \sum_{j=1}^{N-1} \mathbf{a}_k(j)^2 (G_k(z)^2)_{jj}}{z - \sum_{j=1}^{N-1} \mathbf{a}_k(j)^2 G_k(z)_{jj}}.$$

Note first that by Equation (85) of Lemma 6.6 in the appendix, we have, with probability one, for all  $z \in \mathbb{C} \setminus \mathbb{R}$ ,

$$\left| \frac{1 + \sum_j \mathbf{a}_k(j)^2 (G_k(z)^2)_{jj}}{z - \sum_j \mathbf{a}_k(j)^2 G_k(z)_{jj}} \right| \leq 2|\Im z|^{-1}$$

hence

$$(44) \quad |\tilde{Y}_k| \leq \frac{4|\Im z|^{-1}}{\sqrt{N}}.$$

**Lemma 3.1.** *Let  $\tilde{Y}'_k$  be defined in the same way as  $\tilde{Y}_k$  in (43), replacing  $z$  by  $z'$ . Set also*

$$(45) \quad \tilde{C}_N := \sum_{k=1}^N \mathbb{E}_{k-1} [\tilde{Y}_k \tilde{Y}'_k].$$

*Then for  $C_N$  defined as in (41), as  $N \rightarrow \infty$ , we have the convergence*

$$(46) \quad C_N - \tilde{C}_N \xrightarrow{L^1} 0.$$

*Proof.* First note that for both Lévy and exploding moments matrices, for any  $(N - 1) \times (N - 1)$  matrix  $B$  with spectral radius  $\|B\|$ ,

$$\mathbf{a}_k^* B \mathbf{a}_k = \sum_j \mathbf{a}_k(j)^2 (B)_{jj} + \varepsilon_N(B)$$

where there exists  $\delta > 0$  and a finite constant  $C$  so that for all  $t > 0$ , all matrix  $B$ ,

$$(47) \quad \mathbb{P}(|\varepsilon_N(B)| \geq t) \leq \frac{C\|B\|^2}{t^2 N^{2\delta}} + \frac{C}{N^{2\delta}}.$$

By definition,  $\varepsilon_N(B) = \sum_{\ell \neq j} \mathbf{a}_k(\ell) \mathbf{a}_k(j) B_{\ell j}$ . In the case of exploding moments, we have

$$\begin{aligned} \mathbb{E}\left[|\varepsilon_N(B)|^2\right] &= \sum_{i \neq j} \sum_{i' \neq j'} \mathbb{E}[\mathbf{a}_k(i) \mathbf{a}_k(j) b_{ij} \mathbf{a}_k(i') \mathbf{a}_k(j') \bar{b}_{i'j'}] \\ &\leq 2 \sum_{i \neq j} \mathbb{E}[\mathbf{a}_k(i)^2] \mathbb{E}[\mathbf{a}_k(j)^2] |b_{ij}|^2 \\ &\leq \frac{2}{N} (N \mathbb{E}[\mathbf{a}_k(1)^2])^2 \|B\|^2, \end{aligned}$$

from which (47) follows by Tchebychev's inequality with  $\delta = 1/2$ . For Lévy matrices, we put for  $\kappa \in (0, \frac{1}{2(2-\alpha)})$

$$\varepsilon_N(B) = \sum_{j \neq \ell} \mathbf{a}_k(\ell) \mathbf{a}_k(j) B_{\ell j} = (\mathbf{1}_{\max_j |\mathbf{a}_k(j)| \leq N^\kappa} + \mathbf{1}_{\max_j |\mathbf{a}_k(j)| > N^\kappa}) \varepsilon_N(B).$$

By [10, Lemma 4.3], and the same computation as above, for any  $t > 0$  we have

$$\mathbb{P}\left(\{|\varepsilon_N| \geq t\} \cap \{\max_j |a_{kj}| \leq N^\kappa\}\right) \leq \frac{2}{t^2 N^{1-2\kappa(2-\alpha)} \Im z^2}$$

whereas there exists a finite constant  $C$  so that

$$\mathbb{P}\left(\max_j |a_{kj}| \geq N^\kappa\right) \leq \frac{C}{N^{\alpha\kappa}},$$

which prove (47) in this setting with  $2\delta = \min\{\alpha\kappa, 1 - 2\kappa(2 - \alpha)\}$ .

We apply (47) with  $B = G_k(z)$  and  $B = G_k(z)^2$ . Since the function of two complex variables  $\varphi(x_1, x_2) := \frac{1+x_1}{z-x_2}$  has a uniformly bounded gradient on the set  $\{(x_1, x_2) \in \mathbb{C}^2; \Im x_2 \leq 0\}$ , it follows that

$$(48) \quad \frac{1 + \mathbf{a}_k^* G_k(z)^2 \mathbf{a}_k}{z - a_{kk} - \mathbf{a}_k^* G_k(z) \mathbf{a}_k} = \frac{1 + \sum_j \mathbf{a}_k(j)^2 (G_k(z)^2)_{jj}}{z - \sum_j \mathbf{a}_k(j)^2 G_k(z)_{jj}} + \varepsilon'_N$$

where  $|\varepsilon'_N| \leq C' [|\varepsilon_N(G_k(z))| + |\varepsilon_N(G_k(z)^2)|]$  for some finite constant  $C'$ . Now, we can use (81), (82) and (85) to assert that the two first terms, in (48), are bounded by a constant

uniformly in  $N$  and  $k$ . Hence,  $\varepsilon'_N$  is uniformly bounded. It follows that there exists a finite constant  $C'''$  so that for any  $t > 0$ ,

$$\mathbb{E}[|\varepsilon'_N|^2] \leq \frac{C'''}{N^{2\delta}} + C'''t^2 + 4C''' \int_t^{C'''} \frac{s}{s^2 N^{2\delta} \Im z^2} ds.$$

Taking  $t = N^{-\delta}$  we conclude

$$\mathbb{E}[|\varepsilon'_N|^2] \leq \frac{C''' \log N}{N^{2\delta}}$$

To sum up, we have

$$(49) \quad Y_k = \tilde{Y}_k + \gamma_k \quad Y'_k = \tilde{Y}'_k + \gamma'_k,$$

where the error terms  $\gamma_k, \gamma'_k$  are  $O(N^{-\frac{1}{2}-\delta} \times \log N)$  uniformly in  $k$  for the  $L^2$  norm.

But

$$\Delta_k := Y_k Y'_k - \tilde{Y}_k \tilde{Y}'_k = \tilde{Y}_k \gamma'_k + \tilde{Y}'_k \gamma_k + \gamma_k \gamma'_k,$$

so to conclude, it is enough to prove that the RHS is  $o(N^{-1})$  in  $L^1$ , uniformly in  $k$ . We have

$$\mathbb{E}[|\tilde{Y}_k \gamma'_k + \tilde{Y}'_k \gamma_k + \gamma_k \gamma'_k|] \leq \|\tilde{Y}_k\|_{L^2} \|\gamma'_k\|_{L^2} + \|\tilde{Y}'_k\|_{L^2} \|\gamma_k\|_{L^2} + \|\gamma_k\|_{L^2} \|\gamma'_k\|_{L^2},$$

with  $\tilde{Y}_k = O_{L^\infty}(N^{-1/2})$  (and the same for  $\tilde{Y}'_k$ ) by (44) so that (49) implies that

$$\mathbb{E}[|\Delta_k|] \leq C''' \log N N^{-1-\frac{\delta}{2}}$$

which proves the lemma.  $\square$

It remains to prove that the sequence  $\tilde{C}_N$  introduced at (45) converges in probability as  $N$  goes to infinity. Note that we have

$$(50) \quad \tilde{C}_N = \int_{u=0}^1 N \mathbb{E}_{k-1}[\tilde{Y}_{k(u)} \tilde{Y}'_{k(u)}] du \quad \text{with } k(u) := \lceil Nu \rceil.$$

By (44), the integrand is uniformly bounded by  $4|\Im z|^{-1}$ . Hence, by dominated convergence, it is enough to prove that for any fixed  $u \in (0, 1)$ , as  $k, N$  tend to infinity in such a way that  $k/N \rightarrow u$ , we have the convergence in probability of

$$(51) \quad N \mathbb{E}_{k-1}[\tilde{Y}_k \tilde{Y}'_k].$$

Now, set

$$(52) \quad f_k := \frac{1 + \sum_j \mathbf{a}_k(j)^2 (G_k(z))^2_{jj}}{z - \sum_j \mathbf{a}_k(j)^2 G_k(z)_{jj}} \quad f'_k := \frac{1 + \sum_j \mathbf{a}_k(j)^2 (G_k(z'))^2_{jj}}{z - \sum_{i=1}^N \mathbf{a}_k(j)^2 G_k(z')_{jj}}.$$

We have  $\sqrt{N} \tilde{Y}_k = (\mathbb{E}_k - \mathbb{E}_{k-1}) f_k$  and the same for  $\tilde{Y}'_k$ , so

$$N \mathbb{E}_{k-1}(\tilde{Y}_k \tilde{Y}'_k) = \mathbb{E}_{k-1}[\mathbb{E}_k f_k \mathbb{E}_k f'_k] - \mathbb{E}_{k-1} f_k \mathbb{E}_{k-1} f'_k.$$

Let us denote by  $\mathbb{E}_{\mathbf{a}_k}$  the expectation with respect to the randomness of the  $k$ -th column of  $A$  (*i.e.* the conditional expectation with respect to the  $\sigma$ -algebra generated by the  $a_{ij}$ 's such that  $k \notin \{i, j\}$ ). Note that  $\mathbb{E}_{k-1} = \mathbb{E}_{\mathbf{a}_k} \circ \mathbb{E}_k = \mathbb{E}_k \circ \mathbb{E}_{\mathbf{a}_k}$ , hence

$$N\mathbb{E}_{k-1}(\tilde{Y}_k \tilde{Y}'_k) = \mathbb{E}_{\mathbf{a}_k}[\mathbb{E}_k f_k \mathbb{E}_k f'_k] - \mathbb{E}_k \mathbb{E}_{\mathbf{a}_k} f_k \times \mathbb{E}_k \mathbb{E}_{\mathbf{a}_k} f'_k.$$

Now, note that if one introduces (on a possibly enlarged probability space where the conditional expectations  $\mathbb{E}_{k-1}, \mathbb{E}_k, \mathbb{E}_{\mathbf{a}_k}$  keep their definitions as the conditional expectations with respect to the same  $\sigma$ -algebras as above) an  $N \times N$  random symmetric matrix

$$(53) \quad A' = [a'_{ij}]_{1 \leq i, j \leq N}$$

such that:

- the  $a'_{ij}$ 's such that  $i > k$  and  $j > k$  are i.i.d. copies of  $a_{11}$  (modulo the fact that  $A'$  is symmetric), independent of  $A$ ,
- for all other pairs  $(i, j)$ ,  $a'_{ij} = a_{ij}$ ,

then we have

$$\mathbb{E}_k f_k \mathbb{E}_k f'_k = \mathbb{E}_k(f_k \times f''_k),$$

where  $f''_k$  is defined out of  $A'$  in the same way as  $f'_k$  is defined out of  $A$  in (52) (note that the  $k$ -th column is the same in  $A$  and in  $A'$ ). It follows that

$$(54) \quad \begin{aligned} N\mathbb{E}_{k-1}(\tilde{Y}_k \tilde{Y}'_k) &= \mathbb{E}_{\mathbf{a}_k}[\mathbb{E}_k(f_k \times f''_k)] - \mathbb{E}_k \mathbb{E}_{\mathbf{a}_k} f_k \times \mathbb{E}_k \mathbb{E}_{\mathbf{a}_k} f'_k \\ &= \mathbb{E}_k[\mathbb{E}_{\mathbf{a}_k}(f_k \times f''_k)] - \mathbb{E}_k \mathbb{E}_{\mathbf{a}_k} f_k \times \mathbb{E}_k \mathbb{E}_{\mathbf{a}_k} f'_k. \end{aligned}$$

We shall in the sequel prove that as  $N$  tends to infinity, regardless to the value of  $k$ , we have the almost sure convergences

$$(55) \quad \mathbb{E}_{\mathbf{a}_k} f_k \longrightarrow L(z) \quad \mathbb{E}_{\mathbf{a}_k} f'_k \longrightarrow L(z')$$

and that for any fixed  $u \in (0, 1)$ , as  $N, k$  tend to infinity in such a way that  $k/N \longrightarrow u$ , we have the almost sure convergence

$$(56) \quad \mathbb{E}_{\mathbf{a}_k}(f_k \times f''_k) \longrightarrow \Psi^u(z, z').$$

The convergences of (55) and (56) are based on an abstract convergence result stated in next section, where we use the convergence of generalized moments of Proposition 2.1. They are stated in Lemmas 3.4 and 3.5 respectively.

Note that by Lemma 6.6,

$$(57) \quad |f_k| \leq 4|\Im z|^{-1},$$

so once (55) and (56) proved, the convergences will also hold in  $L^2$ , hence by continuity of the conditional expectations, we will have the convergence in  $L^2$

$$\mathbb{E}_k[\mathbb{E}_{\mathbf{a}_k}(f_{k(u)} \times f''_{k(u)})] - \mathbb{E}_k \mathbb{E}_{\mathbf{a}_k} f_k \times \mathbb{E}_k \mathbb{E}_{\mathbf{a}_k} f'_k \longrightarrow \Psi^u(z, z') - L(z)L(z').$$

Thus by (54), we will have proved the convergence of (51), hence concluded the proof of the theorem.

**3.3. An abstract convergence result.** Remember that  $A_k = (a_{ij})$  is the square matrix of size  $N - 1$  obtained by removed the  $k$ -th row and the  $k$ -th column of  $A$ , that  $G_k(z) = (z - A_k)^{-1}$  and that  $A'_k$  is a copy of  $A_k$  where the entries  $(i, j)$  for  $i, j > k$  are independent of  $A_k$  and the other are those of  $A_k$ . We denote  $G'_k(z') = (z' - A'_k)^{-1}$ .

**Lemma 3.2.** *As  $N$  goes to infinity and  $\frac{k}{N}$  tends to  $u$  in  $(0, 1)$ , the random probability measure on  $\mathbb{C}^2$*

$$(58) \quad \nu_{z, z'}^{k, N} := \frac{1}{N-1} \sum_{j=1}^{N-1} \delta_{\{G_k(z)_{jj}, G'_k(z')_{jj}\}}, \quad z, z' \in \mathbb{C} \setminus \mathbb{R}.$$

*converges weakly almost surely to a non random probability measure  $\nu_{z, z'}^u$  on  $\mathbb{C}^2$ .*

*Proof.* By e.g. Theorem C.8 of [1], it is enough to prove that for any bounded and Lipschitz function  $f$ ,  $\nu_{z, z'}^{k, N}(f)$  converges almost surely to  $\nu_{z, z'}^u(f)$ . Moreover, adapting the proof of Lemma 6.4, one can easily see that for any such  $f$ ,

$$\nu_{z, z'}^{k, N}(f) - \mathbb{E}[\nu_{z, z'}^{k, N}(f)]$$

converges almost surely to zero. The only modification of the proof is to complete the resolvent identity by remarking that

$$(z - A_k)^{-2} - (z - A_k^B)^{-2} = (z - A_k^B)^{-2} (z(A_k - A_k^B) - (A_k - A_k^B)^2 - 2A_k^B(A_k - A_k^B)) (z - A_k)^{-2},$$

which gives that this matrix has rank bounded by  $3 \times \text{rank}(A - B)$ . Hence it is enough to prove that the deterministic sequence  $\mathbb{E}[\nu_{z, z'}^{k, N}(\cdot)]$  converges weakly. These measures are supported by the compact subset  $B(0, |\Im z|^{-1})^2$ , hence the sequence is tight and it suffices to prove that it admits at most one accumulation point.

Let us show how the Lévy case can be deduced from the case of Wigner matrices with exploding moments. For any  $B > 0$ , let us introduce the truncated matrices

$$A^B := (a_{i,j} 1_{|a_{i,j}| \leq B})_{i,j=1, \dots, N} \quad A'^B := (a'_{i,j} 1_{|a'_{i,j}| \leq B})_{i,j=1, \dots, N}.$$

These matrices are Wigner matrices with exploding moments with rescaled moments that satisfy (5) (see [31, Section 1.2.1] or [10, Section 9]). Introduce the random probability measure  $\nu_{z, z'}^{k, N, B}$ , defined as  $\nu_{z, z'}^{k, N}$ , except than  $A$  and  $A'$  are replaced by  $A^B$  and  $A'^B$ , and similarly the matrices  $G_{k, B}(z)$  and  $G'_{k, B}(z')$ . By Lemma 2.4 of [10], we know that for any  $\varepsilon > 0$ , there exist  $B(\varepsilon) > 0$  and, for any  $B > B(\varepsilon)$ ,  $\delta(\varepsilon) > 0$  such that with probability at least  $1 - e^{-\delta(\varepsilon, B)N}$ ,

$$\text{rank}(A_k - A_k^B) + \text{rank}(A'_k - A'^B_k) \leq \varepsilon(N - 1).$$

Moreover, we can adapt arguments of [10] (see also Equation (91) of [12]) to get that with probability at least  $1 - e^{-\delta(\varepsilon, B)N}$ , for any 1-Lipschitz function  $f$  on  $B(0, |\Im z|^{-1})^2$ , we have

$$|\nu_{z, z'}^{k, N}(f) - \nu_{z, z'}^{k, N, B}(f)| \leq 2(|\Im z|^{-1} + |\Im z'|^{-1} + 3|\Im z|^{-1} + 3|\Im z'|^{-1})\varepsilon.$$

Hence it suffices to prove the weak convergence of  $\mathbb{E}[\nu_{z,z'}^{k,N}(\cdot)]$  when  $A$  is a Wigner matrix with exploding moments. Since these measures are uniformly compactly supported, it is sufficient to prove the convergence in moments.

We consider a polynomial  $P = x_1^{n_1} \overline{x_1}^{m_1} x_2^{n_2} \overline{x_2}^{m_2}$  and remark that

$$\begin{aligned} \mathbb{E}[\nu_{z,z'}^{k,N}(P)] &= \mathbb{E}\left[\frac{1}{N-1} \sum_{j=1}^{N-1} (G_k(z)_{jj})^{n_1} (G_k(\bar{z})_{jj})^{m_1} (G'_k(z')_{jj})^{n_2} (G'_k(\bar{z}')_{jj})^{m_2}\right] \\ &= \mathbb{E}\left[\frac{1}{N-1} \text{Tr} \left[ G_k(z)^{\circ n_1} \circ G_k(\bar{z})^{\circ m_1} \circ G'_k(z')^{\circ n_2} \circ G'_k(\bar{z}')^{\circ m_2} \right]\right], \end{aligned}$$

where  $\circ$  denotes the Hadamard (entry-wise) product of matrices and

$$M^{\circ n} := \underbrace{M \circ \dots \circ M}_n.$$

We set  $\ell_1 = n_1 + m_1$  and  $\ell_2 = n_2 + m_2$ . Let  $(Y_i, Y'_j)_{i=1,\dots,\ell_1, j=1,\dots,\ell_2}$  be a family of random variables such that for any  $p_i, q_j \geq 0$ ,

$$\begin{aligned} (59) \quad & \mathbb{E}\left[\prod_{i=1}^{\ell_1} Y_i^{p_i} \prod_{j=1}^{\ell_2} Y'_j{}^{q_j}\right] \\ &= \mathbb{E}\left[\frac{1}{N-1} \text{Tr} \left[ A_k^{p_1} \circ \dots \circ A_k^{p_{\ell_1}} \circ A'_k{}^{q_1} \circ \dots \circ A'_k{}^{q_{\ell_2}} \right]\right]. \end{aligned}$$

Such a family exists by Proposition 6.8. By [31, Proposition 3.10], the couple of random matrices  $(A, A')$  satisfies the so-called convergence in distribution of traffics, which implies the convergence in moments of  $(Y_i, Y'_j)_{i,j}$ . For reader's convenience, we give the limiting value of (59), even if we do not use it later. It is obtained by applying the rule of the so-called traffic freeness in [31].

$$(60) \quad \mathbb{E}\left[\prod_{i=1}^{\ell_1} Y_i^{p_i} \prod_{j=1}^{\ell_2} Y'_j{}^{q_j}\right] \xrightarrow{N \rightarrow \infty} \sum_{\pi \in \mathcal{P}(V_T)} \tau^0[T^\pi] \alpha(T^\pi),$$

where

1. we have considered  $T$  the graph whose edges are labelled by indeterminates  $a$  and  $a'$ , obtained by
  - considering the disjoint union of the graphs with vertices  $1, \dots, p_i$  and edges  $\{1, 2\}, \dots, \{p_i - 1, p_i\}, \{p_i, 1\}$  labelled  $a$ ,  $i = 1, \dots, \ell_1$ ,
  - considering the disjoint union of the graphs with vertices  $1, \dots, q_j$  and edges  $\{1, 2\}, \dots, \{q_j - 1, q_j\}, \{q_j, 1\}$  labelled  $a'$ ,  $j = 1, \dots, \ell_2$ ,
  - identifying the vertex 1 of each of these graphs (we get a connected graph, bouquet of cycles),
2. we have denoted by  $V_T$  is the set of vertices of  $T$ ,  $\mathcal{P}(V_T)$  is the set of partitions of  $V_T$  and  $T^\pi$  denotes the graph obtained by identifying the vertices of  $T$  that belong to a same block of  $\pi$ ,

3. the quantity  $\tau^0[T^\pi]$  is as in Proposition 2.1,
4. we have set

$$\alpha(T^\pi) = \sum_{V_\pi = V_1 \sqcup V_2} 1_{\left\{ \begin{array}{l} \text{the edges linking adjacent vertices} \\ \text{of } V_2 \text{ have the same label} \end{array} \right\}} u^{|V_1|} (1-u)^{|V_2|},$$

where the sum is over all partitions of the set  $V_\pi$  of vertices of  $T^\pi$  and  $|E_i|$  is the number of edges between adjacent vertices of  $V_i$ ,  $i = 1, 2$ .

Formula (60) could also be derived by the same techniques than those developed in Section 2. The random variables  $Y_i$  and  $Y'_j$  are distributed according to the limiting eigenvalues distribution of  $A$ . Moreover, following the proof of [38, Proposition 10], the exponential power series of this measure has a positive radius of convergence. So, by a generalization of [11, Theorem 30.1] to the multi-dimensional case, we get that the distribution of  $(Y_i, Y'_j)$  is characterized by its moments. Then, we get that  $(Y_i, Y'_j)$  converges weakly to a family of random variables  $(y_i, y'_j)$ . We set  $f_z(y) = (z - y)^{-1}$ . We then obtain the convergence

$$\mathbb{E}[\nu_{z, z'}^{k, N}(P)] = \mathbb{E}\left[\prod_{i=1}^{n_1} f_z(Y_i) \prod_{j=1}^{n_2} f_{z'}(Y'_j)\right] \xrightarrow{N \rightarrow \infty} \mathbb{E}\left[\prod_{i=1}^{n_1} f_z(y_i) \prod_{j=1}^{n_2} f_{z'}(y'_j)\right]$$

Hence  $\mathbb{E}[\nu_{z, z'}^{k, N}(\cdot)]$  converges weakly.  $\square$

This convergence could also be proven without Proposition 6.8 but with appropriate bounds on the growth of moments, which however we found less elegant.

We have the following Corollary.

**Corollary 3.3.** (1) For  $z \in \mathbb{C} \setminus \mathbb{R}$ ,  $t$  so that  $\Im z t \geq 0$ ,

$$\rho_z^N(t) := \frac{1}{N-1} \sum_{i=1}^{N-1} \Phi(tG_k(z)_{ii}) \xrightarrow{N \rightarrow \infty} \rho_z(t) := \int \Phi(tx) \nu_{z, z'}^0(dx, dx') \quad a.s.$$

(2) For  $z, z' \in \mathbb{C} \setminus \mathbb{R}$ ,  $t$  so that  $\Im z/t \geq 0$ ,  $\Im z'/t' \geq 0$ , and  $k/N$  going to  $u$ ,

$$\rho_{z, z'}^{N, k}(t, t') := \frac{1}{N-1} \sum_{i=1}^{N-1} \Phi(tG_k(z)_{ii} + t'G'_k(z')_{ii})$$

converges almost surely towards

$$(61) \quad \rho_{z, z'}^u(t, t') = \int \Phi(tx + t'x') \nu_{z, z'}^u(dx, dx').$$

(3) The functions  $\rho^{N, k}$  are analytic on  $\Lambda = \{t/\Im z > 0, t'/\Im z' > 0\}$  and uniformly bounded on compact subsets of  $\bar{\Lambda}$ . They have a non positive real part. Their limits are also analytic on  $\Lambda$  and have a non positive real part.



(4) For all  $z \in \mathbb{C}^+$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \operatorname{Tr}((z - A)^{-1}) = i \int_0^\infty e^{itz + \rho_z(t)} dt \quad a.s.$$

*Proof.* Let us first notice that  $\rho_z^N(t) = \rho_{z,z'}^{N,k}(t, 0)$  for all  $k$  so we only need to focus on  $\rho_{z,z'}^{N,k}(t, t')$ . The point wise convergence of the function  $\rho^{N,k}$  is a direct consequence of the continuity of  $\Phi$  (recall we assumed that  $\Phi$  extends continuously to the real line), of the boundedness of  $tG_k(z)_{ii} + t'G'_k(z')_{ii}$  (by  $t/\Im z + t'/\Im z'$ ) and Lemma 3.2. To show analyticity, note that for all  $j \in \{1, \dots, N\}$ ,  $G(z)_{jj}$  is an analytic function on  $\mathbb{C}^+$ , taking its values in  $\mathbb{C}^-$  (and vice versa) almost surely by (83). Hence, on  $\Lambda$ ,  $tG_k(z)_{ii} + t'G'_k(z')_{ii}$  is an analytic function with values in  $\mathbb{C}^-$ . Therefore, as  $\Phi$  is analytic on  $\mathbb{C}^-$ ,  $\rho^{N,k}$  is an analytic function on  $\Lambda$  almost surely. Moreover, as  $\Phi$  extends continuously to the real line, it is uniformly bounded on compact subsets of  $\overline{\mathbb{C}^-}$  and hence  $\rho^{N,k}$  is uniformly bounded on compact subsets of  $\Lambda$ . This implies by Montel's theorem that the limit  $\rho^u$  of  $\rho^{N,k}$  is also analytic. Finally,  $\rho^{N,k}$  has non positive real part as the image of  $\mathbb{C}^-$  by  $\Phi$ . Indeed, as  $\Re(-i\lambda|a_{11}|^2) \leq 0$ , we have

$$(62) \quad |\phi_N(\lambda)| = |\mathbb{E}(e^{-i\lambda|a_{11}|^2})| \leq 1 \quad \Rightarrow \quad \Re\Phi(\lambda) = \lim_{N \rightarrow \infty} N \log |\phi_N(\lambda)| \leq 0.$$

For the last point, first note that by Lemma 6.3, it is enough to prove the result for  $\mathbb{E}[N^{-1} \operatorname{Tr} G(z)]$  instead of  $N^{-1} \operatorname{Tr} G(z)$ . Second, by exchangeability,  $\mathbb{E}[N^{-1} \operatorname{Tr} G(z)] = \mathbb{E}[G(z)_{11}]$ . Remind that  $\mathbf{a}_k$  is the  $k$ -th column of  $A$  (or  $A'$ ) where the  $k$ -th entry has been removed. Using Schur complement formula and getting rid of the off diagonal terms (by arguments similar to those of Section 3.2), we have for any  $z \in \mathbb{C}^+$

$$\begin{aligned} \mathbb{E}[G(z)_{11}] &= \mathbb{E}\left[\frac{1}{z - \sum_{j=1}^{N-1} |\mathbf{a}_1(j)|^2 G_1(z)_{jj} + \varepsilon_N}\right] \\ &= \mathbb{E}\left[\frac{1}{z - \sum_{j=1}^{N-1} \mathbf{a}_1(j)^2 G_1(z)_{jj}}\right] + o(1). \end{aligned}$$

Remember that we have for  $\lambda \in \mathbb{C} \setminus \mathbb{R}$ ,

$$(63) \quad \frac{1}{\lambda} = -i \int_0^{\operatorname{sgn} \lambda \infty} e^{it\lambda} dt,$$

and that  $\phi_N(\lambda) = \mathbb{E}[\exp(-i\lambda|a_{11}|^2)]$ , which gives

$$\begin{aligned} \mathbb{E}[G(z)_{11}] &= i \int_0^\infty \mathbb{E}[e^{i\lambda(z - \sum_{j=1}^{N-1} \mathbf{a}_1(j)^2 G_1(z)_{jj})}] d\lambda + o(1) \\ &= i \int_0^\infty e^{i\lambda z} \mathbb{E}\left[\prod_{j=1}^{N-1} \phi_N(\lambda G_1(z)_{jj})\right] d\lambda + o(1). \end{aligned}$$

We used the exponential decay to switch the integral and the expectation. This also allows to truncate the integral: for any  $M \geq 0$ ,

$$\int_0^\infty e^{i\lambda z} \mathbb{E} \left[ \prod_{j=1}^{N-1} \phi_N(\lambda G_1(z)_{jj}) \right] d\lambda = \int_0^M e^{i\lambda z} \mathbb{E} \left[ \prod_{j=1}^{N-1} \phi_N(\lambda G_1(z)_{jj}) \right] d\lambda + \epsilon_{m,M,z,N},$$

where  $\epsilon_{M,z,N}$  goes to zero as  $M$  tends to infinity, uniformly on  $N$  and on the randomness. Remember that by assumption (8), we have

$$(64) \quad N(\phi_N(\lambda) - 1) \xrightarrow[N \rightarrow \infty]{} \Phi(\lambda), \quad \forall \lambda \in \overline{\mathbb{C}^-}$$

where the convergence is uniform on compact subsets of  $\overline{\mathbb{C}^-}$ . Hence, since for  $|\Im z| \geq \delta > 0$ ,  $|t| \leq M$ ,  $tG_k(z)_{jj}$  belongs to the compact set  $\{\lambda \in \overline{\mathbb{C}^-}; |\lambda| \leq M/\delta\}$ , we have

$$\prod_{j=1}^{N-1} \phi_N(tG_k(z)_{jj}) = \exp \left( \frac{1}{N} \sum_{j=1}^{N-1} \Phi(tG_k(z)_{jj}) \right) + \varepsilon_{t,z,N}^{(1)} = \exp(\rho_z^N(t)) + \varepsilon_{t,z,N}^{(1)}$$

where  $\varepsilon_{t,z,N}^{(1)}$  converges almost surely to zero as  $N$  goes to infinity. By the above,

$$\exp \left( \frac{1}{N} \sum_{j=1}^{N-1} \Phi(tG_k(z)_{jj}) \right) = \exp(\rho_z(t)) + \varepsilon_{t,z,N}^{(2)},$$

where  $\varepsilon_{t,z,N}^{(2)}$  converges to zero almost surely. Hence, we deduce the almost sure convergence

$$(65) \quad e^{itz} \prod_{j=1}^{N-1} \phi_N(tG_k(z)_{jj}) \xrightarrow[N \rightarrow \infty]{} e^{itz + \rho_z(t)}.$$

As  $\rho_z^N$  has non positive real part, we can conclude by dominated convergence theorem and by getting rid of the truncation of the integral.  $\square$

### 3.4. Computation and convergence of $\mathbb{E}_{\mathbf{a}_k} f_k$ and $\mathbb{E}_{\mathbf{a}_k} f'_k$ .

**Lemma 3.4.** *Almost surely, we have the convergence*

$$(66) \quad \mathbb{E}_{\mathbf{a}_k} f_k \xrightarrow[N \rightarrow \infty]{} L(z) := - \int_0^{\text{sgn}_z \infty} \frac{1}{t} \partial_z e^{itz + \rho_z(t)} dt.$$

where  $\text{sgn}_z := \text{sgn}(\Im z)$  and  $\rho_z(t)$  is defined in Corollary 3.3.

*Proof.* Remind that for any  $z \in \mathbb{C} \setminus \mathbb{R}$ ,

$$f_k = \frac{1 + \sum_j \mathbf{a}_k(j)^2 (G_k(z)^2)_{jj}}{z - \sum_j \mathbf{a}_k(j)^2 G_k(z)_{jj}} := \frac{\partial_z \lambda_N(z)}{\lambda_N(z)}.$$

Hence, by (63), since by (83) the sign of the imaginary part of  $\lambda_N(z)$  is  $\text{sgn}_z$ , the random variable  $f_k$  can be written

$$\begin{aligned}
(67) \quad f_k &= -i \int_0^{\text{sgn}_z \infty} \partial_z \lambda_N(z) e^{it\lambda_N(z)} dt. \\
&= -i \int_{\text{sgn}_z m}^{\text{sgn}_z \infty} \partial_z \lambda_N(z) e^{it\lambda_N(z)} dt - i \int_0^{\text{sgn}_z m} \partial_z \lambda_N(z) e^{it\lambda_N(z)} dt \\
&:= \tilde{f}_{k,m} + \eta_m(z),
\end{aligned}$$

where  $m > 0$  and  $\eta_m(z) = \frac{\partial_z \lambda_N(z)}{\lambda_N(z)} (1 - e^{i \text{sgn}_z m \lambda_N(z)})$ . We next show that for all  $\varepsilon > 0$  there exists  $m_0$  small enough so that for all  $m < m_0$ , all  $N$  large enough

$$(68) \quad \mathbb{E}_{\mathbf{a}_k} [|\eta_m(z)|] \leq \varepsilon$$

By (85) and since the sign of the imaginary part of  $\lambda_N(z)$  is  $\text{sgn}_z$ , one has  $|\eta_m(z)| \leq 4|\Im z|^{-1}$ . More precisely, for any  $K > 0$ , we find

$$|\eta_m(z)| \leq \frac{4}{\Im z} (mK |\Im z|^{-1} + 1_{\sum a_k(j)^2 \geq K}).$$

In the exploding moment case, we note that

$$\mathbb{P} \left( \sum a_k(j)^2 \geq K \right) \leq \frac{1}{K} \mathbb{E}[Na_k(1)^2] \leq \frac{C}{K}$$

whereas in the Lévy matrices case, since  $\sum a_k(j)^2$  converges weakly towards a  $\alpha/2$  stable law, we know that for  $N$  large enough

$$\mathbb{P} \left( \sum a_k(j)^2 \geq K \right) \leq \frac{C}{K^{\alpha/2}}.$$

Choosing  $K = 1/\sqrt{m}$  or  $K = m^{-\frac{2}{2+\alpha}}$  we conclude that there exists a  $\gamma > 0$  and a finite constant  $C$  which only depends on  $\Im z > 0$  so that for  $N$  large enough

$$(69) \quad \mathbb{E}_{\mathbf{a}_k} [|\eta_m(z)|] \leq Cm^\gamma.$$

In the following we shall therefore neglect the term  $\eta_m(z)$ . By (83) and (84) we have the following bound: for any  $t \neq 0$  and any  $z \in \mathbb{C} \setminus \mathbb{R}$  such that  $|\Im z| \geq \delta > 0$ ,

$$(70) \quad \left| \sum_{j=1}^{N-1} \mathbf{a}_k(j)^2 (G_k(z)^2)_{jj} e^{-it \sum \mathbf{a}_k(j)^2 G_k(z)_{jj}} \right| \leq \frac{1}{\delta} \sup_{x \geq 0} x e^{-tx} = \frac{e^{-1}}{t\delta}$$

so that for  $M$  large enough and  $|\Im z| \geq \delta > 0$ ,

$$\mathbb{E}_{\mathbf{a}_k} \tilde{f}_{k,m} = -i \int_{\text{sgn}_z m}^{\text{sgn}_z M} \mathbb{E}_{\mathbf{a}_k} [\partial_z \lambda_N(z) e^{it\lambda_N(z)}] dt + \varepsilon_{m,M,z,N}$$

where  $\varepsilon_{m,M,z,N}$  is arbitrary small as  $M$  is large, uniformly on  $N$  and on the randomness. Moreover, by (70) one has

$$(71) \quad \mathbb{E}_{\mathbf{a}_k} [\partial_z \lambda_N(z) e^{it\lambda_N(z)}] = \frac{1}{it} \partial_z (\mathbb{E}_{\mathbf{a}_k} [e^{it\lambda_N(z)}])$$

Recall that  $\phi_N(\lambda) = \mathbb{E}[\exp(-i\lambda|a_{11}|^2)]$ , so we have

$$(72) \quad \mathbb{E}_{\mathbf{a}_k} [e^{it\lambda_N(z)}] = \mathbb{E}_{\mathbf{a}_k} \left[ e^{itz - it \sum_j \mathbf{a}_k(j)^2 G_k(z)_{jj}} \right] = e^{itz} \prod_{j=1}^{N-1} \phi_N(tG_k(z)_{jj}).$$

Remind that in (65) we have shown in the proof of the last item of Corollary 3.3 the almost sure convergence

$$(73) \quad e^{itz} \prod_{j=1}^{N-1} \phi_N(tG_k(z)_{jj}) \xrightarrow{N \rightarrow \infty} e^{itz + \rho_z(t)}.$$

As in the proof of Corollary 3.3, since the left hand side is analytic and uniformly bounded, we deduce by Montel's theorem that its convergence entails the convergence of its derivatives. We then get by (72), for all  $t, z$  so that  $t/\Im z > 0$ , the almost sure convergence

$$\partial_z \left( \mathbb{E}_{\mathbf{a}_k} \left[ e^{itz - it \sum_j \mathbf{a}_k(j)^2 G_k(z)_{jj}} \right] \right) \xrightarrow{N \rightarrow \infty} \partial_z \left( e^{itz + \rho_z(t)} \right).$$

We then obtain by dominated convergence (remember the integrant is uniformly bounded by (71) and estimate (70)),

$$\mathbb{E}_{\mathbf{a}_k} \tilde{f}_{k,m} = - \int_{\operatorname{sgn}_z m}^{\operatorname{sgn}_z M} \frac{1}{t} \partial_z e^{itz + \rho_z(t)} dt + \tilde{\varepsilon}_{m,M,z,N},$$

where  $\tilde{\varepsilon}_{m,M,z,N}$  converges to zero almost surely as  $N$  goes to infinity.

By (71) and estimate (70), we have the estimate

$$\left| \frac{1}{t} \partial_z e^{itz + \rho_z(t)} \right| \leq \left( 1 + \frac{e^{-1}}{t \Im z} \right) e^{-t \Im z}$$

so we can let  $M$  going to infinity to obtain the almost sure convergence

$$\mathbb{E}_{\mathbf{a}_k} \tilde{f}_{m,k} \xrightarrow{N \rightarrow \infty} - \int_{\operatorname{sgn}_z m}^{\operatorname{sgn}_z \infty} \frac{1}{t} \partial_z e^{itz + \rho_z(t)} dt.$$

In the Lévy case, one has  $\rho_z(t) = t^{\frac{\alpha}{2}} \rho_z(1)$ , and in the exploding moments case,  $\left| \frac{1}{t} \partial_z e^{itz + \rho_z(t)} \right| \leq 1 + \frac{C_2}{\Im z^2}$ , so that the integral converges at zero and we obtain

$$\mathbb{E}_{\mathbf{a}_k} f_k \xrightarrow{N \rightarrow \infty} - \int_0^{\operatorname{sgn}_z \infty} \frac{1}{t} \partial_z e^{itz + \rho_z(t)} dt$$

as expected. □

Of course, this convergence is uniform in  $k$  since the law of  $G_k(z)$  does not depend on  $k$  and an analogous formula is true for  $\mathbb{E}_{\mathbf{a}_k} f'_k$ , replacing  $z$  by  $z'$ .

### 3.5. Computation of $\mathbb{E}_{\mathbf{a}_k}(f_k \times f_k'')$ .

**Lemma 3.5.** *Almost surely, we have the convergence*

$$(74) \quad \mathbb{E}_{\mathbf{a}_k} f_k f_k'' \xrightarrow{N \rightarrow \infty} \Psi^u(z, z') := \int_0^{\operatorname{sgn}_z \infty} \int_0^{\operatorname{sgn}_{z'} \infty} \frac{1}{tt'} \partial_{z, z'}^2 e^{itz + it'z' + \rho_{z, z'}^u(t, t')} dt dt',$$

where  $\rho_{z, z'}^u$  is defined in Corollary 3.3.

*Proof.* We shall start again by using Formula (67) for  $f_k$ , and its analogue for

$$f_k'' = \frac{1 + \sum_j \mathbf{a}_k(j)^2 (G_k'(z'))_{jj}}{z - \sum_j \mathbf{a}_k(j)^2 G_k'(z')_{jj}} := \frac{\partial_z \lambda'_N(z')}{\lambda'_N(z')},$$

where  $G_k'(z')$  be defined as  $G_k(z)$ , replacing  $z$  by  $z'$  and the matrix  $A$  by the matrix  $A'$  defined by (53), which gives

$$(75) \quad \begin{aligned} f_k'' &= -i \int_0^{\operatorname{sgn}_{z'} \infty} \partial_z \lambda'_N(z') e^{it' \lambda'_N(z')} dt' \\ &= -i \int_{\operatorname{sgn}_z m}^{\operatorname{sgn}_{z'} \infty} \partial_z \lambda'_N(z') e^{it' \lambda'_N(z')} dt' - i \int_0^{\operatorname{sgn}_z m} \partial_z \lambda'_N(z') e^{it' \lambda'_N(z')} dt' \end{aligned}$$

(76)

The upper bound (70) allows us to bound the first term uniformly by  $\log m^{-1}$  and to truncate the integrals for  $\operatorname{sgn}_z t, \operatorname{sgn}_{z'} t' \leq M$ . Therefore, up to a small error  $\varepsilon_{M, m, z, z', N}$  uniform for  $M$  large,  $m$  small and provided  $|\Im z|, |\Im z'| \geq \delta > 0$ , we have

$$(77) \quad \tilde{f}_k \times \tilde{f}_k'' = - \int_{\operatorname{sgn}_z m}^{\operatorname{sgn}_z M} \int_{\operatorname{sgn}_{z'} m}^{\operatorname{sgn}_{z'} M} \mathbb{E}_{\mathbf{a}_k} \partial_z \lambda_N(z) \partial_z \lambda'_N(z') e^{it \lambda_N(z) + it' \lambda'_N(z')} dt dt' + \varepsilon_{M, m, z, z', N}$$

As in the previous case, the upper bound (70) allows us

$$(78) \quad \mathbb{E}_{\mathbf{a}_k} \left[ \partial_z \lambda_N(z) \partial_z \lambda'_N(z') e^{it \lambda_N(z) + it' \lambda'_N(z')} \right] = -\frac{1}{tt'} \partial_{z, z'}^2 \left( \mathbb{E}_{\mathbf{a}_k} \left[ e^{it \lambda_N(z) + it' \lambda'_N(z')} \right] \right)$$

Remember that  $\phi_N(\lambda) = \mathbb{E}[\exp(-i\lambda|a_{11}|^2)]$ , so we have

$$(79) \quad \mathbb{E}_{\mathbf{a}_k} \left[ e^{it \lambda_N(z) + it' \lambda'_N(z')} \right] = e^{itz + it'z'} \prod_{j=1}^{N-1} \phi_N(tG_k(z)_{jj} + t'G_k'(z')_{jj}).$$

By assumption (8) and Corollary 3.3, we have the following almost sure convergence as  $N$  goes to infinity and  $\frac{k}{N}$  goes to  $u$  in  $(0, 1)$

$$e^{itz + it'z'} \prod_{j=1}^{N-1} \phi_N(tG_k(z)_{jj} + t'G_k'(z')_{jj}) \xrightarrow{N \rightarrow \infty} e^{itz + it'z' + \rho_{z, z'}^u(t, t')}.$$

Almost surely, for any  $t, t'$ , the map  $(z, z') \mapsto e^{itz + it'z'} \prod_{j=1}^{N-1} \phi_N(tG_k(z)_{jj} + t'G_k'(z')_{jj})$  is analytic on  $\mathbb{C}^{\operatorname{sgn} t} \times \mathbb{C}^{\operatorname{sgn} t'}$  and bounded by one. Hence, with the same arguments as in the

previous section, we get almost surely and for any  $t, t'$  so that  $t/\Im z$  and  $t'/\Im z'$  are positive, the uniform convergence for the second derivatives on compact subsets. The truncations of the integrals can be suppressed as in the previous section, we obtain the almost sure convergence

$$\mathbb{E}_{\mathbf{a}_k} f_k \times f_k'' \xrightarrow{N \rightarrow \infty} \int_0^{\operatorname{sgn}_z \infty} \int_0^{\operatorname{sgn}_{z'} \infty} \frac{1}{tt'} \partial_{z, z'}^2 e^{itz + it'z' + \rho_{z, z'}^u(t, t')} dt dt'.$$

□

Hence we have proved the convergences (55) and (56). This completes the proof of Theorem 1.3.

#### 4. FIXED POINT CHARACTERIZATIONS

In this section, we provide characterizations of the functions  $\rho_z$  and  $\rho_{z, z'}^u$  involved in the covariance of the limiting process of Theorem 1.3 as fixed points of certain functions. Lévy matrices and Wigner matrices with exploding moments are considered in a unified way.

**4.1. Fixed point characterization of  $\rho_z(\cdot)$ : proof of Theorem 1.5.** We now prove the fixed point equation for the non random function involved in the Lemma 3.4, given for  $z \in \mathbb{C}^+$  and  $\lambda > 0$  by,

$$\rho_z(\lambda) = \lim_{N \rightarrow \infty} \rho_z^N(\lambda) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N \Phi(\lambda G_k(z)_{jj}),$$

where we have proved that this convergence holds almost surely in Corollary 3.3. Note however that under the assumptions of Theorem 1.5, the arguments below provide another proof of this convergence, see next section for details.

We denote in short  $A$  for  $A_k$ ,  $G$  for  $G_k$  and  $\mathbf{a}$  for  $\mathbf{a}_k$  in the following, and we do not detail the steps of the proof, which are very similar to those in [10] and Corollary 3.3, but outline them. Since we have already seen that  $\rho$  is analytic in Corollary 3.3 we need only to prove the fixed point equation. Let  $A_1$  be the  $(N-2) \times (N-2)$  principal submatrix of  $A$  obtained removing the first row and the first column of  $A$ , and let  $G_1(z) := (z - A_1)^{-1}$ . Let  $\mathbf{a}_1$  be the first column of  $A$  where the first entry has been removed. So using first the concentration lemma 6.4, then exchangeability of the  $G(z)_{jj}$ 's, then Schur complement formula (see [1, Lem. 2.4.6]), and then the fact that we can get rid of the off diagonal terms by the same argument as in the proof of Lemma 3.1 since  $\Phi$  is continuous on  $\overline{\mathbb{C}^-}$ ,

we have for all  $z \in \mathbb{C}^+$

$$\begin{aligned}\mathbb{E}[\rho_z^N(\lambda)] &= \mathbb{E}[\Phi(\lambda G(z)_{11})] \\ &= \mathbb{E}\left[\Phi\left(\frac{\lambda}{z - a_{11} - \mathbf{a}_1^* G_1(z) \mathbf{a}_1}\right)\right] + o(1) \\ &= \mathbb{E}\left[\Phi\left(\frac{\lambda}{z - \sum_{j=1}^{N-1} \mathbf{a}_1(j)^2 G_1(z)_{jj}}\right)\right] + o(1).\end{aligned}$$

Then we use Hypothesis made at (18) to get

$$\begin{aligned}\mathbb{E}[\rho_z^N(\lambda)] &= \int_0^\infty g(y) \mathbb{E}[e^{i\frac{y}{\lambda}(z - \sum_j \mathbf{a}_1(j)^2 G_1(z)_{jj})}] dy + o(1) \\ &= \int_0^\infty g(y) e^{i\frac{y}{\lambda}z} \mathbb{E}\left[\prod_{j=1}^{N-1} \phi_N\left(\frac{y}{\lambda} G_1(z)_{jj}\right)\right] dy + o(1).\end{aligned}$$

Using the definition of  $\Phi$  and the fact that we assumed that it is bounded on every compact subset (since  $\rho^N$  has non positive real part we can cut the integral to keep  $y$  bounded up to a small error, as in the previous sections), we have

$$\begin{aligned}\mathbb{E}[\rho_z^N(\lambda)] &= \int_0^\infty g(y) e^{i\frac{y}{\lambda}z} \mathbb{E}\left[e^{\frac{1}{N} \sum_j \Phi\left(\frac{y}{\lambda} G_1(z)_{jj}\right)}\right] dy + o(1) \\ &= \int_0^\infty g(y) e^{i\frac{y}{\lambda}z} \mathbb{E}\left[e^{\rho_z^{N-1}\left(\frac{y}{\lambda}\right)}\right] dy + o(1) \\ &= \lambda \int_0^\infty g(\lambda y) e^{iyz} \mathbb{E}\left[e^{\rho_z^{N-1}(y)}\right] dy + o(1).\end{aligned}$$

Now, since  $\Re(\rho^{N-1}) \leq 0$  and  $g$  satisfies (19), we may use Corollary 3.3 to take the large  $N$  limit and we get (20) by dominated convergence theorem (by Lemma 6.2,  $\rho_z^{N-1}$  can be replaced by  $\rho_z^N$  in the above formula).

Let us now prove uniqueness of the solutions to this equation. Suppose that there are two solutions  $\rho_z(\lambda), \tilde{\rho}_z(\lambda)$ . For  $z$  fixed, let us define  $\Delta(\lambda) := |\rho_z(\lambda) - \tilde{\rho}_z(\lambda)|$ . Then for all  $\lambda$ , we have, by the hypothesis made on  $g$  in Hypothesis made at (18),

$$\begin{aligned}\Delta(\lambda) &\leq \lambda \int_0^\infty |g(\lambda y)| e^{-y\Im z} \Delta(y) dy \\ &\leq K \left( \lambda^{\gamma+1} \int_0^{\frac{1}{\lambda}} y^\gamma e^{-y\Im z} \Delta(y) dy + \lambda^{\kappa+1} \int_{\frac{1}{\lambda}}^\infty y^\kappa e^{-y\Im z} \Delta(y) dy \right) \\ &\leq K \left( \lambda^{\gamma+1} \int_0^\infty y^\gamma e^{-y\Im z} \Delta(y) dy + \lambda^{\kappa+1} \int_0^\infty y^\kappa e^{-y\Im z} \Delta(y) dy \right)\end{aligned}$$

It follows that  $I_1 := \int_0^\infty \lambda^\gamma e^{-\lambda \Im z} \Delta(\lambda) d\lambda$  and  $I_2 := \int_0^\infty \lambda^\kappa e^{-\lambda \Im z} \Delta(\lambda) d\lambda$  satisfy

$$\begin{aligned} I_1 &\leq K \left( I_1 \int_0^\infty \lambda^{2\gamma+1} e^{-\lambda \Im z} d\lambda + I_2 \int_0^\infty \lambda^{\gamma+\kappa+1} e^{-\lambda \Im z} d\lambda \right), \\ I_2 &\leq K \left( I_1 \int_0^\infty \lambda^{\gamma+\kappa+1} e^{-\lambda \Im z} d\lambda + I_2 \int_0^\infty \lambda^{2\kappa+1} e^{-\lambda \Im z} d\lambda \right). \end{aligned}$$

For  $\Im z$  large enough, the integrals above are strictly less than  $\frac{1}{2K}$ , so for  $\Im z$  large enough,  $I_1 = I_2 = 0$ . It follows that for any fixed  $\lambda$ ,  $\rho_z(\lambda)$  and  $\tilde{\rho}_z(\lambda)$  are analytic functions of  $z$  which coincide for  $\Im z$  large enough, hence are equal.

**4.2. Fixed point characterization of  $\rho_{z,z'}(\cdot, \cdot)$ : proof of Theorem 1.6.** We now find a fixed point system of equations for the non random function of Corollary 3.3. For  $\lambda \Im z \geq 0$  and  $\lambda' \Im z' \geq 0$ , we set

$$\begin{aligned} \rho_{z,z'}^{N,k,1}(\lambda, \lambda') &= \frac{1}{k-1} \sum_{j=1}^{k-1} \Phi(\lambda G_k(z)_{jj} + \lambda' G'_k(z')_{jj}) \\ \rho_{z,z',k}^{N,k,2}(\lambda, \lambda') &= \frac{1}{N-k-1} \sum_{j=k}^{N-1} \Phi(\lambda G_k(z)_{jj} + \lambda' G'_k(z')_{jj}), \end{aligned}$$

where we recall that  $G_k$  and  $G'_k$  are as in (17). To simplify the notations below, as in the previous section, we denote  $(G, G')$  instead of  $(G_k, G'_k)$ , even though their distribution depends on  $k$ .

In the sequel we fix, as in Section 3.3, a number  $u \in (0, 1)$  and will give limits in the regime where  $N \rightarrow \infty$ ,  $k \rightarrow \infty$  and  $k/N \rightarrow u$ . We shall then prove that, under the hypotheses of Theorem 1.6 that we assume throughout this section,  $(\rho_{z,z'}^{N,k,1}, \rho_{z,z'}^{N,k,2})$  converges almost surely and that its limit satisfies a fixed point system of equations which has a unique analytic solution with non positive real part. The convergence could be shown with minor modifications of Lemma 3.2, but we do not need this since we work with stronger assumptions. Using the concentration lemma 6.4 (note that  $\Phi$  is not Lipschitz but can be approximated by Lipschitz functions uniformly on compacts), it is sufficient to prove the fixed point equation for the expectation of these parameters. Moreover, by exchangeability of the  $k$  first entries and  $N - k$  last entries

$$\begin{aligned} \mathbb{E}[\rho_{z,z',k}^{N,1}(t, s)] &= \mathbb{E}[\Phi(tG(z)_{11} + sG'(z')_{11})] + o(1), \\ \mathbb{E}[\rho_{z,z',k}^{N,1}(t, s)] &= \mathbb{E}[\Phi(tG(z)_{NN} + sG'(z')_{NN})] + o(1). \end{aligned}$$

These functions are analytic in  $z, z' \in \mathbb{C}^+$ , smooth and bounded in  $s, t$  on compact subsets and hence tight by Arzela-Ascoli. We let  $\rho_{z,z'}^{u,s}$ ,  $s = 1, 2$  be a limit point. We assume for simplicity that both  $z$  and  $z'$  have positive imaginary parts (in the general case, one only has to replace  $\int_{s=0}^{+\infty} \int_{t=0}^{+\infty}$  below by  $\int_{s=0}^{\operatorname{sgn}(\Im z)\infty} \int_{t=0}^{\operatorname{sgn}(\Im z')\infty}$ ). Under Hypothesis made at (24),



we can write by Schur complement formula and getting rid of the off diagonal terms (note that all integrals are finite as they contain exponentially decreasing terms)

$$\begin{aligned}
\mathbb{E}[\rho_{z,z'}^{N,k,1}(t,s)] &= \int_0^\infty \int_0^\infty e^{i\frac{v}{t}z + i\frac{v'}{s}z'} \mathbb{E}[e^{-i\frac{v}{t}\sum_\ell \mathbf{a}_1(\ell)^2 G_1(z)\ell\ell - i\frac{v'}{s}\sum_\ell \mathbf{a}'_1(\ell)^2 G'_1(z)\ell\ell}] d\tau(v,v') \\
&+ \int e^{i\frac{v}{t}} \mathbb{E}[e^{-i\frac{v}{t}\sum_\ell \mathbf{a}_1(\ell)^2 G_1(z)\ell\ell}] d\mu(v) + \int e^{i\frac{v'}{s}} \mathbb{E}[e^{-i\frac{v'}{s}\sum_\ell \mathbf{a}'_1(\ell)^2 G'_1(z)\ell\ell}] d\mu(v) + o(1) \\
&= \int_0^\infty \int_0^\infty e^{i\frac{v}{t}z + i\frac{v'}{s}z'} e^{u\mathbb{E}\rho_{z,z'}^{N,k,1}(\frac{v}{t}, \frac{v'}{s}) + (1-u)\mathbb{E}\rho_{z,z'}^{N,k,2}(\frac{v}{t}, \frac{v'}{s})} d\tau(v,v') \\
&+ \int_0^\infty e^{i\frac{v}{t}} e^{\rho_z(\frac{v}{t})} d\mu(v) + \int_0^\infty e^{i\frac{v'}{s}} e^{\rho_{z'}(\frac{v'}{s})} d\mu(v) + o(1)
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}[\rho_{z,z'}^{N,k,2}(t,s)] &= \int_0^\infty \int_0^\infty e^{i\frac{v}{t}z + i\frac{v'}{s}z'} \mathbb{E}[e^{-i\frac{v}{t}\sum_\ell \mathbf{a}_N(\ell)^2 G_N(z)\ell\ell - i\frac{v'}{s}\sum_\ell \mathbf{a}'_N(\ell)^2 G'_N(z)\ell\ell}] d\tau(v,v') \\
&+ \int e^{i\frac{v}{t}} \mathbb{E}[e^{-i\frac{v}{t}\sum_\ell \mathbf{a}_N(\ell)^2 G_N(z)\ell\ell}] d\mu(v) + \int e^{i\frac{v'}{s}} \mathbb{E}[e^{-i\frac{v'}{s}\sum_\ell \mathbf{a}'_N(\ell)^2 G'_N(z)\ell\ell}] d\mu(v) + o(1) \\
&= \int_0^\infty \int_0^\infty e^{i\frac{v}{t}z + i\frac{v'}{s}z'} e^{u\mathbb{E}\rho_{z,z'}^{N,k,1}(\frac{v}{t}, \frac{v'}{s}) + (1-u)\rho_z(\frac{v}{t}) + (1-u)\rho_{z'}(\frac{v'}{s})} d\tau(v,v') \\
&+ \int_0^\infty e^{i\frac{v}{t}} e^{\rho_z(\frac{v}{t})} d\mu(v) + \int_0^\infty e^{i\frac{v'}{s}} e^{\rho_{z'}(\frac{v'}{s})} d\mu(v) + o(1)
\end{aligned}$$

where we used that  $\mathbf{a}_1(\ell) = \mathbf{a}'_1(\ell)$  for all  $\ell$  whereas  $\mathbf{a}_N(\ell) = \mathbf{a}'_N(\ell)$  for  $\ell \leq k$  and are independent otherwise, that  $\rho_z^N$  converges towards  $\rho_z$ , and that

$$\begin{aligned}
\frac{1}{k-2} \sum_{j=1}^{k-2} \Phi(\lambda G_1(z)_{jj} + \lambda' G'_1(z')_{jj}) &\sim \rho_{z,z'}^{N,k,1} \\
\frac{1}{N-k-2} \sum_{j=k}^{N-2} \Phi(\lambda G_N(z)_{jj} + \lambda' G'_N(z')_{jj}) &\sim \rho_{z,z'}^{N,k,2},
\end{aligned}$$

by Lemma 6.2 (by continuity of  $\Phi$ , it can be approximated by Lipschitz functions). Hence we find that the limit points  $\rho_{z,z'}^{u,1}, \rho_{z,z'}^{u,2}$  of  $\rho_{z,z'}^{N,k,1}, \rho_{z,z'}^{N,k,2}$  satisfy (25). Moreover,

$$\rho_{z,z'}^{N,k} = u\rho_{z,z'}^{N,k,1} + (1-u)\rho_{z,z'}^{N,k,2} + o(1)$$

gives

$$\rho_{z,z'}^u = u\rho_{z,z'}^{u,1} + (1-u)\rho_{z,z'}^{u,2}.$$

Let  $\rho$  and  $\tilde{\rho}$  be solutions of Equation (25) with non positive real parts (note here that  $\rho_z(u)$  is given and  $\mu$  is so that the above integrals are finite; hence the last two terms in

both equations play the role of a finite given function).

$$\begin{aligned}
\Delta(t, s) &:= |\rho_{z, z'}^{u, 1}(t, s) - \tilde{\rho}_{z, z'}^{u, 2}(t, s)| + |\rho_{z, z'}^{u, 1}(t, s) - \tilde{\rho}_{z, z'}^{u, 2}(t, s)| \\
&\leq 2 \int_0^\infty \int_0^\infty e^{-\Im z v t^{-1} - \Im z' v' s^{-1}} \frac{d|\tau|(v, v')}{dv dv'} \Delta\left(\frac{v}{t}, \frac{v'}{s}\right) dv dv' \\
&\leq 2ts \int_0^\infty \int_0^\infty e^{-\Im z v - \Im z' v'} K((vt)^\gamma 1_{vt \in [0, 1]} + (vt)^\kappa 1_{vt \in [1, \infty[)}) \\
&\quad \times ((v's)^\gamma 1_{v's \in [0, 1]} + (v's)^\kappa 1_{v's \in [1, \infty[)}) \Delta(v, v') dv dv' \\
&\leq 2Kts (t^\gamma s^{\gamma'} I_\Delta(\gamma, \gamma') + t^\gamma s^{\kappa'} I_\Delta(\gamma, \kappa') + t^\kappa s^{\gamma'} I_\Delta(\kappa, \gamma') + t^\kappa s^{\kappa'} I_\Delta(\kappa, \kappa')),
\end{aligned}$$

where for  $(\alpha, \alpha') \in \{\gamma, \kappa\} \times \{\gamma', \kappa'\}$ , we have set

$$I_\Delta(\alpha, \alpha') := \int_0^\infty \int_0^\infty e^{-\Im z v - \Im z' v'} \Delta(v, v') v^\alpha v'^{\alpha'} dv dv'.$$

We put  $I(\alpha, \alpha') = I_1(\alpha, \alpha')$  where 1 denote the constant function equal to one. We get after integrating both sides

$$\begin{aligned}
I_\Delta(\alpha, \alpha') &\leq K(I(\alpha + \gamma + 1, \alpha' + \gamma' + 1)I_\Delta(\gamma, \gamma') + I(\alpha + \gamma + 1, \alpha' + \kappa' + 1)I_\Delta(\gamma, \tilde{\gamma}') + \\
&\quad I(\alpha + \kappa + 1, \alpha' + \gamma' + 1)I_\Delta(\tilde{\gamma}, \gamma') + I(\alpha + \kappa + 1, \alpha' + \tilde{\gamma}' + 1)I_\Delta(\tilde{\gamma}, \tilde{\gamma}')).
\end{aligned}$$

We consider  $\Im z, \Im z'$  large enough so that  $I(\alpha + \beta + 1, \alpha' + \beta' + 1) < \frac{1}{4K}$  for any  $(\alpha, \alpha'), (\beta, \beta') \in \{\gamma, \tilde{\gamma}\} \times \{\gamma', \tilde{\gamma}'\}$  to conclude that  $\Delta(t, s)$  vanishes then and therefore that  $\rho_{z, z'}^{u, s} = \tilde{\rho}_{z, z'}^{u, s}$  for  $\Im z$  and  $\Im z'$  big enough,  $s = 1$  or  $2$ . By analyticity, we conclude that the system of equations (25) has a unique analytic solution with non positive real part. The functions  $\rho^{N, k, 1}$  and  $\rho^{N, k, 2}$  are tight and they limits points are characterized by fixed point equations, so they actually converge.

## 5. PROOF OF COROLLARY 1.4

By linearity and the fact that  $\mathcal{A}$  is stable under complex conjugation, it is enough to consider a single function  $f \in \mathcal{A}$ . By Cauchy formula and Hypothesis (11) on  $f$ , we have for all  $x \in \mathbb{R}$ ,

$$f(x) = \frac{1}{2i\pi} \int_\Gamma \frac{f(z)}{z - x} dz$$

where  $\Gamma = \mathbb{R} \pm i\epsilon$ . Let  $f_L(x) = \frac{1}{2i\pi} \int_{\Gamma \cap \{\Re z \geq L\}} \frac{f(z)}{z - x} dz$  and observe that for  $L$  big enough, the polynomial decay of  $f$  at infinity gives

$$\|f_L\|_{TV} \leq \frac{1}{2\pi} \int_{|x| \geq L} x^{-\gamma} dx \int \frac{dy}{|y - x|^2 + \epsilon^2} \leq C(L^{-\gamma} + L^{1-\gamma}).$$

Therefore, Lemma 6.3 shows that  $Z_N(f_L)$  goes to zero in probability as  $N$  goes to infinity and then  $L$  goes to infinity: for any  $\eta > 0$ ,

$$\lim_{L \rightarrow \infty} \limsup_{N \rightarrow \infty} \mathbb{P}(|Z_N(f_L)| > \eta) = 0.$$

Fix  $L, \delta > 0$  and consider  $x_k = -L + k\delta$  for  $k \in [0, R]$ , with  $R = \lfloor 2L\delta^{-1} \rfloor$ . Let

$$\tilde{f}_L(x) = \frac{1}{2i\pi} \int_{\substack{z=y \pm i\epsilon \\ |y| \leq L}} \frac{f(z)}{z-x} dz, \quad \tilde{f}_L^\delta(x) = \frac{1}{2i\pi} \sum_{s=\pm 1} \sum_{k=0}^R \int_{x_k}^{x_{k+1}} f(y+i s\epsilon) dy \times (x_k+i s\epsilon-x)^{-1}.$$

Note that for each  $L$  and  $\epsilon > 0$  there exists a finite constant  $K(L, \epsilon)$  so that

$$\|\tilde{f}_L - \tilde{f}_L^\delta\|_{TV} \leq K(L, \epsilon)\delta.$$

Hence, Lemma 6.3 implies that  $Z_N(\tilde{f}_L) - Z_N(\tilde{f}_L^\delta)$  goes to zero in probability as  $N$  goes to infinity and  $\delta$  goes to zero for any given  $L, \epsilon > 0$ . Finally, by Theorem 1.3,  $Z_N(\tilde{f}_L^\delta)$  converges in law towards a centered Gaussian variable with covariance

$$\frac{1}{4\pi^2} \sum_{s=\pm 1} \sum_{k=0}^R \sum_{s'=\pm 1} \sum_{k'=0}^R \int_{x_k}^{x_{k+1}} f(y+i s\epsilon) dy \int_{x_{k'}}^{x_{k'+1}} f(y+i s'\epsilon) dy C(x_k+i\epsilon, x_{k'}+i\epsilon)$$

The above covariance converges as  $\delta$  goes to zero by smoothness of the functions  $C$  towards

$$\frac{1}{4\pi^2} \int_{|x| \leq L} \int_{|x'| \leq L} dx dx' f(x+i\epsilon) f(x'+i\epsilon) C(x+i\epsilon, x'+i\epsilon).$$

Therefore, writing

$$Z_N(f) = Z_N(f_L) + [Z_N(\tilde{f}_L) - Z_N(\tilde{f}_L^\delta)] + Z_N(\tilde{f}_L^\delta)$$

and fixing  $L$  large enough and then  $\delta$  small enough so that the first two terms are small with high probability, we conclude that  $Z_N(f)$  is close to  $Z_N(\tilde{f}_L^\delta)$  and therefore converges in law towards the announced centered Gaussian variable. The generalization to the closure of  $\mathcal{A}$  for the total variation norm is a direct consequence of Lemma 6.3.

## 6. APPENDIX

**6.1. Concentration of random matrices with independent rows and linear algebra lemmas.** This section is mostly a reminder of results from [12] and [13].

The total variation norm of  $f : \mathbb{R} \rightarrow \mathbb{R}$  is

$$\|f\|_{TV} := \sup \sum_{k \in \mathbb{Z}} |f(x_{k+1}) - f(x_k)|,$$

where the supremum runs over all sequences  $(x_k)_{k \in \mathbb{Z}}$  such that  $x_{k+1} \geq x_k$  for any  $k \in \mathbb{Z}$ . If  $f = 1_{(-\infty, s]}$  for some real  $s$  then  $\|f\|_{TV} = 1$ , while if  $f$  has a derivative in  $L^1(\mathbb{R})$ , we get

$$\|f\|_{TV} = \int |f'(t)| dt.$$

The next lemma is an easy consequence of Cauchy-Weyl interlacing Theorem. It is an ingredient of the proof of Lemma 6.3.

**Lemma 6.1** (Interlacing of eigenvalues). *Let  $A$  be an  $n \times n$  hermitian matrix and  $B$  a principal minor of  $A$ . Then for any  $f : \mathbb{R} \rightarrow \mathbb{C}$  such that  $\|f\|_{TV} \leq 1$  and  $\lim_{|x| \rightarrow \infty} f(x) = 0$ ,*

$$\left| \sum_{i=1}^n f(\lambda_i(A)) - \sum_{i=1}^{n-1} f(\lambda_i(B)) \right| \leq 1.$$

**Lemma 6.2.** *Let  $A_1, A_2$  be  $n \times n$  random Hermitian matrices and  $\tilde{A}_1, \tilde{A}_2$  be  $(n-1) \times (n-1)$  matrices obtained from  $A_1$  and  $A_2$  respectively by removing the  $\ell$ -th row and column, for some  $\ell \in \{1, \dots, n\}$ . Let  $z, z' \in \mathbb{C}$ ,  $t, t' \in \mathbb{R}$  so that  $\Im z t > 0$  and  $\Im z' t' > 0$  and set  $G = (z - A_1)^{-1}$ ,  $G' = (z' - A_2)^{-1}$  and  $\tilde{G} = (z - \tilde{A}_1)^{-1}$ ,  $\tilde{G}' = (z' - \tilde{A}_2)^{-1}$ . Then, for any function  $f$  on  $B_{z, z', t, t'} := \{g \in \mathbb{C}^-; |g| \leq C(z, z', t, t')\}$  with  $C(z, z', t, t') = t(\Im z)^{-1} + t'(\Im z')^{-1}$ , we have*

$$(80) \quad \left| \frac{1}{n} \sum_{k=1}^n f(tG_{kk} + t'G'_{kk}) - \frac{1}{n} \sum_{k=1}^{n-1} f(t\tilde{G}_{kk} + t'\tilde{G}'_{kk}) \right| \leq \frac{C(z, z', t, t')}{n} \|f\|_{\text{Lip}} + \frac{\|f\|_{\infty}}{n},$$

where  $\|f\|_{\text{Lip}} := \sup_{x \neq y} \frac{|f(y) - f(x)|}{|y - x|}$  and  $\|f\|_{\infty} := \sup_x |f(x)|$ , both supremums running over the elements of  $B_{z, z', t, t'}$ .

*Proof.* The proof is similar to [12, (91)]. We denote by  $\bar{A}_1, \bar{A}_2$  the  $n \times n$  matrices whose entries are the same as  $A_1, A_2$  except for the  $\ell$ th rows and column which have zero entries. We denote  $\bar{G}, \bar{G}'$  the corresponding Stieltjes transform. Then,  $\tilde{G}, \tilde{G}'$  equal  $\bar{G}, \bar{G}'$  except at the  $\ell$ th column and row (where it is equal to  $z^{-1}1_{i=j=k}$ ). Therefore, noting  $\bar{M} = t\bar{G}(z) + t'\bar{G}'(z')$  and  $\tilde{M} = t\tilde{G}(z) + t'\tilde{G}'(z)$ , we conclude that

$$\left| \frac{1}{n} \sum_{k=1}^n f(\tilde{M}_{kk}) - \frac{1}{n} \sum_{k=1}^{n-1} f(\bar{M}_{kk}) \right| \leq \frac{\|f\|_{\infty}}{n}.$$

Moreover, let  $M = tG(z) + t'G'(z')$  and note that  $A_1 - \bar{A}_1$  and  $A_2 - \bar{A}_2$  have rank one so that  $M - \bar{M}$  has rank one. On the other hand it is bounded uniformly by  $C(z, z', t, t') = \frac{t}{\Im z} + \frac{t'}{\Im z'}$ . Hence, we can write  $M - \bar{M} = cuu^*$  with a unit vector  $u$  and  $c$  bounded by  $C(z, z', t, t')$ . Therefore, since  $f$  is Lipschitz,

$$\begin{aligned} \left| \frac{1}{n} \sum_{k=1}^n f(M_{kk}) - \frac{1}{n} \sum_{k=1}^{n-1} f(\bar{M}_{kk}) \right| &\leq \|f\|_{\text{Lip}} \frac{1}{n} \sum_{k=1}^n |M_{kk} - \bar{M}_{kk}| \\ &\leq \|f\|_{\text{Lip}} \frac{1}{n} \sum_{k=1}^n C(z, z', t, t') \langle e_k, u \rangle^2 \\ &= \frac{C(z, z', t, t') \|f\|_{\text{Lip}}}{n}. \end{aligned}$$

(80) follows. □

**Lemma 6.3** (Concentration for spectral measures [14]). *Let  $A$  be an  $n \times n$  random Hermitian matrix. Let us assume that the vectors  $(A_i)_{1 \leq i \leq n}$ , where  $A_i := (A_{ij})_{1 \leq j \leq i} \in \mathbb{C}^i$ , are independent. Then for any  $f : \mathbb{R} \rightarrow \mathbb{C}$  such that  $\|f\|_{\text{TV}} \leq 1$  and  $\mathbb{E} \int f d\mu_A < \infty$ , and every  $t \geq 0$ ,*

$$\mathbb{P} \left( \int f d\mu_A - \mathbb{E} \int f d\mu_A \geq t \right) \leq \exp \left( -\frac{nt^2}{2} \right).$$

**Lemma 6.4** (Concentration for the diagonal of the resolvent [12]). *Let  $A$  be an  $n \times n$  random Hermitian matrix and consider its resolvent matrix  $G(z) = (A - z)^{-1}$ ,  $z \in \mathbb{C}_+$ . Let us assume that the vectors  $(A_i)_{1 \leq i \leq n}$ , where  $A_i := (A_{ij})_{1 \leq j \leq i} \in \mathbb{C}^i$ , are independent. Then for any  $f : \mathbb{C}^- \rightarrow \mathbb{R}$  such that  $\|f\|_{\text{Lip}} \leq 1$ , and every  $t \geq 0$ ,*

$$\mathbb{P} \left( \left| \frac{1}{n} \sum_{k=1}^n f(G(z)_{kk}) - \mathbb{E} \frac{1}{n} \sum_{k=1}^n f(G(z)_{kk}) \right| \geq t \right) \leq 2 \exp \left( -\frac{n\Im(z)^2 t^2}{8} \right).$$

Let  $A'$  be an  $n \times n$  self-adjoint matrices so that  $A'_{ij} = A_{ij}$ ,  $i \wedge j \leq k$  and  $(A'_{ij})_{j \geq k+1, i \geq k+1}$  is independent from  $(A_{ij})_{j \geq k+1, i \geq k+1}$  but with the same distribution. Let  $G(z) = (z - A)^{-1}$  and  $G'(z) = (z - A')^{-1}$  and set for a Lipschitz function  $f$  on  $\overline{\mathbb{C}^-}$ ,

$$\begin{aligned} \rho_{z,z'}^{N,k,1}(\lambda, \lambda')[f] &:= \frac{1}{k} \sum_{\ell=1}^k f(\lambda G(z)_{\ell\ell} + \lambda' G(z')_{\ell\ell}) \\ \rho_{z,z'}^{N,k,2}(\lambda, \lambda')[f] &= \frac{1}{N-k} \sum_{\ell=k+1}^N f(\lambda G(z)_{\ell\ell} + \lambda' G(z')_{\ell\ell}) \end{aligned}$$

Then, for  $\lambda/\Im z \geq 0$ ,  $\lambda'/\Im z' \geq 0$ , we have for all  $\delta \geq 0$ ,  $s \in \{0, 1\}$ ,

$$\mathbb{P} \left( \left| \rho_{z,z'}^{N,k,s+1}(\lambda, \lambda')[f] - \mathbb{E}[\rho_{z,z'}^{N,k,s+1}(\lambda, \lambda')[f]] \right| \geq \delta \right) \leq 2e^{-\frac{\delta^2((k-1)^{1-s} + (N-k-1)^s)}{8\|f\|_{\text{Lip}}^2 C(\lambda, \lambda', t, t')^2}}$$

with

$$C(\lambda, \lambda', t, t') = \frac{2t}{\Im z} + \frac{2t'}{\Im z'}.$$

*Proof.* The first point is proved as in [12, Lemma C.3]. We outline the proof of the second point which is very similar to [12, Lemma C.3]. We concentrate on  $\rho_{\lambda, \lambda'}^{N,k,1}$ , the other case being similar. By Azuma-Hoeffding inequality, it is sufficient to show that

$$X_p := \mathbb{E}[\rho_{z,z'}^{N,k,1}(\lambda, \lambda')[f] | \mathcal{F}_p] - \mathbb{E}[\rho_{z,z'}^{N,k,1}(\lambda, \lambda')[f] | \mathcal{F}_{p-1}]$$

is uniformly bounded by  $\|f\|_{\text{Lip}} C(\lambda, \lambda', t, t') k^{-1}$ . Here  $\mathcal{F}_p$  is the sigma-algebra generated with respect to the  $p$  first column (and row) vectors. Note that  $X_p$  can be written as the conditional expectation of the difference of the parameter  $f$  evaluated at two sets  $A, A'$  and  $\hat{A}, \hat{A}'$  which differ only at the  $p$ th vector column (and row). Hence, we may follow the

proof of Lemma 6.2 to conclude that

$$|X_p| \leq \frac{\|f\|_{\text{Lip}}}{k} \sum_{\ell=1}^k |(M - \tilde{M})_{\ell\ell}| = \frac{|c|\|f\|_{\text{Lip}}}{k} \sum_{\ell=1}^k \langle u, e_k \rangle^2 \leq \frac{|c|\|f\|_{\text{Lip}}}{k}.$$

□

Let  $H = [h_{ij}]$  be an  $N \times N$  Hermitian matrix and  $z \in \mathbb{C} \setminus \mathbb{R}$ . Define  $G := (z - H)^{-1}$ .

**Lemma 6.5** (Difference of traces of a matrix and its major submatrices). *Let  $H_k$  be the submatrix of  $H$  obtained by removing its  $k$ -th row and its  $k$ -th column and set  $G_k := (z - H_k)^{-1}$ . Let also  $\mathbf{a}_k$  be the  $k$ -th column of  $H$  where the  $k$ -th entry has been removed. Then*

$$(81) \quad \text{Tr}(G) - \text{Tr}(G_k) = \frac{1 + \mathbf{a}_k^* G_k^2 \mathbf{a}_k}{z - h_{kk} - \mathbf{a}_k^* G_k \mathbf{a}_k}.$$

Moreover,

$$(82) \quad |\text{Tr}(G) - \text{Tr}(G_k)| \leq \pi |\Im z|^{-1}.$$

*Proof.* For (81), see [4, Th. A.5]. For (82), see Lemma 6.1. □

**Lemma 6.6.** *With the notation introduced above the previous lemma, for each  $1 \leq j \leq N$ ,*

$$(83) \quad \Im z \times \Im G_{jj} < 0,$$

$$(84) \quad |\Im z| \times |(G^2)_{jj}| \leq |\Im G_{jj}| \leq |\Im z|^{-1}$$

and for any  $\mathbf{a} = (\mathbf{a}_1, \dots, \mathbf{a}_N) \in \mathbb{C}^N$ ,

$$(85) \quad \left| \frac{1 + \sum_j |\mathbf{a}_j|^2 (G^2)_{jj}}{z - \sum_j |\mathbf{a}_j|^2 G_{jj}} \right| \leq 2 |\Im z|^{-1}.$$

*Proof.* Set  $z = x + iy$ ,  $x, y \in \mathbb{R}$ . Let  $\lambda_1, \dots, \lambda_N$  be the eigenvalues of  $H$ , associated with the orthonormalized collection of eigenvectors  $\mathbf{u}_1, \dots, \mathbf{u}_N$ . Let also  $\mathbf{e}_j$  denote the  $j$ th vector of the canonical basis. Then (83) and (84) follow directly from the following:

$$(86) \quad |(G^2)_{jj}| = \left| \sum_{k=1}^N \frac{|\langle \mathbf{e}_j, \mathbf{u}_k \rangle|^2}{(z - \lambda_k)^2} \right| \leq \sum_{k=1}^N \frac{|\langle \mathbf{e}_j, \mathbf{u}_k \rangle|^2}{(x - \lambda_k)^2 + y^2} \quad \Im G_{jj} = - \sum_{k=1}^N \frac{|\langle \mathbf{e}_j, \mathbf{u}_k \rangle|^2 y}{(x - \lambda_k)^2 + y^2}$$

Let us now prove (85). By (83), we know that  $\Im z$  and  $-\Im G_{jj}$  have the same sign, so

$$(87) \quad \left| \frac{1}{z - \sum_j |\mathbf{a}_j|^2 G_{jj}} \right| \leq \frac{1}{|\Im(z - \sum_j |\mathbf{a}_j|^2 G_{jj})|} \leq |\Im z|^{-1}.$$

Hence it remains only to prove that

$$(88) \quad \left| \frac{\sum_j |\mathbf{a}_j|^2 (G^2)_{jj}}{z - \sum_j |\mathbf{a}_j|^2 G_{jj}} \right| \leq |\Im z|^{-1},$$

We have

$$(89) \quad \left| \frac{\sum_j |\mathbf{a}_j|^2 (G^2)_{jj}}{z - \sum_j |\mathbf{a}_j|^2 G_{jj}} \right| \leq \frac{\sum_j |\mathbf{a}_j|^2 |(G^2)_{jj}|}{|\Im(z - \sum_j |\mathbf{a}_j|^2 G_{jj})|}.$$

By the first part of (84), the numerator of the RHT of (89) is  $\leq |\Im z|^{-1} \sum_j |\mathbf{a}_j|^2 |\Im G_{jj}|$ . Moreover, as  $\Im z$  and  $-\Im G_{jj}$  have the same sign, the denominator of the RHT of (89) is equal to  $|\Im z| + \sum_j |\mathbf{a}_j|^2 |\Im G_{jj}|$ . As a consequence,

$$\frac{\sum_j |\mathbf{a}_j|^2 |(G^2)_{jj}|}{|\Im(z - \sum_j |\mathbf{a}_j|^2 G_{jj})|} \leq |\Im z|^{-1} \frac{\sum_j |\mathbf{a}_j|^2 |\Im G_{jj}|}{|\Im z| + \sum_j |\mathbf{a}_j|^2 |\Im G_{jj}|} \leq |\Im z|^{-1}$$

and (88) is proved.  $\square$

**6.2. CLT for martingales.** Let  $(\mathcal{F}_k)_{k \geq 0}$  be a filtration such that  $\mathcal{F}_0 = \{\emptyset, \Omega\}$  and let  $(M_k)_{k \geq 0}$  be a square-integrable complex-valued martingale starting at zero with respect to this filtration. For  $k \geq 1$ , we define the random variables

$$Y_k := M_k - M_{k-1} \quad v_k := \mathbb{E}[|Y_k|^2 | \mathcal{F}_{k-1}] \quad \tau_k := \mathbb{E}[Y_k^2 | \mathcal{F}_{k-1}]$$

and we also define

$$v := \sum_{k \geq 1} v_k \quad \tau := \sum_{k \geq 1} \tau_k \quad L(\varepsilon) := \sum_{k \geq 1} \mathbb{E}[|Y_k|^2 1_{|Y_k| \geq \varepsilon}].$$

Let everything depend on a parameter  $N$ , so that  $\mathcal{F}_k = \mathcal{F}_k(N)$ ,  $Y_k = Y_k(N)$ ,  $v = v(N)$ ,  $\tau = \tau(N)$ ,  $L(\varepsilon) = L(\varepsilon, N)$ ,  $\dots$

Then we have the following theorem. It is proved in the real case at [11, Th. 35.12]. The complex case can be deduced noticing that for  $z \in \mathbb{C}$ ,  $\Re(z)^2$ ,  $\Im(z)^2$  and  $\Re(z)\Im(z)$  are linear combinations of  $z^2$ ,  $\bar{z}^2$ ,  $|z|^2$ .

**Theorem 6.7.** *Suppose that for some constants  $v \geq 0, \tau \in \mathbb{C}$ , we have the convergence in probability*

$$v(N) \xrightarrow{N \rightarrow \infty} v \quad \tau(N) \xrightarrow{N \rightarrow \infty} \tau$$

and that for each  $\varepsilon > 0$ ,

$$L(\varepsilon, N) \xrightarrow{N \rightarrow \infty} 0.$$

Then we have the convergence in distribution

$$\sum_{k \geq 1} Y_k(N) \xrightarrow{N \rightarrow \infty} Z,$$

where  $Z$  is a centered complex Gaussian variable such that  $\mathbb{E}(|Z|^2) = v$  and  $\mathbb{E}(Z^2) = \tau$ .

### 6.3. On the Hadamard product of Hermitian matrices.

**Proposition 6.8.** *Let  $A_1, \dots, A_p$  be  $N$  by  $N$  Hermitian random matrices whose entries have all their moments. Then, there exists a family of random variables  $(a_1, \dots, a_p)$  whose joint distribution is given by:*

$$\mathbb{E}[a_1^{n_1} \dots a_p^{n_p}] = \mathbb{E} \left[ \frac{1}{N} \text{Tr}[A_1^{n_1} \circ \dots \circ A_p^{n_p}] \right], \quad \forall n_1, \dots, n_p \geq 0,$$

where  $\circ$  denotes the Hadamard (entry-wise) product.

*Proof. Step 1.* We first assume that the matrices are deterministic and have distinct eigenvalues. By the spectral decomposition, for  $j = 1, \dots, p$ , we have  $A_j = \sum_{i=1}^N \lambda_{j,i} u_{j,i} u_{j,i}^*$  where  $\Lambda_j = (\lambda_{j,i})_{i=1, \dots, N}$  is the family of eigenvalues of  $A_j$  arranged in increasing order, and  $U_j = (u_{j,i})_{i=1, \dots, N}$  is the family of associated eigenvectors. For any  $n_1, \dots, n_p \geq 0$ , one has

$$\begin{aligned} \frac{1}{N} \text{Tr}[A_1^{n_1} \circ \dots \circ A_p^{n_p}] &= \frac{1}{N} \sum_{k=1}^N (A_1^{n_1})(k, k) \dots (A_p^{n_p})(k, k) \\ &= \frac{1}{N} \sum_{k=1}^N \left( \sum_{i_1=1}^N \lambda_{1,i_1}^{n_1} u_{1,i_1} u_{1,i_1}^* \right) (k, k) \dots \left( \sum_{i_p=1}^N \lambda_{p,i_p}^{n_p} u_{p,i_p} u_{p,i_p}^* \right) (k, k) \\ &= \frac{1}{N^p} \sum_{i_1, \dots, i_p=1}^N \lambda_{1,i_1}^{n_1} \dots \lambda_{p,i_p}^{n_p} \times \left( N^{p-1} \sum_{k=1}^N |u_{1,i_1}(k)|^2 \dots |u_{p,i_p}(k)|^2 \right). \end{aligned}$$

For  $j = 1, \dots, p$ , we set  $d\mu_{\Lambda_j} = \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_{j,i}}$  the empirical eigenvalues distributions of  $A_j$ . We denote  $F_{\Lambda_j}(t) = \mu_{\Lambda_j}((-\infty, t])$  the cumulative function of  $d\mu_{\Lambda_j}$ . Since the eigenvalues of the matrices are distinct, one has  $F_{\Lambda_j}(\lambda_{j,i}) = \frac{i}{N}$  for any  $i = 1, \dots, N$ ,  $j = 1, \dots, p$ . Hence, we have

$$\frac{1}{N} \text{Tr}[A_1^{n_1} \circ \dots \circ A_p^{n_p}] = \int_{\mathbb{R}^p} \lambda_1^{n_1} \dots \lambda_p^{n_p} f_N((\lambda_j, \Lambda_j, U_j)_{j=1, \dots, p}) d\mu_{\Lambda_1}(\lambda_1) \dots d\mu_{\Lambda_p}(\lambda_p),$$

where  $f_N((\lambda_j, \Lambda_j, U_j)_{j=1, \dots, p}) = \left( N^{p-1} \sum_{k=1}^N \prod_{j=1}^p |u_{j, (NF_{\Lambda_j}(\lambda_j))}(k)|^2 \right)$ . Hence, a family of random variables  $(a_1, \dots, a_p)$  as in the proposition exists and its joint distribution has density  $f_N((\cdot, \Lambda_j, U_j)_{j=1, \dots, p})$  with respect to  $\mu_{\Lambda_1} \otimes \dots \otimes \mu_{\Lambda_p}$ .

*Step 2.* We now assume that  $(A_1, \dots, A_p)$  are random and that their joint distributions have a density with respect to the Lebesgue measure on  $\mathcal{H}_N^p$ , where  $\mathcal{H}_N$  is the space of Hermitian matrices of size  $N$ . In particular, the matrices have almost surely  $N$  distinct eigenvalues, see [18]. The spectral decompositions of the previous step are measurable (see [18, Section 5.3]) and, with the above notations  $(\Lambda_j, U_j)$  for eigenvalues and eigenvectors of  $A_j$ , we can write the joint distribution of  $(A_1, \dots, A_p)$  in the form  $g_N((\Lambda_j, U_j)_{j=1, \dots, p}) d\mu_{\Delta_N}(\Lambda_1) \dots d\mu_{\Delta_N}(\Lambda_p) d\mu_{\mathcal{U}_N}(U_1) \dots d\mu_{\mathcal{U}_N}(U_p)$ . The symbol  $\mu_{\Delta_N}$  denotes



the Lebesgue measure on  $\Delta_N = \{(x_1, \dots, x_N) | x_1 < \dots < x_N\}$  and  $\mu_{\mathcal{U}_N}$  is the Haar measure on the set  $\mathcal{U}_N$  of unitary matrices of size  $N$ . For any  $n_1, \dots, n_p \geq 0$ , one has

$$\begin{aligned} & \mathbb{E} \left[ \frac{1}{N} \operatorname{Tr} [A_1^{n_1} \circ \dots \circ A_p^{n_p}] \right] \\ &= \int_{\Delta_N^p \times \mathcal{U}_N^p} \frac{1}{N^p} \sum_{i_1, \dots, i_p=1}^N \lambda_{1,i}^{n_1} \dots \lambda_{p,i}^{n_p} f_N((\lambda_{j,i}, \Lambda_j, U_j)_{j=1, \dots, p}) \\ & \quad \times g_N((\Lambda_j, U_j)_{j=1, \dots, p}) d\mu_{\Delta_N}(\Lambda_1) \dots d\mu_{\Delta_N}(\Lambda_p) d\mu_{\mathcal{U}_N}(U_1) \dots d\mu_{\mathcal{U}_N}(U_p) \\ &= \frac{1}{N^p} \sum_{i_1, \dots, i_p=1}^N \int_{\Delta_N^p} \lambda_{1,i}^{n_1} \dots \lambda_{p,i}^{n_p} h_N((\lambda_{j,i}, \Lambda_j)_{j=1, \dots, p}) d\mu_{\Delta_N}(\Lambda_1) \dots d\mu_{\Delta_N}(\Lambda_p), \end{aligned}$$

where  $f_N$  is as in the previous step and

$$\begin{aligned} & h_N((\lambda_{j,i}, \Lambda_j)_{j=1, \dots, p}) \\ &= \int_{\mathcal{U}_N^p} f_N((\lambda_{j,i}, \Lambda_j, U_j)_{j=1, \dots, p}) g_N((\Lambda_j, U_j)_{j=1, \dots, p}) d\mu_{\mathcal{U}_N}(U_1), \dots, d\mu_{\mathcal{U}_N}(U_p). \end{aligned}$$

For any  $i_1, \dots, i_p = 1, \dots, N$ , we have

$$\begin{aligned} & \int_{\Delta_N^p} \lambda_{1,i}^{n_1} \dots \lambda_{p,i}^{n_p} h_N((\lambda_{j,i}, \Lambda_j)_{j=1, \dots, p}) d\mu_{\Delta_N}(\Lambda_1), \dots, d\mu_{\Delta_N}(\Lambda_p) \\ &= \int_{\mathbb{R}^p} \lambda_1^{n_1} \dots \lambda_p^{n_p} h_N^{(i_1, \dots, i_p)}(\lambda_1, \dots, \lambda_p) d\lambda_1, \dots, d\lambda_p, \end{aligned}$$

where  $h_N^{(i_1, \dots, i_p)}(\lambda_1, \dots, \lambda_p)$  is obtained by integrating with respect to the variables  $\lambda_{k_1}, \dots, \lambda_{k_p}$  for  $k_1 \neq i_1, \dots, k_p \neq i_p$ . We finally obtain

$$\mathbb{E} \left[ \frac{1}{N} \operatorname{Tr} [A_1^{n_1} \circ \dots \circ A_p^{n_p}] \right] = \int_{\mathbb{R}^p} \lambda_1^{n_1}, \dots, \lambda_p^{n_p} \bar{h}_N(\lambda_1, \dots, \lambda_p) d\lambda_1, \dots, d\lambda_p,$$

where  $\bar{h}_N = \frac{1}{N^p} \sum_{i_1, \dots, i_p} h_N^{(i_1, \dots, i_p)}$ . Hence, a family of random variables  $(a_1, \dots, a_p)$  as in the proposition exists and its joint distribution has density  $\bar{h}_N$  with respect to the Lebesgue measure on  $\mathbb{R}^p$ .

*Step 3.* We now consider the general case. Let  $(X_1, \dots, X_p)$  be a family of independent random matrices, independent of  $(A_1, \dots, A_p)$ , distributed according to the standard Gaussian measure on  $\mathcal{H}_N$  with respect to the inner product  $\langle A, B \rangle = N \operatorname{Tr}[AB]$ . By the regularizing process of convolution on Hermitian space, for any  $\varepsilon > 0$ , the joint distribution of  $(A_1^\varepsilon, \dots, A_p^\varepsilon) = (A_1 + \varepsilon X_1, \dots, A_p + \varepsilon X_p)$  has a density with respect to the Lebesgue measure. By the previous step, there exists a family of random variables  $(a_1^\varepsilon, \dots, a_p^\varepsilon)$  such that  $\mathbb{E}[(a_1^\varepsilon)^{n_1} \dots (a_p^\varepsilon)^{n_p}] = \mathbb{E}[\frac{1}{N} \operatorname{Tr} [(A_1^\varepsilon)^{n_1} \circ \dots \circ (A_p^\varepsilon)^{n_p}]]$  for any  $n_1, \dots, n_p \geq 0$ . As  $\varepsilon$

goes to zero,  $(a_1^\varepsilon, \dots, a_p^\varepsilon)$  converges in moments to a family of random variables as in the proposition.  $\square$

## REFERENCES

- [1] G. Anderson, A. Guionnet, O. Zeitouni *An Introduction to Random Matrices*. Cambridge studies in advanced mathematics, 118 (2009).
- [2] G. Anderson, O. Zeitouni *A CLT for a band matrix model*, Probab. Theory Rel. Fields, 2005, **134**, 283–338
- [3] Auffinger, Antonio; Ben Arous, Gérard; Péché, Sandrine *Poisson convergence for the largest eigenvalues of heavy tailed random matrices*. Ann. Inst. Henri Poincaré Probab. Stat. 45 (2009), no. 3, 589–610.
- [4] Z.D. Bai, J.W. Silverstein *Spectral analysis of large dimensional random matrices*. Second Edition, Springer, New York, 2009.
- [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, X. Wang, W. Zhou *CLT for linear spectral statistics of Wigner matrices*. Electron. J. Probab. **14** (2009), no. 83, 2391–2417.
- [7] Z.D. Bai, J. Yao *On the convergence of the spectral empirical process of Wigner matrices*. Bernoulli **11** (2005) 1059–1092.
- [8] S. Belinschi, A. Dembo, A. Guionnet *Spectral measure of heavy tailed band and covariance random-matrices*, Comm. Math. Phys., **289**, 2009, 1023–1055.
- [9] G. Ben Arous, K. Dang *On fluctuations of eigenvalues of random permutation matrices*, arXiv:1106.2108, preprint.
- [10] G. Ben Arous, A. Guionnet *The spectrum of heavy tailed random matrices*. Comm. Math. Phys. **278** (2008), no. 3, 715–751.
- [11] P. Billingsley *Probability and measure*, Wiley, third edition.
- [12] C. Bordenave, A. Guionnet *Localization and delocalization of eigenvectors for heavy-tailed random matrices*, arxiv.
- [13] C. Bordenave, P. Caputo, D. Chafaï *Spectrum of large random reversible Markov chains: heavy-tailed weights on the complete graph*, Ann. Probab., **39**, 2011, 1544–1590.
- [14] C. Bordenave, P. Caputo, D. Chafaï *Spectrum of non-Hermitian heavy tailed random matrices*, Comm. Math. Phys., **307**, 2011, 513–560.
- [15] P. Cizeau, J.-P. Bouchaud, *Theory of Lévy matrices* Phys. Rev. E 50 (1994).
- [16] T. Cabanal-Duvillard *Fluctuations de la loi empirique de grande matrices alatoires* Ann. Inst. H. Poincaré Probab. Stat. **37**, 2001, 373–402.
- [17] S. Chatterjee *Fluctuations of eigenvalues and second order Poincaré inequalities*, Probab. Theory Related Fields, **143**, 2009,1–40.
- [18] P. Deift *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*, Courant Lecture Notes in Mathematics, 3, New York, 1999.
- [19] P. Diaconis, S. Evans *Linear functionals of eigenvalues of random matrices*, Trans. Amer. Math. Soc., **353**, 2001, 2615–2633.
- [20] I. Dumitriu, T. Johnson, S. Pal, E. Paquette *Functional limit theorems for random regular graphs*, , Probab. Theory Related Fields, 2012,1–55.
- [21] L. Erdős, B. Schlein, H.T. Yau *Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices*, Ann. Prob. 37 (2009).
- [22] L. Erdős, H.T. Yau, J. Yin *Rigidity of eigenvalues of generalized Wigner matrices*, Adv. Math., **229**, 2012, 1435–1515.

- [23] W. Feller *An introduction to probability theory and its applications*, volume II, second edition, New York London Sydney : J. Wiley, 1966.
- [24] A. Guionnet, *Large random matrices: lectures on macroscopic asymptotics*, Lecture Notes in Mathematics, 1957, Lectures from the 36th Probability Summer School held in Saint-Flour, 2006, Springer-Verlag, Berlin, 2009.
- [25] A. Guionnet *Large deviations upper bounds and central limit theorems for non-commutative functionals of Gaussian large random matrices*. Ann. Inst. H. Poincaré Probab. Stat. **38**, 2002, 341–384.
- [26] A. Guionnet, E. Maurel-Segala *Second order asymptotics for matrix models*, Ann. Probab., **35**, 2007, 2160–2212.
- [27] K. Johansson *On the fluctuations of eigenvalues of random Hermitian matrices*. Duke Math. J. **91** 1998, 151–204.
- [28] D. Jonssohn *Some limit theorems for the eigenvalues of a sample covariance matrix*. J. Mult. Anal. **12**, 1982, 1–38.
- [29] A. Lytova, L. Pastur *Central limit theorem for linear eigenvalue statistics of random matrices with independent entries*, Ann. Probab., **37**, 2009, 1778–1840.
- [30] C. Male *The distribution of traffics and their free product*, arXiv:1111.4662v3 preprint.
- [31] C. Male *The limiting distributions of large heavy Wigner and arbitrary random matrices*, arXiv:1111.4662v3 preprint.
- [32] 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, 226–270.
- [33] G. Samorodnitsky, M. Taqqu *Stable non-Gaussian random processes. Stochastic models with infinite variance. Stochastic Modeling*. Chapman & Hall, New York, 1994.
- [34] M. Shcherbina, B. Tirozzi, *Central limit theorem for fluctuations of linear eigenvalue statistics of large random graphs*, J. Math. Phys., **51**, 2010, 023523, 20
- [35] 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.
- [36] T. Tao, V. Vu *Random matrices: universality of local eigenvalue statistics*, Acta Mathematica, 206 (2011), 127–204.
- [37] E.P. Wigner *On the distribution of the roots of certain symmetric matrices*, Annals Math. , **67**, 1958, 325–327.
- [38] I. Zakharevich *A generalization of Wigner’s law*, Comm. Math. Phys., 268, 2006, 2, 403–414.