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 $\mathcal{O}(1/\sqrt{n})$, which is attained with overwhelming probability up to a logarithmic correction. 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 $m = 1$, corresponding to the Circular Law, was studied in [16] already.

We consider the product

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

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

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

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

Definition 1.1.   *(i) A (complex)* Ginibre matrix X *is a complex non-Hermitian random matrix with independent complex Gaussian entries* $X_{ij} \sim \mathcal{N}_{\mathbb{C}}(0, 1)$.

   (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 $\operatorname{Re} X_{ij}$ and $\operatorname{Im} X_{ij}$ 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 λ the 2-dimensional Lebesgue measure, by $\Rightarrow$ weak convergence of measures and $B_r = B_r(0)$ shall be the open ball of radius $r > 0$ centered at $z = 0$. In [21], Götze and Tikhomirov showed that as $n \to \infty$, $\mathbb{P}$- almost surely we have

$$\mu_n^m \Rightarrow \mu_\infty^m, \text{ where } d\mu_\infty^m(z) = \frac{|z|^{2/m-2}}{\pi m} \mathbb{1}_{B_1}(z) d\lambda(z) \tag{1.1}$$

is the m-th power of the uniform distribution $\mu_\infty = \mu_\infty^1$ on the complex disc, see also [33]. The Gaussian case has been treated in [10, 2], more general models can be found in [25, 17, 7, 1, 22], for the convergence of the singular values see [5] and furthermore for local results we refer to [31, 24, 32, 18, 12].

For $m = 1$, we retrieve the well known *Circular Law* $\mu_n = \mu_n^1 \Rightarrow \mu_\infty$. In the case of Ginibre matrices this has been discovered much earlier in [15]. We are interested in the rate of convergence, more precisely in a uniform Kolmogorov-type distance over balls

$$D(\mu_n^m, \mu_\infty^m) := \sup |\mu_n^m(B_R(z_0)) - \mu_\infty^m(B_R(z_0))|$$

as $n \to \infty$, where the supremum runs over all balls $B_R(z_0) \subseteq \mathbb{C}$. In the sequel, we will also consider the supremum over certain families of balls B. Convergence in the distance D coincides with weak convergence in the case of an absolutely continuous limit distribution, see [16, Lemma A.1]. Using the rotational symmetry of the Circular Law μ_∞ and the mean empirical spectral distribution $\bar{\mu}_n = \mathbb{E}\mu_n$ of the Ginibre ensemble, the following optimal rate of μ_n for $m = 1$ was shown in [16].

Lemma 1.2. *The mean empirical spectral distribution $\bar{\mu}_n = \mathbb{E}\mu_n$ of the Ginibre ensemble satisfies*

$$D(\bar{\mu}_n, \mu_\infty) \sim \frac{1}{\sqrt{2\pi n}} \tag{1.2}$$

and for any fixed $\varepsilon > 0$

$$\sup_{\substack{B_R(z_0) \subseteq \mathbb{C} \setminus B_{1+\varepsilon} \\ \text{or } B_R(z_0) \subseteq B_{1-\varepsilon}}} |\bar{\mu}_n(B_R(z_0)) - \mu_\infty(B_R(z_0))| \lesssim e^{-n\varepsilon^2}. \tag{1.3}$$

Here and in the sequel we denote asymptotic equivalence by $\sim$. We write $\lesssim$, if an inequality holds up to a n-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 $\log^a b = (\log b)^a$.

1.1 The optimal rate of convergence for products of Ginibre matrices

It is natural to ask for a generalization of Lemma 1.2 to products of $m \geq 1$ independent Ginibre matrices. Notably, our analogous result shows that the optimal rate of convergence does not depend on the number m.

Theorem 1.3. *The mean empirical spectral distribution $\bar{\mu}_n^m = \mathbb{E}\, \mu_n^m$ of the Ginibre ensemble satisfies*

$$\sup_{R>0} |\bar{\mu}_n^m(B_R) - \mu_\infty^m(B_R)| \asymp \frac{1}{\sqrt{nm}}. \tag{1.4}$$

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

$$\sup_{R<1-\frac{m}{2}\sqrt{\log n/n}} |\bar{\mu}_n^m(B_R) - \mu_\infty^m(B_R)| \lesssim \frac{\log^{3/2} n}{n} \tag{1.5}$$

and uniformly in $R > 1 + \sqrt{\log n/n}$ we have

$$|\bar{\mu}_n^m(B_R) - \mu_\infty^m(B_R)| \lesssim \exp\left[-n \min((R-1)^2, 1)/3\right]. \tag{1.6}$$

Theorem 1.3 provides the optimal rate of convergence, which however is faster inside and much faster outside of the bulk. The precise constants of the upper and lower bound of (1.4) can be chosen to be $C = \sqrt{\pi}/\sqrt{2}$ and $c = 1/(\sqrt{2\pi})$, coinciding with Lemma 1.2. The constants in (1.5) and (1.6) do depend on m. We will also see that the maximal distance in (1.4) is attained at $R = 1$.

The restriction of Theorem 1.3 to centered balls does not affect its character of quantifying weak convergence: Since both $\bar{\mu}_n^m$ and μ_∞^m have radial symmetric Lebesgue densities, it is easy to see that (1.4) already implies weak convergence $\bar{\mu}_n^m \Rightarrow \mu_\infty^m$. In order to remove this restriction one may exploit monotonicity arguments of the radial part of $\bar{\mu}_n^m$.

While the proof of Lemma 1.2 is an elementary calculation, the proof of Theorem 1.3 is more involved and relies on a saddle-point method of a double contour integral representation for the density of $\bar{\mu}_n^m$, which we will give in Section 2. An idea of the proof is given after the contours are defined, see Figure 2. Figure 1 illustrates the statements of our main results.

Pointwise convergence of the density ρ_n^m of $\bar{\mu}_n$ or $\bar{\mu}_n^m$ has been also discussed in [2, 4, 35]. In particular Akemann and Burda describe the asymptotic

$$\rho_n^m(z) = \frac{|z|^{2/m-2}}{\pi m} \tfrac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{mn}{2}}\frac{|z|^{2/m}-1}{|z|^{1/m}}\right) + o(1)$$

for fixed z without specifying the error. Aside from the error $o(1)$, the appearance of $\mathrm{erfc}(\cdot)$ also hints at exponential convergence, like in (1.3) and (1.6). 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 Maïda studied invariant β-ensembles with external potential V instead of matrices with independent entries, see [11]. Their result implies a rate of convergence to the limiting measure with density $c\Delta V$ of order $\mathcal{O}(\sqrt{\log n/n})$ 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 [34]. Moreover, for determinantal point processes, the fluctuation around the (deterministic) rate has been studied in [13].
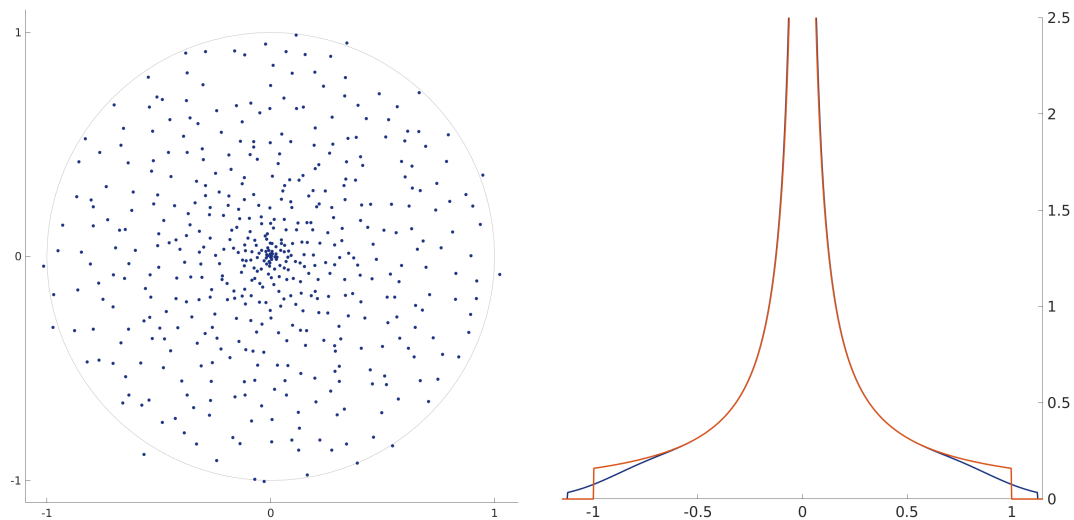
Figure 1: The empirical spectral distribution of the product $\mathbf{X}$ of $m = 2$ Ginibre matrices. Left: The eigenvalues of a sample for $n = 500$ and the unit ball $B_1(0)$ as reference. Theorem 1.6 below shows that gaps (like one can see at the top) and clusters do not significantly differ from 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 μ_∞^m in orange. Clearly the rate of convergence in the bulk is faster than close to the edge, illustrating the statement of Theorem 1.3.

1.2 Non-averaged rate of convergence for products of Ginibre matrices

Note that we cannot expect an exponentially fast rate of convergence like in Lemma 1.2 for the non-averaged empirical spectral distribution μ_n^m, because it is still affected by the individual eigenvalue fluctuations. In particular the rough lower bound $D(\mu_n^m, \mu_\infty^m) \gtrsim 1/n$ follows from placing a disk of radius $cn^{-1/2}$ contained in $B_1(0)$ such that it does not cover any eigenvalue. 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 $n^{-1/2}$ without covering a new eigenvalue. Hence we should also expect the non-averaged distance $D(\mu_n^m, \mu_\infty^m)$ to be of order $\mathcal{O}(n^{-1/2})$ with overwhelming probability.

Theorem 1.4. *For products of m Ginibre matrices and any $\varepsilon, Q > 0$ there exists a constant $c > 0$ such that*

$$\mathbb{P}\left(\sup_{R>0} |\mu_n^m(B_R) - \mu_\infty^m(B_R)| \leq c\sqrt{\frac{\log n}{n}}\right) \geq 1 - n^{-Q}.$$

This rate coincides exactly with the Wasserstein rate obtained in [11]. The proof makes use of the determinantal structure of the eigenvalues of Gaussian matrices in order to apply Bernstein's inequality and will be given in the end of Section 2.

1.3 Rate of convergence for products of non-Gaussian matrices

Based on the ideas of [16], we will also prove a rate of convergence result for products of matrices with independent entries.

Definition 1.5 (Condition (C))**.** *We say $\mathbf{X} = X^{(1)} \cdots X^{(m)} / \sqrt{n^m}$ satisfies condition (C) if*

the matrices $X^{(q)}$, $q = 1, \ldots, m$, have jointly independent entries and satisfy

$$\left| \mathbb{E}\, X_{ij}^{(q)} \right| \leq n^{-1-\varepsilon} \text{ and } \left| 1 - \mathbb{E} \left| X_{ij}^{(q)} \right|^2 \right| \leq n^{-1-\varepsilon}$$

for some $\varepsilon > 0$ independent of n and furthermore

$$\max_{i,j,q,n} \mathbb{E} \left| X_{ij}^{(q)} \right|^{4+\delta} < \infty$$

for some $\delta > 0$.

In Section 3, we will show

Theorem 1.6. *If condition (C) holds, then for every $m \in \mathbb{N}$ and $\tau, Q > 0$ there exists a constant $c > 0$ such that*

$$\mathbb{P} \left(\sup_B |(\mu_n^m - \mu_\infty^m)(B)| \leq c h_m(n) \right) \geq 1 - n^{-Q},$$

where the supremum runs over balls $B \subseteq B_{1-\tau} \cup B_{1+\tau}^c$ and the asymptotic error is given by

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

In the proof of Theorem 1.6 we will see that the m-dependent term is only visible for balls touching the origin and else only the error h_1 remains for all m. To make the statement more comprehensible when comparing with Theorem 1.3, we also state the following result.

Corollary 1.7. *If condition (C) holds, then for every $\tau, Q > 0$ we have*

$$\mathbb{P} \left(\sup_B |(\mu_n^m - \mu_\infty^m)(B)| \lesssim \frac{\log^2 n}{\sqrt{n}} \right) \geq 1 - n^{-Q},$$

where the supremum runs over all balls B such that $\partial B_R(z_0) \subseteq B_{1+\tau}^c \cup B_{1-\tau} \setminus B_\tau$ avoids the edge and the origin.

We already know from the previous discussion that the optimal rate is given by $\mathcal{O}(1/\sqrt{n})$. Corollary 1.7 shows that this rate is (almost) universal in the sense that it continues to hold for matrices with independent entries, if edge and origin are avoided. A weaker rate of convergence for $m = 1$ has already been established in [16, §2.2], where a comparison with similar results can be found as well. We would like to point out a subtle difference between Theorem 1.6 and the Local Law in [18]; 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.

2 Product of Ginibre matrices

We start with the following double contour integral representation for the density of μ_n^m that is essential for Theorem 1.3.

Lemma 2.1. *The density of $\bar{\mu}_n^m$ 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 γ is any closed contour that encircles the numbers $1, \ldots, n$ counter clockwise and no natural number greater than n.

In [26], a similar double contour integral representation for the correlation kernel of the singular values of $\mathbf{X}$ was derived. This was used in [27] to prove bulk universality for singular values of products of independent Ginibre matrices. In general, double contour integrals like in Lemma 2.1 above appear regularly in the theory of products of random matrices, e.g. [26, 14, 23].

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

$$G_{p,q}^{m,n}\left(\begin{array}{c} a_1, \ldots, a_p \\ b_1, \ldots, b_q \end{array}\middle|\xi\right) = \frac{1}{2\pi i}\int_L \frac{\prod_{j=1}^{m}\Gamma(b_j - t)\prod_{j=1}^{n}\Gamma(1 - a_j + t)}{\prod_{j=m+1}^{q}\Gamma(1 - b_j + t)\prod_{j=n+1}^{p}\Gamma(a_j - t)}\xi^t dt, \qquad (2.1)$$

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. Since the Gamma function is the Mellin transform of the exponential function, we have for instance $G_{0,1}^{1,0}\left(\begin{array}{c} - \\ 0 \end{array}\middle|\xi^2\right) = \exp(-\xi^2)$.

In particular the density of $\bar{\mu}_n^m$ is given by

$$\rho_n^m(z) = n^{m-1}\sum_{k=0}^{n-1} \frac{n^{mk}|z|^{2k}}{\pi(k!)^m}G_{0m}^{m0}\left(\begin{array}{c} - \\ 0 \end{array}\middle|n^m|z|^2\right), \qquad (2.2)$$

see [2] and compare to the case $m = 1$, where $G_{01}^{10}\left(\begin{array}{c} - \\ 0 \end{array}\middle|n|z|^2\right) = e^{-n|z|^2}$.

We would like to point out that the Coulomb gas picture partially breaks down here, since $V = -\log G_{0,m}^{m,0}\left(\begin{array}{c} - \\ 0 \end{array}\middle||\cdot|^2\right)$ is not admissible (according to [34]). In particular V is unbounded from below in its singularity and hence $\Delta V \sim \mu_\infty^m$ fails for finite n.

Remark 2.2. The viewpoint of studying products of m matrices and definition (2.2) of ρ_n^m makes sense for $m \in \mathbb{N}$ only. However the representation of Lemma 2.1 makes sense for arbitrary $m > 1, m \in \mathbb{R}$. Furthermore, as we can see from the proof of Theorem 1.3, its statements (1.4) and (1.6) remain true for real $m > 1$, as well as (1.5) for real $m \geq 2$.

Remark 2.3. Since the constant in (1.4) of Theorem 1.3 is explicit in m, it is possible to consider the double scaling limit and let $m = m(n) \to \infty$. A careful inspection of the proof reveals that at most $m = o(n)$ is possible. In this case, the rate will be faster, depending on m, and in particular by setting $m = n$, we expect a rate $\sup_{R>0}|\bar{\mu}_n^n(B_R) - \mu_\infty^n(B_R)| \asymp \frac{1}{n}$. On the other hand, μ_∞^m converges weakly to $\mu_\infty^\infty = \delta_0$ as $m \to \infty$ and for fixed $R > 0$ we see that $\delta_0(B_R) - \mu_\infty^m(B_R) = 1 - R^{2/m}/(2\pi) = \mathcal{O}(1/m)$. Note that $m \sim n$ is also the critical scaling of Lyapunov exponents between deterministic and GUE statistics, see [3, 28].

Proof of Lemma 2.1. For the contour L of the Meijer G-function in (2.2), 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

$$G_{0,m}^{m,0}\left(\begin{array}{c} - \\ 0 \end{array}\middle|n^m|z|^2\right) = \frac{1}{2\pi i}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty}\Gamma(s)^m(n^m|z|^2)^{-s}ds \qquad (2.3)$$

The remaining part of (2.2) 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 γ encircling the numbers $1, \ldots, n$ but no natural number greater than n, we have

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

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

Asymptotic expansions of $G_{0,m}^{m,0}$, like [29, §5.9.1.], together with heuristics for the hypergeometric kernel give rise to pointwise limits in [2]. A rigorous estimation of the error bound uniformly in z seems to be absent in the literature so far. Hence it is reasonable to study the problem by a direct analysis.

Proof of Theorem 1.3. By Lemma 1.2, it is sufficient to consider $m \geq 2$. For $R > 1$ we have $|(\bar\mu_n^m - \mu_\infty^m)(B_R)| < |(\bar\mu_n^m - \mu_\infty^m)(B_1)|$, since $\mathrm{supp}(\mu_\infty^m) = B_1(0)$. Throughout the proof we assume

$$\log^{3m/4} n / n^{m/2} \leq R \leq 1 \tag{2.4}$$

since for smaller values of R it holds

$$|(\bar\mu_n^m - \mu_\infty^m)(B_R(0))| \leq \left|(\bar\mu_n^m - \mu_\infty^m)(B_{\log^{3m/4} n/n^{m/2}}(0))\right| + \mathcal{O}(\log^{3/2} n/n),$$

due to $\mu_\infty^m(B_R) = R^{2/m}$. We first use spherical symmetry of ρ_n^m and Lemma 2.1 in order to calculate

$$
\begin{aligned}
\bar\mu_n^m(B_R) &= \int_0^{R^2} \pi \rho_n^m(\sqrt{r})dr \\
&= \frac{\pi}{n(2\pi i)^2} \oint_\gamma \int_{\frac12 - i\infty}^{\frac12 + i\infty} \left(\frac{\Gamma(s)}{\Gamma(t)}\right)^m \frac{(n^m R^2)^{t-s}}{t-s} \cot(\pi t) ds dt. \tag{2.5}
\end{aligned}
$$

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 2.1 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

$$\frac{\pi}{(2\pi i)^2} \left(\int_{\frac12 - i\infty}^{\frac12 + i\infty} - \int_L\right) \left(\frac{\Gamma(s)}{\Gamma(t)}\right)^m \frac{(n^m R^2)^{t-s}}{t-s} ds = \frac{\pi}{2\pi i} \mathbb{1}_{(1/2,\eta)}(\mathrm{Re}(t)).$$

We temporarily split γ into two parts γ_l and γ_r such that γ_l encircles $\{1, \ldots, \lfloor\eta\rfloor \wedge n\}$ and γ_r encircles $\{\lceil\eta\rceil, \ldots, n\}$. Soon we will make the path of γ more explicit. As in (2.5), continuing the integration of the right hand side of the last equation multiplied with $\cot(\pi t)$ over $\gamma_l \cup \gamma_r$ yields

$$\frac{\pi}{2\pi i} \oint_{\gamma_l} \cot(\pi t) dt = \lfloor\eta\rfloor \wedge n, \tag{2.6}$$

hence we conclude

$$\bar\mu_n^m(B_R) = \frac{\pi}{n(2\pi i)^2} \oint_\gamma \int_L \left(\frac{\Gamma(s)}{\Gamma(t)}\right)^m \frac{(n^m R^2)^{t-s}}{t-s} \cot(\pi t) ds dt + \frac{\lfloor\eta\rfloor}{n} \wedge 1.$$

Choosing $\eta = \lfloor nR^{2/m}\rfloor + 1/2$ we see that the second term is $\mathcal{O}(1/n)$ close to $\mu_\infty^m(B_R) = R^{2/m} \wedge 1$. Moreover, by Cauchy's integral formula, we may artificially add the removed part $\gamma - \gamma_l - \gamma_r$ again as long as γ is symmetric around the x-axis. Let γ 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 of the cotangent. The scaled version $\widetilde\gamma = \gamma/(R^{2/m}n)$ is illustrated below in Figure 2. Furthermore note that the integral exists as we will explicitly show below, see (2.18). Recall Stirling's formula for the Gamma function

$$\log\Gamma(z) = (z - 1/2)\log z - z + \frac12 \log 2\pi + \mathcal{O}(1/\mathrm{Re}\, z),$$

which holds uniformly for $\mathrm{Re}\, z \geq 1/2$, cf. for instance [36, p.249]. Thus, we have

$$\log\left(\frac{\Gamma(s)^m}{(n^m R^2)^s}\right) = m\left(s\left(\log\left(\frac{s}{nR^{2/m}}\right) - 1\right) + \frac{1}{2}\log\left(\frac{2\pi}{s}\right)\right) + \mathcal{O}(1/\mathrm{Re}(s)), \qquad (2.7)$$

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

$$F(z) = z \log z - z$$

to obtain

$$\bar{\mu}_n^m(B_R) - \frac{\lceil nR^{2/m}\rceil}{n}$$
$$= \frac{\pi R^{2/m}}{(2\pi i)^2} \oint_{\widetilde{\gamma}} \int_{\widetilde{L}} \left(\frac{\Gamma(nR^{2/m}s)}{\Gamma(nR^{2/m}t)}\right)^m \frac{(n^m R^2)^{(nR^{2/m})(t-s)}}{t-s} \cot(\pi nR^{2/m}t)\,ds\,dt$$
$$= \frac{\pi}{(2\pi i)^2} \oint_{\widetilde{\gamma}} \int_{\widetilde{L}} \exp\left[nmR^{2/m}\left(F(s)-F(t)\right)\right] \left(\frac{t}{s}\right)^{m/2} \frac{\cot(\pi nR^{2/m}t)}{t-s} \qquad (2.8)$$
$$\cdot \left(R^{2/m} + \mathcal{O}\left(\frac{1}{n\,\mathrm{Re}(s)}\right) + \mathcal{O}\left(\frac{1}{n\,\mathrm{Re}(t)}\right)\right)\,ds\,dt.$$
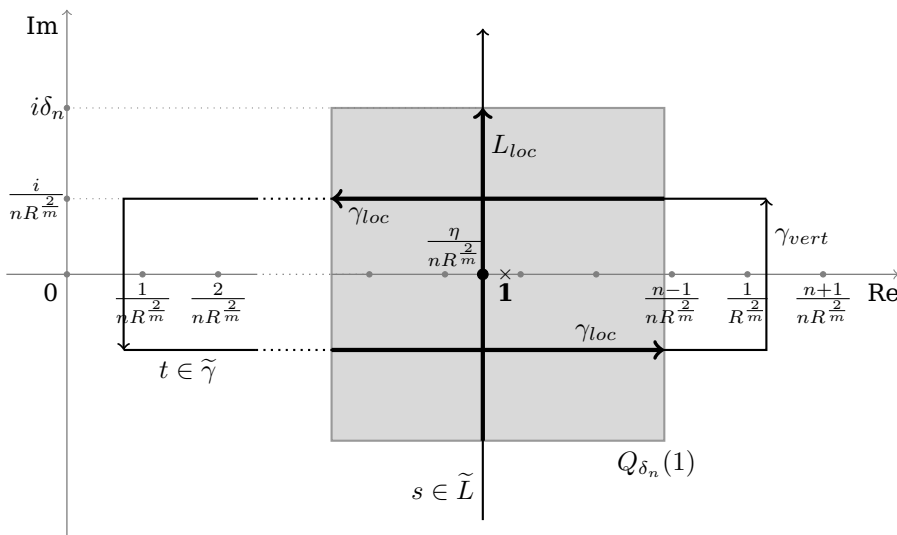

Figure 2: The scaled contours of integration and their local parts in thicker lines. As we will see later, the main contribution comes from γ_{loc} that is in a box δ_n-close to the saddle point at $z = 1$. If $R > 1$, there is no γ_{loc} and the integral vanishes exponentially fast (depending on the distance $|R - 1|$). If $R < 1$, then both horizontal contours of γ_{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 γ_{vert} is small enough.

Observe that $\mathcal{O}(1/(n\,\mathrm{Re}(t))) = \mathcal{O}(R^{2/m})$ and $\mathcal{O}(1/(n\,\mathrm{Re}(s))) = \mathcal{O}(1/n)$. We will analyze this main formula using the method of steepest descent, 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 $z = 1$ only, with $F''(1) = 1 > 0$. Denoting $z = x + iy$, the Cauchy-Riemann equations for F imply

$$\partial_y \operatorname{Re} F(z) = -\operatorname{Im} F'(z) = -\arg(z) > 0 \Leftrightarrow y < 0, \tag{2.9}$$

$$\partial_x \operatorname{Re} F(z) = \operatorname{Re} F'(z) = \log|z| > 0 \Leftrightarrow |z| > 1. \tag{2.10}$$

Hence $\operatorname{Re} F$ 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

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

as $n \to \infty$ by our assumption (2.4). Note that $\widetilde{\gamma}$ is $1/(nR^{2/m}) = \mathcal{O}(\delta_n^2)$-close to the real axis and the vertical path $\widetilde{L}$ is equally close to the saddle point $z = 1$. The local parts of the paths are given by $L_{loc} = \widetilde{L} \cap Q_{\delta_n}(1)$ and $\gamma_{loc} = \widetilde{\gamma} \cap Q_{\delta_n}(1)$, as well as L_{loc}^c and γ_{loc}^c denotes 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 $(s - 1) = i\operatorname{Im}(s) + \mathcal{O}(\delta_n^2)$, hence for $s \in L_{loc}$

$$F(s) = -1 - \operatorname{Im}(s)^2/2 + \mathcal{O}(\delta_n^3), \tag{2.11}$$

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

$$F(t) = -1 + (1 - \operatorname{Re}(t))^2/2 + \mathcal{O}(\delta_n^3), \tag{2.12}$$

since $|\operatorname{Im}(t)| \lesssim \delta_n$. On the other hand for $s \in L_{loc}^c$ by using (2.9) it holds

$$\operatorname{Re} F(s) < \operatorname{Re} F(\eta/(nR^{2/m}) + i\delta_n) = -1 - \delta_n^2/2 + \mathcal{O}(\delta_n^3) \tag{2.13}$$

and for $t \in \gamma_{loc}^c$ we see from (2.10)

$$\operatorname{Re} F(t) > \operatorname{Re} F(1 \pm \delta_n) = -1 + \delta_n^2/2 + \mathcal{O}(\delta_n^3). \tag{2.14}$$

The nonlocal terms are negligible, e.g. we apply (2.12) and (2.13) to obtain

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

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

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

where we used $|\operatorname{Im}(s)| \gtrsim \delta_n$, $|\gamma_{loc}| = \mathcal{O}(\delta_n)$, $t = \mathcal{O}(1)$, $\left| \cot(\pi nR^{2/m} t) \right| \asymp 1$ and $m \geq 2$. Moreover from (2.13), (2.14) and $t = \mathcal{O}(R^{-2/m})$ it follows

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

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

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

where the last step once more follows from the assumption (2.4). Analogously we obtain from (2.11), (2.14)

$$R^{2/m} \int_{\gamma_{loc}^c} \int_{L_{loc}} \exp\left[nmR^{2/m} \left(\operatorname{Re} F(s) - \operatorname{Re} F(t) \right) \right] \left| \frac{t}{s} \right|^{m/2} \frac{\left| \cot(\pi nR^{2/m} t) \right|}{|t - s|} ds \, dt \tag{2.15}$$

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

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

Locally close to $z = 1$, the error term of Stirling's formula (2.8) is $\mathcal{O}(n^{-1})$. Thus, it remains to control

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

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

$$\cdot \left(R^{2/m} + \mathcal{O}\left(R^{1/m}\sqrt{\frac{\log^3 n}{n}}\right)\right) dsdt, \tag{2.16}$$

where we used (2.11), (2.12) and $t/s = 1 + \mathcal{O}(\delta_n)$. 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 (2.16) on γ_{vert} is visible if and only if R is close to 1. In this case, we drop the part in t and the exponential function becomes $e^{-u^2/2}$. Using $R \sim 1$ and $|\cot(\pi/4 + ix)| = 1$ for $x \in \mathbb{R}$, the integration over γ_{vert} can then be bounded by

$$\left|\int_{-\sqrt{m/(nR^{2/m})}}^{\sqrt{m/(nR^{2/m})}} \int_I \frac{e^{-u^2/2}\cot\left(\pi/4 + ivR^{1/m}\sqrt{n/m}\right)}{m(n - \eta + 1/4)/R^{1/m} + i(v - u)\sqrt{nm}}R^{1/m}dudv\right| \tag{2.17}$$

$$\lesssim \int_{-\sqrt{m/(nR^{2/m})}}^{\sqrt{m/(nR^{2/m})}} \int_I \frac{e^{-u^2/2}}{\sqrt{m(n - \eta + 1/4)^2/R^{2/m} + n(v - u)^2}}dudv$$

$$\lesssim \frac{1}{n}\int_{-\infty}^{\infty} \frac{e^{-u^2/2}}{\sqrt{1/n + u^2}}du \lesssim \frac{\log n}{n},$$

where in the second step we shifted u by $v = \mathcal{O}(1/\sqrt{n})$ 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 elementarily by splitting the integration into $|u| \lessgtr 1$. 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 $\gamma_{loc} \setminus \gamma_{vert}$ as horizontal lines

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

where $\widetilde{I}$ is the part of I such that the corresponding contour overlaps $\widetilde{\gamma}$. The inte-

gral (2.16) becomes the sum of two terms, which can be estimated as

$$
\left| \int_{\widetilde{I}} \int_{I} e^{-\frac{u^2+v^2}{2}} \frac{\cot(\pi n R^{\frac{2}{m}} \pm i\pi \mp \pi\sqrt{nR^{\frac{2}{m}}/mv})}{\mp v + (nR^{\frac{2}{m}} - \eta)\sqrt{\frac{m}{nR^{\frac{2}{m}}}} + i\left(\pm\sqrt{\frac{m}{nR^{\frac{2}{m}}}} - u\right)} \right.
$$

$$
\left. \frac{iR^{\frac{1}{m}} + \mathcal{O}(\sqrt{\frac{\log^3 n}{n}})}{\mp\sqrt{nm}} dudv \right| \tag{2.18}
$$

$$
\leq \frac{1}{\sqrt{nm}} \int_{\widetilde{I}} \int_{I} e^{-\frac{u^2+v^2}{2}} \frac{R^{\frac{1}{m}} + \mathcal{O}\left(\sqrt{\frac{\log^3 n}{n}}\right)}{\sqrt{\left(v \mp (nR^{\frac{2}{m}} - \eta)\sqrt{\frac{m}{nR^{\frac{2}{m}}}}\right)^2 + \left(u \mp \sqrt{\frac{m}{nR^{\frac{2}{m}}}}\right)^2}} dudv
$$

$$
\leq \frac{1}{\sqrt{nm}} \int_{\mathbb{R}^2} e^{-\left|z+\mathcal{O}(\sqrt{m/(nR^{2/m})})\right|^2/2} \frac{R^{\frac{1}{m}} + \mathcal{O}\left(\sqrt{\frac{\log^3 n}{n}}\right)}{|z|} d\lambda(z)
$$

where we shifted $z = v + ui \in I + i\widetilde{I}$ by the corresponding values from the first denominator, which are of order $\mathcal{O}(\sqrt{m/(nR^{2/m})}) = \mathcal{O}(\log^{-\frac{3}{4}} n)$, and extended the area of integration. A simple estimate on the errors in the exponential function yields

$$
\frac{1}{\sqrt{nm}} \int_{\mathbb{R}^2} e^{-|z|^2/2} \left(1 + (|z|+1)\mathcal{O}\left(\sqrt{\frac{m}{nR^{\frac{2}{m}}}}\right)\right) \frac{R^{\frac{1}{m}} + \mathcal{O}\left(\sqrt{\frac{\log^3 n}{n}}\right)}{|z|} d\lambda(z)
$$

$$
\leq \frac{1}{\sqrt{nm}} \int_{\mathbb{R}^2} e^{-|z|^2/2} \frac{1 + \mathcal{O}\left(\sqrt{\frac{\log^3 n}{n}}\right)}{|z|} d\lambda(z) \sim \sqrt{\frac{2\pi^3}{nm}}.
$$

Recalling the correct prefactor $c = -1/4\pi$ from (2.8), we conclude

$$
|\mu_\infty^m(B_R) - \bar\mu_n^m(B_R)| \leq \sqrt{\frac{\pi}{2nm}} + o(n^{-1/2}).
$$

In order to obtain the lower bound of the claim, it suffices to consider $R = 1$. Moreover we will only study the sum of (2.18) 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 approximation

$$
\mu_\infty^m(B_1) - \bar\mu_n^m(B_1)
$$
$$
= \frac{1}{4\pi} \left(\int_{-\sqrt{m}/(4\sqrt{n})}^{\sqrt{m\log n}} \int_{-\sqrt{m\log n}}^{+\sqrt{m\log n}} e^{-\frac{u^2+v^2}{2}} \frac{\cot(i\pi - \pi\sqrt{n/m}v)}{v + \frac{\sqrt{m}}{2\sqrt{n}} + i\left(-\sqrt{\frac{m}{n}} + u\right)} \frac{i + o(1)}{\sqrt{nm}} dudv \right.
$$
$$
\left. + \int_{-\sqrt{m\log n}}^{+\sqrt{m}/(4\sqrt{n})} \int_{-\sqrt{m\log n}}^{+\sqrt{m\log n}} e^{-\frac{u^2+v^2}{2}} \frac{\cot(-i\pi + \pi\sqrt{n/m}v)}{v - \frac{\sqrt{m}}{2\sqrt{n}} + i\left(-\sqrt{\frac{m}{n}} - u\right)} \frac{i + o(1)}{\sqrt{nm}} dudv \right)
$$
$$
= \frac{1}{4\pi\sqrt{nm}} \int_{\sqrt{m}/(4\sqrt{n})}^{\sqrt{m\log n}} \int_{-\sqrt{m\log n}}^{+\sqrt{m\log n}} e^{-\frac{u^2+v^2}{2}} \left(1 + \mathcal{O}\left(\sqrt{\frac{\log n}{n}}\right)\right) \frac{-i}{v + ui}
$$
$$
\cdot (\tan(i\pi - \pi\sqrt{n/m}v) + \tan(i\pi + \pi\sqrt{n/m}v))dudv + o(n^{-1/2}).
$$

In the last step we first proceeded similarly to (2.18), i.e. we shifted v by a negligible value $\pm\sqrt{m}/(2\sqrt{n})$ and u by $\pm\sqrt{m/n}$, and then inverted $v \to -v$ in the second integral in order to merge both and put $\cot(\pi/2 - z) = -\tan(-z) = \tan(z)$. Note that

$\sup_{x \in \mathbb{R}} |\tan(i\pi - x) + \tan(i\pi + x) - 2i| =: \varepsilon < 0.01$ and that the left hand side of the previous equation is real, hence

$$
\begin{aligned}
\mu_\infty^m(B_1) &- \bar{\mu}_n^m(B_1) \\
&\geq \frac{1-\varepsilon}{2\pi\sqrt{nm}} \int_{\sqrt{m}/(4\sqrt{n})}^{\sqrt{m\log n}} \int_{-\sqrt{m\log n}}^{+\sqrt{m\log n}} e^{-\frac{u^2+v^2}{2}} \frac{v}{u^2+v^2} du\,dv + o(n^{-1/2}) \\
&\sim \frac{1-\varepsilon}{\sqrt{2\pi mn}}.
\end{aligned}
$$

The same upper bound holds with $1 + \varepsilon$ instead. For a better control of the constant one may vary the distance of γ to the real axis from the start. The above asymptotic yields the first claim and coincides with $m = 1$ from Lemma 1.2.

If we avoid the edge by some distance $|1 - R^{1/m}| \gtrsim \sqrt{\log n/n} > 0$, then $\delta_n < |1 - 1/R^{2/m}|$. Hence γ_{vert} is not a part of γ_{loc} and (2.17) drops out. This is the case for $R < 1 - \frac{m}{2}\sqrt{\log n/n}$, for which we have $I = \widetilde{I}$. As before, we shift u and v by $\mathcal{O}(1/\sqrt{nR^{2/m}})$, now we denote the shifted intervals by I_1, I_2 and obtain

$$
\begin{aligned}
&\int_I \int_I e^{-\frac{u^2+v^2}{2}} \left(\frac{-\cot(\pi n R^{\frac{2}{m}} + i\pi - \pi\sqrt{nR^{\frac{2}{m}}/mv})}{-v + (nR^{\frac{2}{m}} - \eta)\sqrt{m/(nR^{\frac{2}{m}})} + i(\sqrt{m/(nR^{\frac{2}{m}})} - u)} \right. \\
&\left. + \frac{\cot(\pi n R^{\frac{2}{m}} - i\pi + \pi\sqrt{nR^{\frac{2}{m}}/mv})}{v + (nR^{\frac{2}{m}} - \eta)\sqrt{m/(nR^{\frac{2}{m}})} - i(\sqrt{m/(nR^{\frac{2}{m}})} + u)} \right) \frac{iR^{\frac{1}{m}} + \mathcal{O}(\sqrt{\frac{\log^3 n}{n}})}{\sqrt{nm}} du\,dv \\
&= \int_{I_1} \int_{I_2} e^{-\frac{u^2+v^2}{2}} \left(1 + (1 + v + ui)\mathcal{O}\left(\frac{1}{\sqrt{nR^{\frac{2}{m}}}}\right) \right) \frac{iR^{\frac{1}{m}} + \mathcal{O}(\sqrt{\frac{\log^3 n}{n}})}{\sqrt{nm}(v + ui)} \\
&\quad \cdot \left(\tan(i\pi - \pi\sqrt{nR^{\frac{2}{m}}/mv}) - \tan(-i\pi + \pi\sqrt{nR^{\frac{2}{m}}/mv}) \right) du\,dv \\
&= \int_I \int_I e^{-\frac{u^2+v^2}{2}} \frac{iR^{\frac{1}{m}} + \mathcal{O}(\sqrt{\frac{\log^3 n}{n}})}{\sqrt{nm}(v + ui)} \\
&\quad \cdot \left(\tan(i\pi - \pi\sqrt{nR^{\frac{2}{m}}/mv}) - \tan(-i\pi + \pi\sqrt{nR^{\frac{2}{m}}/mv}) \right) du\,dv.
\end{aligned}
$$

As we have already showed and used in (2.18), the error terms of the shifted variables in the exponential function are absorbed in the error term $\mathcal{O}(\sqrt{\log^3 n/n})$. Furthermore, the deviation between the integrals over I instead of I_1, I_2 are obviously negligible as well. From the last line it follows that the horizontal contour integrals over I cancel due to symmetry, thus only the error term remains. Collecting this error, together with the non-local terms, we obtain

$$
\mu_\infty^m(B_R) - \bar{\mu}_n^m(B_R) \lesssim \frac{1}{\sqrt{n}} \int_{-\infty}^\infty \int_{-\infty}^\infty e^{-\frac{u^2+v^2}{2}} \frac{\sqrt{\log^3 n}}{\sqrt{n}\,|v + ui|} du\,dv + \mathcal{O}\left(\frac{1}{n}\right) \lesssim \frac{\log^{3/2} n}{n}.
$$

Lastly we turn to the statement about the 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 (2.10) and $\text{Re}(t) \leq (n+1)/(nR^{2/m})$ for $t \in \widetilde{\gamma}$ yield

$$
\begin{aligned}
\text{Re}\,F(t) &= \text{Re}\,F(\text{Re}(t)) + \mathcal{O}(1/n) \\
&\geq F\left(\frac{n+1}{nR^{2/m}}\right) + \mathcal{O}(1/n) \tag{2.19} \\
&= -R^{-2/m}\left(\log(R^{2/m}) + 1\right) + \mathcal{O}(1/n)
\end{aligned}
$$

for n sufficiently large. Note that in this case of big R, the local area $Q_{\delta_n}(1)$ does intersect $\widetilde{\gamma}$, hence $\gamma_{loc} = \emptyset$. We estimate (2.8) similarly to (2.15), i.e. apply (2.11), (2.13), (2.19) to obtain

$$\int_{\widetilde{\gamma}} \int_{\widetilde{L}} \exp\left[nmR^{2/m}\left(\operatorname{Re} F(s) - \operatorname{Re} F(t)\right)\right] \left|\frac{t}{s}\right|^{m/2} \frac{\left|\cot(\pi n R^{2/m} t)\right|}{|t - s|} ds dt$$
$$\lesssim \int_{\widetilde{\gamma}} \int_{\widetilde{L}} \exp\left[-nm(R^{2/m} - 1 - \log(R^{2/m})) + \mathcal{O}(1)\right] \frac{1}{|s|^{m/2} |\operatorname{Im}(s)|} ds dt.$$

In particular, we used the fact that $|t - s|$ behaves like $|\operatorname{Im}(s)|$ for $s \in L_{loc}^c$ and is bounded from below by $\operatorname{Re}(s - t) \gtrsim \sqrt{\log n/n} \gtrsim |\operatorname{Im}(s)|$ for $s \in L_{loc}$, since $\operatorname{Re}(t)$ is small. The remaining integral over $t \in \widetilde{\gamma}$ is finite, since $\widetilde{\gamma}$ has finite length and the integral over $s \in \widetilde{L}$ is finite, hence we may use Bernoulli's inequality to conclude

$$|\mu_\infty^m(B_R) - \bar{\mu}_n^m(B_R)| \lesssim \exp\left[-2n(R - 1 - \log R)\right] \leq \exp\left[-n(R - 1)^2/3\right],$$

where the last inequality holds for $1 < R < 2$ by a Taylor approximation of the logarithm. Ultimately all claims are proven. $\qquad\square$

It seems to 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 $\mathcal{O}(1/n)$ seems to be achievable due to the discrete nature of the residue calculus, cf. (2.8) above.

The proof of the rate of convergence for the non-averaged ESD makes use of the previous result, the determinantal structure of μ_n^m and Bernstein's inequality.

Proof of Theorem 1.4. Without loss of generality, we consider $0 < R < 1$, since for $R \geq 1$, we have $|\mu_n^m(B_R) - \mu_\infty(B_R)| \leq |\mu_n^m(B_1) - \mu_\infty(B_1)|$. Suppose we had shown already that for any $Q > 0$ there exists a $c > 0$ such that

$$\mathbb{P}\left(|\mu_n^m(B_R) - \mu_\infty^m(B_R)| > t_n/5\right) \leq n^{-Q - \lceil m/4 \rceil} \tag{2.20}$$

for $t_n = c\sqrt{\log n/n}$ and fixed $0 < R < 1$. Set $p = \lceil m/4 \rceil$ and consider the equidistant points $r_k = k/n^p$ for $k = 1, \ldots, n$. We have

$$|\mu_n^m(B_R) - \mu_\infty^m(B_R)| \leq \mu_n^m(B_{r_{k+1}} \setminus B_{r_k}) + \mu_\infty^m(B_{r_{k+1}} \setminus B_{r_k}) + |\mu_n^m(B_{r_k}) - \mu_\infty^m(B_{r_k})|$$
$$\leq \left|\mu_n^m(B_{r_{k+1}}) - \mu_\infty^m(B_{r_{k+1}})\right| + 2\mu_\infty^m(B_{r_{k+1}} \setminus B_{r_k}) + 2\left|\mu_n^m(B_{r_k}) - \mu_\infty^m(B_{r_k})\right|$$

where we chose $k = \lfloor Rn^p \rfloor$. Taking the supremum over R is equivalent to taking the maximum in k, hence the union bound implies

$$\mathbb{P}\left(\sup_{R > 0} |\mu_n^m(B_R) - \mu_\infty^m(B_R)| > t_n\right) \leq \sum_{k=1}^{n^p} \Big(\mathbb{P}\left(\left|\mu_n^m(B_{r_{k+1}}) - \mu_\infty^m(B_{r_{k+1}})\right| > t_n/5\right)$$
$$+ \mathbb{P}\left(\mu_\infty^m(B_{r_{k+1}} \setminus B_{r_k}) > t_n/5\right)$$
$$+ \mathbb{P}\left(|\mu_n^m(B_{r_k}) - \mu_\infty^m(B_{r_k})| > t_n/5\right)\Big).$$

The first and last term are covered by (2.20) and the second term vanishes because of

$$\mu_\infty^m(B_{r_{k+1}} \setminus B_{r_k}) = \frac{(k+1)^{2/m} - k^{2/m}}{n^{2p/m}} \leq n^{-p(1 \wedge 2/m)} \leq n^{-1/2}.$$

Thus, we have shown that for all $Q > 0$ there exists a constant $c > 0$ such that $\sup_{R > 0} |\mu_n^m(B_R) - \mu_\infty^m(B_R)| \leq c\sqrt{\log n/n}$ holds with probability at least $1 - n^{-Q}$.

It remains to show (2.20) for which we follow the ideas of [30, Proposition 2.2] for the case $m = 1$. Fix $R < 1$ and let $\xi_k \sim Ber(\eta_k) \in \{0,1\}$ be independent Bernoulli variables with parameters $\eta_k \in [0,1]$, which shall be defined in the following. According to [6, Corollary 4.2.24], the determinantal point process

$$\#\{\lambda_j(\sqrt{n^m}\mathbf{X}) \in B_{\sqrt{n^m}R}\} = n\mu_n^m(B_R) \overset{\mathcal{D}}{=} \sum_{k=1}^{n} \xi_k$$

has the same distribution as the sum of Bernoulli variables, where the parameters η_k are given by the eigenvalues of the trace class operator

$$\mathcal{K}_R f(z) = \int_{B_{\sqrt{n^m}R}} \sqrt{G_{0,m}^{m,0}\left(\genfrac{}{}{0pt}{}{-}{0}\,\middle|\,|z|^2\right) G_{0,m}^{m,0}\left(\genfrac{}{}{0pt}{}{-}{0}\,\middle|\,|w|^2\right)} \sum_{k=0}^{n-1} \frac{(z\bar{w})^k}{\pi k!^m} f(w)d\lambda(w)$$

for $f \in L^2(B_{\sqrt{n^m}R})$. Due to rotational symmetry, like we argued for the orthogonality of the monomials, $\mathcal{K}_R$ has eigenfunctions $\varphi_k(w) = \sqrt{G_{0,m}^{m,0}\left(\genfrac{}{}{0pt}{}{-}{0}\,\middle|\,|w|^2\right)}\,w^k$ with eigenvalues

$$\eta_k = \int_{B_{\sqrt{n^m}R}} G_{0,m}^{m,0}\left(\genfrac{}{}{0pt}{}{-}{0}\,\middle|\,|w|^2\right) \frac{|w|^{2k}}{\pi k!^m} d\lambda(w) \leq 1.$$

It follows from Theorem 1.3 that

$$\mathbb{P}\left(|\mu_n^m(B_R) - \mu_\infty^m(B_R)| \geq \frac{t}{\sqrt{n}}\right) \leq \mathbb{P}\left(|\mu_n^m(B_R) - \bar{\mu}_n^m(B_R)| \geq \frac{t-c}{\sqrt{n}}\right)$$

$$= \mathbb{P}\left(\left|\sum_{k=1}^{n}\xi_k - \mathbb{E}\left(\sum_{k=1}^{n}\xi_k\right)\right| \geq (t-c)\sqrt{n}\right).$$

Applying Bernstein's inequality, see [8, Equation (2.10)] yields

$$\mathbb{P}\left(|\mu_n^m(B_R) - \mu_\infty^m(B_R)| > \frac{t}{\sqrt{n}}\right) \leq 2\exp\left(-\frac{n(t-c)^2}{2\sum_{k=1}^{n}\mathbb{E}|\xi_k|^2 + \frac{2}{3}(t-c)\sqrt{n}}\right)$$

$$\leq e^{-t^2/3} \tag{2.21}$$

for t sufficiently large, since $\mathbb{E}|\xi_k|^2 \leq \eta_k \leq 1$. In particular for $t = c\sqrt{\log n}/5$ with $c = 5\sqrt{3(Q + \lceil m/4\rceil)}$, we obtain (2.20) as claimed. $\qquad\square$

Note that the logarithmic factor in the rate of convergence is expected to be non-optimal and might be removed by controlling the variance term (2.21), but this will not be pursued here.

3 Matrices with independent entries

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

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

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

$$\mathbf{V}(z) := \begin{pmatrix} 0 & \mathbf{W} - z \\ (\mathbf{W} - z)^* & 0 \end{pmatrix} \tag{3.1}$$

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

Similar to the role of the Stieltjes transform in the theory of Hermitian random matrices, the weak topology of measures μ 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 μ on $\mathbb{C}$ the *logarithmic potential* defined by

$$U_\mu(z) := -\int_{\mathbb{C}} \log|t - z| \, d\mu(t) = (-\log|\cdot| * \mu)(z) \tag{3.2}$$

and it satisfies

$$\Delta U_\mu = -2\pi\mu \tag{3.3}$$

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

$$U_n(z) = -\frac{1}{nm} \sum_{j=1}^{nm} \log|\lambda_j(\mathbf{W}) - z| = -\frac{1}{2nm} \log|\det \mathbf{V}(z)| = -\int_{-\infty}^{\infty} \log|x| \, d\widetilde{\nu}_n^z(x). \tag{3.4}$$

Under the above-mentioned conditions on the matrix entries, the logarithmic potential U_n concentrates around the logarithmic potential U_∞ 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 3.1 ([18]). *If $\mathbf{X}$ obeys (C), then for every $\tau, Q > 0$ there exists a constant $c > 0$ such that*

$$\mathbb{P}\left(|U_n(z) - U_\infty(z)| \leq c\frac{\log^4 n}{n}\right) \geq 1 - n^{-Q} \tag{3.5}$$

holds uniformly in $\{z \in B_{1+\tau^{-1}} : |1 - |z|| \geq \tau\}$.

We remark that the restriction to the bulk in Theorem 1.6 comes directly from the restriction in the previous Proposition 3.1.

Since the statement of Proposition 3.1 is not explicitly worked out in [18], we will derive it now, based on the results proved in this paper. We will directly follow the approach of [18], making use of Girko's Hermitization trick to convert the non-Hermitian problem into a Hermitian one, apply the local Stieltjes transform estimate from [18] and the smoothing inequality from [19]. Let

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

be the Stieltjes transform of $\widetilde{\nu}_n^z$, which converges a.s. to the solution of

$$s(z, w) = -\frac{w + s(z, w)}{(w + s(z, w))^2 - |z|^2} \tag{3.6}$$

see for instance [20]. It is known that $s(z, \cdot)$ is the Stieltjes transform of a measure $\widetilde{\nu}^z$, which appears to be the weak limit of $\widetilde{\nu}_n^z$ as $n \to \infty$, cf. [18]. Moreover, $\widetilde{\nu}^z$ has a symmetric bounded density ρ^z (the bound holds uniformly in z) and has compact support. We refer to [9, §4] for more information on s and ρ^z. Note that s will be unbounded from below for w close to the origin and z close to the edge of the support of μ_∞^m, which is the reason for the bulk constraint of Proposition 3.1.

Proof of Proposition 3.1. 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 (3.4), we may write

$$|U_n(z) - U_\infty(z)| = \left| \int_{\mathbb{R}} \log|x| \, d(\widetilde{\nu}_n^z - \widetilde{\nu}^z)(x) \right|.$$

Due to the singularities of $\log|x|$ at $x = 0$ and $x = \infty$, it is not enough to simply control the convergence of $\widetilde{\nu}_n^z \to \widetilde{\nu}^z$, but it is also necessary to estimate the extremal singular values. The rate of convergence of $\widetilde{\nu}_n^z$ to $\widetilde{\nu}^z$ shall be measured in Kolmogorov distance

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

Introduce the events

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

for some constants $B, B', c > 0$ yet to be chosen. From [33, Theorem 31] it follows that there exists a constant $B > 0$ such that $\mathbb{P}(\Omega_0^c) \lesssim n^{-Q}$. Moreover for any $Q > 0$ it holds

$$\mathbb{P}(s_{\max} \geq n^{(Q+1)/2}) \leq \frac{\mathbb{E} \|(\mathbf{W} - z)\|^2}{n^{Q+1}} \tag{3.7}$$

$$\leq \frac{1}{n^{Q+1}} \sum_{ij}^{nm} \mathbb{E} |\mathbf{W}_{ij}|^2 + \frac{|z|^2}{n^Q} \leq (m + \tau^{-1}) n^{-Q},$$

where the operator norm $\|\cdot\|$ has been estimated by the Hilbert Schmidt norm. Thus, there exists a constant $B' > 0$ with $\mathbb{P}(\Omega_1^c) \lesssim n^{-Q}$. Since $\widetilde{\nu}^z$ has a bounded density, we get

$$\left| \int_{-n^{-B}}^{n^{-B}} \log|x| \, d\widetilde{\nu}^z(x) \right| \lesssim \frac{\log n}{n^B}$$

and furthermore on Ω_2 it holds that

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

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

$$\mathbb{P}\left(|U_n(z) - U_\infty(z)| \geq c \frac{\log^4 n}{n} \right) \leq \mathbb{P}(\Omega_0^c) + \mathbb{P}(\Omega_1^c) + \mathbb{P}(\Omega_2^c).$$

It remains to check $\mathbb{P}(\Omega_2^c) \leq n^{-Q}$, which has been done explicitly in [18, (4.14)-(4.16)], using the Smoothing Inequality [18, Corollary B.3] (originally obtained in [19]) and the Local Law for $\mathbf{V}(z)$ in terms of its Stieltjes transform. $\square$

The core of the proof of the local law for products of non-Hermitian matrices in [18] is the following identity. First, for any function $f \in \mathcal{C}_c^2(\mathbb{C})$ define $\widetilde{f}$ by $\widetilde{f}(z) = f(z^m)$ and note that $\int f \, d\mu_\infty^m = \int \widetilde{f} \, d\mu_\infty^1$, which follows from the definition of μ_∞^m in (1.1). Using the

distributional Poisson equation (3.3) as usual in non-Hermitian random matrix theory and the representation of the eigenvalues of $\mathbf{W}$, we get

$$\int f d(\mu_n^m - \mu_\infty^m) = \frac{1}{nm} \sum_{j=1}^{nm} \widetilde{f}(\lambda_j(\mathbf{W})) - \int \widetilde{f} d\mu_\infty^1 = -\frac{1}{2\pi} \int \Delta \widetilde{f} (U_n - U_\infty) \, d\lambda. \qquad (3.8)$$

In Local Laws [31, 18], locally shrinking functions f_{z_0} have been considered and for fixed global f Gaussian fluctuation has been proven in [24]. As was done in [16], we would like to uniformly approximate all indicator functions $\mathbb{1}_{B_R(z_0)}$ by smooth functions, replace the right hand side of (3.8) by a discrete random sum and use the pointwise estimate from Proposition 3.1. In contrast to [16], we cannot use the smoothing inequality [16, Theorem 2.1], since it would rely on the distance of U_∞ to the logarithmic potential of the matrix $\mathbf{X}$, while instead we can control the logarithmic potential U_n of the linearization matrix $\mathbf{W}$ only. Therefore we will use a direct approach. In this case, we can replace the Monte Carlo approximation of the integral (3.8) by a random grid approximation. This also yields a more precise rate of convergence in Theorem 1.6. Here, and in the sequel we will identify $\mathbb{R}^2 \simeq \mathbb{C}$ via $(u, v) \mapsto u + iv$ in order to simplify notation.

Lemma 3.2. *Let $\alpha, \beta > 0$ and S be a random variable uniformly distributed on $[0,1]^2$. Fix complex numbers $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ with corresponding logarithmic potential U of their empirical distribution $\mu_n = \frac{1}{n} \sum_{j=1}^n \delta_{\lambda_j}$. Furthermore, define the random grid $A = 2\beta n^{-\alpha/2}(\mathbb{Z}^2 + S) \cap [-\beta, \beta]^2$ enumerated by $z_1, \ldots, z_{\lceil n^\alpha \rceil}$. For any function $f \in \mathcal{C}_c^3(\mathbb{C})$ with $\operatorname{supp} f \subseteq (-\beta, \beta)^2$ it holds*

$$\frac{1}{n} \sum_{j=1}^n f(\lambda_j) - \frac{-2\beta^2}{n^\alpha \pi} \sum_{i=1}^{\lceil n^\alpha \rceil} \Delta f(z_i) U(z_i) = \mathcal{O}\big(\big(\|\nabla \Delta f\|_\infty + \|\Delta f\|_\infty \log^2 n \big) n^{-\alpha/2} \big) \qquad (3.9)$$

with overwhelming probability. Moreover, if S is chosen independently of the random matrix $\mathbf{W}$, then (3.9) for $U = U_n$ and

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

hold on an event Ω_* of probability $1 - \mathcal{O}(n^{-\log n})$, which does not depend on f.

Note that (3.9) holds uniformly in $f \in \mathcal{C}_c^3(\mathbb{C})$, hence we could choose a function depending on the positions of λ_j. In order to make the statement more intuitive, suppose we replace the logarithmic potential U by a more regular function $U \in \mathcal{C}^1$. Then (3.9) is nothing but Riemann approximation of the integral

$$\int \Delta f(z) U(z) d\lambda(z) - \frac{(2\beta)^2}{n^\alpha} \sum_{i=1}^{\lceil n^\alpha \rceil} \Delta f(z_i) U(z_i) \lesssim \big(\|\nabla \Delta f\|_\infty + \|\Delta f\|_\infty \big) n^{-\alpha/2}. \qquad (3.10)$$

This follows directly from the mean value theorem, very similarly to what we will do in (3.12) below.

In the Monte Carlo approximation used in [35] and [24], the random points z_i are not ordered in a grid but drawn independently, thus variance bounds are of importance for improving bounds as (3.9). By using reference points or eigenvalue rigidity, the error estimates in [35] and [24] are stronger by a factor of $1/n$ for the same number 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 (3.9) only a single random shift affects all points z_i. Heuristically speaking, this leads to a higher probability than in previous approaches, so that the weaker error bound is negligible.

Proof of Lemma 3.2. Using the definition (3.2) via (3.3), in other words integration by parts, we find

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

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

$$\int \Delta f(z) \log|\lambda - z|\, d\lambda(z) - \frac{4\beta^2}{n^\alpha}\sum_{i=1}^{\lceil n^\alpha \rceil} \Delta f(z_i) \log|\lambda - z_i|$$
$$= \mathcal{O}\big(\big(\|\nabla \Delta f\|_\infty + \|\Delta f\|_\infty \log^2 n\big) n^{-\alpha/2}\big) \tag{3.11}$$

for fixed $\lambda \in \mathbb{C}$, since the claim then follows from summation over j and the union bound.

In the following, we will show that the event, where (3.11) holds, will be $\widetilde{\Omega}_* = \bigcap_{z_i \in A}\{|z_i - \lambda| > 2\beta n^{-(\log n + 1 + \alpha)/2}\}$ which fails if λ is too close to the random grid A. More precisely let $z^* \in 2\beta n^{-\alpha/2}\mathbb{Z}^2 \cap [-\beta, \beta]^2$ be the corner of a box such that $\lambda \in z^* + [0, 2\beta n^{-\alpha/2}]^2$, then

$$\mathbb{P}(\widetilde{\Omega}_*^c) = \mathbb{P}\left(\exists i = 1, \dots, \lceil n^\alpha \rceil : |z_i - \lambda| \le 2\beta n^{-(\log n + 1 + \alpha)/2}\right)$$
$$= \mathbb{P}\left(\text{dist}\left(S, \frac{\lambda - z^*}{2\beta n^{-\alpha/2}}\right) \le n^{(-\log n - 1)/2}\right) = \mathcal{O}(n^{-\log n - 1}),$$

where the distance in $[0, 1]^2$ is measured according to the metric of the quotient space $\mathbb{T}^2$. The situation is illustrated in Figure 3.
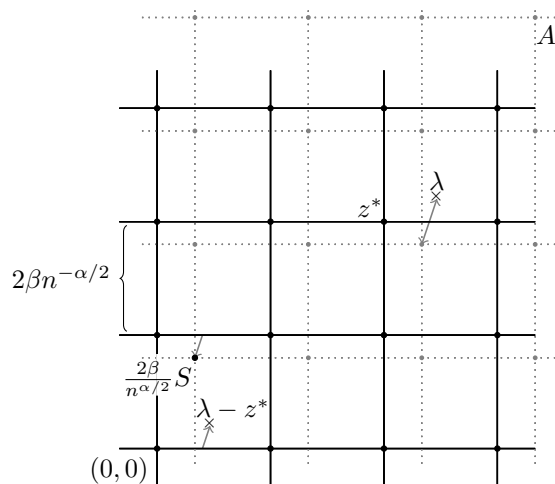

Figure 3: The distance of λ to random grid A is illustrated by the gray doubled arrow. This is equal to the distance of the rescaled random shift $2\beta n^{\alpha/2} S$ to $\lambda - z^*$ (measured in the quotient space), where z^* is the reference corner of the box containing λ.

From now on we will restrict ourselves to this event $\widetilde{\Omega}_*$. Rewrite (3.11) as

$$\sum_{i=1}^{\lceil n^\alpha \rceil}\int_{K_i} \Delta f(z) \log|\lambda - z| - \Delta f(z_i) \log|\lambda - z_i|\, d\lambda(z),$$

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

$$\sum_{i=1}^{\lceil n^\alpha \rceil} \int_{K_i} (\Delta f(z) - \Delta f(z_i)) \log|\lambda - z| \, d\lambda(z) = \mathcal{O}(\|\nabla \Delta f\|_\infty \, n^{-\alpha/2}), \qquad (3.12)$$

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

$$\left| \sum_{i=1}^{\lceil n^\alpha \rceil} \int_{K_i} \Delta f(z_i) \left(\log|\lambda - z| - \log|\lambda - z_i| \right) d\lambda(z) \right|$$
$$\leq \|\Delta f\|_\infty \left(\sum_{i:|z_i - \lambda| \geq n^{-\frac{\alpha}{4}}} + \sum_{i:|z_i - \lambda| < n^{-\frac{\alpha}{4}}} \right) \int_{K_i} \left| \log|\lambda - z| - \log|\lambda - z_i| \right| d\lambda(z).$$

Applying the mean value theorem for $\log$, yields a bound of order $\mathcal{O}(n^{-\alpha/4})$ for the first sum. The second sum can be bounded by performing the integration in polar coordinates

$$\int_0^{2n^{-\alpha/4}} r \left(\log r + \sup_{z_i \in A} \log|\lambda - z_i| \right) dr = \mathcal{O}(n^{-\alpha/2} \log^2 n),$$

where we finally used the event $\widetilde{\Omega}_*$. Putting all estimates together proves the first claim.

The bound for U_n follows from the choice of A and a trivial upper bound on the spectral radius $|\lambda|_{\max}$ of $\mathbf{W}$. On the one hand $|\lambda|_{\max}$ is bounded by the largest singular value $s_{\max}$ and on the other hand we have

$$\mathbb{P}(s_{\max} \geq n^{\log n}) \leq \mathbb{E} \|(\mathbf{W} - z)\|^2 \, n^{-2\log n} \lesssim n^{-2\log n + 1},$$

similar to (3.7). Therefore on the event $\Omega_* = \widetilde{\Omega}_* \cap \{s_{\max} \leq n^{\log n}\}$ with probability at least $1 - \mathcal{O}(n^{-\log n})$, we have (3.9) and

$$\sup_{z_i \in A} |U_n(z_i)| \leq \frac{1}{n} \sum_{j=1}^{n} \sup_{z_i \in A} \left| \log|\lambda_j - z_i| \right|$$
$$\lesssim (\log n + 1 + \alpha) \log n + \log \left| |\lambda|_{\max} + 5 \right| = \mathcal{O}(\log^2 n). \qquad \square$$

Finally, we turn to the

Proof of Theorem 1.6. First, note that we only need to consider $\tau < 1$. In order to restrict ourselves to a bounded region, say $V = B_7(0)$, we separate

$$\sup_{\substack{B_R(z_0) \subseteq \mathbb{C} \setminus B_{1+\tau} \\ \text{or } B_R(z_0) \subseteq B_{1-\tau}}} |(\mu_n^m - \mu_\infty^m)(B_R(z_0))|$$
$$\leq \sup_{\substack{B_R(z_0) \subseteq \mathbb{C} \setminus B_{1+\tau} \\ \text{or } B_R(z_0) \subseteq B_{1-\tau}}} |(\mu_n^m - \mu_\infty^m)(B_R(z_0) \cap V)| + \mu_n^m(V^c). \qquad (3.13)$$

Fix some $\tau, R > 0, z_0 \in \mathbb{C}$ such that $B_R(z_0) \subseteq B_{1-\tau} \cup B_{1+\tau}^c$.

Let $\varphi \in \mathcal{C}^\infty(\mathbb{R})$ be nonnegative with $\operatorname{supp} \varphi \subseteq [-1, 1]$ and $\int \varphi = 1$, and define $\varphi_a(\rho) = a\varphi(a\rho)$ for some large n-dependent parameter $a > 1$ to be determined later. We define the cutoff to V^c by

$$f_0(z) = \left(\mathbb{1}_{(7-1/a, \infty)} * \varphi_a \right)(|z|) \geq \mathbb{1}_{V^c}(z).$$

Moreover, we mollify the indicator function $\mathbb{1}_{B_R(z_0) \cap V}$ appearing in (3.13) via the approximation

$$
\begin{aligned}
f_1(z) := & \left(\mathbb{1}_{(-\infty, R-1/a]} * \varphi_a \right) (|z - z_0|) \cdot \left(\mathbb{1}_{(-\infty, 7-1/a]} * \varphi_a \right) (|z|) \\
& \leq \mathbb{1}_{B_R(z_0) \cap V}(z) \\
& \leq \left(\mathbb{1}_{(-\infty, R+1/a]} * \varphi_a \right) (|z - z_0|) \cdot \left(\mathbb{1}_{(-\infty, 7+1/a]} * \varphi_a \right) (|z|) =: f_2(z),
\end{aligned}
$$

where we choose $f_1 \equiv 0$ if $R \leq 2/a$ for smoothness reasons.

We apply $f_1 \leq \mathbb{1}_{B_R(z_0) \cap V}$ and integration by parts to $\widetilde{f_1} : z \mapsto f_1(z^m)$ as was explained in (3.8) to obtain

$$
\begin{aligned}
\mu_n^m(B_R(z_0) \cap V) &\geq \int f_1 d\mu_n^m = -\frac{1}{2\pi} \int (\Delta \widetilde{f_1}) U_n d\lambda \\
&= -\frac{1}{2\pi} \int \Delta \widetilde{f_1}(U_n - U_\infty) d\lambda - \int (\mathbb{1}_{B_R(z_0) \cap V} - f_1) d\mu_\infty^m + \int \mathbb{1}_{B_R(z_0) \cap V} d\mu_\infty^m. \quad (3.14)
\end{aligned}
$$

Analogous upper bounds hold for f_0 and f_2. A rough estimate of the error of approximation yields for the second term

$$
\int (\mathbb{1}_{B_R(z_0) \cap V} - f_1) d\mu_\infty^m \leq \mu_\infty^m \left(z \in \mathbb{C} : R - 2/a \leq |z - z_0| \leq R \right). \quad (3.15)
$$

Due to the radial monotonicity of μ_∞^m's density, this value increases by bending segments of the given annulus of width $2/a$ into straight rectangles $[-1, 1] \times [-4/a, 4/a]$ for some constant $c > 0$. Let us provide some more details on how a less technical, non-continuous bending procedure can be achieved:

The distribution of the annulus (3.15) is bounded by 10 times the maximal probability of an annulus segment with outer circumferential length $2\pi R/10 \wedge 2$. By radial monotonicity of μ_∞^m, this particular "worst case" $1/10$-th segment S is given by one that is closest to the origin and by radial symmetry we assume it to be symmetric with respect to the x-axis. We want to bend (or rather project) S onto the y-axis. Note that the angle between the y-axis and any tangent to the circumference line of the segment S is always bounded by $\pi/10$. Split our segment S in at most ca many pieces each of outer circumferential length $10/a$ and shift them in x-direction until they touch the y-axis (i.e. "projecting" the pieces). This transformation only increases the probability since we may only decrease the radial part of all points in each piece. Now since the mentioned angle is bounded, at most a constant number of the shifted segments overlap. One can show that this constant is actually equal to 2, or in other words non-neighboring shifted pieces are disjoint. Therefore, the total probability in (3.15) is bounded by $20\mu_\infty^m([-1, 1] \times [10/a, 10/a])$.

The density is bounded in the case of $m = 1$ and hence the term in (3.15) is of order $\mathcal{O}(1/a)$. In general we can bound it by

$$
\begin{aligned}
\mu_\infty^m([-1, 1] \times [-c/a, c/a]) &\leq \frac{1}{\pi m} \int_{-1}^1 \int_{-c/a}^{c/a} (x^2 + y^2)^{1/m-1} dx\, dy \\
&\leq \frac{4c}{\pi m a} \int_{c/(2a)}^1 x^{2/m-2} dx + \frac{2}{m} \int_0^{4c/a} r^{2/m-1} dr \\
&= \frac{8}{\pi(m-2)} \left(\left(\frac{c}{2a} \right)^{2/m} - \left(\frac{c}{2a} \right) \right) + \left(\frac{4c}{a} \right)^{2/m} \lesssim a^{-2/m},
\end{aligned}
$$

where the equation only holds for $m > 2$. For $m = 2$ we get

$$\mu_\infty^2([-1,1] \times [-c/a, c/a]) \leq \frac{1}{2\pi} \int_{-1}^{1} \int_{-c/a}^{c/a} (x^2 + y^2)^{-1/2} dx dy$$

$$= \frac{c}{\pi a} \log\left(\frac{\sqrt{a^2 + c^2} + a}{\sqrt{a^2 + c^2} - a}\right) + \frac{2}{\pi} \log\left(\sqrt{1 + c^2/a^2} + c/a\right)$$

$$\sim a^{-1} \log a$$

and we see that the log-term appears naturally. Define the error function

$$\widetilde{h}_m(a) = \begin{cases} \mathcal{O}(a^{-1}) & \text{for } m = 1, \\ \mathcal{O}(a^{-1} \log a) & \text{for } m = 2, \\ \mathcal{O}(a^{-2/m}) & \text{for } m \geq 3. \end{cases} \tag{3.16}$$

Let us continue to estimate the first term of (3.14) by using our random grid approximation Lemma 3.2. Let $\beta = 7$ and S be a random variable, independent of $\mathbf{X}$ and uniformly distributed on $[0,1]^2$. Conditioned on $\mathbf{X}$, we have with overwhelming probability

$$\int \Delta \widetilde{f}_1(z) U_n(z) d\lambda(z) - \frac{(2\beta)^2}{n^\alpha} \sum_{i=1}^{\lceil n^\alpha \rceil} \Delta \widetilde{f}(z_i) U_n(z_i)$$

$$= \mathcal{O}\big(\big(\|\nabla \Delta \widetilde{f}_1\|_\infty + \|\Delta \widetilde{f}_1\|_\infty \log^2 n\big) n^{-\alpha/2}\big)$$

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 φ_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 (3.10), we conclude that for any matrix $\mathbf{X}$ we have with overwhelming probability

$$\sup_B \left| \int \Delta \widetilde{f}_1(z)(U_n(z) - U_\infty(z)) d\lambda(z) - \frac{(2\beta)^2}{n^\alpha} \sum_{i=1}^{\lceil n^\alpha \rceil} \Delta \widetilde{f}_1(z_i) \left(U_n(z_i) - U_\infty(z_i)\right) \right|$$

$$= \mathcal{O}\big((a^3 + a^2 \log^2 n) n^{-\alpha/2}\big),$$

where the supremum runs over all choices of $B \subseteq B_{1-\tau} \cup B_{1+\tau}^c$. Since we will always choose $a \lesssim n$ (actually we will make it even smaller, cf. (3.17)), 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 $\mathcal{O}(n^{-1})$. It should be emphasized that still no randomness of $\mathbf{X}$ has been used and the only randomness is the shifted grid. Combining the previous steps yields

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

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

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

In the same way, the bound holds for $\mu_n^m(V^c)$ as well. Finally we use the randomness of $\mathbf{X}$ by applying Proposition 3.1. Conditioning on S, i.e. freezing the lattice points z_i, we obtain for any $Q > 0$

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

for each $i = 1, \ldots, \lceil n^\alpha \rceil$. By the union bound this implies that with probability at least $1 - n^{-Q}$ the logarithmic potentials concentrate like $U_n(z_i) - U_\infty(z_i) = \mathcal{O}(\log^4 n/n)$ simultaneously at all lattice points. Therefore, for $k = 0, 1, 2$,

$$\frac{(2\beta)^2}{2\pi n^\alpha}\sum_{i=1}^{\lceil n^\alpha \rceil} \Delta\widetilde{f}_k(z_i)\Big(U_n(z_i) - U_\infty(z_i)\Big) \lesssim \frac{(2\beta)^2 \log^4 n}{n^{1+\alpha}}\sum_{i=1}^{\lceil n^\alpha \rceil} |\Delta\widetilde{f}_k(z_i)|$$

$$= \frac{\log^4 n}{n}\big\|\Delta\widetilde{f}_k\big\|_{L^1} + \mathcal{O}\Big(a^3\frac{\log^4 n}{n^{\alpha/2+1}}\Big)$$

where the integral of the a^3-Lipschitz function $|\Delta\widetilde{f}_k|$ has been approximated by its Riemann sum. Write $\Delta = 4\bar\partial\partial$ in terms of the Wirtinger derivatives $\partial = \frac{1}{2}(\partial_x - i\partial_y)$ and $\bar\partial = \frac{1}{2}(\partial_x + i\partial_y)$. Since $g(z) = z^m$ is holomorphic, i.e. $\bar\partial g = 0$, we obtain by applying the chain rule and changing variables from z to $g(z)$

$$\big\|\Delta\widetilde{f}_k\big\|_{L^1} = \big\|4\bar\partial\partial(f_k \circ g)\big\|_{L^1} = 4\big\|(\bar\partial\partial f_k) \circ g \cdot \bar\partial\bar{g} \cdot \partial g)\big\|_{L^1} = \big\|\Delta f_k\big\|_{L^1}.$$

Since $\Delta f_k \lesssim a^2$ and has support on an area of order a^{-1}, we have

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

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

$$\sup_{B \subseteq B_{1-\tau} \cup B_{1+\tau}^c} |(\mu_n^m - \mu_\infty^m)(B)| \leq ca\frac{\log^4 n}{n} + \widetilde{h}_m(a) + \mathcal{O}(n^{-1}). \tag{3.17}$$

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

In the proof we have seen that the maximal error for the limiting distribution μ_∞^m is by balls, which touch the origin (technically these balls of growing size are not even admissible here). This yields the non-optimal rate in Theorem 1.6 if we do not exclude the origin. Having Theorem 1.3 in mind however, we expect the maximizing ball to appear roughly at $B_1(0)$, where the error would be optimal again.

Proof of Corollary 1.7. As long as the origin is avoided by a fixed distance τ, the density of μ_∞^m is bounded as in the case of $m = 1$. Therefore, the only adjustment to the previous proof of Theorem 1.6 is that the error term in (3.15) is now also given by $\widetilde{h}_m(a) = \mathcal{O}(a^{-1})$. Since the remaining part of the proof remains untouched, Corollary 1.7 follows. $\quad\square$

References

[1] G. Akemann and J.R. Ipsen, *Recent exact and asymptotic results for products of independent random matrices*, Acta Physica Polonica B **46** (2015), no. 9, 1747. MR-3403839

[2] Gernot Akemann and Zdzislaw Burda, *Universal microscopic correlation functions for products of independent ginibre matrices*, Journal of Physics A: Mathematical and Theoretical **45** (2012), no. 46, 465201. MR-2993423

[3] Gernot Akemann, Zdzislaw Burda, and Mario Kieburg, *From integrable to chaotic systems: Universal local statistics of lyapunov exponents*, EPL (Europhysics Letters) **126** (2019), no. 4, 40001.

[4] Gernot Akemann and Milan Cikovic, *Products of random matrices from fixed trace and induced ginibre ensembles*, Journal of Physics A: Mathematical and Theoretical **51** (2018), no. 18, 184002. MR-3787269

[5] Nikita V. Alexeev, Friedrich Götze, and Alexander Tikhomirov, *On the singular spectrum of powers and products of random matrices*, Doklady mathematics, vol. 82, Springer, 2010, pp. 505–507.

[6] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni, *An introduction to random matrices*, vol. 118, Cambridge University Press, 2010. MR-2760897

[7] Charles Bordenave, *On the spectrum of sum and product of non-hermitian random matrices*, Electron. Commun. Probab. **16** (2011), 104–113. MR-2772389

[8] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart, *Concentration inequalities: A nonasymptotic theory of independence*, Oxford University Press, 2013. MR-3185193

[9] Paul Bourgade, Horng-Tzer Yau, and Jun Yin, *Local circular law for random matrices*, Probability Theory and Related Fields **159** (2014), no. 3-4, 545–595. MR-3230002

[10] Zdzisław Burda, Romuald A. Janik, and Bartek Waclaw, *Spectrum of the product of independent random gaussian matrices*, Physical Review E **81** (2010), no. 4, 041132. MR-2736204

[11] Djalil Chafaï, Adrien Hardy, and Mylène Maïda, *Concentration for Coulomb gases and Coulomb transport inequalities*, J. Funct. Anal. **275** (2018), no. 6, 1447–1483. MR-3820329

[12] Natalie Coston and Sean O'Rourke, *Gaussian fluctuations for linear eigenvalue statistics of products of independent iid random matrices*, Journal of Theoretical Probability (2019), 1–72. MR-4125967

[13] Marcel Fenzl and Gaultier Lambert, *Precise deviations for disk counting statistics of invariant determinantal processes*, 2020.

[14] Peter J. Forrester and Dong Wang, *Muttalib-borodin ensembles in random matrix theory – realisations and correlation functions*, Electron. J. Probab. **22** (2017), 43 pp. MR-3666017

[15] Jean Ginibre, *Statistical ensembles of complex, quaternion, and real matrices*, J. Mathematical Phys. **6** (1965), 440–449. MR-0173726

[16] Friedrich Götze and Jonas Jalowy, *Rate of convergence to the circular law via smoothing inequalities for log-potentials*, Random Matrices: Theory and Applications (2020), 2150026.

[17] 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** (2015), no. 02, 1550005. MR-3356883

[18] Friedrich Götze, Alexey Naumov, and Alexander Tikhomirov, *Local laws for non-hermitian random matrices and their products*, Random Matrices: Theory and Applications **9** (2020), no. 04, 2150004. MR-4133077

[19] Friedrich Götze and Alexander Tikhomirov, *Rate of convergence to the semi-circular law*, Probability Theory and Related Fields **127** (2003), no. 2, 228–276. MR-2013983

[20] Friedrich Götze and Alexander Tikhomirov, *The circular law for random matrices*, The Annals of Probability **38** (2010), no. 4, 1444–1491. MR-2663633

[21] Friedrich Götze and Alexander Tikhomirov, *On the asymptotic spectrum of products of independent random matrices*, 2010.

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

[23] Mario Kieburg, Arno B.J. Kuijlaars, and Dries Stivigny, *Singular value statistics of matrix products with truncated unitary matrices*, International Mathematics Research Notices **2016** (2015), no. 11, 3392–3424. MR-3556414

[24] Phil Kopel, Sean O'Rourke, and Van Vu, *Random matrix products: Universality and least singular values*, Annals of Probability **48** (2020), no. 3, 1372–1410. MR-4112718

[25] Holger Kösters and Alexander Tikhomirov, *Limiting spectral distributions of sums of products of non-hermitian random matrices*, 2015.

[26] Arno B.J. Kuijlaars and Lun Zhang, *Singular values of products of ginibre random matrices, multiple orthogonal polynomials and hard edge scaling limits*, Communications in Mathematical Physics **332** (2014), no. 2, 759–781. MR-3257662

[27] Dang-Zheng Liu, Dong Wang, and Lun Zhang, *Bulk and soft-edge universality for singular values of products of ginibre random matrices*, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, vol. 52, Institut Henri Poincaré, 2016, pp. 1734–1762. MR-3573293

[28] Dang-Zheng Liu and Yanhui Wang, *Phase transitions for infinite products of large non-hermitian random matrices*, 2019.

[29] Yudell L Luke, *Mathematical functions and their approximations*, Academic Press, 2014. MR-0501762

[30] Elizabeth S. Meckes and Mark W. Meckes, *A rate of convergence for the circular law for the complex ginibre ensemble*, Ann. Fac. Sci. Toulouse Math. (6) **24** (2015), no. 1, 93–117. MR-3325952

[31] Yuriy Nemish, *Local law for the product of independent non-hermitian random matrices with independent entries*, Electron. J. Probab. **22** (2017), 35 pp. MR-3622892

[32] Yuriy Nemish, *No outliers in the spectrum of the product of independent non-hermitian random matrices with independent entries*, Journal of Theoretical Probability **31** (2018), no. 1, 402–444. MR-3769819

[33] Sean O'Rourke and Alexander Soshnikov, *Products of independent non-hermitian random matrices*, Electronic Journal of Probability **16** (2011), 2219–2245. MR-2861673

[34] Sylvia Serfaty, *Microscopic description of log and Coulomb gases*, 2017. MR-3971159

[35] Terence Tao and Van Vu, *Random matrices: universality of local spectral statistics of non-hermitian matrices*, The Annals of Probability **43** (2015), no. 2, 782–874. MR-3306005

[36] Edmund Taylor Whittaker and George Neville Watson, *A course of modern analysis*, Cambridge University Press, 1996. MR-1424469