Support convergence in the single ring theorem

Alice Guionnet* and Ofer Zeitouni†

December 8, 2010

Keywords Random matrices, non-commutative measure, Schwinger–Dyson equation.
Mathematics Subject of Classification : 15A52 (46L50,46L54)

Abstract

We study the eigenvalues of non-normal square matrices of the form $A_n = U_n T_n V_n$ with U_n, V_n independent Haar distributed on the unitary group and T_n real diagonal. We show that when the empirical measure of the eigenvalues of T_n converges, and T_n satisfies some technical conditions, all these eigenvalues lie in a single ring.

1 The problem

In [6], M. Krishnapur and the authors considered the convergence of the empricial measure of (complex) eigenvalues of matrices of the form $A_n = T_n U_n$, where U_n is Haar distributed on $\mathcal{U}(n)$, the unitary group of $n \times n$ matrices, and independent of the self-adjoint matrix T_n (which therefore can be assumed diagonal, with real non-negative entries $s_i^{(n)}$). That is, with $\lambda_i^{(n)}$ denoting the eigenvalues of A_n, $L_{A_n} = n^{-1} \sum_{i=1}^{n} \delta_{\lambda_i^{(n)}}$ their empirical measure, and with L_{T_n} the empirical measure of

*UMPA, CNRS UMR 5669, ENS Lyon, 46 allée d'Italie, 69007 Lyon, France. aguionne@umpa.ens-lyon.fr. This work was partially supported by the ANR project ANR-08-BLAN-0311-01.

†School of Mathematics, University of Minnesota and Faculty of Mathematics, Weizmann Institute, POB 26, Rehovot 76100, Israel. zeitouni@math.umn.edu. The work of this author was partially supported by NSF grant DMS-0804133 and by a grant from the Israel Science Foundation.

the entries of T_n, the following is part of the main result of [6]. Throughout, for a probability measure μ supported on $\mathbb{R}$ or on $\mathbb{C}$, we write G_μ for its Stieltjes transform, that is

$$G_\mu(z) = \int \frac{\mu(dx)}{z-x}\,.$$

G_μ is analytic off the support of μ. We write G_{T_n} for $G_{\tilde{L}_{T_n}}$, where for any probability measure μ on $\mathbb{R}$ we use $\tilde{\mu}$ to denote the symmetrized of μ, i.e. the probability measure satisfying $\tilde{\mu}(A) = (\mu(A) + \mu(-A))/2$.

Theorem 1. *Assume $\{L_{T_n}\}_n$ converges weakly to a probability measure Θ compactly supported on $\mathbb{R}_+$. Assume further the following.*

1. *There exists a constant $M > 0$ so that*

$$\lim_{n\to\infty} P(\|T_n\| > M) = 0\,. \tag{1}$$

2. *There exist a sequence of events $\{\mathcal{G}_n\}$ with $P(\mathcal{G}_n^c) \to 0$ and constants $\delta, \delta' > 0$ so that for Lebesgue almost any $z \in \mathbb{C}$, with σ_n^z the minimal singular value of $zI - A_n$,*

$$E(\mathbf{1}_{\mathcal{G}_n} \mathbf{1}_{\{\sigma_n^z < n^{-\delta}\}} (\log \sigma_n^z)^2) < \delta'\,. \tag{2}$$

3. *There exist constants $\kappa, \kappa_1 > 0$ such that*

$$|\Im G_{T_n}(z)| \le \kappa_1 \quad \text{on} \quad \{z : \Im(z) > n^{-\kappa}\}\,. \tag{3}$$

Then L_{A_n} converges in probability to a limiting probability measure μ_A, rotationally invariant in $\mathbb{C}$ and supported on the annulus $\{re^{i\theta} : a \le r \le b\}$, where $a = 1/\sqrt{\int x^{-2}\Theta(dx)}$ and $b = \sqrt{\int x^2 \Theta(dx)}$.

The conditions of Theorem 1 were then showed to hold in some examples of interest, and in particular to provide a rigorous proof of the Feinberg-Zee "single ring theorem", see [3]. A version of Theorem 1 was also proved to hold when the Haar measure on $\mathcal{H}_n$ was replaced by the Haar measure on the orthogonal group, see [6, Theorem 18].

Our goal in this paper is to improve the convergence statement in Theorem 1 to a statement concerning the convergence of the support of L_{A_n}. The following is our main theorem.

Theorem 2. *Assume T_n, U_n satisfy the conditions of Theorem 1 and, in addition, assume that*

$$a_n := \frac{1}{\sqrt{\int x^{-2} L_{T_n}(dx)}} \to a = \frac{1}{\sqrt{\int x^{-2}\Theta(dx)}}\,, \tag{4}$$

and

$$b_n := \sqrt{\int x^2 L_{T_n}(dx)} \to b = \sqrt{\int x^2 \Theta(dx)}\,. \tag{5}$$

Further assume if $a > 0$ that $\sup_n \|T_n^{-1}\| < \infty$. Then, the support of L_{A_n} converges to $\mathrm{supp}(\mu_A) = \{z \in \mathbb{C} : |z| \in [a,b]\}$ in probability. If moreover the assumptions of Theorem 1 hold almost surely with respect to the sequence T_n, then the convergence of the support holds almost surely.

When T_n is distributed as the diagonal matrix of singular values of a Ginibre matrix, the conclusion of Theorem 2 follows e.g. from the results in [10].

Remark 3. Recall that μ_A is supported on the annulus $[a,b] \times [0,2\pi)$. An elementary computation using the expression for the density $\rho_A = \rho_A(r)$ of μ_A, see [6, 7], shows that

$$\lim_{r \searrow a} \rho_A(r) = \frac{1}{\pi a^2}\,, \quad \lim_{r \nearrow b} \rho_A(r) = \frac{1}{\pi b^2}\,.$$

It is maybe surprising that in spite of the density having a strictly positive density at the boundary, the eigenvalues still stick to the boundary.

1.1 Background and description of the proof

We recall that the main difficulty in studying the ESD L_{A_n} is that A_n is not a normal matrix, that is $A_n A_n^* \neq A_n^* A_n$, almost surely. For normal matrices, the limit of ESDs can be found by the method of moments or by the method of Stieltjes' transforms. For non-normal matrices, the only known method of proof, which is the one followed in [6], is more indirect and follows an idea of Girko [4]. We recall the general outline and some crucial steps which will be needed in the proof of Theorem 2.

Introduce the $2n \times 2n$ matrix

$$H_n^z := \begin{bmatrix} 0 & zI - A_n \\ (zI - A_n)^* & 0 \end{bmatrix}. \tag{6}$$

Let ν_n^z denote the ESD of H_n^z,

$$\int \frac{1}{y-x} d\nu_n^z(x) = \frac{1}{2n} \mathrm{tr}\left((y - H_n^z)^{-1}\right) ,$$

then, see [6, Eq. (7)],

$$\int \psi(z) dL_{A_n}(z) = \frac{1}{2\pi} \int_{\mathbb{C}} \Delta\psi(z) \int_{\mathbb{R}} \log|x| d\nu_n^z(x) dm(z) . \tag{7}$$

The main advantage of this formulation is that one can reduce attention to the study of the ESD of matrices of the form $(T+U)(T+U)^*$ where T is real diagonal and U is Haar distributed. In the limit (i.e., when T and U are replaced by operators in a C^*-algebra that are freely independent, with T bounded and self adjoint and U unitary), the limit ESD has been identified by Haagerup and Larsen [7]. The Schwinger–Dyson equations give both a characterization of the limit and, more important to us, a discrete approximation that can be used to estimate the discrepancy between the pre-limit ESD and its limit. These will play a crucial role in the study of the support.

Notation

We describe our convention concerning constants. Throughout, by the word *constant* we mean quantities that are independent of n (or of the complex variables z, z_1). Generic constants denoted by the letters C or c, have values that may change from line to line, and they may depend on other parameters. Constants denoted by C_i, K, M, κ and κ' are fixed and do not change from line to line.

2 Preliminaries: evaluation of ν^z and convergence rates

We quickly recall the analysis in [6], assuming throughout that $\|T_n\|$ is uniformly bounded by a constant $M < \infty$. Fix $z \in \mathbb{C}$ and write $\rho = |z|$. With

$$\mathbf{U}_n = \begin{pmatrix} 0 & U_n \\ 0 & 0 \end{pmatrix}, \mathbf{Y}_n = \begin{pmatrix} 0 & B_n \\ B_n^* & 0 \end{pmatrix}, \tag{8}$$

where $B_n = \rho U_n + T_n$, T_n a real, diagonal matrix of uniformly bounded norm and U_n a $\mathcal{H}_n$ unitary matrix, define

$$G^n(z) = E[\frac{1}{2n}\mathrm{tr}\left((z-\mathbf{Y}_n)^{-1}\right)], \quad G_{T_n}(z) = G^n(z)|_{\rho=0}$$

and

$$G^n_U(z) = E[\frac{1}{2n}\mathrm{tr}\left(\mathbf{U}_n(z-\mathbf{Y}_n)^{-1}\right)].$$

Then, see [6, Eq. (35)], the finite n Schwinger-Dyson equations for this problem give

$$\rho(G^n(z_1))^2 = 2G^n_U(z_1)(1+2\rho G^n_U(z_1)) - O_1(n,z_1), \tag{9}$$

where

$$\begin{aligned} O_1(n,z_1) &= 4E\left[(\frac{1}{2n}\mathrm{tr} - E[\frac{1}{2n}\mathrm{tr}]) \otimes (\frac{1}{2n}\mathrm{tr} - E[\frac{1}{2n}\mathrm{tr}])\partial(z_1-\mathbf{Y}_n)^{-1}\mathbf{U}_n\right] \\ &= O\left(\frac{\rho^2}{n^2\Im(z_1)^2(\Im(z_1)\wedge 1)}\right). \end{aligned}$$

In particular, we have

$$G^n_U(z_1) = \frac{1}{4\rho}(-1+\sqrt{1+4\rho^2 G^n(z_1)^2 + 4O_1(n,z_1)}), \tag{10}$$

with the choice of the square root determined by analyticity and behavior at infinity. Further, if one defines

$$z_2 = \psi_n(z_1) := z_1 - \frac{\rho^2 G^n(z_1)}{(1+2\rho G^n_U(z_1))}, \tag{11}$$

then, see [6, Eq. (39)], for all z_1 with $\Im(z_2) > 0$ given by (11),

$$G^n(z_1) = G_{T_n}(\psi_n(z_1)) - \tilde{O}(n,z_1,\psi_n(z_1)), \tag{12}$$

where

$$\tilde{O}(n,z_1,z_2) = \frac{2O(n,z_1,z_2)}{(1+2\rho G^n_U(z_1))}$$

and

$$|O(n,z_1,z_2)| \le \frac{C\rho^2}{n^2|\Im(z_2)|\Im(z_1)^2(\Im(z_1)\wedge 1)}.$$

In particular, for $\Im(z_1)$ large, it holds that $G^n(z_1)$ and $G^n_U(z_1)$ are small, implying that z_2 is well defined with $\Im(z_2) > 0$. This leads (see [6, Lemma 10]) to the following weak convergence statement.

Lemma 4. *If L_{T_n} converges weakly in probability to a probability measure Θ, then for any $z \in \mathbb{C}$, ν_n^z converges weakly in probability to $\nu^z = \tilde{\Theta} \boxplus \lambda_{|z|}$.*

(Recall that $\tilde{\Theta}$ is the symmetrized version of Θ.)

The main work in [6] is then to use the Schwinger-Dyson equation (12) and deduce enough a-priori bounds that allow one to integrate the logarithmic singularity in (7). While we will make use of some of these bounds, at this point we return to our goal, which is to prove Theorem 2.

3 Convergence of the support - proof of Theorem 2

Throughout this section, we are in the setup and assumptions of Theorem 2. We first consider the statement concerning convergence in probability. Recall that $\mathrm{supp}(\mu_A) = \{z \in \mathbb{C} : |z| \in [a,b]\}$. Since the density of μ_A is positive on its support, see [6, Remark 8], we only need to prove that if $z \notin \mathrm{supp}(\mu_A)$ then there exists an $\varepsilon = \varepsilon(z) > 0$ so that, with $B(z,\varepsilon)$ denoting an open ball in $\mathbb{C}$ centered at z with radius ε,

$$P(L_{A_n}(B(z,\varepsilon)) \neq 0) \to_{n\to\infty} 0\,.$$

Let $\overline{\nu}_n^z = \lambda_{|z|} \boxplus \tilde{L}_{T_n}$ (i.e., $\overline{\nu}_n^z$ denotes the free convolution of $\lambda_{|z|}$ with the symmetrized empirical measure of T_n). Since $L_{T_n} \to \Theta$ weakly, we have that $\overline{\nu}_n^z \to \nu^z$ weakly. Write $\overline{G}_n^z$ for the Stieltjes transform of $\overline{\nu}_n^z$. Then, $\overline{G}_n^z(\cdot)$ converges to the Stieltjes transform of ν^z, which is denoted in the sequel by $G(\cdot)$.

The first observation we make reduces the study of the support of L_{A_n} to a question concerning $\overline{\nu}_n^z$.

Lemma 5. *For each $z \notin \mathrm{supp}(\mu_A)$ there exists an $\varepsilon = \varepsilon(z)$ so that $\overline{\nu}_n^{z'}(B(0,\varepsilon)) = 0$ if $|z - z'| < \varepsilon$, for all n large.*

Before bringing the proof of Lemma 5, we provide an a-priori estimate on the spectral radius of certain operators. Throughout, we use $r(A)$ to denote the spectral radius of an operator A. We use the convention that $\|\cdot\|$ denotes the operator norm and $\|\cdot\|_2$ the Hilbert-Schmidt norm. An operator T in a non-commutative probability space is called *R-diagonal* iff it has the same distribution as UH with U unitary, H positive, and the algebras generated by (U,U^*) and H freely independent, see [7, 9].

Lemma 6. *Let A,B be elements of a non-commutative tracial C^*-probability space. Assume that A is R-diagonal and that there exists a constant $c_0 > 0$ so that $\|A\|,\|B\| \leq c_0$. Then, for each $\varepsilon > 0$ there there exists an $\eta = \eta(c_0,\varepsilon) > 0$ so that*

$$r(A+\eta B) \leq \|A\|_2 + \varepsilon.$$

(The case $\eta = 0$ of the lemma is [7, Proposition 4.1].)

Proof. Recall that $r(A+\eta B) = \lim \|(A+\eta B)^n\|^{1/n}$. By [7, Corollary 4.2], we have that $\|A^p\| \leq (1+p)C\|A\|_2^{p-1}$. Therefore, using the sub-additivity of norms, we have, with $C_n = \|(A+\eta B)^n\|$,

$$C_n \leq \|A^n\| + \sum_{k=0}^{n-1} \|A^k\| \cdot \|\eta B\| \cdot C_{n-k-1}\,, \tag{13}$$

where $C_0 = 1$.

For $\gamma > 0$, set $G(\gamma) = \sum_{n\geq 1} \gamma^n C_n$. Clearly $G(\gamma) < \infty$ for γ small enough, and $r(A+\eta B)^{-1} = \sup\{\gamma : G(\gamma) < \infty\}$. Further, $G(\cdot)$ is analytic on $[0, r(A+\eta B)^{-1})$. Define also $F(\gamma) = \sum_{n\geq 1} \gamma^n (1+n)\|A\|_2^{n-1}$ and note that $F(\gamma) < \infty$ whenever $\gamma < \|A\|_2^{-1}$. From (13) we get that whenever $G(\gamma) < \infty$,

$$G(\gamma) \leq C\sum_{n\geq 1} \gamma^n (1+n)\|A\|_2^{n-1} + |\eta| C c_0 \sum_{n=1}^{\infty} \gamma^n \sum_{k=0}^{n-1} (1+k)\|A\|_2^{(k-1)\vee 0} C_{n-k-1}\,. \tag{14}$$

Rearranging, we have that the second sum in the right side of (14) equals

$$\begin{aligned}
& \sum_{n=1}^{\infty} \gamma^n \sum_{k=0}^{n-1} (1+k)\|A_2\|^{(k-1)\vee 0} C_{n-k-1} \\
= \; & \sum_{k=0}^{\infty} \|A_2\|^{(k-1)\vee 0} (k+1)\gamma^{k+1} \sum_{n=k+1}^{\infty} \gamma^{n-k-1} C_{n-k-1} \\
= \; & \gamma \left(1 + \sum_{k=1}^{\infty} \|A_2\|^{k-1} (k+1)\gamma^k \right) (1+G(\gamma))\,.
\end{aligned}$$

It follows that

$$G(\gamma) \leq CF(\gamma) + Cc_0 \eta\gamma(1+F(\gamma))(G(\gamma)+1)\,.$$

Therefore, for all γ with $G(\gamma) < \infty$ and $F(\gamma) < \infty$,

$$(1 - Cc_0\eta\gamma(1+F(\gamma)))G(\gamma) \le CF(\gamma) + Cc_0\eta\gamma(1+F(\gamma))\,.$$

It follows that for $\gamma = (\|A\|_2+\varepsilon)^{-1}$ there exists an $\eta = \eta(\varepsilon, c_0)$ so that $Cc_0\eta\gamma(1+F(\gamma)) < 1/2$ and therefore $G(\gamma) < \infty$. This implies the statement of Lemma 6. $\square$

We can now provide the proof of Lemma 5.

Proof of Lemma 5. Recall that $\overline{\nu}_n^{z'} = \tilde{L}_{T_n} \boxplus \lambda_{|z'|}$, see Theorem 1, and thus possesses the same law as $X+Y_n$ where X, Y_n are freely independent in a non-commutative probability space, the law of X is that of a Bernoulli $\pm|z'|$ variable, and the law of Y_n being $\tilde{L}_{T_n}$.

Assume first that $|z| > b$. We may and will assume that for some $\delta > 0$, $|z'| - b_n > \delta > 0$ for all n large, uniformly in z' with $|z-z'| < \varepsilon$, and consider only such n, ε and δ. We need to check that there exists an ε' such that for all $|\eta| < \varepsilon'$, $X+Y_n-\eta I$ is invertible. Writing $X+Y_n-\eta I = X(I+X^{-1}(Y_n-\eta I))$, we see that $X+Y_n-\eta I$ is invertible iff $I+X^{-1}(Y_n-\eta I)$ is invertible. A sufficient condition for that is that $r(X^{-1}(Y_n-\eta I)) < 1$. Since $\|X^{-1}\| \le |z'|^{-1}$ and $\|Y_n\|$ is uniformly bounded, and since $X^{-1}Y_n$ is R-diagonal with

$$\|X^{-1}Y_n\|_2 \le \|X^{-1}\|_2\|Y_n\|_2 = |z'|^{-1}\|Y_n\|_2 = |z'|^{-1}b_n \le \zeta < 1$$

for some fixed $\zeta = \zeta(b, \varepsilon, \delta)$, the conclusion follows from an application of Lemma 6 with $A = X^{-1}Y_n$ and $B = X^{-1}$.

Similarly, if $|z| \in [0, a)$ (with $a > 0$) and $\|Y_n^{-1}\|$ is uniformly bounded, we repeat the argument, this time writing $X+Y_n-\eta I = Y_n(I+Y_n^{-1}(X-\eta I))$, and then using

$$\|Y_n^{-1}\|_2\|X\|_2 = |z'|/a_n < \zeta < 1\,.$$

$\square$

Let

$$\mathcal{A} = \{z : \exists \varepsilon > 0, \overline{\nu}_n^z(B(0,\varepsilon)) = 0\,, \text{ for all } n \text{ large}\}.$$

Our next step is to prove a control on $G^n(\cdot)$ for $z \in \mathcal{A}$.

Lemma 7. *Fix $z \in \mathcal{A}$, $z \neq 0$. Let $\beta > 0$ be such that for some n_0 large enough,*

$$[-2\beta, 2\beta] \not\in (\cup_{n \ge n_0} \operatorname{supp} \overline{\nu}_z^n)\,.$$

Then, there are constants $\alpha, \gamma, p > 0$ so that for all n large and for all z_1 with $\Im(z_1) > n^{-\gamma}$ and $\Re(z_1) \in [-\beta, \beta]$,

$$|G^n(z_1) - \overline{G}_n^z(z_1)| < \frac{1}{n^{1+\alpha}\Im(z_1)^p}\,. \tag{15}$$

Proof. The proof is divided into several steps. The idea is to use (12) to compare G^n and $\overline{G}_n^z$. To do this up to a small neighborhood of the real axis, an important point is to show that G^n and $\overline{G}_n^z$ do not cross the cut of the square root which enters in the definition of R_ρ. The latter point is first shown at a positive distance of the real axis and then a bootstrap argument is used to approach the real axis.

Step 1. Introduce the set

$$\mathcal{C}_{\varepsilon,\beta} = \{z_1 : \Im(z_1) \in [\varepsilon, 2\varepsilon), \Re(z_1) \in [-\beta, \beta]\}.$$

Since $[-\beta,\beta] \not\in \mathrm{supp}\bar{\nu}_n^z$, we have that $\Im(\overline{G}_n^z(x+i0)) = 0$ for $x \in [-\beta,\beta]$. Moreover $\overline{G}_n^z$ is uniformly Lipschitz on $\cup_{\varepsilon''\le\varepsilon}\mathcal{C}_{\varepsilon'',\beta}$ (with constant only depending on the distance from $[-\beta,\beta]$ to $\mathrm{supp}\bar{\nu}_z^n$, which is uniformly bounded below by β by hypothesis). Therefore, for any fixed $\varepsilon'(=\beta^{-2}\varepsilon)$ (whose value can be taken to be $1/12$ in what follows) we can choose ε small enough such that

$$\text{for all } z_1 \in \cup_{\varepsilon''\le\varepsilon}\mathcal{C}_{\varepsilon'',\beta}, \text{ it holds that } \Im(\overline{G}_n^z(z_1)) < \varepsilon',\ \Im(G(z_1)) < \varepsilon'. \tag{16}$$

By the convergence of G^n to G (which follows from the weak convergence of $L_{\mathbf{Y}_n}$ to μ_Y, see Lemma 4), which can be made uniform by uniform continuity on $\mathcal{C}_{\varepsilon,\beta}$, and replacing ε' by $3\varepsilon'$ if necessary, we get that for all $n > n_0(\varepsilon)$,

$$\text{for all } z_1 \in \mathcal{C}_{\varepsilon,\beta}, \text{ it holds that } \Im(G^n(z_1)) < 3\varepsilon'. \tag{17}$$

Step 2. Consider z_1 with $\Re(z_1) = 0$. In that case, the real part of both $G^n(z_1)$ and $G(z_1)$ vanishes by symmetry (G, G^n are Stieljes transforms of symmetric measures.) Now, with G_U as in [6, Section 3.1], we have, see [6, (22)],

$$G_U(z_1) = \frac{1}{4\rho}(-1+\sqrt{1+4\rho^2 G(z_1)^2})\,.$$

By the analyticity of G, G_U along the imaginary axis, we deduce that $\sqrt{1+4\rho^2 G(z_1)^2}$ can not vanish and since $G(z_1)$ goes to zero at infinity, this implies that $|\Im(G(z_1))| < 1/2$. By continuity for each ε there is a $\delta = \delta(\varepsilon)$ so that with z_1 such that $\Re(z_1) = 0, \Im(z_1) > \varepsilon$, we have $|\Im G(z_1)| \le 1/2 - \delta$. Again by uniform convergence, and reducing δ to $\delta/2$ if necessary, we get the same for G^n and $\overline{G}_n^z$.

Step 3. Define

$$\mathcal{C}'_{\varepsilon,\beta} := \mathcal{C}_{\varepsilon,\beta} \cup \{z_1 : \Re(z_1) = 0, \Im(z_1) > \varepsilon\}.$$

By Steps 1 and 2, there exist $\delta'' = \delta''(\varepsilon) > 0$ such that

$$\text{for all } z_1 \in \cup_{\varepsilon''\le\varepsilon}\mathcal{C}'_{\varepsilon'',\beta}, \text{ it holds that } \Re(1+4G^2(z_1)) > \delta'' \tag{18}$$

and, for all $n > n_0(\varepsilon)$,

$$\text{for all } z_1 \in \mathcal{C}'_{\varepsilon,\beta}, \text{ it holds that } \Re(1+4(G^n)^2(z_1)) > \delta''. \tag{19}$$

In particular, for all $n > n_0(\varepsilon)$, there is a path leading from $+i\infty$ to any point in $\mathcal{C}'_{\varepsilon,\beta}$ along which the choice of the branch of the square-root in (10) (and its version with no error term, see [6, Eq. (22)]) is determined by analyticity (and is the standard one). Denote such a path $\mathcal{P}$. With this, we can improve the statement of boundedness in [6, Lemma 13] to a convergence statement. In what follows, even though at this stage the path $\mathcal{P}$ is bounded away from the real axis (by ε), we make explicit the dependence of bounds on $\Im(z_1)$; this will be useful in Step 4.

We rewrite (12) as

$$\tilde{G}^n(z_1) = G_{T_n}(\psi_n(z_1)) = G^n(z_1) - \tilde{O}(n, z_1, \psi_n(z_1)). \tag{20}$$

With

$$k_n(z_1) = \rho R_\rho(\tilde{G}^n(z_1)) + \psi_n(z_1) - z_1 = \rho R_\rho(\tilde{G}^n(z_1)) - \frac{\rho^2 G^n(z_1)}{(1+2\rho G_U^n(z_1))},$$

we have

$$\tilde{G}^n(z_1) = G_{T_n}\left(z_1 + k_n(z_1) - \rho R_\rho(\tilde{G}^n(z_1))\right). \tag{21}$$

When $\Im(z_1) > 0$ is large, we have that $\Im(\psi_n(z_1))$ is large, and as a consequence, $\tilde{G}^n(z_1)$ is analytic and small in this region. It follows that $k_n(z_1)$ is analytic in that region, and goes to 0 together with its derivative as $\Im(z_1) \to \infty$. Therefore, the map $z_1 \to z_1 + k_n(z_1)$ is invertible in a neighborhood of $+i\infty$ with analytic inverse, denoted $\varphi_n(z_1)$, which is a small perturbation of the identity there. Defining $\hat{G}^n(z_1) = \tilde{G}^n(\varphi_n(z_1))$, we obtain

$$\hat{G}^n(z_1) = G_{T_n}(z_1 - \rho R_\rho(\hat{G}^n(z_1))).$$

Comparing with [6, Equation (29)], we get that in a neighborhood of $+i\infty$, it holds that $\hat{G}^n(z_1) = \overline{G}_n^z(z_1)$, and therefore, in that neighborhood,

$$\tilde{G}^n(z_1) = \overline{G}_n^z(z_1 + k_n(z_1)). \tag{22}$$

On the other hand, from (20), we have that

$$|\tilde{G}^n(z_1) - G^n(z_1)| \le |\tilde{O}(n, z, \psi_n(z))| \le \frac{C\rho^2}{n^2(\Im(z_1)^4 \wedge 1)}. \tag{23}$$

Thus, for $\Im z \geq C_3 n^{-1/4}$, by (19),

$$\text{for all } z_1 \in \mathcal{C}'_{\varepsilon,\beta}, \text{ it holds that } \Re(1+4(\tilde{G}^n)^2(z_1)) > \delta''/2. \tag{24}$$

Therefore R_ρ is continuously differentiable at $\tilde{G}^n(z_1), z_1 \in \mathcal{C}'_{\varepsilon,\beta}$ and we have

$$|\rho R_\rho(\tilde{G}^n(z_1)) - \rho R_\rho(G^n(z_1))| \leq \frac{C}{n^2(\Im(z_1)^4 \wedge 1)}. \tag{25}$$

Moreover, in the proof of [6, Lemma 12], it was shown that $\rho R_\rho(G^n(z_1)) - \frac{\rho^2 G^n(z_1)}{1+2\rho G^n_U(z_1)}$ is small and analytic on $\mathcal{C}'_{\varepsilon,\beta}$ provided $\varepsilon > n^{-1/4}$. Thus, with (24), (25), we deduce that

$$|k_n(z_1)| \leq C_{20}/(n^{3/2}(\Im(z_1)^7 \wedge 1) \tag{26}$$

is smaller than $\Im z_1/2$ and analytic on $\mathcal{C}'_{\varepsilon,\beta}$ provided $\varepsilon > n^{-1/7}$. Hence, (22) extends to $z_1 \in \mathcal{C}'_{\varepsilon,\beta}$ provided $\varepsilon > n^{-1/7}$.

Therefore, again for $z_1 \in \mathcal{C}'_{\varepsilon,\beta}$, $\varepsilon > n^{-1/7}$,

$$\begin{aligned} |G^n(z_1) - \overline{G}^z_n(z_1)| &\leq |\tilde{G}^n(z_1) - \overline{G}^z_n(z_1)| + |G^n(z_1) - \tilde{G}^n(z_1)| \\ &= |\overline{G}^z_n(z_1 + k_n(z_1)) - \overline{G}^z_n(z_1)| + |G^n(z_1) - \tilde{G}^n(z_1)| \\ &\leq \frac{C}{n^{3/2}(\Im(z_1)^8)}. \end{aligned} \tag{27}$$

Step 4 We bootsrap the previous estimate so that one can approach the real axis: recall that if S denotes the Stieltjes transform of a probability measure supported on $\mathbb{R}$, we have that for any $x \in \mathbb{R}$,

$$|\Im(S(x+i\varepsilon/2))| \leq 2|\Im(S(x+i\varepsilon))|.$$

In particular, for all $z_1 = x + iy \in \mathcal{C}_{\varepsilon/2,\beta}$, it holds that

$$\begin{aligned} |\Im(G^n(z_1))| &\leq 2|\Im(G^n(x+2iy))| \\ &\leq 2|\Im(\overline{G}^z_n(x+2iy))| + 2|G^n(x+2iy) - \overline{G}^z_n(x+2iy)| \\ &\leq 2\varepsilon' + \frac{2C}{n^{3/2}(\Im(z_1)^8)}. \end{aligned}$$

In particular, for all $n > n_1(\varepsilon)$, (17) and (19) hold with ε replaced by $\varepsilon/2$.

One now repeats Step 3, and concludes that (27) continues to hold in $\mathcal{C}'_{\varepsilon/2,\beta}$. Iterating this ℓ times so that $\varepsilon 2^{-\ell} \geq n^{-1/7}$(without changing further $n_1(\varepsilon)$ or $\delta''(\varepsilon)$) completes the proof of Lemma 7. □

We have the following corollary of Lemma 7, whose proof is identical to the proof of [1, Lemma 5.5.5].

Corollary 8. *With* β, α *as in Lemma 7, and* φ *any smooth function compactly supported on* $[-\beta, \beta]$*,*

$$\limsup_{n\to\infty} n^{\alpha+1} |E \int \varphi d\nu_z^n| < \infty.$$

In particular,

$$\limsup_{n\to\infty} P(\nu_z^n([-\beta/2, \beta/2]) > 0) = 0\,. \tag{28}$$

We have now prepared all the steps to prove Theorem 2.

Proof of Theorem 2 We only need to consider z in a compact set. We begin by noting that

$$P(A_n \text{ has an eigenvalue in } B(z,\varepsilon)) = P(\nu_n^{z'}(\{0\}) \geq \tfrac{1}{n} \text{ for some } z' \in B(z,\varepsilon))\,. \tag{29}$$

We write $\mathbf{Y}_n(z)$ to emphasize the dependence of $\mathbf{Y}_n$ in z. Let

$$\lambda^*(\mathbf{Y}_n(z)) = \min\{|\lambda_i(\mathbf{Y}(z))|\}.$$

Since $\mathbf{Y_n}(z) - \mathbf{Y_n}(z')$ is Hermitian and of norm bounded by $|z - z'|$, we have that $|\lambda^*(\mathbf{Y}_n(z)) - \lambda^*(\mathbf{Y}_n(z'))| \leq |z - z'|$. Thus, for each $z \notin \text{supp}(\mu_A)$, and with $\beta = \beta(z)$ as in Lemma 7, we can find an $\varepsilon = \varepsilon(z)$ so that by Chebyshev's inequality

$$P(\nu_n^{z'}(\{0\}) \geq \frac{1}{n} \text{ for some } z' \in B(z,\varepsilon)) \leq P(\nu_n^z([-\beta/2, \beta/2]) \geq \frac{1}{n}) \leq Cn^{-\alpha} \to_{n\to\infty} 0\,.$$

Combined with (29), we conclude that

$$P(A_n \text{ has an eigenvalue in } B(z,\varepsilon)) \to_{n\to\infty} 0\,.$$

By a standard covering argument, this implies that for any compact G with $G \cap (\text{supp}\mu_A) = \emptyset$, it holds that

$$P(A_n \text{ has an eigenvalue in } G) \to_{n\to\infty} 0\,.$$

This completes the convergence in probability in the statement of Theorem 2.

We finally prove the almost sure convergence by generalizing the ideas of [8] based on Poincaré inequality. In our case, we shall use concentration of measures on SU(N) [1, Theorem 4.4.27]. Since we now assume that the assumptions of Theorem 1 hold for almost all sequence T_n, we may and will assume the sequence T_n deterministic in the sequel. Recall that for any bounded measurable function φ, $\int \varphi(x) d\nu_n^z(x)$ is a bounded measurable function of the random matrix $W_n = U_n^* V_n^*$. We denote by $E_{U(n)}$ (resp. $E_{SU(n)}$) the expectation over W_n following the Haar measure on $\mathcal{U}(n)$ (resp. $SU(n)$). We also write in the sequel $\mathcal{B} = (\operatorname{supp} \mu_A)^c$.

Lemma 9. *Fix $z \in \mathcal{B}$, α and β as in Lemma 7, and a bounded non negative smooth function φ with support in $[-\beta, \beta]$.*

1. *There exists a finite constant C such that*

$$|E_{U(n)}[\int \varphi(x) d\nu_n^z(x)]| \leq \frac{C}{n^{1+\alpha}} . \tag{30}$$

2. *For all $\delta > 0$, there exists $z' \in \mathcal{B}$ so that $|z - z'| \leq \delta$ and*

$$|E_{SU(n)}[\int \varphi(x) d\nu_n^{z'}(x)]| \leq \frac{C}{n^{1+\frac{\alpha}{2}}} .$$

Moreover there exists $n_0 = n_0(z', \omega)$ so that for almost every ω and all $n > n_0$,

$$|\int \varphi(x) d\nu_n^{z'}(x)| \leq \frac{1}{n^{1+\frac{\alpha}{16}}} . \tag{31}$$

The last point proves the theorem as A_n has an eigenvalue in $B(z, \varepsilon) \subset \mathcal{B}$ for ε small enough only if

$$\nu_n^{z'}([-2\varepsilon, 2\varepsilon]) \geq \frac{1}{n}$$

for all $z' \in B(z, c\varepsilon)$, for an appropriate $c = c(M, z)$. (31) shows that this is impossible for n sufficiently large, almost surely.

Proof. The first point of the lemma is a restatement of the first part of Corollary 8. For the second, recall that any matrix W_n in the unitary group can be decomposed as $W_n = e^{i\theta} S_n$ with S_n in the special unitary group $SU(n)$ and note that multiplying S_n by $e^{i\theta}$ amounts to rotating z by $e^{i\theta}$ in H_n^z. Therefore, by the Chebyshev inequality we deduce from the first point that the set R_n of $\theta \in [0, 2\pi]$ such that

$$|E_{SU(n)}[\int \varphi(x) d\nu_n^{e^{i\theta} z}(x)]| \leq n^{-1-\frac{\alpha}{2}} \tag{32}$$

satisfies $|R_n|/2\pi \geq 1 - Cn^{-\alpha/2}$, where $|R_n|$ denotes the Lebesgue measure of R_n. Thus, in any interval of width $n^{-\alpha/2}$ in the circle of radius $|z|$ there is at least an element of R_n. We finally cover the compact set $\mathcal{B} \cap [0, M]$ (with M as in (1)) with a covering with mesh $\delta/2$ to obtain the existence of a family $(z_i)_{i\geq 0}$ of points of $\mathcal{B}$ so that (32) hold. Repeating this argument with the function $\varphi'(x)^2$, we also have that

$$|E_{SU(n)}[\int \varphi'(x)^2 d\nu_n^{z_i}(x)]| \leq Cn^{-1-\frac{\alpha}{2}}\,. \tag{33}$$

Next, remark that $U_n \to \int \varphi(x) d\nu_n^{z_i}(x)$ is Lipschitz with constant bounded above by $C\left(n^{-1}\int \varphi'(x)^2 d\nu_n^{z_i}(x)\right)^{\frac{1}{2}}$. Set $C_n = \{W_n \in SU(n) : \int \varphi'(x)^2 d\nu_n^{z_i}(x) \leq n^{-\frac{\alpha}{4}}\}$. Then,

$$P(C_n^c) \leq Cn^{-1-\alpha/4}\,. \tag{34}$$

Consequently, using (33),

$$E_{SU(n)}[1_{C_n}\int \varphi(x) d\nu_n^{z_i}(x)] \leq Cn^{-1-\alpha/4}\,.$$

Therefore, we get that for all n large enough,

$$\begin{aligned}
& P\left(\left|\int \varphi(x) d\nu_n^{z_i}(x)\right| \geq n^{-1-\frac{\alpha}{16}}\right) \\
\leq\ & P\left(\left|\int \varphi(x) d\nu_n^{z_i}(x) - E_{SU(n)}[1_{C_n}\int \varphi(x) d\nu_n^{z_i}(x)]\right| \geq \frac{1}{2}n^{-1-\frac{\alpha}{16}}\right) \\
\leq\ & n^{-1-\frac{\alpha}{4}} \\
& + P\left(\left\{\left|\int \varphi(x) d\nu_n^{z_i}(x) - E_{SU(n)}[1_{C_n}\int \varphi(x) d\nu_n^{z_i}(x)]\right| \geq \frac{1}{2}n^{-1-\frac{\alpha}{16}}\right\} \cap C_n\right) \\
\leq\ & Cn^{-1-\frac{\alpha}{4}} + Ce^{-n^{-2-\frac{\alpha}{8}}n^2 n^{\frac{\alpha}{2}}}\,,
\end{aligned}$$

where we have applied [1, Theorem 4.4.27] to the extension of the function $W_n \to g(W_n) = \int \varphi(x) d\nu_n^{z_i}(x)$ outside C_n which is globally Lipschitz with constant $n^{-\frac{1}{2}-\frac{\alpha}{4}}$ and uniformly bounded, see e.g. [5, Section 5.4] for the existence of such extension. Applying the Borel-Cantelli lemma completes the proof. □

Acknowledgement This paper was written while both authors participated in the 2010 MSRI program on random matrix theory, interacting particle systems and integrable systems. We thank MSRI for its hospitality.

References

[1] Anderson, G. W., Guionnet, A. and Zeitouni, O., *An introduction to random matrices*, Cambridge University Press, Cambridge (2010).

[2] Brown, L. G., *Lidskii's theorem in the type II case*, in "Proceedings U.S.–Japan, Kyoto/Japan 1983", Pitman Res. Notes. Math Ser. **123**, 1–35, (1983).

[3] Feinberg, J. and Zee, A., *Non-Gaussian non-Hermitian random matrix theory: phase transition and addition formalism*, Nuclear Phys. B **501**, 643–669, (1997).

[4] Girko, V. L., *The circular law*, Teor. Veroyatnost. i Primenen. **29**, 669–679, (1984).

[5] Guionnet. A, *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.

[6] Guionnet, A., Krishnapur, M. and Zeitouni, O., *The single ring theorem*, arXiv:0909.2214v2 (2009).

[7] Haagerup, U. and Larsen, F., *Brown's spectral distribution measure for R-diagonal elements in finite von Neumann algebras*, J. Funct. Anal. **2**, 331–367, (2000).

[8] Haagerup, U. and Thorbjørnsen, S. *A new application of random matrices:* $\mathrm{Ext}(C^*_{\mathrm{red}}(F_2))$ *is not a group*, Ann. of Math. (2) **162**, 711–775, (2005).

[9] Nica, A. and Speicher, R., *$\mathcal{R}$-diagonal pairs – a common approach to Haar unitaries and circular elements*, Fields Inst. Commun. **12**, 149–188 (1997).

[10] Rider, B., *A limit theorem at the edge of a non-Hermitian random matrix ensemble*, J. Phys. A **36** (2003), pp. 3401–3409.

[11] Voiculescu, D., *Limit laws for random matrices and free products* Inventiones Mathematicae **104**, 201–220, (1991).