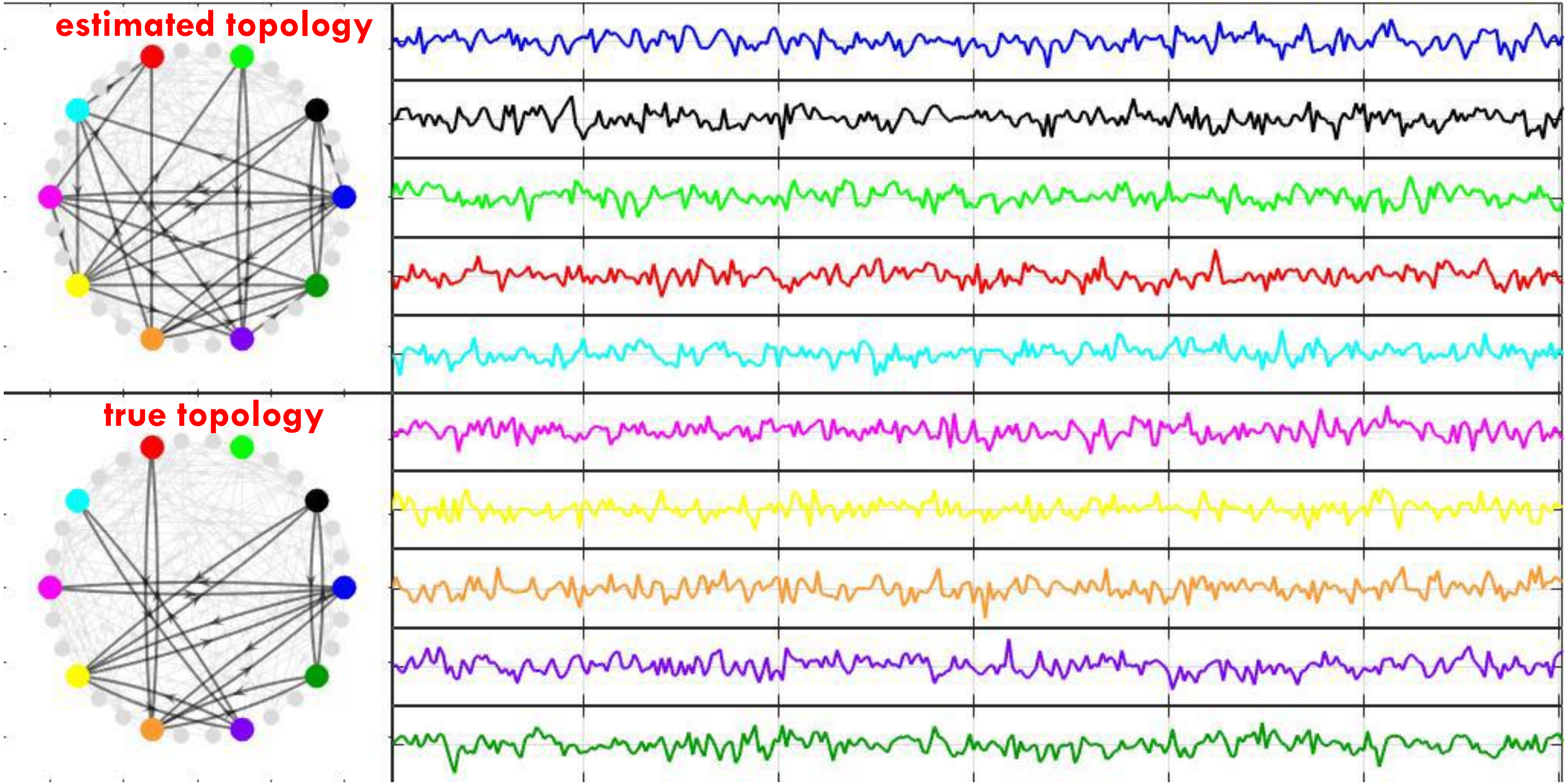
Learning Bollobás-Riordan Graphs under Partial Observability

Michele Cirillo* Vincenzo Matta* Ali H. Sayed**

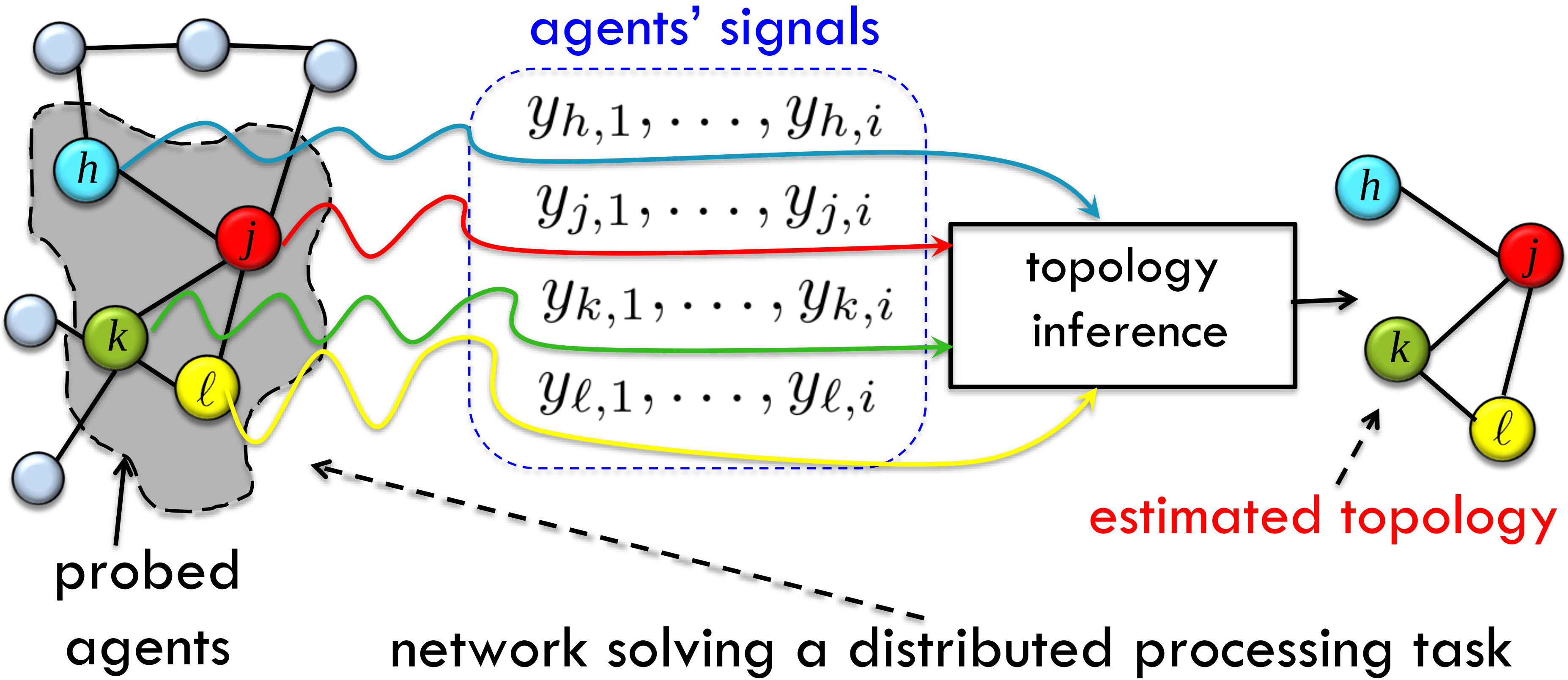
*DIEM University of Salerno **EPFL School of Engineering

ICASSP 2021

The Problem

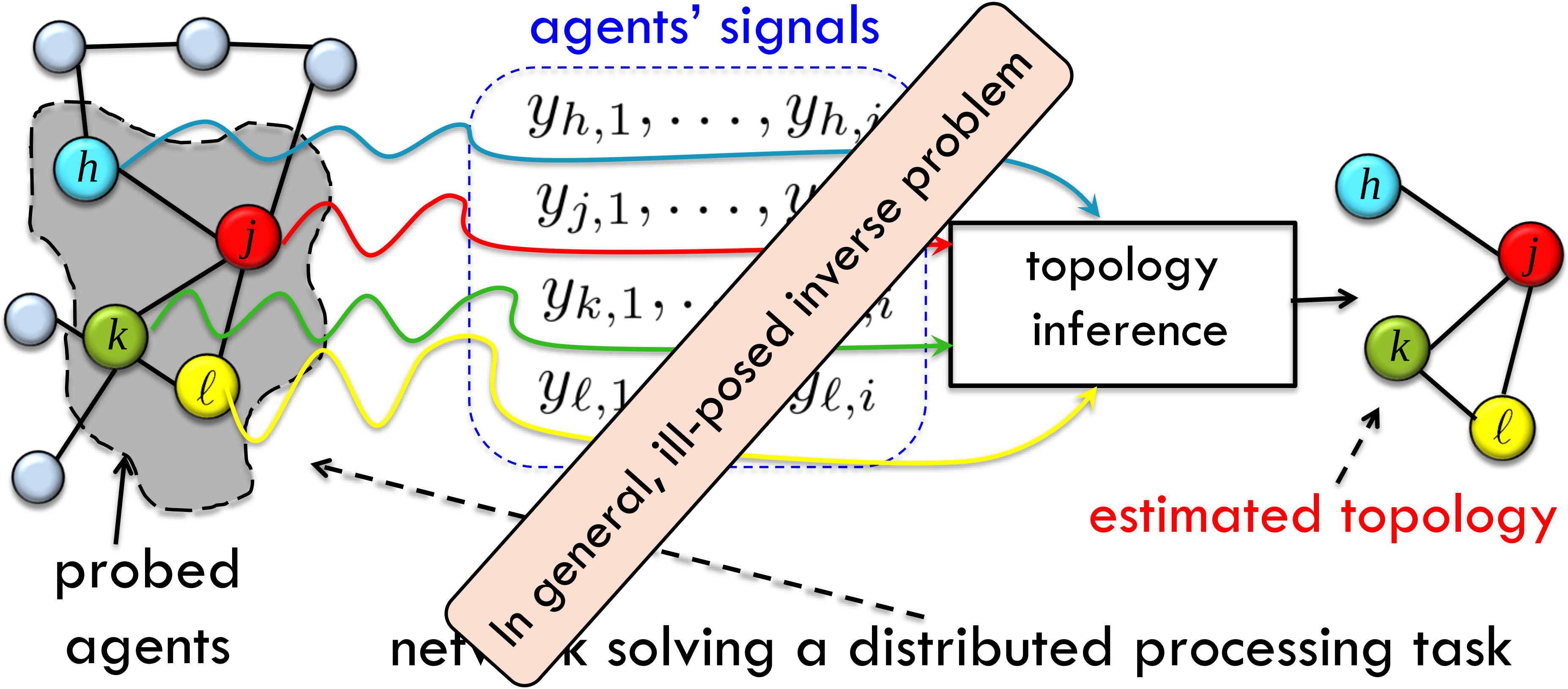


The Problem



network solving a distributed processing task (e.g., distributed detection)

The Problem



agents' signals

$y_{h,1}, \ldots, y_{h,i}$

$y_{j,1}, \ldots, y$

$y_{k,1}, \cdot$ i

$y_{\ell,1}$ $y_{\ell,i}$

topology inference

estimated topology

probed agents

In general, ill-posed inverse problem

net k solving a distributed processing task (e.g., distributed detection)

Fundamental Questions

- **Achievability**: Given a networked system, is it possible to learn the underlying topology from the observable samples?
- **Hardness:** Assuming an unlimited number of samples, how hard is to retrieve the graph? (large matrix inversion, NP-hard search, …)
- **Sample complexity:** How does the performance scale with the number of samples?

Fundamental Questions

- **Achie** Given a networked system, is it possible to learn the unde the observable samples?

Which models can be learned?

- **Hardness:** Assuming an unlim ber of samples, how hard is to retrieve the graph? (large matrix inversion, NP-hard search, …)
- **Sample complexity:** How does the performance scale with the number of samples?

Fundamental Questions

- **Achie** Given a networked system, is it possible to learn the unde the observable samples?

Which models can be learned?

- **Hardness:** Assuming an unlim ber of samples, how hard i retrieve the graph? (large matrix inversion, NP-hard sea

- **Sample complexity:** How does the p e with the number of samples?

Which estimators are good?

Several Disciplines Involved

- ❑ ***Signal Processing.*** G. Mateos, S. Segarra, A. Marques, and A. Ribeiro, "Connecting the dots: Identifying network structure via graph signal processing," IEEE SP Mag., 2019

- ❑ ***Statistics.*** V. Chandrasekaran, P. A. Parrilo, and A. S. Willsky, "Latent variable graphical model selection via convex optimization," *Ann. Statist.*, 2012

- ❑ ***Computer Science.*** A. Anandkumar, V. Y. F. Tan, F. Huang, and A. S. Willsky, "High-dimensional Gaussian graphical model selection: Walk summability and local separation criterion," JMLR, 2012

- ❑ ***Control Theory.*** D. Materassi and M. V. Salapaka, "On the problem of reconstructing an unknown topology via locality properties of the Wiener filter," IEEE Trans. Autom. Control, 2012

- ❑ ***Economics.*** H. Lütkepohl, *New Introduction to Multiple Time Series Analysis*, Springer, 2005

Several Terminologies

Network tomography

Structure learning

Topology inference

Graph learning

Partial observability

Latent variables

…

State of the Art

- Results on achievability available under full observability

- Results on achievability available under partial observability for Graphical Models (no dynamics)

- Recent results on achievability available under partial observability and dynamical systems for Erdős-Rényi random graphs (homogeneous setting, all nodes are equal)

V. Matta, A. Santos, and A. H. Sayed, “Graph learning under partial observability,” Proc. IEEE, 2020

Main Contribution

- ☐ Results on achievability available under full observability

- ☐ Results on achievability available under partial observability for Graphical Models (no dynamics)

- ☐ Recent results on achievability available under partial observability and dynamical systems for Erdős-Rényi random graphs (homogeneous setting, all nodes are equal)

New results on achievability under partial observability and dynamical systems for Bollobás-Riordan random graphs (heterogeneous setting, scale-free behavior)

V. Matta, A. Santos, and A. H. Sayed, “Graph learning under partial observability,” Proc. IEEE, 2020

Dynamical System

VAR model:

(Vector AutoRegressive)

$$\boxed{y_i = Ay_{i-1} + x_i}$$

$$y_{\underbrace{k}_{\text{node}},\underbrace{i}_{\text{time}}} = \sum_{\ell=1}^{N} a_{k\ell} y_{\ell,i-1} + x_{k,i} \qquad \Leftrightarrow \qquad a_{k\ell} = 0$$

node k **not directly** influenced by node ℓ

Note: two nodes are correlated even if they are not directly connected, due to intermediate nodes!

Very Well-Known Problem

Exploiting the diffusion mechanism, we have:

one-lag steady-state covariance

$$R_1 = AR_0$$

steady-state covariance

$$\boxed{A = R_1 R_0^{-1}}$$

best linear predictor (aka Granger predictor)

Very Well-Known Problem

Exploiting the diffusion mechanism, we have:

one-lag steady-state empirical covariance

$$\widehat{R}_1 = \widehat{A}\widehat{R}_0$$

steady-state empirical covariance

$$\boxed{\widehat{A} = \widehat{R}_1\widehat{R}_0^{-1}}$$

best linear predictor (aka Granger predictor)

This is useful since there are many ways to estimate the covariance matrices ***from the output measurements***

Partial Observability

$$\boxed{\widehat{A}_S = [R_1]_S([R_0]_S)^{-1}}$$

probed subset

Granger predictor

given the probed subset

Obviously different from

$$\boxed{A_S = [R_1 R_0^{-1}]_S}$$

Partial Observability

$$\widehat{A}_S = [R_1]_S([R_0]_S)^{-1}$$

probed subset

The action of latent agents can be detrimental!

Granger predictor

given the probed subset

Obviously different from

$$A_S = [R_1 R_0^{-1}]_S$$

Partial Observability

$$\widehat{A}_{\mathcal{S}} = [R_1]_{\mathcal{S}}([R_0]_{\mathcal{S}})^{-1}$$

probed subset

Granger predictor

given the probed subset

Can we still estimate the topology?

Obviously different from

$$A_{\mathcal{S}} = [R_1 R_0^{-1}]_{\mathcal{S}}$$

Considered Class of Matrices

$$a_{k\ell} \propto \frac{1}{d_{\max}}$$

Laplacian combination matrix

connected entries inversely proportional to the maximum degree

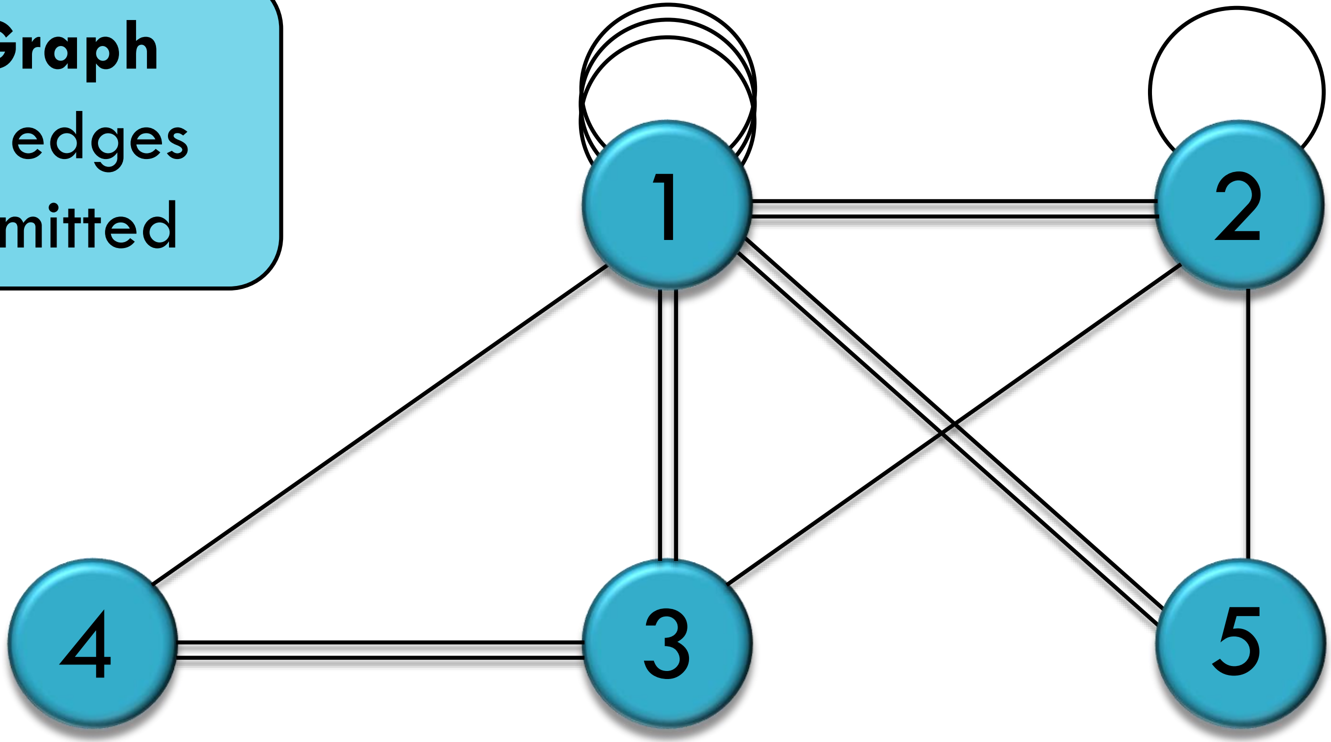
Popular choice for distributed processing algorithms

(e.g., diffusion or consensus algortihms)

A. H. Sayed, *Adaptation, learning, and optimization over networks*, *Found. Trends Mach. Learn.*, 2014

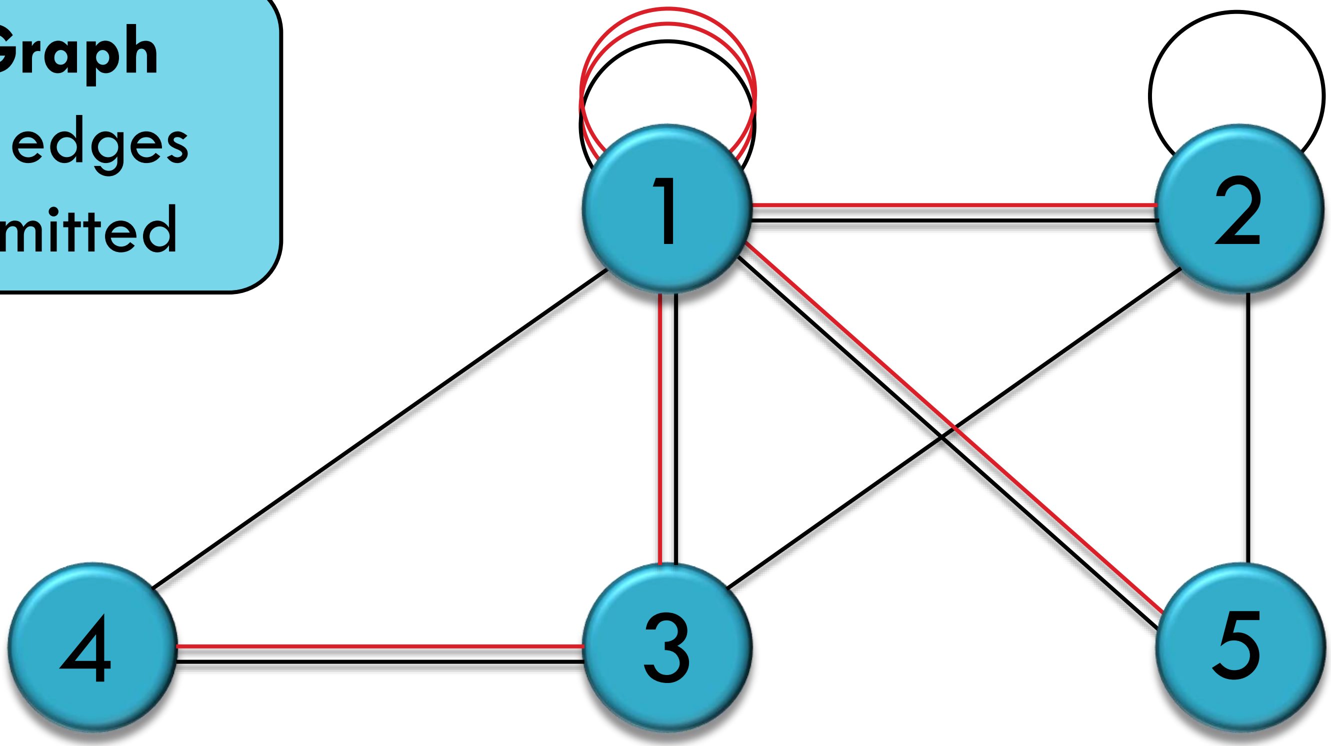
Bollobás-Riordan Model

Multi-Graph
multiple edges
are permitted



Bollobás-Riordan Model

Multi-Graph
multiple edges
are permitted



Bollobás-Riordan Multi-Graph Construction

A Bollobás-Riordan multi-graph of size N and parameter $\eta \in \mathbb{N}$ is obtained by the following iterative procedure, consisting of N steps

Step 1

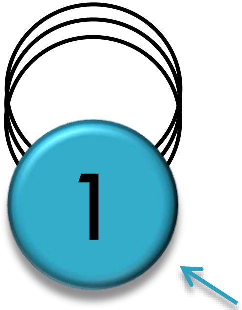
Build a deterministic multi-graph with one node and η self-loops

Steps 2,…,N

Starting from the multi-graph built at the previous step, **add a new node and η new edges** according to a **preferential attachment** rule

Step-by-Step Illustration

$N = 5$



one node, $\eta = 3$ self-loops

Step 1 Build the initial multi-graph

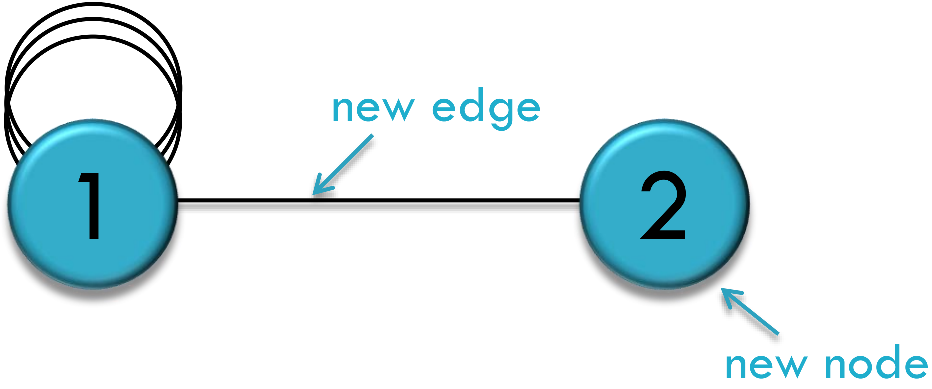
Step-by-Step Illustration

$N = 5$

Step 2 Add a new node and $\eta = 3$ edges

Step-by-Step Illustration

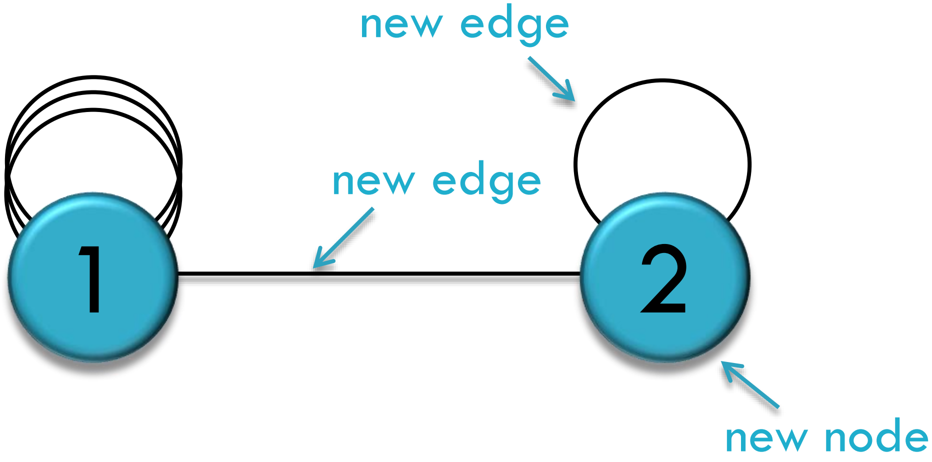
$N = 5$



Step 2 Add a new node and $\eta = 3$ edges

Step-by-Step Illustration

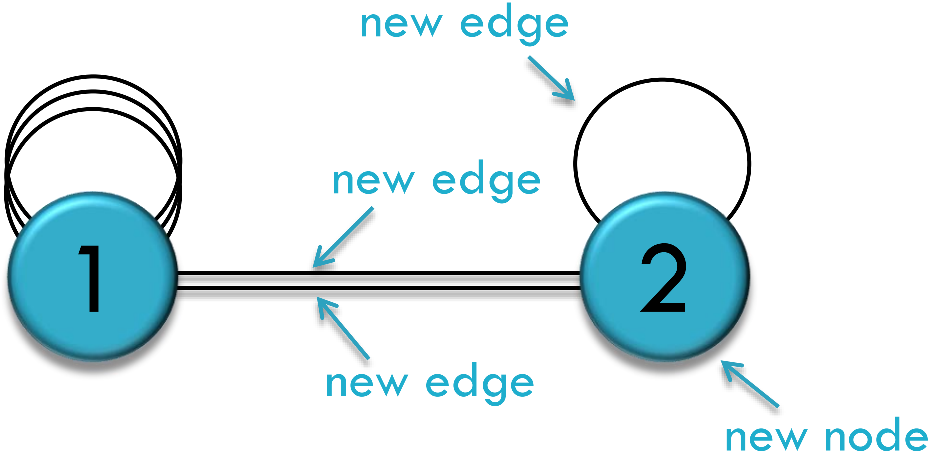
$N = 5$



Step 2 Add a new node and $\eta = 3$ edges

Step-by-Step Illustration

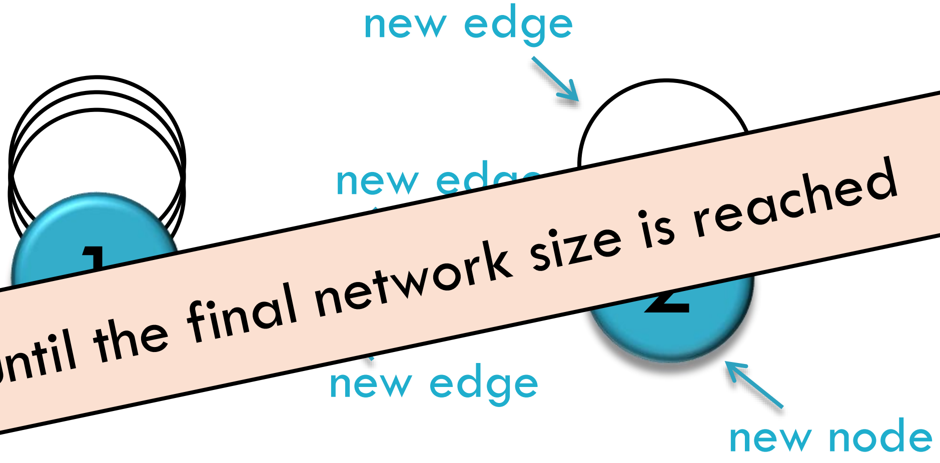
$N = 5$



Step 2 Add a new node and $\eta = 3$ edges

Step-by-Step Illustration

$N = 5$

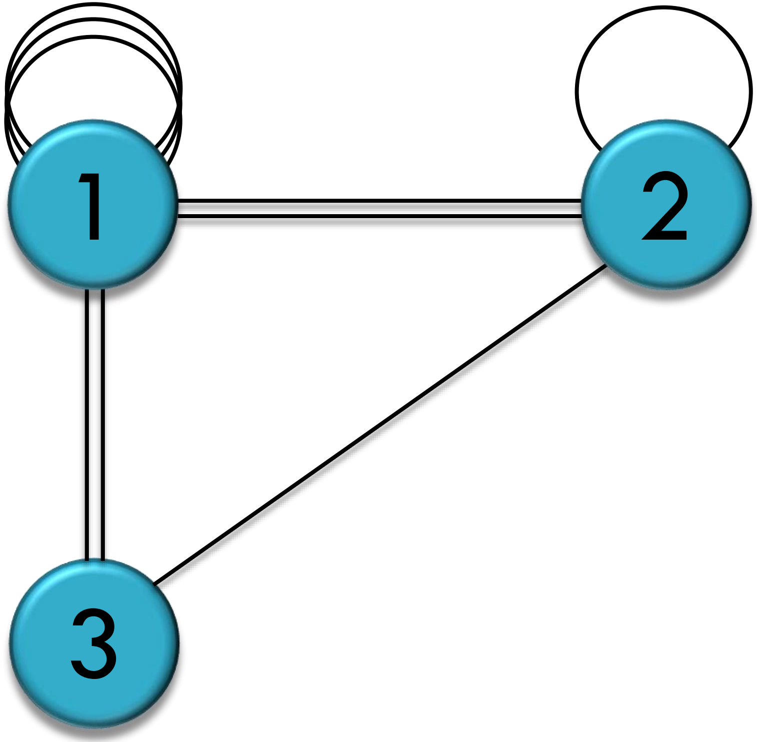


Iterate until the final network size is reached

Step 2 Add a new node and $\eta = 3$ edges

Step-by-Step Illustration

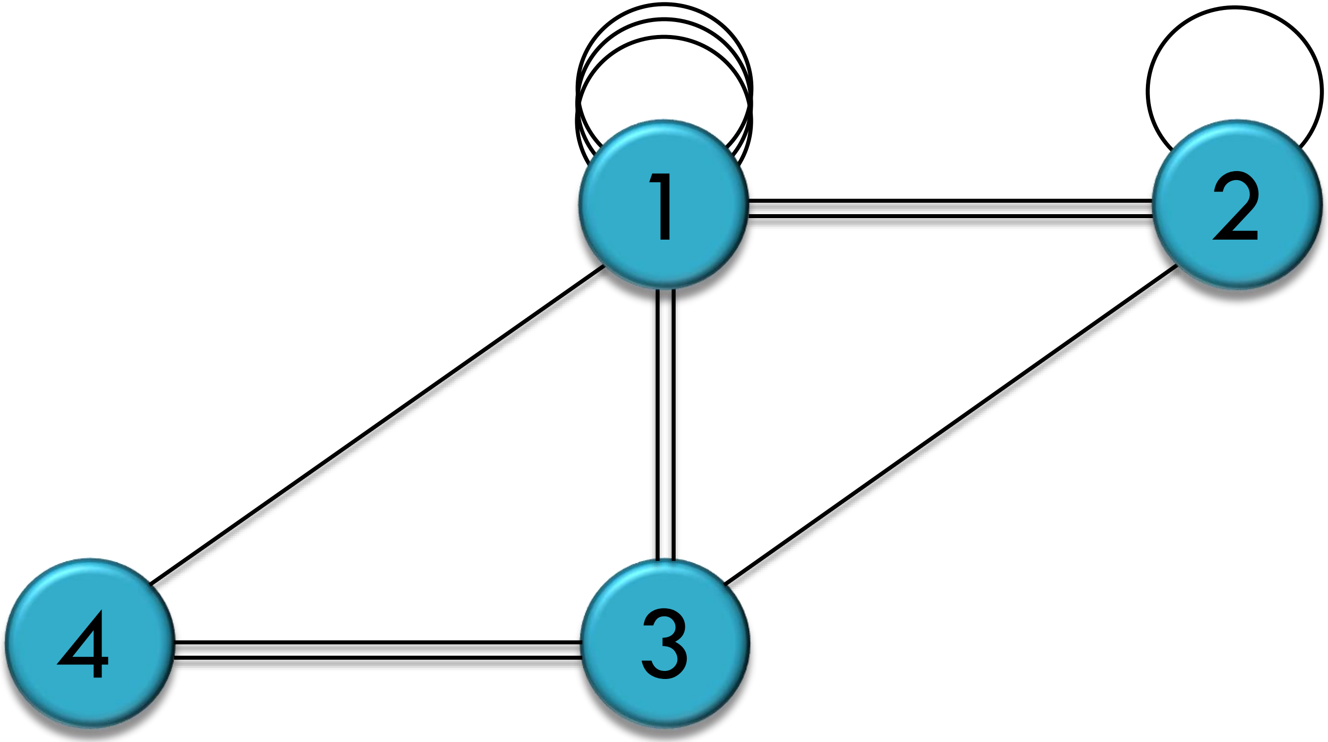
$N = 5$



Step 3

Step-by-Step Illustration

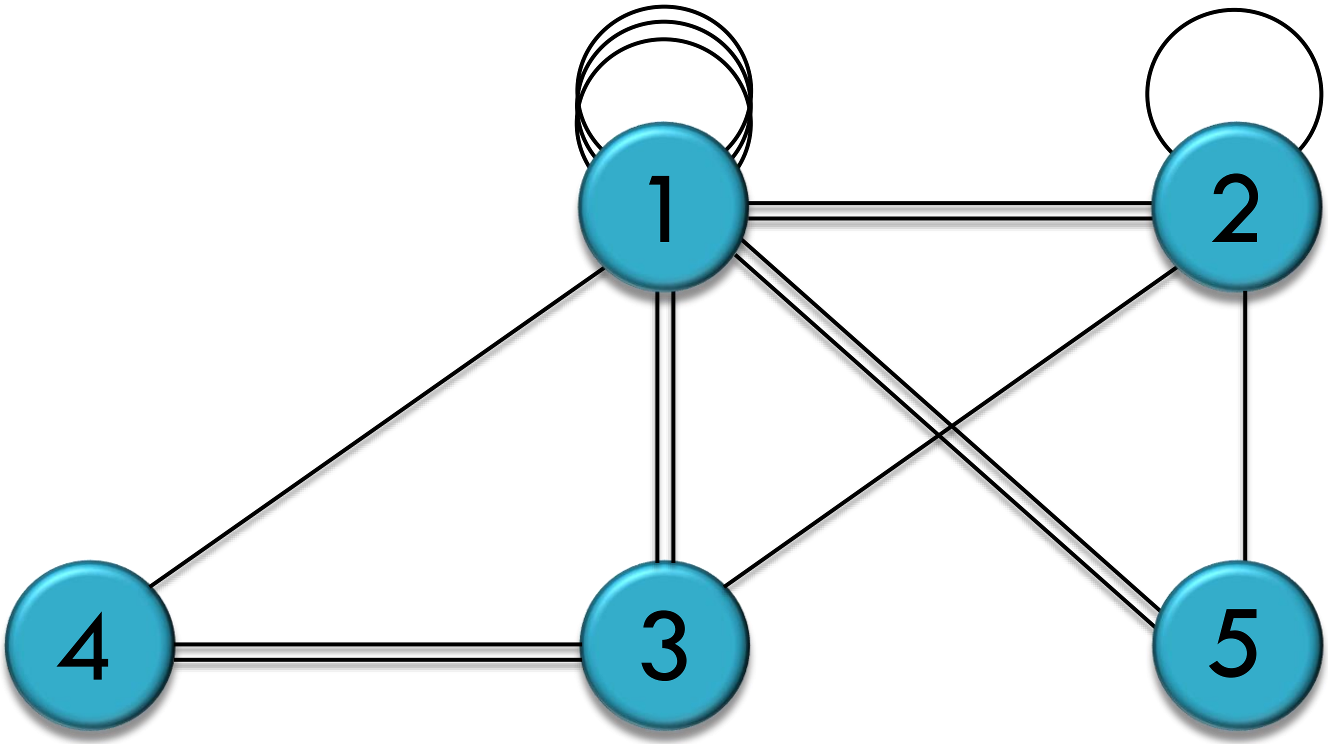
$N = 5$



Step 4

Step-by-Step Illustration

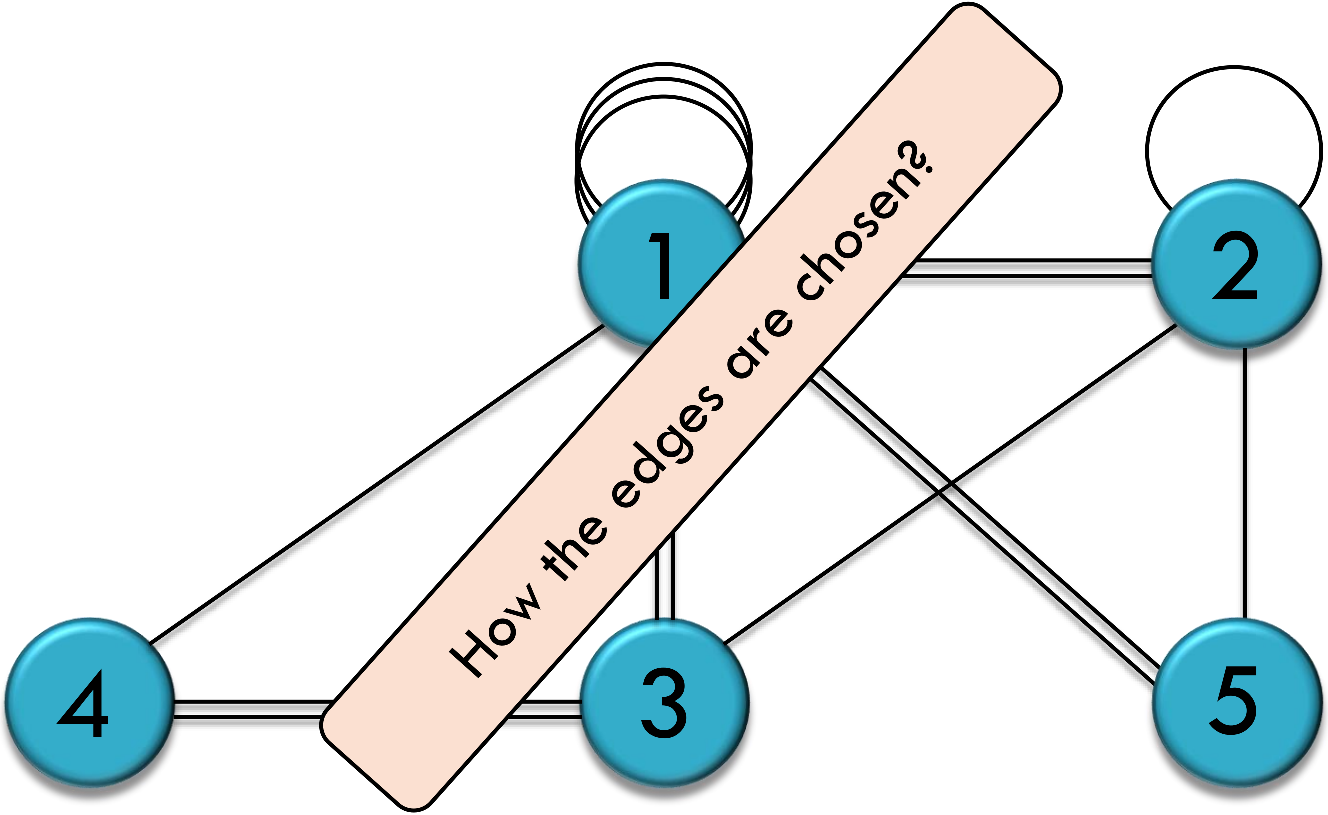
$N = 5$



Step 5

Step-by-Step Illustration

$N = 5$



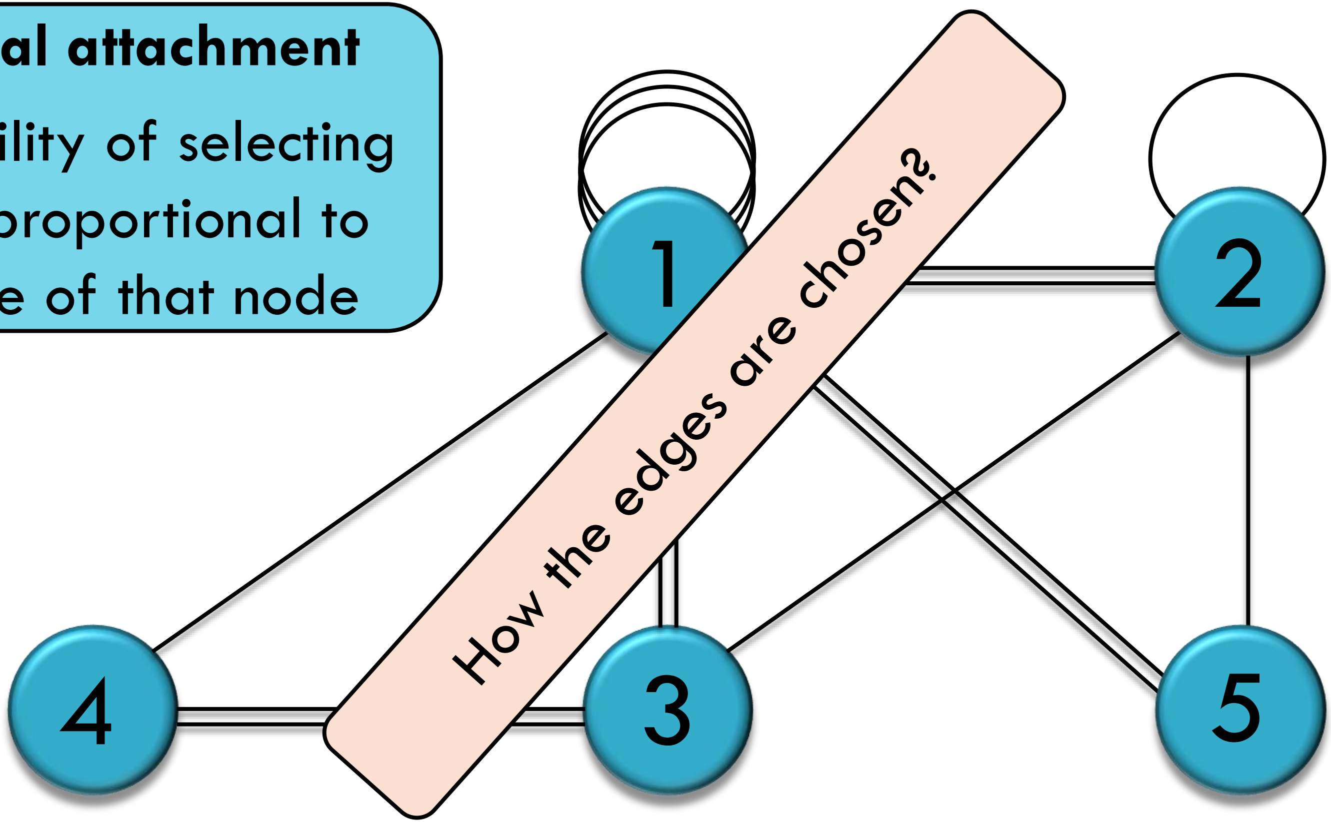
How the edges are chosen?

Step 5

Step-by-Step Illustration

Preferential attachment

the probability of selecting a node is proportional to the degree of that node

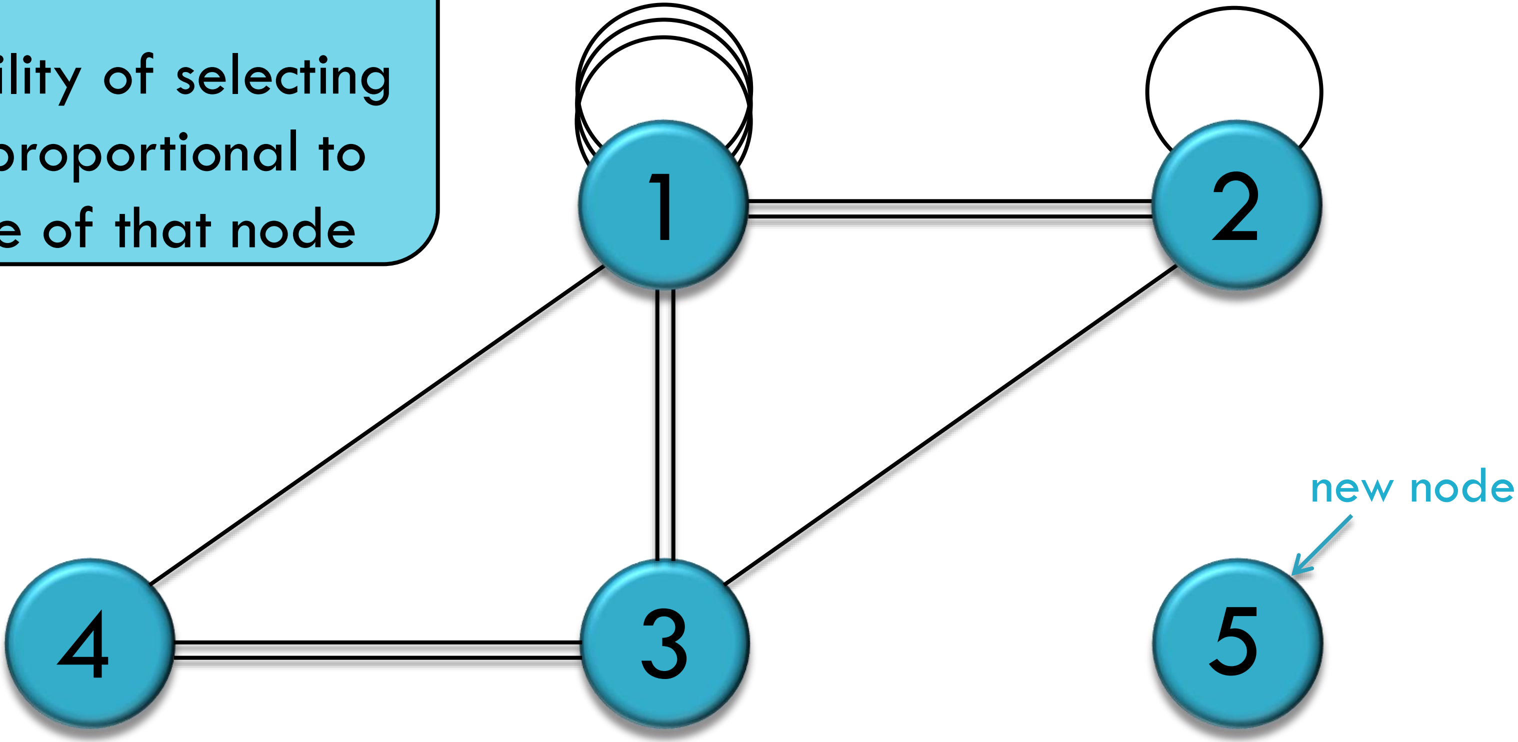


Step 5

Preferential Attachment Rule

Preferential attachment

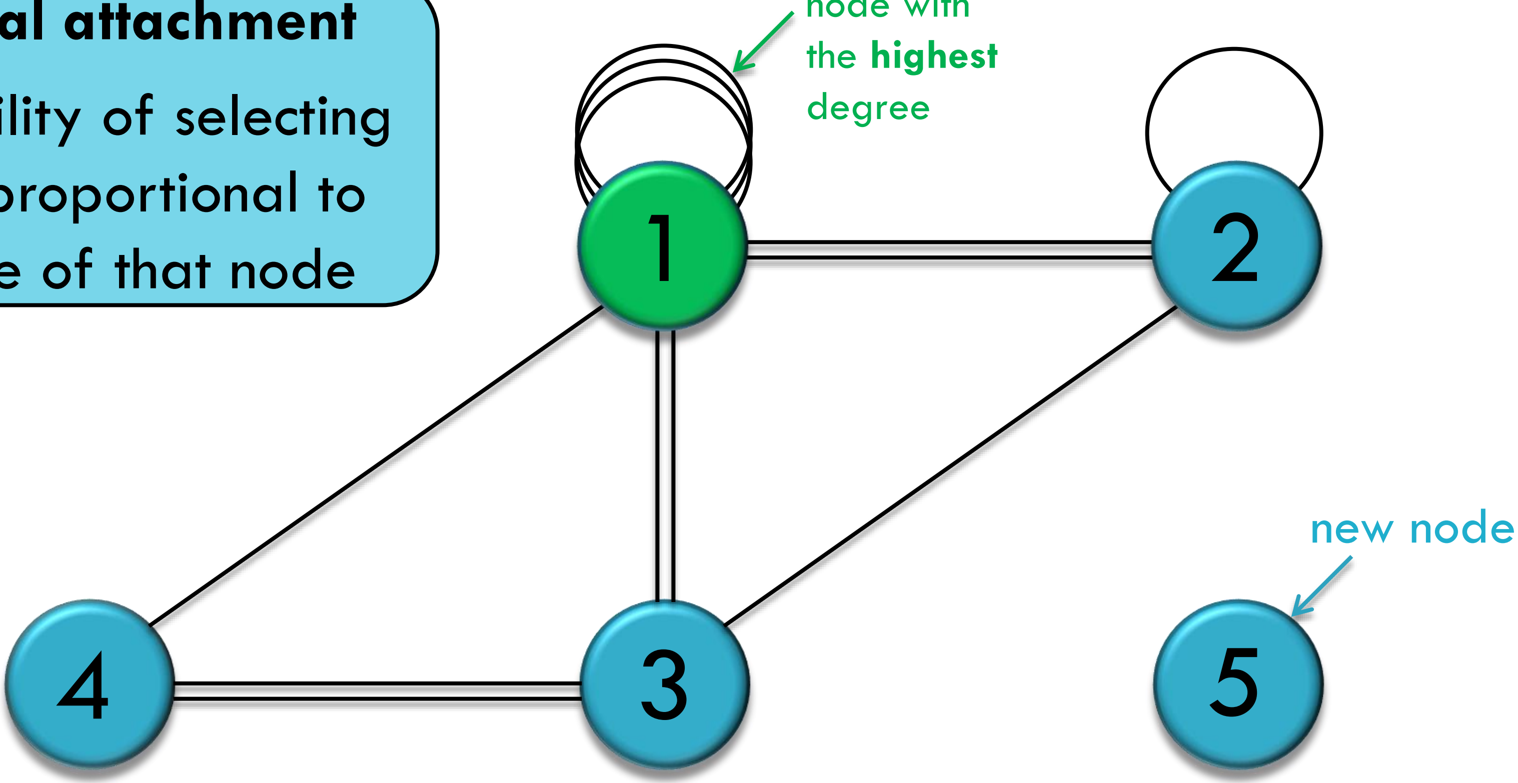
the probability of selecting a node is proportional to the degree of that node



Preferential Attachment Rule

Preferential attachment

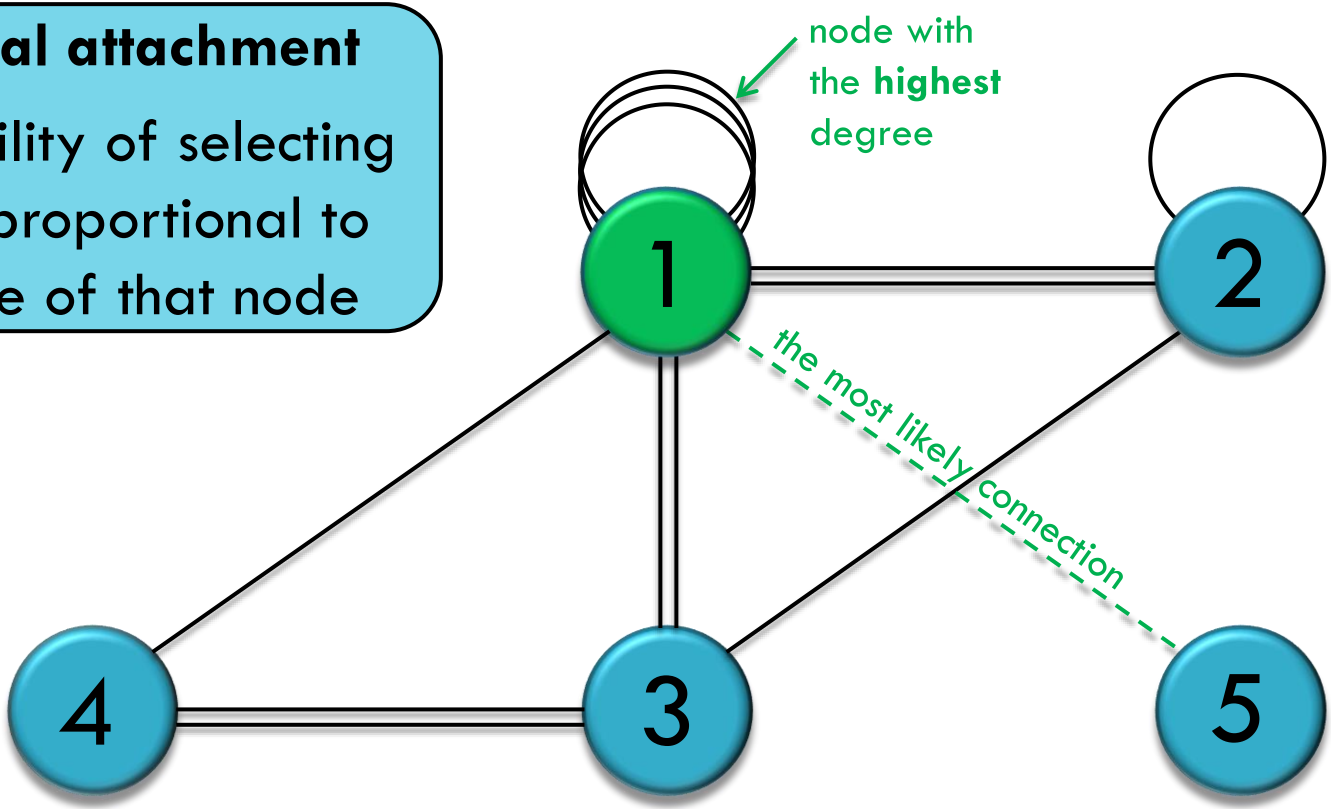
the probability of selecting a node is proportional to the degree of that node



Preferential Attachment Rule

Preferential attachment

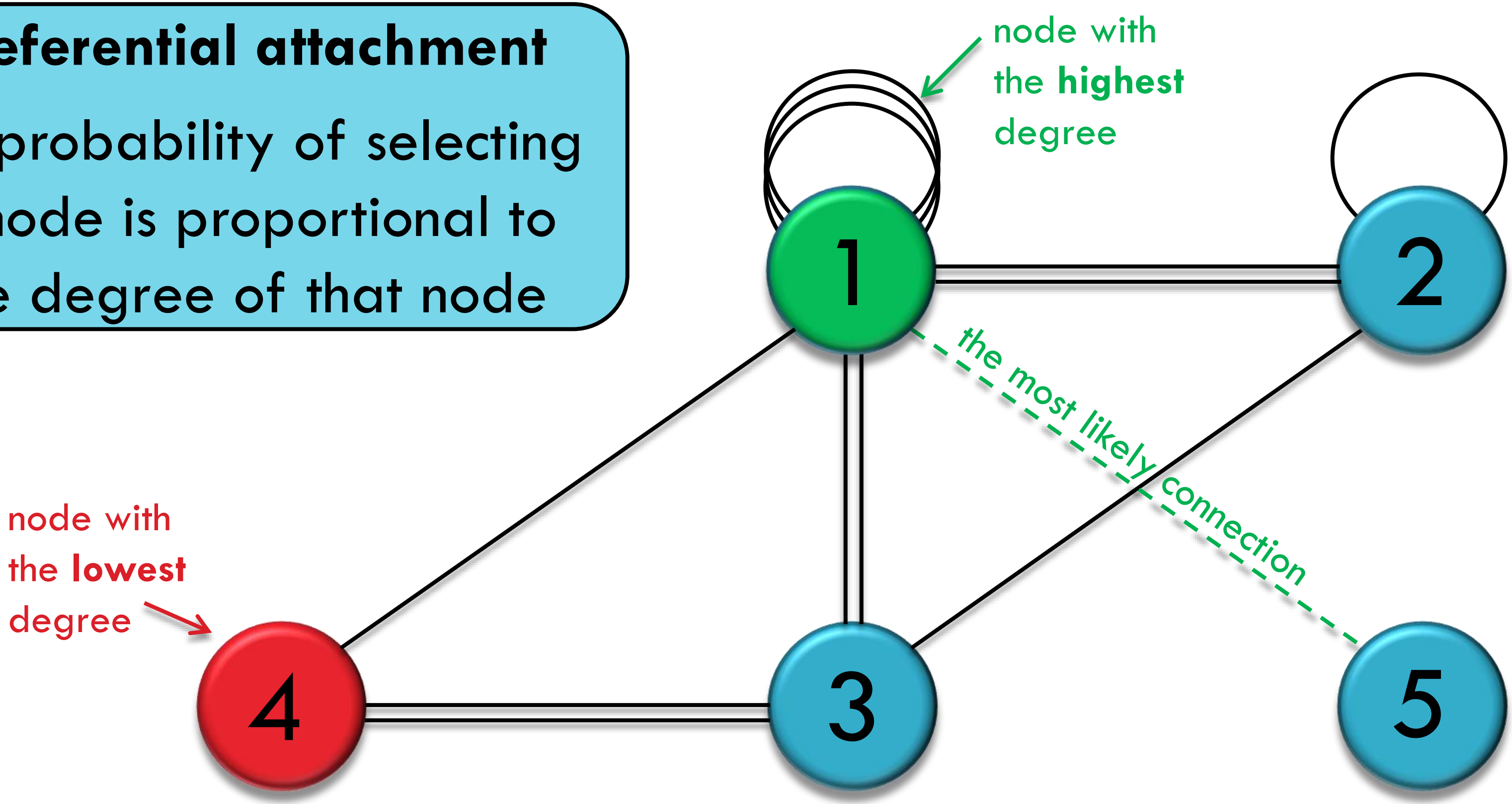
the probability of selecting a node is proportional to the degree of that node



Preferential Attachment Rule

Preferential attachment

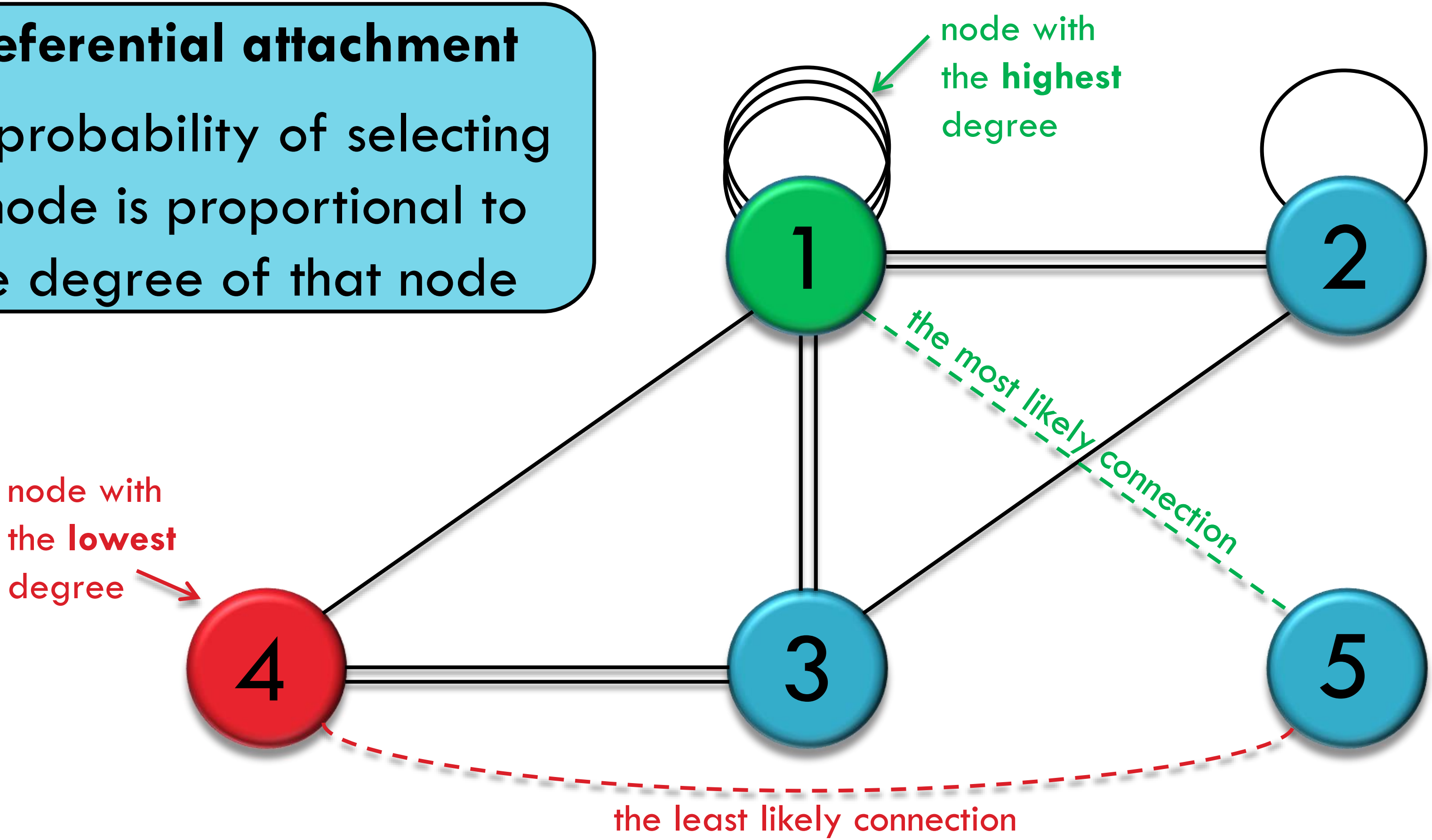
the probability of selecting a node is proportional to the degree of that node



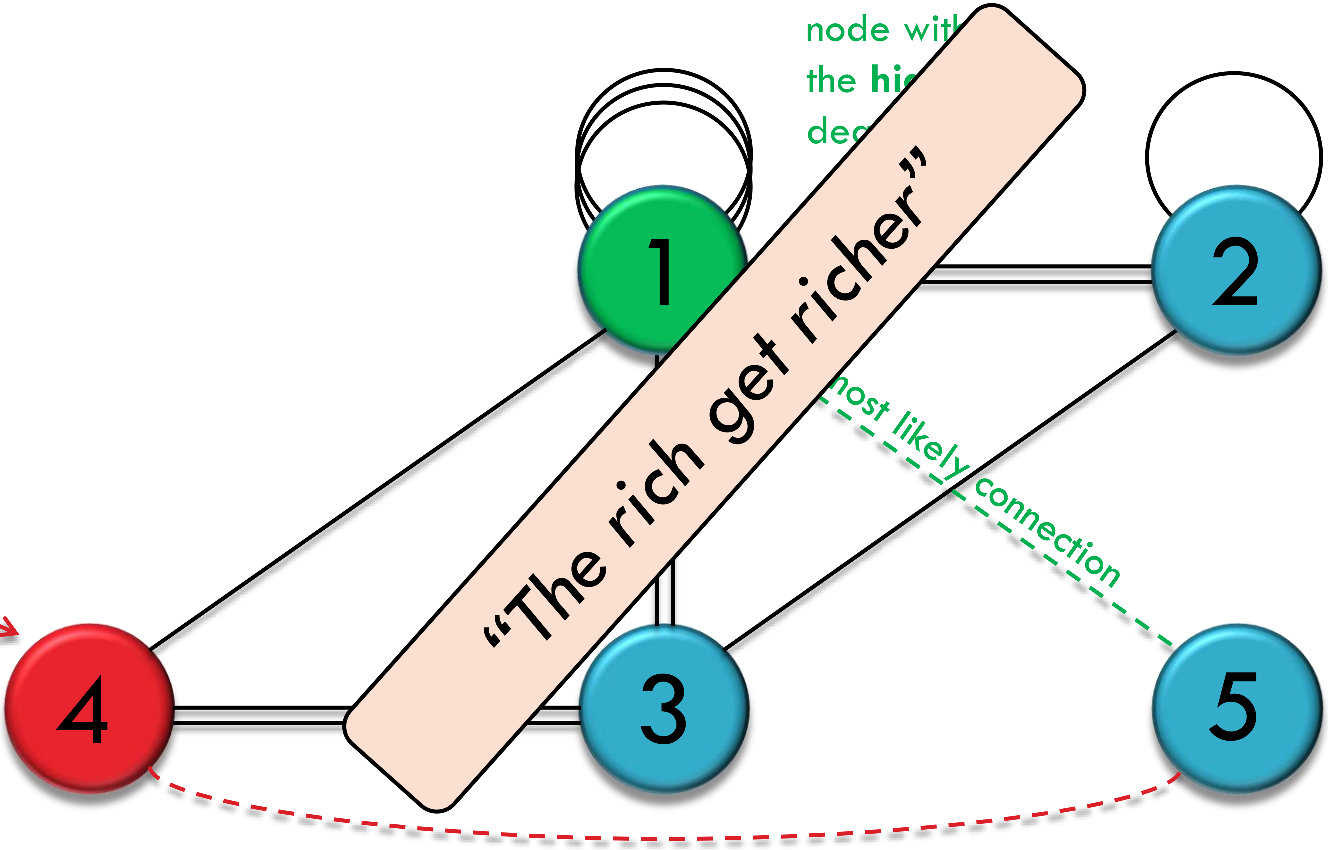
Preferential Attachment Rule

Preferential attachment

the probability of selecting a node is proportional to the degree of that node



Preferential Attachment Rule



node wit
the **hi**
de

“The rich get richer”

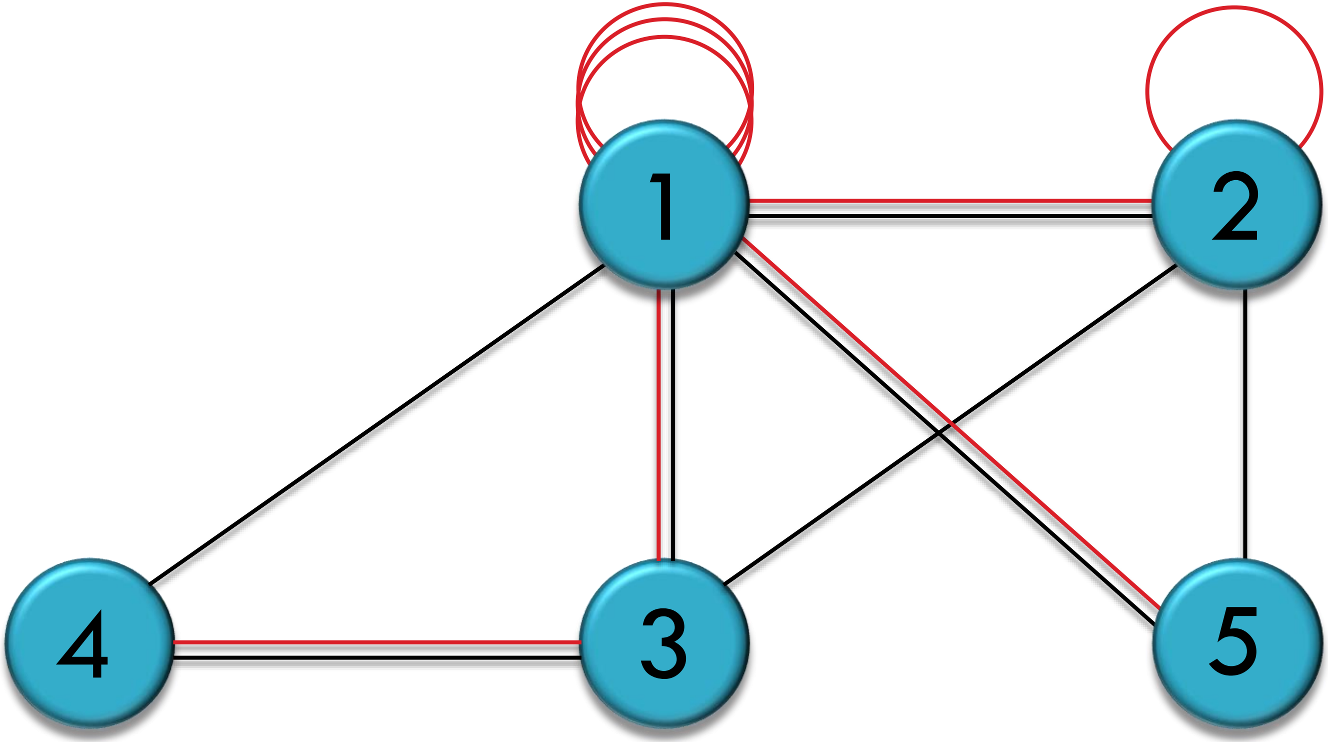
most likely connection

node with
the **lowest**
degree

the least likely connection

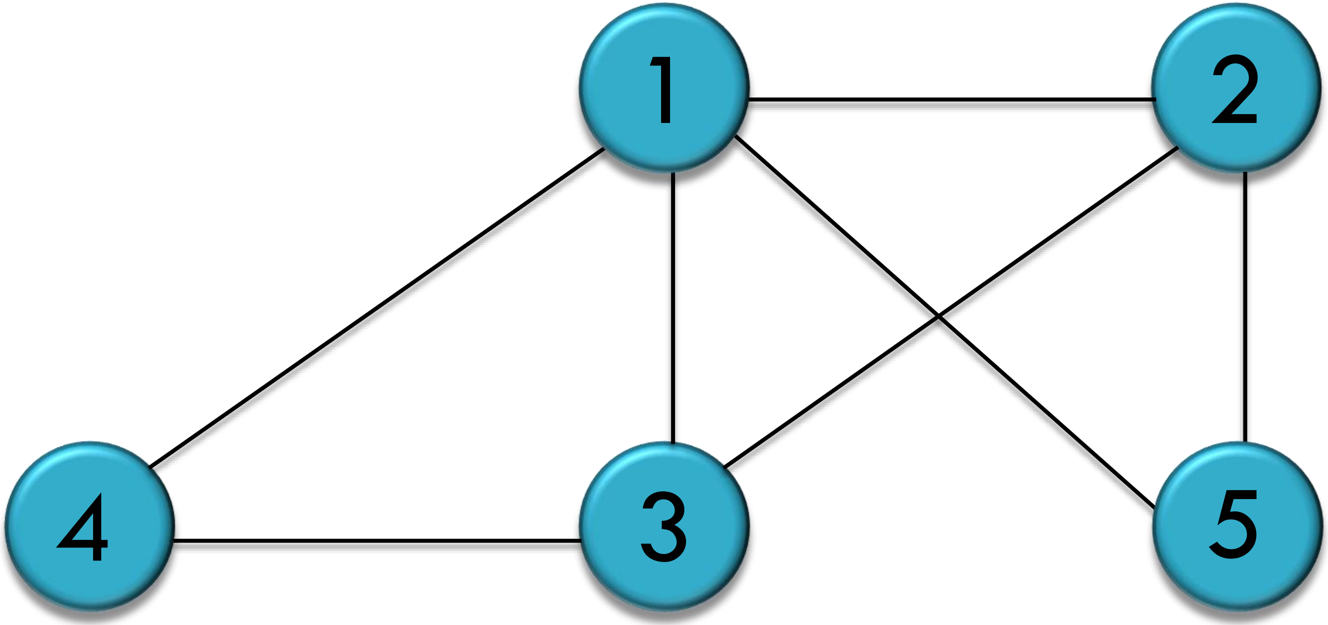
From Multi-Graphs to Graphs

Remove self-loops and redundant edges

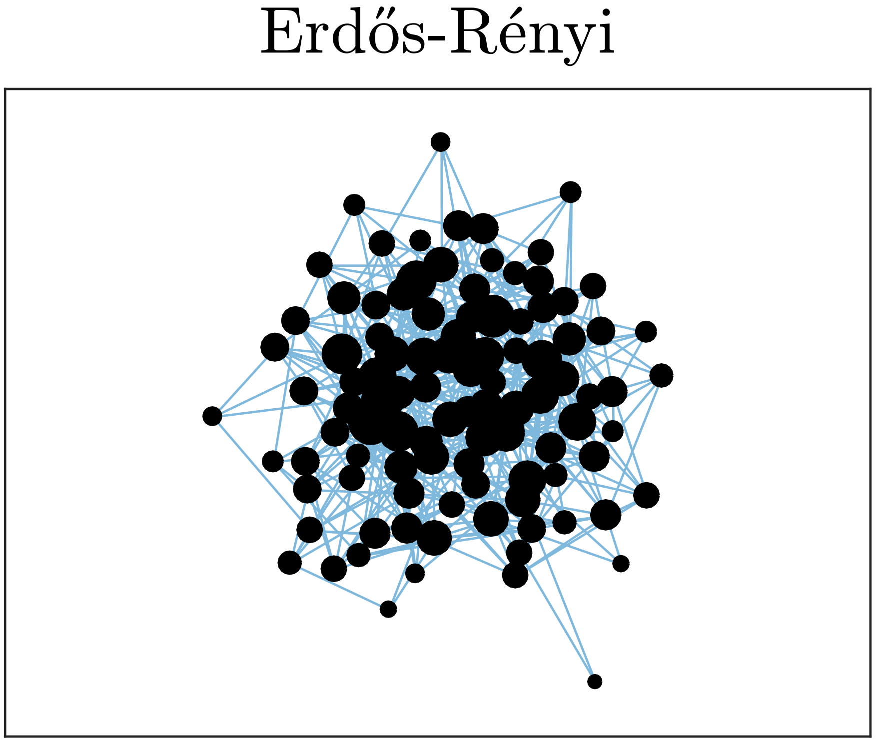


From Multi-Graphs to Graphs

Final Bollobás-Riordan **graph**



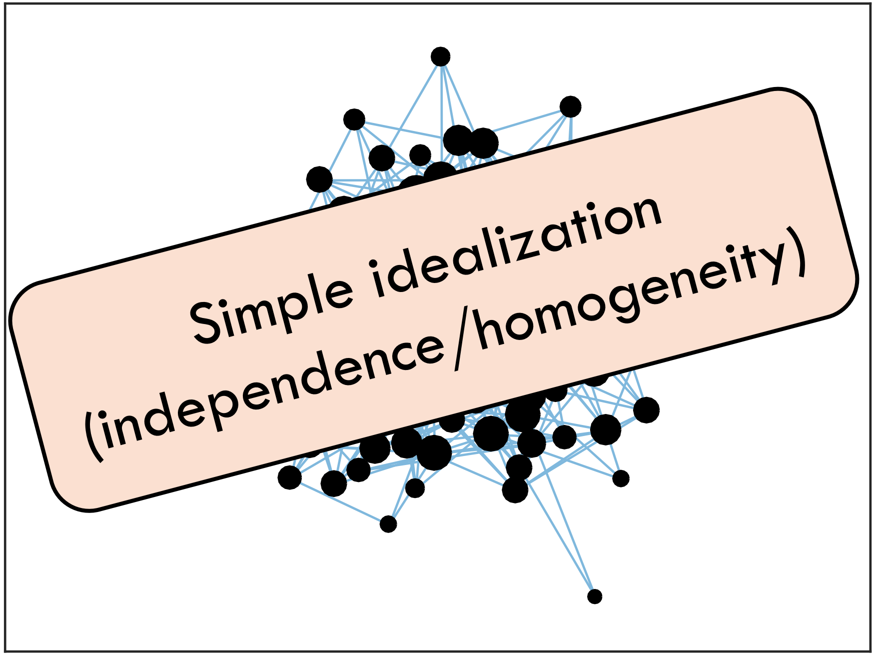
Erdős-Rényi vs. Bollobás-Riordan



B. Bollobás, *Random Graphs*, Cambridge University Press, 2001.

Erdős-Rényi vs. Bollobás-Riordan

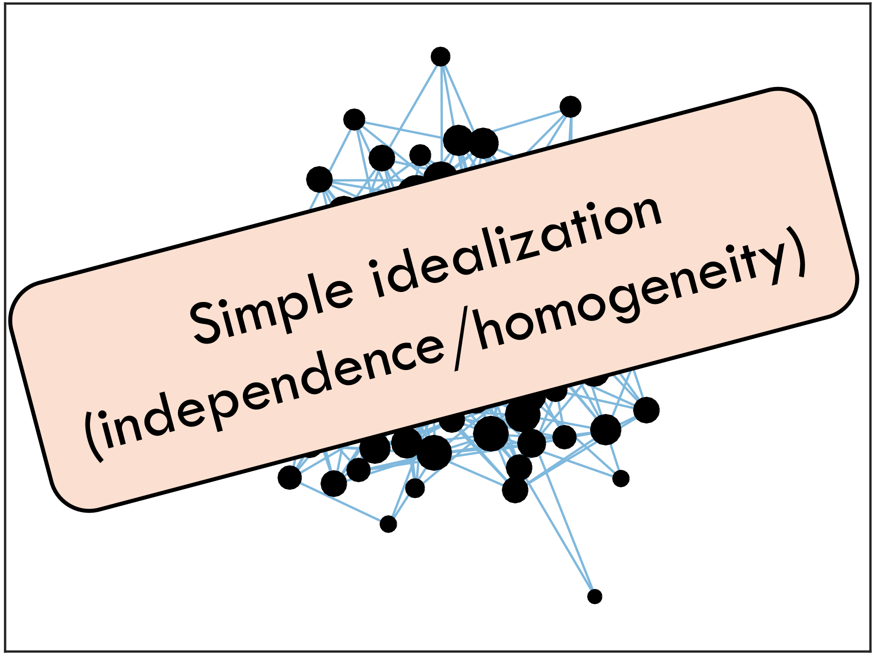
Erdős-Rényi



Simple idealization (independence/homogeneity)

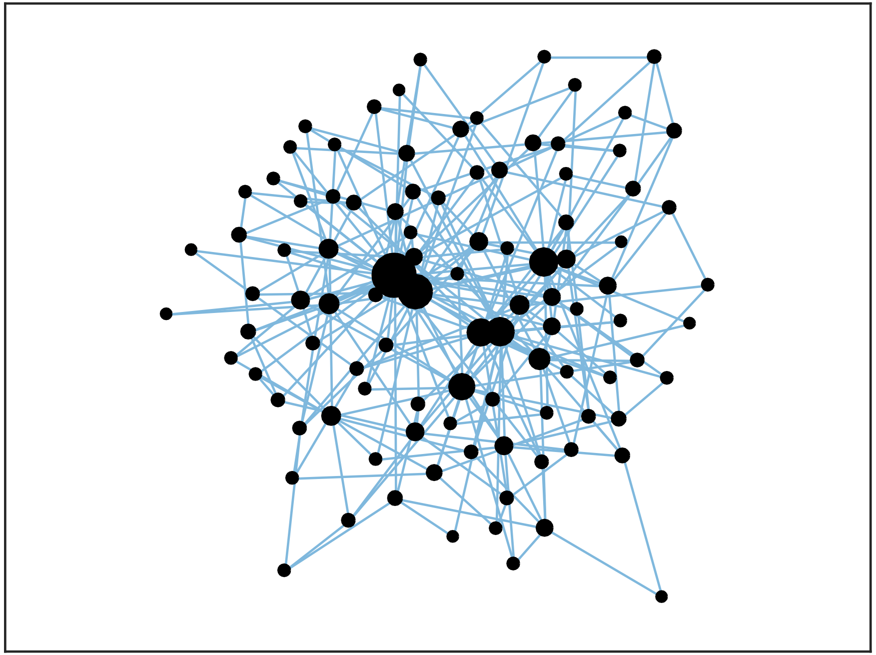
B. Bollobás, *Random Graphs*, Cambridge University Press, 2001.

Erdős-Rényi vs. Bollobás-Riordan



Erdős-Rényi

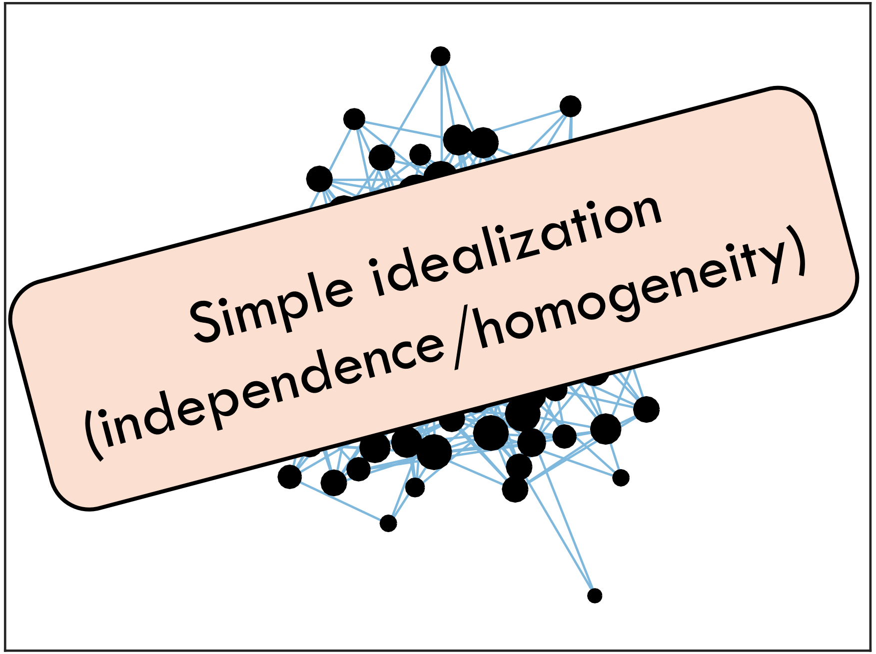
Simple idealization (independence/homogeneity)



Bollobás-Riordan

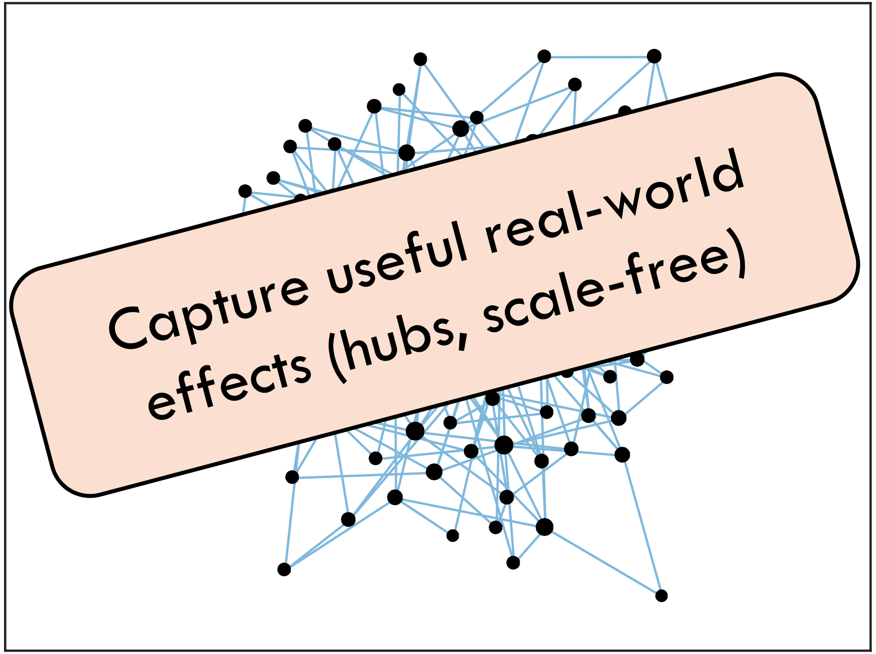
B. Bollobás, *Random Graphs*, Cambridge University Press, 2001.

Erdős-Rényi vs. Bollobás-Riordan



Erdős-Rényi

Simple idealization (independence/homogeneity)



Bollobás-Riordan

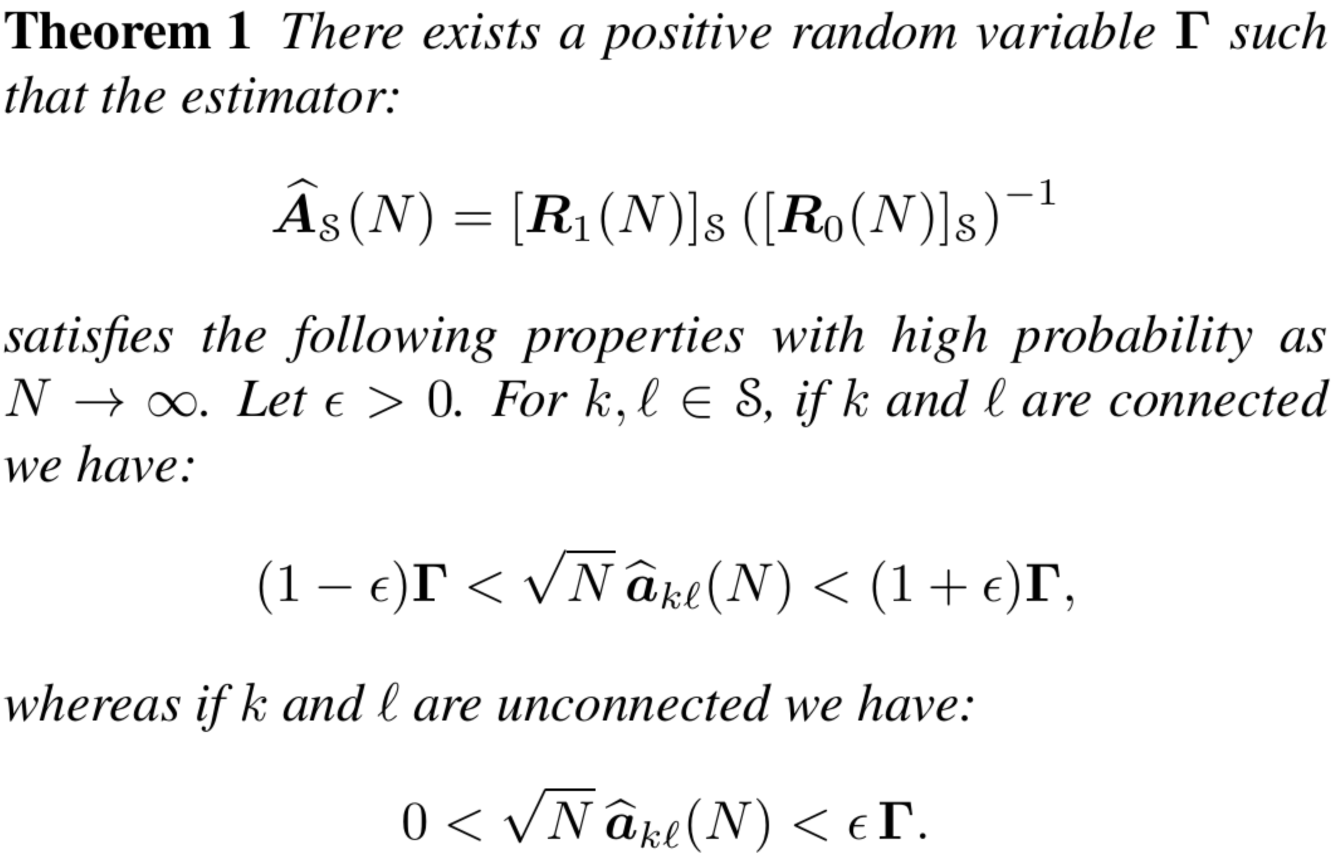
Capture useful real-world effects (hubs, scale-free)

B. Bollobás, *Random Graphs*, Cambridge University Press, 2001.

Technical Tools

- **Random Graphs.** Asymptotic behavior of maximum degree over scale-free graphs
- **Statistical Learning.** Statistical concentration behavior of some relevant network descriptors (challenging issue, due to dependencies introduced by scale-free graphs)

Guarantees of Topology Estimation



Theorem 1 *There exists a positive random variable* $\mathbf{\Gamma}$ *such that the estimator:*

$$\widehat{\boldsymbol{A}}_{\mathcal{S}}(N) = [\boldsymbol{R}_1(N)]_{\mathcal{S}} \left([\boldsymbol{R}_0(N)]_{\mathcal{S}}\right)^{-1}$$

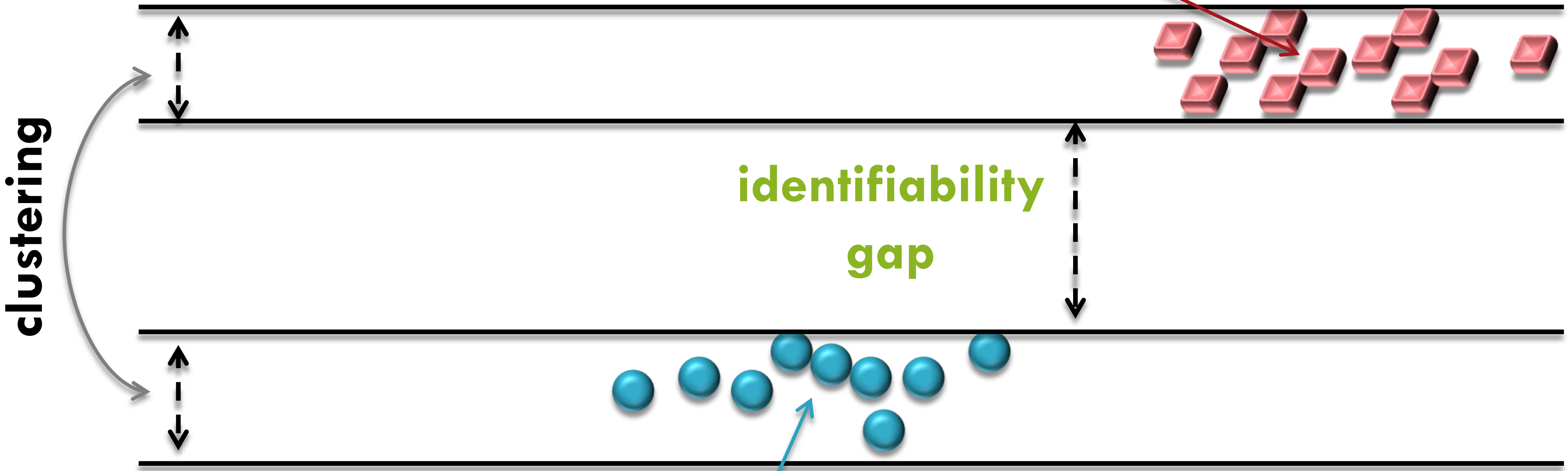
satisfies the following properties with high probability as $N \to \infty$. *Let* $\epsilon > 0$. *For* $k, \ell \in \mathcal{S}$, *if* k *and* ℓ *are connected we have:*

$$(1-\epsilon)\mathbf{\Gamma} < \sqrt{N}\,\widehat{\boldsymbol{a}}_{k\ell}(N) < (1+\epsilon)\mathbf{\Gamma},$$

whereas if k *and* ℓ *are unconnected we have:*

$$0 < \sqrt{N}\,\widehat{\boldsymbol{a}}_{k\ell}(N) < \epsilon\,\mathbf{\Gamma}.$$

Main Result: Identifiability Gap



estimated matrix entries for **connected** pairs

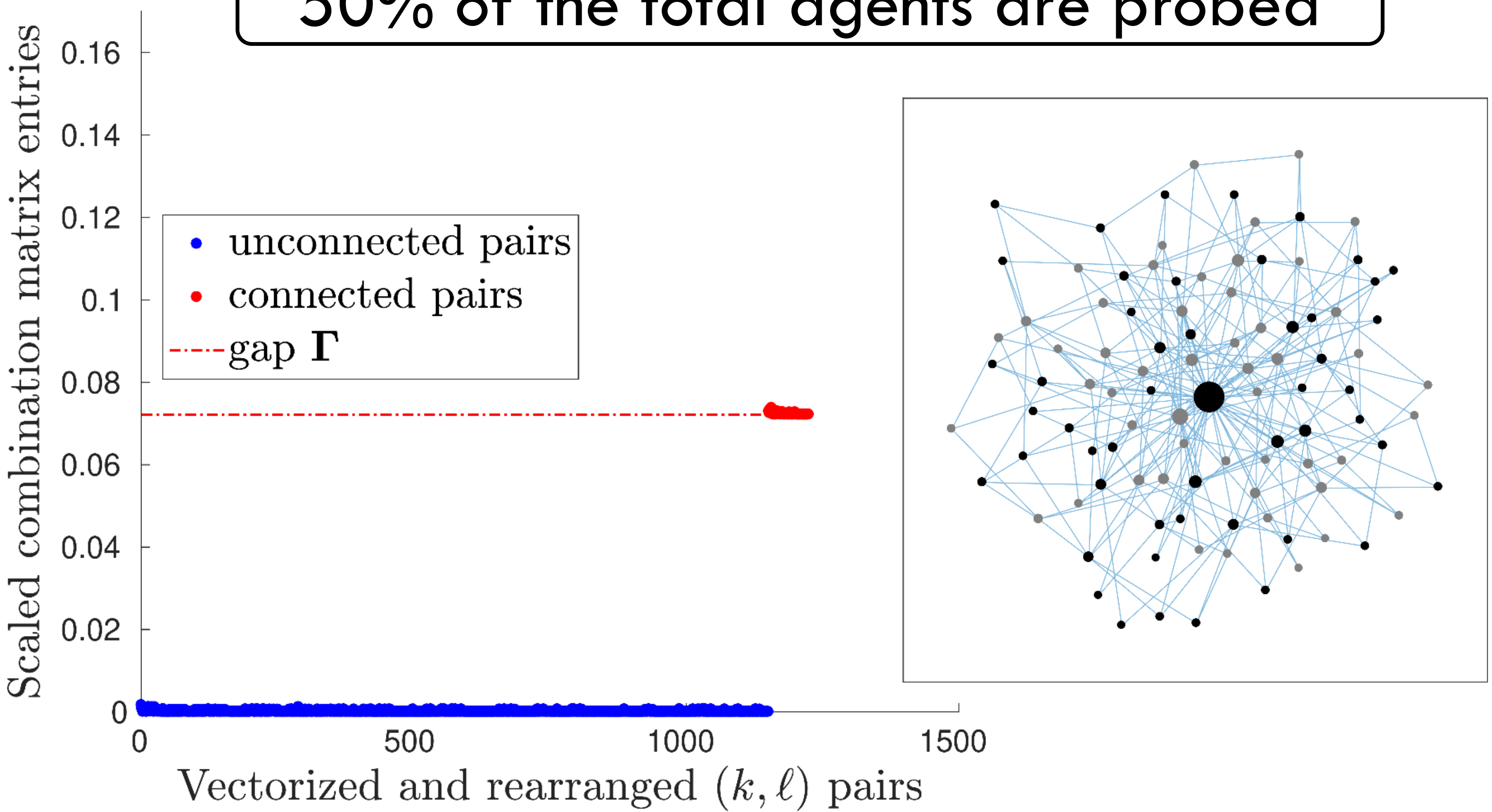
identifiability gap

clustering

estimated matrix entries for **unconnected** pairs

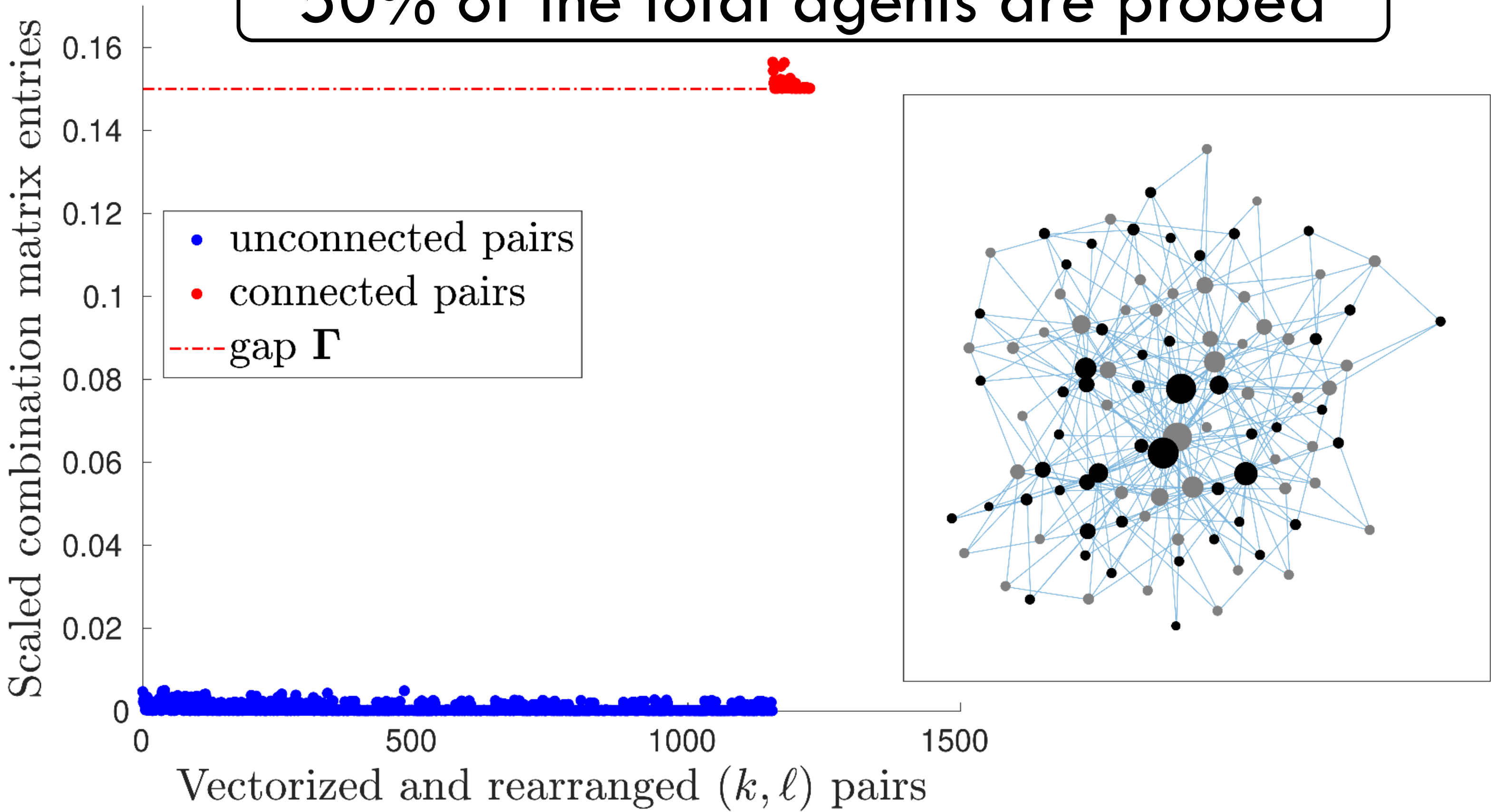
Single Hub

50% of the total agents are probed

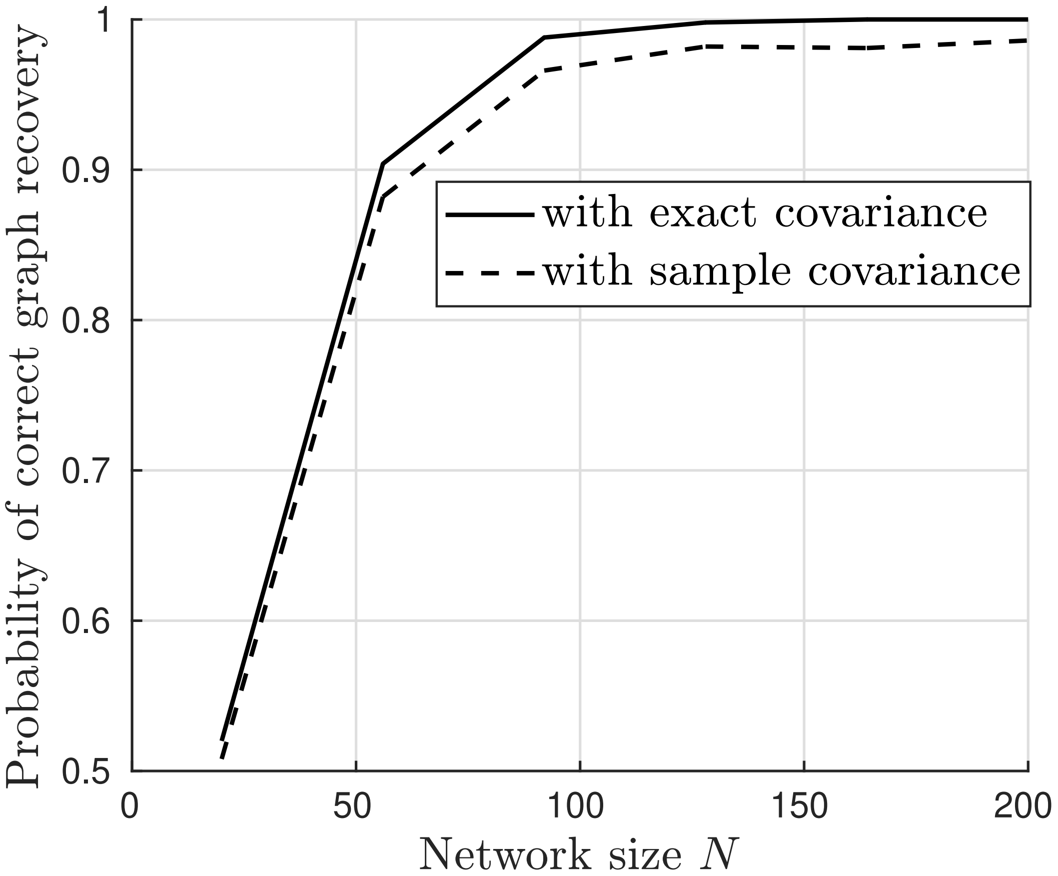


Multiple Hubs

50% of the total agents are probed



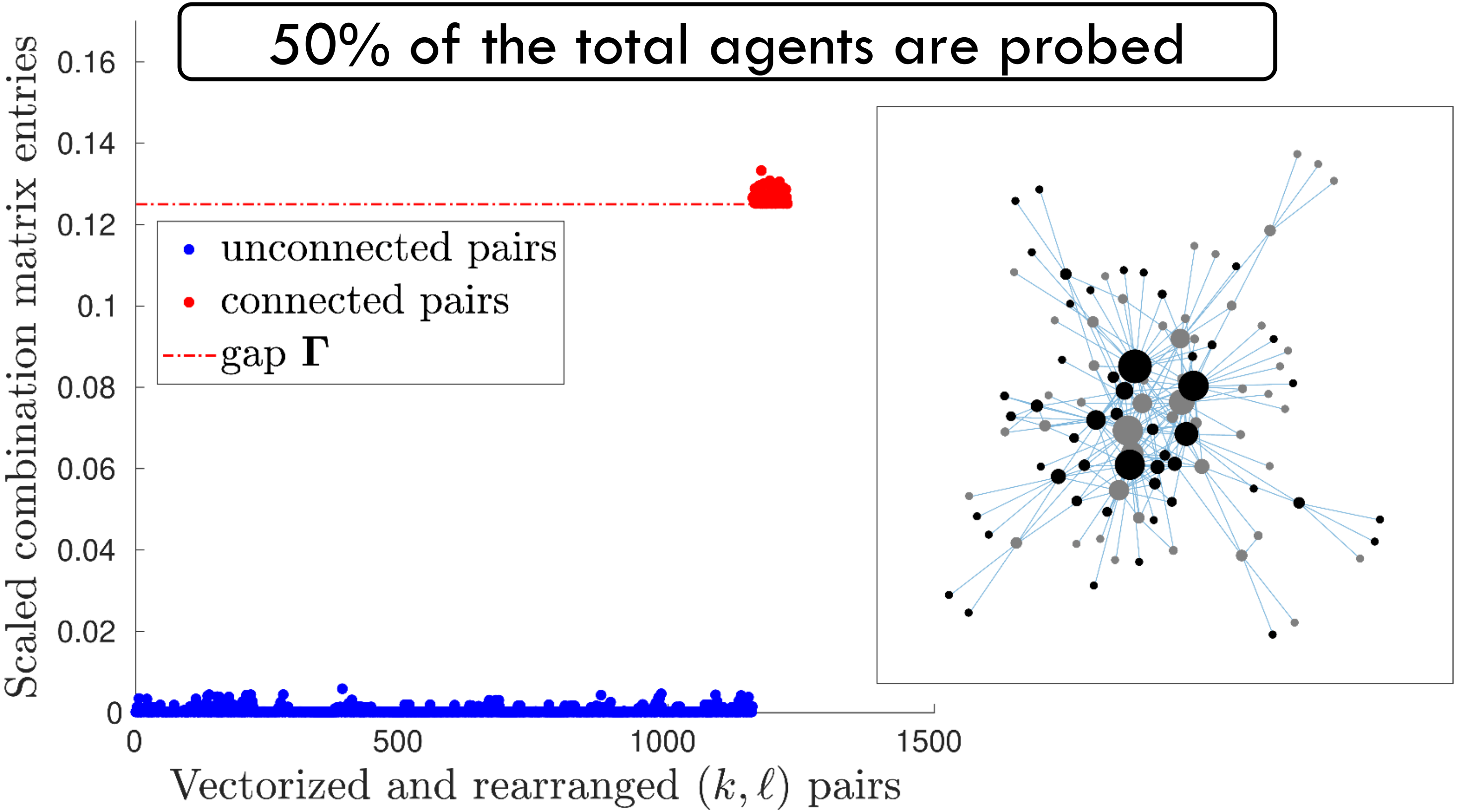
Performance



15% of the total agents are probed

Real Topologies: Network of Routers

50% of the total agents are probed



R. A. Rossi and N. K. Ahmed, “The Network Data Repository with Interactive Graph Analytics and Visualization,” in Proc. AAAI Conference on Artificial Intelligence (AAAI), 2015

Main Advances

- ❑ Results on achievability available under full observability

- ❑ Results on achievability available under partial observability for Graphical Models (no dynamics)

- ❑ Recent results on achievability available under partial observability and dynamical systems for Erdős-Rényi models (homogeneous setting, all nodes are equal)

Main Advances

- Results on achievability available under ~~full~~ observability

we allow for a large number of unobserved agents

- Results on achievability available under partial observability for Graphical Models (no dynamics)

- Recent results on achievability available under partial observability and dynamical systems for Erdős-Rényi models (homogeneous setting, all nodes are equal)

Main Advances

- Results on achievability available under ~~full~~ observability

we allow for a large number of unobserved agents

- Results on achievability available under partial observability for Graphical Models (no ~~dynamics~~)

we introduce dynamics

- Recent results on achievability available under partial observability and dynamical systems for Erdős-Rényi models (homogeneous setting, all nodes are equal)

Main Advances

- Results on achievability available under ~~full~~ observability

we allow for a large number of unobserved agents

- Results on achievability available under partial observability for Graphical Models (no ~~dynamics~~)

we introduce dynamics

- Recent results on achievability available under partial observability and dynamical systems for ~~Erdős-Rényi~~ models (homogeneous setting, all nodes are equal)

we consider preferential attachment models

The End