Topological and Spectral Connectivity in Random Graphs

Nick Rui, Zeyneb Kaya, Hannah Gao

March 23, 2026

Abstract

We study the emergence of connectivity in the Erdős–Rényi random graph model $G(n, p)$, extending the topological perspective on graphs developed in class. We prove that connectivity undergoes a sharp phase transition at the threshold $p = \frac{\ln n}{n}$. We introduce the graph Laplacian and study its second eigenvalue λ_2, the Fiedler value, as a spectral and quantitative measure of connectivity. For connected graphs, we interpret the magnitude of λ_2 through the Rayleigh quotient and minimim degree. Finally, we empirically verify these theoretical frameworks through computational simulations

1 Introduction

In class, we studied graphs as topological and combinatorial objects. This project extends that framework by (1) analyzing the probability of connectnedness in the Erdős–Rényi model and (2) analyzing the eigenvalues of the graph Laplacian as an algebraic measure of connectivity.

1.1 Motivation

We first define graphs, random graphs, and connectedness, drawing from [1].

Definition 1.1. A *graph* $G = (V, E)$ consists of a finite set V of *vertices* and a set E of *edges*, where each edge is an unordered pair of distinct vertices. For this paper, we restrict graphs to have no self-loops or multiple edges. We write $n = |V|$ and $m = |E|$.

Definition 1.2 (Erdős–Rényi Random Graph)**.** The Erdős–Rényi model $G(n, p)$ is a probability distribution on graphs with n labeled vertices in which each of the $\binom{n}{2}$ possible edges is included independently with probability $p \in [0, 1]$. Thus, each graph G on n vertices occurs with probability

$$\mathbb{P}(G) = p^{|E(G)|}(1-p)^{\binom{n}{2}-|E(G)|}.$$

Definition 1.3. A graph G is *connected* if for every pair of vertices $u, v \in V$, there exists a path from u to v. Equivalently, G is connected if the vertex set cannot be partitioned into two nonempty sets with no edges between them. The maximal connected subgraphs of G are called its *connected components*.

Naturally, as p increases we expect samples from $G(n, p)$ to be denser, and so the probability of connectedness increases. To demonstrate this relationship with an initial experiment. For a fixed graph size ($n = 100$), we vary p and construct random Erdős–Rényi graphs $G(n, p)$. We identify if connected graphs and compute $P(\text{connected})$ over m trials (Figure 1). We observe $P(\text{connected})$ increases with p, with a sharp phase transition in the probability of connectivity at a critical p^*.

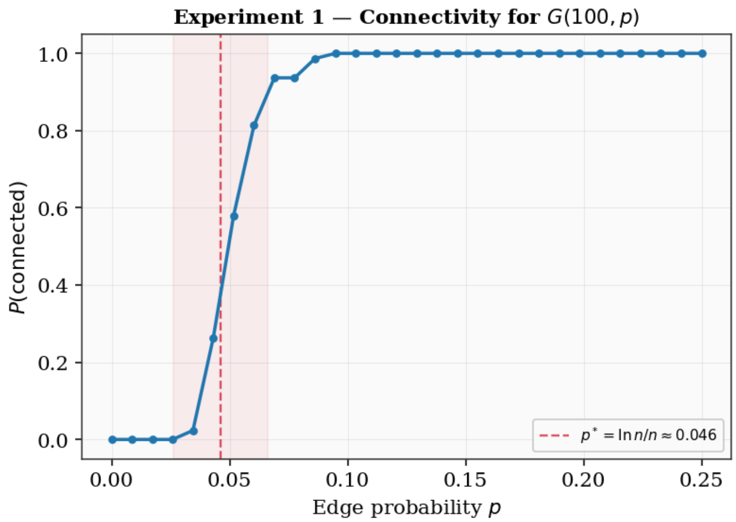

Figure 1: Connectivity of random graphs by p.

2 The Connectivity Threshold

We now prove the existence and precise location of the connectivity threshold for the Erdős–Rényi model. The main result is the following theorem, originally due to Erdős and Rényi [2].

Theorem 2.1 (Connectivity Threshold)**.** *Let $G \sim G(n, p)$ where $p = \frac{\ln n + c(n)}{n}$ for some function $c(n)$.*

(a) If $c(n) \to -\infty$ as $n \to \infty$, then $\mathbb{P}(G$ is connected$) \to 0$.

(b) If $c(n) \to +\infty$ as $n \to \infty$, then $\mathbb{P}(G$ is connected$) \to 1$.

The proof divides into two parts. Below the threshold, we show that isolated vertices appear with high probability, forcing disconnectedness. Above the threshold, we prove that isolated vertices disappear, forcing a high probability of connectivity.

2.1 Preliminary Lemmas

Lemma 2.2. *For all $x \geq 0$ and positive integer n,*

$$(1 - x)^n \leq e^{-nx}. \tag{1}$$

Furthermore, if $0 \le x \le 1$*, then*

$$(1-x)^n \ge e^{-nx-nx^2}. \tag{2}$$

Proof. The first inequality follows from $1 - x \le e^{-x}$ for all $x \ge 0$, verified by noting $f(x) = e^{-x} - (1-x)$ satisfies $f(0) = 0$ and $f'(x) = 1 - e^{-x} \ge 0$.

For the second, use $\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots \ge -x - x^2$ for $0 \le x \le 1$ (since all omitted terms are negative), giving $(1-x)^n = e^{n\ln(1-x)} \ge e^{-nx-nx^2}$. □

Lemma 2.3 (Second Moment Method)**.** *Let* X *be a non-negative integer-valued random variable. If* $\mathbb{E}[X^2]/(\mathbb{E}[X])^2 \to 1$*, then* $\mathbb{P}(X > 0) \to 1$*.*

Proof. By Cauchy-Schwarz, $(\mathbb{E}[X])^2 = (\mathbb{E}[X \cdot \mathbf{1}_{X>0}])^2 \le \mathbb{E}[X^2]\,\mathbb{P}(X > 0)$, giving

$$\mathbb{P}(X > 0) \ge \frac{(\mathbb{E}[X])^2}{\mathbb{E}[X^2]} \to 1.$$ □

2.2 Below the Threshold

Definition 2.4. We say a vertex v in G is *isolated* if $\deg(v) = 0$.

Lemma 2.5. *For* $G \sim G(n,p)$ *with* $p = \frac{\ln n + c}{n}$*, the probability that a fixed vertex* v *is isolated satisfies*

$$\mathbb{P}(v \text{ isolated}) = (1-p)^{n-1} = \frac{e^{-c}}{n}\left(1 + \mathcal{O}\left(\frac{\ln n}{n}\right)\right).$$

Proof. Vertex v is isolated if all $n-1$ potential edges from v are absent, giving $\mathbb{P}(v \text{ isolated}) = (1-p)^{n-1}$. We compute

$$(n-1)p = \left(1 - \frac{1}{n}\right)(\ln n + c) = \ln n + c + \mathcal{O}\left(\frac{\ln n}{n}\right).$$

Since $(n-1)p^2 = \mathcal{O}((\ln n)^2/n) = o(1)$, Lemma 2.2 gives

$$(1-p)^{n-1} = e^{-(n-1)p + \mathcal{O}((n-1)p^2)} = e^{-\ln n - c + \mathcal{O}(\ln n/n)} = \frac{e^{-c}}{n}\left(1 + \mathcal{O}\left(\frac{\ln n}{n}\right)\right).$$ □

Let X count the number of isolated vertices. By linearity of expectation,

$$\mathbb{E}[X] = n \cdot \frac{e^{-c}}{n}\left(1 + \mathcal{O}\left(\frac{\ln n}{n}\right)\right) = e^{-c}\left(1 + \mathcal{O}\left(\frac{\ln n}{n}\right)\right).$$

When $c \to -\infty$, $\mathbb{E}[X] \to \infty$. We formalize this via the second moment method.

Lemma 2.6. *For* $p = \frac{\ln n + c}{n}$ *with* $c \to -\infty$*,* $\mathbb{E}[X^2]/(\mathbb{E}[X])^2 \to 1$*.*

Proof. Let $X_v = \mathbf{1}[v \text{ isolated}]$. Then

$$\mathbb{E}[X^2] = \mathbb{E}[X] + \sum_{v \neq w} \mathbb{P}(v \text{ and } w \text{ both isolated}).$$

For two distinct vertices to both be isolated, the $2(n-1) - 1 = 2n - 3$ potential edges incident to either must all be absent (subtracting 1 for the shared edge $\{v, w\}$). By the same estimates as Lemma 2.5,

$$\mathbb{P}(v, w \text{ both isolated}) = (1-p)^{2n-3} = \frac{e^{-2c}}{n^2}(1 + o(1)).$$

Therefore $\mathbb{E}[X^2] = e^{-c}(1 + o(1)) + e^{-2c}(1 + o(1))$, and

$$\frac{\mathbb{E}[X^2]}{(\mathbb{E}[X])^2} = \frac{e^{-c} + e^{-2c}(1 + o(1))}{e^{-2c}(1 + o(1))} = e^{c}(1 + o(1)) + 1 + o(1) \to 1,$$

since $c \to -\infty$ implies $e^c \to 0$. □

Proof of Theorem 2.1(a). By Lemma 2.6 and the second moment method, $\mathbb{P}(X > 0) \to 1$ when $c \to -\infty$. A graph with an isolated vertex is disconnected, so $\mathbb{P}(G \text{ connected}) \leq \mathbb{P}(X = 0) \to 0$. □

2.3 Above the Threshold

Proposition 2.7. *If* $p = \frac{\ln n + c}{n}$ *with* $c \to +\infty$, *then* $\mathbb{P}(\exists \text{ isolated vertex}) \to 0$.

Proof. By the union bound and Lemma 2.5,

$$\mathbb{P}(\exists \text{ isolated vertex}) \leq n \cdot \frac{e^{-c}}{n}(1 + o(1)) = e^{-c}(1 + o(1)) \to 0.$$ □

Lemma 2.8. *If* $p = \frac{\ln n + c}{n}$ *with* $c \to +\infty$, *then with probability tending to* 1, *G has no connected component of size* $2 \leq k \leq n/2$.

Proof Sketch. Fix $S \subseteq V$ with $|S| = k$. For S to be a component, all $k(n-k) \geq kn/2$ edges between S and $V \setminus S$ must be absent. Using $\binom{n}{k} \leq n^k/k!$ and the estimate $(1-p)^{k(n-k)} \leq e^{-pk(n-k)}$, we have

$$pk(n-k) = \frac{(\ln n + c)k(n-k)}{n} = k \ln n + kc + \mathcal{O}\left(\frac{k^2 \ln n}{n}\right),$$

so $(1-p)^{k(n-k)} \leq n^{-k} e^{-kc}(1 + o(1))$. Taking a union bound over all choices of S,

$$\mathbb{P}(\exists \text{ component of size } k) \leq \frac{n^k}{k!} \cdot n^{-k} e^{-kc}(1 + o(1)) = \frac{e^{-kc}}{k!}(1 + o(1)).$$

For any fixed $k \geq 2$ and $c \to +\infty$, this tends to 0 since $e^{-kc} \to 0$. Summing over fixed $k \in \{2, \ldots, K\}$ for any constant K gives total probability tending to 0. Handling components of size growing with n requires a separate argument; we omit this and refer the reader to [2] for a complete treatment. □

2.4 Proof of Connectivity Threshold

Proof of Theorem 2.1(b). By Proposition 2.7, the minimum degree is at least 1 with high probability. By Lemma 2.8, there are no components of size 2 to $n/2$ with high probability. A disconnected graph with minimum degree at least 1 must contain a component of size between 2 and $n/2$. Since this occurs with probability tending to 0, we conclude $\mathbb{P}(G \text{ connected}) \to 1$. □

Experimentally, we verify this by computing $P(\text{connected})$ for sweeps over p with various sizes n to observe the connectivity thresholds $p^* = \ln(n)/n$ (Figure 2) We can see that the phase transition of connectedness probability aligns with the theoretical threshold.

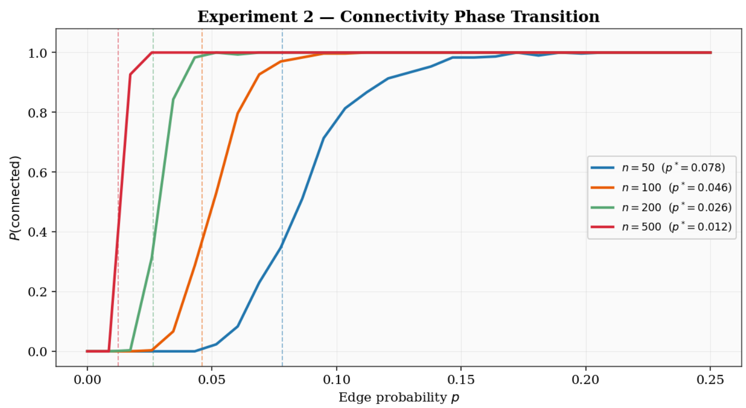

Figure 2: Critical thresholds $p^* = \ln(n)/n$ in connectivity phase transitions for various n.

3 Spectral Properties of the Graph Laplacian

For large n, we've seen that most values of p will have a near-1 probability of connectedness since $\ln(n)/n$ is small. So, nothing interesting occurs when we analyze the probability of connectedness past p^* since it just stays near 1. However, intuitively these graphs should get *more connected* as p increases. In this section, we investigate how we can characterize the strength of graph connectedness.

3.1 The Graph Laplacian

We take an algebraic approach by analyzing the eigenvalues of the graph Laplacian. The following ideas were drawn from [3].

Definition 3.1. Let $G = (V, E)$ be a graph with vertices indexed $1, \dots, n$. The *adjacency matrix* $A \in \mathbb{R}^{n\times n}$ has entries $A_{ij} = 1$ if $\{i, j\} \in E$ and $A_{ij} = 0$ otherwise. The *degree matrix* $D \in \mathbb{R}^{n\times n}$ is diagonal with $D_{ii} = \deg(i)$. The *graph Laplacian* is a matrix defined as

$$L = D - A.$$

An intuitive interpretation of the graph Laplacian comes from analyzing its quadratic form. First, imagine labeling each vertex a real number value. This labeling corresponds with a vector $\mathbf{x} \in \mathbb{R}^n$ where x_i is the label of vertex i. The quadratic form $\mathbf{x}^T L\mathbf{x}$ is then

$$\begin{aligned} \mathbf{x}^T L\mathbf{x} &= \sum_i x_i(D_{ii}x_i) - \sum_{i,j} x_i A_{ij} x_j \\ &= \sum_i \deg(i)\, x_i^2 - 2 \sum_{\{i,j\}\in E} x_i x_j \\ &= \sum_{\{i,j\}\in E} (x_i - x_j)^2. \end{aligned} \tag{3}$$

This shows the Laplacian measures the *total squared variation* of the labeling $\mathbf{x}$ across edges.

Note L is symmetric, and (3) shows L is positive semidefinite. So, we know L has n real non-negative eigenvalues, which we order as

$$0 = \lambda_1 \le \lambda_2 \le \cdots \le \lambda_n.$$

It's easy to see that $\lambda_1 = 0$. This is because $L\mathbf{1} = \mathbf{0}$, summing along the i-th row gives $\deg(i) + \sum_{i=1}^{\deg(i)}(-1) = 0$. The corresponding eigenspace to λ_1 is the space of constant vectors, or $\operatorname{span}(\mathbf{1})$. This agrees with our interpretation of the Laplacian as measuring variation, since labeling each vertex with the same constant results in zero variation.

3.2 Fiedler Value λ_2 and Interpretation

The second eigenvalue λ_2, also called the *Fiedler value* of G carries important properties.

Theorem 3.2. *The multiplicity of the eigenvalue* 0 *in the spectrum of* L *equals the number of connected components of* G*. In particular,* G *is connected if and only if* $\lambda_2 > 0$*.*

Proof. Recall that the multiplicity of the eigenvalue 0 is also the nullity $\dim \operatorname{Ker}(L)$. We prove this equals the number of connected components by proving inequalities both ways.

Let G have connected components $C_1, \dots, C_k$. For each C_i, define an assignment of scalars on the vertices $\mathbf{v}^{(i)} \in \mathbb{R}^n$ by $v_j^{(i)} = 1$ if $j \in C_i$ and $v_j^{(i)} = 0$ otherwise.

We claim $L\mathbf{v}^{(i)} = \mathbf{0}$. For any vertex $j \in C_i$, all neighbors of j lie in C_i, so

$$(L\mathbf{v}^{(i)})_j = \deg(j) - \sum_{\ell\sim j} 1 = 0.$$

For $j \notin C_i$, no neighbor of j lies in C_i, so $(L\mathbf{v}^{(i)})_j = 0 - 0 = 0$. Thus each $\mathbf{v}^{(i)} \in \operatorname{Ker}(L)$.

The vectors $\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(k)}$ are linearly independent (their supports are disjoint), so $\dim \operatorname{Ker}(L) \ge k$.

For the reverse inequality, suppose $\mathbf{v} \in \operatorname{Ker}(L)$. Then $0 = \mathbf{v}^T L\mathbf{v} = \sum_{\{i,j\}\in E}(v_i - v_j)^2$, forcing $v_i = v_j$ for every edge $\{i,j\}$. By transitivity along paths, $\mathbf{v}$ is constant on each connected component. Therefore every element of $\operatorname{Ker}(L)$ is a linear combination of $\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(k)}$, giving $\dim \operatorname{Ker}(L) \le k$.

Hence $\dim \operatorname{Ker}(L) = k$. The final statement follows: $\lambda_2 > 0 \iff k = 1 \iff G$ is connected. □

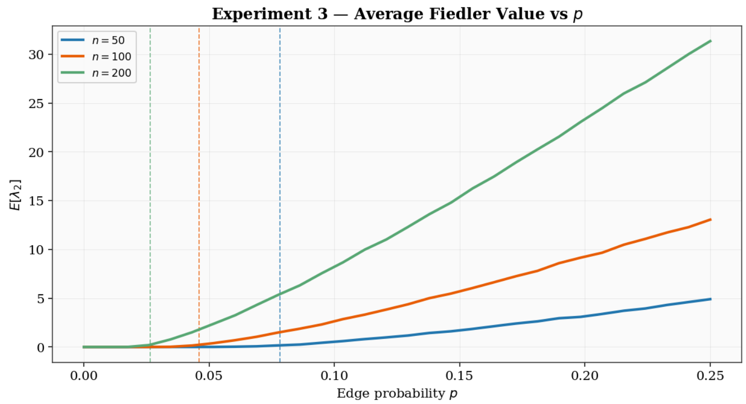

Figure 3: λ_2 and connectivity with p.

Theorem 3.2 establishes that λ_2 is a binary detector of connectivity. But when $\lambda_2 > 0$, its magnitude quantifies *how* connected the graph is. Since L is symmetric, its eigenvectors form an orthogonal basis, so restricting vertex labelings to $\mathbf{x} \perp \mathbf{1}$ yields the following characterization of λ_2.

Theorem 3.3 (Rayleigh Quotient Characterization)**.**

$$\lambda_2 = \min_{\substack{\mathbf{x}\perp\mathbf{1} \\ \mathbf{x}\neq\mathbf{0}}} \frac{\mathbf{x}^T L\mathbf{x}}{\|\mathbf{x}\|^2} = \min_{\substack{\mathbf{x}\perp\mathbf{1} \\ \|\mathbf{x}\|=1}} \sum_{\{i,j\}\in E} (x_i - x_j)^2.$$

Proof. By the spectral theorem, L has an orthonormal eigenbasis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ with $\mathbf{v}_1 = \mathbf{1}/\sqrt{n}$ corresponding to $\lambda_1 = 0$. Any $\mathbf{x} \perp \mathbf{1}$ expands as $\mathbf{x} = \sum_{i=2}^{n} c_i \mathbf{v}_i$, giving

$$\frac{\mathbf{x}^T L\mathbf{x}}{\|\mathbf{x}\|^2} = \frac{\sum_{i=2}^{n} c_i^2 \lambda_i}{\sum_{i=2}^{n} c_i^2} \geq \lambda_2,$$

with equality when $\mathbf{x} = \mathbf{v}_2$. The second expression follows from (3). □

Theorem 3.3 gives an intuitive picture of connectivity: λ_2 is large only when *every* zero-mean labeling $\mathbf{x}$ has high edge variation, meaning no cut can separate the graph into regions with very different labels. Conversely, if λ_2 is small, there exists a labeling with low variation across edges yet high label energy — revealing a bottleneck or near-cut in the graph. Thus λ_2 measures how difficult it is to "pull apart" the graph, and is called the *algebraic connectivity* of G.

We demonstrate experimentally the relationship between p and λ_2 in the Erdős–Rényi model. We sweeps over p for various sizes n and compute the mean λ_2 at each configuration (Figure 3). We observe that prior to the critical threshold, λ_2 is at or near 0, indicating that graphs are disconnected; past the critical threshold, $\lambda_2 > 0$ and increasing, indicating connectivity. We observe a roughly linear trend, with slope increasing with n.

Theorem 3.3 also gives a convenient proof strategy: to show $\lambda_2 \leq c$, it suffices to exhibit a single *test vector* $\mathbf{x} \perp \mathbf{1}$ whose Rayleigh quotient is at most c. One application of this idea gives a simple bound of λ_2 in terms of the minimum degree of any vertex in the graph, which is a much simpler graph property much easier to digest.

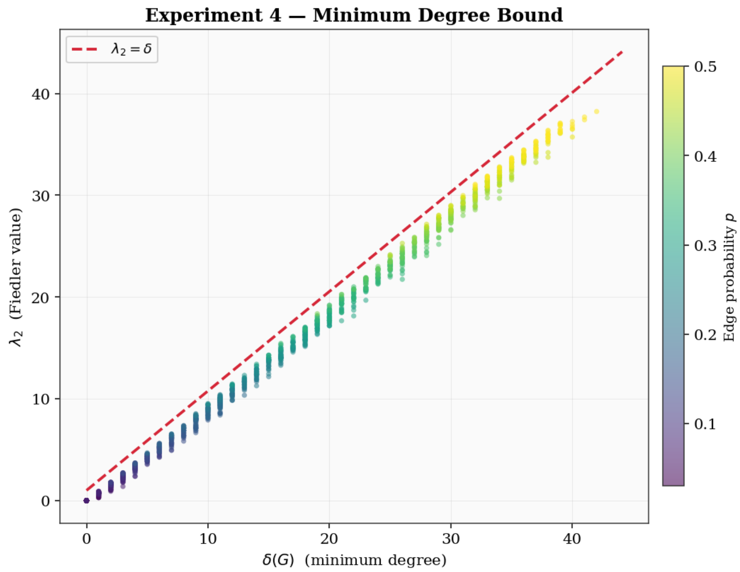

Figure 4: Minimum degree and λ_2 of random graphs with varying p.

Theorem 3.4. *For any graph G, $\lambda_2 \le \delta(G) + 1$ where $\delta(G) = \min_{v \in V(G)} \deg(v)$.*

Proof. Let v be a vertex with $\deg(v) = \delta$ and neighbors $N(v)$. Define the test vector $\mathbf{x}$ as

$$x_i = \begin{cases} \delta & \text{if } i = v \\ -1 & \text{if } i \in N(v) \\ 0 & \text{otherwise.} \end{cases}$$

Then $\sum_i x_i = \delta - \delta = 0$, so $\mathbf{x} \perp \mathbf{1}$. The only edges contributing to $\mathbf{x}^T L \mathbf{x} = \sum_{\{i,j\}} (x_i - x_j)^2$ are those incident to v. For each $\{v, u\}$ with $u \in N(v)$, $(x_v - x_u)^2 = (\delta + 1)^2$, so

$$\mathbf{x}^T L \mathbf{x} = \delta(\delta + 1)^2, \qquad \|\mathbf{x}\|^2 = \delta^2 + \delta = \delta(\delta + 1).$$

By Theorem 3.3, $\lambda_2 \le \delta + 1$. ☐

Remark 3.5. Fiedler [4] showed $\lambda_2 \le \delta(G)$ using a similar but more involved test vector strategy.

Experimentally, we observe the bound on λ_2 and show that it serves as an indicator of 'how connected' a graph is and relates to its minimum degree. We generate random graphs of with fixed size ($n = 100$) and varying p and compute the minimum degree and λ_2 for each (Figure 4). We observe that the minimum degree serves as an upper bound for λ_2 and that λ_2 in general is close to the minimum degree. We can also see that increasing p yields graphs with greater minimum degrees. Further experimental details are provided in Section 4.

4 Experimental Details

We provide additional implementation details on the empirical experiments and results presented.

To sample an Erdős–Rényi random graph $G \sim G(n, p)$, we iterate over all $\binom{n}{2}$ pairs of vertices and include each edge independently with probability p using a uniform random draw. We represent a graph by its number of vertices n and a set of edges containing tuples indicating endpoints of edges in the graph, as well as construct its adjacency list. Connectivity is checked via breadth-first search from vertex 0: if all n vertices are visited, the graph is connected; otherwise, the graph is not connected.

To compute spectral quantities, we construct the Laplacian $L = D - A$ as an $n \times n$ matrix by constructing the adjacency matrix and the diagonal matrix based on the degrees of vertices. We find its eigenvalues using `numpy.linalg.eigvalsh`. The Fiedler value λ_2 is then the second entry of the sorted eigenvalue array. The minimum degree $\delta(G)$ is read directly from the adjacency list.

Experiment 1: Connectivity of random graphs. In Experiment 1, we use $n = 100$ and generate $m = 300$ trials of random graphs for each of 30 p from 0 to 0.25. We determine whether each graph is connected and compute $P(\text{connected})$ as the fraction of the $m = 300$ trials in which $G(n, p)$ is connected.

Experiment 2: Critical connectivity phase transition threshold. In Experiment 2, we follow a similar procedure to Experiment 1 for $n \in \{50, 100, 200, 500\}$. We indicate the theoretical critical $p^* = \frac{\ln n}{n}$ for each n as well.

Experiment 3: Spectral properties and connectivity from Fiedler values. In Experiment 3, we follow a similar procedure to Experiment 2 in constructing random graphs for $n \in \{50, 100, 200\}$. For each $G(n, p)$ we compute λ_2 for each (the second-smallest eigenvalue of $L = D - A$), and average across trials.

Experiment 4: Spectral properties and bounding minimum degree upper bound of Fiedler values. In Experiment 4, We generate 1500 random graphs of fixed size $n = 100$ at 30 different densities $p \in (0.03, 0.5]$. For each graph we compute λ_2 and the minimum degree $\delta(G)$. We plot along with the line $\lambda_2 = \delta$ as the theoretical upper bound.

AI Acknowledgement

We used Claude to (1) help format in LaTeX, (2) help format and draft code, (3) help write matplotlib code for experiment visualizations, (4) proofread notes and presentation slides.

References

[1] Alan Frieze and Michał Karoński. *Introduction to Random Graphs*. Cambridge University Press, 2016.

[2] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5(1):17–61, 1960.

[3] Fan R. K. Chung. *Spectral Graph Theory*, volume 92 of *CBMS Regional Conference Series in Mathematics*. American Mathematical Society, 1997.

[4] Miroslav Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2):298–305, 1973.