Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 126 (2016) 3331–3352
stochastic processes and their applications
www.elsevier.com/locate/spa

Fluctuations of linear statistics of half-heavy-tailed random matrices

Florent Benaych-Georges[a,*], Anna Maltsev[b]

[a] *MAP5, Université Paris Descartes, 45, rue des Saints-Pères 75270 Paris Cedex 06, France*
[b] *Howard House, Departments of Mathematics, University of Bristol, Bristol BS81SN, UK*

Received 21 October 2014; received in revised form 12 April 2016; accepted 27 April 2016
Available online 13 May 2016

Abstract

In this paper, we consider a Wigner matrix A with entries whose cumulative distribution decays as $x^{-\alpha}$ with $2 < \alpha < 4$ for large x. We are interested in the fluctuations of the linear statistics $N^{-1} \operatorname{Tr} \varphi(A)$, for some nice test functions φ. The behavior of such fluctuations has been understood for both heavy-tailed matrices (i.e. $\alpha < 2$) in Benaych-Georges (2014) and light-tailed matrices (i.e. $\alpha > 4$) in Bai and Silverstein (2009). This paper fills in the gap of understanding it for $2 < \alpha < 4$. We find that while linear spectral statistics for heavy-tailed matrices have fluctuations of order $N^{-1/2}$ and those for light-tailed matrices have fluctuations of order N^{-1}, the linear spectral statistics for half-heavy-tailed matrices exhibit an intermediate α-dependent order of $N^{-\alpha/4}$.

MSC: 15A52; 60F05

Keywords: Random matrices; Heavy tailed random variables; Central limit theorem

1. Introduction

Let $A = [a_{ij}]$ be an $N \times N$ Hermitian random matrix whose entries are i.i.d. and let $\lambda_1, \ldots, \lambda_N$ be its eigenvalues. It is well known that if the entries of A are duly renormalized,

* Corresponding author.
E-mail addresses: florent.benaych-georges@parisdescartes.fr (F. Benaych-Georges), annavmaltsev@gmail.com (A. Maltsev).

http://dx.doi.org/10.1016/j.spa.2016.04.030

Table 1
Orders of the fluctuations of the r.v. of (1) around its expectation as a function of the exponent α such that $\mathbb{P}(|a_{ij}| > x) \approx x^{-\alpha}$ for x large.

	$\alpha < 2$	$2 < \alpha < 4$	$\alpha > 4$
Order of the fluctuations of (1)	$N^{-1/2}$	$N^{-\alpha/4}$	N^{-1}

then for any continuous bounded test function φ, the random variable

$$\frac{1}{N}\operatorname{Tr}\varphi(A) = \frac{1}{N}\sum_{i=1}^{N}\varphi(\lambda_i) \tag{1}$$

has a deterministic limit, which is equal to the integral of f with respect to the limit spectral distribution of A, namely the semicircle law when the entries have at least a second moment [2,1] and different distributions depending on α if the entries are heavy-tailed with exponent $\alpha \in (0, 2)$ (see [5,14,9]). The rate of convergence of the random variables of (1) to its limit is not usually $\frac{1}{\sqrt{N}}$, as i.i.d. λ_i's would give. In particular, if the entries of A have a fourth moment, then the fluctuations of $\frac{1}{N}\operatorname{Tr}\varphi(A)$ around its expectation have order $\frac{1}{N}$ (see [2,19,24,18,4,3,21,23]). On the other hand, if the entries are heavy-tailed with exponent $\alpha \in (0, 2)$ or Bernoulli with parameter of order N^{-1}, then the fluctuations of $\frac{1}{N}\operatorname{Tr}\varphi(A)$ around its expectation have order $N^{-1/2}$ [6]. This difference of order in the fluctuations is due to the fact that when the entries of A have enough moments, the eigenvalues of A fluctuate very little, as studied by Erdös, Schlein, Yau, Tao, Vu and their co-authors, who analyzed their rigidity in e.g. [15,16,25]. On the other hand, the heavier the tails the more similar to a sparse matrix the (renormalized) matrix A is, and the more independently its eigenvalues behave.

A finite fourth moment means that for large x, $\mathbb{P}(|a_{ij}| > x) \approx x^{-\alpha}$ with $\alpha > 4$, whereas heavy-tailed entries with exponent $\alpha \in (0, 2)$ correspond precisely to $\mathbb{P}(|a_{ij}| > x) \approx x^{-\alpha}$ with $\alpha \in (0, 2)$. In this text, we fill in the gap of understanding the role of α in the fluctuations linear spectral statistics: when $\alpha \in (2, 4)$, we prove a central limit theorem for $\frac{1}{N}\operatorname{Tr}\varphi(A)$ in the case where φ is a sum of resolvent functions, it appears that the order of the fluctuations, in this case, is $N^{-\alpha/4}$. This completes the picture, summarized in Table 1. Viewed in the light of concentration inequalities for linear spectral functionals of random matrices, random matrices with half-heavy tailed entries interpolate between two extreme regimes, as shown in Table 1:

- Using only the independence of the entries, (see [11, Lem. C.1] or [22,9,10]) for any bounded function $\varphi : \mathbb{R} \to \mathbb{R}$ with finite total variation, we have, for any $\delta > 0$,

$$\mathbb{P}\left(|\operatorname{Tr}\varphi(A) - \mathrm{E}\operatorname{Tr}\varphi(A)| \ge \delta\right) \ \le\ Ce^{-c\frac{\delta^2}{N\|\varphi\|_{\mathrm{TV}}^2}}, \tag{2}$$

  which proves that

$$\sqrt{N}\left(\frac{1}{N}\operatorname{Tr}\varphi(A) - \mathrm{E}[\frac{1}{N}\operatorname{Tr}\varphi(A)]\right)$$

  is bounded in probability and explains why the order of the fluctuations of (1) cannot be larger than $N^{-1/2}$,
- In the case where the entries of A are independent and satisfy a Log-Sobolev inequality (for example in the GO(U)E case), then for any Lipschitz function $\varphi : \mathbb{R} \to \mathbb{R}$, we have, by

[1, Th. 2.3.5] (see also [17,20,12]),

$$\mathbb{P}\left(|\mathrm{Tr}\,\varphi(A)-\mathrm{E}\,\mathrm{Tr}\,\varphi(A)|\geq\delta\right)\ \leq\ Ce^{-c\frac{\delta^2}{\|\varphi'\|_\infty^2}}, \tag{3}$$

which proves that

$$N\left(\frac{1}{N}\mathrm{Tr}\,\varphi(A)-\mathrm{E}[\frac{1}{N}\mathrm{Tr}\,\varphi(A)]\right)$$

is bounded in probability. It explains why the order of the fluctuations of functionals as (1) cannot be larger than N^{-1} in the case of matrices with independent Log-Sobolev entries.

Eq. (2) shows that the case $\alpha<2$ corresponds to the largest possible fluctuations order in (1). On the other hand, (3) proves that in the Gaussian case (this has been extended by Bai et al. to the case $\alpha\geq 4$), the actual order is N^{-1} (to be more precise, (3) only gives an upper-bound for this order, but one can easily check, using, for example, $\varphi(\lambda)=\lambda$ or $\varphi(\lambda)=\lambda^2$, that N^{-1} is actually the right order). The case $2<\alpha<4$ is an intermediate case, where concentration inequalities neither allow to guess the order of the fluctuations, nor allow to extend fluctuation results from a first class of test functions to a wider classes (as was done for example in [6]).

Notation. Here, $A\ll B$ means that $A/B\longrightarrow 0$ as $N\to\infty$, $A\sim B$ means that $A/B\longrightarrow 1$ as $N\to\infty$ (or $x\to\infty$ if x is the underlying variable) and $A=O(B)$ means that A/B stays bounded as $N\to\infty$. If A and B are random variables, similar notations are used in probability. Also, for $\alpha\in\mathbb{R}$, the function $z\mapsto z^\alpha$ is defined on $\mathbb{C}\setminus\mathbb{R}^-$ thanks to the determination of the argument which is null on $(0,+\infty)$. At last, for functions of a complex variable $z=x+\mathrm{i}y$, we use the classical notation $\partial_z=(\partial_x-\mathrm{i}\partial_y)/2$.

2. Main result

Let us consider a random real symmetric or Hermitian matrix

$$A=[a_{ij}]_{1\leq i,j\leq N}=\left[\frac{x_{ij}}{\sqrt{N}}\right]_{1\leq i,j\leq N},$$

where one of two conditions holds: either

(a) (real case) x_{ij}'s, $1\leq i\leq j$, are i.i.d. real random variables with mean 0 and variance 1 such that for a certain $\alpha\in(2,4)$ and a certain $c>0$, as $x\to+\infty$,

$$\mathbb{P}(|x_{ij}|>x)\sim\frac{c}{\Gamma(\alpha+1)}x^{-\alpha}, \tag{4}$$

or

(b) (complex case) $x_{ij}=x_{ij}^R/\sqrt{2}+\mathrm{i}x_{ij}^I/\sqrt{2}$ for $1<i<j$ and $x_{ii}=x_{ii}^R$ where x_{ij}^I and x_{ij}^R are i.i.d. real symmetric random variables with mean 0 and variance 1 that satisfy (4).

Our main theorem is the following.

Theorem 2.1. *For*

$$G(z):=(z-A)^{-1}$$

with A as above the process

$$\left(\frac{1}{N^{1-\alpha/4}}(\mathrm{Tr}\,G(z)-\mathrm{E}\,\mathrm{Tr}\,G(z))\right)_{z\in\mathbb{C}\setminus\mathbb{R}}$$

converges to a complex Gaussian centered process $(X_z)_{z\in\mathbb{C}\setminus\mathbb{R}}$ with covariance defined by the fact that $X_{\overline{z}} = \overline{X_z}$ and that for any $z, z' \in \mathbb{C}\setminus\mathbb{R}$, $\mathrm{E}[X_z X_{z'}] = C(z, z')$, for

$$C(z,z') := -\iint_{t,t'>0} \partial_z \partial_{z'} \Big\{ [(K(z,t) + K(z',t'))^{\alpha/2} - (K(z,t)^{\alpha/2} + K(z',t')^{\alpha/2})] \times \exp\left(\mathrm{sgn}_z \mathrm{i} t z - K(z,t) + \mathrm{sgn}_{z'} \mathrm{i} t' z' - K(z',t')\right) \Big\} \frac{c\,\mathrm{d}t\,\mathrm{d}t'}{2tt'}$$

where c and α are as in (4), $\mathrm{sgn}_z := \mathrm{sgn}(\Im z)$ and $K(z,t) := \mathrm{sgn}_z \mathrm{i} t G_{\mathrm{sc}}(z)$, $G_{\mathrm{sc}}(z)$ being the Stieltjes transform of the semicircle law with support $[-2, 2]$.

Remark 2.2. This theorem proves Gaussian convergence for any random variable of the form

$$\frac{1}{N^{1-\alpha/4}}(\mathrm{Tr}\,\varphi(A) - \mathrm{E}\,\mathrm{Tr}\,\varphi(A)),$$

where φ is a function of the type

$$\varphi(\lambda) = \sum_{j=1}^{p} \frac{c_j}{z_j - \lambda}, \quad (p \geq 1,\ c_1, \ldots, c_p \in \mathbb{C},\ z_1, \ldots, z_p \in \mathbb{C}\setminus\mathbb{R}).$$

The functions φ of this type span (by closure) some larger sets of functions (by, e.g. the Stone–Weierstrass theorem, the Cauchy formula, or the Helffer–Sjöstrand formula). However, the lack of error control in approximating φ (due to the fact that in our case the Log-Sobolev concentration inequality (3) is not true and the general concentration inequality (2) is not sharp enough) prevents us from extending our theorem to a larger class of test functions, as was done from resolvent functions to wider classes in e.g. [2,6,23]. As far as applying [23, Propo. 1] is concerned, the problems come first from the fact that we truncate the entries of the matrices to upper-bound the variance of $\mathrm{Tr}(E + \mathrm{i}\eta - A)^{-1}$ and second from the fact that this variance does not decay enough as $|E|$ grows.

The remainder of the paper consists of the proof of Theorem 2.1. In Section 3.1, we truncate the random variables appropriately, and centralize in the real case (in the complex case, centralization is automatic due to our assumption of symmetry). In Section 3.2, we restate our problem in terms of a martingale approach and cite relevant martingale convergence theorem. In Section 3.3, we show that off-diagonal terms of the resolvent can be neglected in further calculations. Lastly, in Section 3.4 we show that the diagonal terms of the resolvent yield the desired formula for the covariance, using a lemma proved in Section 3.5 that allows us to approximate the diagonal elements of the resolvent by the Stieltjes transform of the spectral measure.

3. Proof of Theorem 2.1

3.1. Truncation, recentralization, and renormalization of the entries

Let A, B be any $N \times N$ matrices, $z \in \mathbb{C}\setminus\mathbb{R}$, and

$$G_B(z) := \frac{1}{z - B}. \tag{5}$$

Then

$$G_B(z) = \frac{1}{z - A} + \frac{1}{z - A}(B - A)\frac{1}{z - B} \tag{6}$$

and we have that

$$|\operatorname{Tr}(G_B(z) - G_A(z))| \le 2|\Im z|^{-1} \operatorname{rank}(B - A).$$

Thus for fixed z,

$$\operatorname{rank}(B - A) \ll N^{1-\alpha/4} \Longrightarrow |\operatorname{Tr}(G_B(z) - G_A(z))| \ll N^{1-\alpha/4}.$$

Let us consider the case of real symmetric matrices and let $\mu_N = \mathrm{E}\, x_{ij} \mathbb{1}_{|x_{ij}| \le N^\beta}$ (for an exponent β which will be specified later). First we estimate the absolute value of μ_N. As x_{ij} is centered,

$$\mu_N := \mathrm{E}\, x_{ij} \mathbb{1}_{|x_{ij}| > N^\beta}, \tag{7}$$

so for N large enough,

$$0 \le |\mu_N| \le \mathrm{E}\, |x_{ij}| \mathbb{1}_{|x_{ij}| \ge N^\beta} = \int_{N^\beta}^{\infty} \mathbb{P}(|x_{ij}| > x) \mathrm{d}x \le \frac{2c}{\Gamma(\alpha+1)(\alpha-1)} (N^\beta)^{-\alpha+1}. \tag{8}$$

Let $B = [a_{ij} \mathbb{1}_{|x_{ij}| \le N^\beta} - \mu_N/\sqrt{N}]$. Subtracting $\mu_N/\sqrt{N}$ from each matrix entry is a rank 1 perturbation. Then, as

$$\mathbb{P}(|x_{ij}| > N^\beta) \le C N^{-\alpha\beta},$$

we have

$$\operatorname{rank}(B - A) \le 1 + 2 \sum_{i=1}^{N} X_i$$

where the X_i's are independent Bernoulli r.v. with parameters

$$\mathbb{P}(X_i = 1) = 1 - (1 - C N^{-\alpha\beta})^i$$

and 1 is added for the rank 1 perturbation of shifting each entry by $\mu_N/\sqrt{N}$. In order to upper-bound $\operatorname{rank}(B - A)$ with high probability thanks to Bennett's inequality or Lemma 5.7 in [7], we compute the sum of these parameters:

$$\sum_{i=1}^{N} 1 - (1 - C N^{-\alpha\beta})^i = N - (1 - C N^{-\alpha\beta}) \frac{1 - (1 - C N^{-\alpha\beta})^N}{C N^{-\alpha\beta}}.$$

If $\alpha\beta > 1$,

$$\begin{aligned} (1 - C N^{-\alpha\beta})^N &= \exp N \log(1 - C N^{-\alpha\beta}) \\ &= 1 - C N^{1-\alpha\beta} + \frac{C^2}{2} N^{2(1-\alpha\beta)} + o(N^{2(1-\alpha\beta)}) \end{aligned}$$

so

$$\begin{aligned} \sum_{i=1}^{N} 1 - (1 - C N^{-\alpha\beta})^i &= N - (1 - C N^{-\alpha\beta}) \frac{C N^{1-\alpha\beta} - \frac{C^2}{2} N^{2(1-\alpha\beta)} + o(N^{2(1-\alpha\beta)})}{C N^{-\alpha\beta}} \\ &= N - (1 - C N^{-\alpha\beta})(N - \frac{C}{2} N^{2-\alpha\beta} + o(N^{2-\alpha\beta})) \\ &= \frac{C}{2} N^{2-\alpha\beta} + o(N^{2-\alpha\beta}). \end{aligned}$$

Thus, by Bennett's inequality or Lemma 5.7 in [7], as soon as $\alpha\beta > 1$, we know that $\operatorname{rank}(B-A)$ has order at most $N^{(2-\alpha\beta)_+}$ (*i.e.* for any $\varepsilon > 0$, $N^{-(2-\alpha\beta)_+-\varepsilon}\operatorname{rank}(B-A)$ tends in probability to zero). So one can replace A by B as long as

$$(2-\alpha\beta)_+ < 1-\alpha/4,$$

i.e.

$$\beta > \frac{2-(1-\alpha/4)}{\alpha} = \frac{1}{4}\left(1+\frac{\alpha}{4}\right).$$

Furthermore we want to renormalize our new truncated centered random variables to have variance 1, so we let

$$\sigma_N := \sqrt{\mathrm{E}(x_{ij}\mathbb{1}_{|x_{ij}|<N^\beta} - \mu_N)^2}. \tag{9}$$

Noting that

$$\mathrm{E}(\mathbb{1}_{|x_{ij}|>N^\beta}x_{ij}^2) = \int_{N^{2\beta}}^{\infty}\mathbb{P}(|x_{ij}|^2 > x)\mathrm{d}x \sim \frac{c}{\Gamma(\alpha+1)(1-\alpha/2)N^{\beta(\alpha-2)}},$$

we see that

$$\sigma_N^2 - 1 = O(N^{\beta(2-\alpha)}).$$

Then we can replace B as above by B/σ_N since

$$|\operatorname{Tr}(G_B(z) - G_{B/\sigma_N}(z))| = O(N^{\beta(2-\alpha)}).$$

In the complex case, subtracting the mean from each matrix entry is no longer a rank 1 perturbation, so this argument will no longer work. This is the reason why, in the complex case, we only consider random variables which are symmetric so that we can truncate and still retain a 0 mean.

Lemma 3.1. *Let* $\epsilon > 0$, $\beta = \frac{1}{4}\left(1+\frac{\alpha}{4}\right)+\epsilon$ *and*

$$\tilde{a}_{ij} := (x_{ij}\mathbb{1}_{|x_{ij}|<N^\beta} - \mu_N)/\left(\sigma_N\sqrt{N}\right)$$

in the real case and

$$\tilde{a}_{ij} := (x_{ij}^R\mathbb{1}_{|x_{ij}^R|<N^\beta} + \mathrm{i}x_{ij}^I\mathbb{1}_{|x_{ij}^I|<N^\beta})/\left(\sigma_N\sqrt{N}\right)$$

in the complex case. Then there is a constant C depending only on the distribution of the x_{ij}'s such that

(i) *the $\tilde{a}_{i,j}$'s are i.i.d., centered, with variance $1/N$,*

(ii) $N^{3/2}\,\mathrm{E}[|\tilde{a}_{i,j}|^3] \le CN^{\left(\frac{1}{4}(1+\frac{\alpha}{4})+\epsilon\right)(3-\alpha)_+}$,

(iii) $N^2\,\mathrm{E}[|\tilde{a}_{i,j}|^4] \le CN^{\beta(4-\alpha)} = CN^{1-\frac{\alpha^2}{16}+\epsilon(4-\alpha)}$,

(iv) *for any $\lambda\in\mathbb{C}$ such that $\Im\lambda \le 0$,*

$$\phi_N(\lambda) := \mathrm{E}[e^{-\mathrm{i}\lambda|\tilde{a}_{ij}|^2}] = 1 - \frac{\mathrm{i}\lambda}{N} - c\frac{(\mathrm{i}\lambda)^{\frac{\alpha}{2}}}{N^{\frac{\alpha}{2}}} + \frac{|\lambda|^{\frac{\alpha}{2}}}{N^{\frac{\alpha}{2}}}\varepsilon_N(\mathrm{i}\lambda/N), \tag{10}$$

where the function $\varepsilon_N(z)$ is analytic in z on $\{\Re(z) > 0\}$, bounded uniformly in $(N,z)\in\mathbb{N}\times\mathcal{K}$ for any compact $\mathcal{K}\subset\{\Re(z)\ge 0\}$ and $\lim_{z\to 0}\varepsilon_N(z) = 0$ uniformly in N.

Proof. (i) is true by the definition of μ_N and σ_N at (7) and (9), (ii) and (iii) are easy computations, relying on the fact that for $f : \mathbb{R}^+ \to \mathbb{R}^+$ increasing and X a positive random variable, $\mathrm{E}[f(X)] = f(0) + \int_0^{+\infty} f'(t)\mathbb{P}(X \geq t)\mathrm{d}t$. Lastly, (iv) follows from [13, Th. 8.1.6] for non-truncated entries. To centralize, note that

$$\mathbb{P}(|x_{ij} - \mu_N| \geq x) = \mathbb{P}(x_{ij} \geq x + \mu_N) + \mathbb{P}(x_{ij} \leq -x + \mu_N \text{ and } x_{ij} \leq \mu_N).$$

. Recalling that $\mu_N \to 0$ we get from (4) that for a certain constant $C > 0$,

$$\frac{c}{C\Gamma(\alpha+1)}(x-\mu_N)^{-\alpha} \leq \mathbb{P}(|x_{ij} - \mu_N| > x) \leq \frac{Cc}{\Gamma(\alpha+1)}(x+\mu_N)^{-\alpha}$$

Since $\mu_N \to 0$, the random variable $x_{ij} - \mu_N$ will also satisfy (4) for large N, and therefore (iv) holds for the shifted entry $x_{ij} - \mu_N$. □

From now on, we suppose that each a_{ij} has been replaced by the $\tilde{a}_{ij}$ of the previous lemma, for

$$\beta := \frac{1}{4}\left(1 + \frac{\alpha}{4}\right) + \epsilon, \tag{11}$$

for $\epsilon > 0$ that can be chosen as small as needed. By a slight abuse of notation, we still denote this random variable by a_{ij} and we henceforth assume the conclusions of Lemma 3.1 to be true for the a_{ij}'s.

3.2. Martingale approach

We want to prove that for certain test functions φ (namely the linear combinations of functions of the type $\lambda \in \mathbb{R} \longmapsto \frac{1}{z-\lambda}$, with $z \in \mathbb{C} \setminus \mathbb{R}$),

$$M(\varphi, N) := \frac{1}{N^{1-\alpha/4}}(\operatorname{Tr}\varphi(A) - \mathrm{E}[\operatorname{Tr}\varphi(A)])$$

converges in distribution to a certain Gaussian distribution. We will use Theorem A.3 for $M(\varphi, N)$, with $M_N(N) = M(\varphi, N)$ and $\mathcal{F}_k(N) := \sigma(x_{i,j}; i \leq k \text{ and } j \leq k)$.

Then, denoting $\mathrm{E}[\,\cdot\,|\mathcal{F}_k]$ by E_k, the random variable $Y_k(N)$ of Theorem A.3 is

$$Y_k = Y_k(N) = \frac{1}{N^{1-\alpha/4}}(\mathrm{E}_k - \mathrm{E}_{k-1})(\operatorname{Tr}\varphi(A)).$$

Let $A^{(k)}$ be the $N-1$ by $N-1$ matrix obtained by removing the kth row and column of A. Then $(\mathrm{E}_k - \mathrm{E}_{k-1})(\operatorname{Tr}\varphi(A^{(k)})) = 0$, hence

$$Y_k = \frac{1}{N^{1-\alpha/4}}(\mathrm{E}_k - \mathrm{E}_{k-1})(\operatorname{Tr}\varphi(A) - \operatorname{Tr}\varphi(A^{(k)})).$$

Note first that by the interlacing property between the spectrums of A and $A^{(k)}$, when φ has finite total variation, we have

$$|\operatorname{Tr}\varphi(A) - \operatorname{Tr}\varphi(A^{(k)})| \leq \|\varphi\|_{\mathrm{TV}}. \tag{12}$$

As a consequence, $|Y_k| \leq \frac{\|\varphi\|_{\mathrm{TV}}}{N^{1-\alpha/4}}$ and the $L(\varepsilon, N)$ of Theorem A.3 is null for N large enough.

Hence it remains to prove that

$$\sum_{k=1}^{N} \mathrm{E}_{k-1}(Y_k^2)$$

and

$$\sum_{k=1}^{N} \mathrm{E}_{k-1}(|Y_k|^2)$$

have finite deterministic limits that agree with the limit covariance of Theorem 2.1.

By linear combination, it suffices to prove that for any $z, z' \in \mathbb{C} \setminus \mathbb{R}$, for

$$Y_k := \frac{1}{N^{1-\alpha/4}}(\mathrm{E}_k - \mathrm{E}_{k-1}) \left(\mathrm{Tr} \frac{1}{z - A} - \mathrm{Tr} \frac{1}{z - A^{(k)}} \right)$$

and

$$Y_k' := \frac{1}{N^{1-\alpha/4}}(\mathrm{E}_k - \mathrm{E}_{k-1}) \left(\mathrm{Tr} \frac{1}{z' - A} - \mathrm{Tr} \frac{1}{z' - A^{(k)}} \right),$$

we have

$$\sum_{k=1}^{N} \mathrm{E}_{k-1}(Y_k Y_k')$$

converges in probability to a deterministic constant which agrees with the limit covariance of Theorem 2.1.

Note that

$$\sum_{k=1}^{N} \mathrm{E}_{k-1}(Y_k Y_k') = \frac{1}{N} \sum_{k=1}^{N} N\, \mathrm{E}_{k-1}(Y_k Y_k') = \int_{u=0}^{1} N\, \mathrm{E}_{\lceil Nu \rceil - 1}(Y_{\lceil Nu \rceil} Y'_{\lceil Nu \rceil}) \mathrm{d}u,$$

hence we shall prove that for any $u \in (0, 1)$, as $N \to \infty$ and $k \to \infty$ with $k/N \to u$, we have

$$N\, \mathrm{E}_{k-1}(Y_k Y_k') \longrightarrow 2uC(z, z'), \tag{13}$$

with $C(z, z')$ the function defined in Theorem 2.1.

Note also that for $G^{(k)}(z) := (z - A^{(k)})^{-1}$, by (31),

$$Y_k = \frac{1}{N^{1-\alpha/4}}(\mathrm{E}_k - \mathrm{E}_{k-1}) \frac{1 + \mathbf{a}_k^*(G^{(k)}(z))^2 \mathbf{a}_k}{z - a_{kk} - \mathbf{a}_k^* G^{(k)}(z) \mathbf{a}_k} \tag{14}$$

where $\mathbf{a}_k$ is the kth column of A without the diagonal term.

3.3. Removing the off-diagonal terms

Proposition 3.2. *Let us define*

$$\tilde{Y}_k := \frac{1}{N^{1-\alpha/4}}(\mathrm{E}_k - \mathrm{E}_{k-1}) \frac{1 + \mathbf{a}_k^*(G^{(k)}(z))^2_{\mathrm{diag}} \mathbf{a}_k}{z - \mathbf{a}_k^* G^{(k)}(z)_{\mathrm{diag}} \mathbf{a}_k}$$

and

$$\tilde{Y}_k' := \frac{1}{N^{1-\alpha/4}}(\mathrm{E}_k - \mathrm{E}_{k-1}) \frac{1 + \mathbf{a}_k^*(G^{(k)}(z'))^2_{\mathrm{diag}} \mathbf{a}_k}{z' - \mathbf{a}_k^* G^{(k)}(z')_{\mathrm{diag}} \mathbf{a}_k},$$

where for a matrix M, M_{diag} denotes the diagonal matrix obtained from M by setting all its non-diagonal entries to zero. Then

$$\sum_{k=1}^{N} \mathrm{E}_{k-1}[Y_k Y_k'] - \mathrm{E}_{k-1}[\tilde{Y}_k \tilde{Y}_k'] \tag{15}$$

converges in probability to 0.

Proof. We define, for $z \in \mathbb{C} \setminus \mathbb{R}$,

$$F_k := \log|z - a_{kk} - \mathbf{a}_k^* G^{(k)}(z)\mathbf{a}_k|^2 \qquad F_k' := \log|z' - a_{kk} - \mathbf{a}_k^* G^{(k)}(z')\mathbf{a}_k|^2$$
$$\tilde{F}_k := \log|z - \mathbf{a}_k^* G^{(k)}(z)_{\mathrm{diag}}\mathbf{a}_k|^2 \qquad \tilde{F}_k' := \log|z' - \mathbf{a}_k^* G^{(k)}(z')_{\mathrm{diag}}\mathbf{a}_k|^2.$$

These functions are well defined because for any $z \in \mathbb{C} \setminus \mathbb{R}$,

$$\Im z \times \Im(-\mathbf{a}_k^* G^{(k)}(z)\mathbf{a}_k) > 0; \qquad \Im z \times \Im(-\mathbf{a}_k^* G^{(k)}(z)_{\mathrm{diag}}\mathbf{a}_k) > 0$$

which implies that

$$\begin{cases} \Im z > 0 \Longrightarrow \Im(z - a_{kk} - \mathbf{a}_k^* G^{(k)}(z)\mathbf{a}_k), \Im(z - a_{kk} - \mathbf{a}_k^* G^{(k)}(z)_{\mathrm{diag}}\mathbf{a}_k) > \Im z \\ \Im z < 0 \Longrightarrow \Im(z - a_{kk} - \mathbf{a}_k^* G^{(k)}(z)\mathbf{a}_k), \Im(z - a_{kk} - \mathbf{a}_k^* G^{(k)}(z)_{\mathrm{diag}}\mathbf{a}_k) < \Im z \end{cases}, \tag{16}$$

so that the argument of the log cannot vanish.

Using the fact that, for an analytic function f defined on $\mathbb{C} \setminus \mathbb{R}$ and taking values in $\mathbb{C} \setminus \mathbb{R}$ such that $f(\overline{z}) = \overline{f(z)}$,

$$\partial_z \log|f(z)|^2 = \partial_z \log(f(z)f(\overline{z})) = \partial_z \log(f(z)) = \frac{f'(z)}{f(z)},$$

we have

$$Y_k = \frac{1}{N^{1-\alpha/4}}(\mathrm{E}_k - \mathrm{E}_{k-1})\partial_z F_k = \partial_z \frac{1}{N^{1-\alpha/4}}(\mathrm{E}_k - \mathrm{E}_{k-1})F_k,$$

where in the second equality, we used (16) to notice the F_k is bounded uniformly in the randomness and in z as z varies in any compact subset of $\mathbb{C} \setminus \mathbb{R}$. Of course analogous formulas hold for Y_k', $\tilde{Y}_k$ and $\tilde{Y}_k'$. Thus, commuting conditional expectation and derivative again for the same reason, we have

$$\sum_{k=1}^{N} \mathrm{E}_{k-1}[Y_k Y_k'] - \mathrm{E}_{k-1}[\tilde{Y}_k \tilde{Y}_k'] = \partial_z \partial_{z'} \sum_{k=1}^{N} N^{-2+\alpha/2}\{\underbrace{\mathrm{E}_{k-1}[(\mathrm{E}_k - \mathrm{E}_{k-1})F_k(\mathrm{E}_k - \mathrm{E}_{k-1})F_k']}_{:=\varphi_k}$$
$$- \underbrace{\mathrm{E}_{k-1}[(\mathrm{E}_k - \mathrm{E}_{k-1})\tilde{F}_k(\mathrm{E}_k - \mathrm{E}_{k-1})\tilde{F}_k']}_{:=\tilde{\varphi}_k}\}$$

hence by the Cauchy inequalities for holomorphic functions, it suffices to prove that uniformly on k, z, z' (as z, z' stay at a macroscopic distance from the real line) we have

$$|\varphi_k - \tilde{\varphi}_k| \ll N^{-1}N^{2-\alpha/2}. \tag{17}$$

Let

$$\eta_k := a_{kk} + \mathbf{a}_k^* G^{(k)}(z)\mathbf{a}_k - \mathbf{a}_k^* G^{(k)}(z)_{\mathrm{diag}}\mathbf{a}_k = a_{kk} + \sum_{i \neq j} G^{(k)}(z)_{ij}\overline{\mathbf{a}_k(i)}\mathbf{a}_k(j),$$

and

$$\varepsilon_k := F_k - \tilde{F}_k = \log|1 - \eta_k(z - \mathbf{a}_k^* G^{(k)}(z)_{\mathrm{diag}}\mathbf{a}_k)^{-1}|^2.$$

We also define η_k' and ε_k' in the same way with z instead of z'.

By Lemma A.4, we have

$$\varphi_k = \mathrm{E}_{k-1}[\mathrm{E}_k(F_k)\,\mathrm{E}_k(F_k')] - \mathrm{E}_{k-1}(F_k)\,\mathrm{E}_{k-1}(F_k')$$

and the analogous equality holds for $\tilde{\varphi}_k$. Thus using the formulas $F_k = \tilde{F}_k + \varepsilon_k$ and $F_k' = \tilde{F}_k' + \varepsilon_k'$, we have

$$\begin{aligned}\varphi_k - \tilde{\varphi}_k &= \mathrm{E}_{k-1}[\mathrm{E}_k(\tilde{F}_k)\,\mathrm{E}_k(\varepsilon_k')] + \mathrm{E}_{k-1}[\mathrm{E}_k(\varepsilon_k)\,\mathrm{E}_k(\tilde{F}_k')] + \mathrm{E}_{k-1}[\mathrm{E}_k(\varepsilon_k)\,\mathrm{E}_k(\varepsilon_k')] \\ &\quad - \mathrm{E}_{k-1}(\tilde{F}_k)\,\mathrm{E}_{k-1}(\varepsilon_k') - \mathrm{E}_{k-1}(\varepsilon_k)\,\mathrm{E}_{k-1}(\tilde{F}_k) - \mathrm{E}_{k-1}(\varepsilon_k)\,\mathrm{E}_{k-1}(\varepsilon_k'). \end{aligned} \tag{18}$$

Let $\mathrm{E}_{\mathbf{a}_k}$ denote the expectation with respect to the randomness of the kth row $\mathbf{a}_k$ of A. We have

$$\mathrm{E}_{k-1}(\,\cdot\,) = \mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}(\,\cdot\,) = \mathrm{E}_{\mathbf{a}_k}\,\mathrm{E}_k(\,\cdot\,), \tag{19}$$

so that (18) can be rewritten as

$$\begin{aligned}\varphi_k - \tilde{\varphi}_k &= \mathrm{E}_{\mathbf{a}_k}[\mathrm{E}_k(\tilde{F}_k)\,\mathrm{E}_k(\varepsilon_k')] + \mathrm{E}_{\mathbf{a}_k}[\mathrm{E}_k(\varepsilon_k)\,\mathrm{E}_k(\tilde{F}_k')] \\ &\quad + \mathrm{E}_{\mathbf{a}_k}[\mathrm{E}_k(\varepsilon_k)\,\mathrm{E}_k(\varepsilon_k')] - \mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}(\tilde{F}_k)\,\mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}(\varepsilon_k') \\ &\quad - \mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}(\varepsilon_k)\,\mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}(\tilde{F}_k) - \mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}(\varepsilon_k)\,\mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}(\varepsilon_k') \\ &=: T_1 + T_2 + T_3 + T_4 + T_5 + T_6. \end{aligned} \tag{20}$$

From now on, C will denote a finite constant (that will change from line to line) depending uniformly in k, z, z' as z, z' stay at any positive distance away from the real line.

Lemma 3.3. *We have:*

(i) $\mathrm{E}_{\mathbf{a}_k}[|\eta_k|^2] \le CN^{-1}$,

(ii) $\mathrm{E}_{\mathbf{a}_k}[|\eta_k|^4] \le CN^{-1-\frac{\alpha^2}{16}+4\epsilon}$, *for ϵ as in (11).*

Besides, the same bounds hold if one replaces η_k by η_k' or by $\mathrm{E}_k(\eta_k)$ or $\mathrm{E}_k(\eta_k')$.

Proof. Let us denote $G = G^{(k)}(z)$ and write E instead of $\mathrm{E}_{\mathbf{a}_k}$ for short. Note that

$$\operatorname{Tr} GG^* \le CN; \qquad \|G\| \le C. \tag{21}$$

Inequality (i) follows from Lemma A.1, which allows to claim that

$$\mathrm{E}_{\mathbf{a}_k}[|\eta_k|^2] \le CN^{-1}.$$

Let us now prove (ii). We denote $\mathbf{a}_k(i)$ by a_i and set

$$Q := \sum_{i \ne j} G_{ij}\overline{a_i}a_j.$$

We have

$$\mathrm{E}_{\mathbf{a}_k}[|Q|^4] = \sum_{i_1 \ne i_2, i_3 \ne i_4, i_5 \ne i_6, i_7 \ne i_8} G_{i_1i_2}\overline{G_{i_3i_4}}\,\overline{G_{i_5i_6}}G_{i_7i_8}\,\mathrm{E}[\overline{a_{i_1}}a_{i_2}a_{i_3}\overline{a_{i_4}}\,\overline{a_{i_5}}a_{i_6}a_{i_7}\overline{a_{i_8}}].$$

As the a_i's are independent and centered, for a term in this sum to be non zero, one needs the partition of level sets of the function $\ell \in \{1, \ldots, 8\} \mapsto i_\ell$ to have no block of cardinality 1.

Moreover, i_1 and i_2 cannot be in the same block, i_3 and i_4 neither, etc. The cardinalities of these blocks can be $(2, 2, 2, 2)$, $(4, 2, 2)$, $(4, 4)$, or $(3, 3, 2)$. The last possibility only arises when the distribution of the entries is not symmetric. We estimate the above sum in each of the four cases, using (21) and the estimates for the moments given in Lemma 3.1:

- The sum of the terms corresponding to $(2, 2, 2, 2)$ has absolute value at most
$$8^4 (\operatorname{Tr} GG^*)^2 \, \mathrm{E}[|a_1|^2]^4 \le CN^{-2}.$$
- The sum of the terms corresponding to $(4, 2, 2)$ has absolute value at most
$$8^3 (\operatorname{Tr} GG^*)^2 \, \mathrm{E}[|a_1|^2]^2 \, \mathrm{E}[|a_1|^4] \le CN^{-4} N^2 N^{1-\frac{\alpha^2}{16}+\epsilon(4-\alpha)} \le CN^{-1-\frac{\alpha^2}{16}+2\epsilon}.$$
- The sum of the terms corresponding to $(4, 4)$ has absolute value at most
$$8^2 (\Im z)^{-2} \operatorname{Tr} GG^* \, \mathrm{E}[|a_1|^4]^2 \le CN^{-1-\frac{\alpha^2}{8}+4\epsilon}.$$
- The sum of the terms corresponding to $(3, 3, 2)$ has absolute value at most
$$8^4 (\operatorname{Tr} GG^*)^2 \, \mathrm{E}[|a_1|^3]^2 \, \mathrm{E}[|a_2|^2]. \tag{22}$$

  If $\alpha \le 3$, then (22) $\le CN^{-4} N^2 \|GG^*\|^2 \le CN^{-2}$.

  If $\alpha > 4$, then
$$\begin{aligned} (22) &\le CN^{-4} N^2 \|GG^*\|^2 \left(N^{\frac{3}{4}-\frac{\alpha}{16}-\frac{\alpha^2}{16}+\epsilon(3-\alpha)} \right)^2 \\ &\le CN^{-\frac{1}{2}-\frac{\alpha}{8}-\frac{\alpha^2}{8}+2\epsilon} \\ &\le CN^{-1-\frac{\alpha^2}{16}+2\epsilon}. \end{aligned}$$
  (in the last step, we used the elementary fact that $\frac{\alpha}{8} + \frac{\alpha^2}{16} > \frac{1}{2}$).

It follows that
$$\mathrm{E}[|Q|^4] \le CN^{-1-\frac{\alpha^2}{16}+4\epsilon}.$$
As $\mathrm{E}_{\mathbf{a}_k}[|a_{kk}|^4] \le CN^{-1-\frac{\alpha^2}{16}+4\epsilon}$, we deduce that
$$\mathrm{E}_{\mathbf{a}_k}[|\eta_k|^4] \le 2^4 (\mathrm{E}[|a_{kk}|^4] + \mathrm{E}_{\mathbf{a}_k}[|Q|^4]) \le CN^{-1-\frac{\alpha^2}{16}+4\epsilon}.$$
To prove the last assertion, one only needs to change G into $G^{(k)}(z')$ or $\mathrm{E}_k[G^{(k)}(z)]$ or $\mathrm{E}_k[G^{(k)}(z')]$. $\square$

To conclude the proof of Proposition 3.2, we need to prove the upper bound (17). We will use the expression of $\varphi_k - \tilde{\varphi}_k$ given at (20), denoting the six terms of its RHS by $T_1, \ldots, T_6$. We will show that $|T_1|$, $|T_2|$, $|T_3|$, $|T_4|$, $|T_5|$, and $|T_6| \ll N^{1-\alpha/2}$.

To make subsequent calculations less cumbersome to write, we introduce the notation
$$J_k = \frac{1}{z - a_{kk} - \mathbf{a}_k^* G^{(k)}(z) \mathbf{a}_k}$$
and
$$J_{k,\mathrm{diag}} = \frac{1}{z - \mathbf{a}_k^* G^{(k)}_{\mathrm{diag}}(z) \mathbf{a}_k},$$

and correspondingly J'_k, $J'_{k,\text{diag}}$ with z' instead of z. Let furthermore

$$J_{k,\text{Tr}} := \frac{1}{-z - \frac{1}{N}\operatorname{Tr} G^{(k)}(z)}$$

and E be given by

$$E := \mathbf{a}_k^* G^{(k)}(z)_{\text{diag}}\mathbf{a} - \frac{1}{N}\operatorname{Tr} G^{(k)}(z) = \sum_j G_{jj}^{(k)}\left(|\mathbf{a}_k(j)|^2 - \frac{1}{N}\right)$$

so that

$$J_{k,\text{diag}} = J_{k,\text{Tr}} + E J_{k,\text{diag}} J_{k,\text{Tr}}.$$

To find a bound on $|\operatorname{E}_{\mathbf{a}_k}[\varepsilon_k]|$, we bound $\operatorname{E}_{\mathbf{a}_k}[\varepsilon_k]$ and $\operatorname{E}_{\mathbf{a}_k}[-\varepsilon_k]$ separately. We notice that $|1-\eta_k J_{k,\text{diag}}|$ and $|1+\eta_k J_k|$ are reciprocals. Using Jensen's inequality, we find a bound on the former:

$$\begin{aligned}
\operatorname{E}_{\mathbf{a}_k}[\varepsilon_k] &= \operatorname{E}_{\mathbf{a}_k} \log|1-\eta_k J_{k,\text{diag}}|^2\\
&\le \log \operatorname{E}_{\mathbf{a}_k}|1-\eta_k J_{k,\text{diag}}|^2\\
&= \log \operatorname{E}_{\mathbf{a}_k}(1 - 2\Re(\eta_k J_{k,\text{diag}}) + |\eta_k J_{k,\text{diag}}|^2)\\
&= \log \operatorname{E}_{\mathbf{a}_k}(1 - 2\Re(\eta_k J_{k,\text{Tr}} + \eta_k E J_{k,\text{diag}} J_{k,\text{Tr}}) + |\eta_k J_{k,\text{diag}}|^2)\\
&\le \log(1 + O(\operatorname{E}_{\mathbf{a}_k}|\eta_k||E|) + O(\operatorname{E}_{\mathbf{a}_k}|\eta_k|^2)).
\end{aligned}$$

Here in the last line we have used that $J_{k,\text{Tr}}$ is independent of $\mathbf{a}_k$ and that $\operatorname{E}_{\mathbf{a}_k}\eta_k = 0$ which gives that $\operatorname{E}_{\mathbf{a}_k}\Re(\eta_k J_{k,\text{Tr}}) = 0$. We have also used that J_k, $J_{k,\text{diag}}$, and $J_{k,\text{Tr}}$ are uniformly bounded to claim that $\operatorname{E}_{\mathbf{a}_k}\eta_k E J_{k,\text{diag}} J_{k,\text{Tr}} \le O(\operatorname{E}_{\mathbf{a}_k}|\eta_k||E|)$ and $\operatorname{E}_{\mathbf{a}_k}|\eta_k J_{k,\text{diag}}|^2 \le O(\operatorname{E}_{\mathbf{a}_k}|\eta_k|^2)$. Similarly, we show the same bound for $\operatorname{E}_{\mathbf{a}_k}[-\varepsilon_k]$:

$$\begin{aligned}
\operatorname{E}_{\mathbf{a}_k}[-\varepsilon_k] &= \operatorname{E}_{\mathbf{a}_k}\log|1+\eta_k J_k|^2\\
&\le \log \operatorname{E}_{\mathbf{a}_k}|1 + \eta_k J_{k,\text{Tr}} + \eta_k(\eta_k + E) J_{k,\text{diag}} J_k|^2\\
&= \log(1 + 2\operatorname{E}_{\mathbf{a}_k}\Re(\eta_k J_{k,\text{Tr}}) + O(\operatorname{E}_{\mathbf{a}_k}|\eta_k||E|) + O(\operatorname{E}_{\mathbf{a}_k}|\eta_k|^2))\\
&\le \log(1 + O(\operatorname{E}_{\mathbf{a}_k}|\eta_k||E|) + O(\operatorname{E}_{\mathbf{a}_k}|\eta_k|^2)).
\end{aligned}$$

By Cauchy–Schwarz we get that

$$\begin{aligned}
\operatorname{E}_{\mathbf{a}_k}[|\eta_k||E|] &\le \sqrt{\operatorname{E}_{\mathbf{a}_k}|\operatorname{E}_k(\eta_k)|^2 \operatorname{E}_{\mathbf{a}_k}|\operatorname{E}_k E|^2}\\
&\le CN^{-\frac{1}{2}} N^{-\frac{\alpha^2}{32} + \frac{1}{2}\epsilon(4-\alpha)}\\
&\le CN^{(1-\frac{\alpha}{2}) - \frac{4-\alpha}{4} - \frac{(4-\alpha)^2}{32} + \frac{1}{2}\epsilon(4-\alpha)}.
\end{aligned}$$

This and Lemma 3.3(i) yield the bound

$$|\operatorname{E}_{\mathbf{a}_k}[\varepsilon_k]| \le O(\operatorname{E}_{\mathbf{a}_k}|\eta_k||E|) \le CN^{(1-\frac{\alpha}{2}) - \frac{4-\alpha}{4} - \frac{(4-\alpha)^2}{32} + \frac{1}{2}\epsilon(4-\alpha)}$$

since $O(\operatorname{E}_{\mathbf{a}_k}|\eta_k||E|) \ll 1$ and therefore the logarithm is given by its Taylor series. Hence $|T_4|$, $|T_5|$ and $|T_6| \ll CN^{1-\alpha/2}$.

Next, after applying Cauchy–Schwarz to T_3, we look for a bound on $\operatorname{E}_{\mathbf{a}_k}[|\operatorname{E}_k(\varepsilon_k)|^2]$. We note first that by Jensen's inequality

$$\operatorname{E}_{\mathbf{a}_k}[|\operatorname{E}_k(\varepsilon_k)|^2] \le \operatorname{E}_k \operatorname{E}_{\mathbf{a}_k}[|(\varepsilon_k)|^2].$$

Let us partition the space of matrices as follows. We define the events

$$S := \{A : |1 + \eta_k J_k|^2 \geq 1\}; \qquad \tilde{S} := \{A : |1 - \eta_k J_{k,\mathrm{diag}}|^2 \geq 1\} \tag{23}$$

and note that since $|1 - \eta_k J_{k,\mathrm{diag}}|$ and $|1 + \eta_k J_k|$ are reciprocals, we have that

$$|\log |1 - \eta_k J_{k,\mathrm{diag}}|^2| = \mathbb{1}_{\tilde{S}} \log |1 - \eta_k J_{k,\mathrm{diag}}|^2 + \mathbb{1}_S \log |1 + \eta_k J_k|^2.$$

Recall that for $x > 0$, $\log(1 + x) < x$, so that on $\tilde{S}$

$$0 \leq (\log |1 - \eta_k J_{k,\mathrm{diag}}|^2)\mathbb{1}_{\tilde{S}} \leq (-2\Re(\eta_k J_{k,\mathrm{diag}}) + |\eta_k J_{k,\mathrm{diag}}|^2)\mathbb{1}_{\tilde{S}}.$$

Using the definition of S and similar reasoning we get a similar bound on $\mathbb{1}_S \log |1+\eta_k J_k|^2$ with J_k instead of $-J_{k,\mathrm{diag}}$, which yields

$$|\log |1 - \eta_k J_{k,\mathrm{diag}}|^2| \leq -\mathbb{1}_{\tilde{S}}(2\Re(\eta_k J_{k,\mathrm{diag}}) + |\eta_k J_{k,\mathrm{diag}}|^2) + \mathbb{1}_S(2\Re(\eta_k J_k) + |\eta_k J_k|^2). \tag{24}$$

Taking the expectation of $|\varepsilon_k|^2$, and using that $|J_k|$ and $|J_{k,\mathrm{diag}}|$ are absolutely bounded we get that

$$\mathrm{E}_k(\log |1 - \eta_k J_{k,\mathrm{diag}}|^2)^2 \leq C\, \mathrm{E}_k\, |\eta_k|^2 < CN^{-1}$$

where the last inequality follows by the previous Lemma 3.3(i). Here the ϵ of (11) is chosen small enough. By Cauchy–Schwartz, it proves that $|T_3| \leq CN^{-1}$.

Let us now treat T_1 and T_2. We have

$$\begin{aligned} T_2 &= \mathrm{E}_{\mathbf{a}_k}[\mathrm{E}_k(\varepsilon_k)\,\mathrm{E}_k(J'_{k,\mathrm{diag}})] \\ &= \mathrm{E}_{\mathbf{a}_k}[\mathrm{E}_k(\varepsilon_k)\,\mathrm{E}_k(J'_{k,\mathrm{Tr}} + E J'_{k,\mathrm{diag}} J'_{k,\mathrm{Tr}})] \\ &= (\mathrm{E}_k\, J'_{k,\mathrm{Tr}})\,\mathrm{E}_{\mathbf{a}_k}\,\mathrm{E}_k(\varepsilon_k) + \mathrm{E}_{\mathbf{a}_k}[\mathrm{E}_k(\varepsilon_k)\,\mathrm{E}_k(E J'_{k,\mathrm{diag}} J'_{k,\mathrm{Tr}})] \\ &\leq C\, \mathrm{E}_{\mathbf{a}_k}[|\,\mathrm{E}_k(\varepsilon_k)\,\mathrm{E}_k\, E|] \end{aligned}$$

where in the last two lines we used that $J'_{k,\mathrm{Tr}}$ is independent of $\mathbf{a}_k$ and that $\mathrm{E}_{\mathbf{a}_k}\, \varepsilon_k = 0$. By Cauchy–Schwartz and our previous bounds, we get that

$$\begin{aligned} |T_2| &\leq C\sqrt{\mathrm{E}_{\mathbf{a}_k}\, |\,\mathrm{E}_k(\varepsilon_k)|^2\, \mathrm{E}_{\mathbf{a}_k}\, |\,\mathrm{E}_k\, E|^2} \leq CN^{-\frac{1}{2}} N^{-\frac{\alpha^2}{32}+\frac{1}{2}\epsilon(4-\alpha)} \\ &\leq N^{(1-\frac{\alpha}{2})-\frac{4-\alpha}{4}-\frac{(4-\alpha)^2}{32}+\frac{1}{2}\epsilon(4-\alpha)}. \end{aligned}$$

The same bound holds for T_1. This concludes the proof of Proposition 3.2. □

3.4. Computation of the limit

By what precedes, to prove (13), it suffices to prove that the random variables

$$f_k(z) := \frac{1 + \sum_j |\mathbf{a}_k(j)|^2 (G^{(k)}(z))^2_{jj}}{z - \sum_j |\mathbf{a}_k(j)|^2 G^{(k)}(z)_{jj}} \tag{25}$$

satisfy the convergence in probability

$$N^{-1+\alpha/2}\, \mathrm{E}_{k-1}[(\mathrm{E}_k - \mathrm{E}_{k-1})(f_k(z))(\mathrm{E}_k - \mathrm{E}_{k-1})(f_k(z'))] \longrightarrow 2uC(z, z') \tag{26}$$

as $N, k \to \infty$ with $k/N \to u$.

Let $\widetilde{f}_k(z)$ be defined as $f_k(z)$, but with the matrix A replaced by a matrix $\widetilde{A}$ whose entries $\widetilde{a}_{ij}$ are the ones of A if $i \le k$ and $j \le k$ and independent random variables with the same distribution as the entries of A if $i > k$ or $j > k$. Let $\widetilde{G}^{(k)}(z)$ denote the resolvent of $\widetilde{A}^{(k)}$. This notation is convenient as we can then express a product of integrals as integrals over different sets of variables. Let also $\mathrm{E}_{\mathbf{a}_k,\widetilde{\mathbf{a}}_k}$ denote the expectation with respect to the randomness of the kth columns of A and $\widetilde{A}$. Furthermore, E_k still denotes the conditional expectation with respect to the σ-algebra generated by the $k \times k$ upper left corner of A (or of $\widetilde{A}$, as they share the same $k \times k$ upper left corner).

Lemma 3.4. *We have*

$$\begin{aligned}&\mathrm{E}_{k-1}[(\mathrm{E}_k - \mathrm{E}_{k-1})(f_k(z))(\mathrm{E}_k - \mathrm{E}_{k-1})(f_k(z'))]\\&\quad = \mathrm{E}_k[\mathrm{E}_{\mathbf{a}_k,\widetilde{\mathbf{a}}_k}(f_k(z)\widetilde{f}_k(z'))] - \mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}\, f_k(z)\,\mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}\, f_k(z').\end{aligned}$$

Proof. On the σ-algebra generated by A and $\widetilde{A}$, we have $\mathrm{E}_{k-1} = \mathrm{E}_{\mathbf{a}_k,\widetilde{\mathbf{a}}_k}\,\mathrm{E}_k = \mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k,\widetilde{\mathbf{a}}_k}$, hence by Lemma A.4,

$$\begin{aligned}&\mathrm{E}_{k-1}[(\mathrm{E}_k - \mathrm{E}_{k-1})(f_k(z))(\mathrm{E}_k - \mathrm{E}_{k-1})(f_k(z'))]\\&\quad = \mathrm{E}_{\mathbf{a}_k,\widetilde{\mathbf{a}}_k}[\mathrm{E}_k\, f_k(z)\,\mathrm{E}_k\, f_k(z')] - \mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}\, f_k(z)\,\mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}\, f_k(z').\end{aligned}$$

Now, note that $\mathrm{E}_k\, f_k(z') = \mathrm{E}_k\, \widetilde{f}_k(z')$, hence

$$\begin{aligned}&\mathrm{E}_{k-1}[(\mathrm{E}_k - \mathrm{E}_{k-1})(f_k(z))(\mathrm{E}_k - \mathrm{E}_{k-1})(f_k(z'))]\\&\quad = \mathrm{E}_{\mathbf{a}_k,\widetilde{\mathbf{a}}_k}[\mathrm{E}_k\, f_k(z)\,\mathrm{E}_k\, \widetilde{f}_k(z')] - \mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}\, f_k(z)\,\mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}\, f_k(z').\end{aligned}$$

Lastly,

$$\mathrm{E}_k\, f_k(z)\,\mathrm{E}_k\, \widetilde{f}_k(z') = \mathrm{E}_k[f_k(z)\widetilde{f}_k(z')]$$

thus

$$\begin{aligned}&\mathrm{E}_{k-1}[(\mathrm{E}_k - \mathrm{E}_{k-1})(f_k(z))(\mathrm{E}_k - \mathrm{E}_{k-1})(f_k(z'))]\\&\quad = \mathrm{E}_{\mathbf{a}_k,\widetilde{\mathbf{a}}_k}[\mathrm{E}_k[f_k(z)\widetilde{f}_k(z')]] - \mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}\, f_k(z)\,\mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}\, f_k(z')\\&\quad = \mathrm{E}_k[\mathrm{E}_{\mathbf{a}_k,\widetilde{\mathbf{a}}_k}[f_k(z)\widetilde{f}_k(z')]] - \mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}\, f_k(z)\,\mathrm{E}_k\,\mathrm{E}_{\mathbf{a}_k}\, f_k(z'). \qquad \square\end{aligned}$$

For $w \in \mathbb{C} \setminus \mathbb{R}$, $\frac{1}{w} = -\mathrm{i}\,\mathrm{sgn}_w \int_0^{+\infty} e^{\mathrm{sgn}_w\, \mathrm{i}tw}\mathrm{d}t$, (recall that $\mathrm{sgn}_w = \mathrm{sgn}(\Im w)$). Hence, letting $w = z - \sum_j |\mathbf{a}_k(j)|^2 G^{(k)}(z)_{jj}$ and using (16), we get that

$$\begin{aligned}f_k(z) &= \frac{1 + \sum\limits_j |\mathbf{a}_k(j)|^2 (G^{(k)}(z))^2_{jj}}{z - \sum\limits_j |\mathbf{a}_k(j)|^2 G^{(k)}(z)_{jj}}\\&= -\mathrm{i}\,\mathrm{sgn}_z \int_0^{+\infty} \left(1 + \sum_j |\mathbf{a}_k(j)|^2 (G^{(k)}(z)^2)_{jj}\right) e^{\mathrm{sgn}_z \mathrm{i}t(z - \sum\limits_j |\mathbf{a}_k(j)|^2 G^{(k)}(z)_{jj})}\mathrm{d}t.\end{aligned}$$

By (16) the above integral in t and in the randomness of $\mathbf{a}_k$ is absolutely convergent and thus we can interchange $\mathrm{E}_{\mathbf{a}_k}$ and $\int_{t=0}^{+\infty}$. Thus

$$\mathrm{E}_{\mathbf{a}_k}\, f_k(z) = -\mathrm{i}\,\mathrm{sgn}_z \int_0^{+\infty} \mathrm{E}_{\mathbf{a}_k} \left(1 + \sum_j |\mathbf{a}_k(j)|^2 (G^{(k)}(z)^2)_{jj}\right) \times e^{\mathrm{sgn}_z \mathrm{i}t(z - \sum_j |\mathbf{a}_k(j)|^2 G^{(k)}(z)_{jj})} \mathrm{d}t.$$

Then, for any $t > 0$, we have that

$$\mathrm{E}_{\mathbf{a}_k} \left(1 + \sum_j |\mathbf{a}_k(j)|^2 (G^{(k)}(z)^2)_{jj}\right) e^{\mathrm{sgn}_z \mathrm{i}t(z - \sum_j |\mathbf{a}_k(j)|^2 G^{(k)}(z)_{jj})} = \mathrm{E}_{\mathbf{a}_k} \frac{1}{\mathrm{sgn}_z \mathrm{i}t} \partial_z \{e^{\mathrm{sgn}_z \mathrm{i}t(z - \sum_j |\mathbf{a}_k(j)|^2 G^{(k)}(z)_{jj})}\}$$

and that $\mathrm{E}_{\mathbf{a}_k}$ and ∂_z can be permuted by (16) again. Hence, we have

$$\mathrm{E}_{\mathbf{a}_k}\, f_k(z) = -\mathrm{i}\,\mathrm{sgn}_z \int_0^{+\infty} \frac{1}{\mathrm{sgn}_z \mathrm{i}t} \partial_z\, \mathrm{E}_{\mathbf{a}_k} \{e^{\mathrm{sgn}_z \mathrm{i}t(z - \sum_j |\mathbf{a}_k(j)|^2 G^{(k)}(z)_{jj})}\} \mathrm{d}t$$

and for $\phi_N(\lambda) = \mathrm{E}\big[\exp(-\mathrm{i}\lambda |a_{11}|^2)\big]$ as defined at (10),

$$\mathrm{E}_{\mathbf{a}_k}\, f_k = -\int_0^{+\infty} \partial_z \frac{1}{t} e^{\mathrm{sgn}_z \mathrm{i}tz} \prod_j \phi_N(\mathrm{sgn}_z t G^{(k)}(z)_{jj}) \mathrm{d}t. \tag{27}$$

Let us now use (10):

$$\phi_N(\lambda) = 1 - \frac{\mathrm{i}\lambda}{N} - c\frac{(\mathrm{i}\lambda)^{\frac{\alpha}{2}}}{N^{\frac{\alpha}{2}}} + \frac{|\lambda|^{\frac{\alpha}{2}}}{N^{\frac{\alpha}{2}}} \varepsilon_N(\mathrm{i}\lambda/N).$$

Hence

$$\mathrm{E}_{\mathbf{a}_k}\, f_k(z) = -\int_0^{+\infty} \partial_z \frac{1}{t} e^{\mathrm{sgn}_z \mathrm{i}tz} \prod_j \left(1 + \frac{u_j(z)}{N}\right) \mathrm{d}t$$

for

$$\begin{aligned} u_j(z) &:= N\left(\Phi_N(\mathrm{sgn}_z t G^{(k)}(z)_{jj}) - 1\right) \\ &= -\mathrm{i}\,\mathrm{sgn}_z t G^{(k)}(z)_{jj} - c\frac{(\mathrm{i}\,\mathrm{sgn}_z t G^{(k)}(z)_{jj})^{\frac{\alpha}{2}}}{N^{\frac{\alpha-2}{2}}} \\ &\quad + \frac{|tG^{(k)}(z)_{jj}|^{\frac{\alpha}{2}}}{N^{\frac{\alpha-2}{2}}} \varepsilon_N(\mathrm{i}\,\mathrm{sgn}_z t G^{(k)}(z)_{jj}/N). \end{aligned} \tag{28}$$

Let

$$\begin{aligned} \delta(z,t) &:= \frac{1}{t} e^{\mathrm{sgn}_z \mathrm{i}tz} \prod_j \phi_N(\mathrm{sgn}_z t G^{(k)}(z)_{jj}) \\ &\quad - \frac{1}{t} e^{\mathrm{sgn}_z \mathrm{i}tz - \frac{\mathrm{sgn}_z \mathrm{i}t}{N} \mathrm{Tr}\, G^{(k)}(z)} \left(1 - \frac{c}{N^{\alpha/2}} \sum_j (\mathrm{i}\,\mathrm{sgn}_z t G^{(k)}(z)_{jj})^{\alpha/2}\right). \end{aligned}$$

We want to show that

$$\left|\int_0^\infty \partial_z \delta(z,t)\mathrm{d}t\right| = o(N^{-\alpha/2+1}).$$

Since $\delta(z,t)$ is analytic in $z \in \mathbb{C}\setminus\mathbb{R}$, by the Cauchy inequality

$$|\partial_z \delta(z,t)| < \frac{2M_t}{\Im z}$$

where $M_t = \max_{B(z,\Im z/2)} |\delta(z,t)|$. Let z_t be the maximizer of $\delta(w,t)$ on $B(z,\Im z/2)$. Then

$$\left|\int_0^\infty \partial_z \delta(z,t)\mathrm{d}t\right| \le \frac{2}{\Im z}\int_0^\infty |\delta(z_t,t)|\mathrm{d}t.$$

Let $0 < \gamma < (\alpha-2)/(2\alpha) < 1/2$. Then we split the above integral into two parts:

$$\int_0^\infty |\delta(z_t,t)|\mathrm{d}t = \left(\int_0^{N^\gamma} + \int_{N^\gamma}^\infty\right)|\delta(z_t,t)|\mathrm{d}t.$$

It is easy to see that

$$\int_{N^\gamma}^\infty |\delta(z_t,t)| \le Ce^{-\frac{N^\gamma|\Im z|}{2}}$$

so we focus on the first integral.

Recalling that $\phi_N(\mathrm{sgn}_z tG^{(k)}(z)_{jj}) = 1 + \frac{u_j(z)}{N}$, we write $\int_0^{N^\gamma}|\delta(z_t,t)|\mathrm{d}t$ as the integral of a sum of three errors $\delta_1,\delta_2,\delta_3$:

$$\begin{aligned}
\int_0^{N^\gamma}|\delta(z_t,t)|\mathrm{d}t &\le \int_0^{N^\gamma}\left|\frac{1}{t}e^{\mathrm{sgn}_z \mathrm{i}tz_t}\left(\prod_j\left(1+\frac{u_j(z_t)}{N}\right) - e^{\frac{1}{N}\sum_j u_j(z_t)}\right)\right| \\
&\quad + \int_0^{N^\gamma}\left|\frac{1}{t}e^{\mathrm{sgn}_z \mathrm{i}tz_t}\left(e^{\frac{1}{N}\sum_j u_j(z_t)} - e^{-\frac{\mathrm{sgn}_z \mathrm{i}t}{N}\operatorname{Tr}G^{(k)}(z_t)}\left(1-\frac{c}{N^{\alpha/2}}\right.\right.\right. \\
&\quad \times \left.\left.\left.\sum_j(\mathrm{i}\,\mathrm{sgn}_z tG^{(k)}(z_t)_{jj})^{\alpha/2} + \sum_j\frac{|tG^{(k)}(z_t)_{jj}|^{\frac{\alpha}{2}}}{N^{\frac{\alpha}{2}}}\varepsilon_N(\mathrm{i}\,\mathrm{sgn}_z tG^{(k)}(z_t)_{jj}/N)\right)\right)\right| \\
&\quad + \int_0^{N^\gamma}\left|\frac{1}{t}e^{\mathrm{sgn}_z \mathrm{i}tz}e^{-\frac{\mathrm{sgn}_z \mathrm{i}t}{N}\operatorname{Tr}G^{(k)}(z_t)}\left(\sum_j\frac{|tG^{(k)}(z_t)_{jj}|^{\frac{\alpha}{2}}}{N^{\frac{\alpha}{2}}}\varepsilon_N(\mathrm{i}\,\mathrm{sgn}_z tG^{(k)}(z_t)_{jj}/N)\right)\right| \\
&=: \delta_1+\delta_2+\delta_3.
\end{aligned}$$

To get a bound on δ_1 we use Lemma A.5 for each t with

$$\begin{aligned}
M_t &:= \max_i |u_i(z_t)| \\
&= \max_i\left|-\mathrm{i}\,\mathrm{sgn}_z tG^{(k)}(z_t)_{jj} - c\frac{(\mathrm{i}\,\mathrm{sgn}_z tG^{(k)}(z_t)_{jj})^{\frac{\alpha}{2}}}{N^{\frac{\alpha-2}{2}}}\right. \\
&\quad \left. + \frac{|tG^{(k)}(z_t)_{jj}|^{\frac{\alpha}{2}}}{N^{\frac{\alpha-2}{2}}}\varepsilon_N(\mathrm{i}\,\mathrm{sgn}_z tG^{(k)}(z_t)_{jj}/N)\right| \\
&\le Ct
\end{aligned}$$

for a constant C. Besides, $|\phi_N(\mathrm{sgn}_z t G^{(k)}(z)_{jj}| \le 1$, hence by (28), $\Re(u_j(z_t)) \le 0$. Hence

$$\delta_1 \le \int_0^{N^\gamma} \left| \frac{1}{t} e^{\mathrm{sgn}_z \mathrm{i} t z_t} \frac{C^2 t^2}{N} e^{\frac{1}{N}\sum_j \Re(u_j(z_t)) + \frac{C^2 t^2}{N}} \right| \mathrm{d}t \le \frac{C^2}{N} \int_0^{N^\gamma} \left| t e^{\mathrm{sgn}_z \mathrm{i} t z_t} \right| e^{\frac{C^2 t^2}{N}} \mathrm{d}t$$
$$\le \frac{C}{N^{1-2\gamma}},$$

where we used the fact that $\gamma < 1/2$.

To get a bound on δ_2, we can use the Taylor series expansion to check that for $|x| \le 1/2$

$$|e^x - (1+x)| = \left| x^2 \sum_{k=0}^{\infty} \frac{x^k}{(k+2)!} \right| \le |x|^2 \sum_{k=0}^{\infty} \frac{|x|^k}{(k+2)!} \le \frac{|x|^2}{1-|x|} \le 2|x|^2.$$

This yields that

$$\exp\left(-c \sum_j \frac{(\mathrm{i}\,\mathrm{sgn}_z t G^{(k)}(z_t)_{jj})^{\frac{\alpha}{2}}}{N^{\alpha/2}} + \sum_j \frac{|t G^{(k)}(z_t)_{jj}|^{\frac{\alpha}{2}}}{N^{\alpha/2}} \varepsilon_N(\mathrm{i}\,\mathrm{sgn}_z t G^{(k)}(z_t)_{jj}/N) \right)$$
$$- \left(1 - c \sum_j \frac{(\mathrm{i}\,\mathrm{sgn}_z t G^{(k)}(z_t)_{jj})^{\frac{\alpha}{2}}}{N^{\alpha/2}} + \sum_j \frac{|t G^{(k)}(z_t)_{jj}|^{\frac{\alpha}{2}}}{N^{\alpha/2}} \varepsilon_N(\mathrm{i}\,\mathrm{sgn}_z t G^{(k)}(z_t)_{jj}/N) \right)$$
$$\le 2 \left| -c \sum_j \frac{(\mathrm{i}\,\mathrm{sgn}_z t G^{(k)}(z_t)_{jj})^{\frac{\alpha}{2}}}{N^{\alpha/2}} + \sum_j \frac{|t G^{(k)}(z_t)_{jj}|^{\frac{\alpha}{2}}}{N^{\alpha/2}} \varepsilon_N(\mathrm{i}\,\mathrm{sgn}_z t G^{(k)}(z_t)_{jj}/N) \right|^2$$

so that, as $\gamma < (\alpha-2)/(2\alpha)$ implies that $\gamma\alpha - \alpha + 2 < -\alpha/2 + 1$,

$$\delta_2 = \int_0^{N^\gamma} \frac{1}{t} e^{-\frac{\mathrm{sgn}_z \mathrm{i} t}{N} \operatorname{Tr} G^{(k)}(z_t)}$$
$$\times \left| -c \sum_j \frac{(\mathrm{i}\,\mathrm{sgn}_z t G^{(k)}(z_t)_{jj})^{\frac{\alpha}{2}}}{N^{\alpha/2}} + \sum_j \frac{|t G^{(k)}(z_t)_{jj}|^{\frac{\alpha}{2}}}{N^{\alpha/2}} \varepsilon_N(\mathrm{i}\,\mathrm{sgn}_z t G^{(k)}(z_t)_{jj}/N) \right|^2 \mathrm{d}t$$
$$\le C N^{\gamma\alpha - \alpha + 2} \ll N^{-\alpha/2+1}.$$

To get a bound on δ_3 we do a dyadic decomposition of the integral. We integrate on $[2^k, 2^{k+1}]$ with k such that $2^k < N^\gamma$. This yields

$$\int_{2^k}^{2^{k+1}} \left| \frac{1}{t} e^{\mathrm{sgn}_z \mathrm{i} t z} e^{-\frac{\mathrm{sgn}_w \mathrm{i} t}{N} \operatorname{Tr} G^{(k)}(z_t)} \left(\sum_j \frac{|t G^{(k)}(z_t)_{jj}|^{\frac{\alpha}{2}}}{N^{\frac{\alpha}{2}}} \varepsilon_N(\mathrm{i}\,\mathrm{sgn}_z t G^{(k)}(z_t)_{jj}/N) \right) \right|$$
$$\le C 2^k e^{-|\Im z| t 2^k} \frac{(2^k)^{\alpha/2-1}}{N^{\alpha/2-1}} \max_{t \in [0, N^\gamma]} \varepsilon_N(\mathrm{i}\,\mathrm{sgn}_z t G^{(k)}(z_t)_{jj}/N).$$

Noting that $\sum_k 2^{k\alpha/2} e^{-|\Im z| 2^k}$ is convergent we get that

$$\delta_3 \le C \frac{\max_{t \in [0, N^\gamma]} \varepsilon_N(\mathrm{i}\,\mathrm{sgn}_z t G^{(k)}(z_t)_{jj}/N)}{N^{\alpha/2-1}} \ll N^{-\alpha/2+1}.$$

As a consequence,

$$\mathrm{E}_{\mathbf{a}_k}\, f_k(z) = -\int_0^{+\infty} \partial_z \frac{1}{t} e^{\operatorname{sgn}_z itz - \frac{\operatorname{sgn}_z it}{N}\operatorname{Tr} G^{(k)}(z)} \left(1 - \frac{c}{N^{\alpha/2}} \sum_j (\mathrm{i}\operatorname{sgn}_z t G^{(k)}(z)_{jj})^{\alpha/2}\right) \mathrm{d}t$$
$$+ o(N^{-\alpha/2+1}).$$

Then Lemma 3.5, whose statement and proof are postponed until the next section, implies that the diagonal terms $G^{(k)}(z)_{jj}$ in the previous expression are close to the Stieltjes transform $G_{\mathrm{sc}}(z)$ of the semicircle law with support $[-2, 2]$. It also implies that $N^{-1}\operatorname{Tr} G^{(k)}(z)$ is close to $G_{\mathrm{sc}}(z)$. As before, the factor $e^{\operatorname{sgn}_z itz}$ allows us to take the integral $\mathrm{d}t$ to infinity. It follows that

$$\mathrm{E}_{\mathbf{a}_k}\, f_k(z) = -\int_0^{+\infty} \partial_z \frac{1}{t} e^{\operatorname{sgn}_z itz - \operatorname{sgn}_z it G_{\mathrm{sc}}(z)} \left(1 - N^{-\alpha/2+1} c K(z,t)^{\alpha/2}\right) \mathrm{d}t$$
$$+ o(N^{-\alpha/2+1}), \tag{29}$$

where $o(1)$ is for the convergence in probability.

Of course, the same holds for $\mathrm{E}_{\mathbf{a}_k}\, \widetilde{f}_k(z')$.

Let us now compute $\mathrm{E}_{\mathbf{a}_k, \widetilde{\mathbf{a}}_k}(f_k(z)\widetilde{f}_k(z'))$. First, as above,

$$f_k(z)\widetilde{f}_k(z') = \int_{t,t'>0} \partial_z \partial_{z'} \frac{1}{tt'} \{e^{\operatorname{sgn}_z it(z - \sum_j |\mathbf{a}_k(j)|^2 G^{(k)}(z)_{jj}) + \operatorname{sgn}_{z'} it'(z' - \sum_j |\widetilde{\mathbf{a}}_k(j)|^2 \widetilde{G}(z')^{(k)}_{jj})}\} \mathrm{d}t\mathrm{d}t'.$$

Hence as $\mathrm{E}_{\mathbf{a}_k}$ denotes the integration with respect to the kth columns of A and A' and as these columns are identical up to the kth entry and independent from the $k+1$th entry on, we have

$$\mathrm{E}_{\mathbf{a}_k, \widetilde{\mathbf{a}}_k}(f_k(z)\widetilde{f}_k(z')) = \int_{t,t'>0} \partial_z \partial_{z'} \frac{1}{tt'} \{e^{\operatorname{sgn}_z itz + \operatorname{sgn}_{z'} it'z'}$$
$$\times \prod_{j<k} \phi_N(\operatorname{sgn}_z t G^{(k)}(z)_{jj} + \operatorname{sgn}_{z'} t' \widetilde{G}(z')^{(k)}_{jj})$$
$$\times \prod_{j>k} \phi_N(\operatorname{sgn}_z t G^{(k)}(z)_{jj}) \phi_N(\operatorname{sgn}_{z'} t' \widetilde{G}(z')^{(k)}_{jj})\} \mathrm{d}t\mathrm{d}t'.$$

Then, proceeding as above when we computed $\mathrm{E}_{\mathbf{a}_k}\, f_k(z)$ (from (27) to (29)), we get that

$$\mathrm{E}_{\mathbf{a}_k, \widetilde{\mathbf{a}}_k}(f_k(z)\widetilde{f}_k(z')) = \int_{t,t'>0} \partial_z \partial_{z'} \frac{1}{tt'} e^{\operatorname{sgn}_z itz - K(z,t)} e^{\operatorname{sgn}_{z'} it'z' - K(z',t')}$$
$$\times \Big(1 - cuN^{-\alpha/2+1}(K(z,t) + K(z',t'))^{\alpha/2}$$
$$- c(1-u)N^{-\alpha/2+1}(K(z,t)^{\alpha/2} + K(z',t')^{\alpha/2}) + o(N^{-\alpha/2+1})\Big) \mathrm{d}t\mathrm{d}t'$$

This equation, together with (29) and Lemma 3.4, imply (26). This concludes the proof.

3.5. *Concentration of the diagonal terms of the resolvent*

Lemma 3.5. *For any fixed $z \in \mathbb{C} \setminus \mathbb{R}$, for any $p \geq 1$, and any $j \in \{1, \ldots, N\}$, the sequence*

$$\|G(z)_{jj} - G_{\mathrm{sc}}(z)\|_{L^p},$$

tends to zero.

Proof. As $|G(z)_{jj} - G_{\mathrm{sc}}(z)| \le 2(\Im z)^{-1}$, it suffices to prove the result for $p = 2$. By the Schur complement formula (see [2, Th. 11.4]), we know that

$$G(z)_{jj} = \frac{1}{z - \frac{x_{jj}}{\sqrt{N}} - \frac{1}{N}\sum_{i,k:i\neq j,k\neq j} G^{(j)}(z)_{ik} x_{ji} x_{kj}}, \tag{30}$$

where $G^{(j)}(z) = (z - A^{(j)})^{-1}$ and $A^{(j)}$ is the matrix obtained after removing the jth row and the jth column of A.

By Lemma A.1 of the Appendix and the fact that $|\operatorname{Tr} G^{(j)}(z) - \operatorname{Tr} G(z)| \le C$ (see Lemma A.6), the denominator of the RHS of (30) can be written

$$z - \frac{1}{N}\operatorname{Tr} G(z) + O_{L^2}(N^{-\eta})$$

for $\eta := \frac{\alpha^2}{32} - \epsilon(2 - \alpha/2)$.

As the function $f_z : x \mapsto \frac{1}{z-x}$ has uniformly bounded gradient on the half plane $\{x \in \mathbb{C};\ \Im z \Im x > 0\}$, we get that

$$G(z)_{jj} = \frac{1}{z - \frac{1}{N}\operatorname{Tr} G(z)} + O_{L^2}(N^{-\eta}).$$

Besides, by [11, Lem. C.1], we know that

$$\frac{1}{N}\operatorname{Tr} G(z) - \mathrm{E}\left[\frac{1}{N}\operatorname{Tr} G(z)\right] = O_{L^2}(N^{-1/2}),$$

hence

$$G(z)_{jj} = \frac{1}{z - \mathrm{E}\left[\frac{1}{N}\operatorname{Tr} G(z)\right]} + O_{L^2}(N^{-\eta}).$$

At last, by [2, Th. 2.5], we know that for any fixed $z \in \mathbb{C} \setminus \mathbb{R}$,

$$\mathrm{E}\left[\frac{1}{N}\operatorname{Tr} G(z)\right] - G_{\mathrm{sc}}(z) \longrightarrow 0.$$

We conclude by using the fact that $\frac{1}{z - G_{\mathrm{sc}}(z)} = G_{\mathrm{sc}}(z)$. □

Acknowledgment

A.M. acknowledges the support of the Leverhulme Trust Early Career Fellowship (ECF 2013-613).

Appendix

A.1. Quadratic forms in heavy-tailed variables

Lemma A.1. *Let* $\mathbf{a} = (a_1, \ldots, a_N)^T$ *be a column vector whose entries are i.i.d., centered and satisfy* (ii) *and* (iii) *of Lemma* 3.1. *Then for any deterministic matrix* G, *the random variables*

$$X := \sum_{i\neq j} G_{ij}\overline{a_i} a_j \qquad E := \sum_i G_{ii}|a_i|^2 - \frac{1}{N}\operatorname{Tr} G$$

satisfy

$$\mathrm{E}[|X|^2] \leq 2N^{-2}\operatorname{Tr}(GG^*) \leq 2N^{-1}\|G\|^2 \qquad \mathrm{E}[|E|^2] \leq 10C\|G\|^2 N^{-\frac{\alpha^2}{16}+\epsilon(4-\alpha)}.$$

Proof. Direct computations, the second one using $\mathrm{E}[|E|^2] = \mathrm{E}[|E - \mathrm{E}[E]|^2] + |\mathrm{E}[E]|^2$. □

Remark A.2. We shall sometimes use this lemma after removal of the kth row and column of B and of the kth entry of $\mathbf{a}$, but it suffices to apply the lemma with the matrix deduced from B by setting its kth row and column to zero.

A.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

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

and we also define

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

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

Then we have the following theorem. It is proved in the real case at [8, 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 $z^2, \overline{z}^2, |z|^2$.

Theorem A.3. *Suppose that for some constants $v \geq 0, \tau \in \mathbb{C}$, we have the convergence in probability*

$$v(N) \underset{N\to\infty}{\longrightarrow} v \qquad \tau(N) \underset{N\to\infty}{\longrightarrow} \tau$$

and that for each $\varepsilon > 0$,

$$L(\varepsilon, N) \underset{N\to\infty}{\longrightarrow} 0.$$

Then we have the convergence in distribution

$$M_N(N) \underset{N\to\infty}{\longrightarrow} Z,$$

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

To apply this theorem, we shall use the following lemma.

Lemma A.4. *Let E_{k-1} and E_k denote conditional expectations given some σ-algebras $\mathcal{F}_{k-1} \subset \mathcal{F}_k$. then for any pair A, B of L^2 r.v., we have*

$$\mathrm{E}_{k-1}[(\mathrm{E}_k - \mathrm{E}_{k-1})(A)(\mathrm{E}_k - \mathrm{E}_{k-1})(B)] = \mathrm{E}_{k-1}[\mathrm{E}_k(A)\,\mathrm{E}_k(B)] - \mathrm{E}_{k-1}(A)\,\mathrm{E}_{k-1}(B).$$

Proof. We have

$$\begin{aligned}
\mathrm{E}_{k-1}[(\mathrm{E}_k - \mathrm{E}_{k-1})(A)(\mathrm{E}_k - \mathrm{E}_{k-1})(B)] &= \mathrm{E}_{k-1}[\mathrm{E}_k\, A(\mathrm{E}_k - \mathrm{E}_{k-1})(B)] \\
&\quad - \mathrm{E}_{k-1}[\mathrm{E}_{k-1}(A)(\mathrm{E}_k - \mathrm{E}_{k-1})(B)] \\
&= \mathrm{E}_{k-1}[\mathrm{E}_k\, A\, \mathrm{E}_k\, B] \\
&\quad - \underbrace{\mathrm{E}_{k-1}[\mathrm{E}_k\, A\, \mathrm{E}_{k-1}(B)]}_{=\mathrm{E}_{k-1}(B)\,\mathrm{E}_{k-1}[\mathrm{E}_k\, A]=\mathrm{E}_{k-1}(A)\,\mathrm{E}_{k-1}(B)} \\
&\quad - \mathrm{E}_{k-1}(A)\underbrace{\mathrm{E}_{k-1}[(\mathrm{E}_k - \mathrm{E}_{k-1})B]}_{=0}
\end{aligned}$$

which concludes the proof. □

A.3. A lemma about large products and the exponential function

Lemma A.5. *Let u_i, $i = 1, \dots, N$, be some complex numbers and set*

$$P := \prod_{i=1}^{N}\left(1 + \frac{u_i}{N}\right) \qquad S := \frac{1}{N}\sum_{i=1}^{N} u_i \qquad M := \max_i |u_i|.$$

There is a universal constant $R > 0$ (independent of N and of M) such that

$$\frac{M}{N} \le R \implies |P - e^S| \le \frac{M^2}{N} e^{\Re(S) + \frac{M^2}{N}}.$$

Proof. Let $L(z)$ be defined on $B(0, 1)$ by $\log(1 + z) = z + z^2 L(z)$ and $R > 0$ be such that on $B(0, R)$, $|L(z)| \le 1$. If $\frac{M}{N} \le R$, we have

$$P = \prod_i \exp\left\{\frac{u_i}{N} + \frac{u_i^2}{N^2} L\left(\frac{u_i}{N}\right)\right\} = e^S \exp\left\{\sum_i \frac{u_i^2}{N^2} L\left(\frac{u_i}{N}\right)\right\},$$

so that

$$P - e^S = e^S\left(\exp\left\{\sum_i \frac{u_i^2}{N^2} L\left(\frac{u_i}{N}\right)\right\} - 1\right).$$

Since for any z, $|e^z - 1| \le |z|e^{|z|}$, the conclusion follows. □

A.4. Linear algebra

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 A.6 (*Difference of Traces of a Matrix and Its Major Submatrices*)**.** *Let H_k be the submatrix of H obtained by removing its kth row and kth column and set $G_k := (z - H_k)^{-1}$. Let also $\mathbf{a}_k$ be the kth column of H where the kth entry has been removed. Then*

$$\mathrm{Tr}(G) - \mathrm{Tr}(G_k) = \frac{1 + \mathbf{a}_k^* G_k^2 \mathbf{a}_k}{z - h_{kk} - \mathbf{a}_k^* G_k \mathbf{a}_k}. \tag{31}$$

Moreover,

$$|\mathrm{Tr}(G) - \mathrm{Tr}(G_k)| \le \pi |\Im z|^{-1}. \tag{32}$$

References

[1] G. Anderson, A. Guionnet, O. Zeitouni, An Introduction to Random Matrices, in: Cambridge Studies in Advanced Mathematics, vol. 118, 2009.
[2] Z.D. Bai, J.W. Silverstein, Spectral Analysis of Large Dimensional Random Matrices, second ed., Springer, New York, 2009.
[3] Z.D. Bai, X. Wang, W. Zhou, CLT for linear spectral statistics of Wigner matrices, Electron. J. Probab. 14 (83) (2009) 2391–2417.
[4] Z.D. Bai, J. Yao, On the convergence of the spectral empirical process of Wigner matrices, Bernoulli 11 (2005) 1059–1092.
[5] G. Ben Arous, A. Guionnet, The spectrum of heavy tailed random matrices, Comm. Math. Phys. 278 (3) (2008) 715–751.
[6] F. Benaych-Georges, A. Guionnet, C. Male, Central limit theorems for linear statistics of heavy tailed random matrices, Comm. Math. Phys. 329 (2) (2014) 641–686.
[7] F. Benaych-Georges, S. Péché, Localization and delocalization for heavy tailed band matrices, Ann. Inst. Henri Poincaré Probab. Stat. 50 (4) (2014) 1385–1403.
[8] P. Billingsley, Probability and Measure, Wiley, third ed.
[9] 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.
[10] C. Bordenave, P. Caputo, D. Chafaï, Spectrum of non-Hermitian heavy tailed random matrices, Comm. Math. Phys. 307 (2) (2011) 513–560.
[11] C. Bordenave, A. Guionnet, Localization and delocalization of eigenvectors for heavy-tailed random matrices, Probab. Theory Related Fields 157 (3–4) (2013) 885–953.
[12] S. Boucheron, G. Lugosi, P. Massart, Concentration Inequalities, Oxford, 2013.
[13] N.H. Bingham, C.M. Goldie, J.L. Teugels, Regular Variation, Cambridge University Press, 1989.
[14] P. Cizeau, J.-P. Bouchaud, Theory of Lévy matrices, Phys. Rev. E 50 (1994).
[15] L. Erdös, B. Schlein, H.T. Yau, Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices, Ann. Probab. 37 (2009).
[16] L. Erdös, H.T. Yau, J. Yin, Rigidity of eigenvalues of generalized Wigner matrices, Adv. Math. 229 (2012) 1435–1515.
[17] A. Guionnet, O. Zeitouni, Concentration of the spectral measure for large matrices, Electron. Commun. Probab. 5 (2000) 119–136.
[18] K. Johansson, On the fluctuations of eigenvalues of random Hermitian matrices, Duke Math. J. 91 (1998) 151–204.
[19] A.M. Khorunzhy, B.A. Khoruzhenko, L.A. Pastur, Asymptotic properties of large random matrices with independent entries, J. Math. Phys. 37 (1996) 5033–5060.
[20] M. Ledoux, The Concentration of Measure Phenomenon, AMS, Providence, RI, 2001.
[21] A. Lytova, L. Pastur, Central limit theorem for linear eigenvalue statistics of random matrices with independent entries, Ann. Probab. 37 (2009) 1778–1840.
[22] L. Pastur, M. Shcherbina, Eigenvalue Distribution of Large Random Matrices, in: Mathematical Surveys and Monographs, vol. 171, American Mathematical Society, Providence, RI, 2011.
[23] M. Shcherbina, Central limit theorem for linear eigenvalue statistics of the Wigner and sample covariance random matrices, J. Math. Phys. Anal. Geom. 7 (2) (2011) 176–192.
[24] 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.
[25] T. Tao, V. Vu, Random matrices: universality of local eigenvalue statistics, Acta Math. 206 (2011) 127–204.

Further reading

[1] L. Erdös, B. Schlein, H.-T. Yau, J. Yin, The local relaxation flow approach to universality of the local statistics for random matrices, Ann. Inst. Henri Poincaré Probab. Stat. 48 (1) (2012) 1–46.
[2] W. Feller, An Introduction to Probability Theory and Its Applications, Vol. II, second ed., J. Wiley, New York, London, Sydney, 1966.