Spectral Analysis of Erdős–Rényi Graphs Using the Lanczos Algorithm

Harris Bubalo

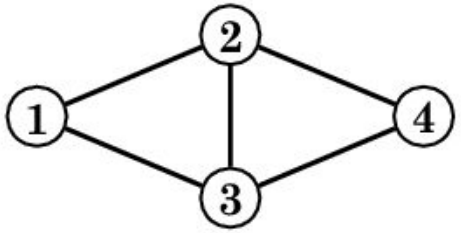
MATH 7203 Numerical Analysis

February 26, 2026

Graph Theory Basics

- A graph G is said to be **connected** if for all node pairs (u, v), there exists a path from u to v
- In an undirected graph, the **degree** of a node is the number of edges incident to the node
- The **adjacency matrix** (A) is a symmetric matrix where element (i, j) is 1 if node i and node j share an edge and 0 otherwise
- The **degree matrix** (D) is a diagonal matrix formed from the degrees of each node
- The **Laplacian** is a symmetric positive semidefinite matrix defined as $L = D - A$

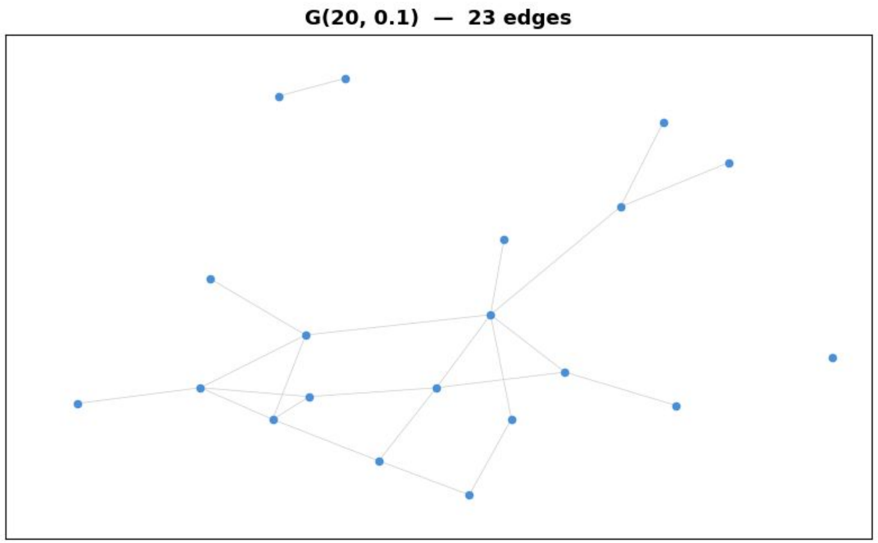
$$D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

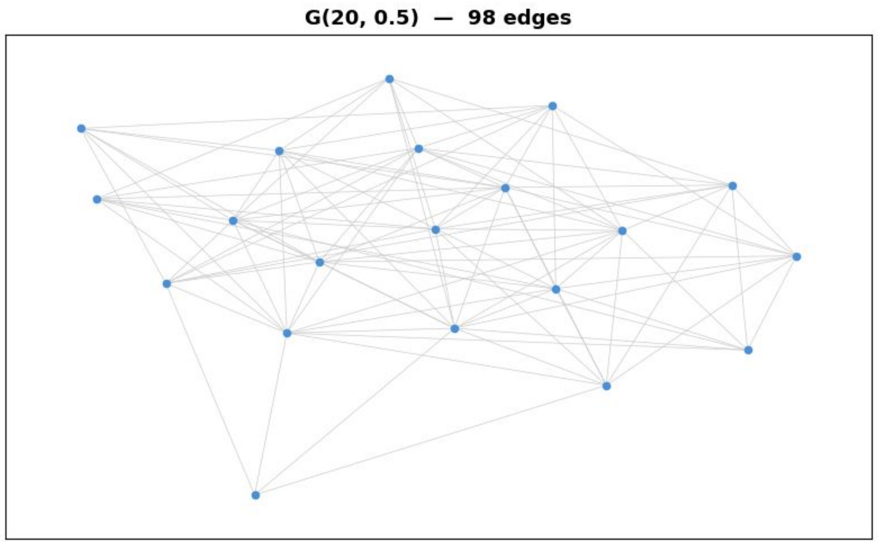


$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \qquad L = \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ 0 & -1 & -1 & 2 \end{bmatrix}$$

Erdős–Rényi Graphs

- The **Erdős–Rényi model** describes a model for generating random graphs
- $G(n, p)$ is a graph of n nodes, where each of the $n(n-1)/2$ possible edges has an independent probability of p of existing in the graph
- When can we expect $G(n, p)$ to be connected? How "connected" will it be?
- Real-world application: **Network robustness**





The Algebraic Connectivity

- Each row of the Laplacian sums to zero, so λ_1 will always be 0 and $(1,1,\ldots,1)^T$ is the corresponding eigenvector (constant value for each node)
- **λ_2 (Fiedler Value)** is the second smallest eigenvalue of the Laplacian… when is it 0?
- $Lx = 0 \Rightarrow x^T Lx = 0$
  - If the graph is **connected**, x must have equal values for all nodes $\Rightarrow$ **$\lambda_2 > 0$**
  - If the graph is **disconnected**, we can create an eigenvector that is constant for all nodes in a component and zero everywhere else $\Rightarrow$ **$\lambda_2 = 0$**
- Rayleigh quotient is monotonic, so λ_2 also shows robustness of connection

$$x^T L x = \sum_{\{i,j\} \in E} (x_i - x_j)^2 \qquad R(L, x) = \frac{\sum_{\{i,j\} \in E} (x_i - x_j)^2}{\sum_{i=1}^{n} x_i^2}$$

The Arnoldi Iteration

- A **Krylov subspace** is defined as

$$\mathcal{K}_k(A, b) = \text{span}\{b, Ab, A^2b, \ldots, A^{k-1}b\}$$

- **Goal:** For an n x n Laplacian L, find $L = QHQ^T$, where H is upper Hessenberg and Q is orthogonal
- n can be huge for large graphs, so consider just the $m << n$ columns of $LQ = QH$, starting from an arbitrary unit vector q_1

$$L \begin{bmatrix} | & & | \\ q_1 & \cdots & q_m \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ q_1 & \cdots & q_{m+1} \\ | & & | \end{bmatrix} \begin{bmatrix} h_{11} & \cdots & h_{1m} \\ h_{21} & \cdots & h_{2m} \\ & \ddots & \vdots \\ & & h_{m+1,m} \end{bmatrix}$$

$$Lq_m = h_{1,m}q_1 + \cdots + h_{m,m}q_m + h_{m+1,m}q_{m+1}$$

$$\boxed{q_{m+1} = (Lq_m - h_{1m}q_1 - \cdots - h_{mm}q_m)/h_{m+1,m}}$$

- Arnoldi iteration is analogous to the **Gram-Schmidt** process ⇒ Constructs the orthonormal basis for the Krylov subspace

$$\mathcal{K}_m(L, b) = \text{span}\{b, Lb, L^2b, \ldots, L^{m-1}b\} = \text{span}\{q_1, q_2, \ldots q_m\}$$

How does it find eigenvalues?

- Compute the eigenvalues of $H_m = Q_m^T L Q_m$ to find m approximations (**Ritz values**)
- **Theorem:** The characteristic polynomial of H_m is the unique solution of

$$||p(A)b|| = \text{minimum}$$

- We can write any vector from the Krylov subspace as

$$\begin{aligned} p(L)b &= (c_0 I + c_1 L + c_2 L^2 + \cdots + L^m)b \\ &= \sum_{j=1}^{n} \alpha_j (c_0 I + c_1 L + c_2 L^2 + \cdots + L^m) v_j \\ &= \sum_{j=1}^{n} \alpha_j (c_0 + c_1 \lambda_j + c_2 \lambda_j^2 + \cdots + \lambda_j^m) v_j \\ &= \sum_{j=1}^{n} \alpha_j p(\lambda_j) v_j \end{aligned}$$

- p^* has the true eigenvalues as its roots when $m = n$ since $||p^*(L)b|| = 0$
- Ritz values after $m << n$ iterations approximate “extreme” eigenvalues

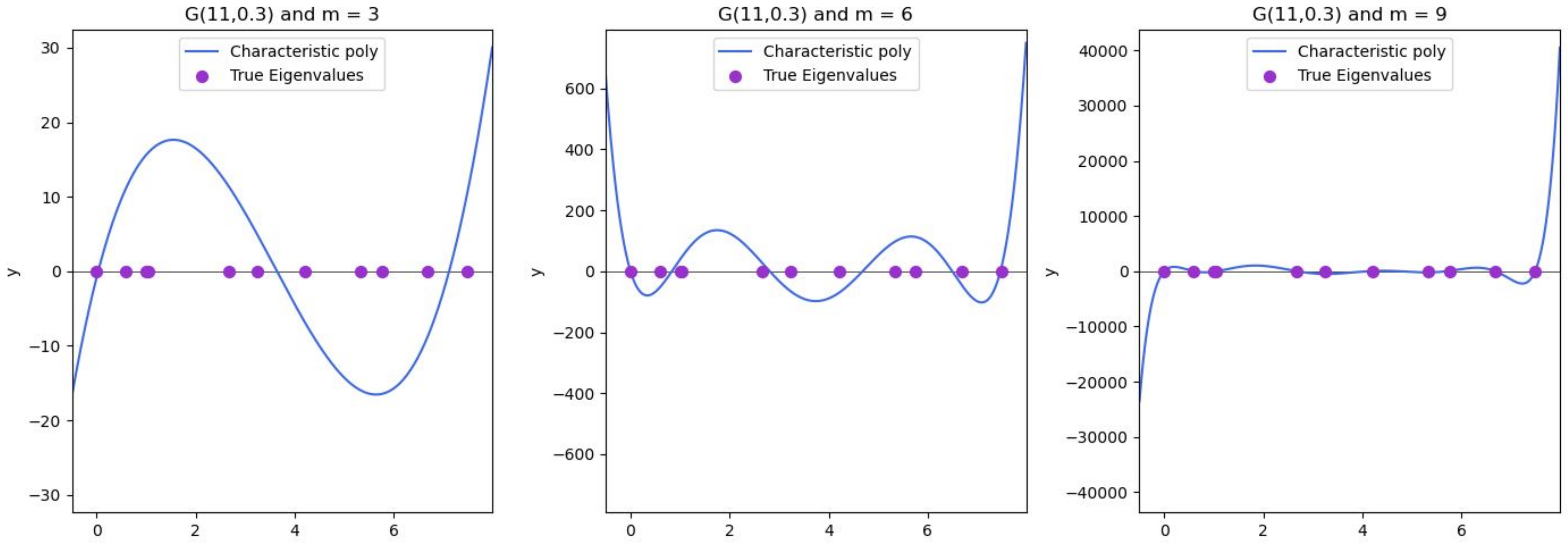
The Lanczos Algorithm

- Special case of the Arnoldi Iteration where the matrix is **symmetric**
- Instead of the upper Hessenberg matrix H_m, we end up with tridiagonal matrix T_m
- Consequently, we only need to do a three term recurrence, rather than a full reorthogonalization

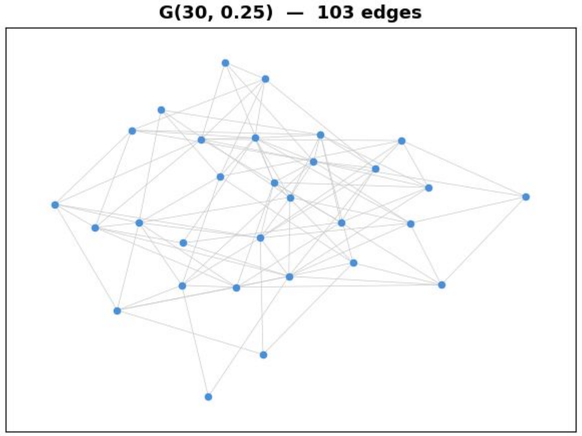
$$T_m = \begin{bmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \beta_2 & \alpha_3 & \ddots & \\ & & \ddots & \ddots & \beta_{m-1} \\ & & & \beta_{m-1} & \alpha_m \end{bmatrix}$$

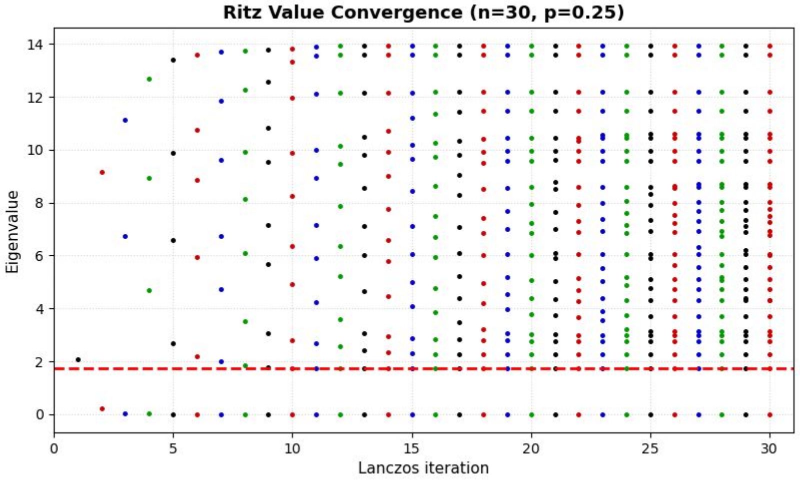
1. $\beta_0 = 0, \; q_0 = 0$
2. Choose b arbitrarily, then $q_1 = \dfrac{b}{\|b\|_2}$
3. **for** $m = 1, 2, 3, \ldots$ **do**
   (a) $v = Lq_m$
   (b) $\alpha_m = q_m^T v$
   (c) $v = v - \beta_{m-1} q_{m-1} - \alpha_m q_m$
   (d) $\beta_m = \|v\|_2$
   (e) $q_{m+1} = \dfrac{v}{\beta_m}$

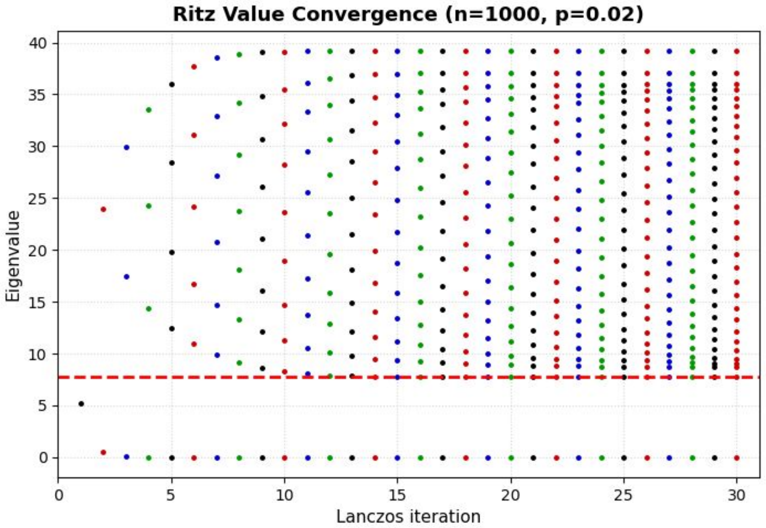
Results

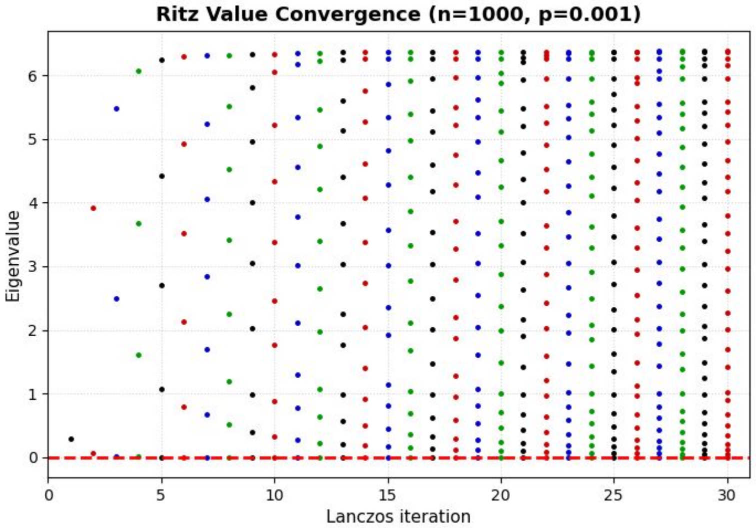

G(11,0.3) and m = 3
G(11,0.3) and m = 6
G(11,0.3) and m = 9
Characteristic poly
True Eigenvalues
y
30
20
10
0
−10
−20
−30
600
400
200
−200
−400
−600
40000
30000
20000
10000
−10000
−20000
−30000
−40000
0
2
4
6

Results (2)

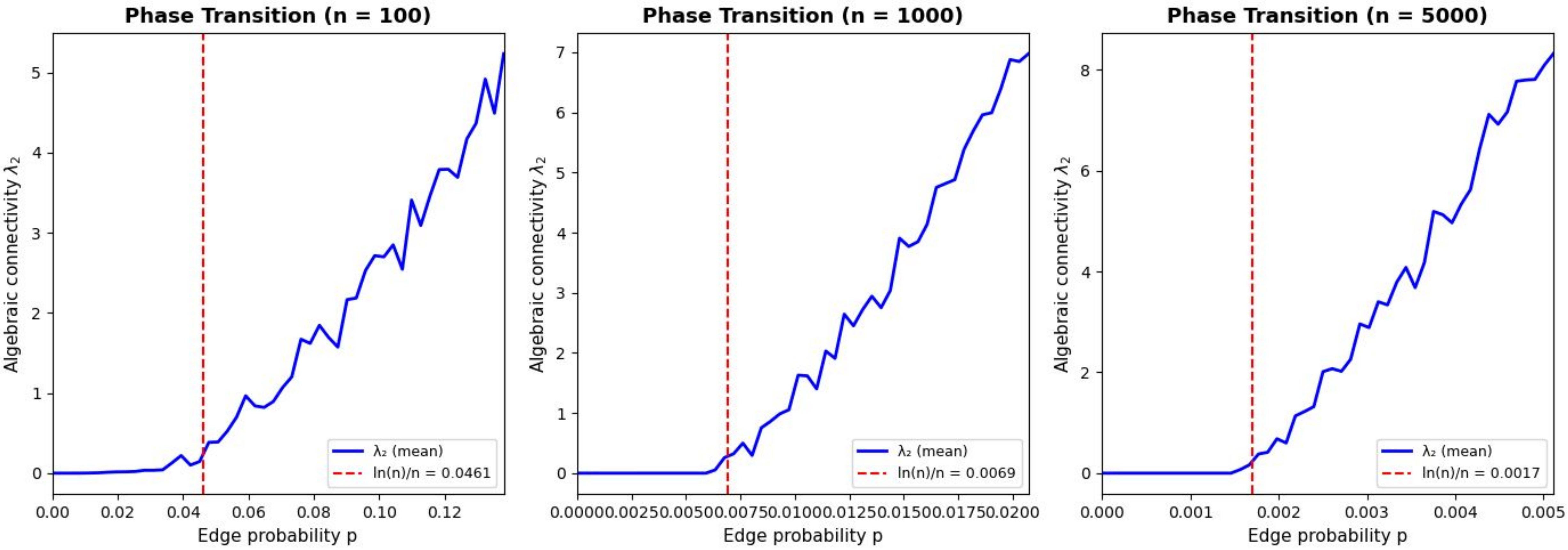

G(30, 0.25) — 103 edges


Ritz Value Convergence (n=30, p=0.25)
Eigenvalue
Lanczos iteration
0
2
4
6
8
10
12
14
0
5
10
15
20
25
30


Ritz Value Convergence (n=1000, p=0.02)
Eigenvalue
Lanczos iteration
0
5
10
15
20
25
30
35
40
0
5
10
15
20
25
30


Ritz Value Convergence (n=1000, p=0.001)
Eigenvalue
Lanczos iteration
0
1
2
3
4
5
6
0
5
10
15
20
25
30

Results (3)



Phase Transition (n = 100)
Phase Transition (n = 1000)
Phase Transition (n = 5000)
Algebraic connectivity λ_2
Edge probability p
λ₂ (mean)
ln(n)/n = 0.0461
ln(n)/n = 0.0069
ln(n)/n = 0.0017
0.00 0.02 0.04 0.06 0.08 0.10 0.12
0.00000.00250.00500.00750.01000.01250.01500.01750.0200
0.000 0.001 0.002 0.003 0.004 0.005

Thank you!