# Spectral stability under random perturbations

Jorge Garza Vargas
Joint work with Jess Banks, Archit Kulkarni and Nikhil Srivastava

Berkeley
UNIVERSITY OF CALIFORNIA

# About this project

- This is a non-asymptotic analysis of non-Hermitian random matrices.
- *Model:* We consider the model $A + M$ where $A \in \mathbb{R}^{n \times n}$ is any deterministic matrix, and $M$ is a real random matrix with independent entries.
- *Motivation*: Applications in numerical linear algebra (approximate diagonalization), study of overlaps in physics, understanding non-Hermitian random matrices better.
- *Quantities we care about:* Eigenvalue condition numbers (overlaps), pseudospectrum and eigenvalue gaps.
- *Tools:* Tail bounds for singular values and anti-concentration results.

# Eigenvalue condition numbers

Let $A$ be an $n \times n$ diagonalizable matrix (possibly non-Hermitian).

- Consider the spectral decomposition of $A = \sum_{i=1}^{n} \lambda_i v_i u_i^*$.
  - Here $\lambda_i$ are the eigenvalues, $v_i$ the right eigenvectors and $u_i$ the left eigenvectors.
  - Normalize such that and $u_i^* v_i = 1$.
- Eigenvalue condition numbers: $\kappa(\lambda_i) \stackrel{\text{def}}{=} \|u_i\| \|v_i\|$.
  - If $A$ is a normal matrix then $\kappa(\lambda_i) = 1$ for all $i = 1, \ldots, n$. If $A$ is a Jordan block $\kappa(\lambda_i) = \infty$.
  - If $A(t)$ smooth trajectory in the space of diagonalizable matrices and $\sum_{i=1}^{n} \lambda_i(t) v_i(t) u_i^*(t)$ then

$$|\lambda_i'(t_0)| = |u_i^* A'(t_0) v_i| \leq \kappa(\lambda_i(t_0)) ||A'(t_0)||$$

# Pseudospectrum

Let $A$ be an $n \times n$ matrix, the $\epsilon$-pseudospectrum of $A$ is defined by

$$\Lambda_\epsilon(A) \stackrel{\text{def}}{=} \left\{ z \in \mathbb{C} : \|(z-A)^{-1}\| \geq \frac{1}{\epsilon} \right\}.$$

- *Fact 1.* $\Lambda_\epsilon(A) = \bigcup_{\|E\| \leq \epsilon} \mathrm{Spec}(A+E)$.
- *Fact 2.* If $A$ has distinct eigenvalues, then as $\epsilon \to 0$ we have

$$\Lambda_\epsilon(A) \subset \bigcup_{i=1}^{n} \mathbb{D}(\lambda_i, r_i),$$

  where $r_i = \epsilon\kappa(\lambda_i) + O(\epsilon^2)$.
- *Fact 3.* If $A$ has distinct eigenvalues, then for any open $\mathcal{B} \subset \mathbb{C}$

$$\lim_{\epsilon \to 0} \frac{\mathrm{vol}(\Lambda_\epsilon(A) \cap \mathcal{B})}{\epsilon^2} = \pi \sum_{\lambda_i \in \mathcal{B}} \kappa(\lambda_i)^2.$$

# Eigenvalue gap

Let $A$ be an $n \times n$ matrix and define

$$\mathrm{gap}(A) \stackrel{\text{def}}{=} \min_{i \neq j} |\lambda_i(A) - \lambda_j(A)|.$$

- If $A(t)$ smooth trajectory in the space of diagonalizable matrices with distinct eigenvalues and $\sum_{i=1}^{n} \lambda_i(t) v_i(t) u_i^*(t)$ is the spectral decomposition, then

$$\frac{1}{\sqrt{n}} \frac{\|v_i'(t_0)\|}{\|v_i(t_0)\|} \leq \frac{\sqrt{\sum_{j=1}^{n} \kappa(\lambda_j(t_0))^2} \|A'(t_0)\|}{\min_{j:j \neq i} |\lambda_j(t_0) - \lambda_i(t_0)|} \leq \frac{\sqrt{\sum_{j=1}^{n} \kappa(\lambda_j)^2} \|A'(t_0)\|}{\mathrm{gap}(A(t_0))}.$$

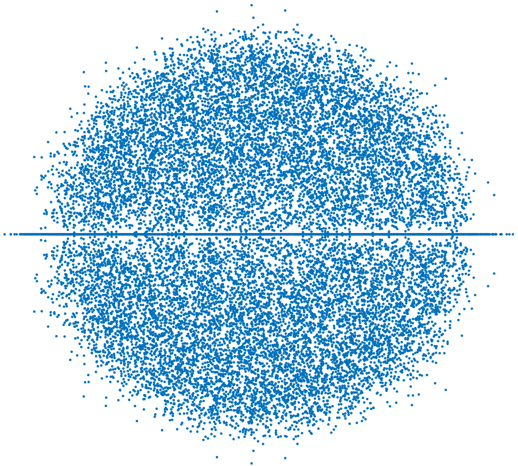
# Overlaps

Given $A = \sum_{i=1}^{n} \lambda_i v_i u_i^*$ define $\mathcal{O}_{ij} \stackrel{\text{def}}{=} (v_j^* v_i)(u_j^* u_i)$. Note that $\mathcal{O}_{ii} = \kappa(\lambda_i)^2$.

- Complex Ginibre matrices (i.i.d complex Gaussian entries):
  - (Chalker, Mehlig 98) Formulas for the limiting expected overlaps.
  - (Fyodorov 2018) Non-asymptotic formula for the joint distribution of an eigenvalue and its diagonal overlap.
  - (Bourgade, Dubach, 2019) Asymptotic formula for the limiting distribution of the diagonal overlaps. Asymptotic correlations between overlaps.
- Real Ginibre matrices (i.i.d real Gaussian entries):
  - (Fyodorov 2018) Non-asymptotic formula for the joint distribution of a real eigenvalue and its diagonal overlap.

# Real random matrices

- (Edelman 94) The expected number of real eigenvalues in an $n \times n$ real Ginibre matrix is $\sqrt{\frac{2n}{\pi}}\left(1 + O\left(\frac{1}{n}\right)\right)$



# Eigenvalue gaps

Let $M$ be an $n \times n$ random matrix with **i.i.d.** real entries of unit variance and bounded fourth moment.

- (Ge 2017) If the entries of $M$ are centered, then for any $s = o(n^{-4+o(1)})$ then the event $\{\text{gap}(M) \geq s\}$ happens with high probability.
- (Ge 2017) If the entries of $M$ are centered, for $s = o(n^{-2+o(1)})$ the event that $M$ has no truly complex eigenvalues in the band $\mathcal{B}_s = \{z \in \mathbb{C} : |\text{Im}(z)| \leq s\}$ happens with high probability.
- (Luh, O'Rourke 2020) Refined the results of Ge and dropped the mean 0 assumption, but still require i.i.d entries.

# Complex Gaussian perturbations

Let $G_n$ be a normalized complex Ginibre matrix (i.e. $(G_n)_{ij} \sim N\left(0, \frac{1_{\mathbb{C}}}{n}\right)$.
Let $A \in \mathbb{C}^{n\times n}$ with $\|A\| \leq 1$ deterministic.Take any $\gamma > 0$ and let

$$\lambda_i \stackrel{\text{def}}{=} \lambda_i(A + \gamma G_n).$$

- (Banks, Kulkarni, Mukherjee, Srivastava, 2019) With high probability

$$\sum_{i=1}^{n} \kappa(\lambda_i)^2 \leq \mathrm{poly}\left(n, \gamma^{-1}\right).$$

- (Banks, JGV, Kulkarni, Srivastava 2019) With high probability

$$\mathrm{gap}(A + \gamma G_n) \geq \mathrm{poly}(n^{-1}, \gamma)$$

# Summary of previous results

| Complex | $\kappa(\lambda_i)$ | gap |
|---|---|---|
| $G_n$ | [CM 98, Fyo 18, BD 19] | [Ge 17, LO 20] |
| $M$ i.i.d | | [Ge 17, LO 20] |
| $A + \gamma G_n$ | [BKMS 19] | [BGVKS 19] |

| Real | $\kappa(\lambda_i), \lambda_i \in \mathbb{R}$ | $\kappa(\lambda_i), \lambda_i \in \mathbb{C} \setminus \mathbb{R}$ | gap |
|---|---|---|---|
| $G_n$ | [Fyo 18] | | [Ge 17, LO 20] |
| $M$ i.i.d | | | [Ge 17, LO 20] |
| $A + \gamma G_n$ | | | |

# Our results: $\kappa(\lambda_i)$

Let $M$ be an $n \times n$ random matrix with independent real entries. Take any $A \in \mathbb{R}^{n\times n}$ with $\|A\| \le 1$.

**Assumption.** The entries of $M$ are absolutely continuous, and there exists $K > 0$, such that the density function of each $M_{ij}$ is pointwise bounded by $K$ and $\mathbb{E}[\|M\|] < \infty$.

## Theorem (Banks, JGV, Kulkarni, Srivastava 2020)

*Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $A + M$, then with high probability*

$$\sum_{\lambda_i \in \mathbb{R}} \kappa(\lambda_i) \le \mathrm{poly}_1(n, K, \mathbb{E}[\|M\|]),$$

*and*

$$\sum_{\lambda_i \in \mathbb{C}\setminus\mathbb{R}} \kappa(\lambda_i)^2 \le \mathrm{poly}_2(n, K, \mathbb{E}[\|M\|])$$

# Our results: Eigenvalue gaps

Let $M$ be an $n \times n$ random matrix with independent real entries. Take any $A \in \mathbb{R}^{n \times n}$ with $\|A\| \leq 1$.

**Assumption.** The entries of $M$ are absolutely continuous with a density function bounded pointwise by $K$ and $\mathbb{E}[\|M\|^8] < \infty$.

## Theorem (Banks, JGV, Kulkarni, Srivastava 2020)

*With high probability*

$$\mathrm{gap}(A + M) \geq \mathrm{poly}\left(n^{-1}, K^{-1}, \mathbb{E}[\|M\|^8]^{-1/8}\right)$$

**Note.** When $M = \gamma G_n$ then $K = \frac{1}{\sqrt{2\pi}\gamma}$ and $\mathbb{E}[\|M\|^8]^{1/8} \leq 7\gamma^8$.

# Summary

| Real | $\kappa(\lambda_i), \lambda_i \in \mathbb{R}$ | $\kappa(\lambda_i), \lambda_i \in \mathbb{C} \setminus \mathbb{R}$ | gap |
|---|---|---|---|
| $G_n$ | [Fyo 18, **BGVKS 19** ] | **[BGVKS 19]** | [Ge 17, LO 20, **BGVKS 19**] |
| $M$ i.i.d | | | [Ge 17, LO 20] |
| $A + M$ a.c. entries | **[BGVKS 19]** | **[BGVKS 19]** | **[BGVKS 19]** |

# Left tails of singular values

Many problems regarding non-Hermitian random matrices can be reduced to studying the singular values of a family of matrices.

- (Szarek 91) Let $G_n$ be a normalized Ginibre matrix (real or complex) and $\sigma_1(G_n) \geq \cdots \geq \sigma_n(G_n)$ its singular values.
  - For real Ginibre:
$$\mathbb{P}\left[\sigma_{n-k+1}(G_n) \leq \frac{k}{n}\epsilon\right] \leq C_1 \epsilon^{k^2}$$
  - For complex Ginibre:
$$\mathbb{P}\left[\sigma_{n-k+1}(G_n) \leq \frac{k}{n}\epsilon\right] \leq C_2 \epsilon^{2k^2}$$
- (Śniady 2001) One can replace $G_n$ by $A + G_n$ for any $A \in \mathbb{R}^{n\times n}$ for real Ginibre and $A \in \mathbb{C}^{n\times n}$ for complex Ginibre.

# Controlling $\kappa(\lambda_i)$ for complex Ginibre

(Banks, Kulkarni, Mukherjee, Srivastava, 2019). *Let $G_n$ be a normalized complex Ginibre, $A \in \mathbb{C}$ with $\|A\| \leq 1$ and $1 > \gamma > 0$. Let $\lambda_1, \cdots \lambda_n$ be the eigenvalues of $A + \gamma G_n$, then with high probability*

$$\sum_{i=1}^{n} \kappa(\lambda_i)^2 \leq \mathrm{poly}(n, \gamma^{-1}).$$

Proof strategy. Recall that with exponentially high probability all eigenvalues of $A + \gamma G_n$ are contained in $\mathbb{D}(0, 4)$.The gist of the proof is in showing that

$$\mathbb{E}\left[\sum_{\lambda_i \in \mathbb{D}(0,4)} \kappa(\lambda_i)^2\right] \leq \mathrm{poly}_1(n, \gamma^{-1})$$

# Controlling $\kappa(\lambda_i)$ for complex Ginibre

Proof sketch. Fix an open set $\mathcal{B} \subset \mathbb{C}$ and recall that

$$\lim_{\epsilon \to 0} \frac{\mathrm{vol}(\Lambda_\epsilon(A + \gamma G_n) \cap \mathcal{B})}{\epsilon^2} = \pi \sum_{\lambda_i \in B} \kappa(\lambda_i)^2.$$

On the other hand, for fixed $\epsilon > 0$

$$\begin{aligned}
\mathbb{E}\left[\mathrm{vol}(\Lambda_\epsilon(A + \gamma G_n) \cap \mathcal{B}\right] &= \mathbb{E}\left[\int_{\mathcal{B}} 1\left\{\|(z - A - \gamma G_n)^{-1}\| \geq \epsilon^{-1}\right\} dz\right] \\
&= \mathbb{E}\left[\int_{\mathcal{B}} 1\left\{\sigma_n(z - A - \gamma G_n) \leq \epsilon\right\} dz\right] \\
&= \int_{\mathcal{B}} \mathbb{P}[\sigma_n(z - A - \gamma G_n) \leq \epsilon] dz \\
&\leq \int_{\mathcal{B}} \frac{n^2\epsilon^2}{\gamma^2} dz = \mathrm{vol}(\mathcal{B}) \frac{n^2\epsilon^2}{\gamma^2}.
\end{aligned}$$

# Controlling $\kappa(\lambda_i)$ for complex Ginibre

Proof sketch. Fix an open set $\mathcal{B} \subset \mathbb{C}$ and recall that

$$\lim_{\epsilon \to 0} \frac{\mathrm{vol}(\Lambda_\epsilon(A + \gamma G_n) \cap \mathcal{B})}{\epsilon^2} = \pi \sum_{\lambda_i \in \mathcal{B}} \kappa(\lambda_i)^2.$$

On the other hand, for fixed $\epsilon > 0$

$$\begin{aligned}
\mathbb{E}\left[\mathrm{vol}(\Lambda_\epsilon(A + \gamma G_n) \cap \mathcal{B}\right] &= \mathbb{E}\left[\int_{\mathcal{B}} 1\left\{\|(z - A - \gamma G_n)^{-1}\| \geq \epsilon^{-1}\right\} dz\right] \\
&= \mathbb{E}\left[\int_{\mathcal{B}} 1\left\{\sigma_n(z - A - \gamma G_n) \leq \epsilon\right\} dz\right] \\
&= \int_{\mathcal{B}} \mathbb{P}[\sigma_n(z - A - \gamma G_n) \leq \epsilon] dz \\
&\leq \int_{\mathcal{B}} \frac{n^2 \epsilon^2}{\gamma^2} dz = \mathrm{vol}(\mathcal{B}) \frac{n^2 \epsilon^2}{\gamma^2}. \quad \square
\end{aligned}$$

# Can we extend this argument to real Ginibre matrices?

Let $G_n$ be a normalized real Ginibre matrix.

- The results of Szarek and Śniady only yield bounds for $\mathbb{P}[\sigma_n(z - A - \gamma G_n) \leq \epsilon]$ when $z \in \mathbb{R}$. This enough to control

$$\sum_{\lambda_i \in \mathbb{R}} \kappa_i(\lambda_i)$$

# Can we extend this argument to analyze real Ginibre matrices?

Let $G_n$ be a normalized real Ginibre matrix.

- The results of Szarek and Śniady only yield bounds for $\mathbb{P}[\sigma_n(z - A - \gamma G_n) \le \epsilon]$ when $z \in \mathbb{R}$.
- (Ge 2017) Let $M$ be a real random matrix with i.i.d. **centered** entries of variance $\frac{1}{n}$ and $z \in \mathbb{C}$, then

$$\mathbb{P}\left[\sigma_n(z - M) \le \frac{\epsilon}{n}\right] \le \frac{C\epsilon^2}{|\mathrm{Im}(z)|} + c^n \quad \text{for some } C > 0, c < 1.$$

- Complications:
  - The entries of $z - A - \gamma G_n$ are not centered.
  - Ge's bound blows up for $z$ close to the real line.
  - Ge's bound doesn't go to 0 as $\epsilon$ goes to 0.

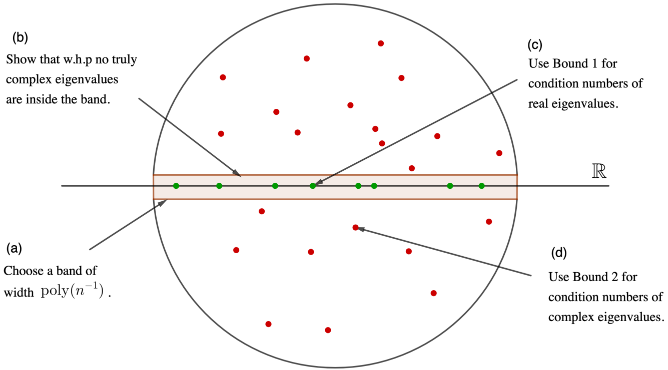
# General strategy

Let $M$ be an $n \times n$ random matrix satisfying the assumptions. The entries of $M$ are independent and absolutely continuous with a density function bounded pointwise by $K$ and $\mathbb{E}[\|M\|] < \infty$.

# General strategy

Let $M$ be an $n \times n$ random matrix satisfying the assumption.

- Step 1: Get a bound of the sort $\mathbb{P}[\sigma_n(z - A - M) \leq \epsilon] \leq \mathbf{poly}(n, K)\epsilon$ for $z \in \mathbb{R}$ and $\mathbb{P}[\sigma_n(z - A - M) \leq \epsilon] \leq \mathbf{poly}(n, K)\frac{\epsilon^2}{|\mathrm{Im}(z)|}$ for $z \in \mathbb{C}$.
- Step 2:



# Singular value bounds for non-i.i.d. random matrices

There were many existing bounds for the left tails of the singular values of $z - A - M$ when $z \in \mathbb{R}$.

- (Tikhomirov 2017) Only assuming that the rows of $M$ are independent and under a technical assumption about the density of each row
$$\mathbb{P}(\sigma_n(z - A - M) \leq \epsilon) \leq C\sqrt{n}K\epsilon.$$

- (Nguyen 2016) Making the same assumptions on $M$ as us, for the $k$-th smallest singular value one has:
$$\mathbb{P}(\sigma_{n-k+1}(z - A - M) \leq \epsilon) \leq \mathrm{poly}_k(n)(K\epsilon)^{k(k-1)}.$$

# Our results

Let $A \in \mathbb{R}^{n \times n}$, and let $M$ be a real $n \times n$ random matrix satisfying the assumption.

## Theorem (Banks, JGV, Kulkarni, Srivastava 2020)

*For all* $1 \leq k \leq n$ *and for all* $z \in \mathbb{R}$

$$\mathbb{P}[\sigma_{n-k+1}(z - A - M) \leq \epsilon] \leq \mathrm{poly}_k(n)(K\epsilon)^{k^2}.$$

*For all* $1 \leq k \leq \sqrt{n} - 2$ *and all* $z \in \mathbb{C} \setminus \mathbb{R}$

$$\mathbb{P}[\sigma_{n-k+1}(z - A - M) \leq \epsilon] \leq \mathrm{poly}_k(n, B_{2k^2}) \left( \frac{K^3 \epsilon^2}{|\mathrm{Im}(z)|} \right)^{k^2}$$

*where* $B_{2k^2} = \mathbb{E}[\|M\|^{2k^2}]^{1/(2k^2)}$.

We also showed that with high probability there are no truly complex eigenvalues of $A + M$ inside a band of width $\mathrm{poly}(n^{-1})$.

# Technical ingredients

- Anticoncentration of bilinear forms (Banks, JGV, Kulkarni, Srivastava 2020): Assume that $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{n\times k}$ are random matrices with independent entries, each with density on $\mathbb{R}$ bounded by $K > 0$. Let $Z \in \mathbb{R}^{n\times n}$, $U, V \in \mathbb{R}^{n\times k}$, and $W \in \mathbb{R}^{k\times k}$ be deterministic, and write $q(\mathbf{X}, \mathbf{Y}) \stackrel{\text{def}}{=} \mathbf{X}^T Z\mathbf{Y} + \mathbf{X}^T U + V^T\mathbf{Y} + W$.

  We gave a bound on the density of $q(\mathbf{X}, \mathbf{Y}) \stackrel{\text{def}}{=} \mathbf{X}^T Z\mathbf{Y} + \mathbf{X}^T U + V^T\mathbf{Y} + W$ in $\mathbb{R}^{k\times k}$ as a function of $k, K$ and the singular values of $Z$.

- Restricted invertibility lemma (Banks, JGV, Kulkarni, Srivastava 2020): Let $X \in \mathbb{C}^{n\times n} \setminus \{0\}$ be positive semidefinite.Then for every $1 \le k \le n$, there exists an $k \times k$ principal submatrix $X_{S,S}$ such that

$$\lambda_k(X_{S,S}) \ge \frac{\lambda_k(X)}{k(n-k+1)}.$$

# Further questions

## Conjecture

*Let $G_n$ be an $n \times n$ normalized real Ginibre matrix. Then for any $\epsilon > 0$ and $z \in \mathbb{C} \setminus \mathbb{R}$ with $|z| \leq 4$ it holds that*

$$\mathbb{P}\left[\sigma_n(z - G_n) \leq \epsilon\right] \leq \frac{Cn^2\epsilon^2}{|\mathrm{Im}(z)|}.$$

## Problem

*Let $M$ be an $n \times n$ random matrix with independent Rademacher/Bernoulli entries. What can you prove about $A + \gamma M$? and how small can $\gamma$ be?*

# Concurrent work

- (Jain, Sah, Sawhney 2020): Also studied $A + M$ under essentially the same assumptions and obtained a high probability bound on

$$\sum_{\lambda_i \in \mathbb{R}} \kappa(\lambda_i)^2, \quad \sum_{\lambda_i \in \mathbb{C} \setminus \mathbb{R}} \kappa(\lambda_i)^2 \quad \text{and} \quad \mathrm{gap}(A + M).$$

Thank you!