Pseudospectrum and spectral anti-concentration: beyond mean field models

Lake Arrowhead, June 21, 2018

Nick Cook, UCLA

Partly based on joint work with Walid Hachem, Jamal Najim and David Renfrew

Outline of the talk

1. Background: Limiting ESDs and Controlling the Pseudospectrum
2. Spectral anti-concentration via the Schwinger–Dyson equations
3. Spectral anti-concentration: geometric approach

Background: Limiting ESDs and Controlling the Pseudospectrum

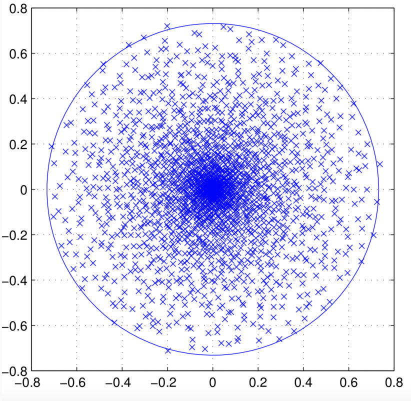
Structured non-Hermitian random matrices

- We study ESDs $\mu_M = \frac{1}{n}\sum_{j=1}^{n}\delta_{\lambda_j(M)}$ for $n \times n$ matrices

$$M = \frac{1}{\sqrt{n}} A \odot X = (\frac{1}{\sqrt{n}} a_{ij}\xi_{ij})$$

  with ξ_{ij} iid, standardized, and deterministic weights $a_{ij} \in [0,1]$.
  For this talk take $a_{ij} \in \{0,1\}$.
- Ex: For $A \equiv \mathbf{1}\mathbf{1}^{\mathsf{T}}$, $\mu_M \to$ circular law (almost surely).
- Applications: stability analysis for large dynamical systems (ecology, neuroscience...). There is interest in allowing A to be structured and sparse.

Simulated ESDs for $M = (\frac{1}{\sqrt{n}} a_{ij} \xi_{ij})$

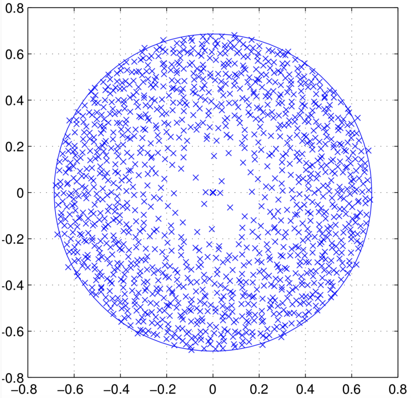


$n = 2000$

$\xi \in \{\pm\frac{1}{\sqrt{2}} \pm i\frac{1}{\sqrt{2}}\}$ uniform.

$a_{ij} = \sigma(\frac{i}{n}, \frac{j}{n})$, with
$\sigma(x, y) = (x + y)^2 1(|x - y| \leq 0.1)$

Simulated ESDs for $M = (\frac{1}{\sqrt{n}} a_{ij} \xi_{ij})$



$n = 2001$

$\xi \in \{\pm \frac{1}{\sqrt{2}} \pm i \frac{1}{\sqrt{2}}\}$ uniform.

$$A_n = (a_{ij}) = \begin{pmatrix} 0 & \mathbf{1}_{n/3} & \mathbf{1}_{n/3} \\ \mathbf{1}_{n/3} & 0 & 0 \\ \mathbf{1}_{n/3} & 0 & 0 \end{pmatrix}.$$

Log-characteristic polynomials and Hermitization

To show ESDs of M converge weakly (a.s. or in probability) to μ, it's enough to show

$$\frac{1}{n}\log|\det(M-z)| = \frac{1}{n}\sum_{j=1}^{n}\log|\lambda_j(M)-z| \to \int_{\mathbb{C}}\log|\lambda-z|d\mu(\lambda) \quad a.e.\ z\in\mathbb{C}.$$

(Recover μ by taking Laplacian.)

Left hand side can be re-expressed as

$$\frac{1}{2n}\log\det[(M-z)^*(M-z)] = \int_0^\infty \log(s)d\mu_{|M-z|}(s).$$

(This is the template for the construction of Brown's spectral measure for non-normal elements of a W^* probability space.)

We need to be worried about very small singular values of $M-z$, i.e. large singular values of the resolvent $(M-z)^{-1}$.

(ε-)*pseudospectrum*: $\Lambda_\varepsilon(M) = \Lambda(M)\cup\left\{z\in\mathbb{C} : \|(M-z)^{-1}\|_{op}\geq 1/\varepsilon\right\}$

Structured random matrices have small pseudospectrum

Theorem (C. '16)
Let $M = \frac{1}{\sqrt{n}} A \odot X = (\frac{1}{\sqrt{n}} a_{ij}\xi_{ij})$ *with* $a_{ij} \in [0,1]$ *deterministic,* ξ_{ij} *iid, centered, unit variance and* $\mathbb{E}|\xi_{ij}|^{4+\varepsilon} < \infty$*. There are constants* $C(\varepsilon, r, R), c(\varepsilon) > 0$ *such that for any* $z \in \mathbb{C}$ *with* $|z| \in [r, R] \subset (0, \infty)$*,*

$$\mathbb{P}(\|(M-z)^{-1}\|_{op} \geq n^C) \lesssim_{\varepsilon,r,R} n^{-c}$$

- Earlier works on invertibility make strong pseudorandomness assumptions on A, don't use the shift (optimal condition found by Rudelson–Zeitouni '13, see also C. '16).
- *Question:* Proof gives tower dependence of C on $1/r$. Can we do better?

Spectral anti-concentration via the Schwinger–Dyson equations

Controlling other large singular values of $(M-z)^{-1}$

- To prove uniform integrability of $\log(s)$ by $\mu_{|M-z|}$, need more than just $\|(M-z)^{-1}\|_{op} = n^{O(1)}$. Would like a *spectral anti-concentration* estimate

$$\mu_{|M-z|}([0,\eta]) \lesssim \eta \qquad w.h.p.\ \forall \eta \geq n^{-c}.$$

- It's enough to bound

$$\frac{1}{\eta}\mu_{\boldsymbol{M}_z}([-\eta,\eta]) \lesssim \frac{1}{2n}\sum_{j=1}^{2n} \frac{\eta}{\lambda_j(\boldsymbol{M}_z)^2+\eta^2} = \frac{1}{2n}\mathrm{Im}\ \mathrm{Tr}(\boldsymbol{M}_z - i\eta)^{-1} \lesssim 1$$

for all $\eta \geq n^{-c}$, where $\boldsymbol{M}_z = \begin{pmatrix} 0 & M-z \\ M^* - \bar{z} & 0 \end{pmatrix}$.

Schwinger–Dyson equations

- In the mean field setting $A = \mathbf{1}\,\mathbf{1}^\mathsf{T}$, Stieltjes transform satisfies a (cubic) scalar polynomial equation.
- In non-mean field case, we have $\operatorname{Im}(\boldsymbol{M}_z - i\eta)^{-1} \approx \operatorname{diag}(\boldsymbol{u}, \boldsymbol{v})$, where $\boldsymbol{u}, \boldsymbol{v} : [n] \to \mathbb{R}_+$ solve the cubic polynomial system

$$
\begin{aligned}
-\frac{1}{\boldsymbol{u}} &= \eta + S\boldsymbol{v} - \frac{|z|^2}{\eta + S^\mathsf{T}\boldsymbol{u}} \\
-\frac{1}{\boldsymbol{v}} &= \eta + S\boldsymbol{u} - \frac{|z|^2}{\eta + S^\mathsf{T}\boldsymbol{v}}
\end{aligned}
$$

and $S(i,j) = \frac{1}{n} a_{ij}^2$.

Theorem (C., Hachem, Najim, Renfrew '16)

Assume $A = (a_{ij})$ is robustly irreducible. Then for any fixed $z \neq 0$,

$$
\|\boldsymbol{u}\|_\infty, \|\boldsymbol{v}\|_\infty \lesssim_{|z|} 1
$$

uniformly for $\eta > 0$.

Boundedness of solutions to the S–D equations

Idea: (Recall $a_{ij} \in \{0,1\}$.) View A as the adjacency matrix of a digraph. Up to constants, the S–D equations give the local constraints

$$\boldsymbol{u}(j) \asymp \min\left\{\langle \boldsymbol{u}\rangle_{\mathcal{N}_{in}(j)},\ \frac{1}{\langle \boldsymbol{v}\rangle_{\mathcal{N}_{out}(j)}}\right\}, \qquad \boldsymbol{v}(j) \quad \asymp \min\left\{\langle \boldsymbol{v}\rangle_{\mathcal{N}_{out}(j)}, \frac{1}{\langle \boldsymbol{u}\rangle_{\mathcal{N}_{in}(j)}}\right\}.$$

We also have a global trace identity

$$\sum_{j\in[n]} \boldsymbol{u}(j) = \sum_{j\in[n]} \boldsymbol{v}(j).$$

1. Assume $\boldsymbol{u}(j_0)$ is large for some $j_0 \in [n]$.
2. Use the robust irreducibility assumption and the local constraints to "propagate" this property to almost all other indices j.
3. On the other hand, we have $\boldsymbol{u}(j)\boldsymbol{v}(j) \lesssim 1$ for all j, so we conclude $\boldsymbol{v}(j)$ is small for almost all indices j. But this contradicts the trace identity.

Spectral anti-concentration: geometric approach

Bounding the density of states

Consider random symmetric matrices of the form $W = A \odot X = (a_{ij}\xi_{ij})$ where $\{\xi_{ij}\}_{1\le i\le j\le n}$ are independent, standardized, sub-Gaussian variables, $a_{ij} \in \{0, 1\}$ are deterministic weights.

For an interval $\mathcal{I} \subset \mathbb{R}$, when can we show

$$\mu_{\frac{1}{\sqrt{n}}W}(\mathcal{I}) \lesssim |\mathcal{I}| \qquad w.h.p.?$$

Theorem (Local semicircle law, Erdős–Schlein–Yau '08) Let $A = \mathbf{1}\mathbf{1}^{\mathsf{T}}$. Then with high probability, for any interval $\mathcal{I} \subset \mathbb{R}$ with $|\mathcal{I}| \ge n^{-1+\varepsilon}$,

$$|\mu_H(\mathcal{I}) - \mu_{sc}(\mathcal{I})| = o(|\mathcal{I}|).$$

Recent extensions to non-mean field models satisfying some technical hypotheses by Erdös et al., using careful analysis of associated Schwinger–Dyson equations (a.k.a. the quadratic vector equations).

Spectral anti-concentration under expanding support

Write $\mathcal{N}(i) = \{j \in [n] : a_{ij} = 1\}$, and define the δ-dense boundary of a set $J \subset [n]$ as

$$\mathcal{D}_\delta(J) = \{i \in J^c : |\mathcal{N}(i) \cap J| \geq \delta |J|\}.$$

Say that A is a (δ, ε)-*expander* if

$$|\mathcal{N}(i)| \geq \delta n \quad \forall i \in [n] \qquad \text{and} \qquad |\mathcal{D}_\delta(J)| \geq \min(\varepsilon |J|, |J^c|) \quad \forall J \subset [n].$$

Theorem (C. '18)

Let A be a (δ, ε)-expander for some $\delta, \varepsilon \in (0,1)$. Fix $\tau > 0$. With high probability, for any $\mathcal{I} \subset \mathbb{R} \setminus (-\tau, \tau)$ with $|\mathcal{I}| \geq Cn^{-1} \log n$, we have

$$\mu_{\frac{1}{\sqrt{n}} W}(\mathcal{I}) \lesssim |\mathcal{I}|.$$

*Note the assumptions on A permit an atom at 0.

*Could reach intervals of length $n^{-1}\sqrt{\log n}$ using recent ideas of Nguyen '17.

Spectral anti-concentration: ideas of the proof

Fix $\mathcal{I} \subset \mathbb{R}$ away from 0 of length η. We want to bound the spectral concentration event

$$Conc(W, n_0) = \{W \text{ has } \geq n_0 \text{ eigenvalues in } \sqrt{n}\mathcal{I}\}.$$

Idea: can use interlacing and weak delocalization of eigenvectors to reduce to bounding the probability that

$$s_{\min}(\Pi W_{J^c,J}) \leq C\eta n/\sqrt{m}.$$

where J is a random set of size $m \ll n_0$ and Π is a spectral projectionf or W_{J^c} of rank $\geq n_0 - m$.

Key: Π and $W_{J^c,J}$ are independent. We've reduced to invertibility of an $(n_0 - m) \times m$ (tall) projection $\Pi W_{J^c,J}$ of a (very tall) submatrix $W_{J^c,J}$.

Main technical step: show no-gaps delocalization holds for the eigenvectors of W_{J^c}.