Uniformity of Eigenvalues of Some Random Matrices

Elizabeth Meckes

Northeast Probability Seminar

November 21, 2014

The empirical spectral measure

The empirical spectral measure

Suppose that M is an $n \times n$ random matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$.

The empirical spectral measure μ of M is the (random) measure

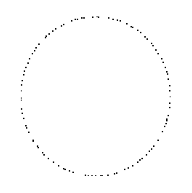
$$\mu := \frac{1}{n} \sum_{k=1}^{n} \delta_{\lambda_k}.$$

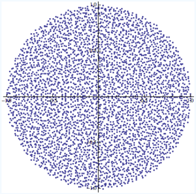
The empirical spectral measure

Suppose that M is an $n \times n$ random matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$.

The empirical spectral measure μ of M is the (random) measure

$$\mu := \frac{1}{n} \sum_{k=1}^{n} \delta_{\lambda_k}.$$





Wigner's Theorem

Wigner's Theorem

For each $n \in \mathbb{N}$, let $\{Y_i\}_{1 \le i}, \{Z_{ij}\}_{1 \le i < j}$ be independent collections of i.i.d. random variables, with

$$\mathbb{E} Y_1 = \mathbb{E} Z_{12} = 0 \qquad \mathbb{E} Z_{12}^2 = 1 \qquad \mathbb{E} Y_1^2 < \infty.$$

Let M_n be the symmetric random matrix with diagonal entries Y_i and off-diagonal entries Z_{ij} or Z_{ji}.

Wigner's Theorem

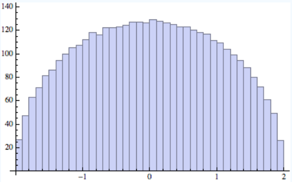
For each $n \in \mathbb{N}$, let $\{Y_i\}_{1 \le i}, \{Z_{ij}\}_{1 \le i < j}$ be independent collections of i.i.d. random variables, with

$$\mathbb{E} Y_1 = \mathbb{E} Z_{12} = 0 \qquad \mathbb{E} Z_{12}^2 = 1 \qquad \mathbb{E} Y_1^2 < \infty.$$

Let M_n be the symmetric random matrix with diagonal entries Y_i and off-diagonal entries Z_{ij} or Z_{ji}.

The empirical spectral measure μ_n of $\frac{1}{\sqrt{n}} M_n$ converges, weakly in probability, to the semi-circular law:

$$\frac{1}{2\pi} \sqrt{4 - x^2} \mathbb{1}_{|x| \le 2} dx.$$



Some random matrix ensembles

Some random matrix ensembles

- Wigner matrices: symmetric with i.i.d. entries

Some random matrix ensembles

- Wigner matrices: symmetric with i.i.d. entries
  - GOE, GUE

Some random matrix ensembles

- Wigner matrices: symmetric with i.i.d. entries
  - GOE, GUE
- Wishart (random covariance) matrices: take X an $m \times n$ ($m \geq n$) random matrix with i.i.d entries, and define $M_n := X^* X$.

Some random matrix ensembles

- Wigner matrices: symmetric with i.i.d. entries
  - GOE, GUE
- Wishart (random covariance) matrices: take X an $m \times n$ ($m \geq n$) random matrix with i.i.d entries, and define $M_n := X^* X$.
- The Ginibre ensemble: M_n has i.i.d. (real or complex) Gaussian entries.

Some random matrix ensembles

- Wigner matrices: symmetric with i.i.d. entries
  - GOE, GUE
- Wishart (random covariance) matrices: take X an $m \times n$ ($m \geq n$) random matrix with i.i.d entries, and define $M_n := X^* X$.
- The Ginibre ensemble: M_n has i.i.d. (real or complex) Gaussian entries.
- Uniform random matrices: random matrices chosen according to Haar measure on a compact matrix group

Some random matrix ensembles

- Wigner matrices: symmetric with i.i.d. entries
  - GOE, GUE
- Wishart (random covariance) matrices: take X an $m \times n$ ($m \geq n$) random matrix with i.i.d entries, and define $M_n := X^* X$.
- The Ginibre ensemble: M_n has i.i.d. (real or complex) Gaussian entries.
- Uniform random matrices: random matrices chosen according to Haar measure on a compact matrix group
  - $\mathbb{O}(n), \mathbb{SO}(n), \mathbb{U}(n), \mathbb{SU}(n), \mathbb{Sp}(2n)$

Determinantal structure

Determinantal structure

The set of eigenvalues of many types of random matrices are determinantal point processes with symmetric kernels:

Determinantal structure

The set of eigenvalues of many types of random matrices are determinantal point processes with symmetric kernels:

	$K_N(x,y)$	Λ
GUE	$\sum_{j=0}^{n-1} h_j(x)h_j(y)e^{-\frac{(x^2+y^2)}{2}}$	$\mathbb{R}$
$\mathbb{U}(N)$	$\sum_{j=0}^{N-1} e^{ij(x-y)}$	$[0,2\pi)$
Complex Ginibre	$\frac{1}{\pi}\sum_{j=0}^{N-1} \frac{(z\overline{w})^j}{j!}e^{-\frac{(\lvert z\rvert^2+\lvert w\rvert^2)}{2}}$	$\{\lvert z\rvert = 1\}$

The gift of determinantal point processes

The gift of determinantal point processes

Theorem (Hough/Krishnapur/Peres/Virág)

Let $K : \Lambda \times \Lambda \to \mathbb{C}$ be the kernel of a determinantal point process, and suppose the corresponding integral operator is self-adjoint, nonnegative, and locally trace-class.

For $D \subseteq \Lambda$, let $\mathcal{N}_D$ denote the number of particles of the point process in D. Then

$$\mathcal{N}_D \stackrel{d}{=} \sum_k \xi_k,$$

where $\{\xi_k\}$ is a collection of independent Bernoulli random variables.

Concentration of the counting function

Concentration of the counting function

Since $\mathcal{N}_D$ is a sum of i.i.d. Bernoullis, Bernstein's inequality applies:

$$\mathbb{P}\left[|\mathcal{N}_D - \mathbb{E}\mathcal{N}_D| > t\right] \leq 2\exp\left(-\min\left\{\frac{t^2}{4\sigma_D^2}, \frac{t}{2}\right\}\right),$$

where $\sigma_D^2 = \operatorname{Var}\mathcal{N}_D$.

Concentration of individual eigenvalues

Concentration of individual eigenvalues

Dallaporta's argument: Let $\lambda_1 < \lambda_2 < \cdots < \lambda_N$ be the eigenvalues of a GUE matrix, and define their predicted locations γ_k by

$$\rho_{sc}((-\infty, \gamma_k]) = \frac{1}{2\pi} \int_{-2}^{\gamma_k} \sqrt{4 - x^2} dx = \frac{k}{N}.$$

Concentration of individual eigenvalues

Dallaporta's argument: Let $\lambda_1 < \lambda_2 < \cdots < \lambda_N$ be the eigenvalues of a GUE matrix, and define their predicted locations γ_k by

$$\rho_{sc}((-\infty, \gamma_k]) = \frac{1}{2\pi} \int_{-2}^{\gamma_k} \sqrt{4 - x^2} dx = \frac{k}{N}.$$

For $1 \leq k \leq N$,

$$\mathbb{P}\left[\lambda_k - \gamma_k \geq \frac{u}{N}\right] = \mathbb{P}\left[\mathcal{N}_{\gamma_k + \frac{u}{N}} < k\right],$$

but

$$\mathbb{E}\left[\mathcal{N}_{\gamma_k + \frac{u}{N}}\right] \approx k + Cu,$$

and (for a large range of t) $\mathcal{N}_t$ concentrates around its mean.

Proposition (Dallaporta)

Fix $\eta \in \left(0, \frac{1}{2}\right]$, *and suppose that* $\eta N \leq k \leq (1-\eta)N$. *There exist constants* C, c, c', δ *(all depending on* η*) such that for* $c \leq u \leq c'N$,

$$\mathbb{P}\left[|\lambda_k - \gamma_k| \geq \frac{u}{N}\right] \leq 4\exp\left[-\frac{C^2u^2}{2c\delta\log(N) + Cu}\right].$$

Expected distance to the semi-circle law

Expected distance to the semi-circle law

The L_p-Kantorovich distance between probability measures μ and ν on a nice metric space $\mathcal{X}$ is

$$W_p(\mu,\nu) := \inf\left\{ \left[\int_{\mathcal{X}^2} d(x,y)^p d\pi(x,y)\right]^{\frac{1}{p}} \middle| \begin{array}{c} \pi(A\times\mathcal{X})=\mu(A) \\ \pi(\mathcal{X}\times B)=\nu(B) \end{array} \right\}.$$

Expected distance to the semi-circle law

The L_p-Kantorovich distance between probability measures μ and ν on a nice metric space $\mathcal{X}$ is

$$W_p(\mu,\nu) := \inf\left\{ \left[\int_{\mathcal{X}^2} d(x,y)^p d\pi(x,y)\right]^{\frac{1}{p}} \middle| \begin{array}{c} \pi(A\times\mathcal{X})=\mu(A) \\ \pi(\mathcal{X}\times B)=\nu(B) \end{array} \right\}.$$

$$\Longrightarrow \quad \text{If } \mu_N := \frac{1}{N}\sum_{k=1}^N \delta_{\lambda_k} \quad \text{and} \quad \nu_N := \frac{1}{N}\sum_{k=1}^N \delta_{\gamma_k},$$

then

$$W_p^p(\mu_N,\nu_N) \le \frac{1}{N}\sum_{k=1}^N |\lambda_k - \gamma_k|^p.$$

$$W_p^p(\mu_N, \nu_N) \leq \frac{1}{N} \sum_{k=1}^{N} |\lambda_k - \gamma_k|^p.$$

$$W_p^p(\mu_N, \nu_N) \le \frac{1}{N} \sum_{k=1}^{N} |\lambda_k - \gamma_k|^p.$$

Concentration of λ_k near γ_k gives good bounds on $\mathbb{E}|\lambda_k - \gamma_k|^p$.

$$W_p^p(\mu_N, \nu_N) \leq \frac{1}{N} \sum_{k=1}^{N} |\lambda_k - \gamma_k|^p.$$

Concentration of λ_k near γ_k gives good bounds on $\mathbb{E}|\lambda_k - \gamma_k|^p$.

In particular, Dallaporta's estimate for the concentration of λ_k about γ_k gives that

$$\mathbb{E} W_2(\mu_N, \rho_{sc}) \leq C \frac{\sqrt{\log(N)}}{N}.$$

Eigenvalue concentration for other ensembles

Eigenvalue concentration for other ensembles

If U is a random unitary matrix, then U has eigenvalues

$$\{e^{i\theta_k}\}_{k=1}^{N},$$

for $0 \leq \theta_1 < \theta_2 < \cdots < \theta_N < 2\pi$.

Eigenvalue concentration for other ensembles

If U is a random unitary matrix, then U has eigenvalues

$$\{e^{i\theta_k}\}_{k=1}^{N},$$

for $0 \leq \theta_1 < \theta_2 < \cdots < \theta_N < 2\pi$.

We define the predicted locations to be

$$\{e^{\frac{ik}{N}}\}_{k=1}^{N}.$$

Eigenvalue concentration for other ensembles

If U is a random unitary matrix, then U has eigenvalues

$$\{e^{i\theta_k}\}_{k=1}^{N},$$

for $0 \leq \theta_1 < \theta_2 < \cdots < \theta_N < 2\pi$.

We define the predicted locations to be

$$\{e^{\frac{ik}{N}}\}_{k=1}^{N}.$$

concentration of $\mathcal{N}_{[0,\theta]}$ $\rightsquigarrow$ concentration of $e^{i\theta_k}$ about $e^{\frac{ik}{N}}$

Eigenvalue concentration for other ensembles

If U is a random unitary matrix, then U has eigenvalues

$$\{e^{i\theta_k}\}_{k=1}^{N},$$

for $0 \leq \theta_1 < \theta_2 < \cdots < \theta_N < 2\pi$.

We define the predicted locations to be

$$\{e^{\frac{ik}{N}}\}_{k=1}^{N}.$$

concentration of $\mathcal{N}_{[0,\theta]}$ $\rightsquigarrow$ concentration of $e^{i\theta_k}$ about $e^{\frac{ik}{N}}$

$$\rightsquigarrow \quad \mathbb{E}W_p(\mu_N, \nu) \leq \frac{Cp\sqrt{\log(N)+1}}{N},$$

where ν is the uniform distribution on $\mathbb{S}^1 \subseteq \mathbb{C}$.

Eigenvalue concentration for other ensembles

Eigenvalue concentration for other ensembles

Theorem (Rains)

For any $m \geq 1$,

$$[\mathbb{U}(N)]^m \overset{e.w.}{\sim} \bigoplus_{0 \leq k < m} \mathbb{U}\left(\left\lceil \frac{N-k}{m} \right\rceil\right).$$

Eigenvalue concentration for other ensembles

Theorem (Rains)

For any $m \geq 1$,

$$[\mathbb{U}(N)]^m \overset{e.w.}{\sim} \bigoplus_{0 \leq k < m} \mathbb{U}\left(\left\lceil \frac{N-k}{m} \right\rceil\right).$$

$\Longrightarrow$ If $U \sim Haar(\mathbb{U}(N))$ and $\mathcal{N}_\theta^{(m)}$ is the number of eigenvalues $e^{i\phi_k}$ of U^m with $0 \leq \phi_k \leq \theta$, then

$$\mathcal{N}_\theta^{(m)} \overset{d}{=} \mathcal{N}_{1,\theta} + \cdots + \mathcal{N}_{m,\theta},$$

where the $\mathcal{N}_{k,\theta}$ are the counting functions for m independent random matrices from $\mathbb{U}\left(\left\lceil \frac{N-k}{m} \right\rceil\right)$.

Eigenvalue concentration for other ensembles

Then $\mathcal{N}_\theta^{(m)}$ is also equal in distribution to a sum of independent Bernoullis, and the earlier route applies:

Eigenvalue concentration for other ensembles

Then $\mathcal{N}_\theta^{(m)}$ is also equal in distribution to a sum of independent Bernoullis, and the earlier route applies:

concentration of $\mathcal{N}_{[0,\theta]}^{(m)}$ $\rightsquigarrow$ concentration of $e^{i\theta_k}$ about $e^{\frac{ik}{N}}$

$$\rightsquigarrow \qquad \mathbb{E} W_p(\mu_N, \nu) \le \frac{Cp\sqrt{m \log\left(\frac{N}{m}\right) + 1}}{N},$$

where ν is the uniform distribution on $\mathbb{S}^1 \subseteq \mathbb{C}$, and $m \in \{1, \ldots, N\}$.

Eigenvalue concentration for other ensembles

The eigenvalues of the complex Ginibre ensemble are also a determinantal point process, but on $\mathbb{C}$.

Eigenvalue concentration for other ensembles

The eigenvalues of the complex Ginibre ensemble are also a determinantal point process, but on $\mathbb{C}$.

concentration of $\mathcal{N}_D$ $\rightsquigarrow$?

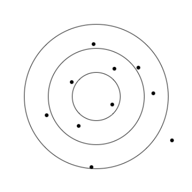
Eigenvalue concentration for other ensembles

The eigenvalues of the complex Ginibre ensemble are also a determinantal point process, but on $\mathbb{C}$.

concentration of $\mathcal{N}_D$ $\rightsquigarrow$?

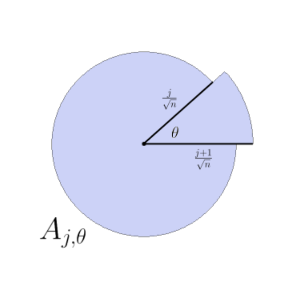
We define the spiral order $\prec$ on $\mathbb{C}$: say $w \prec z$ if

- $\lfloor \sqrt{n}|w| \rfloor < \lfloor \sqrt{n}|z| \rfloor$;
  or
- $\lfloor \sqrt{n}|w| \rfloor = \lfloor \sqrt{n}|z| \rfloor$ and $\arg w < \arg z$;
  or
- $\lfloor \sqrt{n}|w| \rfloor = \lfloor \sqrt{n}|z| \rfloor$, $\arg w = \arg z$, and $|w| \geq |z|$.

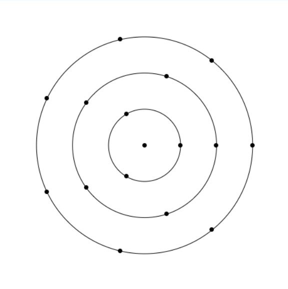


Eigenvalue concentration for other ensembles

We apply the concentration of the counting function to the sets:

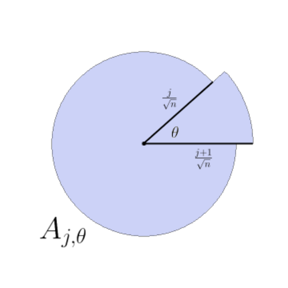


and define predicted eigenvalue locations:

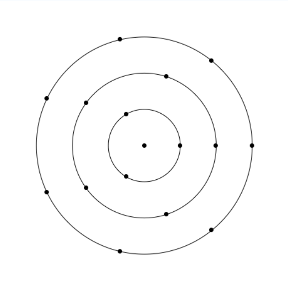


Eigenvalue concentration for other ensembles

We apply the concentration of the counting function to the sets:



and define predicted eigenvalue locations:



$$\rightsquigarrow \mathbb{E} W_2(\mu_N, \nu) \leq C \left(\frac{\log(N)}{N} \right)^{\frac{1}{4}},$$

where ν is the uniform distribution on the circle

Almost sure convergence rates

Almost sure convergence rates

Many random matrix ensembles satisfy the following concentration property:

Let $F : S \subseteq \mathbb{M}_N \to \mathbb{R}$ be 1-Lipschitz with respect to $\| \cdot \|_{H.S.}$. Then

$$\mathbb{P}\Big[\big|F(M) - \mathbb{E}F(M)\big| > t\Big] \leq Ce^{-cNt^2}.$$

Almost sure convergence rates

Many random matrix ensembles satisfy the following concentration property:

Let $F : S \subseteq \mathbb{M}_N \to \mathbb{R}$ be 1-Lipschitz with respect to $\| \cdot \|_{H.S.}$. Then

$$\mathbb{P}\Big[\big|F(M) - \mathbb{E}F(M)\big| > t\Big] \leq Ce^{-cNt^2}.$$

For a normal matrix, the Hoffman-Wieland inequality implies that $W_1(\underbrace{\mu_M}_{\text{spectral measure of } M},\ \underbrace{\mu}_{\text{reference measure}})$

is a $\frac{1}{\sqrt{N}}$-Lipschitz function of M.

Almost sure convergence rates

Almost sure convergence rates

Concentration in an ensemble of normal matrices

Good estimates for $\mathbb{E}W_p(\mu_M, \mu)$ from d.p.p. structure

Ensembles with the concentration phenomenon

Ensembles with the concentration phenomenon

- GUE

Ensembles with the concentration phenomenon

- GUE
- Wigner matrices in which the entries satisfy a quadratic transportation cost inequality with constant $\frac{c}{\sqrt{N}}$.

Ensembles with the concentration phenomenon

- GUE
- Wigner matrices in which the entries satisfy a quadratic transportation cost inequality with constant $\frac{c}{\sqrt{N}}$.
- The compact classical groups: $\mathbb{O}\left(N\right)$, $\mathbb{SO}\left(N\right)$, $\mathbb{U}\left(N\right)$, $\mathbb{SU}\left(N\right)$, $\mathbb{Sp}\left(2N\right)$

Ensembles with the concentration phenomenon

- GUE
- Wigner matrices in which the entries satisfy a quadratic transportation cost inequality with constant $\frac{c}{\sqrt{N}}$.
- The compact classical groups: $\mathbb{O}(N)$, $\mathbb{SO}(N)$, $\mathbb{U}(N)$, $\mathbb{SU}(N)$, $\mathbb{Sp}(2N)$
- Ensembles with matrix density $\propto e^{-N \operatorname{Tr}(u(M))}$, with $u''(x) \geq c > 0$.

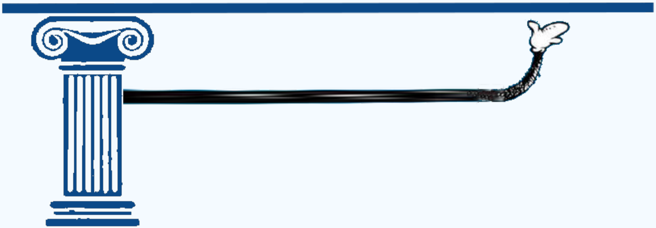
Ensembles with the concentration phenomenon and d.p.p. structure

- GUE

- The compact classical groups: $\mathbb{O}(N)$, $\mathbb{SO}(N)$, $\mathbb{U}(N)$, $\mathbb{SU}(N)$, $\mathbb{Sp}(2N)$

Without determinantal structure

Almost sure convergence rates



Concentration in an ensemble of normal matrices

Good estimates for $\mathbb{E}W_p(\mu_M, \mu)$ from ~~d.p.p. structure~~ concentration

Average distance to average without determinantal structure

Define the centered stochastic process

$$X_f := \int f d\mu_M - \mathbb{E} \int f d\mu_M,$$

indexed by $\{f : \|f\|_{BL} \leq 1\}$.

Average distance to average without determinantal structure

Define the centered stochastic process

$$X_f := \int f d\mu_M - \mathbb{E} \int f d\mu_M,$$

indexed by $\{f : \|f\|_{BL} \leq 1\}$.

The concentration phenomenon implies that

$$\mathbb{P}\Big[\big|X_f - X_g\big| \geq t\Big] \leq Ce^{-\frac{cNt^2}{\|f-g\|_{BL}^2}};$$

that is, $\{X_f\}$ is a sub-Gaussian process with respect to

$$d_N(f,g) := \frac{\|f-g\|_{BL}}{\sqrt{N}}.$$

Theorem (Dudley)

If a stochastic process $\{X_t\}_{t\in T}$ satisfies the a sub-Gaussian increment condition

$$\mathbb{P}\left[\left|X_t - X_s\right| > \epsilon\right] \leq Ce^{-\frac{\epsilon^2}{2\delta^2(s,t)}} \qquad \forall\epsilon > 0,$$

then

$$\mathbb{E}\sup_{t\in T} X_t \leq C\int_0^\infty \sqrt{\log N(T,\delta,\epsilon)}d\epsilon,$$

where $N(T,\delta,\epsilon)$ is the ϵ-covering number of T with respect to the distance δ.

Theorem (Dudley)

If a stochastic process $\{X_t\}_{t \in T}$ satisfies the a sub-Gaussian increment condition

$$\mathbb{P}\left[\left|X_t - X_s\right| > \epsilon\right] \leq C e^{-\frac{\epsilon^2}{2\delta^2(s,t)}} \qquad \forall \epsilon > 0,$$

then

$$\mathbb{E} \sup_{t \in T} X_t \leq C \int_0^\infty \sqrt{\log N(T, \delta, \epsilon)} d\epsilon,$$

where $N(T, \delta, \epsilon)$ is the ϵ-covering number of T with respect to the distance δ.

For us, $T = \{f : \|f_{BL} \leq 1\}$ and $\delta(f, g) = \dfrac{\|f - g\|_{BL}}{\sqrt{N}}$.

So: what is $N\left(T, \frac{\|\cdot\|_{BL}}{\sqrt{N}}, \epsilon\right)$?

So: what is $N\left(T, \frac{\|\cdot\|_{BL}}{\sqrt{N}}, \epsilon\right)$?

Bad news: $N\left(T, \frac{\|\cdot\|_{BL}}{\sqrt{N}}, \epsilon\right) = \infty$.

So: what is $N\left(T, \frac{\|\cdot\|_{BL}}{\sqrt{N}}, \epsilon\right)$?

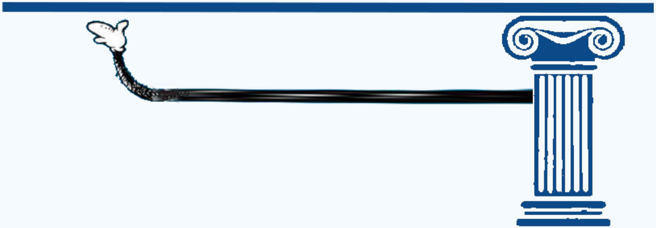
Bad news: $N\left(T, \frac{\|\cdot\|_{BL}}{\sqrt{N}}, \epsilon\right) = \infty$.

This is not that big a deal:

approximating Lipschitz functions by piecewise affine functions and using volumetric estimates in the resulting finite-dimensional normed space of approximating functions does the job.

Almost sure rates without *a priori* concentration of distance

Almost sure convergence rates



Concentration ~~in an ensemble of normal matrices~~ of individual eigenvalues

Good estimates for $\mathbb{E}W_p(\mu_M, \mu)$ from d.p.p. structure

The best of our worlds: $\mathbb{U}(N)$ and friends, GUE

Almost sure convergence rates

Concentration in an ensemble of normal matrices

Good estimates for $\mathbb{E}W_p(\mu_M, \mu)$ from d.p.p. structure

Thank you.