Convergence of the spectral measure of non normal matrices

Alice Guionnet[*] Philip Wood Ofer Zeitouni[†]

October 11, 2011. Revised February 28, 2012

Abstract

We discuss regularization by noise of the spectrum of large random non-normal matrices. Under suitable conditions, we show that the regularization of a sequence of matrices that converges in $*$-moments to a regular element a, by the addition of a polynomially vanishing Gaussian Ginibre matrix, forces the empirical measure of eigenvalues to converge to the Brown measure of a.

1 Introduction

We discuss in this paper the changes in the spectrum of matrices when they are perturbed by noise. The behavior of the spectrum of matrices under (small) perturbation is not well understood when the matrices involved are non-normal, see [7] for background. A standard example of the issues in-

[*]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.

volved (described e.g. in [6]), is the following. Consider the nilpotent matrix

$$T_N = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ 0 & \ddots & \ddots & 0 & 1 \\ 0 & \ddots & \ddots & \ddots & 0 \end{pmatrix}. \tag{1}$$

The eigenvalues of T_N all vanish. However, adding to T_N the matrix P_N whose entries are all 0 except for the bottom left, which is taken as $\varepsilon_N \to 0$, changes the eigenvalues drastically - as we will see below, as N increases, if ε_N decays polynomially, the empirical measure of eigenvalues converges to the uniform measure on the unit circle in the complex plane. A natural question is, then, what happens when a small random matrix is added to T_N. One of the consequences of our results, see Corollaries 6 and 8, is that the spectrum of $T_N + G_N$, where G_N is an additive Gaussian random perturbation that vanishes polynomially, behaves like the spectrum of $T_N + P_N$, while this is false for matrices $\tilde{T}_N$ that are themselves a low-rank perturbation of T_N. The distinction involves the study of *regular elements* in W^* probability spaces.

To state our results we need to introduce some terminology. Consider a sequence $(A_N)_{N\geq 1}$, where A_N is a $N \times N$ matrice of uniformly bounded operator norm, and assume that A_N converges in $*$-moments toward an element a in a W^* probability space $(\mathcal{A}, \|\cdot\|, *, \varphi)$, that is, for any non-commutative polynomial P,

$$\frac{1}{N}\mathrm{tr}P(A_N, A_N^*) \xrightarrow{N\to\infty} \varphi(P(a,a^*)).$$

We assume throughout that the tracial state φ is faithful; this does not represent a loss of generality. If A_N is a sequence of Hermitian matrices, this is enough in order to conclude that the empirical measure of eigenvalues of A_N, that is the measure

$$L_N^A := \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i(A_N)},$$

where $\lambda_i(A_N), i = 1 \ldots N$ are the eigenvalues of A_N, converges weakly to a limiting measure μ_a, the spectral measure of a, supported on a compact subset of $\mathbb{R}$. (See [1, Corollary 5.2.16, Lemma 5.2.19] for this standard result and further background.) Significantly, in the Hermitian case, this convergence is stable under small bounded perturbations: with $B_N = A_N + E_N$ and E_N Hermitian with $\|E_N\| < \varepsilon$, any subsequential limit of L_N^B will belong to $B_L(\mu_a, \delta(\varepsilon))$, with $\delta(\varepsilon) \xrightarrow{\varepsilon\to 0} 0$ and $B_L(\nu_a, r)$ is the ball (in say, the Lévy metric) centered at ν_a and of radius r.

Both these statements fail when A_N is not normal. Recall the matrix T_N, see (1). Obviously, $L_N^T = \delta_0$, while a simple computation reveals that T_N converges in $*$-moments to a unitary Haar element of $\mathcal{A}$, that is

$$\frac{1}{N}\mathrm{tr}(T_N^{\alpha_1}(T_N^*)^{\beta_1}\cdots T_N^{\alpha_k}(T_N^*)^{\beta_k}) \xrightarrow{N\to\infty} \begin{cases} 1, & \text{if} \sum_{i=1}^k \alpha_i = \sum_{i=1}^k \beta_i, \\ 0, & \text{otherwise}. \end{cases} \tag{2}$$

Further, adding to T_N the matrix whose entries are all 0 except for the bottom left, which is taken as ε, changes the empirical measure of eigenvalues drastically — as already mentioned, when N increases, the empirical measure converges to the uniform measure on the unit circle in the complex plane.

We explore this phenomenon in the context of small random perturbations of matrices. We recall some notions. For $a \in \mathcal{A}$, the *Brown measure* ν_a on $\mathbb{C}$ is the measure satisfying

$$\log\det(z-a) = \int \log|z-z'|d\nu_a(z'), \quad \mathbf{z} \in \mathbb{C},$$

where det is the Fuglede-Kadison determinant; we refer to [2, 4] for definitions. We have in particular that

$$\log\det(z-a) = \int \log x \, d\nu_a^z(x) \quad z \in \mathbb{C},$$

where ν_a^z denotes the spectral measure of the operator $|z-a|$. In the sense of distributions, we have

$$\nu_a = \frac{1}{2\pi}\Delta \log\det(z-a).$$

That is, for smooth compactly supported function ψ on $\mathbb{C}$,

$$\begin{aligned}\int \psi(z)d\nu_a(z) &= \frac{1}{2\pi}\int \Delta\psi(z)\left(\int \log|z-z'|\,d\nu_a(z')\right)dz \\ &= \frac{1}{2\pi}\int \Delta\psi(z)\left(\int \log x\, d\nu_a^z(x)\right)dz.\end{aligned}$$

A crucial assumption in our analysis is the following.

Definition 1 (Regular elements)**.** An element $a \in \mathcal{A}$ is *regular* if

$$\lim_{\varepsilon \to 0} \int_{\mathbb{C}} \Delta\psi(z) \left(\int_0^{\varepsilon} \log x \, d\nu_a^z(x) \right) dz = 0\,, \tag{3}$$

for all smooth functions ψ on $\mathbb{C}$ with compact support.

Note that regularity is a property of a, not merely of its Brown measure ν_a. It is easy to check that the unitary Haar element in $\mathcal{A}$ is regular, see Section 4, while elements with dense, purely atomic spectrum are not.

We next introduce the class of Gaussian perturbations we consider.

Definition 2 (Polynomially vanishing Gaussian matrices)**.** A sequence $(G_N)_{N\geq 1}$ with G_N an N-by-N random Gaussian matrix is called *polynomially vanishing* if the entries $(G_N(i,j))$ are independent centered complex Gaussian variables, and there exist $\kappa > 0$, $\kappa' \geq 1+\kappa$ so that

$$N^{-\kappa'} \leq \mathbf{E}|G_{ij}|^2 \leq N^{-1-\kappa}.$$

Remark 3. As will be clear below, see the beginning of the proof of Lemma 10, the Gaussian assumption only intervenes in obtaining a uniform lower bound on singular values of certain random matrices. As pointed out to us by R. Vershynin, this uniform estimate extends to other situations, most notably to the polynomial scaling of matrices whose entries are i.i.d. and possess a bounded density. We do not discuss such extensions here.

Our first result is a stability, with respect to polynomially vanishing Gaussian perturbations, of the convergence of spectral measures for non-normal matrices. Throughout, we denote by $\|M\|_{op}$ the operator norm of a matrix M.

Theorem 4. *Assume that the uniformly bounded (in the operator norm) sequence of N-by-N matrices A_N converges in $*$-moments to a regular element a. Assume further that L_N^A converges weakly to the Brown measure ν_a. Let G_N be a sequence of polynomially vanishing Gaussian matrices, and set $B_N = A_N + G_N$. Then, $L_N^B \to \nu_a$ weakly, in probability.*

Theorem 4 puts rather stringent assumptions on the sequence A_N. In particular, its assumptions are not satisfied by the sequence of nilpotent matrices T_N in (2). Our second result corrects this defficiency, by showing that small Gaussian perturbations "regularize" matrices that are close to matrices satisfying the assumptions of Theorem 4.

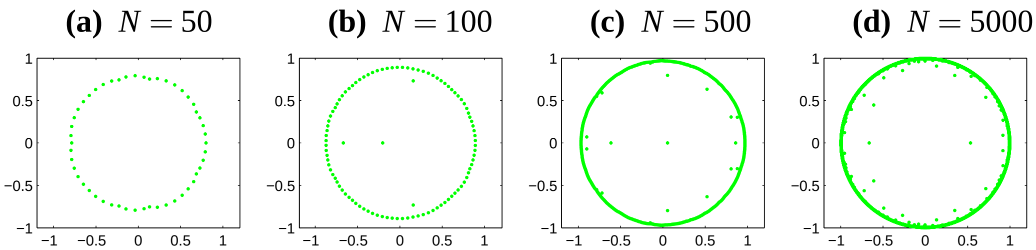

Figure 1: The eigenvalues of $T_N + N^{-3-1/2} G_N$, where G_N is iid complex Gaussian with mean 0, variance 1 entries.

Theorem 5. *Let A_N, E_N be a sequence of bounded (for the operator norm) N-by-N matrices, so that A_N converges in $*$-moments to a regular element a. Assume that $\|E_N\|_{op}$ converges to zero polynomially fast in N, and that $L_N^{A+E} \to \nu_a$ weakly. Let G_N be a sequence of polynomially vanishing Gaussian matrices, and set $B_N = A_N + G_N$. Then, $L_N^B \to \nu_a$ weakly, in probability.*

Theorem 5 should be compared to earlier results of Sniady [6], who used stochastic calculus to show that a perturbation by an asymptotically vanishing Ginibre Gaussian matrix regularizes arbitrary matrices. Compared with his results, we allow for more general Gaussian perturbations (both structurally and in terms of the variance) and also show that the Gaussian regularization can decay as fast as wished in the polynomial scale. On the other hand, we do impose a regularity property on the limit a as well as on the sequence of matrices for which we assume that adding a polynomially small matrix is enough to obtain convergence to the Brown measure.

A corollary of our general results is the following.

Corollary 6. *Let G_N be a sequence of polynomially vanishing Gaussian matrices and let T_N be as in* (2). *Then L_N^{T+G} converges weakly, in probability, toward the uniform measure on the unit circle in $\mathbb{C}$.*

In Figure 1, we give a simulation of the setup in Corollary 6 for various N.

We will now define a class of matrices $T_{b,N}$ for which, if b is chosen correctly, adding a polynomially vanishing Gaussian matrix G_N is not sufficient to regularize $T_{b,N} + G_N$. Let b be a positive integer, and define $T_{b,N}$ to be an N by N block diagonal matrix which each $b+1$ by $b+1$ block on the diagonal equal T_{b+1} (as defined in (2). If $b+1$ does not divide N evenly, a block of zeros is inserted at bottom of the diagonal. Thus, every entry of $T_{b,N}$ is zero except for entries on the superdiagonal (the superdiagonal is the list of entries with coordinates $(i, i+1)$ for $1 \le i \le N-1$), and the superdiagonal of $T_{b,N}$ is equal to

$$(\underbrace{1,1,\ldots,1}_{b},0,\underbrace{1,1,\ldots,1}_{b},0,\ldots,\underbrace{1,1,\ldots,1}_{b},\underbrace{0,0,\ldots,0}_{\leq b}).$$

Recall that the spectral radius of a matrix is the maximum absolute value of the eigenvalues. Also, we will use $\|A\| = \mathrm{tr}(A^*A)^{1/2}$ to denote the Hilbert-Schmidt norm.

Proposition 7. *Let $b = b(N)$ be a sequence of positive integers such that $b(N) \geq \log N$ for all N, and let $T_{b,N}$ be as defined above. Let R_N be an N by N matrix satisfying $\|R_N\| \leq g(N)$, where for all N we assume that $g(N) < \frac{1}{3b\sqrt{N}}$. Then*

$$\rho(T_{b,N} + R_N) \leq (Ng(N))^{1/b} + o(1),$$

where $\rho(M)$ denotes the spectral radius of a matrix M, and $o(1)$ denotes a small quantity tending to zero as $N \to \infty$.

Note that $T_{b,N}$ converges in $*$-moments to a Unitary Haar element of $\mathcal{A}$ (by a computation similar to (2)) if $b(N)/N$ goes to zero, which is a regular element. The Brown measure of the Unitary Haar element is uniform measure on the unit circle; thus, in the case where $(Ng(N))^{1/b} < 1$, Proposition 7 shows that $T_{b,N} + R_N$ does not converge to the Brown measure for $T_{b,N}$.

Corollary 8. *Let R_N be an iid Gaussian matrix where each entry has mean zero and variance one. Set $b = b(N) \geq \log N$ be a sequence of integers, and let $\gamma > 5/2$ be a constant. Then, with probability tending to 1 as $N \to \infty$, we have*

$$\rho(T_{b,N} + \exp(-\gamma b)R_N) \leq \exp\left(-\gamma + \frac{2\log N}{b}\right) + o(1),$$

where ρ denotes the spectral radius and where $o(1)$ denotes a small quantity tending to zero as $N \to \infty$. Note in particular that the bound on the spectral radius is strictly less than $\exp(-1/2) < 1$ in the limit as $N \to \infty$, due to the assumptions on γ and b.

Corollary 8 follows from Proposition 7 by noting that, with probability tending to 1, all entries in R_N are at most $C \log N$ in absolute value for some constant C, and then checking that the hypotheses of Proposition 7 are satisfied for $g(N) = \exp(-\gamma b)CN(\log N)^{1/4}$. There are two instances of Corollary 8 that are particularly interesting: when $b = N-1$, we see that a exponentially decaying Gaussian perturbation does not regularize $T_N = T_{N-1,N}$,

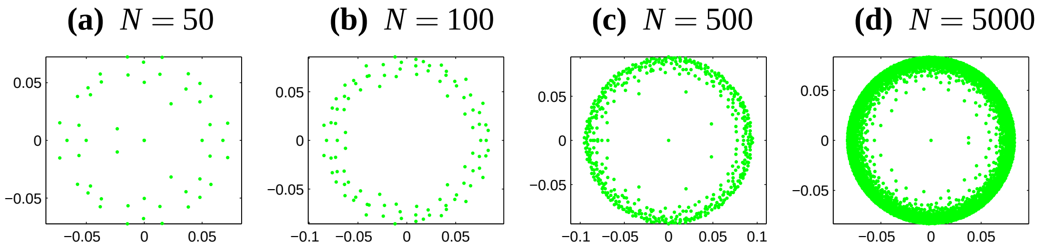

Figure 2: The eigenvaules of $T_{\log N,N} + N^{-3-1/2}G_N$, where G_N is iid complex Gaussian with mean 0, variance 1 entries. The spectral radius is roughly 0.07, and the bound from Corollary 8 is $\exp(-1) \approx 0.37$.

and when $b = \log(N)$, we see that polynomially decaying Gaussian perturbation does not regularize $T_{\log N,N}$ (see Figure 2).

We will prove Proposition 7 in Section 5. The proof of our main results (Theorems 4 and 5) borrows from the methods of [3].

We finally introduce some notation that will be used throughout. For any N-by-N matrix C_N, let

$$\widetilde{C}_N = \begin{pmatrix} 0 & C_N \\ C_N^* & 0 \end{pmatrix}.$$

We denote by G_C the Cauchy-Stieltjes transform of the spectral measure of the matrix $\widetilde{C}_N$, that is

$$G_C(z) = \frac{1}{2N}\mathrm{tr}(z - \widetilde{C}_N)^{-1}, \quad z \in \mathbb{C}_+ .$$

The following useful estimate is immediate from the definition and the resolvent identity:

$$|G_C(z) - G_D(z)| \leq \frac{\|C - D\|_{op}}{|\Im z|^2} . \tag{4}$$

2 Proof of Theorem 4

We keep throughout the notation and assumptions of the theorem. The following is a crucial simple observation.

Proposition 9. *For all complex number ξ, and all z so that $\Im z \geq N^{-\delta}$ with $\delta < \kappa/4$,*

$$\mathbf{E}|\Im G_{B_N+\xi}(z)| \leq \mathbf{E}|\Im G_{A_N+\xi}(z)| + 1 .$$

Proof. Noting that

$$\mathbf{E}\|B_N - A_N\|_{op}^k = \mathbf{E}\|G_N\|_{op}^k \le C_k N^{-\kappa k/2}, \tag{5}$$

the conclusion follows from (4) and Hölder's inequality. $\square$

We continue with the proof of Theorem 4. Let $\nu^z_{A_N}$ denote the empirical measure of the eigenvalues of the matrix $\widetilde{A_N - z}$. We have that, for smooth test functions ψ,

$$\int \Delta\psi(z)\left(\int \log|x|\, d\nu^z_{A_N}(x)\right) dz = \frac{1}{2\pi}\int \psi(z) dL^A_N(z)\,.$$

In particular, the convergence of L^A_N toward ν_a implies that

$$\begin{aligned}&\mathbf{E}\int \Delta\psi(z)\left(\int \log|x| d\nu^z_{A_N}(x)\right) dz \\ &\to \int \psi(z) d\nu_a(z) = \int \Delta\psi(z)\left(\int \log x d\nu^z_a(x)\right) dz\,.\end{aligned}$$

On the other hand, since $x \mapsto \log x$ is bounded continuous on compact subsets of $(0,\infty)$, it also holds that for any continuous bounded function $\zeta : \mathbb{R}_+ \mapsto \mathbb{R}$ compactly supported in $(0,\infty)$,

$$\mathbf{E}\int \Delta\psi(z)\left(\int \zeta(x)\log x d\nu^z_{A_N}(x)\right) dz \to \int \Delta\psi(z)\left(\int \zeta(x)\log x d\nu^z_a(x)\right) dz\,.$$

Together with the fact that a is regular and that A_N is uniformly bounded, one concludes therefore that

$$\lim_{\varepsilon\downarrow 0}\lim_{N\to\infty}\mathbf{E}\int\left(\int_0^\varepsilon \log|x| d\nu^z_{A_N}(x)\right) dz = 0\,.$$

Our next goal is to show that the same applies to B_N. In the following, we let $\nu^z_{B_N}$ denote the empirical measure of the eigenvalues of $\widetilde{B_N - z}$.

Lemma 10.

$$\lim_{\varepsilon\downarrow 0}\lim_{N\to\infty}\int \mathbf{E}\left[\int_0^\varepsilon \log|x|^{-1} d\nu^z_{B_N}(x)\right] dz = 0$$

Because $\mathbf{E}\|B_N - A_N\|_{op}^k \to 0$ for any $k > 0$, we have for any fixed smooth w compactly supported in $(0,\infty)$ that

$$\mathbf{E}\left|\int \Delta\psi(z)\left(\int w(x)\log x d\nu^z_{A_N}(x)\right) dz - \int \Delta\psi(z)\left(\int w(x)\log x d\nu^z_{B_N}(x)\right) dz\right| \xrightarrow{N\to\infty} 0\,,$$

Theorem 4 follows at once from Lemma 10.

Proof of Lemma 10: Note first that by [5, Theorem 3.3] (or its generalization in [3, Proposition 16] to the complex case), there exists a constant C so that for any z, the smallest singular value σ_N^z of $B_N + zI$ satisfies

$$P(\sigma_N^z \le x) \le C\left(N^{\frac{1}{2}+\kappa'}x\right)^{\beta}$$

with $\beta = 1$ or 2 according whether we are in the real or the complex case. Therefore, for any $\zeta > 0$, uniformly in z

$$\begin{aligned}\mathbf{E}\left[\int_0^{N^{-\zeta}} \log|x|^{-1} d\nu^z_{B_N}(x)\right] &\le \mathbf{E}[\log(\sigma_N^z)^{-1} 1_{\sigma_N^z \le N^{-\zeta}}] \\ &= C\left(N^{\frac{1}{2}+\kappa'-\zeta}\right)^{\beta}\log(N^{\zeta}) + \int_0^{N^{-\zeta}} \frac{1}{x} C\left(N^{\frac{1}{2}+\kappa'}x\right)^{\beta} dx\end{aligned}$$

goes to zero as N goes to infinity as soon as $\zeta > \frac{1}{2} + \kappa'$. We fix hereafter such a ζ and we may and shall restrict the integration from $N^{-\zeta}$ to ε. To compare the integral for the spectral measure of A_N and B_N, observe that for any probability measure P, with P_γ the Cauchy law with parameter γ

$$P([a,b]) \le \left(P * P_\gamma\right)([a-\eta, b+\eta]) + P_\gamma([-\eta,\eta]^c) \le \left(P * P_\gamma\right)([a-\eta, b+\eta]) + \frac{\gamma}{\eta} \tag{6}$$

whereas for $b - a > \eta$

$$P([a,b]) \ge \left(P * P_\gamma\right)([a+\eta, b-\eta]) - \frac{\gamma}{\eta}. \tag{7}$$

Recall that

$$\left(P * P_\gamma\right)([a,b]) = \int_a^b |\Im G(x+i\gamma)| dx. \tag{8}$$

Set $\gamma = N^{-\kappa/5}$, $\kappa'' = \kappa/2$ and $\eta = N^{-\kappa''/5}$. We have, whenever $b - a \ge 4\eta$,

$$\begin{aligned}\mathbf{E}\nu^z_{B_N}([a,b]) &\le \int_{a-\eta}^{b+\eta} \mathbf{E}|\Im G_{B_n+z}(x+i\gamma)| dx + N^{-(\kappa-\kappa'')/5} \\ &\le (b-a+2N^{-\kappa''/5}) + \nu^z_{A_N} * P_{N^{-\kappa/5}}([a-N^{-\kappa/10}, b+N^{-\kappa/10}]) + N^{-\kappa/10} \\ &\le (b-a+2N^{-\kappa/10}) + \nu^z_{A_N}([(a-2N^{-\kappa/10})_+, (b+2N^{-\kappa/10})]) + 2N^{-\kappa/10},\end{aligned}$$

where the first inequality is due to (6) and (8), the second is due to Proposition 9, and the last uses (7) and (8). Therefore, if $b - a = CN^{-\kappa/10}$ for some fixed C larger than 4, we deduce that there exists a finite constant C' which only depends on C so that

$$\mathbf{E}\nu^z_{B_N}([a,b]) \le C'(b-a) + \nu^z_{A_N}([(a-2N^{-\kappa/10})_+,(b+2N^{-\kappa/10})]) .$$

As a consequence, as we may assume without loss of generality that $\kappa' > \kappa/10$,

$$\begin{aligned} & \mathbf{E}[\int_{N^{-\zeta}}^{\varepsilon} \log|x|^{-1} d\nu^z_{B_N}(x)] \\ \le & \sum_{k=0}^{[N^{\kappa/10}\varepsilon]} \log(N^{-\zeta}+2CkN^{-\kappa/10})^{-1} \\ & \cdot \mathbf{E}[\nu^z_{B_N}]([N^{-\zeta}+2CkN^{-\kappa/10}, N^{-\zeta}+2C(k+1)N^{-\kappa/10}]) . \end{aligned}$$

We need to pay special attention to the first term that we bound by noticing that

$$\begin{aligned} & \log(N^{-\zeta})^{-1}\mathbf{E}[\nu^z_{B_N}([N^{-\zeta}, N^{-\zeta}+2CN^{-\kappa/10}])] \\ \le & \frac{10\zeta}{\kappa}\log(N^{-\kappa/10})^{-1}\mathbf{E}[\nu^z_{B_N}([0,2(C+1)N^{-\kappa/10}])] \\ \le & \frac{10\zeta}{\kappa}\log(N^{-\kappa/10})^{-1}(2C'N^{-\kappa/10}+\nu^z_{A_N}([0,(C+2)N^{-\kappa/10}])) \\ \le & \frac{20C'\zeta}{\kappa}\log(N^{-\kappa/10})^{-1}N^{-\kappa/10} + C''\int_0^{2(C+2)N^{-\kappa/10}} \log|x|^{-1} d\nu^z_{A_N}(x) . \end{aligned}$$

For the other terms, we have

$$\begin{aligned} & \sum_{k=1}^{[N^{\kappa/10}\varepsilon]} \log(N^{-\zeta}+2CkN^{-\kappa/10})^{-1} \\ & \cdot \mathbf{E}[\nu^z_{B_N}]([N^{-\zeta}+2CkN^{-\kappa/10}, N^{-\zeta}+2C(k+1)N^{-\kappa/10}]) \\ \le & 2C' \sum_{k=1}^{[N^{\kappa/10}\varepsilon]} \log(CkN^{-\kappa/10})^{-1} CN^{-\kappa/10} \\ & + \sum_{k=1}^{[N^{\kappa/10}\varepsilon]} \log(CkN^{-\kappa/10})^{-1} \nu^z_{A_N}([2C(k-1)N^{-\kappa/10}, 2C(k+2)N^{-\kappa/10}]) . \end{aligned}$$

Finally, we can sum up all these inequalities to find that there exists a finite constant C''' so that

$$\mathbf{E}\left[\int_{N^{-\zeta}}^{\varepsilon} \log|x|^{-1} d\nu^z_{B_N}(x)\right] \le C'''\int_0^{\varepsilon} \log|x|^{-1} d\nu^z_{A_N}(x) + C'''\int_0^{\varepsilon} \log|x|^{-1} dx$$

and therefore goes to zero when n and then ε goes to zero. This proves the claim. $\square$

3 Proof of Theorem 5

From the assumptions, it is clear that (A_N+E_N) converges in $*$-moments to the regular element a. By Theorem 4, it follows that L_N^{A+E+G} converges (weakly, in probability) towards ν_a. We can now remove E_N. Indeed, by (4) and (5), we have for any $\chi<\kappa'/2$ and all $\xi\in\mathbb{C}$

$$|G^N_{A+G+\xi}(z)-G^N_{A+G+E+\xi}(z)|\leq\frac{N^{-\chi}}{\Im z^2}$$

and therefore for $\Im z\geq N^{-\chi/2}$,

$$|\Im G^N_{A+G+\xi}(z)|\leq|\Im G^N_{A+G+E+\xi}(z)|+1.$$

Again by [5, Theorem 3.3] (or its generalization in [3, Proposition 16]) to the complex case), for any z, the smallest singular value σ_N^z of A_N+G_N+z satisfies

$$P(\sigma_N^z\leq x)\leq C\left(N^{\frac{1}{2}+\kappa'}x\right)^\beta$$

with $\beta=1$ or 2 according whether we are in the real or the complex case. We can now rerun the proof of Theorem 4, replacing A_N by $A'_N=A_N+E_N+G_N$ and B_N by A'_N-E_N. $\square$

4 Proof of Corollary 6

We apply Theorem 5 with $A_N=T_N$, E_N the N-by-N matrix with

$$E_N(i,j)=\begin{cases}\delta_N=N^{-(1/2+\kappa')}, & i=1,j=N,\\ 0, & \text{otherwise},\end{cases}$$

where $\kappa'>\kappa$. We check the assumptions of Theorem 5. We take a to be a unitary Haar element in $\mathcal{A}$, and recall that its Brown measure ν_a is the uniform measure on $\{z\in\mathbb{C}:|z|=1\}$. We now check that a is regular. Indeed, $\int x^k d\nu_a^z(x)=0$ if k is odd by symmetry while for k even,

$$\int x^k d\nu_a^z(x)=\varphi([(z-a)(z-a)^*]^{k/2})=\sum_{j=1}^{k/2}(|z|^2+1)^{k-j}\left(\begin{array}{c}k\\2j\end{array}\right)\left(\begin{array}{c}2j\\j\end{array}\right),$$

and one therefore verifies that for k even,

$$\int x^k d\nu_a^z(x) = \frac{1}{2\pi}\int (|z|^2 + 1 + 2|z|\cos\theta)^{k/2} d\theta\,.$$

It follows that

$$\int_0^{\varepsilon} \log x d\nu_a^z(x) = \frac{1}{4\pi}\int_0^{2\pi} \log(|z|^2+1+2|z|\cos\theta)\mathbf{1}_{\{|z|^2+1+2|z|\cos\theta<\varepsilon\}} d\theta \xrightarrow{\varepsilon\to 0} 0\,,$$

proving the required regularity.

Further, we claim that L_N^{A+E} converges to ν_a. Indeed the eigenvalues λ of $A_N + E_N$ are such that there exists a non-vanishing vector u so that

$$u_N \delta_N = \lambda u_1, u_{i-1} = \lambda u_i\,,$$

that is

$$\lambda^N = \delta_N.$$

In particular, all the N-roots of δ_N are (distinct) eigenvalues, that is the eigenvalues λ_j^N of A_N are

$$\lambda_j^N = |\delta_N|^{1/N} e^{2i\pi j/N}, \quad 1 \le j \le N\,.$$

Therefore, for any bounded continuous g function on $\mathbb{C}$,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} g(\lambda_j^N) = \frac{1}{2\pi}\int g(\theta) d\theta\,,$$

as claimed. □

5 Proof of Proposition 7

In this section we will prove the following proposition:

Proposition 11. *Let $b = b(N)$ be a sequence of positive integers, and let $T_{b,N}$ be as in Proposition 7. Let R_N be an N by N matrix satisfying $\|R_N\| \le g(N)$, where for all N we assume that $g(N) < \frac{1}{3b\sqrt{N}}$. Then*

$$\rho(T_{b,N} + R_N) \le \left(O\left(\sqrt{Nb}\left(2N^{1/4}g^{1/2}\right)^b\right)\right)^{1/(b+1)} + \left(b^2 N g\right)^{1/(b+1)}\,.$$

Proposition 7 follows from Proposition 11 by adding the assumption that $b(N) \geq \log(N)$ and then simplifying the upper bound on the spectral radius.

Proof of Proposition 11: To bound the spectral radius, we will use the fact that $\rho(T_{b,N} + R_N) \leq \left\|(T_{b,N} + R_N)^k\right\|^{1/k}$ for all integers $k \geq 1$. Our general plan will be to bound $\left\|(T_{b,N} + R_N)^k\right\|$ and then take a k-th root of the bound. We will take $k = b+1$, which allows us to take advantage of the fact that $T_{b,N}$ is $(b+1)$-step nilpotent. In particular, we make use of the fact that for any positive integer a,

$$\|T_{b,N}^a\| = \begin{cases} (b-a+1)^{1/2} \left\lfloor \frac{N}{b+1} \right\rfloor^{1/2} & \text{if } 1 \leq a \leq b \\ 0 & \text{if } b+1 \leq a. \end{cases} \tag{9}$$

We may write

$$\begin{aligned} \left\|(T_{b,N} + R_N)^{b+1}\right\| &\leq \sum_{\lambda \in \{0,1\}^{b+1}} \left\| \prod_{i=1}^{b+1} T_{b,N}^{\lambda_i} R_N^{1-\lambda_i} \right\| \\ &= \sum_{\ell=0}^{b+1} \sum_{\substack{\lambda \in \{0,1\}^{b+1} \\ \lambda \text{ has } \ell \text{ ones}}} \left\| \prod_{i=1}^{b+1} T_{b,N}^{\lambda_i} R_N^{1-\lambda_i} \right\|. \end{aligned}$$

When ℓ is large, we will make use of the following lemma.

Lemma 12. *If $\lambda \in \{0,1\}^k$ has ℓ ones and $\ell \geq (k+1)/2$, then*

$$\left\| \prod_{i=1}^{k} T_{b,N}^{\lambda_i} R_N^{1-\lambda_i} \right\| \leq \left\| T_{b,N}^{\left\lfloor \frac{\ell}{k-\ell+1} \right\rfloor} \right\|^{k-\ell+1} \|R_N\|^{k-\ell}.$$

We will prove Lemma 12 in Section 5.1. Using Lemma 12 with $k = b+1$

along with the fact that $\|AB\| \leq \|A\|\,\|B\|$, we have

$$\begin{aligned}
\left\|(T_{b,N}+R_N)^{b+1}\right\| &\leq \sum_{\ell=0}^{\left\lfloor\frac{b+2}{2}\right\rfloor} \binom{b+1}{\ell} \|T_{b,N}\|^{\ell}\, \|R_n\|^{b-\ell+1} \\
&\qquad + \sum_{\ell=\left\lceil\frac{b+2}{2}\right\rceil}^{b+1} \binom{b+1}{\ell} \left\|T_{b,N}^{\left\lfloor\frac{\ell}{b-\ell+2}\right\rfloor}\right\|^{b-\ell+2} \|R_N\|^{b-\ell+1} \\
&\leq \sum_{\ell=0}^{\left\lfloor\frac{b+2}{2}\right\rfloor} \binom{b+1}{\ell} \|T_{b,N}\|^{\ell}\, g^{b-\ell+1} \qquad (10) \\
&\qquad + \sum_{\ell=\left\lceil\frac{b+2}{2}\right\rceil}^{b+1} \binom{b+1}{\ell} \left\|T_{b,N}^{\left\lfloor\frac{\ell}{b-\ell+2}\right\rfloor}\right\|^{b-\ell+2} g^{b-\ell+1}, \qquad (11)
\end{aligned}$$

where the second inequality comes from the assumption $\|R_N\| \leq g = g(N)$.

We will bound (10) and (11) separately. To bound (10) note that

$$\begin{aligned}
\sum_{\ell=0}^{\left\lfloor\frac{b+2}{2}\right\rfloor} \binom{b+1}{\ell} \|T_{b,N}\|^{\ell}\, g^{b-\ell+1} &\leq \sum_{\ell=0}^{\left\lfloor\frac{b+2}{2}\right\rfloor} \binom{b+1}{\ell} \left((b+1)\left\lfloor\frac{N}{b+1}\right\rfloor\right)^{\ell/2} g^{b-\ell+1} \\
&\leq \frac{b+4}{2}\binom{b+1}{\lfloor (b+1)/2\rfloor} N^{(b+2)/4} g^{b/2} \\
&= O\left(\sqrt{Nb}(2N^{1/4}g^{1/2})^{b}\right). \qquad (12)
\end{aligned}$$

Next, we turn to bounding (11). We will use the following lemma to show that the largest term in the sum (11) comes from the $\ell = b$ term. Note that when $\ell = b+1$, the summand in (11) is equal to zero by (9).

Lemma 13. *If* $\left\|T_{b,N}^{\left\lfloor\frac{\ell+1}{b-\ell+1}\right\rfloor}\right\| > 0$ *and* $\ell \leq b-1$ *and*

$$g \leq \frac{2}{e^{3/2}N^{1/2}b},$$

then

$$\binom{b+1}{\ell} \left\|T_{b,N}^{\left\lfloor\frac{\ell}{b-\ell+2}\right\rfloor}\right\|^{b-\ell+2} g^{b-\ell+1} \leq \binom{b+1}{\ell+1} \left\|T_{b,N}^{\left\lfloor\frac{\ell+1}{b-\ell+1}\right\rfloor}\right\|^{b-\ell+1} g^{b-\ell}.$$

We will prove Lemma 13 in Section 5.1.

Using Lemma 13 we have

$$\begin{aligned}
\sum_{\ell=\lceil\frac{b+2}{2}\rceil}^{b+1}\binom{b+1}{\ell}\left\|T_{b,N}^{\lfloor\frac{\ell}{b-\ell+2}\rfloor}\right\|^{b-\ell+2}g^{b-\ell+1} &\leq \frac{b}{2}(b+1)\left\|T_{b,N}^{\lfloor\frac{b}{2}\rfloor}\right\|^{2}g^{1}\\
&\leq \frac{b}{2}(b+1)(b-\lfloor b/2\rfloor+1)\frac{N}{b+1}g\\
&\leq b^{2}Ng. \qquad (13)
\end{aligned}$$

Combining (12) and (13) with (10) and (11), we may use the fact that $(x+y)^{1/(b+1)}\leq x^{1/(b+1)}+y^{1/(b+1)}$ for positive x,y to complete the proof of Proposition 11. It remains to prove Lemma 12 and Lemma 13, which we do in Section 5.1 below. $\square$

5.1 Proofs of Lemma 12 and Lemma 13

Proof of Lemma 12. Using (9), it is easy to show that

$$\left\|T_{b,N}^{a}\right\|\left\|T_{b,N}^{c}\right\|<\left\|T_{b,N}^{a-1}\right\|\left\|T_{b,N}^{c+1}\right\| \text{ for integers } 3\leq c+2\leq a\leq b. \qquad (14)$$

It is also clear from (9) that

$$\left\|T_{b,N}^{a}\right\|\leq\left\|T_{b,N}^{a-1}\right\| \text{ for all positive integers } a. \qquad (15)$$

Let $\lambda\in\{0,1\}^k$ have ℓ ones. Then, using the assumption that $\ell\geq k-\ell+1$, we may write

$$\prod_{i=1}^{k}T_{b,N}^{\lambda_i}R_N^{1-\lambda_i}=T_{b,N}^{a_1}R_N^{b_1}T_{b,N}^{a_2}R_N^{b_2}\cdots T_{b,N}^{a_{k-\ell}}R_N^{b_{k-\ell}}T_{b,N}^{a_{k-\ell+1}},$$

where $a_i\geq 1$ for all i and $b_i\geq 0$ for all i. Thus

$$\left\|\prod_{i=1}^{k}T_{b,N}^{\lambda_i}R_N^{1-\lambda_i}\right\|\leq\|R_N\|^{k-\ell}\prod_{i=1}^{k-\ell+1}\left\|T_{b,N}^{a_i}\right\|.$$

Applying (14) repeatedly, we may assume that two of the a_i differ by more than 1, all without changing the fact that $\sum_{i=1}^{k-\ell+1}a_i=\ell$. Thus, some of the a_i are equal to $\lfloor\frac{\ell}{k-\ell+1}\rfloor$ and some are equal to $\lceil\frac{\ell}{k-\ell+1}\rceil$. Finally, applying (15), we have that

$$\prod_{i=1}^{k-\ell+1}\left\|T_{b,N}^{a_i}\right\|\leq\left\|T_{b,N}^{\lfloor\frac{\ell}{k-\ell+1}\rfloor}\right\|^{k-\ell+1}.$$

$\square$

Proof of Lemma 13. Using (9) and rearranging, it is sufficient to show that

$$\frac{\ell+1}{b-\ell+1}\left(b-\left\lfloor\frac{\ell}{b-\ell+2}\right\rfloor+1\right)^{1/2}\left\lfloor\frac{N}{b+1}\right\rfloor^{1/2}g\leq\left(\frac{b-\lfloor\frac{\ell+1}{b-\ell+1}\rfloor+1}{b-\lfloor\frac{\ell}{b-\ell+2}\rfloor+1}\right)^{\frac{b-\ell+1}{2}}.$$

Using a variety of manipulations, it is possible to show that

$$\begin{aligned}&\left(\frac{b-\lfloor\frac{\ell+1}{b-\ell+1}\rfloor+1}{b-\lfloor\frac{\ell}{b-\ell+2}\rfloor+1}\right)^{\frac{b-\ell+1}{2}}\\ \geq\quad&\exp\left(-\frac{(b-\ell+2)(b-\ell+1)}{(b+2)(b-\ell+2)-\ell}-\frac{b+2}{(b+2)(b-\ell+2)-\ell}\right)\\ \geq\quad&\exp(-3/2).\end{aligned}$$

Thus, it is sufficient to have

$$\frac{b}{2}N^{1/2}g\leq\exp(-3/2),$$

which is true by assumption. $\square$

References

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

[2] Brown, L. G., *Lidskiı'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] Guionnet, A., Krishnapur, M. and Zeitouni, O., *The single ring theorem*, Annals Math. **174**, 1189–1217, (2011).

[4] 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).

[5] Sankar, A., Spielman, D. A. and Teng, S.-H., *Smoothed analysis of the conditioning number and growth factor of matrices*, SIAM J. Matrix Anal. **28**, 446–476, (2006).

[6] Śniady, P., *Random regularization of Brown spectral measure*, J. Funct. Anal. **193** (2002), pp. 291–313.

[7] Trefethen, L. N. and Embree, M., *Spectra and pseudospectra: the behavior of nonnormal matrices and operators*, Princeton University Press, Princeton N.J. (2005).

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