ETH zürich

AUTOMATIC CONTROL LABORATORY

Modelling, Analysis, and Inference in Social Network Formation

Nicolò Pagan

PhD candidate at Automatic Control Laboratory, ETH

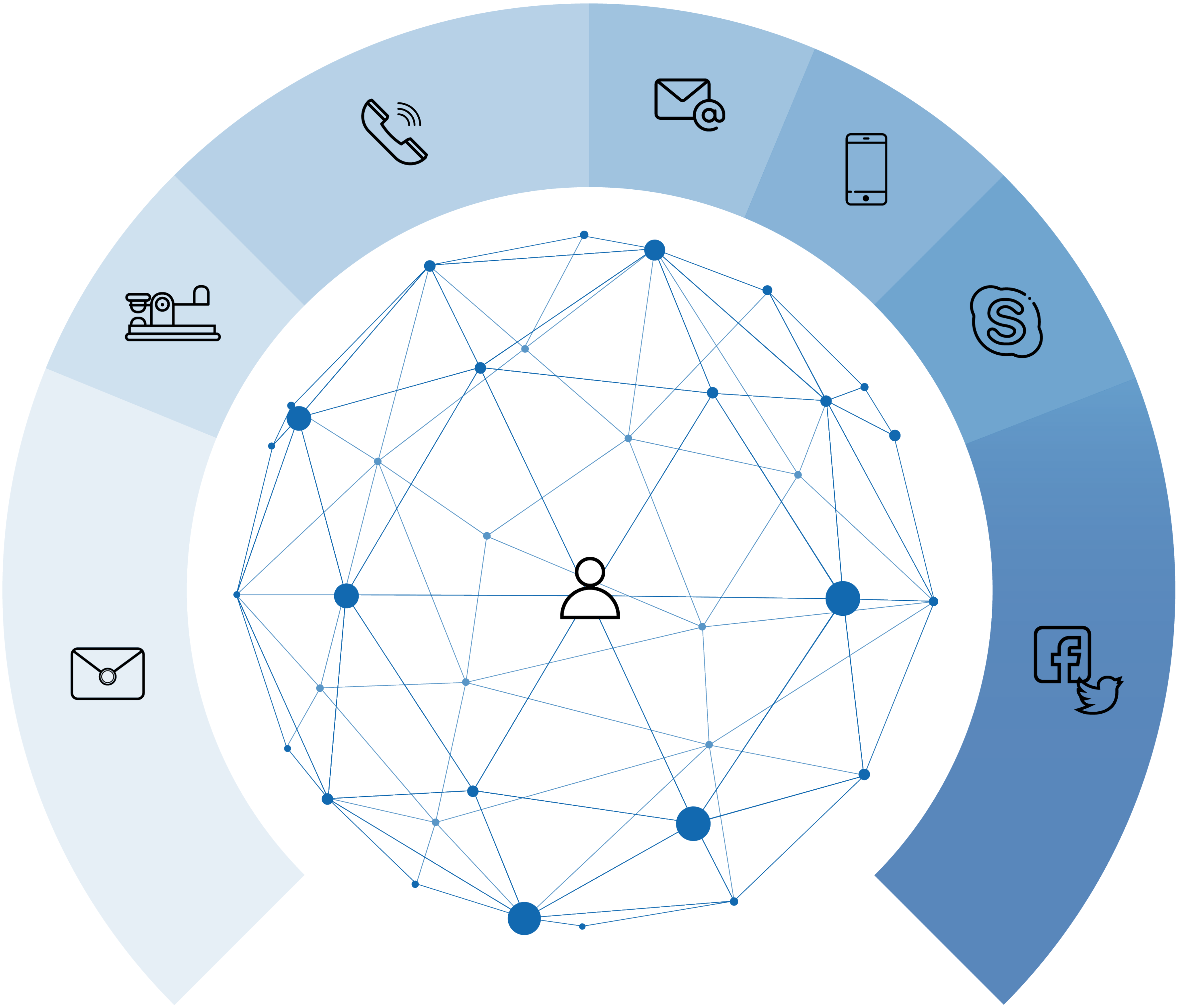
April 08, 2021, Zürich

Social Networks in Human History

Man is by nature a social animal; an individual who is unsocial naturally and not accidentally is either beneath our notice or more than human. Society is something that precedes the individual. Anyone who either cannot lead the common life or is so self-sufficient as not to need to, and therefore does not partake of society, is either a beast or a god.

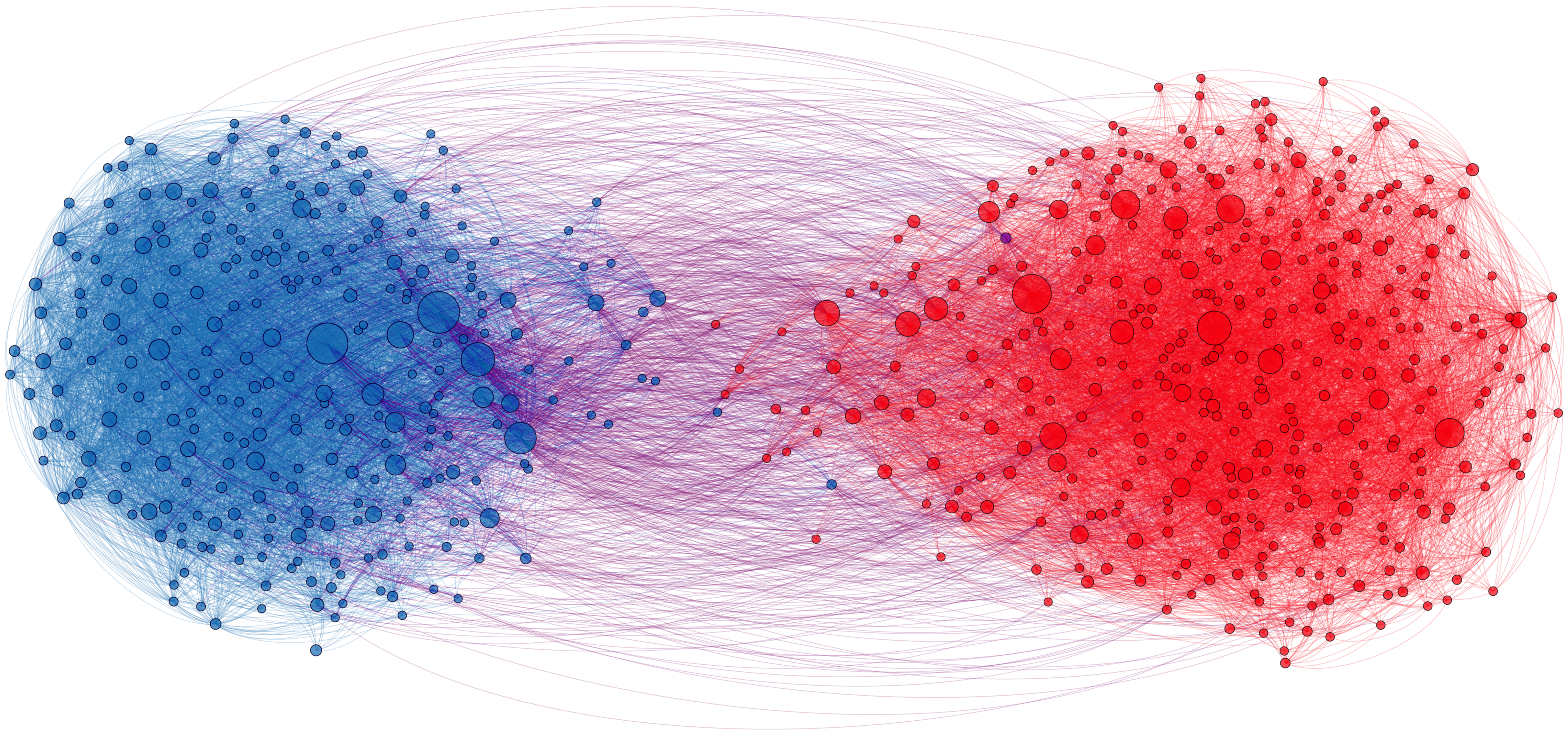
Aristotle, Politics

Social Networks in Human History



Impacts and Effects of Social Networks

- Political Polarization

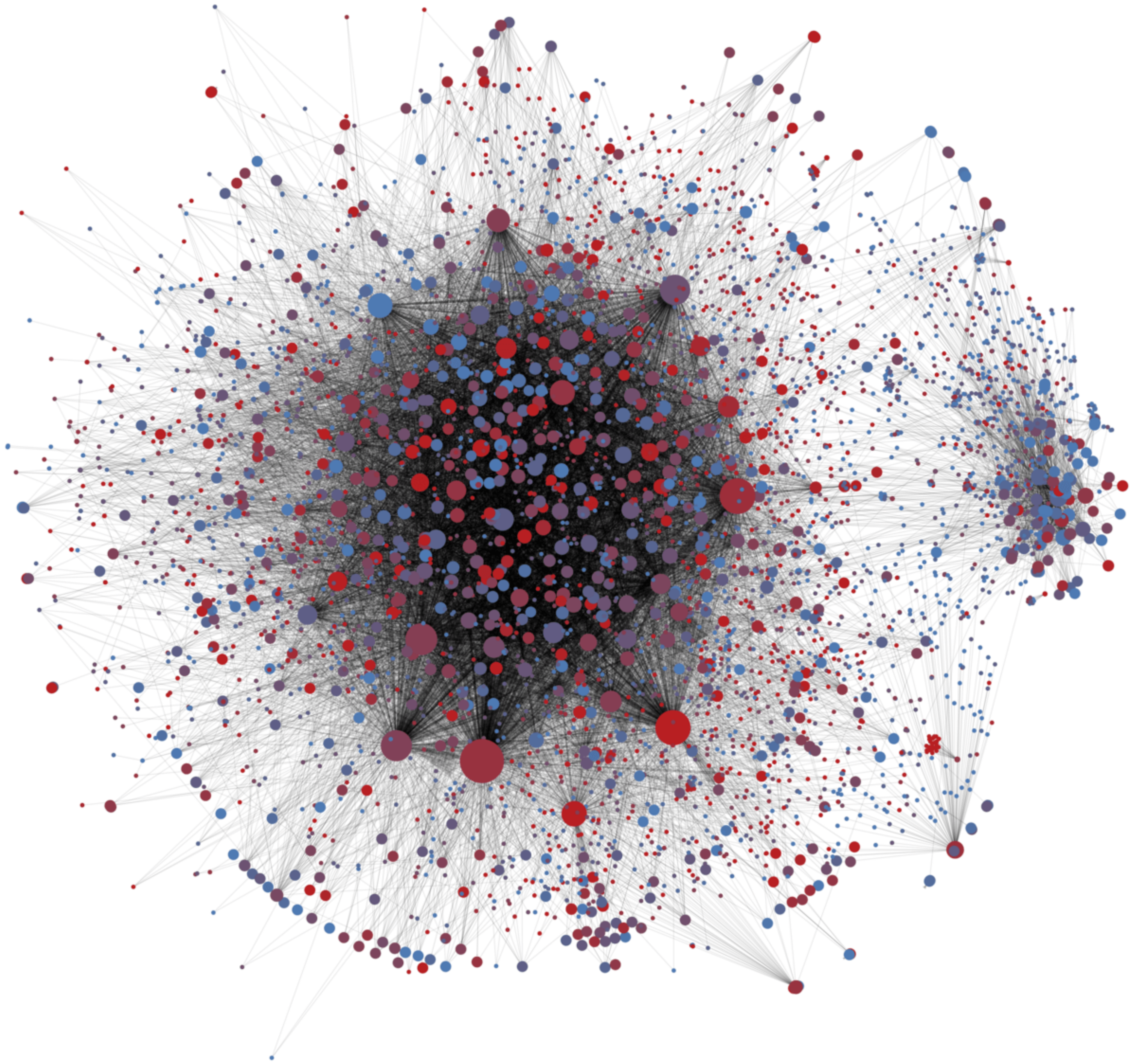


Source: https://allthingsgraphed.com/

Impacts and Effects of Social Networks

- Political Polarization

- (Mis)information spreading



Source: https://www.govtech.com/social/Misinformation-on-Social-Media-Can-Technology-Save-Us.html

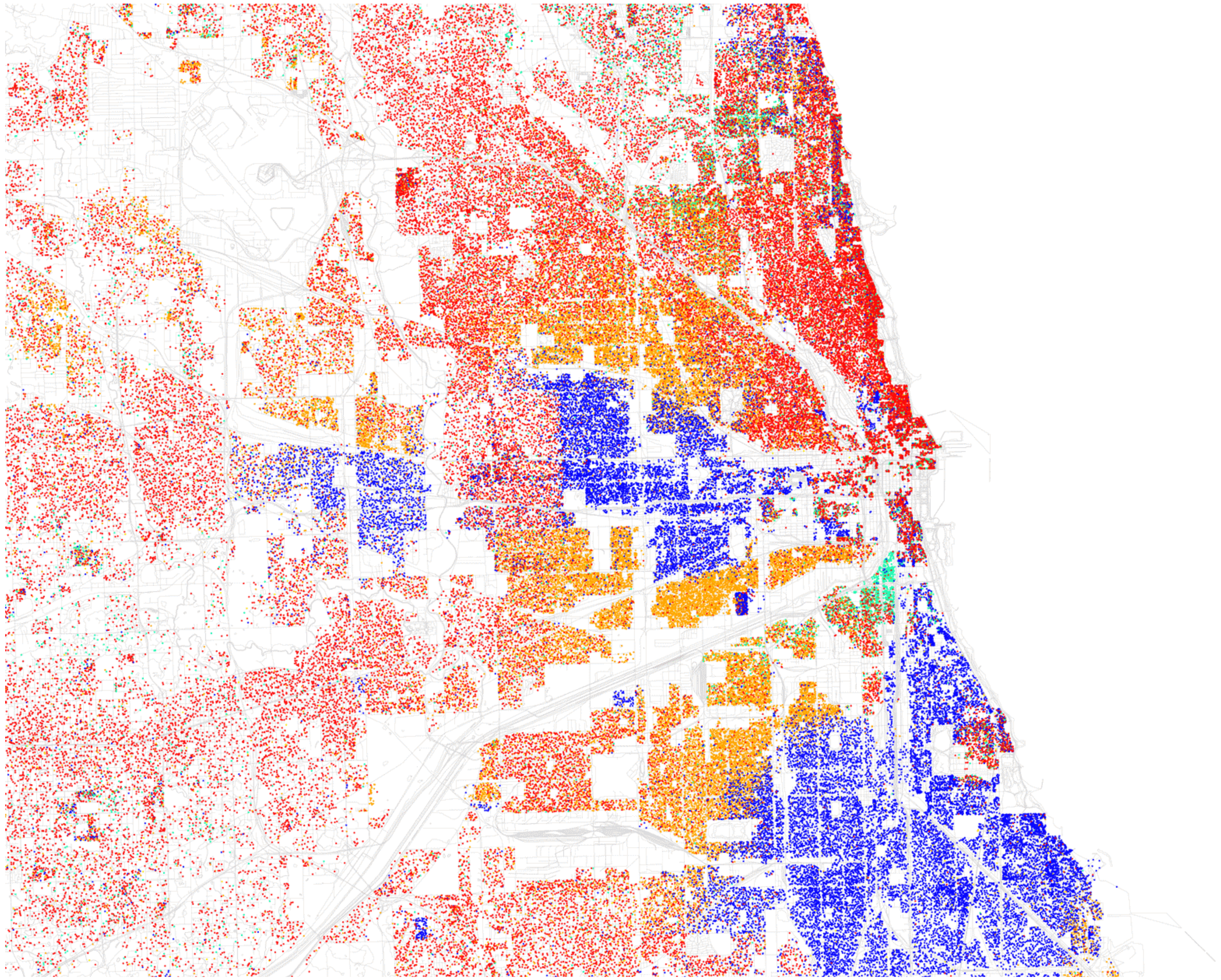
Impacts and Effects of Social Networks

- Political Polarization
- (Mis)information spreading
- Disease spreading

Source: https://storymaps.arcgis.com/stories/4fdc0d03d3a34aa485de1fb0d2650ee0

Impacts and Effects of Social Networks

- Political Polarization
- (Mis)information spreading
- Disease spreading
- Racial segregation
- Adoption of new technologies
- Diffusion of healthy behaviour, …



Source: http://chicagostories.org/black-chicago/

Social Networks
Complexity

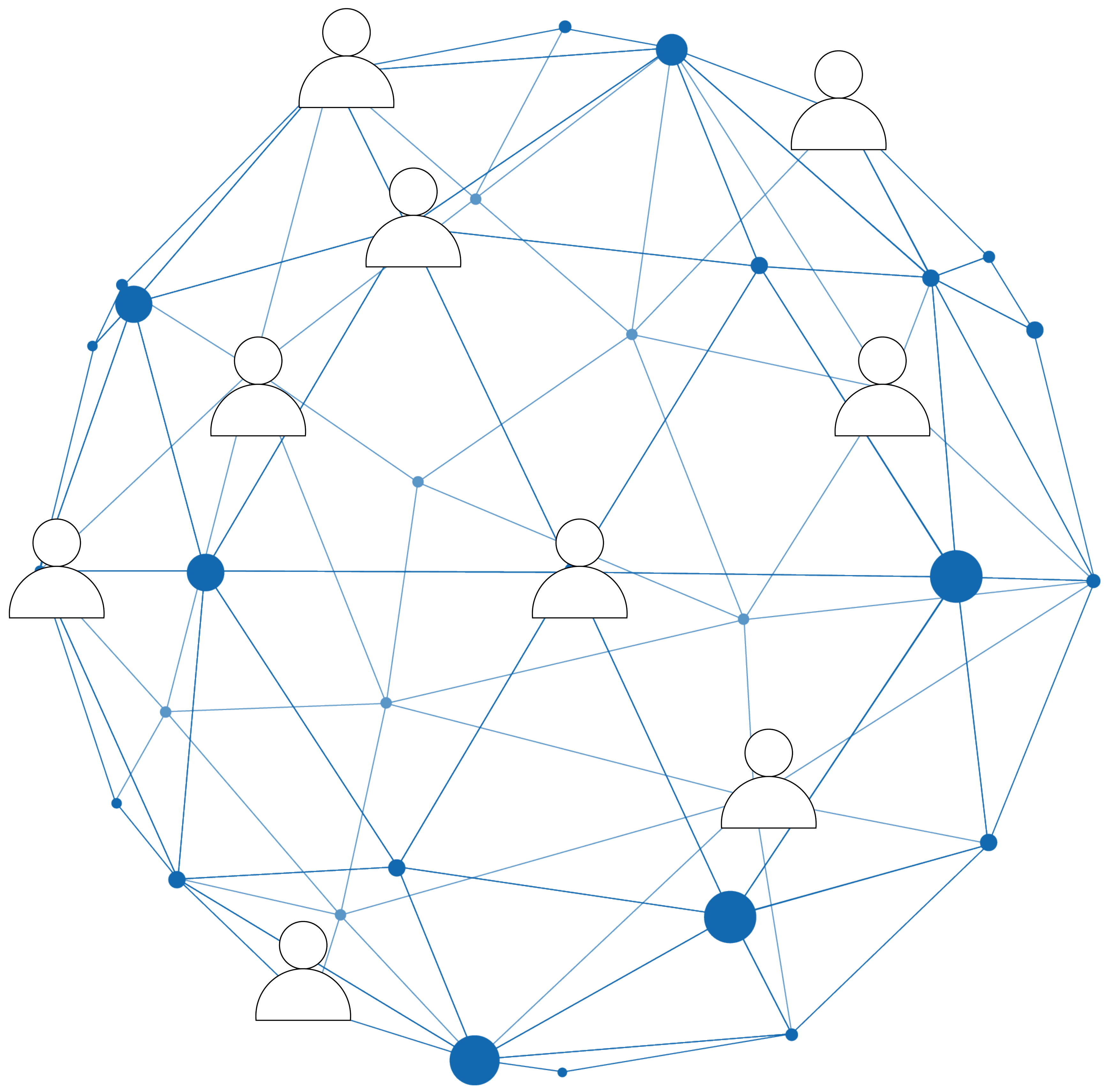
Large-Scale Systems

Human Behaviour

Interactions

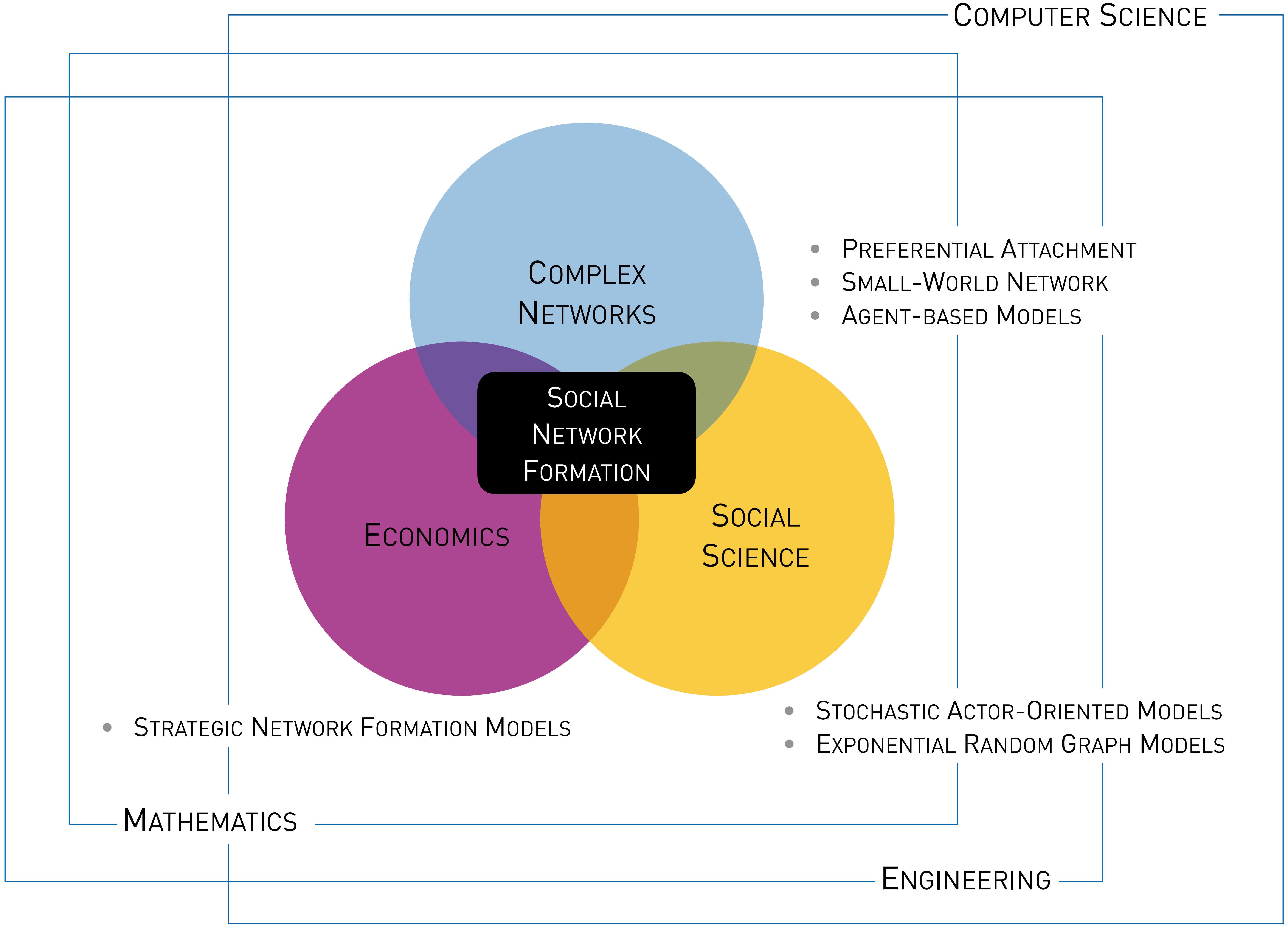
Objective

With the objective of advancing our understanding of such complex systems, we focus on the modelling, analysis, and inference of the social network formation process.



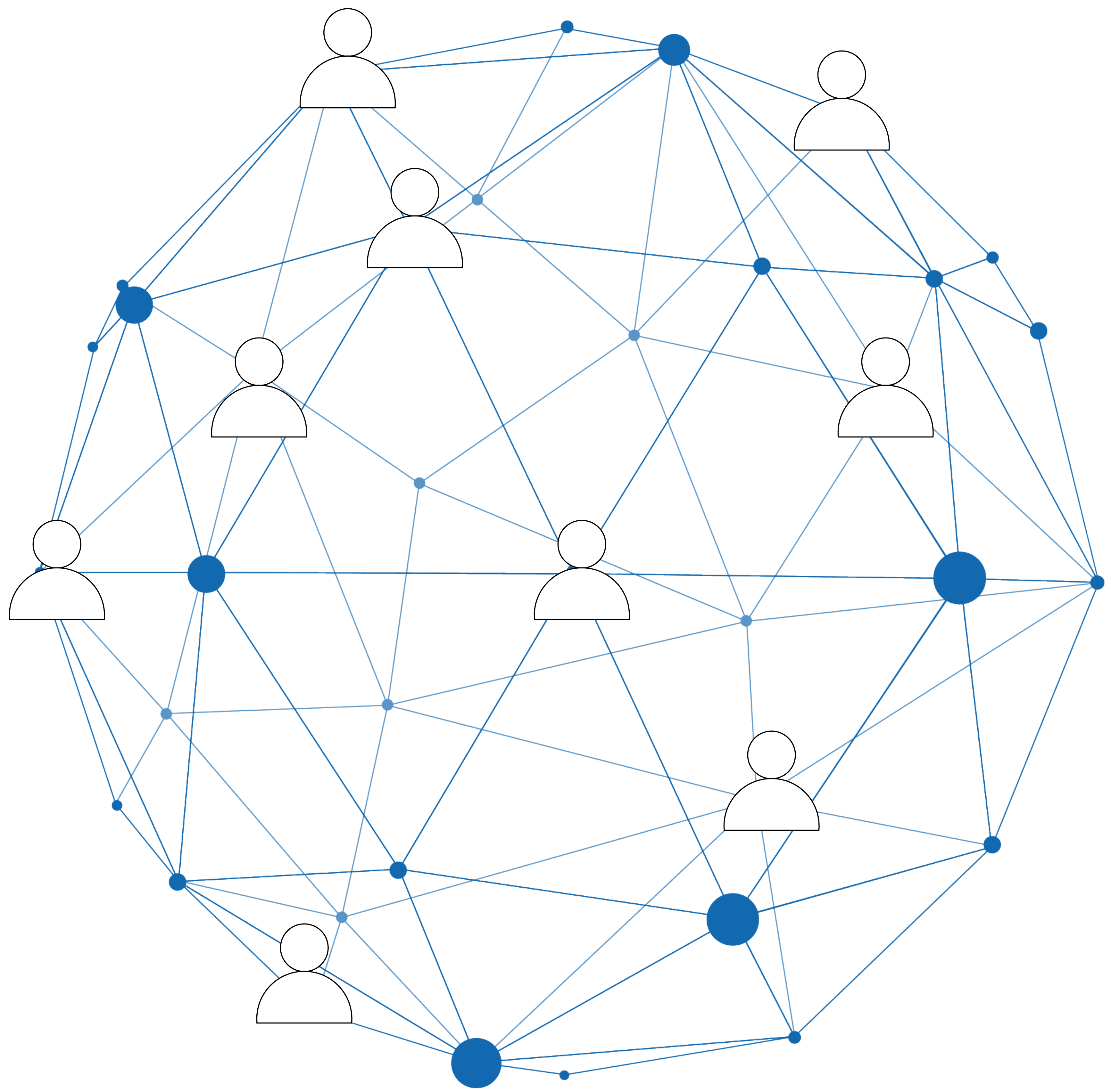
Social Network Formation Problem

Researchers from different communities study the problem of Social Network Formation.



Preliminaries

- A social network is composed of nodes, representing actors/agents, and of ties or links, which correspond to their relationships.



Preliminaries

- A social network is composed of nodes, representing actors/agents, and of ties or links, which correspond to their relationships.

- Actors decide who they want to follow.

Preliminaries

- A social network is composed of nodes, representing actors/agents, and of ties or links, which correspond to their relationships.

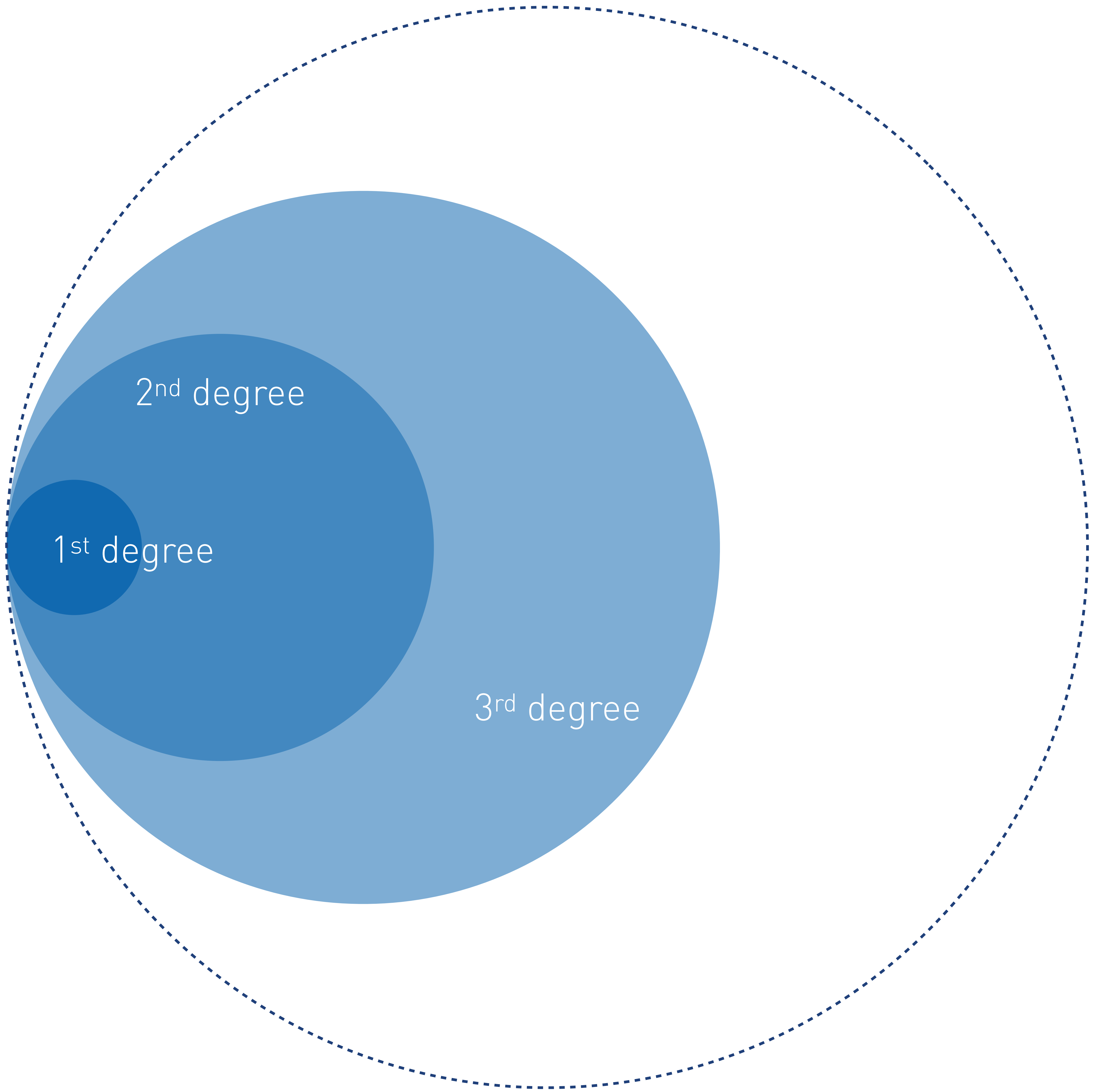
- Actors decide who they want to follow.

- Ties can have different importance (weight).

Source: https://homefamily.net

Preliminaries

- A social network is composed of nodes, representing actors/agents, and of ties or links, which correspond to their relationships.
- Actors decide who they want to follow.
- Ties can have different importance (weight).
- Limited information is available.



Preliminaries

- A social network is composed of nodes, representing actors/agents, and of ties or links, which correspond to their relationships.

- Actors decide who they want to follow.

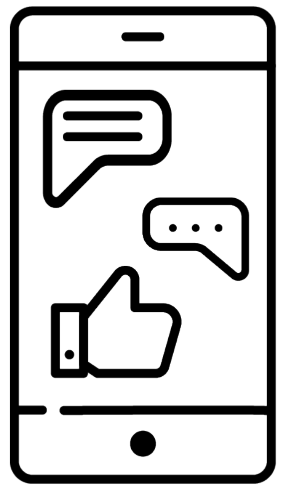
- Ties can have different importance (weight).

- Limited information is available.

- Network positions provide benefits/costs to the actors.

Offline vs Online Social Networks

Offline Social Networks

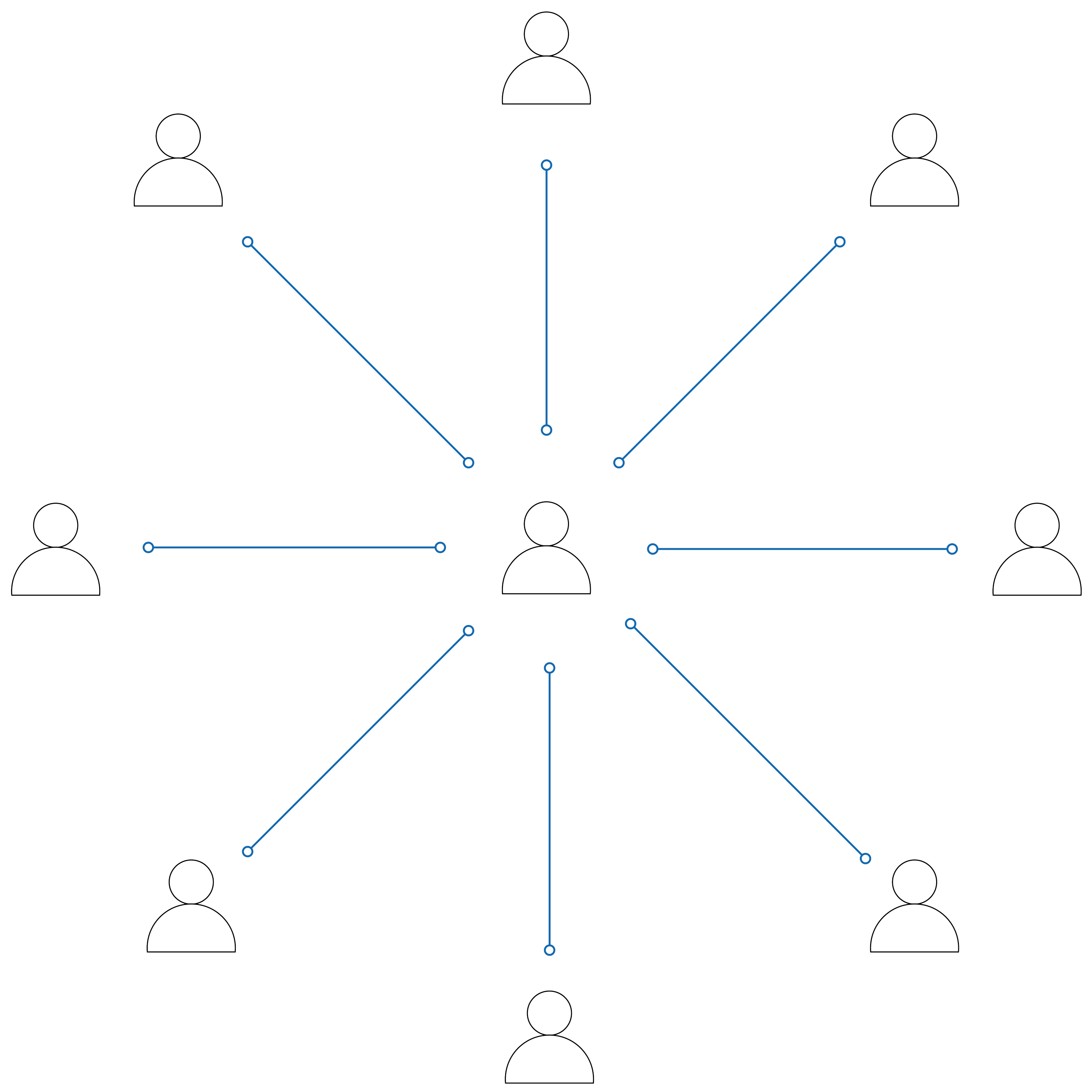


Online Social Networks

PART 1.
STRATEGIC AGENTS:
GAME THEORETICAL APPROACH

Social Network Positions' Benefits

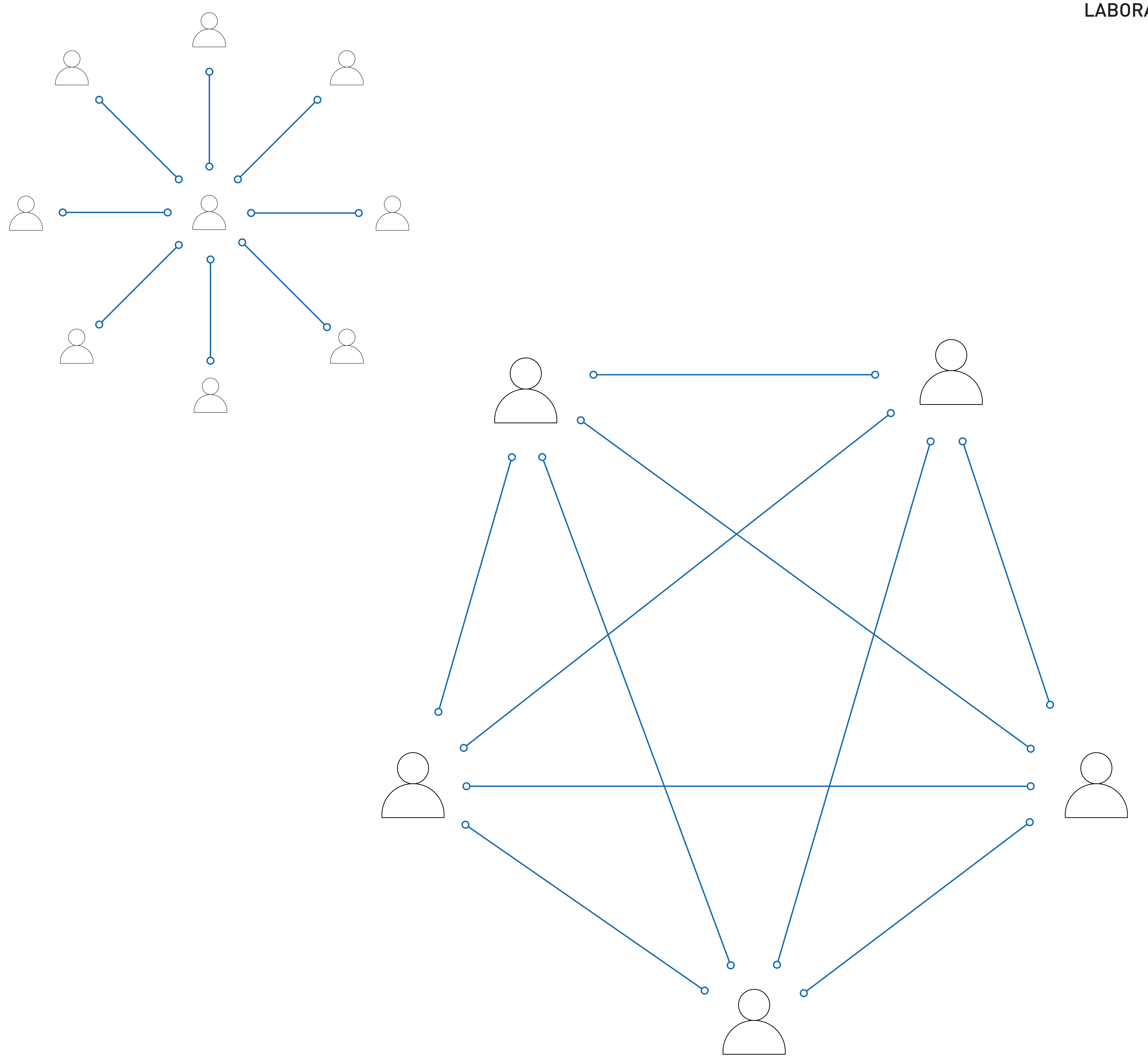
- Social Influence:
  the more people we are connected to, the more influence we have in the network.
  [Robins, G. "Doing social network research", 2015]



Social Network Positions' Benefits

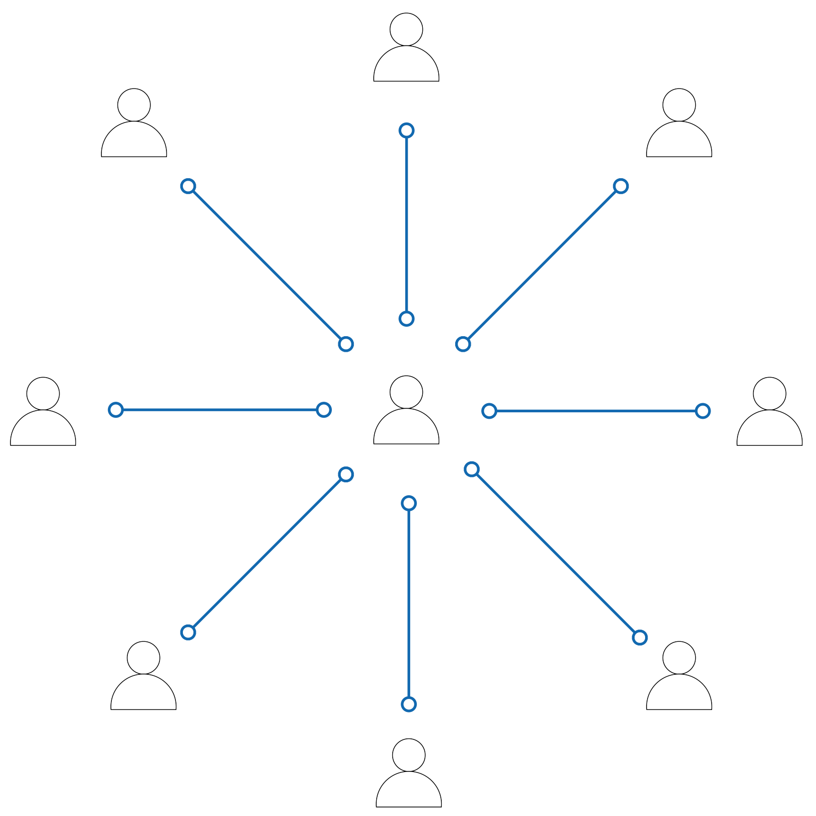
- Social Influence:
  the more people we are connected to, the more influence we have in the network.
  [Robins, G. "Doing social network research", 2015]

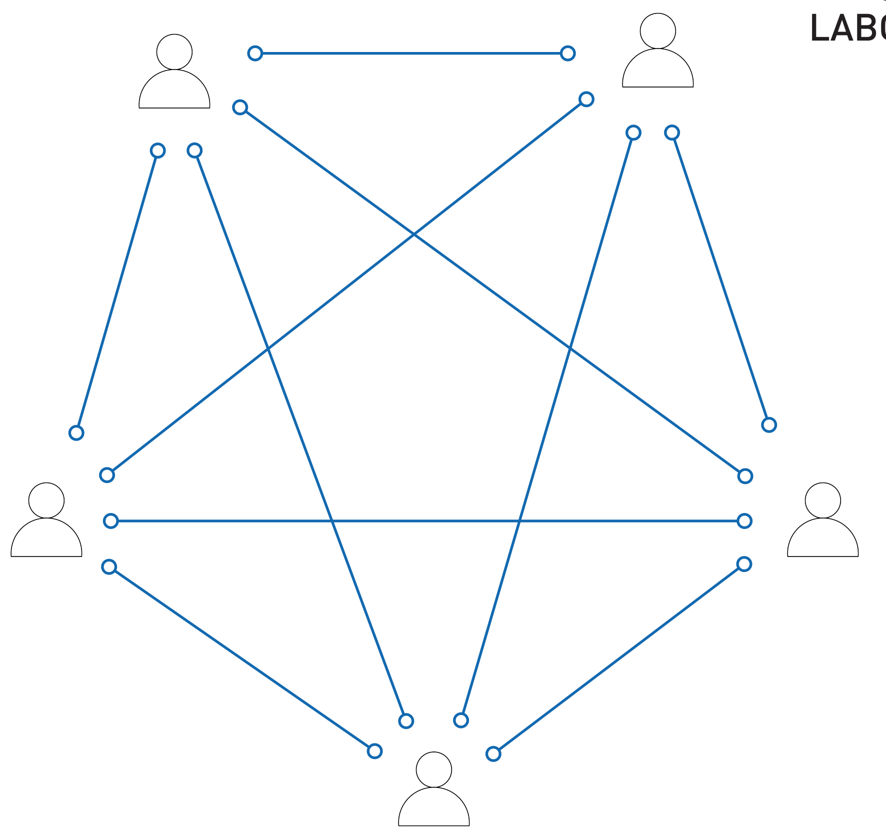
- Social Support:
  the more our friends' friends are our friends, the safer we feel.
  [Coleman, J. "Foundations of social theory", 1990]



Social Network Positions' Benefits

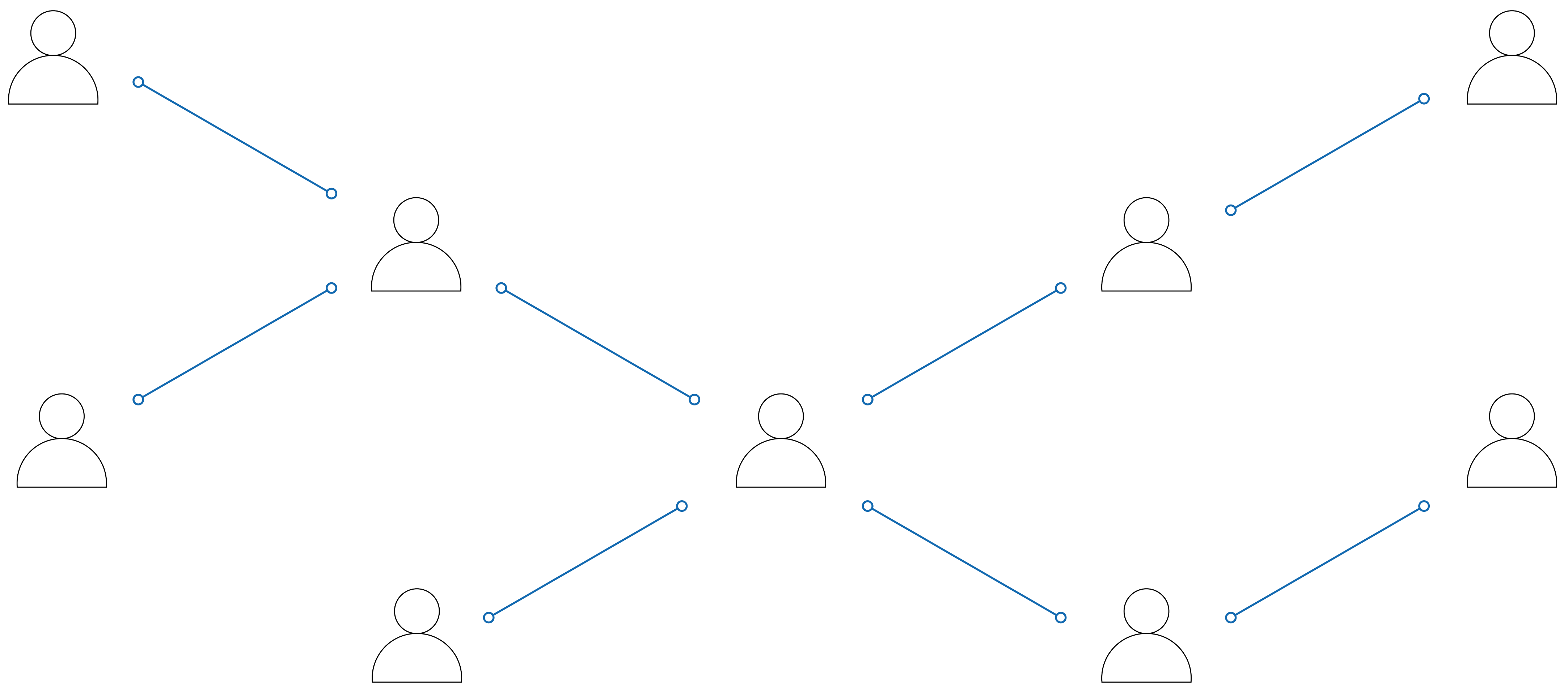
- **Social Influence**:
  the more people we are connected to, the more influence we have in the network.
  [Robins, G. "Doing social network research", 2015]





- **Social Support**:
  the more our friends' friends are our friends, the safer we feel.
  [Coleman, J. "Foundations of social theory", 1990]

- **Brokerage**:
  the more we are on the path between people, the more we can control.
  [Burt, R. S. "Structural Hole", 1992]

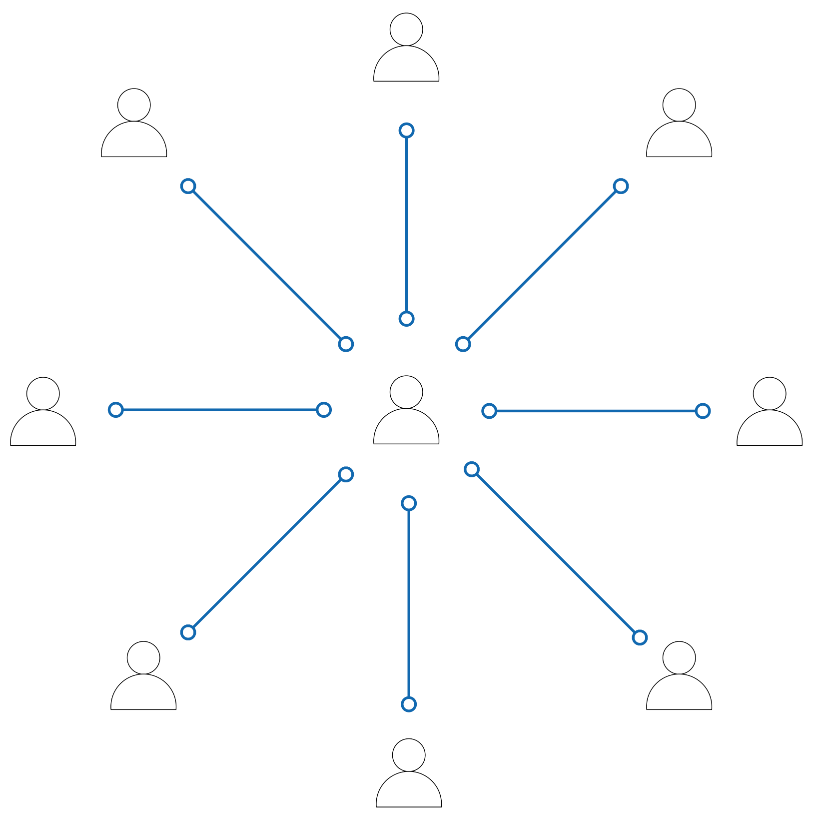


Social Network Positions' Benefits

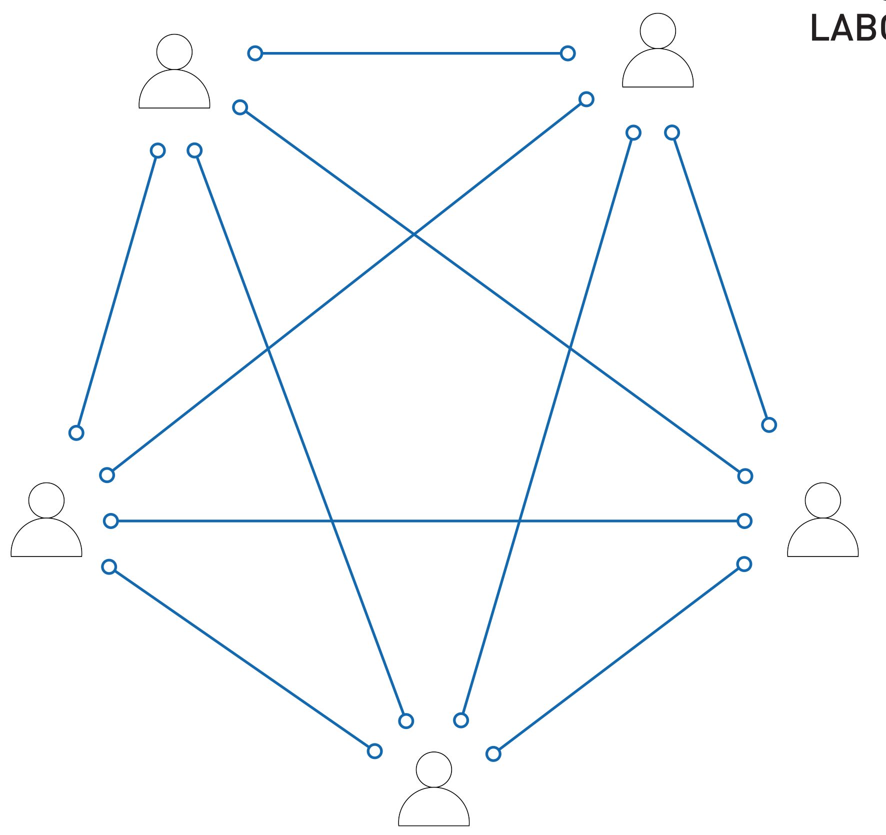
- **Social Influence**:
  the more people we are connected to, the more influence we have in the network.
  [Robins, G. "Doing social network research", 2015]

- **Social Support**:
  the more our friends' friends are our friends, the safer we feel.
  [Coleman, J. "Foundations of social theory", 1990]

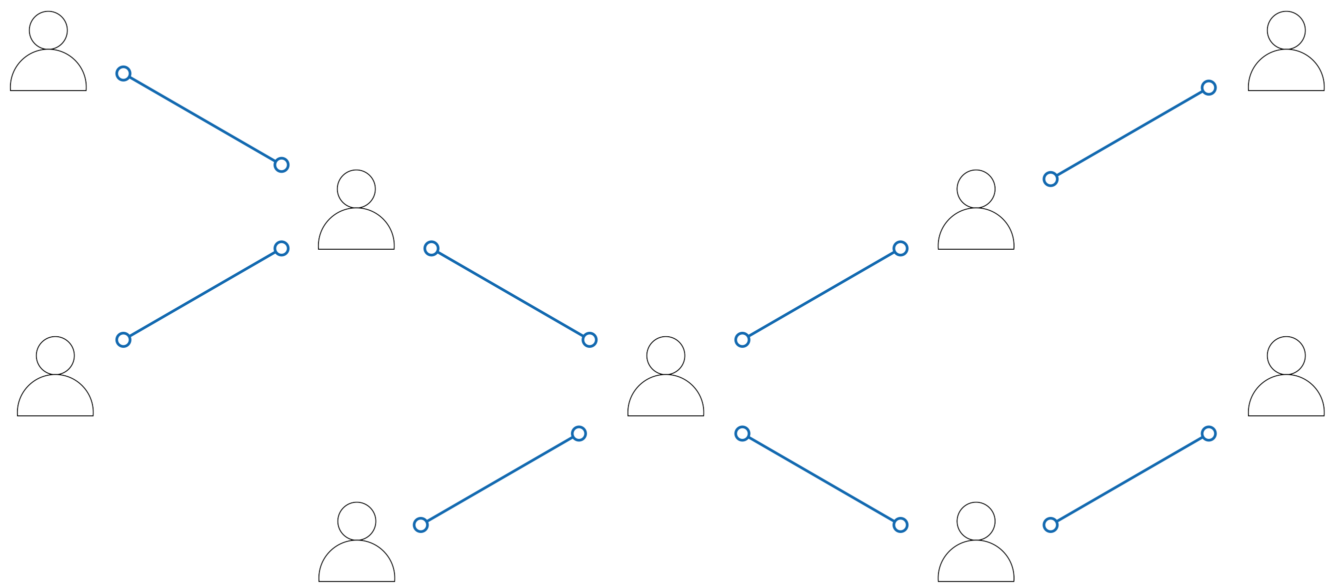
- **Brokerage**:
  the more we are on the path between people, the more we can control.
  [Burt, R. S. "Structural Hole", 1992]



Degree Centrality

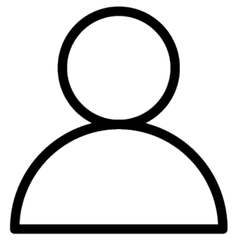


Transitivity



Betweenness Centrality

Social Network Formation Model



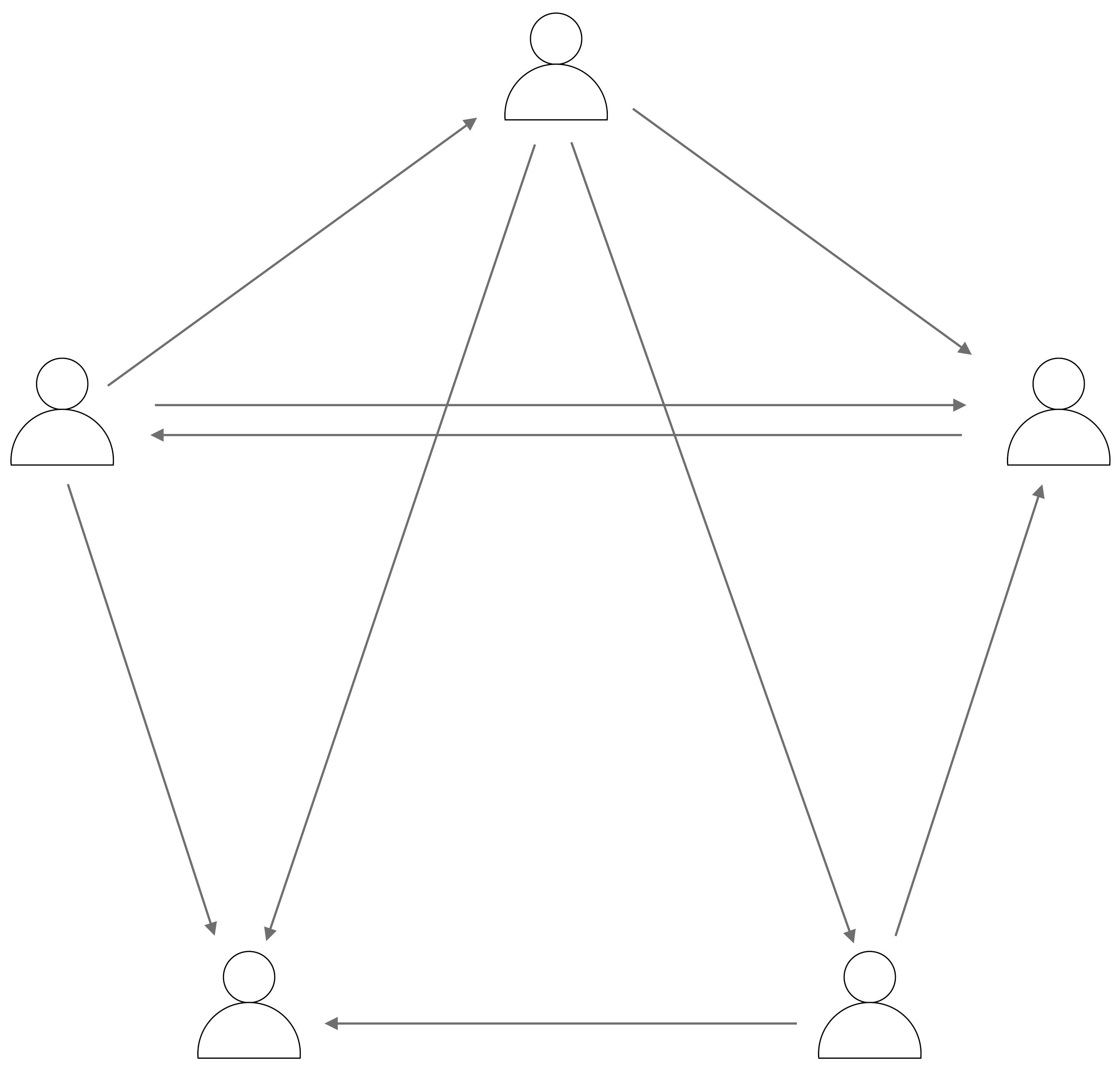
Actors/Agents/Players

Actions

Utility Functions

Social Network Formation Model

Directed weighted network $\mathcal{G}$ with $\mathcal{N} := \{1, \ldots, N\}$ agents. The weight $a_{ij} \in [0,1]$ quantifies the importance of the friendship among i and j from i's point of view.

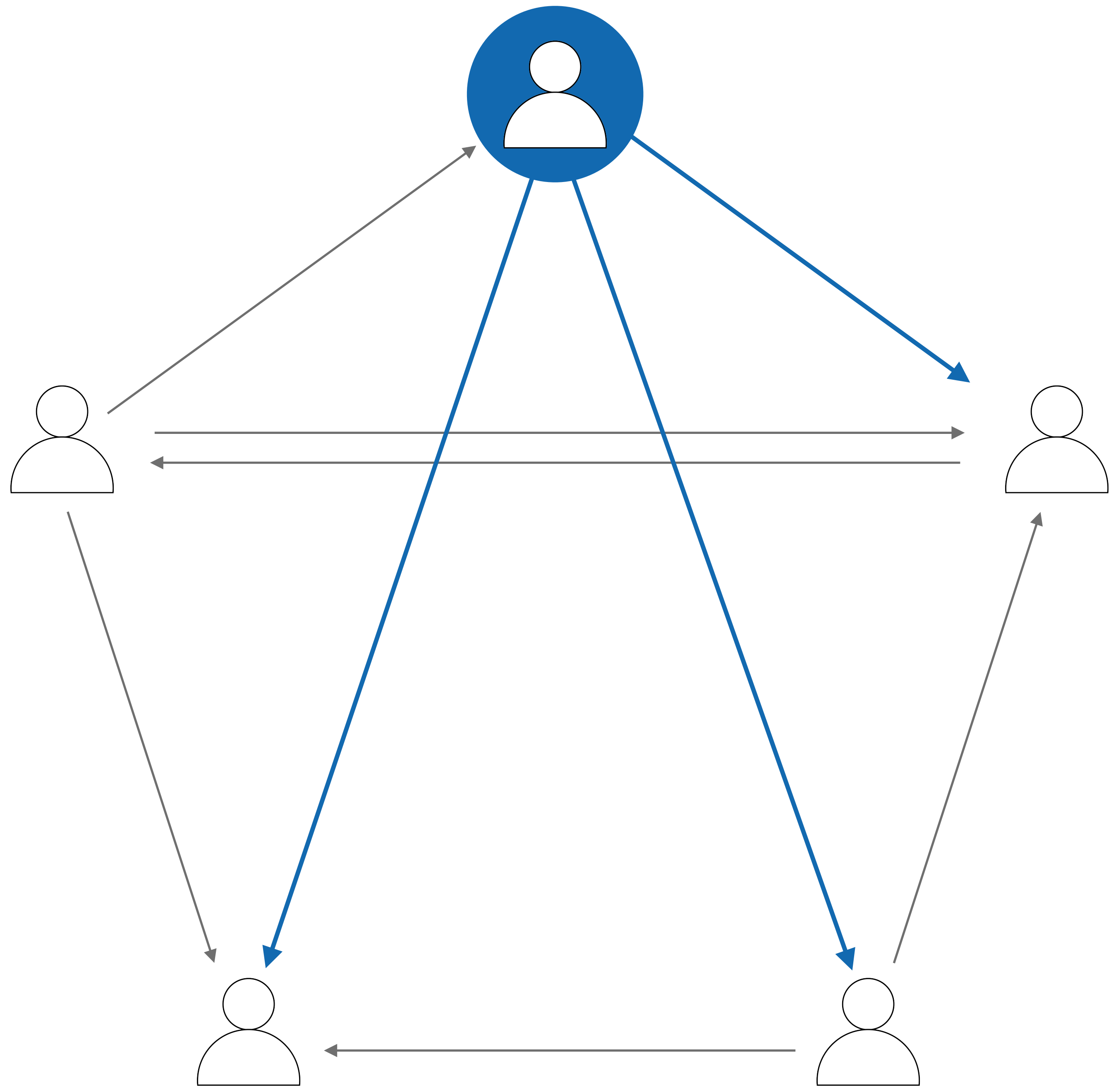


Social Network Formation Model

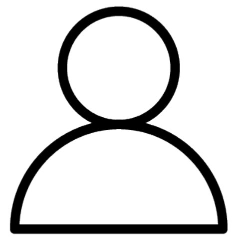
Directed weighted network $\mathcal{G}$ with $\mathcal{N} := \{1,\ldots,N\}$ agents. The weight $a_{ij} \in [0,1]$ quantifies the importance of the friendship among i and j from i's point of view.

A typical action of agent i:

$$a^i = [a_{i1}, \ldots, a_{i,i-1}, a_{i,i+1}, \ldots, a_{iN}] \in \mathcal{A} = [0,1]^{N-1}.$$



Social Network Formation Model



Directed weighted network $\mathscr{G}$ with $\mathscr{N} := \{1,\ldots,N\}$ agents. The weight $a_{ij} \in [0,1]$ quantifies the importance of the friendship among i and j from i's point of view.

A typical action of agent i:
$a^i = [a_{i1}, \ldots, a_{i,i-1}, a_{i,i+1}, \ldots, a_{iN}] \in \mathscr{A} = [0,1]^{N-1}$.

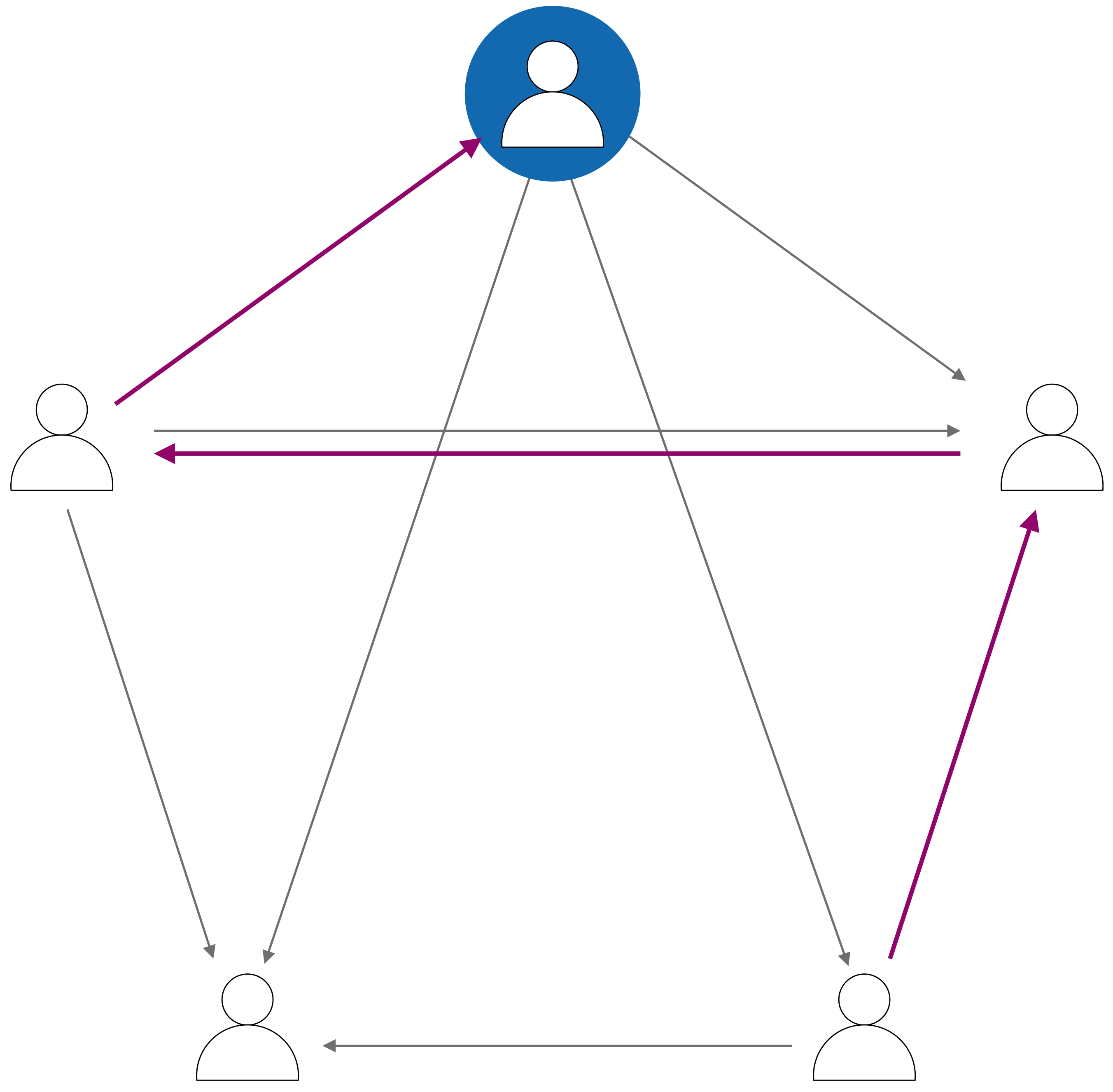
Every agent i is endowed with a payoff function $V^i\left(a^i, a^{-i}\right)$, representing the network position's benefits (minus costs) to agent i, that he/she tries to maximize (utilitarian principle).

Social Network Formation Model

Directed weighted network $\mathcal{G}$ with $\mathcal{N} := \{1,\ldots,N\}$ agents. The weight $a_{ij} \in [0,1]$ quantifies the importance of the friendship among i and j from i's point of view.

A typical action of agent i:
$a^i = [a_{i1}, \ldots, a_{i,i-1}, a_{i,i+1}, \ldots, a_{iN}] \in \mathcal{A} = [0,1]^{N-1}$.

Every agent i is endowed with a payoff function $V^i\left(a^i, a^{-i}\right)$, representing the network position's benefits (minus costs) to agent i, that he/she tries to maximize (utilitarian principle).



Diffusion centrality

$$s_i(a^i, a^{-i}, \delta_i) := \sum_k a_{ki} + \delta_i \sum_l \sum_k a_{lk}a_{ki} + \delta_i^2 \sum_m \sum_l \sum_k a_{ml}a_{lk}a_{ki}$$

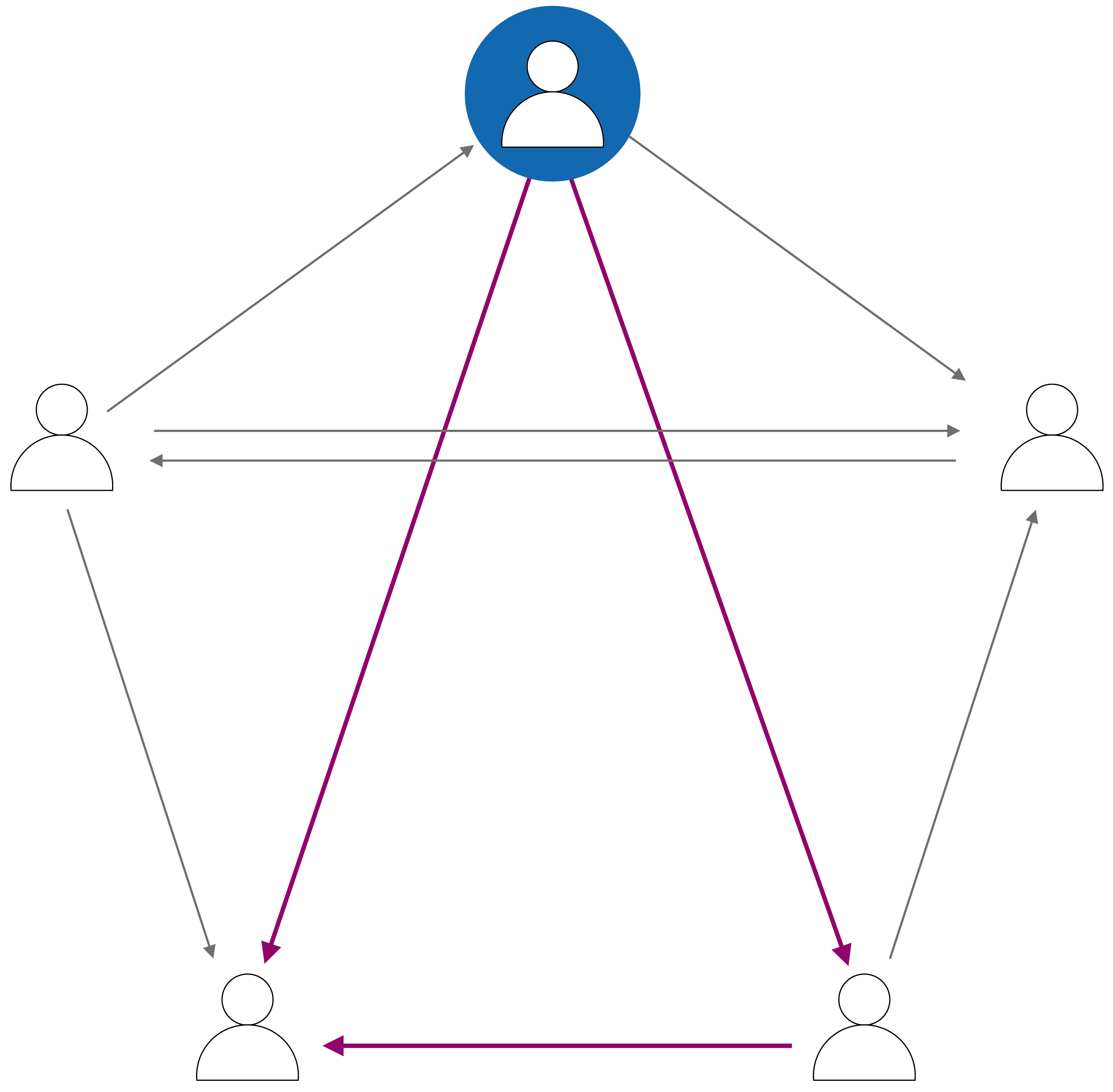
Part 1: Model

Social Network Formation Model

Directed weighted network $\mathcal{G}$ with $\mathcal{N} := \{1,\ldots,N\}$ agents. The weight $a_{ij} \in [0,1]$ quantifies the importance of the friendship among i and j from i's point of view.

A typical action of agent i:
$a^i = [a_{i1}, \ldots, a_{i,i-1}, a_{i,i+1}, \ldots, a_{iN}] \in \mathcal{A} = [0,1]^{N-1}$.

Every agent i is endowed with a payoff function $V^i\left(a^i, a^{-i}\right)$, representing the network position's benefits (minus costs) to agent i, that he/she tries to maximize (utilitarian principle).



Transitivity

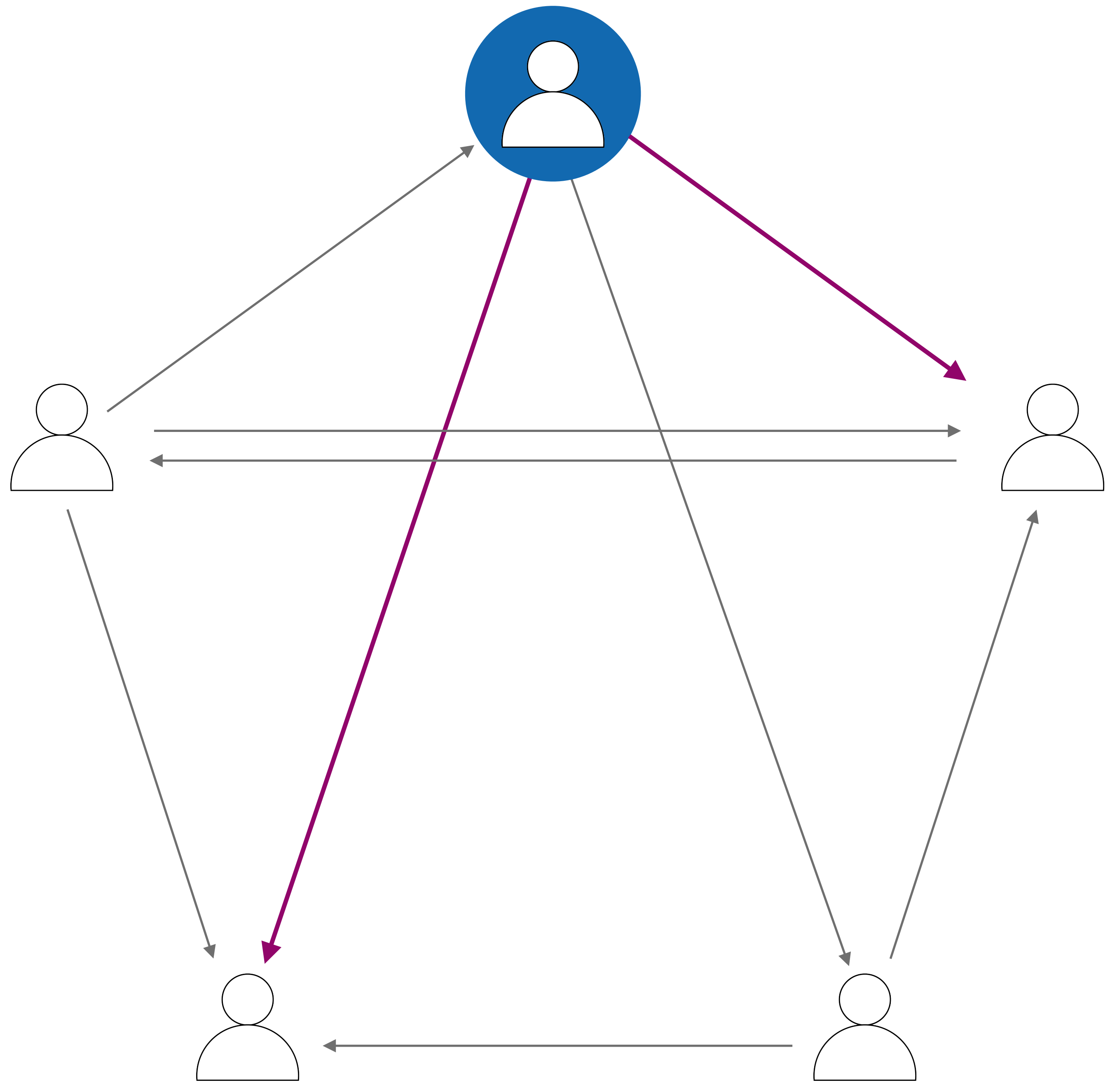
$$t_i(a^i, a^{-i}) := \sum_k a_{ik} \left(\sum_l a_{il} a_{lk} \right)$$

Social Network Formation Model

Directed weighted network $\mathscr{G}$ with $\mathscr{N} := \{1, \ldots, N\}$ agents. The weight $a_{ij} \in [0,1]$ quantifies the importance of the friendship among i and j from i's point of view.

A typical action of agent i:
$a^i = [a_{i1}, \ldots, a_{i,i-1}, a_{i,i+1}, \ldots, a_{iN}] \in \mathscr{A} = [0,1]^{N-1}$.

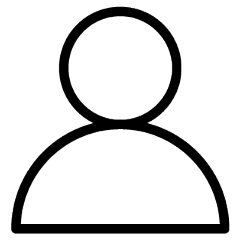
Every agent i is endowed with a payoff function $V^i\left(a^i, a^{-i}\right)$, representing the network position's benefits (minus costs) to agent i, that he/she tries to maximize (utilitarian principle).



Brokerage

$$-t_i(a^i, a^{-i}) := -\sum_k a_{ik} \left(\sum_l a_{il} a_{lk} \right)$$

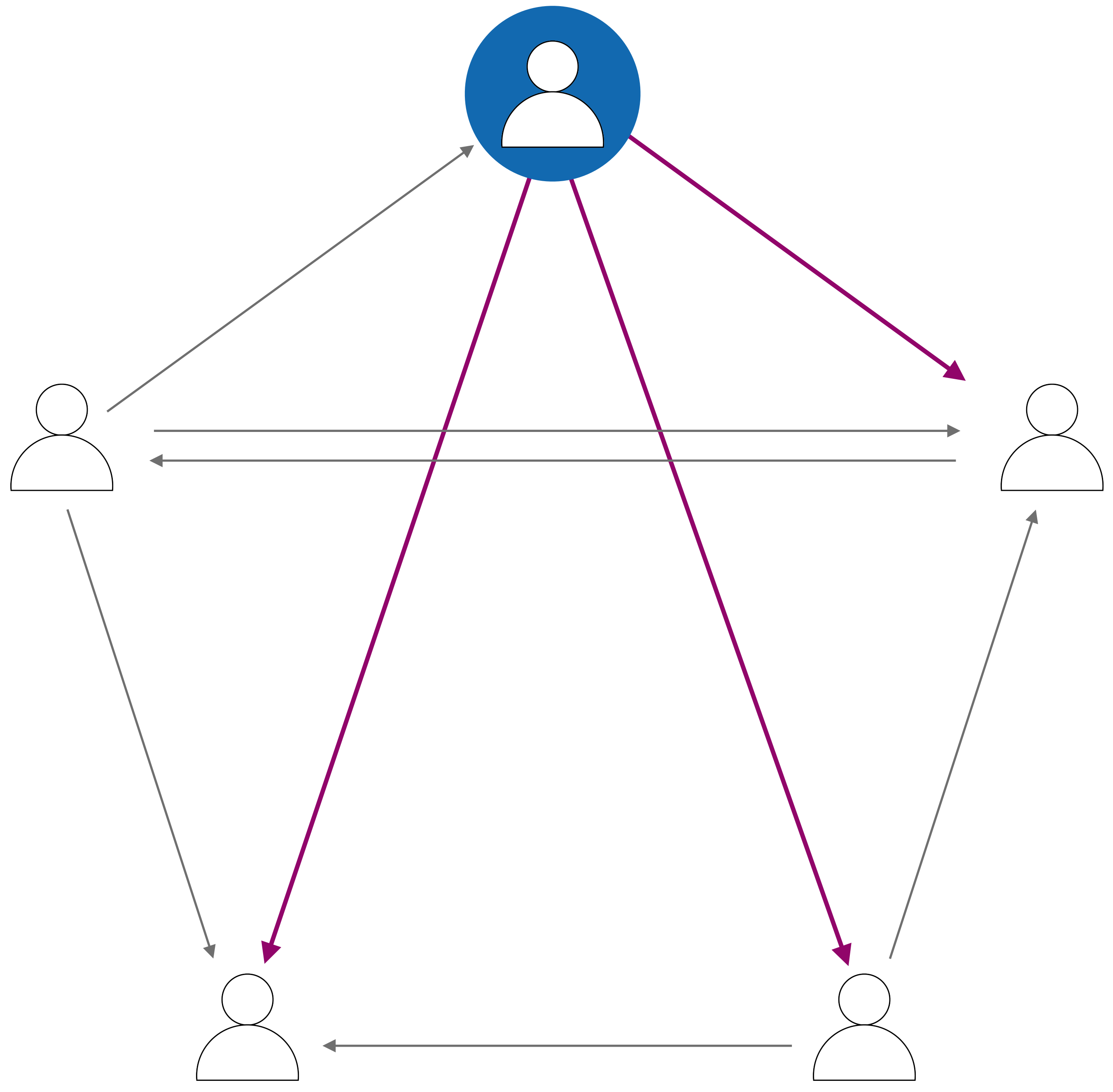
Social Network Formation Model



Directed weighted network $\mathscr{G}$ with $\mathscr{N} := \{1,\ldots,N\}$ agents. The weight $a_{ij} \in [0,1]$ quantifies the importance of the friendship among i and j from i's point of view.

A typical action of agent i:
$a^i = [a_{i1}, \ldots, a_{i,i-1}, a_{i,i+1}, \ldots, a_{iN}] \in \mathscr{A} = [0,1]^{N-1}$.

Every agent i is endowed with a payoff function $V^i\left(a^i, a^{-i}\right)$, representing the network position's benefits (minus costs) to agent i, that he/she tries to maximize (utilitarian principle).



Cost

$$c_i(a^i) := -\sum_k a_{ik}$$

Social Network Formation Model

Directed weighted network $\mathcal{G}$ with $\mathcal{N} := \{1,\ldots,N\}$ agents. The weight $a_{ij} \in [0,1]$ quantifies the importance of the friendship among i and j from i's point of view.

A typical action of agent i:
$a^i = [a_{i1}, \ldots, a_{i,i-1}, a_{i,i+1}, \ldots, a_{iN}] \in \mathcal{A} = [0,1]^{N-1}$.

Every agent i is endowed with a payoff function $V^i\left(a^i, a^{-i}\right)$, representing the network position's benefits (minus costs) to agent i, that he/she tries to maximize (utilitarian principle).

Payoff function

$$V^i(a^i, a^{-i}, \theta^i) := \alpha_i \left(\sum_k a_{ki} + \delta_i \sum_l \sum_k a_{lk} a_{ki} + \delta_i^2 \sum_m \sum_l \sum_k a_{ml} a_{lk} a_{ki} \right) +$$
$$+\beta_i \left(\sum_k a_{ik} \left(\sum_l a_{il} a_{lk} \right) \right) +$$
$$-\gamma_i \left(\sum_k a_{ik} \right).$$

Parameters

$$\theta^i = [\alpha_i, \beta_i, \gamma_i, \delta_i] \in \Theta := \mathbb{R}_{\geq 0} \times \mathbb{R} \times \mathbb{R}_{>0} \times [0,1]$$

Part 1: Model

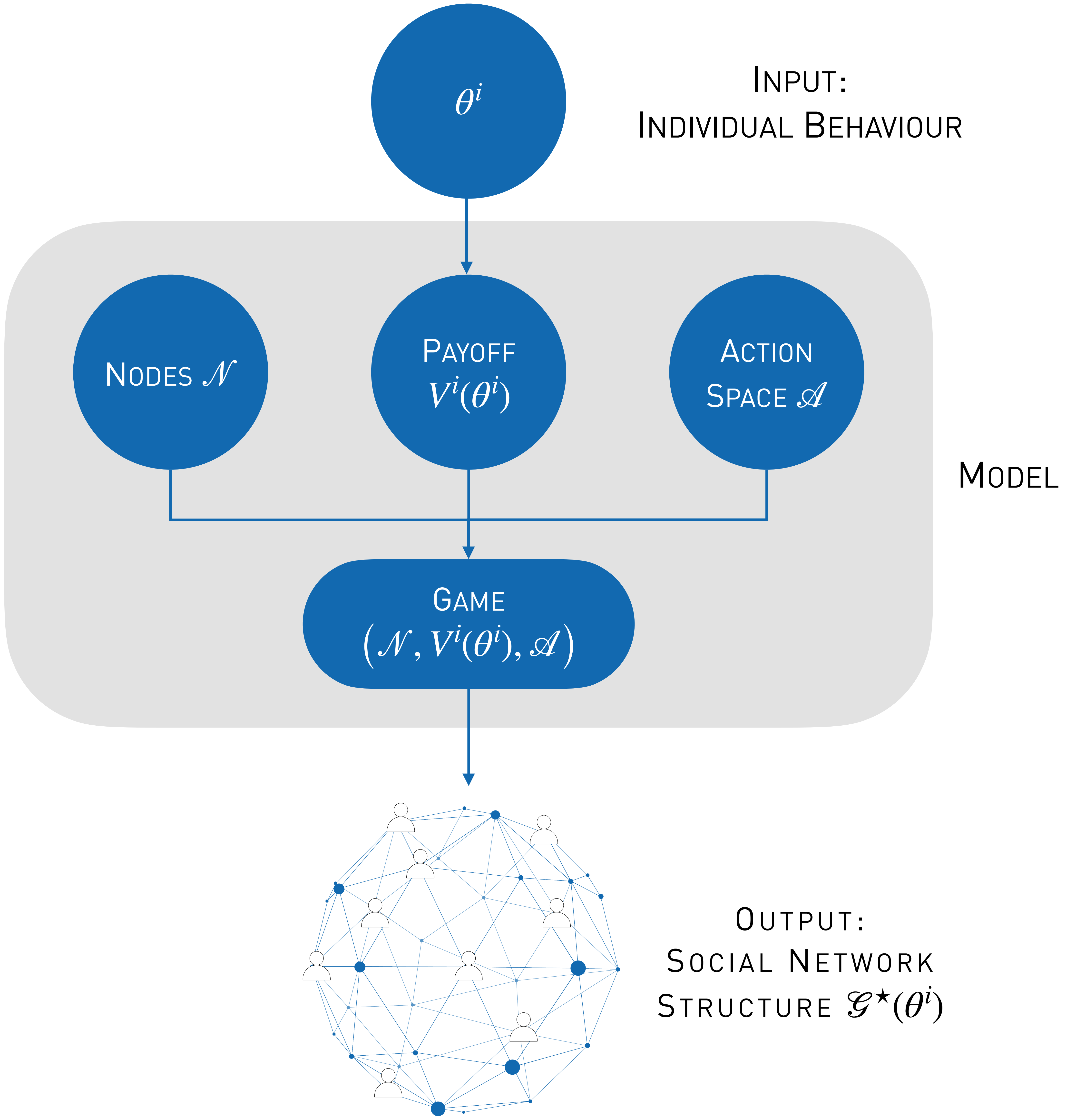
Contributions

1. Model. We propose a parametric network formation model rooted on solid socio-economic theories.

2. Analysis. We perform the theoretical analysis of four network motifs of homogeneous and rational agents.

3. Inference. We construct a game-theoretical inference method based on Nash equilibrium conditions to perform individual behaviour estimation of heterogeneous agents.

4. Validation. We validate our estimation results against sociological and historical observations.



N. Pagan and F. Dörfler, "Game theoretical inference of human behaviour in social networks", Nature Communication, 2019.

Part 1: Results

Contributions

1. Model. We propose a parametric network formation model rooted on solid socio-economic theories.

2. Analysis. We perform the theoretical analysis of four network motifs of homogeneous and rational agents.

3. Inference. We construct a game-theoretical inference method based on Nash equilibrium conditions to perform individual behaviour estimation of heterogeneous agents.

4. Validation. We validate our estimation results against sociological and historical observations.

N. Pagan and F. Dörfler, "Game theoretical inference of human behaviour in social networks", Nature Communication, 2019.

Definition (Nash equilibrium).

The network $\mathscr{G}^\star$ is a Nash equilibrium if for all agents $i \in \mathscr{N}$,

$$V^i\left(a^i, a^{-i,\star}\right) \leq V^i\left(a^{i,\star}, a^{-i,\star}\right), \quad \forall a^i \in \mathscr{A}.$$

Definition (Pairwise-Nash equilibrium).

The network $\mathscr{G}^\star$ is a Pairwise-Nash equilibrium if

1. for all pairs of distinct agents $(i,j) \in \mathscr{N}^2$,

$$V^i\left(a_{ij}, a_{-j}^{i,\star}, a^{-i,\star}\right) \leq V^i\left(a_{ij}^\star, a_{-j}^{i,\star}, a^{-i,\star}\right), \forall a_{ij} \in [0,1],$$
$$V^j\left(a_{ji}, a_{-i}^{j,\star}, a^{-j,\star}\right) \leq V^j\left(a_{ji}^\star, a_{-i}^{j,\star}, a^{-j,\star}\right), \forall a_{ji} \in [0,1],$$

2. for all pairs of distinct agent $(i,j) \in \mathscr{N}^2$ and for all pairs $\left(a_{ij}, a_{ji}\right) \in [0,1]^2$,

$$V^i\left(a_{ij}, a_{ji}, a^{-(i,j),\star}\right) > V^i\left(a_{ij}^\star, a_{ji}^\star, a^{-(i,j),\star}\right)$$
$$\Downarrow$$
$$V^j\left(a_{ij}, a_{ji}, a^{-(i,j),\star}\right) < V^j\left(a_{ij}^\star, a_{ji}^\star, a^{-(i,j),\star}\right).$$

Contributions

1. Model. We propose a parametric network formation model rooted on solid socio-economic theories.

2. Analysis. We perform the theoretical analysis of four network motifs of homogeneous and rational agents.

3. Inference. We construct a game-theoretical inference method based on Nash equilibrium conditions to perform individual behaviour estimation of heterogeneous agents.

4. Validation. We validate our estimation results against sociological and historical observations.

N. Pagan and F. Dörfler, "Game theoretical inference of human behaviour in social networks", Nature Communication, 2019.

Theorem (Empty Network).

Let $\mathscr{G}^{EN}$ be a complete network of $N \geq 3$ agents.

Then

- $\mathscr{G}^{EN}$ is always a Nash equilibrium,
- $\mathscr{G}^{EN}$ is a Pairwise Nash equilibrium if and only if $\gamma \geq \alpha\left(1+\delta+\delta^2\right)$.

Contributions

1. Model. We propose a parametric network formation model rooted on solid socio-economic theories.

2. Analysis. We perform the theoretical analysis of four network motifs of homogeneous and rational agents.

3. Inference. We construct a game-theoretical inference method based on Nash equilibrium conditions to perform individual behaviour estimation of heterogeneous agents.

4. Validation. We validate our estimation results against sociological and historical observations.

N. Pagan and F. Dörfler, "Game theoretical inference of human behaviour in social networks", Nature Communication, 2019.

Theorem (Empty Network).

Let $\mathscr{G}^{EN}$ be a complete network of $N \geq 3$ agents.

Then

- $\mathscr{G}^{EN}$ is always a Nash equilibrium,
- $\mathscr{G}^{EN}$ is a Pairwise-Nash equilibrium if and only if $\gamma \geq \alpha\left(1+\delta+\delta^2\right)$.

Theorem (Complete Network).

Let $\mathscr{G}^{CN}$ be a complete network of $N \geq 3$ agents.
Define

$$\bar{\gamma}_{NE} := \begin{cases} \alpha\delta\left(1+\delta(2N-3)\right)+\beta(N-2), & \text{if } \beta > 0 \\ \alpha\delta\left(1+\delta(2N-3)\right)+2\beta(N-2), & \text{if } \beta \leq 0, \end{cases}$$

$$\bar{\gamma}_{PNE} := \alpha\delta\left(1+\delta(2N-3)\right)+2\beta(N-2),$$

then

- $\mathscr{G}^{CN}$ is a Nash equilibrium if and only if $\gamma \leq \bar{\gamma}_{NE}$,
- $\mathscr{G}^{CN}$ is a Pairwise-Nash equilibrium if and only if $\gamma \leq \bar{\gamma}_{PNE}$.

Contributions

1. Model. We propose a parametric network formation model rooted on solid socio-economic theories.

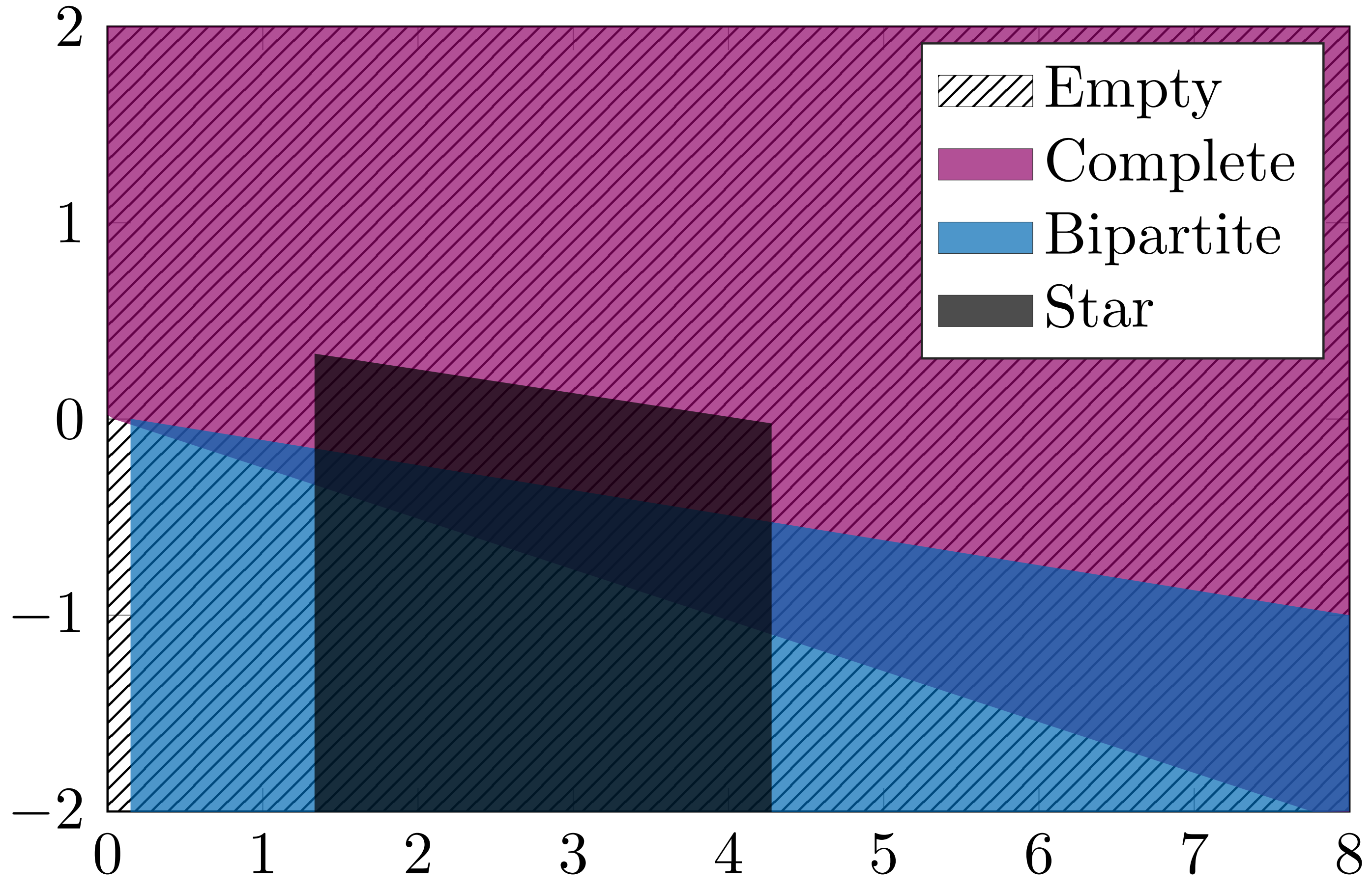
2. Analysis. We perform the theoretical analysis of four network motifs of homogeneous and rational agents.

3. Inference. We construct a game-theoretical inference method based on Nash equilibrium conditions to perform individual behaviour estimation of heterogeneous agents.

4. Validation. We validate our estimation results against sociological and historical observations.

N. Pagan and F. Dörfler, "Game theoretical inference of human behaviour in social networks", Nature Communication, 2019.

Social Support



Social Influence

Brokerage

Fig. Nash equilibrium phase diagram.

Part 1: Results

Contributions

1. Model. We propose a parametric network formation model rooted on solid socio-economic theories.

2. Analysis. We perform the theoretical analysis of four network motifs of homogeneous and rational agents.

3. Inference. We construct a game-theoretical inference method based on Nash equilibrium conditions to perform individual behaviour estimation of heterogeneous agents.

4. Validation. We validate our estimation results against sociological and historical observations.

N. Pagan and F. Dörfler, "Game theoretical inference of human behaviour in social networks", Nature Communication, 2019.

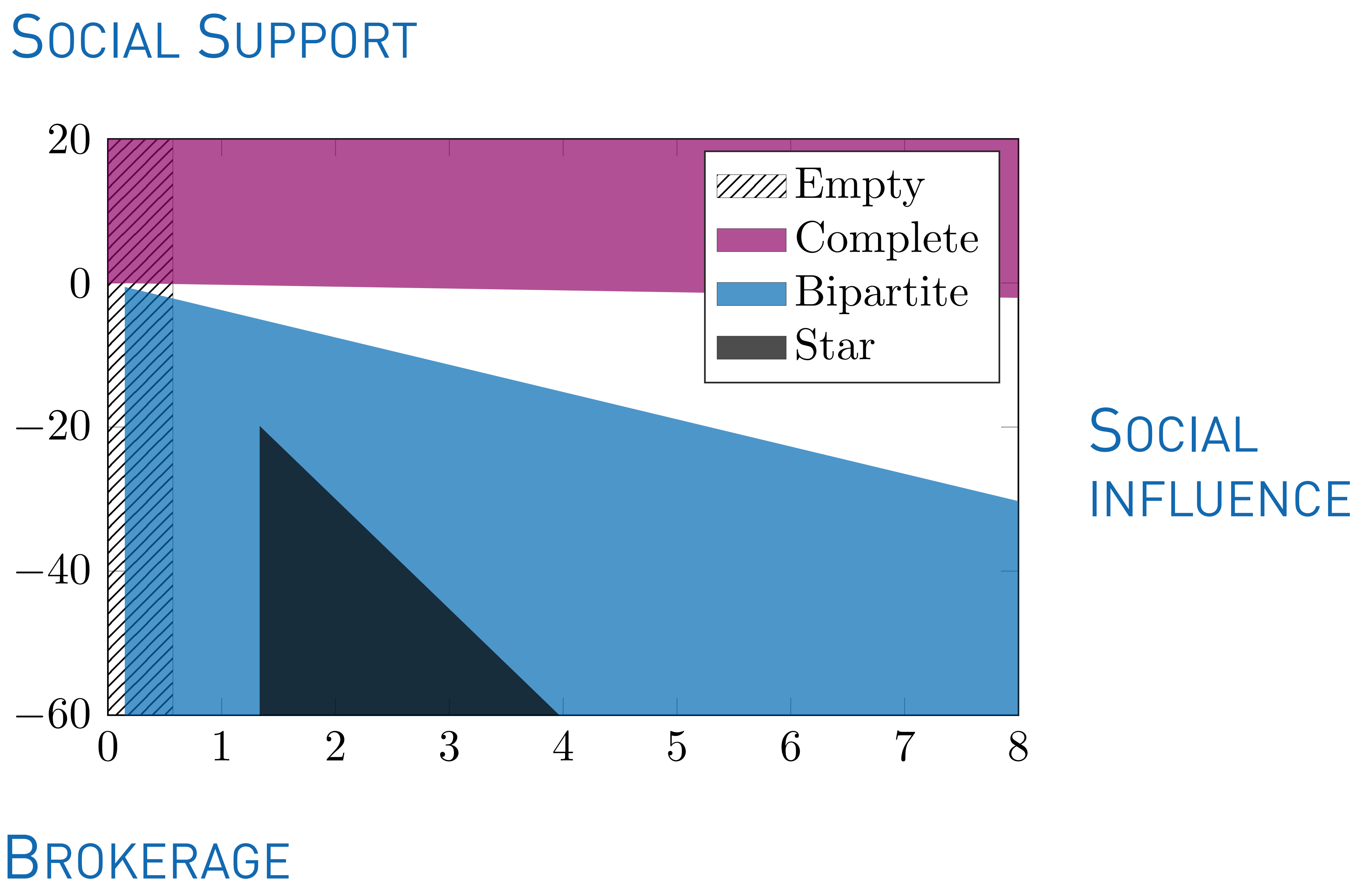


Fig. Pairwise-Nash equilibrium phase diagram.

Contributions

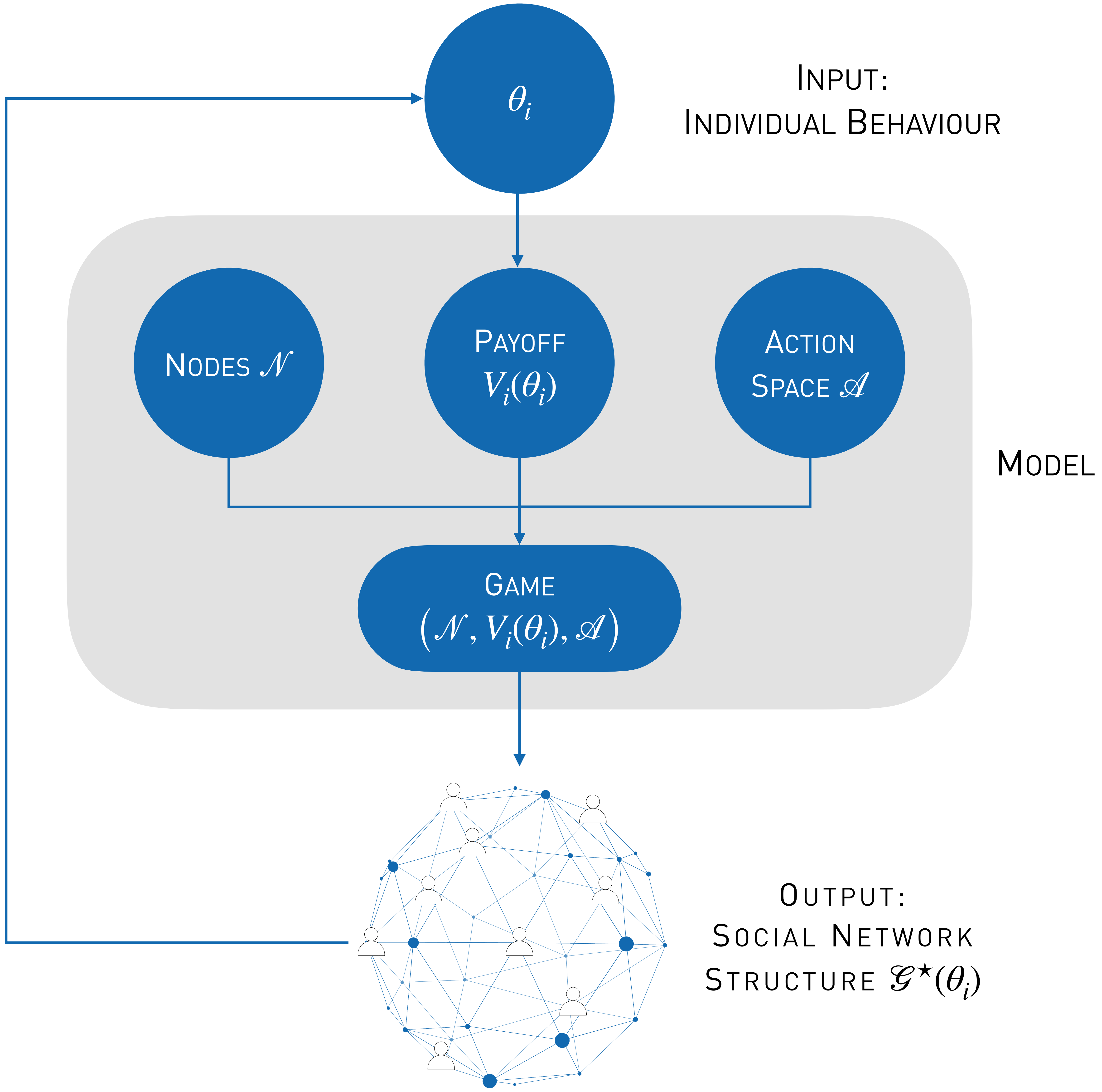
1. Model. We propose a parametric network formation model rooted on solid socio-economic theories.

2. Analysis. We perform the theoretical analysis of four network motifs of homogeneous and rational agents.

3. Inference. We construct a game-theoretical inference method based on Nash equilibrium conditions to perform individual behaviour estimation of heterogeneous agents.

4. Validation. We validate our estimation results against sociological and historical observations.

N. Pagan and F. Dörfler, "Game theoretical inference of human behaviour in social networks", Nature Communication, 2019.



Contributions

1. Model. We propose a parametric network formation model rooted on solid socio-economic theories.
2. Analysis. We perform the theoretical analysis of four network motifs of homogeneous and rational agents.
3. Inference. We construct a game-theoretical inference method based on Nash equilibrium conditions to perform individual behaviour estimation of heterogeneous agents.
4. Validation. We validate our estimation results against sociological and historical observations.

N. Pagan and F. Dörfler, "Game theoretical inference of human behaviour in social networks", Nature Communication, 2019.

Problem [Discrete Minimum NE-Distance Problem].

Given a network $\mathcal{G}^\star$ of N agents, for all agents i find the vectors of preferences $\hat{\theta}_i$ such that

$$\hat{\theta}_i \in \arg\min_{\theta_i \in \Theta} \|\mathbf{e}_i^+(\theta_i)\|^2,$$

where $e_i^+ = \max\{0, \mathbf{X_i}\theta - y\}$ is the vector of positive errors. Then

$$\hat{\theta}_i = \left(\mathbf{X}_i^T(\hat{\theta}_i)\mathbf{X}_i(\hat{\theta}_i)\right)^{-1}\mathbf{X}_i^T(\hat{\theta}_i)\mathbf{y}_i(\hat{\theta}_i).$$

Problem [Ordinary Least Square Regression].

Given a set of n samples $\{x_i\}_{i=1}^n$ and n observations $\{y_i\}_{i=1}^n$, where each scalar y_i is the response to the row vector x_i of values of p predictors (regressors) x_{ij} for $j = 1,\ldots,p$, find $\hat{\theta} \in \mathbb{R}^p$ such that

$$\hat{\theta} = \arg\min_{\theta} \|e(\theta)\|_2^2,$$

where $e = \mathbf{X}\theta - y$ is the vector of residuals. Then

$$\hat{\theta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}.$$

Contributions

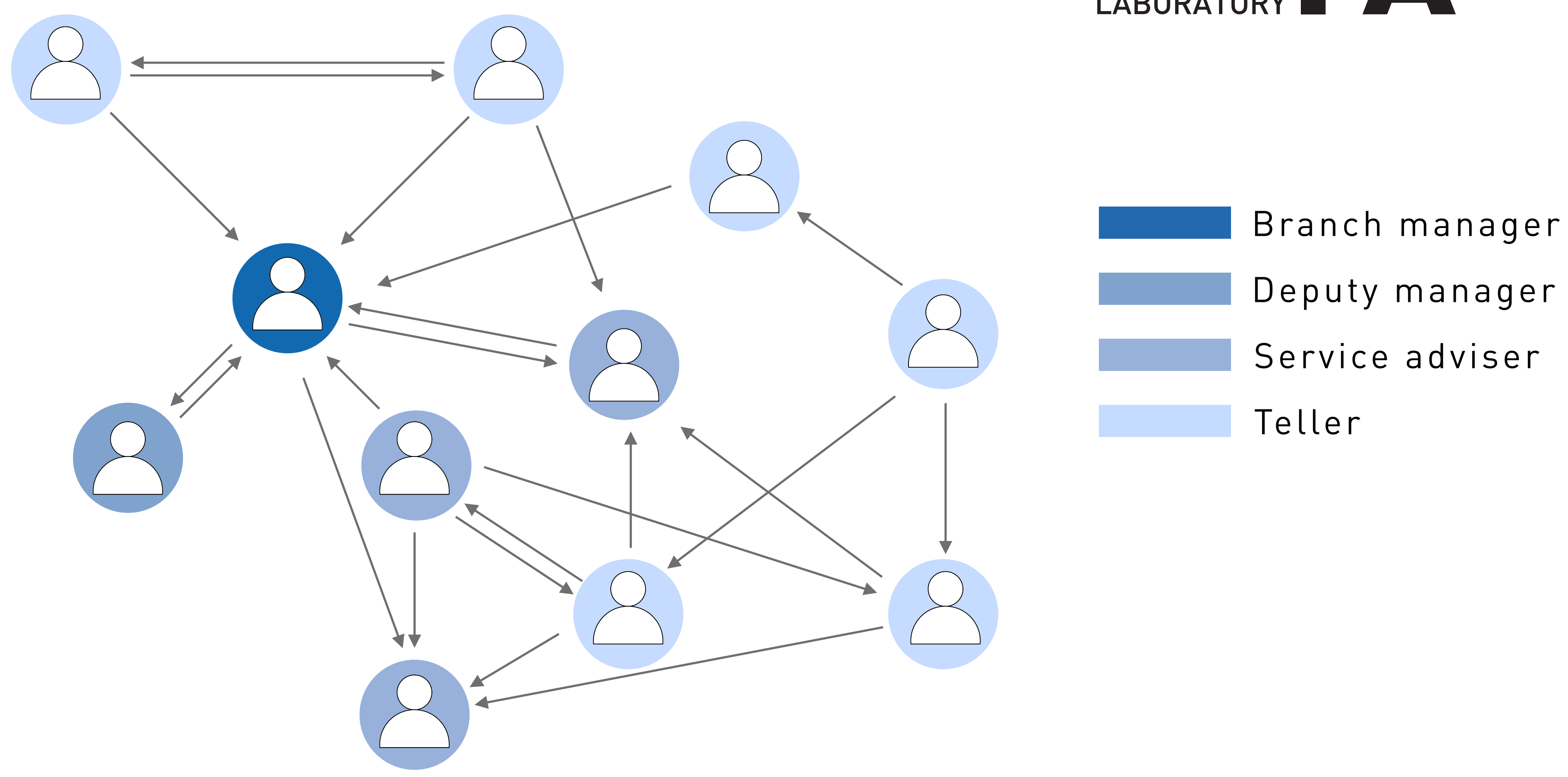
1. Model. We propose a parametric network formation model rooted on solid socio-economic theories.

2. Analysis. We perform the theoretical analysis of four network motifs of homogeneous and rational agents.

3. Inference. We construct a game-theoretical inference method based on Nash equilibrium conditions to perform individual behaviour estimation of heterogeneous agents.

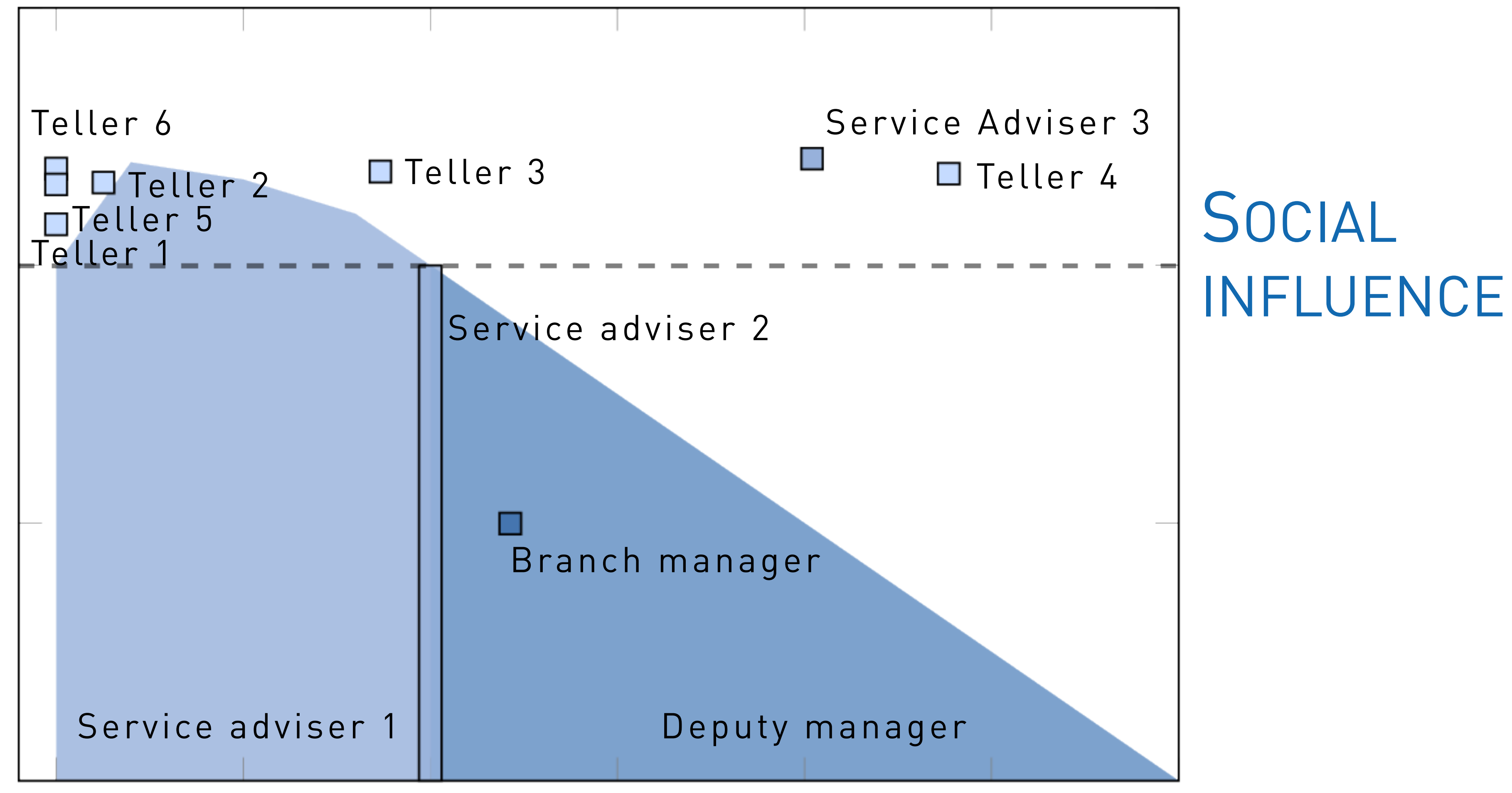
4. Validation. We validate our estimation results against sociological and historical observations.

N. Pagan and F. Dörfler, "Game theoretical inference of human behaviour in social networks", Nature Communication, 2019.



[Pattison, P., Wasserman, S., Robins, G. & Kanfer, A. M. Statistical evaluation of algebraic constraints for social networks. J. Math. Psychology 44, (2000)]

Social Support



Part 1: Results

PART 2.
QUALITY-BASED NETWORK FORMATION MODEL

UGC-BASED PLATFORMS

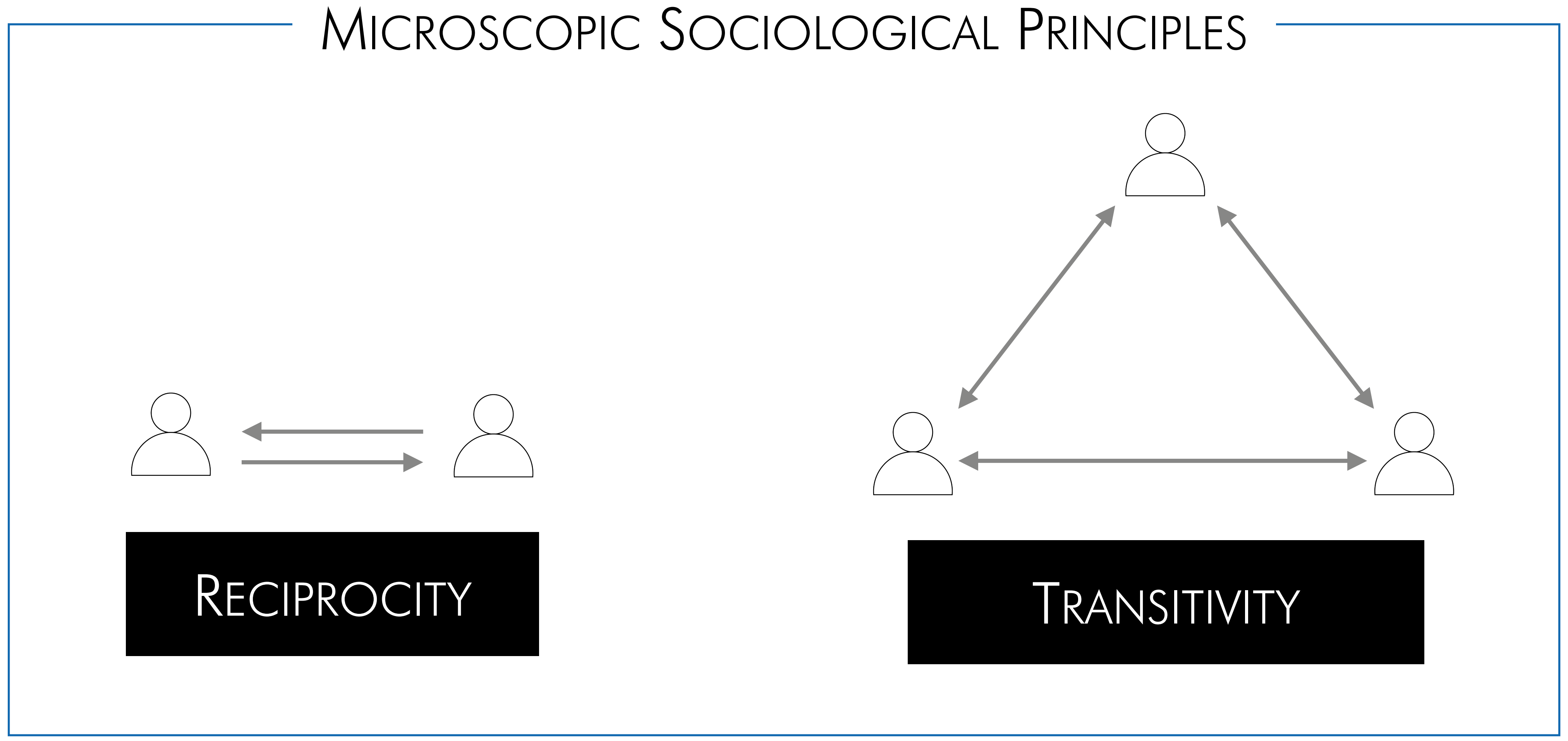
- Online social networks have become predominant in last couple of decades. Billions of people are virtually connected.
- On many of today's most popular platforms, users can contribute with their User-Generated Content (UGC): Tweets, Instagram photos, Youtube videos, Twitch video-streaming, …
- Through the integrated search engines, users can search for their preferred content.

TikTok

Feb 2003 — May 2004 — Feb 2005 — March 2006 — Oct 2010 — Jan 2011 — June 2011 — Sep 2016 →

MERITOCRATIC PRINCIPLE

- Unlike Facebook or LinkedIn, on Twitter, Instagram, YouTube, TikTok, or Twitch, users can follow real-life strangers, e.g., famous people. Hence, connections are not necessarily driven by sociological principles, e.g., reciprocity or transitivity.



RECIPROCITY	100%	22%	14%
TRANSITIVITY	40%	23%	8%

[S. Teng, M. Yeh and K. Chuang, "Toward understanding the mobile social properties: An analysis on instagram photo-sharing network", 2015]

MERITOCRATIC PRINCIPLE

- Unlike Facebook or LinkedIn, on Twitter, Instagram, YouTube, TikTok, or Twitch, users can follow real-life strangers, e.g., famous people. Hence, connections are not necessarily driven by sociological principles, e.g., reciprocity or transitivity.

- Rather, they are driven by interest similarity (homophily) and by the quality of the UGC (meritocratic principle).

Quality-Based Network Formation Model

- Each agent i is attributed with a quality parameter $q_i \in [0,1]$ describing the likelihood her content will be liked by others.

Quality-Based Network Formation Model

- Each agent i is attributed with a quality parameter $q_i \in [0,1]$ describing the likelihood her content will be liked by others.

- Each agent i sequentially meets another distinct agent j (uniformly at random) and decides whether to start following j or not according to:

$$a_{ij}(t+1) = \begin{cases} 1, & \text{if } q_j > V_i(t) := \max_{j \in \mathcal{F}^{i,\text{in}}(t)} q_j, \\ a_{ij}(t), & \text{otherwise,} \end{cases}$$

where $\mathcal{F}^{i,\text{in}}(t) = \left\{ j, \text{ s.t. } a_{ij}(t) = 1 \right\}$ denotes the set of i's followees.

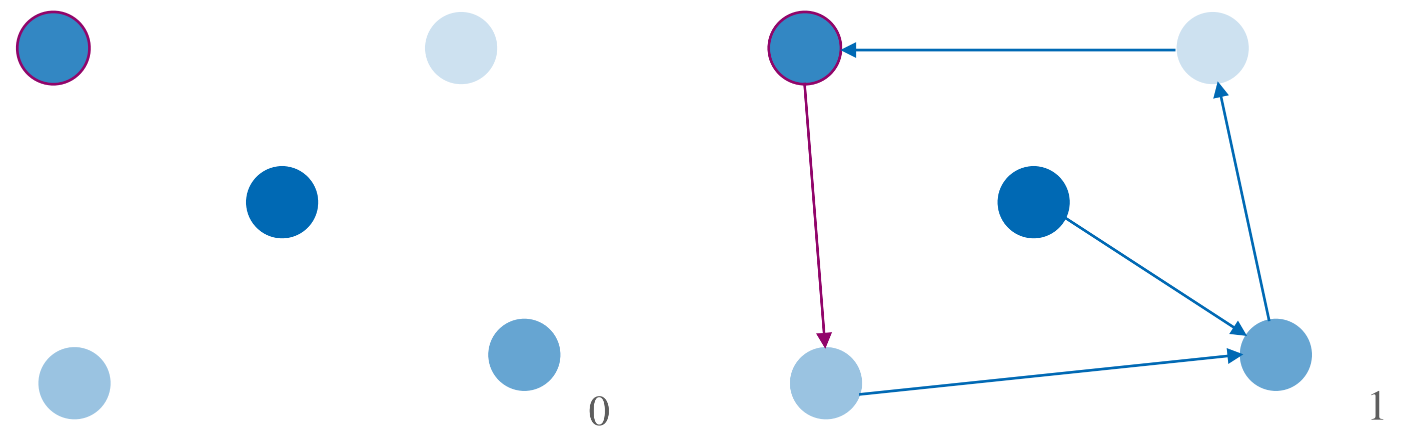
Quality-Based Network Formation Model

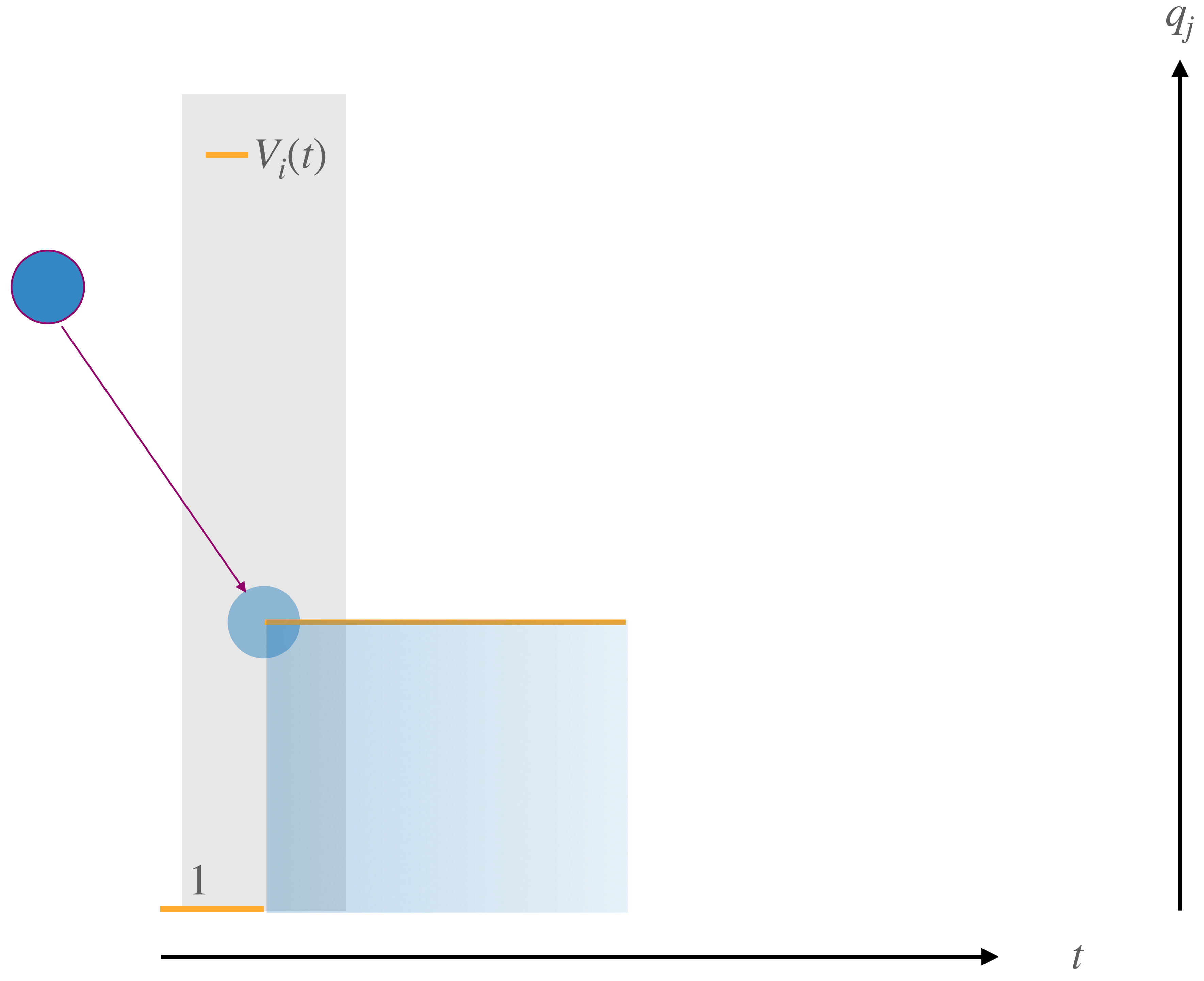
- Each agent i is attributed with a quality parameter $q_i \in [0,1]$ describing the likelihood her content will be liked by others.

- Each agent i sequentially meets another distinct agent j (uniformly at random) and decides whether to start following j or not according to:

$$a_{ij}(t+1) = \begin{cases} 1, & \text{if } q_j > V_i(t) := \max_{j \in \mathcal{F}^{i,\text{in}}(t)} q_j, \\ a_{ij}(t), & \text{otherwise,} \end{cases}$$

where $\mathcal{F}^{i,\text{in}}(t) = \left\{ j, \text{ s.t. } a_{ij}(t) = 1 \right\}$ denotes the set of i's followees.





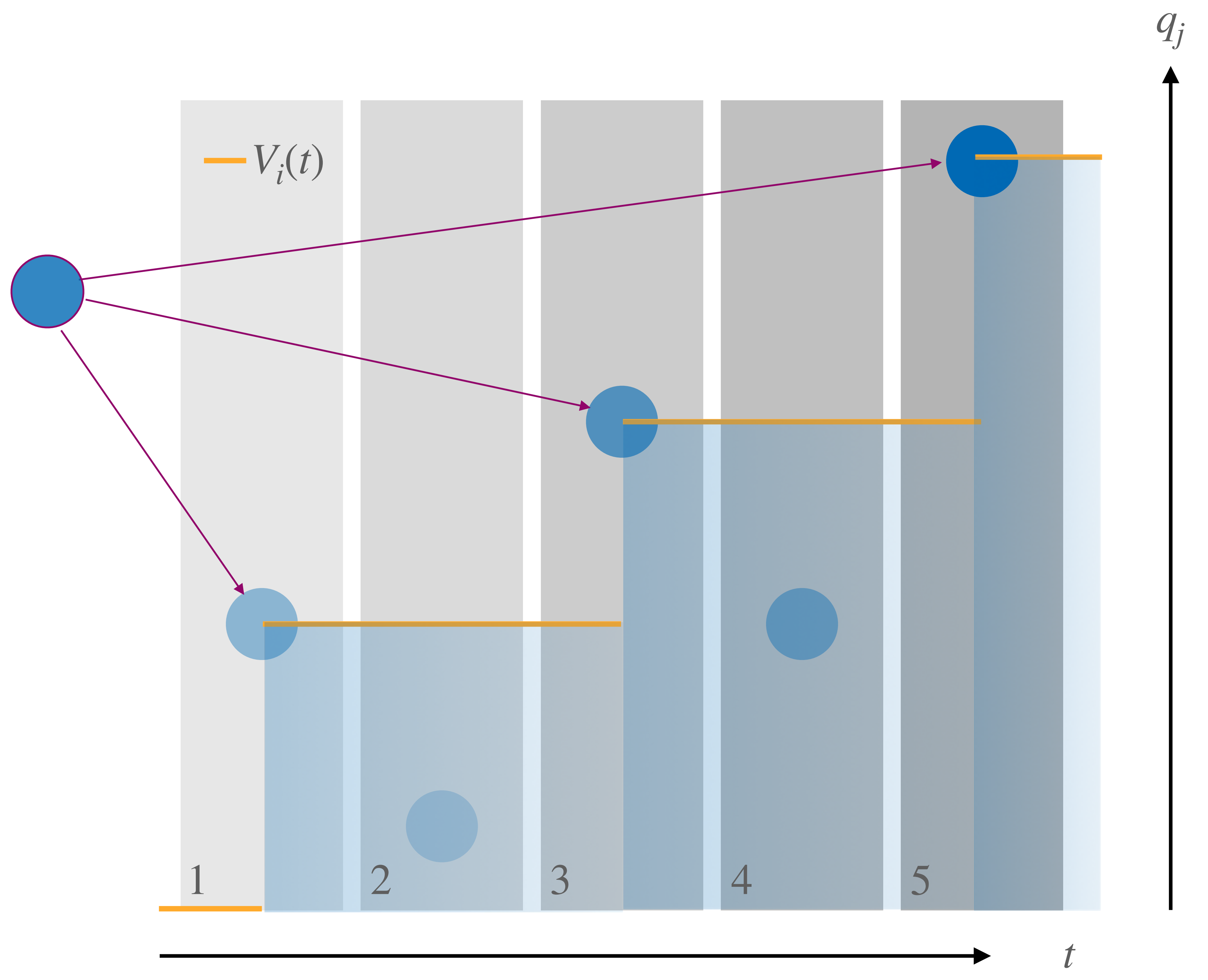
Quality-Based Network Formation Model

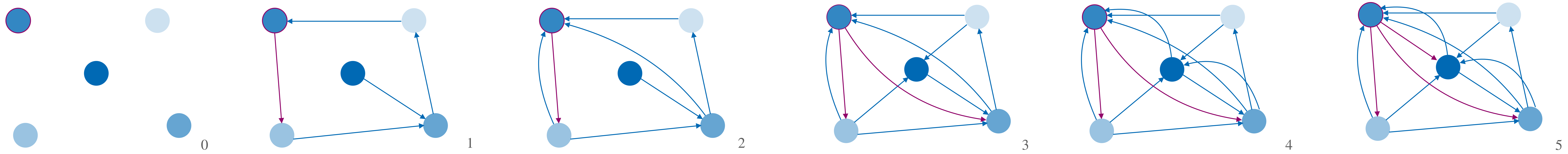
- Each agent i is attributed with a quality parameter $q_i \in [0,1]$ describing the likelihood her content will be liked by others.

- Each agent i sequentially meets another distinct agent j (uniformly at random) and decides whether to start following j or not according to:

$$a_{ij}(t+1) = \begin{cases} 1, & \text{if } q_j > V_i(t) := \max_{j \in \mathcal{F}^{i,\text{in}}(t)} q_j, \\ a_{ij}(t), & \text{otherwise,} \end{cases}$$

where $\mathcal{F}^{i,\text{in}}(t) = \left\{ j, \text{ s.t. } a_{ij}(t) = 1 \right\}$ denotes the set of i's followees.





Contributions

1. Inference. We collect and analyze a longitudinal Twitter data-set that allows to support a meritocratic principle.

2. Model. We propose network formation process based on the quality of the UGC and on the meritocratic principle.

3. Analysis. We analysed the properties of the resulting network in terms of indegree and outdegree distributions, as well as with a newly defined audience overlap index.

4. Validation. We validate our estimation results against data-sets collected on Twitch.

N. Pagan, W. Mei, C. Li, and F. Dörfler, "A meritocratic network formation model for the rise of social media influencers", submitted, 2021.

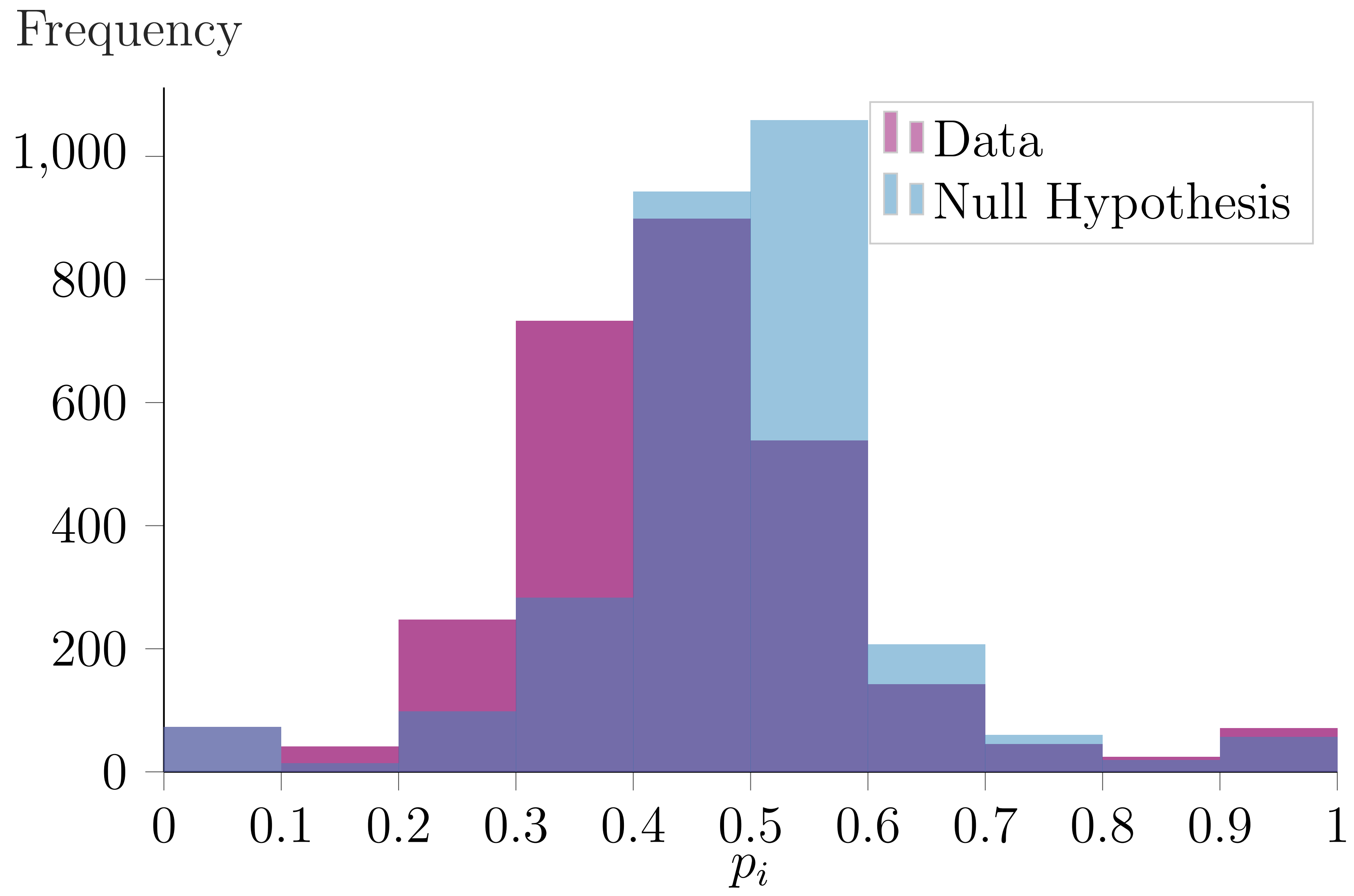


Fig. We compare the ordered temporal sequence (purple) of outgoing connections in a Twitter data-set on a network of social network scientists with a random temporal sequence (blue). The difference between the distributions is statistical significant.

Contributions

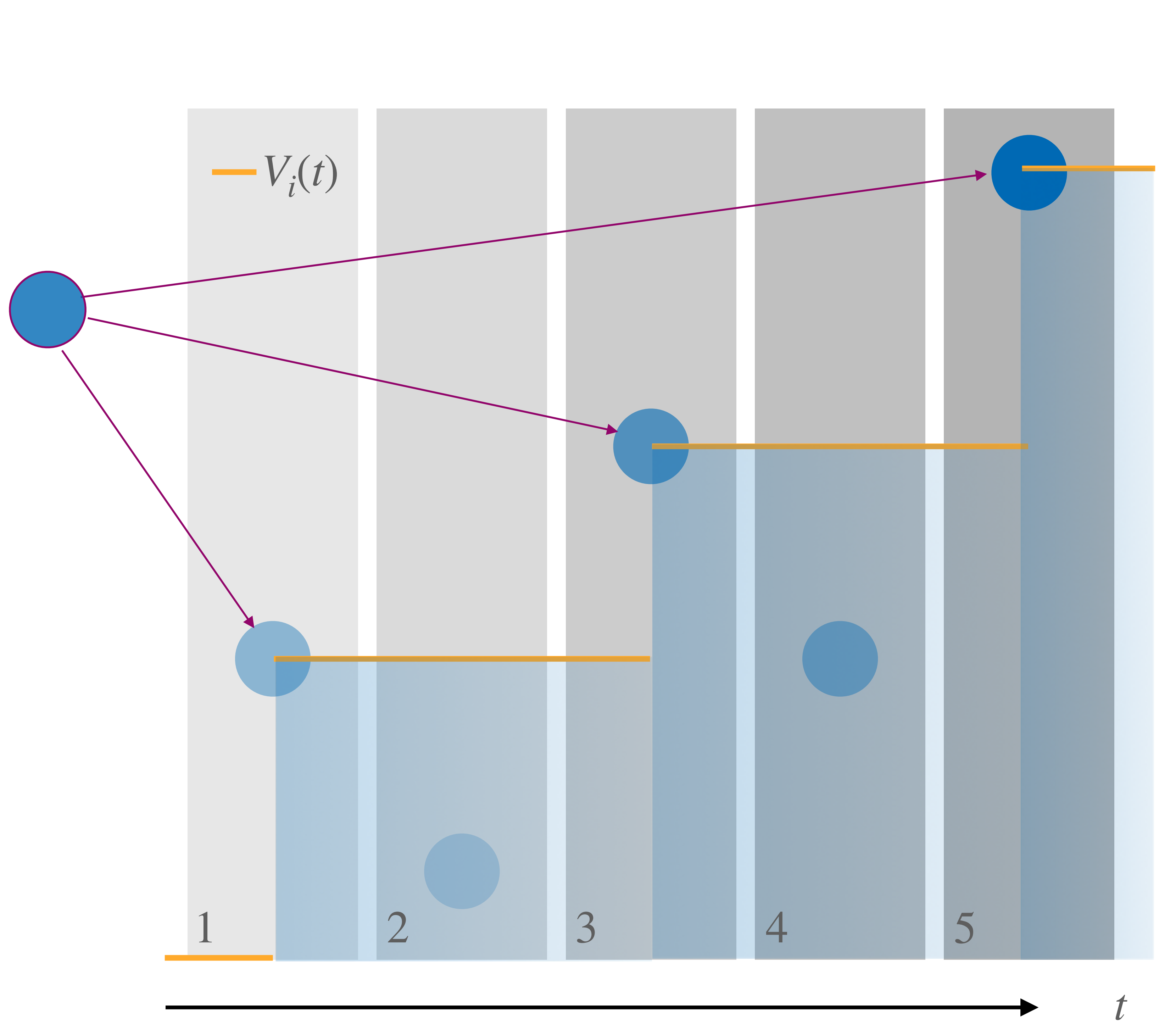
1. Inference. We collect and analyze a longitudinal Twitter data-set that allows to infer a meritocratic principle.

2. Model. We propose network formation process based on the quality of the UGC and on the meritocratic principle.

3. Analysis. We analysed the properties of the resulting network in terms of indegree and outdegree distributions, as well as with a newly defined audience overlap index.

4. Validation. We validate our estimation results against data-sets collected on Twitch.

N. Pagan, W. Mei, C. Li, and F. Dörfler, "A meritocratic network formation model for the rise of social media influencers", submitted, 2021.



Contributions

1. Inference. We collect and analyze a longitudinal Twitter data-set that allows to infer a meritocratic principle.

2. Model. We propose network formation process based on the quality of the UGC and on the meritocratic principle.

3. Analysis. We analysed the properties of the resulting network in terms of indegree and outdegree distributions, as well as with a newly defined audience overlap index.

4. Validation. We validate our estimation results against data-sets collected on Twitch.

N. Pagan, W. Mei, C. Li, and F. Dörfler, "A meritocratic network formation model for the rise of social media influencers", submitted, 2021.

Theorem (Convergence).

For any set of qualities $\{q_1, \ldots, q_N\}$, the network reaches an equilibrium almost surely.

Contributions

1. Inference. We collect and analyze a longitudinal Twitter data-set that allows to infer a meritocratic principle.

2. Model. We propose network formation process based on the quality of the UGC and on the meritocratic principle.

3. Analysis. We analysed the properties of the resulting network in terms of indegree and outdegree distributions, as well as with a newly defined audience overlap index.

4. Validation. We validate our estimation results against data-sets collected on Twitch.

N. Pagan, W. Mei, C. Li, and F. Dörfler, "A meritocratic network formation model for the rise of social media influencers", submitted, 2021.

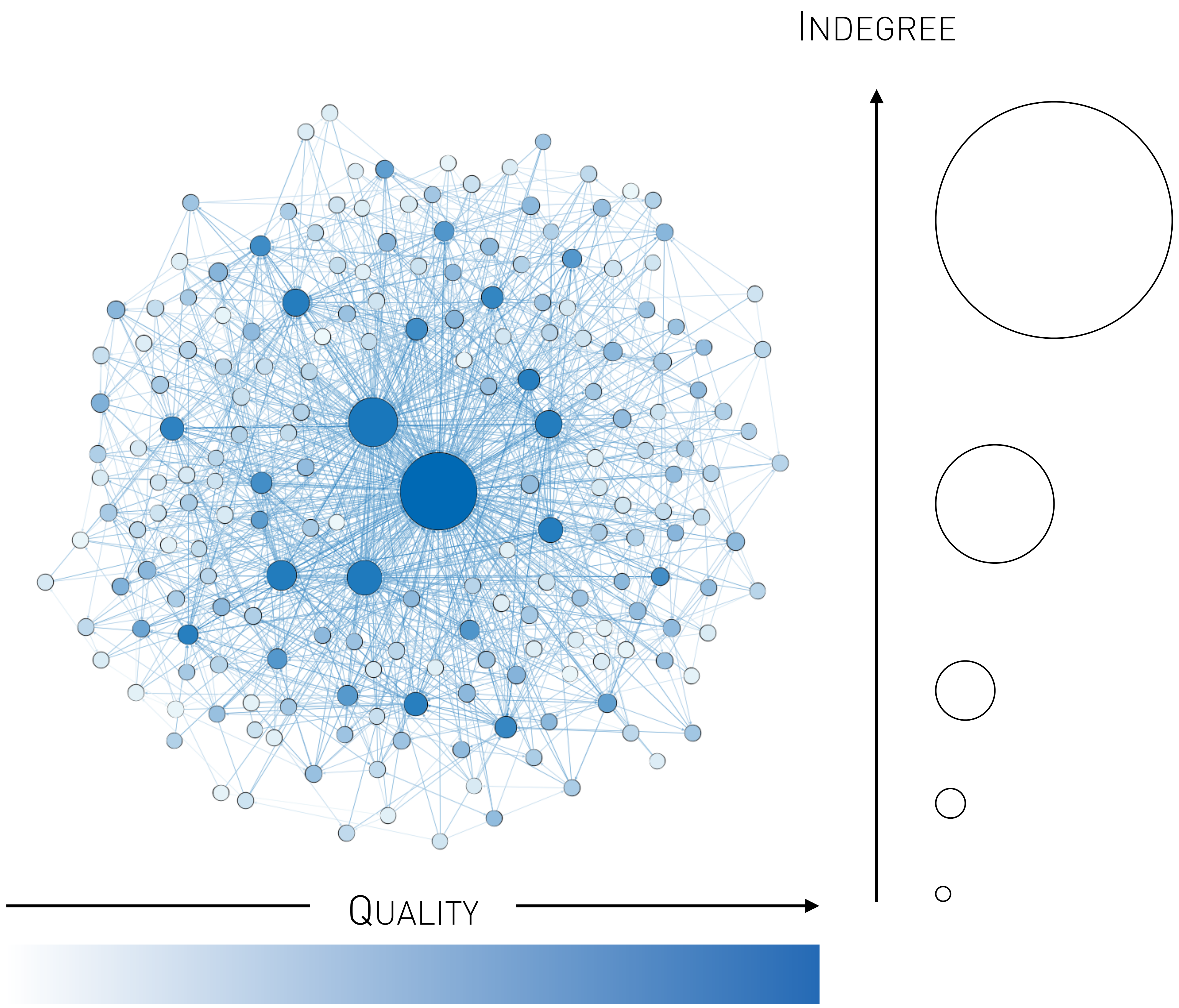


Fig. Numerical simulation results of a network of N=200 agents until equilibrium.

Contributions

1. Inference. We collect and analyze a longitudinal Twitter data-set that allows to infer a meritocratic principle.

2. Model. We propose network formation process based on the quality of the UGC and on the meritocratic principle.

3. Analysis. We analysed the properties of the resulting network in terms of indegree and outdegree distributions, as well as with a newly defined audience overlap index.

4. Validation. We validate our estimation results against data-sets collected on Twitch.

N. Pagan, W. Mei, C. Li, and F. Dörfler, "A meritocratic network formation model for the rise of social media influencers", submitted, 2021.

Theorem (Indegree Distribution).

At equilibrium, the expected indegree follows a Zipf's law:
$$\mathbb{E}(d_i^{in,\star}) = \begin{cases} N-1, & \text{if } i = 1, \\ \frac{N}{i}, & \text{otherwise.} \end{cases}$$

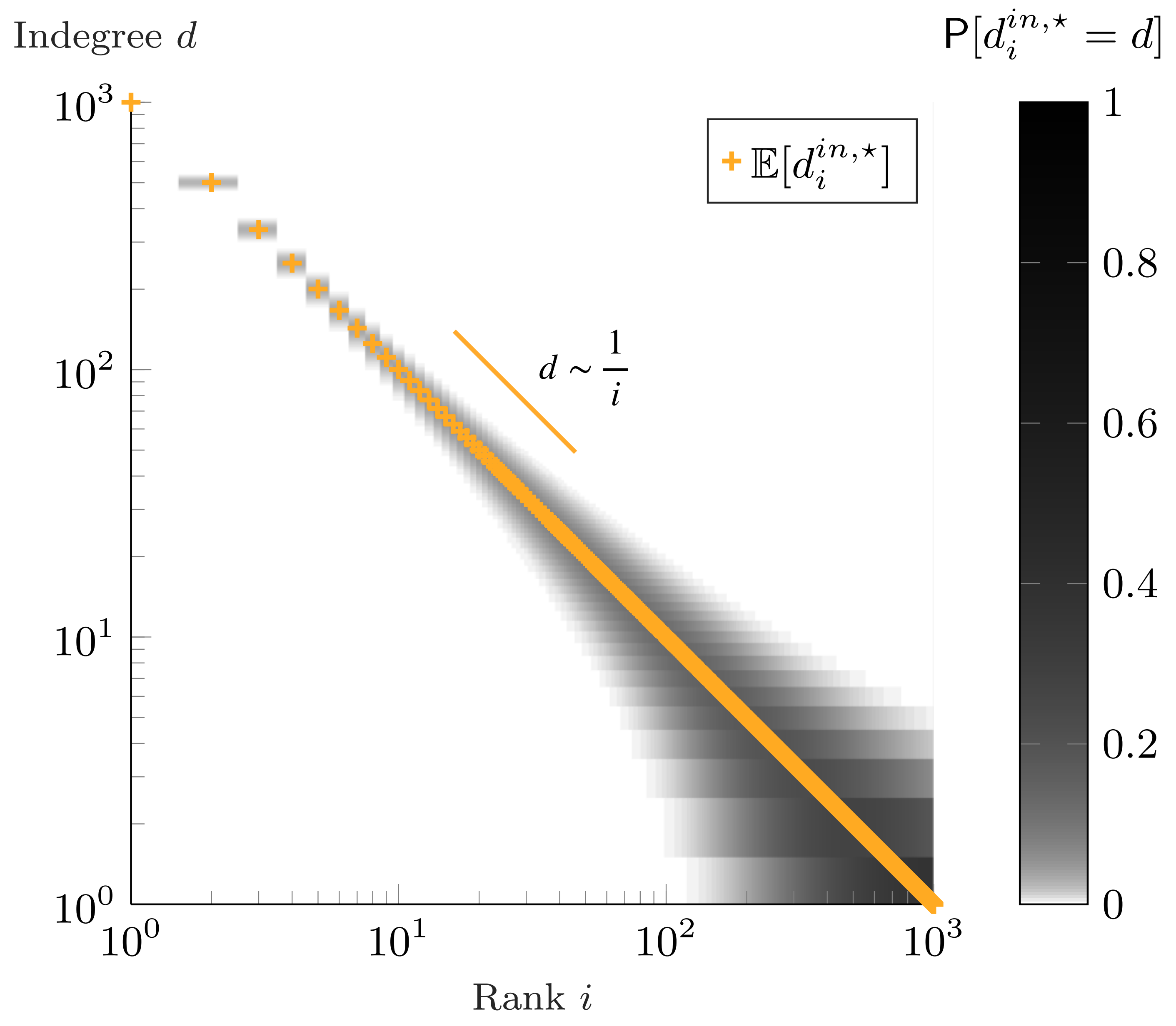


Fig. Theoretical probability density function of the nodes' indegree in a network of N=1000 agents.

CONTRIBUTIONS

1. Inference. We collect and analyze a longitudinal Twitter data-set that allows to infer a meritocratic principle.

2. Model. We propose network formation process based on the quality of the UGC and on the meritocratic principle.

3. Analysis. We analysed the properties of the resulting network in terms of indegree and outdegree distributions, as well as with a newly defined audience overlap index.

4. Validation. We validate our estimation results against data-sets collected on Twitch.

N. Pagan, W. Mei, C. Li, and F. Dörfler, "A meritocratic network formation model for the rise of social media influencers", submitted, 2021.

THEOREM (OUTDEGREE DISTRIBUTION).

The nodes' expected outdegree $\mathbb{E}\left(d_N^{\text{out},\star}\right)$ in a network of $N \geq 2$ agents equals the $(N-1)$-th harmonic number:

$$\mathbb{E}\left(d_N^{\text{out},\star}\right) = \sum_{k=1}^{N-1} \frac{1}{k}.$$

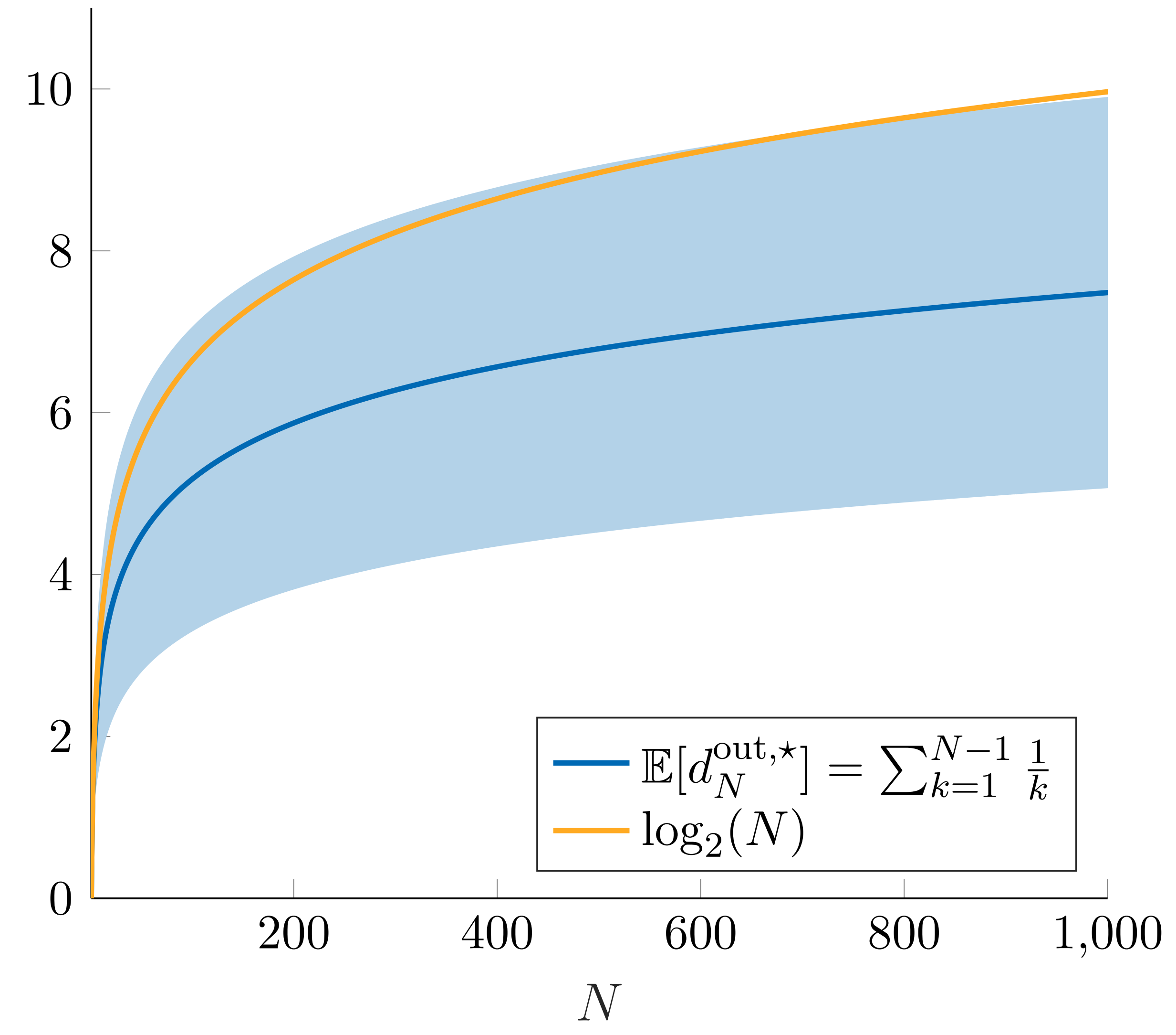


Fig. Theoretical expected outdegree as a function of the network size.

Contributions

1. Inference. We collect and analyze a longitudinal Twitter data-set that allows to infer a meritocratic principle.

2. Model. We propose network formation process based on the quality of the UGC and on the meritocratic principle.

3. Analysis. We analysed the properties of the resulting network in terms of indegree and outdegree distributions, as well as with a newly defined audience overlap index.

4. Validation. We validate our model results against data-sets collected on Twitch.

N. Pagan, W. Mei, C. Li, and F. Dörfler, "A meritocratic network formation model for the rise of social media influencers", submitted, 2021.

Definition (Audience Overlap).

We define the Audience Overlap index as follows:

$$O(i,j) := \frac{\left|\mathcal{F}^{i,\text{in}} \cap \mathcal{F}^{j,\text{in}}\right|}{\left|\mathcal{F}^{i,\text{in}}\right|} \in [0,1], \quad \text{if } \left|\mathcal{F}^{i,\text{in}}\right| > 0,$$

and 0 otherwise, where $\mathcal{F}^{i,\text{in}}$ denotes the set of followers of i.

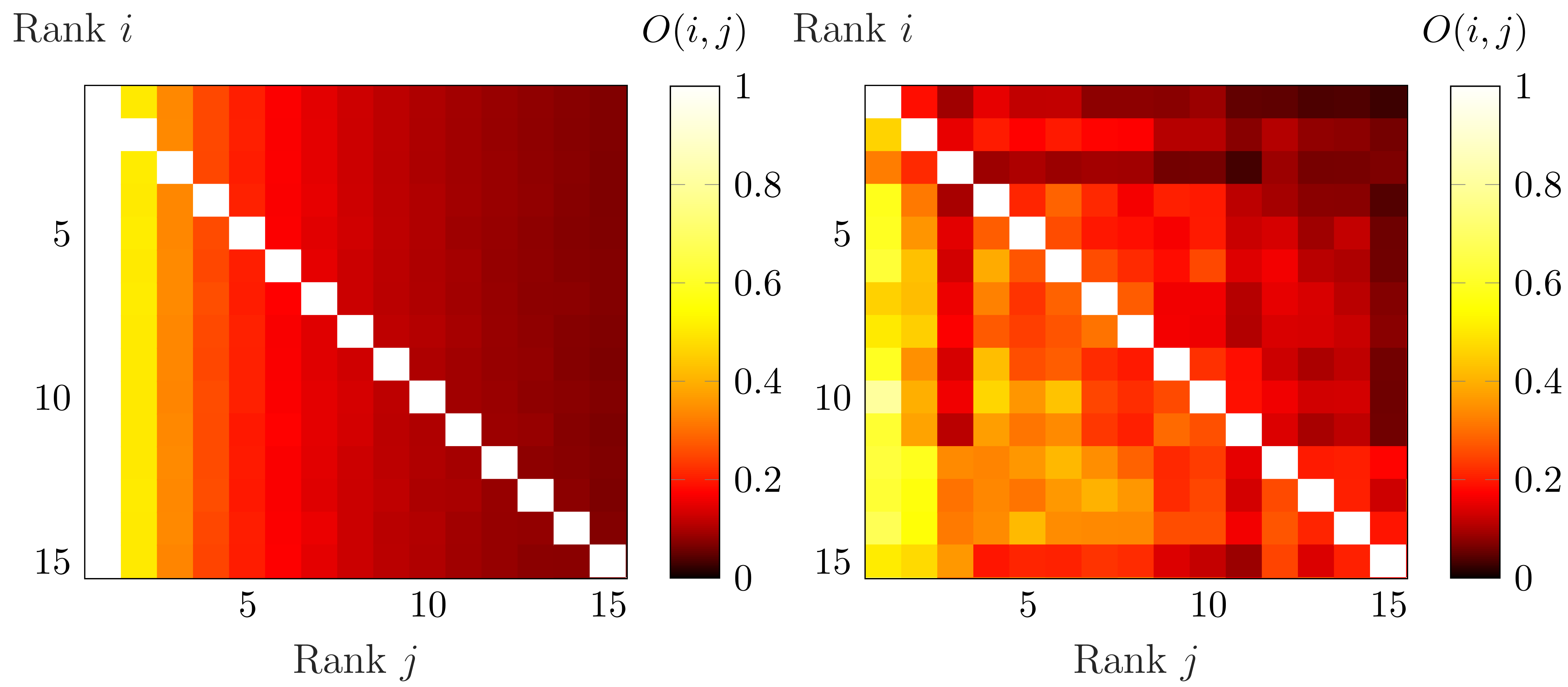
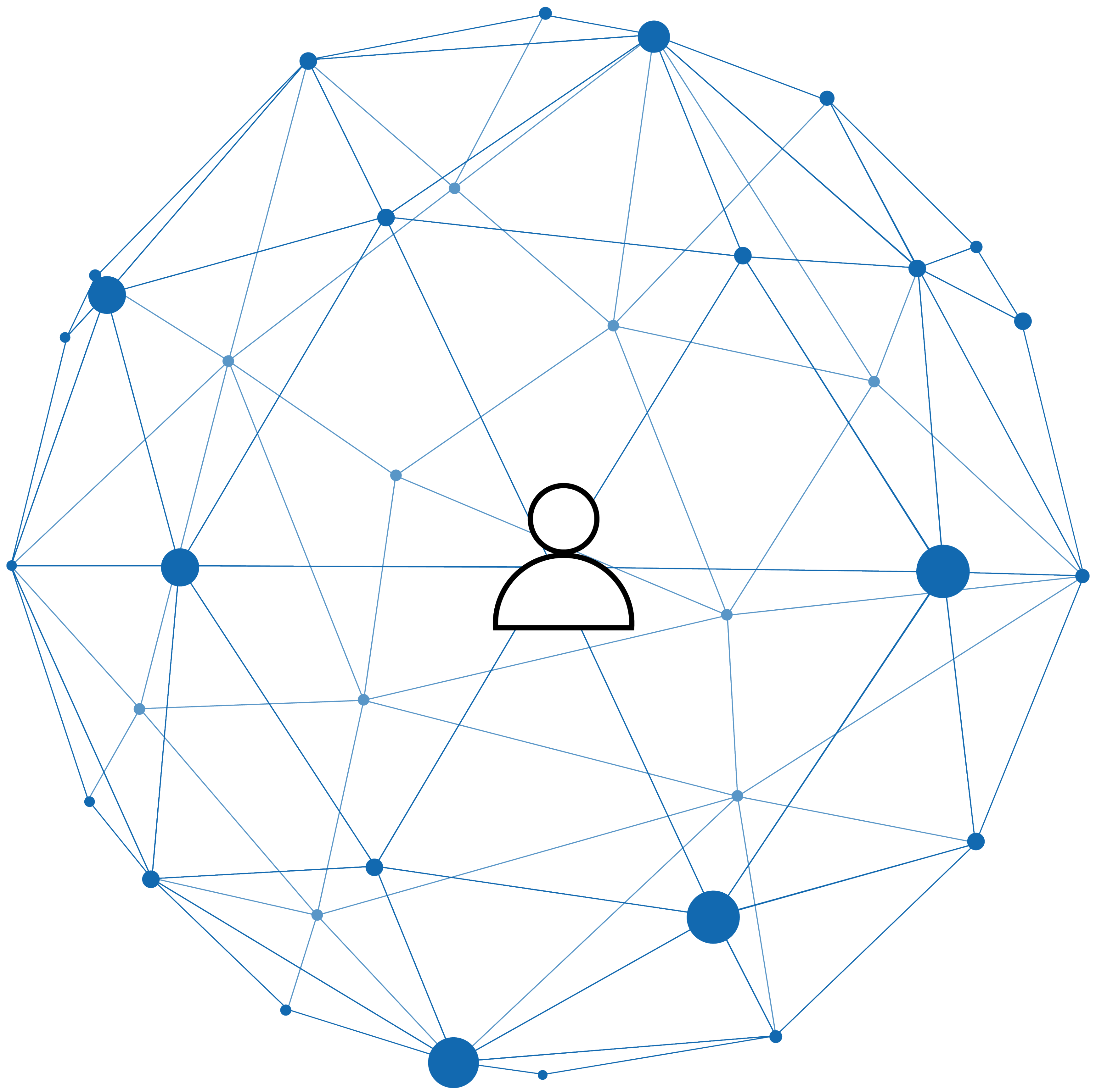


Fig. Left: simulation results. Right: empirical data analysis from Twitch data-set.

Outlook

- Social networks will remain a prevalent aspect of our lives.
- Although complex, social networks show similar patterns, and are ultimately predictable.
- Quantitative studies are now widely accessible thanks to the growing amount of available data, especially from online platforms.
- It is important to continue developing mathematical models, statistical analysis, and computational tools, to better understand the complexity behind human networks.



Outlook

- Part 1:
  - Draw a rigorous connection between the inverse game-theoretical approach and classical statistical regression methods as well as with machine learning methods.
  - Extend the behaviour estimation method to dynamical cases.

- Part 2:
  - Further investigate and validate the meritocratic principle acting on UGC-based online social networks.
  - Analyze the properties of the resulting networks in terms of spreading of information, or mis-information.
  - Study the effect of the recommendation systems and the interplay between humans and social network algorithms.

Acknowledgments

Prof. Dr. Florian Dörfler

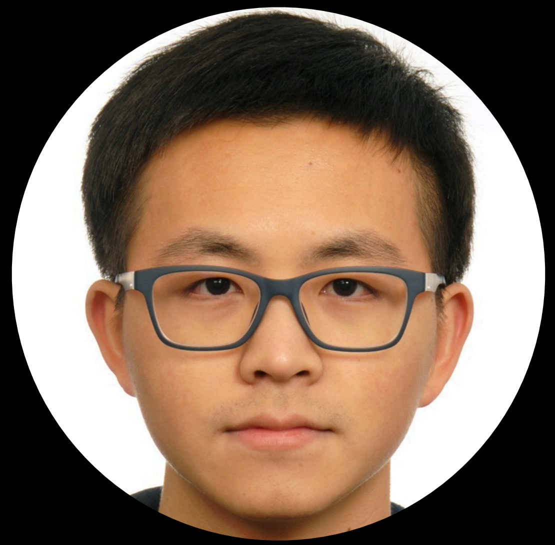
Prof. Dr. Christoph Stadtfeld

Dr. Wenjun Mei

Marco Gallana

Tomer Gidron

Marco Buob

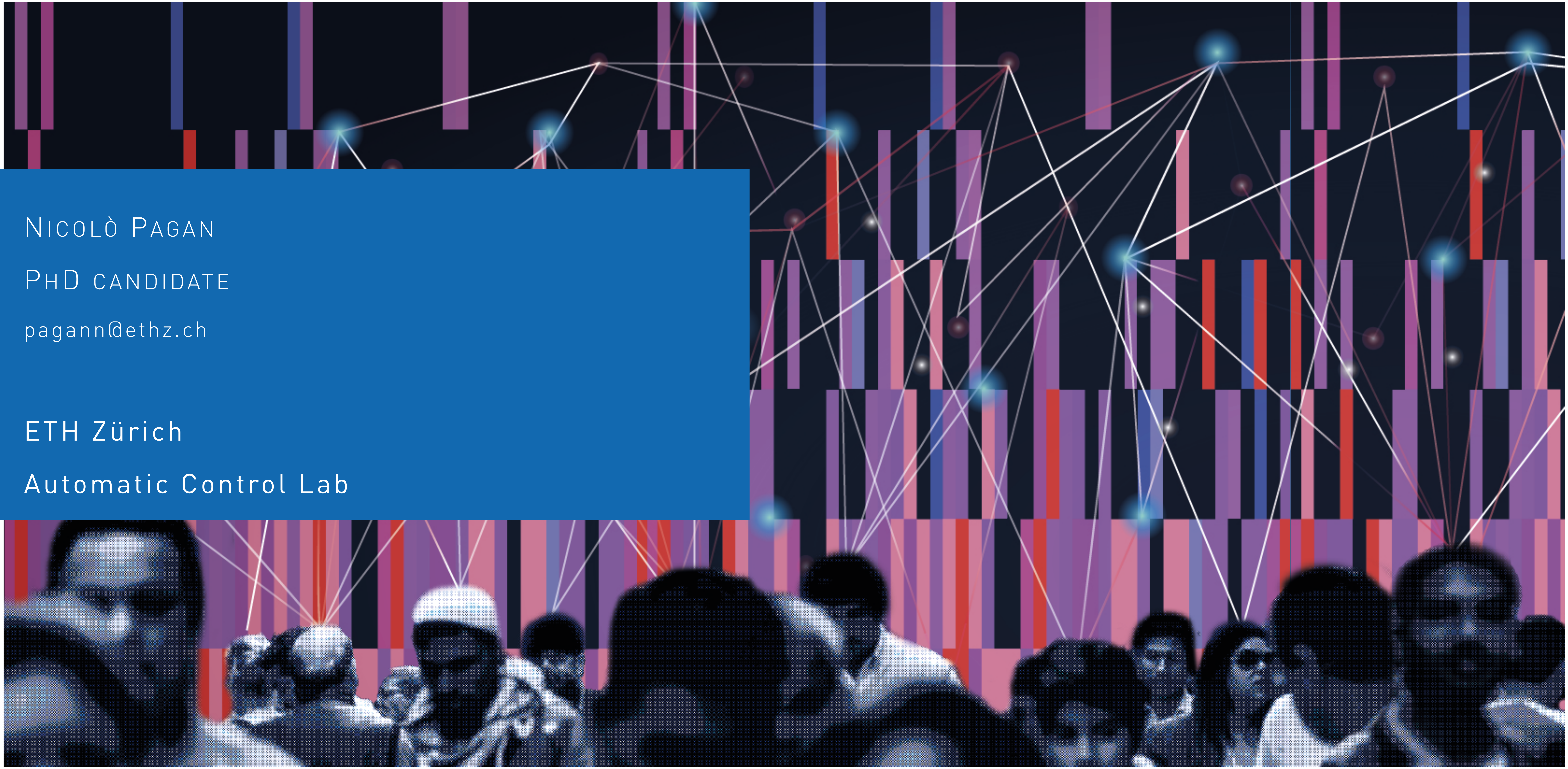


Cheng Li

Anna Maddux

ETH zürich

AUTOMATIC CONTROL LABORATORY



Nicolò Pagan

PhD candidate

pagann@ethz.ch

ETH Zürich

Automatic Control Lab