GENEVA GRADUATE INSTITUTE

INSTITUT DE HAUTES ÉTUDES INTERNATIONALES ET DU DÉVELOPPEMENT

GRADUATE INSTITUTE OF INTERNATIONAL AND DEVELOPMENT STUDIES

SAOMs

James Hollway

Feedback on midterms

- Generally very well done

- Grades /30, where 15 $\approx$ 4.00

- Reminder: choice of centrality measure should be well motivated, not just an index of all

- Reminder: which community detection algorithm produces highest modularity and/or most interpretable/sensible results

- Reminder: nodes in structural holes are called *brokers*; ties linking communities are called *bridges*

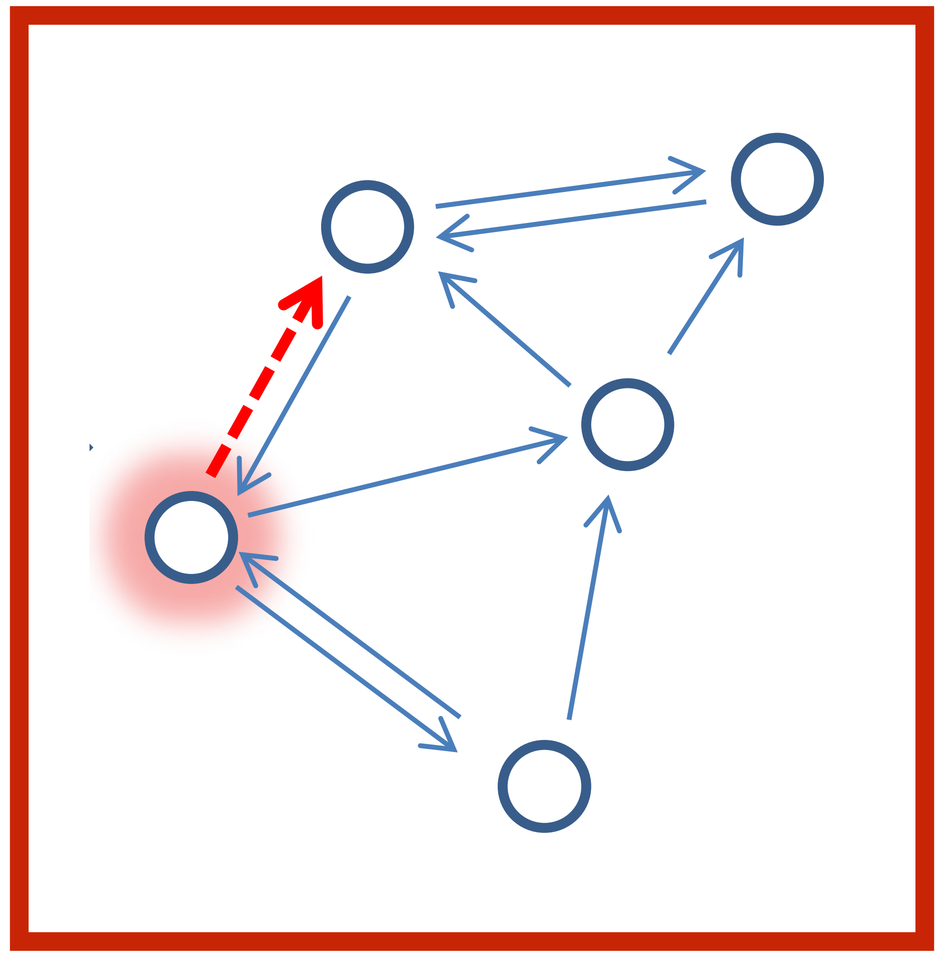
- Reminder: may need to tweak/play with graphs until they illustrate clearly the message you've decided they convey

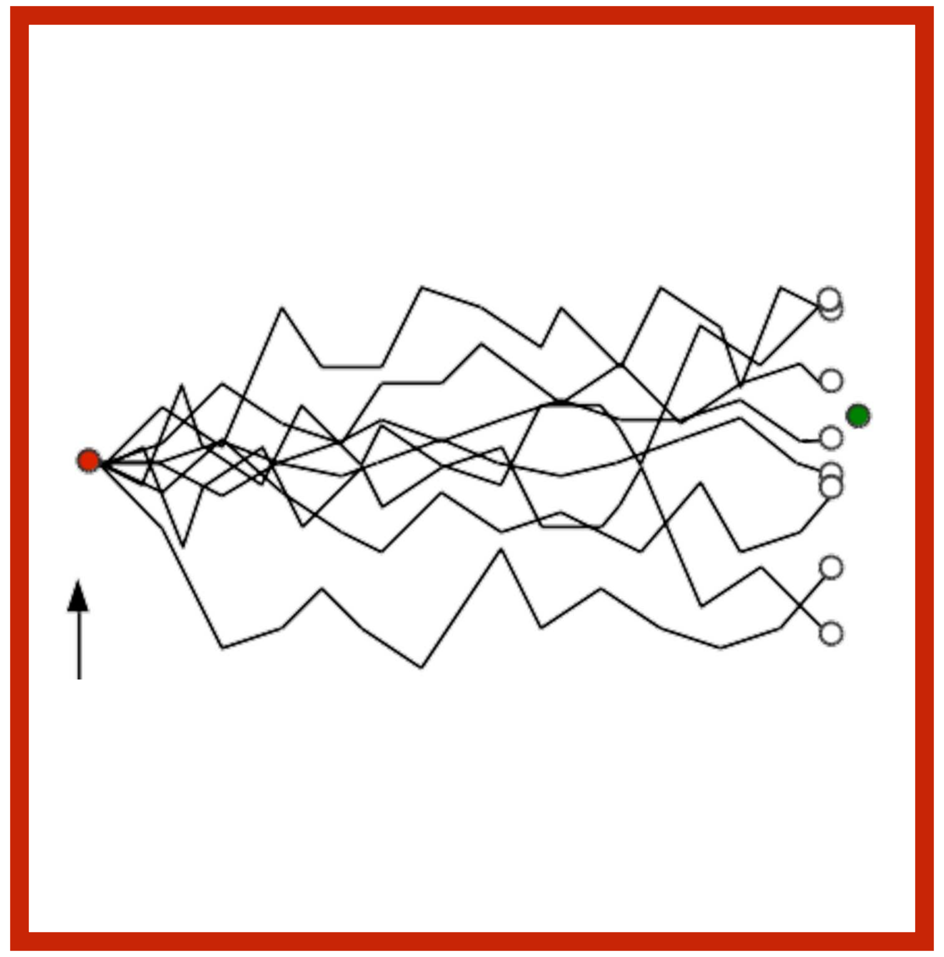
SAOM

Model

Estimation

Influence





Actor vs tie models

MOM vs MLE

Selection vs Influence

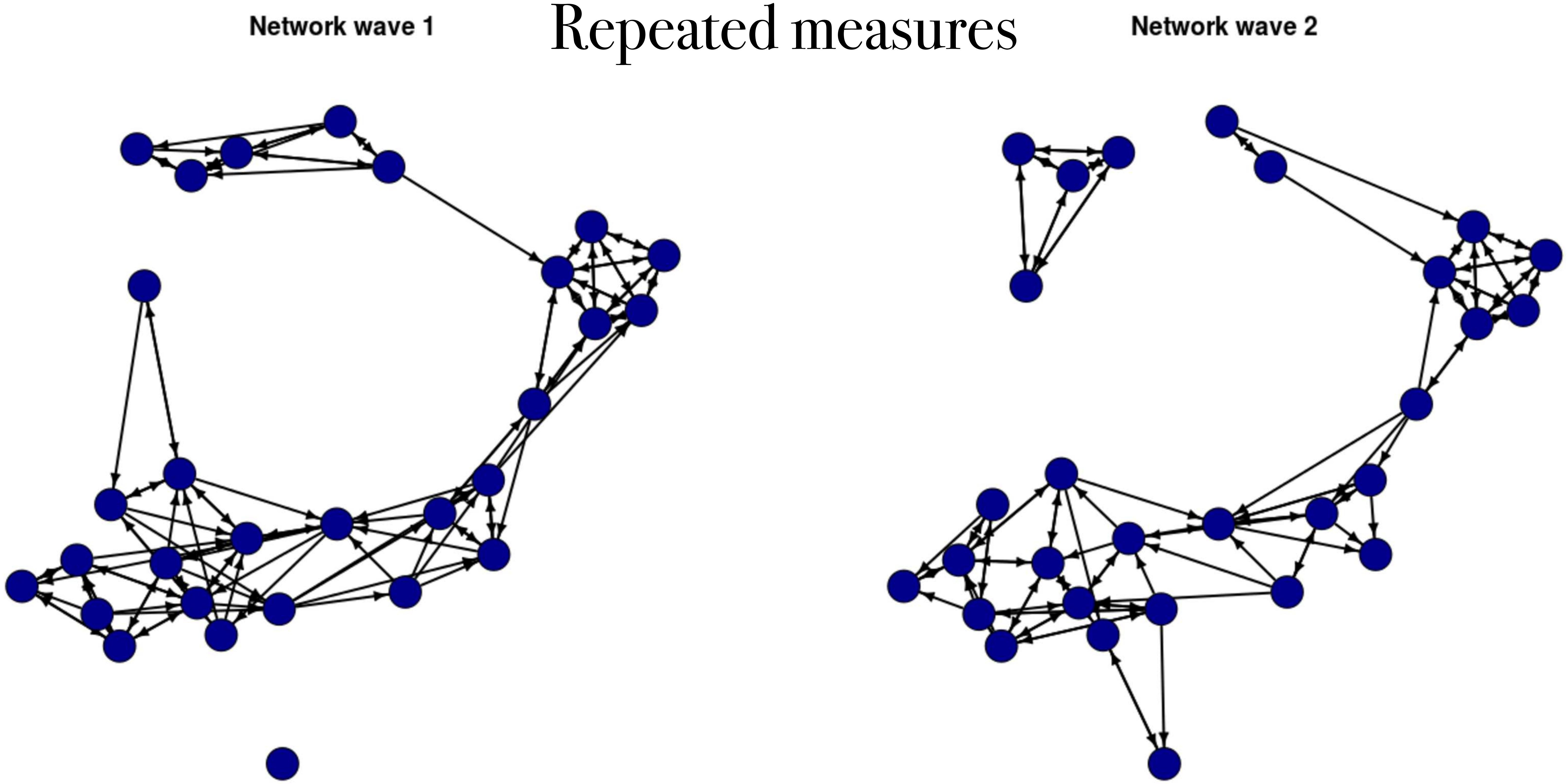
Why Network *Dynamics*

- Because we want to know *why* there are associations
  - E.g. why are depressed people more likely to have depressed friends?
- *Competing* explanations tend to involve *dynamic* mechanisms:
  - because depressed people prefer depressed friends
  - because non-depressed people avoid the depressed
  - because the depressed withdraw from friendly interactions which destroys all other friendships
  - because depression is contagious along friendships

Schaefer et al 2012

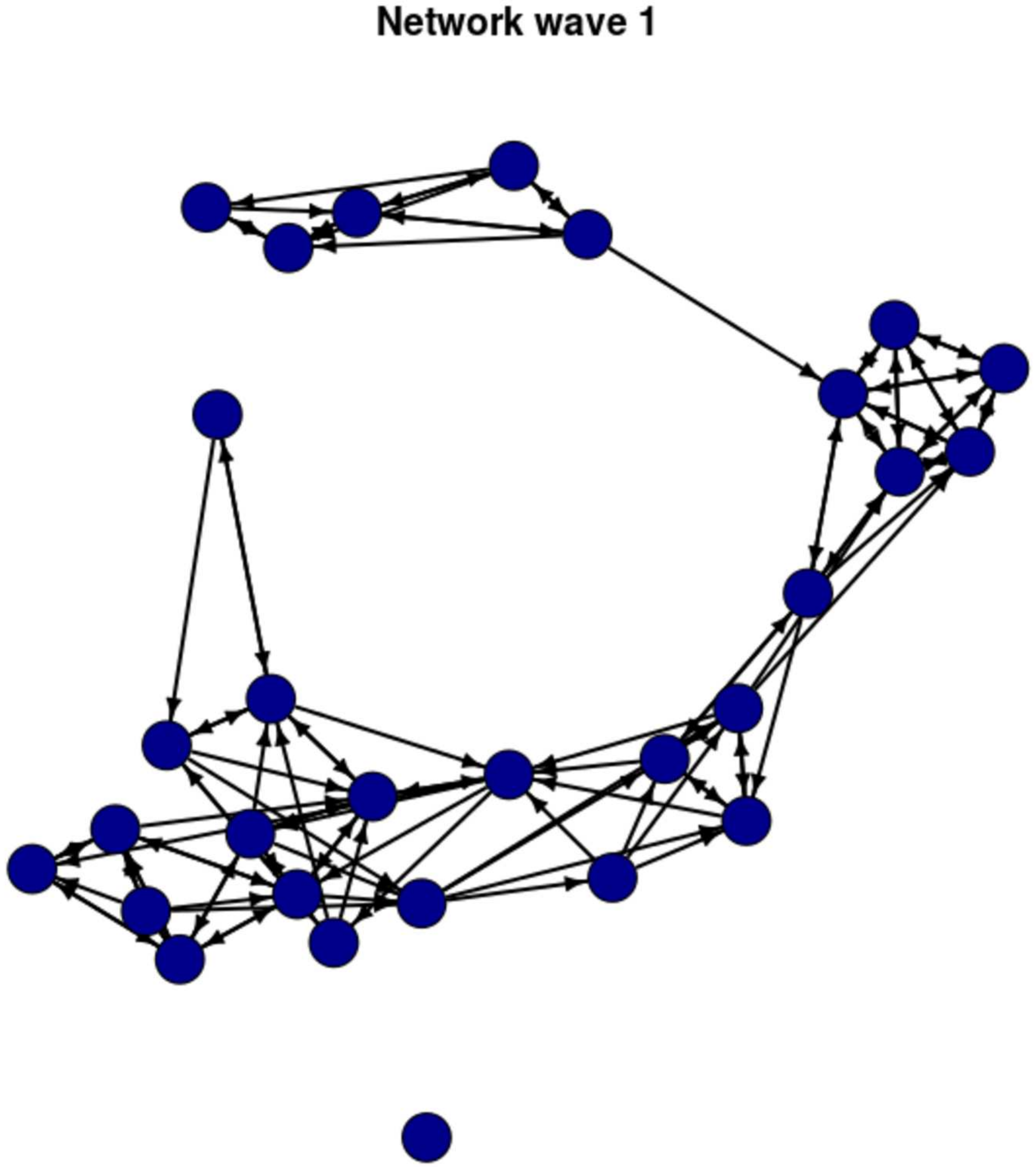
Typical data: panel

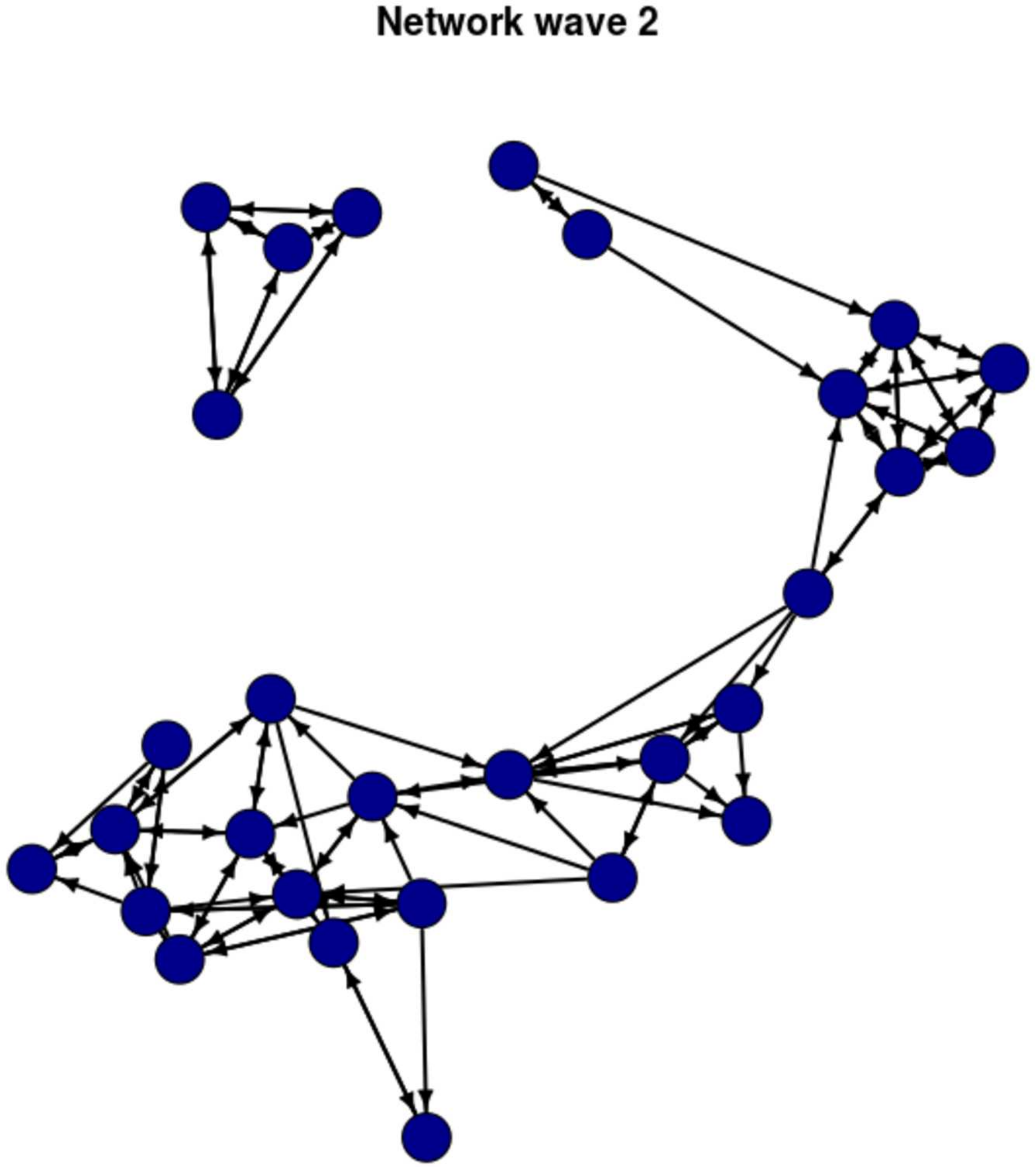
Repeated measures



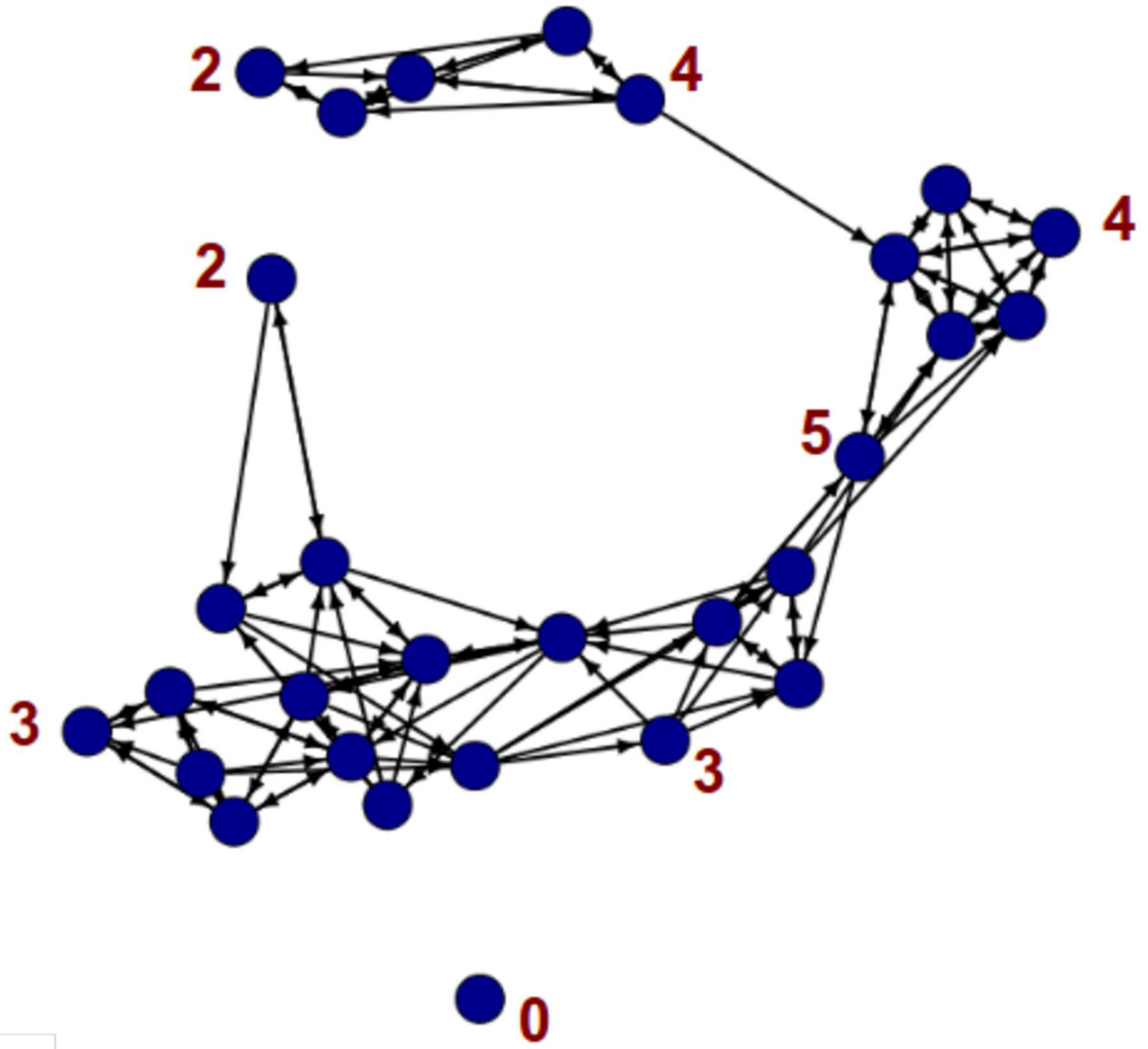
1. Same group of actors (some composition change allowed)
2. Same relational variable (states not events)
3. Some, but not too much change

Which forces shape this social network's evolution?

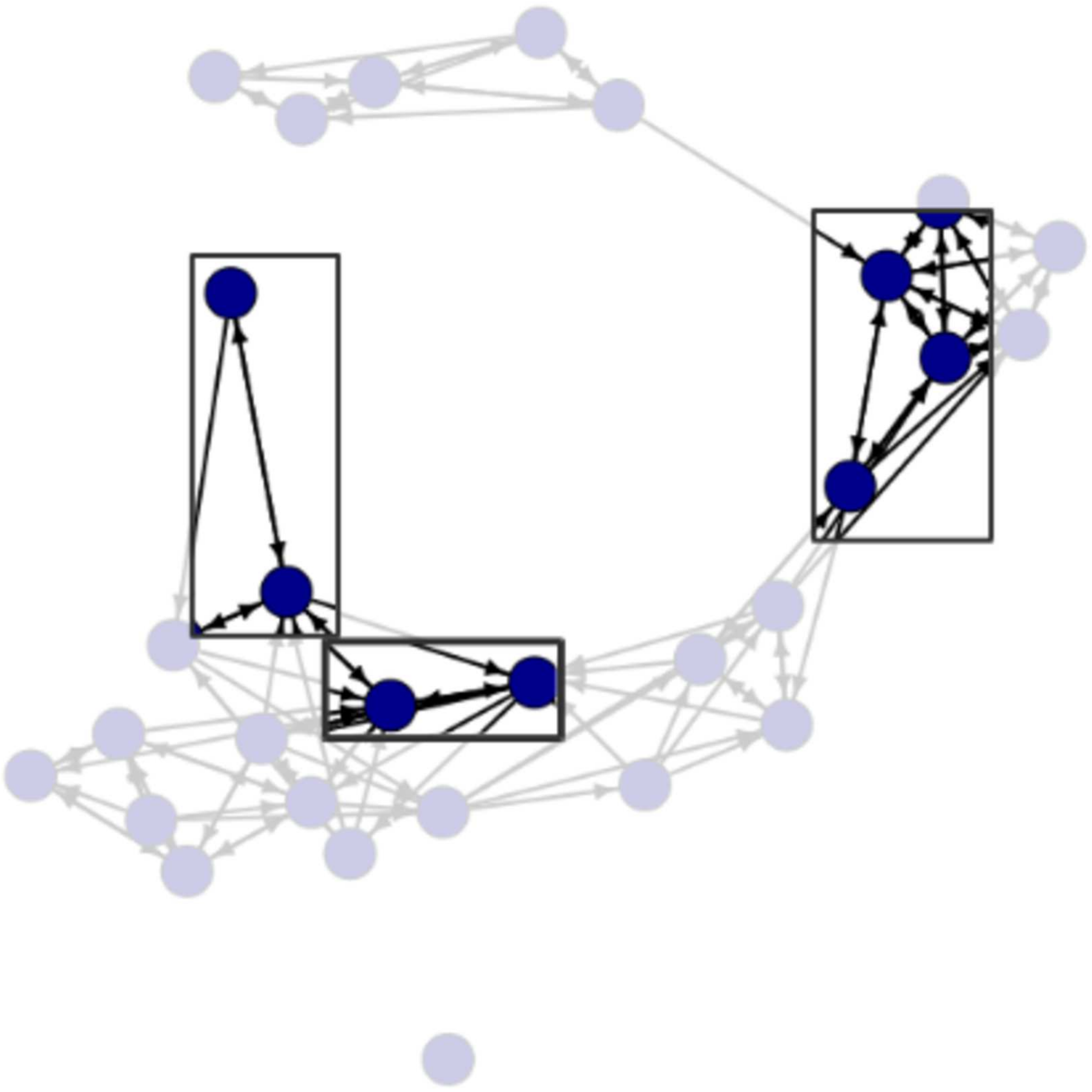




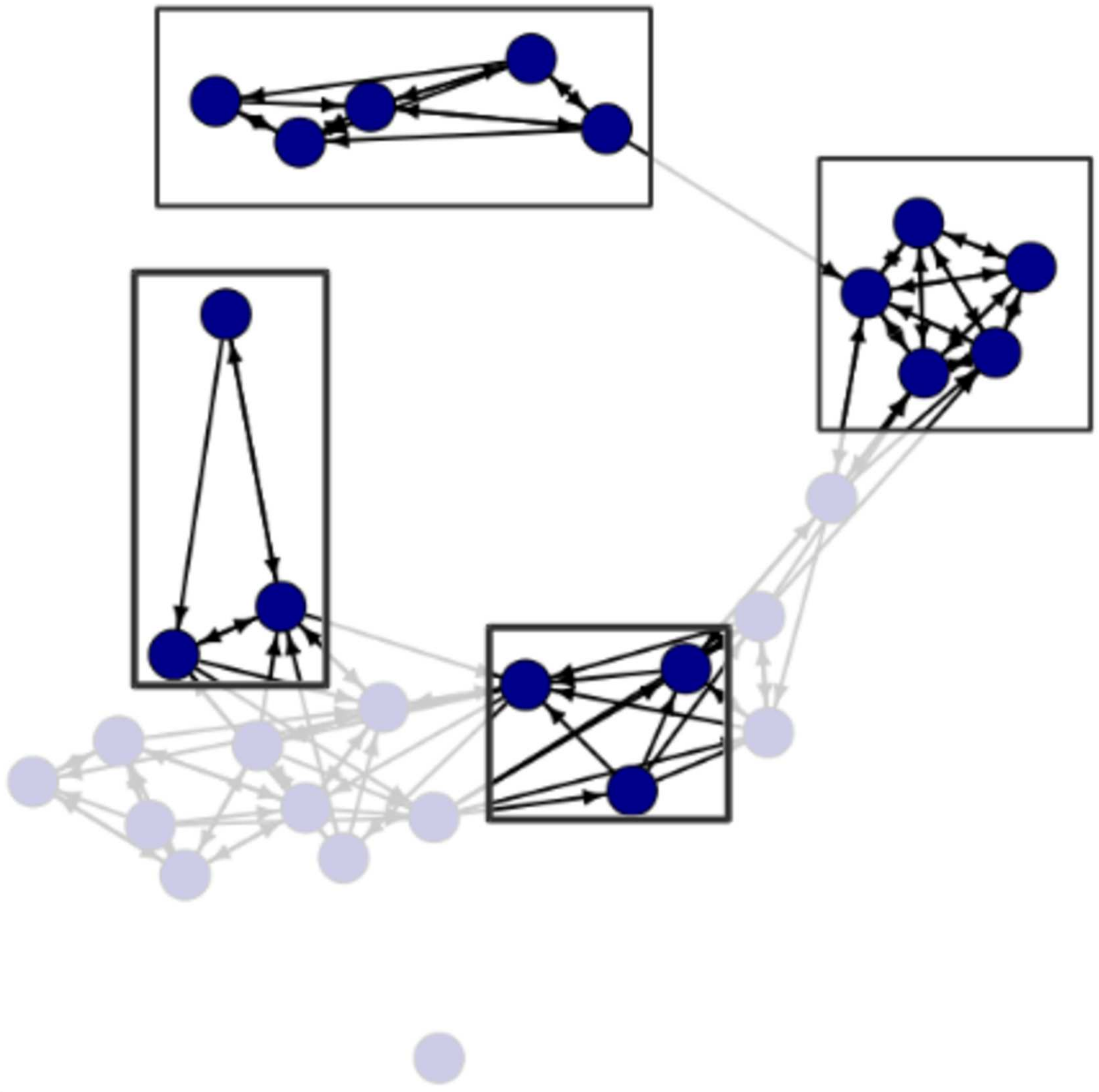
Social network ties are costly



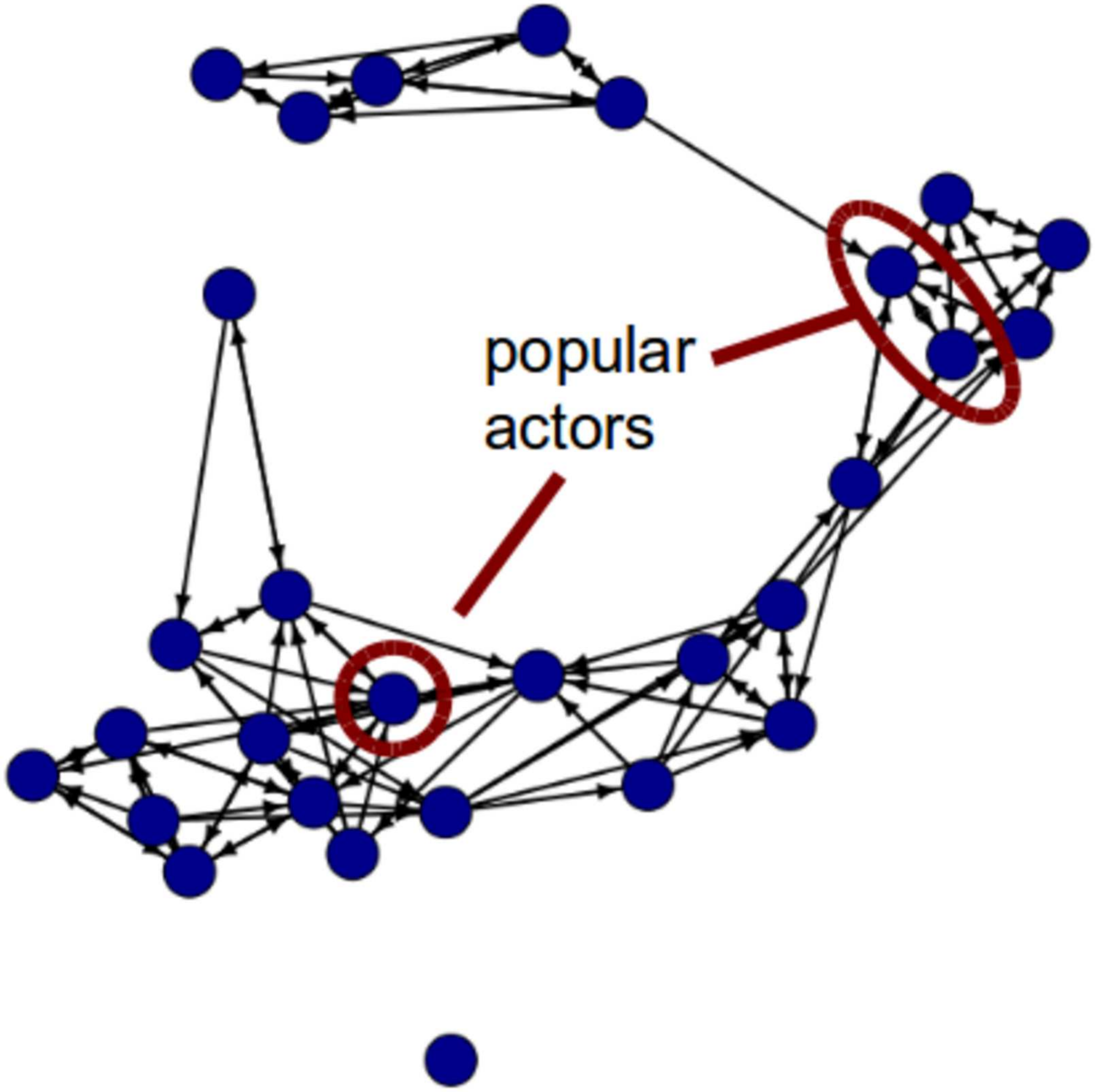
Individuals form and maintain reciprocal ties



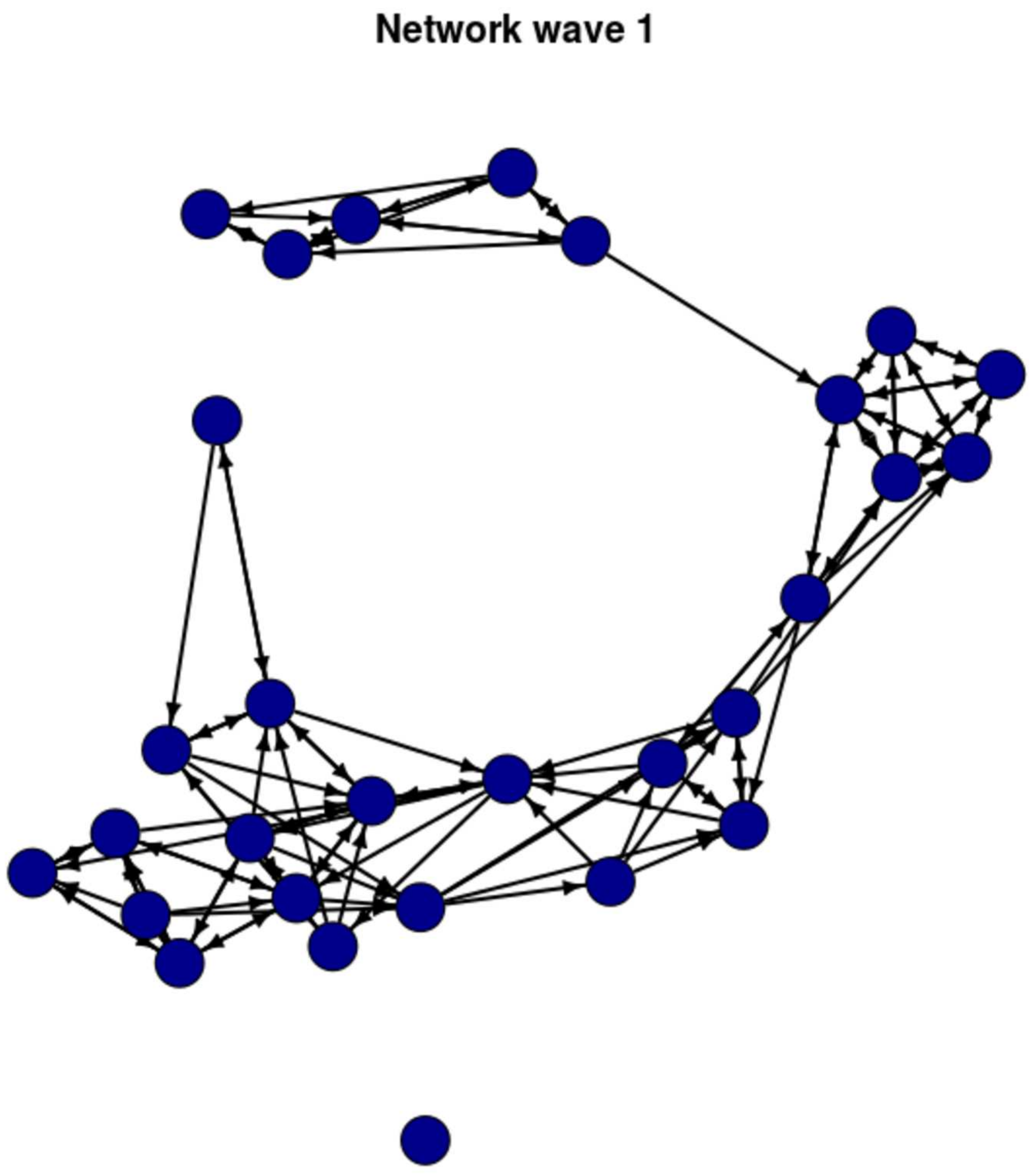
Transitivity leads to clustering

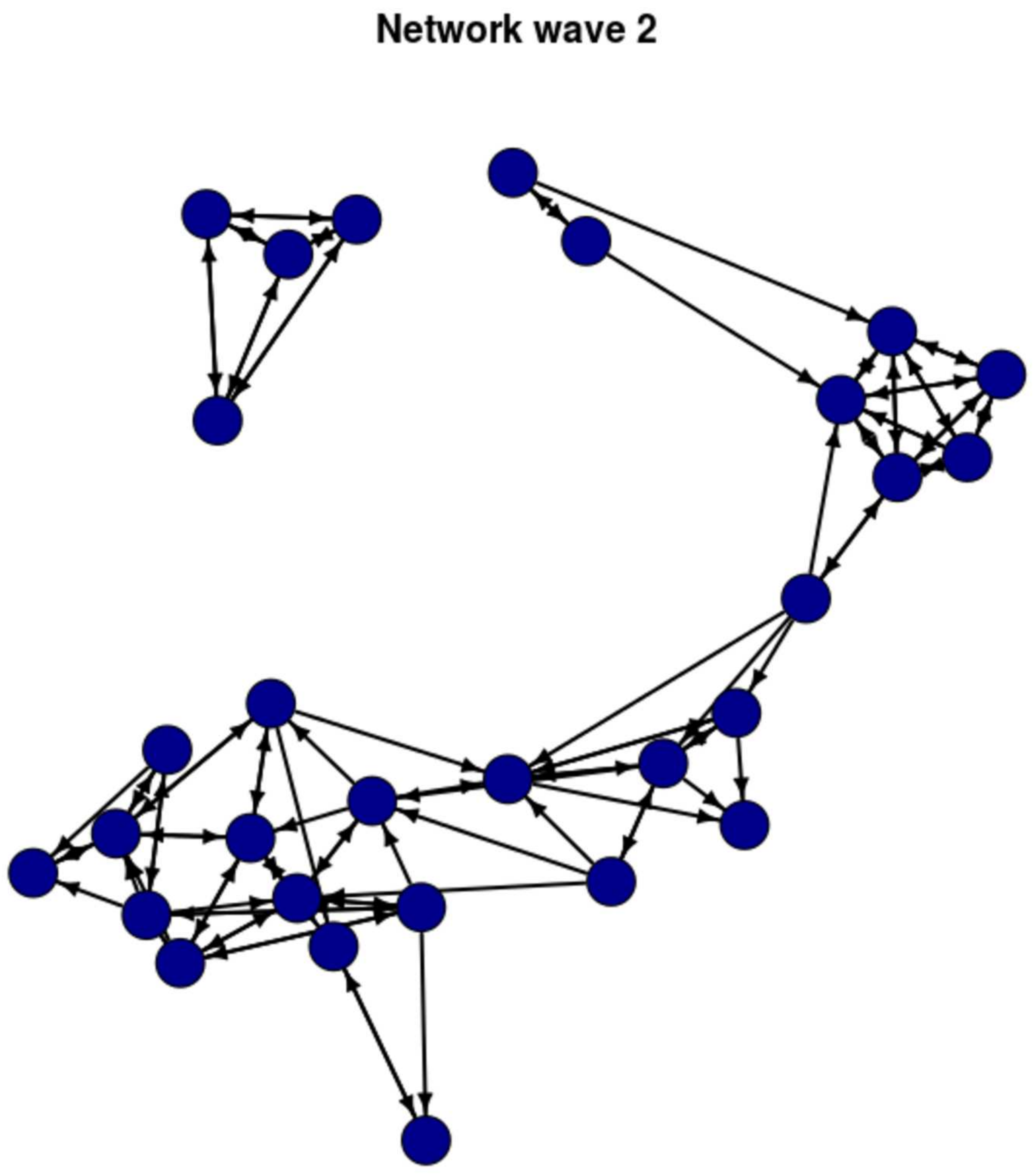


Status hierarchy shapes friendship networks

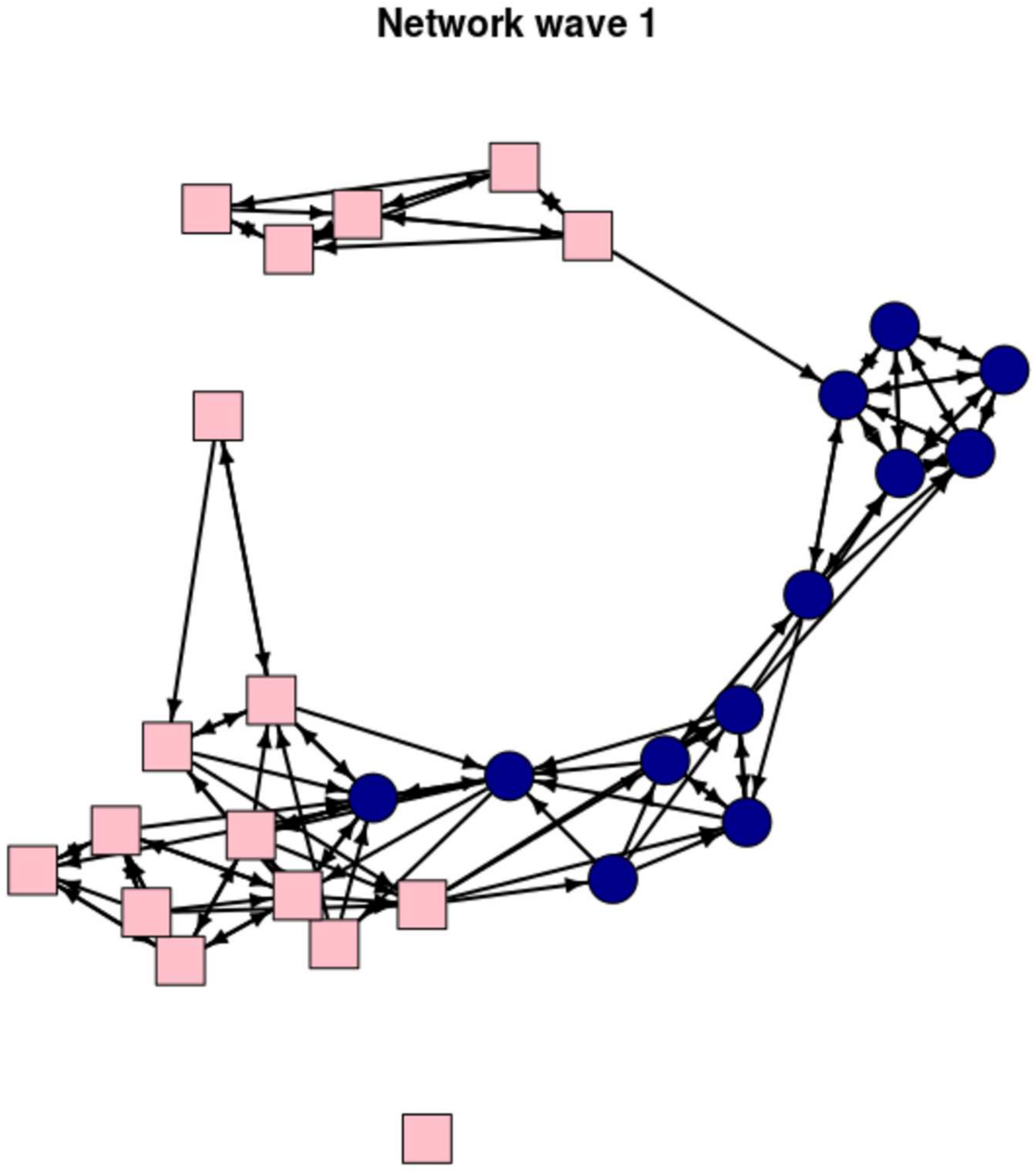


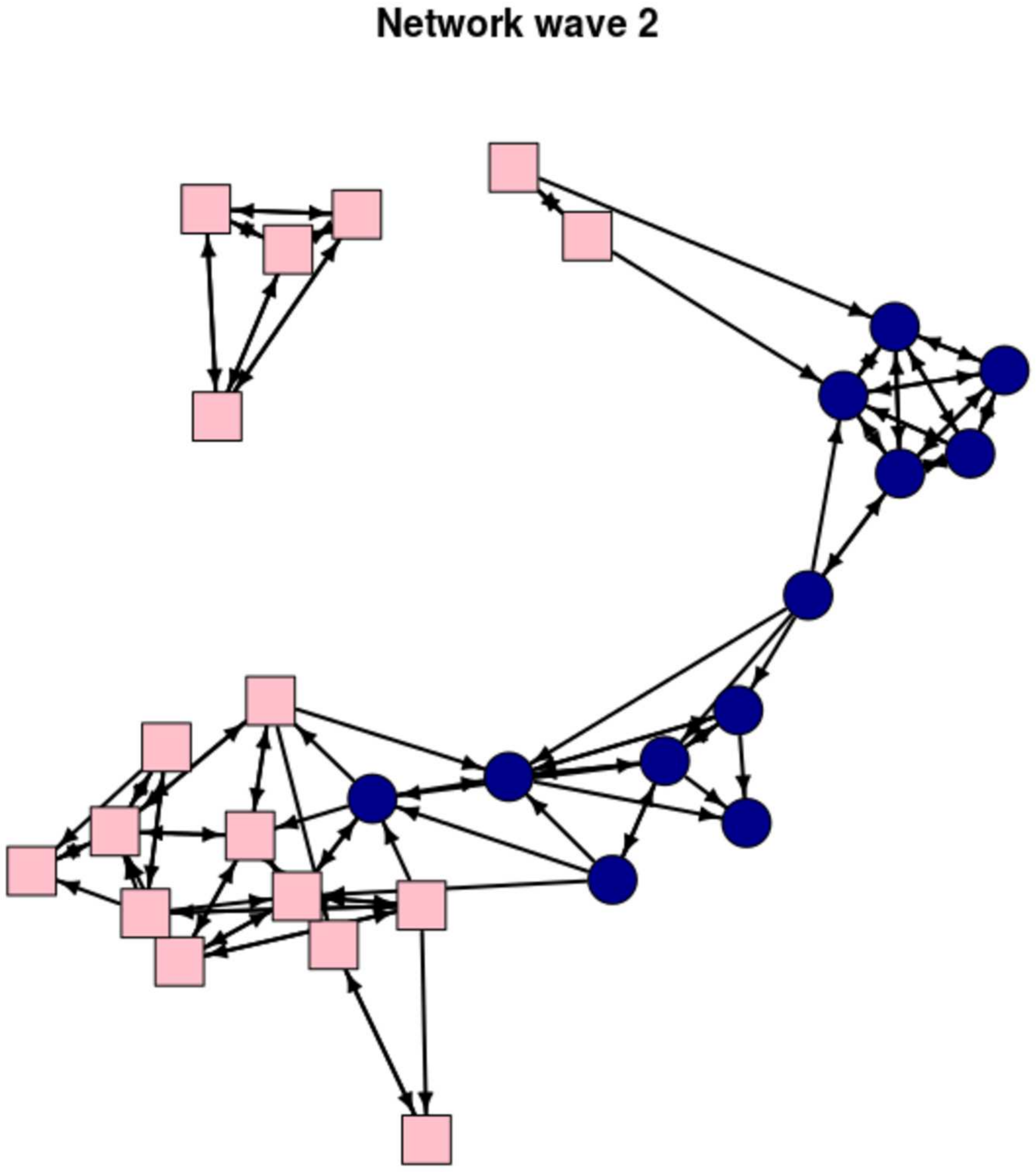
What else?



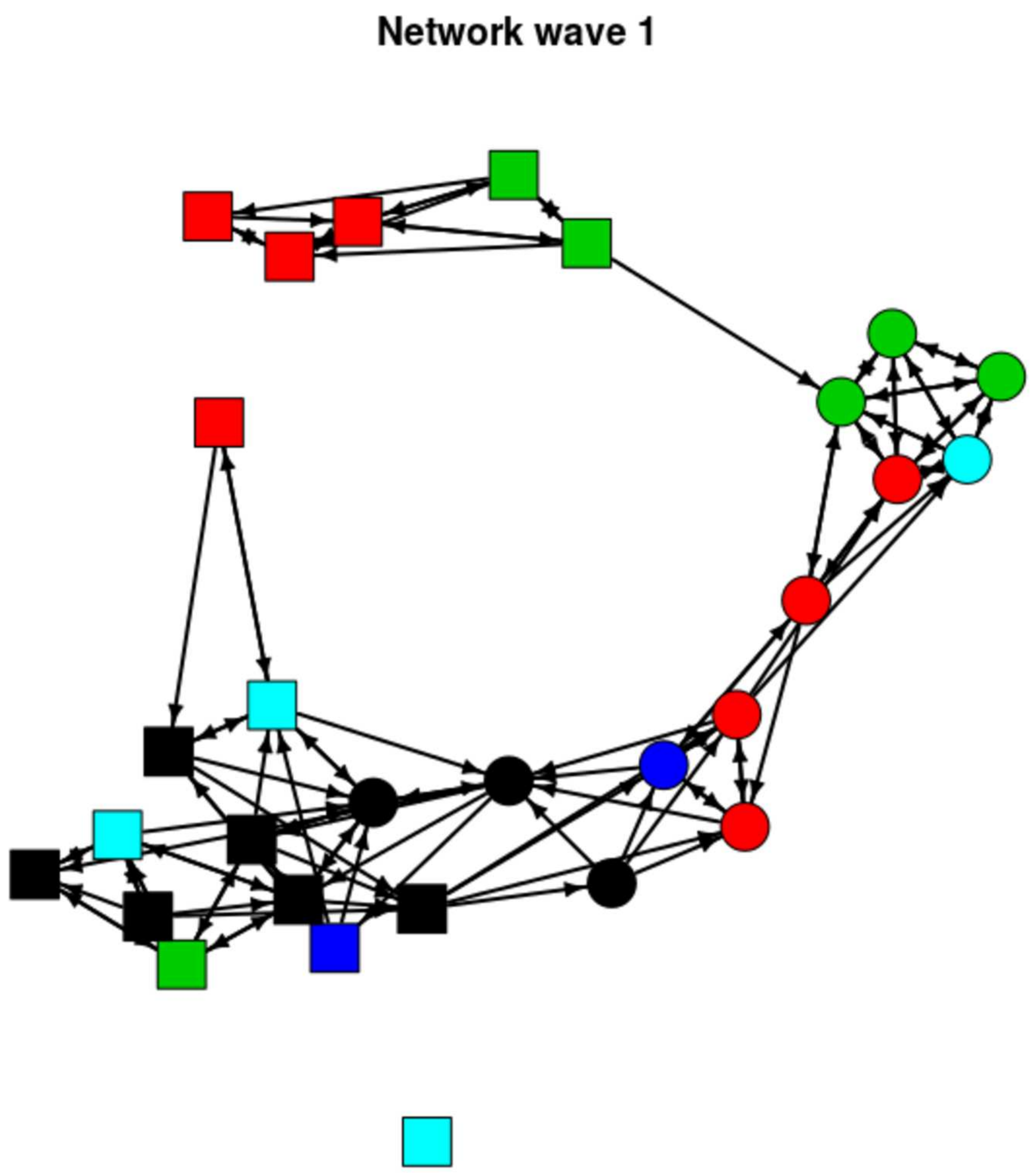


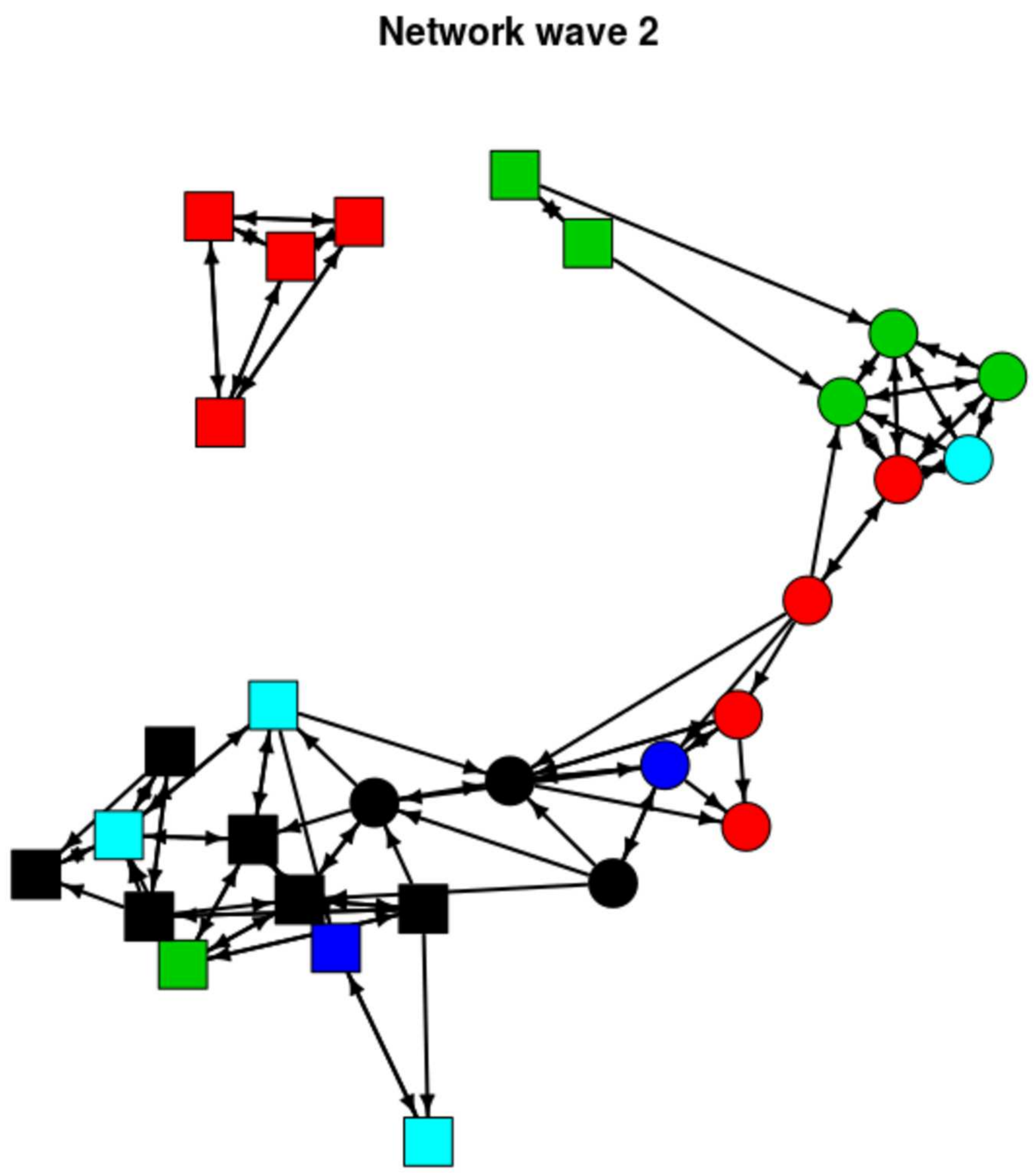
Gender homophily?





Ethnic homophily?





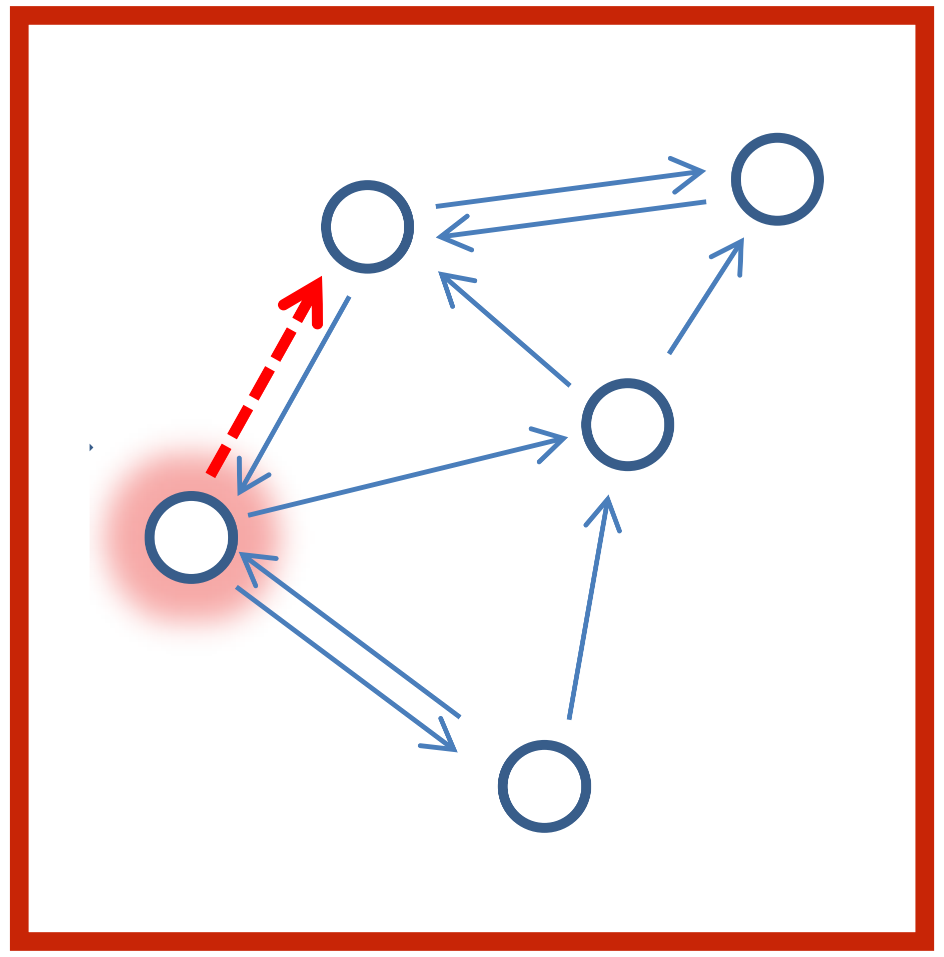
Modelling thoughts

- A statistical approach is necessary to control for alternative explanations
- A complete network approach is necessary because selection can only be studied when the complete pool of candidates is known
- A longitudinal approach is necessary to link antecedents with consequences
- A (weak) methodologically individualist approach is useful to bring the model close to theory

See Udehn 2002

SAOM

Model



Estimation

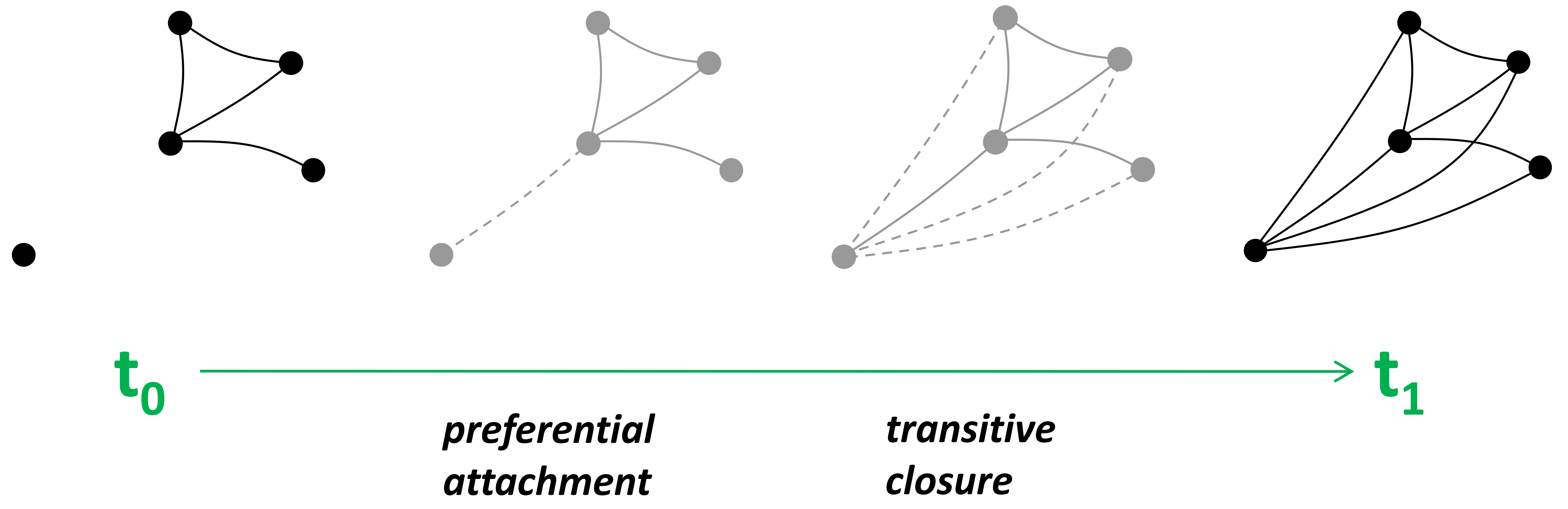
Influence

SAOMs are not ERGMs

- SAOMs are a **continuous-time** network model
  - They model change in social networks in continuous-time using empirical panel data with SIENA (Simulation Investigation for Empirical Network Analysis) (see Block et al 2018)
- SAOMs are an **actor-oriented** network model
  - They model change as a function of individuals' choices about whom they want to relate to and how they want to behave (see Block et al 2019)

Why Continuous-Time?

- Because complex patterns emerge from simple(r) mechanisms

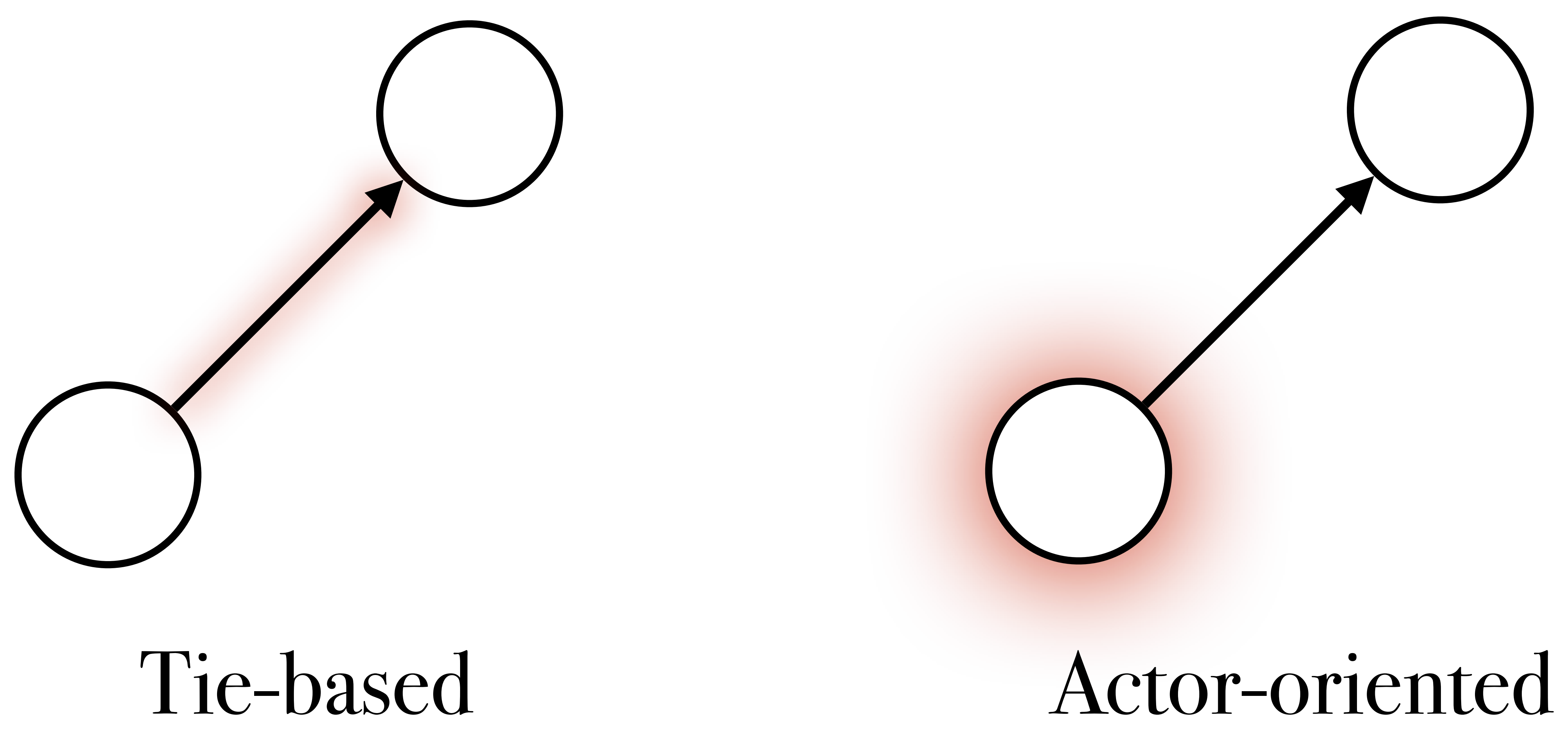


- New ties may be *realisation-contingent* on other new ties.
  - Cannot easily model compound emergence in discrete-time.

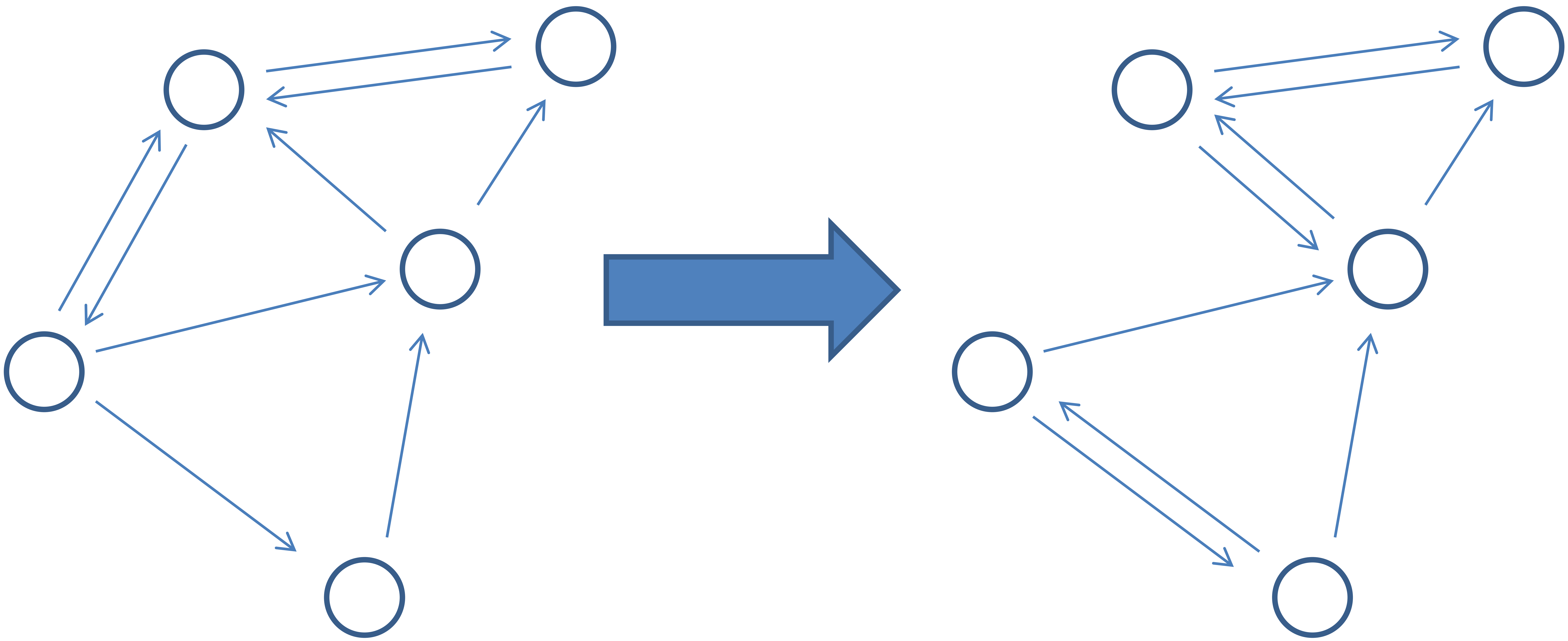
Why Actor-Oriented?

- All social network change is brought about by individual or collective agents that decide to send or drop a tie (homophily, withdrawal, avoidance, etc)

- As the actor is the locus of control, we should model the tie changes from its perspective

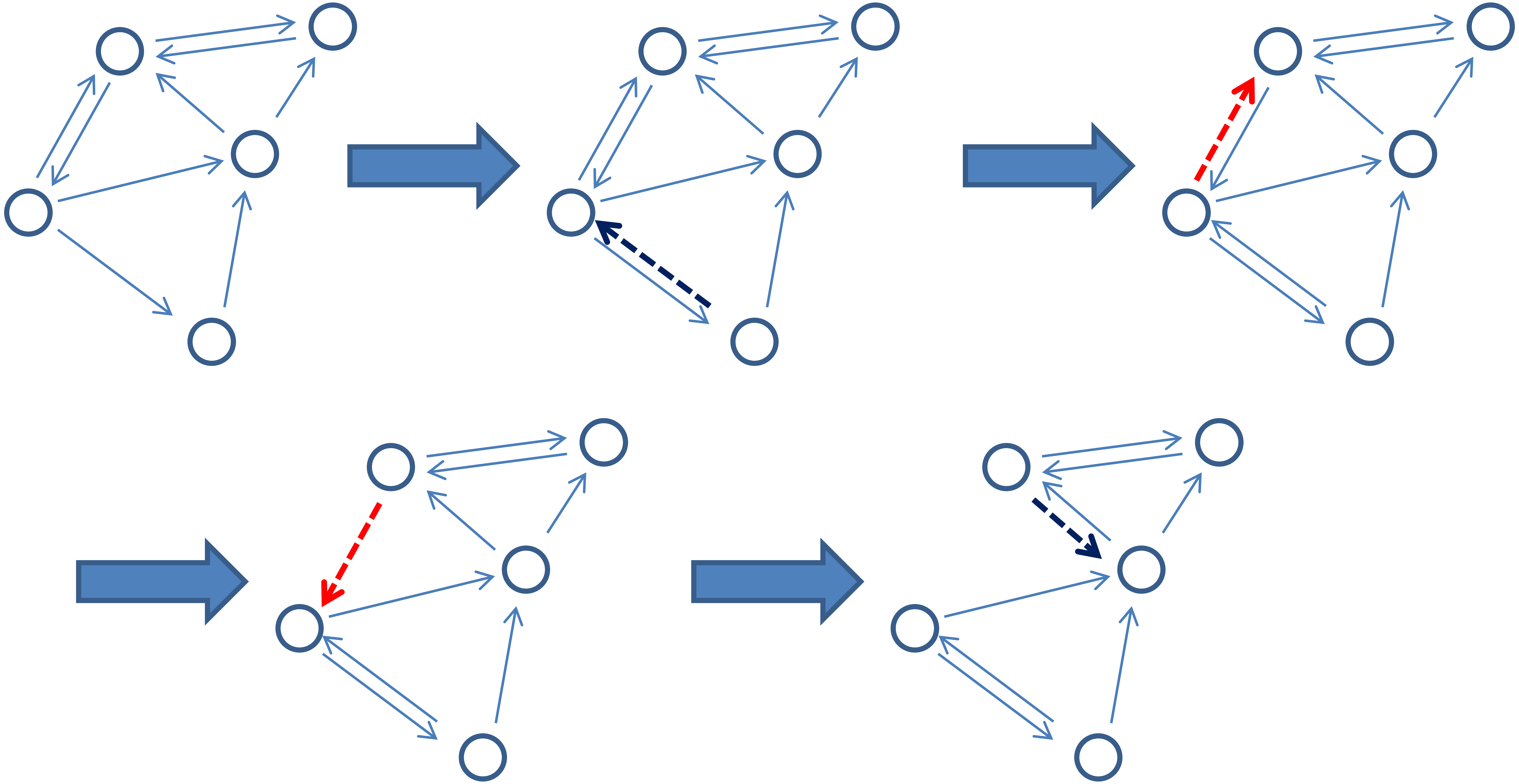


Intuition

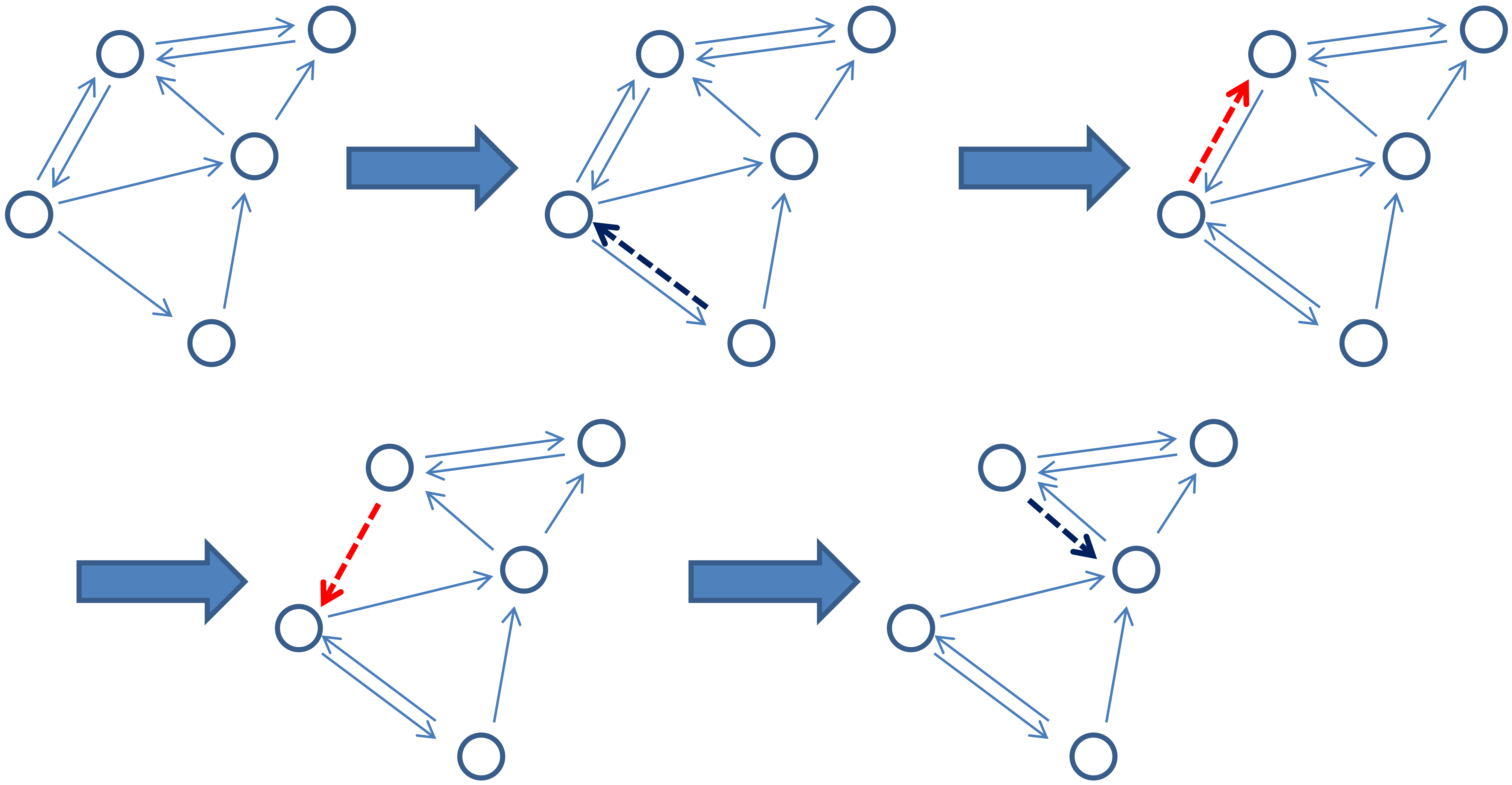


Continuous-Time

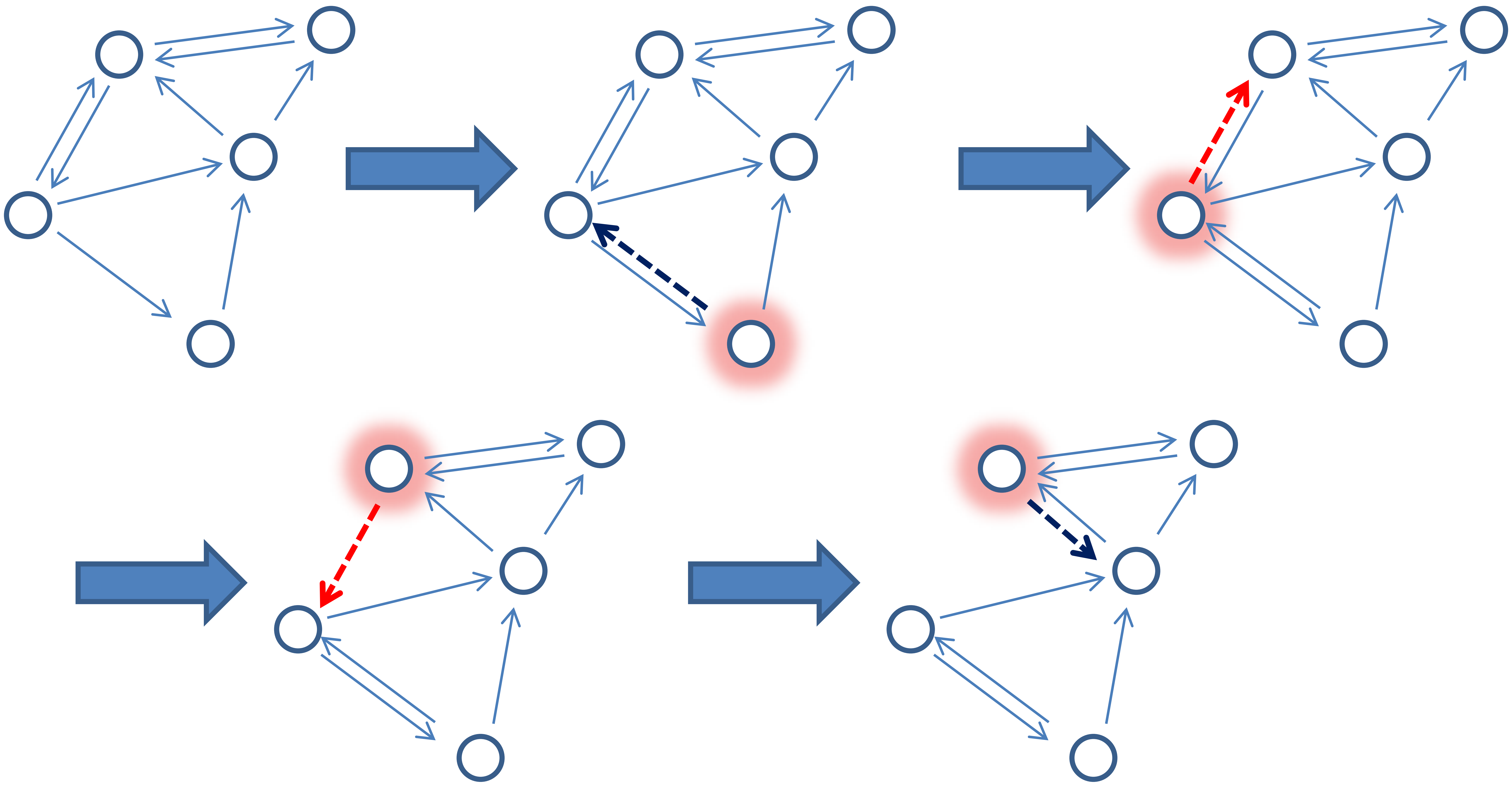
This is one potential path how the network develops from t_1 to t_2



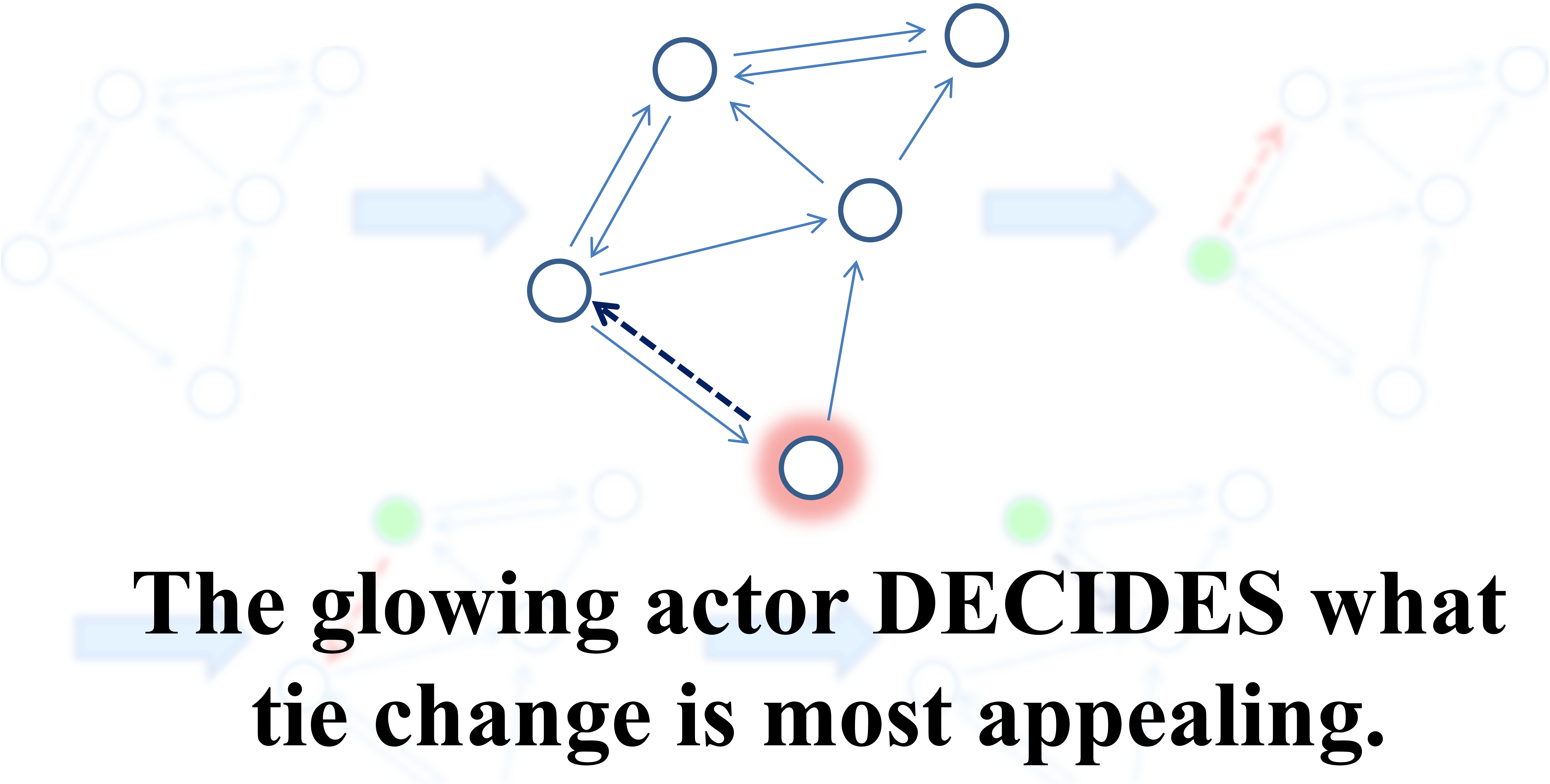
Mini-Step



Actor-Oriented

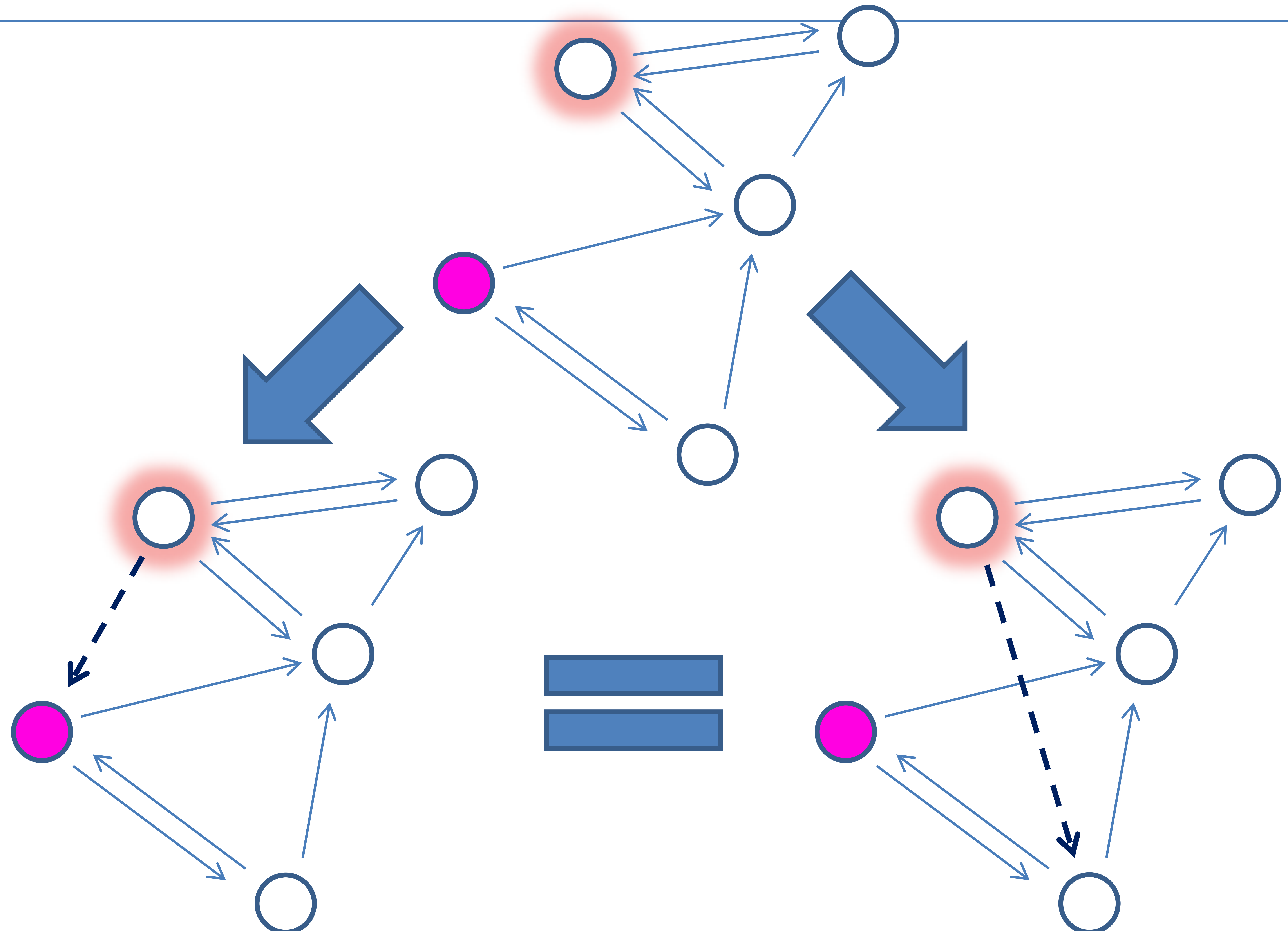


Actor-Oriented

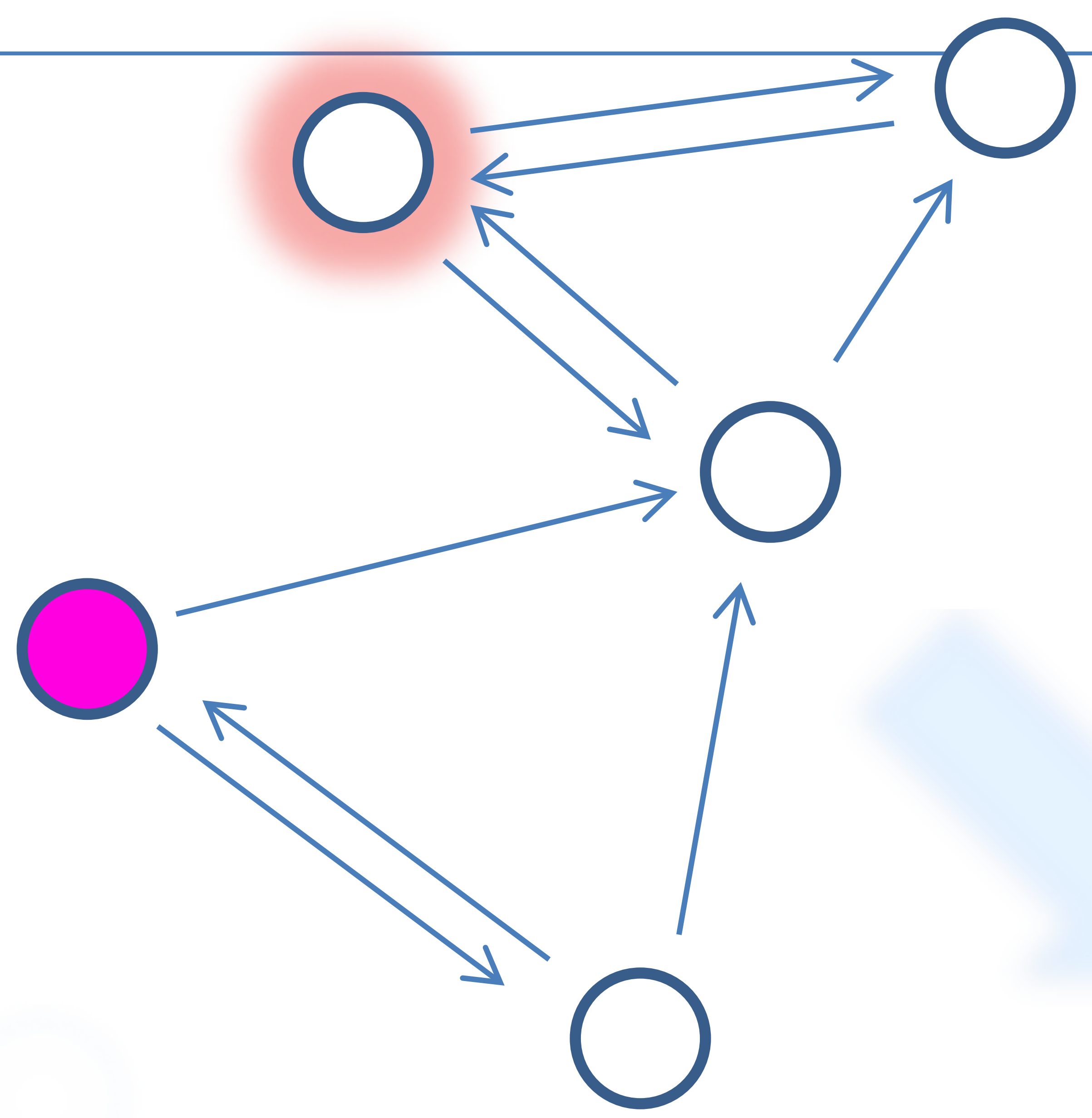


The glowing actor DECIDES what tie change is most appealing.

Markov Assumption

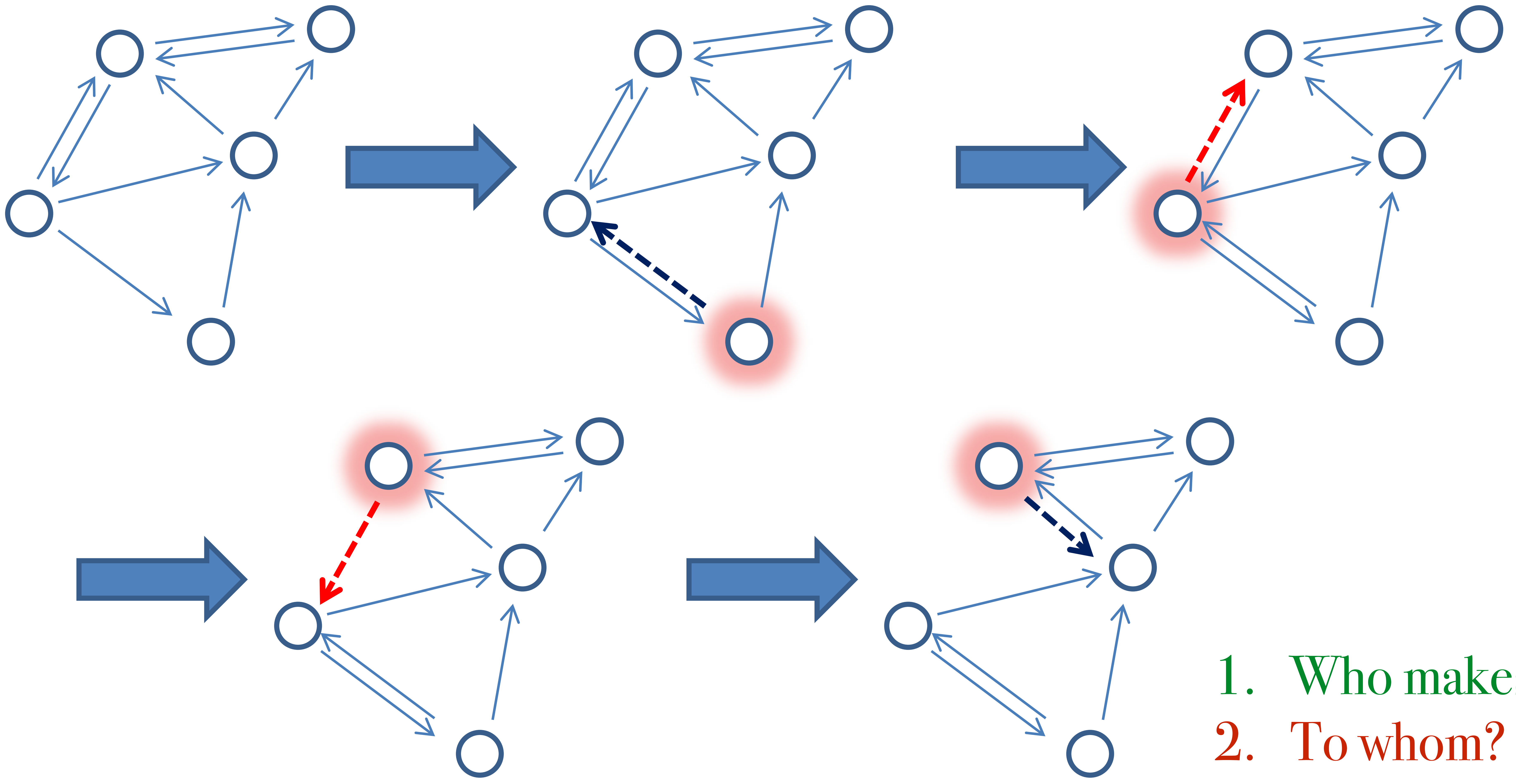


Markov Assumption



The glowing actor does not remember the betrayal by the pink actor

Two Processes in Each Ministep



1. Who makes changes?
2. To whom?

the Secret Sauce

F1 F2

The Two Functions

- Who gets a choice?
  - This is the first part of the ministep
  - A person (*ego* or the *focal actor*) is chosen to consider a change
- Who/what do they choose?
  - Once an ego is chosen, we model which change she makes from her point of view
  - In the case of a network tie, the candidates are people (*alters*)

Rate Function

$$\lambda_i(x) = \exp\left(\sum_k \rho_k r_{ik}(x)\right)$$

Evaluation Function

$$f_i(x) = \sum_k \beta_k s_{ik}(x)$$

The Rate Function

$$\lambda_i(x) = \exp\left(\sum_k \rho_k r_{ik}(x)\right)$$

- Models how much change there is between t_1 and t_2
  - Higher rates mean more change
    - More ministeps necessary to provide actors with more opportunities to make more changes
  - This can mean more ministeps than changes
    - Some actors, when given opportunity to make a tie change, may decide they are actually satisfied
    - Some actors may revert earlier tie changes once local neighbourhood changes as a result of others' choices

The Rate Function

$$\lambda_i(x) = \exp\left(\sum_k \rho_k r_{ik}(x)\right)$$

- Models how many opportunities each actor receives in a time period (between waves)
  - Statistics $r_{ik}(x)$ of i's neighbourhood in x are weighted by parameters ρ_k
    - These weights express whether actors in those configurations correlate with more ($\rho_k > 0$) or less ($\rho_k < 0$) change
- ((Technically, $\lambda_i(x)$ is part of a (non-homogenous) Poisson process))
- Current studies typically assume a periodwise constant rate

But see Hollway 2020

The Evaluation Function

$$f_i(x) = \sum_k \beta_k s_{ik}(x)$$

- Models attractiveness of different network states x to actor i reachable within one step of the current network
  - Statistics $s_{ik}(x)$ of i's neighbourhood in x are weighted by parameters β_k
    - These weights express whether such configurations are desired ($\beta_k>0$) or avoided ($\beta_k<0$)

The Evaluation Function

$$f_i(x) = \sum_k \beta_k s_{ik}(x)$$

- Models actors' choices
  - A value is calculated for each potential alter
  - The model: The alter that increases the evaluation function most is chosen
  - The estimation: Ties must have increased an evaluation function
- ((Technically, $f_i(x)$ is part of a multinomial logit model for discrete, probabilistic choice))
- This is where the action is. It helps us answer questions like whether we prefer happy friends or avoid depressed people.

Statistics and Effects

- By finding out how effects are weighted (the parameters), we can answer our research questions
- Each effect ("IV") has an effect statistic which defines it
  - Are the popular popular?
    - Indegree popularity effect:

$$s_i(x) = \sum_j x_{ij} \sum_k x_{kj}$$

  - Are non-depressed people popular?
    - Alter attribute effect:

$$s_i(x) = \sum_j x_{ij} v_j$$

  - Are the depressed choosing to hang out together?
    - Homophily effect:

$$s_i(x) = \sum_j x_{ij} I\{v_i = v_j\}$$

- They can depend on network configurations (i.e. the position of *j* in the network), or attributes (i.e. a characteristic of *j* or whether it is the same as *i*), or both

<u>Myriad</u> Effects

Table 2
Selection of possible effects for modeling network evolution

effect	network statistic	effective transitions in network*	verbal description
1. outdegree	x_{ij}		preference for ties to arbitrary others
2. reciprocity	$x_{ij}x_{ji}$		preference for reciprocated ties
3. transitive triplets	$x_{ij}\sum_h x_{ih}x_{hj}$		preference for being friend of the friends' friends
4. balance	$x_{ij}\,\mathrm{strsim}_{ij}$		preference for ties to structurally similar others
5. actors at distance two	$\begin{cases}1 & \text{if between}(h;ij)=1 \text{ for some } h\\ 0 & \text{else}\end{cases}$	(the *number* of intermediaries is irrelevant)	preference for keeping others at social distance two
6. popularity alter	$x_{ij}\sum_h x_{hj}$		preference for attaching to popular others, i.e., others who are often named as friend ('preferential attachment')
7. activity alter	$x_{ij}\sum_h x_{jh}$		preference for attaching to active others, i.e., others who name many friends
8. 3-cycles	$x_{ij}\sum_h x_{jh}x_{hi}$		preference for forming relationship cycles (negative indicator for hierarchical relations)
9. betweenness	$\sum_h \mathrm{between}(i;hj)$	(*no direct link* from the left to the right actor)	preference for being be in an intermediary position between unrelated others
10. dense triads	$\sum_h \mathrm{group}(ijh)$		preference for being part of cohesive subgroups
11. peripheral	$\sum_{hk} \mathrm{peripheral}(i;jhk)$		preference for unilaterally attaching to cohesive subgroups
12. similarity	$x_{ij}\,\mathrm{sim}_{ij}$		preference for ties to similar others (selection)
13. behavior alter	$x_{ij}z_j$		main effect of alter's behavior on tie preference
14. behavior ego	$x_{ij}z_i$		main effect of ego's behavior on tie preference
15. similarity × reciprocity	$x_{ij}x_{ji}\,\mathrm{sim}_{ij}$		preference for reciprocated ties to similar others
16. between dis-similar alters	$\sum_h (1-\mathrm{sim}_{jh})\mathrm{between}(i;jh)$		preference for being in an intermediary position between unrelated, dissimilar others (brokerage potential)
17. similarity × dense triads	$\sum_h \mathrm{group}(ijh)\left(\mathrm{sim}_{ij}+\mathrm{sim}_{ih}\right)$		preference for being part of behaviorally similar cohesive subgroups
18. behavior × peripheral	$z_i\sum_{hk} \mathrm{peripheral}(i;jhk)$		behavior-specific preference for unilaterally attaching to cohesive subgroups
19. similarity × peripheral	$\sum_{hk}\left(\mathrm{peripheral}(i;jhk)\times(\mathrm{sim}_{ij}+\mathrm{sim}_{ih}+\mathrm{sim}_{ik})\right)$		preference for unilaterally attaching to behaviorally similar cohesive subgroups

* In the *effective transitions* illustrations, it is assumed that the behavioral dependent variable is dichotomous and centered at zero; the color coding is ○ = low score (negative), ● =high score (positive), ○ =arbitrary score. The tie x_{ij} from actor i to actor j is the one that changes in the transition indicated by the double arrow. Illustrations are not exhaustive.

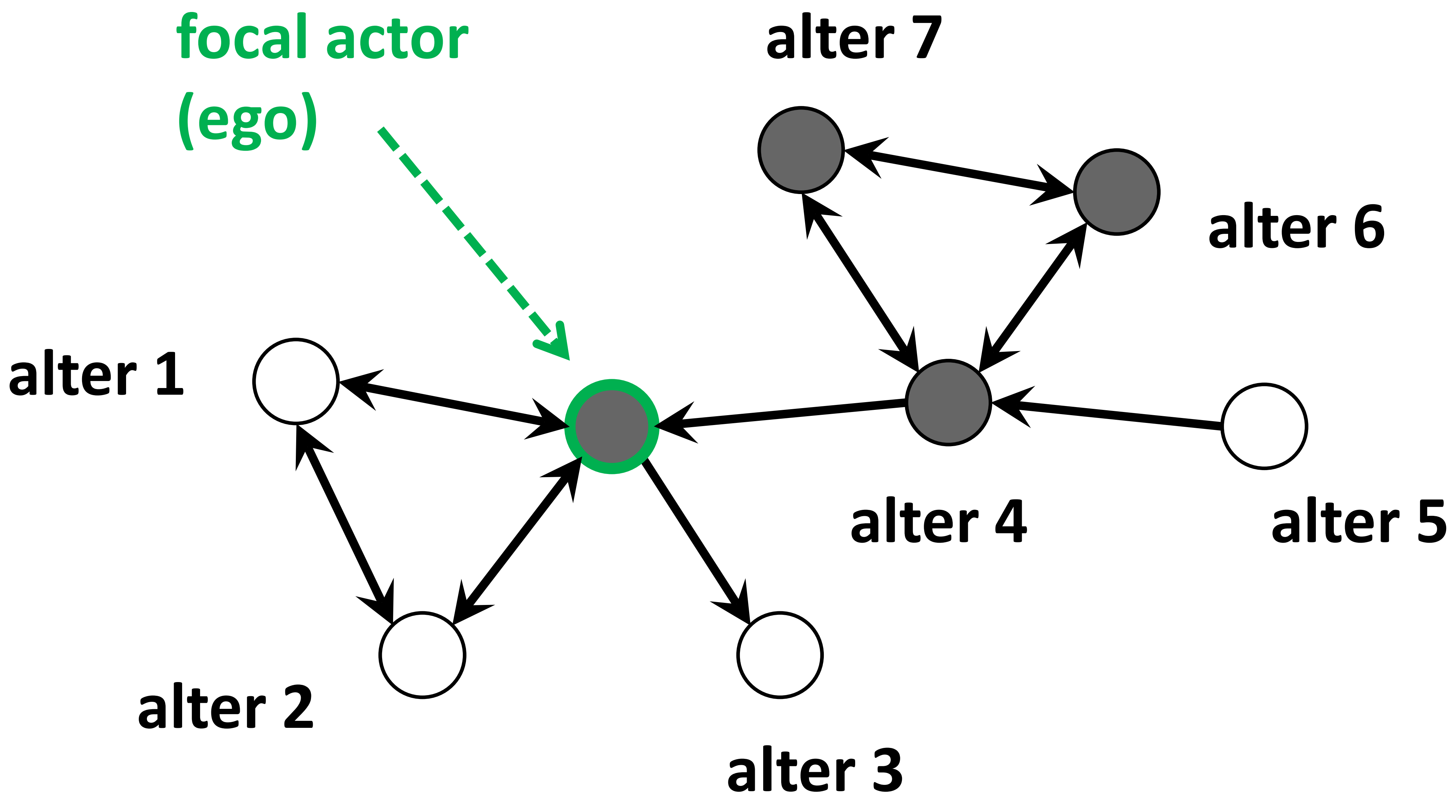
Covariates

- Some effects rely on exogenous information
- There are four types:

Covariates	Monadic	Dyadic
Constant	coCovar	coDyadCovar
Changing	varCovar	varDyadCovar

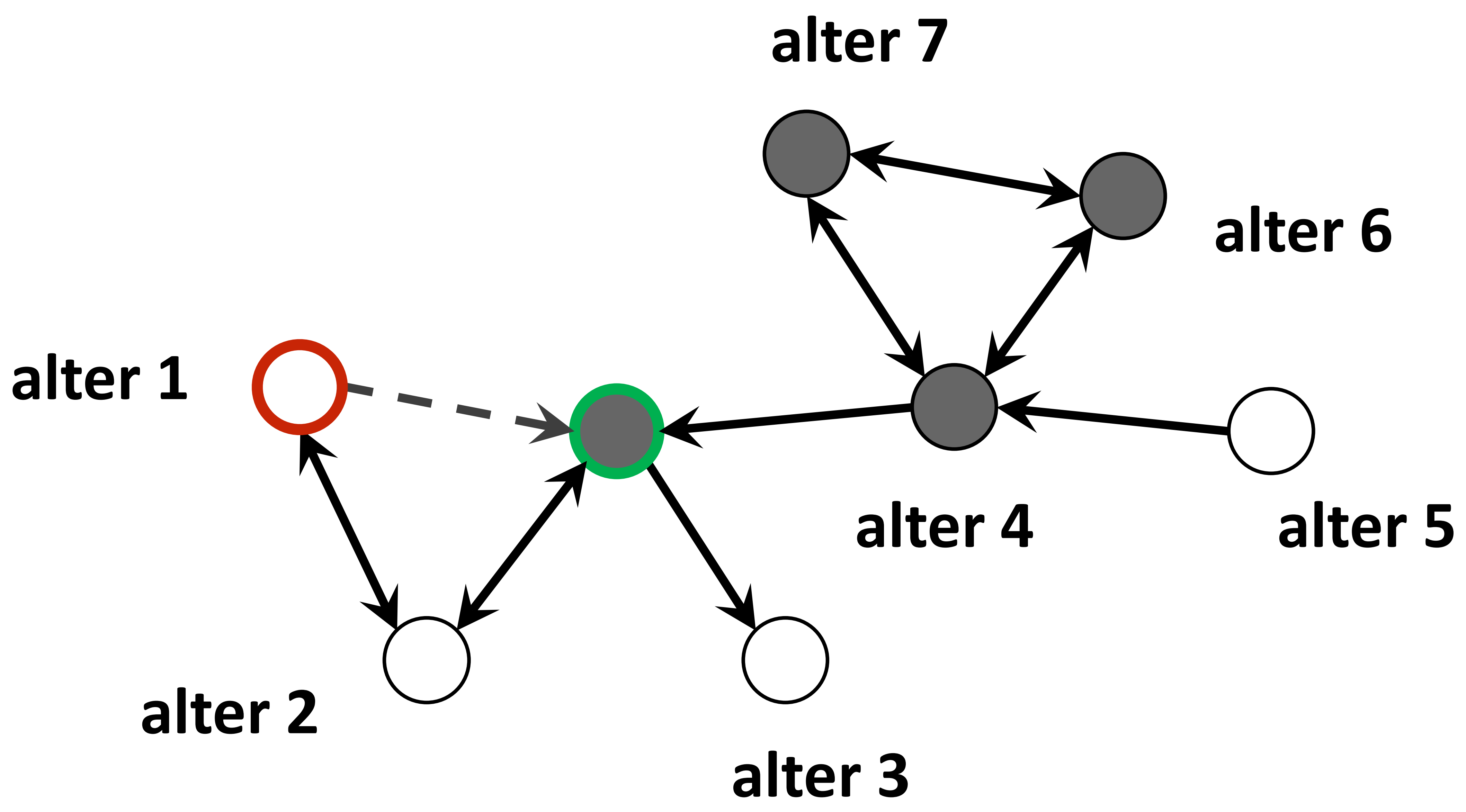
- For each type, multiple effects can be specified

Example of an actor's decision



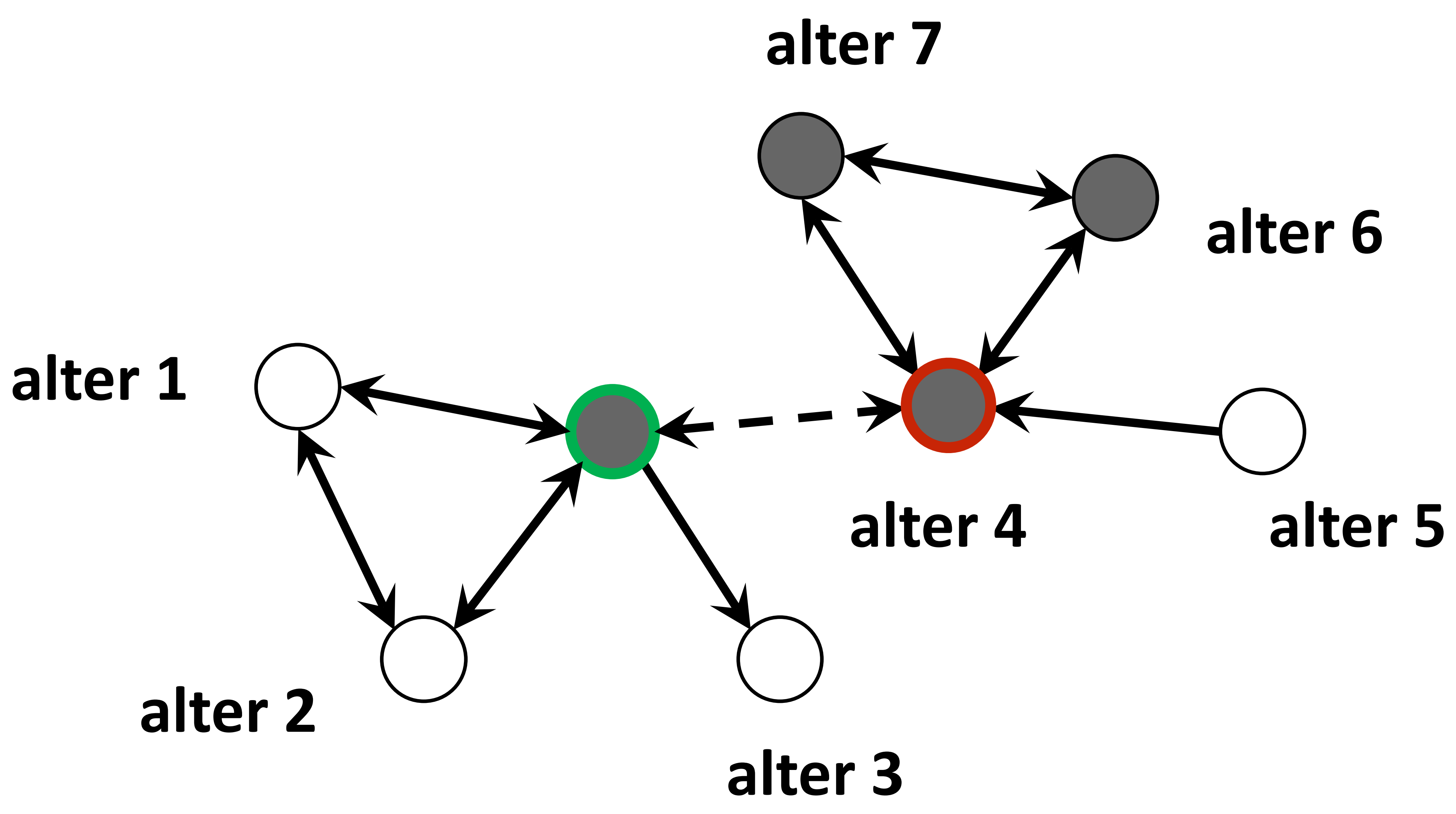
- Options
  - drop tie to 1
  - drop tie to 2
  - drop tie to 3
  - create tie to 4
  - create tie to 5
  - create tie to 6
  - create tie to 7
  - keep status quo

Statistics for dropping tie to 1



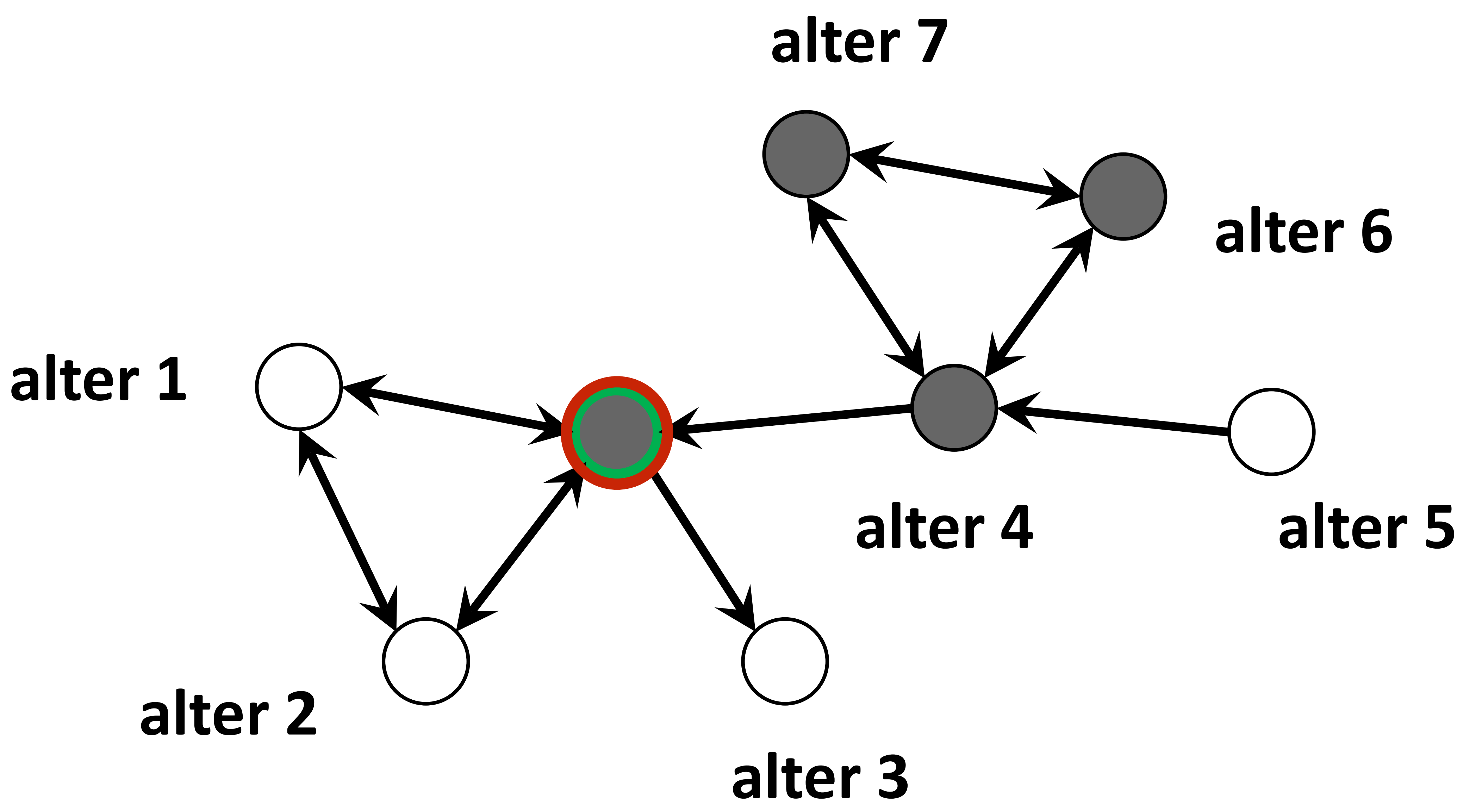
- 2 outgoing ties
- 1 reciprocated tie
- 0 transitive triplets
- 1 three-cycle
- 0 same colour

Statistics for creating tie to 4



- 4 outgoing ties
- 3 reciprocated tie
- 2 transitive triplets
- 2 three-cycles
- 1 same colour

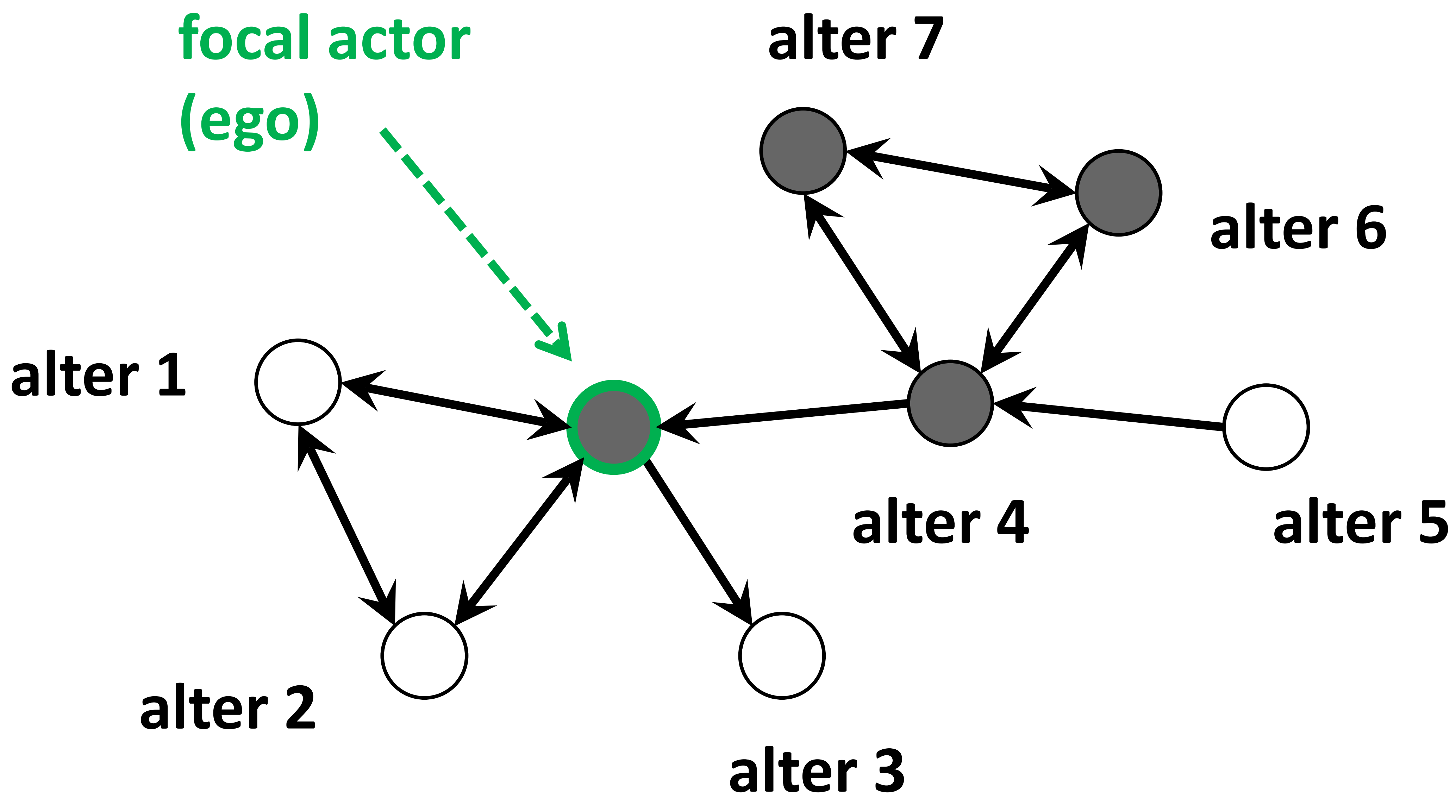
Statistics for status quo



These calculations are done for all possible choices

- 3 outgoing ties
- 2 reciprocated tie
- 2 transitive triplets
- 2 three-cycles
- 0 same colour

Statistics for all options



	#degree	#mutual	#trans	#3cycles	#same col.
Drop 1	2	1	0	1	0
Drop 2	2	1	0	1	0
Drop 3	2	2	2	2	0
Create 4	4	3	2	2	1
Create 5	4	2	2	3	0
Create 6	4	2	2	3	1
Create 7	4	2	2	3	1
Status quo	3	2	2	2	0

Evaluating the options

$$f_i(x) = \sum_k \beta_k s_{ik}(x)$$

	β_{degree}	β_{mutual}	β_{trans}	β_{3cycles}	β_{same}
	-2.6	1.8	0.4	-0.7	0.8

		#degree	**#mutual**	**#trans**	**#3cycles**	**#same col.**
$f_i(\text{drop1})$	**Drop 1**	2	1	0	1	0
$f_i(\text{drop2})$	**Drop 2**	2	1	0	1	0
$f_i(\text{drop3})$	**Drop 3**	2	2	2	2	0
$f_i(\text{create4})$	**Create 4**	4	3	2	2	1
$f_i(\text{create5})$ =	**Create 5**	4	2	2	3	0
$f_i(\text{create6})$	**Create 6**	4	2	2	3	1
$f_i(\text{create7})$	**Create 7**	4	2	2	3	1
$f_i(\text{statusquo})$	**Status quo**	3	2	2	2	0

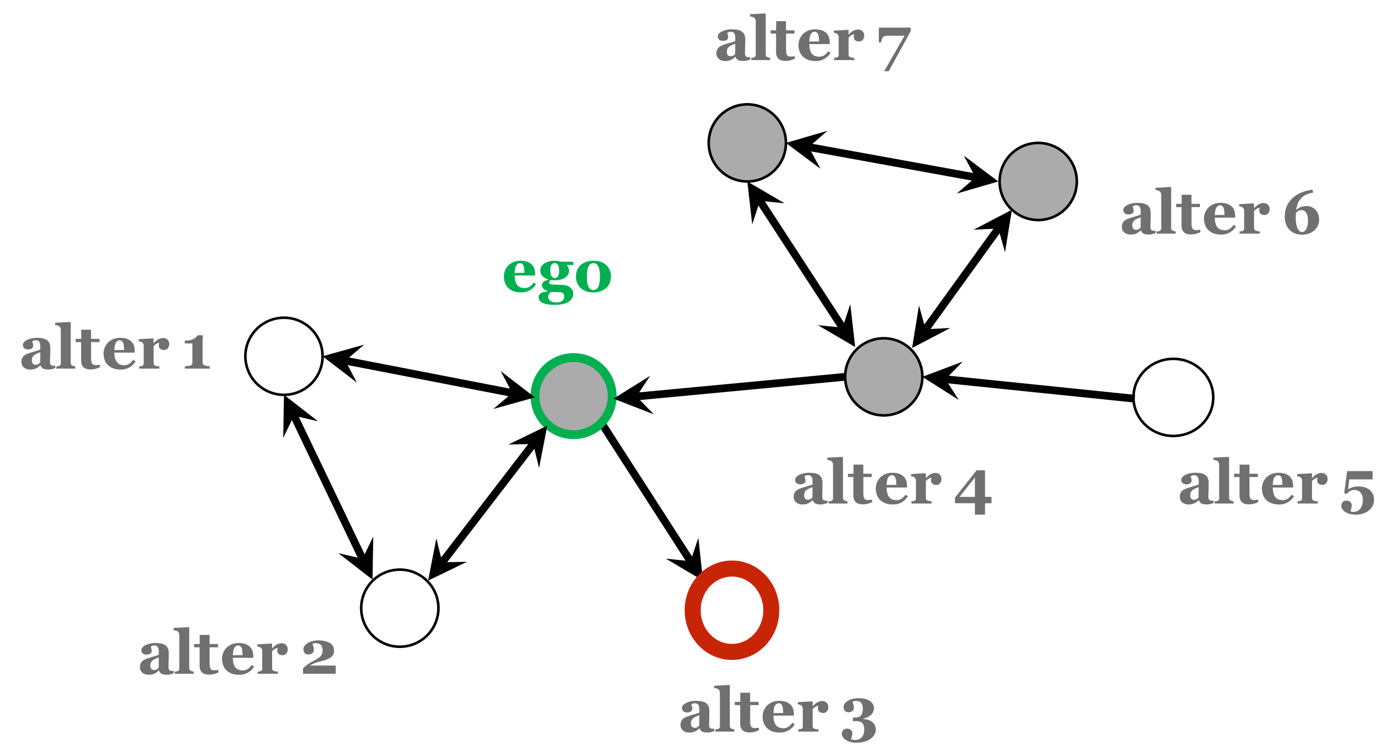
Transforming to probabilities

Using underlying multinomial:

$$p_{i \rightsquigarrow j}(x, \beta) = \frac{exp\left(f\left(x^{i \rightsquigarrow j}, \beta\right)\right)}{\sum_{k=1}^{n} exp\left(f\left(x^{i \rightsquigarrow j}, \beta\right)\right)}$$

	Evaluation	Exponent.	Prob.
Drop 1	-4.1	0.017	10%
Drop 2	-4.1	0.017	10%
Drop 3	-2.2	0.111	68%
Create 4	-4.8	0.008	5%
Create 5	-8.1	0.000	0%
Create 6	-7.3	0.001	1%
Create 7	-7.3	0.001	1%
Status quo	-4.8	0.008	5%

Dropping tie to alter 3 is the most likely choice for ego

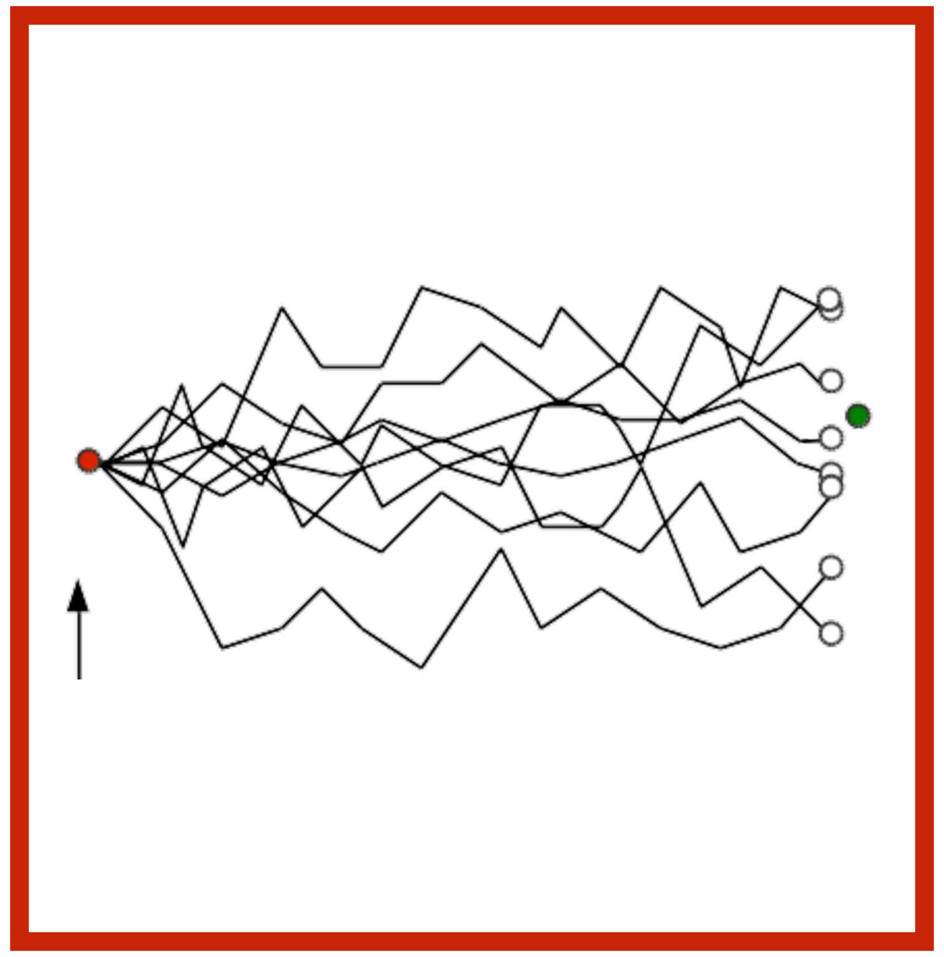


SAOM

Model

Estimation

Influence



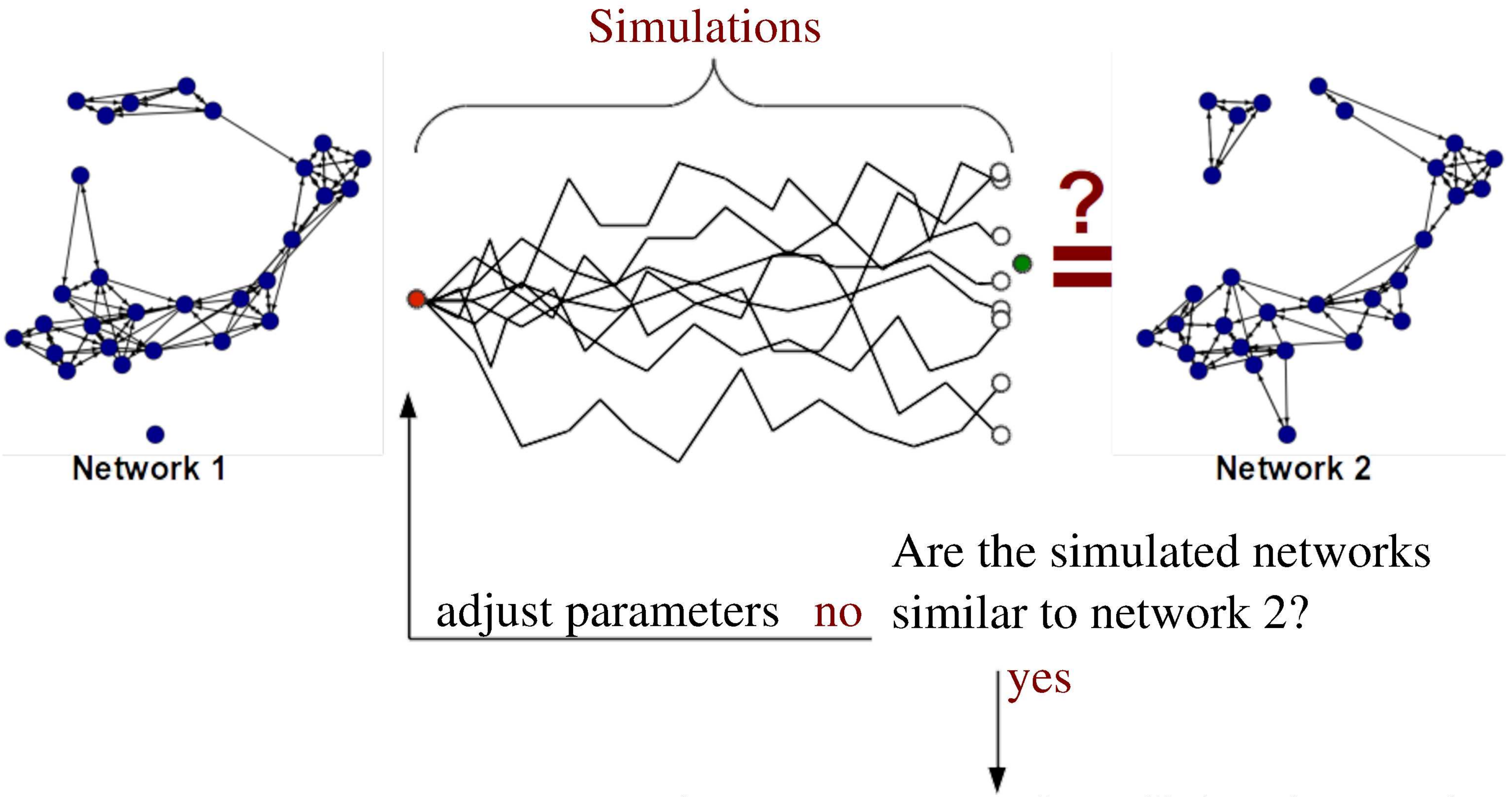
Estimation

- So we now have a well-defined probability model, from which we can simulate networks using defined parameters (β)

- But what we usually want to do is *estimate* parameters from *observed* data!

- We do this using RSiena ("SIENA" = Simulation Investigation for Empirical Network Analysis)

SIENA estimates SAOMs through simulations



The parameters are "good" descriptors of the social processes shaping network 2

Three Estimation Methods

- **Method of Moments (MoM)**
  - Take the network at the first time point and simulate a certain number of mini-steps with some initial β values
  - Compare the simulated networks to the observed network at the second time point
  - According to the differences between observed and simulated networks, update β values
  - Rinse and repeat until the simulated networks "closely" resemble the observed one
- Maximum Likelihood (ML)
  - Actually connects two observations by chains of ministeps and estimates parameters from these chains
- Bayesian (Bayes)
  - For multilevel analysis of networks and enthusiasts

Estimation Results

- While the model is more complicated, RSiena spits out a table at the end, the second part of which can be interpreted like that of a multinomial regression

- Each parameter estimate has a standard error

- If the t-ratio ($= \beta / se$) ≥ 2, then we can say that we can reject the null hypothesis of there being no effect

		Model 1	Model 3
Rate function friendship			
Rate of change $t_1 \rightarrow t_2$		7,54 (0,97)	10,87 (2,63)
Rate of change $t_2 \rightarrow t_3$		2,73 (0,45)	3,04 (0,52)
Rate of change $t_3 \rightarrow t_4$		3,29 (0,49)	3,80 (0,65)
Objective function friendship			
Outdegree		-1,92 (0,17) ***	-2,19 (0,16) ***
Reciprocity		—	0,84 (0,17) ***
Transitive triplets		—	0,18 (0,03) ***
primary school friendship		0,54 (0,21) *	0,40 (0,20) *
Male alter		0,30 (0,18)	0,05 (0,17)
Male ego	strongly	0,11 (0,19)	-0,17 (0,18)
Same sex	biased	1,70 (0,18) ***	0,93 (0,18) ***

Model Specification

- Researchers usually come with *theory* or at least *hypotheses*
  - SAOMs are not for exploration
- Beware spuriousness…
  - Attribute vs centrality (popularity)
  - Homophily vs cohesion (reciprocity, transitivity)

	Model 1		Model 3	
Rate function friendship				
Rate of change $t_1 \rightarrow t_2$	7,54 (0,97)		10,87 (2,63)	
Rate of change $t_2 \rightarrow t_3$	2,73 (0,45)		3,04 (0,52)	
Rate of change $t_3 \rightarrow t_4$	3,29 (0,49)		3,80 (0,65)	
Objective function friendship				
Outdegree	-1,92 (0,17)	***	-2,19 (0,16)	***
Reciprocity	—		0,84 (0,17)	***
Transitive triplets	—		0,18 (0,03)	***
primary school friendship	0,54 (0,21)	*	0,40 (0,20)	*
Male alter	0,30 (0,18)		0,05 (0,17)	
Male ego	0,11 (0,19)		-0,17 (0,18)	
Same sex	1,70 (0,18)	***	0,93 (0,18)	***

strongly biased

Parameter Interpretation

- Estimated parameters need to be interpreted as *within ministeps* and *against other choices*

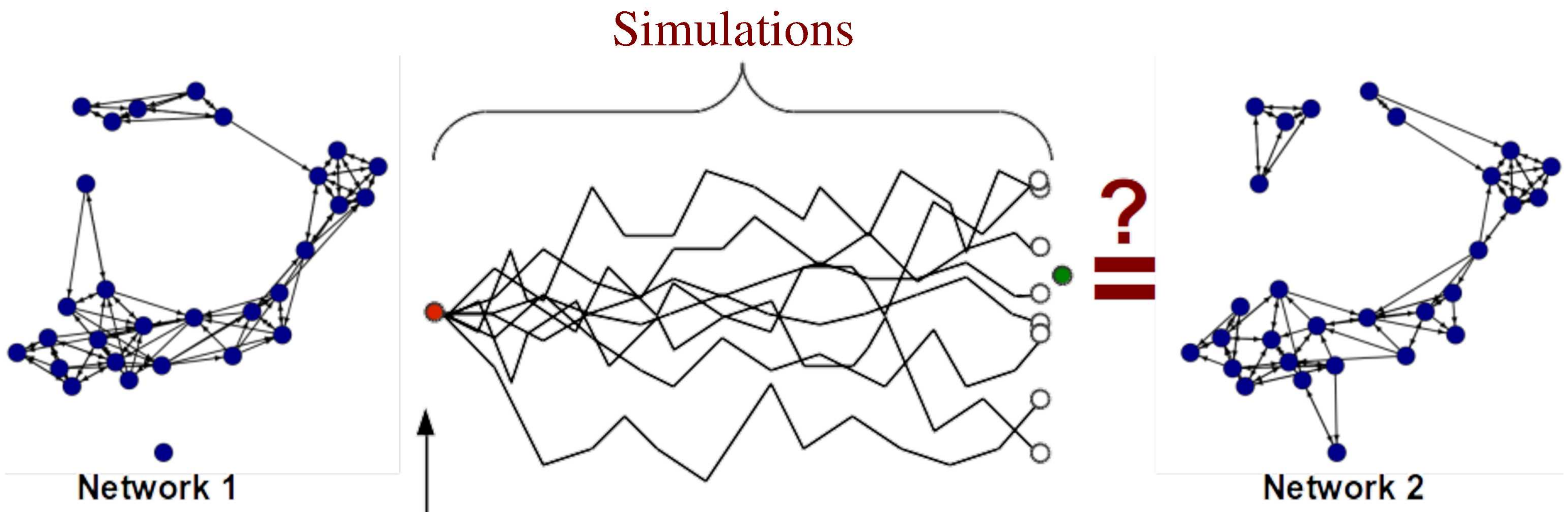
- So we interpret the parameters as: when a chosen ego i is faced with a decision to form a tie to either of two alters, j_1 or j_2, that differ only on one statistic value, then the odds ratio is as follows:

$$\frac{p_{i \rightsquigarrow j_1}}{p_{i \rightsquigarrow j_2}} = \frac{exp\left(f(x^{i \rightsquigarrow j_1}, \beta)\right)}{exp\left(f(x^{i \rightsquigarrow j_2}, \beta)\right)} = \frac{\exp(\beta s_{j_1})}{\exp(\beta s_{j_2})}$$

- So, say i can send a tie to j_1 or j_2, which only differ in that j_1 sends a tie to i and j_2 does not, then given a reciprocity parameter of 2, $\dfrac{\exp(2 \times 1)}{\exp(2 \times 0)} = 7.39$
    - i is 7.39 times more likely to send a tie to j_1 than j_2

Diagnostics

But what does “good” mean?



Are the simulated networks similar to network 2?

no — adjust parameters

yes

The parameters are "good" descriptors of the social processes shaping network 2

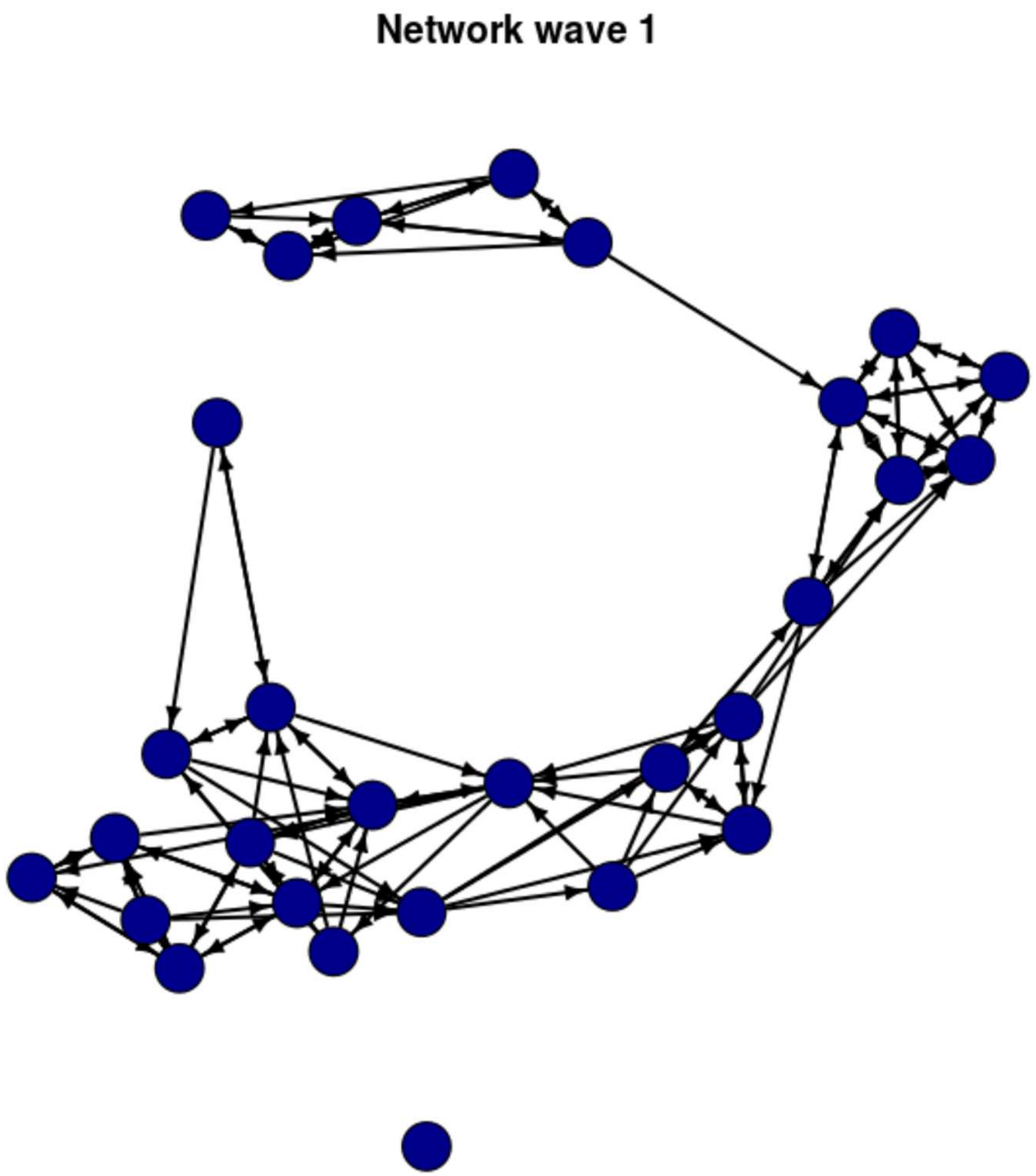
Target statistics Z are listed in the SIENA output file

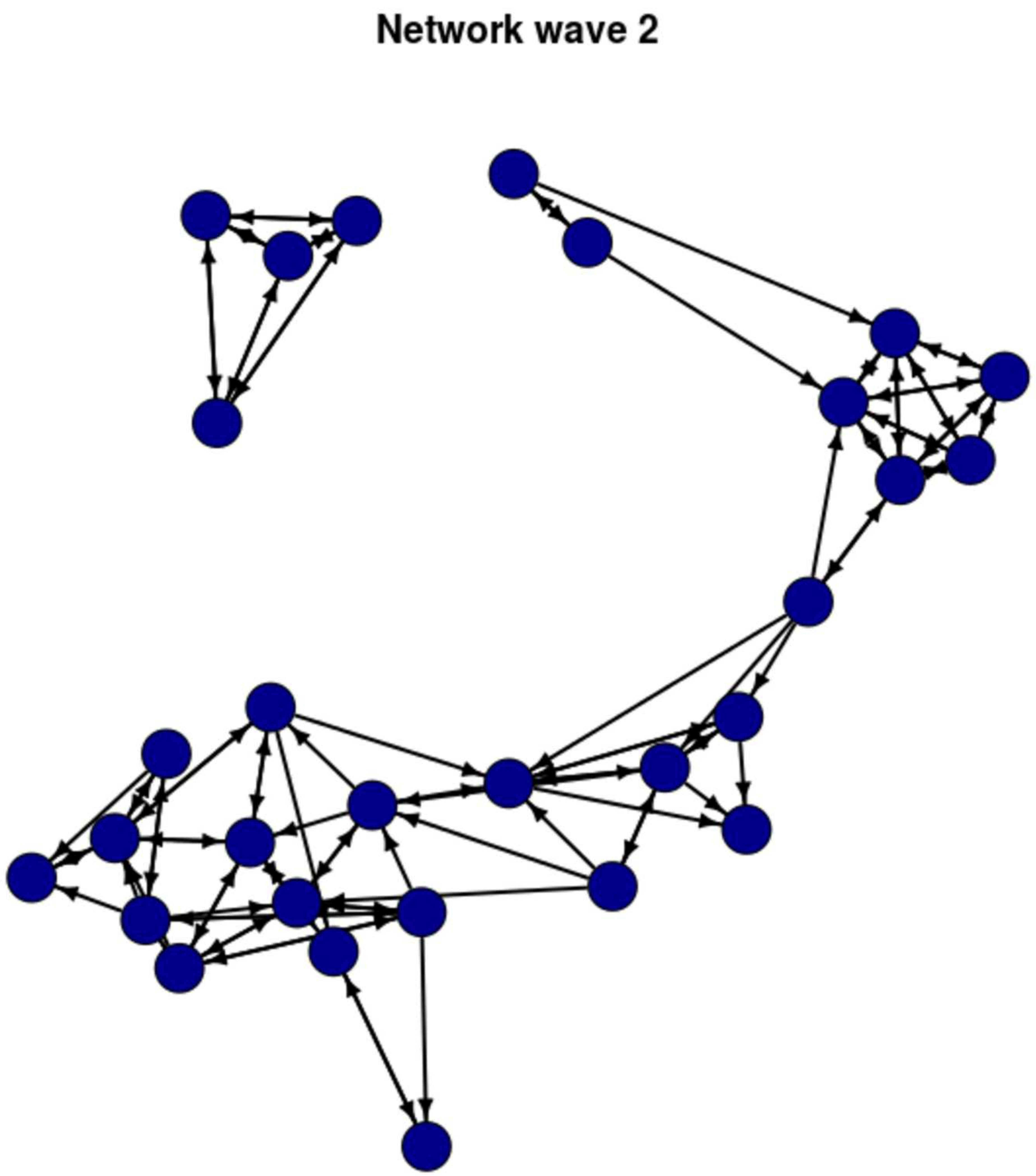
```
Observed values of target statistics are
  1. Number of ties                                      99.0000
  2. Number of reciprocated ties                         72.0000
  3. Number of transitive triplets                      164.0000
  4. 3-cycles                                            47.0000
  5. Sum of squared indegrees                           403.0000
  6. Same values on coo.coCovar                          47.0000
  7. Sum of indegrees x gender.coCovar                   -5.0345
  8. Sum of outdegrees x gender.coCovar                  -4.0345
  9. Same values on gender.coCovar                       90.0000
```

- MoM aims at creating networks that have statistics close to the ones above
  - More formally, parameters $\theta = \{\varrho, \beta\}$ that generate networks for which $E_\theta = \{Z\}$ and are stable have converged

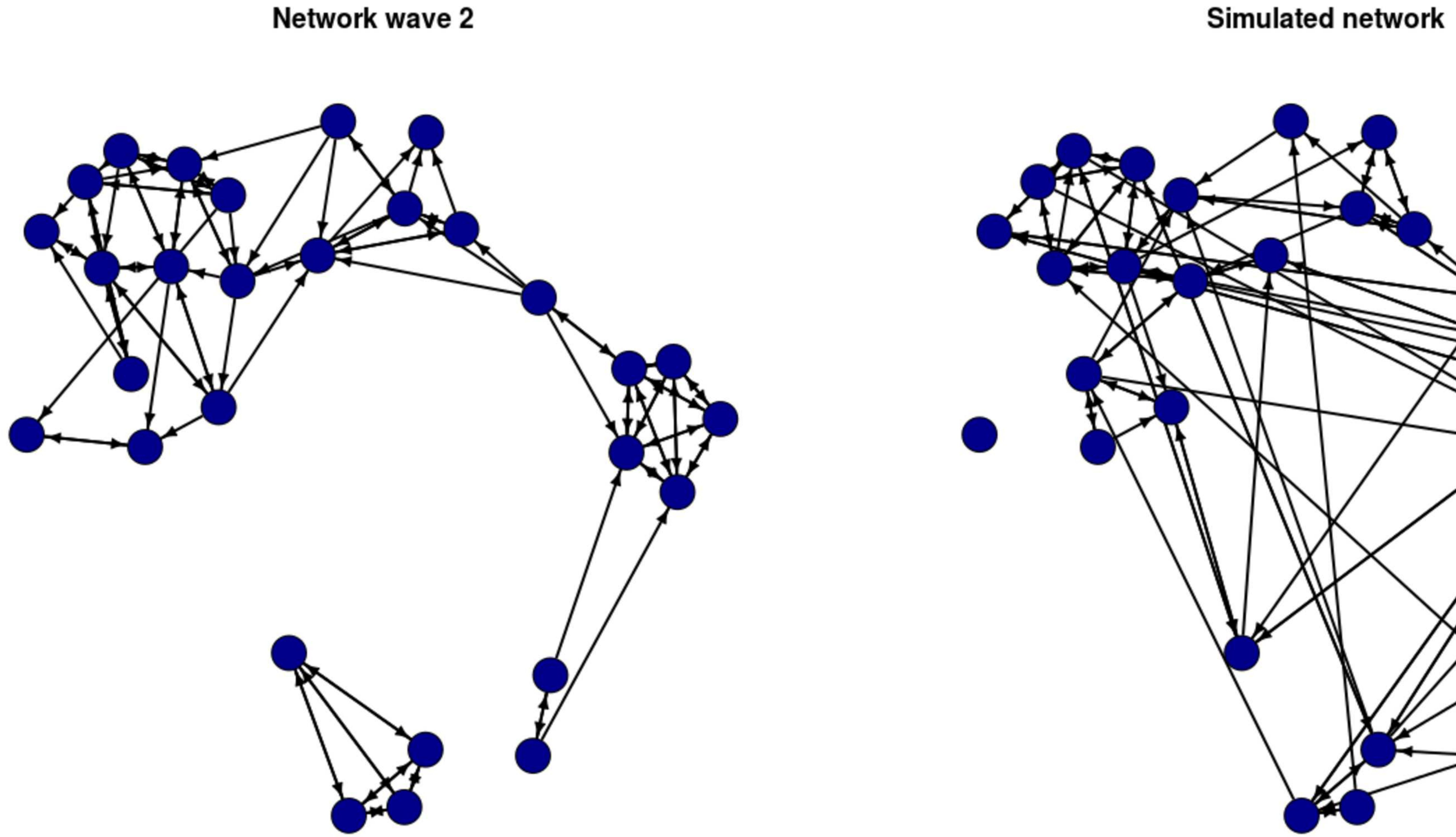
- But do these simulated networks resemble *other, non-modelled macro features* of the network such as the degree distribution, the triad census, etc? (i.e. goodness of fit)

So, which forces shape this social network's evolution?



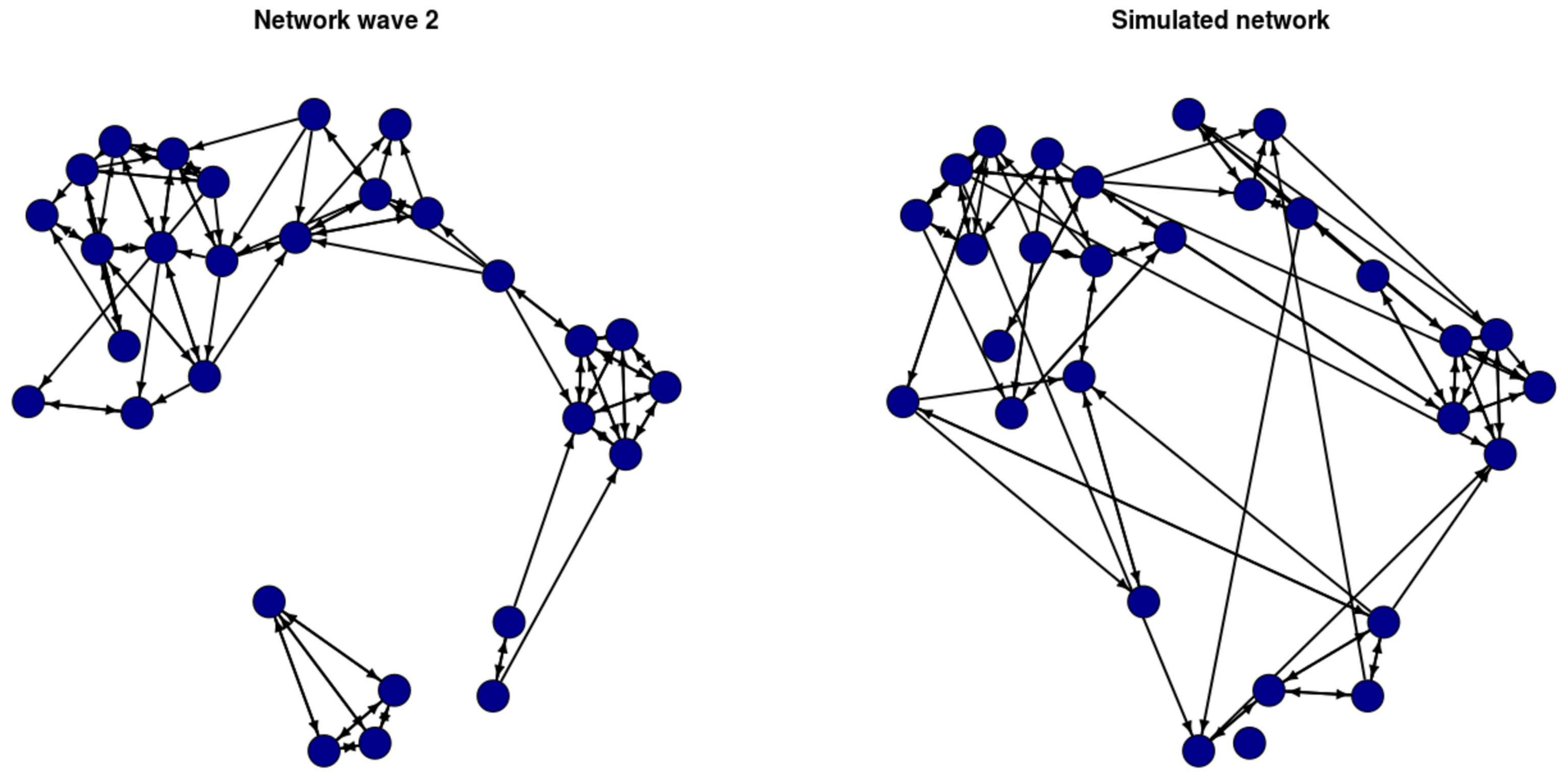


Degree + Reciprocity



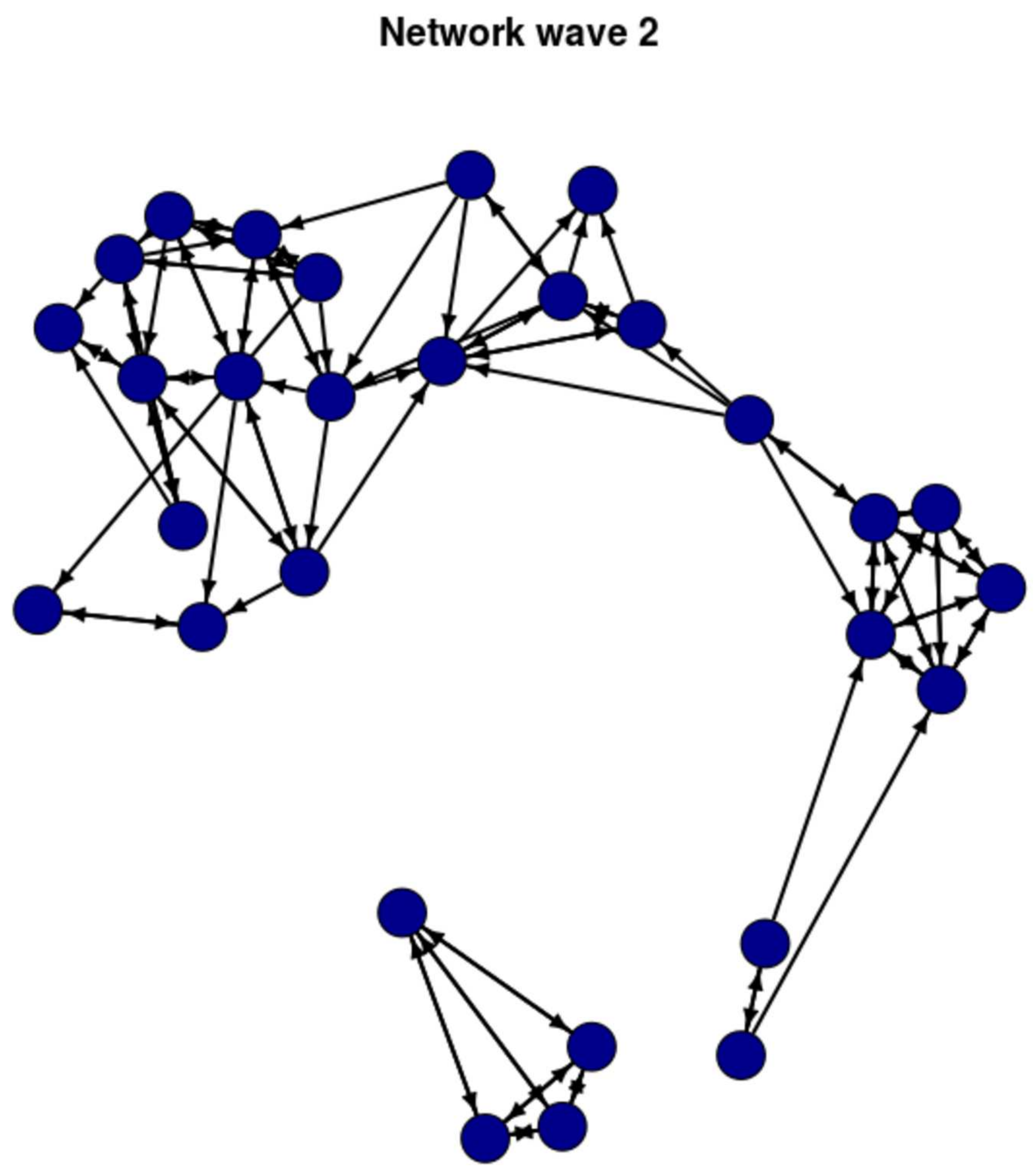
- While the model has converged and the two parameters are highly significant, the model does not represent groups very well…

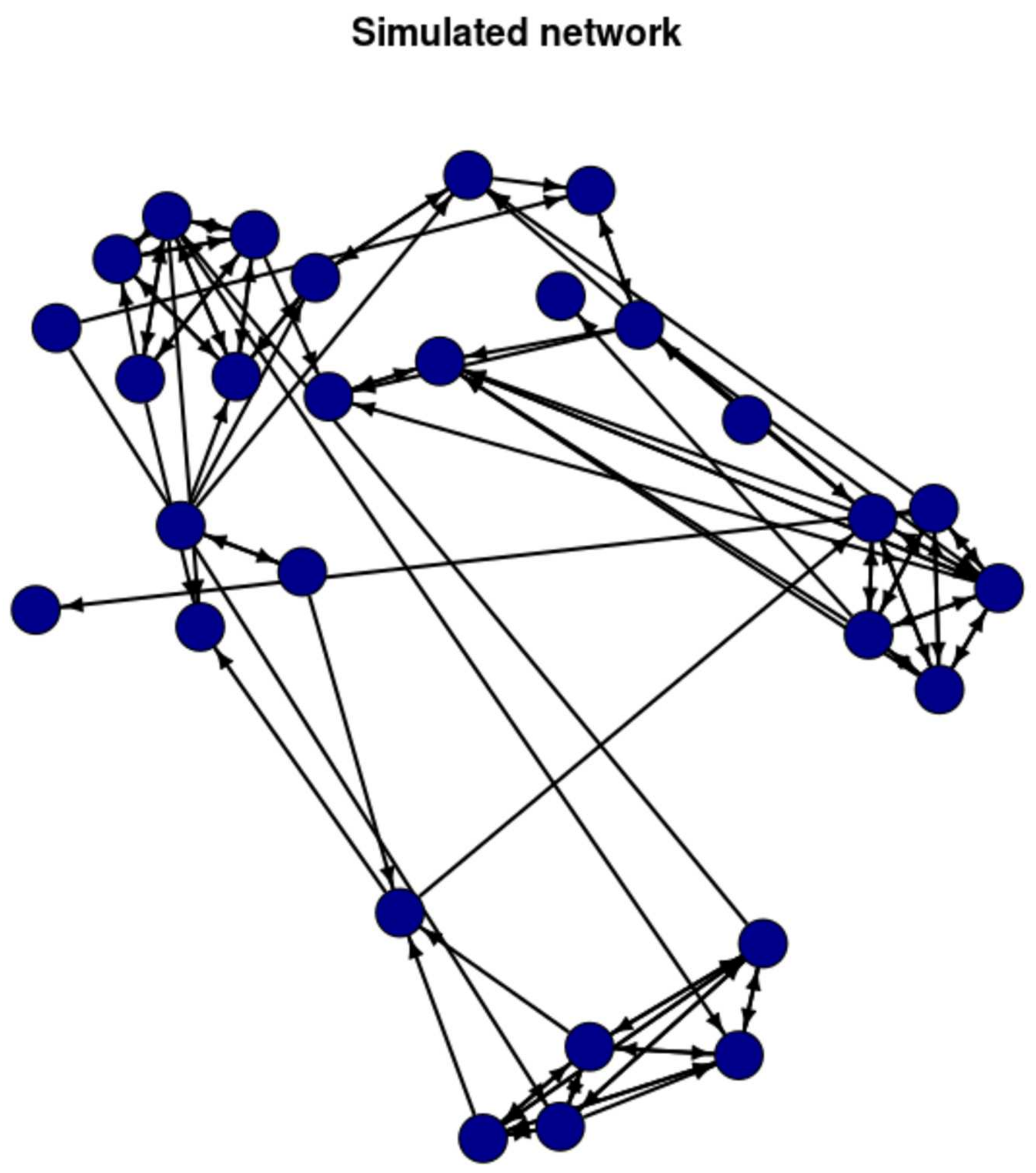
+ Transitivity and 3-Cycles



- Group boundaries are clearer but there are still too many connections between groups

+ Gender Homophily



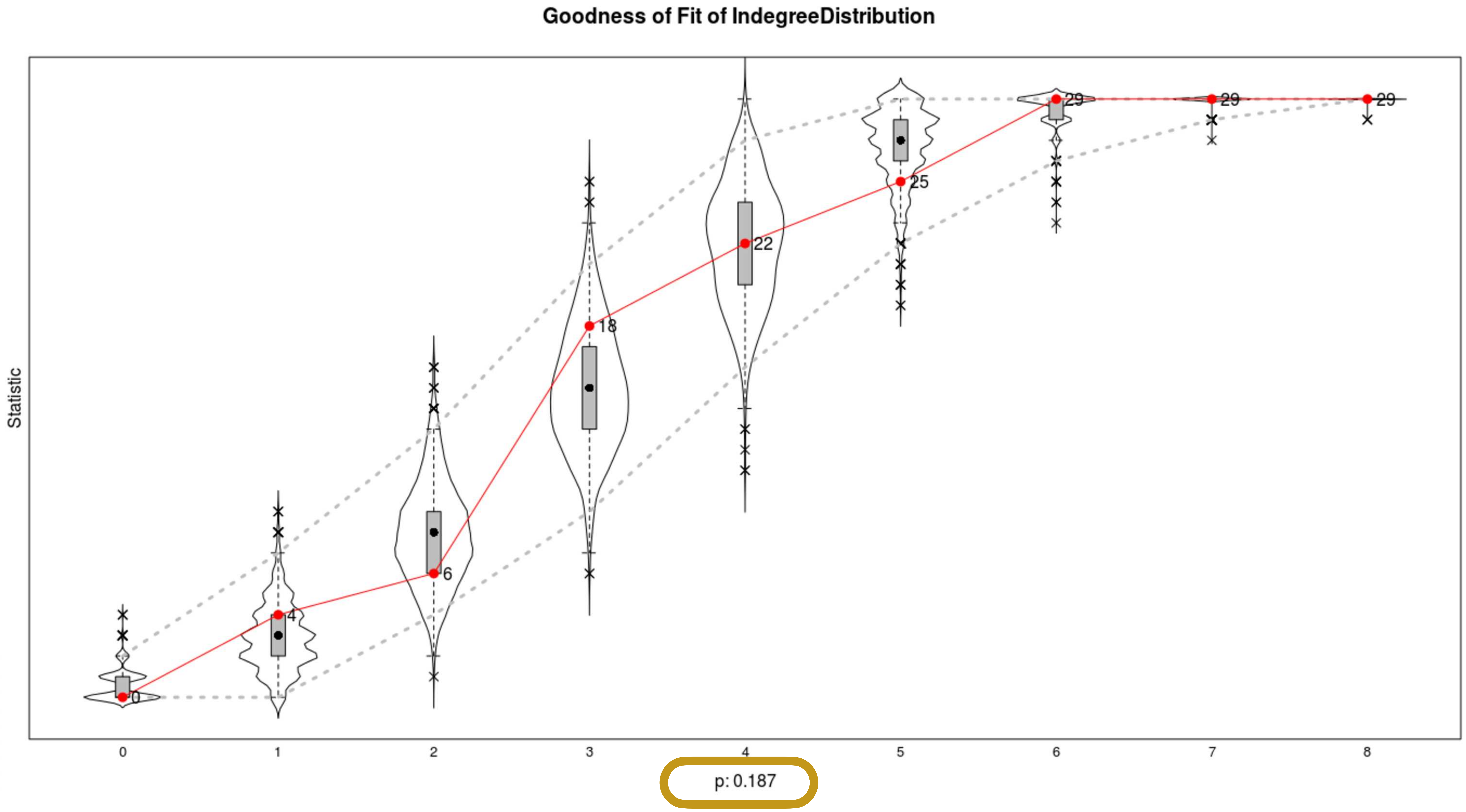


- Fewer links between groups of different gender but still many between-group ties within a gender
- One could try further structural and attribute-related effects

sienaGOF() does this comparison more systematically

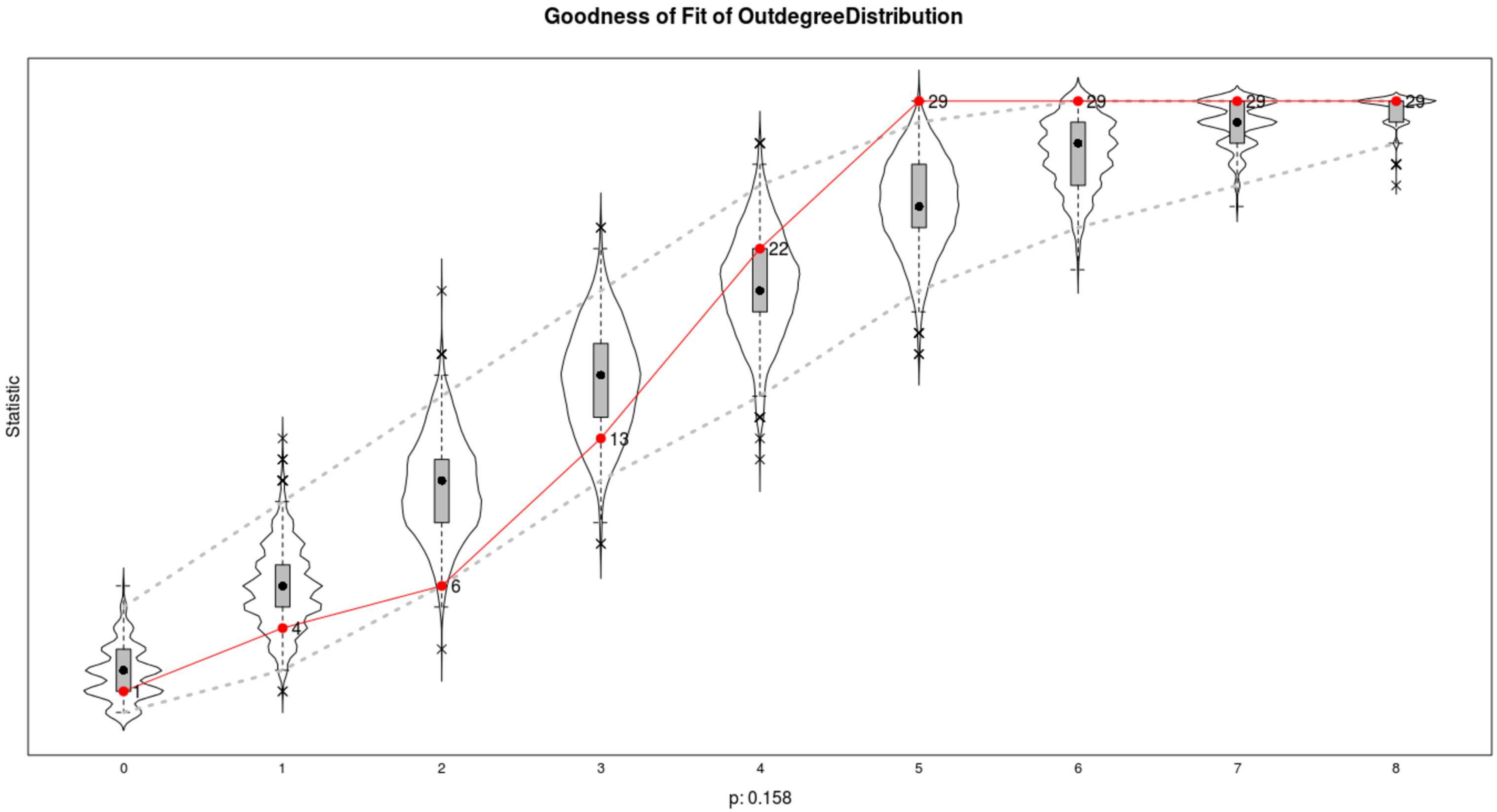
- sienaGOF() tests particular macro features of the simulated social networks and compares them to the empirically observed networks
  - Degree distribution
  - Geodesic distances
  - Triad census
- sienaGOF() takes all simulated networks into account, as opposed to the visual inspection, where we only looked at one

Indegree GOF

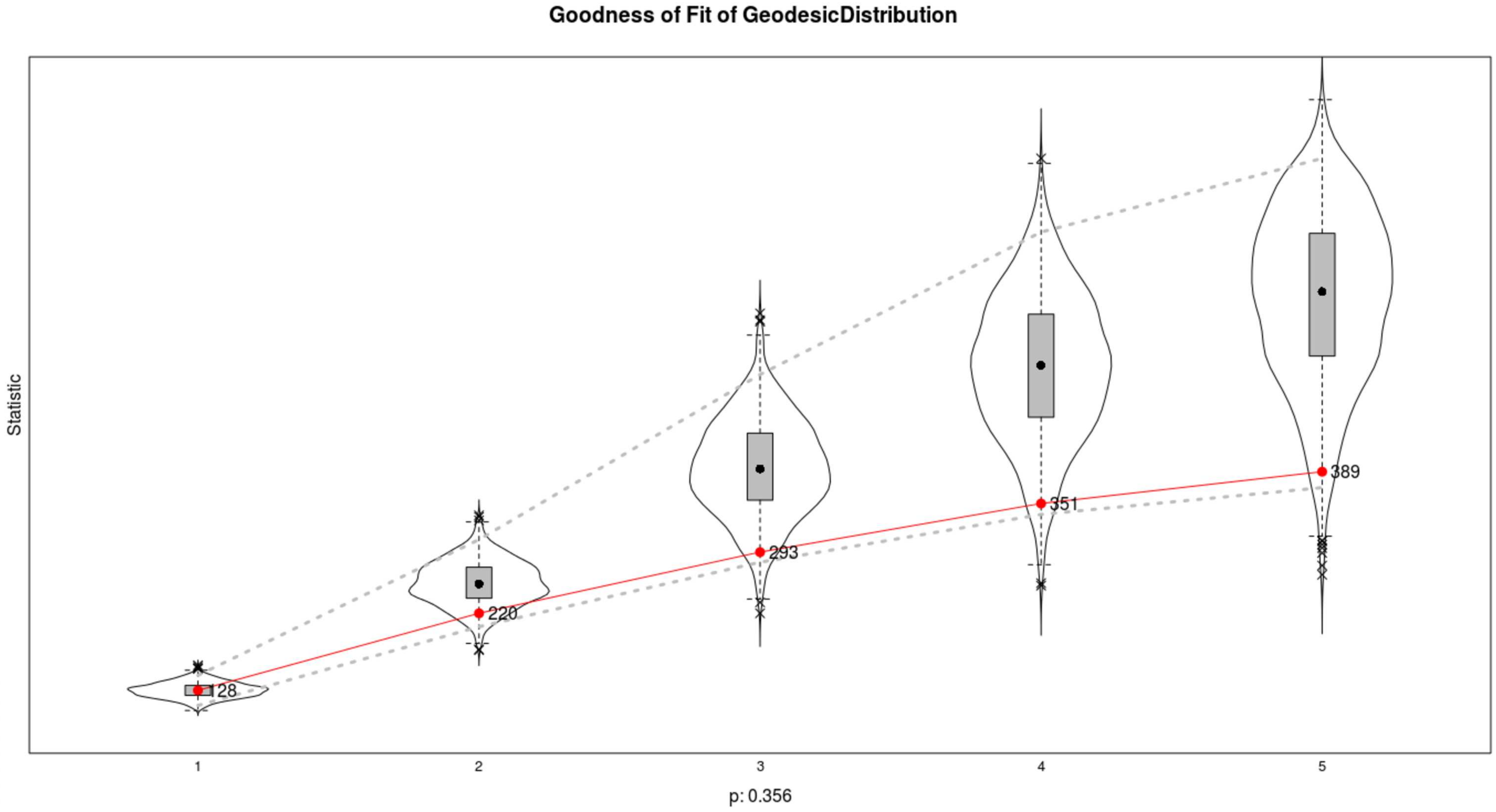


p-value *over* .05 suggests reasonable fit

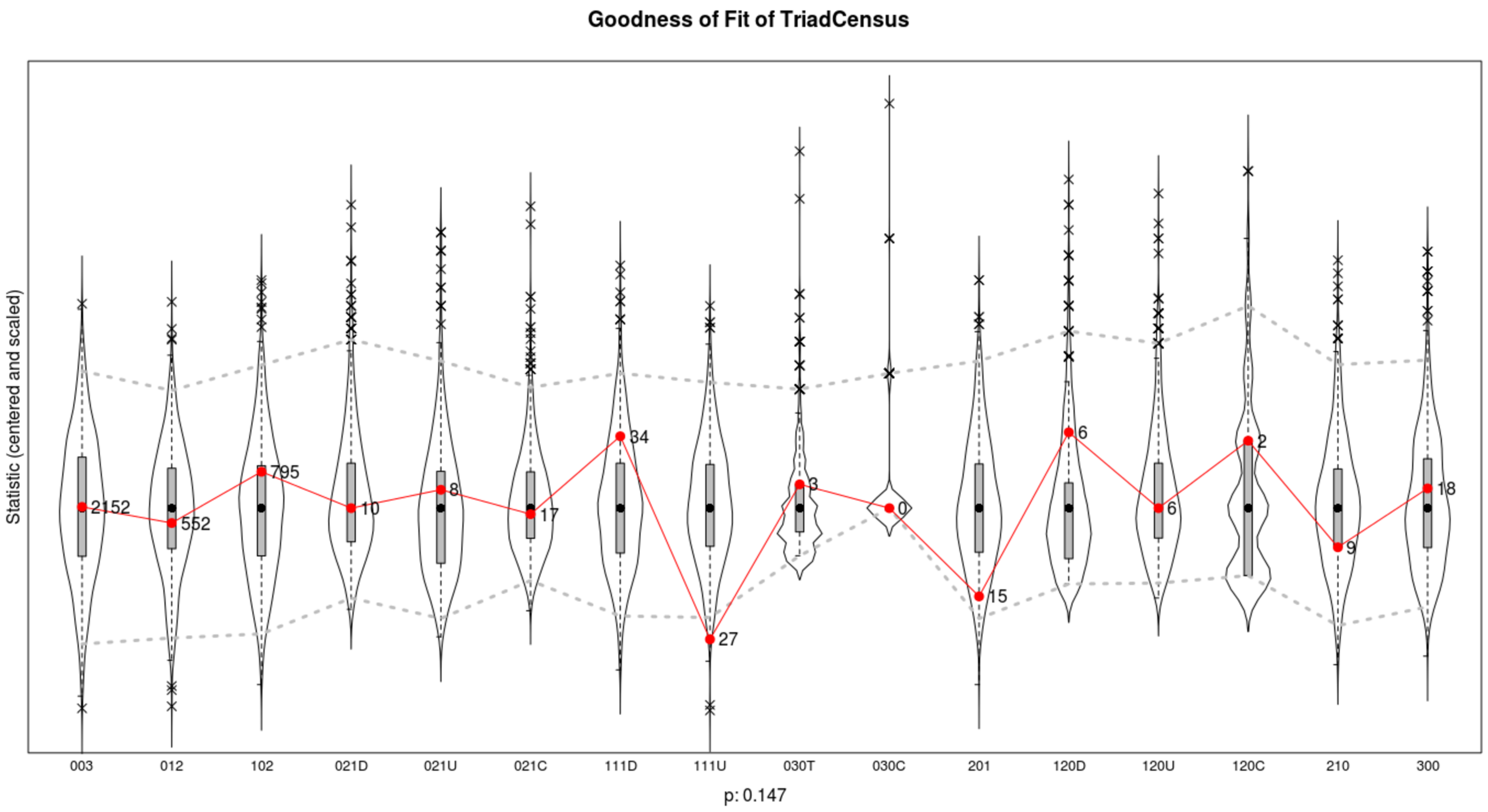
Outdegree GOF



Geodesic GOF



Triad census GOF



Standard Model

Single Binary Directed 1-mode Network Change

Model Extensions

Single	Binary	Directed	1-mode	Network	Change
Multiplex	Ordered	Undirected	2-mode	Behaviour	Creation Deletion

Model Extensions

Single	Binary	Directed	1-mode	Network	Change
Multiplex	Ordered	Undirected	2-mode	Behaviour	Creation Deletion

SAOM

Model

Estimation

Influence

A mother's view of social influence: "if your friend jumped off a bridge, would you?"

An example from my childhood friend Jael…

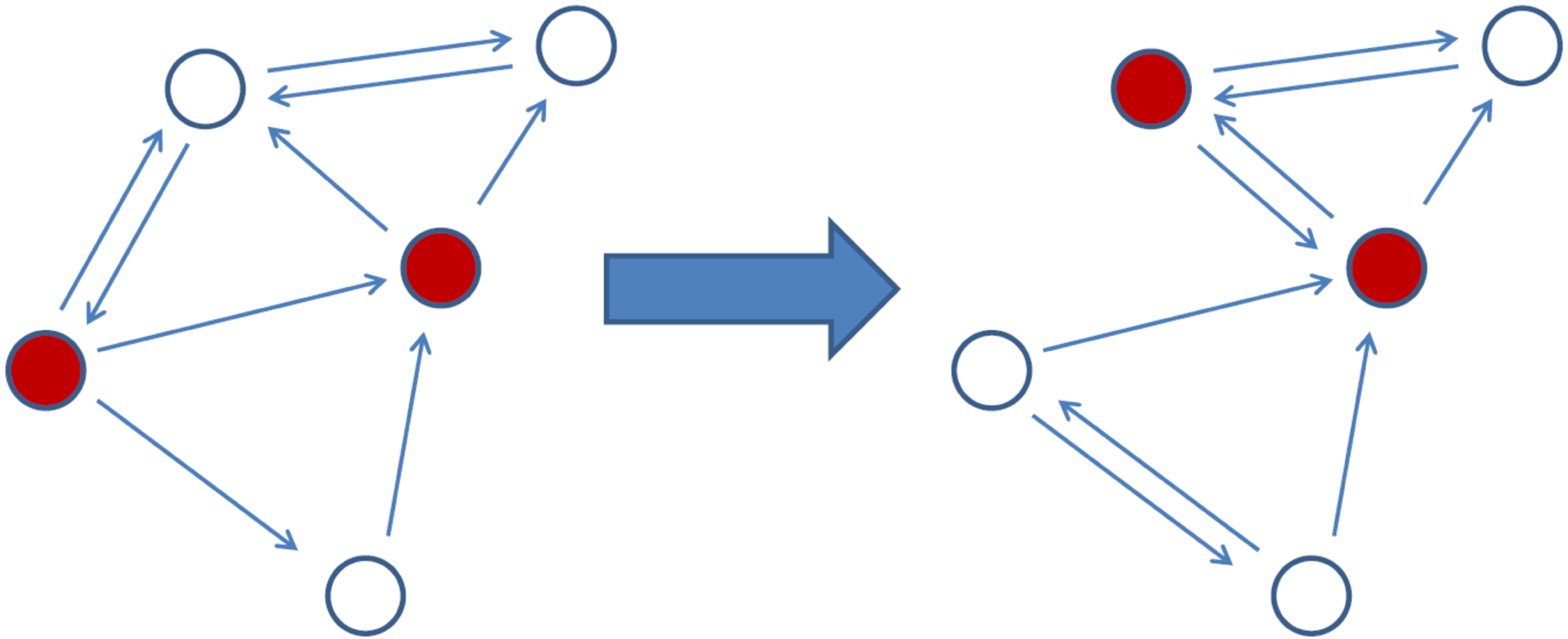
Selection

- Manifest homophily: Jael and I are friends because we both jump off bridges
- Secondary homophily, observable: Jael and I are friends because we are in the same travelling and thrill-seeking club
- Latent homophily, unobservable: Jael and I both like going on rollercoasters

Influence

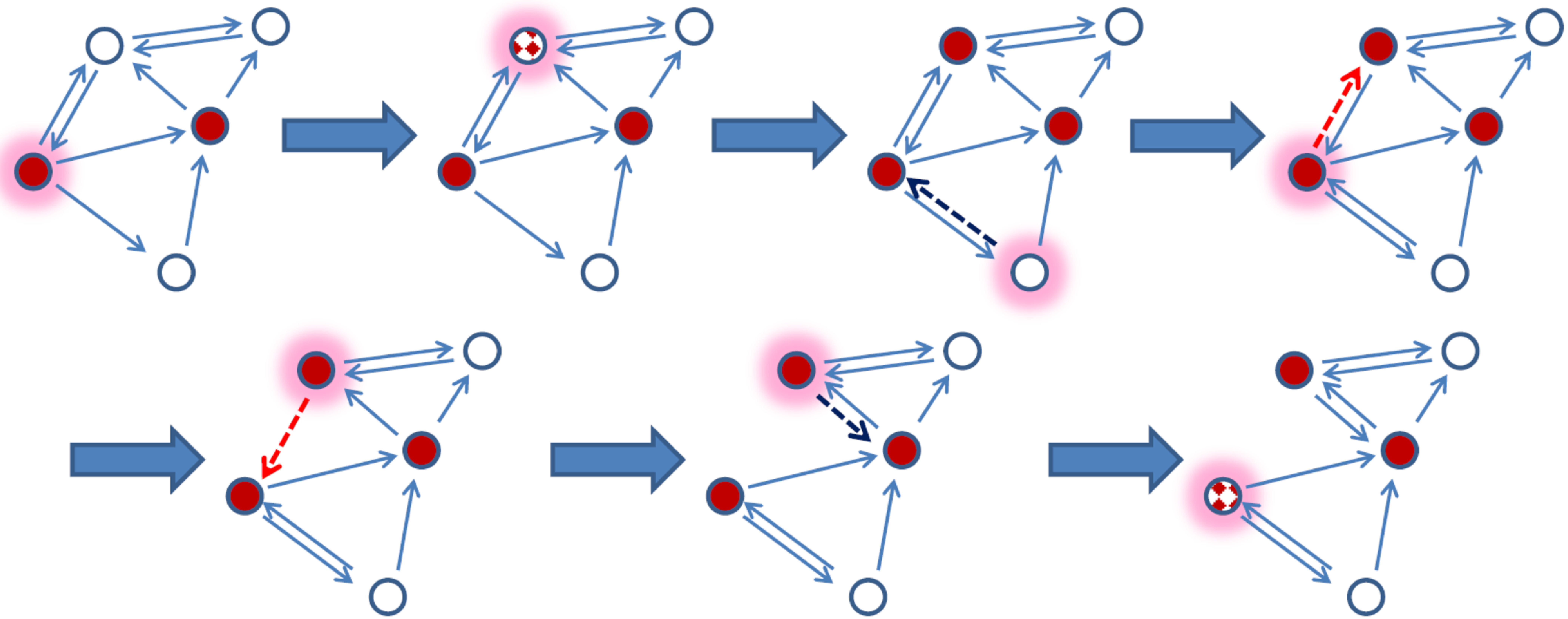
- Common external causation: Jael and I are on the Stari Most on 9 November 1993 and jumping is safer than staying on a bridge that is being destroyed by Croat forces
- Biological contagion: Jael infected me with a virus that makes people jump off bridges
- Social influence: Jael inspired me

Now networks or behaviour may change at ministeps



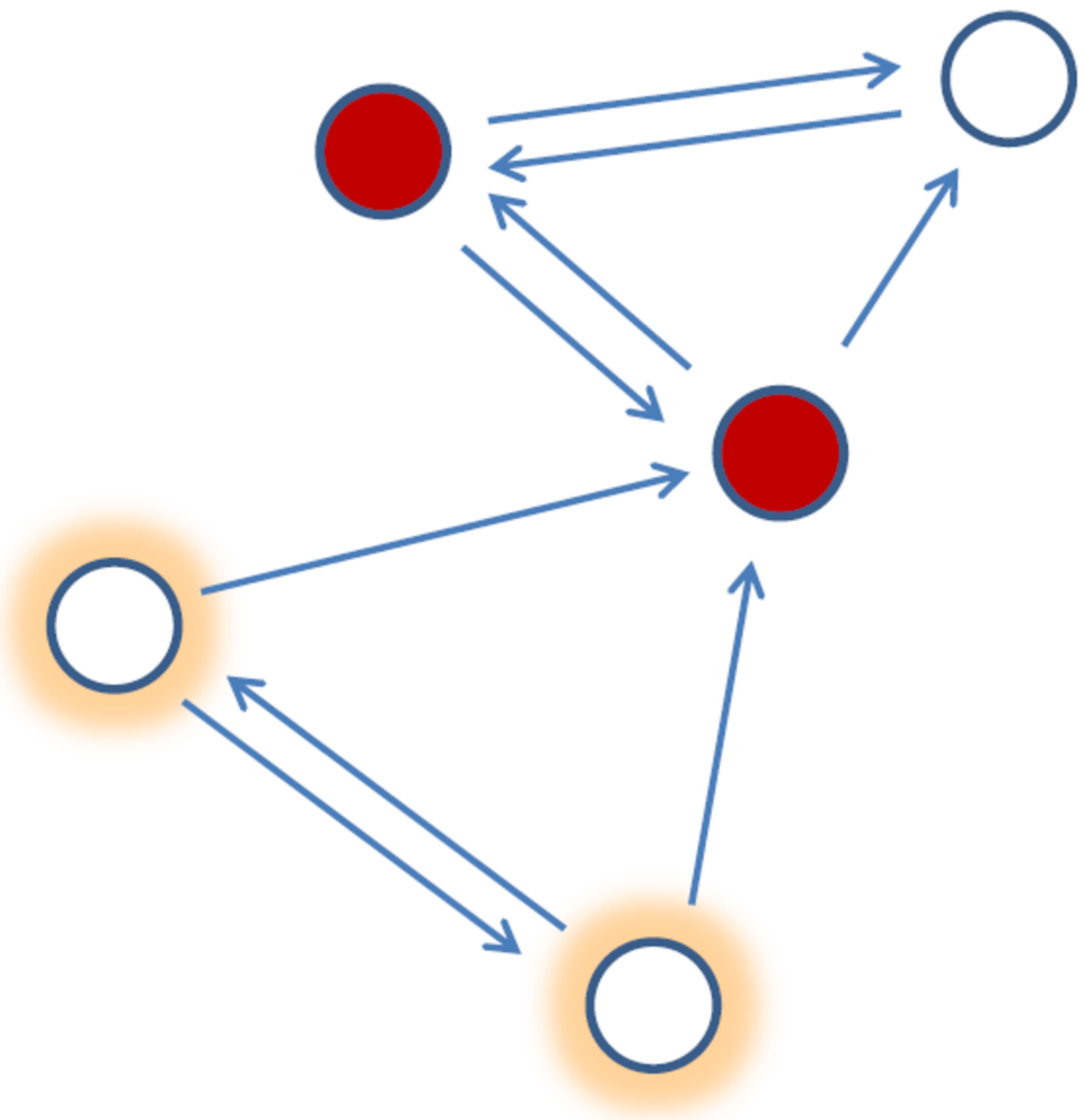
- Still two discrete observations
- Still assume continuous process of change, but now interpolates network-tie changes with behavioural changes

SAOM allows discrete changes on both levels



- Changes are actor-oriented: individuals can decide to change their outgoing ties or their behaviour
- Two Poisson processes determine time intervals between subsequent changes in each dependent variable

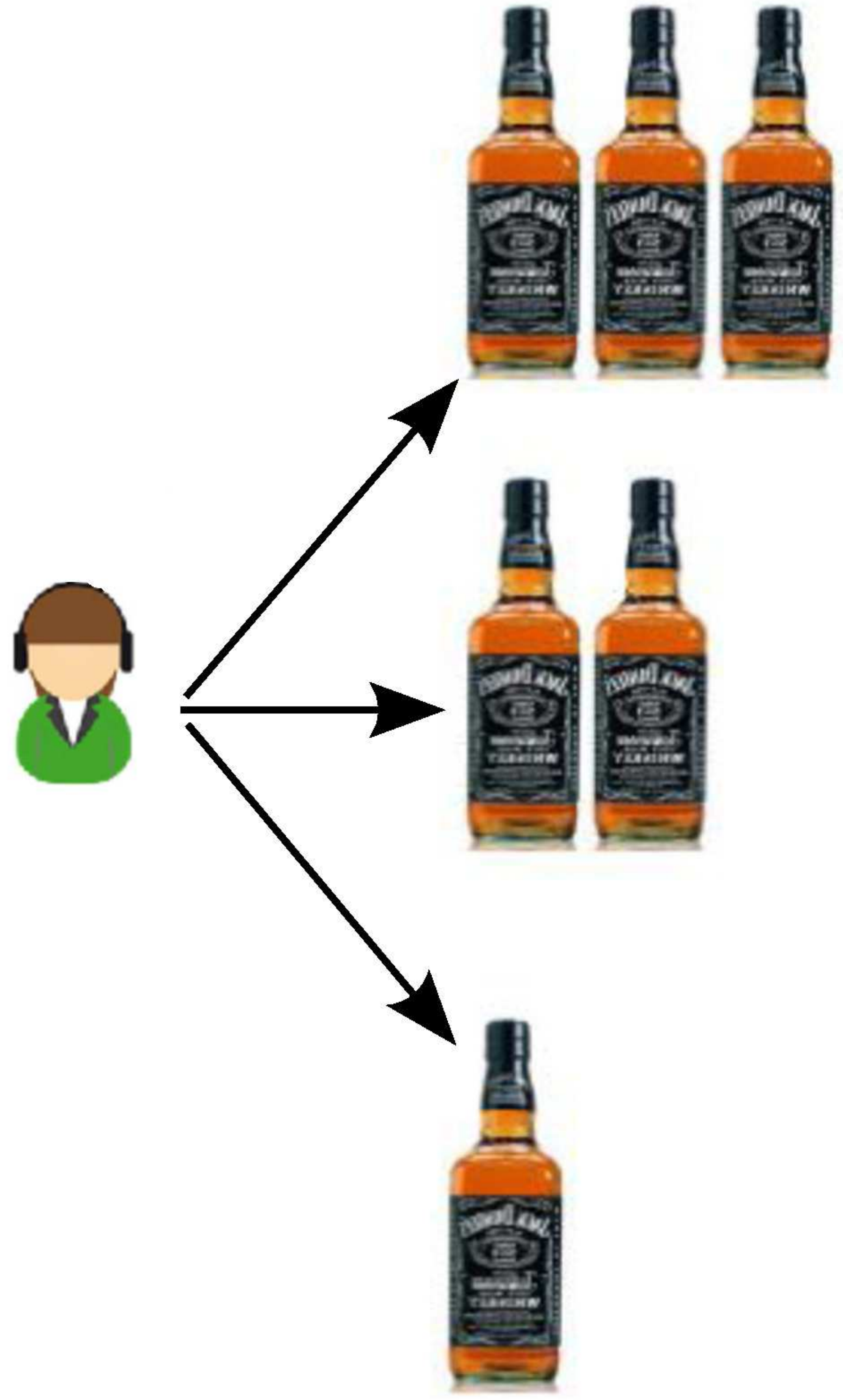
Process Markovian (and thus myopic)



- Both highlighted individuals have the same probability to change their behaviour.
- If social influence is present, they might have an increased likelihood to become red.

Behaviour change can be discrete or continuous

- Once individuals reconsider their behaviour, they can increase, decrease, or maintain it
- Actual choice modelled with a multinomial probability (up, down, stay)
- This means successive opportunities are required for large-scale behavioural changes
- Model is very similar to network change model



Niezink et al 2019

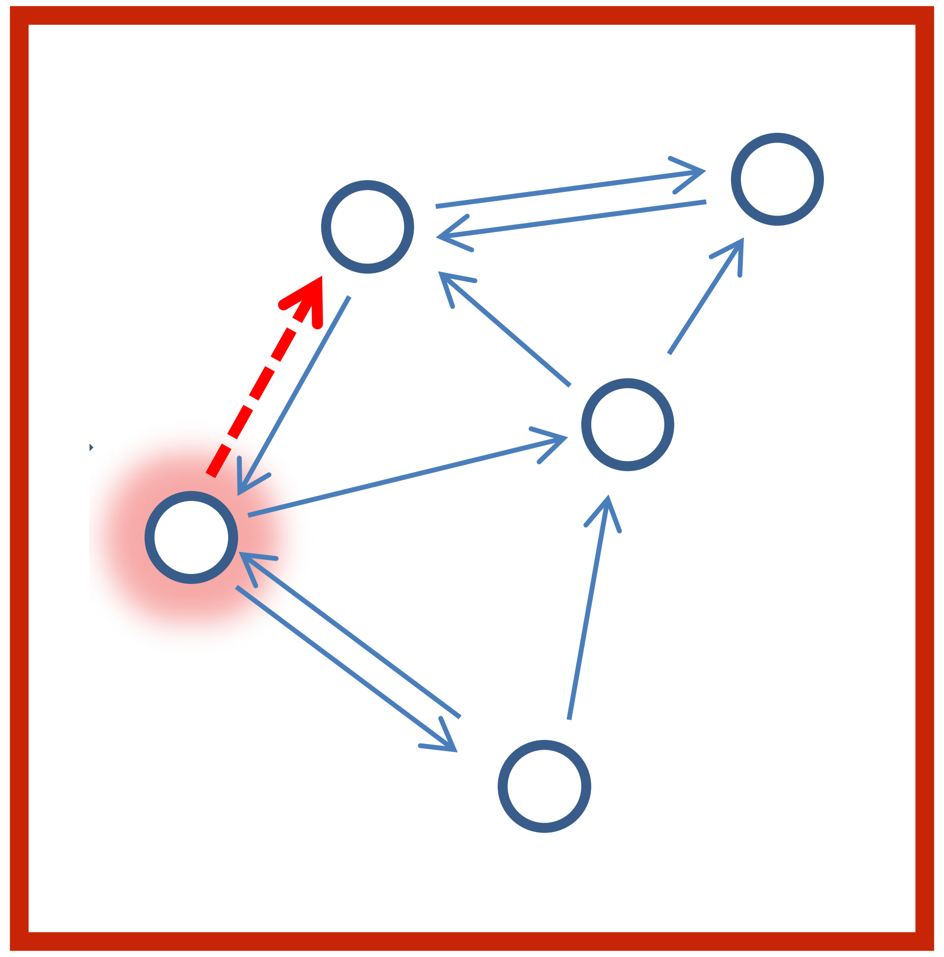
SAOMmary

- Network (and behaviour) change is observed across repeated measures
- This discrete change decomposed into continuous-time ministeps and modelled from an actor-oriented perspective
  - The frequency of these ministeps and which actors are offered an opportunity to change their ties/behaviour is modelled by the rate function
  - What happens during these ministeps/opportunities is modelled by an evaluation function, and the effects included here tend to be most related to research questions
- The Method of Moments estimation procedure seeks to find stable parameter values that simulate networks that match the target statistics of the effects included and are stable (convergence) and also replicate salient macro-structural features (goodness-of-fit)

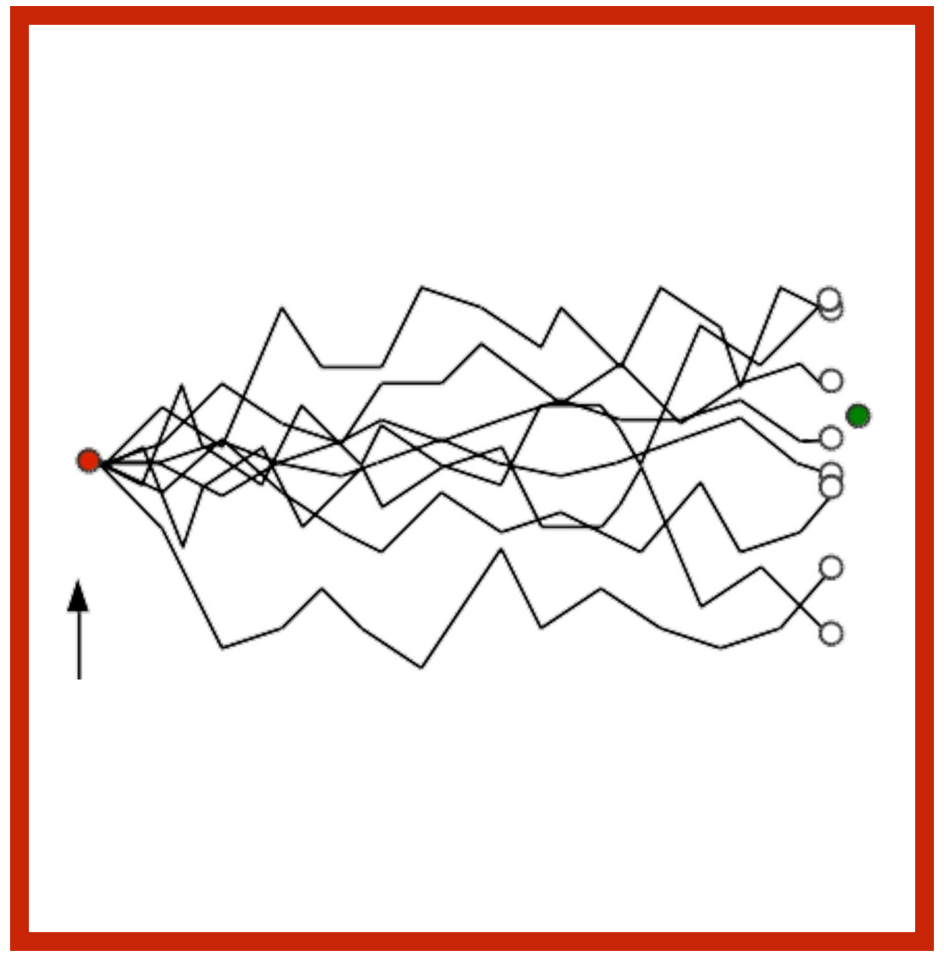
	Tie-Oriented	**Actor-Oriented**
Cross-Sectional /Panel Data	(T)ERGMs	SAOMs
Time-Stamped Data	REMs	DyNAMs

SAOM

Model



Estimation



Influence