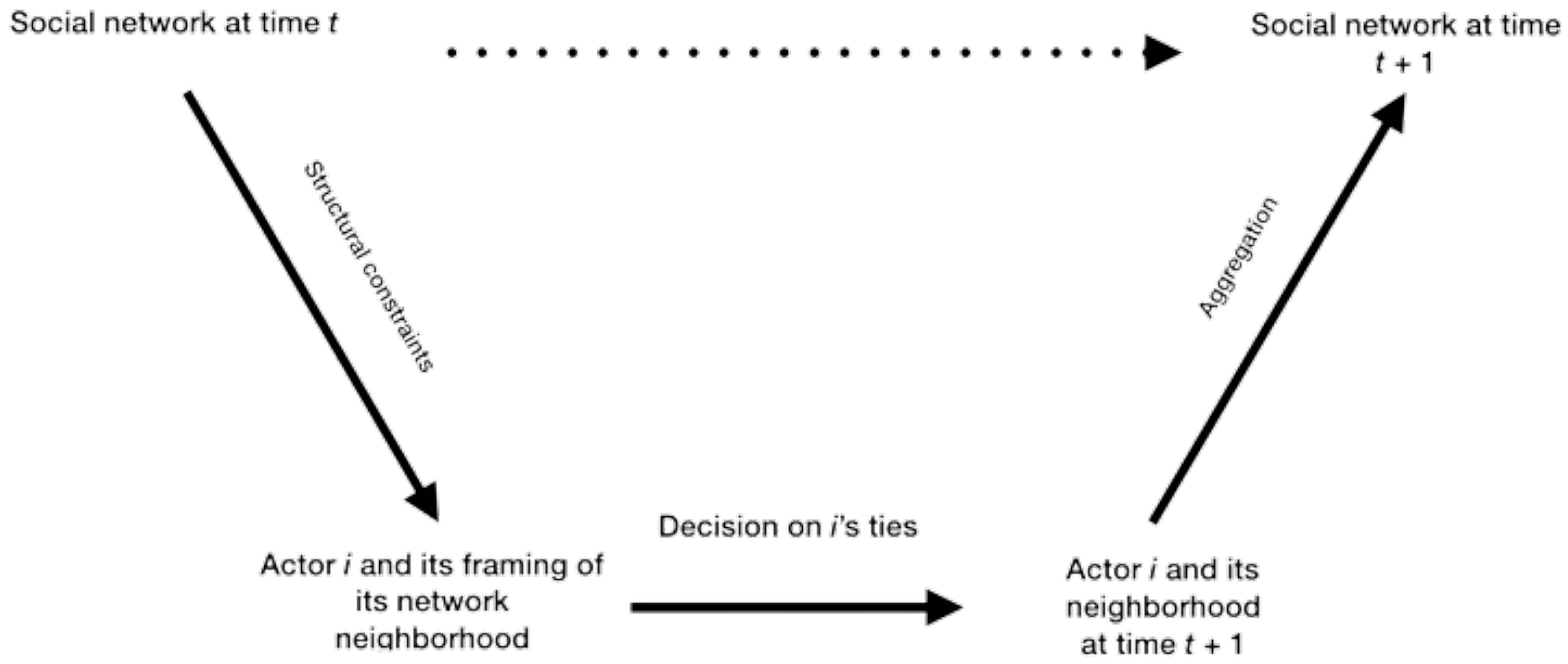
Statistical modelling of social networks

Social Network Analysis - ESOL AY 2025/26

8 May, 2026

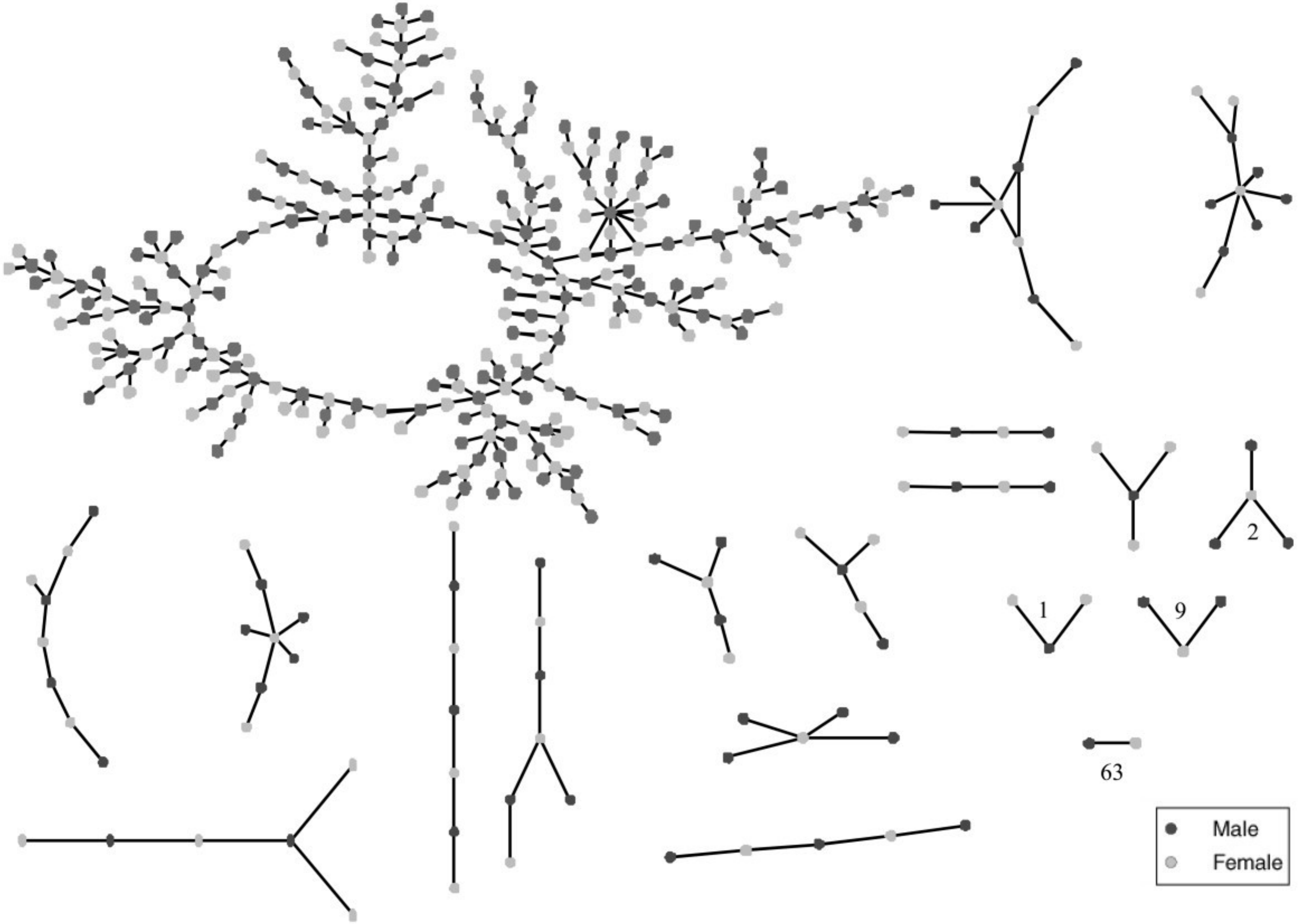
12 May, 2026

Federico Bianchi
Department of Social and Political Sciences

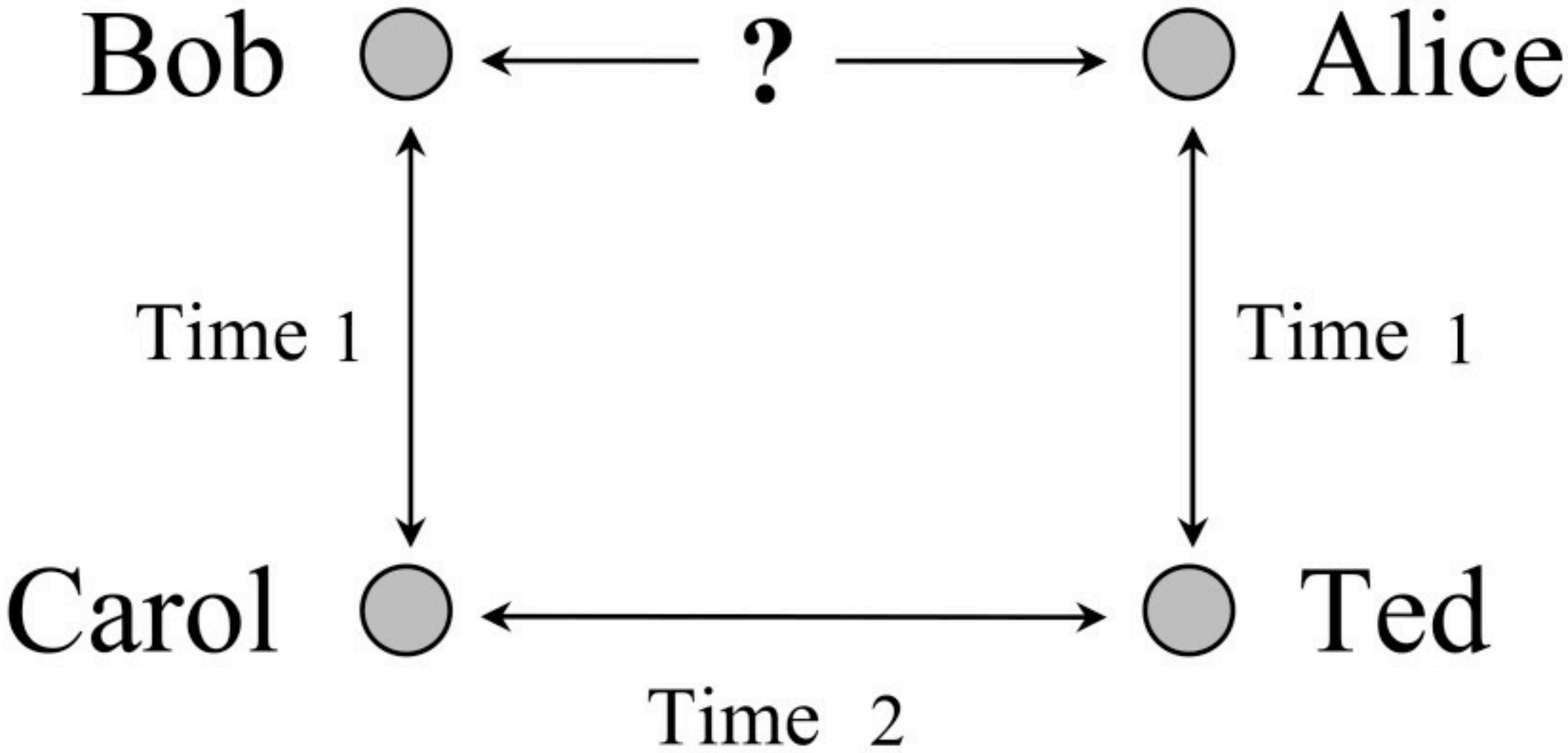


Causal mechanisms

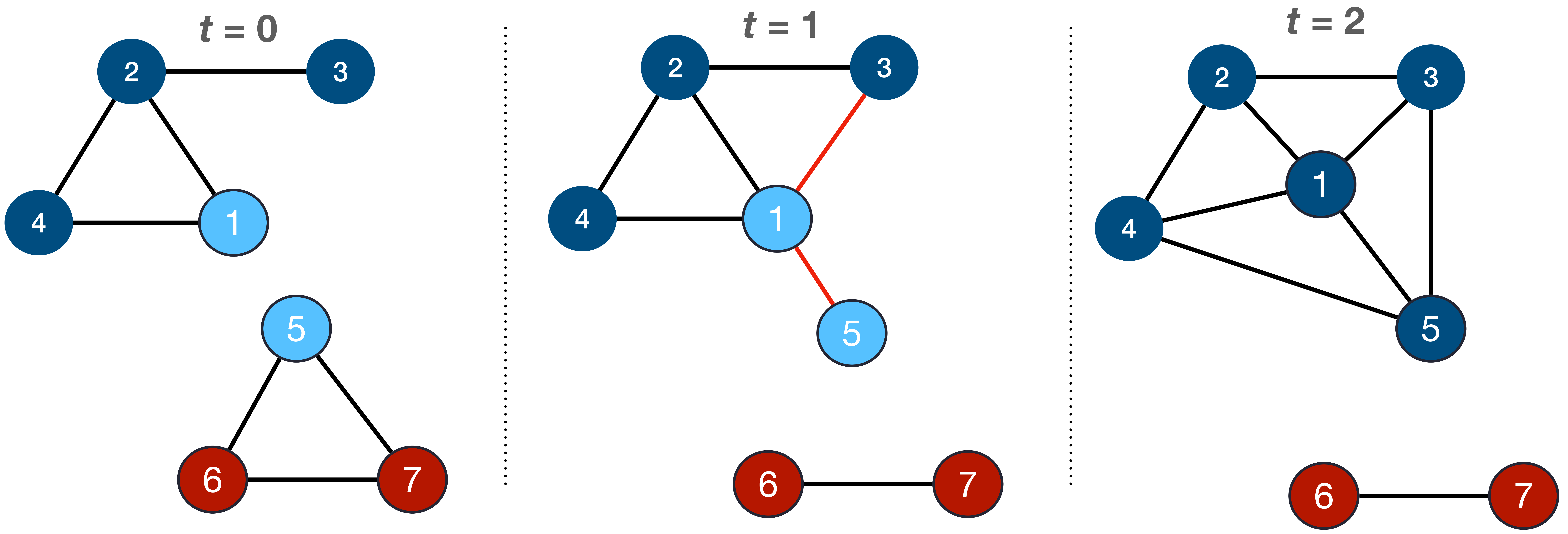
- **Explaining** vs. describing a social network
- Identifying **causal mechanisms** (i.e., a causal chain of events involving social actors' decisions; Hedström, 2005; Elster, 2015)
  - Formation of the network **structure**
  - **Composition** (diffusion of certain attributes)



Bearman, P., Moody, J., & Stovel, K. (2004). Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks. *American Journal of Sociology*, 110(1)

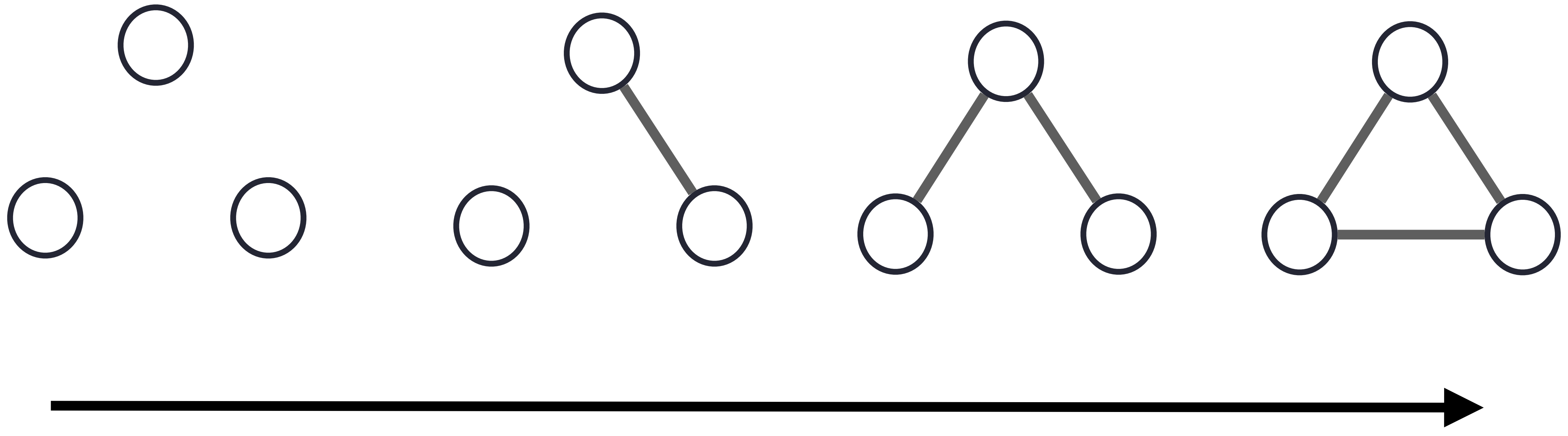


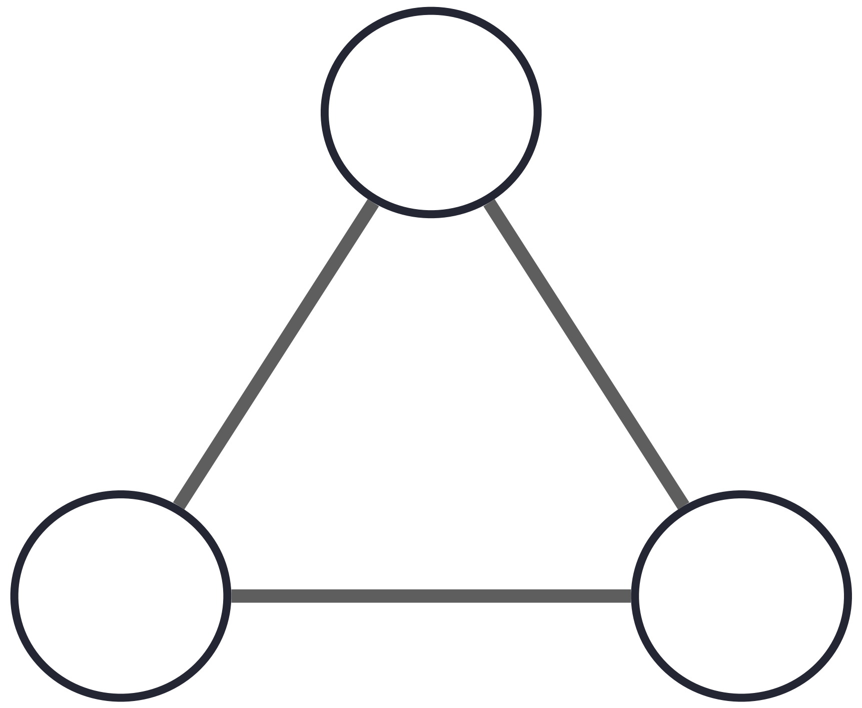
Bearman, P., Moody, J., & Stovel, K. (2004). Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks. *American Journal of Sociology*, 110(1)

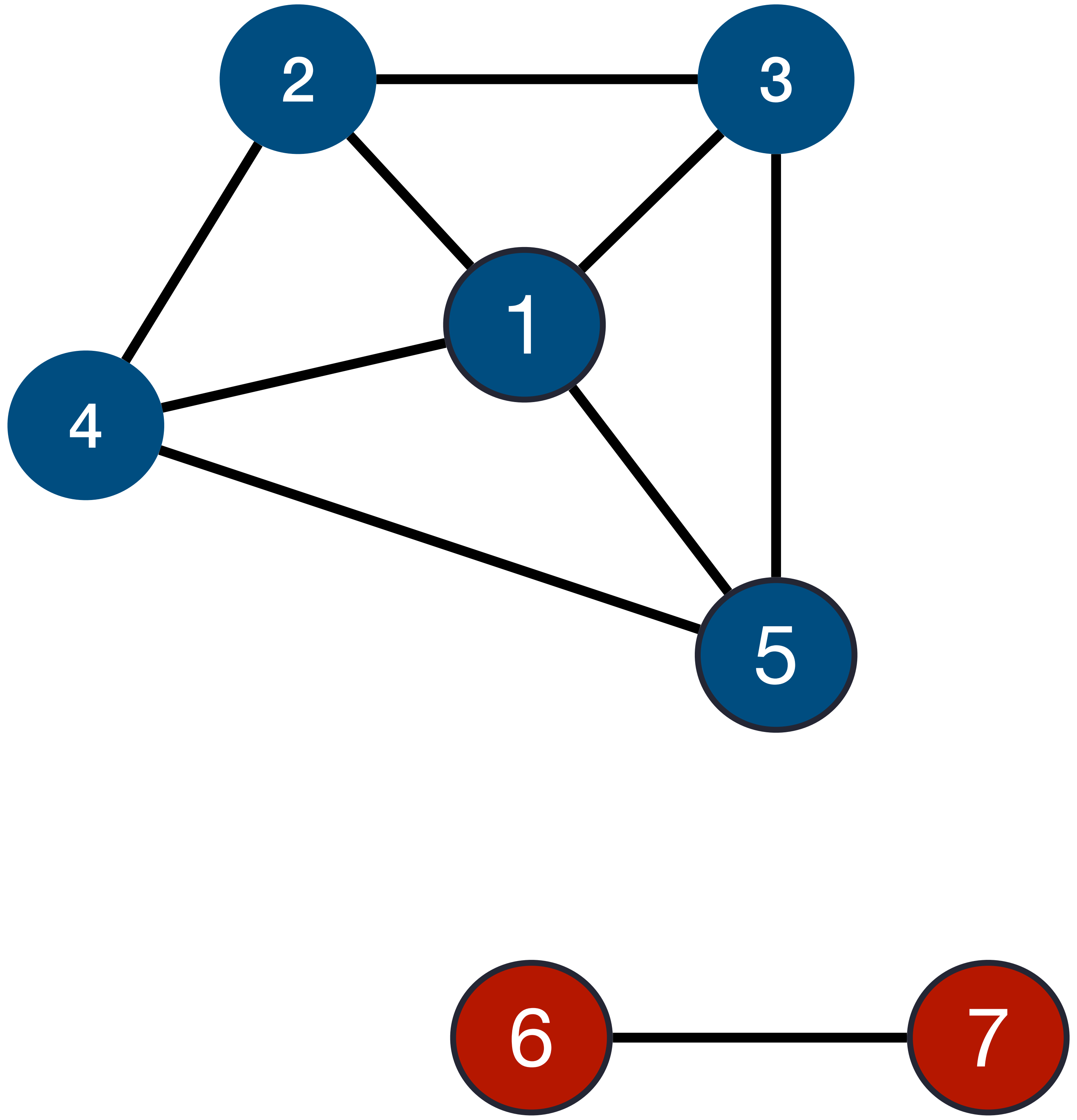


Example: a friendship network and musical tastes in a workplace

Identifying a **network mechanism** —> describing a regular pattern of ties and attributes







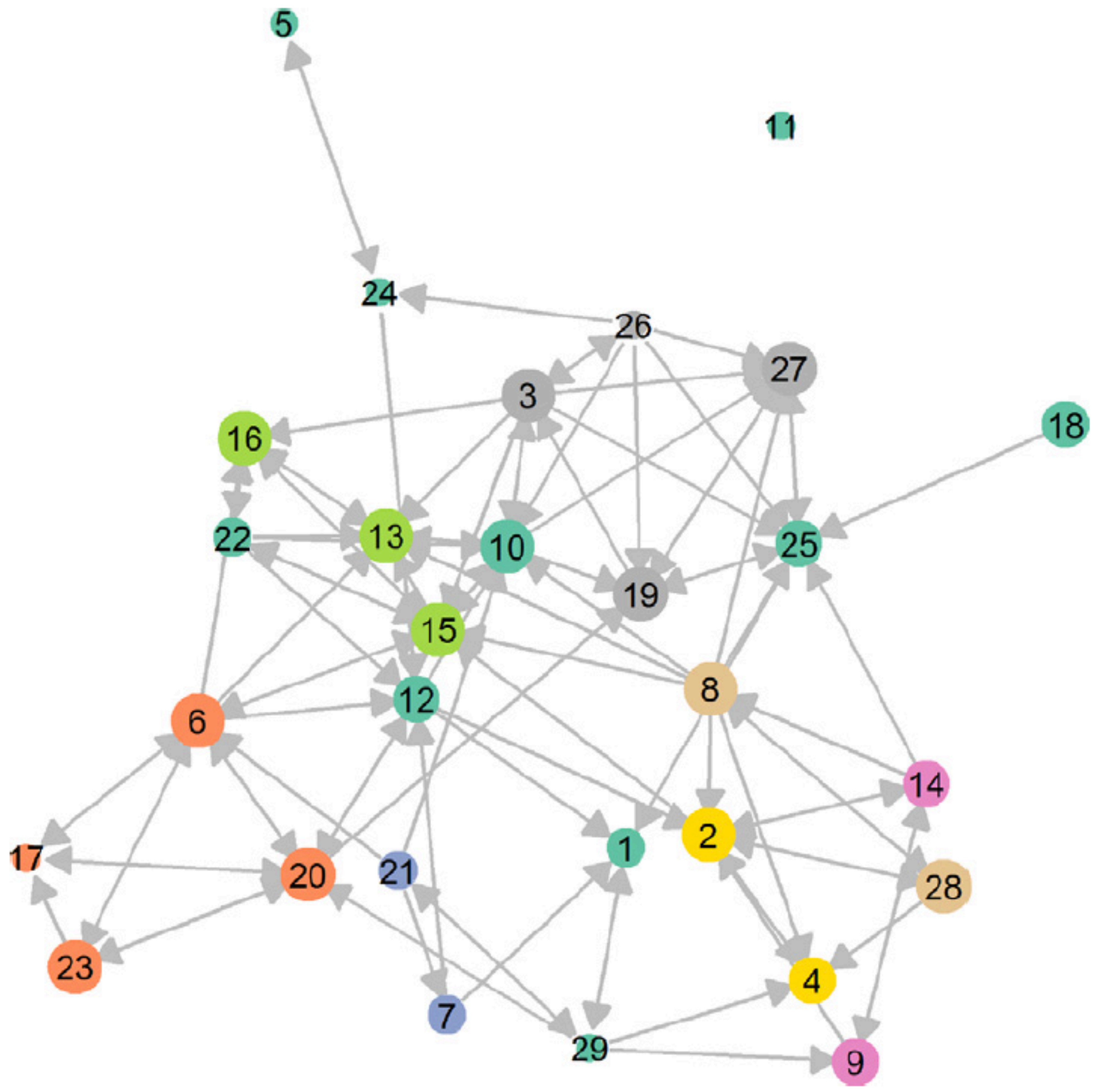
- Inferring the effect of **unobserved**, dynamic relational processes on the evolution of a network from the **prevalence** or **incidence** of certain **local configurations**

Statistical models of social networks

- Network local configurations as “archeological traces” left by causal mechanisms (White, 1970; Lusher et al., 2013)

(Statistical inference)

- The relative effect size of these processes can be estimated by computing **statistics of empirical network data** —> Maximum likelihood or method of moments (numerical simulations)

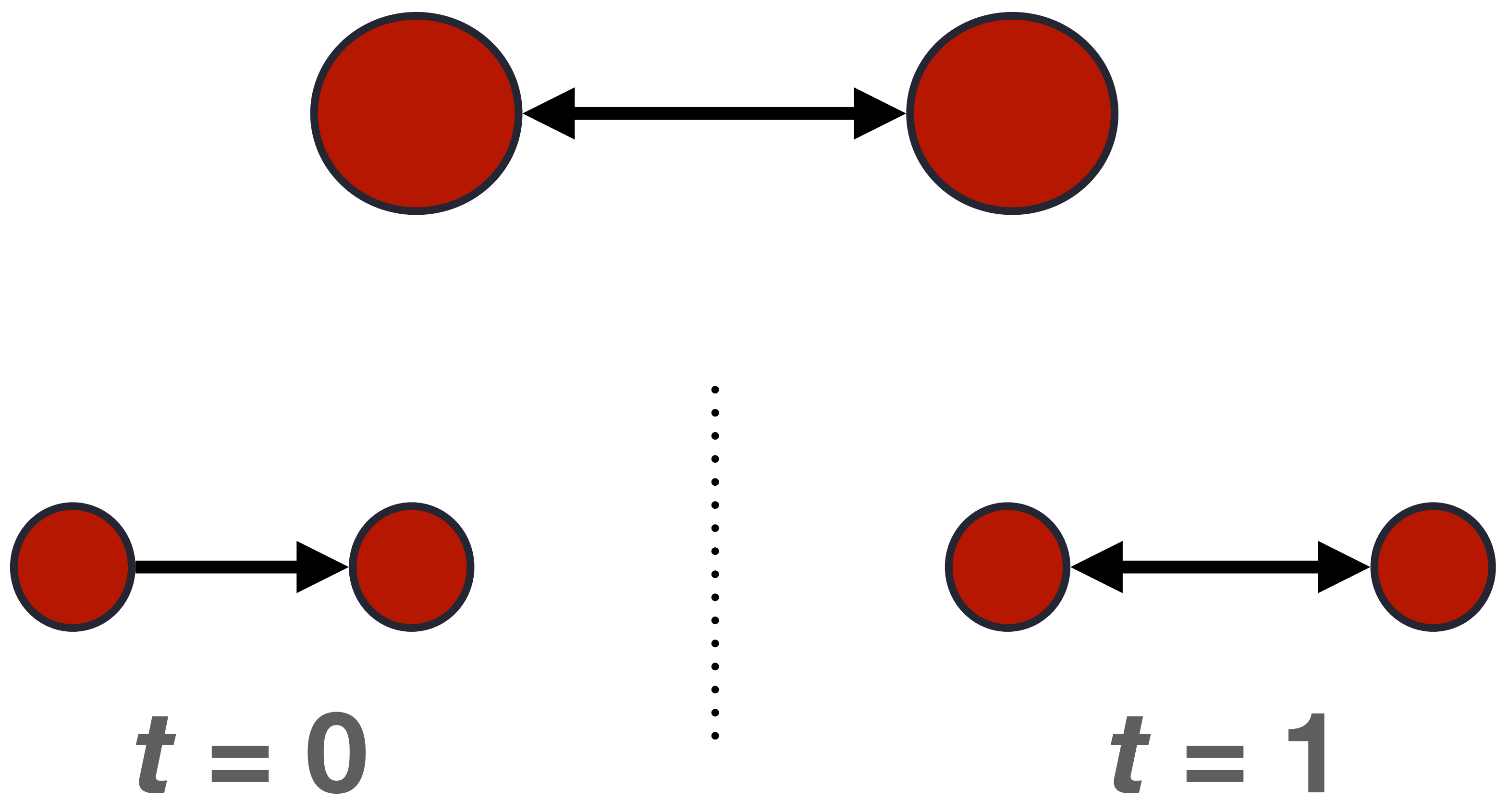


Practical example:

Reciprocity in a coworking space

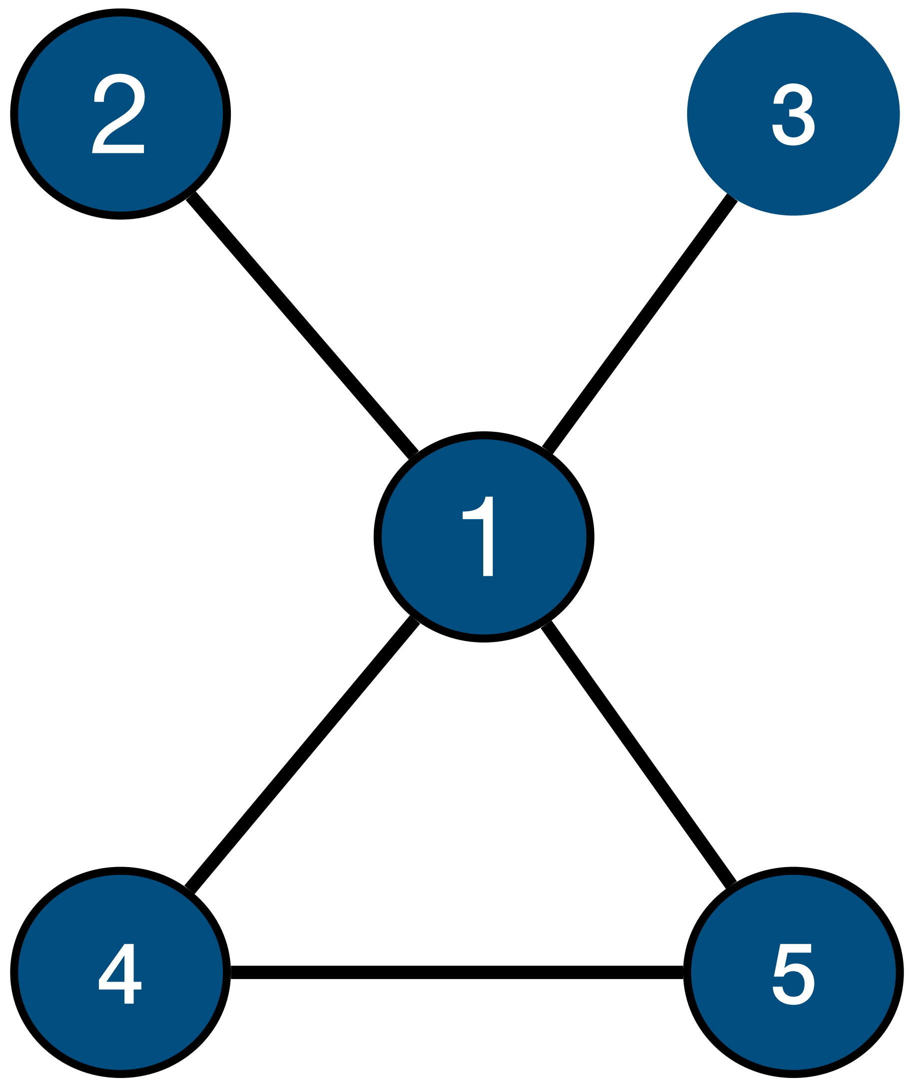
Name generator:

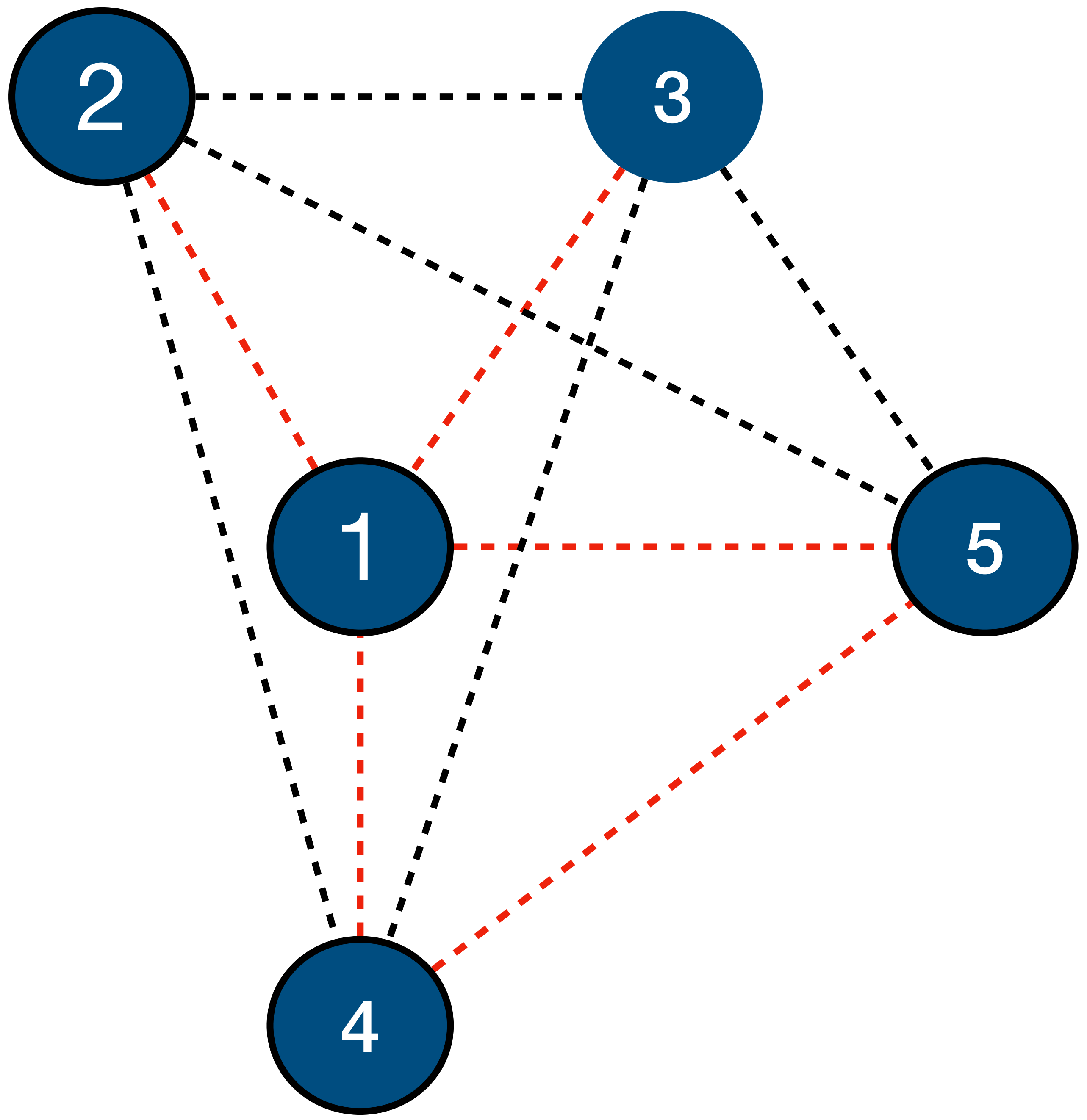
Who do you usually turn to if you need emotional or material support?



Statnet

- Suite of packages for statistical analysis of network data
- Used primarily in the sociological/anthropological/psychological community of SNA
- It includes:
    - `network`: data structure for handling and manipulating network objects in R
    - `sna`: tools for descriptive statistics (connectivity, centrality, clustering, etc.)
    - `ergm`: Exponential Random Graph Models (next session)
    - Other (`tergm`, `stergm`, `latentnet`, `ergm.ego`, etc.)

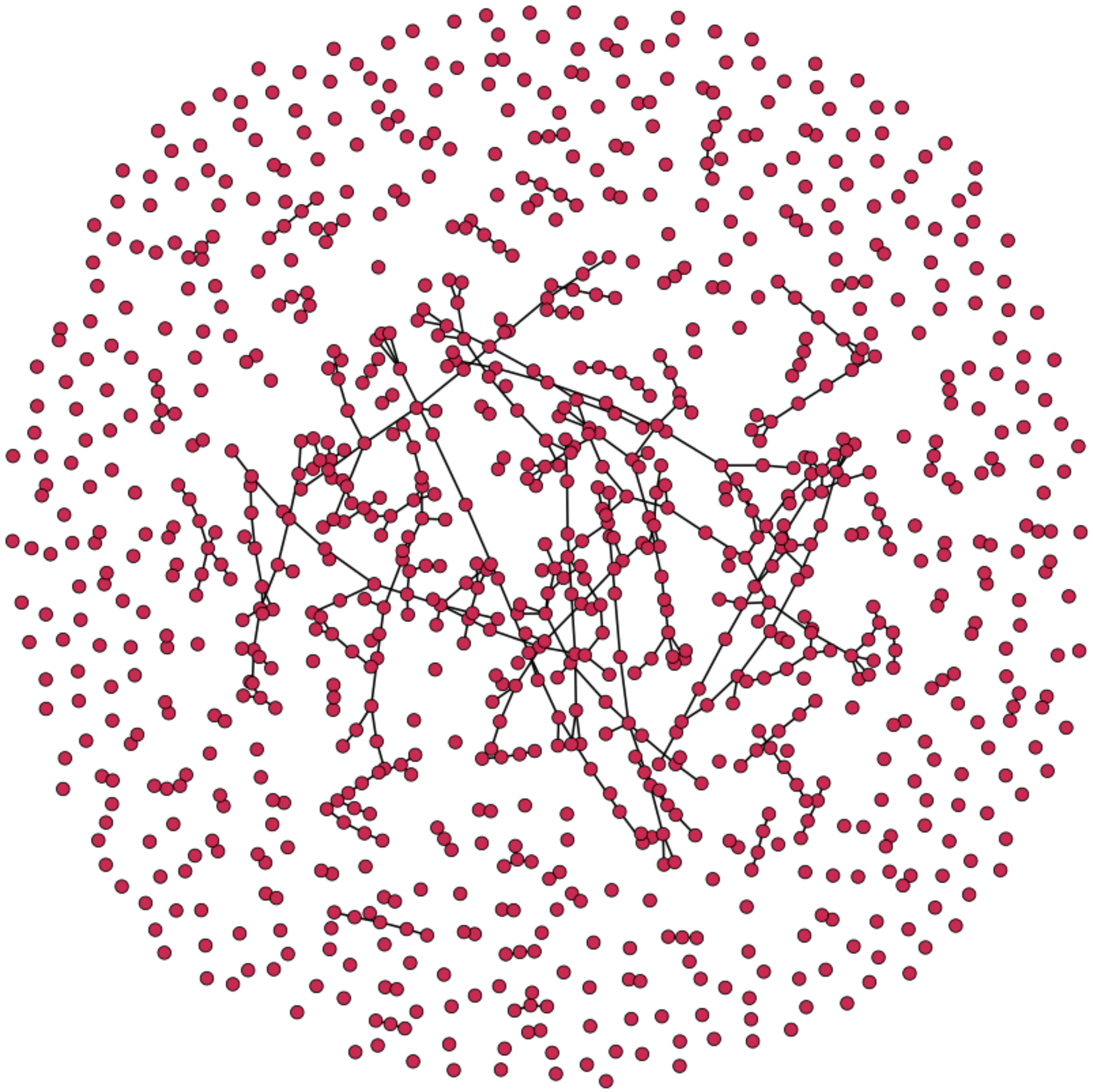




Random

graph models

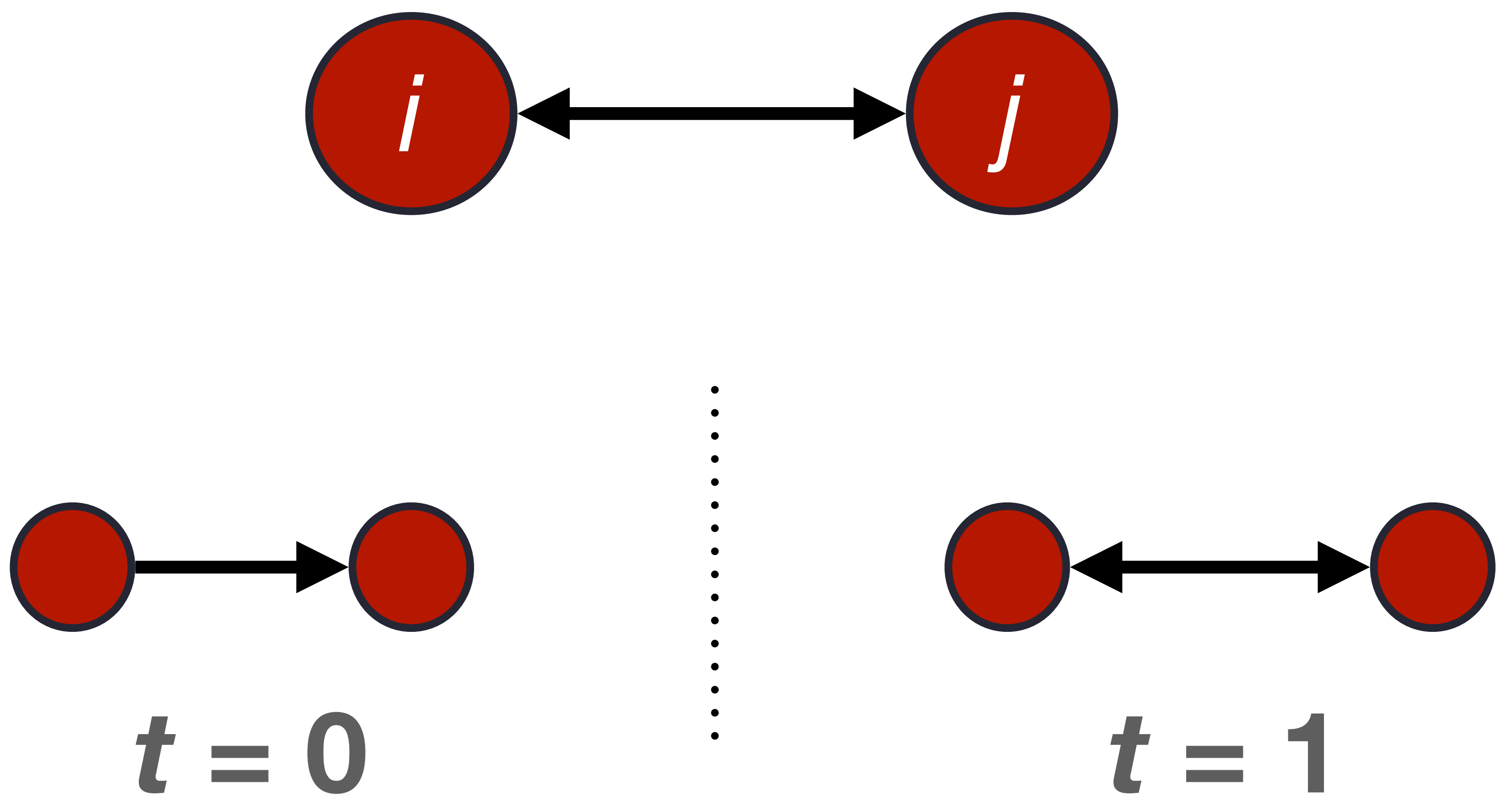
- Random (stochastic) graph model: a family of random tie-variables with a fixed number of nodes *n*
- The observed ties are only a subset of the set of all possible ties
- For each pair of nodes *i* and *j*, *X_ij* is a **random tie variable**
- An observed network is just a realization of all the possible of a random graph



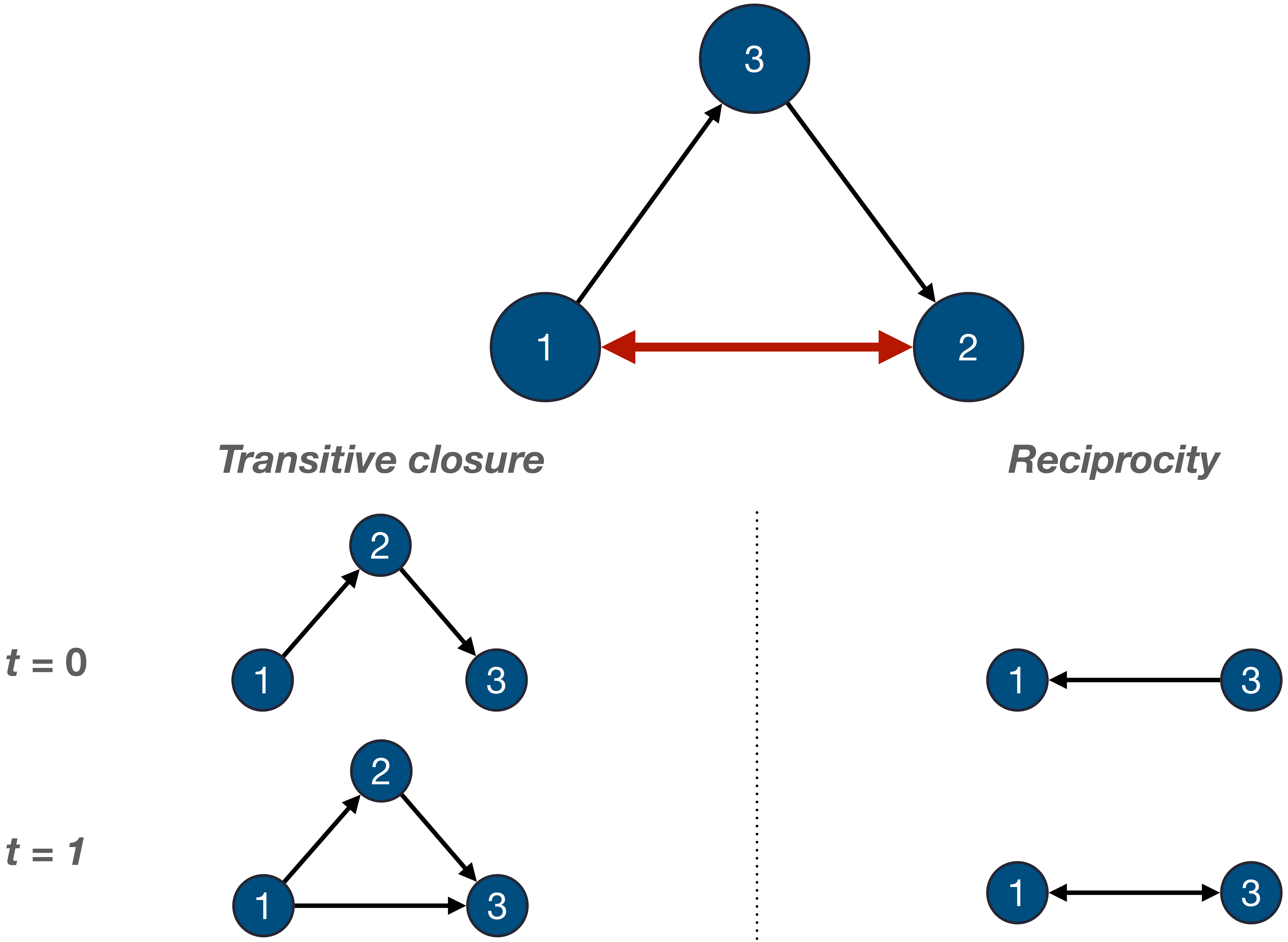
Erdős-Rényi-Gilbert Model

- Random graph *G(n, p)*
- *n*: number of nodes
- *p*: probability that *X_ij* = 1
- Tie-variables are identically distributed and independent
- Bernouilli process

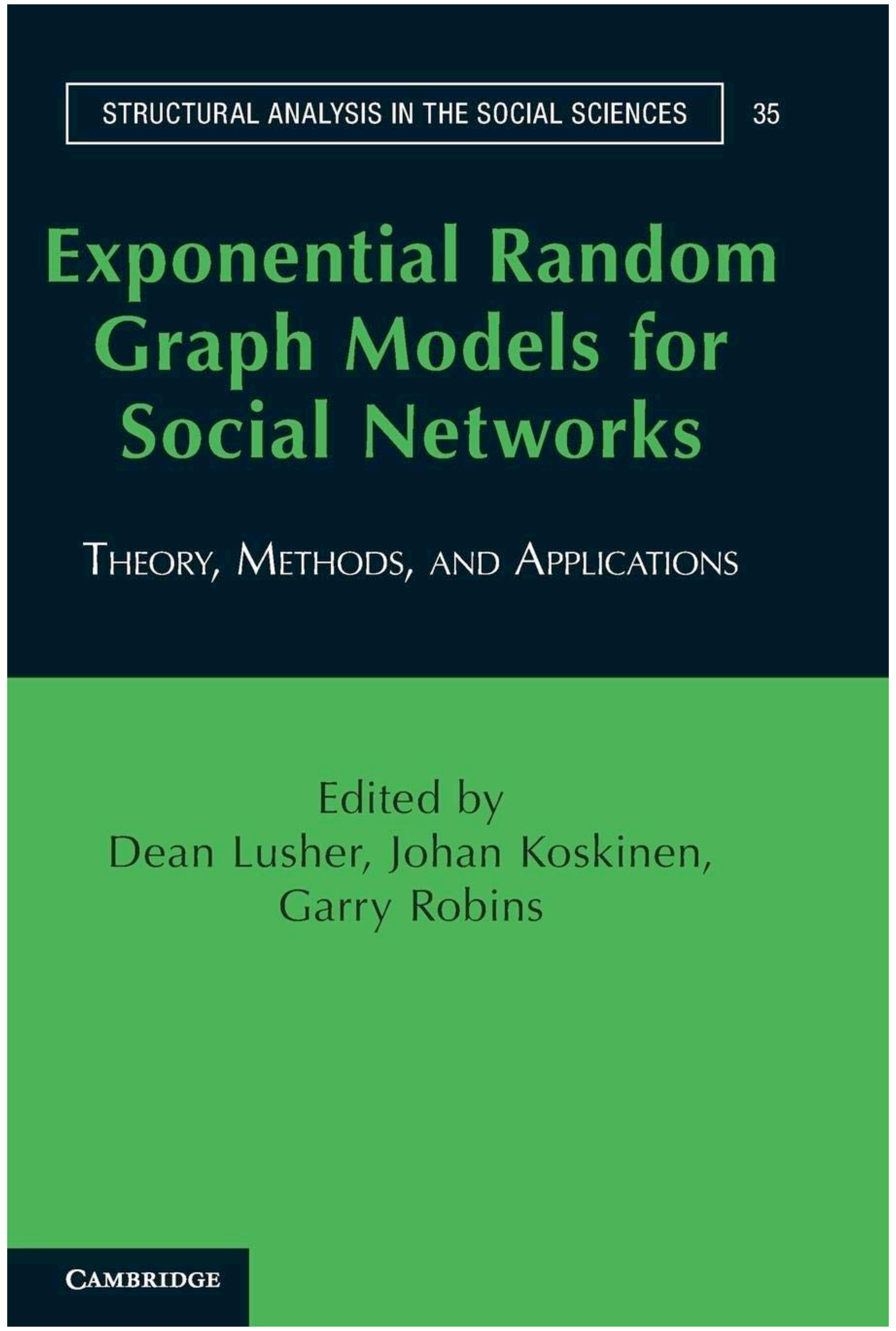
Stochastic dependence of observations



- E-R model assumes independence of observations
- Pr(X_ji = 1) stochastically depends on Pr(X_ij = 1)

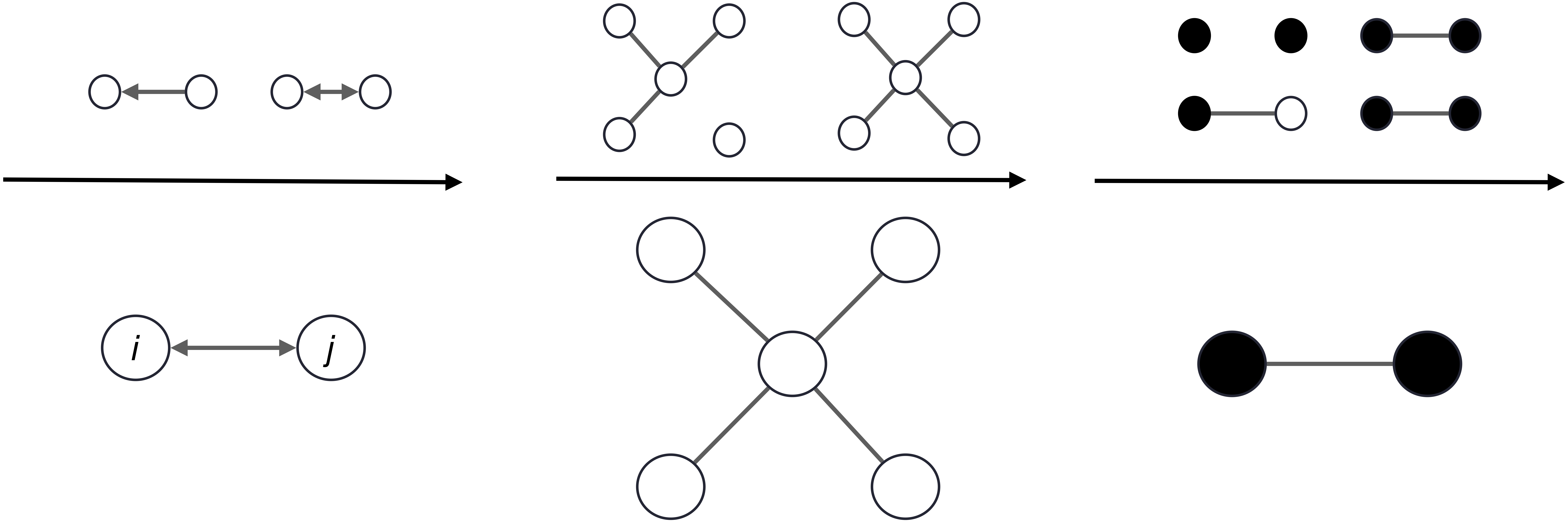

3
1
2
Transitive closure
Reciprocity
t = 0
t = 1
2
1
3
1
3
2
1
3
1
3

ERGM: Exponential Random Graph Models



$$Pr(X = x \,|\, \theta) = \frac{1}{\kappa(\theta)} \exp \theta_1 z_1(x) + \theta_2 z_2(x) + \ldots + \theta_p z_p(x)$$

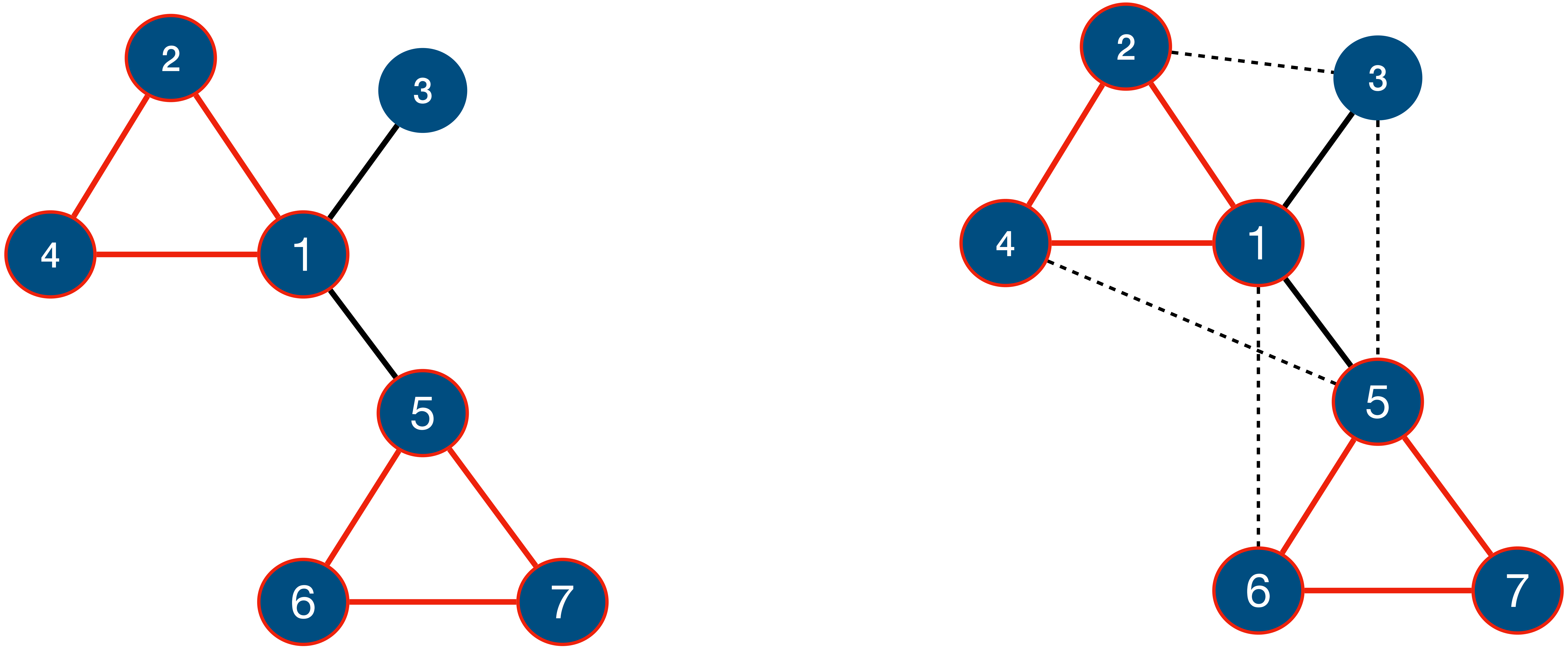
- $Pr(X = x \,|\, \theta)$: the likelihood to observe a given graph x (realization of the random graph X)
- The functions $z_k(x)$ are count statistics of local graph configurations (traces of the processes assumed to have generated x)
- The parameters θ_k weight the relative importance of the count statistics, thereby expressing their effect size
- Maximum Likelihood Estimation



Statistical models of social networks:

local configurations and stochastic dependency assumptions

- A relational process can be linked to a **local configuration**, of which **count statistics** can be computed
- Observations are **not independent**
- Each local configuration comes with a **stochastic dependency assumption**: es., $P(x_{ij}) \cap P(x_{ji}) = P(x_{ij} | x_{ji}) \cdot P(x_{ji})$

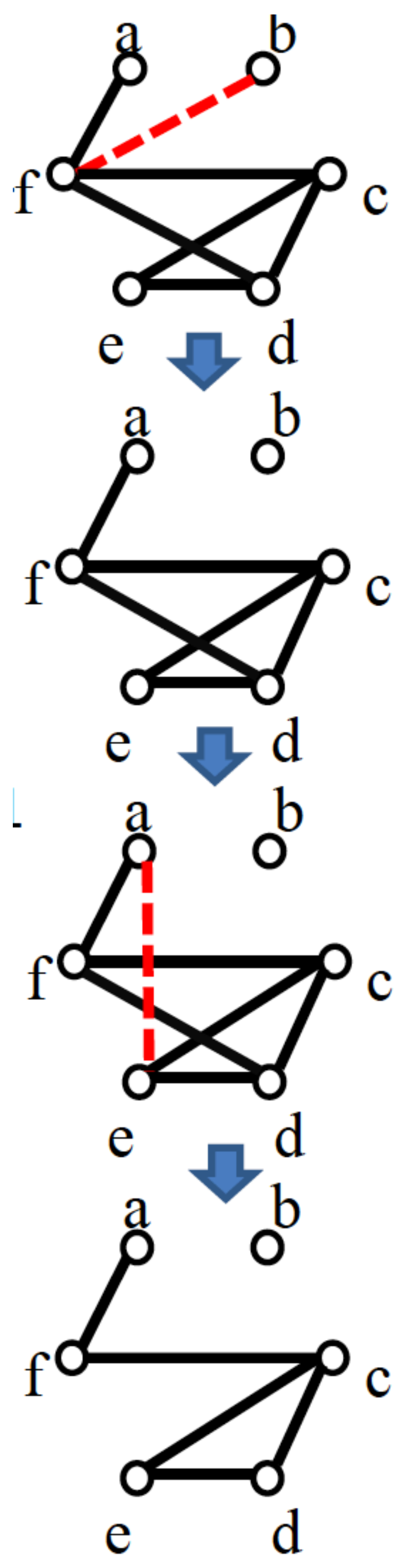


Statistical models of social networks:

hypothesis testing

- Generating (simulating) a random graph distribution centred on the observed statistics
- Identifying a parameter vector
- Computing uncertainty measures (hypothesis testing)

Markov Chain Monte Carlo Maximum Likelihood Estimation



1. Choose a parameter vector (i.e., assign a random value to specified parameters)

2. Start with a random network with the given number of nodes

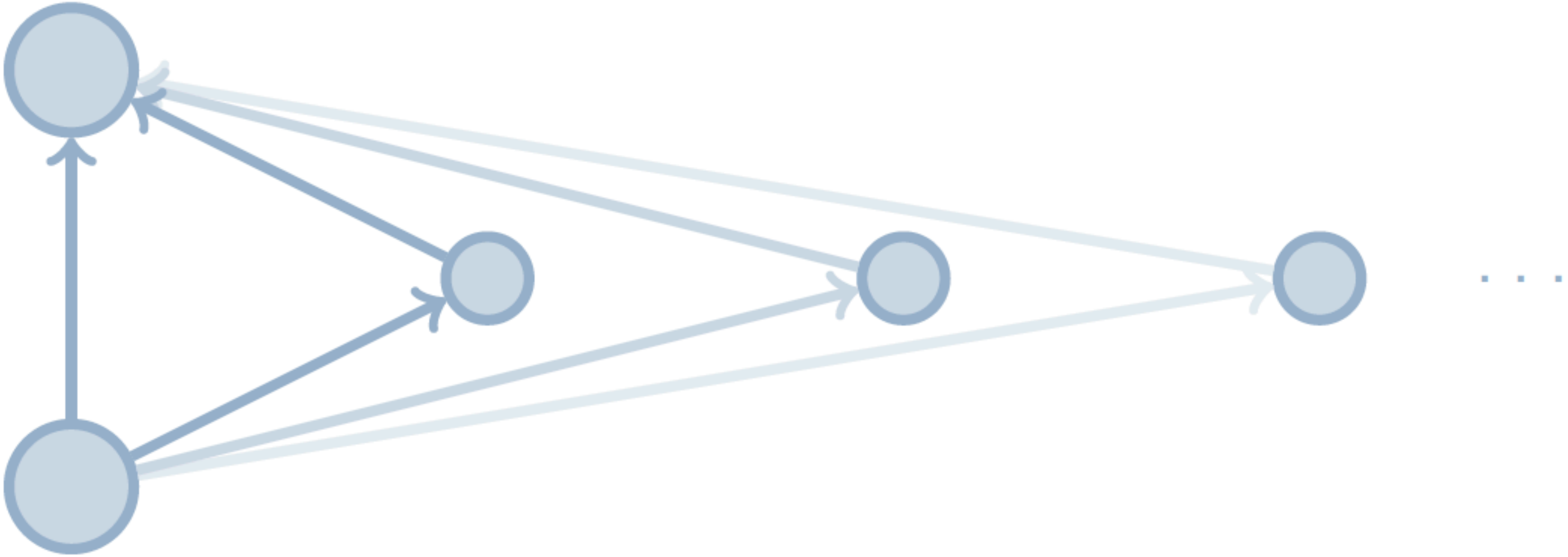
3. Select a random dyad

4. Stochastically update the value of the selected dyad according to the parameter vector at 1.

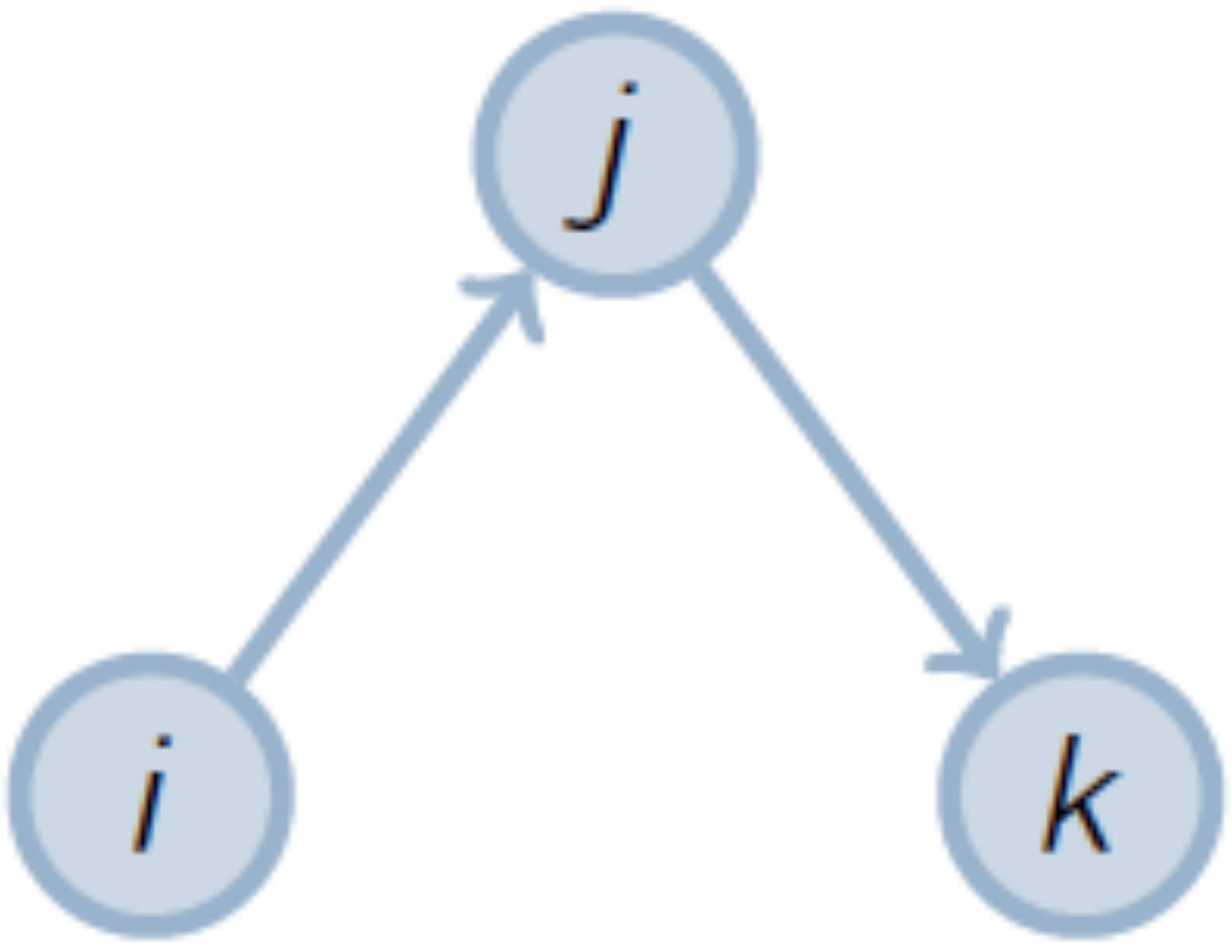
5. Repeat 3. and 4.

Output: The process will eventually converge (**Markov chain**) to a random graph distribution that has the count statistics of the observed network as a central tendency (**maximum likelihood**)

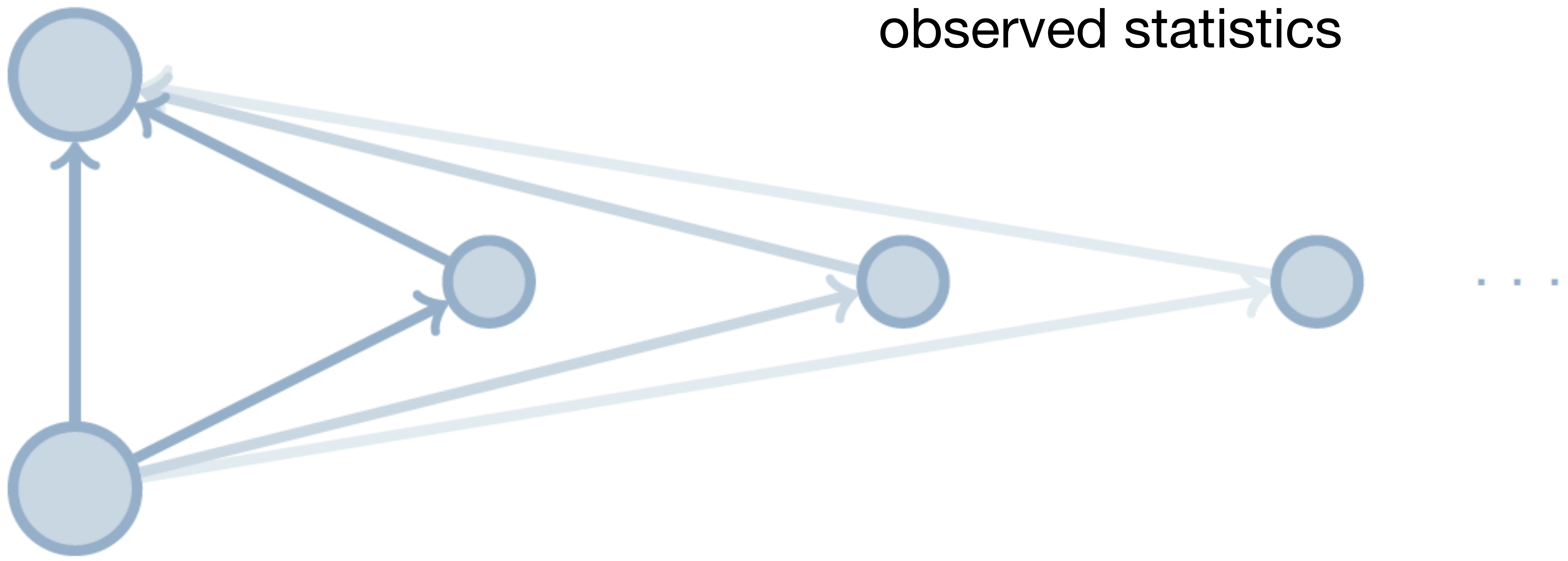
GWESP: Geometrically Weighted Edgewise Shared Partners



Counts the number of closed triads for each pair of nodes (with a marginally decreasing effect α

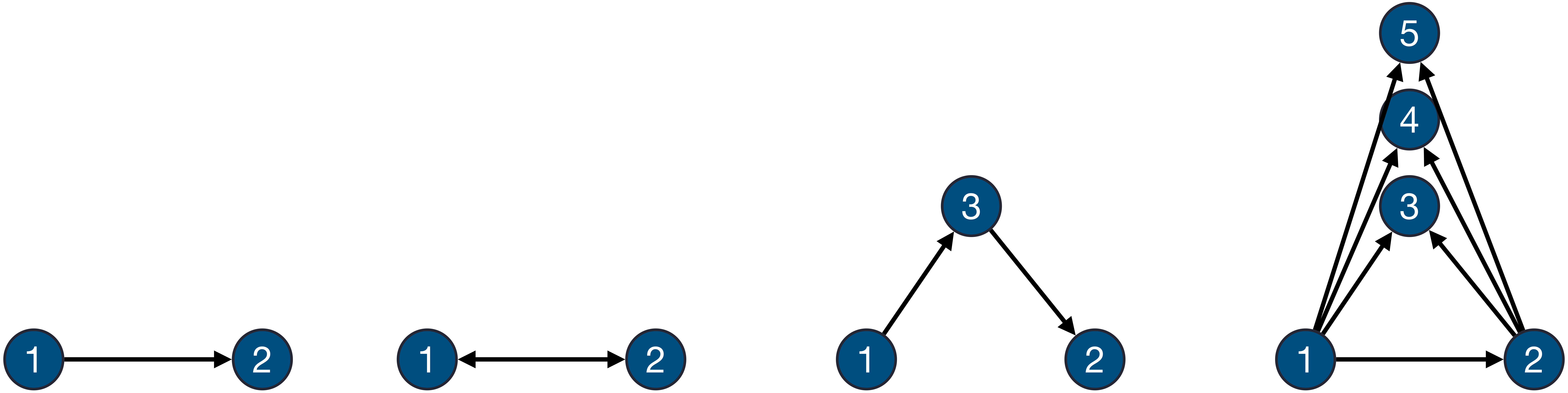


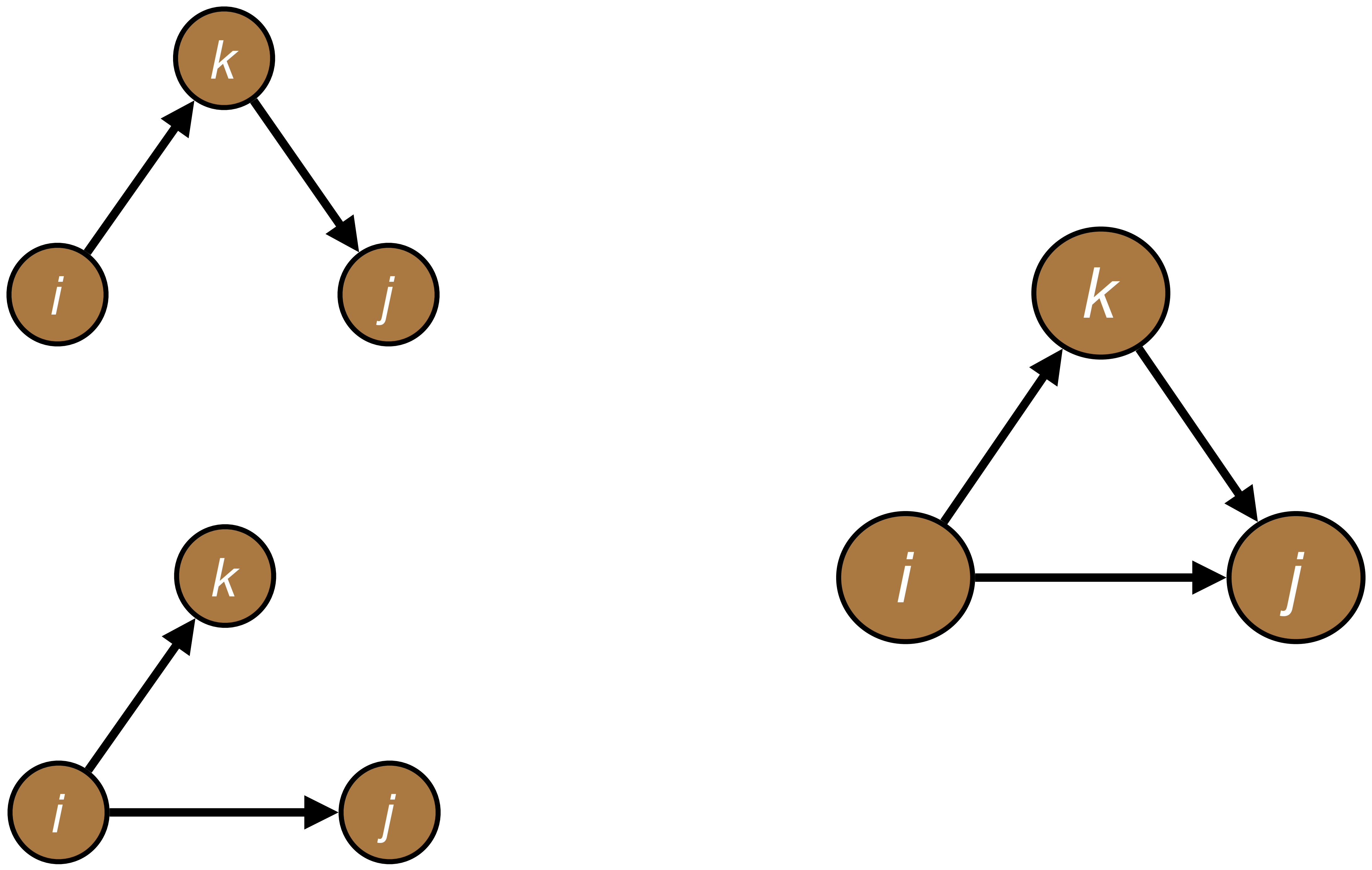
We need to specify one more statistic: the number of open triads, otherwise the model might not converge because the estimation algorithm does not have enough information to generate a random graph distribution centred on the observed statistics



Our model

$$Pr(X = x \mid \theta) = \frac{1}{\kappa(\theta)} \exp \theta_1 \cdot \text{edges} + \theta_2 \cdot \text{mutual dyads} + \theta_3 \cdot \text{open triads} + \theta_4 \cdot \text{GWESP}$$





ERGM:

Limits

- **Tie-based models** (ERGM-family; Lusher et al., 2013): it samples ties, not nodes (agents)

- the occurrence of a tie is assessed independently on **agents' multinomial choice**, typical of many decision-making contexts

- are **indifferent to the specific tie sequences** through which particular configurations emerge (Block et al., 2019)

Background reading

Lusher et al. (2013), Ch. 2-4

Koskinen (2024)

References

Bearman, P., Moody, J., & Stovel, K. (2004). Chains of affection: the structure of adolescent romantic and sexual networks. *American Journal of Sociology*, 110(1): 44-91. doi: 10.1086/386272

Block, P., Stadtfeld, C., & Snijders, T.A.B. (2019). Forms of dependence: comparing SAOMs and ERGMs from basic principles. *Sociological Methods & Research*, 48(1): 202-239. doi: 10.1177/0049124116672680

Elster, J. (2007). *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press

Hedström, P. (2005). *Dissecting the Social: On the Principles of Analytical Sociology*. Cambridge: Cambridge University Press

Koskinen, J. (2024). Exponential Random Graph Models. In McLevey, J., Scott, J., & Carrington, P.J. (Eds.), *The Sage Handbook of Social Network Analysis*, 2nd ed. (pp. 474-500). London: Sage.

Lusher, D., Koskinen, J., & Robins, G. (2013). *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*. Cambridge: Cambridge University Press

White, H.C. (1970). Matching, vacancies, and mobility. *Journal of Political Economy*. 78(1): 97-105. doi: 10.1086/259604