Electron. Commun. Probab. **18** (2013), no. 69, 1–19.
DOI: 10.1214/ECP.v18-2466
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Limiting spectral distribution of sums of unitary and orthogonal matrices*

Anirban Basak† Amir Dembo‡

Abstract

We show that the empirical eigenvalue measure for sum of d independent Haar distributed n-dimensional unitary matrices, converge for $n \to \infty$ to the Brown measure of the free sum of d Haar unitary operators. The same applies for independent Haar distributed n-dimensional orthogonal matrices. As a byproduct of our approach, we relax the requirement of uniformly bounded imaginary part of Stieltjes transform of T_n that is made in [7, Thm. 1].

Keywords: Random matrices, limiting spectral distribution, Haar measure, Brown measure, free convolution, Stieltjes transform, Schwinger-Dyson equation..
AMS MSC 2010: 46L53; 60B10; 60B20.
Submitted to ECP on November 26, 2012, final version accepted on April 29, 2013.

1 Introduction

The method of moments and the Stieltjes transform approach provide rather precise information on asymptotics of the Empirical Spectral Distribution (in short ESD), for many *Hermitian* random matrix models. In contrast, both methods fail for *non-Hermitian* matrix models, and the only available general scheme for finding the limiting spectral distribution in such cases is the one proposed by Girko (in [6]). It is extremely challenging to rigorously justify this scheme, even for the matrix model consisting of i.i.d. entries (of zero mean and finite variance). Indeed, after rather long series of partial results (see historical references in [3]), the *circular law* conjecture, for the i.i.d. case, was only recently established by Tao and Vu [17] in full generality. Barring this simple model, very few results are known in the non-Hermitian regime. For example, nothing is known about the spectral measure of *random oriented d-regular* graphs. In this context, it was recently conjectured in [3] that, for $d \geq 3$, the ESD for the adjacency matrix of a uniformly chosen random oriented d-regular graph converges to a measure μ_d on the complex plane, whose density with respect to Lebesgue measure $m(\cdot)$ on $\mathbb{C}$ is

$$h_d(v) := \frac{1}{\pi} \frac{d^2(d-1)}{(d^2 - |v|^2)^2} \mathbb{I}_{\{|v| \leq \sqrt{d}\}} . \tag{1.1}$$

This conjecture, due to the observation that μ_d is the *Brown measure* of the *free sum* of $d \geq 2$ *Haar unitary operators* (see [9, Example 5.5]), motivated us to consider the related

*Support: Melvin & Joan Lane endowed Stanford Graduate Fellowship Fund, and NSF grant DMS-1106627.
†Department of Statistics, Stanford University, USA.
E-mail: anirbanb@stanford.edu
‡Department of Mathematics, and Department of Statistics, Stanford University, USA.
E-mail: amir@math.stanford.edu

problem of sum of d independent Haar distributed, unitary or orthogonal matrices, for which we prove such convergence of the ESD in Theorem 1.2. To this end, using hereafter the notation $\langle \mathrm{Log}, \mu \rangle_a^b := \int_a^b \log|x| d\mu(x)$ for any $a < b$ and probability measure μ on $\mathbb{R}$ (for which such integral is well defined), with $\langle \mathrm{Log}, \mu \rangle := \int_{\mathbb{R}} \log|x| d\mu(x)$, we first recall the definition of Brown measure for a bounded operator (see [9, Page 333], or [2, 4]).

Definition 1.1. *Let $(\mathcal{A}, \tau)$ be a non-commutative W^*-probability space, i.e. a von Neumann algebra $\mathcal{A}$ with a normal faithful tracial state τ (see [1, Defn. 5.2.26]). For h a positive element in $\mathcal{A}$, let μ_h denote the unique probability measure on $\mathbb{R}^+$ such that $\tau(h^n) = \int t^n d\mu_h(t)$ for all $n \in \mathbb{Z}^+$. The Brown measure μ_a associated with each bounded $a \in \mathcal{A}$, is the Riesz measure corresponding to the $[-\infty, \infty)$-valued sub-harmonic function $v \mapsto \langle \mathrm{Log}, \mu_{|a-v|} \rangle$ on $\mathbb{C}$. That is, μ_a is the unique Borel probability measure on $\mathbb{C}$ such that*

$$d\mu_a(v) = \frac{1}{2\pi} \Delta_v \langle \mathrm{Log}, \mu_{|a-v|} \rangle \, dm(v), \tag{1.2}$$

where Δ_v denotes the two-dimensional Laplacian operator (with respect to $v \in \mathbb{C}$), and the identity (1.2) holds in distribution sense (i.e. when integrated against any test function $\psi \in C_c^\infty(\mathbb{C})$).

Theorem 1.2. *For any $d \geq 1$, and $0 \leq d' \leq d$, as $n \to \infty$ the* ESD *for sum of d' independent, Haar distributed, n-dimensional unitary matrices $\{U_n^i\}$, and $(d - d')$ independent, Haar distributed, n-dimensional orthogonal matrices $\{O_n^i\}$, converges weakly, in probability, to the Brown measure μ_d of the free sum of d Haar unitary operators (whose density is given in (1.1)).*

Recall that as $n \to \infty$, independent Haar distributed n-dimensional unitary (or orthogonal) matrices converge in $\star$-moments (see [16] for a definition), to the collection $\{u_i\}_{i=1}^d$ of $\star$-*free* Haar unitary operators (see [1, Thm. 5.4.10]). However, convergence of $\star$-moments, or even the stronger convergence in distribution of traffics (of [11]), do not necessarily imply convergence of the corresponding Brown measures[1] (see [16, §2.6]). While [16, Thm. 6] shows that if the original matrices are perturbed by adding small Gaussian (of *unknown variance*), then the Brown measures do converge, removing the Gaussian, or merely identifying the variance needed, are often hard tasks. For example, [8, Prop. 7 and Cor. 8] provide an example of ensemble where no Gaussian matrix of polynomially vanishing variance can regularize the Brown measures (in this sense). Theorem 1.2 shows that sums of independent Haar distributed unitary/orthogonal matrices are *smooth* enough to have the convergence of ESD-s to the corresponding Brown measures *without adding any* Gaussian.

Guionnet, Krishnapur and Zeitouni show in [7] that the limiting ESD of $U_n T_n$ for non-negative definite, diagonal T_n of limiting spectral measure Θ, that is independent of the Haar distributed unitary (or orthogonal) matrix U_n, exists, is supported on a *single ring* and given by the Brown measure of the corresponding bounded (see [7, Eqn. (1)]), limiting operator. Their results, as well as our work, follow Girko's method, which we now describe, in brief.

From Green's formula, for any polynomial $P(v) = \prod_{i=1}^n (v - \lambda_i)$ and test function $\psi \in C_c^2(\mathbb{C})$, we have that

$$\sum_{j=1}^n \psi(\lambda_j) = \frac{1}{2\pi} \int_{\mathbb{C}} \Delta\psi(v) \log|P(v)| dm(v)\,.$$

[1] The Brown measure of a matrix is its ESD (see [16, Propn. 1])

Considering this identity for the characteristic polynomial $P(\cdot)$ of a matrix S_n (whose ESD we denote hereafter by L_{S_n}), results with

$$\int_{\mathbb{C}} \psi(v) dL_{S_n}(v) = \frac{1}{2\pi n} \int_{\mathbb{C}} \Delta\psi(v) \log |\det(vI_n - S_n)| dm(v)$$
$$= \frac{1}{4\pi n} \int_{\mathbb{C}} \Delta\psi(v) \log \det[(vI_n - S_n)(vI_n - S_n)^*] dm(v).$$

Next, associate with any n-dimensional non-Hermitian matrix S_n and every $v \in \mathbb{C}$ the $2n$-dimensional *Hermitian* matrix

$$H_n^v := \begin{bmatrix} 0 & (S_n - vI_n) \\ (S_n - vI_n)^* & 0 \end{bmatrix} . \tag{1.3}$$

It can be easily checked that the eigenvalues of H_n^v are merely ± 1 times the singular values of $vI_n - S_n$. Therefore, with ν_n^v denoting the ESD of H_n^v, we have that

$$\frac{1}{n} \log \det[(vI_n - S_n)(vI_n - S_n)^*] = \frac{1}{n} \log |\det H_n^v| = 2\langle \mathrm{Log}, \nu_n^v \rangle ,$$

out of which we deduce the key identity

$$\int_{\mathbb{C}} \psi(v) dL_{S_n}(v) = \frac{1}{2\pi} \int_{\mathbb{C}} \Delta\psi(v) \langle \mathrm{Log}, \nu_n^v \rangle dm(v) \tag{1.4}$$

(commonly known as Girko's formula). The utility of Eqn. (1.4) lies in the following general recipe for proving convergence of L_{S_n} per given family of non-Hermitian random matrices $\{S_n\}$ (to which we referred already as Girko's method).

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

Step 2: Justify that $\langle \mathrm{Log}, \nu_n^v \rangle \to \langle \mathrm{Log}, \nu^v \rangle$ in probability (which is the main technical challenge of this approach).

Step 3: A uniform integrability argument allows one to convert the v-a.e. convergence of $\langle \mathrm{Log}, \nu_n^v \rangle$ to the corresponding convergence for a suitable collection $\mathcal{S} \subseteq C_c^2(\mathbb{C})$ of (smooth) test functions. Consequently, it then follows from (1.4) that for each fixed, non-random $\psi \in \mathcal{S}$,

$$\int_{\mathbb{C}} \psi(v) dL_{S_n}(v) \to \frac{1}{2\pi} \int_{\mathbb{C}} \Delta\psi(v) \langle \mathrm{Log}, \nu^v \rangle dm(v) , \tag{1.5}$$

in probability.

Step 4: Upon checking that $f(v) := \langle \mathrm{Log}, \nu^v \rangle$ is smooth enough to justify the integration by parts, one has that for each fixed, non-random $\psi \in \mathcal{S}$,

$$\int_{\mathbb{C}} \psi(v) dL_{S_n}(v) \to \frac{1}{2\pi} \int_{\mathbb{C}} \psi(v) \Delta f(v) dm(v) , \tag{1.6}$$

in probability. For $\mathcal{S}$ large enough, this implies the convergence in probability of the ESD-s L_{S_n} to a limit which has the density $\frac{1}{2\pi}\Delta f$ with respect to Lebesgue measure on $\mathbb{C}$.

Employing this method in [7] requires, for **Step 2**, to establish suitable asymptotics for singular values of $T_n + \rho U_n$. Indeed, the key to the proofs there is to show that uniform boundedness of the imaginary part of Stieltjes transform of $\boldsymbol{T}_n$ (of the form assumed in [7, Eqn. (3)]), is inherited by the corresponding transform of $\boldsymbol{T}_n + \rho \boldsymbol{U}_n$ (see (1.12) for a definition of $\boldsymbol{U}_n$ and $\boldsymbol{T}_n$). In the context of Theorem 1.2 (for $d' \geq 1$), at the start $d = 1$, the expected ESD for $|vI_n - U_n|$ has unbounded density (see Lem. 4.1), so the imaginary parts of relevant Stieltjes transforms are *unbounded*. We circumvent this problem by localizing the techniques of [7], whereby we can follow the development of unbounded regions of the resolvent via the map $\boldsymbol{T}_n \mapsto \boldsymbol{T}_n + \rho(\boldsymbol{U}_n + \boldsymbol{U}_n^*)$ (see Lem. 1.5), so as to achieve the desired convergence of integral of the logarithm near zero, for Lebesgue almost every z. We note in passing that Rudelson and Vershynin showed in [15] that the condition of [7, Eqn. (2)] about minimal singular value can be dispensed off (see [15, Cor. 1.4]), but the remaining uniform boundedness condition [7, Eqn. (3)] is quite rigid. For example, it excludes atoms in the limiting measure Θ (so does not allow even $T_n = I_n$, see [7, Rmk 2]). As a by product of our work, we relax below this condition about Stieltjes transform of T_n (compare (1.8) with [7, Eqn. (3)]), thereby generalizing [7, Thm. 1].

Proposition 1.3. *Suppose the* ESD *of* $\mathbb{R}^+$*-valued, diagonal matrices* $\{T_n\}$ *converge weakly, in probability, to some probability measure* Θ *such that* $\Theta(\{0\}) = 0$*. Assume further that:*

1. *There exists finite constant* M *so that*

$$\lim_{n\to\infty} \mathbb{P}(\|T_n\| > M) = 0. \tag{1.7}$$

2. *There exists a closed set* $K \subseteq \mathbb{R}$ *of zero Lebesgue measure such that for every* $\varepsilon > 0$*, some* $\kappa_\varepsilon > 0$*,* M_ε *finite and all* n *large enough,*

$$\{z : \Im(z) > n^{-\kappa_\varepsilon}, |\Im(G_{\boldsymbol{T}_n}(z))| > M_\varepsilon\} \subset \{z : z \in \bigcup_{x\in K} B(x,\varepsilon)\}, \tag{1.8}$$

where $G_{\boldsymbol{T}_n}(z)$ *is the Stieltjes transform of the symmetrized version of the* ESD *of* T_n*, as defined in (1.13).*

If Θ *is not a (single) Dirac measure, then the following hold:*

(a) The ESD *of* $A_n := U_n T_n$ *converges, in probability, to limiting probability measure* μ_A*.*

(b) The measure μ_A *possesses a radially-symmetric density* $h_A(v) := \frac{1}{2\pi}\Delta_v\langle \mathrm{Log}, \nu^v\rangle$ *with respect to Lebesgue measure on* $\mathbb{C}$*, where* $\nu^v := \tilde{\Theta} \boxplus \lambda_{|v|}$ *is the free convolution (c.f. [1, §5.3.3]), of* $\lambda_r = \frac{1}{2}(\delta_r + \delta_{-r})$ *and the symmetrized version* $\tilde{\Theta}$ *of* Θ*.*

(c) The support of μ_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* $\int x^{-2} d\Theta(x) = \infty$*.*

(d) The same applies if U_n *is replaced by a Haar distributed orthogonal matrix* O_n*.*

This extension accommodates Θ with atoms, unbounded density, or singular part, as long as (1.8) holds (at the finite n-level). For example, Proposition 1.3 applies for T_n diagonal having $[np_i]$ entries equal $x_i \neq 0$, for $p_i > 0$, $i = 1, 2, \ldots, k \geq 2$, whereas the case of $T_n = \alpha I_n$ for some $\alpha > 0$ is an immediate consequence of Theorem 1.2.
Our presentation of the proof of Theorem 1.2 starts with detailed argument for $d' = d$, namely, the sum of independent Haar distributed unitary matrices. That is, we first prove the following proposition, deferring to Section 5 its extension to all $0 \leq d' < d$.

Proposition 1.4. *For any $d \geq 1$, as $n \to \infty$ the* ESD *of sum of d independent, Haar distributed, n-dimensional unitary matrices $\{U_n^i\}_{i=1}^d$, converges weakly, in probability, to the Brown measure μ_d of free sum of d Haar unitary operators.*

To this end, for any $v \in \mathbb{C}$ and i.i.d. Haar distributed unitary matrices $\{U_n^i\}_{1\leq i\leq d}$, and orthogonal matrices $\{O_n^i\}_{1\leq i\leq d}$, let

$$\boldsymbol{U}_n^{1,v} := \begin{bmatrix} 0 & (U_n^1 - vI_n) \\ (U_n^1 - vI_n)^* & 0 \end{bmatrix}, \tag{1.9}$$

and define $\boldsymbol{O}_n^{1,v}$ analogously, with O_n^1 replacing U_n^1. Set $\boldsymbol{V}_n^{1,v} := \boldsymbol{U}_n^{1,v}$ if $d' \geq 1$ and $\boldsymbol{V}_n^{1,v} := \boldsymbol{O}_n^{1,v}$ if $d' = 0$, then let

$$\boldsymbol{V}_n^{k,v} := \boldsymbol{V}_n^{k-1,v} + \boldsymbol{U}_n^k + (\boldsymbol{U}_n^k)^* := \boldsymbol{V}_n^{k-1,v} + \begin{bmatrix} 0 & U_n^k \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ (U_n^k)^* & 0 \end{bmatrix}, \text{ for } k = 2, \ldots, d', \tag{1.10}$$

and replacing $\boldsymbol{U}_n^k$ by $\boldsymbol{O}_n^k$, continue similarly for $k = d'+1, \ldots, d$. Next, let $G_n^{d,v}$ denote the expected Stieltjes transform of $\boldsymbol{V}_n^{d,v}$. That is,

$$G_n^{d,v}(z) := \mathbb{E}\Big[\frac{1}{2n} \operatorname{Tr}(zI_{2n} - \boldsymbol{V}_n^{d,v})^{-1}\Big], \tag{1.11}$$

where the expectation is over *all* relevant unitary/orthogonal matrices $\{U_n^i, O_n^i, i = 1, \ldots, d\}$. Part (ii) of the next lemma, about the relation between unbounded regions of $G_n^{d,v}(\cdot)$, and $G_n^{d-1,v}(\cdot)$ summarizes the key observation leading to Theorem 1.2 (with part (i) of this lemma similarly leading to our improvement over [7]). To this end, for any $\rho > 0$ and *arbitrary* n-dimensional matrix T_n (possibly random), which is independent of the unitary Haar distributed U_n, let

$$\boldsymbol{Y}_n := \boldsymbol{T}_n + \rho(\boldsymbol{U}_n + \boldsymbol{U}_n^*) := \begin{bmatrix} 0 & T_n \\ T_n^* & 0 \end{bmatrix} + \rho \begin{bmatrix} 0 & U_n \\ 0 & 0 \end{bmatrix} + \rho \begin{bmatrix} 0 & 0 \\ U_n^* & 0 \end{bmatrix} \tag{1.12}$$

and consider the following two functions of $z \in \mathbb{C}^+$,

$$G_{\boldsymbol{T}_n}(z) := \frac{1}{2n} \operatorname{Tr}(zI_{2n} - \boldsymbol{T}_n)^{-1}, \tag{1.13}$$

$$G_n(z) := \mathbb{E}\Big[\frac{1}{2n} \operatorname{Tr}(zI_{2n} - \boldsymbol{Y}_n)^{-1} \,|\, \boldsymbol{T}_n\Big]. \tag{1.14}$$

Lemma 1.5. *(i) Fixing R finite, suppose that $\|T_n\| \leq R$ and the* ESD *of $\boldsymbol{T}_n$ converges to some $\tilde{\Theta}$. Then, there exist $0 < \kappa_1 < \kappa$ small enough, and finite $M_\varepsilon \uparrow \infty$ as $\varepsilon \downarrow 0$, depending only on R and $\tilde{\Theta}$, such that for all n large enough and $\rho \in [R^{-1}, R]$,*

$$\Im(z) > n^{-\kappa_1} \;\&\; |\Im(G_n(z))| > 2M_\varepsilon \implies \exists \psi_n(z) \in \mathbb{C}^+,\ \Im(\psi_n(z)) > n^{-\kappa} \;\&\; |\Im(G_{\boldsymbol{T}_n}(\psi_n(z))| > M_\varepsilon \\ \&\; z - \psi_n(z) \in B(-\rho, \varepsilon) \cup B(\rho, \varepsilon). \tag{1.15}$$

The same applies when U_n is replaced by Haar orthogonal matrix O_n (possibly with different values of $0 < \kappa_1 < \kappa$ and $M_\varepsilon \uparrow \infty$).
(ii) For any R finite, $d \geq 2$ and $d' \geq 0$, there exist $0 < \kappa_1 < \kappa$ small enough and finite $M_\varepsilon \uparrow \infty$, such that (1.15) continues to hold for $\rho = 1$, all n large enough, any $|v| \leq R$ and some $\psi_n(\cdot) := \psi_n^{d,v}(\cdot) \in \mathbb{C}^+$, even when G_n and $G_{\boldsymbol{T}_n}$, are replaced by $G_n^{d,v}$ and $G_n^{d-1,v}$, respectively.

Section 2 is devoted to the proof of Lemma 1.5, building on which we prove Proposition 1.4 in Section 3. The other key ingredients of this proof, namely Lemmas 3.1 and 3.2, are established in Section 4. Finally, short outlines of the proofs of Theorem 1.2 and of Proposition 1.3, are provided in Sections 5 and 6, respectively.

2 Proof of Lemma 1.5

This proof uses quite a few elements from the proofs in [7]. Specifically, focusing on the case of unitary matrices, once a particular choice of $\rho \in [R^{-1}, R]$ and $\boldsymbol{T}_n$ is made in part (i), all the steps appearing in [7, pp. 1202-1203] carry through, so all the equations obtained there continue to hold here (with a slight modification of bounds on error terms in the setting of part (ii), as explained in the sequel). Since this part follows [7], we omit the details. It is further easy to check that the same applies for the estimates obtained in [7, Lem. 11, Lem. 12], which are thus also used in our proof (without detailed re-derivation).

Proof of (i): We fix throughout this proof a fixed realization of the matrix $\boldsymbol{T}_n$, so expectations are taken only over the randomness in the unitary matrix U_n. Having done so, first note that from [7, Eqn. (37)-(38)] we get

$$G_n(z) = G_{\boldsymbol{T}_n}(\psi_n(z)) - \widetilde{O}(n, z, \psi_n(z)), \tag{2.1}$$

for

$$\psi_n(z) := z - \frac{\rho^2 G_n(z)}{1 + 2\rho G_U^n(z)}, \tag{2.2}$$

and

$$G_U^n(z) := \mathbb{E}\Big[\frac{1}{2n} \operatorname{Tr}\big\{\boldsymbol{U}_n(zI_{2n} - \mathbf{Y}_n)^{-1}\big\} \,|\, \boldsymbol{T}_n\Big],$$

where for all $z_1, z_2 \in \mathbb{C}^+$

$$\widetilde{O}(n, z_1, z_2) = \frac{2O(n, z_1, z_2)}{1 + 2\rho G_U^n(z_1)}, \tag{2.3}$$

with $O(n, z_1, z_2)$ as defined in [7, pp. 1202]. Thus, (2.1) and (2.2) provide a relation between G_n and $G_{\boldsymbol{T}_n}$ which is very useful for our proof. Indeed, from [7, Lem. 12] we have that there exists a constant $C_1 := C_1(R)$ finite such that, for all large n, if $\Im(z) > C_1 n^{-1/4}$ then

$$\Im(\psi_n(z)) \geq \Im(z)/2. \tag{2.4}$$

Additionally, from [7, Eqn. (34)] we have that

$$\rho(G_n(z))^2 = 2G_U^n(z)(1 + 2\rho G_U^n(z)) - O_1(n, z), \tag{2.5}$$

where $O_1(\cdot, \cdot)$ is as defined in [7, pp. 1203]. To this end, denoting

$$F(G_n(z)) := \frac{\rho^2 G_n(z)}{1 + 2\rho G_U^n(z)}, \tag{2.6}$$

and using (2.5), we obtain after some algebra the identity

$$G_n(z)\Big[\rho^2 - F^2(G_n(z))\Big] = F(G_n(z))\Big[1 + \frac{\rho O_1(n, z)}{1 + 2\rho G_U^n(z)}\Big]. \tag{2.7}$$

Since

$$1 + 2\rho G_U^n(z) = \frac{1}{2}\Big(1 + \sqrt{1 + 4\rho^2 G_n(z)^2 + 4\rho O_1(n, z)}\Big), \tag{2.8}$$

where the branch of the square root is uniquely determined by analyticity and the known behavior of $G_U^n(z)$ and $G_n(z)$ as $|z| \to \infty$ (see [7, Eqn. (35)]), we further have that

$$\begin{aligned} F(G_n(z)) &= \frac{2\rho^2 G_n(z)}{1 + \sqrt{1 + 4(\rho G_n(z))^2 + 4\rho O_1(n, z)}} \\ &= \frac{1}{2}\Big[\frac{\rho^2 G_n(z)\sqrt{1 + 4(\rho G_n(z))^2 + 4\rho O_1(n, z)}}{(\rho G_n(z))^2 + \rho O_1(n, z)} - \frac{\rho^2 G_n(z)}{(\rho G_n(z))^2 + \rho O_1(n, z)}\Big]. \end{aligned} \tag{2.9}$$

The key to our proof is the observation that if $|\Im(G_n(z))| \to \infty$ and $O_1(n,z)$ remains small, then from (2.9), and (2.2) necessarily $F(G_n(z)) = z - \psi_n(z) \to \pm\rho$. So, if $\widetilde{O}(n,z,\psi_n(z))$ remains bounded then by (2.1) also $|\Im(G_{\boldsymbol{T}_n}(\psi_n(z)))| \to \infty$, yielding the required result.
To implement this, fix $M = M_\varepsilon \geq 10$ such that $6M_\varepsilon^{-1} \leq \varepsilon^2$ and recall that by [7, Lem. 11] there exists finite constant $C_2 := C_2(R)$ such that, for all large n, if $\Im(z) > C_1 n^{-1/4}$ then

$$|1 + 2\rho G_U^n(z)| > C_2\rho[\Im(z)^3 \wedge 1]. \tag{2.10}$$

Furthermore, we have (see [7, pp. 1203]),

$$|O(n,z_1,z_2)| \leq \frac{C\rho^2}{n^2|\Im(z_2)|\Im(z_1)^2(\Im(z_1) \wedge 1)}\,. \tag{2.11}$$

Therefore, enlarging C_1 as needed, by (2.3), (2.4), and (2.10) we obtain that, for all large n,

$$|\widetilde{O}(n,z,\psi_n(z))| \leq \frac{C\rho}{n^2|\Im(\psi_n(z))|\Im(z)^2(\Im(z)^4 \wedge 1)} \leq M_\varepsilon$$

whenever $\Im(z) > C_1 n^{-1/4}$. This, together with (2.1), shows that if $|\Im(G_n(z))| > 2M_\varepsilon$, then we have that $|\Im(G_{\boldsymbol{T}_n}(\psi_n(z)))| > M_\varepsilon$. Now, fixing $0 < \kappa_1 < \kappa < 1/4$ we get from (2.4) that $\Im(\psi_n(z)) > n^{-\kappa}$. It thus remains to show only that $F(G_n(z)) \in B(-\rho,\varepsilon) \cup B(\rho,\varepsilon)$. To this end, note that

$$|O_1(n,z)| \leq \frac{C\rho^2}{n^2\Im(z)^2(\Im(z) \wedge 1)} \tag{2.12}$$

(c.f. [7, pp. 1203]). Therefore, $O_1(n,z) = o(n^{-1})$ whenever $\Im(z) > C_1 n^{-1/4}$, and so the rightmost term in (2.9) is bounded by M_ε^{-1} whenever $|\Im(G_n(z))| > 2M_\varepsilon$. Further, when $\Im(z) > C_1 n^{-1/4}$, $|\Im(G_n(z))| > 2M_\varepsilon$ and n is large enough so $|O_1(n,z)| \leq 1$, we have that for any choice of the branch of the square root,

$$\left|\frac{\rho G_n(z)\sqrt{1 + 4(\rho G_n(z))^2 + 4\rho O_1(n,z)}}{(\rho G_n(z))^2 + \rho O_1(n,z)}\right| \leq \frac{\sqrt{1 + 4|\rho G_n(z)|^2 + 4|\rho O_1(n,z)|}}{|\rho G_n(z)| - 1} \leq 4\,,$$

resulting with $|F(G_n(z))| \leq 3\rho$. Therefore, using (2.10) and (2.12), we get from (2.7) that if $\Im(z) > C_1 n^{-1/4}$ and $|\Im(G_n(z))| > 2M_\varepsilon$, then

$$\left|F^2(G_n(z)) - \rho^2\right| \leq 6|G_n(z)|^{-1} \leq 6M_\varepsilon^{-1} \leq \varepsilon^2\,.$$

In conclusion, $z - \psi_n(z) = F(G_n(z)) \in B(\rho,\varepsilon) \cup B(-\rho,\varepsilon)$, as stated. Further, upon modifying the values of $\kappa_1 < \kappa$ and M_ε, this holds also when replacing U_n by a Haar distributed orthogonal matrix O_n. Indeed, the same analysis applies except for adding to $O(n,z_1,z_2)$ of [7, pp. 1202] a term which is uniformly bounded by $n^{-1}|\Im(z_2)|^{-1}(\Im(z_1) \wedge 1)^{-2}$ (see [7, proof of Thm. 18]), and using in this case [1, Cor. 4.4.28] to control the variance of Lipschitz functions of O_n (instead of U_n).
Proof of (ii): Consider first the case of $d' = d$. Then, setting $\rho = 1$, $\boldsymbol{T}_n = \boldsymbol{V}_n^{d-1,v}$, and $\boldsymbol{Y}_n = \boldsymbol{V}_n^{d,v}$, one may check that following the derivation of [7, Eqn. (37)-(38)], now with *all* expectations taken also *over* $\mathbf{T_n}$, we get that

$$G_n^{d,v}(z) = G_n^{d-1,v}(\psi_n^{d,v}(z)) - \widetilde{O}(n,z,\psi_n^{d,v}(z))\,, \tag{2.13}$$

for some $K < \infty$ and all $\{z \in \mathbb{C}^+ : \Im(z) \geq K\}$, where

$$\psi_n^{d,v}(z) := z - \frac{G_n^{d,v}(z)}{1 + 2G_{U_n}^{d,v}(z)}\;, \tag{2.14}$$

$$G^{d,v}_{U_n}(z) := \mathbb{E}\Big[\frac{1}{2n}\operatorname{Tr}\big\{\boldsymbol{U}^d_n(zI_{2n}-\boldsymbol{V}^{d,v}_n)^{-1}\big\}\Big],$$

and for any $z_1, z_2 \in \mathbb{C}^+$,

$$\widetilde{O}(n,z_1,z_2) := \frac{2O(n,z_1,z_2)}{1+2G^{d,v}_{U_n}(z_1)}.$$

Next, note that for some $C<\infty$ and any $\mathbb{C}$-valued function $f_d(U^1_n,\dots,U^d_n)$ of i.i.d. Haar distributed $\{U^i_n\}$

$$\mathbb{E}[(f_d-\mathbb{E}[f_d])^2] \le dC\|f_d\|^2_L, \tag{2.15}$$

where $\|f_d\|_L$ denotes the relevant coordinate-wise Lipschitz norm, i.e.

$$\|f_d\|_L := \max_{j=1}^{d} \sup_{U^1_n,\dots,U^d_n,\widetilde{U}_n\neq U^j_n} \frac{|f_d(U^1_n,\dots,U^d_n)-f_d(U^1_n,\dots,U^{j-1}_n,\widetilde{U}_n,U^{j+1}_n,\dots)|}{\|U^j_n-\widetilde{U}_n\|_2}.$$

Indeed, we bound the variance of f_d by the (sum of d) second moments of martingale differences $D_jf_d := \mathbf{E}[f_d|U^1_n,\dots,U^j_n]-\mathbf{E}[f_d|U^1_n,\dots,U^{j-1}_n]$. By the independence of $\{U^i_n\}$ and definition of $\|f_d\|_L$, conditional upon $(U^1_n,\dots,U^{j-1}_n)$, the $\mathbb{C}$-valued function $U^j_n \mapsto D_jf_d$ is Lipschitz of norm at most $\|f_d\|_L$ in the sense of [1, Ineq. (4.4.31)]. It then easily follows from the concentration inequalities of [1, Cor. 4.4.28], that the second moment of this function is at most $C\|f_d\|^2_L$ (uniformly with respect to $(U^1_n,\dots,U^{j-1}_n)$).
In the derivation of [7, Lem. 10], the corresponding error term $O(n,z_1,z_2)$ is bounded by a sum of finitely many variances of Lipschitz functions of the form $\frac{1}{2n}\operatorname{Tr}\{H(U^d_n)\}$, each of which has Lipschitz norm of order $n^{-1/2}$, hence controlled by applying the concentration inequality (2.15). We have here the same type of bound on $O(n,z_1,z_2)$, except that each variance in question is now with respect to some function $\frac{1}{2n}\operatorname{Tr}\{H(U^1_n,\dots,U^d_n)\}$ having coordinate-wise Lipschitz norm of order $n^{-1/2}$ (and with respect to the joint law of the i.i.d. Haar distributed unitary matrices). Collecting all such terms, we get here instead of (2.11), the slightly worse bound

$$|O(n,z_1,z_2)| = O\bigg(\frac{1}{n|\Im(z_2)|\Im(z_1)^2(\Im(z_1)\wedge 1)^2(\Im(z_2)\wedge 1)}\bigg) \tag{2.16}$$

(with an extra factor $(\Im(z_2)\wedge 1)^{-1}$ due to the additional randomness in $(z_2I_{2n}-\boldsymbol{T}_n)^{-1}$). Using the modified bound (2.16), we proceed as in the proof of part (i) of the lemma, to first bound $\widetilde{O}(n,z,\psi^{d,v}_n(z))$, $O_1(n,z)$, and derive the inequalities replacing (2.4) and (2.10). Out of these bounds, we establish the stated relation (1.15) between $G^{d,v}_n$ and $G^{d-1,v}_n$ upon following the same route as in our proof of part (i). Indeed, when doing so, the only effect of starting with (2.16) instead of (2.11) is in somewhat decreasing the positive constants κ_1, κ, while increasing each of the finite constants $\{M_\varepsilon, \varepsilon>0\}$.
Finally, with [1, Cor. 4.4.28] applicable also over the orthogonal group, our proof of (2.15) extends to any $\mathbb{C}$-valued function $f_d(U^1_n,\dots,U^{d'}_n,O^{d'+1}_n,\dots,O^d_n)$ of independent Haar distributed unitary/orthogonal matrices $\{U^i_n, O^i_n\}$. Hence, as in the context of part (i), the same argument applies for $0\le d'<d$ (up to adding $n^{-1}|\Im(z_2)|^{-1}(\Im(z_1)\wedge 1)^{-2}$ to (2.16), c.f. [7, proof of Thm. 18]).

3 Proof of Proposition 1.4

It suffices to prove Proposition 1.4 only for $d\ge 2$, since the easier case of $d=1$ has already been established in [12, Cor. 2.8]. We proceed to do so via the four steps of Girko's method, as described in Section 1. The following two lemmas (whose proof is deferred to Section 4), take care of **Step 1** and **Step 2** of Girko's method, respectively.

Lemma 3.1. *Let $\lambda_1 = \frac{1}{2}(\delta_{-1}+\delta_1)$ and $\Theta^{d,v} := \Theta^{d-1,v} \boxplus \lambda_1$ for all $d \geq 2$, starting at $\Theta^{1,v}$ which for $v \neq 0$ is the symmetrized version of the measure on $\mathbb{R}^+$ having the density $f_{|v|}(\cdot)$ of (4.1), while $\Theta^{1,0} = \lambda_1$. Then, for each $v \in \mathbb{C}$ and $d \in \mathbb{N}$, the* ESD-*s $L_{\boldsymbol{V}_n^{d,v}}$ of the matrices $\boldsymbol{V}_n^{d,v}$ (see (1.10)), converge weakly as $n \to \infty$, in probability, to $\Theta^{d,v}$.*

Lemma 3.2. *For any $d \geq 2$ and Lebesgue almost every $v \in \mathbb{C}$,*

$$\langle \mathrm{Log}, L_{\boldsymbol{V}_n^{d,v}} \rangle \to \langle \mathrm{Log}, \Theta^{d,v} \rangle, \tag{3.1}$$

in probability. Furthermore, there exist closed $\Lambda_d \subset \mathbb{C}$ of zero Lebesgue measure, such that

$$\int_{\mathbb{C}} \phi(v) \langle \mathrm{Log}, L_{\boldsymbol{V}_n^{d,v}} \rangle dm(v) \to \int_{\mathbb{C}} \phi(v) \langle \mathrm{Log}, \Theta^{d,v} \rangle dm(v), \tag{3.2}$$

in probability, for each fixed, non-random $\phi \in C_c^\infty(\mathbb{C})$ whose support is disjoint of Λ_d. That is, the support of ϕ is contained for some $\gamma > 0$, in the bounded, open set

$$\Gamma_\gamma^d := \left\{ v \in \mathbb{C} : \gamma < |v| < \gamma^{-1},\ \inf_{u \in \Lambda_d} \{ |v-u| \} > \gamma \right\}. \tag{3.3}$$

We claim that the convergence result of (3.2) provides us already with the conclusion (1.5) of **Step 3** in Girko's method, for test functions in

$$\mathcal{S} := \{ \psi \in C_c^\infty(\mathbb{C}),\ \text{supported within } \Gamma_\gamma^d \text{ for some } \gamma > 0 \}.$$

Indeed, fixing $d \geq 2$, the Hermitian matrices $\boldsymbol{V}_n^{d,v}$ of (1.10) are precisely those H_n^v of the form (1.3) that are associated with $S_n := \sum_{i=1}^d U_n^i$ in Girko's formula (1.4). Thus. combining the latter identity for $\psi \in \mathcal{S}$ with the convergence result of (3.2) for $\phi = \Delta\psi$, we get the following convergence in probability as $n \to \infty$,

$$\int_{\mathbb{C}} \psi(v) dL_{S_n}(v) = \frac{1}{2\pi} \int_{\mathbb{C}} \Delta\psi(v) \langle \mathrm{Log}, L_{\boldsymbol{V}_n^{d,v}} \rangle dm(v) \to \frac{1}{2\pi} \int_{\mathbb{C}} \Delta\psi(v) \langle \mathrm{Log}, \Theta^{d,v} \rangle dm(v). \tag{3.4}$$

Proceeding to identify the limiting measure as the Brown measure $\mu_d := \mu_{s_d}$ of the sum $s_d := u_1 + u_2 + \cdots + u_d$ of $\star$-free Haar unitary operators u_i, recall [14] that each (u_i, u_i^*) is $\mathbf{R}$-diagonal. Hence, by [9, Propn. 3.5] we have that $\Theta^{d,v}$ is the symmetrized version of the law of $|s_d - v|$, and so by definition (1.2) we have that for any $\psi \in C_c^\infty(\mathbb{C})$,

$$\frac{1}{2\pi} \int_{\mathbb{C}} \Delta\psi(v) \langle \mathrm{Log}, \Theta^{d,v} \rangle dm(v) = \int_{\mathbb{C}} \psi(v) \mu_{s_d}(dv). \tag{3.5}$$

In parallel with **Step 4** of Girko's method, it thus suffices for completing the proof, to verify that the convergence in probability

$$\int_{\mathbb{C}} \psi(v) dL_{S_n}(v) \to \int_{\mathbb{C}} \psi(v) d\mu_{s_d}(v), \tag{3.6}$$

for each *fixed* $\psi \in \mathcal{S}$, yields the weak convergence, in probability, of L_{S_n} to μ_{s_d}.

To this end, suppose first that (3.6) holds almost surely for each fixed $\psi \in \mathcal{S}$, and recall that for any $\gamma > 0$ and each open $G \subset \Gamma_\gamma^d$ there exist $\psi_k \in \mathcal{S}$ such that $\psi_k \uparrow 1_G$. Consequently, a.s.

$$\liminf_{n\to\infty} L_{S_n}(G) \geq \sup_k \liminf_{n\to\infty} \int_{\mathbb{C}} \psi_k(v) dL_{S_n}(v) = \sup_k \int_{\mathbb{C}} \psi_k(v) d\mu_{s_d}(v) = \mu_{s_d}(G).$$

Further, from [9, Example 5.5] we know that μ_{s_d} has, for $d \geq 2$, a bounded density with respect to Lebesgue measure on $\mathbb{C}$ (given by $h_d(\cdot)$ of (1.1)). In particular, since

$m(\Lambda_d) = 0$, it follows that $\mu_{s_d}(\Lambda_d) = 0$ and hence $\mu_{s_d}(\Gamma^d_\gamma) \to 1$ when $\gamma \to 0$. Given this, fixing some $\gamma_\ell \downarrow 0$ and open $G \subset \mathbb{C}$, we deduce that a.s.

$$\liminf_{n\to\infty} L_{S_n}(G) \ge \lim_{\ell\to\infty} \liminf_{n\to\infty} L_{S_n}(G \cap \Gamma^d_{\gamma_\ell}) \ge \lim_{\ell\to\infty} \mu_{s_d}(G \cap \Gamma^d_{\gamma_\ell}) = \mu_{s_d}(G)\,. \tag{3.7}$$

This applies for any countable collection $\{G_i\}$ of open subsets of $\mathbb{C}$, with the reversed inequality holding for any countable collection of closed subsets of $\mathbb{C}$. In particular, fixing any countable convergence determining class $\{f_j\} \subset C_b(\mathbb{C})$ and countable dense $\widehat{\mathbb{Q}} \subset \mathbb{R}$ such that $\mu_{s_d}(f_j^{-1}(\{q\})) = 0$ for all j and $q \in \widehat{\mathbb{Q}}$, yield the countable collection $\mathcal{G}$ of μ_{s_d}-continuity sets (consisting of interiors and complement of closures of $f_j^{-1}([q, q'))$, $q, q' \in \widehat{\mathbb{Q}}$), for which $L_{S_n}(\cdot)$ converges to $\mu_{s_d}(\cdot)$. The stated a.s. weak convergence of L_{S_n} to μ_{s_d} then follows as in the usual proof of Portmanteau's theorem, under our assumption that (3.6) holds a.s.

This proof extends to the case at hand, where (3.6) holds in probability, since convergence in probability implies that for every subsequence, there exists a further subsequence along which a.s. convergence holds, and the whole argument uses only countably many functions $\psi_{k,\ell,i} \in \mathcal{S}$. Specifically, by a Cantor diagonal argument, for any given subsequence n_j, we can extract a further subsequence $j(l)$, such that (3.7) holds a.s. for $L_{S_{n_{j(l)}}}$ and all G in the countable collection $\mathcal{G}$ of μ_{s_d}-continuity sets. Therefore, a.s. $L_{S_{n_{j(l)}}}$ converges weakly to μ_{s_d} and by the arbitrariness of $\{n_j\}$ we have that, in probability, L_{S_n} converges to μ_{s_d} weakly.

4 Proofs of Lemma 3.1 and Lemma 3.2

We start with a preliminary result, needed for proving Lemma 3.1.

Lemma 4.1. *For Haar distributed U_n and any $r > 0$, the expected* ESD *of $|U_n - rI_n|$ has the density*

$$f_r(x) = \frac{2}{\pi} \frac{x}{\sqrt{(x^2 - (r-1)^2)((r+1)^2 - x^2)}}\,, \quad |r-1| \le x \le r+1 \tag{4.1}$$

with respect to Lebesgue's measure on $\mathbb{R}^+$ (while for $r = 0$, this ESD *consists of a single atom at $x = 1$).*

Proof: It clearly suffices to show that the expected ESD of $(U_n - rI_n)(U_n - rI_n)^*$ has for $r > 0$ the density

$$g_r(x) = \frac{1}{\pi} \frac{1}{\sqrt{(x - (r-1)^2)((r+1)^2 - x)}}, \quad (r-1)^2 \le x \le (r+1)^2\,. \tag{4.2}$$

To this end note that by the invariance of the Haar unitary measure under multiplication by $e^{i\theta}$, we have that

$$\mathbb{E}[\frac{1}{n}\,\mathrm{Tr}\{U_n^k\}] = \mathbb{E}[\frac{1}{n}\,\mathrm{Tr}\{(U_n^*)^k\}] = 0\,, \tag{4.3}$$

for all positive integers k and n. Thus,

$$\mathbb{E}\Big[\frac{1}{n}\,\mathrm{Tr}\left\{(U_n + U_n^*)^k\right\}\Big] = \binom{k}{k/2} \text{ for } k \text{ even and } 0 \text{ otherwise.}$$

Therefore, by the moment method, the expected ESD of $U_n + U_n^*$ (denoted $\bar{L}_{U_n+U_n^*}$), satisfies

$$\bar{L}_{U_n+U_n^*} \overset{d}{=} 2\cos\theta = e^{i\theta} + e^{-i\theta}, \text{ where } \theta \sim \mathrm{Unif}(0, 2\pi).$$

Consequently, we get the formula (4.2) for the density $g_r(x)$ of the expected ESD of

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

by applying the change of variable formula for $x = (1+r^2)-2r\cos\theta$ (and $\theta \sim \mathrm{Unif}(0, 2\pi)$). □

Proof of Lemma 3.1: Recall [1, Thm. 2.4.4(c)] that for the claimed weak convergence of $L_{\boldsymbol{V}_n^{d,v}}$ to $\Theta^{d,v}$, in probability, it suffices to show that per fixed $z \in \mathbb{C}^+$, the corresponding Stieltjes transforms

$$f_n^{d,v}(z) := \frac{1}{2n}\operatorname{Tr}\{(zI_{2n} - \boldsymbol{V}_n^{d,v})^{-1}\}$$

converge in probability to the Stieltjes transform $G_\infty^{d,v}(z)$ of $\Theta^{d,v}$. To this end, note that each $f_n^{d,v}(z)$ is a point-wise Lipschitz function of $\{U_n^i\}$, whose expected value is $G_n^{d,v}(z)$ of (1.11), and that $\|f_n\|_L \to 0$ as $n \to \infty$ (per fixed values of d, v, z). It thus follows from (2.15) that as $n \to \infty$,

$$\mathbb{E}[(f_n^{d,v}(z) - G_n^{d,v}(z))^2] \to 0$$

and therefore, it suffices to prove that per fixed d, $v \in \mathbb{C}$ and $z \in \mathbb{C}^+$, as $n \to \infty$,

$$G_n^{d,v}(z) \to G_\infty^{d,v}(z)\,. \tag{4.4}$$

Next observe that by invariance of the law of U_n^1 to multiplication by scalar $e^{i\theta}$, the expected ESD of $\boldsymbol{V}_n^{1,v}$ depends only on $r = |v|$, with $\Theta^{1,v} = \mathbb{E}[L_{\boldsymbol{V}_n^{1,v}}]$ (see Lem. 4.1). Hence, (4.4) trivially holds for $d = 1$ and we proceed to prove the latter pointwise (in z, v), convergence by an induction on $d \geq 2$. The key ingredient in the induction step is the (finite n) Schwinger-Dyson equation in our set-up, namely Eqn. (2.13)-(2.14). Specifically, from (2.13)-(2.14) and the induction hypothesis it follows that for some non-random $K < \infty$, any limit point, denoted $(G^{d,v}, G_U^{d,v})$, of the uniformly bounded, equi-continuous functions $(G_n^{d,v}, G_{U_n}^{d,v})$ on $\{z \in \mathbb{C}^+ : \Im(z) \geq K\}$, satisfies

$$G^{d,v}(z) = G_\infty^{d-1,v}(\psi(z)), \text{ with } \psi(z) := z - \frac{G^{d,v}(z)}{1 + 2G_U^{d,v}(z)}\,. \tag{4.5}$$

Moreover, from the equivalent version of (2.5) in our setting, we obtain that

$$4G_U^{d,v}(z) = -1 + \sqrt{1 + 4G^{d,v}(z)^2}\,,$$

for a suitable branch of the square root (uniquely determined by analyticity and decay to zero as $|z| \to \infty$ of $z \mapsto (G^{d,v}(z), G_U^{d,v}(z))$). Thus, $G(z) = G^{d,v}(z)$ satisfies the relation

$$G(z) - G_\infty^{d-1,v}\Big(z - \frac{2G(z)}{1 + \sqrt{1 + 4G(z)^2}}\Big) = 0\,. \tag{4.6}$$

Since $\Theta^{d,v} = \Theta^{d-1,v} \boxplus \lambda_1$, it follows that (4.6) holds also for $G(\cdot) = G_\infty^{d,v}(\cdot)$ (c.f. [7, Rmk. 7]). Further, $z \mapsto G_\infty^{d-1,v}(z)$ is analytic on $\mathbb{C}^+$ with derivative of $O(z^{-2})$ at infinity, hence by the implicit function theorem the identity (4.6) uniquely determines the value of $G(z)$ for all $\Im(z)$ large enough. In particular, enlarging K as needed, $G^{d,v} = G_\infty^{d,v}$ on $\{z \in \mathbb{C}^+ : \Im(z) \geq K\}$, which by analyticity of both functions extends to all of $\mathbb{C}^+$. With (4.4) verified, this completes the proof of the lemma. □

The proof of Lemma 3.2 requires the control of $\Im(G_n^{d,v}(z))$ as established in Lemma 4.3. This is done inductively in d, with Lemma 4.2 providing the basis $d = 1$ of the induction.

Lemma 4.2. *For some C finite, all $\varepsilon \in (0,1)$ and $v \in \mathbb{C}$,*

$$\Big\{z \in \mathbb{C}^+ : \ |\Im G_n^{1,v}(z)| \geq C\varepsilon^{-2}\Big\} \subseteq \Big\{E+i\eta : \eta \in (0,\varepsilon^2), E \in \big(\pm(1\pm|v|)-2\varepsilon, \pm(1\pm|v|)+2\varepsilon\big)\Big\}.$$

Proof: It is trivial to confirm our claim in case $v = 0$ (as $G_n^{1,0}(z) = z/(z^2-1)$). Now, fixing $r = |v| > 0$, let $\widetilde{f}_r(\cdot)$ denote the symmetrized version of the density $f_r(\cdot)$, and note that for any $\eta > 0$,

$$\begin{aligned} |\Im G_n^{1,v}(E+i\eta)| &= \int_{|x-E|>\sqrt{\eta}} \frac{\eta}{(x-E)^2+\eta^2}\widetilde{f}_r(x)dx + \int_{|x-E|\leq\sqrt{\eta}} \frac{\eta}{(x-E)^2+\eta^2}\widetilde{f}_r(x)dx \\ &\leq 1 + \Big[\sup_{\{x:|x-E|\leq\sqrt{\eta}\}} \widetilde{f}_r(x)\Big]\int_{|x-E|\leq\sqrt{\eta}} \frac{\eta}{(x-E)^2+\eta^2}dx \\ &\leq 1+\pi\Big[\sup_{\{x:|x-E|\leq\sqrt{\eta}\}} \widetilde{f}_r(x)\Big]. \end{aligned} \tag{4.7}$$

With Γ_ε denoting the union of open intervals of radius ε around the four points $\pm 1 \pm r$, it follows from (4.1) that for some C_1 finite and any $r, \varepsilon > 0$,

$$\sup_{x \notin \Gamma_\varepsilon} \{\widetilde{f}_r(x)\} \leq C_1 \varepsilon^{-2}.$$

Thus, from (4.7) it follows that

$$\sup_{\{E,\eta:(E-\sqrt{\eta},E+\sqrt{\eta})\subset\Gamma_\varepsilon^c\}} |\Im G_n^{1,v}(E+i\eta)| \leq C\varepsilon^{-2},$$

for some C finite, all $\varepsilon \in (0,1)$ and $r > 0$. To complete the proof simply note that

$$\{(E,\eta) : E \in \Gamma_{2\varepsilon}^c, \eta \in (0,\varepsilon^2)\} \subseteq \{(E,\eta) : (E-\sqrt{\eta}, E+\sqrt{\eta}) \subseteq \Gamma_\varepsilon^c\},$$

and

$$\sup_{E\in\mathbb{R},\eta\geq\varepsilon^2} |\Im G_n^{1,v}(E+i\eta)| \leq \varepsilon^{-2}. \qquad \square$$

Since the density $\widetilde{f}_{|v|}(\cdot)$ is unbounded at $\pm 1 \pm |v|$, we can not improve Lemma 4.2 to show that $\Im G_n^{1,v}(z)$ is uniformly bounded. The same applies for $d \geq 2$ so a result such as [7, Lem. 13] is not possible in our set-up. Instead, as we show next, inductively applying Lemma 1.5(ii) allows us to control the region where $|\Im(G_n^{d,v}(z))|$ might blow up, in a manner which suffices for establishing Lemma 3.2 (and consequently Proposition 1.4).

Lemma 4.3. *For $r \geq 0$, $\gamma > 0$ and integer $d \geq 1$, let $\Gamma_\gamma^{d,r} \subset \mathbb{C}$ denote the union of open balls of radius γ centered at $\pm m \pm r$ for $m = 0,1,2,\ldots,d$. Fixing integer $d \geq 1$, $\gamma \in (0,1)$ and R finite, there exist M finite and $\kappa > 0$ such that for all n large enough and any $v \in B(0,R)$,*

$$\sup\{|\Im(G_n^{d,v}(z))| : \ \Im(z) > n^{-\kappa}, \ z \notin \Gamma_\gamma^{d,|v|}\} \leq M. \tag{4.8}$$

Proof: For any $d \geq 1$, $v \in \mathbb{C}$, positive κ and finite M, set

$$\Gamma_n^{d,v}(M,\kappa) := \{z : \Im(z) > n^{-\kappa}, |\Im(G_n^{d,v}(z))| > M\},$$

so our thesis amounts to the existence of finite M and $\kappa > 0$, depending only on R, $d \geq 2$ and $\gamma \in (0,1)$, such that for all n large enough,

$$\Gamma_n^{d,v}(M,\kappa) \subset \Gamma_\gamma^{d,|v|}, \qquad \forall v \in B(0,R). \tag{4.9}$$

Indeed, for $d = 1$ this is a direct consequence of Lemma 4.2 (with $\gamma = 2\varepsilon$, $M = C\varepsilon^{-2}$), and we proceed to confirm (4.9) by induction on $d \geq 2$. To carry out the inductive step from $d-1$ to d, fix R finite and $\gamma \in (0,1)$, assuming that (4.9) applies at $d-1$ and $\gamma/2$, for some finite $M_\star$ and positive $\kappa_\star$ (both depending only on d, R and γ). Then, let $\varepsilon \in (0, \gamma/2)$ be small enough such that Lemma 1.5(ii) applies for some $M_\varepsilon \geq M_\star$ and $0 < \kappa_1 < \kappa \leq \kappa_\star$. From Lemma 1.5(ii) we know that for any n large enough, $v \in B(0,R)$ and $z \in \Gamma_n^{d,v}(2M_\varepsilon, \kappa_1)$, there exists $w := \psi_n^{d,v}(z)$ for which

$$z - w \in B(-1,\varepsilon) \cup B(1,\varepsilon) \qquad \& \qquad w \in \Gamma_n^{d-1,v}(M_\varepsilon, \kappa) \subseteq \Gamma_n^{d-1,v}(M_\star, \kappa_\star) \subset \Gamma_{\gamma/2}^{d-1,|v|},$$

where the last inclusion is due to our choice of $M_\star$ and $\kappa_\star$. With $\varepsilon \leq \gamma/2$, it is easy to check that $z - w \in B(-1,\varepsilon) \cup B(1,\varepsilon)$ and $w \in \Gamma_{\gamma/2}^{d-1,r}$ result with $z \in \Gamma_\gamma^{d,r}$. That is, we have established the validity of (4.9) at d and arbitrarily small γ, for $M = 2M_\varepsilon$ finite and κ_1 positive, both depending only on R, d and γ. □

Proof of Lemma 3.2: Recall [15, Thm 1.1] the existence of universal constants $0 < c_1$ and $c_2 < \infty$, such that for any non-random matrix D_n and Haar distributed unitary matrix U_n, the smallest singular value $s_{\min}$ of $U_n + D_n$ satisfies,

$$\mathbb{P}(s_{\min}(U_n + D_n) \leq t) \leq t^{c_1} n^{c_2}. \tag{4.10}$$

The singular values of $\boldsymbol{V}_n^{d,v}$ are clearly the same as those of $S_n - vI_n = U_n^1 + D_n$ for $D_n = \sum_{i=2}^d U_n^i - vI_n$, which is independent of the Haar unitary U_n^1. Thus, applying (4.10) conditionally on D_n, we get that

$$\mathbb{P}(s_{\min}(\boldsymbol{V}_n^{d,v}) \leq t) \leq t^{c_1} n^{c_2}\,, \tag{4.11}$$

for every $v \in \mathbb{C}$, $t > 0$ and n. It then follows that for any $\delta > 0$ and $\alpha < c_1$,

$$\mathbb{E}\Big[(s_{\min}(\boldsymbol{V}_n^{d,v}))^{-\alpha} \mathbb{I}_{\{s_{\min}(\boldsymbol{V}_n^{d,v}) \leq n^{-\delta}\}}\Big] \leq \frac{c_1}{c_1 - \alpha} n^{c_2 - \delta(c_1-\alpha)}\,. \tag{4.12}$$

Setting hereafter $\alpha = c_1/2$ positive and $\delta = 4c_2/c_1$ finite, the right side of (4.12) decays to zero as $n \to \infty$. Further, for any n, d and v,

$$\mathbb{E}\Big[\langle|\mathrm{Log}|, L_{\boldsymbol{V}_n^{d,v}}\rangle_0^{n^{-\delta}}\Big] \leq \mathbb{E}\Big[|\log s_{\min}(\boldsymbol{V}_n^{d,v})| \mathbb{I}_{\{s_{\min}(\boldsymbol{V}_n^{d,v}) \leq n^{-\delta}\}}\Big]\,. \tag{4.13}$$

Hence, with $|x|^\alpha \log|x| \to 0$ as $x \to 0$, upon combining (4.12) and (4.13) we deduce that

$$\limsup_{n\to\infty} \sup_{v \in \mathbb{C}} \mathbb{E}\Big[\langle|\mathrm{Log}|, L_{\boldsymbol{V}_n^{d,v}}\rangle_0^{n^{-\delta}}\Big] = 0\,. \tag{4.14}$$

Next, consider the collection of sets Γ_γ^d as in (3.3), that corresponds to the compact

$$\Lambda_d := \big\{v \in \mathbb{C} : |v| \in \{0, 1, \ldots, d\}\big\}$$

(such that $m(\Lambda_d) = 0$). In this case, $v \in \Gamma_\gamma^d$ implies that $\{iy : y > 0\}$ is disjoint of the set $\Gamma_\gamma^{d,|v|}$ of Lemma 4.3. For such values of v we thus combine the bound (4.8) of Lemma 4.3 with [7, Lem. 15], to deduce that for any integer $d \geq 1$ and $\gamma \in (0,1)$ there exist finite n_0, M and positive κ (depending only on d and γ), for which

$$\mathbb{E}\big[L_{\boldsymbol{V}_n^{d,v}}(-y,y)\big] \leq 2M(y \vee n^{-\kappa}) \qquad \forall n \geq n_0,\, y > 0,\, v \in \Gamma_\gamma^d\,. \tag{4.15}$$

Imitating the derivation of [7, Eqn. (49)], we get from (4.15) that for some finite $C = C(d,\gamma,\delta)$, any $\varepsilon \leq e^{-1}$, $n \geq n_0$ and $v \in \Gamma_\gamma^d$,

$$\mathbb{E}\Big[\langle|\mathrm{Log}|, L_{\boldsymbol{V}_n^{d,v}}\rangle_{n^{-\delta}}^{\varepsilon}\Big] \leq C\varepsilon|\log\varepsilon|\,. \tag{4.16}$$

Thus, combining (4.14) and (4.16) we have that for any $\gamma > 0$,

$$\lim_{\varepsilon \downarrow 0} \limsup_{n\to\infty} \sup_{v\in\Gamma_\gamma^d} \mathbb{E}\Big[\langle |\mathrm{Log}|, L_{\boldsymbol{V}_n^{d,v}}\rangle_0^\varepsilon\Big] = 0\,. \tag{4.17}$$

Similarly, in view of (4.4), the bound (4.8) implies that

$$|\Im(G_\infty^{d,v}(z))| \le M\,, \qquad \forall z \in \mathbb{C}^+\backslash\Gamma_\gamma^{d,|v|},\, v \in B(0,R)\,,$$

which in combination with [7, Lem. 15], results with

$$\Theta^{d,v}(-y,y) \le 2My \qquad \forall y > 0,\, v\in\Gamma_\gamma^d$$

and consequently also

$$\lim_{\varepsilon\downarrow 0} \sup_{v\in\Gamma_\gamma^d} \{\langle |\mathrm{Log}|, \Theta^{d,v}\rangle_0^\varepsilon\} = 0\,. \tag{4.18}$$

Next, by Lemma 3.1, the real valued random variables $X_n^{(\varepsilon)}(\omega,v) := \langle \mathrm{Log}, L_{\boldsymbol{V}_n^{d,v}}\rangle_\varepsilon^\infty$ converge in probability, as $n\to\infty$, to the non-random $X_\infty^{(\varepsilon)}(v) := \langle \mathrm{Log}, \Theta^{d,v}\rangle_\varepsilon^\infty$, for each $v\in\mathbb{C}$ and $\varepsilon > 0$. This, together with (4.17) and (4.18), results with the stated convergence of (3.1), for each $v \in \Gamma_\gamma^d$, so considering $\gamma\to 0$ we conclude that (3.1) applies for all $v\in\Lambda_d^c$, hence for m-a.e. v.
Turning to prove (3.2), fix $\gamma > 0$ and non-random, uniformly bounded ϕ, supported within Γ_γ^d. Since $\{L_{\boldsymbol{V}_n^{d,v}}, v\in\Gamma_\gamma^d\}$ are all supported on $B(0,\gamma^{-1}+d)$, for each fixed $\varepsilon>0$, the random variables $Y_n^{(\varepsilon)}(\omega,v) := \phi(v)X_n^{(\varepsilon)}(\omega,v)m(\Gamma_\gamma^d)$ with respect to the product law $\overline{\mathbb{P}} := \mathbb{P}\times m(\cdot)/m(\Gamma_\gamma^d)$ on (ω,v) are bounded, uniformly in n. Consequently, their convergence in $\mathbb{P}$-probability, for m-a.e. v, to $Y_\infty^{(\varepsilon)}(v)$ (which we have already established), implies the corresponding L_1-convergence. Furthermore, by (4.17) and Fubini's theorem,

$$\overline{\mathbb{E}}[|Y_n^{(0)} - Y_n^{(\varepsilon)}|] \le m(\Gamma_\gamma^d)\|\phi\|_\infty \sup_{v\in\Gamma_\gamma^d} \mathbb{E}[|X_n^{(0)}(\omega,v) - X_n^{(\varepsilon)}(\omega,v)|] \to 0\,,$$

when $n\to\infty$ followed by $\varepsilon\downarrow 0$. Finally, by (4.18), the non-random $Y_\infty^{(\varepsilon)}(v) \to Y_\infty^{(0)}(v)$ as $\varepsilon \downarrow 0$, uniformly over Γ_γ^d. Consequently, as $n\to\infty$ followed by $\varepsilon\downarrow 0$,

$$\overline{\mathbb{E}}[|Y_n^{(0)} - Y_\infty^{(0)}|] \le \overline{\mathbb{E}}[|Y_n^{(0)} - Y_n^{(\varepsilon)}|] + \overline{\mathbb{E}}[|Y_n^{(\varepsilon)} - Y_\infty^{(\varepsilon)}|] + \sup_{v\in\Gamma_\gamma^d}\{|Y_\infty^{(0)} - Y_\infty^{(\varepsilon)}|\}$$

converges to zero and in particular

$$\int_{\mathbb{C}} \phi(v)X_n^{(0)}(\omega,v)dm(v) \to \int_{\mathbb{C}} \phi(v)X_\infty^{(0)}(v)dm(v)\,,$$

in L_1, hence in $\mathbb{P}$-probability, as claimed. □

5 Proof of Theorem 1.2

Following the proof of Proposition 1.4, it suffices for establishing Theorem 1.2, to extend the validity of Lemmas 3.1 and 3.2 in case of $S_n = \sum_{i=1}^{d'} U_n^i + \sum_{i>d'}^d O_n^i$. To this end, recall that Lemma 1.5(ii) applies regardless of the value of d'. Hence, Lemmas 3.1 and 3.2 hold as soon as we establish Lemma 4.2, the bound (4.11) on $s_{\min}(\boldsymbol{V}_n^{d,v})$, and the convergence (4.4) for $d=1$. Examining Section 4, one finds that our proof of the latter three results applies as soon as $d'\ge 1$ (i.e. no need for new proofs if we start with U_n^1).
In view of the preceding, we set hereafter $d'=0$, namely consider the sum of (only) i.i.d Haar orthogonal matrices and recall that suffices to prove our theorem when $d\ge 2$

(for the case of $d=1$ has already been established in [12, Cor. 2.8]). Further, while the Haar orthogonal measure is *not invariant* under multiplication by $e^{i\theta}$, it is not hard to verify that nevertheless

$$\lim_{n\to\infty} \mathbb{E}[\frac{1}{n}\operatorname{Tr}\{O_n^k\}] = \mathbb{E}[\frac{1}{n}\operatorname{Tr}\{(O_n^*)^k\}] = 0\,,$$

for any positive integer k. Replacing the identity (4.3) by the preceding and thereafter following the proof of Lemma 4.1, we conclude that $\mathbb{E}[L_{\boldsymbol{O}_n^{1,v}}] \Rightarrow \Theta^{1,v}$ as $n\to\infty$, for each fixed $v\in\mathbb{C}$. This yields of course the convergence (4.4) of the corresponding Stieltjes transforms (and thereby extends the validity of Lemma 3.1 even for $d'=0$). Lacking the identity (4.3), for the orthogonal case we replace Lemma 4.2 by the following.

Lemma 5.1. *The Stieltjes transform* $G_n^{1,v}$ *of the* ESD $\mathbb{E}[L_{\boldsymbol{O}_n^{1,v}}]$ *is such that*

$$\begin{aligned}\big\{z\in\mathbb{C}^+ :\ |\Im G_n^{1,v}(z)|\ge C\varepsilon^{-2}\big\} \subset \Big\{& E+i\eta : \eta\in(0,\varepsilon^2),\\ & E\in\big(\pm(1\pm|v|)-2\varepsilon,\pm(1\pm|v|)+2\varepsilon\big)\cup\Big(\pm(|1\pm v|-2\varepsilon,\pm(|1\pm v|)+2\varepsilon\Big)\Big\}\,,\end{aligned}$$

for some C *finite, all* $\varepsilon\in(0,1)$ *and any* $v\in\mathbb{C}$.

Proof: We express $G_n^{1,v}(z)$ as the expectation of certain additive function of the eigenvalues of O_n^1, whereby information about the marginal distribution of these eigenvalues shall yield our control on $|\Im(G_n^{1,v}(z))|$. To this end, set $g(z,r) := z/(z^2-r)$ for $z\in\mathbb{C}^+$, $r\ge 0$, and let $\phi(O_n^1) := \frac{1}{2n}\operatorname{Tr}\{(zI_{2n}-\boldsymbol{O}_n^{1,v})^{-1}\}$. Clearly,

$$\phi(O_n^1) = \frac{1}{n}\sum_{k=1}^n g(z,s_k^2)\,, \tag{5.1}$$

where $\{s_k\}$ are the singular values of $O_n^1 - vI_n$. For any matrix A_n and orthogonal matrix $\widetilde{O}_n$, the singular values of A_n are the same as those of $\widetilde{O}_n A_n \widetilde{O}_n^*$. Considering $A_n = O_n^1 - vI_n$, we thus deduce from (5.1) that $\phi(\widetilde{O}_n O_n^1 \widetilde{O}_n^*) = \phi(O_n^1)$, namely that $\phi(\cdot)$ is a *central function* on the orthogonal group (see [1, pp. 192]).

The group of n-dimensional orthogonal matrices partitions into the classes $\mathcal{O}^+(n)$ and $\mathcal{O}^-(n)$ of orthogonal matrices having determinant $+1$ and -1, respectively. In case $n=2\ell+1$ is odd, any $O_n\in\mathcal{O}^\pm(n)$ has eigenvalues $\{\pm1, e^{\pm i\theta_j}, j=1,\ldots,\ell\}$, for some $\underline{\theta}=(\theta_1,\ldots,\theta_\ell)\in[-\pi,\pi]^\ell$. Similarly, for $n=2\ell$ even, $O_n\in\mathcal{O}^+(n)$ has eigenvalues $\{e^{\pm i\theta_j}, j=1,\ldots,\ell\}$, whereas $O_n\in\mathcal{O}^-(n)$ has eigenvalues $\{-1,1,e^{\pm i\theta_j}, j=1,\ldots,\ell-1\}$. Weyl's formula expresses the expected value of a central function of Haar distributed orthogonal matrix in terms of the joint distribution of $\underline{\theta}$ under the probability measures $\mathbb{P}_n^\pm$ corresponding to the classes $\mathcal{O}^+(n)$ and $\mathcal{O}^-(n)$. Specifically, it yields the expression

$$\begin{aligned}G_n^{1,v}(z) = \mathbb{E}[\phi(O_n^1)] &= \frac{1}{2}\mathbb{E}_n^+[\phi(\operatorname{diag}(+1,R_\ell(\underline{\theta}))] + \frac{1}{2}\mathbb{E}_n^-[\phi(\operatorname{diag}(-1,R_\ell(\underline{\theta}))]\,, \quad \text{for } n=2\ell+1,\\ &= \frac{1}{2}\mathbb{E}_n^+[\phi(\operatorname{diag}(R_\ell(\underline{\theta}))] + \frac{1}{2}\mathbb{E}_n^-[\phi(\operatorname{diag}(-1,1,R_{\ell-1}(\underline{\theta}))]\,, \quad \text{for } n=2\ell,\end{aligned} \tag{5.2}$$

where $R_\ell(\underline{\theta}) := \operatorname{diag}(R(\theta_1),R(\theta_2),\cdots,R(\theta_\ell))$ for the two dimensional rotation matrix

$$R(\theta) = \begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}$$

(see [1, Propn. 4.1.6], which also provides the joint densities of $\underline{\theta}$ under $\mathbb{P}_n^\pm$).

In view of (5.1) and (5.2), to evaluate $G_n^{1,v}(z)$ we need the singular values of $R_\ell(\underline{\theta}) - vI_\ell$. Since this is a block-diagonal matrix, its singular values are those of the 2×2 block diagonal parts $R(\theta_j) - vI_2$ for $1 \le j \le \ell$. Setting $v := |v|e^{i\psi}$ it is easy to check that the singular values of $R(\theta) - vI_2$ are precisely square-root of the eigenvalues of $(1+|v|^2)I_2 - |v|(e^{-i\psi}R(\theta) + e^{i\psi}R^*(\theta))$, which turn out to be $1 + |v|^2 - 2|v|\cos(\theta \pm \psi)$. Combining this with (5.1) and (5.2) we obtain in case $n = 2\ell + 1$, that

$$G_n^{1,v}(z) = \frac{1}{2n}\Big\{g(z,|1-v|^2) + \sum_{k=0}^{1}\sum_{j=1}^{\ell} \mathbb{E}_n^+[g(z, 1+|v|^2 - 2|v|\cos(\theta_j + (-1)^k\psi))] \\ + g(z,|1+v|^2) + \sum_{k=0}^{1}\sum_{j=1}^{\ell} \mathbb{E}_n^-[g(z, 1+|v|^2 - 2|v|\cos(\theta_j + (-1)^k\psi))]\Big\} . \quad (5.3)$$

The same expression applies for $n = 2\ell$, except for having the latter sum only up to $j = \ell - 1$. Next, recall that under $\mathbb{P}_n^\pm$ the random variables $\{\theta_j\}$ are exchangeable, each having the same density $q_n^\pm(\cdot)$ which is bounded, uniformly in n (see the diagonal terms in [5, Propn. 5.5.3]; for example, $q_{2\ell+1}^\pm(\theta) = \frac{1}{2\pi}(1 \mp \sin(2\ell\theta)/(2\ell\sin\theta))$, is bounded by $1/\pi$, uniformly over θ and ℓ). Further, $g(z,r) \in \mathbb{C}^-$ for all $r \ge 0$ and $z \in \mathbb{C}^+$. Hence, for some C finite, all $n \ge 3$, $v \in \mathbb{C}$ and $z \in \mathbb{C}^+$,

$$|\Im(G_n^{1,v}(z))| \le \frac{1}{2n}|\Im(g(z,|1-v|^2))| + \frac{1}{2n}|\Im(g(z,|1+v|^2))| \\ + C\Big|\Im\Big\{\frac{1}{2\pi}\int_{-\pi}^{\pi} g(z, 1+|v|^2 - 2|v|\cos(\theta \pm \psi))d\theta\Big\}\Big| . \quad (5.4)$$

The last expression in (5.4) does not depend on $\pm\psi$ and is precisely the imaginary part of the Stieltjes transform of the symmetrization of the probability measure $|e^{i\theta} - |v||$, where $\theta \sim \text{Unif}(0, 2\pi)$. While proving Lemma 4.1 we saw that the expected ESD of $\boldsymbol{U}_n^{1,v}$ has the latter law, hence the conclusion of Lemma 4.2 applies for the last expression in (5.4). To complete the proof, simply note that $\Im(g(E+i\eta, s^2)) \le 1$ as soon as $|E \pm s| \ge \sqrt{\eta}$ (and consider $s = |v \pm 1|$). □

Now, using Lemma 5.1 for the basis $d = 1$ of an induction argument (instead of Lemma 4.2), and with Lemma 1.5(ii) serving again for its inductive step, we obtain here the same conclusion as in Lemma 4.3, except for replacing $\Gamma_\gamma^{d,|v|}$ by the union $\widetilde{\Gamma}_\gamma^{d,v}$ of open balls of radius γ centered at the points $\pm m \pm 1 \pm |v|$ *and* $\pm m \pm |1 \pm v|$ for $m = 0, \ldots, d-1$. Turning to prove Lemma 3.2, this translates to taking in this case the sets Γ_γ^d which correspond via (3.3) to the compact

$$\Lambda_d := \big\{ v \in \mathbb{C} : |v| \in \{0, 1, \ldots, d\}, \quad \text{or} \quad |v \pm 1| \in \{0, 1, \ldots, d-1\} \big\}$$

(of zero Lebesgue measure), thereby assuring that $\{iy : y > 0\}$ is disjoint of $\widetilde{\Gamma}_\gamma^{d,v}$ whenever $v \in \Gamma_\gamma^d$. One may then easily check that the proof of Lemma 3.2 (and hence of the theorem), is completed upon establishing the following weaker form of (4.11).

Lemma 5.2. *For some $c_1 > 0$, $c_2 < \infty$, the sum S_n of $d \ge 2$ independent Haar orthogonal matrices and any $\gamma \in (0,1)$, there exist $C' = C'(d,\gamma)$ finite and events $\{\mathcal{G}_n\}$ determined by the minimal and maximal singular values of S_n, such that $\mathbb{P}(\mathcal{G}_n^c) \to 0$ as $n \to \infty$, and for any $n, t \ge 0$,*

$$\sup_{v \in \Gamma_\gamma^d} \mathbb{P}\Big(\mathcal{G}_n \cap \{s_{\min}(\boldsymbol{V}_n^{d,v}) \le t\}\Big) \le C' t^{c_1} n^{c_2} . \quad (5.5)$$

Proof: We use here [15, Thm. 1.3] (instead of [15, Thm. 1.1] which applies only for Haar unitary matrices), and introduce events $\mathcal{G}_n$ under which the condition [15, Eqn. (1.2)] holds. Specifically, let $D_n = \mathrm{diag}(r_1, r_2, \ldots, r_n)$ denote the diagonal matrix of singular values of S_n, ordered so that $r_1 \geq r_2 \geq \ldots \geq r_n$ and

$$\mathcal{G}_n := \{r_n \leq \frac{1}{2} \quad \text{and} \quad r_1 \geq 1\}.$$

Let O_n be Haar distributed n-dimensional orthogonal matrix, independent of $\{O_n^i, i = 1, \ldots, d\}$, noting that O_n is independent of $-O_n S_n$, with the latter having the same law and singular values as S_n. Further, the singular values of $\boldsymbol{V}_n^{d,v}$ equal to those of $vI_n - S_n = O_n^*(vO_n - O_n S_n)$, hence for any n and $t \geq 0$,

$$q_{n,v}(t) := \mathbb{P}\Big(\mathcal{G}_n \cap \{s_{\min}(\boldsymbol{V}_n^{d,v}) \leq t\}\Big) = \mathbb{P}\Big(\mathcal{G}_n \cap \{s_{\min}(vO_n + S_n) \leq t\}\Big).$$

Next, by the singular value decomposition $S_n = (O_n^{'})^* D_n (O_n^{''})^*$ for some pair of orthogonal matrices O_n' and $O_n^{''}$. Conditional on D_n, $O_n^{'}$ and $O_n^{''}$, the matrix $O_n^{'} O_n O_n^{''}$ is again Haar distributed, hence independent of D_n (and of $\mathcal{G}_n$). Consequently, for any $v \neq 0$,

$$q_{n,v}(t) = \mathbb{P}\Big(\mathcal{G}_n \cap \{s_{\min}(vO_n^{'} O_n O_n^{''} + D_n) \leq t\}\Big) = \mathbb{P}\Big(\mathcal{G}_n \cap \{|v| s_{\min}(O_n + v^{-1} D_n) \leq t\}\Big).$$

Now from [15, Thm. 1.3] we know that for some absolute constants $c_1 > 0$ and $c_2 < \infty$,

$$\mathbb{P}(|v| s_{\min}(O_n + v^{-1} D_n) \leq t \,|\, D_n) \leq \Big(\frac{t}{|v|}\Big)^{c_1} \Big(\frac{Kn}{\delta}\Big)^{c_2}, \tag{5.6}$$

provided [15, Eqn. (1.2)] holds for $v^{-1} D_n$, some $K \geq 1$ and $\delta \in (0, 1)$. That is, when

$$r_1 \leq K|v| \text{ and } r_1^2 \geq r_n^2 + \delta |v|^2. \tag{5.7}$$

In our setting the singular values of S_n are uniformly bounded by d and $|v| \in (\gamma, \gamma^{-1})$ throughout Γ_γ^d. Hence, the event $\mathcal{G}_n$ implies that (5.7) holds for $K = d/\gamma$ and $\delta = \gamma^2/2$. Thus, multiplying both sides of (5.6) by $\mathbb{I}_{\mathcal{G}_n}$ and taking the expectation over D_n yields the inequality (5.5) for some finite $C' = C'(d, \gamma)$.
Proceeding to verify that $\mathbb{P}(\mathcal{G}_n^c) \to 0$ as $n \to \infty$, recall [9, Propn. 3.5] that $\Theta^{d,0}$ is the symmetrization of the law $\mu_{|s_d|}$, for the sum $s_d = u_1 + \cdots + u_d$ of $\star$-free Haar unitary operators $u_1, \ldots, u_d$, and [9, Eqn. (5.7)] that for $d \geq 2$ the measure $\mu_{|s_d|}$ on $\mathbb{R}^+$ has the density

$$\frac{d\mu_{|s_d|}}{dx} = \frac{d\sqrt{4(d-1) - x^2}}{\pi(d^2 - x^2)} \mathbb{I}_{[0, 2\sqrt{d-1}]}(x), \tag{5.8}$$

so in particular both $\mu_{|s_d|}((0, 1/2))$ and $\mu_{|s_d|}((1, 3/2))$ are strictly positive. Further, from Lemma 3.1 we already know that the symmetrization of the ESD $\nu_{|S_n|}$ of D_n, converges weakly, in probability, to $\Theta^{d,0}$ and consequently, $\nu_{|S_n|}$ converges weakly to $\mu_{|s_d|}$, in probability. From the preceding we deduce the existence of $g \in C_b(\mathbb{R}^+)$ supported on $[0, 1/2]$, such that $\langle g, \mu_{|s_d|} \rangle \geq 1$ and that for such g,

$$\mathbb{P}(r_n > 1/2) \leq \mathbb{P}(\langle g, \nu_{|S_n|} \rangle = 0) \leq \mathbb{P}\Big(|\langle g, \nu_{|S_n|} \rangle - \langle g, \mu_{|s_d|} \rangle| > 1/2\Big) \to 0, \tag{5.9}$$

as $n \to \infty$. Similarly, considering $g \in C_b(\mathbb{R}^+)$ supported on $[1, 3/2]$ for which $\langle g, \mu_{|s_d|} \rangle \geq 1$, we get that $\mathbb{P}(r_1 < 1) \to 0$, from which we conclude that $\mathbb{P}(\mathcal{G}_n^c) \to 0$. □

6 Proof of Proposition 1.3

The main task here is to show that for m-a.e.$v \in \mathbb{C}$, the logarithm is uniformly integrable with respect to the ESD of $|U_n T_n - vI_n|$. As shown in [7], setting $\rho = |v|$, this is equivalent to such uniform integrability for the ESD ν_n^v of the matrix $\boldsymbol{Y}_n^v$ (per (1.12)).

The key for the latter is to show that $\Im(G_n(\cdot))$ is uniformly bounded on $\{i\eta : \eta > n^{-\kappa_1}\}$ for some $\kappa_1 > 0$ and Lebesgue almost every ρ (see proof of [7, Propn. 14 (i)]). In [7], this was done under the assumption of [7, Eqn. (3)], whereas here we show that the same holds under the weaker condition (1.8).
To this end, [7, Lem. 10] yields (analogously to Lemma 3.1), the weak convergence, in probability, of ν_n^v to ν^v, as well as the identities and bounds [7, Eqn. (34)–(38)], without ever using [7, Eqn. (2) or Eqn. (3)]. The same applies for [7, Lem. 11 and Lem. 12] which validate the Schwinger-Dyson equation [7, Eqn. (38)] for all n large enough, any $\Im(z) > C_1 n^{-1/4}$ and $\rho \in (0, R]$. We then use Lemma 1.5(i) to bypass [7, Lem. 13]. Specifically, from (1.8), using Lemma 1.5(i) we have that for every $\varepsilon \in (0, 1/2)$ and finite R, there exist finite M_1 and $\kappa_1 > 0$ depending only on R and ε such that for every $\rho \in [R^{-1}, R]$,

$$\{z : \Im(z) > n^{-\kappa_1}, |\Im(G_n(z))| > M_1\} \subset \Gamma_{2\varepsilon}^{\rho} \tag{6.1}$$

where Γ_γ^ρ denotes the union of open balls of radius $\gamma > 0$ centered at points from the symmetric subset $K \pm \rho$ of $\mathbb{R}$. Having (6.1) instead of the bound (4.8) of Lemma 4.3, we consider here the closed set $\Lambda_K := \{v \in \mathbb{C} : |v| \in K\}$ such that $m(\Lambda_K) = 0$, the bounded, open sets Γ_γ, $\gamma > 0$, associated with Λ_K via (3.3), and the corresponding collection $\mathcal{S} \subset C_c^\infty(\mathbb{C})$ of test functions. Using this framework and following the proof of Lemma 3.2, we deduce that

$$\langle \mathrm{Log}, \nu_n^v \rangle \to \langle \mathrm{Log}, \nu^v \rangle,$$

in probability for each $v \in \Gamma_\gamma$, and consequently for m-a.e. $v \in \mathbb{C}$. Then, utilizing our assumption (1.7) on the uniformly bounded support of the relevant ESD-s, we have that further, for any fixed $\phi \in \mathcal{S}$,

$$\int_{\mathbb{C}} \phi(v) \langle \mathrm{Log}, \nu_n^v \rangle dm(v) \to \int_{\mathbb{C}} \phi(v) \langle \mathrm{Log}, \nu^v \rangle dm(v),$$

in probability. Since $\Theta(\{0\}) = 0$ and Θ is not a Dirac measure, we know from [9, Thm. 4.4 and Cor. 4.5] and [7, Rmk. 8] that μ_A has a density with respect to the Lebesgue measure on $\mathbb{C}$. Consequently $\mu_A(\Lambda_K) = 0$, and following the same argument as in the proof of Proposition 1.4, we get part (a) of Proposition 1.3.
For parts (b) and (c) of the proposition see [7, Rmk 8] (which does not involve [7, Eqn. (2) or Eqn. (3)]). For part (d) recall that Lemma 1.5(i) applies even in case U_n is replaced by a Haar distributed orthogonal matrix O_n, as does the relevant analysis from [7] (c.f. proof of [7, Thm. 18]). Hence, following the same argument as in the unitary case, the proof is complete once we establish the analog of Lemma 5.2. That is, specify events $\mathcal{G}_n$ determined by T_n, such that $\mathbb{P}(\mathcal{G}_n^c) \to 0$ as $n \to \infty$ and

$$\sup_{v \in \Gamma_\gamma} \mathbb{P}\Big(\mathcal{G}_n \cap \big\{|v| s_{\min}(O_n + v^{-1} T_n) \le t\big\}\Big) \le C' t^{c_1} n^{c_2}, \tag{6.2}$$

for any $\gamma > 0$, some $C' = C'(\gamma)$ finite and all t, n. To this end, with Θ non-degenerate, there exist $\xi > 0$ and $b_+^2 \ge b_-^2 + \xi$, such that both $\Theta([0, b_-))$ and $\Theta((b_+, M])$ are positive. Consequently, setting $T_n = \mathrm{diag}(r_1, \ldots, r_n)$ with $r_1 \ge r_2 \ge \ldots \ge r_n$, it follows from the weak convergence of L_{T_n} to Θ (in probability), that $\mathbb{P}(\mathcal{G}_n^c) \to 0$ for $\mathcal{G}_n := \{r_n \le b_-$ and $r_1 \in [b_+, M]\}$ (by the same reasoning as in the derivation of (5.9)). Further, (6.2) follows by an application of [15, Thm. 1.3] conditional upon T_n (where [15, Eqn. (1.2)] holds under $\mathcal{G}_n$ for $v^{-1} T_n$, $v \in \Gamma_\gamma$, $K = M/\gamma$ and $\delta = \xi\gamma^2$, see (5.6)-(5.7)).

Acknowledgment We thank Ofer Zeitouni for suggesting to look at the sum of d unitary matrices and for pointing our attention to preliminary versions of [3, 15]. We also thank the anonymous referees whose valuable comments on an earlier draft helped improving the presentation of this work.

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. MR-2760897

[2] P. Biane, and F. Lehner. Computation of some examples of Brown's spectral measure in free probability. *Coll. Math.*, 90, 181-211, 2001. MR-1876844

[3] C. Bordenave, and D. Chafaï. Around the circular law. *Probability surveys*, 9, 1-89, 2012. MR-2908617

[4] L. G. Brown. Lidskii's theorem in the type II case. *Geometric methods in operator algebras* (Kyoto,1983), Longman Sci. Tech., Harlow, 1986, pp. 135. MR-0866489

[5] P. J. Forrester. *Log-gases and random matrices* (LMS-34), Princeton Univ. Press, Princeton, 2010. MR-2641363

[6] V. L. Girko. The circular law. *Teor. Veroyatnost. i Primenen*, 29, 669-679, 1984. MR-0773436

[7] A. Guionnet, M. Krishnapur, and O. Zeitouni. The single ring theorem. *Ann. Math.*, 174, 1189-1217, 2011. MR-2831116

[8] A. Guionnet, P. Wood, and O. Zeitouni. Convergence of spectral measure of non normal matrices. *Proc. AMS*, to appear, 2012. MR-3000558

[9] 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. MR-1784419

[10] F. Hiai, and D. Petz. Asymptotic freeness almost everywhere for random matrices. *Acta Sci. Math. (Szeged)*, 66(3-4), 809-834, 2000. MR-1804226

[11] C. Male. The distribution of traffics and their free product: an asymptotic freeness theorem for random matrices and a central limit theorem. *Preprint*, arXiv:1111.4662v3.

[12] E. S. Meckes, and M. W. Meckes. Concentration and convergence rates for spectral measures of random matrices. *Probab. Theory Related Fields*, to appear, 2012. MR-3055255

[13] M. L. Mehta. *Random matrices*, Pure and Appl. Math. 142, Elsevier, Academic Press, Amsterdam, 2004. MR-2129906

[14] A. Nica, and R. Speicher. R-diagonal pairs–a common approach to Haar unitaries and circular elements. *in Free Probability Theory* (Waterloo, ON, 1995), Fields Inst. Commun. 12, Amer. Math. Soc., Providence, RI, 1997, pp. 149-188. MR-1426839

[15] M. Rudelson, and R. Vershynin. Invertibility of random matrices: unitary and orthogonal transformation. *Journal of the AMS*, to appear.

[16] P. Śniady. Random regularization of Brown spectral measure. *J. Funct. Anal.*, 193, 291-313, 2002. MR-1929504

[17] 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 M. Krishnapur). MR-2722794