Network models

Seminars in Social Networks and Markets

Generative models of networks

- Can we have mathematical models that describe how a network "came into existence"?
- They may help us understand **network formation** processes
- Can be used to **forecast** interesting measures, such as:
  - Degree distribution
  - Diameter
  - Component sizes
  - Clustering coefficient
  - etc.

Observed measures of real social networks

- Average degree is **small**: $O((\log n)^r)$
- Degree distribution is a **power law**: $\Theta(k^{-\alpha})$
- (Effective) diameter is **small**: $O((\log n)^{q)}$
- One **giant** component with $\Theta(n)$ nodes
- Clustering coefficient is **large**: $\geq$ constant

Random graph ensembles

- A **probability distribution** over a set of graphs

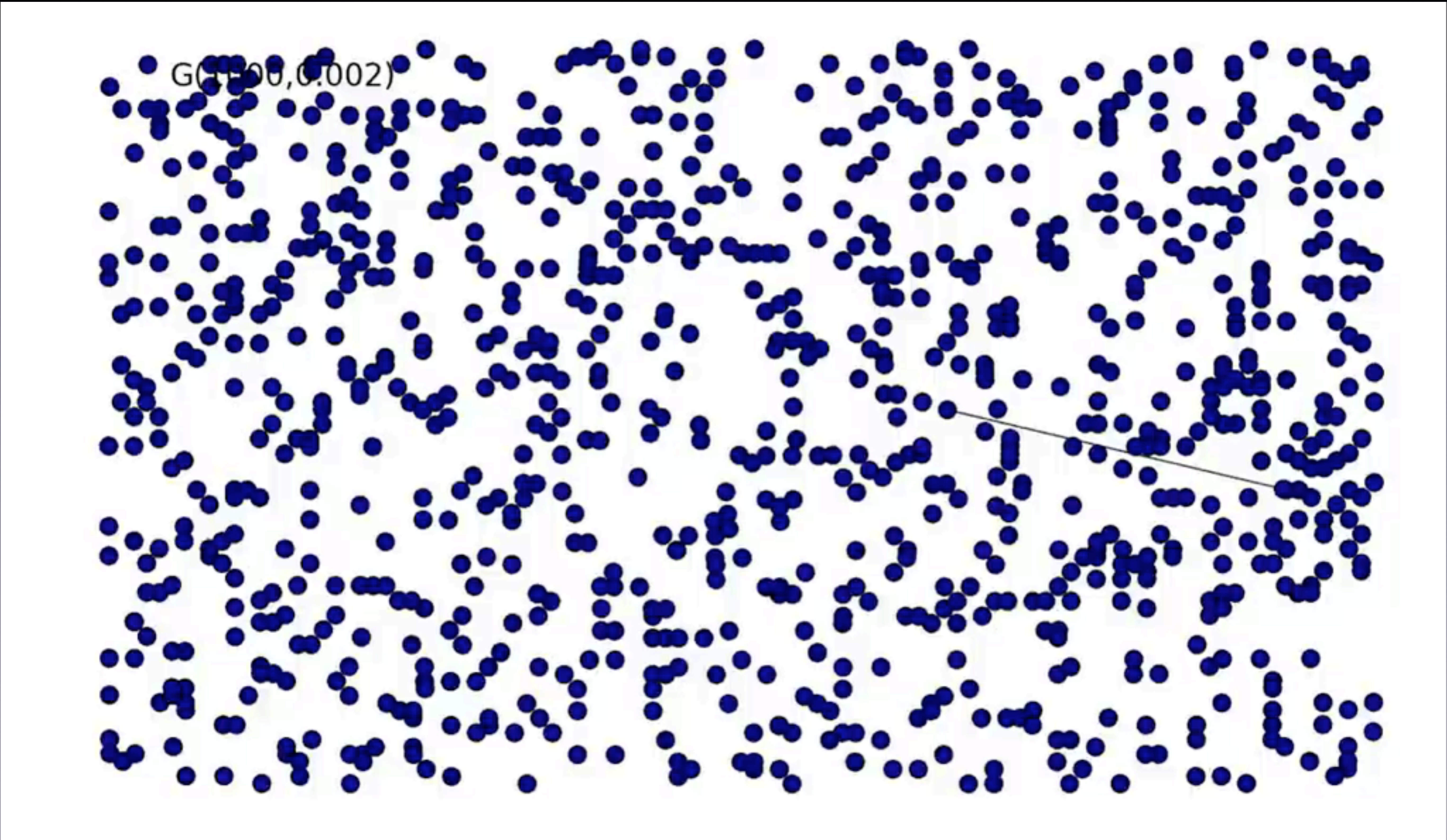
The Erdős-Rényi (ER) model, version 1

- *G(n, m)* model:
  - *n* is the number of nodes
  - *m* is the number of edges
  - Each graph with *n* nodes and *m* edges is assigned **equal probability**
- Natural, but not always easy to work with

The Erdős-Rényi (ER) model, version 2

- *G(n, p)* model:
  - *n* is the number of nodes
  - *p* is a parameter between 0 and 1
  - Each pair of nodes is **joined with independent probability** *p*

The evolution of *G(n,p)*



Properties of *G(n,p)*

✓ Expected average degree: $\mathbf{E}[c] = p\cdot(n\text{-}1)$

Reasonable, as long as p is (say) $O((\log n)/n)$

✗ Degree distribution is **Poisson**: $p_k = c^k\, e^{-c} / k!$

✗ Expected clustering coefficient: $p = O(c/n)$

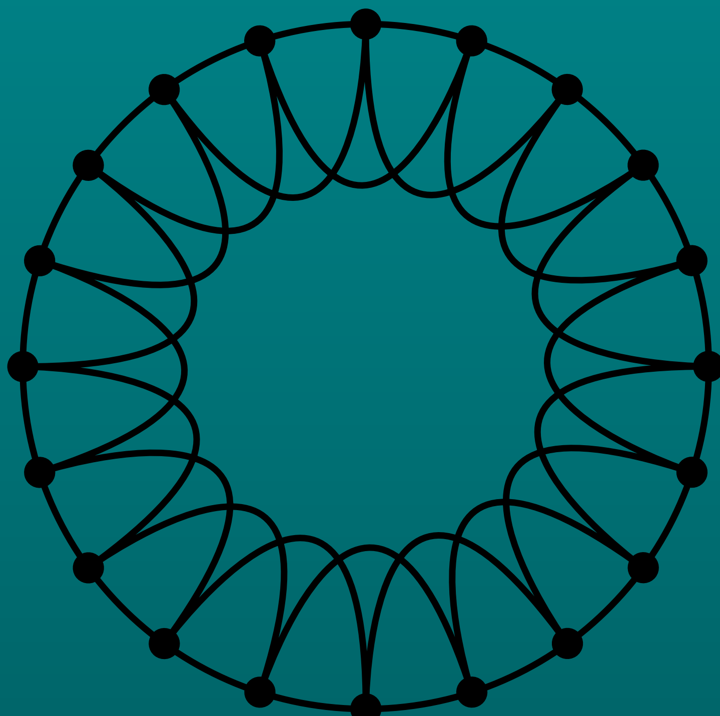
✓ Giant component size: $\Theta(n)$ if $p \geq (1+\varepsilon)/n$

✓ Average shortest distance: $O(\log n / \log c)$

The Watts-Strogatz “small-world” model

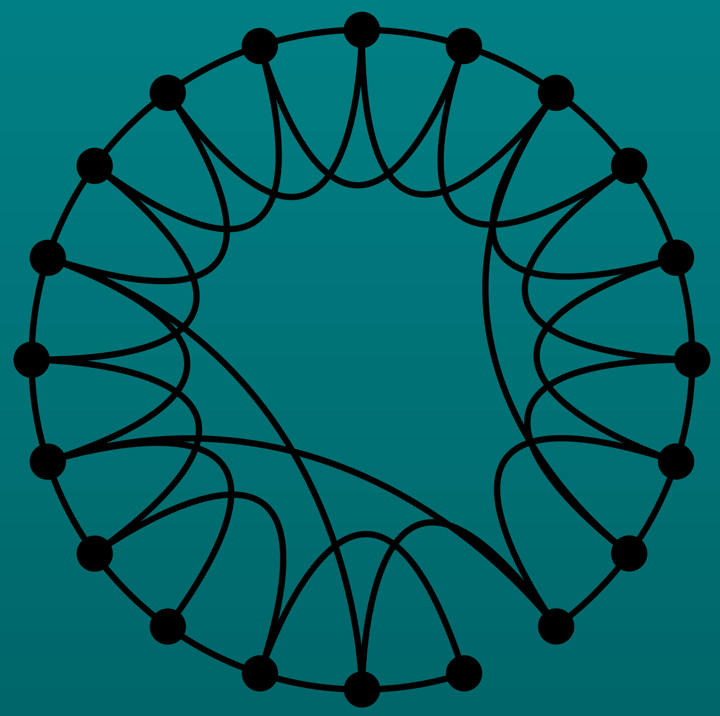
- Can we model **large** transitivity (clustering coefficient) and **small** average distances?

1. Start with a ring lattice with n nodes and degree c

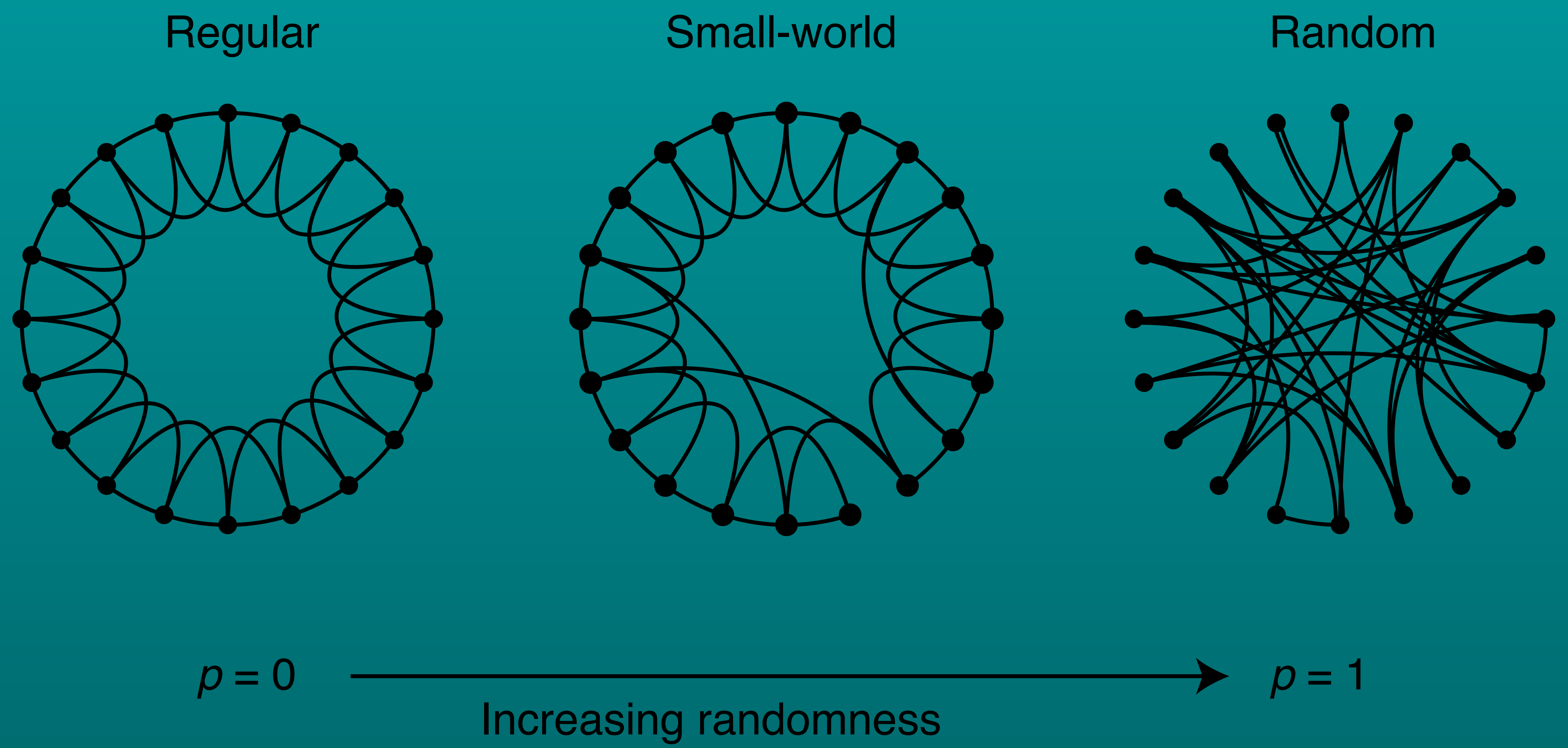


The Watts-Strogatz “small-world” model

2. **Rewire** randomly each edge with probability *p*

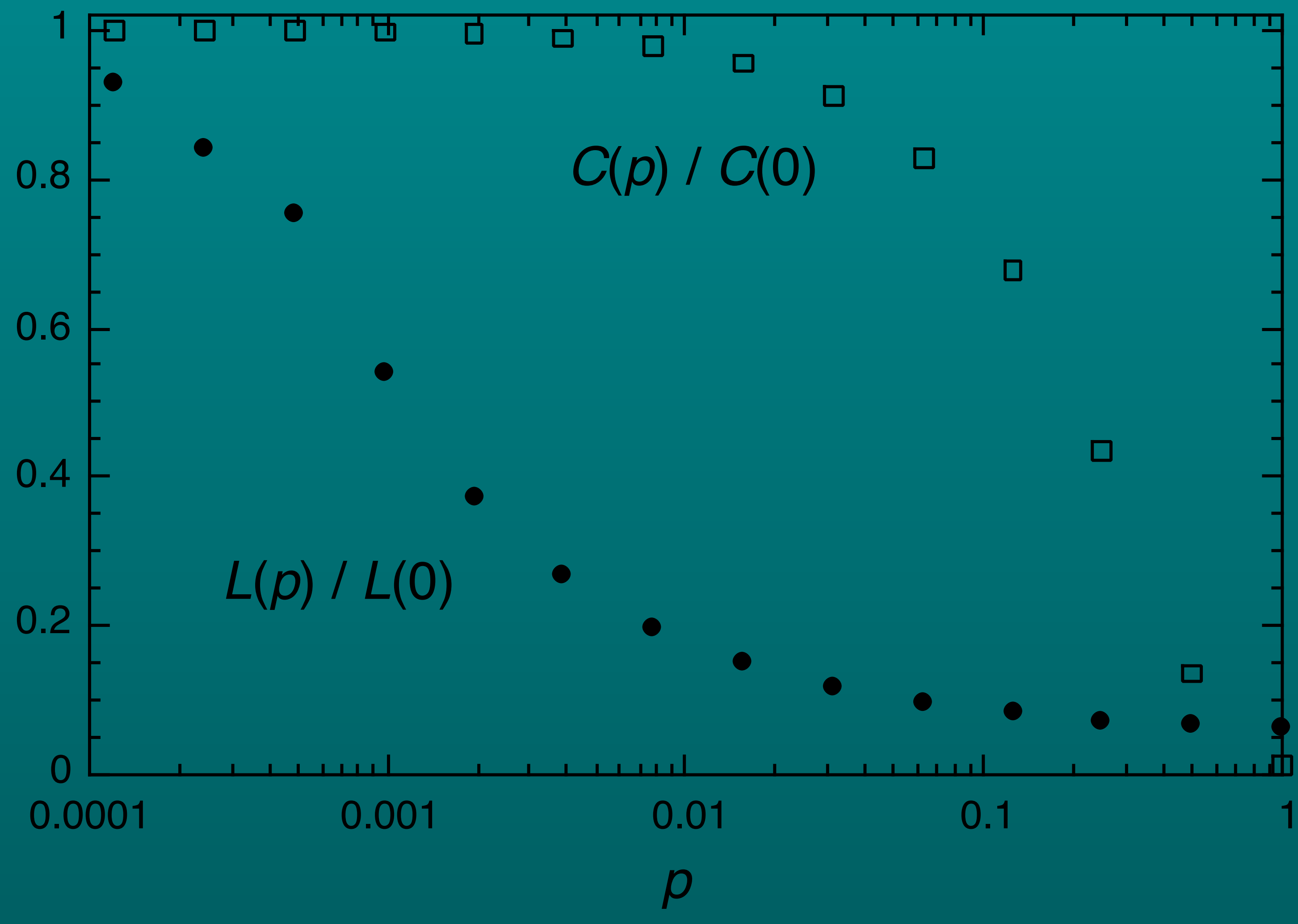


The Watts-Strogatz "small-world" model



The Watts-Strogatz "small-world" model

- L(p) = average path length
- C(p) = clustering coefficient



The Barabási-Albert Evolving Network model

- What about a simple network model yielding a **power-law** degree distribution?

The Barabási-Albert Evolving Network model

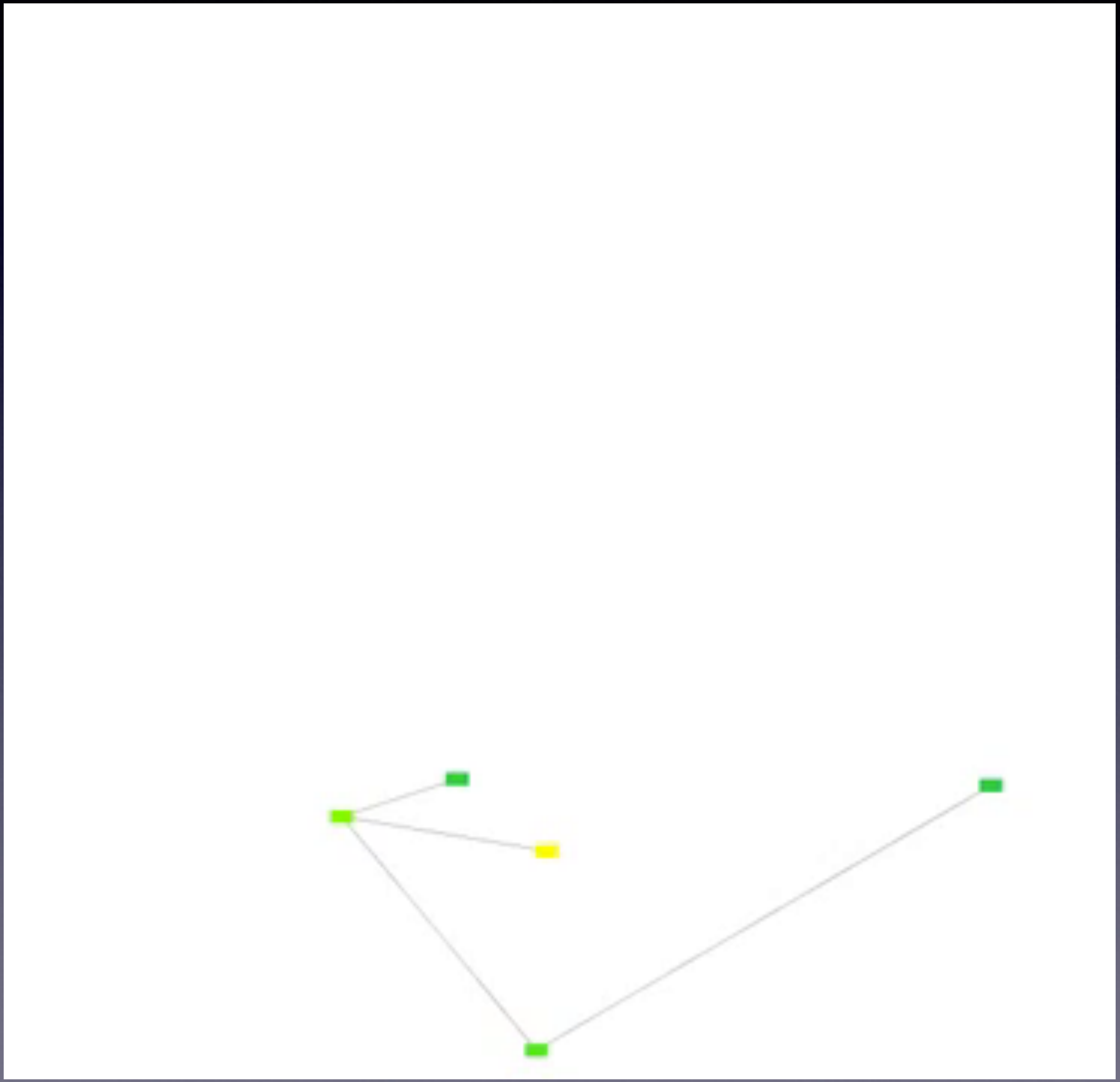
1. Nodes are created in order: 1, 2, …, n

2. When node j is created, it will link to K previous nodes:

   - the probability that it links to node i is **proportional to the degree of** i (**preferential attachment** rule, or "**rich-get-richer**")

- The model exhibits a **power-law** degree distribution (as observed in many real networks)

The Barabási-Albert Evolving Network model



Another rich-get-richer model: the Copying model

1. Nodes are created in order: 1, 2, …, n

2. When node *j* is created:

   A. With prob. *p*, node *j* links to a node $i < j$ uniformly at random

   B. With prob. *1-p*, node *j* selects a node $i < j$ uniformly at random **and copies its (only) outgoing link**

Optional: step 2 can be repeated to obtain higher out-degree (but we'll assume out-degree 1)

Copying model: equivalent description

1. Nodes are created in order: 1, 2, …, n
2. When node *j* is created:
   A. With prob. *p*, node *j* links to a node $i < j$ uniformly at random
   B. With prob. *1-p*, node *j* links to a node $\ell < j$ **with probability proportional to the in-degree of** ℓ

- This model also exhibits a **power-law** degree distribution