ANNALES DE LA FACULTÉ DES SCIENCES DE TOULOUSE Mathématiques

Tom Carroll, Jordi Marzo, Xavier Massaneda and Joaquim Ortega-Cerdà

Equidistribution and β-ensembles

Tome XXVII, nº 2 (2018), p. 377-387.

<http://afst.cedram.org/item?id=AFST_2018_6_27_2_377_0>

Annales de la faculté des sciences de Toulouse Volume XXVII, n° 2, 2018
pp. 377-387

Equidistribution and β-ensembles

Tom Carroll [1], Jordi Marzo [2],
Xavier Massaneda [3] and Joaquim Ortega-Cerdà [4]

Abstract. — We find the precise rate at which the empirical measure associated to a β-ensemble converges to its limiting measure. In our setting the β-ensemble is a random point process on a compact complex manifold distributed according to the β power of a determinant of sections in a positive line bundle. A particular case is the spherical ensemble of generalized random eigenvalues of pairs of matrices with independent identically distributed Gaussian entries.

Résumé. — On trouve le taux précis où la mesure empirique associée à un β-ensemble converge vers sa mesure limite. Le β-ensemble est un processus de points aléatoires sur une variété complexe compacte répartis selon la puissance β d'un déterminant de sections d'un fibré de ligne positif. Un cas particulier est l'ensemble sphérique de valeurs propres généralisés de paires de matrices aléatoires avec entrées gaussiennes identiquement distribuées et independantes.

1. Background and setting

Let (X, ω) be an n-dimensional compact complex manifold endowed with a smooth Hermitian metric ω. Let (L, ϕ) be a holomorphic line bundle with a positive Hermitian metric ϕ. This has to be understood as a collection of smooth functions ϕ_i defined in trivializing neighborhoods U_i of the line bundle. If $e_i(x)$ is a frame in U_i, then $|e_i(x)|^2_\phi = e^{-\phi_i(x)}$. Thus ϕ_i must satisfy the compatibilty condition $\phi_i - \phi_j = \log|g_{ij}|$, where g_{ij} are the transition functions.

[1] University College Cork — t.carroll@ucc.ie
[2] Universitat de Barcelona, BGSMath — jmarzo@ub.edu
[3] Universitat de Barcelona, BGSMath — xavier.massaneda@ub.edu
[4] Universitat de Barcelona, BGSMath — jortega@ub.edu

The last three authors have been partially supported by grants MTM2014-51834-P and MTM2017-83499-P by the Ministerio de Economía y Competitividad, Gobierno de España and by the Generalitat de Catalunya (project 2017 SGR 359).

As usual we denote by $H^0(X,L)$ the global holomorphic sections. If $s \in H^0(X,L)$ we will denote by $|s(x)|_\phi$ the pointwise norm on the fiber induced by ϕ. If we have any other line bundles (like L^k) with a natural metric induced by ϕ we will still denote by $|s(x)|_\phi$ the corresponding norm.

If L is a line bundle over X and M is a line bundle over Y, we denote by $L \boxtimes M$ the line bundle over the product manifold $X \times Y$ defined as $L \boxtimes M = \pi_X^*(L) \otimes \pi_Y^*(M)$, where $\pi_X : X \times Y \to X$ is the projection onto the first factor and $\pi_Y : X \times Y \to Y$ is the projection onto the second. The line bundle $L \boxtimes M$ carries a metric induced by that of L and M.

Given a basis $s_1, \dots, s_N$ of $H^0(X,L)$, we define $\det(s_i(x_j))$ as a section of $L^{\boxtimes N}$ over X^N by the identities $\det(s_i(x_j)) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \bigotimes_{i=1}^N s_i(x_{\sigma_i})$.

We fix a probability measure on X, given by the normalized volume form ω^n, that we denote by σ.

Definition 1.1. — *Let $\beta > 0$. A β-ensemble is an N point random process on X which has joint density given by*

$$\frac{1}{Z_N} |\det s_i(x_j)|_\phi^\beta \, \mathrm{d}\sigma(x_1) \otimes \cdots \otimes \mathrm{d}\sigma(x_N), \tag{1.1}$$

where $Z_N = Z_N(\beta)$ is chosen so that this is a probability distribution in X^N and $|\cdot|_\phi$ denotes the norm measured using the induced metric in $(L^k)^{\boxtimes N_k}$.

Observe that the random point process is independent of the choice of basis $(s_j)_j$.

A particularly interesting case is when $\beta = 2$, since then the process is determinantal. Let K denote the Bergman kernel of the Hilbert space $H^0(X,L)$ endowed with the norm $\|s\|^2 = \int_X |s(x)|_\phi^2 \, \mathrm{d}\sigma(x)$. Then

$$|\det(s_i(x_j))|_\phi^2 = |\det(K(x_i,x_j))|_\phi.$$

Another interesting situation occurs when $\beta \to \infty$. In this case the probability charges the maxima of the function $|\det(s_i(x_j))|$. A set of points $\{x_j\}_j$ with cardinality $\dim H^0(X,L)$ and maximizing this determinant is known as a Fekete sequence. The distribution of these sequences has been studied in [9], [10], and [2] and we will draw some ideas from there to study general β-ensembles.

We consider now the situation where we replace L by a power L^k, $k \in \mathbb{N}$, and let k tend to infinity. We denote by N_k the dimension of $H^0(X,L^k)$. It is well-known, by the Riemann–Roch theorem and the Kodaira vanishing theorem, that

$$\dim H^0(X,L^k) = \frac{c_1(L)^n}{n!} k^n + O(k^{n-1}) \sim k^n,$$

where $c_1(L)$ denotes the first Chern class of L.

For each k we consider a collection of N_k points chosen randomly according to the law (1.1). For each k the collection is picked independently of the previous ones.

Given points $x_1^{(k)}, \dots, x_{N_k}^{(k)}$ chosen according to (1.1), consider its associated empirical measure $\mu_k = \frac{1}{N_k} \sum \delta_{x_i^{(k)}}$. For convenience we will drop the superindex (k) hereafter. We are interested in understanding the limiting distribution of the measures μ_k.

The following result is well known; see [2].

THEOREM (Berman, Boucksom, Witt Nyström). — *Let μ_k be the empirical measure associated to a Fekete sequence for the bundle $H^0(X, L^k)$. Then, as $k \to \infty$,*

$$\mu_k \longrightarrow \nu := \frac{(i\partial\bar{\partial}\phi)^n}{\int_X (i\partial\bar{\partial}\phi)^n}$$

in the weak-$*$ topology.

The measure ν is called the equilibrium measure.

There is a counterpart of this result for empirical measures of general β-ensembles (see [3], which gives an estimate for the large deviations of the empirical measure from the equilibrium measure).

Our aim is to obtain a different quantitative version of the weak convergence of the empirical measure to the equilibrium measure, measured in terms of the Kantorovich–Vaserstein distance between mesaures. We have chosen the compact setting since it is technically simpler than the non compact case as the Ginibre ensemble studied in [15].

This sort of quantification has also been studied, with different tools, in the context of random matrix models, (see for instance [11, 12, 13]), where similar determinantal point processes arise.

In fact some of the β-ensembles we are considering admit random matrix models, at least when $\dim_{\mathbb{C}}(M) = 1$. For instance, Krishnapur studied in [8] the following point process: let A, B be $k \times k$ random matrices with i.i.d. complex Gaussian entries. He proved that the generalized eigenvalues associated with the pair (A, B), i.e. the eigenvalues of $A^{-1}B$, have joint probability density:

$$\frac{1}{Z_k} \prod_{l=1}^{k} \frac{1}{(1 + |x_l|^2)^{k+1}} \prod_{i<j} |x_i - x_j|^2, \tag{1.2}$$

with respect to the Lebesgue measure in the plane.

It was also observed in [8] that, using the stereographic projection

$$\begin{aligned} \pi : \mathbb{S}^2 &\longrightarrow \mathbb{C} \\ P_j &\mapsto x_j, \end{aligned}$$

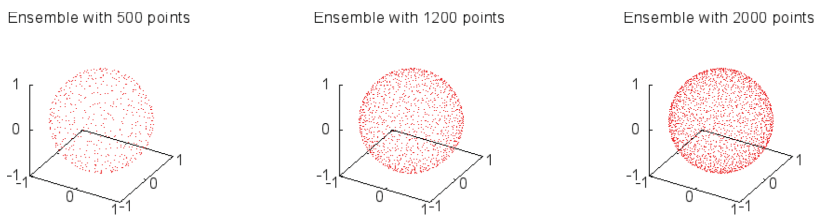
the joint density (1.2) (with respect to the product area measure in the product of spheres) is

$$\frac{1}{Z_k} \prod_{i<j} \|P_i - P_j\|^2_{\mathbb{R}^3}.$$

Since this is invariant under rotations of the sphere, the point process is called the spherical ensemble.

A point process with this law had been considered earlier (without a random matrix model) by Caillol [6] as the model of one-component plasma.

One typical instance of the process is as in the picture.



The spherical ensemble has received much attention. We mention a couple of properties related to our results. In [5], Bordenave proves the universality of the spectral distribution of the $k \times k$-matrix $A^{-1}B$ with respect to other i.i.d. random distribution of entries. As an outcome, he proves that the weak-* limit of the spectral measures $\mu_k = \frac{1}{k}\sum_i \delta_{x_i}$, where x_i are the generalized eigenvalues, is the normalized area measure in the sphere. This convergence is rather uniform: in [1] Alishahi and Zamani estimate the discrepancy of the empirical measure with respect to its limit and give precise estimates of the Riesz and the logarithmic energies.

Our main result is a quantification of the equidistribution of the empirical measure associated to a β-ensemble in terms of the Kantorovich–Vaserstein distance.

Theorem 1.2. — *Let $\beta \geqslant 1$ and consider the empirical measure μ_k associated to the β-ensemble given in Definition 1.1 and let $\nu = \frac{(i\partial\bar\partial\phi)^n}{\int_X (i\partial\bar\partial\phi)^n}$ be the equilibrium measure. Then the expected Kantorovich–Vaserstein distance from μ_k to μ can be estimated by*

$$\mathbb{E}W(\mu_k, \nu) \leqslant C/\sqrt{k}.$$

1.1. The Kantorovich–Vaserstein distance

To measure the uniformity and speed of convergence of the empirical measures μ_k to the limiting measure ν we use the Kantorovich–Vaserstein distance W. Given probability measures μ and ν, it is defined as

$$W(\mu,\nu) = \inf_{\rho} \iint_{X\times X} d(x,y)\,\mathrm{d}\rho(x,y),$$

where $d(x,y)$ is the distance associated to the metric ω and the infimum is taken over all admissible transport plans ρ, i.e., all probability measures in $X \times X$ with marginal measures μ and ν respectively.

In general, the Kantorovich–Vaserstein distance is defined on probability measures over a compact metric space X, and it metrizes the weak-$*$ convergence of measures.

It was observed in [9] that in the definition of W it is possible to enlarge the class of admissible transport plans to complex measures ρ that have marginals μ and ν respectively. We include the argument for the sake of completeness.

Let

$$\widetilde{W}(\mu,\nu) = \inf_{\rho\in S} \iint_{X\times X} d(x,y)\,\mathrm{d}|\rho(x,y)|, \tag{1.3}$$

where the infimum is now taken over the set S of all complex measures ρ on $X \times X$ with marginals $\rho(\,\cdot\,,X) = \mu$ and $\rho(X,\cdot\,) = \nu$.

In order to see that $\widetilde{W}(\mu,\nu) = W(\mu,\nu)$, we recall the dual formulation of W (see [17, (6.3)]):

$$W(\mu,\nu) = \sup\left\{\left|\int_X f\,d(\mu-\nu)\right| : f \in \mathrm{Lip}_{1,1}(X)\right\}, \tag{1.4}$$

where $\mathrm{Lip}_{1,1}(X)$ is the collection of all functions f on X satisfying $|f(x) - f(y)| \leqslant d(x,y)$.

For any complex measure ρ with marginals μ and ν and any $f \in \mathrm{Lip}_{1,1}(X)$ we have

$$\left|\int_X f\,d(\mu-\nu)\right| = \left|\iint_{X\times X} (f(x)-f(y))\,\mathrm{d}\rho(x,y)\right| \leqslant \iint_{X\times X} d(x,y)\,\mathrm{d}|\rho(x,y)|.$$

Hence

$$W(\mu,\nu) \leqslant \inf_{\rho\in S} \iint_{X\times X} d(x,y)\,\mathrm{d}|\rho(x,y)| = \widetilde{W}(\mu,\nu).$$

The remaining inequality ($\widetilde{W}(\mu,\nu) \leqslant W(\mu,\nu)$) is trivial.

A standard reference for basic facts on Kantorovich–Vaserstein distances is the book [17].

1.2. Lagrange sections

We fix now a basis of sections $s_1, \dots s_{N_k}$ of $H^0(X, L^k)$. Given any collection of points $(x_1, \dots, x_{N_k})$ we define the Lagrange sections informally as:

$$\ell_j(x) = \frac{\begin{vmatrix} s_1(x_1) & \cdots & s_1(x) & \cdots & s_1(x_{N_k}) \\ \vdots & & \vdots & & \vdots \\ s_{N_k}(x_1) & \cdots & s_{N_k}(x) & \cdots & s_{N_k}(x_{N_k}) \end{vmatrix}}{\begin{vmatrix} s_1(x_1) & \cdots & s_1(x_j) & \cdots & s_1(x_{N_k}) \\ \vdots & & \vdots & & \vdots \\ s_{N_k}(x_1) & \cdots & s_{N_k}(x_j) & \cdots & s_{N_k}(x_{N_k}) \end{vmatrix}}$$

Clearly $\ell_j \in H^0(X, L^k)$ and $\ell_j(x_i) = 0$ if $i \neq j$ and $|\ell_j(x_j)| = 1$.

More formally, we proceed as in [9]: if $e_j(x)$ is a frame in a neighborhood U_j of the point x_j, then the sections $s_i(x)$ are represented on each U_j by scalar functions f_{ij} such that $s_i(x) = f_{ij}(x)e_j(x)$. Similarly, the metric $k\phi$ is represented on U_j by a smooth real-valued function $k\phi_j$ such that $|s_i(x)|^2 = |f_{ij}(x)|^2 e^{-k\phi_j(x)}$.

To construct the Lagrange sections we denote by A the matrix

$$\left(e^{-\frac{k}{2}\phi_j(x_j)} f_{ij}(x_j)\right)_{i,j},$$

and define

$$\ell_j(x) := \frac{1}{\det(A)} \sum_{i=1}^{N_k} (-1)^{i+j} A_{ij} s_i(x),$$

where A_{ij} is the determinant of the submatrix obtained from A by removing the i-th row and j-th column. Clearly $\ell_j \in H^0(X, L^k)$, and it is not difficult to check that $|\ell_j(x_i)|_\phi = \delta_{ij}$, $1 \leqslant i, j \leqslant N_k$.

Notice that if we denote by $\rho_k(x_1, \dots, x_{N_k}) = \frac{1}{Z_{N_k}} |\det s_i(x_j)|_\phi^\beta$ then

$$|\ell_j(x)|_\phi^\beta = \frac{\rho_k(x_1, \dots, x, \dots, x_{N_k})}{\rho_k(x_1, \dots, x_j, \dots, x_{N_k})}, \tag{1.5}$$

and thus $\mathbb{E}(\|\ell_j\|_\beta) \leqslant 1$, because

$$\begin{aligned} \mathbb{E}(\|\ell_j\|_\beta)^\beta \leqslant \mathbb{E}(\|\ell_j\|_\beta^\beta) &= \mathbb{E}\left(\int_X |\ell_j(x)|_\phi^\beta \mathrm{d}\sigma(x)\right) \\ &= \int_{X^{N_k+1}} \rho_k(x_1, \dots, x, \dots, x_{N_k}) \mathrm{d}\sigma(x) \mathrm{d}\sigma(x_1) \cdots \mathrm{d}\sigma(x_{N_k}) = 1. \end{aligned}$$

In the case of the Fekete points ($\beta = \infty$), $\sup_X |\ell_j(x)|_\phi = 1$ by definition.

2. Proof of the main result

Before proving the main result a couple of remarks on the sharpness of the result are in order.

Remark. — The rate of convergence cannot be improved. Let σ be any nowhere vanishing smooth probability distribution on X. Let E_k be any discrete set on X with cardinality $\#E_k \simeq k^n \simeq N_k$, and let $\mu_k = \frac{1}{\#E_k}\sum_{y\in E_k}\delta_y$. Then the distance $W(\mu_k,\sigma) \gtrsim 1/\sqrt{k}$.

To obtain a lower bound for $W(\mu_k,\sigma)$ we use the dual formulation of the Kantorovich–Vaserstein distance (1.4) and the function $f(x) = d(x,E_k)$, which is in $\mathrm{Lip}_{1,1}(X)$. Since $d(x,E_k) = 0$ on the support of μ_k we obtain

$$W(\mu_k,\sigma) \geqslant \int_X d(x,E_k)\,\mathrm{d}\sigma.$$

Vitali's covering lemma ensures that for each k and for some ε small enough, independent of k, there are at least $2\#E_k$ pairwise disjoint balls of radius $\varepsilon/\sqrt{k}$. Since the number of balls is twice the number of points in E_k, at least half the balls contain no point of E_k. We consider one such ball, $B(y_i,\varepsilon/\sqrt{k})$. In the smaller ball $B(y_i,0.5\varepsilon/\sqrt{k})$ we have $d(x,E_k) \geqslant 0.5\varepsilon/\sqrt{k}$. Thus

$$\begin{aligned}\int_X d(x,E_k)\,\mathrm{d}\sigma &\geqslant \sum_i \int_{B(y_i,\varepsilon/\sqrt{k})} d(x,E_k)\,\mathrm{d}\sigma \gtrsim \sum_i \frac{1}{\sqrt{k}}\sigma\big(B(y_i,\varepsilon/\sqrt{k})\big)\\ &\gtrsim \#E_k \frac{1}{\sqrt{k}} k^{-n} \simeq \frac{1}{\sqrt{k}}.\end{aligned}$$

Remark. — Once we have observed that the rate of convergence is optimal we may consider what is the value of the constant C that appears on the speed of convergence. This constant depends on the off-diagonal estimate of the Bergman kernel (2.2). Thus the positivity of the holomorphic line bundle plays an important role in the speed of convergence.

As a final remark we observe that the techniques that we use are modelled after the proof of the speed of the equidistribution of the Fekete points that appears in [9].

Proof of Theorem 1.2. — To prove this we provide a (complex) transport plan between the probability measure $b_k(x) = \frac{1}{N_k}K_k(x,x)$ (b_k stands for *Bergman measure*) and the empirical measure μ_k. We are going to prove that

$$\mathbb{E}W(\mu_k,b_k) = O\Big(\frac{1}{\sqrt{k}}\Big).$$

Standard estimates for the Bergman kernel provide:

$$W(b_k, \nu) = O\Big(\frac{1}{\sqrt{k}}\Big).$$

Actually one can prove that the total variation (which dominates the Kantorovich–Vaserstein distance) satisfies:

$$\left\| \frac{K_k(x,x)}{N_k} - \nu \right\| \leqslant \frac{C}{\sqrt{k}}. \tag{2.1}$$

This follows for instance from the expansion in powers of $1/k$ of the Bergman kernel. In this context this is due to Tian, Catlin and Zelditch, [7, 16, 18].

In the particular case of the spherical ensemble, the kernel is explicit and invariant under rotations, and the estimate is even better: the Bergman measure is the equilibrium measure, i.e. $b_k = \nu$.

Consider the transport plan

$$p(x,y) = \frac{1}{N_k} \sum_{j=1}^{N_k} \delta_{x_j}(y) \langle K_k(x, x_j), \ell_j(x) \rangle \, \mathrm{d}\sigma(x).$$

It has the correct marginals (b_k and μ_k respectively) and thus

$$\begin{aligned} W(b_k, \mu_k) &\leqslant \iint_{X \times X} d(x,y) \, \mathrm{d}|p|(x,y) \\ &\leqslant \frac{1}{N_k} \sum_{j=1}^{N_k} \int_X d(x, x_j) |\ell_j(x)| |K_k(x, x_j)| \, \mathrm{d}\sigma(x). \end{aligned}$$

Now, letting β' be the conjugate exponent of β (so that $1/\beta + 1/\beta' = 1$), we have

$$\begin{aligned} (\mathbb{E}W)^\beta &\leqslant \int_{X^{N_k}} \frac{1}{N_k} \sum_{j=1}^{N_k} \left(\int_X d(x, x_j) |\ell_j(x)| |K_k(x, x_j)| \mathrm{d}\sigma(x) \right)^\beta \\ &\qquad\qquad\qquad\qquad \times \rho_k(x_1, \ldots, x_{N_k}) \mathrm{d}\sigma(x_1) \cdots \mathrm{d}\sigma(x_{N_k}) \\ &\leqslant \int_{X^{N_k}} \frac{1}{N_k} \sum_{j=1}^{N_k} \left(\int_X d(x, x_j) |K_k(x, x_j)| \mathrm{d}\sigma(x) \right)^{\beta/\beta'} \\ &\qquad\qquad \times \left(\int_X |\ell_j(x)|^\beta |K_k(x, x_j)| d(x, x_j) \mathrm{d}\sigma(x) \right) \\ &\qquad\qquad\qquad\qquad \times \rho_k(x_1, \ldots, x_{N_k}) \mathrm{d}\sigma(x_1) \cdots \mathrm{d}\sigma(x_{N_k}). \end{aligned}$$

Assume for the moment that the following off-diagonal decay of the Bergman kernel holds:

$$\sup_{y\in X}\int_X d(x,y)|K_k(x,y)|\,\mathrm{d}\sigma(x)\leqslant \frac{C}{\sqrt{k}}. \tag{2.2}$$

Then, by (1.5), we obtain:

$$\begin{aligned}(\mathbb{E}W)^\beta &\leqslant \left(\frac{C}{\sqrt{k}}\right)^{\beta/\beta'}\int_{X^{N_k}}\frac{1}{N_k}\sum_{j=1}^{N_k}\int_X|\ell_j(x)|^\beta|K_k(x,x_j)|d(x,x_j)\\ &\qquad\qquad\times\rho_k(x_1,\dots,x_j,\dots,x_{N_k})\mathrm{d}\sigma(x)\mathrm{d}\sigma(x_1)\cdots\mathrm{d}\sigma(x_{N_k})\\ &=\left(\frac{C}{\sqrt{k}}\right)^{\beta/\beta'}\int_{X^{N_k}}\frac{1}{N_k}\sum_{j=1}^{N_k}\int_X|K_k(x,x_j)|d(x,x_j)\\ &\qquad\qquad\times\rho_k(x_1,\dots,x,\dots,x_{N_k})\mathrm{d}\sigma(x)\mathrm{d}\sigma(x_1)\cdots\mathrm{d}\sigma(x_{N_k}).\end{aligned}$$

Finally, integrating first in x_j and applying again (2.2) we obtain

$$(\mathbb{E}W)^\beta\leqslant\Big(\frac{C}{\sqrt{k}}\Big)^{\beta/\beta'}\Big(\frac{C}{\sqrt{k}}\Big)=\mathrm{O}\Big(\frac{1}{\sqrt{k}}\Big)^\beta,$$

as desired.

Estimate (2.2) follows from the pointwise estimate for the Bergman kernel

$$|K_k(x,y)|\leqslant CN_ke^{-C\sqrt{k}\,d(x,y)}, \tag{2.3}$$

which holds when the line bundle is positive, see [4].

Indeed, consider the function $h(s)=se^{-C\sqrt{k}s}$ that is strictly decreasing in $\big[\frac{1}{C\sqrt{k}},+\infty\big)$. For any $y\in X$ we bound the integral in (2.2) as

$$\begin{aligned}\int_X d(x,y)|K_k(x,y)|\,\mathrm{d}\sigma(x)&\lesssim\int_0^{+\infty}\sigma\left(\{x\in X: h(d(x,y))>s\}\right)\mathrm{d}s\\ &\lesssim N_k\int_{(C\sqrt{k})^{-1}}^{+\infty}|h'(s)|\sigma\left(\{x\in X: d(x,y)<s\}\right)\mathrm{d}s\lesssim\frac{1}{\sqrt{k}},\end{aligned}$$

where the last estimate follows from $\sigma(B(y,s))\lesssim s^{2n}$ and $N_k\sim k^n$.

In the particular case of the spherical ensemble, the kernel is explicit and the decay is even faster:

$$\begin{aligned}|K_k(z,w)|^2&=k^2\left(1-\frac{|z-w|^2}{(1+|z|^2)(1+|w|^2)}\right)^{k-1}\\ &\leqslant Kk^2\exp\left(-Ck\frac{|z-w|^2}{(1+|z|^2)(1+|w|^2)}\right)\\ &=Kk^2\exp\left(-Ck\,d(z,w)^2\right),\end{aligned}$$

where here $d(z,w)$ coincides with the chordal metric. $\square$

3. The determinantal setting

Now we turn our attention to the almost sure convergence of the empirical measure. Using the fact that Lipschitz functionals of determinantal process concentrate the measure around the mean, we prove the following result.

COROLLARY 3.1. — *If μ_k is the empirical measure associated with the determinantal point process given by (1.1) with $\beta = 2$, and ν denotes the equilibrium measure, then:*

- *If $\dim_{\mathbb{C}}(X) > 1$ then $W(\mu_k,\nu) = O(1/\sqrt{k})$ almost surely.*
- *If $\dim_{\mathbb{C}}(X) = 1$ then $W(\mu_k,\nu) = O(\log k/\sqrt{k})$ almost surely.*

In particular, any realization of the spherical ensemble satisfies $W(\mu_k,\nu) = O(\log k/\sqrt{k})$ almost surely.

Let ν be, as before, the normalized equilibrium measure. Let us define the functional f on the set of measures of the form $\sigma = \sum_{i=1}^n \delta_{x_i}$ by

$$f(\sigma) = nW\Big(\frac{\sigma}{n},\nu\Big).$$

As the Kantorovich–Vaserstein distance is controlled by the total variation, f is a Lipschitz functional with Lipschitz norm one with respect to the total variation distance. Here we use the following result of Pemantle and Peres [14, Theorem 3.5].

THEOREM (Pemantle-Peres). — *Let Z be a determinantal point process of N points. Let f be a Lipschitz-1 functional defined in the set of finite counting measures (with respect to the total variation distance). Then*

$$\mathbb{P}(f - \mathbb{E}f \geqslant a) \leqslant 3\exp\left(-\frac{a^2}{16(a+2N)}\right)$$

Take now $a = 10\alpha_k N_k/\sqrt{k}$, where $\alpha_k = C\sqrt{\log k}$ for $n = 1$ and $\alpha_k = C$ for $n > 1$ (C is the constant that appears in Theorem 1.2). Then

$$\begin{aligned}\mathbb{P}\Big(W(\mu_k,\nu) > \frac{11\alpha_k}{\sqrt{k}}\Big) &\leqslant \mathbb{P}\Big(N_k W(\mu_k,\nu) > N_k\mathbb{E}W(\mu_k,\nu) + 10\alpha_k\frac{N_k}{\sqrt{k}}\Big)\\ &\leqslant 3\exp\left(-\frac{100\alpha_k^2 N_k^2/k}{16(10\alpha_k N_k/\sqrt{k} + 2N_k)}\right)\\ &\lesssim \exp(-\alpha_k^2 N_k/k) \lesssim \frac{1}{k^2}.\end{aligned}$$

Finally, a standard application of the Borel–Cantelli lemma shows that, with probability one, for all k large enough,

$$W(\mu_k, \nu) \leqslant \frac{11\sqrt{\alpha_k}}{\sqrt{k}}.$$

Bibliography

[1] K. Alishahi & M. Zamani, "The spherical ensemble and uniform distribution of points on the sphere", *Electron. J. Probab.* **20** (2015), p. Article ID 23, 27 p.

[2] R. Berman, S. Boucksom & D. Witt Nyström, "Fekete points and convergence towards equilibrium measures on complex manifolds", *Acta Math.* **207** (2011), no. 1, p. 1-27.

[3] R. J. Berman, "Determinantal point processes and fermions on complex manifolds: large deviations and bosonization", *Commun. Math. Phys.* **327** (2014), no. 1, p. 1-47.

[4] B. Berndtsson, "Bergman kernels related to Hermitian line bundles over compact complex manifolds", in *Explorations in complex and Riemannian geometry*, Contemporary Mathematics, vol. 332, American Mathematical Society, 2003, p. 1-17.

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

[6] J.-M. Caillol, "Exact results for a two-dimensional one-component plasma on a sphere", *J. Physique Lett.* **42** (1981), no. 12, p. 245-247.

[7] D. Catlin, "The Bergman kernel and a theorem of Tian", in *Analysis and geometry in several complex variables*, Trends in Mathematics, Birkhäuser, 1999, p. 1-23.

[8] M. Krishnapur, "From random matrices to random analytic functions", *Ann. Probab.* **37** (2009), no. 1, p. 314-346.

[9] N. Lev & J. Ortega-Cerdà, "Equidistribution estimates for Fekete points on complex manifolds", *J. Eur. Math. Soc.* **18** (2016), no. 2, p. 425-464.

[10] J. Marzo & J. Ortega-Cerdà, "Equidistribution of Fekete points on the sphere", *Constr. Approx.* **32** (2010), no. 3, p. 513-521.

[11] E. S. Meckes & M. W. Meckes, "Concentration and convergence rates for spectral measures of random matrices", *Probab. Theory Relat. Fields* **156** (2013), no. 1-2, p. 145-164.

[12] ———, "Spectral measures of powers of random matrices", *Electron. Commun. Probab.* **18** (2013), p. Article ID 78, 13 p.

[13] ———, "A rate of convergence for the circular law for the complex Ginibre ensemble", *Ann. Fac. Sci. Toulouse, Math.* **24** (2015), no. 1, p. 93-117.

[14] R. Pemantle & Y. Peres, "Concentration of Lipschitz functionals of determinantal and other strong Rayleigh measures", *Combin. Probab. Comput.* **23** (2014), no. 1, p. 140-160.

[15] B. Rider & B. Virág, "The noise in the circular law and the Gaussian free field", *Int. Math. Res. Not.* (2007), no. 2, p. Article ID rnm006, 32 .

[16] G. Tian, "On a set of polarized Kähler metrics on algebraic manifolds", *J. Differ. Geom.* **32** (1990), no. 1, p. 99-130.

[17] C. Villani, *Optimal transport. Old and new*, Grundlehren der Mathematischen Wissenschaften, vol. 338, Springer, 2009, xxii+973 pages.

[18] S. Zelditch, "Szegő kernels and a theorem of Tian", *Int. Math. Res. Not.* (1998), no. 6, p. 317-331.