

Spectral Statistics of Sparse Random Graphs

Jiaoyang Huang (University of Pennsylvania)

Based on joint works with Roland Bauerschmidt, Paul Bourgade, Antti Knowles,
Benjamin Landon, Theo McKenzie and Horng-Tzer Yau

ICM 2026, Probability Section, Philadelphia

- 1 Random Matrices and Local Statistics
- 2 Spectral statistics of Sparse Random Graphs
- 3 Loop Equations Characterize Random Matrix Statistics

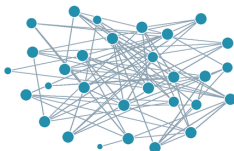
Two Sparse Graph Ensembles

Erdős–Rényi graph $G(N, p_N)$

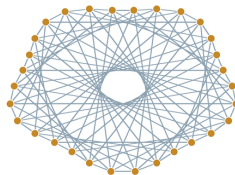
Random graphs with N vertices, where each edge is included independently with probability p_N .

Random d -regular graph $G_{N,d}$

Uniform probability measure on set of simple d -regular graphs with N vertices.



Erdős–Rényi graph



random d -regular graph

For a random graph $\mathcal{G}$ on N vertices, denote its adjacency matrix by

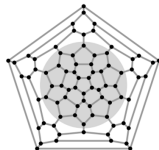
$$A_{ij} = \mathbf{1}_{\{i \sim j\}}.$$

Why Care About Spectral Statistics?

- **Networks and inference:** Outlier eigenvalues and eigenvectors reveal low-rank signals, planted subgraphs, and community effects.
- **Expansion:** A family of d -regular graphs is an **expander** family if its spectral gap is uniformly bounded away from zero. The spectral gap controls expansion and mixing.
- **Quantum chaos and geometry:** Random d -regular graphs provide discrete models for studying the spectral statistics of chaotic systems and random hyperbolic surfaces [Monk–Naud's talk].



WHAT IS . . .



an Expander?

Peter Sarnak

Random Matrices

- **Wigner matrices** were introduced by the physicist Eugene Wigner in the 1950s to model the energy levels of heavy atomic nuclei.

Gaussian Orthogonal Ensemble (GOE)

$$H = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1N} \\ h_{21} & h_{22} & \dots & h_{2N} \\ \vdots & \vdots & & \vdots \\ h_{N1} & h_{N2} & \dots & h_{NN} \end{pmatrix}$$

where $h_{jk} = h_{kj}$ (for $j \leq k$) are real independent **normal** random variables

$$\mathbb{E}[h_{jk}] = 0, \quad \mathbb{E}[|h_{jk}|^2] = \frac{1 + \delta_{jk}}{N}.$$

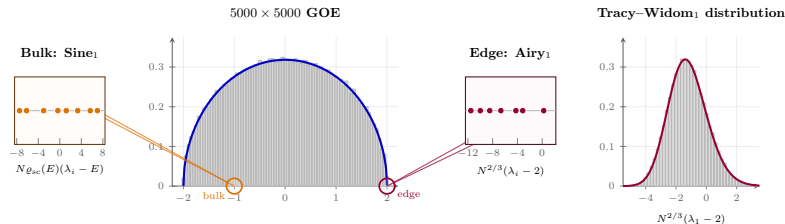


Figure: Bulk and edge eigenvalue statistics

*We picture a complex nucleus as a
“black box” in which a large num-
ber of particles are interacting ac-
cording to unknown laws.*

—Freeman Dyson (1962)



Universality for Wigner Matrices

Universality

After rescaling by the local eigenvalue spacing, the limiting statistics depend only on the symmetry class, not on the entry distribution:

$$\beta = 1 \text{ (real symmetric)}, \quad \beta = 2 \text{ (complex Hermitian)}.$$

- **Bulk.** For every $E \in (-2, 2)$,

$$\sum_{i=1}^N \delta_{N \varrho_{\text{sc}}(E)(\lambda_i - E)} \implies \text{Sine}_{\beta}.$$

Erdős–Schlein–Yau (2009–11); Tao–Vu (2009); Erdős–Yau–Yin (2012);
Bourgade–Erdős–Yau–Yin (2016), Aggarwal (2016), Aggarwal–Lopatto–Yau (2018),...

- **Edge.** The rescaled edge spectrum satisfies

$$\sum_{i=1}^N \delta_{N^{2/3}(\lambda_i - 2)} \implies \text{Airy}_{\beta}.$$

In particular,

$$N^{2/3}(\lambda_1 - 2) \implies \text{TW}_{\beta}.$$

Tracy–Widom (1994, 1996); Soshnikov (1999); Tao–Vu (2010); Erdős–Yau–Yin (2012);
Knowles–Yin (2013); Lee–Yin (2014).

- **Challenges** for sparse random graphs: **Sparsity** and degree constraints force **long-range correlations**.

- 1 Random Matrices and Local Statistics
- 2 Spectral statistics of Sparse Random Graphs
- 3 Loop Equations Characterize Random Matrix Statistics

Random d -regular Graph

- Normalize the adjacency matrix: $H = A/\sqrt{d-1}$.
- The eigenvalues of H , $\lambda_1 = d/\sqrt{d-1} \geq \lambda_2 \geq \dots \geq \lambda_N$.
- The empirical eigenvalue distributions of random d -regular graphs converge to the **Kesten-McKay** distribution:

$$\frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i} \rightarrow \varrho_d(x) = \left(\frac{d}{d-1} - \frac{x^2}{d} \right)^{-1} \frac{\sqrt{4-x^2}}{2\pi}.$$

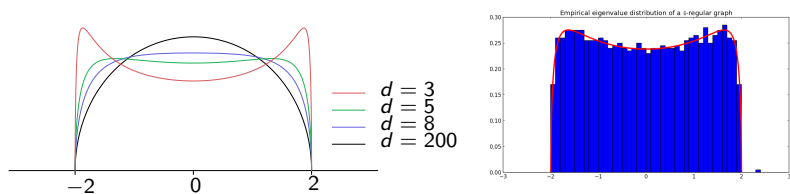


Figure: The Kesten-McKay distribution and the empirical spectral distribution of a random 4-regular graph on 1000 vertices.

Theorem (Alon-Boppana (1986))

For every connected d -regular graph $\mathcal{G}$, there exists a constant C_d ,

$$\lambda_2 \geq 2 - \frac{C_d}{\text{diam}(\mathcal{G})}.$$

- The diameter of a d -regular graph on N vertices is at least $O(\log N)$. For a sequence of d -regular graphs on N vertices, $\liminf_{N \rightarrow \infty} \lambda_2 \geq 2$.

Ramanujan Graph

Theorem (Alon-Boppana (1986))

For every connected d -regular graph $\mathcal{G}$, there exists a constant C_d ,

$$\lambda_2 \geq 2 - \frac{C_d}{\text{diam}(\mathcal{G})}.$$

- The diameter of a d -regular graph on N vertices is at least $O(\log N)$. For a sequence of d -regular graphs on N vertices, $\liminf_{N \rightarrow \infty} \lambda_2 \geq 2$.

Ramanujan graph (Lubotzky-Phillips-Sarnak (1988))

A d -regular graph on N vertices is called a Ramanujan graph if

$$\max_{|\lambda_i| < d/\sqrt{d-1}} |\lambda_i| \leq 2.$$

- Ramanujan graphs are the **best possible expanders**.
- Let p be a prime with $p \equiv 1 \pmod{4}$. Lubotzky, Phillips and Sarnak (1988) gave explicit constructions of Ramanujan graphs with $d = p + 1$. Also independently by Margulis (1988).
- Marcus, Spielman and Srivastava (2013) proved the existence of infinite families of **bipartite** d -regular Ramanujan graphs, using the method of interlacing polynomials.
- From numerical simulation, it is observed that the distribution of λ_2, λ_N is well-modeled by the **Tracy-Widom₁ distribution**, the asymptotic distribution of the largest eigenvalue of GOE.

Conjecture (Sarnak (2004), Miller-Novikoff-Sabelli (2006))

Let $\mathcal{G}$ be a random d -regular graph on N vertices, then

$$C_1^d N^{2/3} (\lambda_2 - 2) - C_2^d \rightarrow TW_1.$$

The analogous statement holds for the smallest eigenvalue.

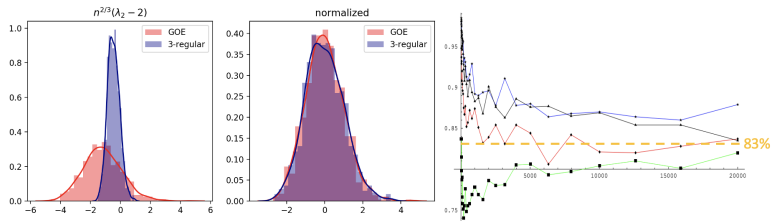


Figure: The second largest eigenvalue of random 3-regular graphs on 3000 vertices, and GOE; Percentage that $\lambda_2 \leq 2$ for random 3-regular graphs (from various models).

Alon's conjecture

It was first a conjecture by Alon (1986) that random d -regular graphs have nearly the largest possible spectral gaps, $\lambda_2 = 2 + o_N(1)$.

Theorem (Friedman (2004), Bordenave (2015))

Let $\mathcal{G}$ be a random d -regular graphs on N vertices, then with probability $1 - o_N(1)$,

$$\lambda_2 = 2 + o_N(1).$$

- In the proof of Friedman (2004) and Bordenave (2015), the error term is polylog small, i.e. $(\log \log N / \log N)^2$.
- Improved to $\lambda_2 = 2 + O(N^{-c})$ for small $c > 0$ by H.-Yau (2022).
- Chen–Garza–Vargas–Tropp–Van Handel (2024), Chen–Garza–Vargas–van Handel (2024): new approach to strong convergence [van Handel's Talk], for random permutation models of d -regular graphs with even $d \geq 4$, with probability $1 - o_N(1)$,

$$\lambda_2 \leq 2 + \left(\frac{\log N}{N} \right)^{1/6}.$$

Theorem (H.-McKenzie-Yau (2024))

Fix $d \geq 3$. Define the classical location γ_i of the i -th eigenvalue by

$$\int_{\gamma_i}^2 \varrho_d(x) dx = \frac{i}{N}, \quad 2 \leq i \leq N.$$

Then with probability $1 - O(N^{-c})$ eigenvalue rigidity holds:

$$|\lambda_i - \gamma_i| \leq N^{-2/3+o(1)} (\min\{i, N-i+1\})^{-1/3}, \quad 2 \leq i \leq N.$$

Theorem (H.-McKenzie-Yau (2024))

Fix $d \geq 3$. Let $\mathcal{G}$ be a random d -regular graph on N vertices, then

$$(\mathcal{A}N)^{2/3}(\lambda_2 - 2) \rightarrow TW_1, \quad \mathcal{A} = \frac{d(d-1)}{(d-2)^2}.$$

The analogous statement holds for the smallest eigenvalue.

- For any $d \geq 3$, about $(83\%)^2 \approx 69\%$ d -regular graphs on N vertices have all nontrivial eigenvalues bounded in absolute value by 2, i.e. they are Ramanujan graphs!

- Edge universality for growing d :

$$N^\varepsilon \leq d \leq N^{1/3-\varepsilon} \quad \text{or} \quad N^{2/3+\varepsilon} \leq d \leq N/2.$$

Bauerschmidt–H.–Knowles–Yau (2019); H.–Yau (2023); He (2022).

- Bulk universality for growing d :

$$(\log N)^{24} \ll d \leq N^{2/3-\varepsilon}.$$

Bauerschmidt–H.–Knowles–Yau (2017); Landon–Sosoe–Yau (2019); Bourgade–H. (2026).

Fixed-degree bulk conjecture (Jakobson–Miller–Rivin–Rudnick (1999))

For every fixed $d \geq 3$, the bulk local statistics converge to the Sine_1 process.

Let $D := Np_N$, and consider the normalized adjacency matrix.

- **Bulk universality.** For $D \geq N^\varepsilon$, the local bulk statistics converge to Sine_1 .
Erdős–Knowles–Yau–Yin (2013); H.–Landon–Yau (2015); Landon–Sosoe–Yau (2019).
- **Edge fluctuations.** For the largest nontrivial eigenvalue, after the appropriate scaling,

$$\begin{aligned} D \gg N^{1/3} &\implies \text{TW}_1, \\ D \asymp N^{1/3} &\implies \text{TW}_1 + \text{Gaussian}, \\ N^\varepsilon \leq D \ll N^{1/3} &\implies \text{Gaussian}. \end{aligned}$$

After an explicit random recentering L , for all $D \geq N^\varepsilon$,

$$\sum_i \delta_{N^{2/3}(\lambda_i - L)} \implies \text{Airy}_1.$$

Erdős–Knowles–Yau–Yin (2013); Lee–Schnelli (2018); H.–Landon–Yau (2020); He–Knowles (2021); Lee (2022); H.–Yau (2022).

Erdős–Rényi Graphs $G(N, p_N)$: Very Sparse Regime

Let $D := Np_N$, and consider the normalized adjacency matrix.

Poisson statistics and localization at the edge

For

$$(\log N)^{-1/15} \leq D \leq (b_* - o(1)) \log N, \quad b_* = \frac{1}{\log 4 - 1}$$

the edge eigenvalues have Poisson statistics, and the corresponding eigenvectors are exponentially localized. Alt–Ducatez–Knowles (2019–2023); Hiesmayr–McKenzie (2023).

Fixed-average-degree bulk conjecture

For sufficiently large fixed D , there exists a nonempty interval I such that, after restricting $G(N, p_N)$ to its giant component and removing compactly supported eigenmodes, the rescaled local eigenvalue statistics in I converge to Sine_1 .

Fyodorov–Mirlin (1991), supersymmetry; Evangelou–Economou (1992), Klefogiannis–Amanatidis (2024), numerics.

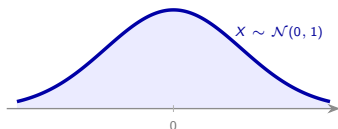
- 1 Random Matrices and Local Statistics
- 2 Spectral statistics of Sparse Random Graphs
- 3 Loop Equations Characterize Random Matrix Statistics**

- **Integrable invariant ensembles:** determinantal/Pfaffian formulas, followed by orthogonal-polynomial or Riemann–Hilbert asymptotics. Mehta–Gaudin (1960); Dyson (1962); Deift–Kriecherbauer–McLaughlin–Venakides–Zhou (1999); Deift–Gioev (2007).
- **Moment comparison:** Lindeberg replacement for Wigner matrices whose entries match moments through order four. Tao–Vu (2010, 2011).
- **Change of variables / approximate transport:** compare a general β -log gas directly with a reference ensemble. Shcherbina (2014); Bekerman–Figalli–Guionnet (2015); Figalli–Guionnet (2016).
- **Dynamical three-step approach:** i) local law and rigidity, ii) short-time Dyson Brownian motion, iii) remove the Gaussian perturbation. Erdős–Schlein–Yau (2009, 2011); Erdős–Yau–Yin (2012); Bourgade–Erdős–Yau–Yin (2016),...**[Yau's talk]**

Can one characterize random matrix statistics intrinsically?

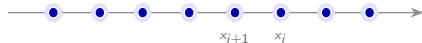
Stein's Identity and Loop Equations

Standard Normal variable



Local eigenvalue statistics

Sine $_{\beta}$ in the bulk, Airy $_{\beta}$ at the edge



Stein's identity

Standard Normal distribution is characterized by the integration by parts identity

$$X \sim \mathcal{N}(0, 1) \iff \mathbb{E}[Xf(X)] = \mathbb{E}[f'(X)] \text{ for a suitable class } f.$$

OBSERVABLES

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$

Suitable test functions

OBSERVABLES

$$\mathbb{E} \left[\prod_{j=1}^p s(z_j) \right]$$

Stieltjes-transform moments

Stein's identity

$$\mathbb{E}[Xf(X)] = \mathbb{E}[f'(X)]$$

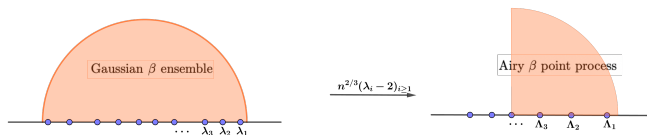
One identity, valid for every suitable f .

Loop equations

$$\mathcal{L}_p(z_1, \dots, z_p) = 0, \quad p \geq 1$$

A hierarchy of equations.

- **Loop equations or Schwinger–Dyson equations**, are integration-by-parts identities for Gibbs measures. They form a hierarchy relating moments of Stieltjes transforms in matrix models. Makeenko–Migdal (1979); Johansson (1998); Eynard (2004); Borot–Guionnet (2013)...
- Rescaled extreme eigenvalues of GOE converge to the **Airy₁ point process** $\Lambda_1 \geq \Lambda_2 \geq \Lambda_3 \geq \dots$, where Λ_1 has Tracy–Widom₁ distribution. Tracy–Widom (1993, 1994), Ramírez–Rider–Virág (2006).



- Denote the normalized Stieltjes transform of the Airy _{β} point process $\Lambda_1 \geq \Lambda_2 \geq \Lambda_3 \geq \dots$: $s(w) \sim \sqrt{w}$ as $w \rightarrow \infty$

$$s(w) = \sum_{i \geq 1} \frac{1}{\Lambda_i - w} - \text{“normalized sum”}, \quad w \in \mathbb{C}^+.$$

The Airy points are precisely the poles of $s(w)$.

Edge Loop Equations

The first loop equation: for any $w \in \mathbb{C} \setminus \mathbb{R}$

$$\mathbb{E} \left[\frac{2-\beta}{\beta} \partial_w s(w) + s(w)^2 - w \right] = 0,$$

Edge loop equations at rank p : for any $w_1, w_2, \dots, w_p \in \mathbb{C} \setminus \mathbb{R}$

$$\begin{aligned} & \mathbb{E} \left[\left(\frac{2-\beta}{\beta} \partial_{w_1} s(w_1) + s(w_1)^2 - w_1 \right) \prod_{j=2}^p s(w_j) \right] \\ & + \frac{2}{\beta} \mathbb{E} \left[\sum_{j=2}^p \partial_{w_j} \frac{s(w_1) - s(w_j)}{w_1 - w_j} \prod_{\substack{\ell=2 \\ \ell \neq j}}^p s(w_\ell) \right] = 0. \end{aligned} \tag{1}$$

Theorem (Bourgade-H. (2026))

Assume that $\beta > 0$ is rational. The full hierarchy of edge loop equations (1) uniquely characterizes Airy_β point process.

- Equivalent to the BBGKY hierarchy at equilibrium, for the **logarithmic interaction**.
- Inside the bulk, analogous **bulk loop equations** characterize Sine_β point process.
- These uniqueness results provide a direct route to universality: it suffices to verify the corresponding approximate loop equations for the ensemble.

Local Law to Loop Equations

- Given a random d -regular graph $\mathcal{G}$. Let H be the normalized adjacency matrix,

$$H = A/\sqrt{d-1},$$

with eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$.

- The **Stieltjes transform** of the empirical eigenvalue distribution and the **Kesten–McKay** distribution ϱ_d :

$$m_N(z) = \frac{1}{N} \sum_{i=1}^N \frac{1}{\lambda_i - z}, \quad m_d(z) = (d-1) \frac{(d-2)z + d\sqrt{z^2 - 4}}{2((d-1)z^2 - d^2)}$$

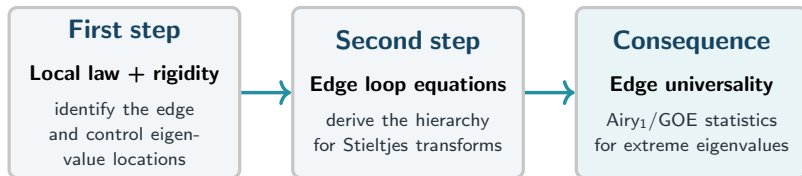
Proposition (H.-McKenzie-Yau (2024))

Let $\Delta(z) = (m_N(z) - m_d(z))$ and $\mathcal{A} = d(d-1)/(d-2)^2$. Then for any $p \geq 1$, and $z_1, z_2, \dots, z_p \in \mathbb{C} \setminus \mathbb{R}$ close to the right edge 2,

$$\begin{aligned} & \mathbb{E} \left[\left(\Delta(z_1)^2 + 2\mathcal{A}\sqrt{z_1 - 2}\Delta(z_1) + \frac{\partial_{z_1} m_N(z_1)}{N} \right) \prod_{j=2}^p \Delta(z_j) \right] \\ & + \frac{2}{N^2} \sum_{2 \leq j \leq p} \mathbb{E} \left[\partial_{z_j} \left(\frac{m_N(z_1) - m_N(z_j)}{z_1 - z_j} \right) \prod_{\substack{\ell=2 \\ \ell \neq j}}^p \Delta(z_\ell) \right] \ll N^{-(p+1)/3}. \end{aligned}$$

- The **scaling limit** $N^{1/3} \Delta(2 + w/(\mathcal{A}N)^{2/3})/\mathcal{A}^{2/3}$ converges to $s(w) - \sqrt{w}$ in the microscopic version of loop equations.

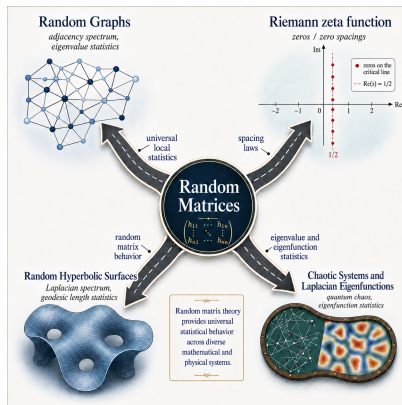
Local Law to Loop Equations



Local law and rigidity makes the microscopic scale accessible, and the edge loop-equation hierarchy identifies the Tracy-Widom statistics.

Summary

- **Random regular graphs.** Edge universality holds for random d -regular graphs with fixed degree. Asymptotically, 69% are Ramanujan.
- **An intrinsic characterization.** Loop equations characterize the random matrix statistics.
- **Beyond random matrices.** Zeta zeros, quantum chaos, and random surfaces—universality is still wide open.



Thank you for listening!