RATE OF CONVERGENCE FOR PRODUCTS OF INDEPENDENT NON-HERMITIAN RANDOM MATRICES

JONAS JALOWY

ABSTRACT. We study the rate of convergence of the empirical spectral distribution of products of independent non-Hermitian random matrices to the power of the Circular Law. The distance to the deterministic limit distribution will be measured in terms of a uniform Kolmogorov-like distance. First, we prove that for products of Ginibre matrices, the optimal rate is given by $O(1/\sqrt{n})$. Avoiding the edge, the rate of convergence of the mean empirical spectral distribution is even faster. Second, we show that also products of matrices with independent entries attain this optimal rate in the bulk up to a logarithmic factor. In the case of Ginibre matrices, we apply a saddlepoint approximation to a double contour integral representation of the density and in the case of matrices with independent entries we make use of techniques from local laws.

1. Introduction

The Circular Law states that the empirical spectral distribution of a non-Hermitian random matrix with i.i.d. entries converges to the uniform distribution on the complex disc as the size of the matrix tends to infinity. Interestingly, for the product of $m$ independent matrices of such type, the limit distribution will be the $m$-th power of the Circular Law. Here we investigate the question: How fast does it converge? The case $m = 1$ has been studied in [GJ18] already.

We consider the product

$$X = \frac{1}{\sqrt{n^m}} \prod_{q=1}^{m} X^{(q)}$$

of $m$ independent random matrices $X^{(1)}, \ldots, X^{(m)}$, each of size $n \times n$. For fixed $m \in \mathbb{N}$, the asymptotic in $n \to \infty$ will be of interest. Its empirical spectral distribution is given by

$$\mu_n^m = \frac{1}{n} \sum_{j=1}^{n} \delta_{\lambda_j^m(X)},$$

where $\delta_{\lambda}$ are Dirac measures in the eigenvalues $\lambda_j$ of the matrix $X$. In this note we are interested in two different classes of random matrices $X^{(q)}$ that appear in the product.

Date: December 20, 2019.

2010 Mathematics Subject Classification. 60B20 (Primary); 41A25 (Secondary).

Key words and phrases. products of non-Hermitian random matrices, Ginibre matrices, Meijer G function, saddlepoint approximation, logarithmic potential, circular law, rate of convergence.

Supported by the German Research Foundation (DFG) through the IRTG 2235.
**Definition 1.** (i) A (complex) Ginibre matrix $X$ is a complex non-Hermitian random matrix with independent complex Gaussian entries $X_{ij} \sim \mathcal{N}_C(0,1)$.

(ii) A non-Hermitian random $n \times n$-matrix $X$ is said to have independent entries if $X_{ij}$ are independent complex or real random variables, and in the complex case we additionally assume $\Re X_{ij}$ and $\Im X_{ij}$ to be independent.

If $X^{(1)}, \ldots, X^{(m)}$ have independent entries satisfying $\mathbb{E} X_{ij}^{(q)} = 0$ and $\mathbb{E}|X_{ij}^{(q)}|^2 = 1$, then the empirical spectral distribution converges weakly to a deterministic probability measure on the complex plane as the matrix size grows. We denote by $\mu$ the 2-dimensional Lebesgue measure, by $\Rightarrow$ weak convergence of measures and $B_r = B_r(0)$ shall be the centered open ball of radius $r > 0$. In [GT10b], Götze and Tikhomirov showed that as $n \to \infty$, $\mathbb{P}$-almost surely we have

$$
\mu_n^m \Rightarrow \mu_\infty^m, \text{ where } d\mu_\infty^m(z) = \frac{|z|^{2/m-2}}{\pi m} \mathbf{1}_{B_1^m}(z) d\mu(z)
$$

(1)

is the $m$-th power of the uniform distribution $\mu_\infty = \mu_\infty^1$ on the complex disc, see also [OS11]. The Gaussian case has been treated in [BJW10, AB12], more general models can be found in [KT15, GKT15, Bor11, AI15, IK14], for the convergence of the singular values see [AGT10] and furthermore for local results we refer to [Nem17, KOV18, Nem18, GNT17]. For $m = 1$, we retrieve the well known circular law $\mu_n = \mu_1^n \Rightarrow \mu_\infty$. In the case of Ginibre matrices this has been discovered much earlier in [Gin65]. We are interested in the rate of convergence, more precisely in the Kolmogorov distances over balls

$$
D(\mu_n^m, \mu_\infty^m) := \sup_{z_0 \in \mathbb{C}, R > 0} |\mu_n^m(B_R(z_0)) - \mu_\infty^m(B_R(z_0))|
$$

as $n \to \infty$. Convergence in this distance coincides with weak convergence in the case of an absolutely continuous limit distribution, see [GJ18]. Using the rotational invariant property of the Circular Law $\mu_\infty$ and the mean empirical spectral distribution $\bar{\mu}_n = \mathbb{E} \mu_n$ of the Ginibre ensemble, the following optimal rate of $\mu_n$ for $m = 1$ has been shown in [GJ18].

**Lemma 2.** The mean empirical spectral distribution $\bar{\mu}_n = \mathbb{E} \mu_n$ of the Ginibre ensemble satisfies

$$
D(\bar{\mu}_n, \mu_\infty) \sim \frac{1}{\sqrt{2\pi n}}
$$

(2)

and for any fixed $\varepsilon > 0$

$$
\sup_{B_R(z_0) \subseteq \mathbb{C} \setminus B_{1+\varepsilon} \text{ or } B_R(z_0) \subseteq B_{1-\varepsilon}} |\bar{\mu}_n(B_R(z_0)) - \mu_\infty(B_R(z_0))| \lesssim e^{-n\varepsilon^2}.
$$

(3)

Here and in the sequel we denote asymptotic equivalence by $\sim$. We write $\lesssim$, if an inequality holds up to a parameter-independent constant $c > 0$ and $A \asymp B$ if $c |B| \leq |A| \leq C |B|$ for some constants $0 < c < C$. These constants $c, C$ may differ in each occurrence. Moreover we abbreviate $\log^a b = \log(b)^a$.

We will show the following analogous result for products of $m \geq 1$ independent Ginibre matrices.

**Theorem 3.** The mean empirical spectral distribution $\bar{\mu}_n^m = \mathbb{E} \mu_n^m$ satisfies

$$
\sup_{R > 0} |\bar{\mu}_n^m(B_R) - \mu_\infty^m(B_R)| \asymp \frac{1}{\sqrt{mn}}.
$$

(4)
The following more detailed estimates hold as long as the boundary of the complex disk is avoided

\[
\sup_{R<1-\frac{m}{2}\sqrt{\log n/n}} |\bar{\mu}_n^m(B_R) - \mu^m_\infty(B_R)| \lesssim \frac{\log^{3/2} n}{n} \tag{5}
\]

and uniformly in \( R > 1 + \sqrt{\log n/n} \)

\[
|\bar{\mu}_n^m(B_R) - \mu^m_\infty(B_R)| \lesssim e^{-n(R-1)^2}. \tag{6}
\]

The precise constants of the upper and lower bound of (4) can be chosen to be \( C = \sqrt{\pi}/\sqrt{2} \) and \( c = 1/(\sqrt{2\pi}) \), coinciding with Lemma 2. We will see that the maximal distance is attained at \( R = 1 \). The rate of convergence is faster inside and much faster outside of the bulk. However, it might be an artifact of the method of proof that we do not obtain an exponential rate of convergence inside the bulk in the case of products of Gaussian random matrices. Only the rate \( O(1/n) \) seems to be achievable due to the discrete nature of the residue calculus, cf. (13) below. While the proof of Lemma 2 is an elementary calculation, the proof of Theorem 3 is more involved and relies on a saddle-point method of a double contour integral representation for the density of \( \mu^m_\infty \), see Lemma 7. An idea of the proof is given after the contours are defined, see Figure 2. Figure 1 illustrates the statements of our main results.

Heuristically, the typical distance of \( n \) uniformly distributed eigenvalues in the Circular Law is \( n^{-1/2} \), therefore one may vary \( B_R(z_0) \) up to a magnitude of \( n^{-1/2} \)

---

**Figure 1.** The empirical spectral distribution of the product \( \mathbf{X} \) of \( m = 2 \) Ginibre matrices

Left: The eigenvalues of a sample for \( n = 500 \) and the unit ball \( B_1(0) \) as reference. Theorem 5 below shows that gaps (like one can see at the top) and clusters do not significantly differ from the limit distribution.

Right: The radial part of the densities of \( \bar{\mu}_n^m \) for \( n = 15 \) in blue and of the limit distribution \( \mu^m_\infty \) in orange. Clearly the rate of convergence in the bulk is faster than close to the edge, illustrating the statement of Theorem 3.
without covering a new eigenvalue. Hence we also expect \( D(\mu^m_n, \mu^m_\infty) \) to be of order \( O(n^{-1/2}) \).

Pointwise convergence of the density \( \rho^m_n \) of \( \bar{\mu}_n \) or \( \bar{\mu}^m_n \) has been also discussed in [AB12, AC18, TV15]. In particular Akemann and Burda describe the asymptotic

\[
\rho^m_n(z) = \frac{|z|^{2/m-2}}{\pi m} \left( \sqrt{\frac{mn}{2}} \frac{|z|^{2/m} - 1}{|z|^{1/m}} \right) + o(1)
\]

for fixed \( z \) without specifying the error. Aside from the error \( o(1) \), the appearance of \( \text{erfc}(\cdot) \) also hints at exponential convergence, like in (3) and (6). Note that Akemann and Cikovic mention an algebraic rate of convergence for the fixed trace ensemble and conclude that the exponential rate of convergence for Ginibre matrices is rather special. Chafaï, Hardy and Maida studied invariant \( \beta \)-ensembles with external potential \( V \) instead of matrices with independent entries, see [CHM18]. Their result implies a rate of convergence to the limiting measure with density \( c \Delta V \) of order \( O(\sqrt{\log n/n}) \) with respect to the bounded Lipschitz metric and the 1-Wasserstein distance. Similar questions in this context of log-gases, but for the non-uniform variant of \( D \) (the discrepancy) have been addressed in [Ser17].

Based on the ideas of [GJ18], we will also prove a rate of convergence result for products of matrices with independent entries.

**Definition 4 (Condition (C)).** We say \( X = X^{(1)} \cdots X^{(m)} / \sqrt{mn} \) satisfies condition (C) if the matrices \( X^{(q)}, q = 1, \ldots, m, \) have jointly independent entries and satisfy

\[
\left| \mathbb{E} X^{(q)}_{ij} \right| \leq n^{-1-\epsilon} \text{ and } \left| 1 - \mathbb{E} \left| X^{(q)}_{ij} \right|^2 \right| \leq n^{-1-\epsilon}
\]

for some \( \epsilon > 0 \) independent of \( n \) and furthermore

\[
\max_{i,j,q,n} \mathbb{E} \left| X^{(q)}_{ij} \right|^{4+\delta} < \infty
\]

for some \( \delta > 0 \).

In Section 3, we will show

**Theorem 5.** If condition (C) holds, then for every \( \tau, Q > 0 \) there exist a constant \( c > 0 \) such that

\[
P \left( \sup_{B \subseteq B_{1-\tau} \cup B_{1+\tau}} \left| (\mu^m_n - \mu^m_\infty)(B) \right| \leq ch_m(n) \right) \geq 1 - n^{-Q},
\]

where the asymptotic error is given by

\[
h_m(n) = \begin{cases} 
  n^{-1/2} \log^2 n & \text{for } m = 1, \\
  n^{-1/2} \log^3 n & \text{for } m = 2, \\
  n^{-2/(m+2)} \log^{8/(m+2)} n & \text{for } m \geq 3.
\end{cases}
\]

Theorem 3 provides the optimal rate of convergence. In the proof of Theorem 5 we will see that the \( m \)-dependent term is only visible for balls touching the origin. To make the statement more comprehensible when comparing with Theorem 3, we also state the following result.

**Corollary 6.** If condition (C) holds, then for every \( \tau, Q > 0 \) we have

\[
P \left( \sup_{B} \left| (\mu^m_n - \mu^m_\infty)(B) \right| \lesssim \frac{\log^2 n}{\sqrt{n}} \right) \geq 1 - n^{-Q},
\]
where the supremum runs over all balls $B$ such that $\partial B_R(z_0) \subseteq B_{R+} \cup B_{1-} \setminus B_{\tau}$ avoids the edge and the origin.

We already know from Theorem 3 that the optimal rate is given by $O(1/\sqrt{n})$, hence Corollary 6 shows that this rate is also satisfied for matrices with independent entries, if edge and origin are avoided. A weaker rate of convergence for $m = 1$ has already been established in [GJ18], where a comparison with similar results can be found as well. We would like to point out a subtle difference between Theorem 5 and the Local Law in [GNT17]; the latter compares an integral over a smooth function with the limiting distribution on a shrinking support for a fixed point $z_0$, while the former allows to choose the “worst ball” $B_R(z_0)$ depending on the random sample of eigenvalues.

Note that we cannot expect an exponentially fast rate of convergence like in Lemma 2 for the non-averaged empirical spectral distribution $\mu_n^m$, because it is still affected by the individual eigenvalue fluctuations. In particular the rough lower bound $D(\mu_n^m, \mu_\infty^m) \geq 1/n$ follows from placing a disk of radius $cn^{1/2}$ contained in $B_1(0)$ such that it does not cover any eigenvalue.

2. Product of Ginibre matrices

We start with the following double contour integral representation for the density of $\mu_n^m$ that is essential for Theorem 3.

**Lemma 7.** The density of $\mu_n^m$ satisfies

$$
\rho_n^m(z) = \frac{1}{n(2\pi i)^2} \oint_\gamma \int_{\frac{1}{2}+i\infty}^{\frac{1}{2}-i\infty} \left( \frac{\Gamma(s)}{\Gamma(t)} \right)^m n^{m(t-s)} |z|^{2(t-s)-1} \cot(\pi t) ds dt,
$$

where $\gamma$ is any closed contour that encircles the numbers $1, \ldots, n$ counter clockwise and no natural number greater than $n$.

In [KZ14], a similar double contour integral representation for the correlation kernel of the singular values of $X$ has been derived. This was used in [LWZ16] to prove bulk universality for singular values of products of independent Ginibre matrices. In general, double contour integrals like in Lemma 7 above appear regularly in the theory of products of random matrices, e.g. [KZ14, FW17, KKS15].

In the sequel we will make use of the Meijer G-function, for $\xi \in \mathbb{C} \setminus \{0\}$ defined as the Mellin inverse

$$
G_{pq}^{mn} \left( \frac{a_1, \ldots, a_p}{b_1, \ldots, b_q}, \xi \right) = \frac{1}{2\pi i} \int_L \frac{\prod_{j=1}^m \Gamma(b_j - t) \prod_{j=1}^n \Gamma(1 - a_j + t) \prod_{j=m+1}^p \Gamma(1 - b_j + t) \prod_{j=n+1}^q \Gamma(a_j - t)}{\prod_{j=1}^n \Gamma(1 - j - t)} \xi^t dt,
$$

where $0 \leq m \leq q$, $0 \leq n \leq p$, $a_k - b_j \notin \mathbb{N}$ for $k = 1, \ldots, n$ and $j = 1, \ldots, m$. The contour $L$ goes from $-i\infty$ to $i\infty$, but can be chosen arbitrarily as long as the poles of $\Gamma(b_j - t)$ are on the right hand side of the path and the poles of $\Gamma(1 - a_j - t)$ are on the left hand side.

In particular the density of $\mu_n^m$ is given by

$$
\rho_n^m(z) = n^{m-1} \sum_{k=0}^{n-1} \frac{n^{mk} |z|^{2k}}{\pi(k)!^m} G_{0m}^{m0} \left( \frac{-n^m |z|^2}{0} \right), \quad (7)
$$

see [AB12] and compare to the case $m = 1$, where $G_{01}^{10} \left( -n |z|^2 \right) = e^{-n|z|^2}$. 
Remark 8. The viewpoint of studying products of $m$ matrices and definition (7) of $\tau_n^m$ makes sense for $m \in \mathbb{N}$ only. However the representation of Lemma 7 makes sense for arbitrary $m > 1, m \in \mathbb{R}$. Furthermore, as we can see from the proof of Theorem 3, its statements (4) and (6) remain true for real $m > 1$, as well as (5) for real $m \geq 2$.

Remark 9. Since the constants and errors in Theorem 3 are explicit in $m$, it is possible to consider the double scaling limit and let $m = m(n) \to \infty$. In this case, the rate will be faster, depending on $m$ and in particular by setting $m = n$ we obtain a rate

$$\sup_{R > 0} |\tilde{\mu}_n^m(B_R) - \delta_0(B_R)| \sim \frac{1}{n}.$$  

This follows from the fact that $\mu_n^m$ converges to its weak limit $\mu_\infty^m = \delta_0$ at rate $O(1/n)$. Note that $m \sim n$ is also the critical scaling of Lyapunov exponents between deterministic and GUE statistics, see [ABK19].

Proof of Lemma 7. For the contour $L$ of the Meijer G-function in (7), we choose the straight vertical line $L = [-1/2 - i\infty, -1/2 + i\infty]$ that after a simple change of variables $-t = s$ leads to

$$G_{0m}^m \left( \begin{array}{c} -n^m |z|^2 \\ 0 \end{array} \right) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \Gamma(s)(n^m |z|^2)^{-s} ds$$

(8)

The remaining part of (7) is the kernel of the (monic) orthogonal polynomials with respect to the Meijer-G-weight. It can be rewritten with the help of the residue theorem. For any closed curve $\gamma$ encircling the numbers $1, \ldots, n$ but no natural number greater than $n$, we have

$$\frac{1}{2\pi i} \int_\gamma \left( \frac{(n^m |z|^2)^{t-1}}{\Gamma(t)^m} \right) \frac{\cot(\pi t)}{\pi} dt = \sum_{k=0}^{n-1} \frac{n^{mk} |z|^{2k}}{\pi (k!)^m},$$

since the integrand is holomorphic except for its simple poles in $\mathbb{N}$ with residues $1/\pi$ each. Combining both contour integrals proves the claim. □

Asymptotic expansions of $G_{0m}^m$, like §5.9.1. in [Luk14], together with heuristics for the hypergeometric kernel give rise to pointwise limits in [AB12]. A rigorous estimation of the error bound uniformly in $z$ seems to be absent in the literature so far. Even the precise asymptotic formulas for $G_{0m}^m$ are valid only for fixed $z$, hence integration around $z = 0$ is impossible. Integrating $\tau_n^m$ first leads to $n$-dependent coefficients $a_j$ or $b_j$ in the Meijer G-function, but its asymptotics, e.g. from [Luk14], are not uniform in these coefficients. Hence it is reasonable to study the problem by a direct analysis.

Proof of Theorem 3. By Lemma 2, it is sufficient to consider $m \geq 2$. For $R > 1$ we have $|\tilde{\mu}_n^m(B_R)| < |(\tilde{\mu}_n^m - \mu_\infty^m)(B_1)|$, since $\text{supp}(\mu_\infty^m) = B_1(0)$. Throughout the proof we assume

$$\log^{3m/4} n/n^{m/2} \leq R \leq 1$$

(9)

since for smaller values of $R$ it would hold

$$|(\tilde{\mu}_n^m - \mu_\infty^m)(B_R)| \leq |(\tilde{\mu}_n^m - \mu_\infty^m)(B_{\log^{3m/4} n/n^{m/2}})| + O(\log^{3/2} n/n),$$

for $R > 1$.
due to $\mu^m_n(B_R) = R^{2/m}$. We first use spherical symmetry of $\rho^m_n$ and Lemma 7 in order to calculate

$$\tilde{\mu}_n^m(B_R) = \int_0^R 2\pi \rho^m_n(\sqrt{r}) dr$$

$$= \frac{\pi}{n(2\pi i)^2} \oint_\gamma \frac{1}{s - i\infty} \left( \frac{\Gamma(s)}{\Gamma(t)} \right)^{m} \frac{(n^m R^2)^{t-s}}{t-s} \cot(\pi t) ds dt. \quad (10)$$

This holds in the case where $s$ and $t$ have distance bounded from below, which is what we will choose in the following. We will now show that shifting the vertical contour in Lemma 7 to $L = [\eta - i\infty, \eta + i\infty]$, for another real part $\eta \geq 1/2$, $\eta \neq 1, \ldots, n$, produces an additional term. Cauchy’s integral formula implies

$$\frac{\pi}{(2\pi i)^2} \left( \int_{\frac{1}{2}+i\infty}^{\frac{1}{2}-i\infty} - \int_L \right) \left( \frac{\Gamma(s)}{\Gamma(t)} \right)^{m} \frac{(n^m R^2)^{t-s}}{t-s} ds = \frac{\pi}{2\pi i} \mathbb{1}_{(1/2, \eta)}(\text{Re}(t)).$$

We temporarily split $\gamma$ into two parts $\gamma_l$ and $\gamma_r$ such that $\gamma_l$ encircles $\{1, \ldots, [\eta]/n\}$ and $\gamma_r$ encircles $\{[\eta], \ldots, n\}$. Soon we will make the path of $\gamma$ more explicit. Continuing the integration of the right hand side of the last equation over $\gamma_l \cup \gamma_r$ as in (10) yields

$$\frac{\pi}{2\pi i} \oint_\gamma \cot(\pi t) dt = [\eta] \wedge n,$$

hence we conclude

$$\tilde{\mu}_n^m(B_R) = \frac{\pi}{n(2\pi i)^2} \oint_\gamma \int_L \left( \frac{\Gamma(s)}{\Gamma(t)} \right)^{m} \frac{(n^m R^2)^{t-s}}{t-s} \cot(\pi t) ds dt + \frac{[\eta]}{n} \wedge 1$$

Choosing $\eta = [n R^{2/m}] + 1/2$ we see that the second term is $O(1/n)$ close to $\mu^m_n(B_R) = R^{2/m} \wedge 1$. Moreover, by Cauchy’s integral formula, we may artificially add the removed part $\gamma - \gamma_l - \gamma_r$ again as long as $\gamma$ is symmetric around the $x$-axis. Let $\gamma$ be the rectangular contour connecting the vertices $3/4 - i, n + 1/4 - i, n + 1/4 + i$ and $3/4 + i$. This ensures a constant distance to the singularities. The scaled version $\tilde{\gamma} = \gamma/(R^{2/m}n)$ is illustrated below in Figure 2. Furthermore note that the integral exists as we will explicitly show below, see (23). Recall Stirling’s formula for the Gamma function

$$\log \Gamma(z) = (z - 1/2) \log z - z + \frac{1}{2} \log 2\pi + O(1/\text{Re}z),$$

which holds uniformly for $\text{Re}z \geq 1/2$, cf. for instance [WW96] p.249. Thus, we have

$$\log \left( \frac{\Gamma(s)^m}{(n^m R^2)^{s}} \right) = m \left( s \log \left( \frac{s}{n R^{2/m}} \right) - 1 \right) + \frac{1}{2} \log \left( \frac{2\pi}{s} \right) + O(1/\text{Re}(s)), \quad (12)$$

where for all $s \in L$ (and analogously for $t \in \gamma$) the error term is at most $O(1)$. We rescale the integration by $n R^{2/m}$ and denote $\tilde{\gamma} = \gamma/(n R^{2/m})$, $L = L/n R^{2/m}$ as well as

$$F(z) = z \log z - z.$$
Taylor approximation around $z$ to be the remaining part of the path (under slight abuse of notation).

Observe that $O(\frac{1}{n\text{Re}(s)}) = O(R^{2/m})$ and $O(\frac{1}{n\text{Re}(s)}) = O(1/n)$. We will analyze this main formula using the method of steepest decent, hence we are interested in the saddle points of $F$. The saddle point equation simply reads

$$F'(z) = \log z = 0$$

and is obviously satisfied only for $z = 1$ with $F''(1) = 1 > 0$. Denoting $z = x + iy$, the Cauchy-Riemann equations for $F$ imply

$$\partial_y \text{Re} F(z) = -\text{Im} F'(z) = -\arg(z) > 0 \iff y < 0 \quad \text{(14)}$$
$$\partial_x \text{Re} F(z) = \text{Re} F'(z) = \log |z| > 0 \iff |z| > 1, \quad \text{(15)}$$

hence $\text{Re} F$ attains its local maximum $F(1) = -1$ in $y$-direction and its minimum in $x$-direction. Define the box $Q_{\delta_n}(1) = [1 - \delta_n, 1 + \delta_n] \times [-\delta_n, \delta_n]$ around $z = 1$ of range

$$\delta_n = \sqrt{\frac{\log n}{nR^{2/m}}} \leq \log^{-1/4} n \to 0$$

by our assumption (9). Note that $\tilde{\gamma}$ is $1/nR^{2/m} \in O(\delta_n^2)$-close to the real axis and the vertical path $\tilde{L}$ is equally close to the saddle point $z = 1$. The local part of the paths are given by $L_{loc} = \tilde{L} \cap Q_{\delta_n}(1)$ and $\gamma_{loc} = \tilde{\gamma} \cap Q_{\delta_n}(1)$ as well as $\tilde{L}_{loc}$ and $\gamma_{loc}$ to be the remaining part of the path (under slight abuse of notation).

Let us collect the necessary bounds for each part of the contour by applying a Taylor approximation around $z = 1$. We have $(s - 1) = i\text{Im}(s) + O(\delta_n^2)$, hence for $s \in L_{loc}$

$$F(s) = -1 - \text{Im}(s)^2/2 + O(\delta_n^3), \quad \text{(16)}$$

and similarly for $t \in \gamma_{loc}$

$$F(t) = -1 + (1 - \text{Re}(t))^2/2 + O(\delta_n^3), \quad \text{(17)}$$

since $|\text{Im}(t)| \leq \delta_n$. On the other hand for $s \in L_{loc}'$ by using (14)

$$\text{Re} F(s) < \text{Re} F(\eta/nR^{2/m} + i\delta_n) = -1 - \delta_n^2/2 + O(\delta_n^3) \quad \text{(18)}$$

and for $t \in \gamma_{loc}'$ we see from (15)

$$\text{Re} F(t) > \text{Re} F(1 \pm \delta_n) = -1 + \delta_n^2/2 + O(\delta_n^3). \quad \text{(19)}$$
Figure 2. The scaled contours of integration and their local parts in thicker lines. As we will see later, the main contribution comes from $\gamma_{loc}$ that is in a box $\delta_n$-close to the saddle point at $z = 1$. If $R > 1$, there is no $\gamma_{loc}$ and the integral vanishes exponentially fast (depending on the distance $|R - 1|$). If $R < 1$, then both horizontal contours of $\gamma_{loc}$ will cancel, because of their symmetry. In this case we will obtain a rate of convergence of microscopic order $1/n$ due to the discrete nature of the residues. The maximal rate will be attained for $R = 1$, where the integrals do not cancel, yet the vertical part $\gamma_{vert}$ is small enough.

The nonlocal terms are negligible, e.g. we apply (17) and (18) to obtain

$$R^{2/m} \int_{\gamma_{loc}} \int_{L_{loc}} \exp \left[ nmR^{2/m} (\Re F(s) - \Re F(t)) \right] \left| \frac{t}{s} \right|^{m/2} \left| \frac{\cot(\pi nR^{2/m}t)}{|t - s|} \right| ds dt$$

$$\lesssim R^{2/m} \int_{\gamma_{loc}} \int_{L_{loc}} \exp \left[ - \frac{m}{2} \log n + \mathcal{O}(\delta_n) \right] \frac{1}{|s|^{m/2} |\Im(s)|} ds dt$$

$$\lesssim R^{2/m} n^{-m/2} \lesssim n^{-1},$$

where we used $|\Im(s)| \gtrsim \delta_n$, $|\gamma_{loc}| \in \mathcal{O}(\delta_n)$, $t \in \mathcal{O}(1)$, $|\cot(\pi nR^{2/m}t)| \approx 1$ and $m \geq 2$. Moreover from (18), (19) and $t \in \mathcal{O}(R^{-2/m})$ it follows

$$R^{2/m} \int_{\gamma_{loc}} \int_{L_{loc}} \exp \left[ nmR^{2/m} (\Re F(s) - \Re F(t)) \right] \left| \frac{t}{s} \right|^{m/2} \left| \frac{\cot(\pi nR^{2/m}t)}{|t - s|} \right| ds dt$$

$$\lesssim R^{2/m} \int_{\gamma_{loc}} \int_{L_{loc}} \exp \left[ - m \log n + \mathcal{O}(\delta_n) \right] \frac{R^{-1}}{|s|^{m/2} |\Im(s)|} ds dt$$

$$\lesssim R^{-1} n^{-m} \lesssim n^{-1},$$
where the last step once more follows from the assumption (9). Analogously we obtain from (16), (19)

\[ R^{2/m} \int_{\gamma_{loc}} \int_{L_{loc}} \exp \left[ \frac{nmR^{2/m}}{2} (Re(s) - Re(t)) \right] \left| \frac{t}{s} \right|^{m/2} \frac{1}{|t-s|} dsdt \]

\[ \lesssim R^{2/m} \int_{\gamma_{loc}} \int_{L_{loc}} \exp \left[ -\frac{m}{2} \log n + O(\delta_n) \right] \left| \frac{t}{s} \right|^{m/2} \frac{1}{|Re(t) - 1|} dsdt \]

\[ \lesssim \delta_n R^{-1+2/m} n^{-m/2} \lesssim n^{-1}. \]

Locally close to \( z = 1 \), the error term of Stirling’s formula (13) is \( O(n^{-1}) \). Thus it remains to control

\[ \int_{\gamma_{loc}} \int_{L_{loc}} \exp \left[ \frac{nmR^{2/m}}{2} (F(s) - F(t)) \right] \left( \frac{t}{s} \right)^{\frac{m}{2}} \frac{\cot(\pi n R^{2/m} t)}{t-s} \left( R^{2/m} + O(n^{-1}) \right) dsdt \]

\[ = \int_{\gamma_{loc}} \int_{L_{loc}} \exp \left[ -\frac{nmR^{2/m}}{2} (\Im(s)^2 + (1 - \Re(t))^2) \right] \frac{\cot(\pi n R^{2/m} t)}{t-s} \left( R^{2/m} + O(R^{1/m} \sqrt{\log^3 n/n}) \right) dsdt \]

where we used (16), (17) and \( t/s = 1 + O(\delta_n) \). We parameterize \( L_{loc} \) as the straight line

\[ s = \frac{\eta}{nR^{2/m}} + \frac{i}{\sqrt{nmR^{2/m}}} u, \quad u \in I = (-\sqrt{m \log n}, +\sqrt{m \log n}). \]

The vertical microscopic part

\[ \gamma_{vert} = [3/4 + i, 3/4 - i]/nR^{2/m} \cup ([n + 1/4 - i, n + 1/4 + i]/nR^{2/m}) \]

receives the same scaling, e.g. for the right part we choose

\[ t = R^{-2/m} + \frac{1}{4nR^{2/m}} + \frac{i}{\sqrt{nmR^{2/m}}} v, \quad v \in (-\sqrt{m/nR^{2/m}}, \sqrt{m/nR^{2/m}}). \]

This part of the integral (21) on \( \gamma_{vert} \) is visible if and only if \( R \) is close to 1. The exponential function becomes \( e^{u^2/2} \) after dropping the negligible part in \( t \). Using \( R \sim 1 \) and \( |\cot(\pi/4 + ix)| = 1 \) for \( x \in \mathbb{R} \), the integration over \( \gamma_{vert} \) can then be bounded by

\[ \left| \int_{\sqrt{m/nR^{2/m}}}^{\sqrt{m/nR^{2/m}}} \int_{I} e^{-u^2/2} \cot \left( \frac{\pi}{4} + ivR^{1/m} \sqrt{n/m} \right) \frac{m(n - \eta + 1/4)/R^{1/m} + iv - u}{\sqrt{nm}} R^{1/m} mdu \right| \]

\[ \lesssim \int_{\sqrt{m/nR^{2/m}}}^{\sqrt{m/nR^{2/m}}} \int_{I} e^{-u^2/2} \frac{m(n - \eta + 1/4)^2/R^{2/m} + n(v - u)^2}{\sqrt{nm}} mdu \]

\[ \lesssim \frac{1}{n} \int_{-\infty}^{\infty} e^{-u^2/2} du \lesssim \log n/n, \]

where in the second step we shifted \( u \) by \( v = O(1/\sqrt{n}) \) and used \( m(n - \eta + 1/4)^2/R^{2/m} \gtrsim 1 \). The last step follows from the asymptotics of the modified Bessel function \( K_0(1/4n) \) or more elementary by splitting the integration into \( |u| \leq 1 \).
From $\delta_n \to 0$ and $1/nR^{2/m} \to 0$ it follows that the left vertical path is not contained in $Q_{\delta_n}(1)$.

We parameterize the remaining path of $\gamma_{\text{loc}} \setminus \gamma_{\text{vert}}$ as horizontal lines

$$t = 1 + \frac{1}{\sqrt{nmR^{2/m}}}v \pm \frac{i}{nR^{2/m}}, \quad v \in \tilde{I} \subseteq I,$$

where $\tilde{I}$ is the part of $I$ such that the corresponding contour overlaps $\tilde{\gamma}$. The integral (21) becomes the sum of

$$\int_I \int \frac{e^{-\frac{u^2 + v^2}{2}}}{\mp v + (nR^{2/m} - \eta)\sqrt{m/nR^{2/m}} + i \left(\pm \sqrt{m/nR^{2/m}} - u\right)} \, du \, dv + O(\sqrt{\log n/n}) \, d\lambda \sim \frac{\pi}{\sqrt{nm}},$$

where we shifted $u, v \in I$ by $O(\sqrt{m/nR^{2/m}}) = O(1/\log n)$ and extended the area of integration. Recalling the correct prefactor $c = -1/4\pi$ from (13), we conclude

$$|\tilde{\mu}_n(B_R) - \tilde{\mu}_n(B_R)| \leq \frac{\pi}{2nm} + o(n^{-1/2}).$$

In order to obtain the lower bound of the claim, it suffices to consider $R = 1$. Moreover we will only study the sum of (23) keeping the phase factor of the integrand, since all the other parts of the double contour integral are proven to be of strictly lower order than $o(n^{-1/2})$. We have the asymptotic

$$\tilde{\mu}_n(B_1) - \tilde{\mu}_n(B_1)$$

$$= \frac{1}{4\pi} \left( \int_{-\sqrt{m}/4\sqrt{n}}^{\sqrt{m}/4\sqrt{n}} \int_{-\sqrt{m}/4\sqrt{n}}^{\sqrt{m}/4\sqrt{n}} e^{-\frac{u^2 + v^2}{2}} \frac{\cot(i\pi - \pi \sqrt{m/n})}{\sqrt{nm}} \, du \, dv + \frac{1}{4\pi} \right)$$

$$+ \frac{1}{4\pi} \left( \int_{-\sqrt{m}/4\sqrt{n}}^{\sqrt{m}/4\sqrt{n}} \int_{-\sqrt{m}/4\sqrt{n}}^{\sqrt{m}/4\sqrt{n}} e^{-\frac{u^2 + v^2}{2}} \frac{\cot(-i\pi + \pi \sqrt{m/n})}{\sqrt{nm}} \, du \, dv \right)$$

$$= \frac{1}{4\pi} \left( \int_{-\sqrt{m}/4\sqrt{n}}^{\sqrt{m}/4\sqrt{n}} \int_{-\sqrt{m}/4\sqrt{n}}^{\sqrt{m}/4\sqrt{n}} e^{-\frac{u^2 + v^2}{2}} \frac{1}{\sqrt{nm}} \, du \, dv \right)$$

$$\sim \frac{1}{\sqrt{2\pi mn}}.$$
The same upper bound holds with $1 + \varepsilon$ instead. For a better control of the constant one may vary the distance of $\gamma$ to the real axis from the start. The above asymptotic yields the first claim and coincides with $m = 1$ from Lemma 2.

If we avoid the edge by some distance $|1 - R^{1/m}| \geq \sqrt{\log n / n} > 0$, then $|1 + 1/R^{2/m}| > \delta_n$. Hence $\gamma_{\text{vert}}$ is not part of $\gamma_{\text{loc}}$ and (22) drops out. This is the case for $R < 1 - \frac{\pi}{2} \sqrt{\log n / n}$, for which we have $I = \tilde{I}$. As before, we shift $u$ and $v$ by $O(1/\sqrt{nR^{2/m}})$ and obtain

$$
\int_I \int_I e^{-x^2 - y^2} \left( \frac{-\cot(\pi n R^{2/m} + i \pi - \sqrt{nR^{2/m}}/mv)}{-v + (nR^{2/m} - \eta)\sqrt{m/nR^{2/m} + i(\sqrt{m/nR^{2/m}} - u)}} 
+ \frac{\cot(\pi n R^{2/m} - i \pi + \sqrt{nR^{2/m}}/mv)}{v + (nR^{2/m} - \eta)\sqrt{m/nR^{2/m} - i(\sqrt{m/nR^{2/m}} + u)}} \right) \frac{iR^{1/m} + O(\sqrt{nR^{2/m}}/mv)}{\sqrt{nR}} dudv
$$

$$
= \int_I \int_I e^{-x^2 - y^2} \frac{iR^{1/m} + O(\sqrt{nR^{2/m}}/mv)}{\sqrt{nR nR^{2/m}}} dudv
$$

$$
\tan(i \pi - \pi\sqrt{nR^{2/m}/mv}) - \tan(-i \pi + \pi\sqrt{nR^{2/m}/mv}) dudv.
$$

From the last line its obvious that the horizontal contour integrals (23) cancel, hence

$$
\mu_{\infty}^m(B_R) - \mu_{n}^m(B_R)
\lesssim \frac{\log^{3/2} n}{n}
$$

due to symmetry.

Lastly we turn to the statement about exponential decay for $R > 1 + \sqrt{\log n / n}$. As before it is sufficient to consider $R \leq 2$, because of $\text{supp}(\mu_{\infty}^m) = B_1$. The position of the minimum in $x$-direction in (15) and $\text{Re}(t) \leq (n + 1)/n R^{2/m}$ for $t \in \tilde{\gamma}$ yield

$$
\text{Re}F(t) = \text{Re}F(\text{Re}(t)) + O(1/n)
\geq F \left( \frac{n + 1}{n R^{2/m}} \right) + O(1/n)
$$

$$
= -R^{-2/m} \left( \log(R^{2/m}) + 1 \right) + O(1/n)
$$

for $n$ sufficiently large. Since $\gamma_{\text{loc}} = \emptyset$, we estimate (13) similar to (20), hence apply (16), (18), (24) to obtain

$$
\int_\tilde{\gamma} \int_L \exp \left[ nmR^{2/m} \left( \text{Re}F(s) - \text{Re}F(t) \right) \right] \left| \frac{t}{s} \right|^{m/2} \frac{\left| \cot(\pi n R^{2/m} t) \right|}{|t - s|} dsdt
\lesssim \int_\tilde{\gamma} \int_L \exp \left[ -nm(R^{2/m} - 1 - \log(R^{2/m})) + O(1) \right] \frac{1}{|s|^{m/2} \left| \text{Im}(s) \right|} dsdt
\lesssim \exp \left[ -2n(R - 1 - \log R) \right]
\leq \exp \left[ -(n + 1)^2 \right],
$$
where again Bernoulli’s inequality was used and the last inequality holds for $R > 1$. Ultimately all claims are proven. □

3. Matrices with independent entries

Let us begin with some necessary notations. We define the linearization matrix as the $mn \times mn$-block matrix

$$W := \frac{1}{\sqrt{n}} \begin{pmatrix} 0 & X^{(1)} & 0 & \cdots & 0 \\ \vdots & 0 & X^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 \\ X^{(m)} & 0 & \cdots & \cdots & 0 \end{pmatrix}$$

and note that $W^m$ is a block diagonal matrix with cyclic products of $X^{(1)}, \ldots, X^{(m)}$. Consequently, its eigenvalues consist of the eigenvalues $\lambda_j(X)$ of $X$ with multiplicity $m$. Furthermore define its shifted Hermitization

$$V(z) := \begin{pmatrix} 0 & W - z \\ (W - z)^* & 0 \end{pmatrix}$$

for $z \in \mathbb{C}$. The eigenvalues of $V(z)$ are given by $\pm s_j(W - z)$, where $s_{\max} = s_1 \geq \cdots \geq s_{mn} = s_{\min}$ are the singular values of $W - z$. Its ESD shall be denoted as $\tilde{\nu}_n^z$.

Similar to the role of the Stieltjes transform in the theory of Hermitian random matrices, the weak topology of measures $\mu$ on $\mathbb{C}$ can be expressed in terms of the so-called logarithmic potential $U$, which is the solution of the distributional Poisson equation. More precisely for every finite Radon measure $\mu$ on $\mathbb{C}$ the logarithmic potential defined by

$$U_\mu(z) := -\int_{\mathbb{C}} \log |t - z| \, d\mu(t) = (-\log |\cdot| * \mu)(z) \quad \text{satisfies} \quad \Delta U = -2\pi \mu$$

in the sense of distributions. Let $U_n$ denote the logarithmic potential of the ESD of $W$. The advantage of the logarithmic potentials $U_n$ of $\mu_n$ in non-Hermitian random matrix theory is the following identity known as Girko’s Hermitization trick

$$U_n(z) = -\frac{1}{nm} \sum_{j=1}^{nm} \log |\lambda_j(W) - z| = -\frac{1}{2nm} \log |\det V(z)| = -\int_{-\infty}^{\infty} \log |x| \, d\tilde{\nu}_n^z(x).$$

Under the above-mentioned conditions on the matrix entries, the logarithmic potential $U_n$ concentrates around the logarithmic potential $U_\infty$ of the Circular Law given by

$$U_\infty(z) = \begin{cases} -\log |z| & \text{if } |z| > 1 \\ \frac{1}{2}(1 - |z|^2) & \text{if } |z| \leq 1. \end{cases}$$

**Proposition 10 ([GNT17]).** If $X$ obeys $(C)$, then for every $\tau, Q > 0$ there exist a constant $c > 0$ such that

$$\mathbb{P} \left( |U_n(z) - U_\infty(z)| \leq c \frac{\log^4 n}{n} \right) \geq 1 - n^{-Q}$$

holds uniformly in $\{z \in B_{1+\tau^{-1}} : |1 - |z|| \geq \tau\}$. 
We remark that the restriction to the bulk in Theorem 5 comes directly from the restriction in the previous Proposition 10.

Since this is not explicitly worked out in [GNT17], we will derive it now, based on the results proved in this paper. We will directly follow the approach of [GNT17], making use of Girko’s Hermitization trick to convert the non-Hermitian problem into a Hermitian one, apply the local Stieltjes transform estimate from [GNT17] and the smoothing inequality from [GT03]. Let \( \tilde{\nu}_n \) be the symmetrized empirical singular value distribution of the shifted linearized matrices \( W - z \), defined in (27) and

\[
m_n(z, \cdot) : \mathbb{C} \setminus \mathbb{R} \to \mathbb{C}, w \mapsto \int_{\mathbb{R}} \frac{1}{w - t} d\tilde{\nu}_n^z(t)
\]

be the Stieltjes transform which converges a.s. to the solution of

\[
s(z, w) = -\frac{w + s(z, w)}{(w + s(z, w))^2 - |z|^2}
\]

see for instance [GT10a]. It is known that \( s(z, \cdot) \) corresponds to a limiting measure \( \tilde{\nu}_z \) which has a symmetric bounded density \( \rho_z \) (the bound holds uniformly in \( z \)) and has compact support. Note that \( s \) will be unbounded for \( z \) close to the edge, which is the reason for the bulk constraint of Proposition 10.

**Proof of Proposition 10.** Fix some arbitrary \( Q, \tau > 0 \) and \( z \in B_{1+\tau^{-1}} \) satisfying \( |1 - |z|| \geq \tau \). As is explained in Girko’s Hermitization trick (27),

\[
|U_n(z) - U_\infty(z)| = \left| \int_{\mathbb{R}} \log |x| d(\tilde{\nu}_n^z - \tilde{\nu}^z)(x) \right|
\]

and therefore it is necessary to estimate the extremal singular values as well as the rate of convergence of \( \tilde{\nu}_n^z \) to \( \tilde{\nu}^z \) in Kolmogorov distance

\[
d_n^*(z) = \sup_{x \in \mathbb{R}} |(\tilde{\nu}_n^z - \nu_n^z)(-\infty, x)|.
\]

Introduce the events

\[
\Omega_0 := \{s_{\text{min}} \geq n^{-B}\}, \quad \Omega_1 := \{s_{\text{max}} \leq n^{B'}\}, \quad \Omega_2 := \{d_n^*(z) \leq c \log^3 n/n\}
\]

for some constants \( B, B', c > 0 \) yet to be chosen. Theorem 31 in [OS11] states that there exists a constant \( B > 0 \) such that \( P(\Omega_0) \lesssim n^{-Q} \) and analogously to what has been shown in (35) there exists a constants \( B' > 0 \) with \( P(\Omega_1) \lesssim n^{-Q} \). Since \( \tilde{\nu}^z \) has a bounded density, we get

\[
\left| \int_{-n^{-B}}^{n^{-B}} \log |x| d\tilde{\nu}^z(x) \right| \lesssim \frac{\log n}{n^B}
\]

and furthermore on \( \Omega_2 \) it holds that

\[
\left| \int_{n^{-B} \leq |x| \leq n^{B'}} \log |x| d(\tilde{\nu}_n^z - \tilde{\nu}^z)(x) \right| \lesssim d_n^*(z) \log n \lesssim \frac{\log^4 n}{n}.
\]

Hence the claimed concentration of \( U_n \) holds on \( \Omega_0 \cap \Omega_1 \cap \Omega_2 \), implying

\[
P \left( |U_n(z) - U_\infty(z)| \geq c \frac{\log^4 n}{n} \right) \leq P(\Omega_0) + P(\Omega_1) + P(\Omega_2)
\]
and it remains to check $\mathbb{P}(\Omega_2^n) \leq n^{-Q}$, which has been done explicitly in [GNT17], (4.14)-(4.16) using the smoothing inequality [Corollary B.3] from [GT03] and the local law for $d_n^*(z)$ in terms of their Stieltjes transforms.

The core of the proof of the local law for products of non-Hermitian matrices in [GNT17] is the following identity. First, for any function $f \in C^2_c(\mathbb{C})$ define $\bar{f}$ by $\bar{f}(z) = f(z^*)$ and note that $\int f d\mu_n^m = \int \bar{f} d\mu^1_\infty$, which follows from definition of $\mu_n^m$ in (1). Using the distributional Poisson equation (26) as usual in non-Hermitian random matrix theory and the representation of the eigenvalues of $W$, we get

$$
\int f d(\mu_n^m - \mu^m_\infty) = \frac{1}{nm} \sum_{j=1}^m \bar{f}(\lambda_j(W)) - \int \bar{f} d\mu^1_\infty = -\frac{1}{2\pi} \int \Delta \bar{f} (U_n - U_\infty) d\lambda.
$$

(30)

As was done in [GJ18], we would like to uniformly approximate all indicator functions $\mathbb{1}_{B_n(z_0)}$ by smooth functions, replace the right hand side of (30) by a discrete random sum and use the pointwise estimate from Proposition 10. In contrast to what was done in [GJ18], we cannot use the smoothing inequality, since there is no control of the difference of logarithmic potentials of $\mu_n^m$, but of the matrix $W$. Therefore we will use a direct approach. In this case, we can replace the Monte Carlo approximation of the integral (30) by a random grid approximation. This also yields a more precise rate of convergence in Theorem 5.

Lemma 11 (Random grid approximation). Let $\alpha, \beta > 0$, $S$ be a random variable uniformly distributed on $[0, 1]^2$, and $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ be fixed with corresponding logarithmic potential $U$. Define the random grid $A = 2\beta n^{-\alpha/2}(\mathbb{Z}^2 + S) \cap [-\beta, \beta]^2$ enumerated by $z_1, \ldots, z_{[n^\alpha]}$. For any function $f \in C^3_c(\mathbb{C})$ with supp $f \subseteq (-\beta, \beta)^2$ it holds

$$
\frac{1}{n} \sum_{j=1}^n f(\lambda_j) - \frac{2\beta^2}{n^\alpha \pi} \sum_{i=1}^{[n^\alpha]} \Delta f(z_i) U(z_i)
$$

(31)

$$
= \mathcal{O}(\|\nabla \Delta f\|_\infty n^{-\alpha/2}) + \mathcal{O}(\|\Delta f\|_\infty \log(n)^2 n^{-\alpha/4}).
$$

with overwhelming probability. More precisely if $S$ is chosen independently of the random matrix elements, then (31) and

$$
\sup_i |U(z_i)| = \mathcal{O}(\log^2 n)
$$

holds on an event $\Omega_*$ of probability $1 - \mathcal{O}(n^{-\log n})$ that does not depend on $f$.

Note that (31) holds uniformly in $f \in C^3_c(\mathbb{C})$, hence we could choose a function dependent on the positions of $\lambda_j$. In order to make the statement more intuitive, suppose we replace the logarithmic potential $U$ by a more regular function $U \in C^1$. Then (31) is nothing but Riemann approximation of the integral

$$
\int \Delta f(z) U(z) d\lambda(z) - \frac{(2\beta)^2}{n^\alpha} \sum_{i=1}^{[n^\alpha]} \Delta f(z_i) U(z_i) \lesssim (\|\nabla \Delta f\|_\infty + \|\Delta f\|_\infty) n^{-\alpha/2}.
$$

(32)

This follows directly from the mean value theorem, very similar to what we will do in (34) below.
In the Monte Carlo approximation used in [TV15] and [KOV18], the random points \(z_i\) are not ordered in a grid but drawn independently, thus variance bounds are of importance for improving bounds as (31). By using reference points or eigenvalue rigidity, the error estimates in [TV15] and [KOV18] are stronger by a factor of \(1/n\) for the same amount of points \(z_i\). On the other hand, in order to control the singularities of \(U_n\), one has to handle many random effects of all \(z_i\), whereas in (31) only a single random shift effects all points \(z_i\). Heuristically speaking, this leads to a higher probability then in previous approaches, so that the weaker error bound is negligible.

**Proof of Lemma 11.** Using the (26), in other words integration by parts, we find

\[
\frac{1}{n} \sum_{j=1}^{n} f(\lambda_j) = \int f(d\mu_n) = -\frac{1}{2\pi} \int \Delta f(z) U(z) dz.
\]

It suffices to show that with probability at least \(1 - n^{-\log n - 1}\) we have

\[
\int \Delta f(z) \log |\lambda - z| dz - \frac{4\beta^2}{\alpha^2} \sum_{i=1}^{[n\alpha]} \Delta f(z_i) \log |\lambda - z_i| = \mathcal{O}(\|\nabla \Delta f\|_\infty n^{-\alpha/2}) + \mathcal{O}(\|\Delta f\|_\infty \log(n) n^{-\alpha/4})
\]

for fixed \(\lambda \in \mathbb{C}\), since the claim then follows from freezing the eigenvalues, i.e. conditioning on \(X\), summation and the union bound. The event, where (33) holds, will be \(\bigcap_{z_i \in A} \{|z_i - \lambda| > 2\beta n^{-(\log n + 1 + \alpha)/2}\}\) which fails if \(S = n^{-(\log n - 1)/2}\) close to \(\lambda\) shifted by the grid. More precisely let \(z^* \in 2\beta n^{-\alpha/2} \mathbb{Z}^2 \cap [-\beta, \beta]^2\) be the corner of this box with \(\lambda \in z^* + [0, 2\beta n^{-\alpha/2}]^2\), then

\[
\mathbb{P} \left( \exists i = 1, \ldots, [n\alpha] : |z_i - \lambda| \leq 2\beta n^{-(\log n + 1 + \alpha)/2} \right)
= \mathbb{P} \left( \text{dist}(S, \frac{\lambda - z^*}{2\beta n^{-\alpha/2}}) \leq n^{-(\log n - 1)/2} \right) = \mathcal{O}(n^{-\log n - n}),
\]

where the distance in \([0, 1]^2\) is measured according to the metric of the quotient space \(\mathbb{T}^2\). From now on we will restrict ourselves to this event. Rewrite (33) as

\[
\sum_{i=1}^{[n\alpha]} \int_{K_i} \Delta f(z) \log |\lambda - z| - \Delta f(z_i) \log |\lambda - z_i| dz,
\]

where we denoted the boxes with corner \(z_i\) by \(K_i = z_i + [0, 2\beta n^{-\alpha/2}]^2\). Adding and removing \(\Delta f(z_i) \log |\lambda - z|\), we obtain one error of order

\[
\sum_{i=1}^{[n\alpha]} \int_{K_i} (\Delta f(z) - \Delta f(z_i)) \log |\lambda - z| dz = \mathcal{O}(\|\nabla \Delta f\|_\infty n^{-\alpha/2}),
\]

where we used the mean value theorem and local integrability of \(\log \) in \(\mathbb{C}\). The second term can be bounded by

\[
\sum_{i=1}^{[n\alpha]} \int_{K_i} \Delta f(z_i) (\log |\lambda - z| - \log |\lambda - z_i|) dz
\leq \|\Delta f\|_\infty \left( \sum_{i:|z_i - \lambda| \geq n^{-\frac{\alpha}{2}}} + \sum_{i:|z_i - \lambda| < n^{-\frac{\alpha}{2}}} \right) \int_{K_i} (\log |\lambda - z| - \log |\lambda - z_i|) dz.
\]
Applying the mean value theorem for log, yields a bound of order \( O(n^{-\alpha/4}) \) for the first sum. The second sum can be bounded by performing the integration

\[
\int_0^{2n^{-\alpha/4}} r \log r + \sup_{z \in A} \log |\lambda - z_i| \, dr = O(n^{-\alpha/2} \log n) + O(n^{-\alpha/4} \log(n)^2),
\]

where we finally used the prescribed event. Putting all estimates together proves the first claim.

The bound for \( U \) follows from the choice of \( A \) and a trivial upper bound on \( |\lambda|_{\text{max}} \). On the one hand \( |\lambda|_{\text{max}} \) is bounded by the largest singular value \( s_{\text{max}} \) and on the other hand we have

\[
\mathbb{P}(s_{\text{max}} \geq n^{\log n}) \leq \frac{1}{n^{2 \log n}} \mathbb{E} \left[ |X/\sqrt{n}|^2 \right] \leq \frac{1}{n^{2 \log n + 1}} \sum_{i,j} \mathbb{E} |X_{ij}|^2 \leq n^{-2 \log n + 1},
\]

where the operator norm \( \|\cdot\| \) has been estimated by the Hilbert Schmidt norm. Therefore on an event \( \Omega_\star \) with probability at least \( 1 - O(n^{-\log n}) \), we have

\[
\sup_{z \in A} |U(z_i)| \leq \frac{1}{n} \sum_{j=1}^n \sup_{z \in A} |\log \lambda_j - z_i| \lesssim (\log n + 1 + \alpha) \log n + \log |\lambda|_{\text{max}} + 5 = O(\log^2 n).
\]

Ultimately we turn to the

**Proof of Theorem 5.** First, note that we only need to consider \( \tau < 1 \) and a bounded region, say \( V = B_\tau(0) \), i.e.

\[
D^m(\mu_n^m, \mu_\infty^m) = \sup_{B_R(z_0) \subseteq \mathbb{C} \setminus B_{1+\tau}} |(\mu_n^m - \mu_\infty^m)(B_R(z_0))| \leq \sup_{B_R(z_0) \subseteq \mathbb{C} \setminus B_{1+\tau}} |(\mu_n^m - \mu_\infty^m)(B_R(z_0) \cap V)| + \mu_n^m(V^c). \tag{36}
\]

But \( \mu_n^m(V^c) = (\mu_\infty^m - \mu_n^m)(V^c) \) is again bounded by 3 times (say) the left hand side of (36), hence we only have to estimate this first term. Fix some \( \tau, R > 0, z_0 \in \mathbb{C} \) such that \( B_R(z_0) \subseteq B_{1-\tau} \cup B_{1+\tau} \).

Let \( \varphi \in C^\infty(\mathbb{R}) \) be nonnegative with \( \text{supp} \varphi \subseteq [-1,1] \) and \( \int \varphi = 1 \), and define \( \varphi_a(\rho) = a \varphi(a \rho) \) for some \( a > 1 \) to be determined later. We mollify the indicator function appearing in (36) via the approximation

\[
f_1(z) := \left( 1_{(-\infty, R-1/a]} * \varphi_a \right)(|z - z_0|) \cdot \left( 1_{(-\infty, 7-1/a]} * \varphi_a \right)(|z|).
\]

where we choose \( f_1 \equiv 0 \) if \( R \leq 2/a \) for smoothness reasons. We apply \( f_1 \leq 1_{B_R(z_0) \cap V} \) and integration by parts to \( \widetilde{f}_1 : z \mapsto f_1(z^m) \) as was forecast in (30) to
obtain
\[
\mu^m_n((B_R(z_0) \cap V) \geq \int f_1 d\mu^m_n = -\frac{1}{2\pi} \int (\Delta \tilde{f}_1) U_n d\lambda
\]
\[
= -\frac{1}{2\pi} \int \Delta \tilde{f}_1(U_n - U_\infty) d\lambda - \int (\mathbb{1}_{B_R(z_0) \cap V} - f_1) d\mu^m_\infty + \int \mathbb{1}_{B_R(z_0) \cap V} d\mu^m_\infty \quad (37)
\]
where \( \lambda \) denotes the Lebesgue measure on \( \mathbb{C} \). A rough estimate of the error of approximation yields for the second term
\[
\int (\mathbb{1}_{B_R(z_0) \cap V} - f_1) d\mu^m_\infty \leq \mu^m_\infty (z \in \mathbb{C} : R = 2/a \leq |z - z_0| \leq R) \quad (38)
\]
Due to the radial monotonicity of \( \mu^m_\infty \)'s density, this value increases by bending two halves of the given annulus of width \( 2/a \) into two straight rectangles \([-1, 1] \times [-4/a, 4/a]\). The density is bounded in the case of \( m = 1 \) or for Corollary 6, where the origin is avoided, and hence the term in (38) is of order \( O(1/a) \). In general we can bound it by
\[
\mu^m_\infty([-1, 1] \times [-4/a, 4/a]) \leq \frac{1}{\pi m} \int_{-1}^{1} \int_{-4/a}^{4/a} (x^2 + y^2)^{1/m-1} dx dy
\]
\[
\leq \frac{4}{\pi ma} \int_{-a/2}^{a/2} x^2 dx + \frac{2}{m} \int_{0}^{a/2} 2 x^2/m-1 dx
\]
\[
= \frac{2}{\pi(m - 2)} \left( \frac{2}{a^2} 2/m - 2/a \right) + \frac{4}{a^2} 2/m \leq a^{-2/m},
\]
where the equation only holds for \( m > 2 \). For \( m = 2 \) we get
\[
\mu^2_\infty([-1, 1] \times [-4/a, 4/a]) \leq \frac{1}{2\pi} \int_{-1}^{1} \int_{-4/a}^{4/a} (x^2 + y^2)^{1/2} dx dy
\]
\[
= \frac{8}{\pi a} \log \left( \sqrt{1 + 16/a^2} + 1 \right) + \frac{2}{\pi} \log \left( \sqrt{1 + 16/a^2} + 4/a \right) + \frac{8}{\pi a} \log(a/4)
\]
\[
\sim a^{-1} \log a
\]
and we see that the log-term appears naturally. Define the error function
\[
\tilde{h}_m(a) = \begin{cases} O(a^{-1}) & \text{for } m = 1 \text{ or Corollary 6}, \\ O(a^{-1/2} \log a) & \text{for } m = 2, \\ O(a^{-2/m}) & \text{for } m \geq 3. \end{cases}
\]
Let us continue to estimate the first term of (37) by using our random grid approximation Lemma 11. Let \( \beta = 7 \) and \( S \) be a random variable, independent of \( X \) and uniformly distributed on \([0, 1]^2\). Conditioned on \( X \), we have with overwhelming probability
\[
\int \Delta \tilde{f}_1(z) U_n(z) d\lambda(z) - \frac{(2\beta)^2}{n^a} \sum_{i=1}^{[n^a]} \Delta \tilde{f}_1(z_i) U_n(z_i)
\]
\[
= O(\| \nabla \tilde{f}_1 \|_\infty n^{-a/2}) + O(\| \Delta \tilde{f}_1 \|_\infty \log(n) n^{-a/4}).
\]
Due to our explicit choice of functions \( f_1 \) and \( f_2 \) as product of shifted radial symmetric functions, the partial derivatives become fairly simple. Each derivative that
hits one of the $\varphi_a$ produces a factor of $a$, more precisely any $k$-th directional derivative satisfies $\|\partial^{(k)} f_1(z)\|_\infty \lesssim a^k$. This estimate, again, is independent on the choice of the ball $B_R(z_0)$.

Together with the Riemann approximation (32), we conclude that for any matrix $X$ we have with overwhelming probability

$$\sup_B \left| \int \Delta \tilde{f}_1(z)(U_n(z) - U_\infty(z))d\lambda(z) - \frac{(2\beta)^2}{2\pi n^\alpha} \sum_{i=1}^{[n^n]} \Delta \tilde{f}_1(z_i)(U_n(z_i) - U_\infty(z_i)) \right| = O(a^3 n^{-\alpha/2}) + O(a^2 \log(n)^2 n^{-\alpha/4}),$$

where the supremum runs over all choices of $B \subseteq B_{1-\tau} \cup B_{1+\tau}$. Since we will always choose $a \lesssim n$ (actually we will make it even smaller, cf. (40)), it is possible to freely choose $\alpha > 0$ sufficiently big such that the error is arbitrarily small. For instance $\alpha = 13$ is more than enough to ensure that the error is of order $O(n^{-1})$. It should be emphasized that still no randomness of $X$ has been used and the only randomness is the shifted grid. Combining the previous steps yield

$$(\mu_n^m - \mu_\infty^m)(B \cap V) \geq -\frac{(2\beta)^2}{2\pi n^\alpha} \sum_{i=1}^{[n^n]} \Delta \tilde{f}_1(z_i)(U_n(z_i) - U_\infty(z_i)) - \tilde{h}_m(a) - O(n^{-1})$$

(39)

uniformly in $B \subseteq B_{1-\tau} \cup B_{1+\tau}$ with overwhelming probability. Noting $\mu_n^m(B_R(z_0) \cap V) \leq \int d\mu_n^m$ and taking the same route for $f_2$ as for $f_1$, we obtain the same upper bound

$$(\mu_n^m - \mu_\infty^m)(B \cap V) \leq -\frac{(2\beta)^2}{2\pi n^\alpha} \sum_{i=1}^{[n^n]} \Delta \tilde{f}_2(z_i)(U_n(z_i) - U_\infty(z_i)) + \tilde{h}_m(a) + O(n^{-1}).$$

Finally we use the randomness of $X$ by applying Proposition 10. Conditioning on $S$, i.e. freezing the lattice points $z_i$, we obtain for any $Q > 0$

$$\mathbb{P}\left( |U_n(z_i) - U_\infty(z_i)| \geq c \frac{\log^4(n)}{n} |S| \right) \leq n^{-Q-\alpha}$$

for each $i = 1, \ldots, [n^n]$. By the union bound this implies that with probability at least $1 - n^{-Q}$ the logarithmic potentials concentrate like $U_n(z_i) - U_\infty(z_i) = O(\log^4 n/n)$ simultaneously at all lattice points. Therefore, for $k = 1, 2$

$$\frac{(2\beta)^2}{2\pi n^\alpha} \sum_{i=1}^{[n^n]} \Delta \tilde{f}_k(z_i)(U_n(z_i) - U_\infty(z_i)) \leq \frac{(2\beta)^2 \log^4 n}{n^{1+\alpha}} \sum_{i=1}^{[n^n]} |\Delta \tilde{f}_k(z_i)|$$

$$= \frac{\log^4 n}{n} \|\Delta \tilde{f}_k\|_{L^1} + O\left( a^3 \frac{\log^4}{n^{\alpha/2+1}} \right)$$

where the integral of the $a^3$-Lipschitz function $|\Delta \tilde{f}_k|$ has been approximated by its Riemann sum. Write $\Delta = 4\partial \bar{\partial}$ in terms of the Wirtinger derivatives $\partial = \frac{1}{i}(\partial_x - i\partial_y)$ and $\bar{\partial} = \frac{1}{i}(\partial_x + i\partial_y)$. Since $g(z) = z^n$ is holomorphic, i.e. $\partial g = 0$, we obtain by applying the chain rule and changing variables from $z$ to $g(z)$

$$\|\Delta f_k\|_{L^1} = \|4\partial \bar{\partial}(f_k \circ g)\|_{L^1} = 4\|\partial \partial f_k \circ g \cdot \partial \bar{\partial} \circ g\|_{L^1} = 4\|\Delta f_k\|_{L^1}.$$
Since $\Delta f_k \lesssim a^2$ and has support on an area of order $a^{-1}$, we have

$$\sup_B \|\Delta f_k\|_{L^1} \lesssim a.$$  

So overall we have proven that for all $Q$ there exists a constant $c > 0$ such that with $1 - n^{-Q}$-high probability we have

$$\sup_{B \subseteq B_{1-r} \cup B_{1+r}} |(\mu_m^m - \mu_{\infty}^\infty)(B)| \leq c a \frac{\log^4 n}{n} + \tilde{h}_m(a) + O(n^{-1}).$$

Optimizing in $a$ yields $a = \sqrt{n}/\log^2 n$ for Corollary 6 and $m = 1$, as well as $h_2(n) = \log^3 n/\sqrt{n}$. The asymptotic $h_m(n)$ for higher $m$ follows from choosing $a = n^{m/m+2} \log^{-4m/(m+2)} n$.  

\[\Box\]

In the proof we have seen that the maximal error for the limiting distribution $\mu_{\infty}^\infty$ is approached by infinitely big balls that touch the origin (technically these balls are not even admissible here). This yields the non-optimal rate in Theorem 5 if we do not exclude the origin. Having Theorem 3 in mind however, we expect the “worst ball” to appear roughly at $B_1(0)$, where the error would be optimal again.

**Acknowledgements**

Financial support by the German Research Foundation (DFG) through the IRTG 2235 is gratefully acknowledged. The author would like to thank Friedrich Götze for valuable comments and suggestions and Mario Kieburg for helpful discussions regarding the saddle-point analysis.

**References**

[AB12] Gernot Akemann and Zdzislaw Burda. Universal microscopic correlation functions for products of independent ginibre matrices. *Journal of Physics A: Mathematical and Theoretical*, 45(46), 2012.

[ABK19] Gernot Akemann, Zdzislaw Burda, and Mario Kieburg. From integrable to chaotic systems: Universal local statistics of lyapunov exponents. *EPL (Europhysics Letters)*, 126(4):40001, 2019.

[AC18] Gernot Akemann and Milan Cikovic. Products of random matrices from fixed trace and induced ginibre ensembles. *Journal of Physics A: Mathematical and Theoretical*, 51(18):184002, 2018.

[AGT10] NV Alexeev, Friedrich Götze, and AN Tikhomirov. On the singular spectrum of powers and products of random matrices. In *Doklady mathematics*, volume 82, pages 505–507. Springer, 2010.

[AI15] Gernot Akemann and Jesper R Ipsen. Recent exact and asymptotic results for products of independent random matrices. *arXiv preprint arXiv:1502.01667*, 2015.

[BJW10] Z Burda, RA Janik, and B Waclaw. Spectrum of the product of independent random gaussian matrices. *Physical Review E*, 81(4):041132, 2010.

[Bor11] Charles Bordenave. On the spectrum of sum and product of non-hermitian random matrices. *Electron. Commun. Probab.*, 16:104–113, 2011.

[CHM18] Djalil Chafaï, Adrien Hardy, and Mylène Maïda. Concentration for Coulomb gases and Coulomb transport inequalities. *J. Funct. Anal.*, 275(6):1447–1483, 2018.

[FW17] Peter J. Forrester and Dong Wang. Muttalib-borodin ensembles in random matrix theory - realisations and correlation functions. *Electron. J. Probab.*, 22:43 pp., 2017.

[Gin65] Jean Ginibre. Statistical ensembles of complex, quaternion, and real matrices. *J. Mathematical Phys.*, 6:440–449, 1965.

[GJ18] Friedrich Götze and Jonas Jalowy. Rate of convergence to the circular law via smoothing inequalities for log-potentials. *arXiv preprint arXiv:1807.00489*, 2018.
[GKT15] Friedrich Götze, Holger Kösters, and Alexander Tikhomirov. Asymptotic spectra of matrix-valued functions of independent random matrices and free probability. *Random Matrices: Theory and Applications*, 4(02):1550005, 2015.

[GNT17] Friedrich Götze, Alexey Naumov, and Alexander Tikhomirov. On local laws for non-hermitian random matrices and their products. *Random Matrices: Theory and Applications*, 2017. accepted.

[GT03] Friedrich Götze and Alexander Tikhomirov. Rate of convergence to the semi-circular law. *Probability Theory and Related Fields*, 127(2):228–276, 2003.

[GT10a] Friedrich Götze and Alexander Tikhomirov. The circular law for random matrices. *The Annals of Probability*, 38(4):1444–1491, 2010.

[GT10b] Friedrich Götze and Alexander Tikhomirov. On the asymptotic spectrum of products of independent random matrices. *arXiv preprint arXiv:1012.2710*, 2010.

[IK14] Jesper R. Ipsen and Mario Kieburg. Weak commutation relations and eigenvalue statistics for products of rectangular random matrices. *Phys. Rev. E*, 89:032106, 3 2014.

[KKS15] Mario Kieburg, Arno B.J. Kuijlaars, and Dries Stivigny. Singular Value Statistics of Matrix Products with Truncated Unitary Matrices. *International Mathematics Research Notices*, 2016(11):3392–3424, 08 2015.

[KOV18] Phil Kopel, Sean O’Rourke, and Van Vu. Random matrix products: Universality and least singular values. *arXiv preprint arXiv:1802.03064*, 2018.

[KT15] Holger Kösters and Alexander Tikhomirov. Limiting spectral distributions of sums of products of non-hermitian random matrices. *arXiv preprint arXiv:1506.04436*, 2015.

[KZ14] Arno B.J. Kuijlaars and Lun Zhang. Singular values of products of g
dibre random matrices, multiple orthogonal polynomials and hard edge scaling limits. *Communications in Mathematical Physics*, 332(2):759–781, 2014.

[Luk14] Yudell L. Luke. *Mathematical functions and their approximations*. Academic Press, 2014.

[LWZ+16] Dang-Zheng Liu, Dong Wang, Lun Zhang, et al. Bulk and soft-edge universality for singular values of products of g
dibre random matrices. In *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, volume 52, pages 1734–1762. Institut Henri Poincaré, 2016.

[Nem17] Yuriy Nemish. Local law for the product of independent non-hermitian random matrices with independent entries. *Electron. J. Probab.*, 22:35 pp., 2017.

[Nem18] Yuriy Nemish. No outliers in the spectrum of the product of independent non-hermitian random matrices with independent entries. *Journal of Theoretical Probability*, 31(1):402–444, 03 2018.

[OS11] Sean O’Rourke and Alexander Soshnikov. Products of independent non-hermitian random matrices. *Electronic Journal of Probability*, 16:2219–2245, 2011.

[Ser17] Sylvia Serfaty. Microscopic description of log and coulomb gases. *arXiv preprint arXiv:1709.01089*, 2017.

[TV15] Terence Tao and Van Vu. Random matrices: universality of local spectral statistics of non-hermitian matrices. *The Annals of Probability*, 43(2):782–874, 2015.

[WW96] Edmund Taylor Whittaker and George Neville Watson. *A course of modern analysis*. Cambridge university press, 1996.

*Jonas Jalowy, Faculty of Mathematics, Bielefeld University, Germany
E-mail address: jjalowy@math.uni-bielefeld.de*