Spectral statistics of random d-regular graphs

Horng-Tzer Yau

Harvard University

with Jiaoyang Huang and Theo McKenzie

Visions in Arithmetic and Beyond: Celebrating Peter Sarnak's Work and Impact

Outline

1. Spectral statistics of random graphs.
   - Random d-regular and Erdős Rényi graphs.
   - Expander and Ramanujan graphs.
2. Eigenvalue rigidity and eigenvector delocalization.
   - Optimal rigidity for random d-regular graphs for $d \geq 3$.
   - Completely delocalization for $d \geq 3$.
   - Tracy-Widom distribution for $d \geq N^{\varepsilon}$.
3. Some ideas of the proof.
   - Local resampling for random d-regular graphs.
   - Kesten-McKay law vs semicircle law.
   - Generalized local resampling.
   - Expansion with Woodbury formula.
   - Iterative generalized local resampling.
4. Edge universality of random d-regular graphs.
   - Noise dominating phenomenon.
   - Comparison of local resampling with matrix Brownian motions.
5. Summary

Outline

Random d-regular graph

Random d-regular graphs $G_{N,d}$:

Uniform distribution on simple d-regular graphs with N vertices

- Denote by $A_{ij} = \mathbf{1}_{\{i \sim j\}}$ the adjacency matrices and the normalized adjacency matrices by $H = (h_{ij}) = A/\sqrt{d-1}$
- d-regular graphs are highly correlated matrices.
- Denote the eigenvalues of H by $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_N$. Notice that $\lambda_1 = d/\sqrt{d-1}$ with the eigenvector $\mathbf{1}$.

Erdős Rényi graphs

- Erdős-Rényi graphs $G(N, d/N)$: each edge is selected with probability $p = d/N$ independently.
- Denote the normalized adjacency matrices with mean subtracted by $H = (h_{ij})$. Then the matrix entries are independent and

$$\mathbb{E}[h_{ij}] = 0, \mathbb{E}[h_{ij}^2] = 1/N.$$

Expander Graph

An expander graph is a sparse graph that has strong connectivity properties. Expander graphs have many applications in mathematics and computer science. We will focus on spectral expander and in particular on the second largest eigenvalue.

Theorem (Alon-Boppana)

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

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

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

Random d-regular graph

Definition (Lubotzky-Phillips-Sarnak)

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

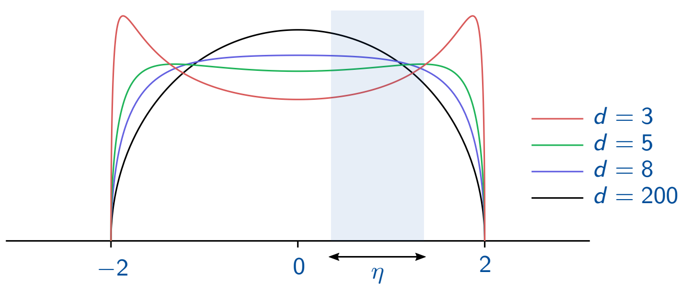
$$\max_{|\lambda_i| < d/\sqrt{d-1}} |\lambda_i| \leqslant 2.$$

- Ramanujan graphs are the best possible spectral expanders.
- Let p be a prime with $p \equiv 1 (\mathit{mod}\ 4)$. Lubotzky, Phillips and Sarnak (1988) gave explicit constructions of Ramanujan graphs with $d = p + 1$.
- Marcus, Spielman and Srivastava (2013) proved the existence of infinite families of bipartite Ramanujan graphs, using the method of interlacing polynomials.

Random d-regular graphs

The empirical eigenvalue distribution of the random d-regular graphs converges to the **Kesten-McKay** distribution:

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



- Kesten[1957]: d-regular tree. McKay[1981]: d-regular graphs.
- $\lambda_1 = d/\sqrt{d-1}$ was not shown on the KM law.
- Kesten-McKay law becomes the semicircle law (with density ρ_{sc}) as $d \to \infty$, i.e., $\rho_d \to \rho_{sc}$ as $d \to \infty$.

Random d-regular graphs

From numerical simulation, it was observed that the distribution of λ_2, λ_N is well-modeled by TW, the Tracy-Widom $\beta = 1$ distribution (possibly up to a constant shift).

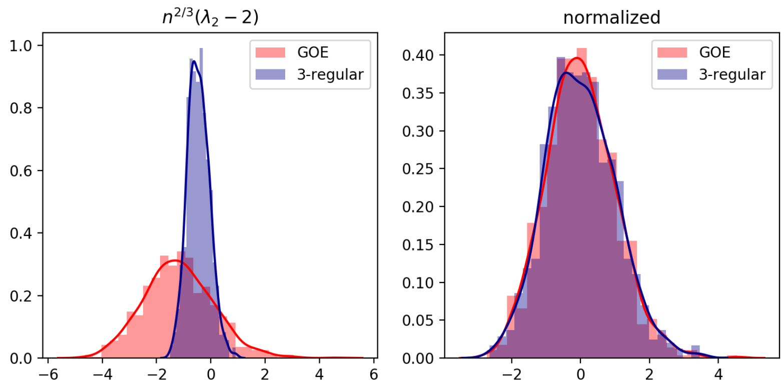
TW can be characterized as the asymptotic distribution of the largest eigenvalue of the Gaussian orthogonal ensemble (symmetric matrices with Gaussian distributions):

Conjecture (Miller-Novikoff-Sabelli (2006))

The second largest eigenvalue of random d-regular graph with $d \geq 3$ satisfies

$$C_1 N^{2/3}(\lambda_2 - 2) - C_d \to TW.$$

Also asked in Sarnak (2004).



Extremal eigenvalues of random 3-regular graphs on 3000 vertices, and GOE.

Outline

Known results for edge eigenvalues for fixed d

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

Fix $d \geqslant 3$. For random d-regular graphs on N vertices, we have with probability $1 - o_N(1)$,

$$\lambda_2 = 2 + O(\log\log N / \log N)^2)$$

There are many results on random surfaces using ideas parallel to random graphs.

Main Results: Optimal rigidity estimate for fixed d

Theorem (Huang-McKenzie-Y, 2024 (Optimal rigidity estimate))

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

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

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

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

In particular, with probability $1 - N^{-(1-o_N(1))}$,

$$|\lambda_2 - 2| \leqslant N^{-2/3+o(1)}$$

Chen, Garza-Vargas, Tropp and Van Handel (preprint last week). For random permutation models of d-regular graphs with even $d \geq 4$,

$$\mathbb{P}(\lambda_2 \geq 2+\varepsilon) \lesssim \frac{1}{N}\left(\frac{d}{\varepsilon}\right)^8 \text{poly}(\log d, \log \varepsilon)$$

Eigenvalue universality for $d \geq N^{\varepsilon}$

Theorem (Huang-Y(2023))

Let $\mathcal{G}$ be a random d-regular graph on N vertices with $N^{\varepsilon} \leqslant d \ll N^{1/3}$, then

$$N^{2/3}(\lambda_2 - 2) \to TW$$

The analogous statement holds for all edge eigenvalues, and in particular, the smallest eigenvalue.

- The result states that in the range $N^{\varepsilon} \leqslant d \ll N^{1/3}$, about 69% d-regular graphs on N vertices are Ramanujan graphs, i.e., all nontrivial eigenvalues are bounded in absolute value by 2.
- Similar results for $N^{2/9} \ll d \ll N^{1/3}$ was proved by Bauerschmidt-Huang-Knowles-Y (2019) and for $d \gg N^{2/3}$ by He (2022).
- Eigenvalue distributions in the bulk are universal (i.e., bulk universality) and were proved in Bauerschmidt-Huang-Knowles-Y (2016). Bulk universality turns out to be easier than edge universality.
- N^{ε} can be replaced by poly$(\log N)$.

Theorem (Huang-Y 2023 (Erdos, Knowles, Yin, Lee, Schnelli, Landon))

Suppose that there is $\varepsilon > 0$ such that $d \geq N^{\varepsilon}$ (it can be relaxed to $d \geq (\log N)^C$). Then there is a scalar $L = 2 + O(1/d)$ and a random variable

$$\mathcal{X} = \tilde{\mathcal{X}} + \text{lower order terms}, \quad \tilde{\mathcal{X}} = \frac{1}{N}\sum_{ij}\left(h_{ij}^2 - 1/N\right) \sim \frac{\text{gaussian}}{\sqrt{Nd}}$$

such that

$$\{N^{2/3}(\lambda_i - L - \mathcal{X})\}_{1 \leqslant i \leqslant k} \to \text{TW}.$$

In particular,

$$\text{Var}(N^{2/3}(\lambda_i - L)) \sim \frac{N^{4/3}}{Nd} + O(1) = \frac{N^{1/3}}{d} + O(1) \gg 1 \quad \text{if } d \ll N^{1/3}$$

Hence Gaussian fluctuations dominate TW for $d \ll N^{1/3}$ and TW dominates for $d \gg N^{1/3}$.

Previous results on edge statistics for Erdős-Rényi graphs

- $d \geq N^{2/3}$; Erdos-Knowles-Y-Yin [2011] proved that the leading behavior of extremal eigenvalues are Tracy-Widom.
- $d \geq N^{1/3}$; Lee-Schnelli [2016] proved that the leading behavior of extremal eigenvalues are Tracy-Widom.
- $N^{2/9} \ll d \ll N^{1/3}$; Huang-Landon-Y [2017]

$$\tilde{\mathcal{X}} + N^{-2/3}\xi; \qquad \tilde{\mathcal{X}} \sim \frac{\text{gaussian}}{\sqrt{Nd}}, \quad \xi \sim \text{Tracy-Widom law}$$

- $N^{\varepsilon} \leq d \leq N^{2/9}$; He-Knowles [2021] proved that the leading behavior of extremal eigenvalues are Gaussian.
- $(\log \log N)^4 \ll d < b_* \log N$ Alt-Ducatez-Knowles [2021] and $(\log N)^{-c} < d < (\log \log N)^4$ Hiesmayr-McKenzie [2023] proved that the leading behavior are Poisson.
- $N^{\varepsilon} \leq d \leq N^{2/9}$; Jaehun Lee [2021] proved that after subtracting a random term χ, the rigidity of extremal eigenvalues holds.

Green's function (resolvent) estimates

Stieltjes transform of a measure ϱ:

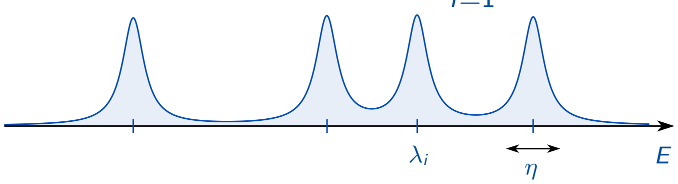
$$m_\varrho(z) = \int \frac{\varrho(x)dx}{x - z},$$
$$\varrho(E) = \frac{1}{\pi} \lim_{\eta \to 0+} \Im[m_\rho(E + i\eta)].$$

- The Stieltjes transform m_{sc} of the semicircle distribution satisfies a quadratic equation
$$m_{sc}(z)^2 + z m_{sc}(z) + 1 = 0.$$
- The Stieltjes transform m_{KM} of the Kesten–McKay law satisfies
$$m_{KM}(z) = \frac{1}{-z - \frac{d}{d-1} m_{sc}(z)}.$$
- The self-consistent equation for m_{KM} is quite a bit more complicated.

Green's function estimates

- Define the Green's function $G(z) = \frac{1}{H-z}$ and $z = E + i\eta$.
- η specifies the spectral resolution.

$$\operatorname{Im} m_N := \operatorname{Im} \frac{1}{N} \operatorname{Tr} G(z) \;=\; \frac{1}{N} \sum_{i=1}^{N} \frac{\eta}{(\lambda_i - E)^2 + \eta^2}.$$



- $\eta \gtrsim N^{-1}$ is the optimal scale; below this scale, m_N is random.
- Our goal is to prove that $m_N(z) \to m_{KM}(z)$ up to the optimal scale. To this end, we need to show that m_N satisfies the self-consistent equation of m_{KM}.
- $m_N(z) \to m_{KM}(z)$ for $\eta \sim O(1)$ is called the global law; for $\eta \gtrsim N^{-1}$ is called the local law.
- Local law with an optimal error bound implies eigenvalue rigidity (Erods-Y-Yin).

Theorem (Huang-McKenzie-Y (2024))

Fix $d \geqslant 3$. Then with probability $1 - N^{-(1-o_N(1))}$ we have

$$|\frac{1}{N}\operatorname{Tr} G(z) - m_{KM}(z)| \leqslant N^{o(1)} \begin{cases} \frac{1}{N\eta}, & -2 \leqslant E \leqslant 2, \\ \frac{1}{\sqrt{\kappa+\eta}}\left(\frac{1}{N\eta^{1/2}} + \frac{1}{(N\eta)^2}\right), & |E| \geqslant 2, \end{cases}$$

for $\eta \geq \frac{N^{-1+\varepsilon}}{\sqrt{\kappa+\eta}}$ where $\kappa = \min\{|E-2|, |E+2|\}$.

Estimates on matrix elements $G_{ij}(z)$ with non-optimal error bounds were previously proven by Huang-Y (2023).
Denote the eigenvector by u_j. Then

$$\operatorname{Im} G_{aa}(z) = \sum_j \frac{\eta}{(\lambda_j - E)^2 + \eta^2} |u_j(a)|^2$$

$$\operatorname{Im} G_{aa}(z) \leq C \Longrightarrow |u_j(a)|^2 \leq \eta, \quad \text{with the optimal } \eta \gtrsim N^{-1}.$$

Definition: An eigenvector is completely delocalized if $\|\mathbf{u}_i\|_\infty^2 = O((\log N)^C/N)$.

Global and local laws ($\eta \ll 1$) for d regular graphs:

	d	scale η
Dumitriu-Pal [2009]	$(\log N)^C$	$(\log N)^{-1}$
Tran-Vu-Wang [2010]	$\gg 1$	$d^{-1/10}$
Anatharaman-Le Masson* [2013]	fixed	$(\log N)^{-c}$
Geisinger [2014]	fixed	$(\log N)^{-1}$
Brooks–Lindenstrauss [2013]*; eigenvector	fixed	$(\log N)^{-c}$
Huang-Y** (2015-2023); complete delocalization	$d \geq 3$	N^{-1}

* Proof does not use randomness, but uses some local tree-like conditions.

** Earlier papers also with Knowles and Bauerschmidt.

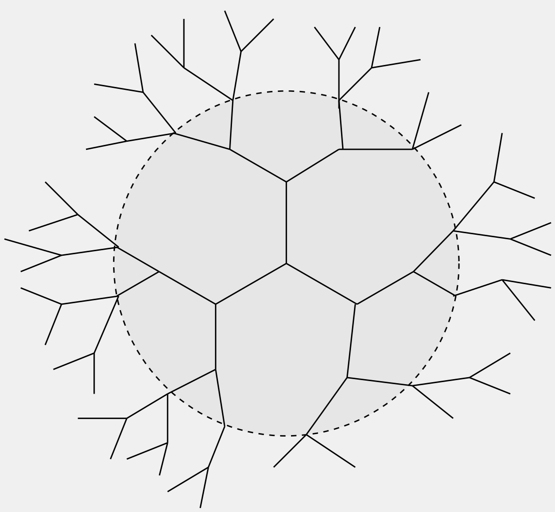
- Most random matrix methods were invented for Wigner matrices and are more suitable when there are lots of edges. Sparsity makes it harder to apply random matrix methods.

Outline

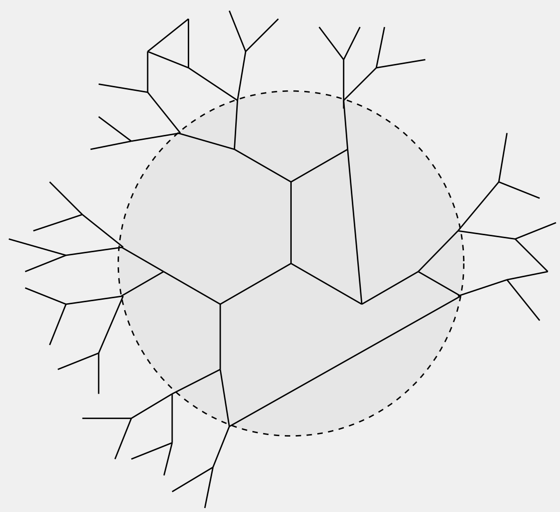
Some ideas of the proof

Local geometric structure of random regular graphs:

In a random d-regular graph, up to radius $R = c \log_d N$, with high probability:



Most R-neighborhoods have no cycles.



All R-neighborhoods have few cycles.

A Self-consistent Equation

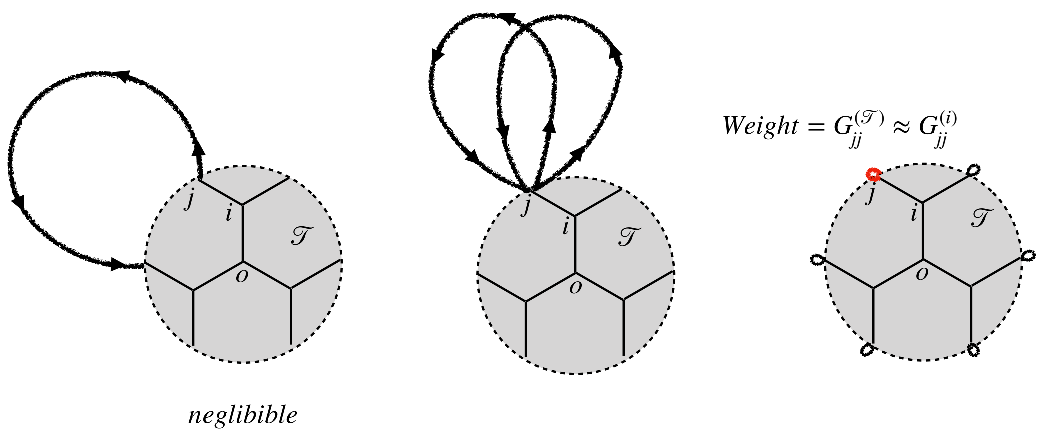
Entries of the Green's function can be interpreted as sum over weighted path:

$$G(z) = (z - H)^{-1} = \frac{1}{z} + \frac{H}{z^2} + \frac{H^2}{z^3} + \frac{H^3}{z^4} + \cdots$$

The i, j-th entry

$$G_{ij}(z) = (z - H)^{-1}_{ij} = \text{sum over weighted paths from } i \text{ to } j.$$

To compute G_{oo}, we need to sum over weighted paths from o to itself.

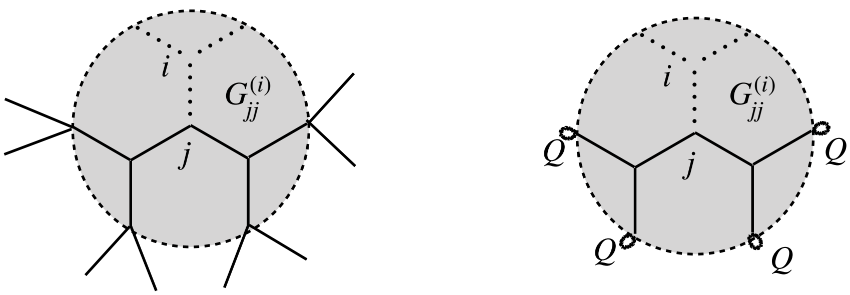


Sum over all the paths leaving from j and coming back at j, which are the same as all the paths from j to itself outside $\mathcal{T}$. We can approximate G_{oo} by the Green's function of $\mathcal{T}$ with extra weights at boundary vertices.

A self-consistent equation

$G_{jj}^{(i)}$ = Green's function with the i vertex removed and $Q = \frac{1}{Nd}\sum_{i\sim j} G_{jj}^{(i)}$.

To compute $G_{jj}^{(i)}$, we can approximate it by the Green's function of a neighborhood of j (with vertex i removed) with extra weights Q at boundary vertices.



Let $Y_\ell(Q)$ be the Green's function at the root vertex of a $(d-1)$-ary tree of depth ℓ with boundary weight Q. Then we have the fact $m_{sc} = Y_\ell(m_{sc})$ for all ℓ . If we can prove that

$$Q = \frac{1}{Nd}\sum_{i\sim j} G_{jj}^{(i)} \approx Y_\ell(Q),$$

then together with the stability of this equation, we get $Q \approx m_{sc}(z)$.

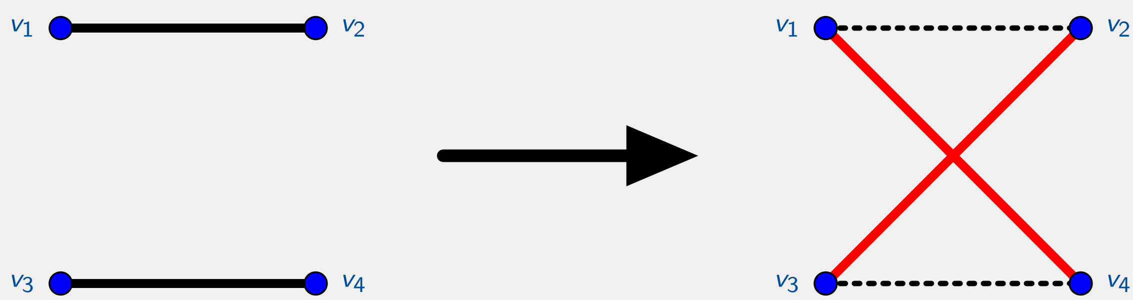
Ideas of the Proof: 1. Local resampling

Our goal is to bound $\mathbb{E}[|Q - Y_\ell(Q)|^{2p}]$ for any fixed integer p.

How to resolve the dependence of matrix entries of d-regular graphs?

Simple Switching

Replace two randomly sampled edges $\{v_1, v_2\}, \{v_3, v_4\}$ by $\{v_1, v_4\}, \{v_2, v_3\}$:



Uniform random d-regular graph is invariant under simple switching.

McKay [1981] introduced simple switchings to the random regular graph.

The resolvent expansion for local resampling

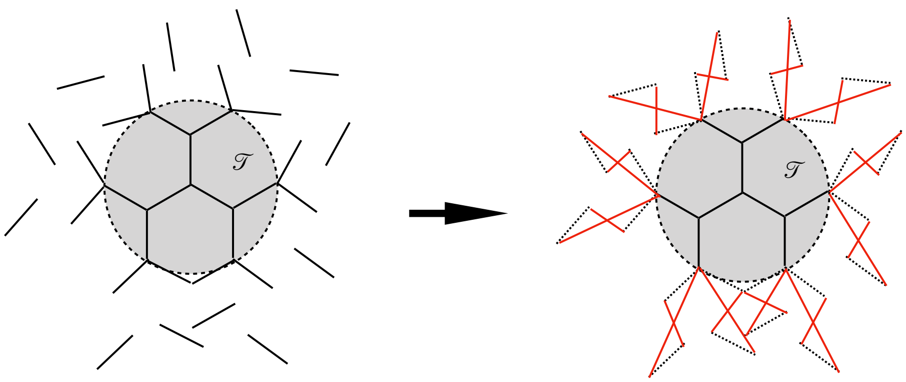
- Let H and $\widetilde{H}$ be the adjacency matrices of the original and the switched graphs.
- Denote $\xi := \widetilde{H} - H$, which is a 4×4 matrix of a simple switching.
- The resolvent identity implies

$$\widetilde{G}(z) - G(z) = \frac{1}{\tilde{H} - z} - \frac{1}{H - z} = -G(z)\xi G(z) + \cdots$$

- By our convention, $\xi \sim \frac{1}{\sqrt{d-1}}$ and it is barely smaller than one.
- This expansion is difficult to use for small d.

Generalized local resampling

- Generalized local resampling switches the $d(d-1)^{\ell-1}$ boundary edges of $\mathcal{T}$ with independently uniformly sampled edges in the remaining of the graph.



- The law of the (d-regular) graphs is invariant under the generalized switching.
- $A =$ adjacency matrix of the local tree; $B =$ the switching matrix. Schur's formula:

$$M = \begin{pmatrix} A - z & B \\ B^* & D - z \end{pmatrix}, \quad (M^{-1})_{ll} = [A - z - B(D-z)^{-1}B^*]^{-1}.$$

Generalized local resampling

- $A =$ adjacency matrix of the local tree; $B =$ the switching matrix. Schur formula:

$$M = \begin{pmatrix} A - z & B \\ B^* & D - z \end{pmatrix}, \quad (M^{-1})_{II} = [A - z - B(D-z)^{-1}B^*]^{-1}.$$

- Recall the semicircle consistent equation $m_{sc}(z) = -\frac{1}{z + m_{sc}(z)}$. If $A \to 0$ and

$$B(D-z)^{-1}B^* \to \frac{1}{N}\operatorname{Tr}(D-z)^{-1} \sim m(z), \qquad (*)$$

then Schur's formula (after averaging over the center of resampling) converges to the semicircle equation.

- For Wigner or ER graphs, the entries of B are independent and (*) is easy. But the entries are correlated for d-regular graphs.

- In addition, m_{KM} satisfies a different equation if d is fixed.

A new formula for generalized local resampling

- The Woodbury formula

$$\tilde{G} - G = (H - z + UCU)^{-1} - (H - z)^{-1} = -GU(C^{-1} + UGU)^{-1}UG.$$

  Here $\tilde{G}$ is the Green's function of the switched graph, U is the projection onto the switching and boundary edges (of $\mathcal{T}$) and C is the matrix of the switching in the finite dimensional space of switching edges.
- The right hand side depends only on G while Schur's formula involves some other quantities on the right hand side.
- The middle term on the right hand side involves only UGU which is the projection of the original Green's function onto the switching and boundary edges.
- Heuristically, UGU can be approximated by the Green's function on the forest consists of $\mathcal{T}$ and switching edges (which is basically just the Green's function on a tree).

Multi (iterative) local resampling: an example

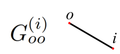
- Local resampling exchanges the boundary edge $\{\ell_\alpha, a_\alpha\}$ and a randomly sampled edge $\{b_\alpha, c_\alpha\}$ with $\{\ell_\alpha, c_\alpha\}, \{a_\alpha, b_\alpha\}$.
- Recall our goal is to bound $\mathbb{E}[|Q - Y_\ell(Q)|^{2p}]$.
- The law of $G_{oo}^{(i)}$ is the same as the law of $\widetilde{G}_{oo}^{(i)}$ for the switched graph.
- Schur and Woodbury formulas imply

$$\widetilde{G}_{oo}^{(i)} - Y_\ell(Q) = c_1 \sum_\alpha (G_{c_\alpha c_\alpha}^{(b_\alpha)} - Q) + c_2 \sum_\alpha G_{c_\alpha c_\beta}^{(b_\alpha b_\beta)} + \text{high order terms}.$$

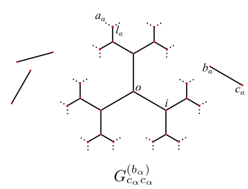
- Using this formula in one $(Q - Y_\ell(Q))$-factor in $|Q - Y_\ell(Q)|^{2p}$, we choose an edge $\{b_\alpha, c_\alpha\}$ and preform a generalized local resampling centered at $\{b_\alpha, c_\alpha\}$.
- Each time we apply a resampling, we gain a factor $(d-1)^{-\ell/2}$.
- Iterate this procedure K times until we gain $(d-1)^{-K\ell/2} \ll (N\eta)^{-1}$.
- Note that local resampling affects all quantities in any expression and the whole procedure is much more complicated.

Example of Iterated Sampling Procedure

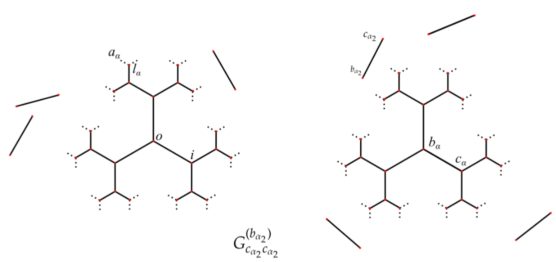
- We start with the term $G_{oo}^{(i)}$, which depends explicitly on a single edge (o, i).



- After performing one switch, our function now depends on the edges (b_α, c_α), which were randomly selected in the graph, and randomly switched with edges (l_α, a_α) at some fixed distance ℓ from o.



- We then perform another switch around a **new** edge (b_α, c_α), replacing the dependence on this edge with new randomly selected edges $(b_{\alpha_2}, c_{\alpha_2})$.



Outline

Edge universality (i.e., the TW law) for $d \geq N^{\varepsilon}$

- Noise dominating phenomenon at the edge: Define the matrix Brown motion

$$dH_t = \frac{1}{\sqrt{N}} d\mathcal{B}(t) - H_t dt \quad \mathcal{B}_{ij}(t) : \text{symm. indep. BM;}$$

$$H_t \sim e^{-t/2} H_0 + \sqrt{1 - e^{-t}} \text{GOE} \sim \sqrt{1-t} H_0 + \sqrt{t} \text{GOE} \ \text{ in law.}$$

  If the edge eigenvalue fluctuation for $H_0 = H$ is $O(N^{-2/3+\varepsilon})$, then the edge statistics of H_t with $t \sim N^{-2/3+\varepsilon}$ is given by TW possibly up to a shift (Landon-Y[2017], Adhikari-Huang[2018], ...).
- Edge universality follows by comparing the edge statistics of H and H_t .
- For random d-regular graphs, this is difficult. We compared a (constraint) matrix Brown motion with the local resampling dynamics when $d \geq N^{\varepsilon}$.
- The resampling is an expansion in $\frac{1}{\sqrt{d-1}}$. New ideas will be needed for fixed d.
- Edge universality for Wigner matrices were proved long before bulk universality by trace method. But edge universality can be much harder for two reasons:
  1. There can be (random) shifts at the edges.
  2. It takes time $O(N^{-2/3+\varepsilon})$ to reach edge universality while only $O(N^{-1+\varepsilon})$ for the bulk universality.
- Noise dominating phenomenon is of fundamental importance in large data principal component analysis .

Summary

Summary

- Eigenvalue rigidity was proved for all eigenvalues of random d-regular graphs with $d \geq 3$.
- The fluctuations of λ_2 is of order $N^{-2/3+\varepsilon}$ for $d \geq 3$.
- TW law was proved for $d \geq N^{\varepsilon}$.
- TW law (and the bulk university) for $d \geq 3$ is still open.
- For Erdős-Rényi graphs, there is a Gaussian fluctuation of order $\frac{1}{\sqrt{Nd}}$ in addition to the TW fluctuation of order $N^{-2/3}$. This implies a transition from TW to Gaussian at $d = N^{1/3}$.
- Eigenvector complete delocalization was proved for d-regular graphs with $d \geq 3$.
- Random matrix methods also yield results on eigenvector statistics and quantum unique ergodicity for random graphs.


Happy
Birthday
PETER!
RAMANUJAN
GRAPHS