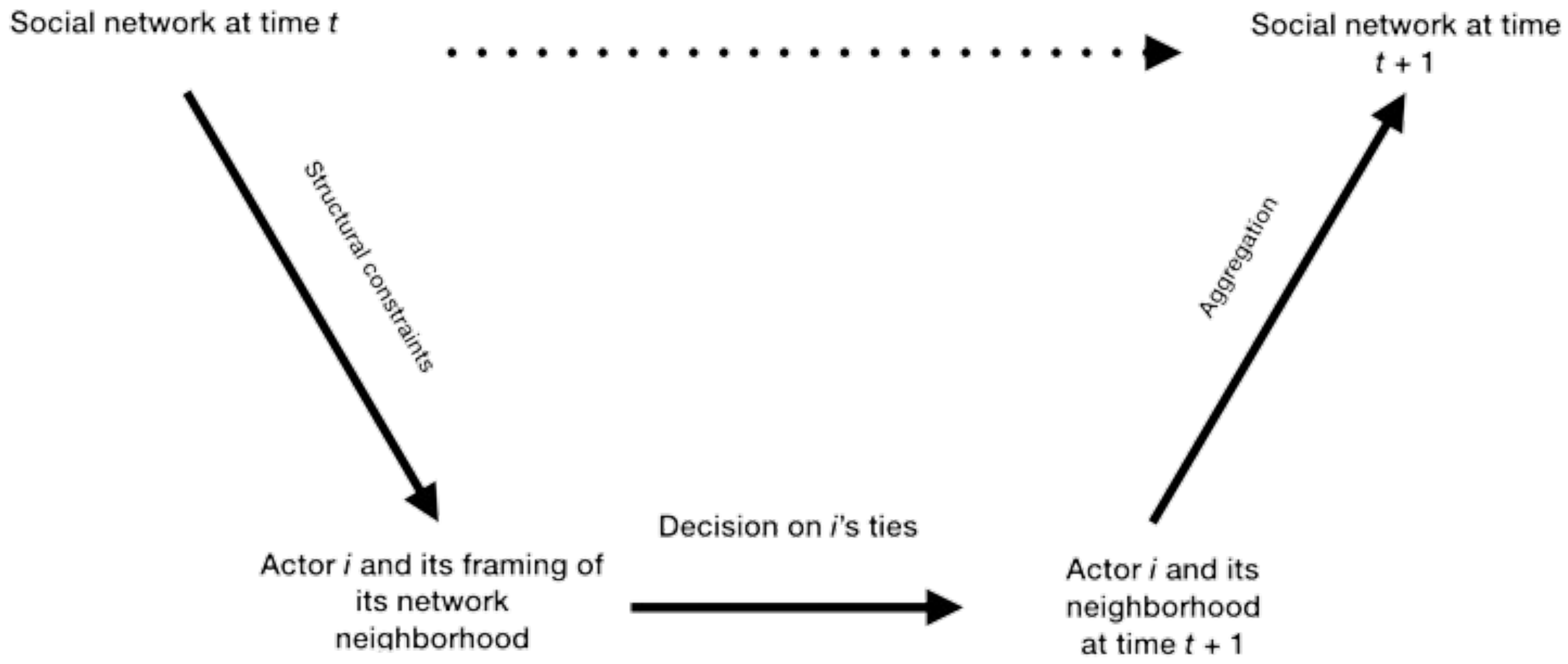
Random graphs

Social Network Analysis - ESOL AY 2024/25
13 May, 2025

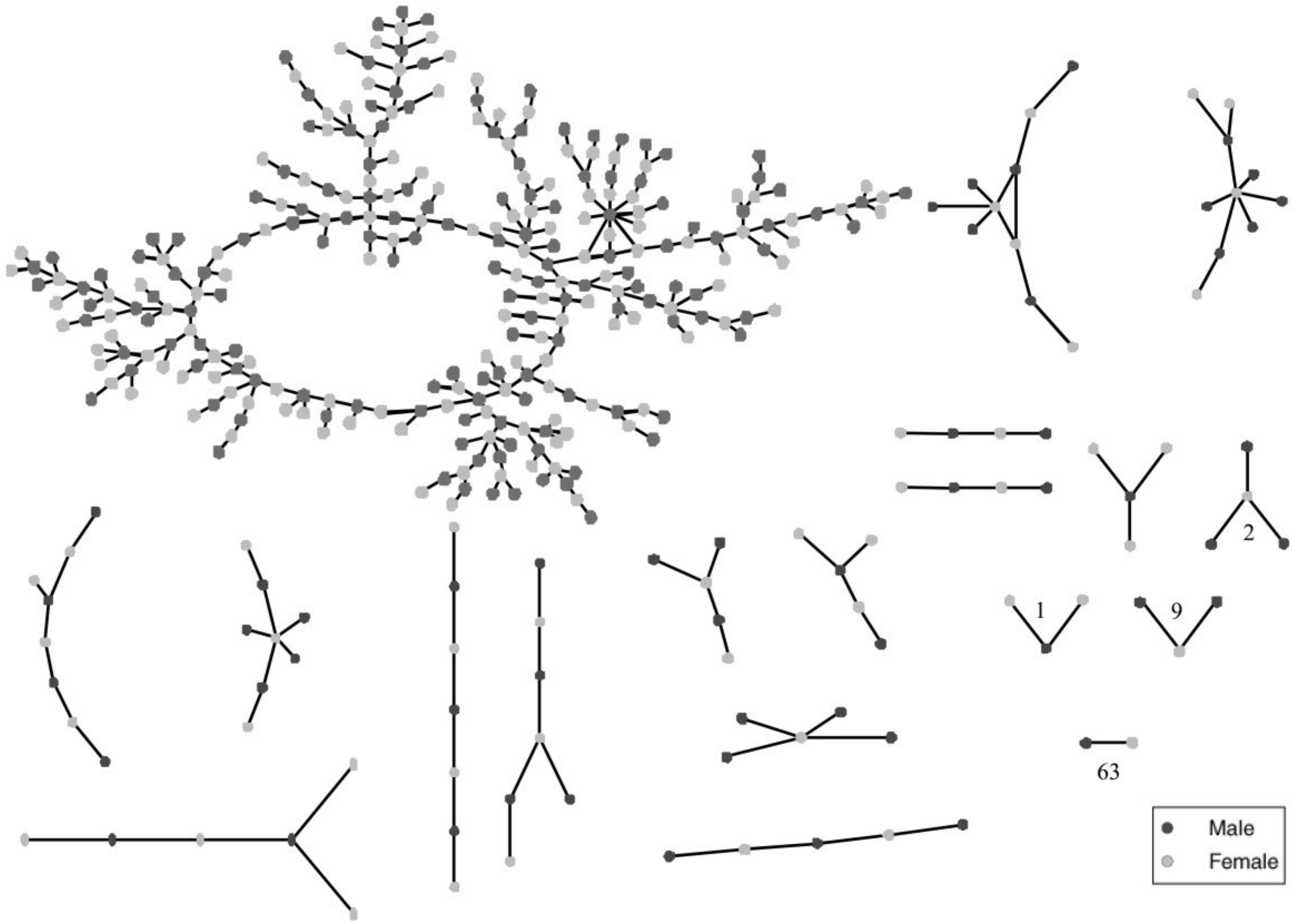
Federico Bianchi
Department of Social and Political Sciences

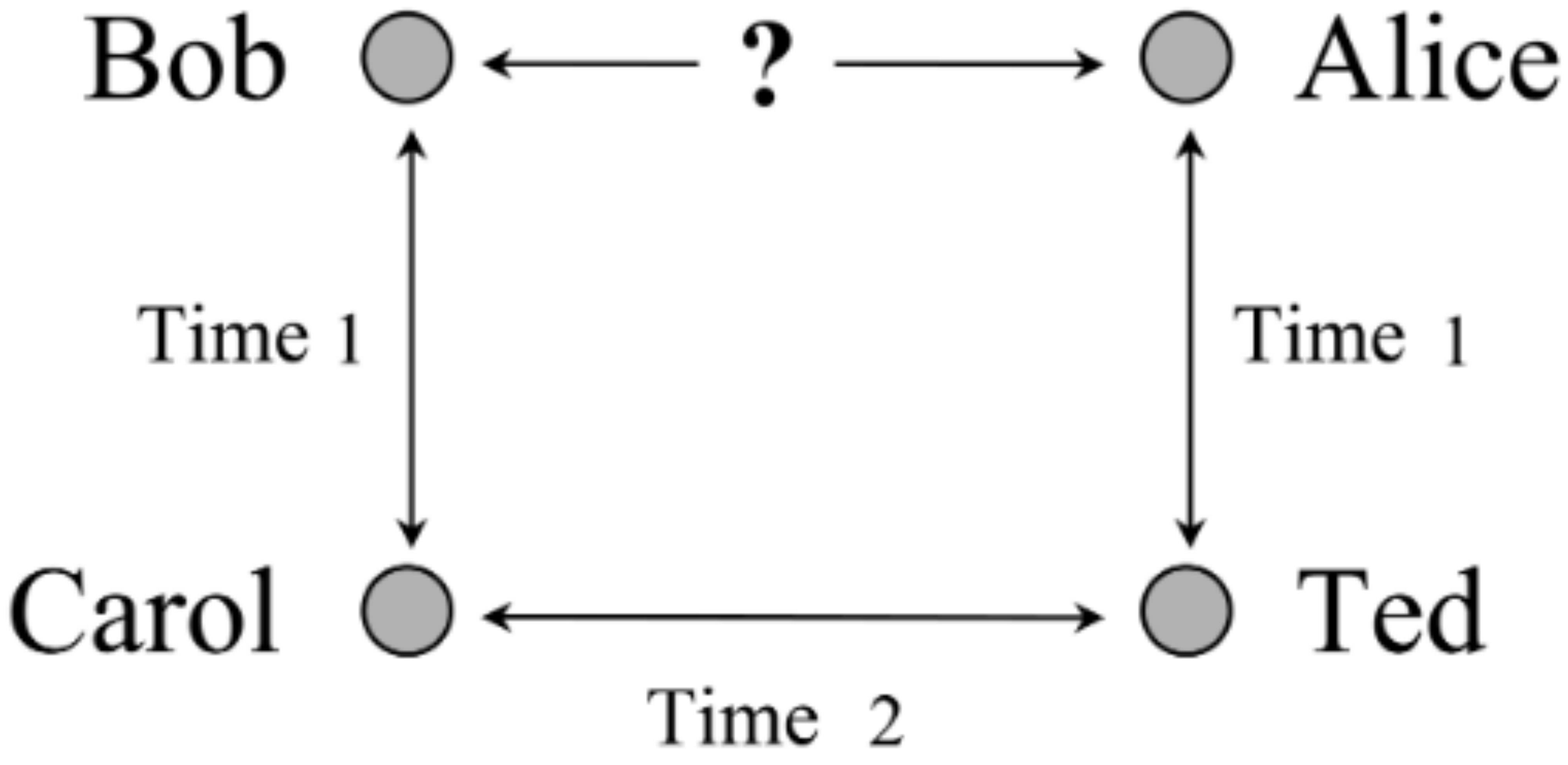


Causal mechanisms

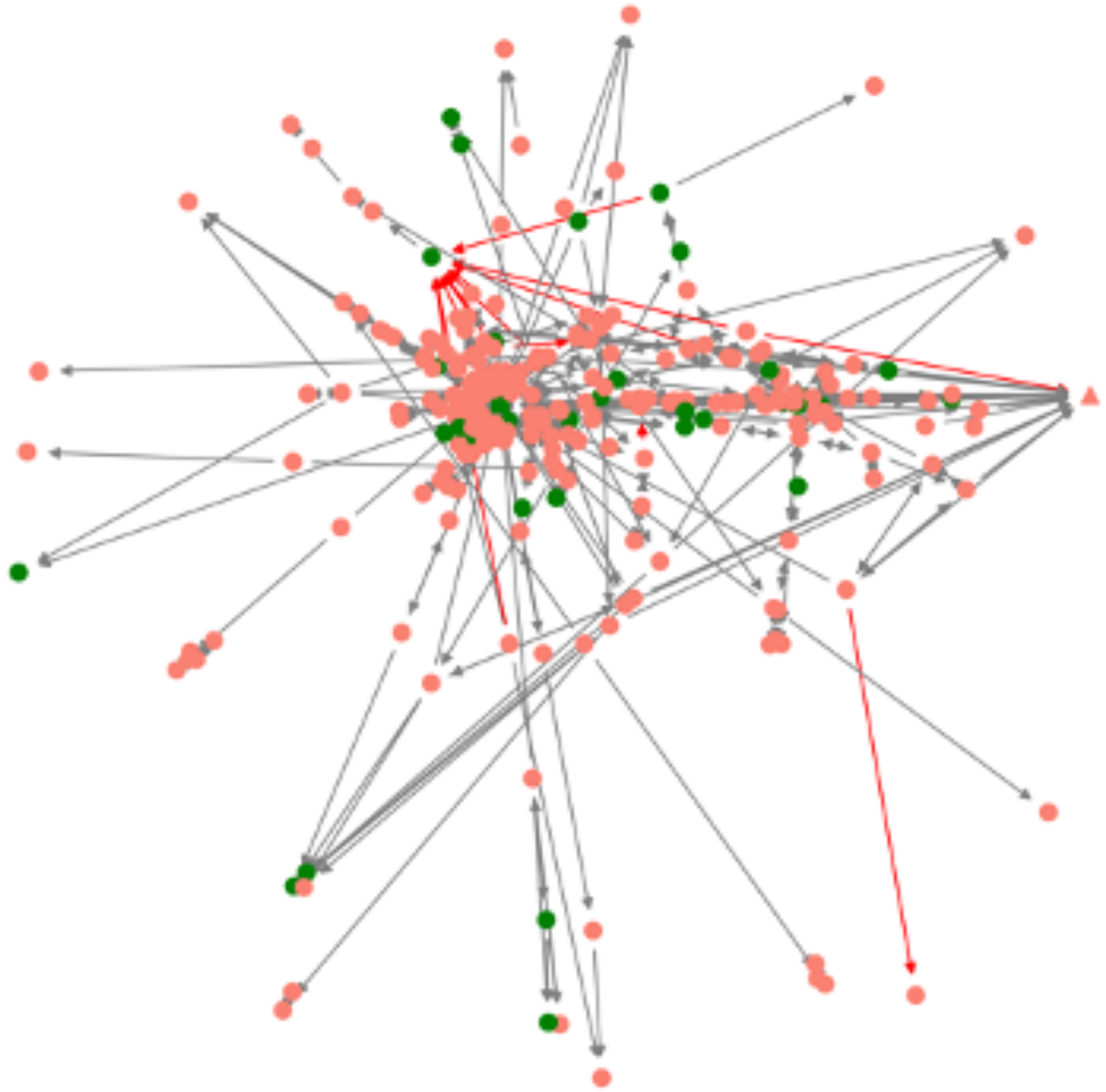
- **Explaining** vs. describing a social network
- Identifying **causal mechanisms** of a social phenomenon (i.e., a causal chain of events involving social actors' decisions under macro-level constraints; Hedström, 2005; Elster, 2015)
  - Formation of the network **structure**
  - **Composition** (diffusion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

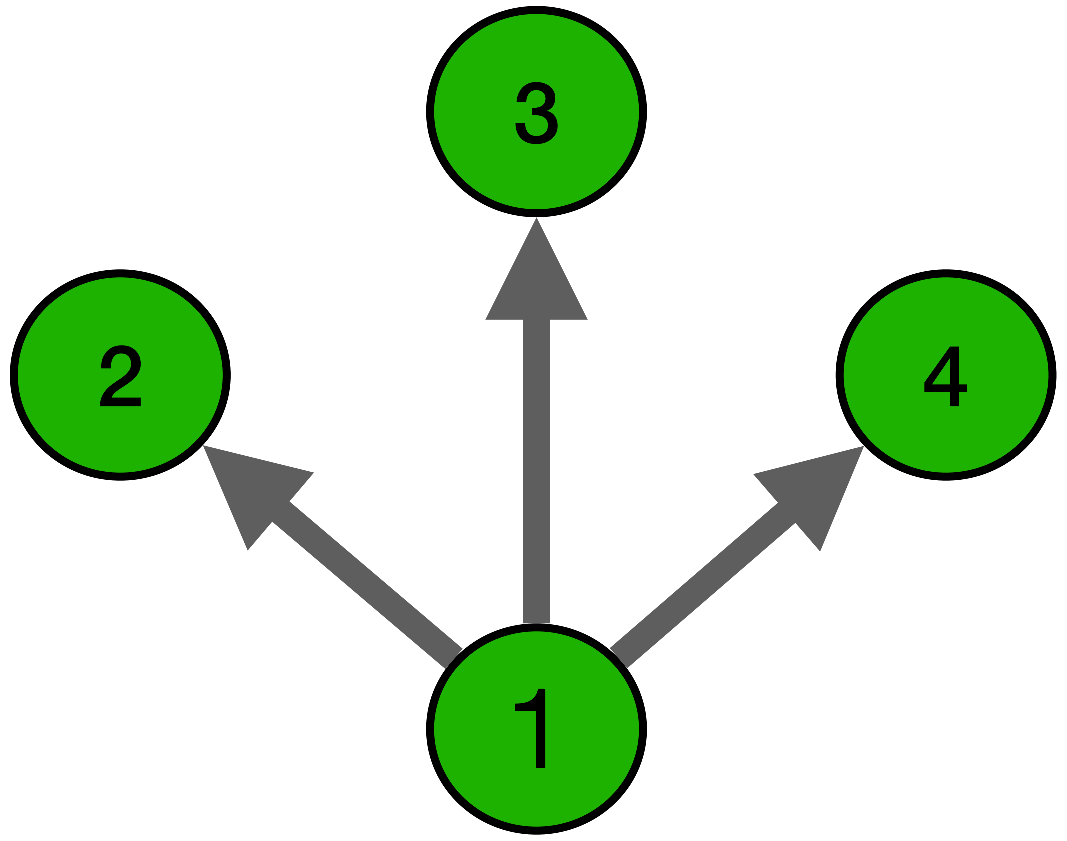
Example: what mechanisms explain the observed network structure (Bearman et al., 2004)?

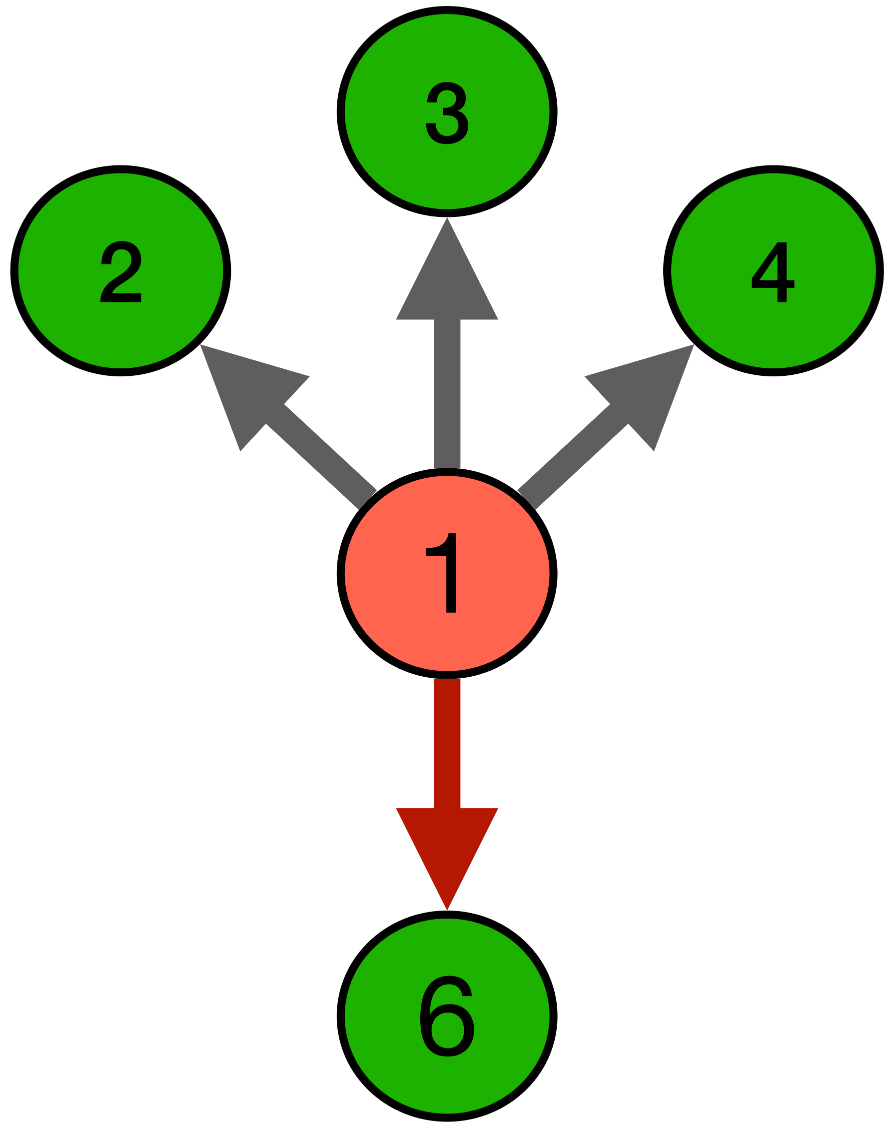


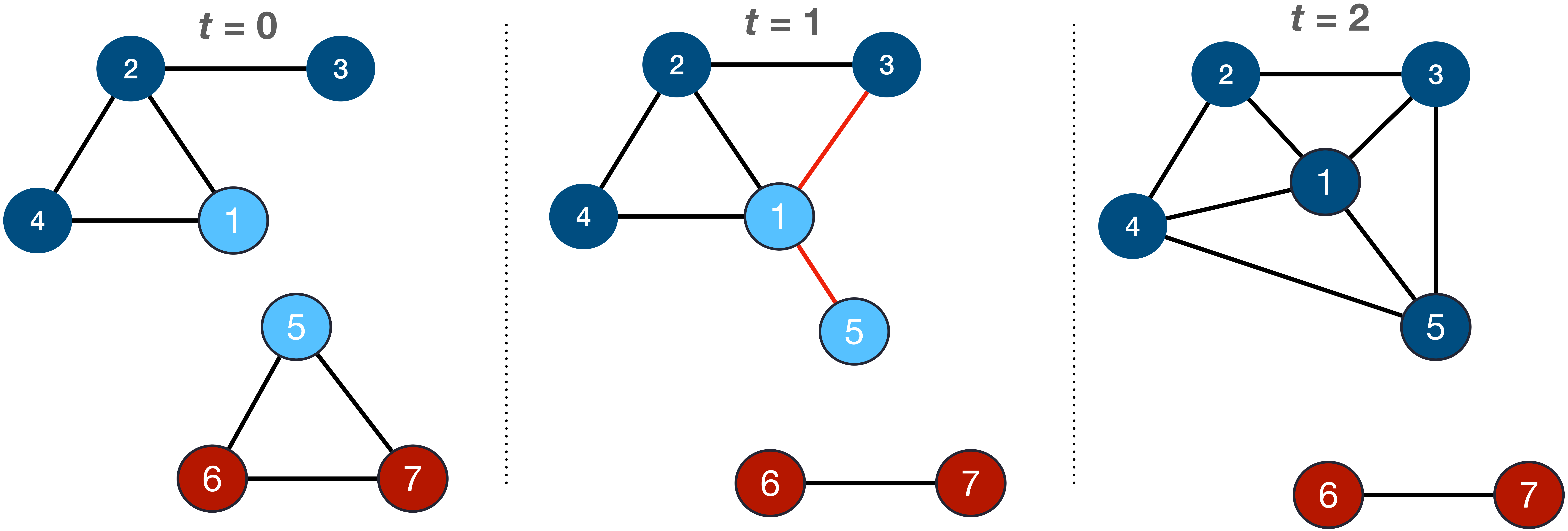


Example: what mechanisms explain the observed network composition (Bianchi et al., work in progress)?



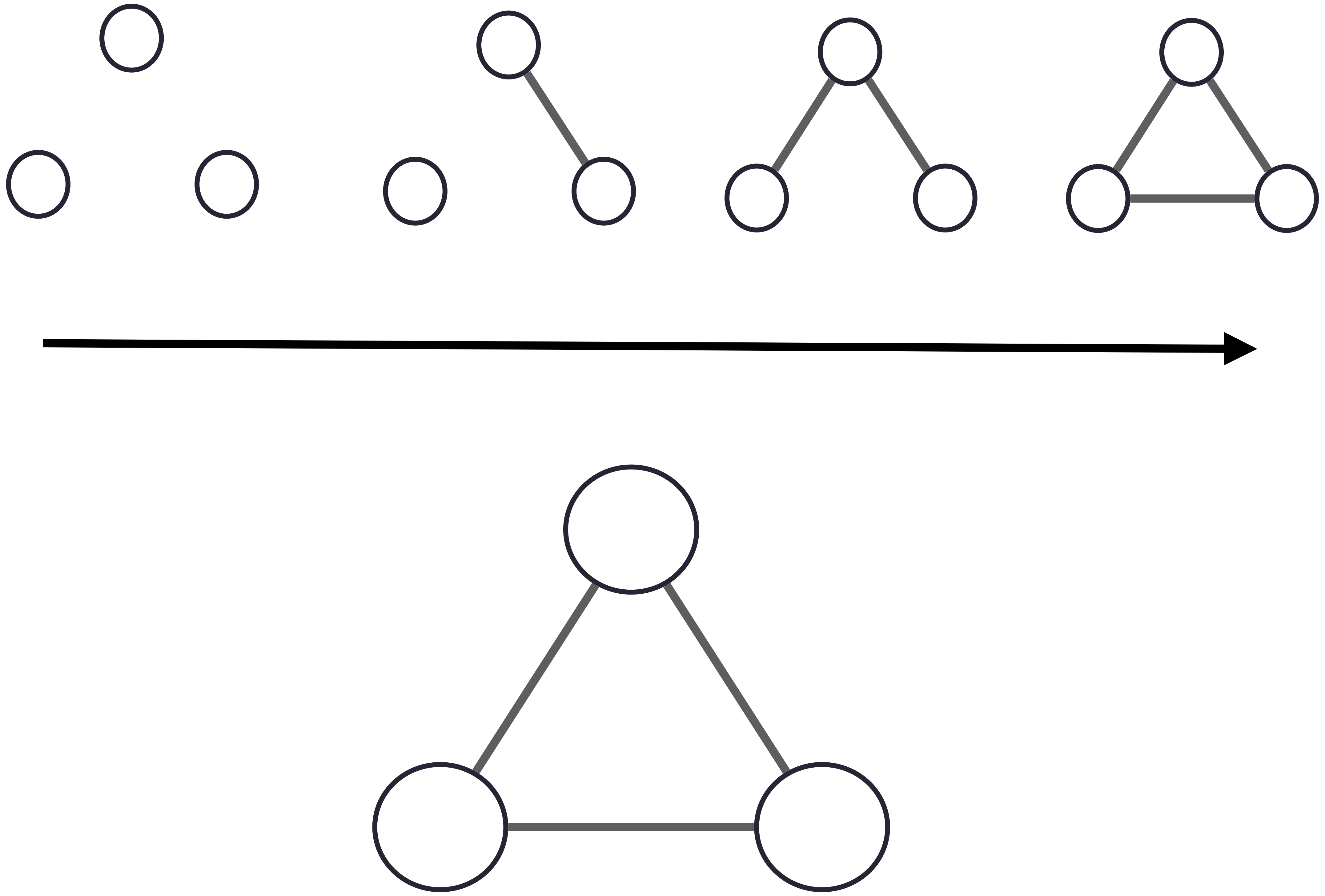


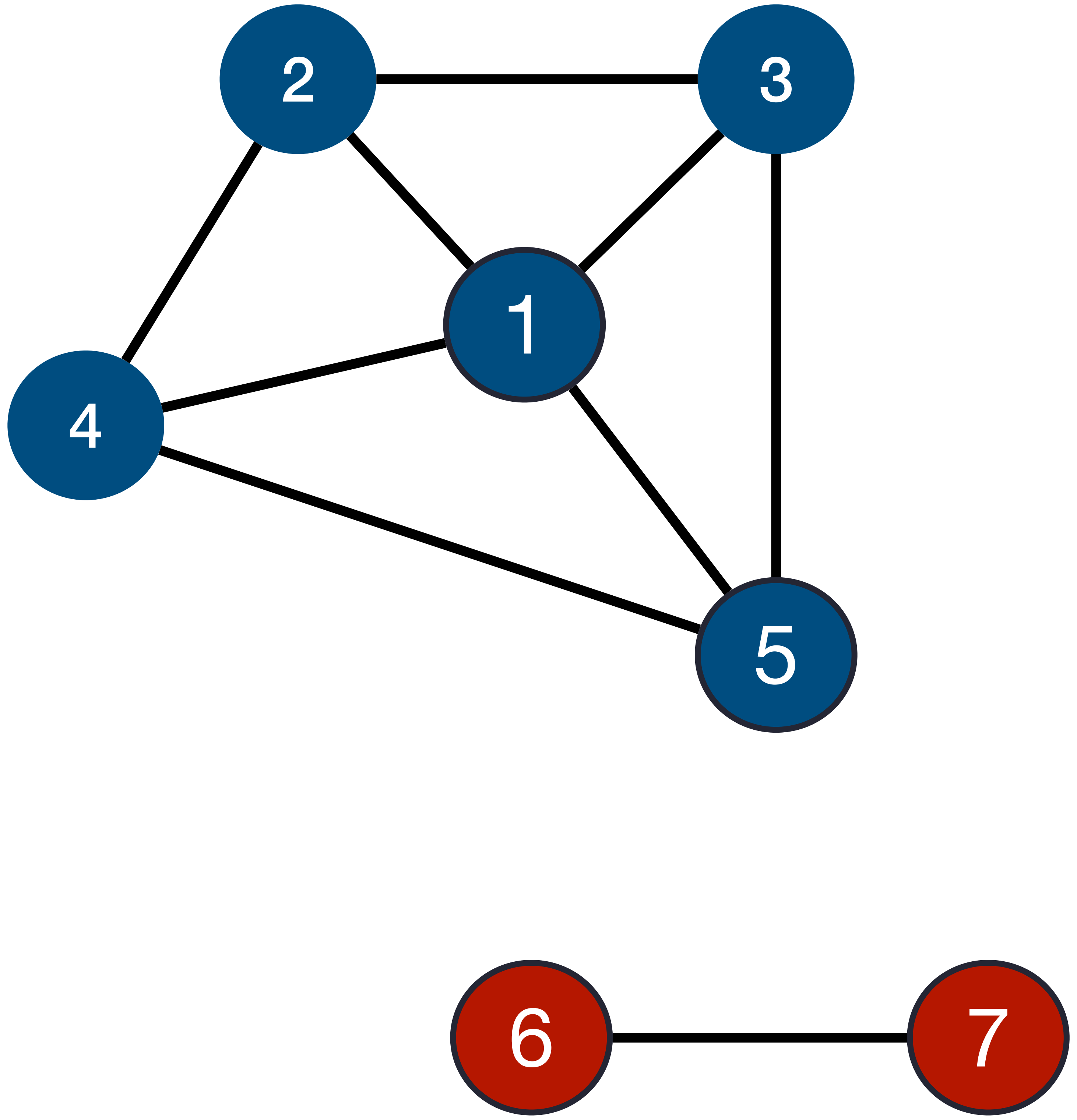




Example: a friendship network and musical tastes in a workplace

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





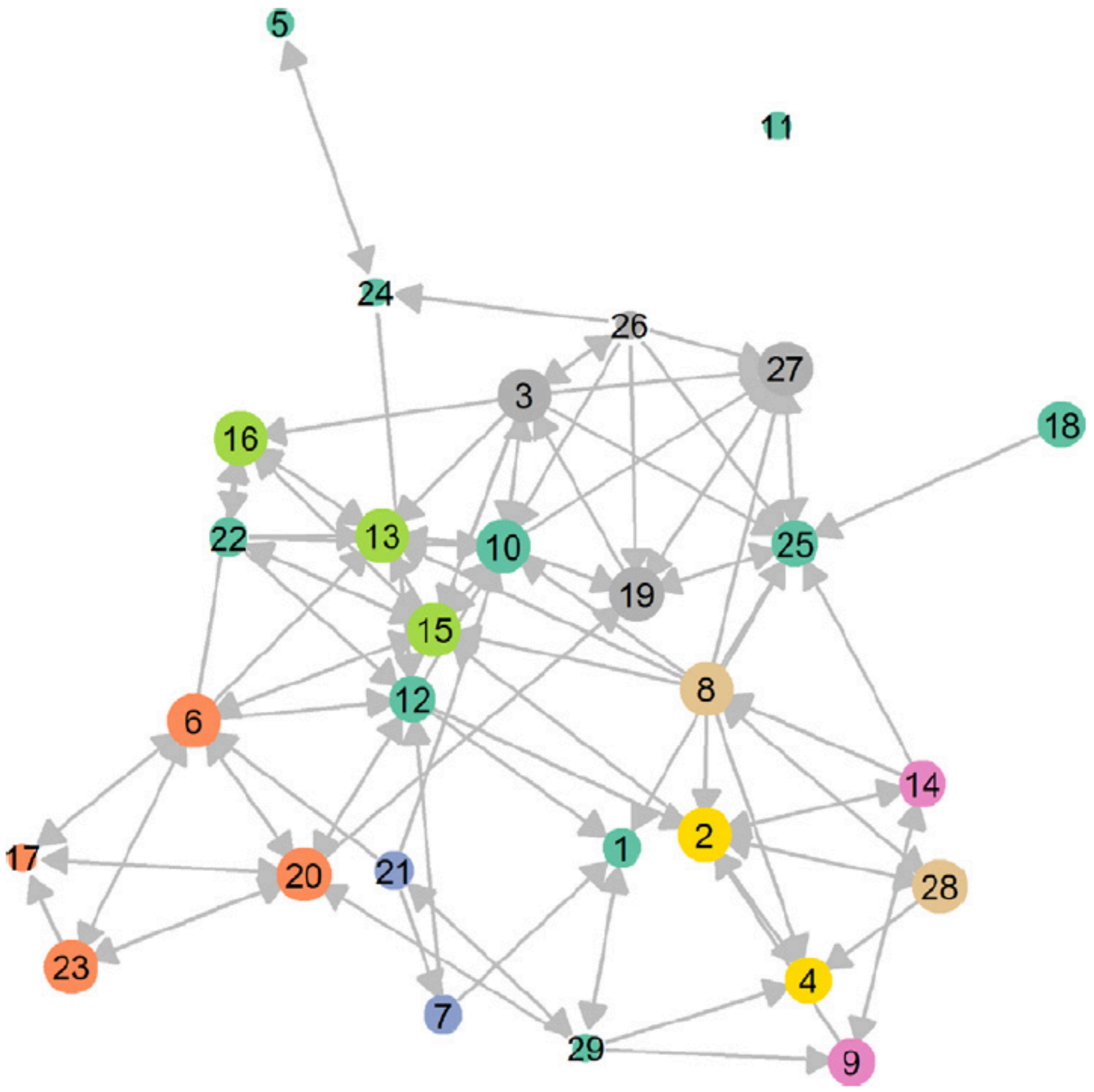
Statistical models of social networks

(Statistical inference)

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

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

- 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)

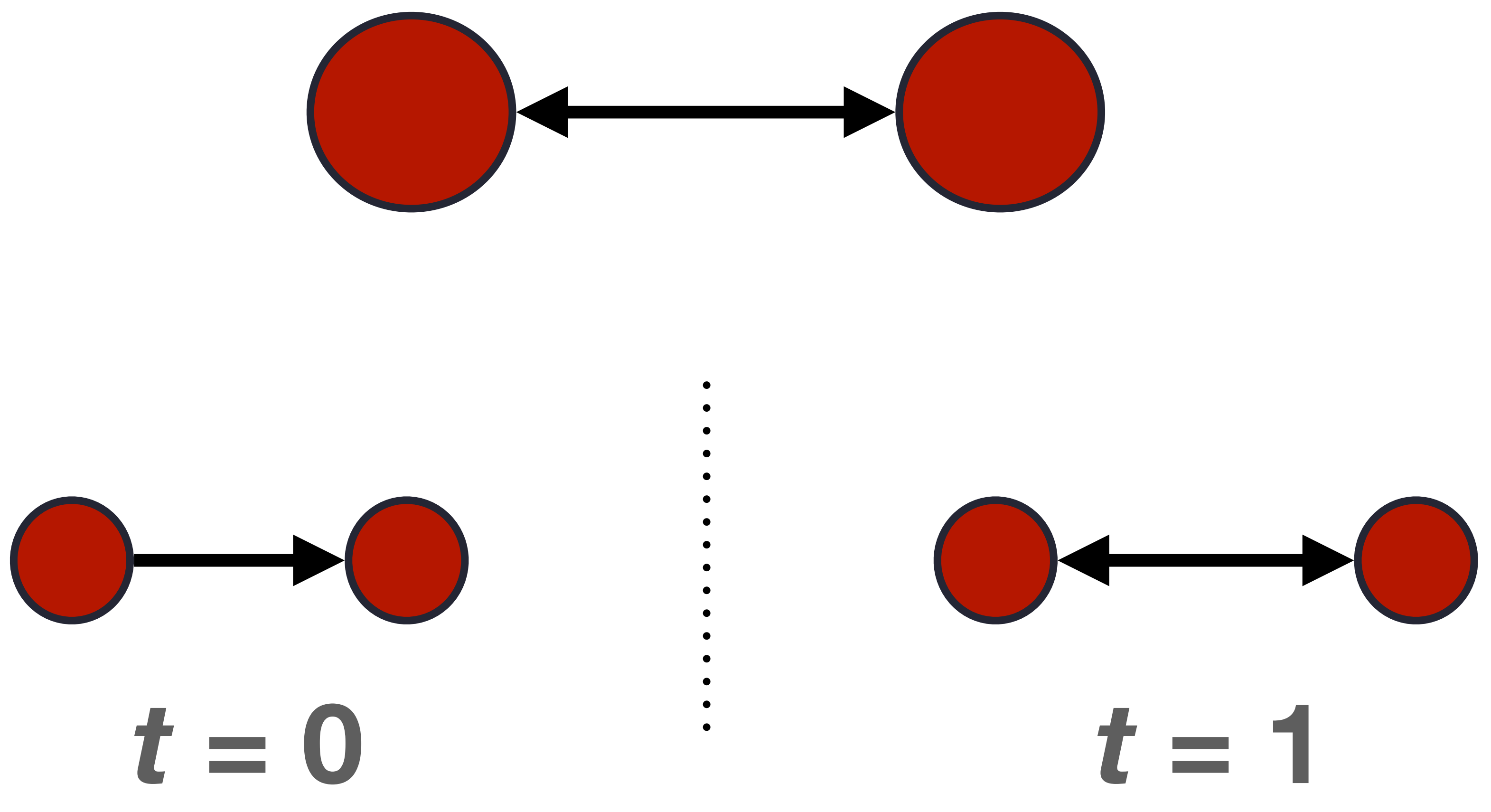


Practical example:

Reciprocity in a coworking space

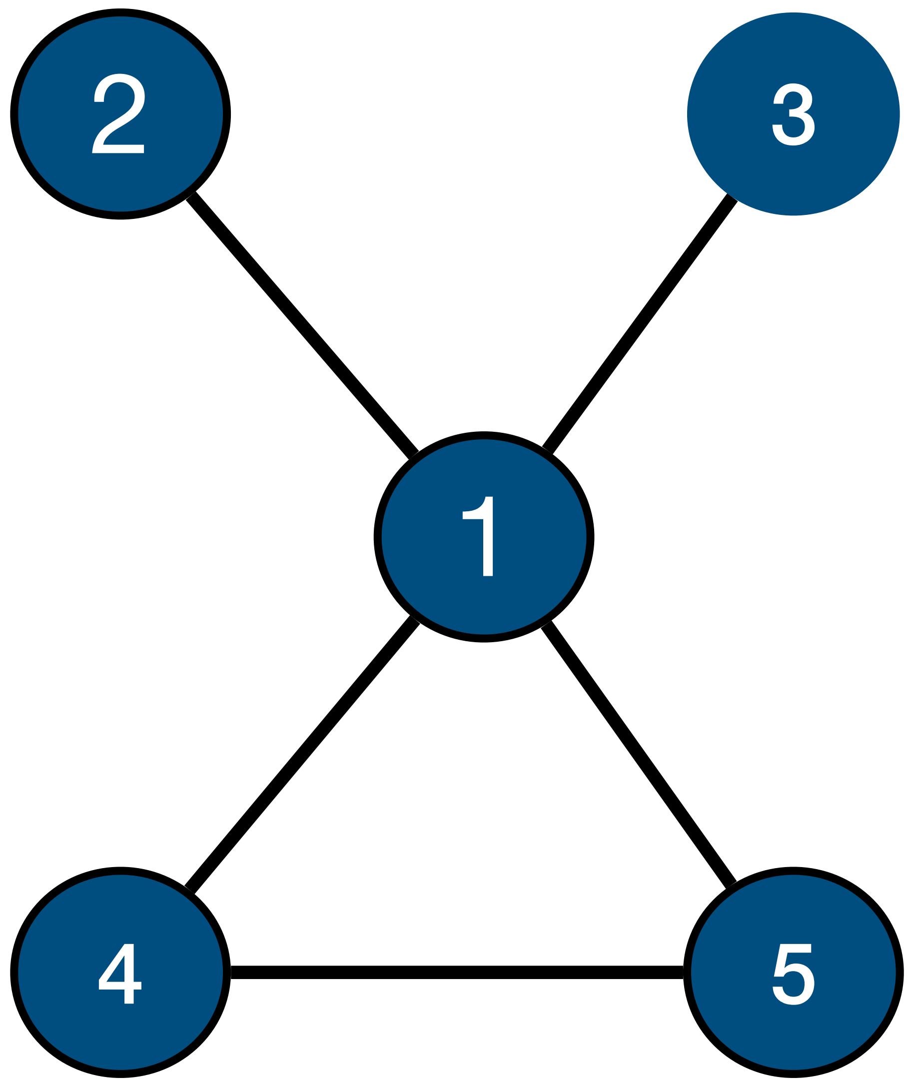
Name generator:

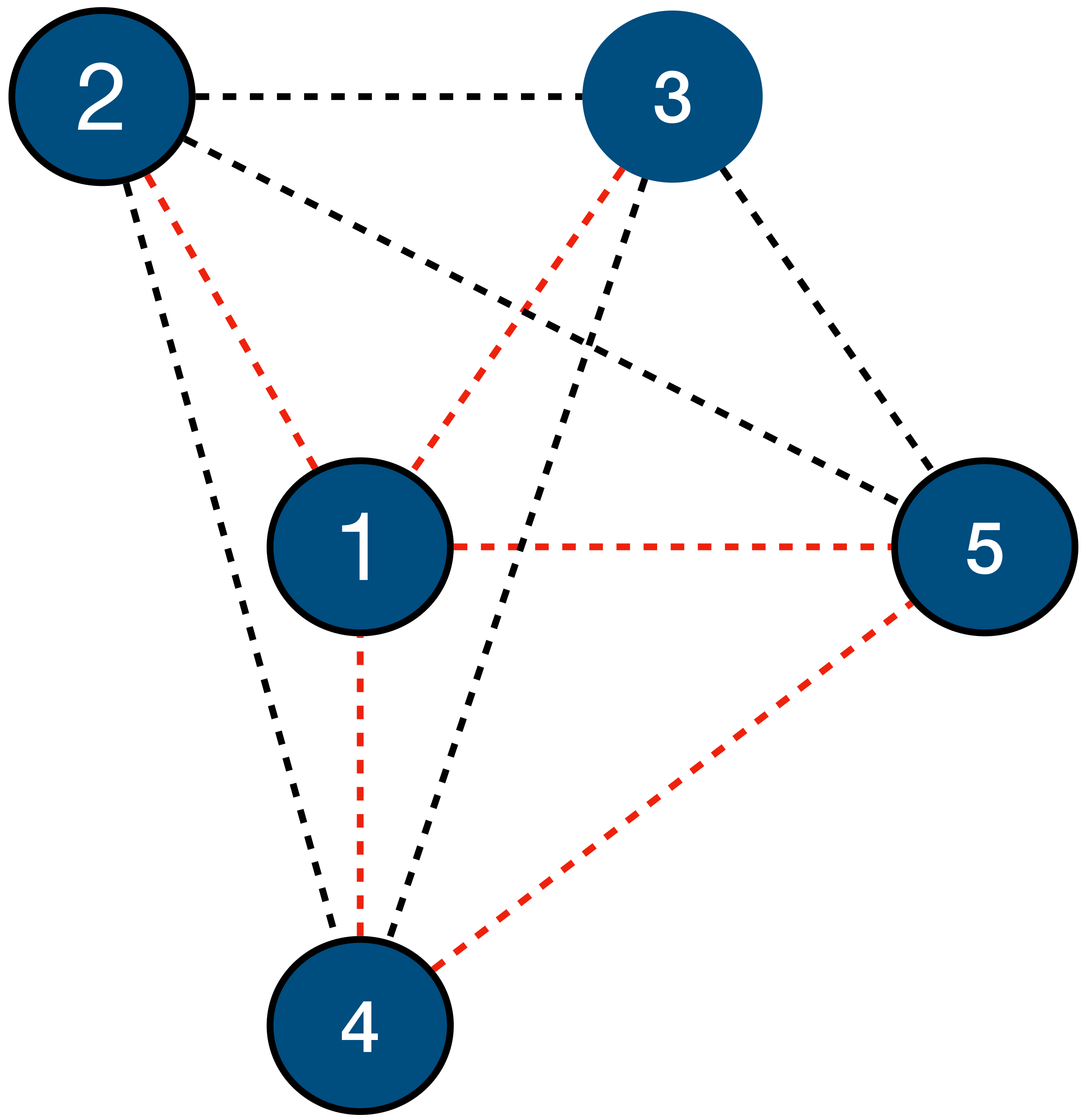
Who do you usually turn to if you need emotional or material support?



Statnet

- Suite of packages for statistical analysis of network data
- Used primarily in the sociological/anthropological/psychological community of SNA
- Comprende:
  - `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.)





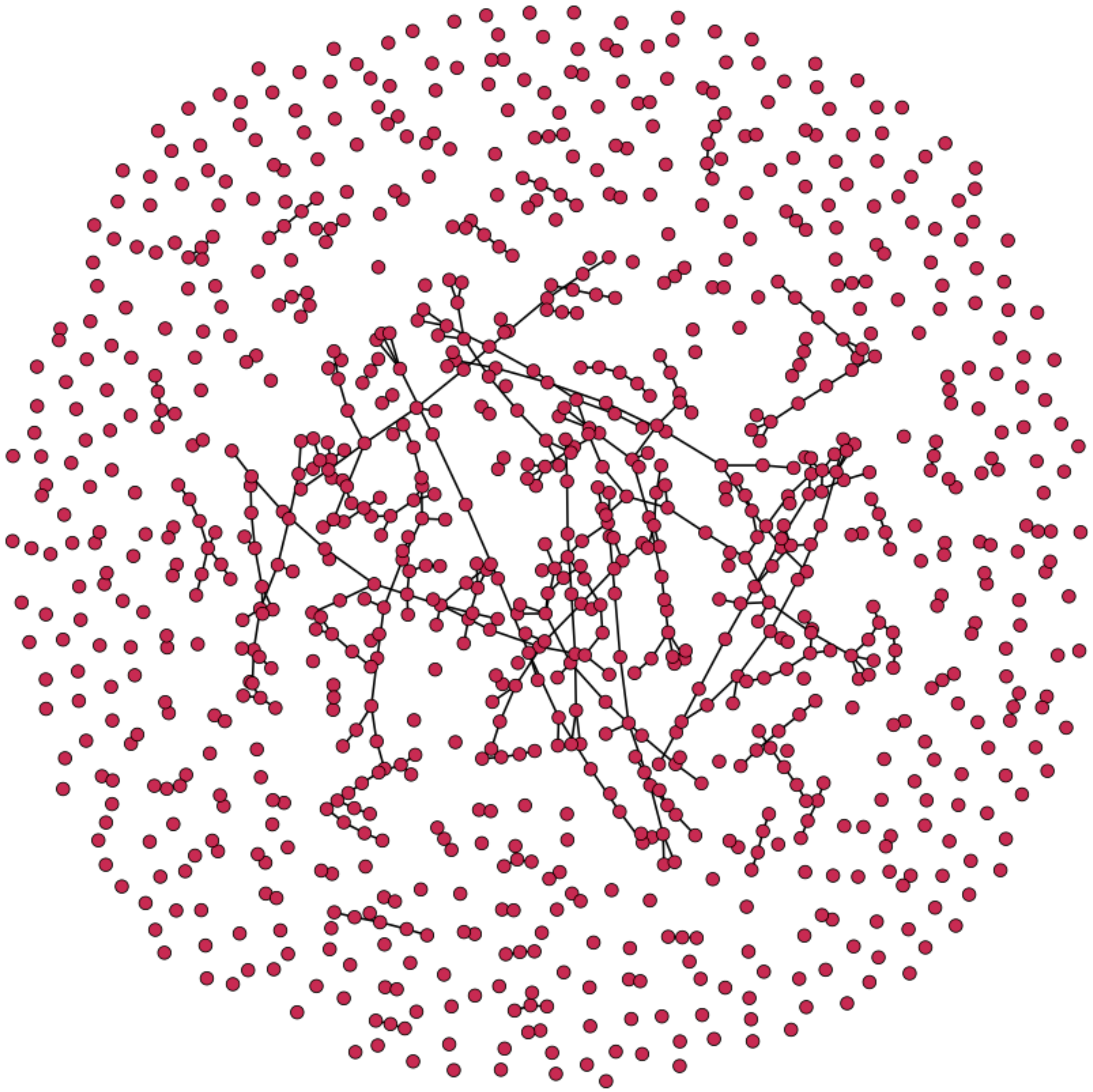
Random

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

Random graph models

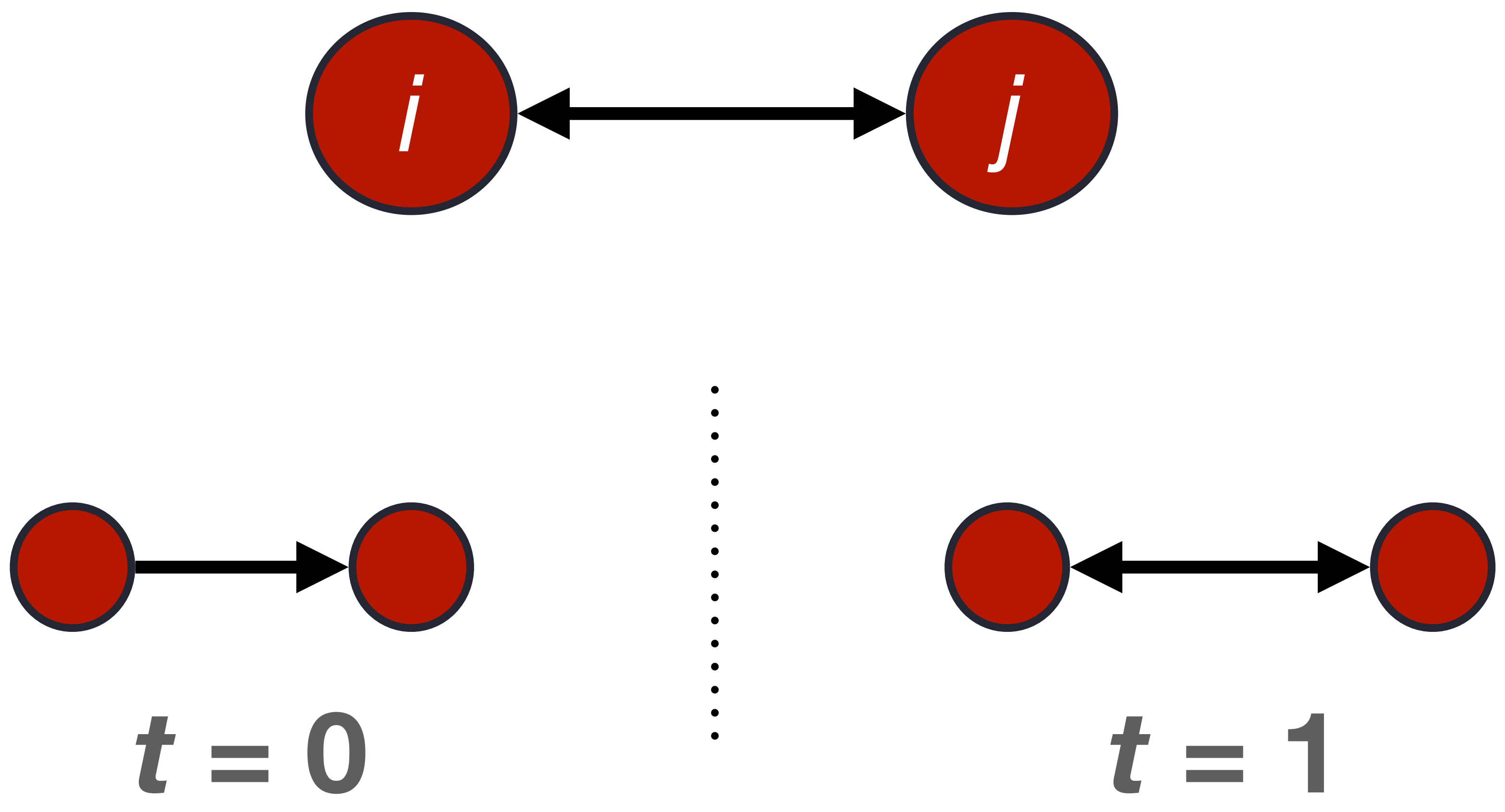
- (Stochastic) models of graphs: defined as a family of random tie-variables
- $N = \{1, \ldots, n\}$ is fixed and predetermined
- Let J be the set of all possible relational ties for N (no self-loops) (cardinality of J is $\frac{n(n-1)}{2}$)
- E (set of ties) is a random subset of J
- For any element of J *(i,j)*, X_{ij} is a **tie-variable** which can be either 0 or 1
- All tie-variables make up a stochastic adjacency matrix $\mathbf{X} = [X_{ij}]$
- A target empirical network is a realization $x = [x_{ij}]$ of $\mathbf{X}$
- *Erdős-Rényi* model (Gilbert): *G(n,p)* (a graph G with n vertices and $Pr(x_{ij} = 1) = p$)



Erdős-Rényi-Gilbert Model

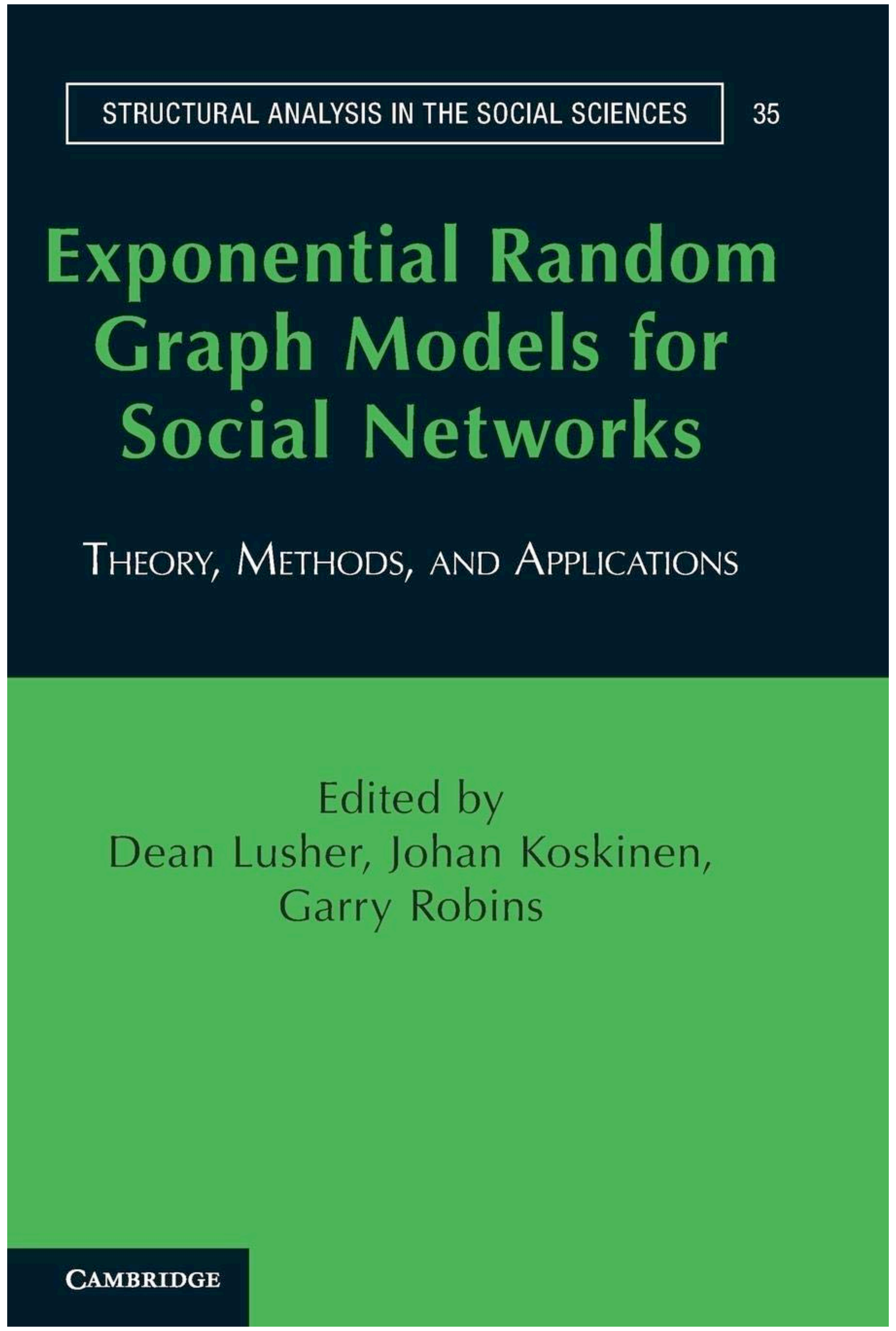
- Random graph *G(n, p)*
- *n*: number of nodes
- *p*: probability that $X\_ij = 1$
- Tie-variables are identically distributed and independent
- Bernouilli process

Stochastic dependence of observations



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

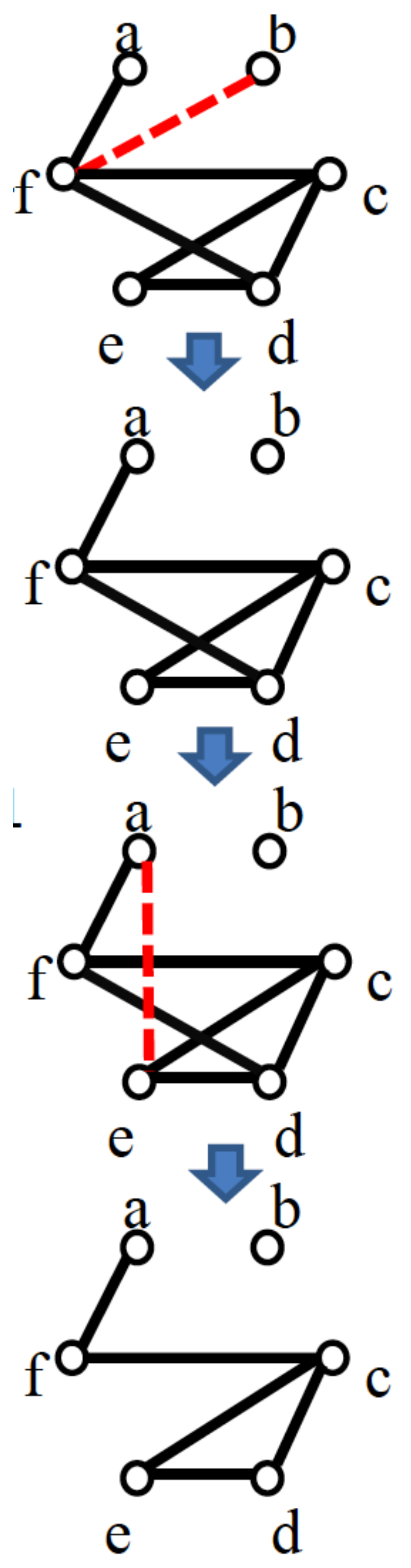
ERGM: Exponential Random Graph Models



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

- $Pr(X = x \mid \theta)$: 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

Markov Chain Monte Carlo Maximum Likelihood Estimation



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

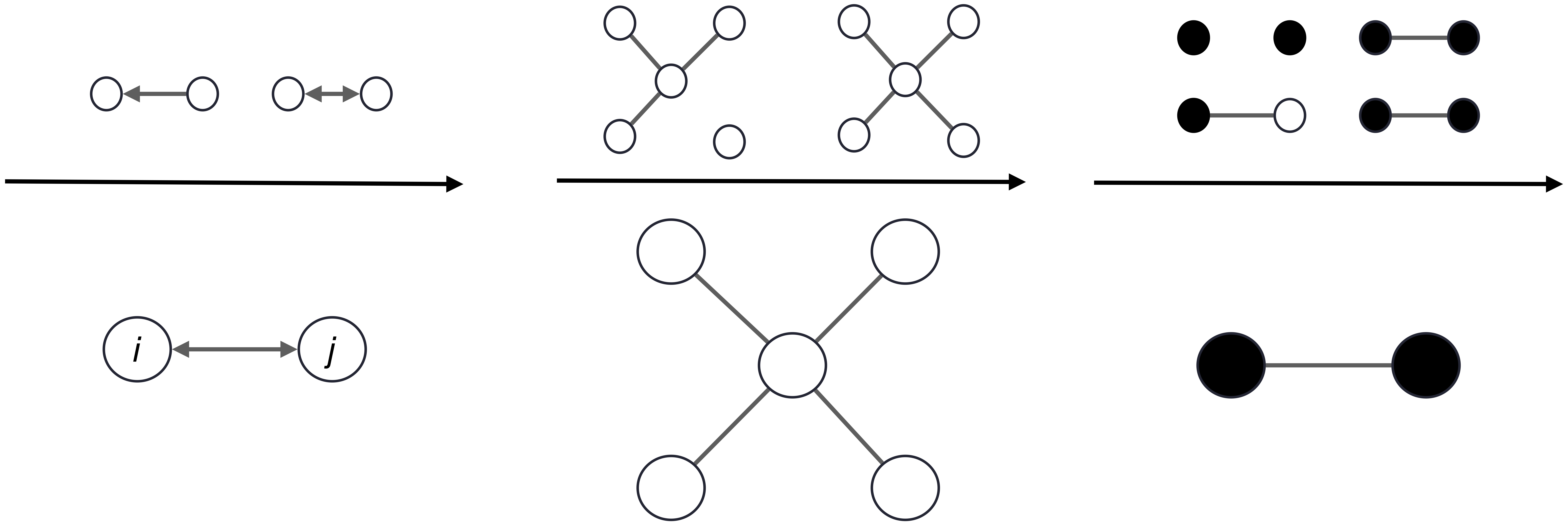
2. Start with a random network with the given number of nodes

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.

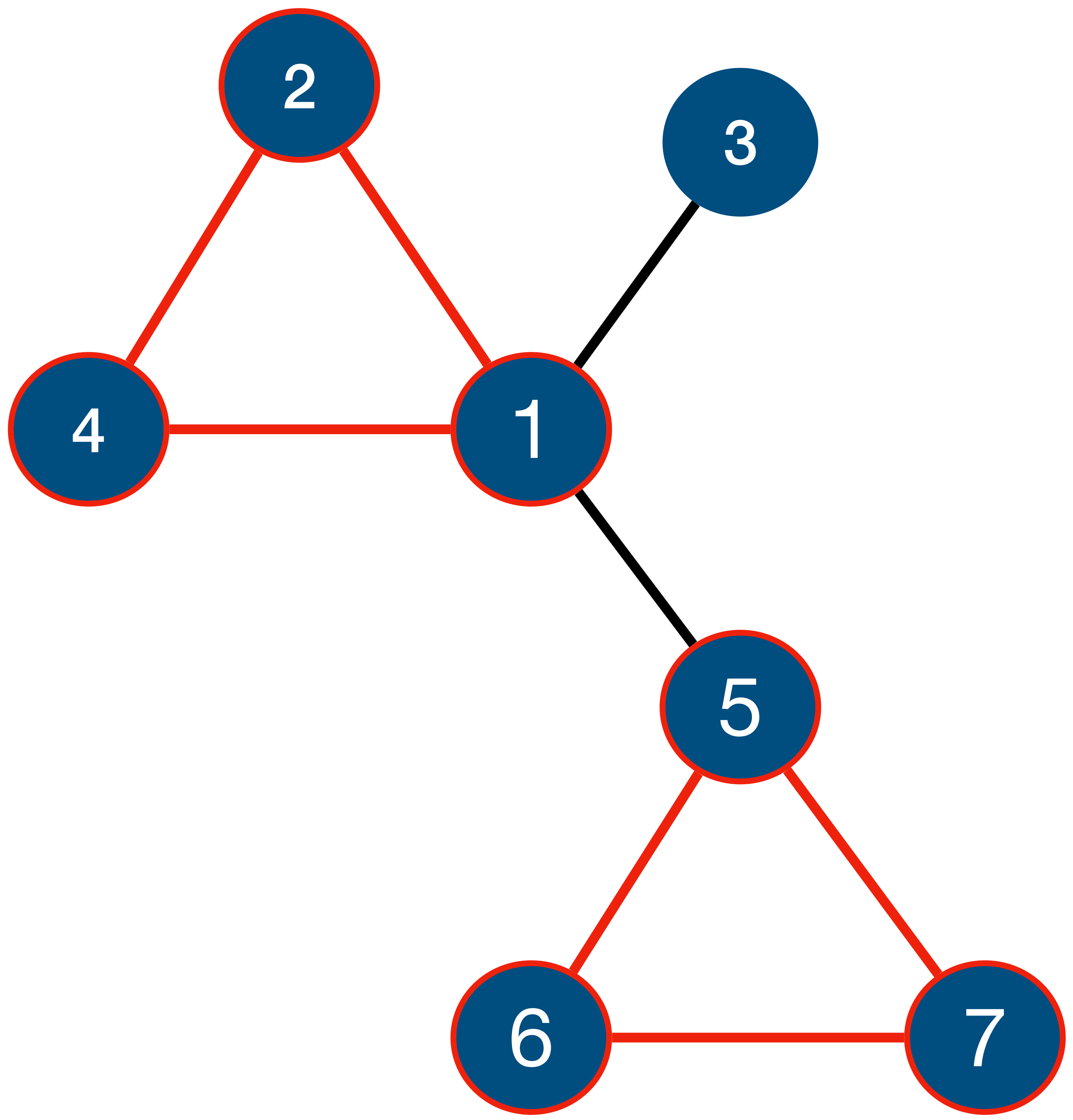
Output: The process will eventually converge (**Markov chain**) to a random graph distribution that has the count statistics of the observed network as a central tendency (**maximum likelihood**)

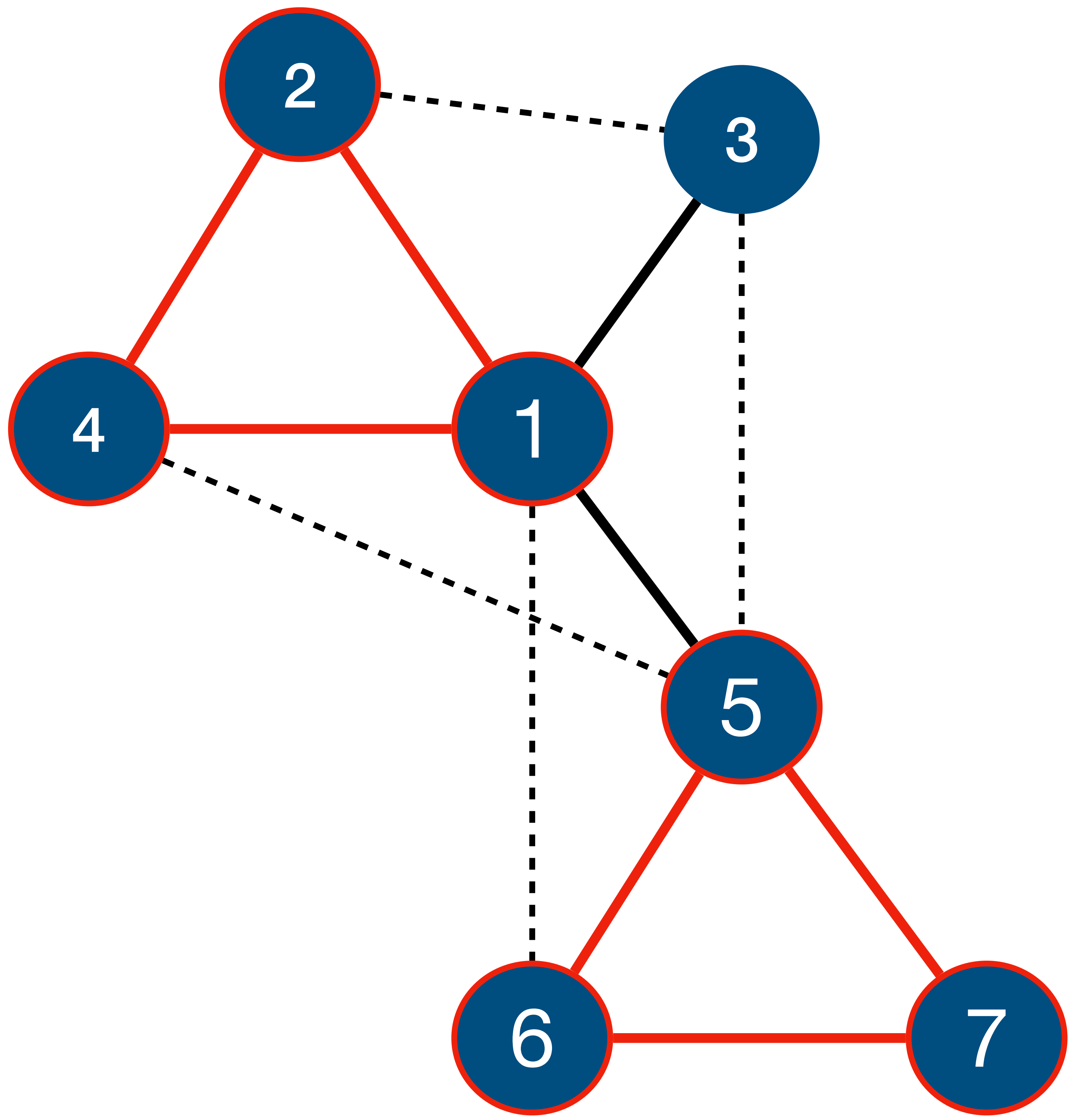


Statistical models of social networks:

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., $P(x_{ij}) \cap P(x_{ji}) = P(x_{ij} | x_{ji}) \cdot P(x_{ji})$

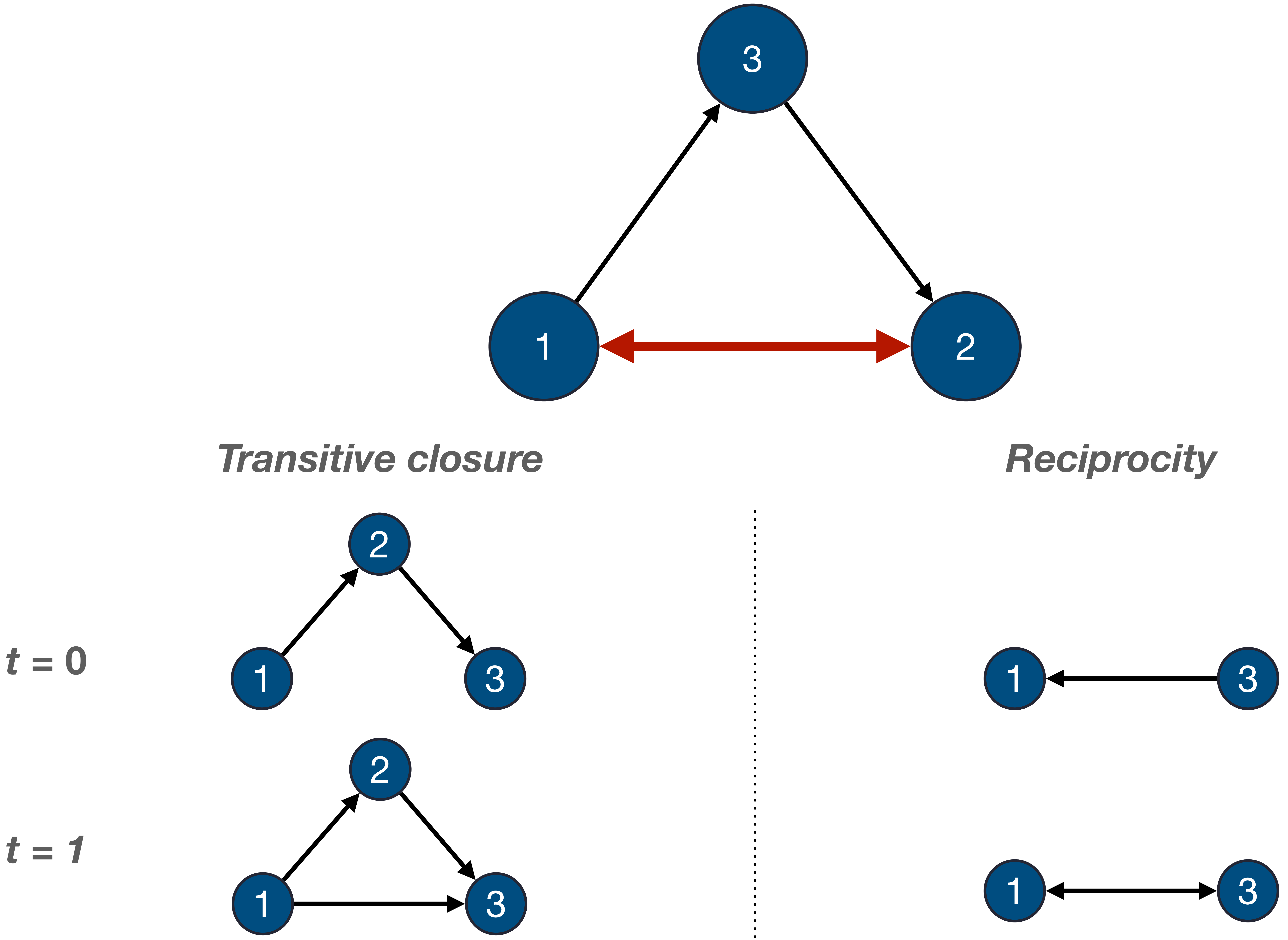




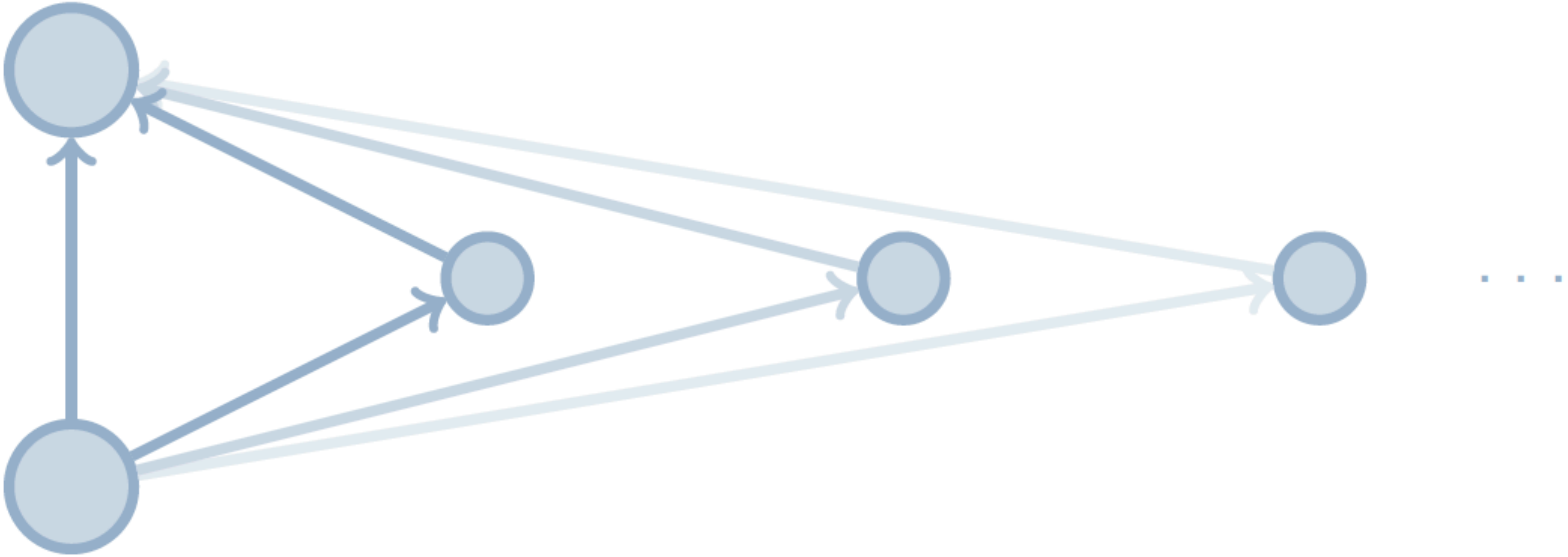
Statistical models of social networks:

hypothesis testing

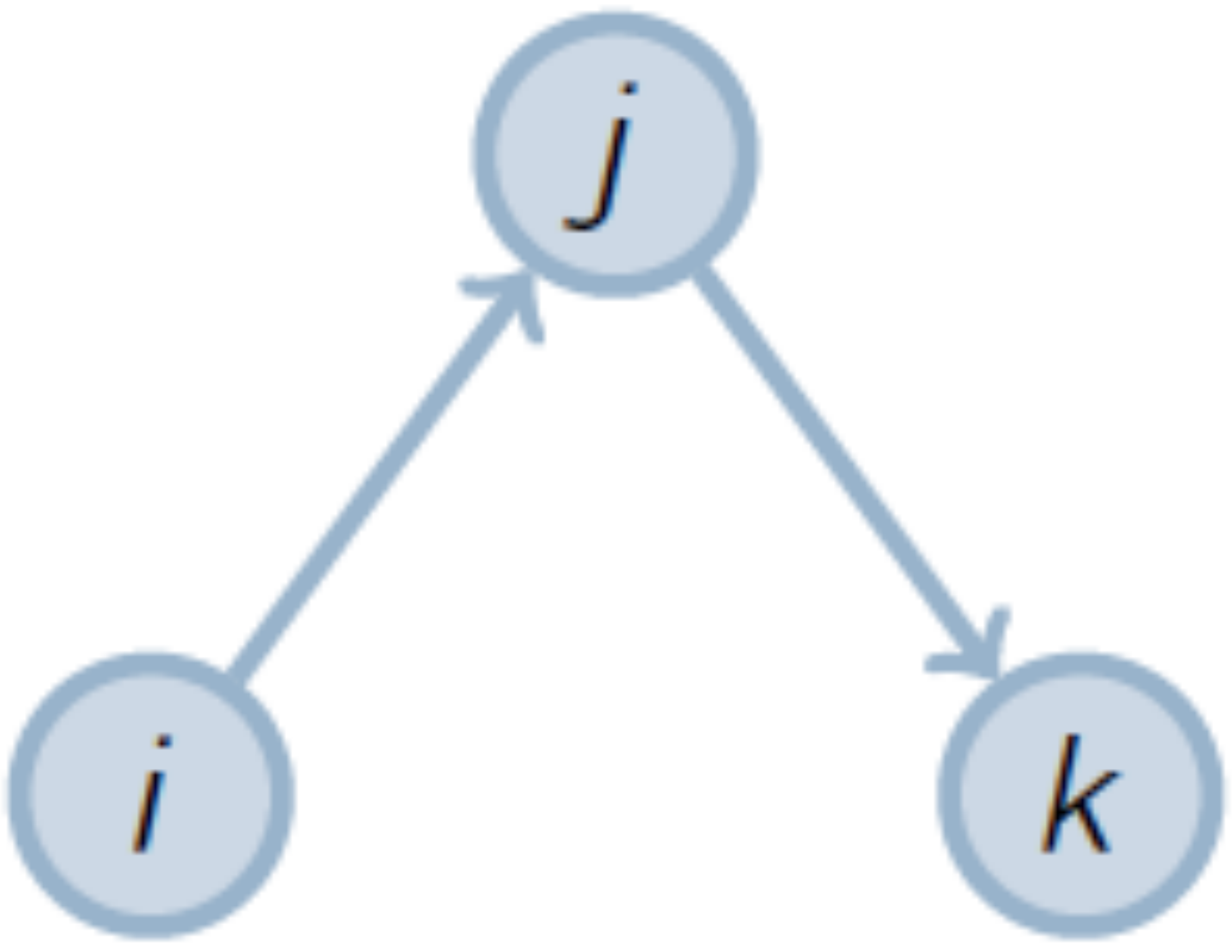
- Generating (simulating) a random graph distribution centred on the observed statistics
- Identifying a parameter vector
- Computing uncertainty measures (hypothesis testing)


Transitive closure
Reciprocity
t = 0
t = 1

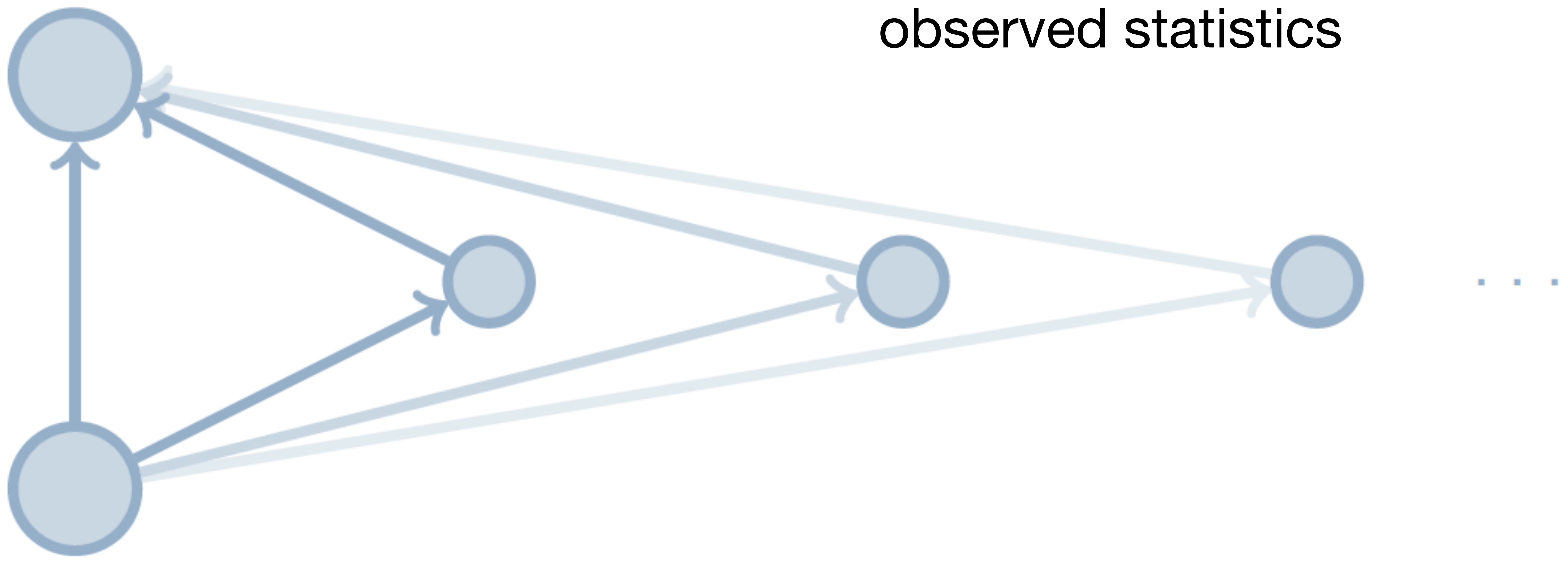
GWESP: Geometrically Weighted Edgewise Shared Partners



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

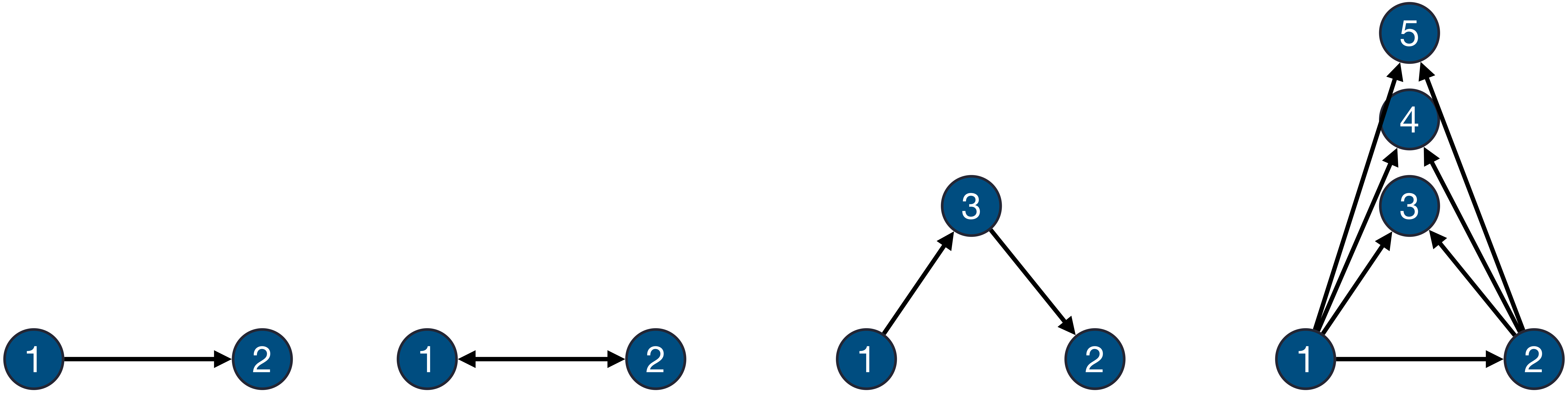


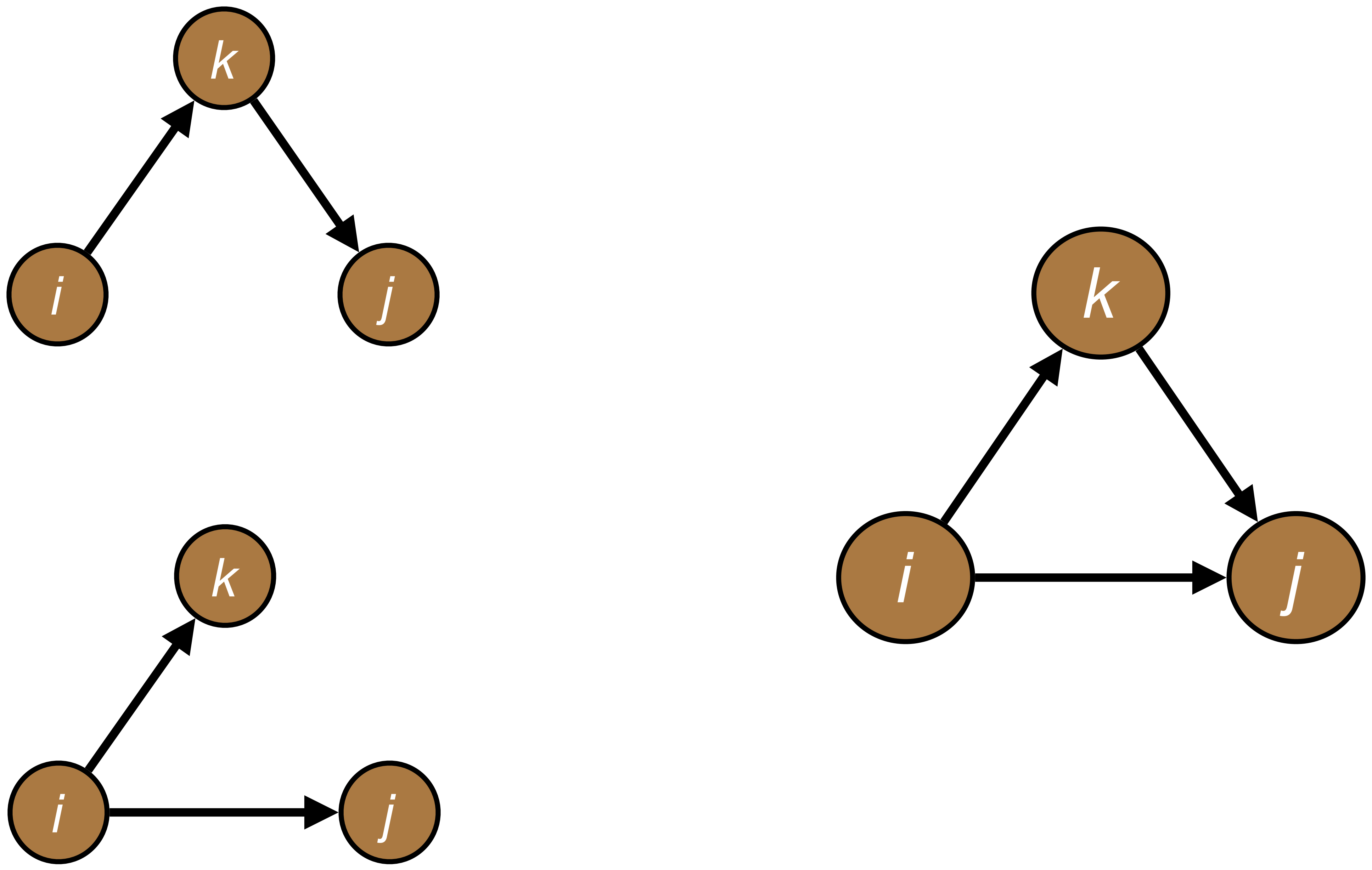
We need to specify one more statistic: the number of open triads, otherwise the model might not converge because the estimation algorithm does not have enough information to generate a random graph distribution centred on the observed statistics



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)

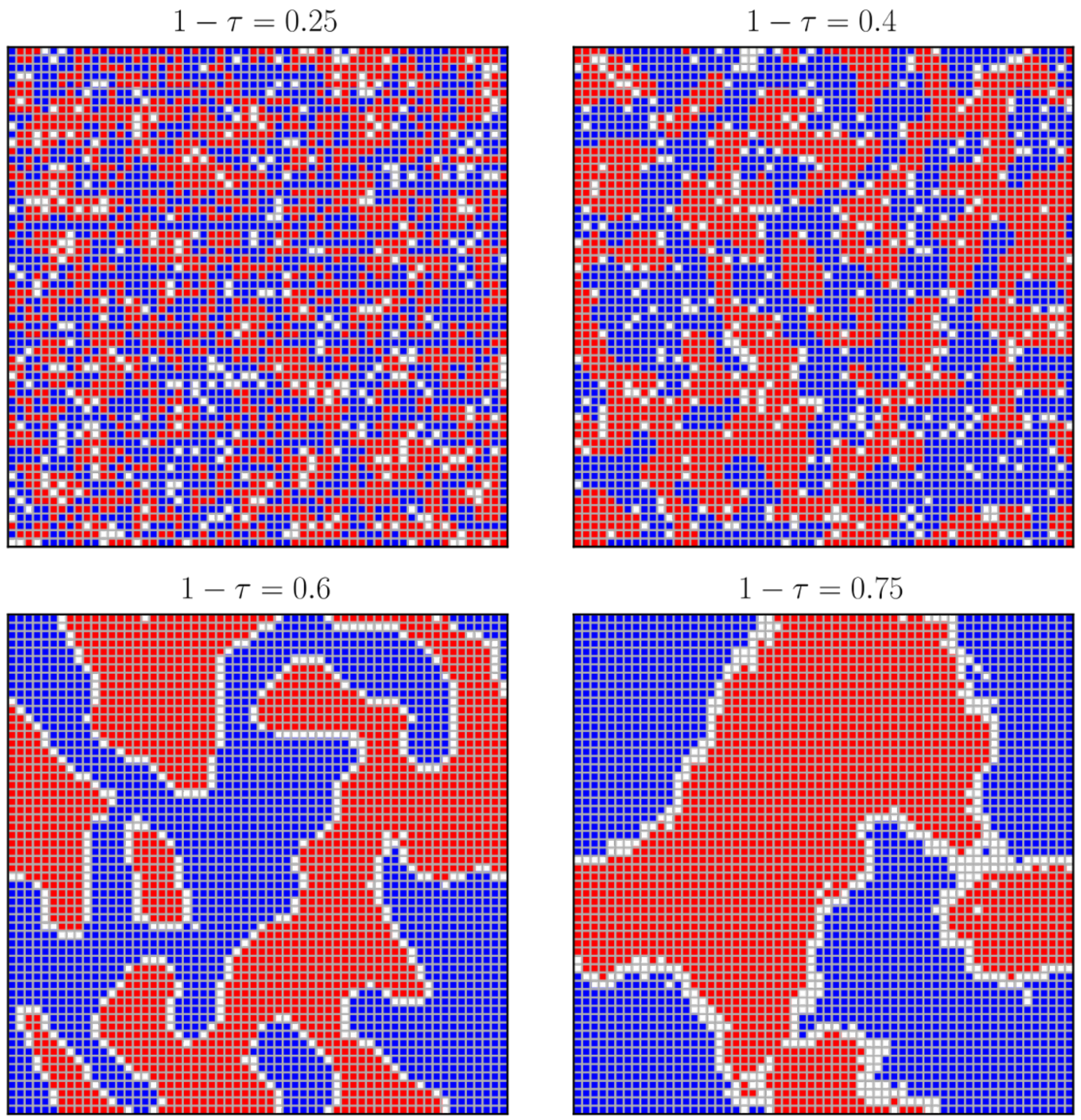
- the occurrence of a tie is assessed independently on **agents' multinomial choice**, typical of many decision-making contexts

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

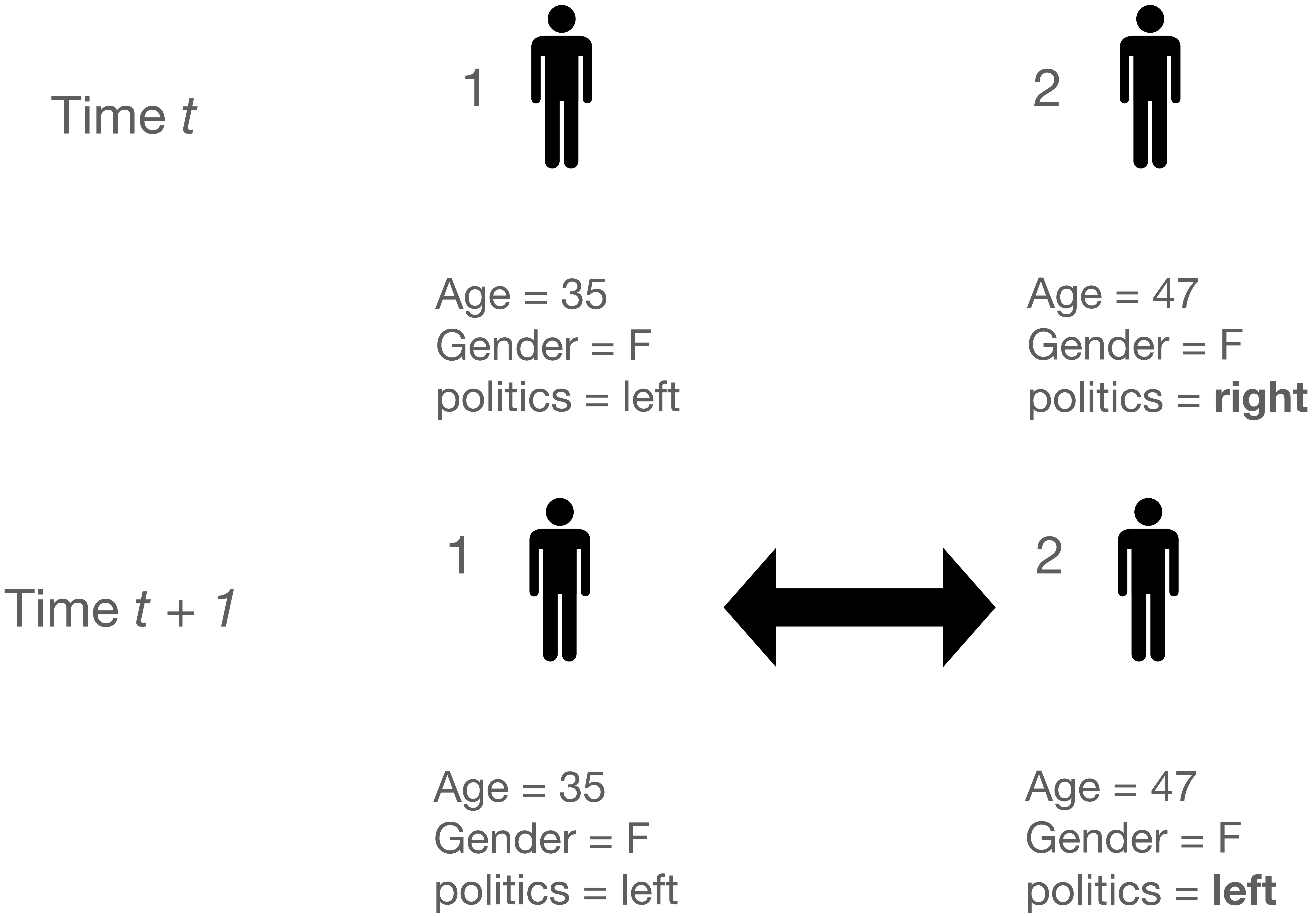
ABM: Agent-Based Models

- **Computational, dynamic models** that formalize a population of **interdependent social actors** (i.e., *agents*) with specific **properties**, interacting according to a set of **behavioural rules** within certain **environmental constraints** (Gilbert & Troitzsch, 2005; Squazzoni, 2012; Hedström & Manzo, 2015)

- Widely applied in the social sciences to explain empirical phenomena (Bianchi & Squazzoni, 2015)



ABMs are models of social interaction



	Age	**Gender**	**Politics**
1	35	F	left
2	47	F	right
...			
n			

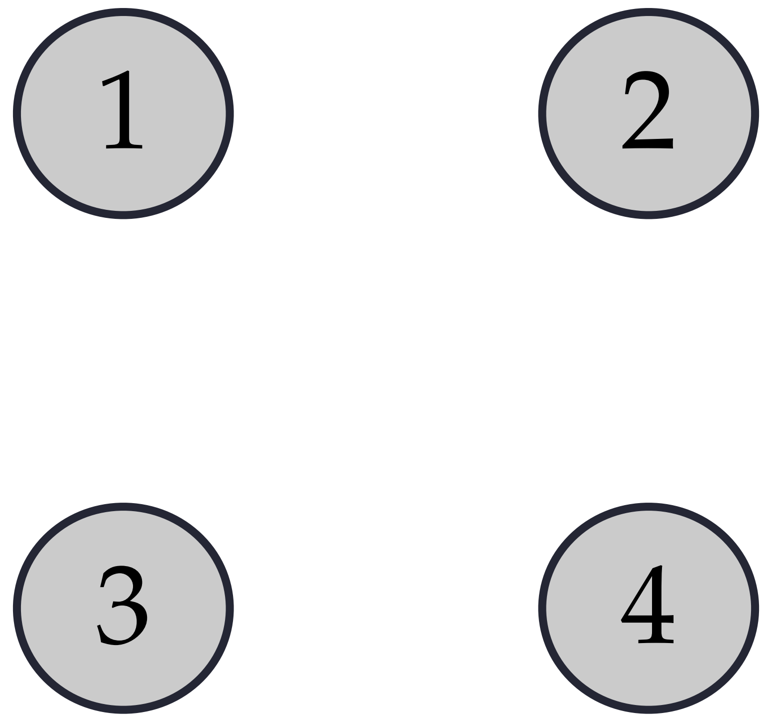
“From factors to actors”
(Macy & Willer, 2002)

ABMs can model social networks

Time t

Age = 35
Gender = F
Neighbours = ( )

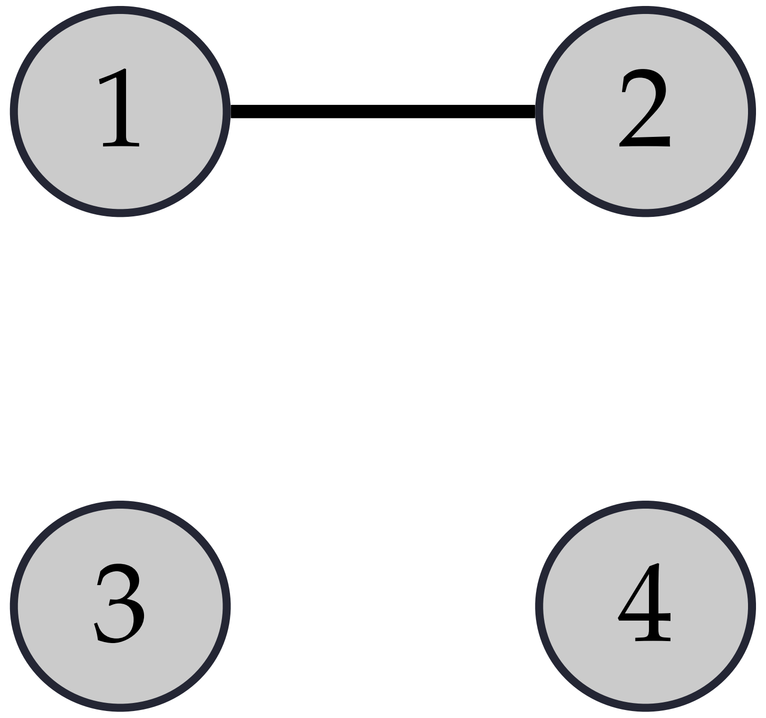
Age = 47
Gender = F
Neighbours = ( )



Time $t + 1$

Age = 35
Gender = F
Neighbours = (2)

Age = 47
Gender = M
Neighbours = (1)



Real mechanism
- Actors
- Actors' properties
- Actors' (inter)actions
- Actors' relationships

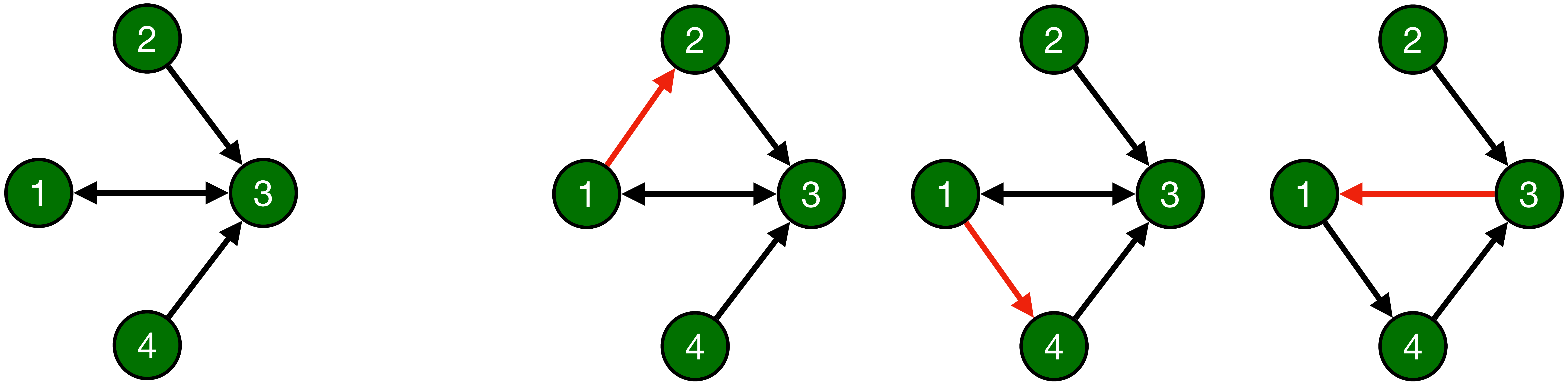
→

Agent-based model
- Agents
- Agents' attributes
- Agents' rules of behaviour
- Agents' structural constraints

ABM:

flexibility and granularity

- "**Structural homology**" with causal mechanisms (Manzo, 2014):
  - **Cognitive** or **cultural** constituents of actors' decisions
  - Social **interactions**
  - **Institutional**, **relational**, or **spatial** constraints
- High **flexibility** —> wide **granularity** range of agent modelling (Wooldridge & Jennings, 1995)
  - **Social** characteristics: autonomy, interdependence, embeddedness, heterogeneity
  - **Cognitive** characteristics: reactivity, proactivity, heuristic-based rationality, adaptiveness



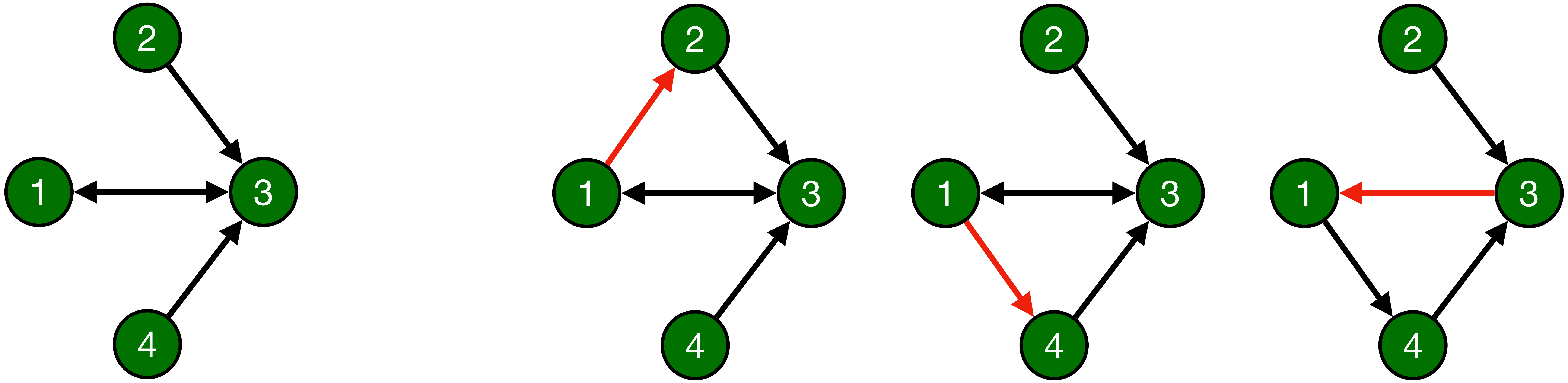
- ***Agent-based*** model: the likelihood of a tie to occurr is assessed as a function of a focal node-agent's neighborhood structure/composition

- Each agent decides whether to change the state of an outgoing dyad through a multinomial experiment (McFadden, 1973), by optimising an objective function $P(x \rightarrow x^{\pm ij}) = \dfrac{exp(f_i(\beta; x^{\pm ij}))}{\sum_{h=1}^{n} exp(\beta; f_i(x^{(ih\pm)}))}$

- The function parameters can be interpreted as the agents' **relative preferences** on the prevalence of certain local configurations

SAOM

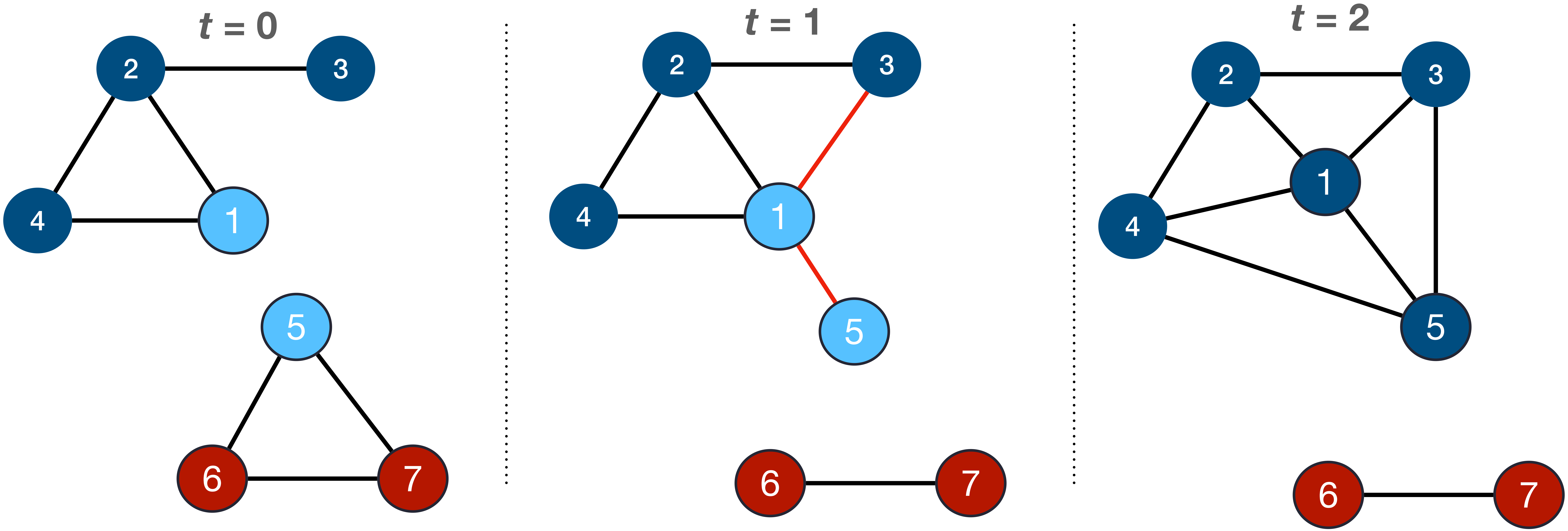
Stochastic Actor-Oriented Models



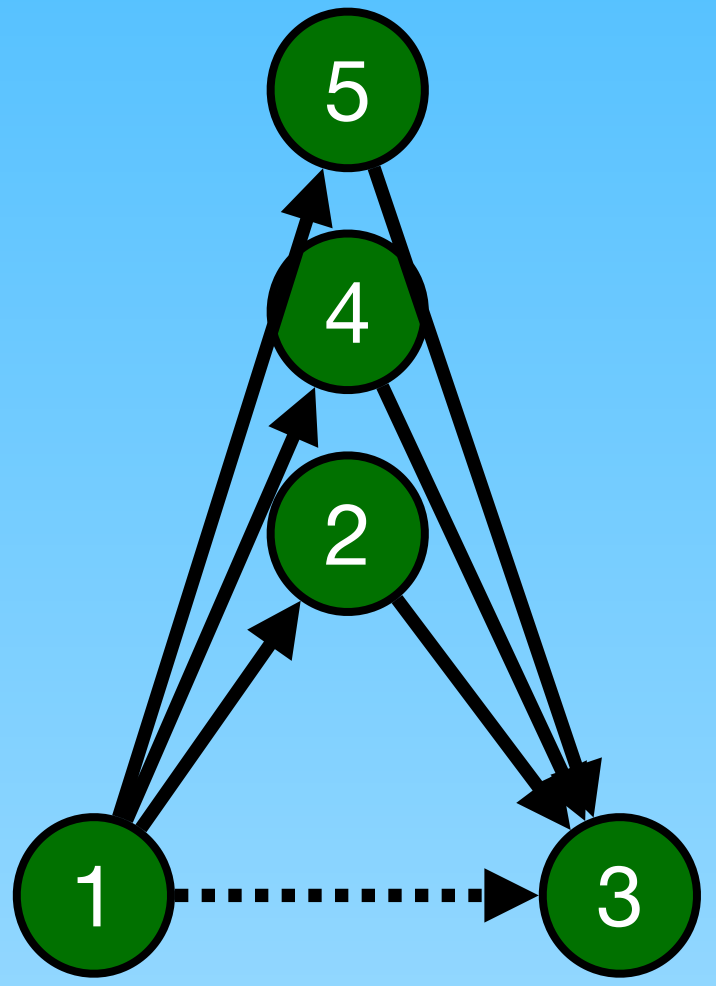
SAOM

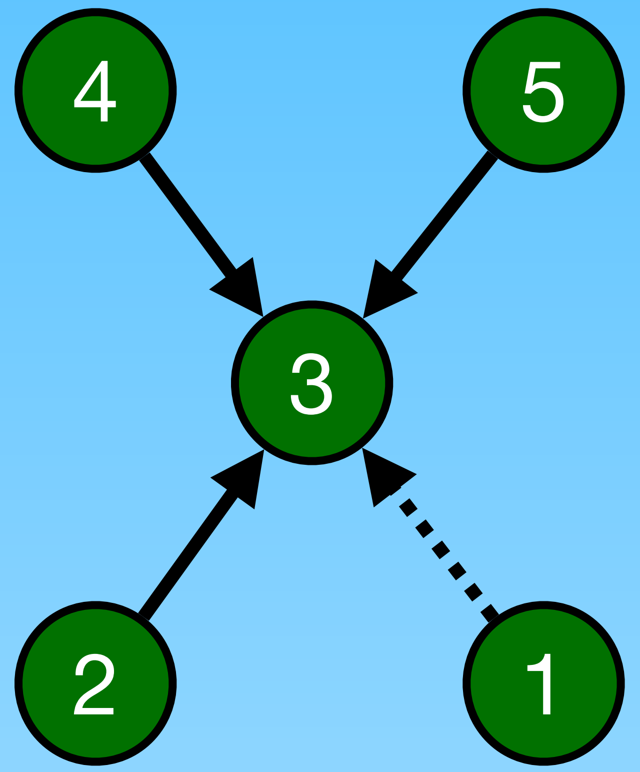
Stochastic Actor-Oriented Models

- Dynamic model: the network evolves as a continuous-time Markov process *X(t)*
- The dynamic process is unobserved except for a finite number of waves —> easily fitted to panel data
- Each agent has an opportunity to make changes in its neighborhood according to a rate function $\lambda_i(\alpha, x)$



Coevolution of ties and node attributes





To be **mathematically tractable**, (most) **SAOMs** (Snijders, 2017) assume agents':

- access to **information about the whole network** (e.g., geometrically weighted configurations): **unplausible for large networks** or **competitive contexts** where information is strategically concealed (e.g., Renzini et al., 2023) —> **idiosyncratic models**

- **changing one tie** at each simulation step: **prevents modelling coordination** and collective action (Leifeld & Cranmer, 2019) and **cascade dynamics** driven by **threshold-based preferences** (Renzini et al., 2023)

SAOM

Stochastic Actor-Oriented Models

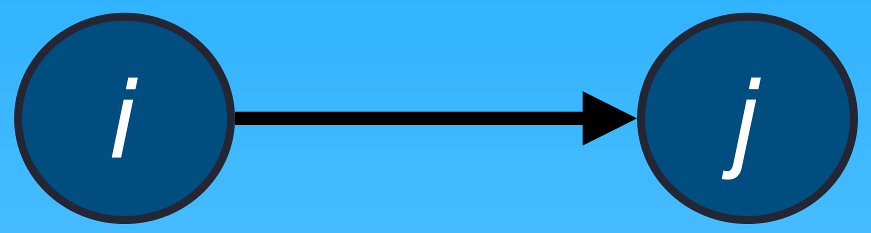
$$P(x \to x^{\pm ij}) = \frac{exp(f_i(\beta; x^{\pm ij}))}{\sum_{h=1}^{n} exp(\beta; f_i(x^{(ih\pm)}))}$$

SAOM

Stochastic Actor-Oriented Models

- tie selection as a **multinomial choice** based on **preference optimization**: **unplausible for cognitive relations** not requiring psychological investment (liking vs. disliking, status attribution)

- **myopia**: **prevents modelling** a) **backward-looking rationality** and **learning** processes; b) **forward-looking rationality** (strategic behaviour in competitive contexts)



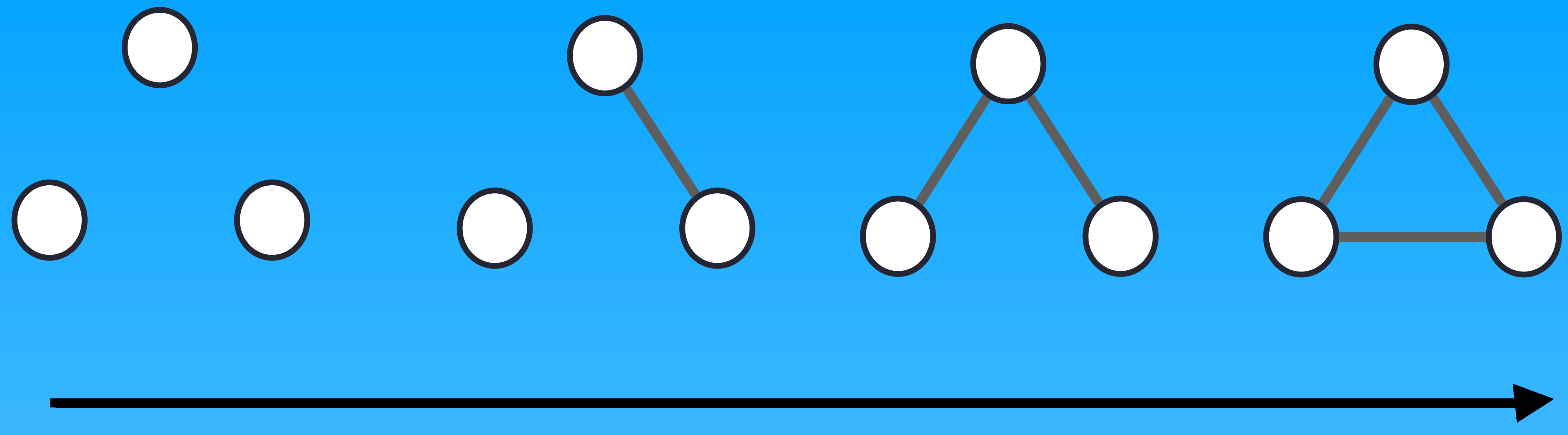
t = 0

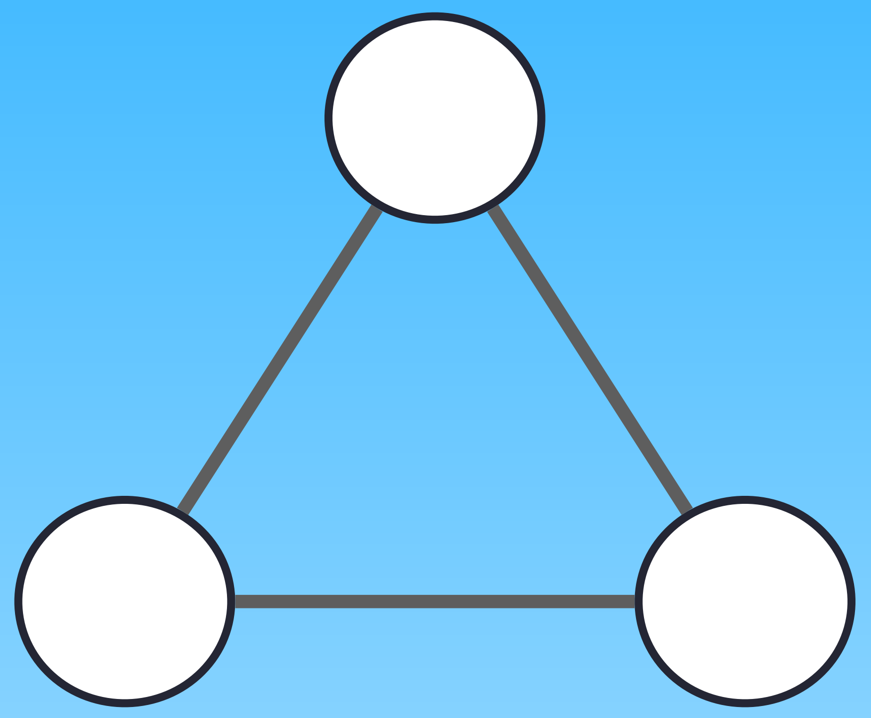
1. Complying to a solidarity norm (Lindenberg, 2015)
2. Strategically investing in a long-term relationship (Coleman, 1991)
3. Controlling one's reputation (Buskens & Raub, 2005)

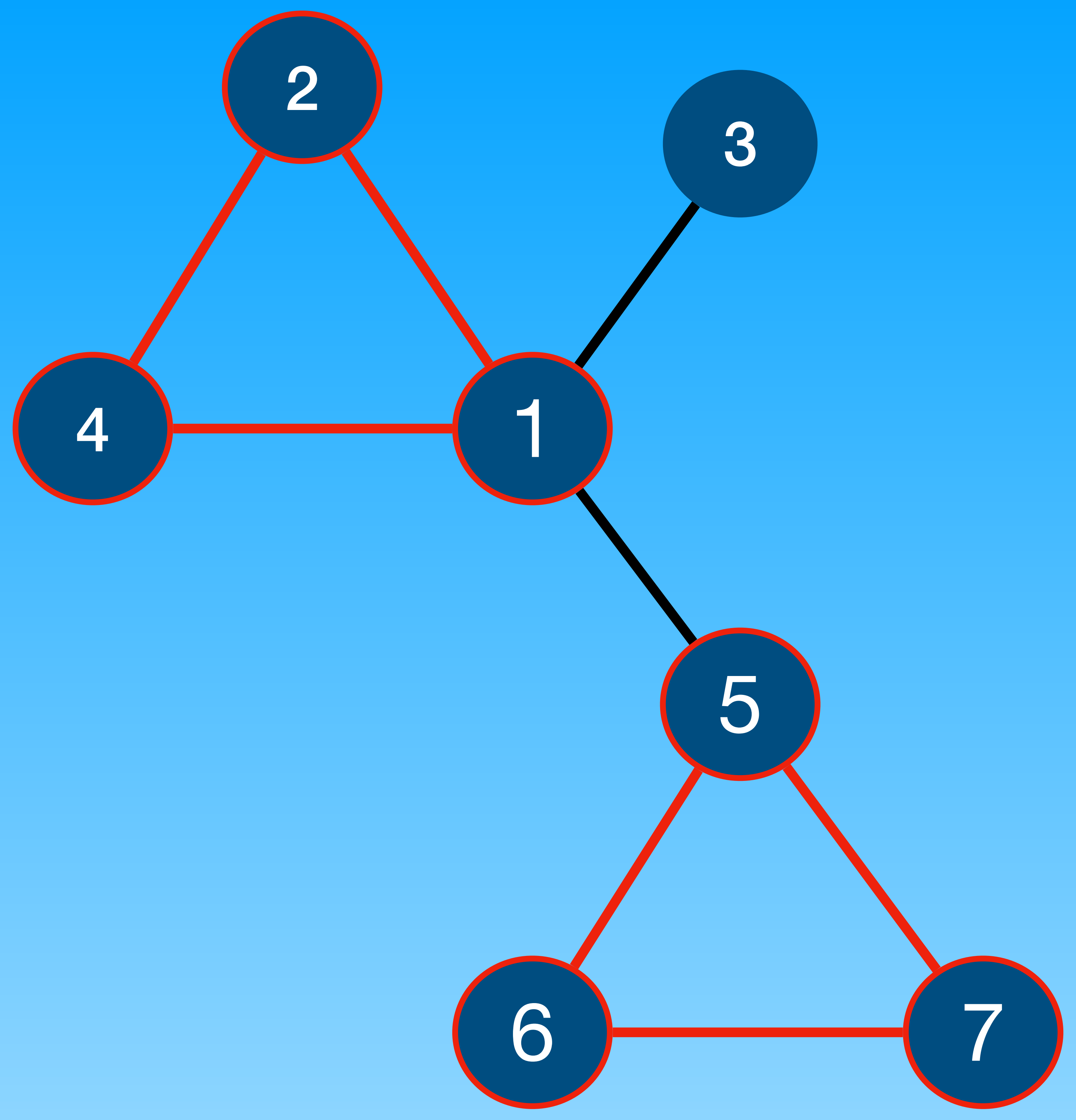
t = 1

Underdetermination of statistical models

- Statistical models of social networks usually provide **underdetermined evidence of causal mechanisms**
- "Network patterns" (Robins, 2015) or "network mechanisms" (Stadtfeld & Amati, 2021) underlie different possible causal mechanisms







Why?

Methodological models

- Prevalence or incidence of the "**archeological traces**" of unobserved, past relational processes (White, 1970, 2008; Lusher et al., 2013)
- **Mathematical tractability**: sufficient statistics of local configurations + parameters estimated via robust algorithms (maximum likelihood or method of moments)
- "**Methodological models**" (Skvoretz, 1991; Sørensen, 1998): finding internal associations within aggregate-level data

ABMs can complement for statistical models' limits concerning:

- **actors' behaviour**
- **tie types**
- **context**

- Tie-based models (e.g., ERGM-family) are indifferent to the specific tie sequences through which particular configurations emerge (Block et al., 2019)

- To be mathematically tractable, (most) SAOMs need assuming agents':

  - access to information about the whole network (e.g., geometrically weighted configurations): unplausible for large networks or competitive contexts where **information is strategically concealed** (e.g., Renzini et al., 2023)

  - tie selection as a multinomial choice based on preference optimization: unplausible for cognitive relations **not requiring psychological investment** (liking vs. disliking, status attribution)

  - myopia: prevents modelling a) **backward-looking rationality** and **learning** processes; b) **forward-looking rationality** (strategic behaviour in competitive contexts)

  - changing one tie at each simulation step: prevents modelling coordination and collective action (Leifeld & Cranmer, 2019) and cascade dynamics driven by **threshold-based preferences** (Renzini et al., 2019)

ABMs can complement for statistical models' limits concerning:

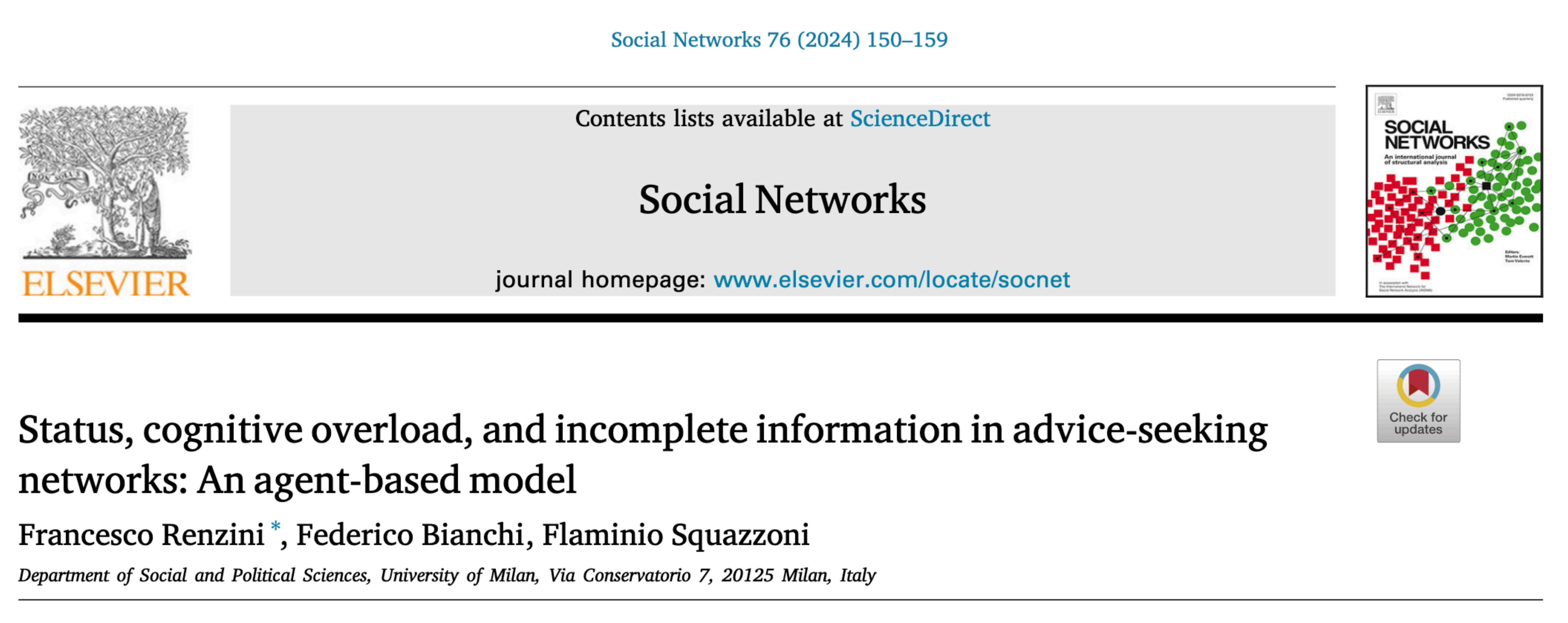
- **actors' behaviour**
- **tie types**
- **context**

- Tie-based models (e.g., ERGM-family) are indifferent to the specific tie sequences through which particular configurations emerge (Block et al., 2019)
- To be mathematically tractable, (most) SAOMs need assuming agents':
  - access to information about the whole network (e.g., geometrically weighted configurations): unplausible for large networks or competitive contexts where information is strategically concealed (e.g., **instrumental ties**, as in Renzini et al., 2023)
  - tie selection as a multinomial choice based on preference optimization: unplausible for **cognitive relations** not requiring psychological investment (liking vs. disliking, status attribution)
  - myopia: prevents modelling a) backward-looking rationality and learning processes; b) forward-looking rationality (strategic behaviour in competitive contexts)
  - changing one tie at each simulation step: prevents modelling coordination and collective action (Leifeld & Cranmer, 2019) and cascade dynamics driven by threshold-based preferences (Renzini et al., 2019)

ABMs can complement for statistical models' limits concerning:

- **actors' behaviour**
- **tie types**
- **context**

- Tie-based models (e.g., ERGM-family) are indifferent to the specific tie sequences through which particular configurations emerge (Block et al., 2019)
- To be mathematically tractable, (most) SAOMs need assuming agents':
  - access to information about the whole network (e.g., geometrically weighted configurations): unplausible for **large networks** or **competitive contexts** where information is strategically concealed (e.g., Renzini et al., 2023)
  - tie selection as a multinomial choice based on preference optimization: unplausible for cognitive relations not requiring psychological investment (liking vs. disliking, status attribution)
  - myopia: prevents modelling a) backward-looking rationality and learning processes; b) forward-looking rationality (strategic behaviour in **competitive contexts**)
  - changing one tie at each simulation step: prevents modelling coordination(Leifeld & Cranmer, 2019) and cascade dynamics driven by threshold-based preferences (Renzini et al., 2023)


Social Networks 76 (2024) 150–159

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

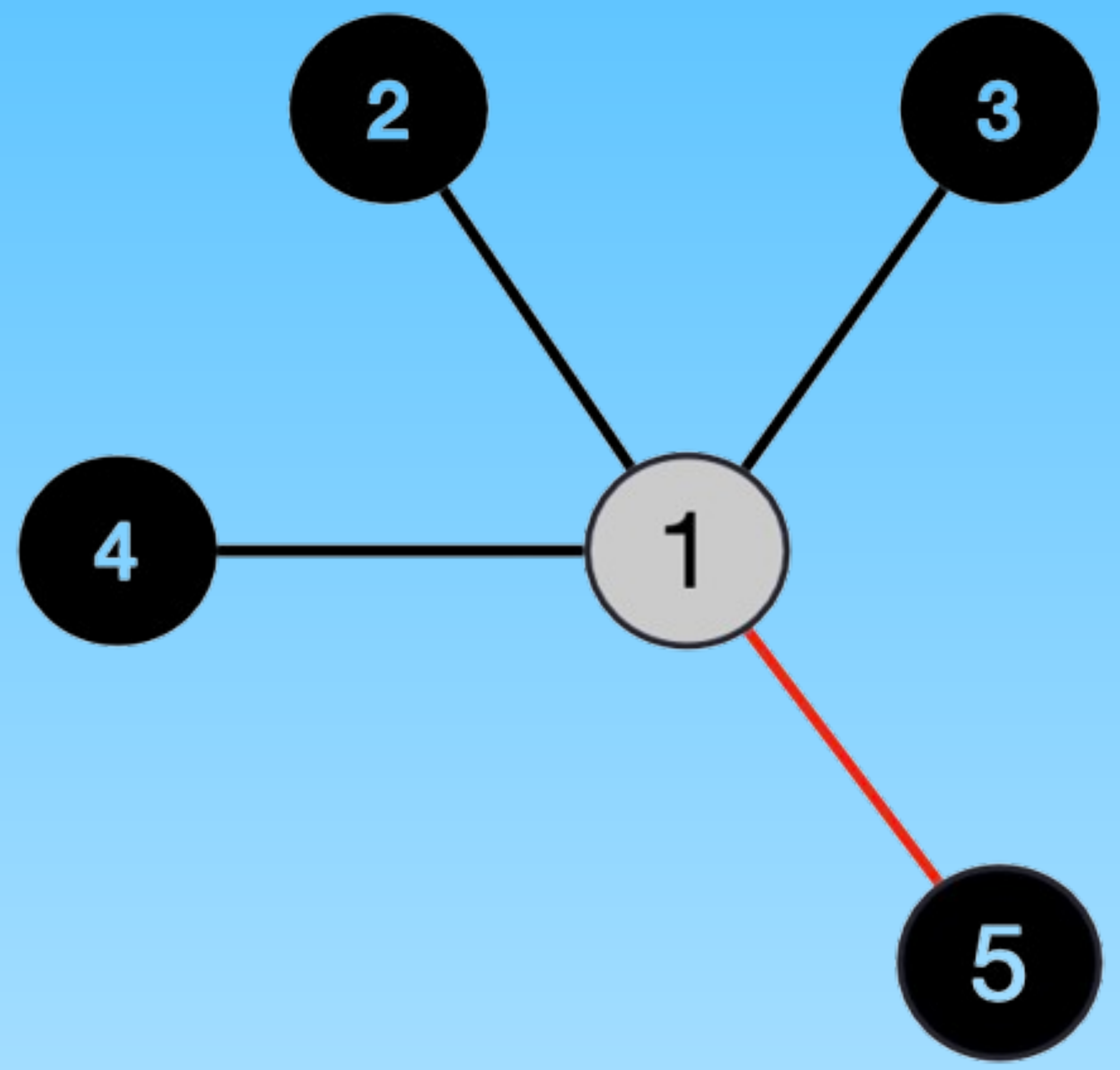
ELSEVIER

Check for updates

Status, cognitive overload, and incomplete information in advice-seeking networks: An agent-based model

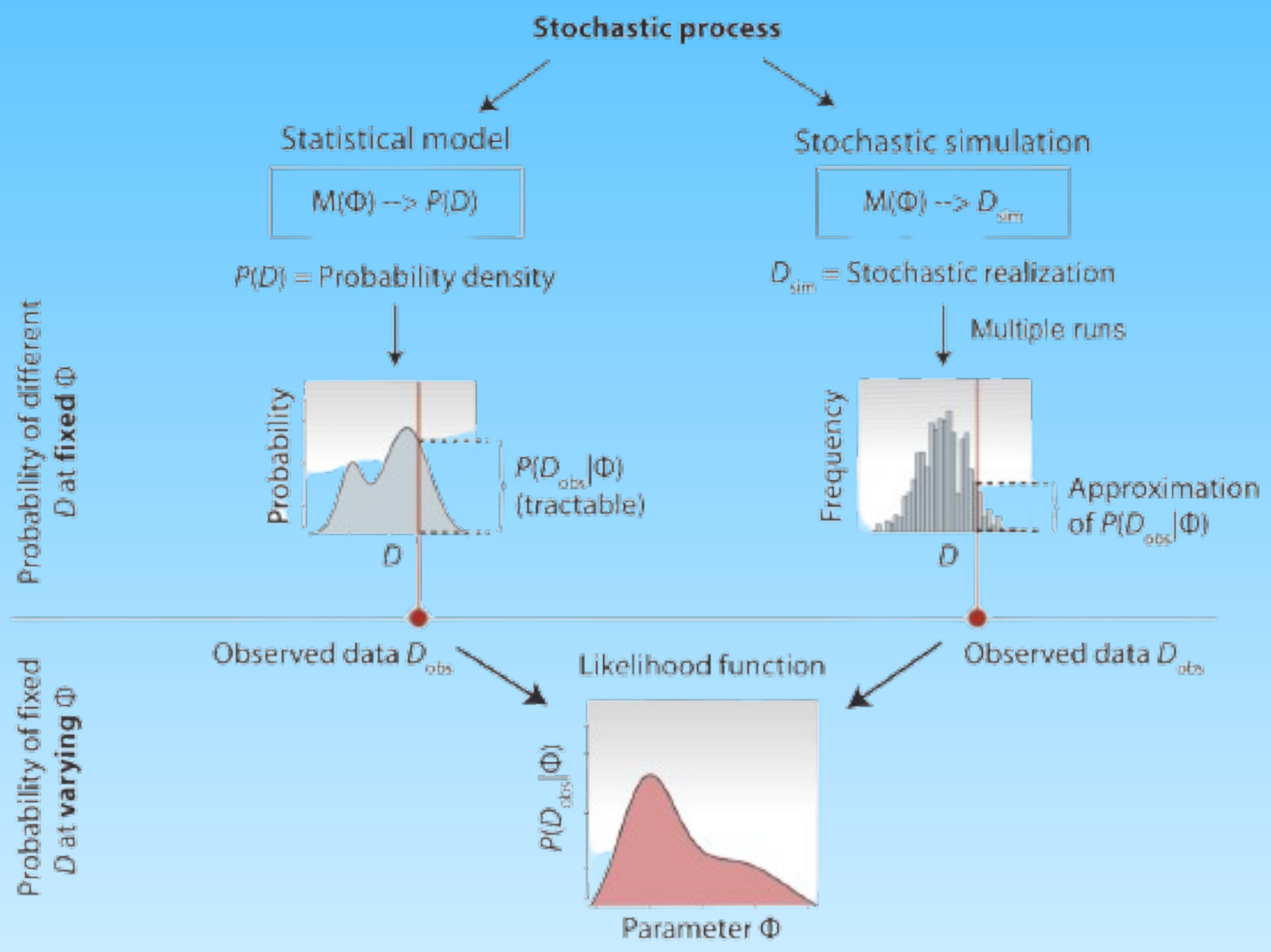
Francesco Renzini *, Federico Bianchi, Flaminio Squazzoni

Department of Social and Political Sciences, University of Milan, Via Conservatorio 7, 20125 Milan, Italy



Examples of ABMs of social networks

- **Renzini, Bianchi, & Squazzoni (2023)**:
  - Explaining advice-seeking network formation as the outcome of request overload (threshold-based)
  - Limited information, local heuristics, plausible and parsimonious model
  - Fitted to classic Lazega’s (2001) network
- **Bianchi, Bellotti, & Renzini (*wip*)**:
  - Explaining low adoption rates of malaria prevemptive practices in tribal villages in Meghalaya (India)
  - Complex contagion via information ties (threshold-based) * negative influence



Theoretical, yet empirical

- **Generativist method** (Epstein, 2006): **sequential complexification** of the modelled mechanism along with computer simulations **until the generated outcome fits the empirical observations** (summary statistics)

- **Testing for unobserved** (unobservable?) **mechanism components** (e.g., thresholds, motives, etc.)

- Simulation-based **point estimates** of parameters and **uncertainty measures** for **untractable likelihood functions** (Hartig et al., 2011; Carrella, 2021)

- **No need to rely on unplausible assumptions** to obtain a tractable likelihood function

Conclusions

- ABM of social networks to estimate **unobserved** or **unobservable** processes
- Bringing back context-dependent **behaviour** and **cognition** (type of ties) to the core of explanations of social phenomena
- Experiment (Brashears & Gladstone, 2020)
- Middle-range social science

References / 1

Amaturo, E. (1997), *Premessa all'edizione italiana*, in J. Scott, *L'analisi delle reti sociali*, a cura di E. Amaturo, Roma, La Nuova Italia Scientifica, pp. 9-21.

Block, P., Stadtfeld, C. e Snijders, T.A.B. (2019), *Forms of Dependence: Comparing SAOMs and ERGMs from Basic Principles*, in «Sociological Methods & Research», 48, n. 1, pp. 202-239.

Borgatti, S.P., Mehra, A., Brass, D.J. e Labianca, G. (2009), *Network Analysis in the Social Sciences*, in «Science», 323, n. 5916, pp. 892-895.

Brandes, U., Robins, G., McCranie, A.N.N. e Wasserman, S. (2013), *What Is Network Science?*, in «Network Science», 1, n. 1, pp. 1-15.

Chiesi, A.M. (1980), *L'analisi dei reticoli sociali: teoria e metodi*, in «Rassegna Italiana di Sociologia», 21, n. 2, pp. 291-310.

Chiesi, A.M. (1981), *L'analisi dei reticoli sociali: un'introduzione alle tecniche*, in «Rassegna Italiana di Sociologia», 22, n. 4, pp. 577-603.

Chiesi, A.M. (1999), *L'analisi dei reticoli*, Milano, Franco Angeli.

Epstein, J.M. (2006), *Generative Social Science: Studies in Agent-Based Computational Modeling*, Princeton, Princeton University Press.

Fagiolo, G., Guerini, M., Lamperti, F., Moneta, A. e Roventini, A. (2019). *Validation of Agent-Based Models in Economics and Finance*, in C. Beisbart e N. Saam (a cura di), *Computer Simulation Validation*, Berlin/Heidelberg, Springer, pp. 763-787.

Leifeld, P., & Cranmer, S.J. (2019). A Theoretical and Empirical Comparison of the Temporal Exponential Random Graph Model and the Stochastic Actor-Oriented Model. Network Science, 7(1), 20-51.

References / 2

Lusher, D., Koskinen, J. e Robins, G. (2013, a cura di), *Exponential Random Graph Models. Theory, Methods, and Applications*, New York, Cambridge University Press.

McFadden, D. (1973), *Conditional Logit Analysis of Qualitative Choice Behavior*, in P. Zarembka (a cura di), *Frontiers in Econometrics*, New York, Academic Press, pp. 105-142.

Piselli, F. (1995, a cura di), *Reti. L'analisi di network nelle scienze sociali*, Roma, Donzelli.

Renzini, F., Bianchi, F., & Squazzoni, F. (2024). Status, cognitive overload and incomplete information in advice-seeking networks: an agent-based model. *Social Networks*, *76*: 150–159.

Robins, G. (2015), *Doing Social Network Research. Network-Based Research Design for Social Scientists*, Sage, London

Scott, J. (1991), *Social Network Analysis*, London, Sage.

Snijders, T.A.B. (2017), *Stochastic Actor-Oriented Models for Network Dynamics*, in «Annual Review of Statistics and Its Applications», 4, pp. 434-363.

Squazzoni, F. (2012), *Agent-Based Computational Sociology*, Chichester, Wiley.

Stadtfeld, C. e Amati, V. (2021), *Network Mechanisms and Network Models*, in G. Manzo (a cura di), *Research Handbook on Analytical Sociology*, Cheltenham, Elgar, pp. 432-452.

Thiele, J.C., Kurth, W.E. e Grimm, V. (2014). *Facilitating Parameter Estimation and Sensitivity Analysis of Agent-Based Models: A Cookbook Using NetLogo and 'R'*. «Journal of Artificial Societies and Social Simulation«, 17, n. 3, 11.

White, H.C. (1970), *Matching, Vacancies, and Mobility*, in «Journal of Political Economy», 78, n. 1, pp. 97-105.