Eigenvectors of random perturbations non-normal Toeplitz matrices

Anirban Basak

International Centre for Theoretical Sciences (ICTS)
Tata Institute of Fundamental Research (TIFR)

May 07, 2024

Random Matrices and Related Topics in Jeju

Normal operator/matrix: $NN^\star = N^\star N$;
Non-normal: $NN^\star \neq N^\star N$.

Non-normal operator and its spectral instability

For any bounded normal operator N

$$\|(N-z)^{-1}\| = \frac{1}{\text{dist}(z, \text{Spec}(N))}, \qquad z \notin \text{Spec}(N).$$

Non-normal operator and its spectral instability

For a non-normal operator N and $z \notin \mathrm{Spec}(N)$ one has

either

$$\|(N-z)^{-1}\| \asymp \frac{1}{\mathrm{dist}(z, \mathrm{Spec}(N))}$$

(zone of spectral stability)

or

$$\|(N-z)^{-1}\| \gg \frac{1}{\mathrm{dist}(z, \mathrm{Spec}(N))}.$$

(zone of spectral instability)

Example: Left shift operator on $\mathbb{C}^N$ / Jordan block

$$J_N := \begin{bmatrix} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & 0 & 1 & \\ & & & & 0 & 1 \\ & & & & & 0 \end{bmatrix}, \qquad \mathrm{Spec}(J_N) = \{0\}.$$

Non-normal operator and its spectral instability

$$J_N - z := \begin{bmatrix} -z & 1 & & & & \\ & -z & 1 & & & \\ & & \ddots & \ddots & & \\ & & & -z & 1 & \\ & & & & -z & 1 \\ & & & & & -z \end{bmatrix}.$$

Zone of spectral instability: For $z \in D(0,1) := \{w \in \mathbb{C} : |w| < 1\}$

$$\|(J_N - z)v\|_2 = |z|^N \Rightarrow \|(J_N - z)^{-1}\| \geqslant |z|^{-N}$$

$$v := \begin{pmatrix} 1 & z & z^2 & \cdots & z^{N-1} \end{pmatrix}^{\mathsf{T}} \quad \Rightarrow \quad \|v\|_2 \asymp 1.$$

Zone of spectral stability: For $z \in \mathbb{C} \setminus \overline{D(0,1)}$.

$$\|(J_N - z)^{-1}\| \asymp 1.$$

Challenges with non-normal matrices

(i) The eigenvalue analysis in many applications turns out to be misleading.

(ii) The eigenvalues are sensitive to perturbations and thereby often yielding unreliable results.

Challenges with non-normal matrices

Example (Jordan Block). Simulate a uniformly random unitary matrix U_N and set $\widehat{J}_N := U_N J_N U_N^*$.
$\mathrm{Spec}(J_N) = \mathrm{Spec}(\widehat{J}_N) = \{0\}$.

$$J_N := \begin{bmatrix} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & 0 & 1 & \\ & & & & 0 & 1 \\ & & & & & 0 \end{bmatrix}$$

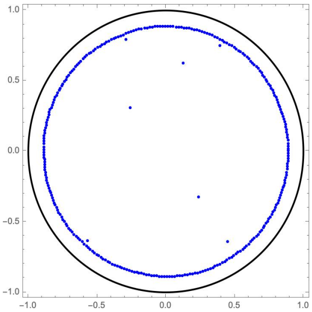
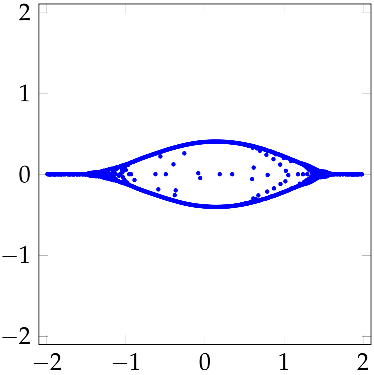


Figure: $N = 1000$. Eigenvalues of $\widehat{J}_N$ computed through Mathematica are plotted in blue and the unit circle $\mathbb{S}^1$ on the complex plane is in black.

Challenges with non-normal matrices

Example. Simulate a Haar U_N. Compute the eigenvalues of $U_N H_N U_N^*$. $N = 1000$.

$$H_N := J_N + D_N$$
$$D_N = \mathrm{diag}(\{X_i\}_{i=1}^N)$$
$$\{X_i\} \text{ i.i.d. } \mathrm{Unif}[-2, 2]$$



Non-periodic one-way model – “limit” of Hatano-Nelson model

(due to Brézin, Feinberg, and Zee)

Eigenvalues move to the ‘Hatano-Nelson bubble’

Pseudospectrum

ε-pseudospectrum $(\varepsilon > 0)$

$$(1).\ \mathrm{Spec}_\varepsilon(A) := \mathrm{Spec}(A) \cup \left\{ z \in \mathbb{C} \backslash \mathrm{Spec}(A) : \|(A-z)^{-1}\| \geqslant \varepsilon^{-1} \right\}$$

$$(2).\ \mathrm{Spec}_\varepsilon(A) = \bigcup_{\|E\| \leqslant \varepsilon} \mathrm{Spec}(A+E)$$

$(3).\ z \in \mathrm{Spec}_\varepsilon(A) \Leftrightarrow z \in \mathrm{Spec}(A)$ or $\exists v_z$ s.t. $\|(A-z)v_z\| \leqslant \varepsilon \, \|v_z\|$

$$(1) \Leftrightarrow (2) \Leftrightarrow (3)$$

[Varah '79], [Trefethen, Embree '05]

Pseudospectrum

For any $A \in \mathbb{C}^{N \times N}$ and any $\varepsilon > 0$

$$\mathrm{Spec}_\varepsilon(A) \supset \mathrm{Spec}(A) + D(0, \varepsilon).$$

If $\| \cdot \| = \| \cdot \|_2$ and $A \in \mathbb{C}^{N \times N}$ then

$$A \text{ normal } \Leftrightarrow \mathrm{Spec}_\varepsilon(A) = \mathrm{Spec}(A) + D(0, \varepsilon) \, \forall \varepsilon > 0.$$

More generally, if $A = V \Lambda V^{-1}$ is diagonalizable then

$$\mathrm{Spec}_\varepsilon(A) \subset \mathrm{Spec}(A) + D(0, \varepsilon \kappa(V)), \quad \kappa(V) := \frac{s_{\max}(V)}{s_{\min}(V)}.$$

Jordan Block Example (revisited): For any $\delta = \delta_N > 0$ let

$$J_N^{(\delta)} := \begin{bmatrix} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & 0 & 1 & \\ & & & & 0 & 1 \\ \delta & & & & & 0 \end{bmatrix}.$$

Observe: Eigenvalues of $J_N^{(\delta)} = \{\delta^{1/N} e^{2\pi \mathrm{i} k/N}, k \in [0, N-1] \cap \mathbb{Z}\}$.
Therefore

- If $\delta = |z|^N$ for some $z \in D(0,1)$ then an exponentially small perturbation of J_N produces eigenvalues that are at a distance $|z|$ from $\mathrm{Spec}(J_N)$. Thus $\mathrm{Spec}_{r^N}(J_N) \supset D(0,r)$ for any $r \in (0,1)$.
- If $\delta \asymp 1$ or if $\delta = O(N^{-\alpha})$ for any $\alpha > 0$ then eigenvalues of $J_N^{(\delta)}$ approaches $\mathbb{S}^1 := \partial D(0,1)$.

Jordan Block Example (continued):

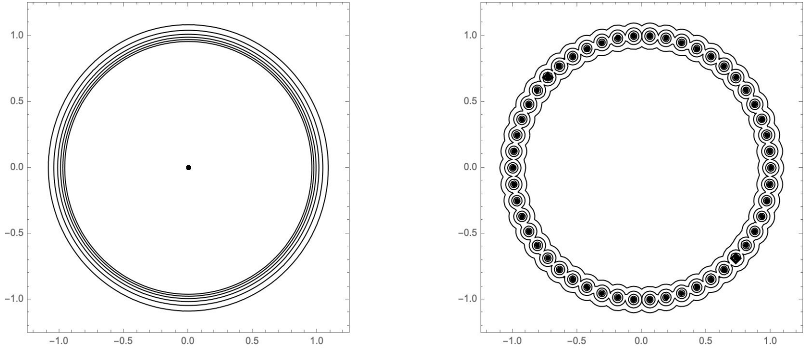


Figure: $N = 50$, $\varepsilon = 10^{-1}, 10^{-1.2}, \ldots, 10^{-2}$. Pseudospectral level lines: J_N on the left panel, $C_N := J_N^{(1)}$ on the right panel.

Move from pseudospectrum to random perturbation

- Pseudospectra are generally harder to characterize and computationally more expensive – random perturbation is an efficient model
- Natural to study spectral features of disordered perturbations (typical perturbation) of a non-normal operators/matrices, compared to the pseudospectrum (worst-case perturbation).
- If the simulated $U_N = \mathcal{U}_N + \Delta_N$, where $\mathcal{U}_N$ is a 'true' unitary and Δ_N captures the machine/rounding error then the spectrum of $\widehat{A}_N := U_N A_N U_N^\star$ is same as that of $A_N + \widehat{\Delta}_N$.
- Non-Hermitian analog of Rosensweig-Porter model.
  - Rosensweig-Porter model: $D_N + \sqrt{t} W_N$.

[Benigni '20], [Landon, Yau '17], [Lee, Schnelli '13]
[Lee, Schnelli, Stelter, Yau '16], [von Soosten, Warzel '18]

Banded Toeplitz matrix with a given symbol:

For $\boldsymbol{a}(\xi) := \sum_{i=-d_-}^{d_+} a_i \xi^i$, with $\xi \in \mathbb{S}^1$, set

$$T_N(\boldsymbol{a}) := \sum_{i \geqslant 0} a_i J_N^i + \sum_{i<0} a_i (J_N^\star)^i.$$

For $\boldsymbol{a}(\xi) = 2\xi^{-3} - \xi^{-2} + 2\iota\xi^{-1} - 4\xi - 2\iota\xi^2$

$$T_N(\boldsymbol{a}) := \begin{bmatrix} 0 & -4 & -2\iota & & & & & \\ 2\iota & 0 & -4 & -2\iota & & & & \\ -1 & 2\iota & 0 & -4 & -2\iota & & & \\ 2 & -1 & 2\iota & 0 & -4 & -2\iota & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & & & & 0 & -4 & -2\iota \\ & & & 2 & -1 & 2\iota & 0 & -4 \\ & & & & 2 & -1 & 2\iota & 0 \end{bmatrix}.$$

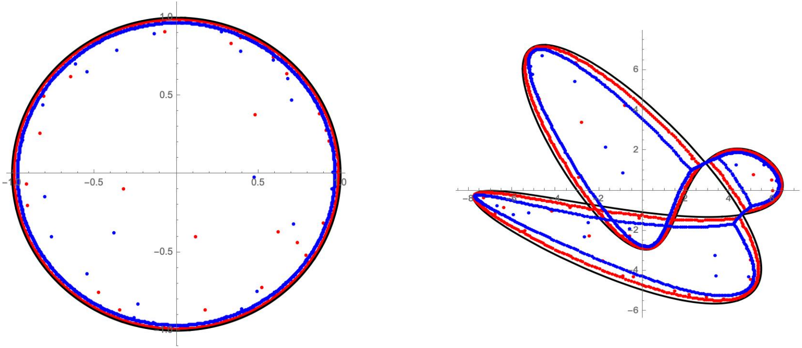


Figure: $N = 1000$. Eigenvalues of $U_N A_N U_N^\star$, U_N a simulated Haar unitary, computed through Mathematica are in blue. Eigenvalues of $A_N + N^{-2} G_N$ are in red, where G_N is the random matrix with i.i.d. standard complex Gaussian entries. Left panel: $A_N = J_N$, and right panel: $A_N = T_N(\boldsymbol{a})$. Symbol curves $\mathbb{S}^1$ (left panel) and $\boldsymbol{a}(\mathbb{S}^1)$ (right panel) in black.

Limit of the bulk of the spectrum

Theorem (B., Paquette, Zeitouni '19, '20)

For any $\gamma > \frac{1}{2}$, if E_N satisfies Assumption (A) then the empirical distribution of the eigenvalues of $T_N(\boldsymbol{a}) + N^{-\gamma} E_N$ converges weakly, in probability, to the law of $\boldsymbol{a}(U)$ where $U \sim \mathrm{Unif}(\mathbb{S}^1)$.

[Hager, Davies '09], [Guionnet, Wood, Zeitouni '14]
[B., Paquette, Zeitouni '19, '20], [Sjöstrand, Vogel '21a, '21b]
[O'Rourke, Wood '22]

Limit of the bulk of the spectrum

Theorem (B., Paquette, Zeitouni '19, '20)

For any $\gamma > \frac{1}{2}$, if E_N satisfies Assumption (A) then the empirical distribution of the eigenvalues of $T_N(\boldsymbol{a}) + N^{-\gamma}E_N$ converges weakly, in probability, to the law of $\boldsymbol{a}(U)$ where $U \sim \mathrm{Unif}(\mathbb{S}^1)$.

Examples.

$$T_N = J_N,\ \boldsymbol{a}(\xi) = \xi.\ L_N \Rightarrow \text{law of } U, \text{ where } U \sim \mathrm{Unif}(\mathbb{S}^1).$$

$$T_N = J_N + J_N^2,\ \boldsymbol{a}(\xi) = \xi + \xi^2.\ L_N \Rightarrow \text{law of } U + U^2.$$

Limit of the bulk of the spectrum

Theorem (B., Paquette, Zeitouni '19, '20)

For any $\gamma > \frac{1}{2}$, if E_N satisfies Assumption (A) then the empirical distribution of the eigenvalues of $T_N(\boldsymbol{a}) + N^{-\gamma} E_N$ converges weakly, in probability, to the law of $\boldsymbol{a}(U)$ where $U \sim \mathrm{Unif}(\mathbb{S}^1)$.

Assumption (A)

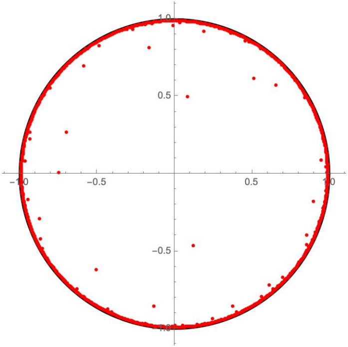
(1)

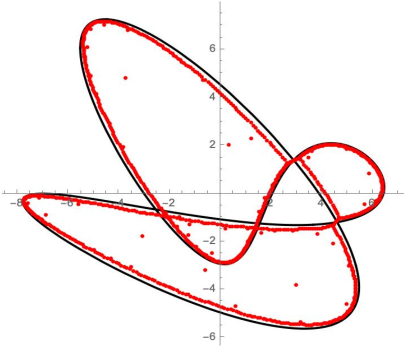
$$\mathbb{E}\left[\|E_N\|_{\mathrm{HS}}^2\right] = \mathbb{E}\left[\sum_{i,j} |e_{i,j}|^2\right] = O(N^2).$$

(2) For every $\alpha > 0 \; \exists \, \beta \in (0, \infty)$, such that for any M_N with $\|M_N\| = O(N^\alpha)$,

$$\mathbb{P}\left(s_{\min}(M_N + E_N) \leqslant N^{-\beta}\right) = o(1).$$

Regions of no outliers





Theorem (B., Zeitouni '20)

The entries of E_N are independent entries with zero mean and unit variance. Then for any $\gamma > \frac{1}{2}$, with probability $\to 1$, there are no outliers in any open set

$$U \subsetneq \mathcal{R}_0 := \{z \in \mathbb{C}\backslash \boldsymbol{a}(\mathbb{S}^1) : \text{wind}_{\boldsymbol{a}}(z) = 0\}.$$

Limit of outliers

Theorem (B., Zeitouni '20)

Additionally assume that E_N be a random matrix with i.i.d. entries having zero mean and unit variance and satisfying some anti-concentration bound (e.g. bounded density). Then for any $\gamma > \frac{1}{2}$, the point processes induced by the outlier eigenvalues converge to the zero set of some non-universal (w.r.t. the distribution of the entries of E_N) random analytic function.

Definition of the limiting random analytic function involves skew semistandard Young Tableaux

Limit of outliers: Jordan block

$T_N = J_N$, entries of E_N are standard complex Gaussian

Limiting random analytic function is a hyperbolic Gaussian analytic function:

$$F(z) = \sum_{\ell=0}^{\infty} g_\ell z^\ell \sqrt{\ell + 1}$$

$\{g_\ell\}$ i.i.d. standard complex Gaussian

Localization/delocalization of eigenvectors

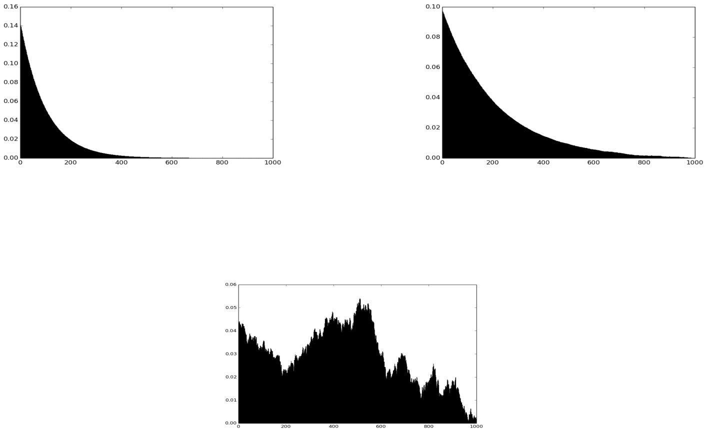


Figure: Moduli of the entries of an eigenvector of $J_N + N^{-\gamma} E_N$: $N = 1000$; top left: $\gamma = 2$, top right: $\gamma = 1.5$, bottom: $\gamma = 1$.

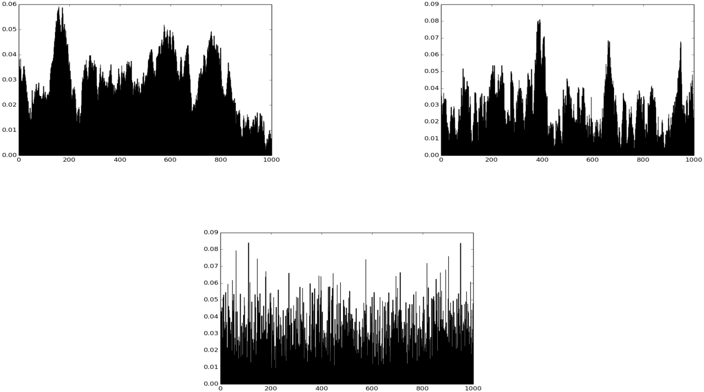


Figure: Moduli of the entries of an eigenvector of $J_N + N^{-\gamma} E_N$: $N = 1000$; `top left:` $\gamma = 0.9$, `top right:` $\gamma = 0.75$, `bottom:` $\gamma = 0.4$.

Localization of eigenvectors for $\gamma > 1$

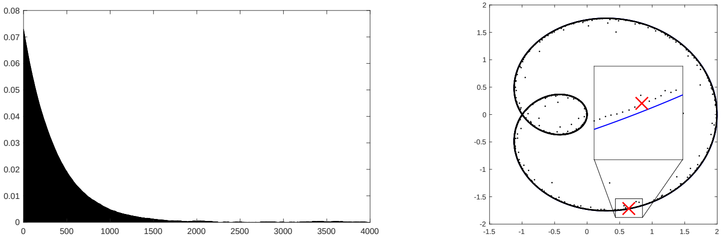


Figure: Eigenvectors (left panel) and eigenvalues (right panel) of $J_N + J_N^2 + N^{-\gamma} E_N$ for $N = 4000$, $\gamma = 1.2$. Plotted are the moduli of the entries of the eigenvector that corresponds to the eigenvalue marked with a red $\times$.

Localization of eigenvectors for $\gamma > 1$

Theorem (B., Vogel, Zeitouni '23)

For most (right)-eigenvectors v, with probability $\to 1$, as $N \to \infty$, (under some assumptions on E_N) the followings hold:

- *Localization at scale $N/\log N$: For any $\ell \in [1, N] \cap \mathbb{Z}$*

$$\|v\|_{\ell^2([1,N-\ell])} \wedge \|v\|_{\ell^2([\ell,N])} \lesssim \exp(-c\ell \log N/N) + N^{-c'}$$

- *Eigenvectors spread out at scale $N/\log N$:*

$$|\mathrm{Supp}(v)| \gtrsim N/\log N$$

$$|\mathrm{Supp}(v)| := \min\{|I| : \|v\|_{\ell^2(I)} \gtrsim 1\}$$

- Long-range correlation: u_z is the right eigenvector corresponding to $z \in \mathbb{C}$.

$$\sum_i u_z(i)u_z(i+L) \to \begin{cases} 1 & \text{if } L \ll N^{2\gamma-1}, \\ 0 & \text{if } L \gg N^{2\gamma-1}. \end{cases}$$

- u_z is approximated by $N^{2(1-\gamma)}$ many $\sin$ and $\cos$ curves and $O(1)$ localized vectors.

Work in progress with Vogel and Zeitouni.

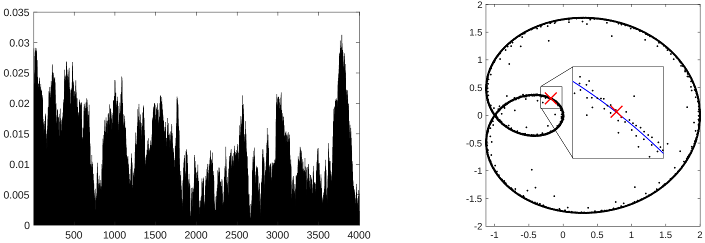


Figure: $T_N = J_N + J_N^2$, $N = 4000$, $\gamma = 0.8$.

Proof ideas: localization of eigenvectors - the Jordan block

- $z \in \mathbb{C}$ an eigenvalue of $J_N^\gamma := J_N + N^{-\gamma} E_N$ with its right-eigenvector u_z.
- $\{e_i(z)\}_{i=1}^N$ right singular vectors of $J_N - z$.

$$u_z = \sum_{i=1}^{m} \langle u_z, e_i(z) \rangle e_i(z) + \sum_{i=m+1}^{N} \langle u_z, e_i(z) \rangle e_i(z).$$

- **Strategy**: For $\gamma > 1$ for a typical eigenvalue z, the projection of u_z on $\mathrm{Span}(e_i(z), i \geqslant 2)$ is negligible, and $e_1(z)$ is localized.

Proof ideas: localization of eigenvectors - the Jordan block

- Schur complement formula/Grushin problem:

$$\Big\| \sum_{i=m+1}^{N} \langle u_z, e_i(z) \rangle e_i(z) \Big\| \lesssim N^{-(\gamma-1/2)} \|E(z)(I+N^{-\gamma}E_N E(z))^{-1}\|$$

$m = 1$ for J_N

- $\|E(z)\| = 1/$ second smallest singular value of $J_N - z$.
- Observe for $z \in D(0,1)$

$$\|(J_N - z)v_z\| = |z|^N, \qquad \|v_z\| \asymp (1-|z|)^{-1/2},$$

$$v_z^{\mathsf{T}} := \begin{pmatrix} 1 & z & z^2 & \cdots & z^{N-1} \end{pmatrix}.$$

- For $z \in \mathbb{C}$ with $\operatorname{dist}(z, \mathbb{S}^1) \gtrsim \log N/N$

$$\|v_z/\|v_z\| - e_1(z)\| = o(1).$$

- Most eigenvalues reside in the domain

$$\{\lambda \in \mathbb{C} : \operatorname{dist}(\lambda, \mathbb{S}^1) \in ((\gamma-1)\log N/N, C\log N/N]\}.$$

Proof ideas: localization of eigenvectors - the general case

For z an eigenvalue with v_z being its eigenvector

- $\exists v_z \in \mathrm{Span}(e_j(z), j, = 1, 2, \ldots, |d(z)|)$ such that $\|v_z - u_z\| = o(1)$. $d(z) = \mathrm{wind}_{\boldsymbol{a}}(z)$
- Most eigenvalues reside in the domain

$$\{\lambda \in \mathbb{C} : \mathrm{dist}(\lambda, \boldsymbol{a}(\mathbb{S}^1)) \in (c_{\gamma-1} \log N/N, C \log N/N]\}.$$

- Vectors approximating $e_j(z)$'s decay exponentially at a speed $r \asymp 1$ for $j < |d(z)|$ and the last one decays at speed $r_N \asymp \log N/N$.

  (gives the upper bound)
- Show that v_z have a non-negligible projection onto the subspace spanned by the vector approximating $e_{|d(z)|}(z)$.

  (gives the lower bound)

Thank you!