RATE OF CONVERGENCE TO THE CIRCULAR LAW VIA SMOOTHING INEQUALITIES FOR LOG-POTENTIALS

FRIEDRICH GÖTZE AND JONAS JALOWY

Abstract. The aim of this note is to investigate the Kolmogorov distance of the circular law to the empirical spectral distribution of non-Hermitian random matrices with independent entries. The optimal rate of convergence is determined by the Ginibre ensemble and is given by $n^{-1/2}$. A smoothing inequality for complex measures that quantitatively relates the uniform Kolmogorov-like distance to the concentration of logarithmic potentials is shown. Combining it with results from local circular laws, we apply it to prove nearly optimal rate of convergence to the circular law in Kolmogorov distance. Furthermore we show that the same rate of convergence holds for the empirical root measure of Weyl random polynomials.

1. Introduction

The (complex) empirical spectral distribution of a non-Hermitian random matrix with i.i.d. entries will converge to the uniform distribution on the complex disc as the size of the matrix tends to infinity. This circular law has a long history going back to Ginibre [Gin65], proving the special case of complex Gaussian entries. Later, Bai [Bai97] used Girko’s famous Hermitization Trick, introduced in [Gir85], to prove the circular law under extra density and moment assumptions. The density assumption was removed by Götze and Tikhomirov [GT07] and several reductions of the moment conditions were made in [GT10, PZ10, TV08]. Significant progress was possible due to the control of the smallest singular values in [RV08]. Ultimately, the circular law was proven under optimal second moment assumption by Tao and Vu (with an appendix by Krishnapur) [TV10]. We recommend the survey [BC12] for further discussions.

Random Matrix Theory is mostly concerned with universality phenomena, like the global universality in the circular law. Here, the limiting spectral distribution remains universal among a big class of entry distributions of the underlying matrix. Its local analogue has recently been investigated in [BYY14a, BYY14b, GNT17, TV15] among others. In this work, we address universality of the rate of convergence, containing local as well as global universality in a uniform and quantitative manner.

Consider a non-Hermitian random matrix $X = (X_{ij})_{1 \leq i,j \leq n}$ having independent real or complex entries $X_{ij}$, where in the complex case we additionally assume...
Below, we prove a for the underlying matrix.

\[ \mu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i(X/\sqrt{n})}, \]

where \( \delta_{\lambda} \) are Dirac measures in the eigenvalues \( \lambda_j \) of the scaled matrix \( X/\sqrt{n} \). The circular law states that if \( E X_{ij} = 0 \) and \( E |X_{ij}|^2 = 1 \), then \( \mathbb{P} \)-a.s. we have

\[ \mu_n \Rightarrow \mu_\infty, \]

where \( d\mu_\infty(z) = \frac{1}{\pi} \frac{1}{1+|z|^2} \, dz \)

is the uniform distribution on the complex disc. We are interested in the rate of convergence, more precisely in the Kolmogorov distances over balls

\[ D_n := D(\mu_n, \mu_\infty) := \sup_{z_0 \in \mathbb{C}, R > 0} |\mu_n(B_R(z_0)) - \mu_\infty(B_R(z_0))| \]

as \( n \to \infty \). Convergence in this distance coincides with weak convergence in the case of an absolutely continuous limit distribution, see Lemma 15 below. For the mean empirical spectral distribution \( \bar{\mu}_n = \mathbb{E} \mu_n \) of the so called Ginibre ensemble, i.e. \( X_{ij} \sim \mathcal{N}(0,1) \), it is easy to compute that the Kolmogorov distance satisfies

\[ D(\bar{\mu}_n, \mu_\infty) \approx 1/\sqrt{n}, \]

which turns out to be the optimal rate of \( \mu_n \) to the circular law. We write \( A \asymp B \) if \( c |B| \leq |A| \leq C |B| \) for some constants \( c, C > 0 \). Interestingly, if one avoids the edge of \( B_1(0) \) by a fixed distance \( \varepsilon \), then the rate of convergence is exponentially fast

\[ \sup_{B_R(z_0) \subseteq \mathbb{C} \setminus B_{1+\varepsilon}(0) \text{ or } B_R(z_0) \subseteq B_{1-\varepsilon}(0)} |\bar{\mu}_n(B_R(z_0)) - \mu_\infty(B_R(z_0))| \lesssim e^{-n\varepsilon^2}. \]

We prove these statements for the Ginibre ensemble in the Appendix A, Lemma 14. Here and in the sequel \( \lesssim \) will always denote an inequality that holds up to a parameter-independent constant \( c > 0 \) that may differ in each occurrence. Nevertheless we cannot expect an exponentially fast rate of convergence for the non-averaged empirical spectral distribution \( \mu_n \), since it is still sensitive to individual eigenvalue fluctuations. In particular, for each fixed set of eigenvalues \( \{\lambda_i\}_{i \leq n} \) we may select a ball of radius \( (10\sqrt{n})^{-1} \) contained in \( B_1(0) \) such that it does not cover any eigenvalue and obtain \( D_n \gtrsim 1/n \). Heuristically, the typical distance of \( n \) uniformly distributed eigenvalues 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 and hence we expect \( D_n \) to be of order \( n^{-1/2} \). In our main result, see Theorem 5 below, we prove a rate of convergence of order \( n^{-1/2+\varepsilon} \) for non-Gaussian entry distributions of the underlying matrix.

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

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

satisfies \( \Delta U = -2\pi \mu \)

in the sense of distributions. Obviously the logarithmic potential of a measure is superharmonic in \( \mathbb{C} \), harmonic outside the support of \( \mu \) and is only unique up to
addition of harmonic functions. 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

$$U_n(z) = -\frac{1}{n} \sum_{j=1}^{n} \log |\lambda_j - z| = -\frac{1}{n} \log \left| \det \left( \frac{1}{\sqrt{n}} X - z \right) \right| = -\frac{1}{n} \log \det \left( \frac{1}{\sqrt{n}} X - z \left| \frac{1}{\sqrt{n}} X - z \right|^* \right) = -\int_{0}^{\infty} \log(x) d\nu^z_n(x), \quad (4)$$

where $\nu^z_n$ is the empirical singular value distribution of the shifted matrix $X/\sqrt{n} - z$.

Due to this fact, all the information on the complex spectrum of $X/\sqrt{n}$ is stored in the real and positive spectra of $(X/\sqrt{n} - z)(X/\sqrt{n} - z)^*$ for all shifts $z$. Note that its symmetrized version around 0 is the empirical eigenvalue distribution of the Hermitian matrix

$$V(z) = \begin{bmatrix} 0 & (X/\sqrt{n} - z)^* \\ (X/\sqrt{n} - z) & 0 \end{bmatrix}.$$ 

Under certain conditions on the matrix entries, the logarithmic potential $U_n$ concentrates around the logarithmic potential $U_\infty$ of the circular law given by

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

Let us fix some notation and the above-mentioned conditions.

**Definition 1.** 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 $\text{Re}X_{ij}$ and $\text{Im}X_{ij}$ to be independent.

(A) Additionally we say $X$ satisfies condition (A) if it has independent entries $X_{ij}$ with mean zero, variance $E|X_{ij}|^2 = 1$, subexponential tails

$$P(|X_{ij}| \geq t) \leq C \exp(-ct)$$

for some fixed $c, C > 0$ and match either the real or complex Gaussian moments up to third order, i.e.

$$E X_{ij} = E \text{Re}(X_{ij})^3 = E \text{Im}(X_{ij})^3 = 0$$

and either $E |\text{Re}X_{ij}|^2 = E |\text{Im}X_{ij}|^2 = 1/2$ or $E |\text{Re}X_{ij}|^2 = 1, E |\text{Im}X_{ij}|^2 = 0.$

(B) We say $X$ satisfies condition (B) if it has independent entries, where

$$\max_{i,j} |E X_{ij}| \leq n^{-1-\varepsilon} \quad \text{and} \quad \max_{i,j} \left| 1 - E |X_{ij}|^2 \right| \leq n^{-1-\varepsilon}$$

for some $\varepsilon > 0$ and furthermore

$$\max_{i,j,n} E |X_{ij}|^{4+\delta} < \infty$$

for some $\delta > 0$.

Note that in contrast to Wigner matrices, the distributions of the entries may be different and clearly, (A) implies (B). The following concentration of the logarithmic potentials has been proven in [TV15], Theorem 25.
\textbf{Theorem 2 (TV15).} If $X$ satisfies (A), then for every $\varepsilon, \tau, Q > 0$ there exist a constant $c > 0$ such that
\[
P \left( |U_n(z) - U_\infty(z)| \leq cn^{-(1-\varepsilon)} \right) \geq 1 - n^{-Q}
\] holds uniformly for $z \in B_{1+\tau}(0)$.

Such results on the concentration of the logarithmic potentials are used to derive local circular laws. In [GNT17], the assumptions have been weakened, the rate has been improved and the result has been generalized to products of independent matrices, but unfortunately the region is restricted to the bulk $\|z\| - 1 \geq \tau$.

\textbf{Theorem 3 ([GNT17]).} If $X$ obeys (B), then for every $\tau, Q > 0$ there exist a constant $c > 0$ such that
\[
P \left( |U_n(z) - U_\infty(z)| \leq c \log^4 n \right) \geq 1 - n^{-Q}
\] holds uniformly in $\{z \in B_{1+\tau^{-1}}(0) : |1 - |z|| \geq \tau\}$.

Since this is not explicitly worked out in [GNT17], we will derive it in Appendix A based on the results proved in this paper.

2. Main Results

Consider a sequence of probability measures $\mu_n$ on $\mathbb{C}$ with logarithmic potentials $U_n$. If $U_n$ converges pointwise to some function $U : \mathbb{C} \to (-\infty, \infty]$ and if $U_n$ is locally uniformly Lebesgue integrable, then (by continuity of $\Delta$ on the space of distributions) there exist a probability measure $\mu = -\frac{1}{2\pi} \Delta U$ on $\mathbb{C}$ such that $\mu_n$ converges weakly to $\mu$. The following smoothing inequality quantifies this statement by relating $D_n$ to the concentration of logarithmic potentials.

\textbf{Proposition 4.} Let $\mu, \nu$ be probability measures on $\mathbb{C}$ with $\text{supp} \nu \subseteq B_K(0)$ for some $K > 0$, let $U_\mu, U_\nu$ be their logarithmic potentials and fix some $1 \leq p \leq \infty$. For any $a \geq 1/2$ we have
\[
D(\mu, \nu) \lesssim a^{1+1/p} \|U_\mu - U_\nu\|_{L^p(B_{K+1/a}(0))} + \sup_{R \geq 0, z_0 \in \mathbb{C}} \nu \left( R \leq |z - z_0| \leq R + 1/a \right).
\]

In the same manner it is possible to show an analogue for the classical Kolmogorov distance between 2-dimensional distribution functions, see Corollary 12. For measures $\mu, \nu$ on $\mathbb{R}$, where $\nu$ has a bounded density, Dinh and Vu showed in [DV17] another direct relation of similar type
\[
|\mu(I) - \nu(I)| \lesssim \|U_\mu - U_\nu\|_{L^\infty(\text{supp} \nu)}^{1/2}
\] for all intervals $I \subseteq \mathbb{R}$ and it was used to show a rate of convergence in Wigner’s semicircular law and the Marchenko-Pastur law. Proposition 4 may be of independent interest, since it can be considered as a complex counterpart of other smoothing inequalities of distributions $\mu, \nu$ on the real line. For instance in the case of Fourier transforms $\varphi_\mu(t) = \int e^{itx} d\mu(x)$, the well known \textit{Berry-Essen inequality}
\[
\sup_{x \in \mathbb{R}} |(\mu - \nu)((-\infty, x]])| \lesssim \int_a^\infty \left| \frac{\varphi_\mu(t) - \varphi_\nu(t)}{t} \right| dt + \sup_{x \in \mathbb{R}} \nu((x, x + c/a])
\] leads to a rate of convergence of order $1/\sqrt{n}$ in the Central Limit Theorem, when choosing $\nu = \mathcal{N}(0,1)$ and $\mu = \mathbb{P}_{S_n}$ for the normalized sum $S_n = n^{-1/2} \sum_{k=1}^n X_k$.
of i.i.d. random variables $X_k$ with $\mathbb{E}X_1 = 0, \mathbb{E}X_1^2 = 1$ and finite third moment $\mathbb{E}X_1^3 < \infty$. In Random Matrix Theory, Bai's inequality is a handy tool to profit from control of Stieltjes' transforms $m_\mu(z) = \int \frac{1}{x-z}d\mu(x)$ that can be simplified to

$$\sup_{x \in \mathbb{R}} |(\mu - \nu)((-\infty, x])| \lesssim \int |m_\mu - m_\nu| (t + i/a)dt + \sup_{x \in \mathbb{R}} \nu((x, x + c/a]). \quad (8)$$

Roughly speaking, [BS10] uses $a \simeq \sqrt{n}$ to show a rate of convergence of order $1/\sqrt{n}$ for the Kolmogorov distance in Wigner's semicircle law under finite sixth moment condition. Using an improved, but more involved smoothing inequality, it is shown in [GT16] that the optimal rate of convergence to the semicircle distribution is given by $\mathcal{O}(1/n)$.

All smoothing inequalities (7), (8) and Proposition 4 are used to derive convergence rates under moment conditions and they share the essential structure of bounding the Kolmogorov distance by the distance of certain integral-transforms and an additional maximal shell probability of width $\mathcal{O}(1/a)$ with respect to the “limit distribution”. Regarding Proposition 4, we consider the distributions $\mu = \mu_n, \nu = \nu_\infty$ from the introduction and choose $a = \sqrt{n}, K = 1$. In this case we see that the remainder term is of order $n^{-1/2}$ and a rate of convergence for $D_n$ follows.

It is important to carefully distinguish between events holding with high probability uniformly in $z$ and uniform events that hold w.h.p.. The former leads to local circular laws like Theorem 20 in [TV15] (see also Corollary 13 below) and hence do not imply the latter, which is an estimate on $D_n$. Contrary to local circular laws, a bound on $D_n$ w.h.p. allows to choose the ball $B_R(z_0)$ depending on the random sample of the eigenvalues $(\lambda_j(X(\omega)/\sqrt{n}))_j$. Similarly, the statement of Theorem 3 should not be confused with an assertion about the uniform term $\sup_{z \in B_K(0)} |U_n(z) - U_\infty(z)|$, since it equals $\infty$ whenever an eigenvalue lies in $B_K(0)$. Due to this fact one cannot simply take $p = \infty$ in Proposition 4 in order to obtain the following result.

**Theorem 5.** If condition (A) holds, then for every (small) $\varepsilon > 0$ and (large) $Q > 0$

$$\mathbb{P}(D_n \leq n^{-1/2+\varepsilon}) \geq 1 - n^{-Q}. \quad (9)$$

holds for sufficiently large $n$, where $D_n = \sup_{z_0 \in \mathbb{C}, R > 0} |(\mu_n - \mu_\infty)(B_R(z_0))|.$

By virtue of Corollary 12, the following Kolmogorov distance analogue holds.

**Theorem 6.** If condition (A) holds, then for every $\varepsilon, Q > 0$

$$\mathbb{P}(d_n \leq n^{-1/2+\varepsilon}) \geq 1 - n^{-Q}. \quad (10)$$

holds for sufficiently large $n$, where $d_n = \sup_{s,t \in \mathbb{R}} |(\mu_n - \mu_\infty)((-\infty, s] \times (-\infty, t])|.$

Invoking Theorem 3, we prove a rate of convergence result weakening the conditions of the last statements at the cost of excluding sets close to the edge.

**Theorem 7.** If condition (B) holds, then for every $\varepsilon, \tau, Q > 0$

$$\mathbb{P}(D_\tau^\circ \leq n^{-1/2+\varepsilon}) \geq 1 - n^{-Q}$$

holds for sufficiently large $n$, where $D_\tau^\circ = \sup_{B_R(z_0) \subseteq B_{1-\tau}(0)} |(\mu_n - \mu_\infty)(B_R(z_0))|.$
Tao and Vu showed in [TV08] that with probability 1 the Kolmogorov distance $d_n$ of the 2-dimensional distribution functions is of order $n^{-\eta}$ for some unknown $\eta > 0$, which holds for finite $2 + \epsilon$-moments of the entries. Comparing this to Corollary 6, we see that a nearly optimal rate of convergence is obtained in (10) which holds with overwhelming probability. On the other hand a much stronger moment assumption for the entries is needed. In particular, this explicit rate of convergence gives a partial answer to an open problem mentioned in [TV09]. In the special case of Gaussian entries, i.e. for the Ginibre ensemble, $P$-a.s. convergence rates of order $\sqrt{\log n}/n^{1/4}$ in $p$-Wasserstein distance for $1 \leq p \leq 2$ have been proven in [MM15].

In [CHM16], Chafaï, Hardy and Maïda studied invariant $\beta$-ensembles with external potential $V$ instead of independent-entry matrices. Their result implies a rate of convergence to the limiting measure with density $c\Delta V$ of order $\sqrt{\log n}/n$ with respect to the bounded Lipschitz metric and the 1-Wasserstein distance. The paper [CHM16] is also based on an inequality between distances of measures to their energy, i.e. integrated logarithmic potential, similar to Proposition 4, however it relies critically on the existence of a confining potential, hence a joint probability density function for the eigenvalues. Note that their result is given for a Coulomb gas point process in arbitrary dimension $d > 1$, yielding a bound of order $n^{-1/d}$ up to logarithmic factors. This coincides with the rate of order $1/n$ for the semicircle law for $d = 1$ as well as the optimal order $1/\sqrt{n}$ in the circular law and can also be interpreted as mentioned in the introduction.

3. Application to Random Polynomials

In this section we will apply the Smoothing Inequality to the empirical distribution of roots of random polynomials in order to obtain the same rate of convergences to the circular law as before. In the previous section we considered the roots of the characteristic polynomial of a random matrix, where the coefficients of the polynomial exhibit specific dependencies. We begin by replacing the independence condition on the matrix entries by independent coefficients in the polynomial.

**Definition 8.** Given $n \in \mathbb{N}$ many complex numbers $c_0, \ldots, c_n$ and i.i.d. centered complex random variables $\xi_0, \ldots, \xi_n$ with $E|\xi_k|^2 = 1$, we define the random polynomial $f_n : \mathbb{C} \to \mathbb{C}$ by

$$f_n(z) = \sum_{k=0}^{n} c_k \xi_k z^k.$$ 

In particular we will work with so called Weyl (or Flat) polynomials $f_n^W$ corresponding to $c_k = \sqrt{n^k/k!}$. By analogy to the Introduction, we associate to a random polynomial $f_n$ its multiset of zeros $\Lambda := \{ \lambda \in \mathbb{C} : f_n(\lambda) = 0 \}$ taking their multiplicities into account and its empirical measure given by

$$\mu_{f_n} = \frac{1}{n} \sum_{\lambda \in \Lambda} \delta_\lambda.$$ 

It should be remarked that $\mu_{f_n}$ is not necessarily normalized, since a random polynomial may have degree $\text{deg}(f_n) < n$. Unsurprisingly this does not affect the large $n$ limit, since $n - \text{deg}(f_n) \in O(1)$ $P$-a.s. and as in [IZ13], we may always assume $P(\xi_0 = 0) = 0$, since otherwise we may restrict ourselves to $\{ \text{deg}(f_n) = k, \min\{j \leq n : \xi_j \neq 0\} = l \}$. 
The circular law for the empirical root measure of Weyl polynomials has been established in [KZ14] by Kabluchko and Zaporozhets, see also [FH99] for the Gaussian case, stating

$$\mu_{f_n}^W \Rightarrow \mu_\infty \text{ \ P-a.s..}$$

Note that their result holds for much more general random analytic functions and under the much weaker condition of the coefficients having finite logarithmic moments $\mathbb{E}\log(1 + |\xi_0|) < \infty$.

We aim to quantify this result by showing a rate of these convergences of order $n^{-1/2+\varepsilon}$ by using results about logarithmic potentials. Since local universality for certain random polynomials has been proven in by Tao and Vu using concentration of logarithmic magnitudes $\log|f_n|$, we can apply the same methods as before.

We denote $U_n = -\frac{1}{n} \log|f_n|$ and rephrase Lemma 12.1 from [TV14]: For every $\varepsilon, \delta, \tau, Q > 0$ there exist a constant $c > 0$ such that

$$\mathbb{P}\left(|U_n^W(z) - U_\infty(z)| + 1/2| \leq cn^{-(1-\varepsilon)}\right) \geq 1 - n^{-Q} \quad (11)$$

holds uniformly for $n^{-1/2+\delta} \leq |z| \leq 1 + \tau$. The origin has to be avoided, since the distribution of $U_n^W(0) = -\frac{1}{n} \log|\xi_0|$ around 0 stays arbitrary. In particular, the bound (11) will not hold if $z = 0$ if $\mathbb{P}(\xi_0 = 0) > 0$. Due to the application of the Monte Carlo method we still need a technical assumption on the concentration of $\xi_0$ near $z = 0$ in the following rate of convergence result which we deduce from a variant of Smoothing inequality 4.

**Theorem 9.** If $\mathbb{E}|1/\xi_0|^\delta < \infty$ for some $\delta > 0$, then for every $\varepsilon, Q > 0$ and sufficiently large $n$ we have

$$\mathbb{P}(D(\mu_{n}^W, \mu_\infty) \leq n^{-1/2+\varepsilon}) \geq 1 - n^{-Q}.$$

It seems likely that other polynomials, like elliptic polynomials, omit the same asymptotics to their corresponding limit root distributions, but we focus on circular laws in this work.

4. **Proofs of the Smoothing Inequalities**

We will proof the following slightly more general statement that covers all variants we need.

**Theorem 10.** Let $\mu, \nu$ be probability measures on $\mathbb{C}$ with logarithmic potentials $U_\mu, U_\nu$ respectively (i.e. the distributional Poisson equation (3) holds), fix $1 \leq p \leq \infty$ and for some $z^* \in \mathbb{C}$, $K > 0$, $\eta \geq 0$ define the rings $V = B_K(z^*) \setminus B_{2\eta/a}(z^*)$ and $V' = B_{K+2/a}(z^*) \setminus B_{\eta/a}(z^*)$. For any $a > 1$

$$D(\mu, \nu) \leq a^{1/p} \|U_\mu - U_\nu\|_{L^p(V')} + \mu(V^c) + \nu(V^c) + \sup_{R \geq 0, z_0 \in \mathbb{C}} \nu(z \in V' : R \leq |z - z_0| \leq R + \max(2, \eta/a)).$$

Here, $\eta \neq 0$ is only needed for the applications to random polynomials, where the logarithmic potential near the origin cannot be controlled.

**Proof.** First, note that

$$\sup_{R \geq 0, z_0 \in \mathbb{C}} |(\mu - \nu)(B_R(z_0))| \leq \sup_{R \geq 0, z_0 \in \mathbb{C}} |(\mu - \nu)(B_R(z_0) \cap V)| + \mu(V^c) + \nu(V^c),$$
hence we have to estimate the first term. Fix some \( a > 1 \), let \( \varphi \in C^\infty(\mathbb{R}) \) be nonnegative with \( \text{supp} \varphi \subseteq [-1,1] \) and \( \int \varphi = 1 \), and define \( \varphi_a(\rho) = a \varphi(\rho/a) \). For arbitrary \( R > 0 \) and \( z_0 \in \mathbb{C} \) we mollify the indicator function appearing in \( D(\mu, \nu) \) via the rotationally invariant approximation

\[
 f_1(z) := (1_{(-\infty, R-1/a]} * \varphi_a)(|z - z_0|) \\
 \leq 1_{B_R(z_0)}(z) \\
 \leq (1_{(-\infty, R+1/a]} * \varphi_a)(|z - z_0|) =: f_2(z),
\]

where we choose \( f_1 \equiv 0 \) if \( R \leq 2/a \) for smoothness reasons. Furthermore we will approximate \( 1_V \) by smooth functions \( h_1 \) from inside and by \( h_2 \) from outside, more precisely define

\[
 h_1(z) := \begin{cases} 
 (1_{[\eta/2a, \infty)} * \varphi_{2a/\eta}) \cdot (1_{(-\infty, K-1/a]} * \varphi_a)(|z - z^*|) & \text{if } \eta > 0 \\
 1_{(\infty, K-1/a]} * \varphi_a(|z - z^*|) & \text{if } \eta = 0
\end{cases}
\]

\[
 h_2(z) := \begin{cases} 
 (1_{[\eta/2a, \infty)} * \varphi_{2a/\eta}) \cdot (1_{(-\infty, K+1/a]} * \varphi_a)(|z - z^*|) & \text{if } \eta > 0 \\
 1_{(\infty, K+1/a]} * \varphi_a(|z - z^*|) & \text{if } \eta = 0
\end{cases}
\]

We apply \( h_1 f_1 \leq 1_{B_R(z_0) \cap V} \) and integration by parts (in other words we use the definition of the distributional Poisson equation \((3)\)) back and forth to obtain

\[
 \mu(B_R(z_0) \cap V) \geq \int h_1 f_1 d\mu = -\frac{1}{2\pi} \int \Delta(h_1 f_1) U_{\mu} d\lambda \\
 = -\frac{1}{2\pi} \int \Delta(h_1 f_1) (U_\mu - U_\nu) d\lambda - \int (1_{B_R(z_0) \cap V} - h_1 f_1) d\nu + \int 1_{B_R(z_0) \cap V} d\nu
\]

where \( \lambda \) denotes the Lebesgue measure on \( \mathbb{C} \). A rough estimate of the error of approximation yields for the second term

\[
 \int (1_{B_R(z_0) \cap V} - h_1 f_1) d\nu \leq \nu(z \in V' : R - 2/a \leq |z - z_0| \leq R) + \nu(V' \setminus V) \\
 \leq 3 \sup_{R \geq 0, z_0 \in \mathbb{C}} \nu(z \in V' : R \leq |z - z_0| \leq R + \max(2, \eta)/a) =: 3M_\nu(a).
\]

We use Hölder’s inequality to estimate the first term, implying

\[
 (\mu - \nu)(B_R(z_0) \cap V) \geq -\frac{1}{2\pi} \|\Delta(h_1 f_1)\|_{L^p} \|U_\mu - U_\nu\|_{L^p} - 3M_\nu(a), \quad (12)
\]

where \( L^p = L^p(V') \), \( L^q = L^q(V') \) (we omit \( V' \) in the sequel), \( 1/p + 1/q = 1 \) and \( R > 0, z_0 \in \mathbb{C} \) are still arbitrary. Noting \( \mu(B_R(z_0) \cap V) \leq \int h_2 f_2 d\mu \) and taking the same route for \( h_2 f_2 \) as for \( h_1 f_1 \), we obtain the same upper bound, i.e.

\[
 -\frac{1}{2\pi} \|\Delta(h_1 f_1)\|_{L^q} \|U_\mu - U_\nu\|_{L^p} - 3M_\nu(a) \\
 \leq (\mu - \nu)(B_R(z_0) \cap V) \\
 \leq -\frac{1}{2\pi} \|\Delta(h_2 f_2)\|_{L^q} \|U_\mu - U_\nu\|_{L^p} + 3M_\nu(a). 
\]

Therefore it remains to control

\[
 \|\Delta(h_2 f_2)\|_{L^q} \leq \|h_j \Delta f_j\|_{L^q} + 2 \|\nabla h_j \cdot \nabla f_j\|_{L^q} + \|f_j \Delta h_j\|_{L^q}.
\]

We see that the supports of all three functions are (unions of) ring-segments, e.g. \( V' \cap (B_{R+2/a}(z_0) \setminus B_R(z_0)) \) for \( h_2 \Delta f_2 \), with length at most \( 2\pi(K + 2/a) \) and the
width equals $\max(2, \eta)/a$. Hence uniformly in $R > 0$ and $z_0 \in \mathbb{C}$, the size of the area of integration is bounded by $cK \max(2, \eta)/a$ and we arrive at
\[ \|\Delta(h_j f_j)\|_{L^p} \leq (cK \max(2, \eta)/a)^{1/q} \left( \|\Delta f_j\|_{L^\infty} + 2 \|\nabla h_j \cdot \nabla f_j\|_{L^\infty} + \|\Delta h_j\|_{L^\infty} \right) \]

With our choice of $f_j$ and $h_j$, the radial derivatives become fairly simple, e.g.
\[ \partial_r(f_2(z + z_0)) = \partial_r \int_{|z| - R - 1/a}^\infty \phi_a(\rho) d\rho = -a \phi(a |z| - aR - 1). \]

Due to the rotational symmetry of $f_2$, we have $\|\nabla f_2\|_{L^\infty} \leq \|\phi_a\|_{L^\infty} \lesssim a$ and again exploiting rotational symmetry it follows that the maximal curvature is attained in radial direction, i.e.
\[ \|\Delta f_2\|_{L^\infty} = \sup_{r > 0} |\partial_r^2 f_2(z_0 + r)| = a^2 \|\phi'\|_{L^\infty}. \]

The same bounds hold for $j = 1$, whereas in the case of $h$ instead of $f$ we replace $a$ by $\max(a, 2a/\eta)$. Finally we conclude
\[ \|\Delta h_j f_j\|_{L^p} \lesssim a^2 (K/a)^{1/q} \lesssim K^{1-1/p} a^{1+1/p}, \quad (14) \]

where the implicit constant in the last $\lesssim$ depends on $p, \eta$ and $\phi$ only. The claim now follows from taking the supremum over $R > 0$ and $z_0 \in \mathbb{C}$ in (13). \hfill $\square$

We retrieve Proposition 3 by taking $\eta = 0$, $\varphi = 0$, $\nu(V^c) = 0$, replacing $a$ by $2a$ for simplicity and noting that for probability distributions
\[ \mu(V^c) = (\nu - \mu)(V) \leq \sup_{R \geq 0, z_0 \in \mathbb{C}} |(\mu - \nu)(B_R(z_0) \cap V)| \]

is what we have estimated in the previous proof. In fact, by setting $\eta = 0$, we get a local smoothing inequality that makes it possible to invoke Theorem 3.

**Corollary 11.** Let $\mu, \nu$ be probability measures on $\mathbb{C}$ with logarithmic potentials $U_\mu, U_\nu$ respectively, and fix some $z^* \in \mathbb{C}$, $K, \tau > 0$ and $1 \leq p \leq \infty$. There exists a constant $c > 0$ such that for any $a > 1 \wedge \tau^{-1}$
\[
\sup_{B_R(z_0) \subseteq B_K(z^*)} |(\mu - \nu)(B_R(z_0))| \leq c a^{1+1/p} \left\|U_\mu - U_\nu\right\|_{L^p(B_K(z^*))} \\
+ \sup_{R \geq 0, z_0 \in \mathbb{C}} |\nu(z \in B_K(z^*): R \leq |z - z_0| \leq R + 2/a|. 
\]

Moreover, the method of proof extends to the case of the classical Kolmogorov distance between 2-dimensional distribution functions.

**Corollary 12.** Let $\mu, \nu$ be probability measures on $\mathbb{C}$ with $\text{supp} \nu \subseteq [-K, K]^2$ for some $K > 0$, let $U_\mu, U_\nu$ be their logarithmic potentials and fix some $\tau > 0$ and $1 \leq p \leq \infty$. There exists a constant $c > 0$ such that for any $a > 1$
\[
\sup_{s, t \in \mathbb{R}} |(\mu - \nu)((-\infty, s] \times (-\infty, t])| \leq c a^{1+1/p} \left\|U_\mu - U_\nu\right\|_{L^p([-K - \tau, K + \tau]^2)} \\
+ 3 \sup_{s, t \in \mathbb{R}} |\nu((s, s + 2/a] \times \mathbb{R}) \cup (\mathbb{R} \times [t, t + 2/a])|. 
\]

**Proof.** We continue with the same notation as in the last proof and exploit the same ideas. Define now
\[ f_1(z) := \mathbb{1}_{(-\infty, -s-1/a]} \ast \phi_a(\text{Re} z) \cdot \mathbb{1}_{(-\infty, -t-1/a]} \ast \phi_a(\text{Im} z) \]
\[ \leq \mathbb{1}_{(-\infty, -s+1/a]} \ast \phi_a(\text{Re} z) \cdot \mathbb{1}_{(-\infty, -t+1/a]} \ast \phi_a(\text{Im} z) =: f_2(z), \]
and \( h(z) = \mathbb{1}_{[-K-\tau/2,K+\tau/2]} \ast \varphi_{\tau/2}(\text{Re}z) \cdot \mathbb{1}_{[-K-\tau/2,K+\tau/2]} \ast \varphi_{\tau/2}(\text{Im}z) \). Here, if \( \nu \) has compact support, we do not need \( h_1 \) in order to restrict ourselves to \( V \). By similar arguments as above, e.g. \( h_{f_1} \leq \mathbb{1}_{[-\infty,a] \times (-\infty,t]} \), we obtain
\[
(\mu - \nu)((-\infty, s] \times (-\infty, t]) \geq \frac{1}{2\pi} \| \Delta(h_{f_1}) \|_{L^p} \| U_{\mu} - U_{\nu} \|_{L^p} - M_\nu(a),
\]
where now \( M_\nu(a) = \sup_{s,t \in \mathbb{R}} \nu((s, s+2/a] \times \mathbb{R}) \cup (\mathbb{R} \times [t, t+2/a]) \) and we abbreviated \( L^p = L^p([-K-\tau, K+\tau]^2), L^q = L^q([-K-\tau, K+\tau]^2) \). For a short moment, consider
\[
f_{f_1}^0(z) = \mathbb{1}_{[-K+1/a, K-1/a]} \ast \varphi_a(\text{Re}z) \cdot \mathbb{1}_{[-K+1/a, K-1/a]} \ast \varphi_a(\text{Im}z)
\]
which analogously to the idea mentioned before Corollary 11 yields
\[
1 - \mu([-K, K]^2) = (\nu - \mu)([-K, K]^2) \leq \frac{1}{2\pi} \| \Delta(f_{f_1}^0) \|_{L^q} \| U_{\mu} - U_{\nu} \|_{L^p} + 2M_\nu(a).
\]
We conclude
\[
\frac{1}{2\pi} \| \Delta(h_{f_1}) \|_{L^p} \| U_{\mu} - U_{\nu} \|_{L^p} - M_\nu(a)
\]
\[
\leq (\mu - \nu)((-\infty, s] \times (-\infty, t])
\]
\[
\leq \frac{1}{2\pi} \left( \| \Delta(h_{f_2}) \|_{L^q} + \| \Delta(f_{f_1}^0) \|_{L^q} \right) \| U_{\mu} - U_{\nu} \|_{L^p} + 3M_\nu(a).
\]
Consequently it remains to derive similar estimates \( \| \Delta(h_{f_j}) \|_{L^q} \leq a^{1+1/p} \) using the same arguments as before. We omit the details here. \( \square \)

5. Proof of the Rates of Convergence

Proof of Theorem 5. Without loss of generality \( \varepsilon < 4 \), we choose \( p > 4/\varepsilon \) and apply Proposition 4 to \( \mu = \mu_n, \nu = \mu_\infty, K = 1 \) and \( a = \sqrt{n} \),
\[
D_n \lesssim n^{1/2+\varepsilon/2} \| U_n - U_\infty \|_{L^p(B_1(0))} + \sup_{R \geq 0, z_0 \in \mathbb{C}} \mu_\infty \left( R \leq |z - z_0| \leq R + 2n^{-1/2} \right).
\]
Since \( \mu_\infty \) has bounded support and bounded density it is clear that the second term is of order \( O(n^{-1/2}) \). In order to obtain a bound of the \( L^p(B_1(0)) \)-norm of the log potentials from the pointwise estimate in Theorem 2, we adapt the Monte Carlo sampling method which was used in [TV15] (in a different form); we approximate
\[
\int I(z)^p dz := \frac{1}{\pi(1+\tau)^2} \int_{B_1+\tau(0)} |U_n(z) - U_\infty(0)|^p dz \approx \frac{1}{m} \sum_{j=1}^m I(z_j)^p =: S_m,
\]
where \( (z_j)_{j=1,...,m} \) are independent random variables (also independent of \( X_{ij} \)) uniformly distributed on \( B_{1+\tau}(0) \). More precisely we will show that for every \( Q > 0 \)
\[
\left| \int I(z)^p dz - S_m \right|^{1/p} \lesssim n^{-1}
\]
(15)
as well as
\[
|S_m|^{1/p} \lesssim n^{-1+\varepsilon/2}
\]
(16)
holds with probability at least $1 - n^{-Q}$ for some large $n$-dependent $m$. Assuming (15) and (16) are true, we would get

$$P(D_n \geq cn^{-1/2+\varepsilon})$$

$$\leq P\left(cn^{1/2+\varepsilon}/\left(\int I(z)^p dz - S_m \right)^{1/p} + |S_m|^{1/p} \right) + cn^{-1/2} \geq cn^{-1/2+\varepsilon}$$

$$\leq P\left(\left|\int I(z)^p dz - S_m \right|^{1/p} \geq cn^{-1/2+\varepsilon}\right) + P\left(|S_m|^{1/p} \geq cn^{-1/2+\varepsilon}\right) \leq n^{-Q}$$

proving the claim.

Let's turn to the proof of (15). First, we restrict ourselves to the set of polynomially bounded eigenvalues. On the one hand the largest absolute value of eigenvalues $|\lambda|_{\text{max}}$ is bounded by the largest singular value $s_{\text{max}}$ and on the other hand for every $Q > 0$ we have

$$P(s_{\text{max}} \geq n^{(Q+1)/2}) \leq \frac{1}{n^{Q+1}} \mathbb{E}\|X/\sqrt{n}\|^2 \leq \frac{1}{n^{Q+2}} \sum_{ij} \mathbb{E}|X_{ij}|^2 \leq n^{-Q}, \quad (17)$$

where the operator norm $\|\cdot\|$ has been estimated by the Hilbert Schmidt norm. We freeze the coefficients $X_{ij}$ and use Chebyshev's inequality for the probability measure conditioned on $X$

$$P\left(|S_m - \int I(z)^p dz \right)^{1/p} \geq \frac{cn}{n} |X| \leq \frac{n^{2p}}{4p} \mathbb{Var}(S_m |X) \leq \frac{n^{2p}}{mc^{2p}} \mathbb{Var}(I^p |X).$$

The variance of $I^p$ given $X$ is given by

$$\mathbb{Var}(I^p |X) \leq \mathbb{E}(I^p |X) \leq \int |U_n(z)|^{2p} + |U_{n+1}(z)|^{2p} dz.$$
It remains to show (16). To this end we use Theorem 2 with an adjusted error probability stating
\[ P(I(z) \geq cn^{-1+\varepsilon/2}) \leq n^{-2p-5Q/2-1} \] (18)
uniformly in \( B_{1+\tau}(0) \). If \( I(z_j) \leq n^{-1+\varepsilon/2} \) for all \( j = 1, \ldots, n \) then \( |S_m|^{1/p} \leq n^{-1+\varepsilon/2} \) which implies
\[ P(|S_m|^{1/p} \geq n^{-1+\varepsilon/2}) \leq \sum_{j=1}^{m} P(I(z_j) \geq n^{-1+\varepsilon/2}) \leq c m n^{-2p-5Q/2-1} = cn^{-Q}. \]
The proof is now complete, since these constants may be absorbed by the \( n^{-Q} \) (respectively \( n^{-\varepsilon} \)) term for some slightly larger \( Q \) (respectively smaller \( \varepsilon \)). \( \square \)

It may be possible to prove similar results by using Riemann sums or by a direct approach without a separated smoothing inequality, but we do not pursue it here. Analogously, Theorem 6 follows from Corollary 12 and Theorem 7 follows from Corollary 11. The details are exactly the same as above and we skip them. Moreover using the same techniques it's possible to show the following version of a local circular law. Compared to [GNT17] it improves the statement to hold with overwhelming probability but replaces the constant \( \|\Delta f\|_{L^q} \) by \( \|\Delta f\|_{L^p} \) and is stated for a single matrix, instead for a product of \( m \) many.

**Corollary 13** (Local circular law). Let \( q > 1 \), \( z_0 \in B_{1+\tau-1}(0) \) with \( 1 - |z_0| \geq \tau \), \( f : \mathbb{C} \to \mathbb{R}_+ \) be a bounded smooth function, which is compactly supported with \( \|f\|_{L^\infty} \leq n^c \) for some constant \( c > 0 \). Define the function \( f_{z_0}(z) := n^{2s}f((z - z_0)n^s) \) which zooms into \( z_0 \) at speed \( s \in (0, 1/2) \). For any \( Q > 0 \) there exist a constant \( c > 0 \) such that
\[ P\left( \frac{1}{n} \sum_{j=1}^{n} f_{z_0}(\lambda_j) - \int_{\mathbb{C}} f_{z_0}(z) d\mu_\infty(z) \right) \leq c \frac{\log n}{n^{1-2s}} \|\Delta f\|_{L^p} \geq 1 - n^{-Q}. \]

Recalling the discussion in section 2, \( z_0 \) and \( f \) are not allowed to depend on \( \omega \) here.

**Proof.** As in the proof of Proposition 4, integration by parts yields
\[ \frac{1}{n} \sum_{j=1}^{n} f_{z_0}(\lambda_j) - \int_{\mathbb{C}} f_{z_0}(z) d\mu_\infty(z) = \frac{n^{2s}}{2\pi} \int_{\mathbb{C}} \Delta f(z) (U_n(z) - U_\infty(z)) \, dz. \]
After applying Hölder’s inequality as was done in (12), it remains to show the estimate \( \|U_n - U_\infty\|_{L^p} \lesssim \log^4 n/n \) which we already showed in the proof of Theorem 5 via Monte Carlo sampling and Theorem 3. \( \square \)

We now turn to an application for random polynomials. The proof does not differ much from those above.
Proof of Theorem 9. As above, we choose \( p > (1 - \varepsilon)/\varepsilon \) large enough and apply Theorem 10 to \( \mu = \mu_n^W, \nu = \mu_\infty, K = 2, \eta = 1, z^* = 0 \) and \( a = n^{1/2 - \varepsilon} \), and obtain

\[
D(\mu_n^W, \mu_\infty) \lesssim n^{1/2} \left\| U_n^W - U_\infty \right\|_{L^p(B(0) \setminus B_{1/2 + \varepsilon}(0))} + \mu_n^W(B_{2n^{-1/2 + \varepsilon}}(0)) + \mu_n^W(B_2(0)^c) \\
+ \mu_\infty(B_{2n^{-1/2 + \varepsilon}}(0)) + \mu_\infty(B_2(0)^c) \\
+ \sup_{R \geq 0, z_0 \in \mathbb{C}} \mu_\infty \left( R \leq |z| \leq R + 2n^{-1/2 + \varepsilon} \right).
\]

Let us consider each term starting with the last one. Obviously the last term is of order \( n^{-1/2 + \varepsilon} \) and the third line equals \( 4n^{-1/2 + \varepsilon} \). From an already existing (non-uniform) local circular law for random polynomials, see [TV14] formula (87), it follows that with overwhelming probability (i.e. \( \geq 1 - n^{-Q/2} \)) the second line of our estimation can also be bounded by \( cn^{-1/2 + \varepsilon} \). Therefore it remains to control the \( L^p \) distance of the logarithmic potentials. The application of Monte Carlo sampling and the pointwise control of the logarithmic potentials from (11) remains unchanged. The only notable difference to the proof of Theorem 5 is the restriction to polynomially bounded moduli of the zeros. From Rouché’s Theorem, we deduce an upper bound for the largest root

\[
|\lambda|_{\max} \leq 1 + \frac{\max\{|c_0|, \ldots, |c_{n-1}|, |\xi_{n-1}|\}}{c_n |\xi_n|}
\]

of any polynomial. Hence for any \( Q > 0 \) we have

\[
\mathbb{P}(|\lambda|_{\max} \geq n^{(Q+1)/\delta}) \leq \mathbb{P} \left( \frac{\max\{|\xi_0|, \ldots, |\xi_{n-1}|\}}{|\xi_n|} \gtrsim n^{(Q+1)/\delta} \right) \\
\leq (n - 1) \mathbb{P}(|\xi_0| \gtrsim n^{(Q+1)/\delta} |\xi_n|) \\
\lesssim \frac{n - 1}{n^{Q+1}} \mathbb{E} |\xi_0|^{\delta} \mathbb{E} |1/\xi_0|^{\delta} \lesssim n^{-Q},
\]

which replaces (17) and the proof is finished. \( \square \)

Appendix A.

Lemma 14. The mean empirical spectral distribution \( \bar{\mu}_n = \mathbb{E} \mu_n \) of the Ginibre ensemble satisfies

\[
\sup_{z_0 \in \mathbb{C}, R > 0} \left| \bar{\mu}_n(B_R(z_0)) - \mu_\infty(B_R(z_0)) \right| \asymp 1/\sqrt{n}
\]

and

\[
\sup_{B_R(z_0) \subseteq \mathbb{C} \setminus B_{1+\varepsilon}(0) \text{ or } B_R(z_0) \subseteq B_{1-\varepsilon}(0)} |\bar{\mu}_n(B_R(z_0)) - \mu_\infty(B_R(z_0))| \lesssim e^{-n\varepsilon^2}.
\]

Proof. Since [Gin65], the density \( p_n \) of \( \bar{\mu}_n \) has been known to be

\[
p_n(z) = \frac{1}{\pi} e^{-|z|^2} \sum_{k=0}^{n-1} \frac{n^k |z|^{2k}}{k!},
\]
which converges to \( p_\infty(z) = \frac{1}{\pi} \mathbb{I}_{B_1(0)}(z) \). In the case of \( z_0 = 0 \), we can explicitly calculate

\[
\bar{\mu}_n(B_R(0)) = \frac{1}{\pi} \int_{B_R(0)} e^{-n|z|^2} \sum_{k=0}^{n-1} \frac{n^k |z|^{2k}}{k!} dz
\]

\[
= \frac{1}{n} \sum_{k=0}^{n-1} \int_0^{nR^2} e^{-r} \frac{r^k}{k!} dr
\]

\[
= \frac{1}{n} \sum_{k=0}^{n-1} \left[ 1 - e^{-nR^2} \sum_{j=0}^{k} \frac{(nR^2)^j}{j!} \right]
\]

\[
= 1 - e^{-nR^2} \sum_{k=0}^{n-1} \frac{(n-k)(nR^2)^k}{nk!}
\]

\[
= 1 - e^{-nR^2} \left( \frac{(nR^2)^n}{n!} + (1-R^2) \sum_{k=0}^{n-1} \frac{(nR^2)^k}{k!} \right)
\]

where we used the substitution \( r = n|z|^2 \) and integration by parts. The function

\[
\tilde{D}_n(R) = \mu_\infty(B_R(0)) - \bar{\mu}_n(B_R(0))
\]

\[
= 1 \wedge R^2 - 1 + e^{-nR^2} \left( \frac{(nR^2)^n}{n!} + (1-R^2) \sum_{k=0}^{n-1} \frac{(nR^2)^k}{k!} \right)
\]

is continuous in \( R \) and differentiable for \( R \neq 1 \) with derivative

\[
2R \left( \mathbb{I}_{[0,1]}(R) - e^{-nR^2} \sum_{k=0}^{n-1} \frac{(nR^2)^k}{k!} \right)
\]

\[
> 0 \quad \text{if } R < 1
\]

\[
< 0 \quad \text{if } R > 1
\]

Hence the maximum is attained at \( R = 1 \) and Stirling’s formula yields

\[
\sup_{R>0} |\bar{\mu}_n(B_R(0)) - \mu_\infty(B_R(0))| = \frac{n^n}{e^n n!} \approx \frac{1}{\sqrt{2\pi n}}.
\]

For arbitrary balls we roughly bound

\[
|\bar{\mu}_n - \mu_\infty|(B_R(z_0)) \leq \int_{B_1(0)} p_\infty(z) - p_n(z) dz + \int_{B_1(0)^c} p_n(z) dz
\]

\[
= 2 \int_{B_1(0)} p_\infty(z) - p_n(z) dz \approx \sqrt{\frac{2}{\pi n}},
\]

hence the first part of the statement is proven. For \( R \leq 1 \) we have

\[
\tilde{D}_n(R) = e^{-nR^2} \left( \frac{(nR^2)^n}{n!} + (1-R^2) \sum_{k=n}^{\infty} \frac{(nR^2)^k}{k!} \right)
\]
and
\[ e^{-nR^2} \sum_{k=n}^{\infty} \frac{n^k R^{2k}}{k!} \leq e^{-nR^2} \left( \frac{nR^2}{n!} \sum_{k=0}^{\infty} \left( \frac{nR^2}{(n+1)} \right)^k \right) \]
\[ = e^{-nR^2} \frac{n+1}{n(1-R^2)+1} \]
\[ \approx \frac{1}{\sqrt{2\pi n}} e^{-n(R^2-1-\log(R^2))} \frac{n+1}{n(1-R^2)+1}. \]

where we applied Stirling's formula again. Consequently
\[ |\tilde{D}_n(R)| \lesssim \frac{1}{\sqrt{n}} e^{-n(R^2-1-\log(R^2))} \left( 1 + (1-R^2) \frac{n+1}{n(1-R^2)+1} \right) \]
\[ \lesssim \frac{1}{\sqrt{n}} e^{-n(R^2-1-\log(R^2))} \]

for \( R \leq 1 \). On the other hand if \( R \geq 1 \), then
\[ \tilde{D}_n(R) = e^{-nR^2} \left( \frac{(nR^2)^n}{n!} - (R^2-1) \sum_{k=0}^{n-1} \frac{(nR^2)^k}{k!} \right), \]

where analogously we have
\[ \sum_{k=0}^{n-1} \frac{(nR^2)^k}{k!} \leq \frac{(nR^2)^{n-1}}{(n-1)!} \sum_{k=0}^{n-1} \frac{n-1}{nR^2} \leq \frac{(nR^2)^{n-1}}{(n-1)!} \frac{1}{(R^2-1)+1} \]

and hence
\[ |\tilde{D}_n(R)| \lesssim \frac{1}{\sqrt{n}} e^{-n(R^2-1-\log(R^2))}. \]

Finally choose \( R = 1 - \varepsilon \) (or \( R = 1 + \varepsilon \), respectively) and note that \( R^2 - 1 - \log R^2 \geq 2\varepsilon^2 + O(\varepsilon^3) \), we conclude
\[ |\tilde{D}_n(1-\varepsilon)| \lesssim e^{-n\varepsilon^2} \]

and the second part of the Lemma follows. \( \Box \)

**Lemma 15.** Convergence of distributions on \( \mathbb{C} \) with respect to the spherical Kolmogorov distance \( D \) implies weak convergence.

For absolutely continuous limit distributions, the converse statement is also true, see for instance [TDHJ76]. Hence \( D \) is a reasonable object for studying the rate of convergence to the circular law. Moreover we justify the term Kolmogorov distance by formally retrieving the 1-dimensional Kolmogorov distance \( d(\mu_j, \nu_j) \) of the marginals \( j = 1, 2 \) in limits such as
\[ (\mu_1 - \nu_1)((-\infty, t]) = \lim_{K \to \infty} (\nu - \mu)(B_K(t + K, 0)). \]

**Proof.** Let \( \mu, \nu \) be distributions on \( \mathbb{C} \), \( f \in C_c(\mathbb{C}) \) be a continuous function with compact support and \( f_r = \frac{1}{\pi r^2} f * 1_{B_r(0)} \) be its ball mean function. Furthermore
and has compact support where the endpoints are given by

\[ \mathbb{J}^z := \begin{cases} [-\lambda_+, -\lambda_-] \cup [\lambda_-, \lambda_+], & \text{if } |z| > 1 \\ [-\lambda_+, \lambda_+], & \text{if } |z| \leq 1 \end{cases} \]

where the endpoints are given by

\[ \lambda_\pm^2 := \frac{(\alpha \pm 3)^2}{8(\alpha \pm 1)} \wedge 0, \quad \alpha := \sqrt{1+8|z|^2}. \]

Note that \( \lambda_- \sim (1-|z|)^{3/2} \) as \( |z| \to 1 \), i.e. a new gap in the support emerges at 0. Therefore \( s \) will be unbounded for \( z \) close to the edge, which is the reason for the bulk constraint of Theorem 3.
Proof of Theorem 3. Fix some arbitrary $Q, \tau > 0$ and $z \in B_{1+\tau^{-1}}(0)$ satisfying $|1 - z| \geq \tau$. As is explained in Girko’s Hermitization trick (4),

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

and therefore it is necessary to estimate the extremal singular values as well as the rate of convergence of $\tilde{\nu}^z_n$ to $\tilde{\nu}^z$ in Kolmogorov distance $d^*_n(z)$. 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. Theorem 2.1 in [TV08] states that there exists a constant $B > 0$ such that $P(\Omega_0) \lesssim n^{-Q}$ and analogously to what has been shown in (17) there exists a constants $B' > 0$ with $P(\Omega_1') \lesssim n^{-Q}$. Since $\tilde{\nu}^z$ has a bounded density, we get

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

and furthermore on $\Omega_2$ it holds that

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

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 $P(\Omega_2') \lesssim n^{-Q}$, 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. \hfill $\Box$

Acknowledgements

Financial support by the German Research Foundation (DFG) through the IRTG 2235 is gratefully acknowledged. We would like to thank A. Tikhomirov and A. Naumov for helpful discussions and valuable suggestions. We are also grateful to A. Sinulis for a careful reading of this note.

References

[Bai97] Zhidong Bai. Circular law. The Annals of Probability, pages 494–529, 1997.

[BC12] Charles Bordenave and Djalil Chafaï. Around the circular law. Probability surveys, 9, 2012.

[BS10] Zhidong Bai and Jack W Silverstein. Spectral analysis of large dimensional random matrices, volume 20. Springer, 2010.

[BYY14a] Paul Bourgade, Horng-Tzer Yau, and Jun Yin. Local circular law for random matrices. Probability Theory and Related Fields, 159(3-4):545–595, 2014.

[BYY14b] Paul Bourgade, Horng-Tzer Yau, and Jun Yin. The local circular law ii: the edge case. Probability Theory and Related Fields, 159(3-4):619–660, 2014.
[CHM16] Djalil Chafaï, Adrien Hardy, and Mylène Maïda. Concentration for coulomb gases and coulomb transport inequalities. arXiv preprint arXiv:1610.00980, 2016.

[DV17] Tien-Cuong Dinh and Duc-Viet Vu. Large deviation theorem for random covariance matrices. arXiv preprint arXiv:1707.07174, 2017.

[FH99] P J Forrester and G Honner. Exact statistical properties of the zeros of complex random polynomials. Journal of Physics A: Mathematical and General, 32(16):2961, 1999.

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

[Gir85] Vyacheslav L Girko. Circular law. Theory of Probability & Its Applications, 29(4):694–706, 1985.

[GNT17] Friedrich Götze, Alexey Naumov, and Alexander Tikhomirov. On local laws for non-hermitian random matrices and their products. arXiv preprint arXiv:1708.06950, 2017.

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

[GT07] Friedrich Götze and Alexander Tikhomirov. On the circular law. arXiv preprint math/0702386, 2007.

[GT10] Friedrich Götze and Alexander Tikhomirov. The circular law for random matrices. The Annals of Probability, 38(4):1444–1491, 2010.

[GT16] Friedrich Götze and Alexander Tikhomirov. Optimal bounds for convergence of expected spectral distributions to the semi-circular law. Probability Theory and Related Fields, 165(1-2):163–233, 2016.

[IZ13] Ildar Ibragimov and Dmitry Zaporozhets. On distribution of zeros of random polynomials in complex plane. In Prokhorov and contemporary probability theory, pages 303–323. Springer, 2013.

[KZ14] Zakhar Kabluchko and Dmitry Zaporozhets. Asymptotic distribution of complex zeros of random analytic functions. The Annals of Probability, 42(4):1374–1395, 2014.

[MM15] 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(1):93–117, 2015.

[PZ10] Guangming Pan and Wang Zhou. Circular law, extreme singular values and potential theory. Journal of Multivariate Analysis, 101(3):645–656, 2010.

[RV08] Mark Rudelson and Roman Vershynin. The littlewood–offord problem and invertibility of random matrices. Advances in Mathematics, 218(2):600–633, 2008.

[TDHJ76] Flemming Topsøe, Richard M Dudley, and Jørgen Hoffmann-Jørgensen. Two examples concerning uniform convergence of measures wrt balls in banach spaces. In Empirical Distributions and Processes, pages 141–146. Springer, 1976.

[TV08] Terence Tao and Van Vu. Random matrices: the circular law. Communications in Contemporary Mathematics, 10(02):261–307, 2008.
[TV09] Terence Tao and Van Vu. From the littlewood-offord problem to the circular law: universality of the spectral distribution of random matrices. *Bulletin of the American Mathematical Society*, 46(3):377–396, 2009.

[TV10] Terence Tao and Van Vu. Random matrices: Universality of esds and the circular law. *The Annals of Probability*, pages 2023–2065, 2010. With an appendix by Manjunath Krishnapur.

[TV14] Terence Tao and Van Vu. Local universality of zeroes of random polynomials. *International Mathematics Research Notices*, 2015(13):5053–5139, 2014.

[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.

Friedrich Götze, Faculty of Mathematics, Bielefeld University, Germany
E-mail address: goetze@math.uni-bielefeld.de

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