LIMITING SPECTRAL DISTRIBUTIONS OF SUMS OF PRODUCTS OF NON-HERMITIAN RANDOM MATRICES*

BY

HOLGER KÖSTERS (BIELEFELD) AND ALEXANDER TIKHOMIROV (SYKTYVKAR)

Abstract. For fixed $l \geqslant 0$ and $m \geqslant 1$, let $\mathbf{X}_n^{(0)}, \mathbf{X}_n^{(1)}, \ldots, \mathbf{X}_n^{(l)}$ be independent random $n \times n$ matrices with independent entries, let $\mathbf{F}_n^{(0)} := \mathbf{X}_n^{(0)}(\mathbf{X}_n^{(1)})^{-1} \ldots (\mathbf{X}_n^{(l)})^{-1}$, and let $\mathbf{F}_n^{(1)}, \ldots, \mathbf{F}_n^{(m)}$ be independent random matrices of the same form as $\mathbf{F}_n^{(0)}$. We show that as $n \to \infty$, the matrices $\mathbf{F}_n^{(0)}$ and $m^{-(l+1)/2}(\mathbf{F}_n^{(1)} + \ldots + \mathbf{F}_n^{(m)})$ have the same limiting eigenvalue distribution.

To obtain our results, we apply the general framework recently introduced in Götze, Kösters, and Tikhomirov (2015) to sums of products of independent random matrices and their inverses. We establish the universality of the limiting singular value and eigenvalue distributions, and we provide a closer description of the limiting distributions in terms of free probability theory.

2010 AMS Mathematics Subject Classification: Primary: 60B20; Secondary: 60E07, 60F05, 46L54.

Key words and phrases: Non-Hermitian random matrices, limiting spectral distributions, free probability theory, stable distributions.

1. INTRODUCTION AND MAIN RESULTS

The investigation of the asymptotic spectral distributions of random matrices is a major topic in random matrix theory. In recent years, sums and products of independent *non-Hermitian* random matrices with independent entries have found increasing attention; see e.g. [1], [3], [11], [13]–[18], [20], [25], [27], [29], [31], [35]–[37] for results on *global* spectral distributions, and also the survey paper [2] and the references therein for results on *local* spectral distributions. In particular, the paper [18] provides a general framework for the investigation of the limiting (global) spectral distributions of products of independent random matrices with independent entries. Furthermore, the paper [37] shows that this approach proves useful for the investigation of sums of products as well. The aim of the present

---

paper is to show that certain products of independent random matrices give rise to random matrices with *stable* limiting eigenvalue distributions, in the sense that the sums of several independent copies of these products have the same limiting eigenvalue distribution after appropriate rescaling.

Throughout this paper, for each $n \geqslant 1$, let $\mathbf{X}_n^{(1)}, \mathbf{X}_n^{(2)}, \mathbf{X}_n^{(3)}, \ldots$ be independent random matrices of size $n \times n$ with independent entries. More precisely, we assume that

$$(1.1) \qquad \mathbf{X}_n^{(q)} = \left(\tfrac{1}{\sqrt{n}} X_{jk}^{(q)} \right)_{j,k=1,\ldots,n},$$

where $(X_{jk}^{(q)})_{j,k,q \in \mathbb{N}}$ is a family of independent real or complex random variables such that

$$(1.2) \qquad \mathbb{E} X_{jk}^{(q)} = 0,\ \mathbb{E}(X_{jk}^{(q)})^2 = 1 \quad \text{in the real case,}$$

and

$$(1.3) \qquad \mathbb{E} X_{jk}^{(q)} = 0,\ \mathbb{E}(X_{jk}^{(q)})^2 = 0,\ \mathbb{E}|X_{jk}^{(q)}|^2 = 1 \quad \text{in the complex case,}$$

and we additionally assume that this family is uniformly square-integrable, i.e.

$$(1.4) \qquad \lim_{a \to \infty} \sup_{j,k,q \in \mathbb{N}} \mathbb{E}\big(|X_{jk}^{(q)}|^2 \, \mathbf{1}_{\{|X_{jk}^{(q)}| \geqslant a\}} \big) = 0.$$

In this case we also say the matrices $\mathbf{X}_n^{(q)}$ are independent *Girko–Ginibre matrices*. In the special case where the entries have real or complex Gaussian distributions, we usually write $\mathbf{Y}_n^{(q)} = \left(\tfrac{1}{\sqrt{n}} Y_{jk}^{(q)} \right)_{j,k=1,\ldots,n}$ instead of $\mathbf{X}_n^{(q)} = \left(\tfrac{1}{\sqrt{n}} X_{jk}^{(q)} \right)_{j,k=1,\ldots,n}$ and call the matrices $\mathbf{Y}_n^{(q)}$ *Gaussian random matrices* or *Ginibre matrices*. Note that the assumption (1.4) is clearly satisfied in this special case, the random variables $Y_{jk}^{(q)}$ being independent and identically distributed (i.i.d.).

We will be interested in the limiting spectral distributions of random matrices $\mathbf{F}_n$ given by sums of products of the matrices $\mathbf{X}_n^{(q)}$ and their inverses. Let $\mathbf{F}_n$ have the singular values $s_1(\mathbf{F}_n) \geqslant \ldots \geqslant s_n(\mathbf{F}_n)$ and eigenvalues $\lambda_1(\mathbf{F}_n), \ldots, \lambda_n(\mathbf{F}_n)$. Then we write $\nu_n := \nu(\mathbf{F}_n) := \tfrac{1}{n} \sum_{j=1}^{n} \delta_{s_j(\mathbf{F}_n)}$ for the (empirical) *singular value distribution* of $\mathbf{F}_n$ and $\mu_n := \mu(\mathbf{F}_n) := \tfrac{1}{n} \sum_{j=1}^{n} \delta_{\lambda_j(\mathbf{F}_n)}$ for the (empirical) *eigenvalue distribution* of $\mathbf{F}_n$. The corresponding weak limits in probability (if existent) will be denoted by $\nu := \nu_{\mathbf{F}}$ and $\mu := \mu_{\mathbf{F}}$, respectively. Note that μ will in general be a probability measure on the complex plane, the random matrices $\mathbf{F}_n$ being non-Hermitian. The density of μ (if existent) will be denoted by $f(z)$, or by $f(r)$ (with $r = |z|$) in case it is rotation-invariant with respect to the origin.

Let us mention some relevant results from the literature.

EXAMPLES 1.1.

(a) (*Circular law*) Let $\mathbf{F}_n = \mathbf{X}_n^{(1)}$. Then $f(r) = \tfrac{1}{\pi} \mathbf{1}_{[0,1]}(r)$, i.e. μ is the uniform distribution on the unit disk.

(b) Let $\mathbf{F}_n = \mathbf{X}_n^{(1)} + \ldots + \mathbf{X}_n^{(m)}$. Then $\mathbf{F}_n$ is a random matrix with independent entries of mean zero and variance m/n, so, by simple rescaling, $f(r) = \frac{1}{m\pi} \mathbf{1}_{[0,\sqrt{m}]}(r)$. In particular, for the rescaled matrices $\frac{1}{\sqrt{m}}\mathbf{F}_n$, the limiting eigenvalue distribution is again the uniform distribution on the unit disk.

(c) Let $\mathbf{F}_n = \mathbf{X}_n^{(1)}\mathbf{X}_n^{(2)}$. Then $f(r) = \frac{1}{2\pi r} \mathbf{1}_{[0,1]}(r)$, i.e. μ is the induced distribution of the uniform distribution on the unit disk under the mapping $z \mapsto z^2$. See e.g. [18], Section 8.2.2, for a "simple" derivation.

(d) Let $\mathbf{F}_n = \mathbf{X}_n^{(1)}\mathbf{X}_n^{(2)} + \ldots + \mathbf{X}_n^{(2m-1)}\mathbf{X}_n^{(2m)}$. Then

$$f(r) = \frac{1}{\pi\sqrt{(m-1)^2 + 4r^2}} \mathbf{1}_{[0,\sqrt{m}]}(r);$$

see [37], Section 2.

(e) (*Spherical law*) Let $\mathbf{F}_n = \mathbf{X}_n^{(1)}(\mathbf{X}_n^{(2)})^{-1}$. Then

$$f(r) = \frac{1}{\pi(1 + r^2)^2},$$

i.e. μ is the spherical distribution on the complex plane.

(f) Let $\mathbf{F}_n = \mathbf{X}_n^{(1)}(\mathbf{X}_n^{(2)})^{-1} + \ldots + \mathbf{X}_n^{(2m-1)}(\mathbf{X}_n^{(2m)})^{-1}$. Then

$$f(r) = \frac{m^2}{\pi(m^2 + r^2)^2};$$

see [37], Section 3. Thus, for the rescaled matrices $\frac{1}{m}\mathbf{F}_n$, the limiting eigenvalue distribution is again the spherical distribution on the complex plane.

In view of examples (b) and (f), it seems natural to ask whether there exist further examples of random matrices $\mathbf{F}_n^{(0)}$ such that for any $m \in \mathbb{N}$, the sums of m independent matrices of the same form as $\mathbf{F}_n^{(0)}$ have the same limiting eigenvalue distribution as the original random matrices $\mathbf{F}_n^{(0)}$, after appropriate rescaling. We will answer this question in the affirmative by proving the following result, which contains examples (b) and (f) as special cases:

THEOREM 1.1. *Fix $m \in \mathbb{N}$ and $l \in \mathbb{N}_0$, let*

$$(1.5) \qquad \mathbf{F}_n^{(0)} := (\mathbf{X}_n^{(0)})(\mathbf{X}_n^{(1)})^{-1} \ldots (\mathbf{X}_n^{(l)})^{-1},$$

where $\mathbf{X}_n^{(0)}, \mathbf{X}_n^{(1)}, \ldots, \mathbf{X}_n^{(l)}$ are independent random matrices as in (1.1)–(1.4), and let $\mathbf{F}_n^{(1)}, \ldots, \mathbf{F}_n^{(m)}$ be independent matrices of the same form as $\mathbf{F}_n^{(0)}$. Then the matrices $m^{-(l+1)/2}(\mathbf{F}_n^{(1)} + \ldots + \mathbf{F}_n^{(m)})$ and $\mathbf{F}_n^{(0)}$ have the same limiting eigenvalue distribution μ. More precisely, we have $\mu = \mathcal{H}(\sigma_s(\frac{2}{l+1}))$, where $\sigma_s(\frac{2}{l+1})$ is the symmetric $\boxplus$-stable distribution with parameter $\frac{2}{l+1}$ (see Section 2.3) and $\mathcal{H}(\sigma_s(\frac{2}{l+1}))$ is the associated rotation-invariant distribution on $\mathbb{C}$ (see Section 2.1).

Moreover, as we will see in Section 3, apart from a possible permutation of the exponents ± 1, the matrices $\mathbf{F}_n^{(0)}$ in Theorem 1.1 are the only examples of products of *independent* Girko–Ginibre matrices and their inverses such that for any $m \in \mathbb{N}$, $\mathbf{F}_n^{(0)}$ and $\mathbf{F}_n^{(1)} + \ldots + \mathbf{F}_n^{(m)}$ have the same limiting eigenvalue distribution after appropriate rescaling. In particular, the matrices

$$(1.6) \qquad \mathbf{F}_n^{(0)} := \mathbf{X}_n^{(1)} \ldots \mathbf{X}_n^{(k)} (\mathbf{X}_n^{(k+1)})^{-1} \ldots (\mathbf{X}_n^{(k+l)})^{-1}$$

with $k > 1$ do not share this property.

However, the same limiting eigenvalue distributions may arise for products involving *powers* of random matrices:

THEOREM 1.2. *Fix* $m \in \mathbb{N}$, $k \in \mathbb{N}_0$ *and* $l_1, \ldots, l_k \in \mathbb{N}$, *let* $l := l_1 + \ldots + l_k$ *and define*

$$(1.7) \qquad \mathbf{F}_n^{(0)} := (\mathbf{X}_n^{(0)})(\mathbf{X}_n^{(1)})^{-l_1} \ldots (\mathbf{X}_n^{(k)})^{-l_k},$$

where $\mathbf{X}_n^{(0)}, \mathbf{X}_n^{(1)}, \ldots, \mathbf{X}_n^{(k)}$ *are independent random matrices as in* (1.1)–(1.4), *and let* $\mathbf{F}_n^{(1)}, \ldots, \mathbf{F}_n^{(m)}$ *be independent matrices of the same form as* $\mathbf{F}_n^{(0)}$. *Then the matrices* $m^{-(l+1)/2}(\mathbf{F}_n^{(1)} + \ldots + \mathbf{F}_n^{(m)})$ *and* $\mathbf{F}_n^{(0)}$ *have the same limiting eigenvalue distribution* μ, *which is the same as in Theorem* 1.1.

Theorem 1.1 will be deduced from a more general result about random matrices of the form

$$(1.8) \qquad \mathbf{F}_n(\mathbf{X}) := \sum_{q=1}^{m} \mathbf{F}_n^{(q)}(\mathbf{X}) := \sum_{q=1}^{m} \prod_{r=1}^{l} (\mathbf{X}_n^{((q-1)l+r)})^{\varepsilon_r},$$

where $m, l \in \mathbb{N}$ and $\varepsilon_1, \ldots, \varepsilon_l \in \{+1, -1\}$ are fixed. (Thus, the matrices $\mathbf{F}_n^{(q)}(\mathbf{X})$ are independent random matrices of the same form as the matrix $\prod_{r=1}^{l} (\mathbf{X}_n^{(r)})^{\varepsilon_r}$.) Let us note that under the assumptions (1.1)–(1.4), each matrix $\mathbf{X}_n^{(r)}$ is invertible with probability $1 + o(1)$ as $n \to \infty$ (see e.g. Lemma 4.9), so that $\mathbf{F}_n(\mathbf{X})$ is defined with probability $1 + o(1)$ as $n \to \infty$. Here we have the following result, which establishes the existence of the limiting singular value and eigenvalue distributions and provides a closer description of them in terms of free probability theory:

THEOREM 1.3. *Let the matrices* $\mathbf{F}_n(\mathbf{X})$ *be defined as in* (1.8). *Then there exist non-random probability measures* ν *and* μ *on* $(0, \infty)$ *and* $\mathbb{C}$, *respectively, such that* $\lim_{n\to\infty} \nu\big(\mathbf{F}_n(\mathbf{X})\big) = \nu$ *and* $\lim_{n\to\infty} \mu\big(\mathbf{F}_n(\mathbf{X})\big) = \mu$ *weakly in probability, and the limiting distributions are the same as those for the matrices* $\mathbf{F}_n(\mathbf{Y})$ *derived from Gaussian random matrices. More precisely, the limiting measures* ν *and* μ *are given by*

$$(1.9) \qquad \mathcal{S}\nu = \big(\mathcal{Q}^{-1}(\gamma^{\varepsilon_1} \boxtimes \ldots \boxtimes \gamma^{\varepsilon_l})\big)^{\boxplus m} \quad and \quad \mu = \mathcal{H}(\mathcal{S}\nu),$$

where γ is the Marchenko–Pastur distribution, γ^{-1} is the inverse Marchenko–Pastur distribution, $\boxplus$ and $\boxtimes$ denote the additive and multiplicative free convolution, and $\mathcal{S}$, $\mathcal{Q}$ and $\mathcal{H}$ are the operators described in Section 2.1 *below.*

In particular, this result shows that the limiting spectral distributions are *universal*, i.e. they do not depend on the distributions of the matrix entries apart from a few moment conditions as in (1.2)–(1.4).

REMARK 1.1. *As will follow from the proof, Theorem* 1.3 *extends to* certain *sums of products of powers of independent Girko–Ginibre matrices and their inverses*, *namely to random matrices of the form*

$$(1.10) \qquad \mathbf{F}_n(\mathbf{X}) := \sum_{q=1}^{m} \mathbf{F}_n^{(q)}(\mathbf{X}) := \sum_{q=1}^{m} \prod_{r=1}^{k} \big((\mathbf{X}_n^{((q-1)k+r)})^{\varepsilon_r} \big)^{l_r},$$

where $m, k \in \mathbb{N}$, $\varepsilon_1, \ldots, \varepsilon_k \in \{-1, +1\}$ and $l_1, \ldots, l_k \in \mathbb{N}$ are fixed, and

$$(1.11) \qquad\qquad\qquad \textit{for some } r = 1, \ldots, k, \textit{ we have } l_r = 1.$$

Here, with the notation as above, the limiting measures ν and μ are given by

$$(1.12) \quad \mathcal{S}\nu = \Big(\mathcal{Q}^{-1}\big((\gamma^{\varepsilon_1})^{\boxtimes l_1} \boxtimes \ldots \boxtimes (\gamma^{\varepsilon_k})^{\boxtimes l_k} \big) \Big)^{\boxplus m} \quad \textit{and} \quad \mu = \mathcal{H}(\mathcal{S}\nu).$$

This will be important for the proof of Theorem 1.2.

To obtain the preceding results, we apply the general framework from [18] for the investigation of (global) limiting spectral distributions to sums of products of independent Girko–Ginibre random matrices and their inverses (see Section 4). Related results for various special cases can be found e.g. in [1], [3], [11], [13], [14], [16], [20], [25], [27], [29], [35], [37]. In particular, in the Gaussian case, the limiting eigenvalue and singular value distributions of the products (1.6) were recently obtained in [1] and [16], respectively.

To apply the framework from [18], we need to verify certain technical conditions, see Conditions A, B and C in Section 4.2 for details. This will be achieved by means of a suitable induction argument, which forms the major part of Section 4 and which represents the main contribution of this work. Furthermore, to identify the limiting spectral distributions, we use tools from free probability theory. Here it is worth emphasizing that for the matrices in Theorems 1.1 and 1.2 the limiting spectral distributions may be described relatively explicitly. It seems that comparable results are available only in a few special cases, see e. g. [10], [21], [22], [26]. Let us mention, however, the very recent work [6], [33], [7] which provides an algorithm for calculating the Brown measures of general polynomials in free non-commutative random variables. This should yield many further examples where the limiting spectral distributions may now be determined.

2. BACKGROUND

In this section we recall some well-known concepts and results from the literature which will be needed later.

2.1. Results from random matrix theory. The derivation of our results on limiting eigenvalue distributions will be based on Girko's *Hermitization method* (see also [12]). Thus, we will first study the limiting eigenvalue distributions of the *Hermitian* matrices

$$(2.1) \qquad \mathbf{V}_n := \begin{bmatrix} \mathbf{O} & \mathbf{F}_n \\ \mathbf{F}_n^* & \mathbf{O} \end{bmatrix} \quad \text{and} \quad \mathbf{W}_n := \mathbf{F}_n \mathbf{F}_n^*.$$

Note that if the singular values of $\mathbf{F}_n$ are given by $s_1, \ldots, s_n$, then the eigenvalues of $\mathbf{V}_n$ and $\mathbf{W}_n$ are given by $\pm s_1, \ldots, \pm s_n$ and $s_1^2, \ldots, s_n^2$, respectively. It is easy to see that knowledge of one of the distributions $\nu(\mathbf{F}_n), \mu(\mathbf{V}_n), \mu(\mathbf{W}_n)$ (or its convergence) implies knowledge of the other two (or their convergence). More precisely, if $\mathcal{S}$ denotes the operator which associates with each distribution ν on $(0, \infty)$ its symmetrization on $\mathbb{R}^*$, and $\mathcal{Q}$ denotes the operator which associates with each symmetric distribution μ on $\mathbb{R}^*$ its induced distribution on $(0, \infty)$ under the mapping $x \mapsto x^2$, the operators $\mathcal{S}$ and $\mathcal{Q}$ are one-to-one, and we have

$$(2.2) \qquad \mu(\mathbf{V}_n) = \mathcal{S}\nu(\mathbf{F}_n) \quad \text{and} \quad \mu(\mathbf{W}_n) = \mathcal{Q}\mu(\mathbf{V}_n).$$

Furthermore, given a symmetric distribution $\mu_{\mathbf{V}}$ on $\mathbb{R}^*$ such that

$$(2.3) \qquad \int \log^+ |t| \, d\mu_{\mathbf{V}}(t) < \infty,$$

we write $\mathcal{H}\mu_{\mathbf{V}}$ for the rotation-invariant distribution on $\mathbb{C}$ (if existent) such that

$$(2.4) \quad U_{\mathbf{V}}(\alpha) := -\int \log |z - \alpha| \, d(\mathcal{H}\mu_{\mathbf{V}})(z) = -\int \log |x| \, d\big(\mu_{\mathbf{V}} \boxplus B(\alpha)\big)(x)$$

for any $\alpha \in \mathbb{C}$. Here, the function $U_{\mathbf{V}}(\alpha)$ is the so-called *logarithmic potential* of the measure $\mathcal{H}\mu_{\mathbf{V}}$, $B(\alpha) := \frac{1}{2}\delta_{-|\alpha|} + \frac{1}{2}\delta_{+|\alpha|}$ denotes the Bernoulli distribution, and $\boxplus$ denotes free additive convolution. It follows from basic results in logarithmic potential theory that such a distribution $\mathcal{H}\mu_{\mathbf{V}}$, if it exists, is uniquely determined by (2.4), see e.g. [32], and also the comments at the end of Section 2.2.

Girko's *Hermitization method* (see also [12]) now states that under appropriate assumptions, the weak convergence of the eigenvalue distributions $\mu(\mathbf{F}_n)$ follows from the weak convergence of the singular value distributions $\nu(\mathbf{F}_n - \alpha \mathbf{I}_n)$ of the *shifted* matrices $\mathbf{F}_n - \alpha \mathbf{I}_n$ for all $\alpha \in \mathbb{C}$. We will only need the following special case:

THEOREM 2.1 ([18], Theorem 7.6). *If the random matrices $\mathbf{F}_n$ satisfy Condition C in Section 4 below and there exists a non-random probability measure $\nu_{\mathbf{F}}$ on $(0, \infty)$ such that for all $\alpha \in \mathbb{C}$, $\mathcal{S}\nu(\mathbf{F}_n - \alpha \mathbf{I}_n) \to (\mathcal{S}\nu_{\mathbf{F}}) \boxplus B(\alpha)$ weakly in*

probability, then $\mu(\mathbf{F}_n) \to \mu_{\mathbf{F}} := \mathcal{H}(\mathcal{S}\nu_{\mathbf{F}})$ *weakly in probability. Moreover, with the notation from* [18] *and under regularity conditions, the measure* $\mu_{\mathbf{F}}$ *has the Lebesgue density*

$$(2.5) \qquad f(u,v) = \frac{1}{2\pi|\alpha|^2}\left(u\,\frac{\partial\psi}{\partial u} + v\,\frac{\partial\psi}{\partial v}\right),$$

where $\alpha = u + iv$ *and* ψ *is a continuous function on* $\mathbb{C}^*$ *taking values in* $[0,1]$ *and satisfying*

$$(2.6) \qquad \psi(\alpha)\big(1 - \psi(\alpha)\big) = -|\alpha|^2\big(1 - \psi(\alpha)\big)^2\Big(S_{\mathbf{V}}\Big(-\big(1 - \psi(\alpha)\big)\Big)\Big)^2.$$

Here, $S_{\mathbf{V}}$ *denotes the* S*-transform of the symmetric probability measure* $\mu_{\mathbf{V}} = \mathcal{S}\nu_{\mathbf{F}}$ (see [30], [4], [18]).

Furthermore, as the starting point for the proof of Theorem 1.3 (which will be by induction on the number of factors and summands in (1.8)), we will rely upon the well-known *Marchenko–Pastur theorem*, which states that when $\mathbf{F}_n = \mathbf{X}_n^{(1)}$, $n \in \mathbb{N}$, then $\mu(\mathbf{W}_n) \to \gamma$ weakly in probability, where

$$(2.7) \qquad \gamma(dx) = \frac{1}{2\pi}\sqrt{\frac{4-x}{x}}\,\mathbf{1}_{(0,4)}(x)\,\lambda(dx)$$

is the *Marchenko–Pastur distribution* (with parameter one). Therefore, when $\mathbf{F}_n = (\mathbf{X}_n^{(1)})^{-1}$, $n \in \mathbb{N}$, we have $\mu(\mathbf{W}_n) \to \gamma^{-1}$ weakly in probability, where γ^{-1} is the induced measure of γ under the mapping $x \mapsto x^{-1}$. We will call this measure the *inverse Marchenko–Pastur distribution*. Finally, let us note that the S-transforms of γ and γ^{-1} are given by

$$(2.8) \qquad S_{\gamma}(z) = \frac{1}{z+1} \quad \text{and} \quad S_{\gamma^{-1}}(z) = -z,$$

respectively, see e.g. Section 8.1.1 in [18].

2.2. Results from free probability theory. To describe the limiting singular value distributions of the random matrices $\mathbf{F}_n$ in Theorem 1.3, we will use various concepts and results from free probability theory. See e.g. [38], [28] for a thorough introduction to free probability theory, or Section 5 in [18] for a brief introduction tailored to our purposes. In particular, we will use the free additive and multiplicative convolutions $\boxplus$ and $\boxtimes$, the associated R and S transforms (also for probability measures with *unbounded* support), and the asymptotic freeness of random matrices. Furthermore, we will frequently use the following result:

PROPOSITION 2.1 (Asymptotic freeness). *For each* $n \in \mathbb{N}$, *let* $\mathbf{A}_n$ *and* $\mathbf{B}_n$ *be independent bi-unitary invariant random matrices of size* $n \times n$ *such that*

$$\sup_{n\in\mathbb{N}} \max\left\{\mathbb{E}\big(\tfrac{1}{n}\operatorname{trace}(\mathbf{A}_n\mathbf{A}_n^*)^k\big), \mathbb{E}\big(\tfrac{1}{n}\operatorname{trace}(\mathbf{B}_n\mathbf{B}_n^*)^k\big)\right\} < \infty$$

for all $k \in \mathbb{N}$, and suppose that there exist compactly supported (deterministic)
probability measures $\mu_{\mathbf{AA}^*}$ and $\mu_{\mathbf{BB}^*}$ on $(0, \infty)$ such that $\mu(\mathbf{A}_n\mathbf{A}_n^*) \to \mu_{\mathbf{AA}^*}$
and $\mu(\mathbf{B}_n\mathbf{B}_n^*) \to \mu_{\mathbf{BB}^*}$ weakly in probability. Then one has the following:
   (a) *The families $\{\mathbf{A}_n, \mathbf{A}_n^*\}$ and $\{\mathbf{B}_n, \mathbf{B}_n^*\}$ are asymptotically free, and*

$$(\mathbf{A}_n\mathbf{B}_n)(\mathbf{A}_n\mathbf{B}_n)^* \to \mu_{\mathbf{AA}^*} \boxtimes \mu_{\mathbf{BB}^*} \text{ in moments.}$$

   (b) *For any $k, l \in \mathbb{N}$, the matrices $(\mathbf{A}_n^k)^*\mathbf{A}_n^k$ and $\mathbf{A}_n^l(\mathbf{A}_n^l)^*$ are asymptoti-*
cally free, and for any $k \in \mathbb{N}$,

$$\mathbf{A}_n^k(\mathbf{A}_n^k)^* \to \mu_{\mathbf{AA}^*}^{\boxtimes k} \text{ in moments.}$$

   (c) *The matrices $\mathbf{V}_n(\mathbf{A}_n)$ and $\mathbf{V}_n(\mathbf{B}_n)$ are asymptotically free, and*

$$\mathbf{V}_n(\mathbf{A}_n) + \mathbf{V}_n(\mathbf{B}_n) \to \mu_{\mathbf{V(A)}} \boxplus \mu_{\mathbf{V(B)}} \text{ in moments.}$$

   (d) *The matrices $\mathbf{V}_n(\mathbf{A}_n)$ and $\mathbf{J}_n(\alpha)$ are asymptotically free, and*

$$\mathbf{V}_n(\mathbf{A}_n) + \mathbf{J}_n(\alpha) \to \mu_{\mathbf{V(A)}} \boxplus B(\alpha) \text{ in moments.}$$

   Here, $\mathbf{V}(\mathbf{A}_n)$ and $\mathbf{V}(\mathbf{B}_n)$ are defined as in equation (2.1), $\mu_{\mathbf{V(A)}}$ and $\mu_{\mathbf{V(B)}}$
denote the corresponding limiting distributions, and

$$(2.9) \qquad \mathbf{J}_n(\alpha) := \begin{bmatrix} \mathbf{O} & -\alpha\mathbf{I}_n \\ -\overline{\alpha}\mathbf{I}_n & \mathbf{O} \end{bmatrix}.$$

Parts (a) and (b) follow from the results in Section 4.3 in [23], part (d) is proved
in Section 5 in [18], and part (c) follows by similar arguments. Also, let us mention
that part (c) is already implicit in [37].

REMARK 2.1. *Observe that Proposition* 2.1 *may be used to establish the weak*
convergence of the mean *singular value distributions of the matrices $\mathbf{A}_n\mathbf{B}_n$, $\mathbf{A}_n^k$*
and $\mathbf{A}_n + \mathbf{B}_n$. However, in most of the situations in which we will use Propo-
sition 2.1 *later, this already implies the weak convergence in probability of the*
singular value distributions of these matrices (see e.g. Section A.1 in [18]).

It is worth mentioning that there is another description of the limiting density
$f(u, v)$ in Theorem 2.1 due to Haagerup and Larsen [21] and Haagerup and Schultz
[22]. Actually, in these papers, the density f is shown to describe the *Brown mea-*
sure of a so-called R-*diagonal element* in a W^*-probability space. Roughly speak-
ing, an R-diagonal element is a non-commutative random variable of the form uh,
where u is Haar unitary and h is a positive element $*$-free from u.

For our purposes, this description of the density f may be summarized as fol-
lows. In the situation of Theorem 2.1, let $\nu_{\mathbf{F}}$ be the limiting singular value distri-
bution of the matrices $\mathbf{F}_n$, set $\mu_{\mathbf{V}} = \mathcal{S}\nu_{\mathbf{F}}$ and $\mu_{\mathbf{W}} = \mathcal{Q}\nu_{\mathbf{V}}$ (which are the limiting

eigenvalue distributions of the matrices $\mathbf{V}_n$ and $\mathbf{W}_n$ in (2.1), respectively), and suppose that $\mu_{\mathbf{W}}$ is not a Dirac measure. Let $S_{\mathbf{W}}$ denote the S-transform of $\mu_{\mathbf{W}}$, and set

$$F(t) := \frac{1}{\sqrt{S_{\mathbf{W}}(t-1)}}.$$

Then F is a smooth bijection from the interval $(0,1)$ to the interval

$$(a,b) := \left(\left(\int x^{-2}\, d\nu_{\mathbf{F}}(x) \right)^{-1/2}, \left(\int x^2\, d\nu_{\mathbf{F}}(x) \right)^{1/2} \right)$$

(where $1/\infty := 0$ and $1/0 := \infty$), and the limiting eigenvalue distribution $\mu_{\mathbf{F}} = \mathcal{H}\mu_{\mathbf{V}}$ of the matrices $\mathbf{F}_n$ has a rotation-invariant density $f(r)$ given by

$$(2.10) \qquad\qquad f(r) = \frac{1}{2\pi r\, F'\big(F^{-1}(r)\big)} \mathbf{1}_{(a,b)}(r)$$

(see [21], Section 4, and [22], Section 4). Clearly, the connection to Theorem 2.1 arises from the fact that $\psi = F^{-1}$ on the interval (a,b). Moreover, equation (2.10) shows that $F^{-1}(r) = \int_0^r 2\pi s\, f(s)\, ds$, which implies that $\mu_{\mathbf{W}}$, and hence $\mu_{\mathbf{V}}$, is uniquely determined by $\mu_{\mathbf{F}}$. Thus, the mapping $\mu_{\mathbf{V}} \mapsto \mu_{\mathbf{F}}$ is one-to-one.

Furthermore, it follows from the results in [21], [22] that the measure $\mathcal{H}\mu_{\mathbf{V}}$ exists for any symmetric probability measure $\mu_{\mathbf{V}}$ on $\mathbb{R}^*$ satisfying (2.3) and that the operator $\mathcal{H}$ thus defined furnishes a one-to-one correspondence between the set of these distributions on $\mathbb{R}^*$ and a certain set H of rotation-invariant distributions on $\mathbb{C}$. Finally, it is easy to see that for any symmetric distribution μ on $\mathbb{R}^*$ satisfying (2.3), we have

$$(2.11) \qquad\qquad \mathcal{H}(\mathcal{D}_c\mu) = \mathcal{D}_c\mathcal{H}(\mu)$$

for all $c > 0$, where $\mathcal{D}_c$ is the scaling operator which maps a probability measure on $\mathbb{R}$ or $\mathbb{C}$ to its induced measure under the mapping $x \mapsto cx$.

2.3. Results on $\boxplus$-stable distributions. Let us collect some results on $\boxplus$-stable distributions which will be needed later. A distribution μ on $\mathbb{R}$ is called (*strictly*) $\boxplus$-*stable* if there exists a constant $\alpha > 0$ such that $\mu^{\boxplus m} = \mathcal{D}_{m^{1/\alpha}}\mu$ for all $m \in \mathbb{N}$. Here, $\mathcal{D}_c$ is defined as in equation (2.11). We will often call the constant α the *stability index* of μ.

The (strictly) $\boxplus$-stable distributions have been investigated in [9], [8] and [4]. First of all, let us recall that for any $\boxplus$-stable distribution, $\alpha \in\,]0,2]$. We will need the following result, which is contained in Appendix A of [8] and in [4]:

PROPOSITION 2.2. *Fix $\alpha \in\,]0,2]$. For a symmetric probability measure μ on $\mathbb{R}^*$, the following are equivalent*:
    (i) *μ is* (*strictly*) *$\boxplus$-stable with stability index α.*
    (ii) *$R_\mu(z) = bz^{\alpha-1}$, where $b \in \mathbb{C}^*$ with $\arg b = -\pi + \alpha\pi/2$.*

(iii) $S_\mu(z) = z^{(1/\alpha)-1}/b^{1/\alpha}$, *where* $b \in \mathbb{C}^*$ *with* $\arg b = -\pi + \alpha\pi/2$.
Moreover, in this case, the constants b in parts (ii) *and* (iii) *are the same.*

Here, for the S-transform $S_\mu(z)$, we make the convention that we take arguments in $]-\pi, +\pi]$ to define powers of b and arguments in $(-2\pi, 0)$ to define powers of z. Then, with i the imaginary unit, we have $S_\mu(z) \in (0, \infty)$i for $z \in (-1, 0)$, in line with the convention in [18].

Henceforward, we write $\sigma_s(\alpha)$ for the (unique) symmetric $\boxplus$-stable distribution with parameters $\alpha \in]0, 2]$ and $b := e^{(-\pi+\alpha\pi/2)i}$. Note that in the special cases $\alpha = 2$ and $\alpha = 1$, we obtain the standard semicircle and Cauchy distribution, respectively. Furthermore, let us recall from [8], Appendix A, that the distribution $\sigma_s(\alpha)$ has a continuous density f_α such that $f_\alpha(x) = \mathcal{O}(|x|^{-\alpha-1})$ as $|x| \to \infty$. Thus, in particular, the distribution $\sigma_s(\alpha)$ satisfies condition (2.3).

3. PROOF OF THEOREMS 1.1 AND 1.2

In this section, we prove Theorems 1.1 and 1.2 using Theorem 1.3 and Remark 1.1, respectively.

P r o o f o f T h e o r e m 1.1. By Theorem 1.3, the limiting eigenvalue distributions of the matrices $\mathbf{F}_n := \mathbf{F}_n^{(0)}$ and $\widetilde{\mathbf{F}}_n := m^{-(l+1)/2}(\mathbf{F}_n^{(1)} + \ldots + \mathbf{F}_n^{(m)})$ in Theorem 1.1 are given by

$$\mu_{\mathbf{F}} = \mathcal{H}\Big(\mathcal{Q}^{-1}\big(\gamma \boxtimes (\gamma^{-1})^{\boxtimes l}\big)\Big)$$

and

$$\mu_{\widetilde{\mathbf{F}}} = \mathcal{H}\Big(\mathcal{D}_{m^{-(l+1)/2}}\Big(\mathcal{Q}^{-1}\big(\gamma \boxtimes (\gamma^{-1})^{\boxtimes l}\big)\Big)^{\boxplus m}\Big),$$

respectively, where $\mathcal{D}_c$ is defined as in equation (2.11). To obtain the description asserted in the theorem, we calculate the S-transform of $\mathcal{Q}^{-1}\big(\gamma \boxtimes (\gamma^{-1})^{\boxtimes l}\big)$. Using (2.8) and the relation $S_{\nu_1 \boxtimes \nu_2}(z) = S_{\nu_1}(z)S_{\nu_2}(z)$, we find that

$$S_{\mathbf{W}}(z) = S_{\gamma \boxtimes (\gamma^{-1})^{\boxtimes l}}(z) = \frac{(-z)^l}{z+1},$$

and therefore

$$S_{\mathbf{V}}(z) = S_{\mathcal{Q}^{-1}(\gamma \boxtimes (\gamma^{-1})^{\boxtimes l})}(z) = \sqrt{\frac{z+1}{z} S_{\mathbf{W}}(z)} = \sqrt{\frac{z+1}{z}\frac{(-z)^l}{z+1}} = i^l z^{(l-1)/2}.$$

By Proposition 2.2, the corresponding distribution is $\mathcal{Q}^{-1}\big(\gamma \boxtimes (\gamma^{-1})^{\boxtimes l}\big) = \sigma_s\big(\frac{2}{l+1}\big)$, the symmetric $\boxplus$-stable distribution of parameter $\frac{2}{l+1}$. Thus, $\mu_{\mathbf{F}} = \mathcal{H}\big(\sigma_s\big(\frac{2}{l+1}\big)\big)$. Also, using the defining property of a $\boxplus$-stable distribution, we get

$$\mathcal{D}_{m^{-(l+1)/2}}\Big(\big(\mathcal{Q}^{-1}\big(\gamma \boxtimes (\gamma^{-1})^{\boxtimes l}\big)\big)^{\boxplus m}\Big) = \mathcal{D}_{m^{-(l+1)/2}}\Big(\big(\sigma_s\big(\tfrac{2}{l+1}\big)\big)^{\boxplus m}\Big) = \sigma_s\big(\tfrac{2}{l+1}\big).$$

Thus, $\mu_{\widetilde{\mathbf{F}}} = \mathcal{H}\big(\sigma_s\big(\frac{2}{l+1}\big)\big)$ as well, and the proof of Theorem 1.1 is complete. $\blacksquare$

Proof of Theorem 1.2. By Theorem 1.3 and Remark 1.1, the limiting eigenvalue distributions of the matrices

$$\mathbf{F}_n := \mathbf{F}_n^{(0)} \quad \text{and} \quad \widetilde{\mathbf{F}}_n := m^{-(l+1)/2}(\mathbf{F}_n^{(1)} + \ldots + \mathbf{F}_n^{(m)})$$

in Theorem 1.2 are given by

$$\mu_{\mathbf{F}} = \mathcal{H}\Big(\mathcal{Q}^{-1}\big(\gamma \boxtimes (\gamma^{-1})^{\boxtimes l_1} \boxtimes \ldots \boxtimes (\gamma^{-1})^{\boxtimes l_k}\big)\Big)$$

and

$$\mu_{\widetilde{\mathbf{F}}} = \mathcal{H}\Big(\mathcal{D}_{m^{-(l+1)/2}}\big(\mathcal{Q}^{-1}\big(\gamma \boxtimes (\gamma^{-1})^{\boxtimes l_1} \boxtimes \ldots \boxtimes (\gamma^{-1})^{\boxtimes l_k}\big)\big)^{\boxplus m}\Big),$$

respectively. But $(\gamma^{-1})^{\boxtimes l_1} \boxtimes \ldots \boxtimes (\gamma^{-1})^{\boxtimes l_k} = (\gamma^{-1})^{\boxtimes l}$, so the assertion follows in the same way as in the previous proof. ∎

REMARK 3.1. *In principle, the density of the limiting distribution $\mu_{\mathbf{F}}$ in Theorems* 1.1 *and* 1.2 *can be found by means of Theorem* 7.6 *in* [18]. *In our situation, it is easy to check that equation* (2.6) *reduces to*

$$\psi(\alpha)\big(1 - \psi(\alpha)\big) = |\alpha|^2\big(1 - \psi(\alpha)\big)^{l+1}.$$

(*Recall from Section* 2.3 *that $S_{\mathbf{V}}(z)$ takes values in $(0,\infty)$i when $z \in (-1,0)$.)* *Thus, since $\psi(\alpha)$ is continuous with values in $[0,1]$ and $\psi(\alpha) \neq 1$ for $\alpha \approx 0$ (see Sections* 6 *and* 7 *in* [18]), *we obtain, for $l = 0,1,2,3$,*

$$\psi_0(r) = 1 \wedge r^2, \quad \psi_1(r) = \frac{r^2}{1+r^2}, \quad \psi_2(r) = 1 - \frac{2}{\sqrt{1+4r^2}+1},$$

$$\psi_3(r) = 1 - \frac{3}{\big(1 + v^2(r) + w^2(r)\big)^2},$$

and therefore

$$f_0(r) = \frac{1}{\pi}\mathbf{1}_{(0,1)}(r), \quad f_1(r) = \frac{1}{\pi(1+r^2)^2},$$

$$f_2(r) = \frac{2}{\pi\sqrt{1+4r^2}(1 + 2r^2 + \sqrt{1+4r^2})},$$

$$f_3(r) = \frac{27\big(v(r) + w(r)\big)}{\pi\sqrt{4+27r^2}\big(1 + v^2(r) + w^2(r)\big)^3},$$

where we have set

$$v(r) := \big(\tfrac{1}{2}\sqrt{4+27r^2} + \tfrac{1}{2}\sqrt{27}r\big)^{1/3} \quad \text{and} \quad w(r) := \big(\tfrac{1}{2}\sqrt{4+27r^2} - \tfrac{1}{2}\sqrt{27}r\big)^{1/3}$$

for abbreviation. ∎

REMARK 3.2. *It seems natural to ask whether there exist further examples of random matrices $\mathbf{F}_n^{(0)}$ such that for any $m \in \mathbb{N}$, $\mathbf{F}_n^{(0)}$ and $\mathbf{F}_n^{(1)} + \ldots + \mathbf{F}_n^{(m)}$ have the same limiting eigenvalue distributions after appropriate rescaling. However, it turns out that within the class of products of independent Girko–Ginibre matrices and their inverses, there exist no further examples beyond those mentioned in Theorem 1.1, apart from possible permutations of the exponents ± 1. Indeed, suppose that $\mathbf{F}_n^{(0)}$ is a product of p factors $\mathbf{Y}_n^{(r)}$ and q factors $(\mathbf{Y}_n^{(r)})^{-1}$ (all of them independent, and in arbitrary order), and let $\mathbf{V}_n$ and $\mathbf{W}_n$ be defined as in (2.1). Then, arguing as in the proof of Theorem 1.1, we find that the corresponding S-transforms $S_{\mathbf{W}}$ and $S_{\mathbf{V}}$ are given by*

$$S_{\mathbf{W}}(z) = \frac{(-z)^q}{(1+z)^p} \quad \text{and} \quad S_{\mathbf{V}}(z) = \frac{\mathrm{i}^q \, z^{(q-1)/2}}{(1+z)^{(p-1)/2}},$$

respectively, and by Proposition 2.2, the latter is the S-transform of a symmetric $\boxplus$-stable distribution if and only if $p = 1$ and $q \in \mathbb{N}_0$. Now use the observation that, by equation (2.11), if $\mu^{\boxplus m}$ is not a rescaled version of μ, then $\mathcal{H}(\mu^{\boxplus m})$ is not a rescaled version of $\mathcal{H}(\mu)$. ■

REMARK 3.3. *The limiting eigenvalue distribution in Theorems 1.1 and 1.2 may be interpreted as a* stable distribution *with respect to an appropriately defined convolution $\oplus$. To define this convolution, suppose that μ_1 and μ_2 are two probability measures which belong to the class H introduced above equation (2.11) and that $\mathbf{A}_n$ and $\mathbf{B}_n$ are independent bi-unitary invariant random matrices with limiting eigenvalue distributions μ_1 and μ_2, respectively. Also, suppose that these matrices satisfy the assumptions of Theorem 2.1. Then, if $\widetilde{\nu}_1$ and $\widetilde{\nu}_2$ are the limiting symmetrized singular value distributions of $\mathbf{A}_n$ and $\mathbf{B}_n$, we have $\mu_1 = \mathcal{H}(\widetilde{\nu}_1)$ and $\mu_2 = \mathcal{H}(\widetilde{\nu}_2)$ by Theorem 2.1. Furthermore, suppose that the matrix sums $\mathbf{A}_n + \mathbf{B}_n$ have the limiting symmetrized singular value distribution $\widetilde{\nu}_1 \boxplus \widetilde{\nu}_2$ (which seems natural in view of Proposition 2.1) and that they also satisfy the assumptions of Theorem 2.1. Then, again by Theorem 2.1, the associated limiting eigenvalue distribution is given by $\mathcal{H}(\widetilde{\nu}_1 \boxplus \widetilde{\nu}_2)$. This motivates the following definition:*

DEFINITION 3.1. Given two probability measures μ_1 and μ_2 of class H, set $\mu_1 \oplus \mu_2 := \mathcal{H}\big(\mathcal{H}^{-1}(\mu_1) \boxplus \mathcal{H}^{-1}(\mu_2)\big)$.

This convolution $\oplus$ may also be interpreted in terms of free probability. Indeed, given μ_1 and μ_2 in H, pick R-diagonal elements x_1 and x_2 (in some W^*-probability space) such that the Brown measure of x_1 is μ_1, the Brown measure of x_2 is μ_2, and x_1 and x_2 are $*$-free. Then $\mu_1 \oplus \mu_2$ is the Brown measure of $x_1 + x_2$, as follows from the results in [21] and [22].

It is now natural to introduce the concept of a (strictly) $\oplus$-stable distribution: A probability measure μ of class H is called (strictly) $\oplus$-stable if there exists a constant $\alpha > 0$ such that $\mu^{\oplus m} = \mathcal{D}_{m^{1/\alpha}}\mu$ for all $m \in \mathbb{N}$. Then, by equation (2.11), $\widetilde{\nu}$ is $\boxplus$-stable if and only if $\mathcal{H}(\widetilde{\nu})$ is $\oplus$-stable. Therefore, the $\oplus$-stable distributions in

H are in one-to-one correspondence with the symmetric $\boxplus$-*stable distributions on* $\mathbb{R}^*$, *and the limiting spectral distributions occurring in Theorems* 1.1 *and* 1.2 *are special examples of this type.* $\blacksquare$

4. PROOF OF THEOREM 1.3

4.1. Overview. In this section we prove Theorem 1.3 using the general framework from [18]. In Subsection 4.2, we summarize the technical conditions and the main universality results from [18] to make the presentation reasonably self-contained. Subsections 4.3–4.5 prepare for the proof of Theorem 1.3 by verifying the technical conditions from [18]. Subsection 4.6 contains the core of the proof of Theorem 1.3, and Subsection 4.7 describes the necessary modifications for Remark 1.1. Some auxiliary results from the literature are collected in Subsection 4.8.

4.2. General framework. A major step in [18] is to prove the *universality* of the limiting singular value and eigenvalue distributions, i.e. to show that these distributions (if existent) do not depend on the distributions of the matrix entries, apart from a few moment conditions as in (1.2)–(1.4). To state this more precisely, we need two sets of random matrices.

To this end, it seems convenient to regard $\mathbf{F}_n$ in (1.8) as a *matrix function* (by slight abuse of notation) and to write

$$(4.1) \qquad \mathbf{F}_n(\mathbf{Z}_n^{(1)}, \ldots, \mathbf{Z}_n^{(ml)}) := \sum_{q=1}^{m} \prod_{r=1}^{l} (\mathbf{Z}_n^{((q-1)l+r)})^{\varepsilon_r},$$

where $m, l \in \mathbb{N}$ and $\varepsilon_1, \ldots, \varepsilon_l \in \{-1, +1\}$ are the same as in (1.8), and $\mathbf{Z}_n^{(q)} = (Z_{jk}^{(q)})_{j,k=1,\ldots,n}$ is a matrix in the indeterminates $Z_{jk}^{(q)}$, $q = 1, \ldots, ml$. Then, we may write $\mathbf{F}_n(\mathbf{X}) := \mathbf{F}_n(\mathbf{X}_n^{(1)}, \ldots, \mathbf{X}_n^{(ml)})$ for the random matrices built from the random matrices $\mathbf{X}_n^{(q)} := \left(\frac{1}{\sqrt{n}} X_{jk}^{(q)}\right)_{j,k=1,\ldots,n}$, and $\mathbf{F}_n(\mathbf{Y}) := \mathbf{F}_n(\mathbf{Y}_n^{(1)}, \ldots, \mathbf{Y}_n^{(ml)})$ for the corresponding random matrices built from the *Gaussian* random matrices $\mathbf{Y}_n^{(q)} := \left(\frac{1}{\sqrt{n}} Y_{jk}^{(q)}\right)_{j,k=1,\ldots,n}$. We always assume that the families $(X_{jk}^{(q)})_{j,k,q \in \mathbb{N}}$ and $(Y_{jk}^{(q)})_{j,k,q \in \mathbb{N}}$ are defined on the same probability space and independent. When the choice of the matrices $\mathbf{X}_n^{(1)}, \ldots, \mathbf{X}_n^{(ml)}$ is clear from the context, we also write $\mathbf{F}_n$ instead of $\mathbf{F}_n(\mathbf{X})$.

REMARK 4.1. *More generally, using the arguments from this section, we might deal with matrix functions of the form*

$$\mathbf{F}_n(\mathbf{Z}_n^{(1)}, \mathbf{Z}_2^{(n)}, \mathbf{Z}_3^{(n)}, \ldots) := \sum_{q=1}^{m} \prod_{r=1}^{l_q} (\mathbf{Z}_n^{(i_{q,r})})^{\varepsilon_{q,r}},$$

where $m, l_1, \ldots, l_m \in \mathbb{N}, \varepsilon_{q,r} \in \{+1, -1\}$, *the indices* $i_{q,r} \in \mathbb{N}$ *are pairwise diffe-*

rent, and all parameters do not depend on n. That is to say, the numbers and the types of the factors in the m summands need not be the same.

In our investigation of the limiting spectral distributions of the matrices $\mathbf{F}_n$, we will also consider the *shifted matrices* $\mathbf{F}_n - \alpha \mathbf{I}_n$, with $\alpha \in \mathbb{C}$, the *regularized matrices* $\mathbf{F}_{n,t}$, with $t > 0$, and their combinations $\mathbf{F}_{n,t} - \alpha \mathbf{I}_n$. Here, the regularized matrices $\mathbf{F}_{n,t}$ arise from the regularized matrix functions

$$(4.2) \qquad \mathbf{F}_{n,t}(\mathbf{Z}_n^{(1)}, \dots, \mathbf{Z}_n^{(ml)}) := \sum_{q=1}^{m} \prod_{r=1}^{l} (\mathbf{Z}_n^{((q-1)l+r)})_t^{\varepsilon_r},$$

where $(\mathbf{Z}_n)_t^{\varepsilon} := \mathbf{Z}_n$ for $\varepsilon = +1$ and $(\mathbf{Z}_n)_t^{\varepsilon} := (\mathbf{Z}_n^* \mathbf{Z}_n + t \mathbf{I}_n)^{-1} \mathbf{Z}_n^*$ for $\varepsilon = -1$. Note that, by definition, the regularization has no effect when $\varepsilon = +1$ and that $\lim_{t \downarrow 0} (\mathbf{Z}_n)_t^{-1} = (\mathbf{Z}_n)^{-1}$ when $\mathbf{Z}_n$ is invertible.

Furthermore, fix a sequence $(\tau_n)_{n \in \mathbb{N}}$ of positive real numbers such that $\tau_n \to 0$ and $\tau_n \sqrt{n} \to \infty$, and for $0 \leqslant \varphi \leqslant \pi/2$, set

$$Z_{jk}^{(q)}(\varphi) := (\cos \varphi) X_{jk}^{(q)} \mathbf{1}_{\{|X_{jk}^{(q)}| \leqslant \tau_n \sqrt{n}\}} + (\sin \varphi) Y_{jk}^{(q)} \mathbf{1}_{\{|Y_{jk}^{(q)}| \leqslant \tau_n \sqrt{n}\}} \qquad (j, k, q \in \mathbb{N}),$$

$$\mathbf{Z}_n^{(q)}(\varphi) := \left(\tfrac{1}{\sqrt{n}} Z_{jk}^{(q)}(\varphi)\right)_{j,k=1,\dots,n} (q \in \mathbb{N}), \ \mathbf{F}_n(\varphi) := \mathbf{F}_n\big(\mathbf{Z}_n^{(1)}(\varphi), \dots, \mathbf{Z}_n^{(ml)}(\varphi)\big).$$

Note that the matrices $\mathbf{Z}_n^{(q)}(\varphi)$ provide an interpolation between truncated versions of the matrices $\mathbf{X}_n^{(q)}$ (for $\varphi = 0$) and $\mathbf{Y}_n^{(q)}$ (for $\varphi = \pi/2$).

With this notation, we have to check the following Conditions A, B and C.

CONDITION A. *For $\mathbf{F}_n = \mathbf{F}_n(\mathbf{X})$ and $\mathbf{F}_n = \mathbf{F}_n(\mathbf{Y})$, the matrices $\mathbf{F}_n$ satisfy the following condition:*

For each $\alpha \in \mathbb{C}$ and $z \in \mathbb{C}^+$, we have $\lim_{t \to 0} \limsup_{n \to \infty} |s_{n,t}(z) - s_n(z)| = 0$ in probability, where $s_n(z)$ and $s_{n,t}(z)$ are the Stieltjes transforms of the Hermitian matrices $(\mathbf{F}_n - \alpha \mathbf{I}_n)(\mathbf{F}_n - \alpha \mathbf{I}_n)^*$ and $(\mathbf{F}_{n,t} - \alpha \mathbf{I}_n)(\mathbf{F}_{n,t} - \alpha \mathbf{I}_n)^*$, respectively.

CONDITION B. *For each $t > 0$, $\alpha \in \mathbb{C}$, $z \in \mathbb{C}^+$, we have*

$$\sup_{\varphi \in [0; \pi/2]} \sup_{q,j,k} \max_{D} \big\| \mathbb{E}\{D g_{j,k}^{(q)}(\varphi) | X_{jk}^{(q)}, Y_{jk}^{(q)}\} \big\|_\infty \leqslant A < \infty,$$

where the maximum is over all partial derivatives D of orders zero, one and two in the matrix entries $\operatorname{Re} Z_{jk}^{(q)}$ and $\operatorname{Im} Z_{jk}^{(q)}$, $g_{j,k}^{(q)}(\varphi)$ may be either

$$g_{j,k}^{(q)}(\varphi) = \left(\frac{\partial}{\partial \operatorname{Re} Z_{jk}^{(q)}} \operatorname{trace} \big(\mathbf{V}_{n,t}(\alpha; \varphi) - z \mathbf{I}_{2n}\big)^{-1} \right)\Big|_{Z_{jk}^{(q)} \to \theta Z_{jk}^{(q)}}$$

or

$$g_{j,k}^{(q)}(\varphi) = \left(\frac{\partial}{\partial \operatorname{Im} Z_{jk}^{(q)}} \operatorname{trace} \big(\mathbf{V}_{n,t}(\alpha; \varphi) - z \mathbf{I}_{2n}\big)^{-1} \right)\Big|_{Z_{jk}^{(q)} \to \theta Z_{jk}^{(q)}},$$

the $\mathbf{V}_{n,t}(\alpha; \varphi)$ are the Hermitizations of the matrices $\mathbf{F}_n(\varphi)$, but with $\mathbf{F}_n$ replaced by $\mathbf{F}_{n,t} - \alpha\mathbf{I}_n$, and θ (the rescaling parameter in the substitution $Z_{j,k}^{(q)} \to \theta Z_{j,k}^{(q)}$) is a random variable which is uniformly distributed on $[0, 1]$ and independent of everything else.

CONDITION C. For $\mathbf{F}_n = \mathbf{F}_n(\mathbf{X})$ and $\mathbf{F}_n = \mathbf{F}_n(\mathbf{Y})$, the matrices $\mathbf{F}_n$ satisfy the following conditions:

(C0) There exists some $p > 0$ such that $\frac{1}{n} \sum_{k=1}^{n} s_k^p(\mathbf{F}_n)$ is bounded in probability as $n \to \infty$.

(C1) For any fixed $\alpha \in \mathbb{C}$, there exists some $Q > 0$ such that

$$\lim_{n\to\infty} \mathbb{P}\big(s_n(\mathbf{F}_n - \alpha\mathbf{I}_n) \leqslant n^{-Q}\big) = 0.$$

(C2) For any fixed $\alpha \in \mathbb{C}$, there exists some $0 < \gamma < 1$ such that for *any* sequence $(\delta_n)_{n\in\mathbb{N}}$ with $\delta_n \to 0$,

$$\lim_{n\to\infty} \mathbb{P}\big(\tfrac{1}{n} \sum_{n_1 \leqslant j \leqslant n_2} |\log s_j(\mathbf{F}_n - \alpha\mathbf{I}_n)| > \varepsilon\big) = 0 \quad \text{for all } \varepsilon > 0,$$

where $n_1 = [n - n\delta_n] + 1$ and $n_2 = [n - n^\gamma]$.

REMARK 4.2 (Condition C$_{\text{simple}}$). *It will be convenient to consider Condition* C *for more general random matrices* $\mathbf{F}_n$ *(with* $\mathbf{F}_n$ *of dimension* $n \times n$*) than in* (1.8). *If a sequence of random matrices* $\mathbf{F}_n$ *satisfies Conditions* (C0), (C1) *and* (C2), *we say that the matrices* $\mathbf{F}_n$ *satisfy* Condition C. *Also, if a sequence of random matrices* $\mathbf{F}_n$ *satisfies Condition* (C0) *as well as Conditions* (C1) *and* (C2) *with* $\alpha = 0$, *we say that the matrices* $\mathbf{F}_n$ *satisfy* Condition C$_{\text{simple}}$.

The following universality result is implicitly contained in [18]:

THEOREM 4.1 (Universality of singular value and eigenvalue distributions). *Let* $\mathbf{F}_n(\mathbf{X})$, $\mathbf{F}_n(\mathbf{Y})$ *be defined as above, and let* $\nu_n(\mathbf{X})$, $\nu_n(\mathbf{Y})$ *and* $\mu_n(\mathbf{X})$, $\mu_n(\mathbf{Y})$ *denote the associated singular value and eigenvalue distributions, respectively.*
   (a) *If Conditions* A *and* B *hold,* $\nu_n(\mathbf{X}) - \nu_n(\mathbf{Y}) \to 0$ *weakly in probability.*
   (b) *If Conditions* A, B *and* C *hold,* $\mu_n(\mathbf{X}) - \mu_n(\mathbf{Y}) \to 0$ *weakly in probability.*

P r o o f. (a) Set $\alpha := 0$. For $\mathbf{Z} = \mathbf{X}$ and $\mathbf{Z} = \mathbf{Y}$, let $m_n(z; \mathbf{Z})$ and $s_n(z; \mathbf{Z})$ denote the Stieltjes transforms of the Hermitian matrices

$$\mathbf{V}_n(\mathbf{Z}) := \begin{bmatrix} \mathbf{O} & \mathbf{F}_n(\mathbf{Z}) \\ \mathbf{F}_n^*(\mathbf{Z}) & \mathbf{O} \end{bmatrix} \quad \text{and} \quad \mathbf{W}_n(\mathbf{Z}) := \mathbf{F}_n(\mathbf{Z})\mathbf{F}_n^*(\mathbf{Z}),$$

and let $m_{n,t}(z; \mathbf{Z})$ and $s_{n,t}(z; \mathbf{Z})$ denote the corresponding Stieltjes transforms when $\mathbf{F}_n(\mathbf{Z})$ is replaced with $\mathbf{F}_{n,t}(\mathbf{Z})$. Fix $t > 0$. By Condition B and Theorem 3.2 in [18], we have, for each $z \in \mathbb{C}^+$, $m_{n,t}(z; \mathbf{X}) - m_{n,t}(z; \mathbf{Y}) \to 0$ in probability,

and therefore $s_{n,t}(z;\mathbf{X}) - s_{n,t}(z;\mathbf{Y}) \to 0$ in probability. (Note that the Lindeberg condition in [18] holds by our assumption (1.4), while the rank condition in [18] follows from basic inequalities for the rank of matrix sums and matrix products.) It therefore follows from Condition A that, for each $z \in \mathbb{C}^+$, $s_n(z;\mathbf{X}) - s_n(z;\mathbf{Y}) \to 0$ in probability, which implies the claim.

(b) By the same argument as in (a), the conclusion of (a) holds not only for the singular value distributions of the matrices $\mathbf{F}_n$, but also for the singular value distributions of the shifted matrices $\mathbf{F}_n - \alpha\mathbf{I}_n$ for any fixed $\alpha \in \mathbb{C}$. Thus, the claim follows from Condition C and Remark 4.2 in [18]. ∎

REMARK 4.3. *As follows from the proof, if one is only interested in the limiting singular value distributions of the matrices* $\mathbf{F}_n$, *it suffices to assume that Conditions* A *and* B *hold with* $\alpha = 0$.

We will use Theorem 4.1 to establish Theorem 1.3. This requires verifying Conditions A, B and C, of course. For this purpose, we provide some auxiliary results in the next three subsections.

4.3. On Condition A. Let $\mathbf{F}_n = \mathbf{F}_n(\mathbf{X})$ be defined as in (1.8). To obtain a matrix function which is smooth in the matrix entries (as needed for Condition B), we replace all inverses $(\mathbf{X}_n^{(q)})^{-1}$ with regularized inverses $(\mathbf{X}_n^{(q)})_t^{-1}$. We do this in a step-by-step fashion. Hence, fix $t > 0$, fix an index Q such that $\varepsilon_Q = -1$, and for all the other indices q with $\varepsilon_q = -1$, fix a choice between $(\mathbf{X}_n^{(q)})^{-1}$ and $(\mathbf{X}_n^{(q)})_t^{-1}$. Then it suffices to consider random matrices of the form

$$(4.3) \qquad \mathbf{F}_n = \mathbf{A}_n(\mathbf{X}_n)^{-1}\mathbf{B}_n + \mathbf{C}_n,$$

where $\mathbf{X}_n \equiv \mathbf{X}_n^{(Q)}$ (we omit the index Q for simplicity) and $\mathbf{A}_n$, $\mathbf{B}_n$ and $\mathbf{C}_n$ depend only on the matrices $\mathbf{X}_n^{(q)}$ with $q \neq Q$.

Fix $\alpha \in \mathbb{C}$, and for $0 \leqslant u \leqslant t$, let

$$(4.4) \qquad \mathbf{F}_{n,u} := \mathbf{A}_n(\mathbf{X}_n)_u^{-1}\mathbf{B}_n + \mathbf{C}_n := \mathbf{A}_n(\mathbf{X}_n^*\mathbf{X}_n + u)^{-1}\mathbf{X}_n^*\mathbf{B}_n + \mathbf{C}_n$$

and

$$(4.5) \qquad s_{n,u}(z) := \tfrac{1}{n}\operatorname{trace}\left((\mathbf{F}_{n,u} - \alpha\mathbf{I}_n)(\mathbf{F}_{n,u} - \alpha\mathbf{I}_n)^* - z\mathbf{I}_n\right)^{-1}.$$

Note that $\mathbf{F}_{n,0}$ coincides with $\mathbf{F}_n$ if $\mathbf{X}_n$ is invertible. Then, by way of induction, it will suffice to prove the following lemma:

LEMMA 4.1. *For each* $n \in \mathbb{N}$, *let* $\mathbf{X}_n = \left(\frac{1}{\sqrt{n}}X_{jk}\right)_{j,k=1,\ldots,n}$ *be as in* (1.1)–(1.4). *Furthermore, for each* $n \in \mathbb{N}$, *let* $\mathbf{A}_n$, $\mathbf{B}_n$ *and* $\mathbf{C}_n$ *be random matrices of dimension* $n \times n$ *such that the singular value distributions of the random matrices* $\mathbf{B}_n$ *and* $\mathbf{C}_n$ *converge weakly in probability to* (*non-random*) *probability measures on*

$(0, \infty)$ *and* $[0, \infty)$, *respectively, and let* $\mathbf{F}_{n,u}$ *and* $s_{n,u}(z)$ *be defined as in* (4.4) *and* (4.5). *Then, for any* $z = u + \mathrm{i}v \in \mathbb{C}^+$, *we have*

(4.6) $$\lim_{t \to 0} \limsup_{n \to \infty} |s_{n,t}(z) - s_{n,0}(z)| = 0 \text{ in probability.}$$

REMARK 4.4. *Let us emphasize that although the matrices* $\mathbf{A}_n$, $\mathbf{B}_n$, $\mathbf{C}_n$ *and* $\mathbf{X}_n$ *in the decomposition* (4.3) *are independent, this is not required in Lemma* 4.1.

REMARK 4.5. *Lemma* 8.16 *in* [18] *contains a similar result for the case* $\mathbf{C}_n = 0$, *although under additional assumptions and with a proof which does not seem to extend to the case* $\mathbf{C}_n \neq 0$. *The main difference in the proof of Lemma* 4.1 (*as compared to that of Lemma* 8.16 *in* [18]) *is that we control the* auxiliary *modifications of the matrices* $\mathbf{B}_n$ *and* $\mathbf{C}_n$ *via the matrix rank, and not via the resolvent.*

REMARK 4.6. *Let us illustrate the way Lemma* 4.1 *will be used later. Consider an* l-*fold product* $\mathbf{F}_n(\mathbf{X}) = (\mathbf{X}_n^{(1)})^{\varepsilon_1} \dots (\mathbf{X}_n^{(l)})^{\varepsilon_l}$, *where* $\varepsilon_1, \dots, \varepsilon_l \in \{-1, +1\}$, *and suppose by way of induction that we have weak convergence for any matrix product with less than* l *factors, possibly regularized. Then, setting* $\mathbf{F}_{n,t_1,\dots,t_l} := (\mathbf{X}_n^{(1)})_{t_1}^{\varepsilon_1} \dots (\mathbf{X}_n^{(l)})_{t_l}^{\varepsilon_l}$ *and*

$$s_n(t_1, \dots, t_l; z) := \tfrac{1}{n} \operatorname{trace} \left((\mathbf{F}_{n,t_1,\dots,t_l} - \alpha \mathbf{I}_n)(\mathbf{F}_{n,t_1,\dots,t_l} - \alpha \mathbf{I}_n)^* - z \mathbf{I}_n \right)^{-1}$$

and writing $\mathbf{t}_k := (t, \dots, t, 0, \dots, 0)$ *for the vector consisting of* k t's *and* $l - k$ 0's, *we have the estimate*

(4.7) $$|s_{n,t}(z) - s_{n,0}(z)| \leqslant \sum_{k=1}^{l} |s_n(\mathbf{t}_k; z) - s_n(\mathbf{t}_{k-1}; z)|.$$

Now, for each $k = 1, \dots, l$, *the* kth *summand on the right-hand side in* (4.7) *satisfies* (4.6), *either trivially (when* $\varepsilon_k = +1$) *or by Lemma* 4.1 *(when* $\varepsilon_k = -1$). *Thus, the left-hand side in* (4.7) *satisfies* (4.6) *as well, and Condition* A *is proved for the* l-*fold product* $\mathbf{F}_n(\mathbf{X})$.

P r o o f   o f   L e m m a   4.1. For the sake of simplicity, we consider only the case $\alpha = 0$ here, the extension to the case $\alpha \neq 0$ being straightforward. We have to show that for any given $\varepsilon > 0$ and $\delta > 0$,

(4.8) $$\limsup_{t \to 0} \limsup_{n \to \infty} \mathbb{P}\big(|s_{n,t}(z) - s_{n,0}(z)| > \varepsilon\big) < \delta.$$

Hence, fix $\varepsilon > 0$ and $\delta > 0$. As in the proof of Lemma 8.16 in [18], we introduce *auxiliary* modifications of the matrices $\mathbf{B}_n$ and $\mathbf{C}_n$ before we regularize the inverse matrices $\mathbf{X}_n^{-1}$.

For an $n \times n$ matrix $\mathbf{M}$, let $s_1(\mathbf{M}) \geqslant \dots \geqslant s_n(\mathbf{M})$ denote the singular values. Since the singular value distributions of $\mathbf{B}_n$ and $\mathbf{C}_n$ converge weakly in probability

to (non-random) probability measures on $(0, \infty)$ and $[0, \infty)$, respectively, we may find $K > 1$ and $N \in \mathbb{N}$ such that for $n \geqslant N$, we have

$$\mathbb{P}\left(\frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{s_k(\mathbf{B}_n) < K^{-1} \text{ or } s_k(\mathbf{B}_n) > K\}} > \frac{\varepsilon v}{24} \text{ or } \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{s_k(\mathbf{C}_n) > K\}} > \frac{\varepsilon v}{24} \right) < \frac{\delta}{2}.$$

Then, the modifications $\widetilde{\mathbf{B}}_n$ and $\widetilde{\mathbf{C}}_n$ are defined as follows. For the matrix $\mathbf{C}_n$, take the singular value decomposition $\mathbf{C}_n = \mathbf{U}\mathbf{\Delta}\mathbf{V}^*$, let $\widetilde{\mathbf{\Delta}}$ be the diagonal matrix obtained from $\mathbf{\Delta}$ by replacing the diagonal elements $\mathbf{\Delta}_{kk}$ with $\widetilde{\mathbf{\Delta}}_{kk} := \mathbf{\Delta}_{kk} \wedge K$, and set $\widetilde{\mathbf{C}}_n := \mathbf{U}\widetilde{\mathbf{\Delta}}\mathbf{V}^*$. For the matrix $\mathbf{B}_n$, take the singular value decomposition $\mathbf{B}_n = \mathbf{U}\mathbf{\Delta}\mathbf{V}^*$, let $\widetilde{\mathbf{\Delta}}$ be the diagonal matrix obtained from $\mathbf{\Delta}$ by replacing the diagonal elements $\mathbf{\Delta}_{kk}$ with $\widetilde{\mathbf{\Delta}}_{kk} := (\mathbf{\Delta}_{kk} \wedge K) \vee K^{-1}$, and set $\widetilde{\mathbf{B}}_n := \mathbf{U}\widetilde{\mathbf{\Delta}}\mathbf{V}^*$. Then we have

$$(4.9) \qquad \|\widetilde{\mathbf{B}}_n\| \leqslant K, \quad \|\widetilde{\mathbf{B}}_n^{-1}\| \leqslant K, \quad \|\widetilde{\mathbf{C}}_n\| \leqslant K,$$

and for $n \geqslant N$, with a probability of at least $1 - \delta/2$, we also have

$$(4.10) \qquad \tfrac{1}{n} \operatorname{rank}(\mathbf{B}_n - \widetilde{\mathbf{B}}_n) \leqslant \varepsilon v/24, \quad \tfrac{1}{n} \operatorname{rank}(\mathbf{C}_n - \widetilde{\mathbf{C}}_n) \leqslant \varepsilon v/24.$$

Furthermore, let $\widetilde{\mathbf{F}}_{n,u}$ and $\widetilde{s}_{n,u}(z)$ be defined as in (4.4) and (4.5), but with $\mathbf{B}_n$ and $\mathbf{C}_n$ replaced by $\widetilde{\mathbf{B}}_n$ and $\widetilde{\mathbf{C}}_n$. It then follows from (4.10) that for $n \geqslant N$, with a probability of at least $1 - \delta/2$, we have

$$\tfrac{1}{n} \operatorname{rank}(\mathbf{F}_{n,u}\mathbf{F}_{n,u}^* - \widetilde{\mathbf{F}}_{n,u}\widetilde{\mathbf{F}}_{n,u}^*) \leqslant \varepsilon v/6,$$

and therefore, by the rank inequality (compare e.g. [5], Lemma 6.9),

$$|s_{n,u}(z) - \widetilde{s}_{n,u}(z)| \leqslant \varepsilon/3.$$

Thus, we have reduced the proof of (4.8) to showing that

$$(4.11) \qquad \lim_{t \to 0} \limsup_{n \to \infty} |\widetilde{s}_{n,t}(z) - \widetilde{s}_{n,0}(z)| = 0 \text{ in probability.}$$

Since we only deal with the modified matrices for the rest of the proof, we omit the tildes and write $\mathbf{B}_n, \mathbf{C}_n, \mathbf{F}_{n,u}$ and $s_{n,u}(z)$ instead of $\widetilde{\mathbf{B}}_n, \widetilde{\mathbf{C}}_n, \widetilde{\mathbf{F}}_{n,u}$ and $\widetilde{s}_{n,u}(z)$, respectively. Moreover, for brevity, we usually omit the index n.

To establish (4.11), we may proceed as in the proof of Lemma 8.16 in [18]. Setting $\mathbf{R}_u := (\mathbf{F}_u\mathbf{F}_u^* - z\mathbf{I})^{-1}$, $0 \leqslant u \leqslant t$, we have the estimates

$$(4.12) \qquad \|\mathbf{R}_u\| \leqslant v^{-1}, \quad \|\mathbf{F}_u^*\mathbf{R}_u\mathbf{F}_u\| \leqslant 1 + |z|v^{-1},$$
$$\|\mathbf{R}_u\mathbf{F}_u\| \leqslant \left(v^{-1}(1 + |z|v^{-1})\right)^{1/2}, \quad \|\mathbf{F}_u^*\mathbf{R}_u\| \leqslant \left(v^{-1}(1 + |z|v^{-1})\right)^{1/2}$$

as well as the representation

$$(4.13) \qquad \mathbf{R}_t - \mathbf{R}_0 = \int_0^t \frac{d\mathbf{R}_u}{du}\, du = -\int_0^t \mathbf{R}_u \frac{d(\mathbf{F}_u\mathbf{F}_u^*)}{du}\mathbf{R}_u\, du.$$

Thus, it is straightforward to check that

$$\left|\tfrac{1}{n}\operatorname{trace}(\mathbf{R}_t - \mathbf{R}_0)\right| \leqslant \int_0^t \left|\tfrac{1}{n}\operatorname{trace}\big(\mathbf{R}_u\mathbf{F}_u\mathbf{B}^{-1}(\mathbf{X}\mathbf{X}^*+u\mathbf{I})^{-1}\mathbf{B}\mathbf{F}_u^*\mathbf{R}_u\big)\right|du$$

$$+\int_0^t \left|\tfrac{1}{n}\operatorname{trace}\big(\mathbf{R}_u\mathbf{C}\mathbf{B}^{-1}(\mathbf{X}\mathbf{X}^*+u\mathbf{I})^{-1}\mathbf{B}\mathbf{F}_u^*\mathbf{R}_u\big)\right|du$$

$$+\int_0^t \left|\tfrac{1}{n}\operatorname{trace}\big(\mathbf{R}_u\mathbf{F}_u\mathbf{B}^*(\mathbf{X}\mathbf{X}^*+u\mathbf{I})^{-1}(\mathbf{B}^*)^{-1}\mathbf{F}_u^*\mathbf{R}_u\big)\right|du$$

$$+\int_0^t \left|\tfrac{1}{n}\operatorname{trace}\big(\mathbf{R}_u\mathbf{F}_u\mathbf{B}^*(\mathbf{X}\mathbf{X}^*+u\mathbf{I})^{-1}(\mathbf{B}^*)^{-1}\mathbf{C}^*\mathbf{R}_u\big)\right|du.$$

Using the inequality $|\operatorname{trace}(\mathbf{M}_1\mathbf{M}_2\mathbf{M}_3)| \leqslant \|\mathbf{M}_1\|\|\mathbf{M}_3\| \operatorname{trace}(\mathbf{M}_2)$ (which holds for any $n \times n$ matrices $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{M}_3$ such that $\mathbf{M}_2$ is positive definite) as well as (4.9) and (4.12), we therefore obtain

$$(4.14) \qquad \left|\tfrac{1}{n}\operatorname{trace}(\mathbf{R}_t - \mathbf{R}_0)\right| \leqslant C(K, z) \int_0^t \tfrac{1}{n}\operatorname{trace}(\mathbf{X}\mathbf{X}^* + u\mathbf{I})^{-1}\, du,$$

where $C(K, z)$ is some constant depending only on K and z. Thus, it remains to show that

$$(4.15) \qquad \lim_{t\to 0}\limsup_{n\to\infty}\int_0^t \tfrac{1}{n}\operatorname{trace}(\mathbf{X}_n\mathbf{X}_n^* + u\mathbf{I}_n)^{-1}\, du = 0 \text{ in probability.}$$

But this follows from the fact that the random matrices $\mathbf{X}_n$ satisfy Condition C; see the proof of Lemma 8.14 in [18] for details. ∎

4.4. On Condition B. Here we have the following result:

LEMMA 4.2. *With $\mathbf{F}_n$ defined as in equation* (1.8), *Condition* B *holds.*

The proof follows from similar estimates to those in Section 8.1 in [18]. Since the required modifications are relatively straightforward, we omit the details.

4.5. On Condition C. Here we provide a number of lemmas which will be helpful in verifying Conditions C and C_{simple}. Recall that Condition C_{simple} was introduced in Remark 4.2.

LEMMA 4.3. *For each $n \in \mathbb{N}$, let $\mathbf{F}_n$ and $\mathbf{G}_n$ be random matrices of dimension $n \times n$. If the matrices $\mathbf{F}_n$ and $\mathbf{G}_n$ satisfy Condition C_{simple}, then the matrix products $\mathbf{F}_n\mathbf{G}_n$ also satisfy Condition C_{simple}.*

Since this result follows from similar arguments to those in the proof of Theorem 8.22 in [18] or to those for Lemma 4.6 below, we omit the proof.

LEMMA 4.4. *For each* $n \in \mathbb{N}$, *let* $\mathbf{X}_n = \left(\frac{1}{\sqrt{n}} X_{jk}\right)_{j,k=1,\dots,n}$ *be as in the assumptions* (1.1)–(1.4). *Then the matrices* $\mathbf{X}_n$ *and* $\mathbf{X}_n^{-1}$ *satisfy Condition* C.

For the matrices $\mathbf{X}_n$, Condition C is checked in [20] (in fact, it follows from the relation $\mathbb{E}\|\mathbf{X}\|_2^2 = n$ and from Lemmas 4.9 and 4.10), and for the matrices $\mathbf{X}_n^{-1}$, Condition C follows from the arguments given in the proof of Theorem 8.22 in [18]. We therefore omit the details.

REMARK 4.7. *A careful analysis of the proof of Theorem* 8.22 *in* [18] *shows that if the matrices* $\mathbf{G}_n$ *satisfy Condition* C_{simple}, *then the inverse matrices* $\mathbf{G}_n^{-1}$ *satisfy Conditions* (C1) *and* (C2) *with* $\alpha = 0$.

LEMMA 4.5. *Let* $\mathbf{F}_n = (\mathbf{X}_n^{(i_1)})^{\varepsilon_1} \dots (\mathbf{X}_n^{(i_l)})^{\varepsilon_l}$, *where* $l \in \mathbb{N}$, $i_1, \dots, i_l \in \mathbb{N}$ (*not necessarily different*), *and* $\varepsilon_1, \dots, \varepsilon_l \in \{-1, +1\}$ *are fixed. Then* $\mathbf{F}_n$ *satisfies Condition* C_{simple}.

P r o o f. By Lemma 4.4, the claim is true (even with the stronger Condition C) for $l = 1$. By Lemma 4.3 and induction, the claim remains true for $l > 1$. ∎

LEMMA 4.6. *For each* $n \in \mathbb{N}$, *let* $\mathbf{X}_n = \left(\frac{1}{\sqrt{n}} X_{jk}\right)_{j,k=1,\dots,n}$ *be as in the assumptions* (1.1)–(1.4). *Furthermore, for each* $n \in \mathbb{N}$, *let* $\mathbf{A}_n$, $\mathbf{B}_n$ *and* $\mathbf{C}_n$ *be random matrices of dimension* $n \times n$ *such that* $\mathbf{A}_n$, $\mathbf{B}_n$, $\mathbf{C}_n$ *and* $\mathbf{X}_n$ *are independent.*

(a) *If the matrices* $\mathbf{A}_n$ *and* $\mathbf{B}_n$ *satisfy Condition* C_{simple} *and the matrices* $\mathbf{C}_n$ *satisfy Condition* (C0), *then the matrices* $\mathbf{A}_n \mathbf{X}_n \mathbf{B}_n + \mathbf{C}_n$ *satisfy Condition* C.

(b) *If the matrices* $\mathbf{A}_n$ *and* $\mathbf{B}_n$ *satisfy Condition* C_{simple} *and the matrices* $\mathbf{C}_n$ *satisfy Condition* C *or* $\mathbf{C}_n = \mathbf{0}$ *for all* $n \in \mathbb{N}$, *then the matrices* $\mathbf{A}_n \mathbf{X}_n^{-1} \mathbf{B}_n + \mathbf{C}_n$ *satisfy Condition* C.

P r o o f. To shorten the notation, we omit the index n throughout this proof. First of all, let us note that if a sequence of random matrices $\mathbf{G}_n$ (with $\mathbf{G}_n$ of dimension $n \times n$) satisfies Condition (C0), there exists some $L_{\mathbf{G}} > 0$ such that

$$(4.16) \qquad \lim_{n \to \infty} \mathbb{P}(\|\mathbf{G}_n\| \geqslant n^{L_{\mathbf{G}}}) = 0.$$

In fact, if $p > 0$ is such that $\frac{1}{n} \sum_{k=1}^{n} s_k^p(\mathbf{G}_n)$ is bounded in probability as $n \to \infty$ and $\varepsilon > 0$ is arbitrary, it follows that

$$\limsup_{n \to \infty} \mathbb{P}\left(s_1(\mathbf{G}_n) \geqslant n^{(1+\varepsilon)/p}\right) \leqslant \limsup_{n \to \infty} \mathbb{P}\left(\frac{1}{n} \sum_{k=1}^{n} s_k^p(\mathbf{G}_n) \geqslant n^\varepsilon\right) = 0,$$

so that the assertion holds for any $L_{\mathbf{G}} > 1/p$.

(a) Condition (C0) follows from Lemmas 4.8 and 4.7, Hölder's inequality, and the fact that the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{X}$ satisfy Condition (C0). To prove Conditions (C1) and (C2), we use the factorization

$$\mathbf{A}\mathbf{X}\mathbf{B} + \mathbf{C} - \alpha\mathbf{I} = \mathbf{A}\big(\mathbf{X} + \mathbf{A}^{-1}(\mathbf{C} - \alpha\mathbf{I})\mathbf{B}^{-1}\big)\mathbf{B}.$$

Then it remains to check that for each of the three factors $\mathbf{M}_n$ on the right-hand side, we have, for some $Q > 0$,

$$\mathbb{P}\big(s_n(\mathbf{M}_n) \leqslant n^{-Q}\big) = o(1) \quad \text{and} \quad \frac{1}{n}\sum_{n_1 \leqslant n \leqslant n_2} \log^- s_j(\mathbf{M}_n) = o_P(1).$$

For $\mathbf{A}$ and $\mathbf{B}$, this is true by assumption. For $\mathbf{X} + \mathbf{A}^{-1}(\mathbf{C} - \alpha\mathbf{I})\mathbf{B}^{-1}$, this follows from Lemmas 4.9 and 4.10. More precisely, if the matrices $\mathbf{A}$ and $\mathbf{B}$ satisfy Condition (C1) with $\alpha = 0$ and $Q > 0$, and the matrices $\mathbf{C}$ satisfy (4.16) with $L_{\mathbf{C}} > 0$, we have $\mathbb{P}\big(s_1\big(\mathbf{A}^{-1}(\mathbf{C} - \alpha\mathbf{I})\mathbf{B}^{-1}\big) > 2n^{2Q+L_{\mathbf{C}}}\big) \to 0$ by Lemma 4.7. Thus, we may use Lemmas 4.9 and 4.10 conditionally on $\mathbf{A}, \mathbf{B}, \mathbf{C}$, and on the set of probability $1 + o(1)$ where $s_1\big(\mathbf{A}^{-1}(\mathbf{C} - \alpha\mathbf{I})\mathbf{B}^{-1}\big) \leqslant 2n^{2Q+L_{\mathbf{C}}}$.

(b) We consider only the case that the matrices $\mathbf{C}$ satisfy Condition C, leaving the simpler case $\mathbf{C} = \mathbf{0}$ to the reader. By reasoning as above, we see that Condition (C0) follows from Lemmas 4.8 and 4.7, Hölder's inequality, and the fact that the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{X}^{-1}$ satisfy Condition (C0). To prove Conditions (C1) and (C2), we use the factorization

$$\mathbf{A}\mathbf{X}^{-1}\mathbf{B} + \mathbf{C} - \alpha\mathbf{I} = \mathbf{A}\mathbf{X}^{-1}\big(\mathbf{B}(\mathbf{C} - \alpha\mathbf{I})^{-1}\mathbf{A} + \mathbf{X}\big)\mathbf{A}^{-1}(\mathbf{C} - \alpha\mathbf{I}).$$

Then it remains to check that for each of the five factors $\mathbf{M}_n$ on the right-hand side, we have, for some $Q > 0$,

$$\mathbb{P}\big(s_n(\mathbf{M}_n) \leqslant n^{-Q}\big) = o(1) \quad \text{and} \quad \frac{1}{n}\sum_{n_1 \leqslant n \leqslant n_2} \log^- s_j(\mathbf{M}_n) = o_P(1).$$

But this is true (i) by assumption, (ii) by Lemma 4.4, (iii) by Lemmas 4.9 and 4.10 (applied conditionally on $\mathbf{A}, \mathbf{B}, \mathbf{C}$), (iv) by Remark 4.7, and (v) again by assumption. ∎

4.6. Proof of Theorem 1.3. After the preparations above, we may turn to the proof of Theorem 1.3. Given a sequence of random matrices $(\mathbf{G}_n)_{n\in\mathbb{N}}$, we write $\nu(\mathbf{G}_n)$ for the singular value distributions, $\mu(\mathbf{G}_n\mathbf{G}_n^*)$ for the squared singular value distributions, $\mathcal{S}\nu(\mathbf{G}_n)$ for the symmetrized singular value distributions, and $\nu_{\mathbf{G}}$, $\mu_{\mathbf{G}\mathbf{G}^*}$ and $\mathcal{S}\nu_{\mathbf{G}}$ for the corresponding weak limits in probability (if existent). Furthermore, for $t > 0$, let $\gamma_t^{+1} := \gamma^{+1} := \gamma$, and let γ_t^{-1} be the induced measure of γ under the mapping $x \mapsto (x + t)^{-1}x(x + t)^{-1}$. These notions are motivated by our regularization procedure in equation (4.2).

Let us start with the singular value distributions. We will first use induction on l to prove the claim for the case $m = 1$ and then use induction on m to prove the claim for the case $m > 1$. More precisely, we will show the following:

(4.17)   The matrices $\mathbf{F}_n(\mathbf{X})$ from (1.8) satisfy Conditions A and B, and for any $t > 0$, the singular value distributions of the matrices $\mathbf{F}_{n,t}(\mathbf{X})$ converge weakly in probability to the probability measure ν_t on $(0, \infty)$ with symmetrization $\mathcal{S}\nu_t = \left(\mathcal{Q}^{-1}(\gamma_t^{\varepsilon_1} \boxtimes \ldots \boxtimes \gamma_t^{\varepsilon_l}) \right)^{\boxplus m}$.

Indeed, by Condition A, we may then let $t \to 0$ get the limiting singular value distribution of the matrices $\mathbf{F}_n(\mathbf{X})$. Note that Condition B has already been established in Lemma 4.2, so that it remains to check Condition A as well as the claim about the limiting distribution.

Products of independent random matrices. For $\mathbf{F}_n(\mathbf{X}) = \mathbf{X}_n$, Condition A holds trivially, and for $\mathbf{F}_n(\mathbf{X}) = \mathbf{X}_n^{-1}$, Condition A holds by Lemma 4.1. Furthermore, the Marchenko–Pastur theorem implies that, for any $t > 0$ and $\varepsilon \in \{-1, +1\}$, we have $\mu\left(\mathbf{X}_{n,t}^{\varepsilon}(\mathbf{X}_{n,t}^{\varepsilon})^* \right) \to \gamma_t^{\varepsilon}$. Thus, (4.17) is true for $l = 1$.

Now let $l > 1$, let $\mathbf{F}_n$ be an l-fold product of independent random matrices, and suppose that (4.17) holds for any product $\mathbf{G}_n$ with less than l factors. It then follows from Lemma 4.1 that the matrices $\mathbf{F}_n(\mathbf{X})$ satisfy Condition A; see Remark 4.6 for details. Now consider the particular decomposition $\mathbf{F}_n(\mathbf{X}) = \mathbf{X}_n^{\varepsilon} \mathbf{G}_n(\mathbf{X})$, where $\varepsilon = +1$ or $\varepsilon = -1$ and $\mathbf{X}_n$ and $\mathbf{G}_n(\mathbf{X})$ are independent. Then, for any $t > 0$, the matrices $\mathbf{Y}_{n,t}^{\varepsilon}$ and $\mathbf{G}_{n,t}(\mathbf{Y})$ are independent bi-unitary invariant matrices with

$$\mu\left(\mathbf{Y}_{n,t}^{\varepsilon}(\mathbf{Y}_{n,t}^{\varepsilon})^* \right) \to \gamma_t^{\varepsilon} \quad \text{and} \quad \mu\left(\mathbf{G}_{n,t}(\mathbf{Y}) \mathbf{G}_{n,t}^*(\mathbf{Y}) \right) \to \mu_{\mathbf{G}(t)\mathbf{G}(t)^*},$$

by the inductive hypothesis in the latter case. Therefore, by asymptotic freeness (see Proposition 2.1 (a)),

$$\mu\left(\mathbf{F}_{n,t}(\mathbf{Y}) \mathbf{F}_{n,t}^*(\mathbf{Y}) \right) \to \gamma_t^{\varepsilon} \boxtimes \mu_{\mathbf{G}(t)\mathbf{G}(t)^*}.$$

Thus, by Theorem 4.1 (a), (4.17) holds for the matrices $\mathbf{F}_n(\mathbf{X})$ as well.

Hence, by induction on l, we come to the conclusion that (4.17) holds for any product of independent matrices (i.e. for the case $m = 1$).

Sums of products of independent random matrices. We have just proved (4.17) for $m = 1$. Now let $m > 1$, let $\mathbf{F}_n$ be an m-fold sum of products of independent random matrices, and suppose that (4.17) holds for any such sum $\mathbf{C}_n$ with less than m summands. It then follows by Lemma 4.1 and a similar argument to that in Remark 4.6 that the matrices $\mathbf{F}_n(\mathbf{X})$ satisfy Condition A. Now consider the particular decomposition $\mathbf{F}_n(\mathbf{X}) = \mathbf{G}_n(\mathbf{X}) + \mathbf{C}_n(\mathbf{X})$, where $\mathbf{G}_n(\mathbf{X})$ is a product, $\mathbf{C}_n(\mathbf{X})$ is an $(m-1)$-fold sum of products, and $\mathbf{G}_n(\mathbf{X})$ and $\mathbf{C}_n(\mathbf{X})$ are independent. Then, for any $t > 0$, the matrices $\mathbf{G}_{n,t}(\mathbf{Y})$ and $\mathbf{C}_{n,t}(\mathbf{Y})$ are independent bi-unitary invariant matrices with

$$\mathcal{S}\nu\left(\mathbf{G}_{n,t}(\mathbf{Y}) \right) \to \mathcal{S}\nu_{\mathbf{G}(t)} \quad \text{and} \quad \mathcal{S}\nu\left(\mathbf{C}_{n,t}(\mathbf{Y}) \right) \to \mathcal{S}\nu_{\mathbf{C}(t)}$$

by the result for the case $m = 1$ and the inductive hypothesis, respectively. Therefore, by asymptotic freeness (see Proposition 2.1 (c)),

$$\mathcal{S}\nu\big(\mathbf{F}_{n,t}(\mathbf{Y})\big) \to \mathcal{S}\nu_{\mathbf{C}(t)} \boxplus \mathcal{S}\nu_{\mathbf{G}(t)}.$$

Thus, by Theorem 4.1 (a), (4.17) holds for the matrices $\mathbf{F}_n(\mathbf{X})$ as well.

Hence, by induction on m, we come to the conclusion that (4.17) holds for any sum of products of independent matrices (i.e. for the case $m > 1$).

Let us now consider the eigenvalue distributions. To begin with, by Lemma 4.6, we may check by induction on m that the matrices $\mathbf{F}_n(\mathbf{X})$ satisfy Condition C, too. Therefore, we may use Theorem 4.1 (b), and it remains to determine the limiting eigenvalue distributions in the Gaussian case, i.e. for the matrices $\mathbf{F}_n(\mathbf{Y})$. Here, it follows by asymptotic freeness (see Proposition 2.1 (d)) that $\mathcal{S}\nu\big(\mathbf{F}_{n,t}(\mathbf{Y}) - \alpha \mathbf{I}_n\big) \to \mathcal{S}\nu_{t,\alpha} := (\mathcal{S}\nu_t) \boxplus B(\alpha)$, with $B(\alpha)$ as in Theorem 2.1. Letting $t \to 0$ and using Condition A, it further follows that $\mathcal{S}\nu\big(\mathbf{F}_n(\mathbf{Y}) - \alpha \mathbf{I}_n\big) \to \mathcal{S}\nu_\alpha := (\mathcal{S}\nu) \boxplus B(\alpha)$, where ν is the probability measure described in the theorem. Now apply Theorem 2.1. ∎

4.7. Proof of Remark 1.1. A slight variation of the preceding arguments shows that Conditions A, B and C continue to hold for random matrices $\mathbf{F}_n$ of the form (1.10), *provided that* the extra condition (1.11) holds:

C o n d i t i o n A. Here we can regularize the matrices $(\mathbf{X}_n^{-1})^l$ by means of $\big((\mathbf{X}_n)_t^{-1}\big)^l$ (i.e. each factor in the power is regularized individually) and invoke Lemma 4.1. For this, it is important that the matrices $\mathbf{A}_n$, $\mathbf{B}_n$ and $\mathbf{C}_n$ in Lemma 4.1 need not be independent of $\mathbf{X}_n$; see Remark 4.4.

C o n d i t i o n B. Here we may extend Lemma 4.2 to products of powers of independent Girko–Ginibre matrices, using similar arguments to those in Sections 8.1.3 and 8.1.4 in [18].

C o n d i t i o n C. Under the extra condition (1.11), it follows from Lemma 4.6 (applied with $\mathbf{X} = \mathbf{X}^{(r)}$) and by induction on m that the matrices $\mathbf{F}_n$ satisfy Condition C. (Unfortunately, without the extra condition (1.11), Lemma 4.6 does not allow us to draw this conclusion in general, even though we would expect that Condition C continues to hold in this case.)

Now, the proof of Remark 1.1 is quite similar to that of Theorem 1.3, which is why we omit the details. ∎

4.8. Auxiliary results. In this subsection we collect several auxiliary results from the literature which we have used to verify Condition C. Let $\mathbf{A}$ and $\mathbf{B}$ be $n \times n$ matrices, and recall that $s_1(\mathbf{M}) \geqslant \ldots \geqslant s_n(\mathbf{M})$ denote the singular values of the $n \times n$ matrix $\mathbf{M}$.

LEMMA 4.7 ([24], Theorem 3.3.14). *For all $p > 0$ and all $k = 1, \ldots, n$, we have $\sum_{j=1}^k \big(s_j(\mathbf{AB})\big)^p \leqslant \sum_{j=1}^k \big(s_j(\mathbf{A}) s_j(\mathbf{B})\big)^p$.*

LEMMA 4.8 ([24], Theorem 3.3.16). *For all $p > 0$, we have $\sum_{j=1}^{n} s_j^p(\mathbf{A} + \mathbf{B})$ $\leqslant C_p\big(\sum_{j=1}^{n} s_j^p(\mathbf{A}) + \sum_{j=1}^{n} s_j^p(\mathbf{B})\big)$, where C_p is a constant depending only on p.*

Since $s_j(\mathbf{M}^{-1}) = s_{n-j+1}^{-1}(\mathbf{M})$, $j = 1, \ldots, n$, it is clear that similar results hold for the smallest singular values.

LEMMA 4.9 ([20], Section 5). *Suppose that the conditions (1.1)–(1.4) hold. Then, for any fixed $K > 0$ and $L > 0$, there exist positive constants A and B such that for any non-random matrices $\mathbf{M}_n$ with $\|\mathbf{M}_n\|_2 \leqslant Kn^L$, we have*

$$\mathbb{P}\big(s_n(\mathbf{X}_n - \mathbf{M}_n) \leqslant n^{-A}\big) \leqslant n^{-B}.$$

LEMMA 4.10 ([20], Section 5). *Suppose that the conditions (1.1)–(1.4) hold. Then, for any fixed $K > 0$ and $L > 0$, there exists a constant $0 < \gamma < 1$ such that for any non-random matrices $\mathbf{M}_n$ with $\|\mathbf{M}_n\|_2 \leqslant Kn^L$ and for any sequence $\delta_n \to 0$, we have*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{n_1 \leqslant j \leqslant n_2} \log^- s_j(\mathbf{X}_n - \mathbf{M}_n) = 0 \text{ almost surely,}$$

where $n_1 = [n - n\delta_n] + 1$ and $n_2 = [n - n^\gamma]$.

REFERENCES

[1] K. Adhikari, N. Kishore Reddy, T. Ram Reddy, and K. Saha, *Determinantal point processes in the plane from products of random matrices*, Ann. Inst. Henri Poincaré Probab. Stat. 52 (1) (2016), pp. 16–46.

[2] G. Akemann and J. R. Ipsen, *Recent exact and asymptotic results for products of independent random matrices*, Acta Phys. Polon. B 46 (9) (2015), pp. 1747–1784.

[3] N. Alexeev, F. Götze, and A. Tikhomirov, *Asymptotic distribution of singular values of powers of random matrices*, Lith. Math. J. 50 (2) (2010), pp. 121–132.

[4] O. Arizmendi E. and V. Pérez-Abreu, *The S-transform of symmetric probability measures with unbounded supports*, Proc. Amer. Math. Soc. 137 (9) (2009), pp. 3057–3066.

[5] Z. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, second edition, Springer, New York 2010.

[6] S. T. Belinschi, T. Mai, and R. Speicher, *Analytic subordination theory of operator-valued free additive convolution and the solution of a general random matrix problem*, J. Reine Angew. Math. 732 (2017), pp. 21–53.

[7] S. T. Belinschi, P. Śniady, and R. Speicher, *Eigenvalues of non-Hermitian random matrices and Brown measure of non-normal operators: Hermitian reduction and linearization method*, Linear Algebra Appl. 537 (2018), pp. 48–83.

[8] H. Bercovici and V. Pata, *Stable laws and domains of attraction in free probability theory*, Ann. of Math. (2) 149 (3) (1999), pp. 1023–1060.

[9] H. Bercovici and D. Voiculescu, *Free convolution of measures with unbounded support*, Indiana Univ. Math. J. 42 (3) (1993), pp. 733–773.

[10] P. Biane and F. Lehner, *Computation of some examples of Brown's spectral measure in free probability*, Colloq. Math. 90 (2) (2001), pp. 181–211.

[11] C. Bordenave, *On the spectrum of sum and product of non-Hermitian random matrices*, Electron. Commun. Probab. 16 (2011), pp. 104–113.

[12] C. Bordenave and D. Chafaï, *Around the circular law*, Probab. Surv. 9 (2012), pp. 1–89.

[13] Z. Burda, *Free products of large random matrices – a short review of recent developments*, J. Phys.: Conf. Ser. 473 (2013), 012002.

[14] Z. Burda, R. A. Janik, and B. Waclaw, *Spectrum of the product of independent random Gaussian matrices*, Phys. Rev. E (3) 81 (4) (2010), 041132.

[15] Z. Burda, A. Jarosz, G. Livan, M. A. Nowak, and A. Swiech, *Eigenvalues and singular values of products of rectangular Gaussian random matrices*, Phys. Rev. E (3) 82 (6) (2010), 061114.

[16] P. J. Forrester, *Eigenvalue statistics for product complex Wishart matrices*, J. Phys. A 47 (2014), 345202.

[17] P. J. Forrester and D. Liu, *Raney distributions and random matrix theory*, J. Stat. Phys. 158 (5) (2015), pp. 1051–1082.

[18] F. Götze, H. Kösters, and A. Tikhomirov, *Asymptotic spectra of matrix-valued functions of independent random matrices and free probability*, Random Matrices Theory Appl. 4 (2) (2015), 1550005.

[19] F. Götze and A. Tikhomirov, *The circular law for random matrices*, Ann. Probab. 38 (4) (2010), pp. 1444–1491.

[20] F. Götze and A. N. Tikhomirov, *On the asymptotic spectrum of products of independent random matrices*, preprint, arXiv:1012.2710, 2010.

[21] U. Haagerup and F. Larsen, *Brown's spectral distribution measure for R-diagonal elements in finite von Neumann algebras*, J. Funct. Anal. 176 (2) (2000), pp. 331–367.

[22] U. Haagerup and H. Schultz, *Brown measures of unbounded operators affiliated with a finite von Neumann algebra*, Math. Scand. 100 (2) (2007), pp. 209–263.

[23] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables and Entropy*, American Mathematical Society, Providence, RI, 2000.

[24] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge 1991.

[25] A. Jarosz, *Summing free unitary random matrices*, Phys. Rev. E 84 (2011), 011146.

[26] F. Lehner, *On the computation of spectra in free probability*, J. Funct. Anal. 183 (2) (2001), pp. 451–471.

[27] W. Młotkowski, M. A. Nowak, K. A. Penson, and K. Życzkowski, *Spectral density of generalized Wishart matrices and free multiplicative convolution*, Phys. Rev. E (3) 92 (1) (2015), 012121.

[28] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability*, Cambridge University Press, Cambridge 2006.

[29] S. O'Rourke and A. Soshnikov, *Products of independent non-Hermitian random matrices*, Electron. J. Probab. 16 (2011), paper no. 81, pp. 2219–2245.

[30] N. R. Rao and R. Speicher, *Multiplication of free random variables and the S-transform: The case of vanishing mean*, Electron. Commun. Probab. 12 (2007), pp. 248–258.

[31] T. Rogers, *Universal sum and product rules for random matrices*, J. Math. Phys. 51 (9) (2010), 093304.

[32] E. B. Saff and V. Totik, *Logarithmic Potentials with External Fields* (with Appendix B by Thomas Bloom), Springer, Berlin 1997.

[33] R. Speicher, *Polynomials in asymptotically free random matrices*, Acta Phys. Polon. B 46 (9) (2015), pp. 1611–1624.

[34] T. Tao and V. Vu, *Random matrices: Universality of ESDs and the circular law*, Ann. Probab. 38 (5) (2010), pp. 2023–2065.

[35] A. N. Tikhomirov, *Asymptotic distribution of the singular numbers for spherical ensemble matrices*, Mat. Tr. 16 (2013), pp. 169–200.

[36] D. A. Timushev and A. N. Tikhomirov, *On the asymptotic distribution of singular val-*

ues of powers products of sparse random matrices, Izvestia Komi Science Center of Ural Division of RAS 13 (2013), pp. 10–17.

[37] D. A. Timushev and A. N. Tikhomirov, *Limit theorems for spectra of sums of random matrices*, Proc. Komi Science Center of Ural Division of RAS 187 (2014), pp. 24–32.

[38] D. V. Voiculescu, K. J. Dykema, and A. Nica, *Free Random Variables: A Noncommutative Probability Approach to Free Products with Applications to Random Matrices, Operator Algebras, and Harmonic Analysis on Free Groups*, American Mathematical Society, Providence, RI, 1992.

Holger Kösters   Alexander Tikhomirov
Department of Mathematics   Institute of Physics and Mathematics
Bielefeld University, Germany   Komi Science Center of Ural Division of RAS
E-mail: hkoesters@math.uni-bielefeld.de   Syktyvkar State University, Russia
                                            E-mail: antikhom51@gmail.com