Spectra of Random Graphs

Linyuan Lu
University of South Carolina

Selected Topics on Spectral Graph Theory (III)
Nankai University, Tianjin, May 29, 2014
Selected Topics on Spectral Graph Theory

1. Graphs with Small Spectral Radius  
   Time: Friday (May 16) 4pm.-5:30p.m.

2. Laplacian and Random Walks on Graphs  
   Time: Thursday (May 22) 4pm.-5:30p.m.

3. Spectra of Random Graphs  
   Time: Thursday (May 29) 4pm.-5:30p.m.

4. Hypergraphs with Small Spectral Radius  
   Time: Friday (June 6) 4pm.-5:30p.m.

5. Laplacian of Random Hypergraphs  
   Time: Thursday (June 12) 4pm.-5:30p.m.
I: Spectral Graph Theory  
II: Random Graph Theory  
III: Random Matrix Theory
Classical random theory: Erdős-Rényi model
Outline

- Classical random theory: Erdős-Rényi model
- Power law graphs
Outline

- Classical random theory: Erdős-Rényi model
- Power law graphs
- Chung-Lu model
Outline

- Classical random theory: Erdős-Rényi model
- Power law graphs
- Chung-Lu model
- Edge-independent random graphs
A graph consists of two sets $V$ and $E$.
- $V$ is the set of vertices (or nodes).
- $E$ is the set of edges, where each edge is a pair of vertices.

The degree of a vertex is the number of edges, which are incident to that vertex.

Diameter: the maximum distance $d(u, v)$, where $u$ and $v$ are in the same connected component.
A graph consists of two sets $V$ and $E$.
- $V$ is the set of vertices (or nodes).
- $E$ is the set of edges, where each edge is a pair of vertices.

The degree of a vertex is the number of edges, which are incident to that vertex.

Diameter: the maximum distance $d(u, v)$, where $u$ and $v$ are in the same connected component.

Average distance: the average among all distance $d(u, v)$ for pairs of $u$ and $v$ in the same connected component.
A random graph is a set of graphs together with a probability distribution on that set.
Random graphs

A random graph is a set of graphs together with a probability distribution on that set.

**Example:** A random graph on 3 vertices and 2 edges with the uniform distribution on it.

- Probability $\frac{1}{3}$
- Probability $\frac{1}{3}$
- Probability $\frac{1}{3}$
Random graphs

A random graph is a set of graphs together with a probability distribution on that set.

**Example:** A random graph on 3 vertices and 2 edges with the uniform distribution on it.

A random graph $G$ *almost surely* satisfies a property $P$, if

$$Pr(G \text{ satisfies } P) = 1 - o_n(1).$$
Erdős-Rényi model $G(n, p)$

- $n$ nodes
Erdős-Rényi model $G(n, p)$

- $n$ nodes
- For each pair of vertices, create an edge independently with probability $p$. 
Erdős-Rényi model $G(n, p)$

- $n$ nodes
- For each pair of vertices, create an edge independently with probability $p$.
- The graph with $e$ edges has the probability $p^e (1 - p)^{\binom{n}{2} - e}$. 
Erdős-Rényi model $G(n, p)$

- $n$ nodes
- For each pair of vertices, create an edge independently with probability $p$.
- The graph with $e$ edges has the probability $p^e(1 - p)^{\binom{n}{2} - e}$.
- $n$ nodes
- For each pair of vertices, create an edge independently with probability $p$.
- The graph with $e$ edges has the probability $p^e (1 - p)^{\binom{n}{2} - e}$.
Erdős-Rényi model $G(n, p)$

- $n$ nodes
- For each pair of vertices, create an edge independently with probability $p$.
- The graph with $e$ edges has the probability $p^e (1 - p)^{\binom{n}{2} - e}$.
Erdős-Rényi model $G(n, p)$

- $n$ nodes
- For each pair of vertices, create an edge independently with probability $p$.
- The graph with $e$ edges has the probability $p^e(1 - p)^{\binom{n}{2} - e}$.

The probability of this graph is

$$p^4(1 - p)^2.$$
Erdős-Rényi 1960s:

- $p \sim c/n$ for $0 < c < 1$: The largest connected component of $G_{n,p}$ is a tree and has about 
  $\frac{1}{\alpha}(\log n - \frac{5}{2} \log \log n)$ vertices, where $\alpha = c - 1 - \log c$. 
Erdős-Rényi 1960s:

- \( p \sim c/n \) for \( 0 < c < 1 \): The largest connected component of \( G_{n,p} \) is a tree and has about \( \frac{1}{\alpha} (\log n - \frac{5}{2} \log \log n) \) vertices, where \( \alpha = c - 1 - \log c \).

- \( p \sim 1/n + c/n^{4/3} \), the largest connected component is \( \Theta(n^{2/3}) \). **Double jump:** \( \Theta(\log n) \rightarrow \Theta(n^{2/3}) \rightarrow \Theta(n) \).
Erdős-Rényi 1960s:

- \( p \sim c/n \) for \( 0 < c < 1 \): The largest connected component of \( G_{n,p} \) is a tree and has about \( \frac{1}{\alpha} (\log n - \frac{5}{2} \log \log n) \) vertices, where \( \alpha = c - 1 - \log c \).

- \( p \sim 1/n + c/n^{4/3} \), the largest connected component is \( \Theta(n^{2/3}) \). Double jump: \( \Theta(\log n) \rightarrow \Theta(n^{2/3}) \rightarrow \Theta(n) \).

- \( p \sim c/n \) for \( c > 1 \): Except for one “giant” component, all the other components are relatively small. The giant component has approximately \( f(c)n \) vertices, where

\[
 f(c) = 1 - \frac{1}{c} \sum_{k=1}^{\infty} \frac{k^{k-1}}{k!} (ce^{-c})^k.
\]
Bollobás (1985): (denser graph)

\[
diam(G(n, p)) = \left\lfloor \frac{\log n}{\log np} \right\rfloor \quad \text{or} \quad \left\lceil \frac{\log n}{\log np} \right\rceil \quad \text{if} \quad np \gg \log n.
\]
Diameter of $G(n, p)$

Bollobás (1985): (denser graph)

$$
diam(G(n, p)) = \left\lfloor \frac{\log n}{\log np} \right\rfloor \quad \text{or} \quad \left\lceil \frac{\log n}{\log np} \right\rceil \quad \text{if } np \gg \log n. $$

Chung Lu, (2000) (Sparser graph)

$$
diam(G(n, p)) = \begin{cases} 
(1 + o(1)) \frac{\log n}{\log np} & \text{if } np \rightarrow \infty \\
\Theta(\frac{\log n}{\log np}) & \text{if } \infty > np > 1. 
\end{cases} $$
Wigner’s semicircle law

(Wigner, 1958)

- $A$ is a real symmetric $n \times n$ matrix.
- Entries $a_{ij}$ are independent random variables.
- $E(a_{ij}^{2k+1}) = 0$.
- $E(a_{ij}^2) = m^2$.
- $E(a_{ij}^{2k}) < M$.

The distribution of eigenvalues of $A$ converges into a semicircle distribution of radius $2m \sqrt{n}$. 
The eigenvalues of an Erdős-Rényi random graph follow the semicircle law. (Füredi and Komlós, 1981)

Laplacian eigenvalues also follow the semicircle law.
Erdős-Rényi model $G(n, p)$ is classical, simple, beautiful..., but not suitable to model complex graphs.
Erdős-Rényi model $G(n, p)$ is classical, simple, beautiful..., but not suitable to model complex graphs.

- What are complex graphs?
Erdős-Rényi model $G(n, p)$ is classical, simple, beautiful..., but not suitable to model complex graphs.

- What are complex graphs?
- How to model these complex graphs by random graphs?
Challenge

Erdős-Rényi model $G(n, p)$ is classical, simple, beautiful..., but not suitable to model complex graphs.

- What are complex graphs?
- How to model these complex graphs by random graphs?
- How to deduce the graph properties of these general random graph models?
Examples of complex graphs

WWW Graphs
Call Graphs
Collaboration Graphs
Gene Regulatory Graphs
Graph of U.S. Power Grid
Costars Graph of Actors

Faculty, Ph.D. students, Postdocs, and visitors to the Combinatorics Group at the University of South Carolina.
An IP Graph (by Bill Cheswick)
Vertex: AS (autonomous system)

Edges: AS pairs in BGP routing table.
Large BGP subgraph

Only a portion of 6400 vertices and 13000 edges is drawn.
Hollywood Graph

Vertex: actors and actress

Edges: co-playing in the same movie

Only 10,000 out of 225,000 are shown.
Protein-interaction network

A subgraph of the Collaboration Graph
Folklore of Erdős numbers

- Erdős has Erdős number 0.
- Erdős’ coauthor has Erdős number 1.
- Erdős’ coauthor’s coauthor has Erdős number 2.
  
  ...
Folklore of Erdős numbers

- Erdős has Erdős number 0.
- Erdős’ coauthor has Erdős number 1.
- Erdős’ coauthor’s coauthor has Erdős number 2.

... 

My Erdős number is 2.
Folklore of Erdős numbers

- Erdős has Erdős number 0.
- Erdős’ coauthor has Erdős number 1.
- Erdős’ coauthor’s coauthor has Erdős number 2.

My Erdős number is 2.

Erdős number is the graph distance to Erdős in the Collaboration graph.
An induced subgraph of the collaboration graph (with Erdos number at most 2).
Made by Fan Chung Graham and Lincoln Lu in 2002.
Characteristics

- Large
Characteristics

- Large
- Sparse
Characteristics

- Large
- Sparse
- Power law degree distribution
Characteristics

- Large
- Sparse
- Power law degree distribution
- Small world phenomenon
The power law

The number of vertices of degree $k$ is approximately proportional to $k^{-\beta}$ for some positive $\beta$. 

![Graph showing the power law distribution](image)
The power law

The number of vertices of degree $k$ is approximately proportional to $k^{-\beta}$ for some positive $\beta$.

A power law graph is a graph whose degree sequence satisfies the power law.
Left: The collaboration graph follows the power law degree distribution with exponent $\beta \approx 3.0$.
Power law distribution

Left: The collaboration graph follows the power law degree distribution with exponent $\beta \approx 3.0$

Right: An IP graph follows the power law degree distribution with exponent $\beta \approx 2.4$
Power law graphs

Left: Part of the collaboration graph (authors with Erdős number 2)

Right: An IP graph (by Bill Cheswick)
Robustness of Power Law

Size  
25,3339

Degree distribution

52,186
Basic questions

- How to model power law graphs?
Basic questions

- How to model power law graphs?

- What graph properties can be derived from the model?
Model \( G(w_1, w_2, \ldots, w_n) \)

Random graph model with given expected degree sequence (Chung-Lu model)

- \( n \) nodes with weights \( w_1, w_2, \ldots, w_n \).
Random graph model with given expected degree sequence (Chung-Lu model)

- $n$ nodes with weights $w_1, w_2, \ldots, w_n$.
- For each pair $(i, j)$, create an edge independently with probability $p_{ij} = w_i w_j \rho$, where $\rho = \frac{1}{\sum_{i=1}^{n} w_i}$. 
Random graph model with given expected degree sequence (Chung-Lu model)

- $n$ nodes with weights $w_1, w_2, \ldots, w_n$.
- For each pair $(i, j)$, create an edge independently with probability $p_{ij} = w_i w_j \rho$, where $\rho = \frac{1}{\sum_{i=1}^{n} w_i}$.
- The graph $H$ has probability
\[
\prod_{ij \in E(\mathcal{H})} p_{ij} \prod_{ij \notin E(\mathcal{H})} (1 - p_{ij}).
\]
Model $G(w_1, w_2, \ldots, w_n)$

Random graph model with given expected degree sequence (Chung-Lu model)

- $n$ nodes with weights $w_1, w_2, \ldots, w_n$.

- For each pair $(i, j)$, create an edge independently with probability $p_{ij} = w_i w_j \rho$, where $\rho = \frac{1}{\sum_{i=1}^{n} w_i}$.

- The graph $H$ has probability

$$\prod_{ij \in E(H)} p_{ij} \prod_{ij \notin E(H)} (1 - p_{ij}).$$

- The expected degree of vertex $i$ is $w_i$. 
An example: $G(w_1, w_2, w_3, w_4)$
An example: $G(w_1, w_2, w_3, w_4)$
An example: $G(w_1, w_2, w_3, w_4)$
An example: $G(w_1, w_2, w_3, w_4)$
An example: $G(w_1, w_2, w_3, w_4)$
An example: \( G(w_1, w_2, w_3, w_4) \)

The probability of the graph is

\[
w_1^3 w_2^2 w_3^2 w_4 \rho^4 (1 - w_2 w_4 \rho) \times (1 - w_3 w_4 \rho) \prod_{i=1}^{4} (1 - w_i^2 \rho).
\]
Chung-Lu model

For $G = G(w_1, \ldots, w_n)$, let

- $d = \frac{1}{n} \sum_{i=1}^{n} w_i$
- $\tilde{d} = \frac{\sum_{i=1}^{n} w_i^2}{\sum_{i=1}^{n} w_i}$
- The volume of $S$: $\text{Vol}(S) = \sum_{i \in S} w_i$. 
For $G = G(w_1, \ldots, w_n)$, let

- $d = \frac{1}{n} \sum_{i=1}^{n} w_i$
- $\tilde{d} = \frac{\sum_{i=1}^{n} w_i^2}{\sum_{i=1}^{n} w_i}$.
- The volume of $S$: $\text{Vol}(S) = \sum_{i \in S} w_i$.

We have

$$\tilde{d} \geq d$$

“=” holds if and only if $w_1 = \cdots = w_n$. 
For $G = G(w_1, \ldots, w_n)$, let

- $d = \frac{1}{n} \sum_{i=1}^{n} w_i$
- $\tilde{d} = \frac{\sum_{i=1}^{n} w_i^2}{\sum_{i=1}^{n} w_i}$.
- The volume of $S$: $\text{Vol}(S) = \sum_{i \in S} w_i$.

We have

$$\tilde{d} \geq d$$

“$=$” holds if and only if $w_1 = \cdots = w_n$.

A connected component $S$ is called a giant component if

$$\text{vol}(S) = \Theta(\text{vol}(G)).$$
Chung and Lu (2001) For $G = G(w_1, \ldots, w_n)$,

- If $\tilde{d} < 1 - \epsilon$, then almost surely, all components have volume at most $O(\sqrt{n} \log n)$. 
Chung and Lu (2001) For $G = G(w_1, \ldots, w_n)$,

- If $\tilde{d} < 1 - \epsilon$, then almost surely, all components have volume at most $O(\sqrt{n \log n})$.
- If $d > 1 + \epsilon$, then almost surely there is a unique giant component of volume $\Theta(\text{Vol}(G))$. All other components have size at most

$$\begin{cases} 
\frac{\log n}{d-1-\log d-\epsilon d} & \text{if } \frac{1}{1-\epsilon} < d < \frac{2}{1-\epsilon} \\
\frac{\log n}{1+\log d-\log 4+2\log(1-\epsilon)} & \text{if } d > \frac{4}{e(1-\epsilon)^2}.
\end{cases}$$
Chung and Lu (2004)
If the average degree is strictly greater than 1, then almost surely the giant component in a graph $G$ in $G(w)$ has volume $(\lambda_0 + O(\sqrt{n \frac{\log n}{\text{Vol}(G)}})) \text{Vol}(G)$, where $\lambda_0$ is the unique positive root of the following equation:

$$\sum_{i=1}^{n} w_i e^{-w_i \lambda} = (1 - \lambda) \sum_{i=1}^{n} w_i.$$
A real application

Apply to the Collaboration Graph (2002 data):
The size of giant component is predicted to be about 177,400 by our theory. This is rather close to the actual value 176,000, within an error bound of less than 1%.
Question: Does the random graph with equal expected degrees generates the smallest giant component among all possible degree distribution with the same volume?
**Question:** Does the random graph with equal expected degrees generates the smallest giant component among all possible degree distribution with the same volume?

**Chung Lu (2004)**

- Yes, for $1 < d \leq \frac{e}{e-1}$.
Question: Does the random graph with equal expected degrees generates the smallest giant component among all possible degree distribution with the same volume?

Chung Lu (2004)

■ Yes, for $1 < d \leq \frac{e}{e-1}$.
■ No, for sufficiently large $d$. 
Question: Does the random graph with equal expected degrees generates the smallest giant component among all possible degree distribution with the same volume?

Chung Lu (2004)

- Yes, for \(1 < d \leq \frac{e}{e-1}\).
- No, for sufficiently large \(d\).
- When \(d \geq \frac{4}{e}\), almost surely the giant component of \(G(w_1, \ldots, w_n)\) has volume at least

\[
\left(\frac{1}{2}(1 + \sqrt{1 - \frac{4}{de}}) + o(1)\right)\text{Vol}(G).
\]

This is asymptotically best possible.
Diameter of $G(w_1, \ldots, w_n)$

Chung Lu (2002)

- For a random graph $G$ with admissible expected degree sequence $(w_1, \ldots, w_n)$, the average distance is almost surely $(1 + o(1)) \frac{\log n}{\log d}$. 
Chung Lu (2002)

- For a random graph $G$ with **admissible** expected degree sequence $(w_1, \ldots, w_n)$, the average distance is almost surely $(1 + o(1)) \frac{\log n}{\log d}$.

- For a random graph $G$ with **strongly admissible** expected degree sequence $(w_1, \ldots, w_n)$, the diameter is almost surely $\Theta\left(\frac{\log n}{\log d}\right)$. 
Diameter of $G(w_1, \ldots, w_n)$

Chung Lu (2002)

- For a random graph $G$ with admissible expected degree sequence $(w_1, \ldots, w_n)$, the average distance is almost surely $(1 + o(1)) \frac{\log n}{\log \tilde{d}}$.

- For a random graph $G$ with strongly admissible expected degree sequence $(w_1, \ldots, w_n)$, the diameter is almost surely $\Theta\left(\frac{\log n}{\log \tilde{d}}\right)$.

These results apply to $G(n, p)$ and random power law graph with $\beta > 3$. 
A random subgraph of the Collaboration Graph.

A Connected component of $G(n, p)$ with $n = 500$ and $p = 0.002$. 
A random subgraph of the Collaboration Graph with $n = 500$ and $p = 0.002$.

- Dense core for non-admissible graphs.
- No dense core for admissible graphs.
Power law graphs with $\beta \in (2, 3)$

Chung, Lu (2002)

- Examples: the WWW graph, Collaboration graph, etc.
Chung, Lu (2002)
- Examples: the WWW graph, Collaboration graph, etc.
- Non-admissible.
Chung, Lu (2002)

- Examples: the WWW graph, Collaboration graph, etc.
- Non-admissible.
- Containing a dense core, with diameter $\log \log n$. 
Chung, Lu (2002)

- Examples: the WWW graph, Collaboration graph, etc.
- Non-admissible.
- Containing a dense core, with diameter $\log \log n$.
- Mostly vertices are within the distance of $O(\log \log n)$ from the core.
Power law graphs with $\beta \in (2, 3)$

Chung, Lu (2002)

- Examples: the WWW graph, Collaboration graph, etc.
- Non-admissible.
- Containing a dense core, with diameter $\log \log n$.
- Mostly vertices are within the distance of $O(\log \log n)$ from the core.
- There are some vertices at the distance of $O(\log n)$. 
Chung, Lu (2002)

- Examples: the WWW graph, Collaboration graph, etc.
- Non-admissible.
- Containing a dense core, with diameter $\log \log n$.
- Mostly vertices are within the distance of $O(\log \log n)$ from the core.
- There are some vertices at the distance of $O(\log n)$.

The diameter is $\Theta(\log n)$, while the average distance is $O(\log \log n)$. 
Experimental results

- **Faloutsos et al. (1999)** The eigenvalues of the Internet graph do not follow the semicircle law.
- **Farkas et. al. (2001), Goh et. al. (2001)** The spectrum of a power law graph follows a “triangular-like” distribution.
- **Mihail and Papadimitriou (2002)** They showed that the large eigenvalues are determined by the large degrees. Thus, the significant part of the spectrum of a power law graph follows the power law.

\[ \mu_i \approx \sqrt{d_i}. \]
Suppose \( w_1 \geq w_2 \geq \ldots \geq w_n \). Let \( \mu_i \) be \( i \)-th largest eigenvalue of \( G(w_1, w_2, \ldots, w_n) \). Let \( m = w_1 \) and \( \tilde{d} = \sum_{i=1}^{n} w_i^2 \rho \). Almost surely we have:

\[
(1-o(1)) \max\{\sqrt{m}, \tilde{d}\} \leq \mu_1 \leq 7\sqrt{\log n} \cdot \max\{\sqrt{m}, \tilde{d}\}.
\]
Suppose $w_1 \geq w_2 \geq \ldots \geq w_n$. Let $\mu_i$ be $i$-th largest eigenvalue of $G(w_1, w_2, \ldots, w_n)$. Let $m = w_1$ and $\tilde{d} = \sum_{i=1}^{n} w_i^2 \rho$. Almost surely we have:

- $(1-o(1)) \max\{\sqrt{m}, \tilde{d}\} \leq \mu_1 \leq 7 \sqrt{\log n} \cdot \max\{\sqrt{m}, \tilde{d}\}.$
- $\mu_1 = (1 + o(1)) \tilde{d}$, if $\tilde{d} > \sqrt{m} \log n.$

Suppose \( w_1 \geq w_2 \geq \ldots \geq w_n \). Let \( \mu_i \) be \( i \)-th largest eigenvalue of \( G(w_1, w_2, \ldots, w_n) \). Let \( m = w_1 \) and \( \tilde{d} = \sum_{i=1}^{n} w_i^2 \rho \). Almost surely we have:

1. \( (1-o(1)) \max \{ \sqrt{m}, \tilde{d} \} \leq \mu_1 \leq 7\sqrt{\log n} \cdot \max \{ \sqrt{m}, \tilde{d} \} \).
2. \( \mu_1 = (1 + o(1))\tilde{d} \), if \( \tilde{d} > \sqrt{m} \log n \).
3. \( \mu_1 = (1 + o(1))\sqrt{m} \), if \( \sqrt{m} > \tilde{d} \log^2 n \).

Suppose $w_1 \geq w_2 \geq \ldots \geq w_n$. Let $\mu_i$ be $i$-th largest eigenvalue of $G(w_1, w_2, \ldots, w_n)$. Let $m = w_1$ and

$$\tilde{d} = \sum_{i=1}^{n} w_i^2 \rho.$$ 

Almost surely we have:

- $(1-o(1)) \max\{\sqrt{m}, \tilde{d}\} \leq \mu_1 \leq 7\sqrt{\log n} \cdot \max\{\sqrt{m}, \tilde{d}\}$.
- $\mu_1 = (1 + o(1))\tilde{d}$, if $\tilde{d} > \sqrt{m} \log n$.
- $\mu_1 = (1 + o(1))\sqrt{m}$, if $\sqrt{m} > \tilde{d} \log^2 n$.
- $\mu_k \approx \sqrt{w_k}$ and $\mu_{n+1-k} \approx -\sqrt{w_k}$, if $\sqrt{w_k} > \tilde{d} \log^2 n$. 

The first $k$ and last $k$ eigenvalues of the random power law graph with $\beta > 2.5$ follows the power law distribution with exponent $2\beta - 1$. It results a “triangular-like” shape.
Random walks on a graph $G$:

$$\pi_{k+1} = AD^{-1}\pi_k.$$ 

$$AD^{-1} \sim D^{-1/2} AD^{-1/2}.$$
Random walks on a graph \( G \):

\[
\pi_{k+1} = AD^{-1}\pi_k.
\]

\[
AD^{-1} \sim D^{-1/2}AD^{-1/2}.
\]

Laplacian spectrum

\[
0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1} \leq 2
\]

are the eigenvalues of \( L = I - D^{-1/2}AD^{-1/2} \).
Random walks on a graph $G$:

$$\pi_{k+1} = AD^{-1}\pi_k.$$  

$$AD^{-1} \sim D^{-1/2} AD^{-1/2}.$$  

Laplacian spectrum

$$0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{n-1} \leq 2$$

are the eigenvalues of $L = I - D^{-1/2} AD^{-1/2}$.

The eigenvalues of $AD^{-1}$ are $1, 1 - \lambda_1, \ldots, 1 - \lambda_{n-1}$. 

Laplacian spectrum
Let

- $w_{\text{min}} = \min\{w_1, \ldots, w_n\}$,
- $d = \frac{1}{n} \sum_{i=1}^{n} w_i$,
- $g(n)$ — a function tending to infinity arbitrarily slowly.


If $w_{\text{min}} \gg \log^2 n$, then almost surely the Laplacian spectrum $\lambda_i$'s of $G(w_1, \ldots, w_n)$ satisfy

$$\max_{i \neq 0} |1 - \lambda_i| \leq (1 + o(1)) \frac{4}{\sqrt{d}} + \frac{g(n) \log^2 n}{w_{\text{min}}}.$$
Let

- \( w_{\text{min}} = \min\{w_1, \ldots, w_n\} \),
- \( d = \frac{1}{n} \sum_{i=1}^{n} w_i \),
- \( g(n) \) — a function tending to infinity arbitrarily slowly.

**Chung, Vu, and Lu (2003)**

- If \( w_{\text{min}} \gg \log^2 n \), then almost surely the Laplacian spectrum \( \lambda_i \)'s of \( G(w_1, \ldots, w_n) \) satisfy

\[
\max_{i \neq 0} |1 - \lambda_i| \leq (1 + o(1)) \frac{4}{\sqrt{d}} + \frac{g(n) \log^2 n}{w_{\text{min}}}.
\]

- If \( w_{\text{min}} \gg \sqrt{d} \), the Laplacian spectrum follows the semi-circle distribution with radius \( r \approx \frac{2}{\sqrt{d}} \).
General edge-independent random graphs:

- \( n \): the number of vertices.
- \( p_{ij} \): a probability for \( ij \) being an edge.
- Edges are mutually independent.

**Question:** What can we say about the spectrum of the adjacency matrix and the Laplacian matrix?
Notation

- $A$: adjacency matrix
- $\bar{A} := (p_{ij})$: the expectation of $A$
- $\Delta$: the maximum expected degree
- $\delta$: the minimum expected degree
- $D$: the diagonal matrix of degrees
- $\bar{D}$: the expectation of $D$
- $L := I - D^{-1/2}AD^{-1/2}$: the normalized Laplacian
- $\bar{L} := I - \bar{D}^{-1/2}\bar{A}\bar{D}^{-1/2}$: the Laplacian of $\bar{A}$
Known results

Oliveira [2010]: For $\Delta \geq C \ln n$, with high probability we have

$$|\lambda_i(A) - \lambda_i(\bar{A})| \leq 4\sqrt{\Delta \ln n}.$$

For $\delta \geq C \ln n$, with high probability we have

$$\lambda_i(L) - \lambda_i(\bar{L}) \leq 14\sqrt{\ln(4n)/\delta}.$$
Known results

Oliveira [2010]: For $\Delta \geq C \ln n$, with high probability we have

$$|\lambda_i(A) - \lambda_i(\bar{A})| \leq 4\sqrt{\Delta \ln n}.$$ 

For $\delta \geq C \ln n$, with high probability we have

$$\lambda_i(L) - \lambda_i(\bar{L}) \leq 14\sqrt{\ln(4n)/\delta}.$$ 

Chung-Radcliffe [2011] reduces the constant coefficient using a new matrix Chernoff inequality.
Our results

Lu-Peng [2012+] : If $\Delta \gg \ln^4 n$, then almost surely

$$|\lambda_i(A) - \lambda_i(\tilde{A})| \leq (2 + o(1))\sqrt{\Delta}.$$
Our results

Lu-Peng [2012+]: If $\Delta \gg \ln^4 n$, then almost surely

$$|\lambda_i(A) - \lambda_i(\bar{A})| \leq (2 + o(1))\sqrt{\Delta}.$$ 

Lu-Peng [2012+]:

Let $\Lambda := \{\lambda_i(\bar{L}): |1 - \lambda_i(\bar{L})| = \omega(1/\sqrt{\ln n})\}$. If $\delta \gg \max\{|\Lambda|, \ln^4 n\}$, then almost surely

$$|\lambda_i(L) - \lambda_i(\bar{L})| \leq \left(2 + \sqrt{\sum_{\lambda \in \Lambda} (1 - \lambda)^2 + o(1)}\right) \frac{1}{\sqrt{\delta}}.$$
Lu-Peng [2012+]: If $\Delta \gg \ln^4 n$, then almost surely

$$|\lambda_i(A) - \lambda_i(\bar{A})| \leq (2 + o(1))\sqrt{\Delta}.$$ 

Lu-Peng [2012+]:
Let $\Lambda := \{\lambda_i(\bar{L}) : |1 - \lambda_i(\bar{L})| = \omega(1/\sqrt{\ln n})\}$. If $\delta \gg \max\{|\Lambda|, \ln^4 n\}$, then almost surely

$$|\lambda_i(L) - \lambda_i(\bar{L})| \leq \left(2 + \sqrt{\sum_{\lambda \in \Lambda} (1 - \lambda)^2 + o(1)}\right) \frac{1}{\sqrt{\delta}}.$$ 

In both case, we remove the multiplicative factor $\sqrt{\ln n}$. 
$B = (b_{i,j})$ is a random symmetric matrix satisfying:

- $b_{i,j}$: independent, but not necessary identical,
- $|b_{i,j}| \leq K$,
- $E(b_{i,j}) = 0$,
- $\text{Var}(b_{i,j}) \leq \sigma^2$. 
\( B = (b_{i,j}) \) is a random symmetric matrix satisfying:

- \( b_{i,j} \): independent, but not necessary identical,
- \( |b_{i,j}| \leq K \),
- \( \mathbb{E}(b_{i,j}) = 0 \),
- \( \text{Var}(b_{i,j}) \leq \sigma^2 \).

**Füredi-Komlós [1981]:**

\[
\| B \| \leq 2\sigma \sqrt{n} + cn^{1/3} \ln n.
\]
$B = (b_{i,j})$ is a random symmetric matrix satisfying:

- $b_{i,j}$: independent, but not necessary identical,
- $|b_{i,j}| \leq K$,
- $E(b_{i,j}) = 0$,
- $\text{Var}(b_{i,j}) \leq \sigma^2$.

**Füredi-Komlós [1981]:**

$$\|B\| \leq 2\sigma \sqrt{n} + cn^{1/3} \ln n.$$

**Vu [2007]:**

$$\|B\| \leq 2\sigma \sqrt{n} + c(K\sigma)^{1/2}n^{1/4} \ln n.$$
Lu-Peng [2012+]: We further assume \( \text{Var}(b_{ij}) \leq \sigma_{ij}^2 \). Let
\[
\Delta := \max_{1 \leq i \leq n} \sum_{j=1}^{n} \sigma_{ij}^2.
\]
If \( \Delta \geq C'K^2 \ln^4 n \), then asymptotically almost surely
\[
\|B\| \leq 2\sqrt{\Delta} + C\sqrt{K}\Delta^{1/4} \ln n.
\]

- It generalizes Vu’s theorem.
- This result is asymptotically tight.
Graph percolation

- $G$: a connected graph on $n$ vertices
- $p$: a probability ($0 \leq p \leq 1$)

$G_p$: a random spanning subgraph of $G$, obtained as follows: for each edge $f$ of $G$, independently,

$$\Pr(f \text{ is an edge of } G_p) = p.$$
Graph percolation

- \( G \): a connected graph on \( n \) vertices
- \( p \): a probability \( (0 \leq p \leq 1) \)

\( G_p \): a random spanning subgraph of \( G \), obtained as follows: for each edge \( f \) of \( G \), independently,

\[
\Pr(f \text{ is an edge of } G_p) = p.
\]
Lu-Peng [2012+]:

If \( p \gg \frac{\ln^4 n}{\Delta} \), then almost surely we have

\[
|\lambda_i(A(G_p)) - p\lambda_i(A(G))| \leq (2 + o(1))\sqrt{p\Delta}.
\]
Lu-Peng [2012+]:

- If \( p \gg \frac{\ln^4 n}{\Delta} \), then almost surely we have

  \[
  |\lambda_i(A(G_p)) - p\lambda_i(A(G))| \leq (2 + o(1))\sqrt{p\Delta}.
  \]

- Suppose that all but \( k \) Laplacian eigenvalues \( \lambda \) of \( G \) satisfies \( |1 - \lambda| = o\left(\frac{1}{\sqrt{\ln n}}\right) \). If \( \delta \gg \max\{k, \ln^4 n\} \), then for \( p \gg \max\left\{ \frac{k}{\delta}, \frac{\ln^4 n}{\delta} \right\} \), almost surely we have

  \[
  |\lambda_i(L(G_p)) - \lambda_i(L(G))| \leq (2 + \sqrt{\sum_{i=1}^{k}(1 - \lambda_i)^2 + o(1)})) \frac{1}{\sqrt{p\delta}}.
  \]
We will illustrate Wigner’s trace method through the sketch proof of the following result.
We will illustrate Wigner’s trace method through the sketch proof of the following result.

**Lu-Peng [2012+]**: If $B = (b_{ij})$ is a random symmetric matrix satisfying:

- $b_{ij}$: independent, but not necessary identical,
- $|b_{ij}| \leq K$,
- $\mathbb{E}(b_{ij}) = 0$,
- $\text{Var}(b_{ij}) \leq \sigma_{ij}^2$.

then almost surely

$$\|B\| \leq 2\sqrt{\Delta} + C\sqrt{K}\Delta^{1/4} \ln n,$$

where $\Delta := \max_{1 \leq i \leq n} \sum_{j=1}^{n} \sigma_{ij}^2$. 
WLOG, we can assume \( K = 1 \) and \( b_{ii} = 0 \). Using Wigner’s trace method, we have

\[
E(\text{Trace}(B^k)) = \sum_{i_1, i_2, \ldots, i_k} E(b_{i_1i_2} b_{i_2i_3} \ldots b_{i_{k-1}i_k} b_{i_ki_1})
\]

\[
= \sum_{p=2}^{\lfloor k/2 \rfloor + 1} \sum_{w \in \mathcal{G}(n, k, p)} \prod_{e \in E(w)} E(b_e^{q_e}).
\]

Here \( \mathcal{G}(n, k, p) \) is the set of “good” closed walks \( w \) in \( K_n \) of length \( k \) on \( p \) vertices, where each edge in \( w \) appears more than once \( (q_e \geq 2) \).
Let $\tilde{G}(k, p)$ be the set of good closed walks $w$ of length $k$ on the complete graph $K_p$ where vertices first appear in $w$ in the order $1, 2, \ldots, p$. 
Let $\tilde{G}(k, p)$ be the set of good closed walks $w$ of length $k$ on the complete graph $K_p$ where vertices first appear in $w$ in the order $1, 2, \ldots, p$.

All walks in $G(n, k, p)$ can be coded by a walk in $\tilde{G}(k, p)$ plus the ordered $p$ distinct vertices. Let $[n]^p := \{(v_1, v_2, \ldots, v_p) \in [n]^p : v_1, v_2, \ldots, v_p$ are distinct$\}$. 
Let $\tilde{G}(k, p)$ be the set of good closed walks $w$ of length $k$ on the complete graph $K_p$ where vertices first appear in $w$ in the order $1, 2, \ldots, p$.

All walks in $G(n, k, p)$ can be coded by a walk in $\tilde{G}(k, p)$ plus the ordered $p$ distinct vertices. Let $[n]^p := \{(v_1, v_2, \ldots, v_p) \in [n]^p : v_1, v_2, \ldots, v_p \text{ are distinct}\}$.

Define a rooted tree $T(w)$ so that the edge $i_ji_{j+1} \in E(T(w))$ if it brings in a new vertex $i_{j+1}$ when it occurs first time.
\[
\sum_{w \in G(n,k,p)} \prod_{e \in E(w)} \sigma_e^2 = \sum_{\tilde{w} \in \tilde{G}(k,p)} \sum_{(v_1, \ldots, v_p) \in [n]^p} \prod_{x y \in E(\tilde{w})} \sigma_{v_x v_y}^2
\]

\[
\leq \sum_{\tilde{w} \in \tilde{G}(k,p)} \sum_{v_1=1}^{n} \sum_{v_2=1}^{n} \cdots \sum_{v_p=1}^{n} \prod_{x y \in E(T)} \sigma_{v_x v_y}^2
\]

\[
= \sum_{\tilde{w} \in \tilde{G}(k,p)} \sum_{v_1=1}^{n} \sum_{v_2=1}^{n} \cdots \sum_{v_{p-1}=1}^{n} \prod_{y = 2}^{p-1} \sigma_{v_{\eta(y)} v_y}^2 \sum_{v_p=1}^{n} \sigma_{v_{\eta(p)} v_p}^2
\]

\[
\leq \Delta \sum_{\tilde{w} \in \tilde{G}(k,p)} \sum_{v_1=1}^{n} \sum_{v_2=1}^{n} \cdots \sum_{v_{p-1}=1}^{n} \prod_{y = 2}^{p-1} \sigma_{v_{\eta(y)} v_y}^2
\]

\[
\leq \cdots
\]

\[
\leq n \Delta^{p-1} |\tilde{G}(k,p)|.
\]
Vu [2007] proved

$$|\tilde{G}(k,p)| \leq \binom{k}{2p-2} 2^{2k-2p+3} p^{k-2p+2} (k - 2p + 4)^{k-2p+2}.$$
Vu [2007] proved

\[ |\tilde{G}(k, p)| \leq \binom{k}{2p - 2} 2^{2k-2p+3} p^{k-2p+2} (k - 2p + 4)^{k-2p+2}. \]

We get

\[ |E(\text{Trace}(B^k))| \leq \sum_{w \in \mathcal{G}(n,k)} \prod_{e \in E(w)} \sigma_e^2 \leq \sum_{p=2}^{k/2+1} n \Delta^{p-1} \left| \tilde{G}(k, p) \right| \]

\[ \leq n \sum_{p=2}^{k/2+1} \Delta^{p-1} \binom{k}{2p - 2} 2^{2k-2p+3} p^{k-2p+2} (k - 2p + 4)^{k-2p+2} \]

\[ := n \sum_{p=2}^{k/2+1} S(n, k, p). \]
One can show

\[ S(n, k, p - 1) \leq \frac{16k^4}{\Delta} S(n, k, p). \]
One can show

\[ S'(n, k, p - 1) \leq \frac{16k^4}{\Delta} S(n, k, p). \]

For any even integer \( k \) such that \( k^4 \leq \frac{\Delta}{32} \), we get

\[
\left| E \left( \text{Trace} \left( B^k \right) \right) \right| \leq \sum_{p=2}^{k/2+1} S(n, k, p) \]

\[
\leq S(n, k, k/2 + 1) \sum_{p=2}^{k/2+1} \left( \frac{1}{2} \right)^{k/2+1-p} \]

\[
< 2S(n, k, k/2 + 1) \]

\[
= n2^{k+2}\Delta^{k/2}.
\]
For even $k$, we have

$$
\Pr(\|B\| \geq 2\sqrt{\Delta} + C\Delta^{1/4} \ln n) = \Pr(\|B\|^k \geq (2\sqrt{\Delta} + C\Delta^{1/4} \ln n)^k) \leq \Pr(\text{Trace}(B^k) \geq (2\sqrt{\Delta} + C\Delta^{1/4} \ln n)^k) \\
\leq \frac{\mathbb{E}(\text{Trace}(B^k))}{(2\sqrt{\Delta} + C\Delta^{1/4} \ln n))^k} \quad (\text{Markov’s inequality}) \\
\leq \frac{n2^{k+2} \Delta^{k/2}}{(2\sqrt{\Delta} + C\Delta^{1/4} \ln n))^k} \\
= 4ne^{-(1+o(1))\frac{C}{2}k\Delta^{-1/4} \ln n}.
$$

Setting $k = \left(\frac{\Delta}{32}\right)^{1/4}$, this probability is $o(1)$ for sufficiently large $C$. \qed
For even $k$, we have

$$\Pr(\|B\| \geq 2\sqrt{\Delta} + C\Delta^{1/4} \ln n)$$

$$= \Pr(\|B\|^k \geq (2\sqrt{\Delta} + C\Delta^{1/4} \ln n)^k)$$

$$\leq \Pr(\text{Trace}(B^k) \geq (2\sqrt{\Delta} + C\Delta^{1/4} \ln n)^k)$$

$$\leq \frac{\mathbb{E}(\text{Trace}(B^k))}{(2\sqrt{\Delta} + C\Delta^{1/4} \ln n))^k}$$

$$\leq \frac{n2^{k+2}\Delta^{k/2}}{(2\sqrt{\Delta} + C\Delta^{1/4} \ln n))^k}$$

$$= 4ne^{-(1+o(1))\frac{C}{2}k\Delta^{-1/4} \ln n}.$$

Setting $k = (\frac{\Delta}{32})^{1/4}$, this probability is $o(1)$ for sufficiently large $C$. $\square$
For $p < p_c$, almost surely there is no giant component.

For $p > p_c$, almost surely there is a giant component.
Motivations

- Graph theory: random graphs
- Theoretical physics: crystals melting
- Sociology: the spread of disease on contact networks
Percolation of $\mathbb{Z}^d$
Percolation of $\mathbb{Z}^d$
Kesten (1980): \( p_c(\mathbb{Z}^2) = \frac{1}{2} \).
**Percolation of $\mathbb{Z}^d$**

Kesten (1980): $p_c(\mathbb{Z}^2) = \frac{1}{2}$.

Lorenz and Ziff (1997, simulation):

$p_c(\mathbb{Z}^3) \approx 0.2488126 \pm 0.0000005$ if it exists.
Kesten (1980): $p_c(\mathbb{Z}^2) = \frac{1}{2}$.

Lorenz and Ziff (1997, simulation): $p_c(\mathbb{Z}^3) \approx 0.2488126 \pm 0.0000005$ if it exists.

Kesten (1990): $p_c(\mathbb{Z}^d) \sim \frac{1}{2d}$ as $d \to \infty$. 
Alon, Benjamini, Stacey (2004): Suppose $d \geq 2$ and let $(G_n)$ be a sequence of $d$-regular expanders with $\text{girth}(G_n) \to \infty$, then

$$p_c = \frac{1}{d-1} + o(1).$$
Bollobás, Borgs, Chayes, and Riordan (2008): Suppose that $G$ is a dense graph (i.e., average degree $d = \Theta(n)$). Let $\mu$ be the largest eigenvalue of the adjacency matrix of $G$. Then

$$p_c \approx \frac{1}{\mu}.$$
Bollobás, Borgs, Chayes, and Riordan (2008): Suppose that $G$ is a dense graph (i.e., average degree $d = \Theta(n)$). Let $\mu$ be the largest eigenvalue of the adjacency matrix of $G$. Then

$$p_c \approx \frac{1}{\mu}.$$ 

Remark: The requirement of “dense graph” is essential. Their methods cannot be extended to sparse graphs.
Chung, Lu, Horn [2008]:

If \( p < \frac{1}{\mu} \), then \( G_p \) has no giant component.
Chung, Lu, Horn [2008]:

- If \( p < \frac{1}{\mu} \), then \( G_p \) has no giant component.

- The condition \( p > \frac{1}{\mu} \) in general does not imply that \( G_p \) has a giant component.
Chung, Lu, Horn [2008]:

- If $p < \frac{1}{\mu}$, then $G_p$ has no giant component.

- The condition $p > \frac{1}{\mu}$ in general does not imply that $G_p$ has a giant component.

- If $p > \frac{1}{\mu}$, $\Delta = O(d)$, and $\sigma = o\left(\frac{1}{\log n}\right)$, then $G_p$ has a giant component.
Bhamidi-van der Hofstad-van Leeuwaarden [2012]: Consider $G(w)$, where $w = (w_1, \ldots, w_n)$ follows the power law of exponent $\beta$. If $\mathbb{E}(\sum_{i=1}^{n} w_i^2)$ converges and is bounded, then the percolation threshold is $(1 + o(1))^{\frac{1}{d}}$.

- For $\beta > 4$, $\mathbb{E}(\sum_{i=1}^{n} w_i^3)$ converges. The largest component has the size $\Theta(n^{2/3})$ at the critical window.
Percolation of $G(w)$

Bhamidi-van der Hofstad-van Leeuwaarden [2012]:
Consider $G(w)$, where $w = (w_1, \ldots, w_n)$ follows the power law of exponent $\beta$. If $E(\sum_{i=1}^{n} w_i^2)$ converges and is bounded, then the percolation threshold is $(1 + o(1))^{\frac{1}{d}}$.

- For $\beta > 4$, $E(\sum_{i=1}^{n} w_i^3)$ converges. The largest component has the size $\Theta(n^{2/3})$ at the critical window.
- For $2 < \beta < 3$, $E(\sum_{i=1}^{n} w_i^3)$ diverges. The largest component has the size $\Theta(n^{\frac{\beta-2}{\beta-1}})$ at the critical window.
References


Homepage: http://www.math.sc.edu/~lu/

Thank You