LIMITING SPECTRAL DISTRIBUTION OF SUM OF UNITARY MATRICES

ANIRBAN BASAK* AND AMIR DEMBO†

Abstract. In this short note we consider the empirical distribution of the eigenvalues of the sum of $d$ independent Haar distributed Unitary matrices, and show that this sequence of measures converge to the Brown measure of free sum of $d$ unitary operators. As a consequence of our proof we improve the uniform bounded condition on the imaginary part of the Stieltjes transform in [4, Theorem 1].

1. Introduction

Random matrix models can be broadly divided into two regimes — Hermitian matrix models, and non-Hermitian matrix models. Many methods, like the method of moments, and the Stieltjes transform work well in the hermitian regime, and provide a good understanding of the behavior of the spectrum for many different matrix models. However both of these methods fail in the non-hermitian regime, and understanding the behavior of the spectral measure in any given model is much more difficult.

In an article by Girko [3], a nice scheme was proposed to find the limiting spectral distribution of non-hermitian matrix model. However carrying out this scheme, even in the simplest model, where all the entries are i.i.d. with mean 0 and finite variance, is extremely challenging. After a long sequence of partial results (see references in [2]), the circular law conjecture, for the i.i.d. case, was recently established by Tao and Vu [10] in full generality. Barring this simple model very few results are known in the non-hermitian regime. For example, nothing is known about the spectrum of random oriented $d$-regular graphs. It was recently conjectured in [2] that the empirical distribution of the adjacency matrix of a random oriented $d$-regular graph sampled uniformly converges to a measure $\mu$ on the complex plane, which has a density

$$\frac{1}{\pi} \frac{d^2(d-1)}{(d^2-|z|^2)^2} \mathbb{1}_{\{|z| \leq \sqrt{d}\}}$$

with respect to the lebesgue measure on $\mathbb{C}$. This conjecture is due to the observation that $\mu$ is the Brown measure of the free sum of $d$ unitary operators (see [6, Example 5.5]), and motivated us to consider a simpler problem, about the sum of $d$ independent unitary matrices. For the latter, we prove the following:

Theorem 1.1. For any positive integer $d$, the empirical distribution of the eigenvalues of the sum of $d$ independent, Haar distributed unitary matrices, converges weakly, in probability, to the Brown measure of free sum of $d$ unitary operators.

Date: August 28, 2012.

2010 Mathematics Subject Classification. 46L53, 60B10, 60B20.

Key words and phrases. Random matrices, limiting spectral distribution, Haar measure, Brown measure, free convolution, Stieltjes transform, Schwinger-Dyson equation.

*Research supported by Melvin and Joan Lane endowed Stanford Graduate Fellowship Fund.
†Research partially supported by NSF grant DMS-1106627.
It is worthy to note that if \( \{ U_n^i \}_{i=1}^d \) is a sequence of independent Haar distributed unitary matrices, then it converges in \( \ast \)-moments (see [9] for a definition) to \( \{ u_i \}_{i=1}^d \), where \( u_i \)'s are free Haar unitary operators (see [1, Theorem 5.4.10]). However Brown measure being discontinuous, convergence in \( \ast \)-moments does not necessarily imply the convergence of Brown measures\(^1\) (see [9, §2.6]). Nevertheless, it is proved in [9, Theorem 6] that if the original matrices are perturbed by adding small Gaussian (of unknown variance), then the Brown measures do converge. But removing the Gaussian, and even identifying the correct variance, is often a very hard task. For example in [5, Proposition 7, Corollary 8] it is shown that for a particular ensemble adding polynomially vanishing Gaussian is not sufficient to regularize the Brown measure. Also generalizing the notion of convergence in \( \ast \)-moments to the notion of convergence in distribution of traffics [7], does not help to remove the Gaussian. We circumvent these obstacles, and in Theorem 1.1 we essentially show that the original matrices are smooth enough, and therefore there is no need to add any Gaussian in this case to get convergence of the corresponding Brown measure.

To prove Theorem 1.1 we adopt the techniques developed in [4]. There one relies on the uniform boundedness of the imaginary part of the Stieltjes transform (see [4, Theorem 1, Eqn. 3]). Here, the expected empirical distribution of the eigenvalues of \( U_n - zI_n \) has unbounded density (see Lemma 3.1), and therefore the imaginary part of its Stieltjes transform is unbounded. Nevertheless, in this short note we show how to control the unbounded regions so that we achieve the desired result, namely the convergence of the integral of the logarithm near zero for Lebesgue almost every \( z \). It is not hard to realize that, using similar argument, same result as in Theorem 1.1 can be obtained, if we replace unitary matrices by Haar distributed orthogonal matrices.

In [4, Theorem 1] it is shown that under assumptions [4, Eqn. 1-3] the empirical distribution of the eigenvalues of \( U_nT_n \) converges, is a single ring, and is related to the Brown measure of the corresponding limiting operator. In a recent paper by Rudelson, and Vershynin [8] it is shown that the condition [4, Eqn. 3] can always be removed (see [8, Corollary 1.3]). It can also be noted that [4, Eqn. 2] is a very rigid condition, and it prevents the existence of atoms in the support of the limiting distribution. For example, it can not include the simplest case where \( T_n = I_n \) (see [4, Remark 2]). But as a corollary of our proof, we improve the uniform boundedness condition of the Stieltjes transform (compare (1.2) with [4, Eqn. 2]) to obtain the following which can accommodate atoms:

**Corollary 1.2.** Let empirical distribution of the eigenvalues of \( \{ T_n \} \) converges weakly to a probability measure \( \Theta \), compactly supported on \( \mathbb{R}_+ \). Assume further

1. There exists a constant \( M > 0 \) so that
   \[
   \lim_{n \to \infty} \mathbb{P}(\| T_n \| > M) = 0. 
   \]

2. There exists a set of isolated points \( \{ x_i \}_{i=1}^K \) such that, for every \( \varepsilon > 0 \), there exists \( 0 < \kappa, M < \infty \) for which the following hold:
   \[
   \{ z : \Im(z) > n^{-\kappa}, |\Im(G_{T_n}(z))| > M \} \subset \{ z : \Re(z) \in \bigcup_{i=1}^K (x_i - \varepsilon, x_i + \varepsilon) \}. 
   \]

If \( \Theta \) is not a Dirac measure, then the following hold.

(a) The empirical distribution of the eigenvalues of \( A_n := U_nT_n \) converges, in probability, to limiting probability measure \( \mu_A \).

(b) The measure \( \mu_A \) possesses a radially-symmetric density with respect to the Lebesgue measure on \( \mathbb{C} \), satisfying \( \rho_A(z) = \frac{1}{2\pi} \Delta_z(\int \log |x| d\nu^z(x)) \), where \( \Delta_z \) denotes the Laplacian with respect to the variable \( z \), \( \nu^z := \Theta \perp \lambda_{|z|} \), \( \lambda_r = \frac{1}{2}(\delta_r + \delta_{-r}) \), and \( \Theta \) is the symmetrized version of \( \Theta \).

\(^1\)For a matrix its Brown measure is its empirical distribution of the eigenvalues (see [9, Proposition 1])
The support of $\mu_A$ is single ring: There exists constants $0 \leq a < b < \infty$ so that
\[ \text{supp}\mu_A = \{re^{i\theta} : a \leq r \leq b\}. \]
Further, $a = 0$ if and only if $\int x^{-2}d\Theta(x) = \infty$.

This extension above can now include the cases where $\Theta$ can have atoms, or unbounded density, as long as, at the finite $n$-level (1.2) satisfied. For example, if $T_n$ is diagonal, with $\alpha_i$ proportion of them are $x_i$, for $i = 1, 2, \ldots, k$, then $\Theta$ is $x_i$ with probability $p_i$, and at the finite $n$-level (1.2) is satisfied. The case when $T_n = \alpha I_n$, for some $\alpha > 0$, is immediate from Theorem 1.1.

2. Proof of Theorem 1.1

Theorem 1.1 is proved via Girko’s method, and for completeness, we start with describing the key steps in Girko’s method (for more details see [4]). To this end, given a matrix $A_n$, let $L_{A_n}$ denote the empirical distribution of its eigenvalues. For every $z \in \mathbb{C}$, define
\[ H_n^z := \begin{bmatrix} 0 & (A_n - zI_n)^* \\ (A_n - zI_n)^* & 0 \end{bmatrix}, \]
and let $\nu_n^z$ be the empirical distribution of the eigenvalues of $H_n^z$. Girko’s method consists of:

Step 1: Show that for (Lebesgue almost) every $z \in \mathbb{C}$, the measures $\nu_n^z$ converge weakly, in probability, to measure $\nu^z$ as $n \to \infty$.

Step 2: Justify that $\int \log |x|\nu_n^z(dx) \to \int \log |x|\nu^z(dx)$.

In the non-hermitian regime, the main (technical) challenge is to verify Step 2. Beyond that completing the proof of Theorem 1.1 requires few relatively standard additional steps (outlined in sequel).

Turning to prove Step 1 and Step 2, for any $z \in \mathbb{C}$, and $\{U_i^d\}_{1 \leq i \leq d}$ i.i.d. unitary matrices, distributed according to Haar measure, define,
\[ V_n^{1,z} := U_n^{1,z} := \begin{bmatrix} 0 & (U_n - zI_n)^* \\ (U_n - zI_n)^* & 0 \end{bmatrix}, \]
and
\[ V_n^{d,z} := V_n^{d-1,z} + U_n^d + (U_n^d)^* := V_n^{d-1,z} + \begin{bmatrix} 0 & U_n^d \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ (U_n^d)^* & 0 \end{bmatrix}, \quad \text{for } d \geq 2. \]

For any two compactly supported measures $\mu_1$ and $\mu_2$ on $\mathbb{R}$, we denote $\mu_1 \boxplus \mu_2$ to be free convolution (see [1, §5.3.3] for more details) of $\mu_1$ and $\mu_2$.

Step 1 and Step 2 of Girko’s method are taken care of by the following two lemmas respectively (whose proofs are deferred to Section 3).

Lemma 2.1. Let $\Theta^{1,z}$ is the symmetrized version the measure having density given by (3.1) with $r = |z|$, and for $d \geq 2$, $\Theta^{d,z} := \Theta^{d-1,z} \boxplus \lambda_1$, where $\lambda_1 = \frac{1}{2}(\delta_{-1} + \delta_1)$. Then for $d \in \mathbb{N}$, the empirical measure $L_{V_n^{d,z}}$ converges weakly, in probability, to $\Theta^{d,z}$.

Lemma 2.2. For Lebesgue almost every $z \in \mathbb{C}$, for every $d \geq 2$,
\[ \int \log |x|dL_{V_n^{d,z}}(x) \to \int \log |x|d\Theta^{d,z}(x), \quad (2.1) \]
in probability. Furthermore, for every $d \in \mathbb{N}$ there exists $\{r_i\}_{i=1}^{K'}$ (depending on $d$), such that for every $\varepsilon > 0$ and $R > 0$, whenever $\phi$ is any smooth function supported on $\{z : |z| \leq R\} \cup \bigcup_{i=1}^{K'} \{z : |z|$
\[ |z| \in (r_i - \varepsilon, r_i + \varepsilon), \] we have
\[
\int \phi(z) \int \log |x| dL_{V_{n}}(x) dm(z) \to \int \phi(z) \int \log |x| d\Theta^{d,z}(x) dm(z), \tag{2.2}
\]
in probability.

**Proof of Theorem 1.1:** First let us assume that (2.2) holds almost surely. Combining Lemma 2.2 with [4, Eqn. 5], we get that for every \( R > 0 \) and \( \varepsilon > 0 \), and any smooth function \( \psi \), supported on
\[
\Gamma^{d}_{R,\varepsilon} := B(0, R) \setminus A^{d}_{\varepsilon} := \{ z : |z| \leq R \} \setminus \cup_{i=1}^{K'} \{ z : |z| \in (r_i - \varepsilon, r_i + \varepsilon) \},
\tag{2.3}
\]
we have
\[
\int \psi(z) dL_{U_{n}^{d}}(z) \to \frac{1}{2\pi} \int \Delta \psi(z) \left[ \int \log |x| d\Theta^{d,z}(x) \right] dm(z)
= \frac{1}{2\pi} \int \psi(z) \Delta \left[ \int \log |x| d\Theta^{d,z}(x) \right] dm(z) := \int \psi(z) h(z) dm(z), \tag{2.4}
\]
where the integration by parts in the last but one step follows from \( \psi \) being compactly supported and smooth. Now by [6, Proposition 3.5] we note that \( \Theta^{d,z} \) is the symmetrized version of the law of \( |u_1 + u_2 + \cdots + u_d - zI| \), and from the definition of Brown measure it follows that \( h(z) dm(z) \) is the Brown measure of \( u_1 + u_2 + \cdots + u_d \) (see [6]). We further note that for \( d \geq 2 \), it has a density (namely, \( h(z) \)) with respect to the Lebesgue measure on \( \mathbb{C} \). Therefore by a standard argument it follows that for any open set \( G \subset \mathbb{C} \), for any \( R > 0 \), and \( \varepsilon > 0 \),
\[
\liminf_{n \to \infty} L_{U_{n}^{d}}(G \cap (\Gamma^{d}_{R,\varepsilon})^{\circ}) \geq \int_{G \cap (\Gamma^{d}_{R,\varepsilon})^{\circ}} h(z) dm(z),
\]
a.s. and consequently
\[
\liminf_{n \to \infty} L_{U_{n}^{d}}(G) \geq \lim_{R \to \infty} \lim_{\varepsilon \to 0} \int_{G \cap (\Gamma^{d}_{R,\varepsilon})^{\circ}} h(z) dm(z)
\geq \lim_{R \to \infty} \int_{G \cap B(0,R)} h(z) dm(z) - \lim_{\varepsilon \to 0} \int_{(A^{d}_{\varepsilon})^{\circ}} h(z) dm(z) = \int_{G} h(z) dm(z), \tag{2.5}
\]
almost surely. Since the same statement remains true for a countable collection of open sets, the proof completes by a use of Portmanteau theorem, if (2.2) holds a.s..

To prove the above when (2.2) holds in probability, we use the fact that the convergence in probability implies that for every subsequence, there exists a further subsequence such that a.s. convergence holds. To this end, for any given subsequence, we can extract a further subsequence (by cantor diagonal argument) such that (2.5) holds for a countable collection of opens sets \( G \subset \mathbb{C} \), and therefore again by a use of Portmanteau theorem, the proof is completed for \( d \geq 2 \).

When \( d = 1 \), from the proof of Lemma 2.1 it is clear that the expected empirical measure of the eigenvalues of \( |U_n - zI_n| \) is same as \( \Theta^{1,z} \), and therefore the expected Brown measure of \( U_n \) is same as that of Haar unitary. Consequently the proof is completed by a use of a standard concentration inequality for Haar measures on unitary group (see [1, Corollary 4.4.30]). \( \square \)

3. Proofs of Lemma 2.1 and Lemma 2.2

Since \( U_{n}^{d} \)'s are invariant under rotation, it is easy to see that \( \Theta^{d,z} \) is a function of \( |z| \). Therefore the lemmas will be proved only for \( z = r > 0 \). Before we prove Lemma 2.1 we will need a preliminary result first:
Lemma 3.1. Let $U_n$ be Haar distributed. For any $r > 0$, let $\tilde{L}_{U_n,r}$ be the expected empirical distribution of the eigenvalues of $[U_n - rI_n]$. Then $\tilde{L}_{U_n,r}$ has a density

$$f_r(x) = \frac{2}{\pi} \frac{x}{\sqrt{(x^2 - (r - 1)^2)((r + 1)^2 - x^2)}}, \quad |r - 1| \leq x \leq r + 1. \quad (3.1)$$

with respect to the Lebesgue measure.

Proof: Note that to prove the above it is enough to show that the expected empirical distribution of the eigenvalues of $(U_n - rI_n)(U_n - rI_n)^*$ has the density

$$g_r(x) = \frac{1}{\pi} \frac{1}{\sqrt{(x - (r - 1)^2)((r + 1)^2 - x)}}, \quad (r - 1)^2 \leq x \leq (r + 1)^2. \quad (3.2)$$

To this end first note that using rotational invariance of the Haar measure, we have $E[\frac{1}{n}\text{Tr}(U_n)] = E[\frac{1}{n}\text{Tr}((U_n)^k)] = 0$ for any positive integer $k$. Thus

$$E[\frac{1}{n}\text{Tr}((U_n + U_n^*)^k)] = \left\{\begin{array}{ll} k/2 & \text{for } k \text{ even and } 0 \text{ otherwise.} \end{array}\right. \quad (3.3)$$

Therefore

$$\tilde{L}_{U_n + U_n^*} = 2 \cos \theta = e^{i\theta} + e^{-i\theta}, \quad \text{where } \theta \sim \text{Unif}(0, 2\pi). \quad (3.4)$$

Now noting that

$$(U_n - rI_n)(U_n - rI_n)^* = (1 + r^2)I_n - r(U_n + U_n^*),$$

the result follows from the change of variable formula. $\square$

Proof of Lemma 2.1: Let $\tilde{L}_{V_n^{d,r}}$ be the expected empirical distribution of the eigenvalues of $V_n^{d,r}$. The fact that $\tilde{L}_{V_n^{d,r}} \equiv \Theta^{1,r}$ follows from Lemma 3.1. We proceed to show that $\tilde{L}_{V_n^{d,r}} = \Theta^{d,r}$, by induction on $d \geq 2$, following the proof of [4, Lemma 10], setting there $\rho = 1$, and replacing $T_n$ by $V_n^{d-1,r}$, to conclude (relying on the Schwinger-Dyson equation (3.5)) that the corresponding Stieltjes transforms converge. The fact that $\tilde{L}_{V_n^{d,r}} \Rightarrow \Theta^{d,r}$, in probability, follows from the standard concentration inequality for Haar measures (see [1, Corollary 4.4.30]). $\square$

Lemma 2.1 completes the proof of Step 1 of Girko’s method. Turning to Step 2 for $d \geq 1$, and $z \in \mathbb{C}^+$, define

$$G_n^{d,r}(z) = E\left[\frac{1}{2n} \text{Tr}(zI_n - V_n^{d,r})^{-1}\right],$$

and

$$G_{U_n}^{d,r}(z) = E\left[\frac{1}{2n} \text{Tr}\left(U_n(zI_n - V_n^{d,r})^{-1}\right)\right].$$

Now imitating the steps leading from [4, Eqn. 30] to [4, Eqn. 38], we obtain the following finite $n$-level Schwinger-Dyson equation:

$$G_n^{d,r}(z) = G_n^{d-1,r}(\psi_n^{d}(z)) - \tilde{O}(n, z, \psi_n^{d}(z)), \quad (3.5)$$

(analogous to [4, Eqn. 38]) where

$$\psi_n^{d}(z) := z - \frac{G_n^{d,r}(z)}{1 + 2G_n^{d,r}(z)}. \quad (3.6)$$

(analogous to [4, Eqn. 37]), for every $z_1, z_2 \in \mathbb{C}^+$

$$\tilde{O}(n, z_1, z_2) = \frac{O(n, z_1, z_2)}{1 + 2G_n^{d,r}(z_1)},$$
and
\[
O(n, z_1, z_2) = E \left[ \left( \frac{1}{2n} \text{Tr} - E \left[ \frac{1}{2n} \text{Tr} \right] \right) \right.
\left. \otimes \left( \frac{1}{2n} \text{Tr} - E \left[ \frac{1}{2n} \text{Tr} \right] \right) \partial (z_1 I_n - V_n^{d,r})^{-1} (z_2 I_n - V_n^{d-1,r})^{-1} U_n^d \right]
\]
\[= O \left( \frac{1}{n^2 |\Im(z_2)| |\Im(z_1)|^2 (|\Im(z_1)| \land 1)} \right). \tag{3.7}\]

We note that the Schwinger-Dyson equation (3.5) is valid whenever the denominator in (3.6) does not vanish. In particular, it holds when \(\Im(z)\) is large. However our aim is to extend the Schwinger-Dyson equation to a region very near to the real line. To do so we need the following fine estimates on the terms appearing in (3.5). To this end similarly as in [4, Eqn. 34] we further obtain,
\[
(G_n^{d,r}(z))^2 = 2G_{U_n}^{d,r}(z)(1 + 2G_{U_n}^{d,r}(z)) - O_1(n, z), \tag{3.8}
\]
where
\[
O_1(n, z) = 4E \left[ \left( \frac{1}{2n} \text{Tr} - E \left[ \frac{1}{2n} \text{Tr} \right] \right) \right.
\left. \otimes \left( \frac{1}{2n} \text{Tr} - E \left[ \frac{1}{2n} \text{Tr} \right] \right) \partial (zI_n - V_n^{d,r})^{-1} U_n^d \right]
\]
\[= O \left( \frac{1}{n^2 |\Im(z)|^2 (|\Im(z)| \land 1)} \right). \tag{3.9}\]

Thus,
\[
G_{U_n}^{d,r}(z) = \frac{1}{4} \left( -1 + \sqrt{1 + 4G_{U_n}^{d,r}(z)^2 + 4O_1(n, z)} \right), \tag{3.10}\]
where the branch of the square root is uniquely determined by analyticity, and the behavior of \(G_{U_n}^{d,r}\) and \(G_n^{d,r}\) when \(|z| \to \infty\). Furthermore as in [4, Lemma 11, Lemma 12] we get:
For every positive integer \(d \geq 2\), and \(0 \leq r \leq R\), there exists constants \(C_1, C_2, C_3\) depending only on \(d\) and \(R\) such that for all \(z \in \mathbb{C}^+\), with \(\Im(z) > C_1 n^{-1/3}\), and all large \(n\),
\[
|1 + 2G_{U_n}^{d,r}(z)| > C_2 [\Im(z)^3 \land 1], \tag{3.11}
\]
and for all \(z \in \mathbb{C}^+\) such that \(\Im(z) > C_3 n^{-1/4}\),
\[
\Im(\psi_n^d(z)) \geq \Im(z)/2. \tag{3.12}
\]

Thus it shows that the finite \(n\)-level Schwinger-Dyson equation (3.5) extends to the region \(\{z : \Im(z) > C_3 n^{-1/4}\}\), for all large \(n\).

The proof of Lemma \ref{lem:finite} consists of controlling the Stieltjes transform of \(G_n^{d,r}(z)\) inductively in \(d\), where Lemma \ref{lem:inf} provides the basis \(d = 1\) of the induction, and Lemma \ref{lem:prev} the inductive step.

**Lemma 3.2.** For any \(\epsilon > 0\) and \(r > 0\) there exists some constant \(C\), such that
\[
\left\{ z : |\Im G_n^{d,r}(z)| \geq C \epsilon^{-2} \right\} \subset \left\{ z : E + i\eta \in \mathbb{C}^+ : \eta \in (0, \epsilon^2), E \in \left( \pm (1 \pm r) - 2\epsilon, \pm (1 \pm r) + 2\epsilon \right) \right\}.
\]
Proof: First note that for $\eta > 0$, 
\[
|\Im G_n^{1,r}(E + i\eta)| = \int_{|x-E|>\sqrt{\eta}} \frac{\eta}{(x-E)^2 + \eta^2} f_r(x) dx + \int_{|x-E|<\sqrt{\eta}} \frac{\eta}{(x-E)^2 + \eta^2} f_r(x) dx
\]
\[
\leq 1 + \sup_{x:|x-E|\leq \sqrt{\eta}} f_r(x) \int_{|x-E|\leq \sqrt{\eta}} \frac{\eta}{(x-E)^2 + \eta^2} dx
\]
\[
\leq 1 + \pi \sup_{x:|x-E|\leq \sqrt{\eta}} f_r(x).
\] (3.13)

Denoting $\Gamma_\varepsilon$ to be the union of open intervals of radius $\varepsilon$ around the four points $\pm 1 \pm r$; from (3.1) it follows that for any $\varepsilon > 0$, 
\[
\sup_{x \notin \Gamma_\varepsilon} \{ f_r(x) \} = C_1 \varepsilon^{-1},
\]
for some constant $C_1$, independent of choice of $r$. Thus from (3.13) it follows that
\[
\sup_{E,\eta: (E - \sqrt{\eta}, E + \sqrt{\eta}) \in \Gamma_\varepsilon} |\Im G_n^{1,r}(E + i\eta)| \leq C \varepsilon^{-1}.
\] (3.14)

Now noting that 
\[
\{(E, \eta): E \in \Gamma^c_2 \varepsilon, \eta \in (0, \varepsilon^2)\} \subset \{(E, \eta): (E - \sqrt{\eta}, E + \sqrt{\eta}) \in \Gamma^c_\varepsilon\},
\]
and
\[
\sup_{E,\eta: \eta \geq \varepsilon^2} |\Im G_n^{1,r}(E + i\eta)| \leq \varepsilon^{-2},
\]
the proof follows from (3.14). $\square$

In Lemma 3.2 we note that $\Im G_n^{1,r}(z)$ is not uniformly bounded, and therefore results similar to [4, Lemma 13] is not possible in our set-up. Instead in Lemma 3.3 we show that if we have some knowledge about the regions where $\Im(G_n^{d,r}(z))$ blows up, using (3.5) we can identify the regions where $\Im(G_n^{d+1,r}(z))$ blows up, which is good enough to establish Lemma 2.2, and consequently Theorem 1.1.

**Lemma 3.3.** Fix any $d \in \mathbb{N}$, and some $r \in (0, R]$. Given any $\varepsilon > 0$ assume that there exists constants $\kappa, M,$ and $K$, depending possibly only on $d$ and $R$ such that
\[
\Gamma_d := \{ z : \Im(z) > n^{-\kappa}, |\Im(G_n^{d,r}(z))| > M \} \subset \{ z : \Re(z) \in \Gamma^{d,r}_\varepsilon \},
\] (3.15)
where $\Gamma^{d,r}_\varepsilon$ is the union of $\varepsilon$ intervals around $\pm m \pm r$, for $m = 1, 2, \ldots, K$.

Then there exists constants $\kappa_1, M_1,$ and $K_1$, depending possibly only on $d$ and $R$ such that
\[
\Gamma_{d+1} := \{ z : \Im(z) > n^{-\kappa_1}, |\Im(G_n^{d+1,r}(z))| > M_1 \} \subset \{ z : \Re(z) \in \Gamma^{d+1,r}_\varepsilon \},
\] (3.16)
where $\Gamma^{d+1,r}_\varepsilon$ is the union of $\varepsilon$ intervals around $\pm m \pm r$, for $m = 1, 2, \ldots, K_1$.

**Proof:** Enlarging $M$ in (3.15) as needed, we assume without loss of generality $6M^{-1} \leq \varepsilon^2$. Fixing any $0 < \kappa_1 < \kappa$, and $M_1 = 2M$ let $z \in \Gamma_{d+1}$. Now note that by (3.7) and (3.11), we get
\[
|\tilde{O}(n, z, \nu_n^{d+1}(z))| \leq \frac{C}{n^2 |\Im(\nu_n^{d+1}(z))| \Im(z)^2 (\Im(z)^4 \wedge 1) \leq M,
\]
whenever $\Im(z) \geq C'n^{-\kappa'}$ for some $C' < \infty, \kappa' > 0$, and using (3.12) in the right most inequality. Thus from (3.5) and (3.12), we get $|\Im(G_n^{d,r}(\nu_n^{d+1}(z)))| \geq M$, and $\Im(\nu_n^{d+1}(z)) > n^{-\kappa}$, and
consequently $\psi_n^{d+1}(z) \in \Gamma_d$. Now for any $z \in \mathbb{C}^+$ define

$$R(G_n^{d+1,r}(z)) := \frac{G_n^{d+1,r}(z)}{1 + 2G_n^{d+1,r}(z)}.$$  \hfill (3.17)

Using (3.8) we note that

$$G_n^{d+1,r}(z)\left[1 - R^2(G_n^{d+1,r}(z))\right] = R(G_n^{d+1,r}(z)) + \frac{G_n^{d+1,r}(z)O_1(n,z)}{(1 + 2G_n^{d+1,r}(z))^2}.$$  \hfill (3.18)

Now using (3.10) we further obtain that,

$$R(G_n^{d+1,r}(z)) = \frac{2G_n^{d+1,r}(z)}{1 + \sqrt{1 + 4(G_n^{d+1,r}(z))^2 + 4O_1(n,z)}} = \frac{1}{2} \frac{G_n^{d+1,r}(z)}{(G_n^{d+1,r}(z))^2 + O_1(n,z)} - \frac{1}{2} \frac{G_n^{d+1,r}(z)}{(G_n^{d+1,r}(z))^2 + O_1(n,z)}.$$ \hfill (3.19)

Noting that, from (3.9), $O_1(n,z) = O(n^{-1})$ for $\Im(z) \geq n^{-1/3}$ the rightmost term in the RHS of (3.19) is bounded by $M^{-1}$ whenever $|\Im(G_n^{d+1,r}(z))| \geq M$. Also note that if both $\Im(z) \geq n^{-1/3}$, and $|\Im(G_n^{d+1,r}(z))| > M$, then for any choice of the branch of the square root,

$$|\frac{1}{2} \frac{G_n^{d+1,r}(z)}{(G_n^{d+1,r}(z))^2 + O_1(n,z)}| \leq 2 \sqrt{\frac{|1 + 4(G_n^{d+1,r}(z))^2 + 4O_1(n,z)|}{4(G_n^{d+1,r}(z))^2}} \leq 4.$$  \hfill (3.19)

Therefore

$$|R(G_n^{d+1,r}(z))| \leq 5,$$

whenever $\Im(z) \geq n^{-1/3}$, and $|\Im(G_n^{d+1,r}(z))| > M$. Furthermore, from (3.9) and (3.11), there exists constant $\kappa_2 > 0$ such that

$$\Im(z) > n^{-\kappa_2} \Rightarrow \frac{G_n^{d+1,r}(z)O_1(n,z)}{(1 + 2G_n^{d+1,r}(z))^2} = O(n^{-1}).$$

Therefore, from (3.18) we get that for $\Im(z) > n^{-(\kappa_2 \wedge 1/3)}$, and $|\Im(G_n^{d+1,r}(z))| > M$,

$$|R^2(G_n^{d+1,r}(z)) - 1| \leq 6|G_n^{d+1,r}(z)|^{-1} \leq 6M^{-1} \leq \varepsilon^2.$$  \hfill (3.20)

This further shows that $R(G_n^{d+1,r}(z)) \in B(1,\varepsilon) \cup B(-1,\varepsilon)$ \hfill (3.21)

Now collecting all the arguments above we see that if we choose $\kappa_1$ small enough (in particular any $\kappa_1 < \kappa \wedge \kappa_2 \wedge 1/3$ will do) , and $M_1 = 2M$ then for any $z \in \Gamma_{d+1}$, we have $\psi_n^{d+1}(z) \in \Gamma_d$, and from (3.6) it follows that

$$R(z) = R(\psi_n^{d+1}(z)) + R(R(G_n^{d+1,r}(z))) \in \Gamma_{2\varepsilon}^{d+1,r}.$$  \hfill (3.22)

Thus the proof is completed. \hfill \Box

Now we are ready to prove Lemma 2.2.

Proof of Lemma 2.2: First note that by Lemma 3.2, Lemma 3.3, and using induction, for every $\varepsilon > 0$, and $R > 0$ fixed, we get that for $d \in \mathbb{N}$, and $r \in (0, R]$, there exists constants $\kappa, M,$ and $K$,
depending only on $d$ and $R$, such that for every
\[ z \in \{ z : \Im(z) > n^{-\kappa_1}, \Re(z) \notin \Gamma_{R,\varepsilon}^d \}, \]
\[ |\Im(G_{n}^{d,r}(z))| \leq M, \]
where $\Gamma_{R,\varepsilon}^d$ is the union of $\varepsilon$ intervals around $\pm m \pm r$, $m = 0, 1, 2, \ldots, K$. Now we note that $0 \in \Gamma_{R,\varepsilon}^d \Leftrightarrow r \in \cup_{i=1}^{K'} (r_i - \varepsilon, r_i + \varepsilon)$ for some sequence $\{r_i\}_{i=1}^{K'}$. Therefore using [4, Lemma 15], and imitating the steps leading to [4, Eqn. 49], for any $\alpha \in [1, 2]$, and for any $\varepsilon' > 0$, we obtain
\[ E \left[ \int_{n-\kappa}^{n-\kappa} |\log x|^{\alpha} dL_{\nu_{n}^{d,s}}(x) \right] \leq C \varepsilon'|\log(\varepsilon')|^\alpha, \tag{3.20} \]
whenever $z \in \Gamma_{R,\varepsilon}^d$ (defined in (2.3)), and the constant $C$ depends only on $R$. Now from [8, Theorem 1.1] we know that there exists constants $0 < c_1, c_2 < \infty$ such that for every $z \in \mathbb{C}$,
\[ \mathbb{P}(s_{\min}(\nu_{n}^{d,z}) \leq t) \leq t^{c_1} n^{c_2}, \]
and thus upon choosing a $\delta > c_2/c_1$, there exists a $\delta'$ such that for all large $n$,
\[ E \left[ \int_{0}^{n-\delta} |\log x|^{\alpha} dL_{\nu_{n}^{d,s}}(x) \right] < \delta'. \tag{3.21} \]
We further note that without loss of generality we can assume $\kappa < \delta$ and, hence imitating the steps in the proof of [4, Proposition 14(i)] we obtain (2.1).

Now to prove (2.2) we first note that the bounds in (3.21), (3.20) are uniform for $z \in \Gamma_{R,\varepsilon}^d$, and then the proof is completed by adapting the steps in [4, Proposition 14(ii)].

Proof of Corollary 1.2: The main step of the proof is to show that for Lebesgue almost every $z$, logarithm is uniformly integrable with respect to the empirical distribution of the eigenvalues of $|U_{n}T_{n} - zI_{n}|$. As shown in [4], this problem is equivalent to that of $Y_{n}^r := T_{n} + \rho(U_{n} + U_{n}^*)$, where $\rho = |z|$. We note that for this latter problem, one key-step is to show that the imaginary part of the Stieltjes transform of the expected empirical distribution of the eigenvalues of $T_{n} + \rho(U_{n} + U_{n}^*)$ is uniformly bounded on the imaginary axis, provided the imaginary part is bigger than $n^{-\kappa}$, for some $\kappa > 0$, for Lebesgue almost every $\rho$ (see [4, Proposition 14(i)]). In [4] this was shown under the assumption [4, Eqn. 2]. Here we show that same can be achieved under the modified assumption (1.2).

To this end note that the Schwinger-Dyson equation obtained here, and that of [4], are similar, it is not hard to realize that, adapting the proof of Lemma 3.3, from (1.2), we obtain that for every $\varepsilon > 0$, and $R > 0$ there exists some constants $M_1$, and $\kappa_1$ depending only on $R$ and $\varepsilon$ such that, for every $\rho \in (0, R]$,
\[ \{ z : \Im(z) > n^{-\kappa_1}, |\Im(G_{T_{n} + \rho(U_{n} + U_{n}^*)}(z))| > M_1 \} \subset \{ z : \Re(z) \in \cup_{i=1}^{K}(x_i \pm \rho - 2\varepsilon, x_i \pm \rho + 2\varepsilon) \}. \]

Now following the proof of Lemma 2.2, and [4, Corollary 1.3] we obtain that for Lebesgue almost every $z$
\[ \int \log |x| d\nu_{n}^{r}(x) \to \int \log |x| d\nu^{r}(x), \]
in probability, where $\nu_{n}^{r}$ is the empirical measure of the eigenvalues of $Y_{n}^{r}$. Moreover the same proof shows that, for any small $\varepsilon > 0$ (any $\varepsilon < \min\{x_i : x_i \neq 0\}$ will do), any $R > 0$, and any smooth function $\phi$, supported on $B(0, R) \setminus \cup_{i=1}^{K} \{ z : |z| \in (x_i - \varepsilon, x_i + \varepsilon) \}$, we have
\[ \int \phi(z) \int \log |x| d\nu_{n}^{r}(x) dm(z) \to \int \phi(z) \int \log |x| d\nu^{r}(x) dm(z). \]
Since $\Theta$ not a Dirac measure, from [6, Theorem 4.4], and [4, Remark 8] we note that $\mu_A$ has a density with respect to the Lebesgue measure on $\mathbb{C}$. Hence, combining the proof of Theorem 1.1, and [4, Theorem 1], we obtain the desired result.

Acknowledgements. We thank Ofer Zeitouni for suggesting to look at the sum of $d$ unitary matrices, and pointing our attention to preliminary versions of [2, 8].

References

[1] G. W. Anderson, A. Guionnet, and O. Zeitouni. An Introduction to Random Matrices, Cambridge Stud. Adv. Math. 118, Cambridge Univ. Press, Cambridge, 2010.
[2] C. Bordenave, and D. Chafaï. Around the circular law. Probability surveys, 9, 1-89, 2012.
[3] V. L. Girko. The circular law. Teor. Veroyatnost. i Primenen, 29, 669-679, 1984.
[4] A. Guionnet, M. Krishnapur, and O. Zeitouni. The single ring theorem. Ann. Math., 174, 1189-1217, 2011.
[5] A. Guionnet, P. Wood, and O. Zeitouni. Convergence of spectral measure of non normal matrices. Proc. AMS, to appear, 2012.
[6] U. Haagerup, and F. Larsen. Brown’s spectral distribution measure for R-diagonal elements in finite von Neumann algebras. J. Funct. Anal., 176, 331-367, 2000.
[7] C. Male. The distribution of traffic and their free product: an asymptotic freeness theorem for random matrices and a central limit theorem. Preprint, arXiv:1111.4662v3.
[8] M. Rudelson, and R. Vershynin. Invertibility of random matrices: unitary and orthogonal transformation. Preprint, arXiv:1206.5180v1.
[9] P. Śniady. Random regularization of Brown spectral measure. J. Funct. Anal., 193, 291-313, 2002.
[10] T. Tao, and V. Vu. Random matrices: universality of the ESDs and the circular law. Ann. Probab., 38, 2023-2065, 2010. With an appendix by Manjunath Krishnapur.

†Department of Mathematics, Stanford University
Building 380, Sloan Hall, Stanford, California 94305

∗†Department of Statistics, Stanford University
Sequoia Hall, 390 Serra Mall, Stanford, California 94305