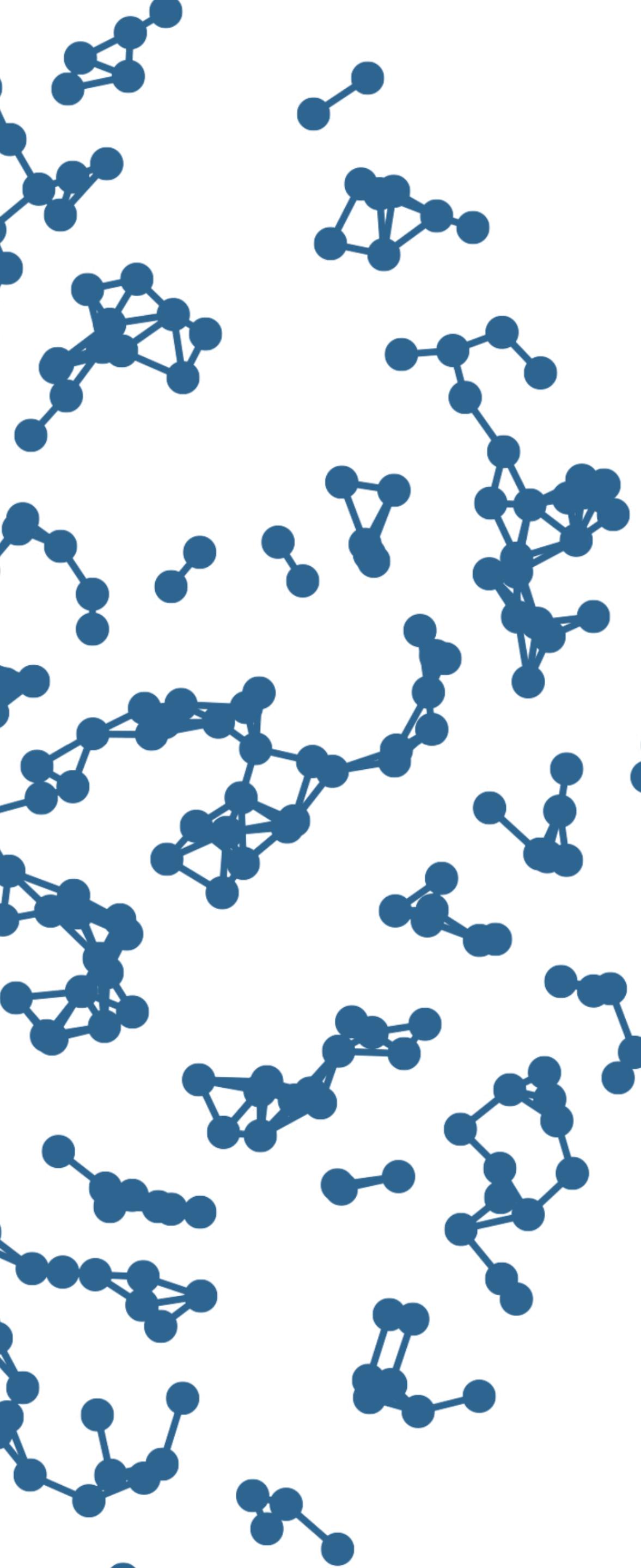


STOCHASTIC SPREADING ON COMPLEX NETWORKS

Gerrit Großmann

12.12.2022





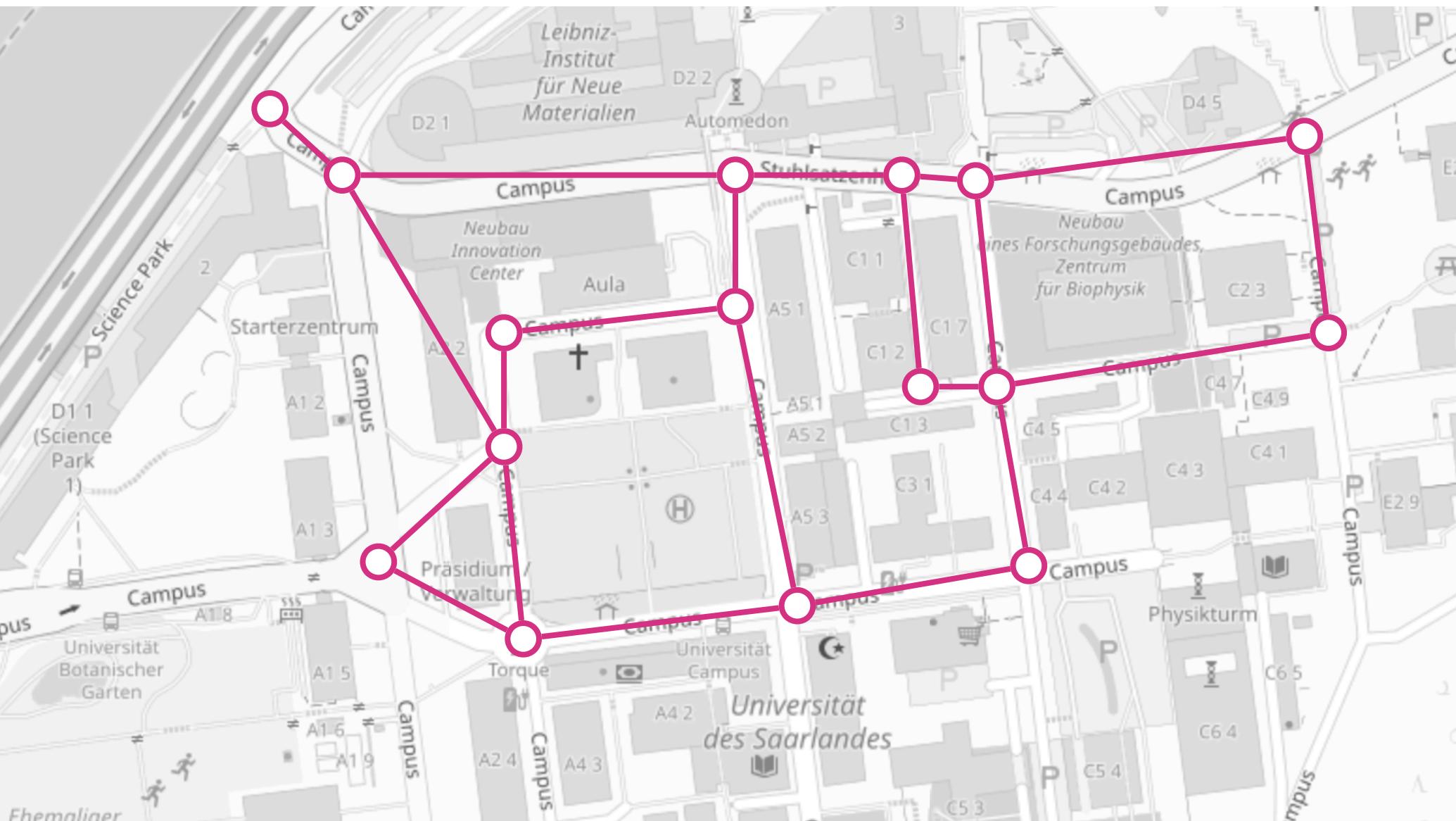
Prologue

NETWORKS ARE EVERYWHERE



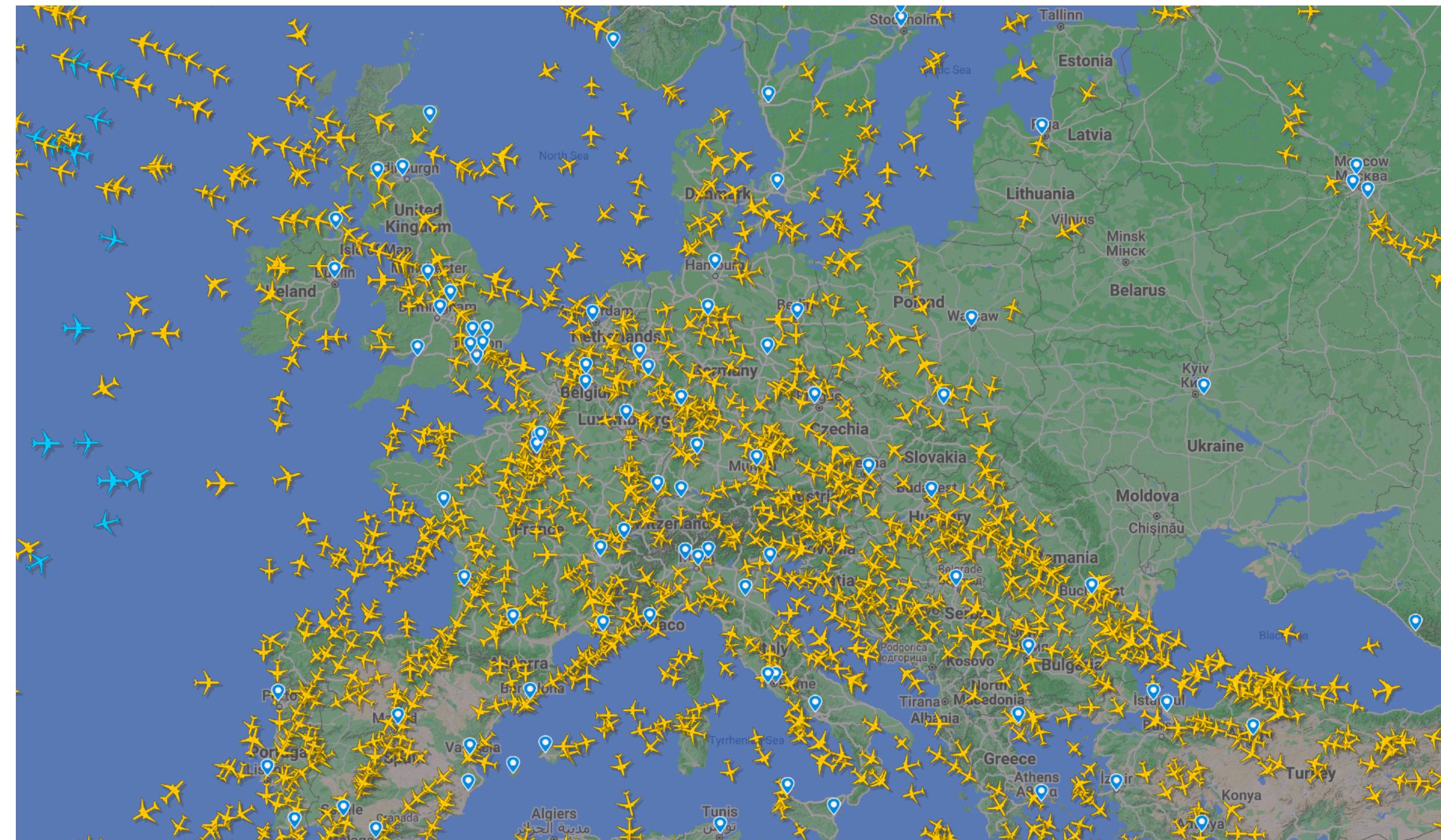
In many ways, graphs are the main modality
of data we receive from nature.

- Petar Veličković



In many ways, graphs are the main modality
of data we receive from nature.

- Petar Veličković



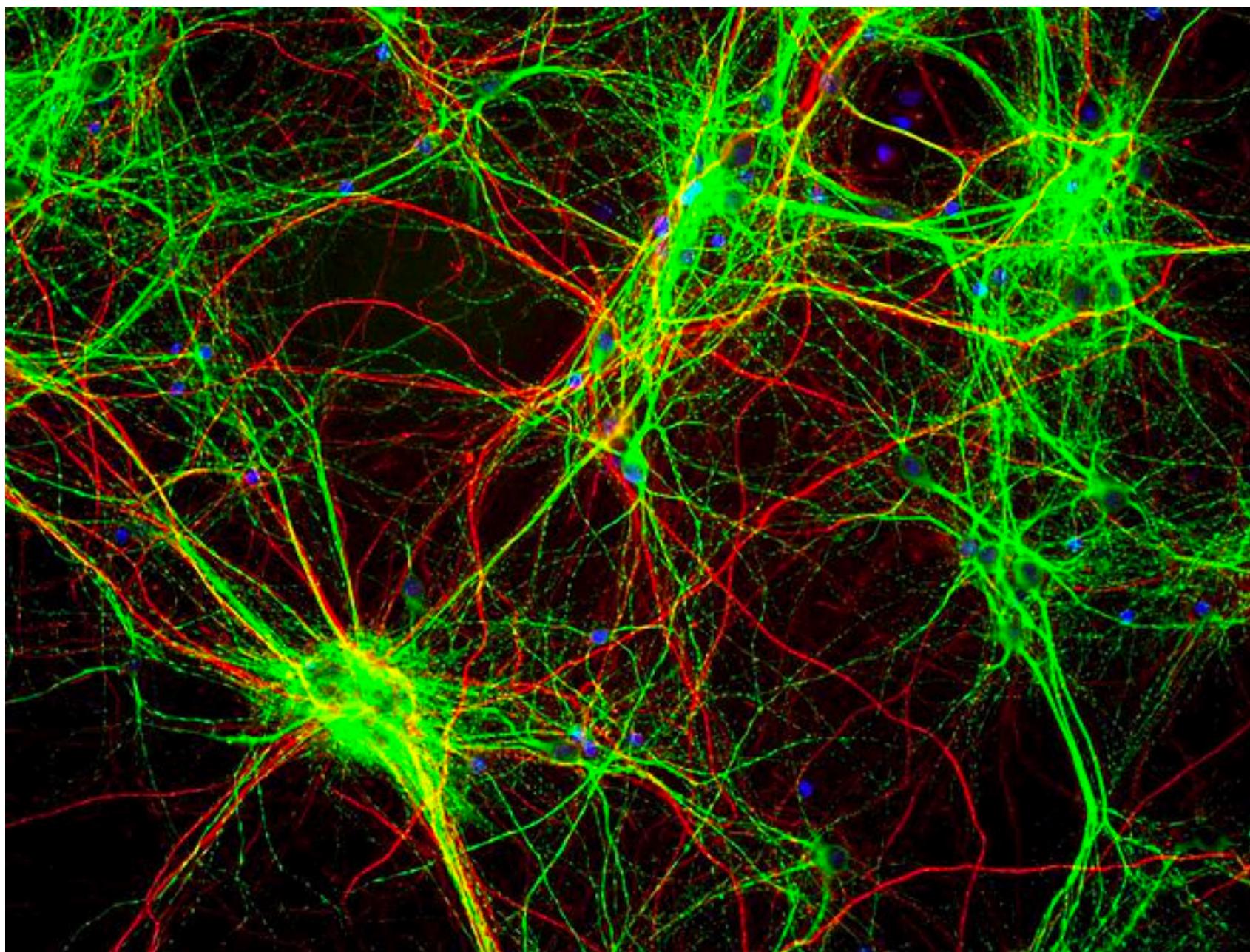
In many ways, graphs are the main modality
of data we receive from nature.

- Petar Veličković



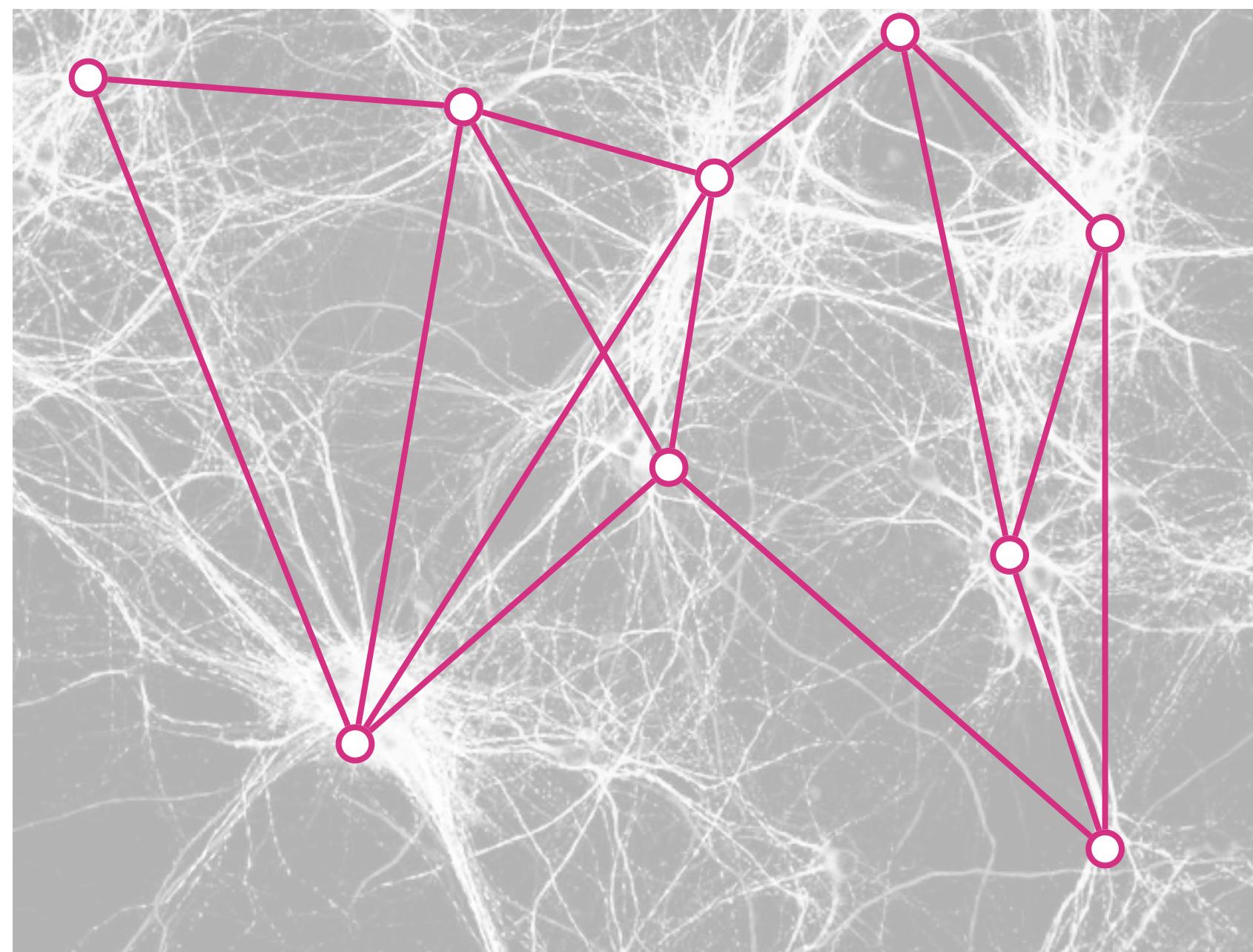
In many ways, graphs are the main modality
of data we receive from nature.

- Petar Veličković



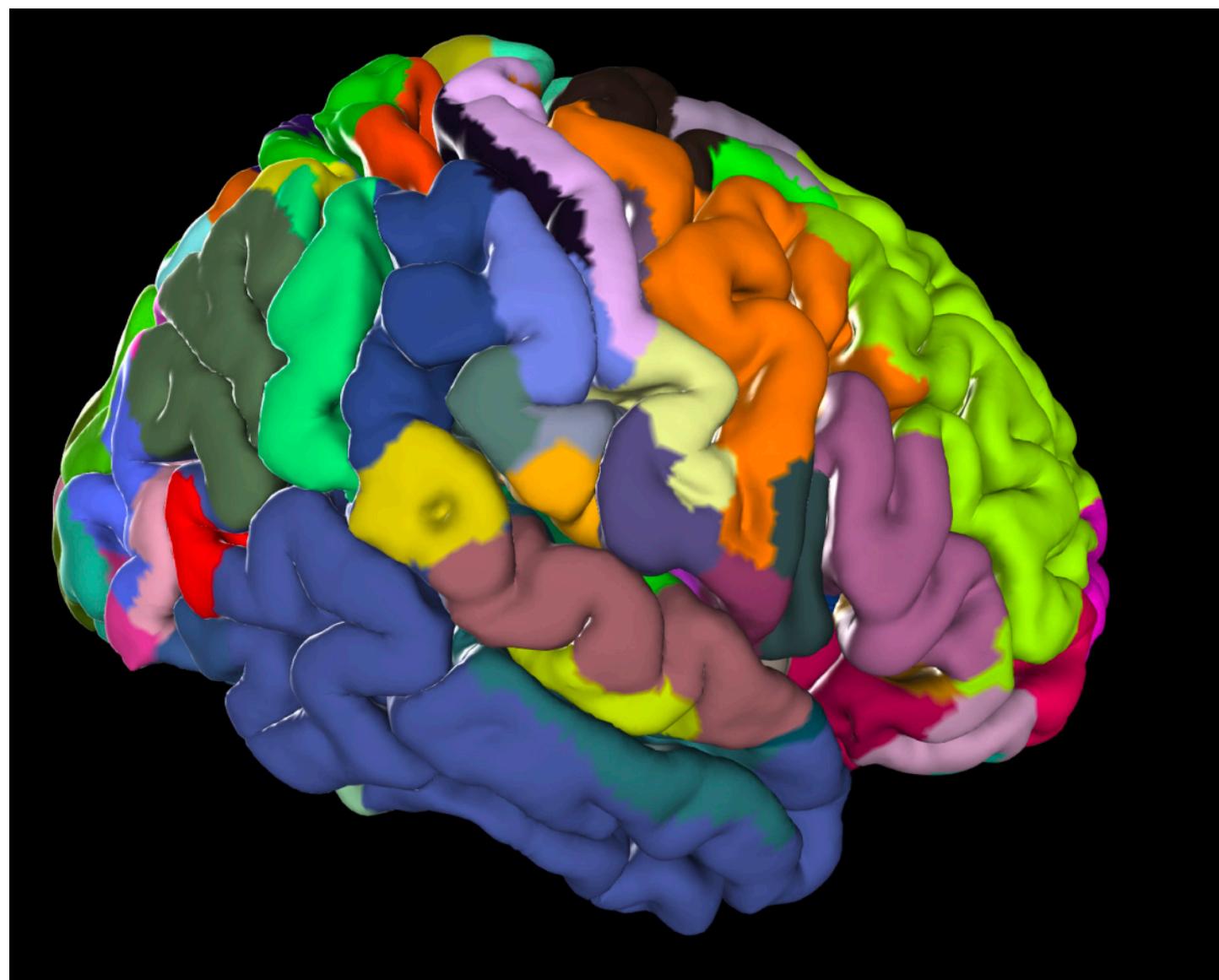
In many ways, graphs are the main modality
of data we receive from nature.

- Petar Veličković



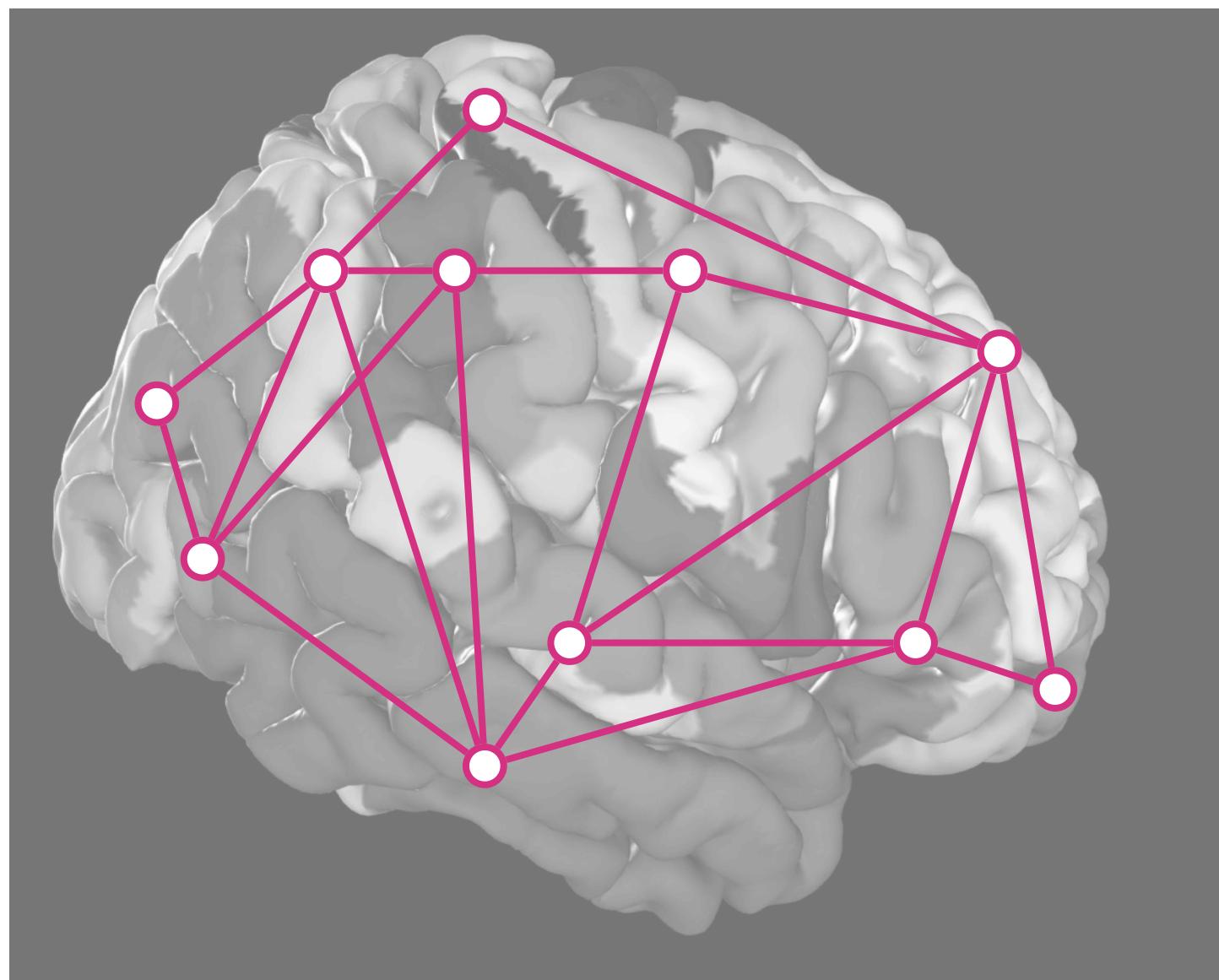
In many ways, graphs are the main modality
of data we receive from nature.

- Petar Veličković



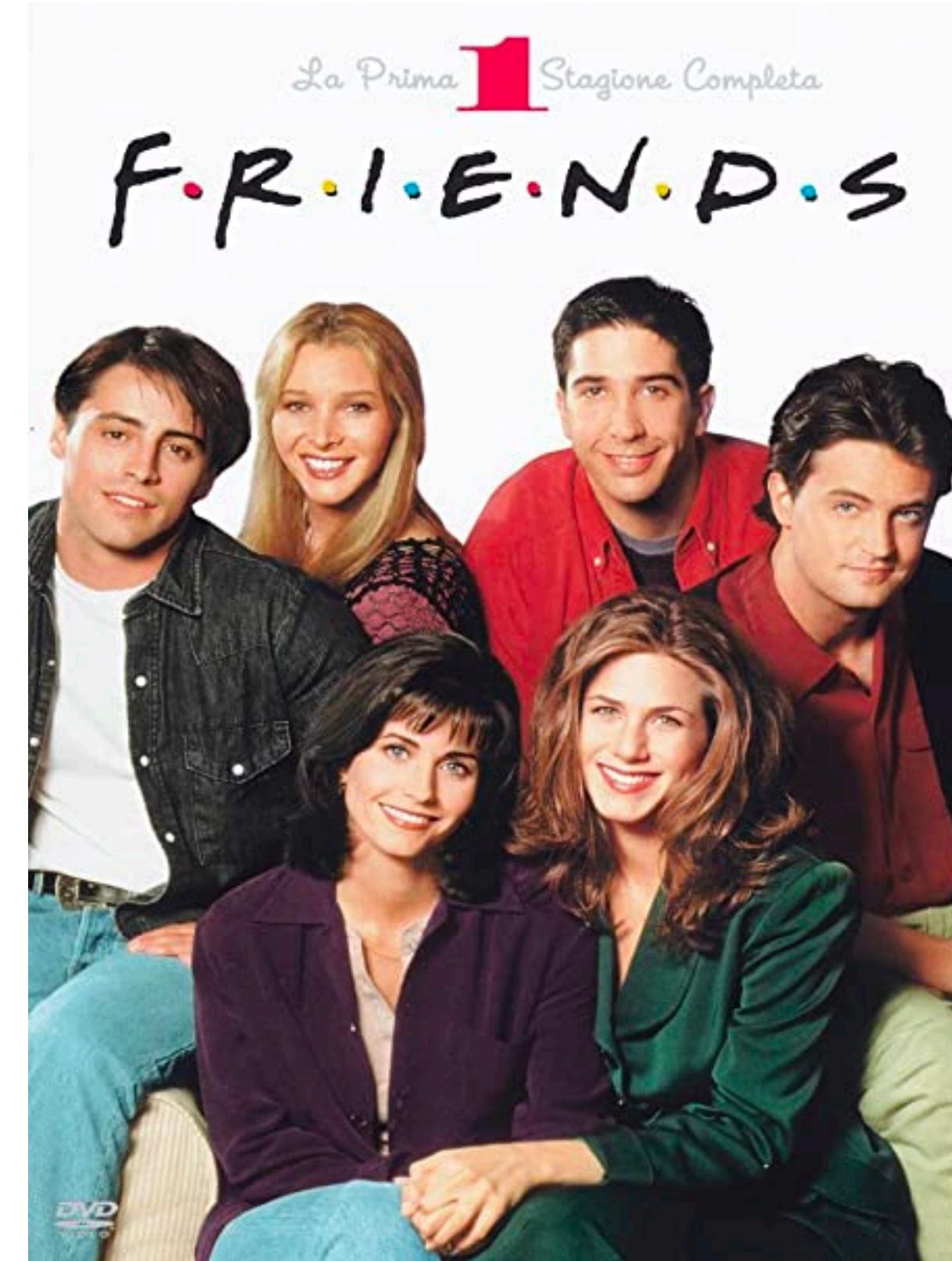
In many ways, graphs are the main modality
of data we receive from nature.

- Petar Veličković



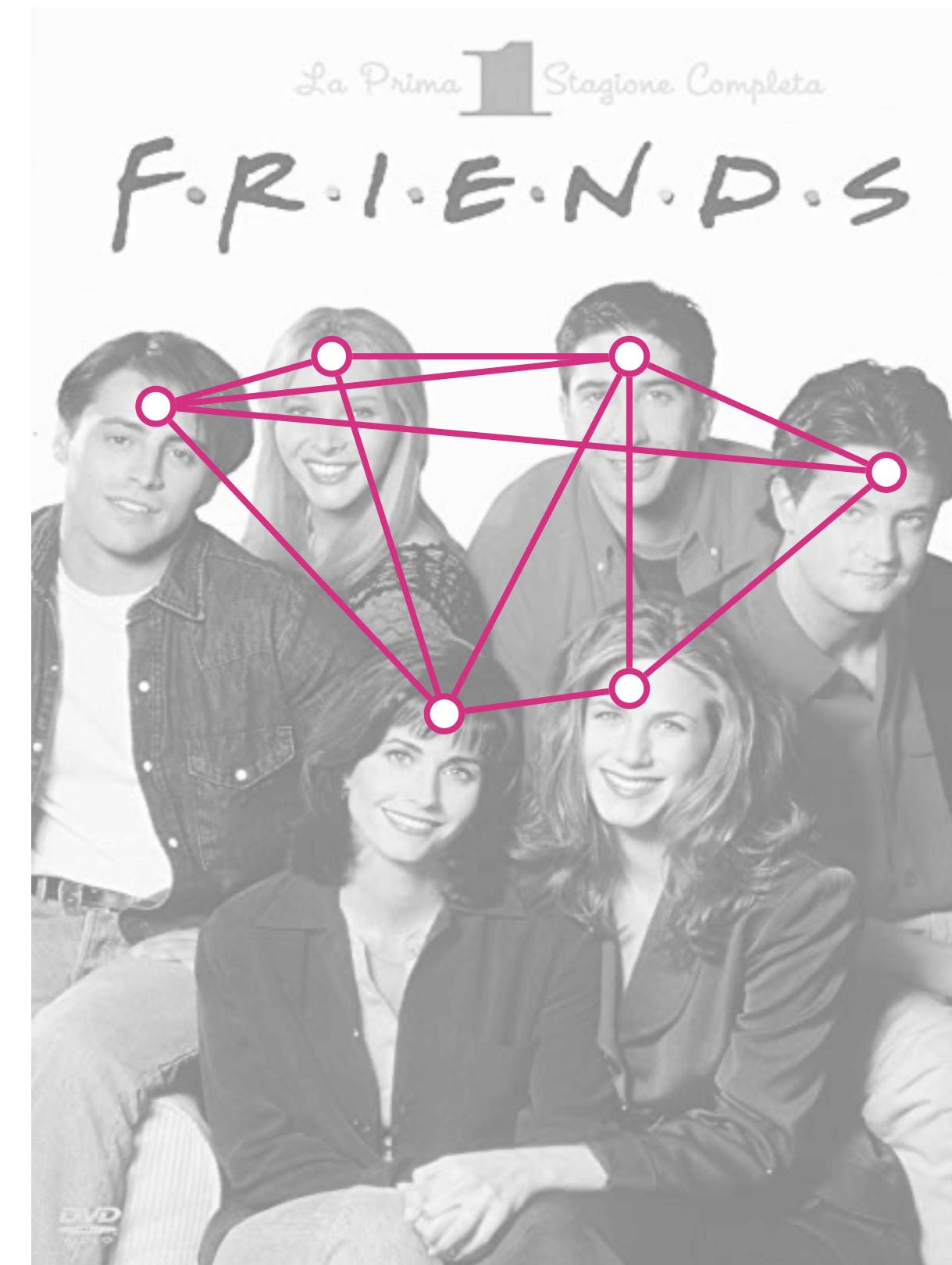
In many ways, graphs are the main modality
of data we receive from nature.

- Petar Veličković



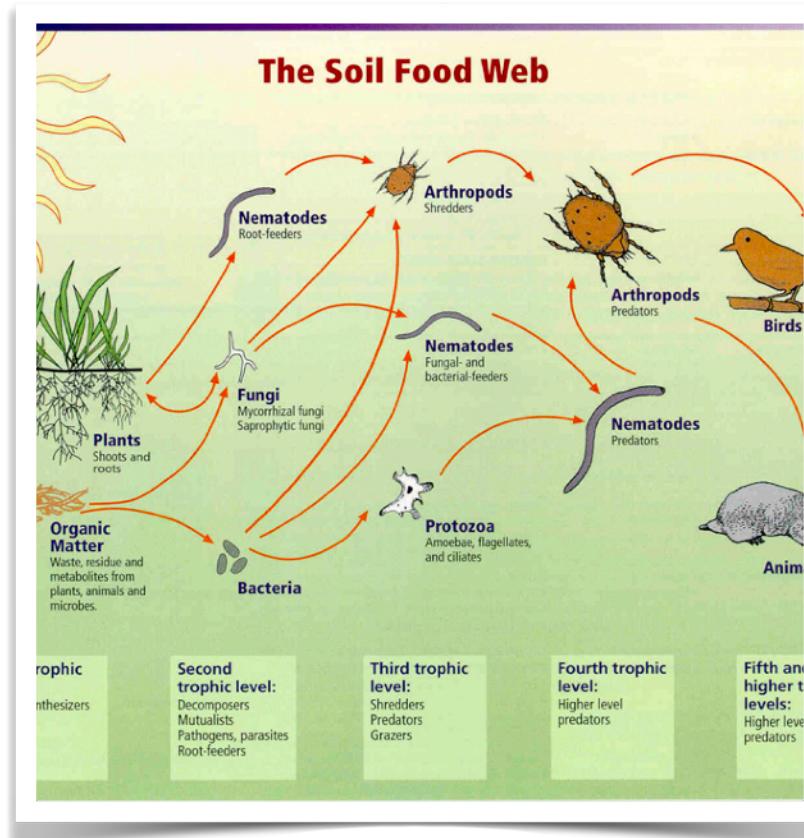
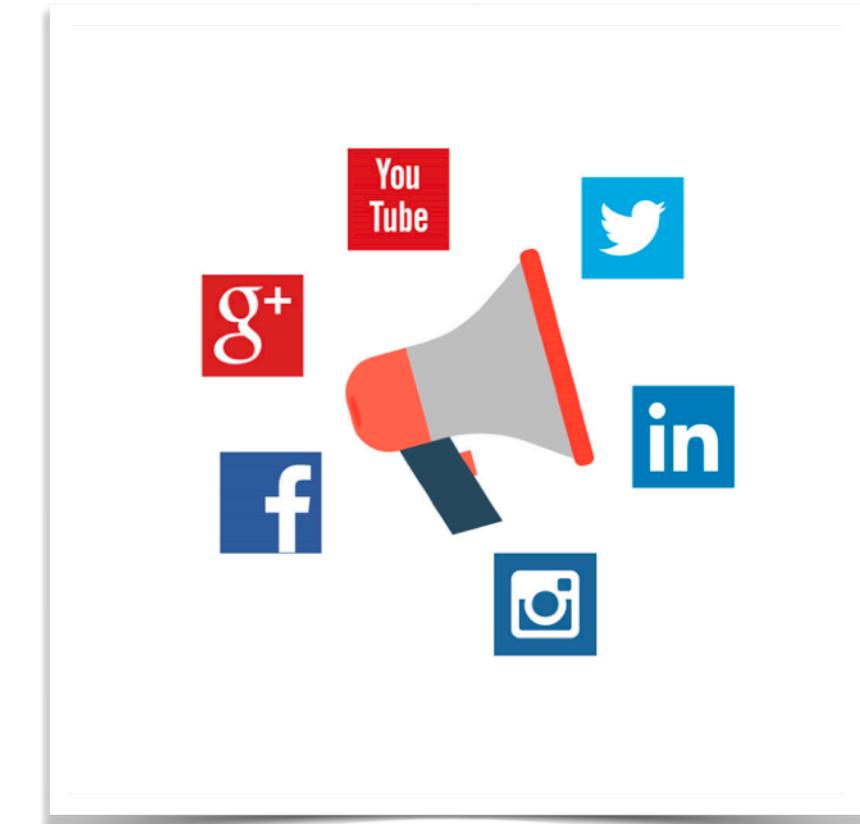
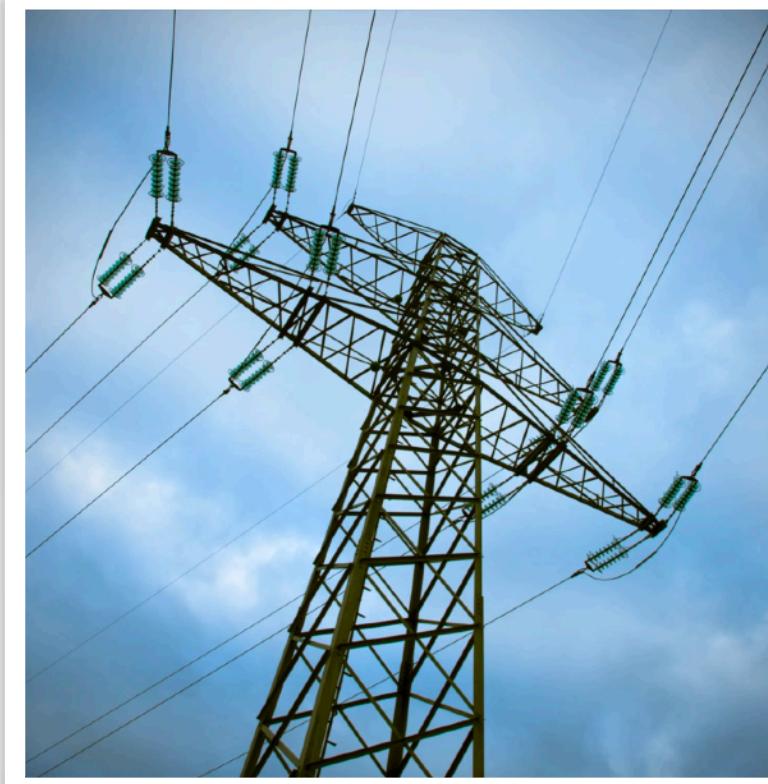
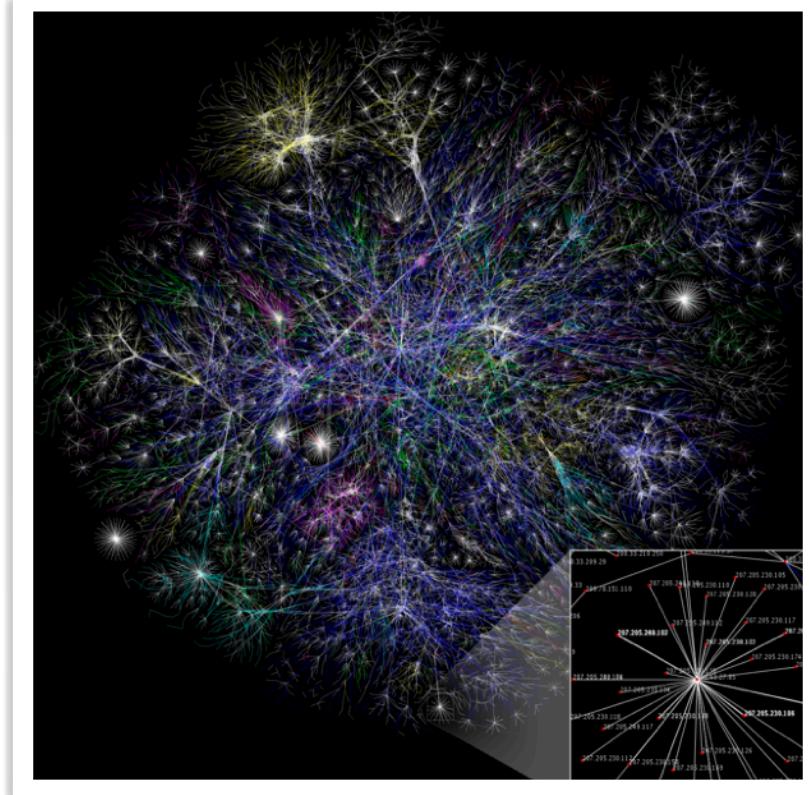
In many ways, graphs are the main modality
of data we receive from nature.

- Petar Veličković



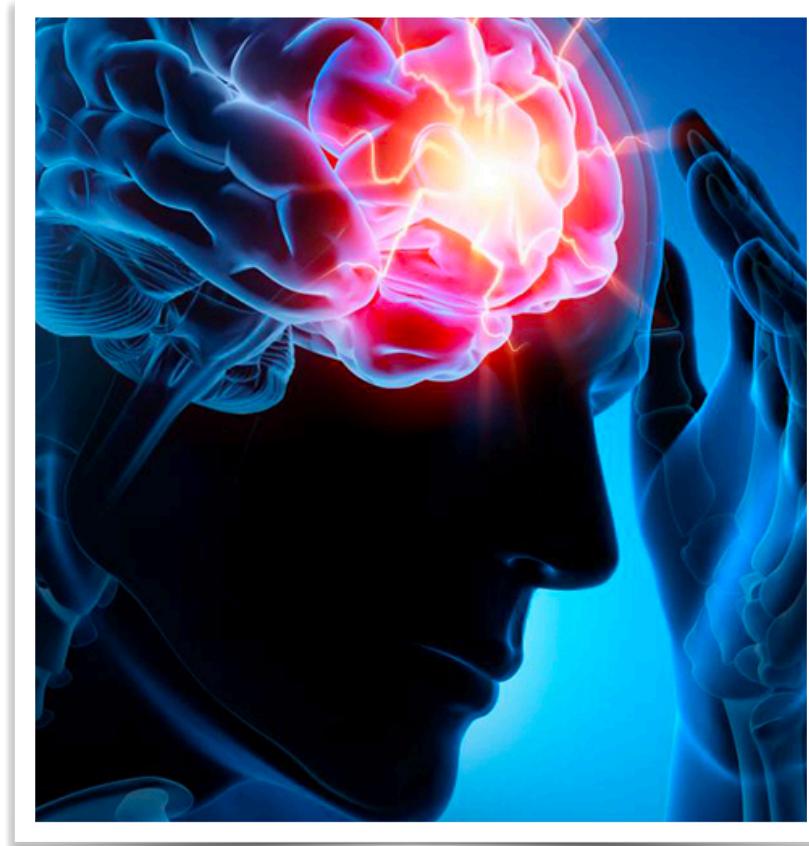
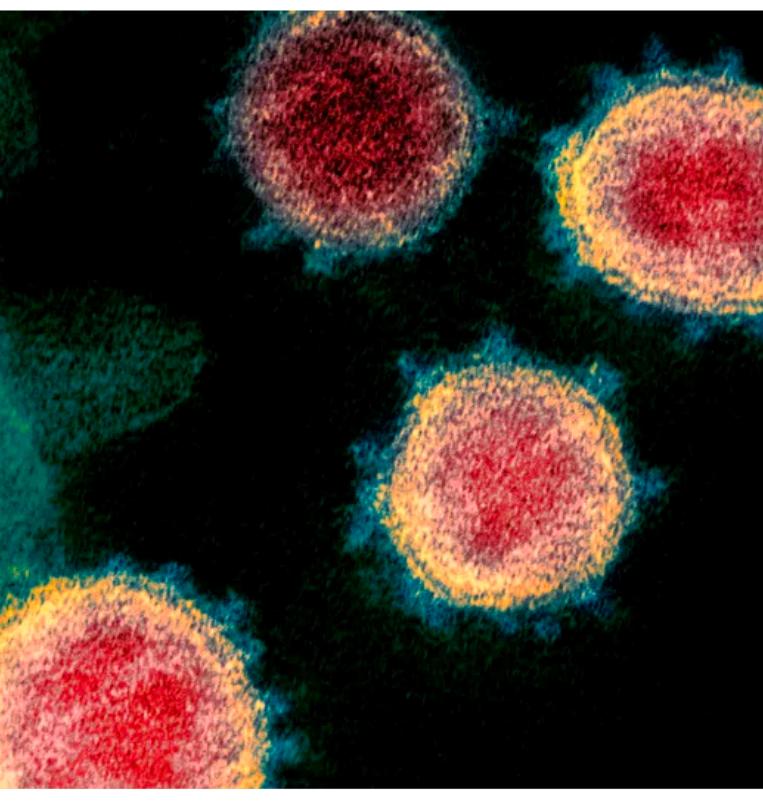
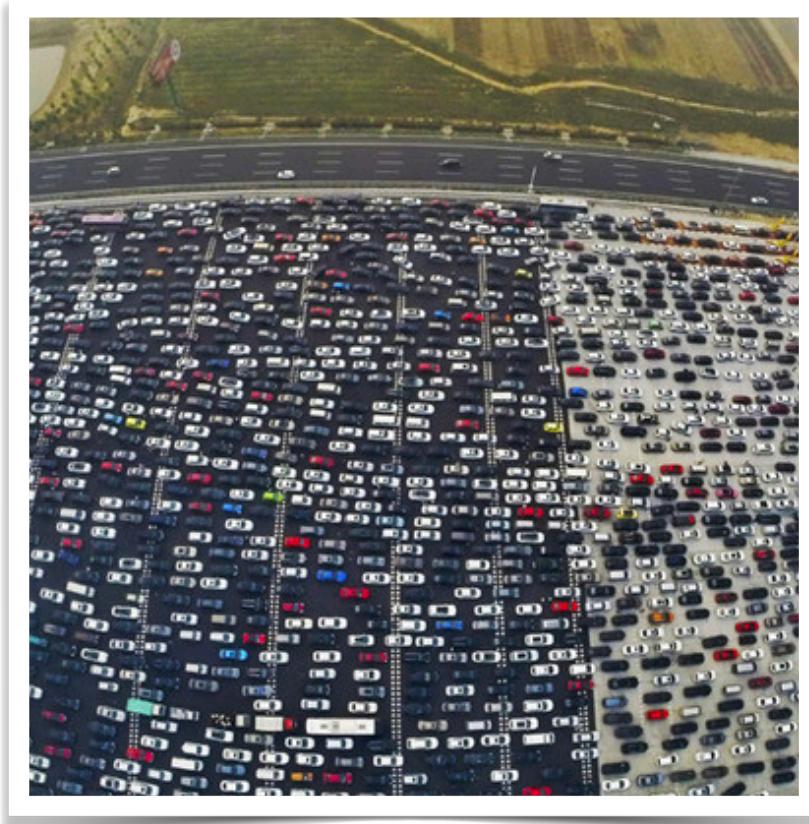
In many ways, graphs are the main modality
of data we receive from nature.

- Petar Veličković



In many ways, graphs are the main modality
of data we receive from nature.

- Petar Veličković

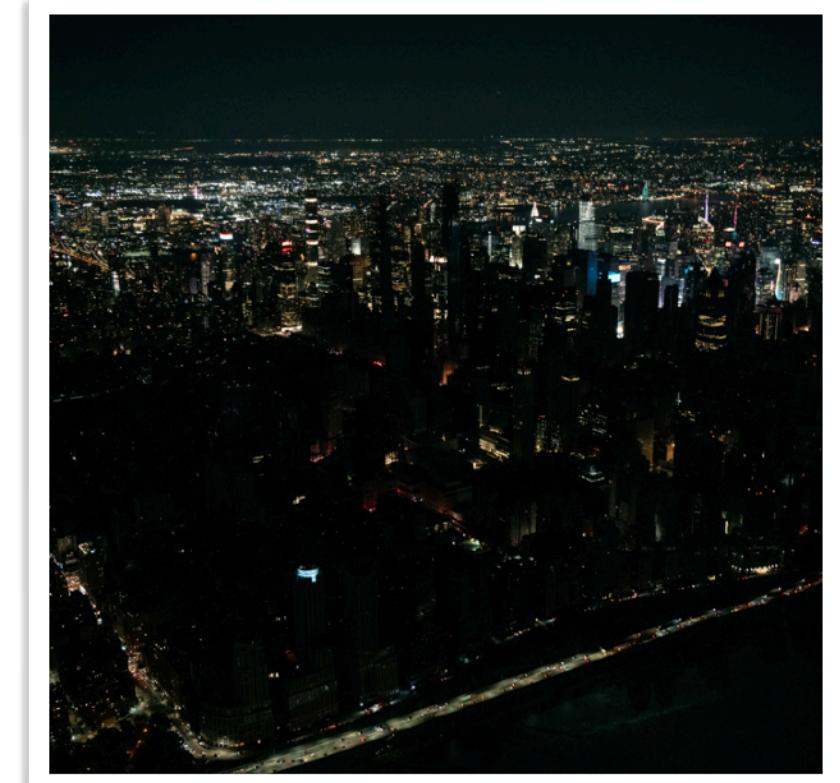


Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement

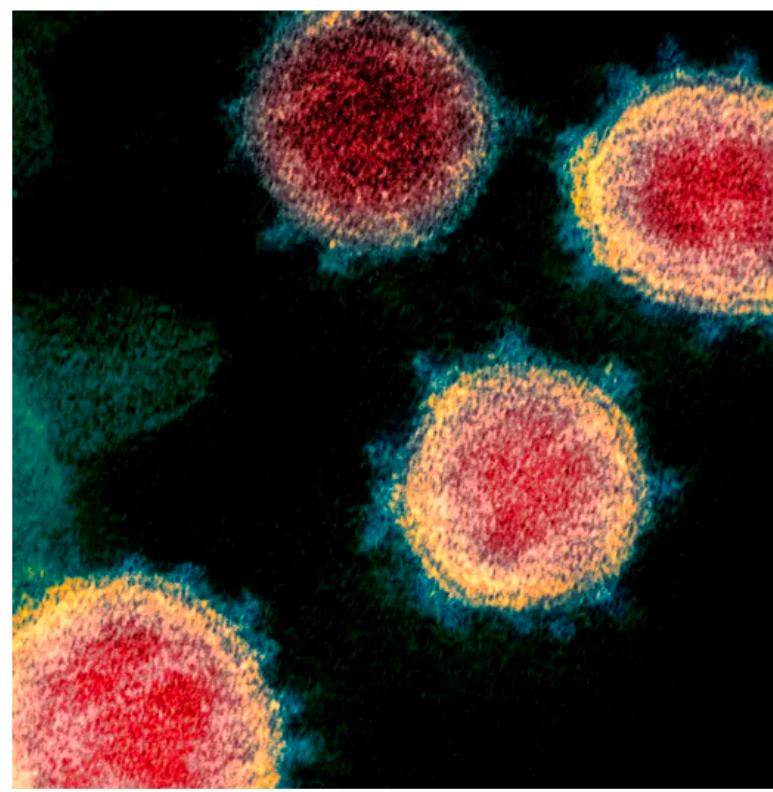
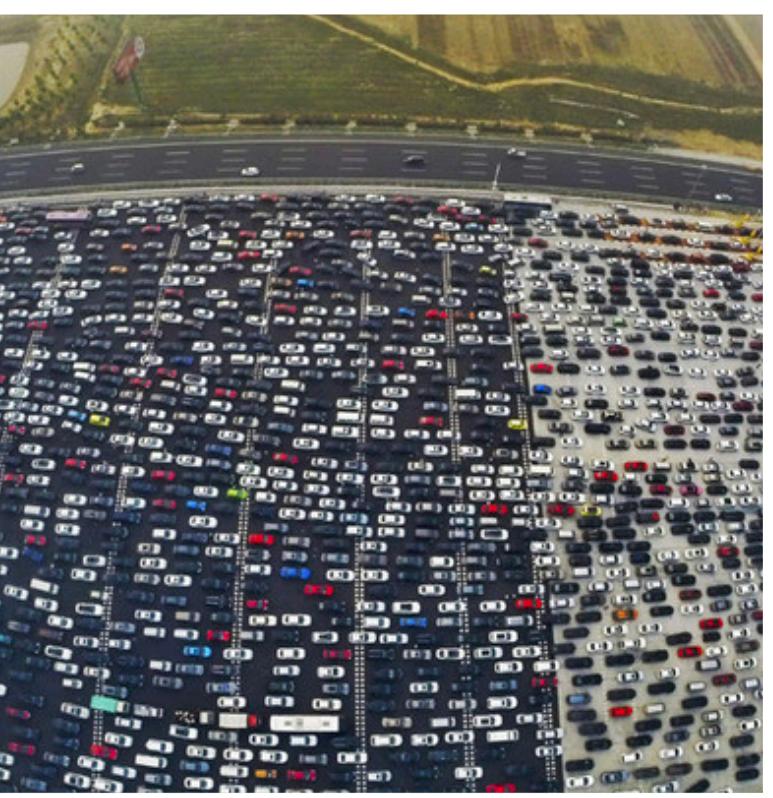
TOPICS: Pope Francis Endorses Donald Trump

Wrong again. WTOES

Facebook engagements: 961,300



When there is a network, things will spread.



Pope Francis Shocks World,
Endorses Donald Trump for
President, Releases Statement

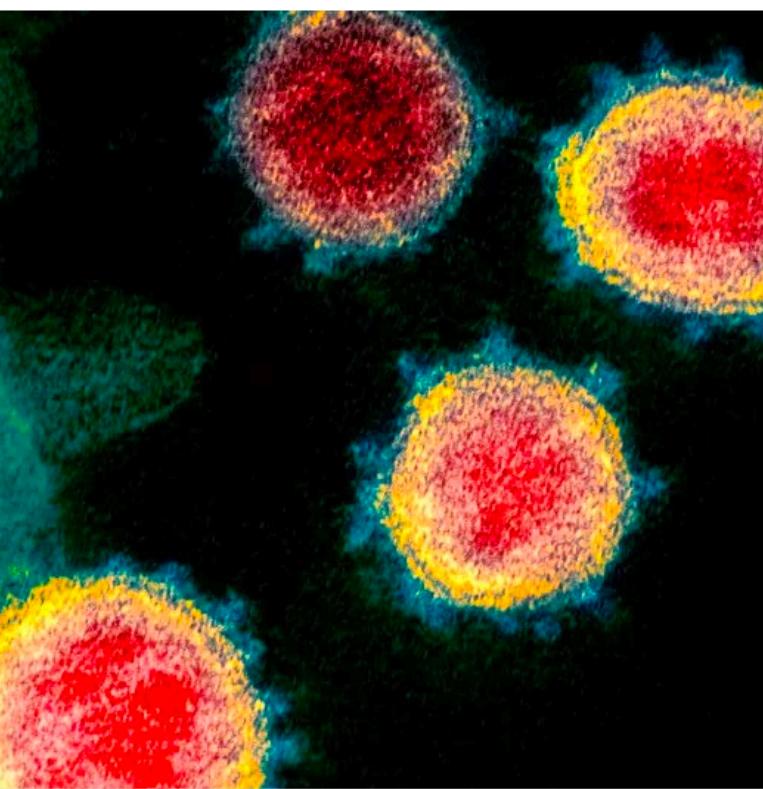
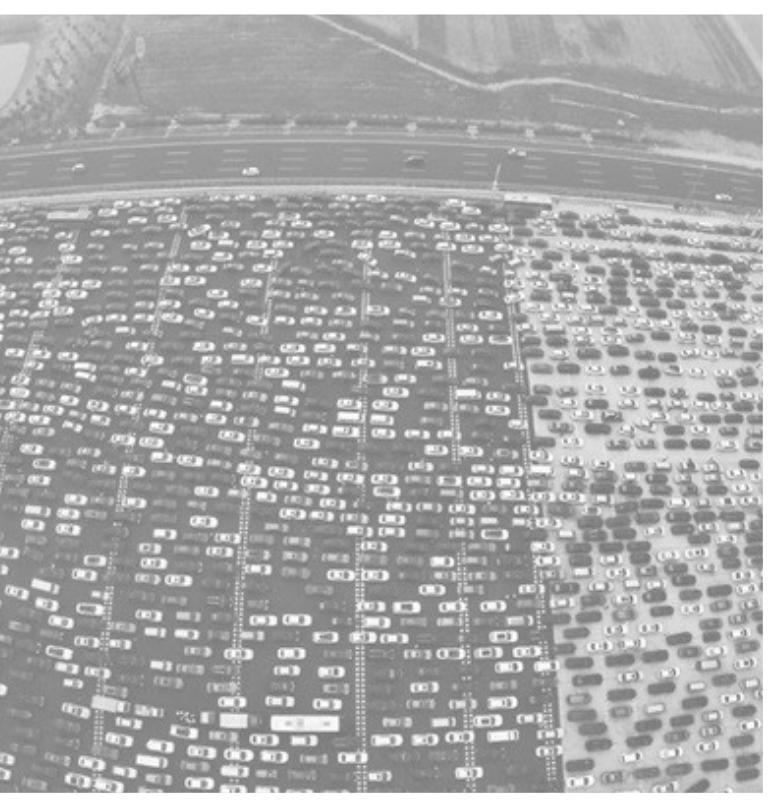
TOPICS: Pope Francis Endorses Donald Trump

Wrong again. WTOES

Facebook engagements: 961,300



Goal: Build a mathematical model of information diffusion.



Pope Francis Shocks World,
Endorses Donald Trump for
President, Releases Statement

TOPICS: Pope Francis Endorses Donald Trump



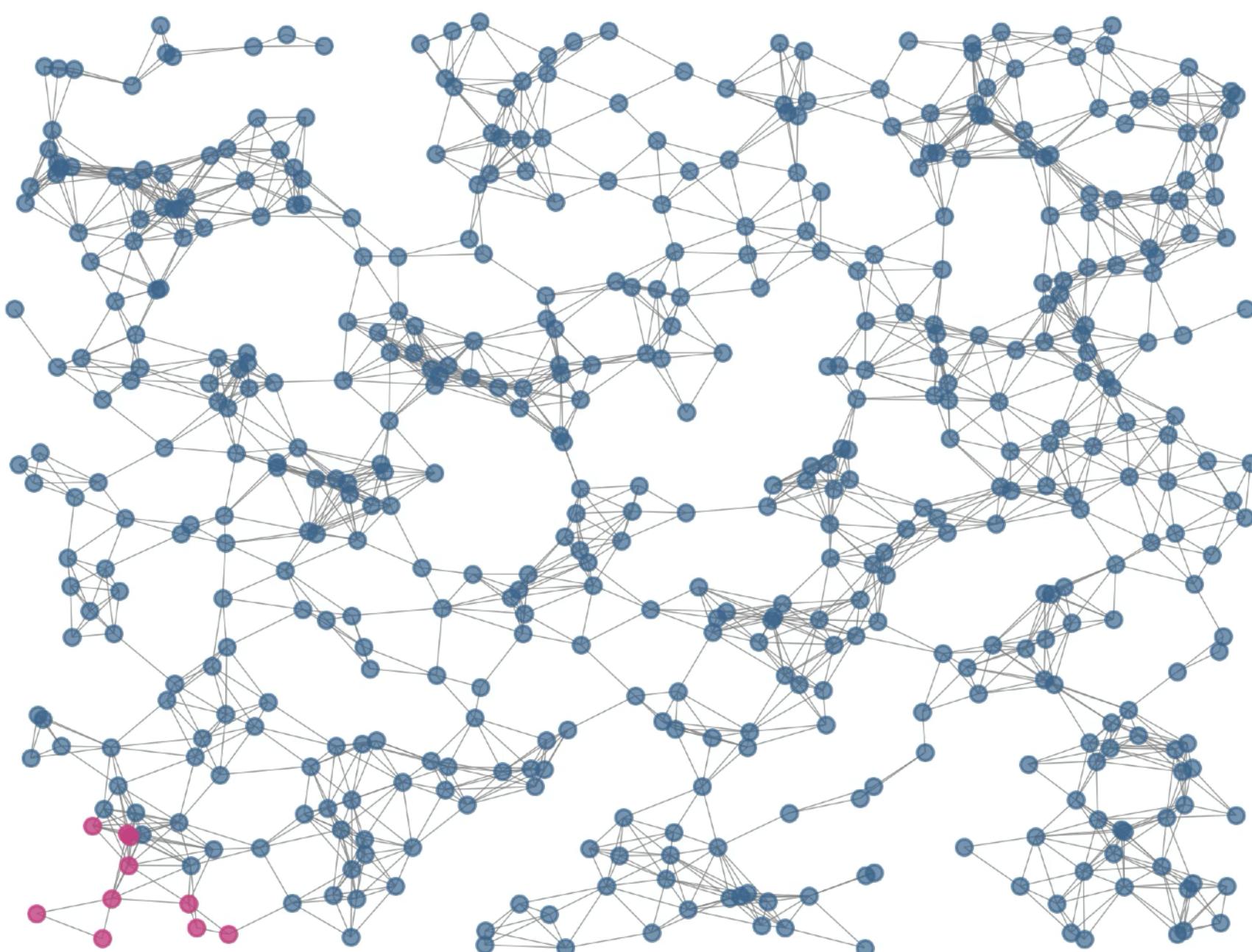
Wrong again. WTOES

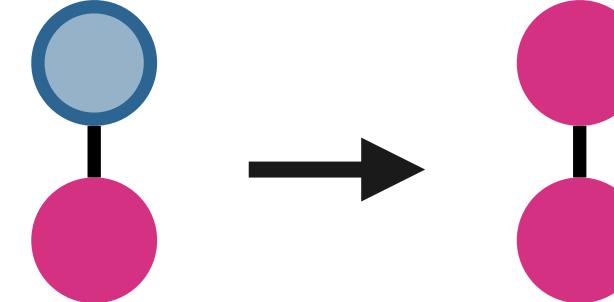
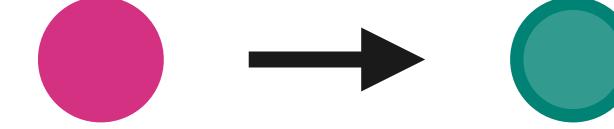
Facebook engagements: 961,300



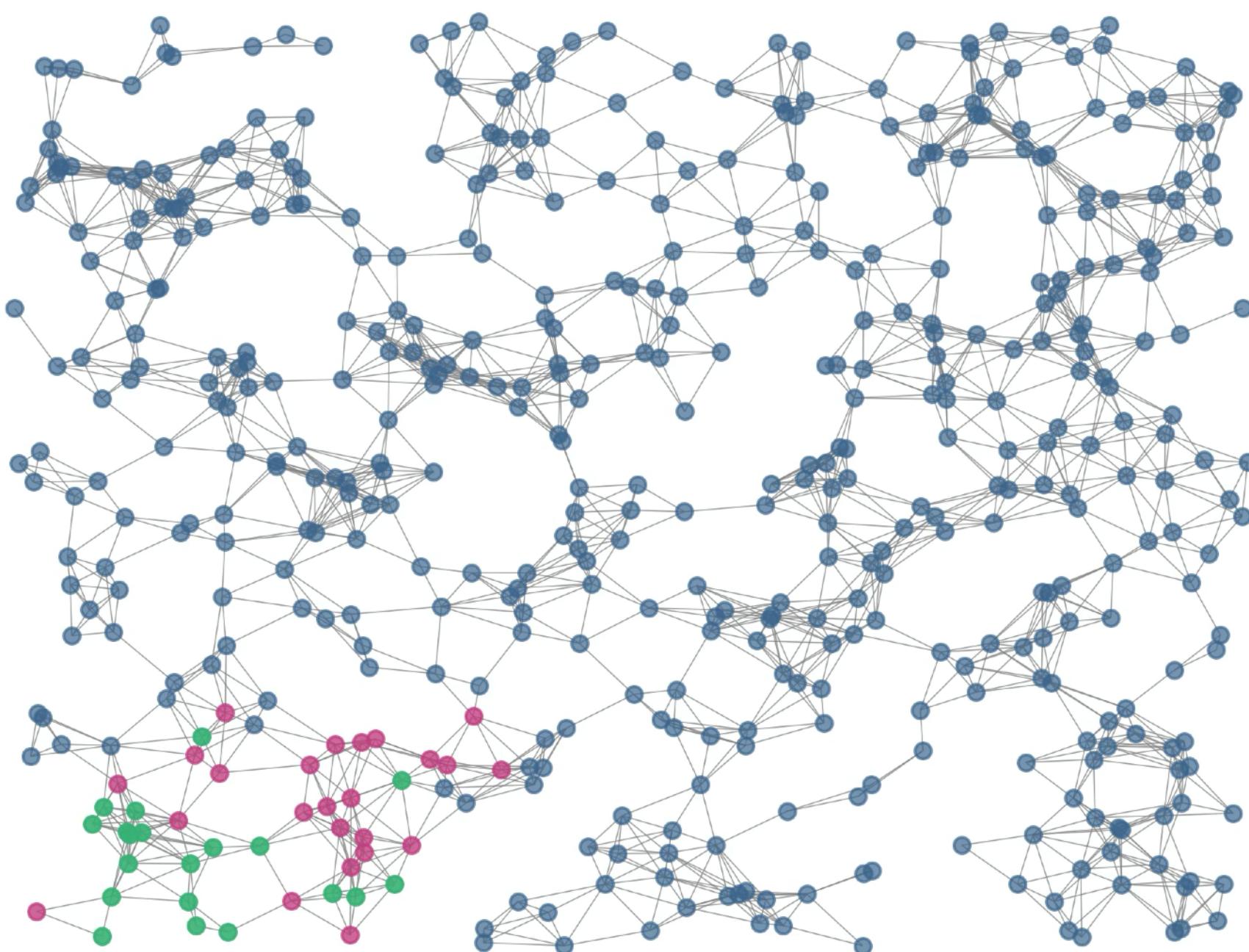
Everything that spreads is an epidemic.

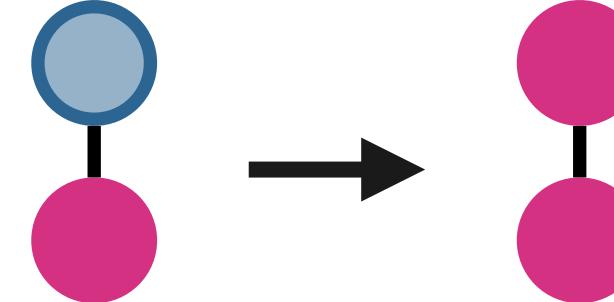
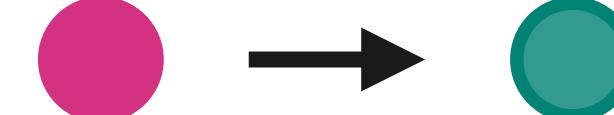
SIR



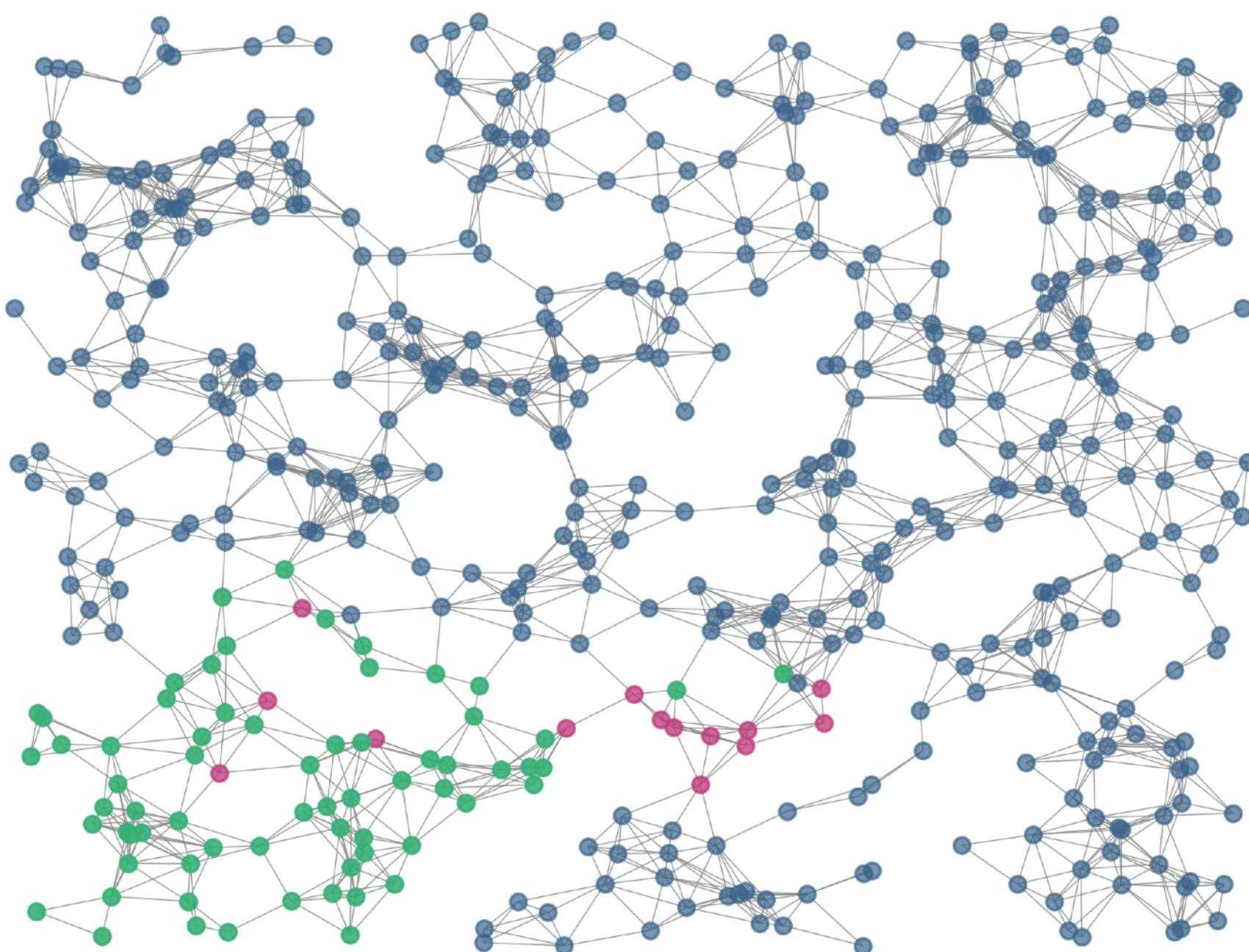
- ▶ Nodes are **susceptible**, **infected**, **recovered**.
- ▶  **Infected** nodes infect **susceptible** neighbors.
- ▶  **Infected** nodes recover (gain immunity).

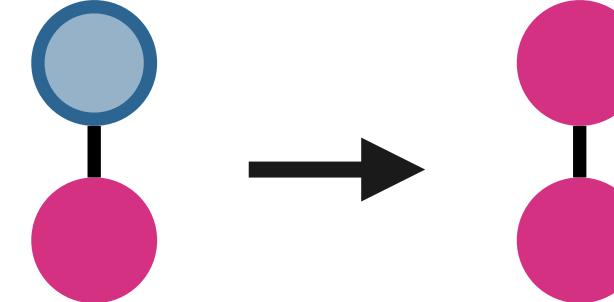
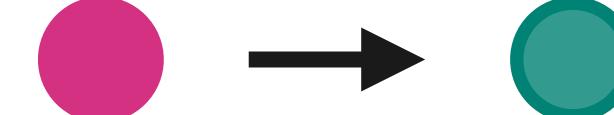
SIR



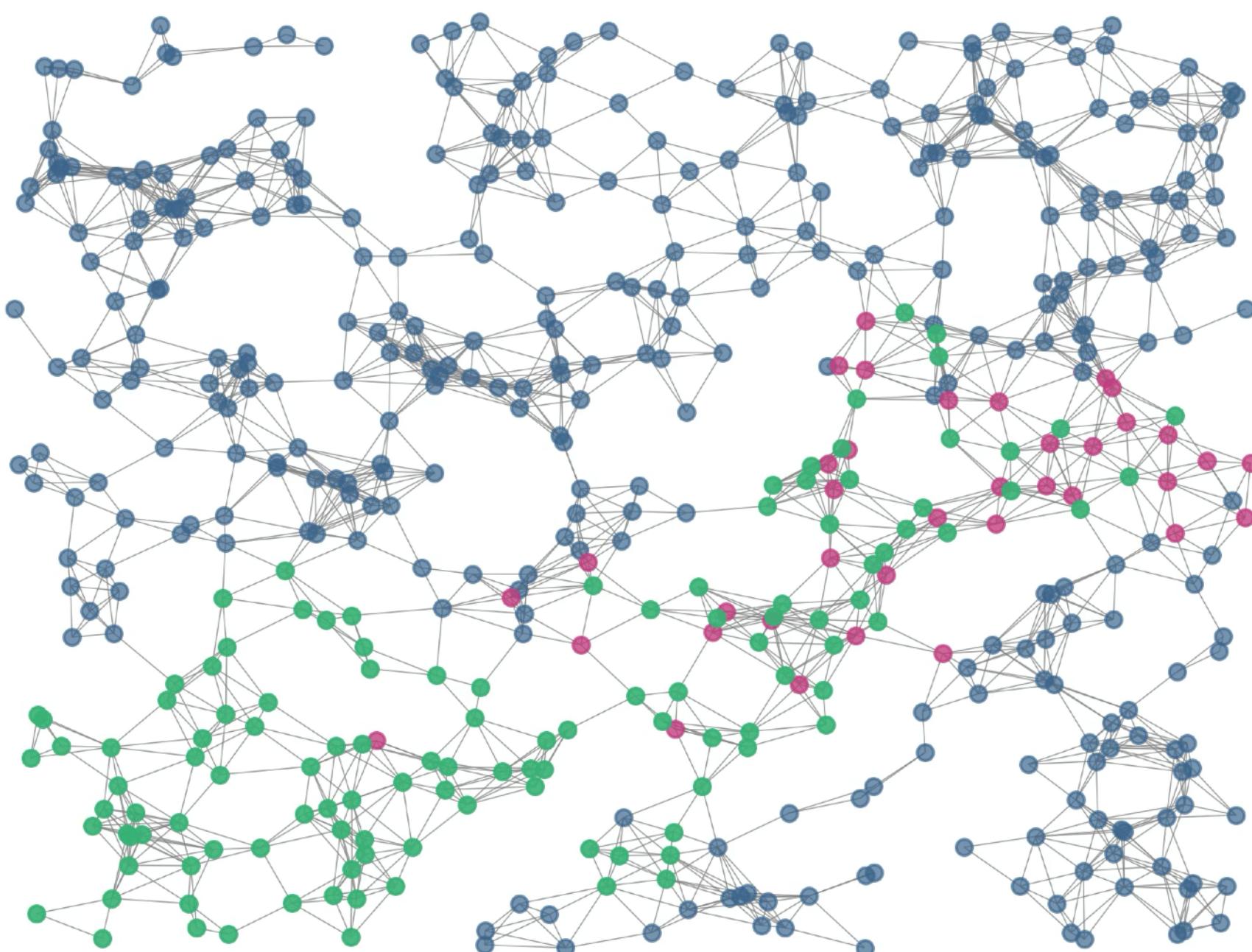
- ▶ Nodes are **susceptible**, **infected**, **recovered**.
- ▶ 
- ▶ **Infected** nodes infect **susceptible** neighbors.
- ▶ 
- ▶ **Infected** nodes recover (gain immunity).

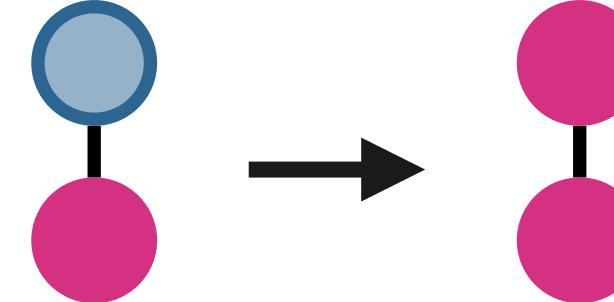
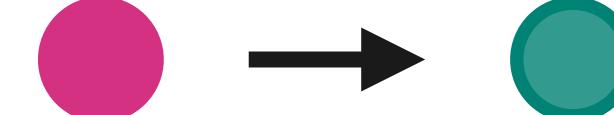
SIR



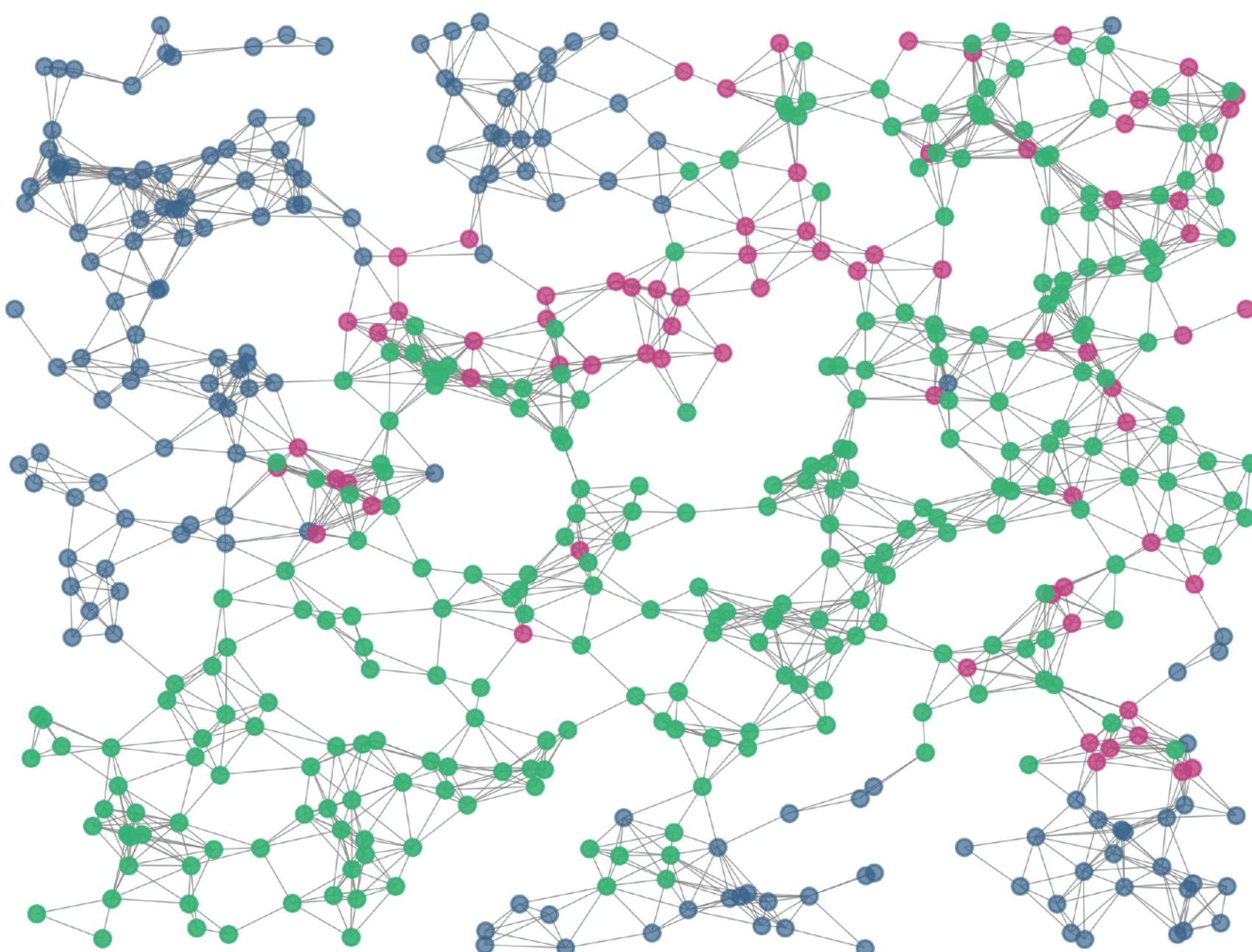
- ▶ Nodes are **susceptible**, **infected**, **recovered**.
- ▶ 
- ▶ **Infected** nodes infect **susceptible** neighbors.
- ▶ 
- ▶ **Infected** nodes recover (gain immunity).

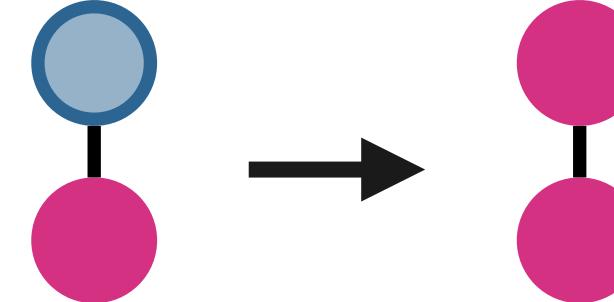
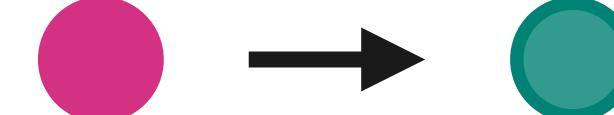
SIR



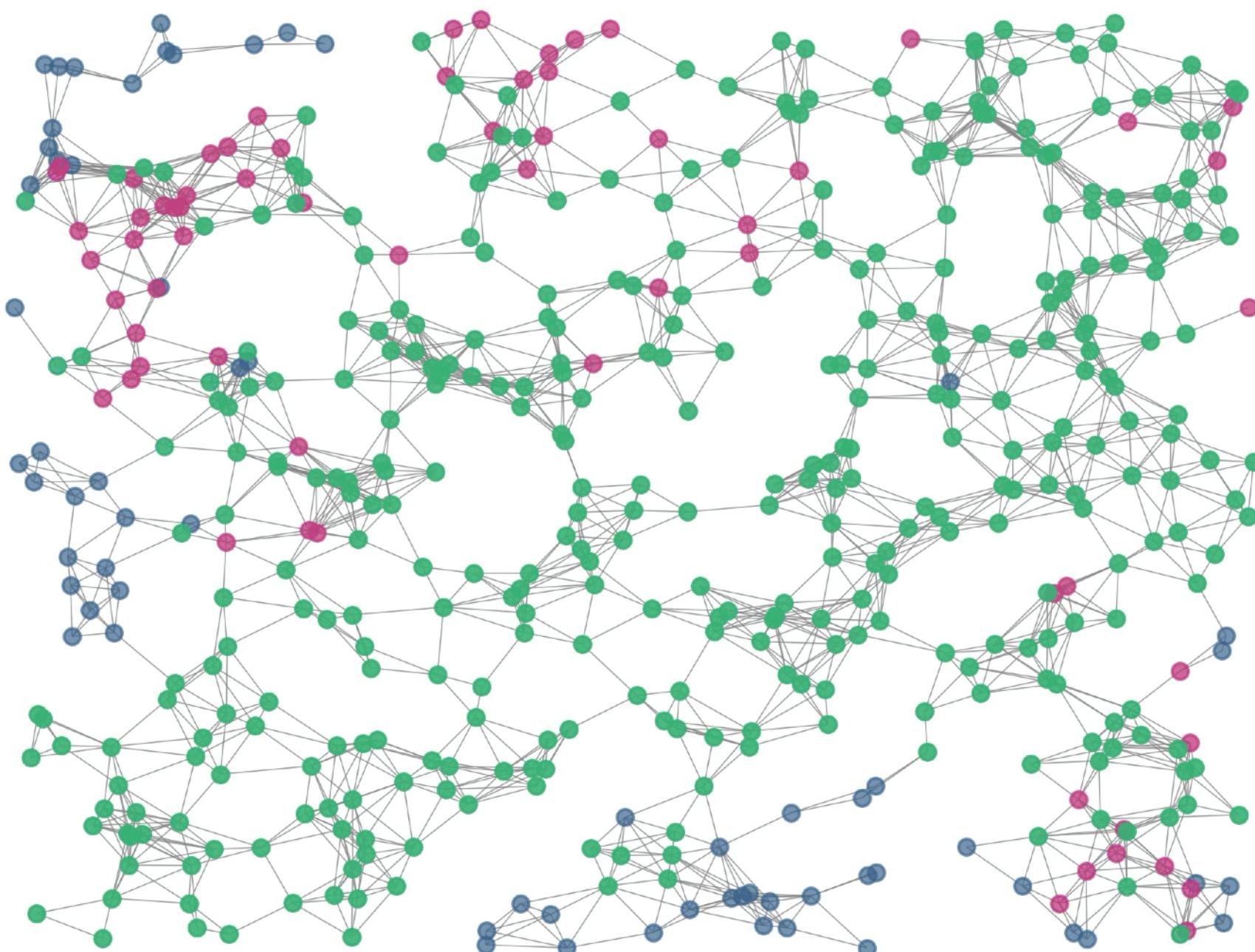
- ▶ Nodes are **susceptible**, **infected**, **recovered**.
- ▶  **Infected** nodes infect **susceptible** neighbors.
- ▶  **Infected** nodes recover (gain immunity).

SIR

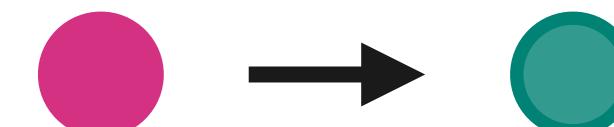
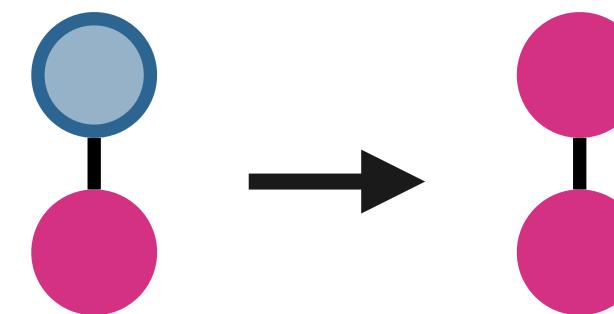


- ▶ Nodes are **susceptible**, **infected**, **recovered**.
- ▶  **Infected** nodes infect **susceptible** neighbors.
- ▶  **Infected** nodes recover (gain immunity).

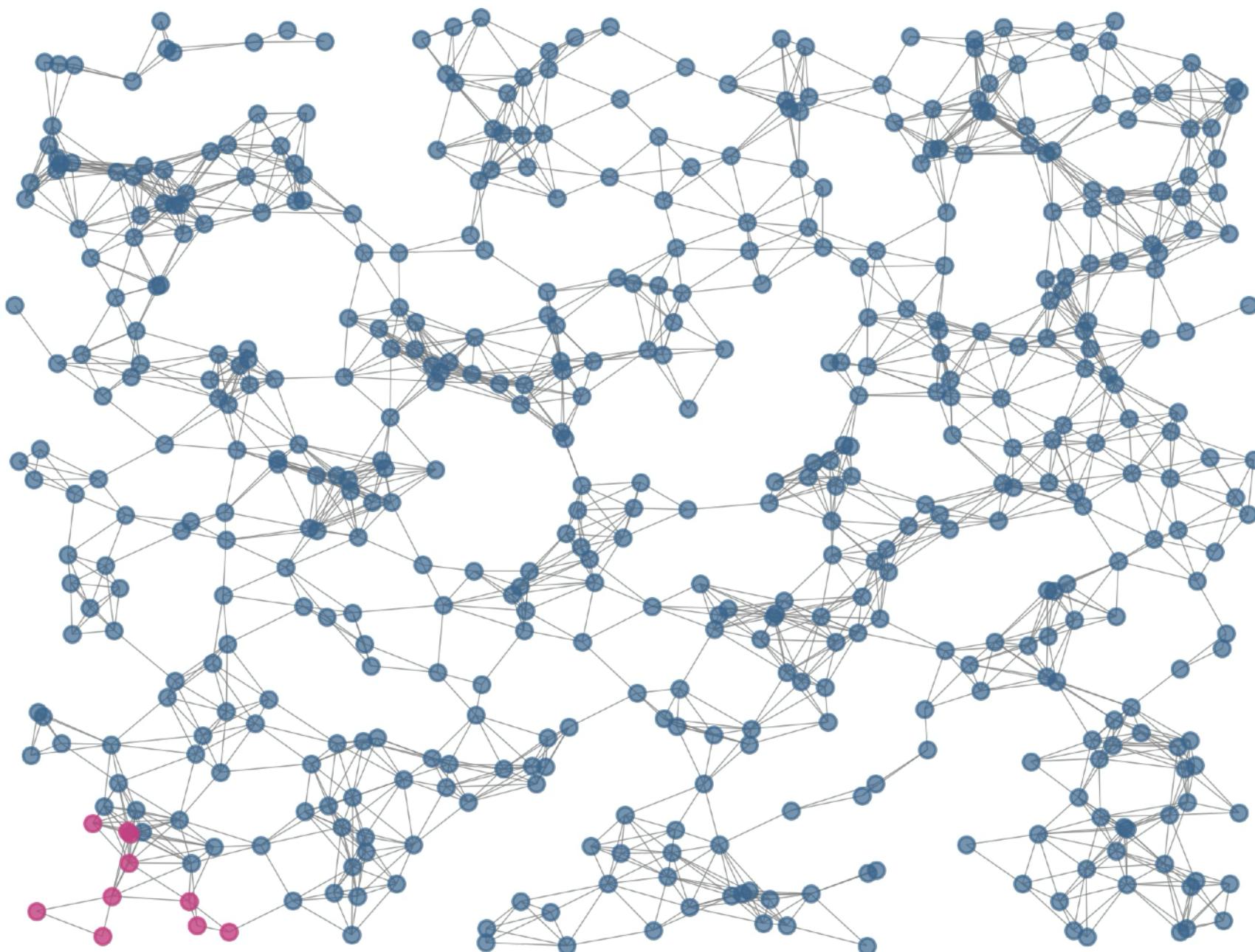
SIR



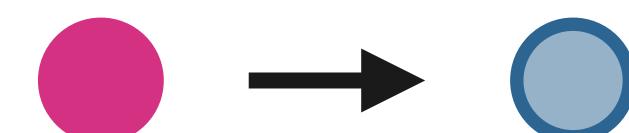
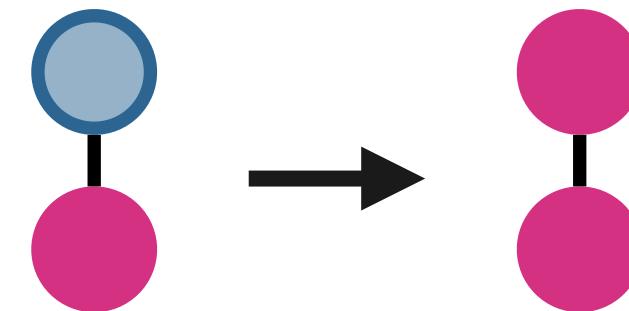
- ▶ Nodes are **susceptible**, **infected**, **recovered**.
- ▶ **Infected** nodes infect **susceptible** neighbors.
- ▶ **Infected** nodes recover (gain immunity).



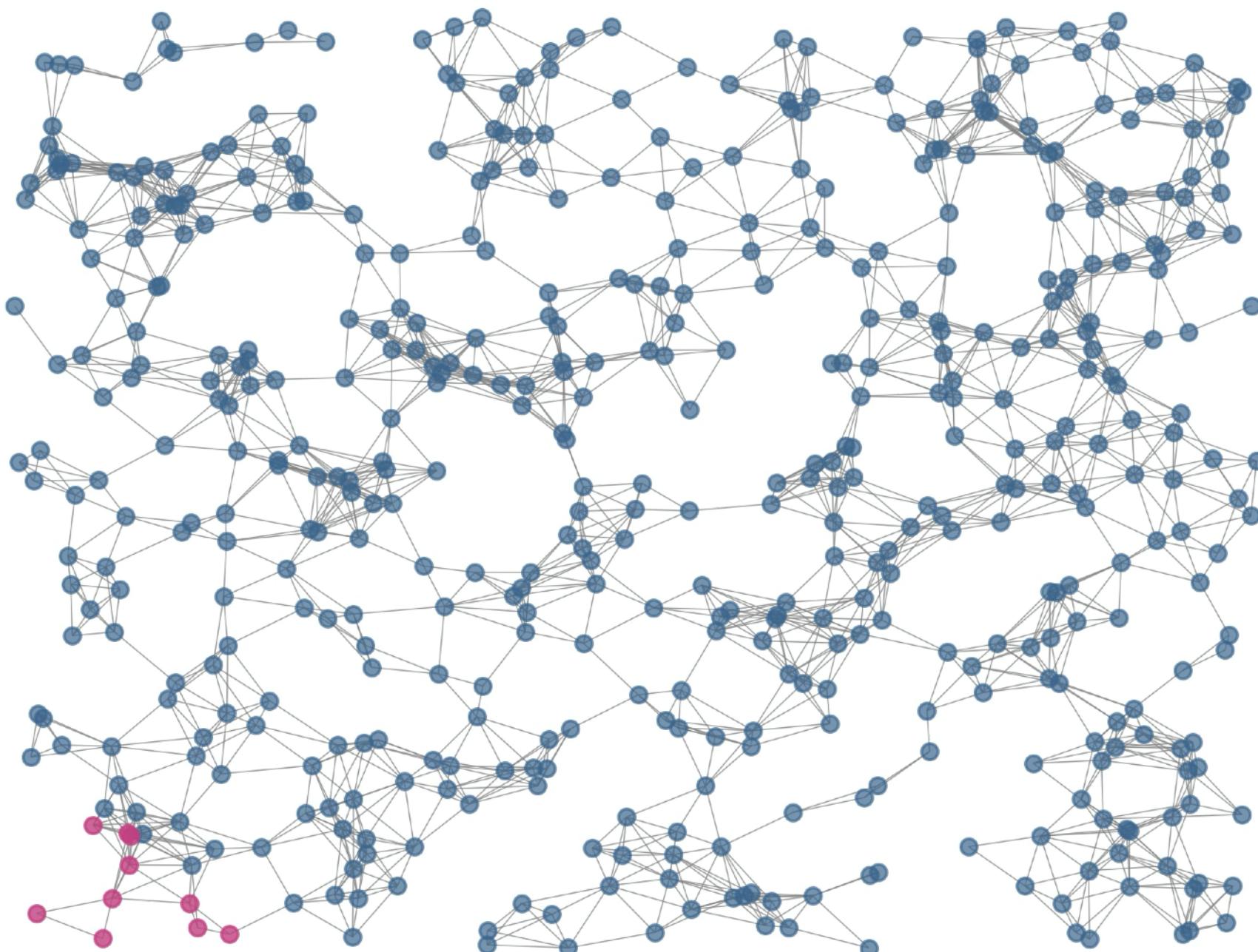
SIS



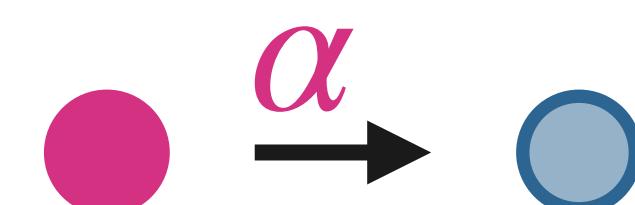
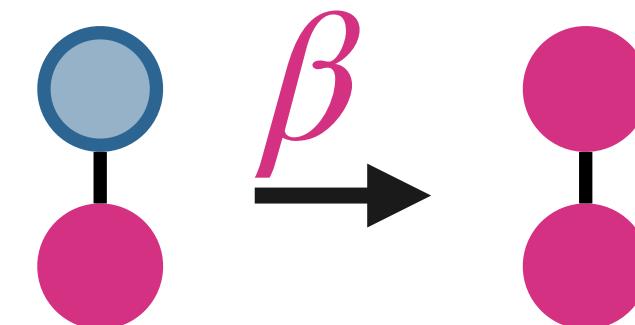
- ▶ Nodes are **infected** or **susceptible**.
- ▶ **Infected** nodes infect **susceptible** neighbors.
- ▶ **Infected** nodes recover (no immunity).



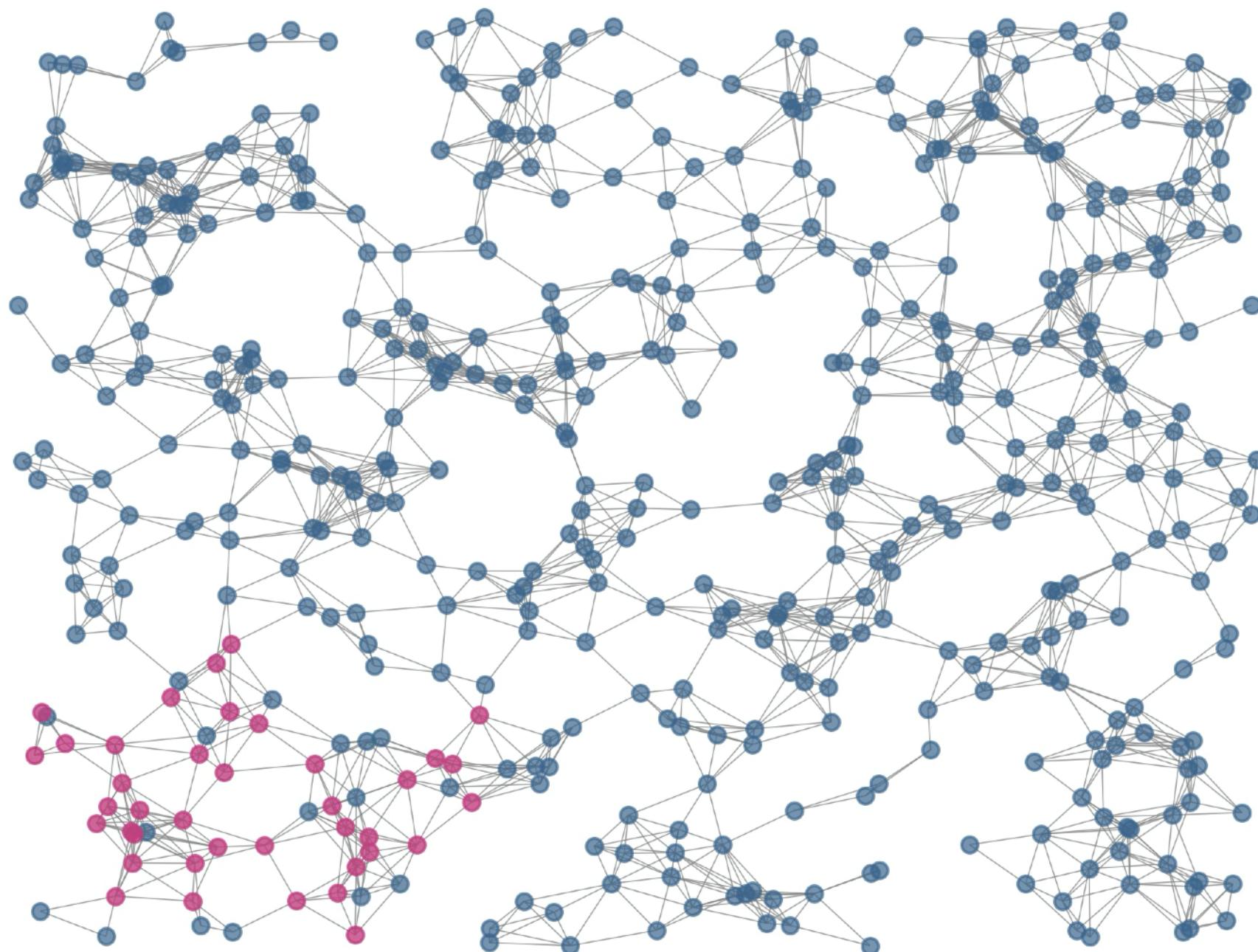
SIS



- ▶ Nodes are infected or susceptible.
- ▶ Infected nodes infect susceptible neighbors.
- ▶ Infected nodes recover (no immunity).
- ▶ **Continuous-time stochastic jump process**

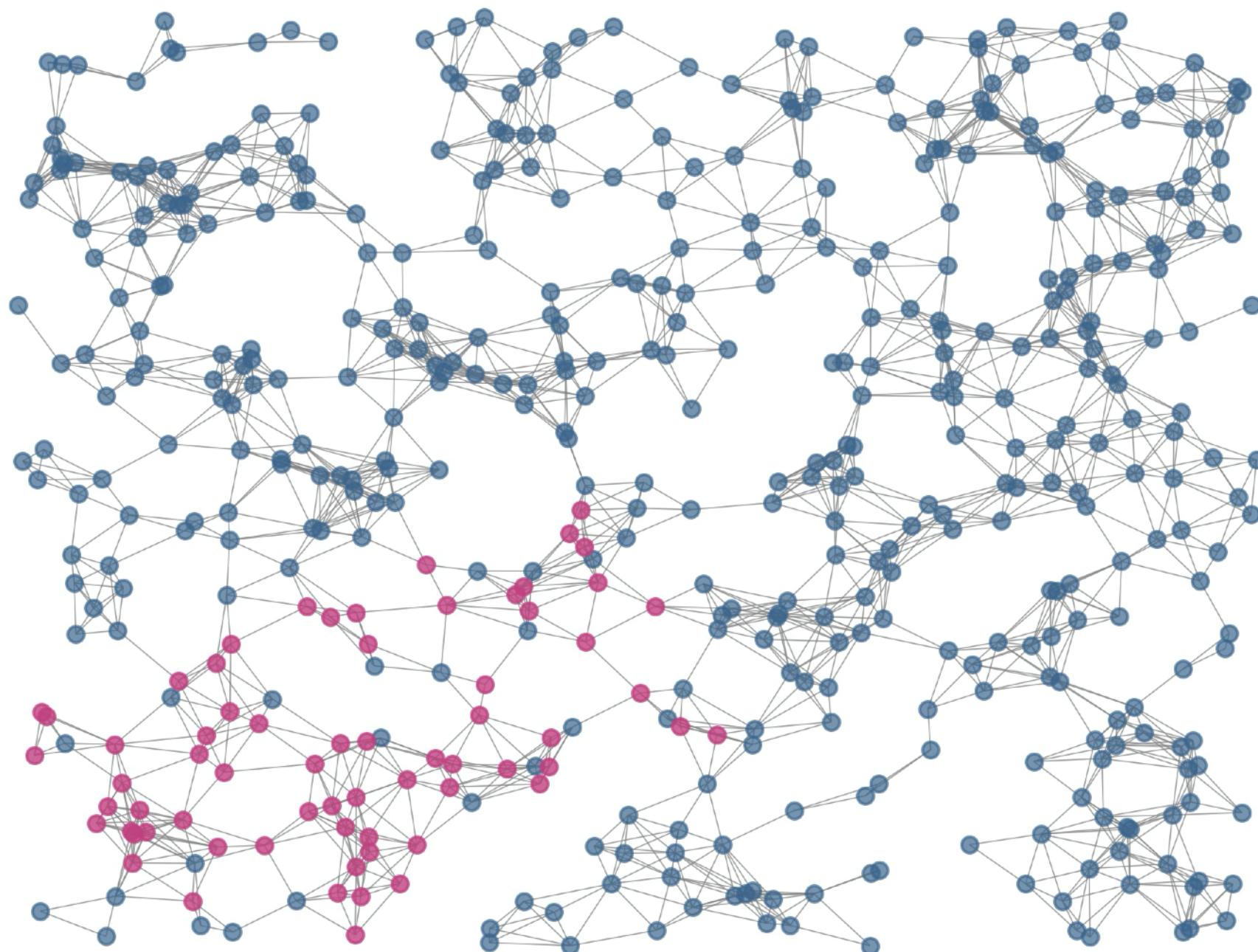


SIS

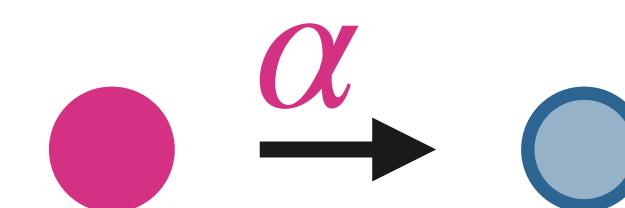
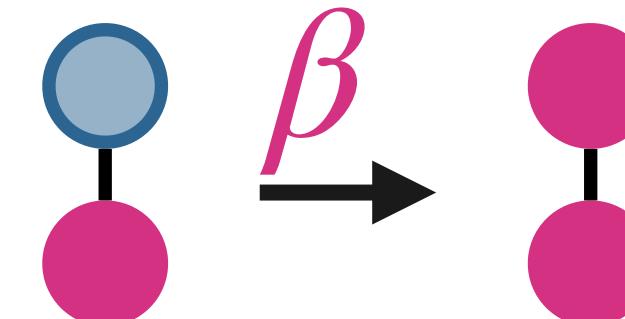


- ▶ Nodes are infected or susceptible.
 - ▶ Infected nodes infect susceptible neighbors.
 - ▶ Infected nodes recover (no immunity).
 - ▶ **Continuous-time stochastic jump process**
-

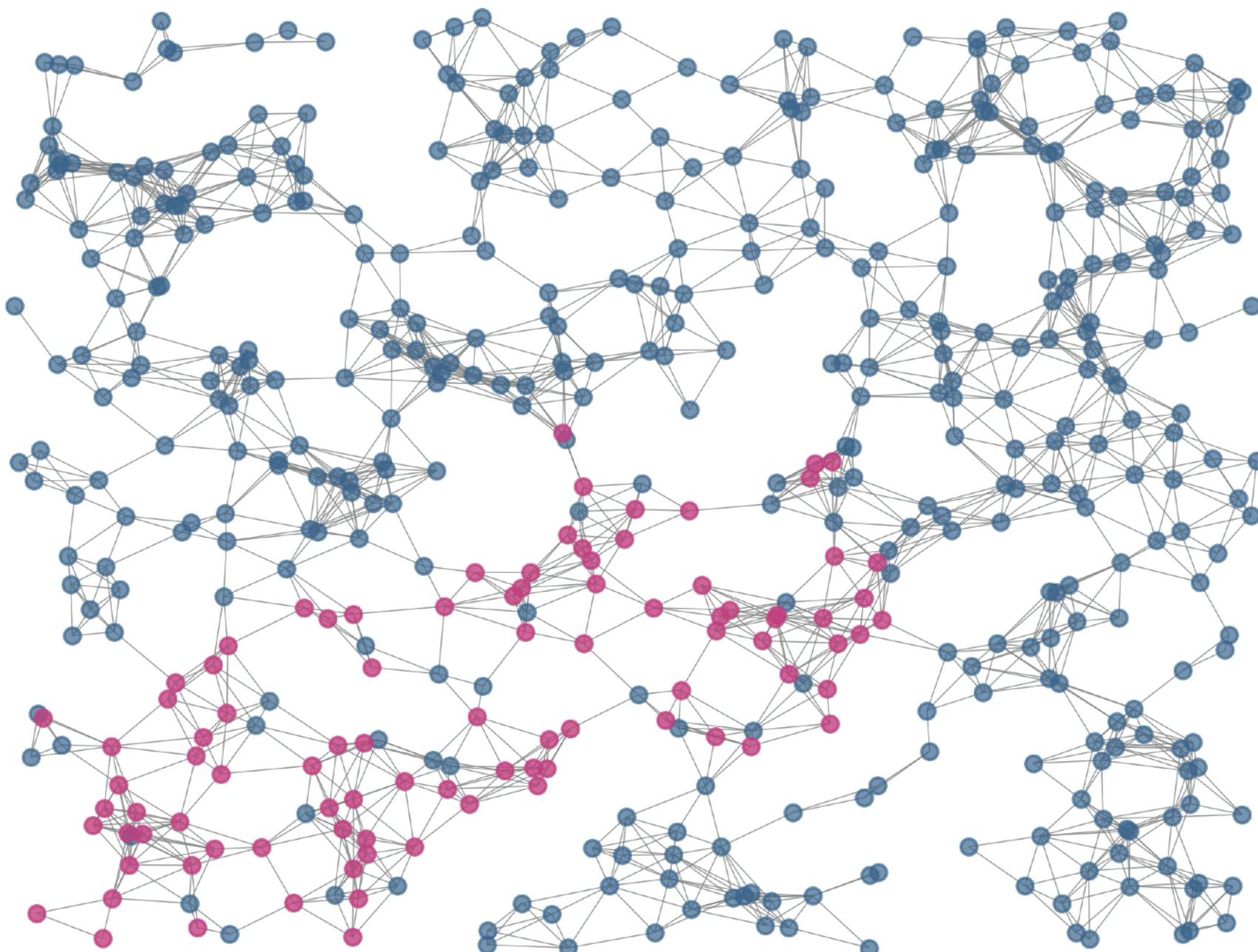
SIS



- ▶ Nodes are infected or susceptible.
- ▶ Infected nodes infect susceptible neighbors.
- ▶ Infected nodes recover (no immunity).
- ▶ **Continuous-time stochastic jump process**

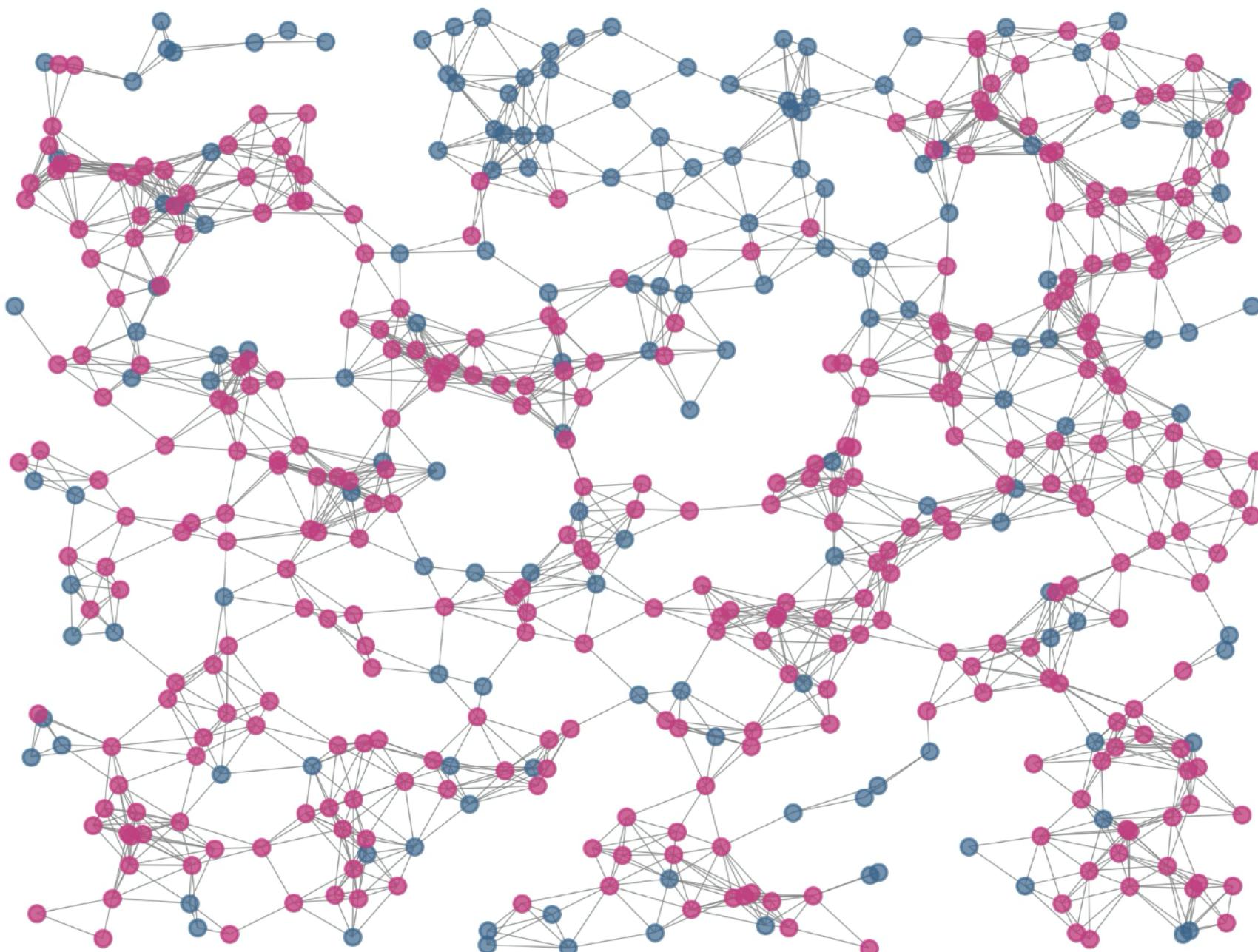


SIS

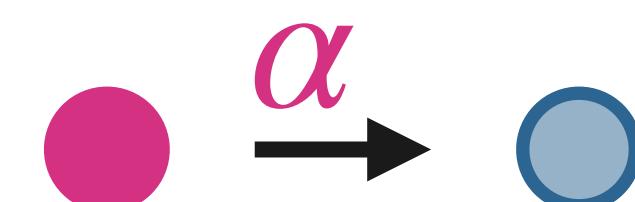
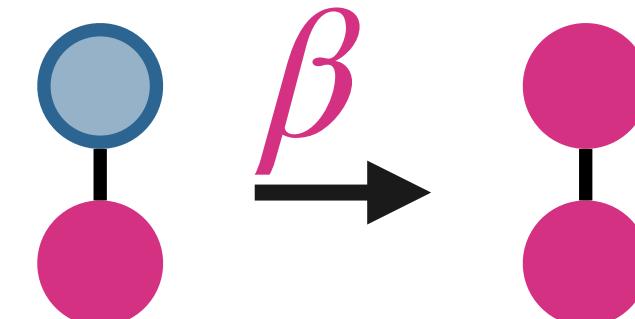


- ▶ Nodes are infected or susceptible.
 - ▶ Infected nodes infect susceptible neighbors.
 - ▶ Infected nodes recover (no immunity).
 - ▶ **Continuous-time stochastic jump process**
- β
- α

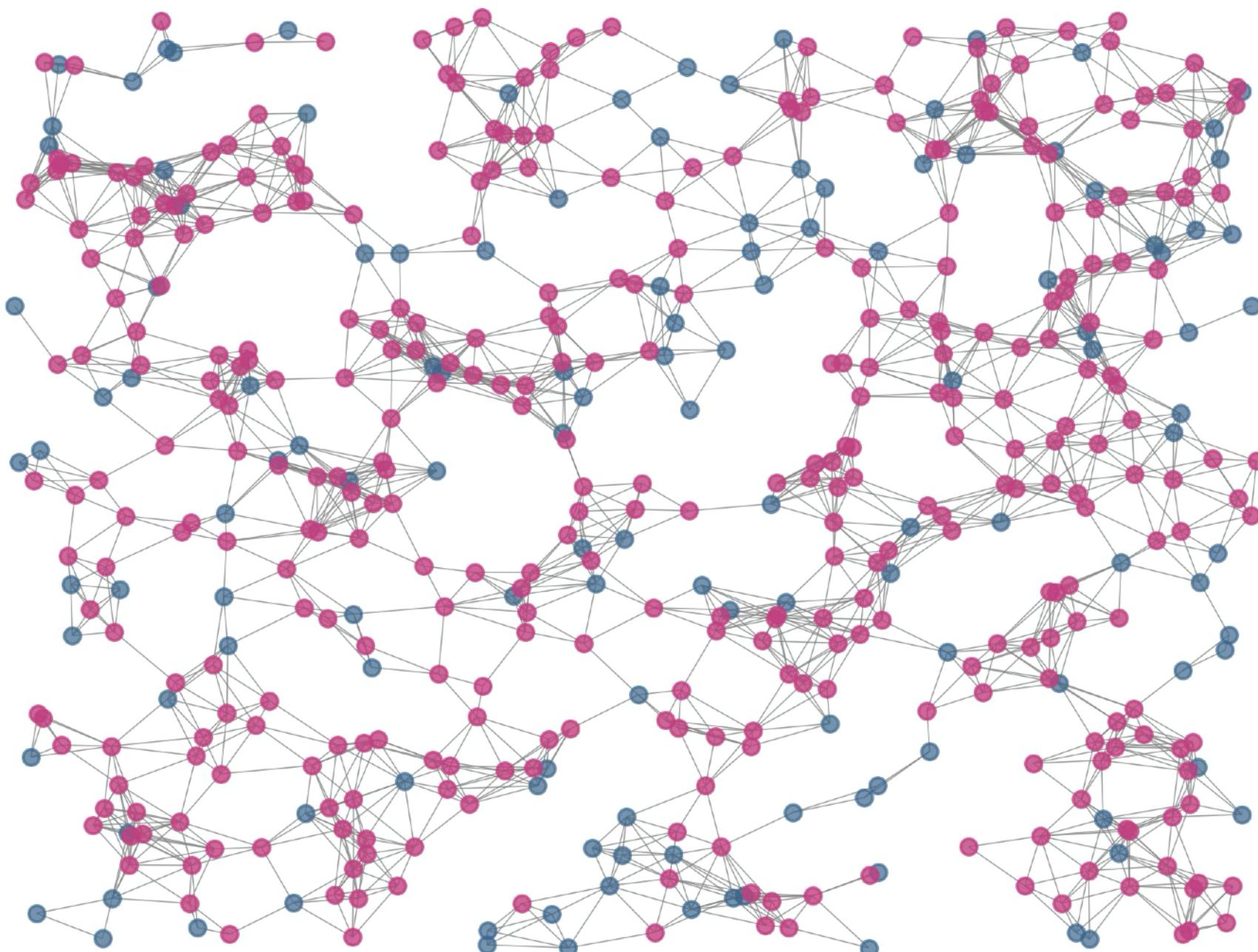
SIS



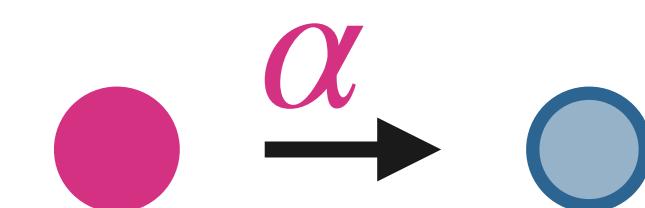
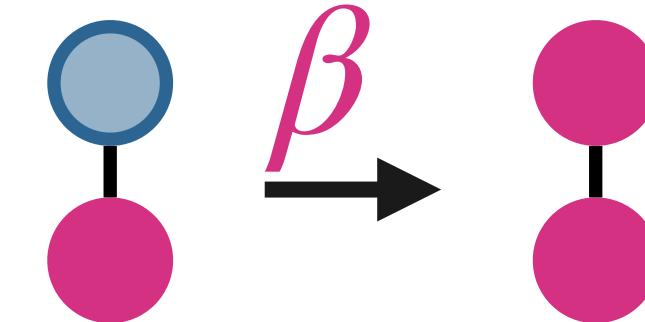
- ▶ Nodes are infected or susceptible.
- ▶ Infected nodes infect susceptible neighbors.
- ▶ Infected nodes recover (no immunity).
- ▶ **Continuous-time stochastic jump process**



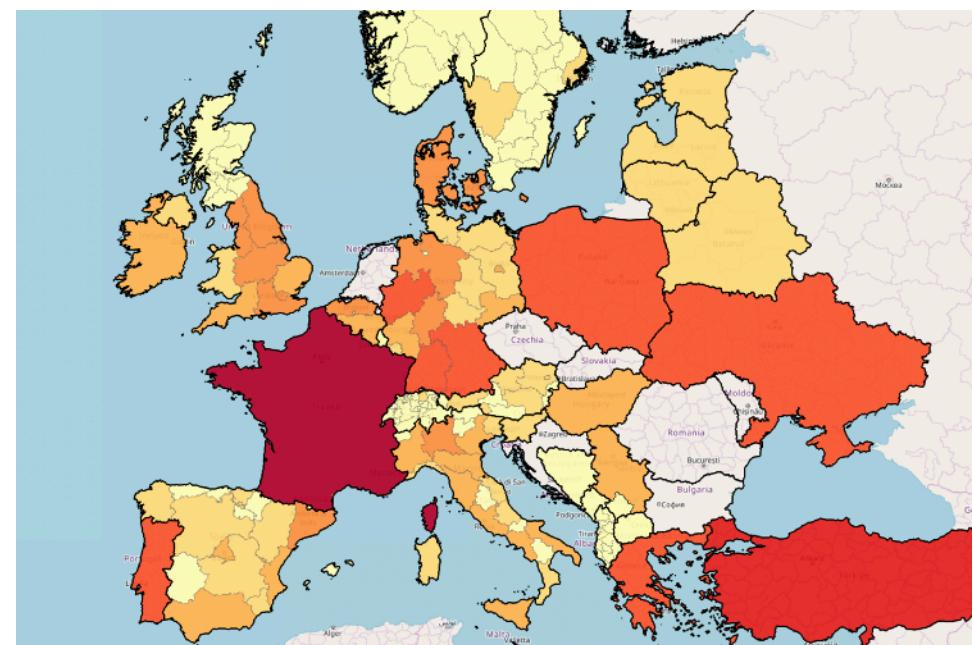
SIS



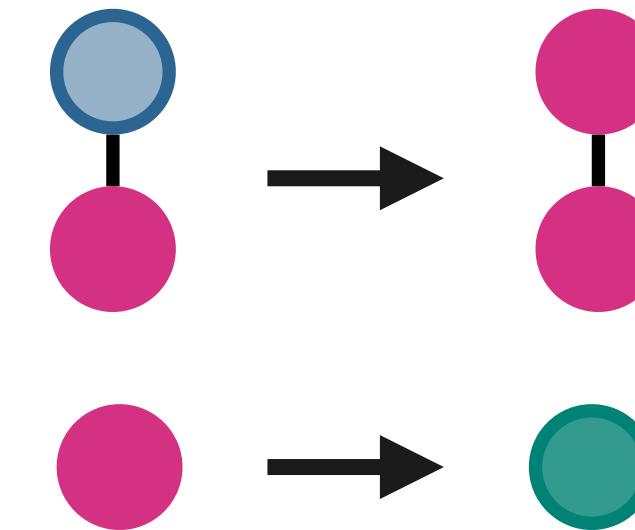
- ▶ Nodes are infected or susceptible.
- ▶ Infected nodes infect susceptible neighbors.
- ▶ Infected nodes recover (no immunity).
- ▶ **Continuous-time stochastic jump process**



Pipeline



Observations

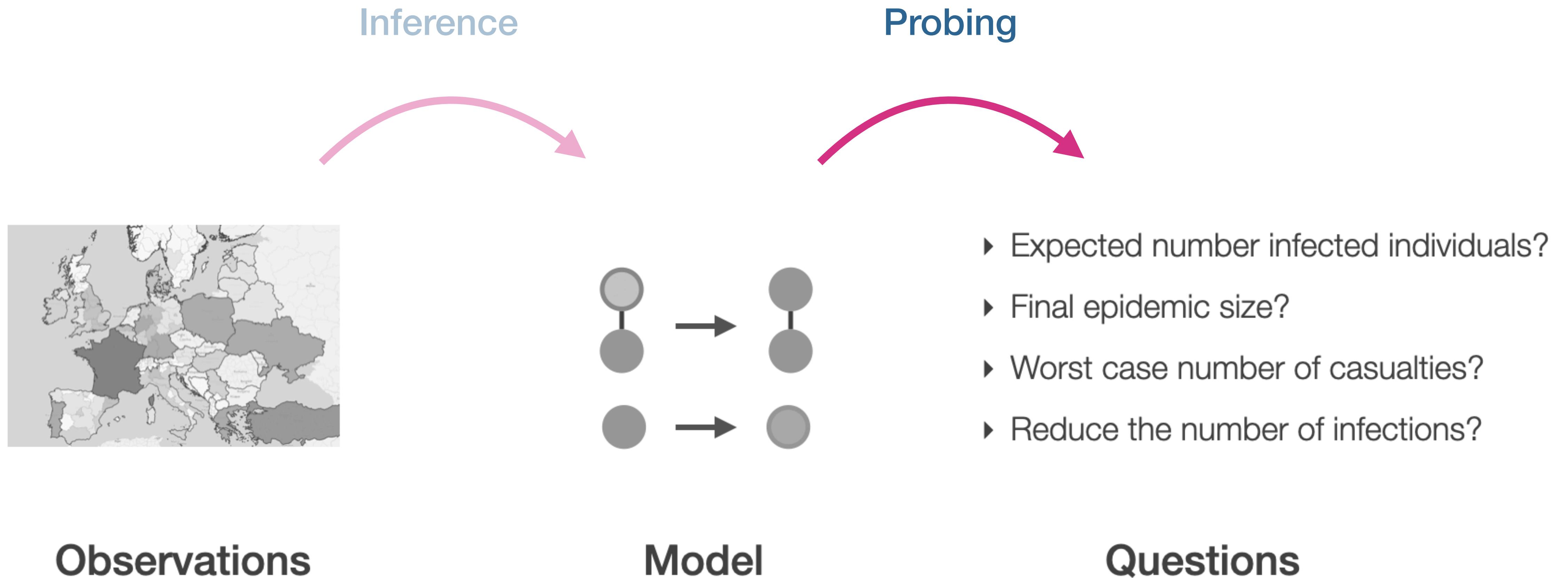


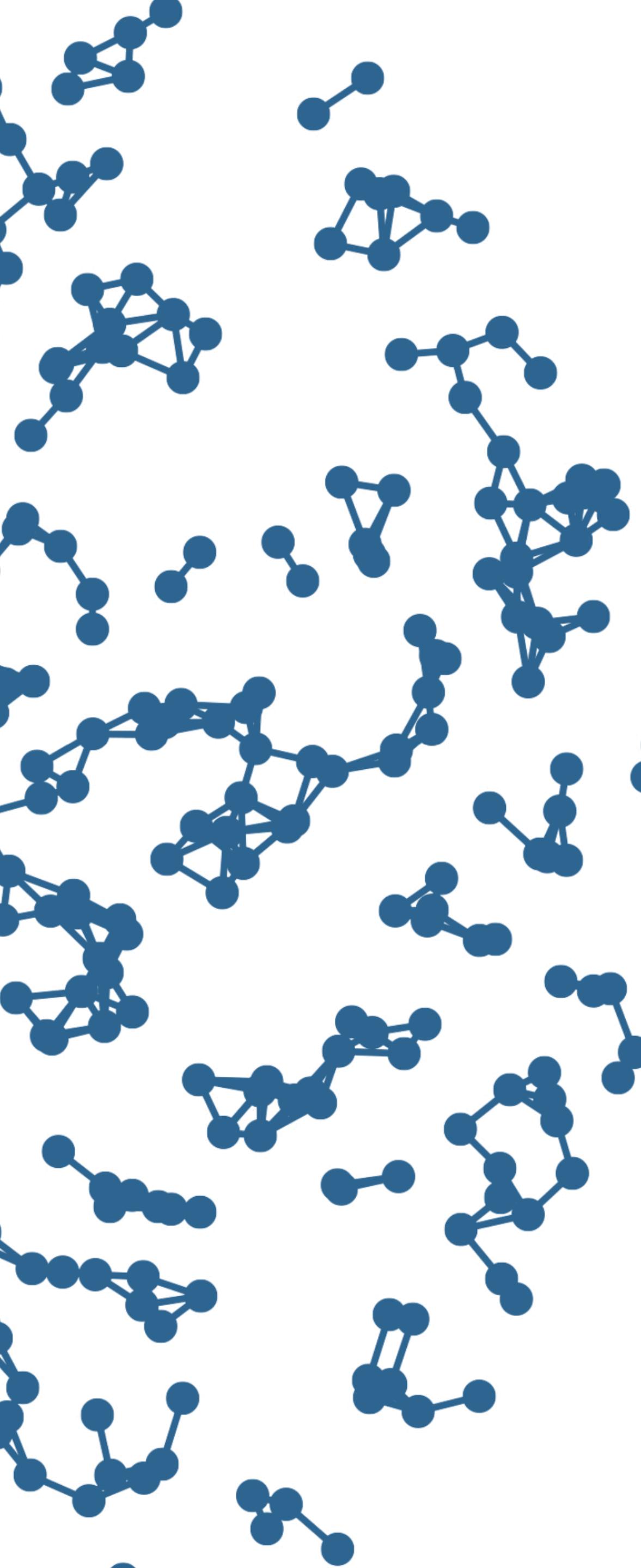
Model

- ▶ Expected number infected individuals?
- ▶ Final epidemic size?
- ▶ Worst case number of casualties?
- ▶ Reduce the number of infections?

Questions

Pipeline





Agenda

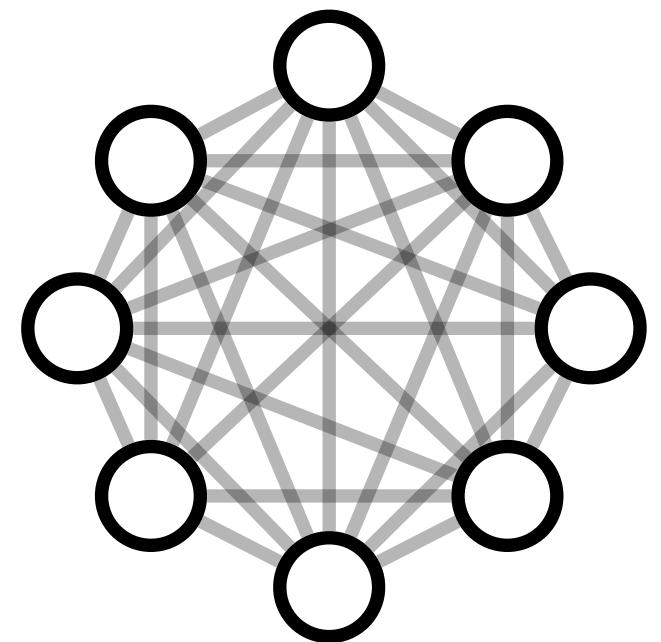
METHODS

COVID

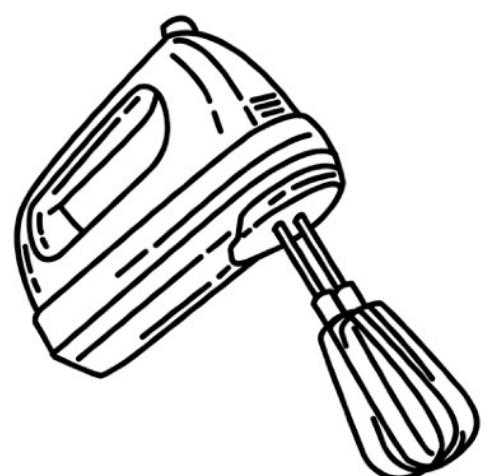
(NON-)MARKOVIAN SIMULATION
CONTROL
INFERENCE

Networks - Why Do We Care?

Homogenous connectivity:



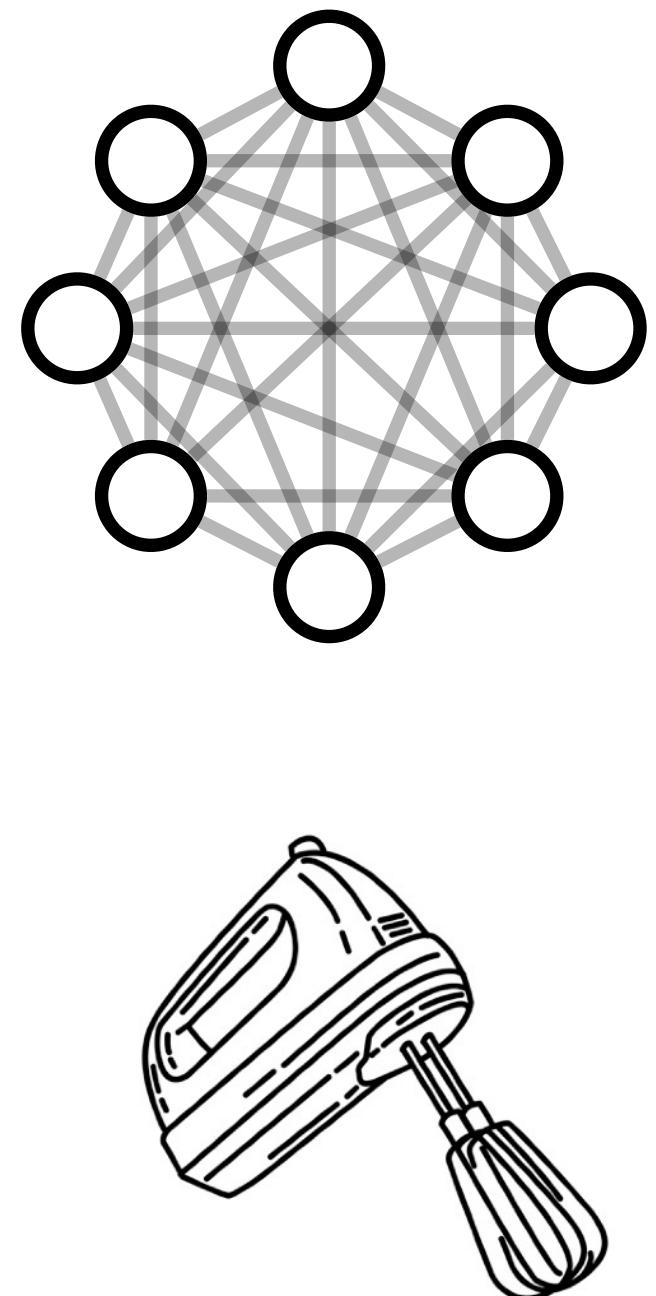
Every one is connected with everyone else “a little bit”.



Well-stirred analogy

Networks - Why Do We Care?

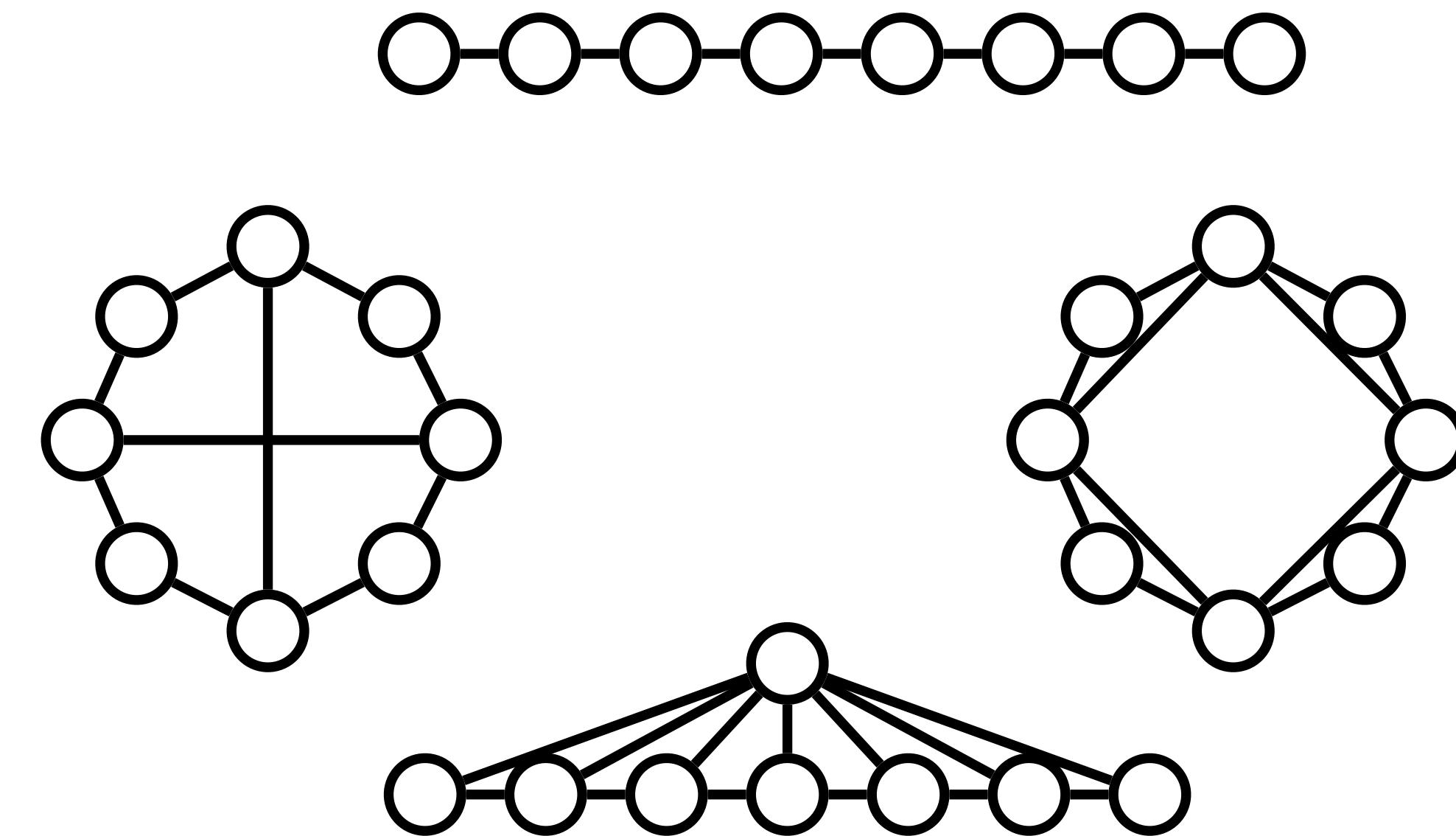
Homogenous connectivity:



Every one is connected with everyone else “a little bit”.

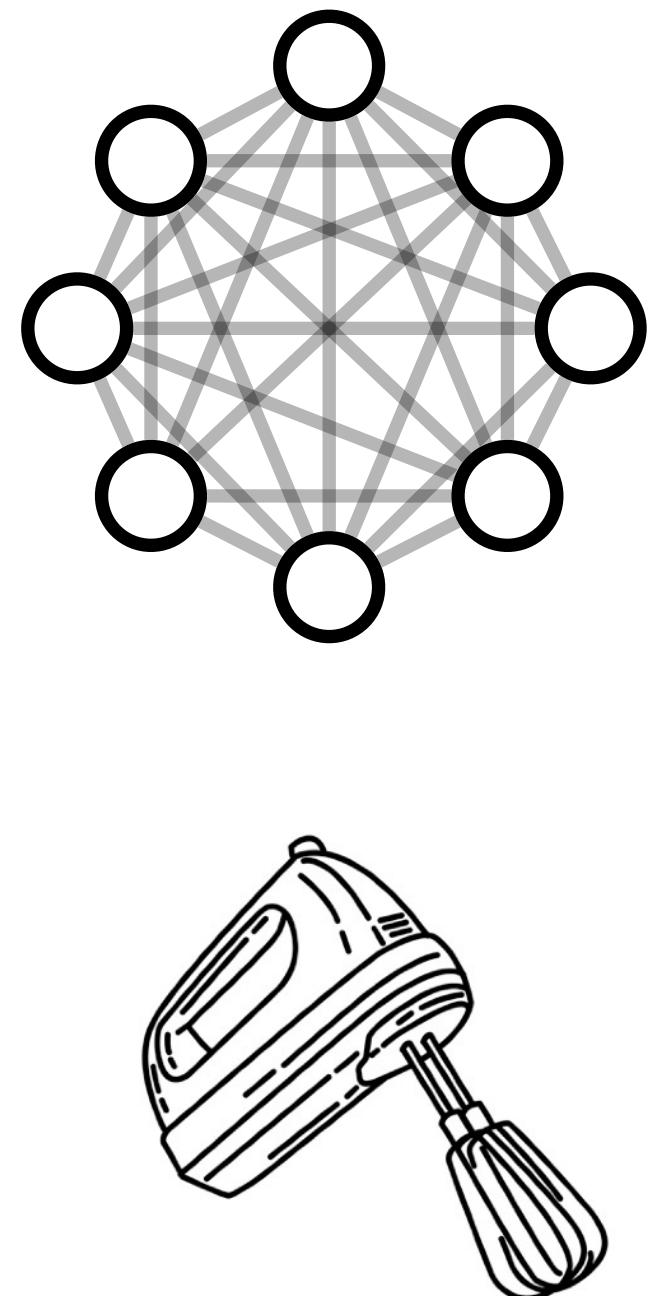
Well-stirred analogy

Network-constrained connectivity:



Networks - Why Do We Care?

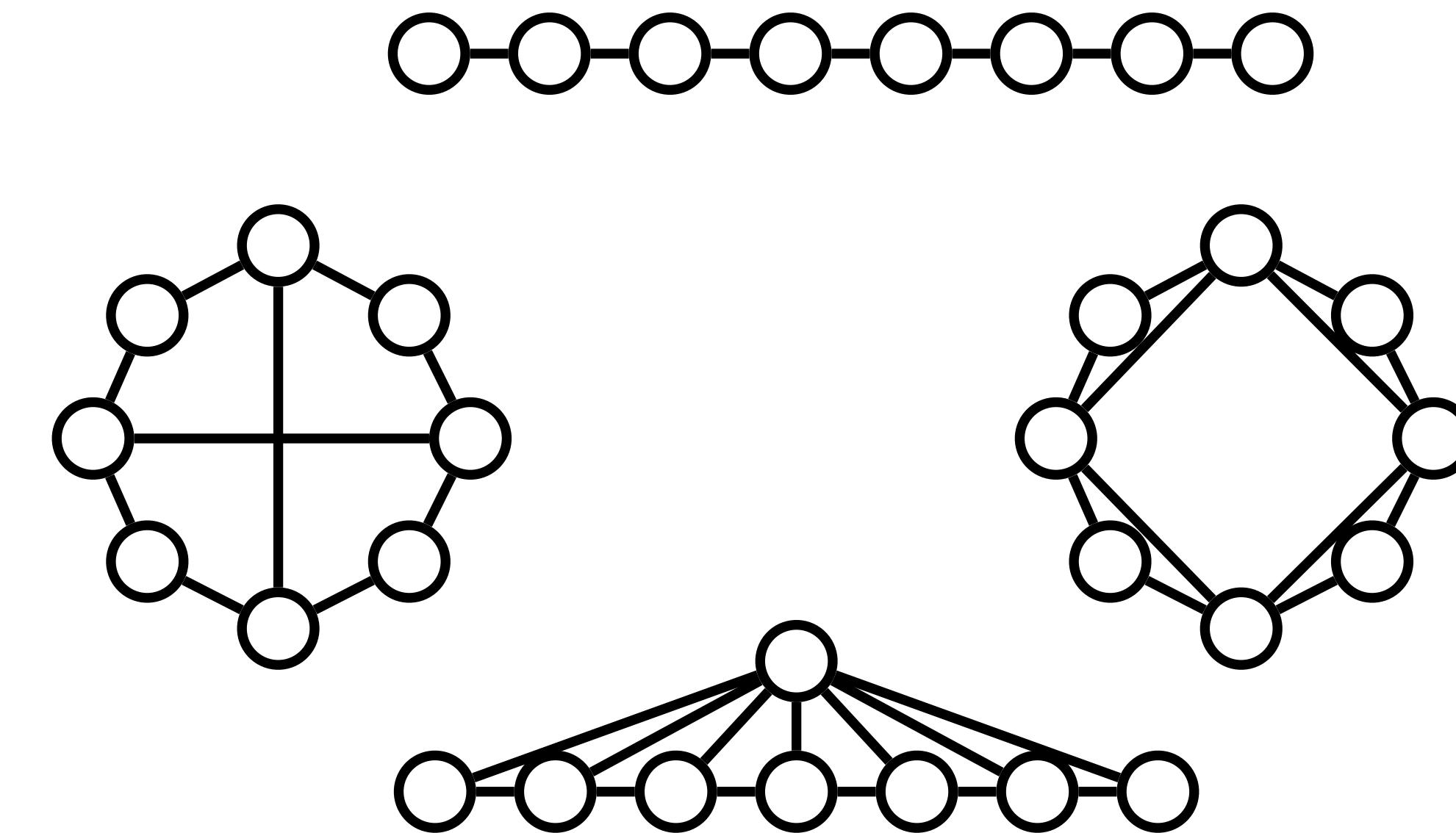
Homogenous connectivity:



Every one is connected with everyone else “a little bit”.

Well-stirred analogy

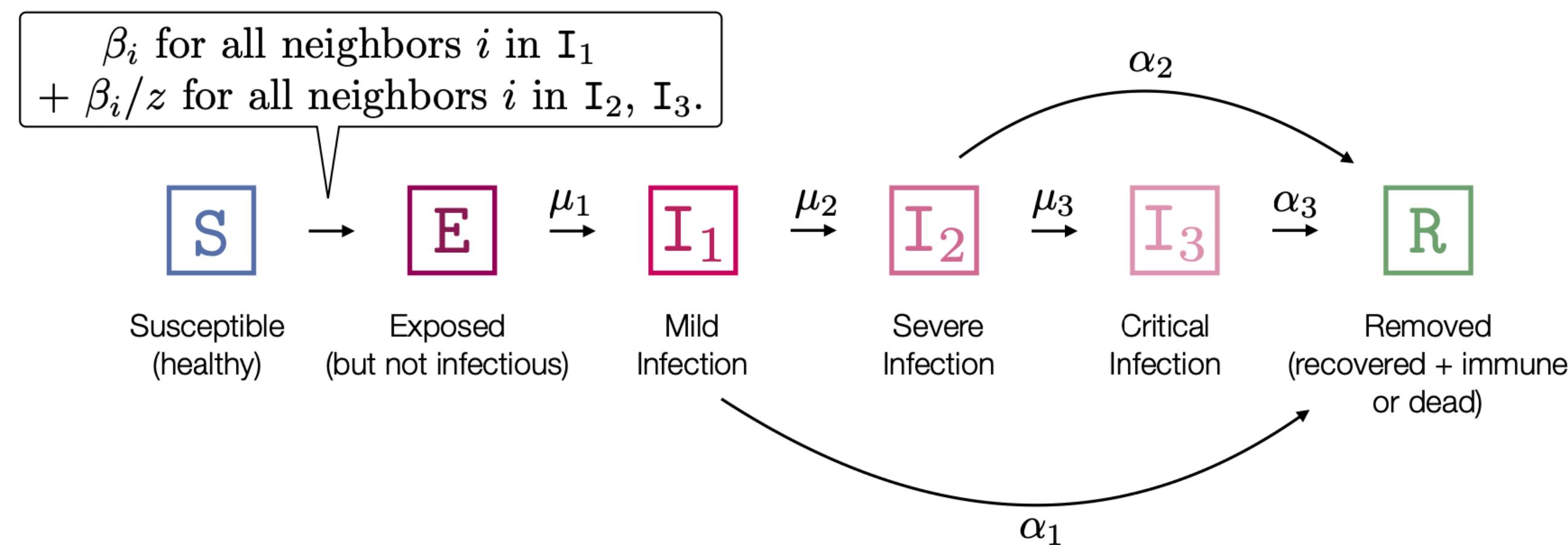
Network-constrained connectivity:



Spreading on network is very different from spreading in a homogenous population.

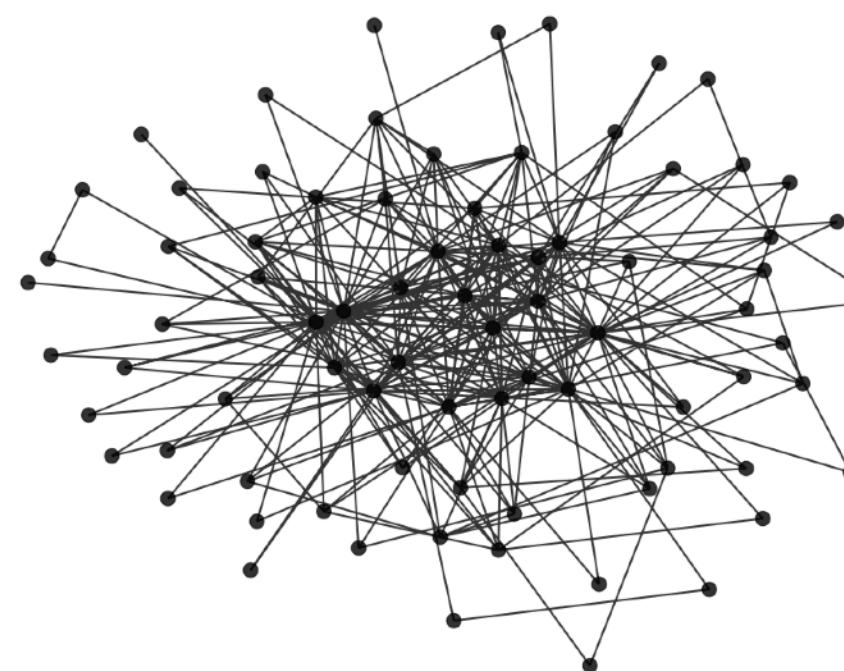
Covid-19

Adopt **SIR** model for Covid-19:



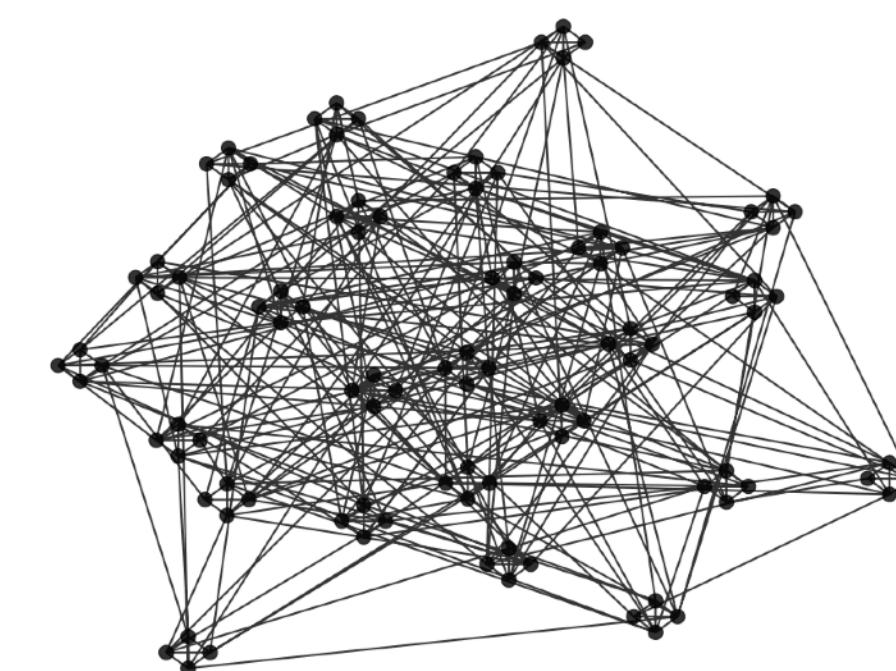
Covid-19

Test different networks types (fixed mean degree):



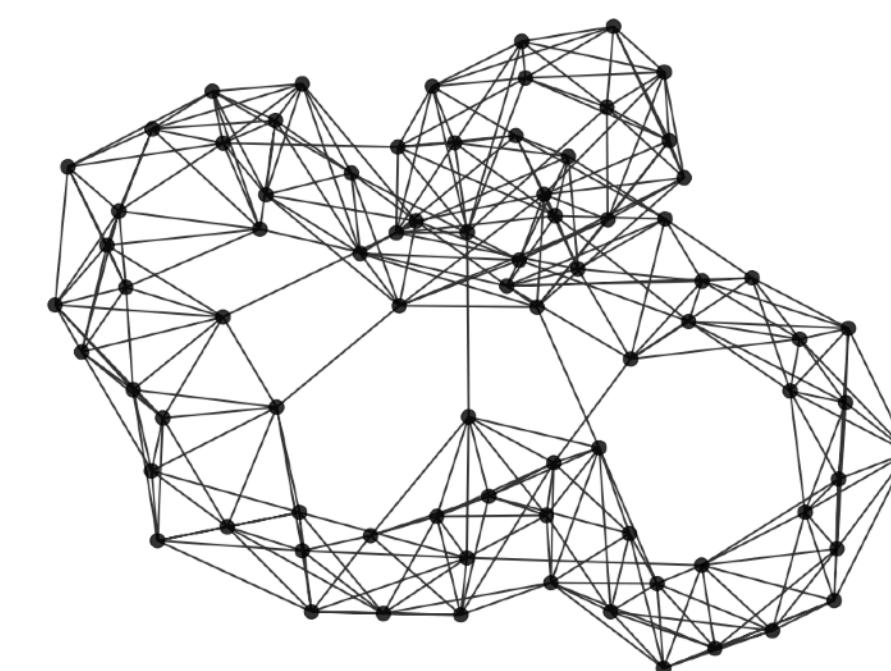
Power-law

- ▶ Small number of hubs
- ▶ Many nodes with very few neighbors



Household

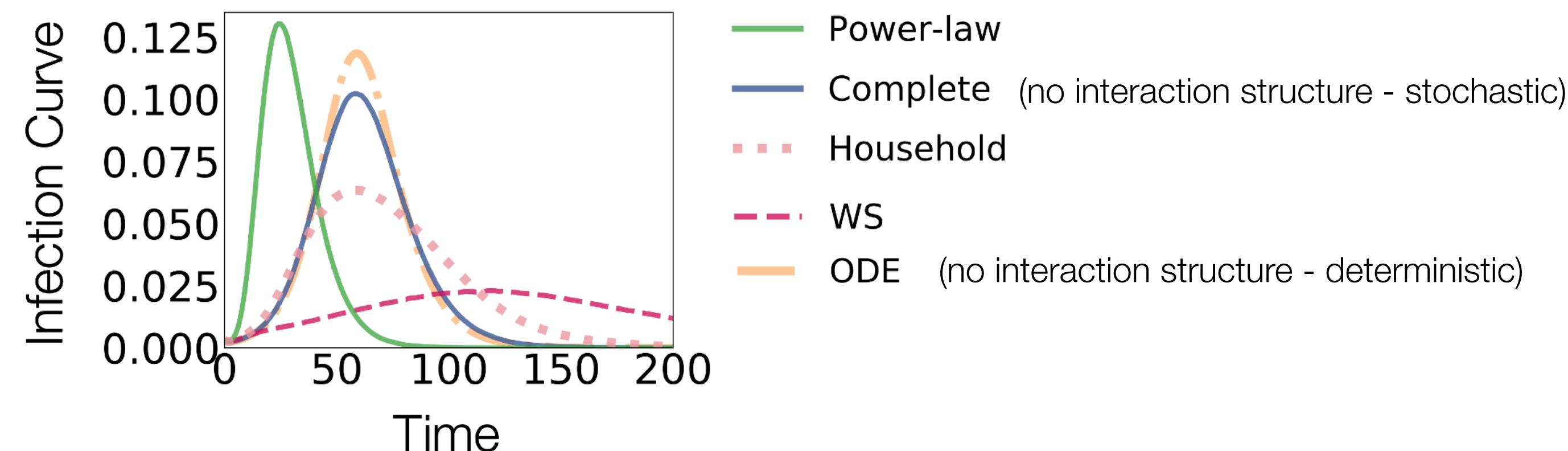
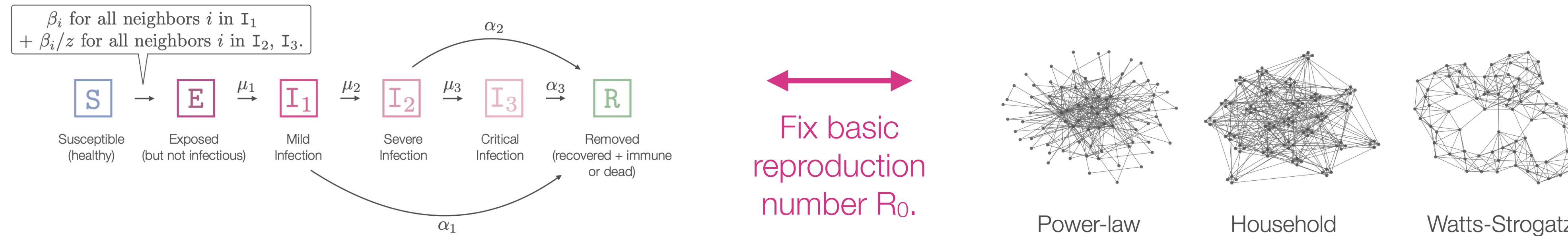
- ▶ Household size = 4
- ▶ Connected via a Geometric network



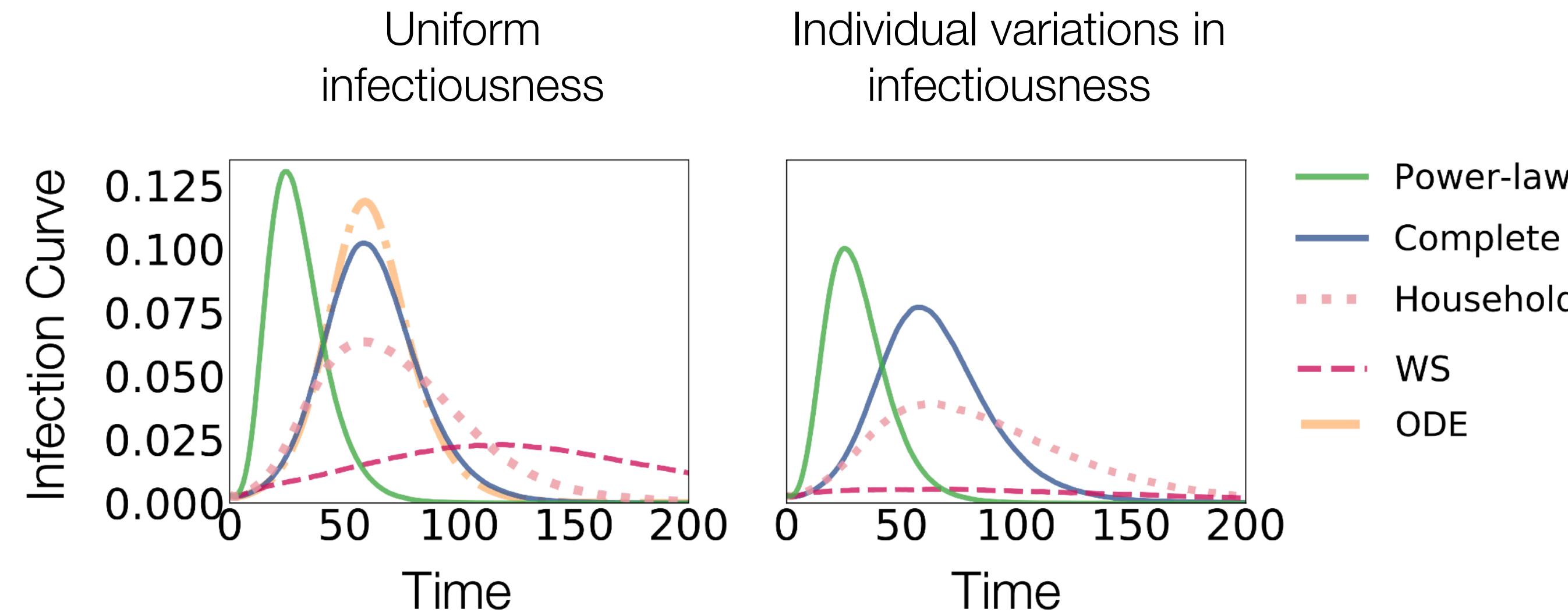
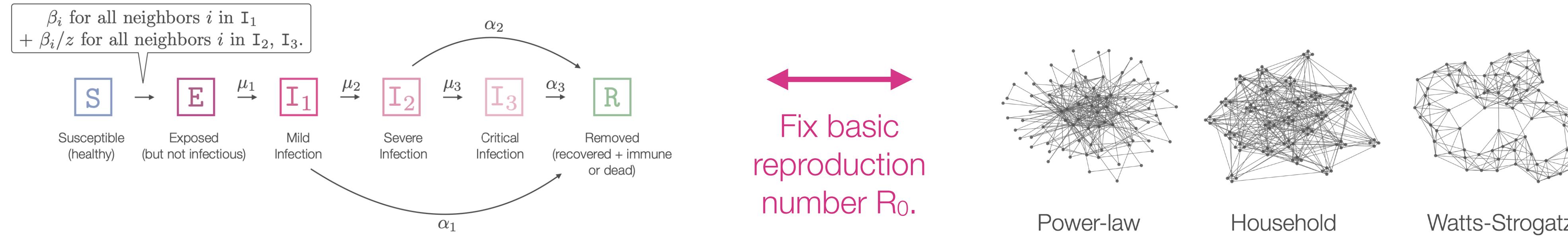
Watts-Strogatz

- ▶ Fixed number of contacts
- ▶ Small-world

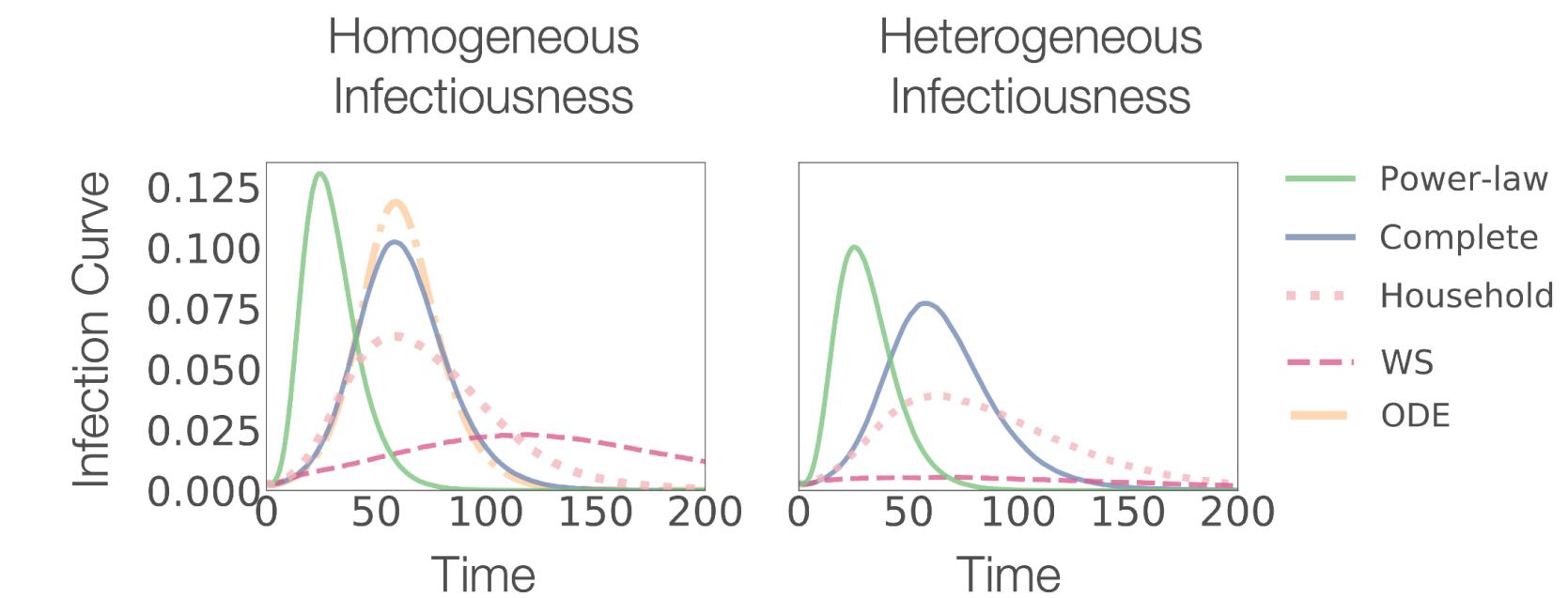
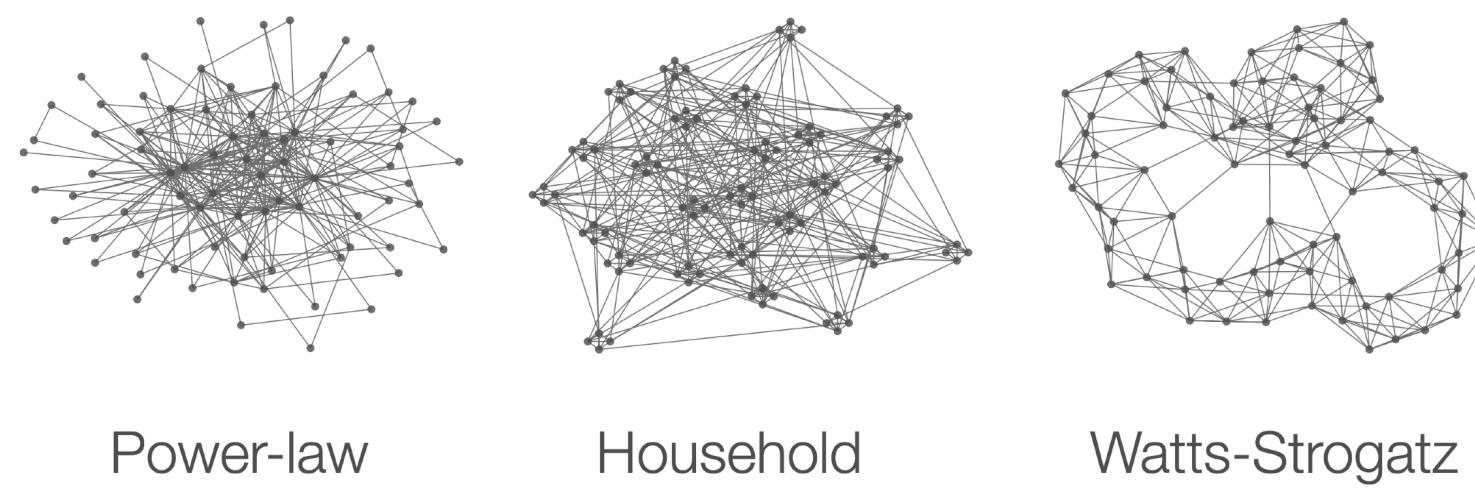
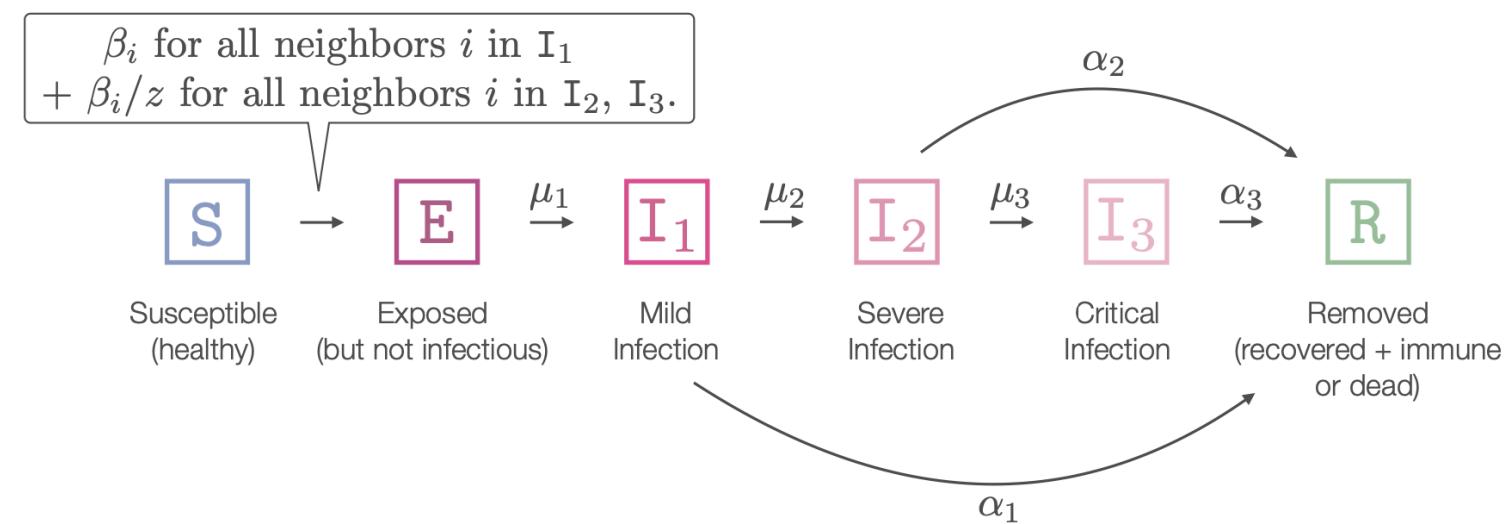
Covid-19



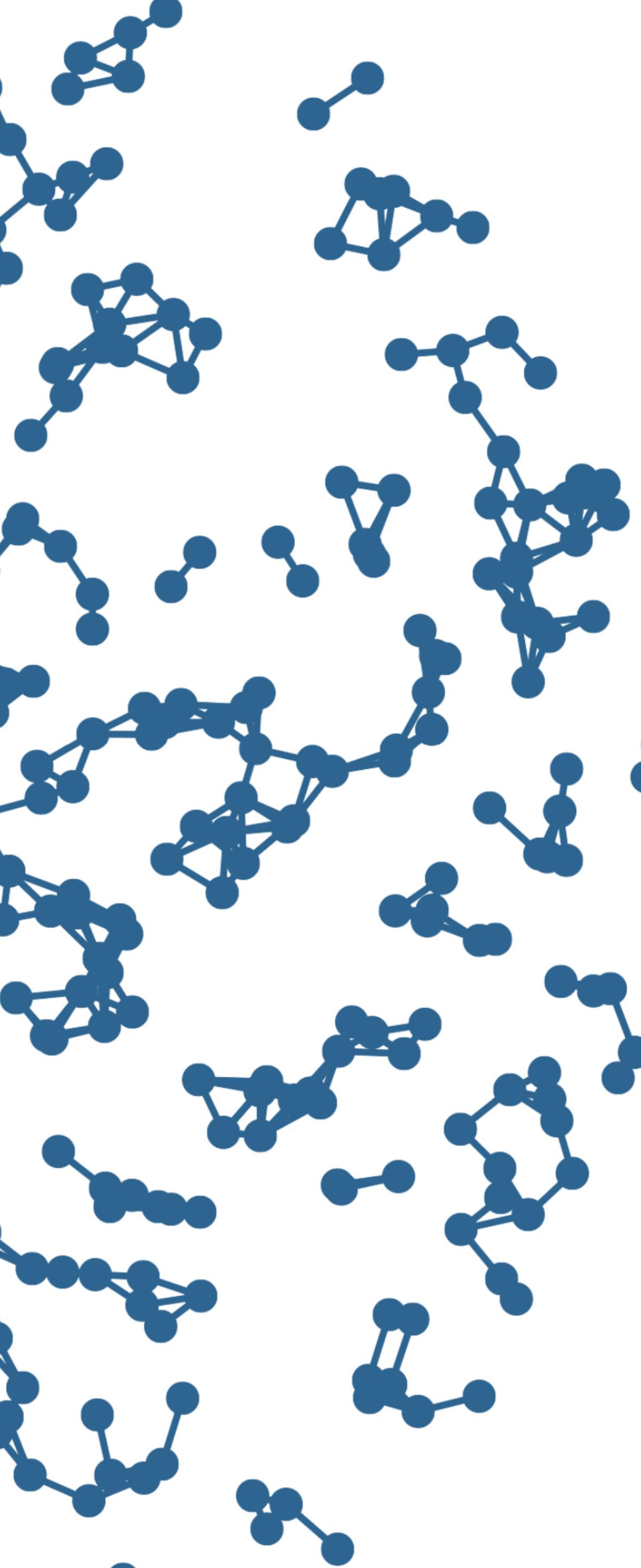
Covid-19



Covid-19



- ▶ **Do not use population averages** to describe a population (mean number of contacts, mean infectiousness).
- ▶ **Vast differences** in all relevant properties among networks (infection curve, R_t, dispersion, deaths)
- ▶ **Network structure** vs **individual differences** vs **environment** (complex interplay hinders analysis)



Agenda

METHODS

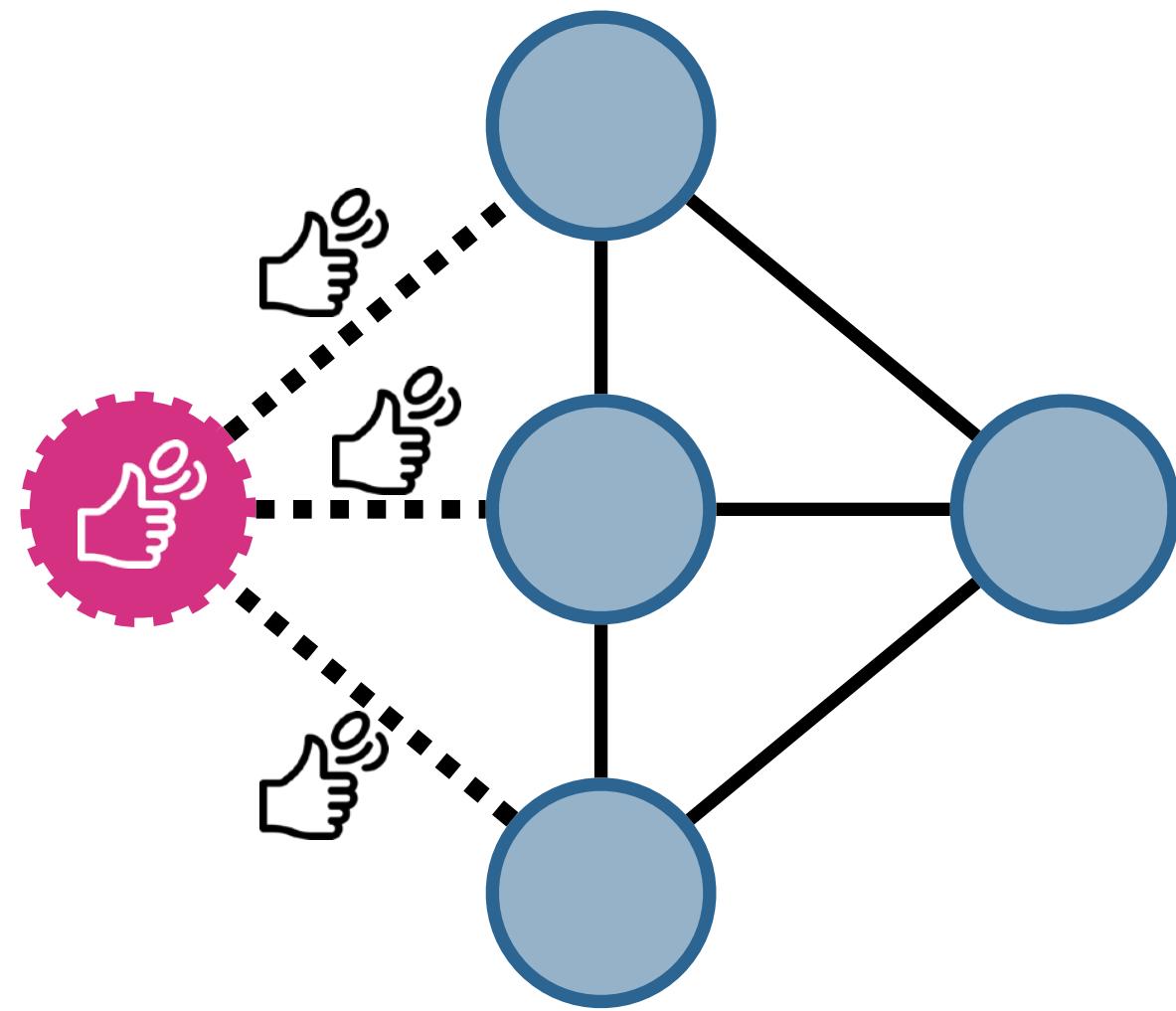
COVID

(NON-)MARKOVIAN SIMULATION

CONTROL
INFERENCE

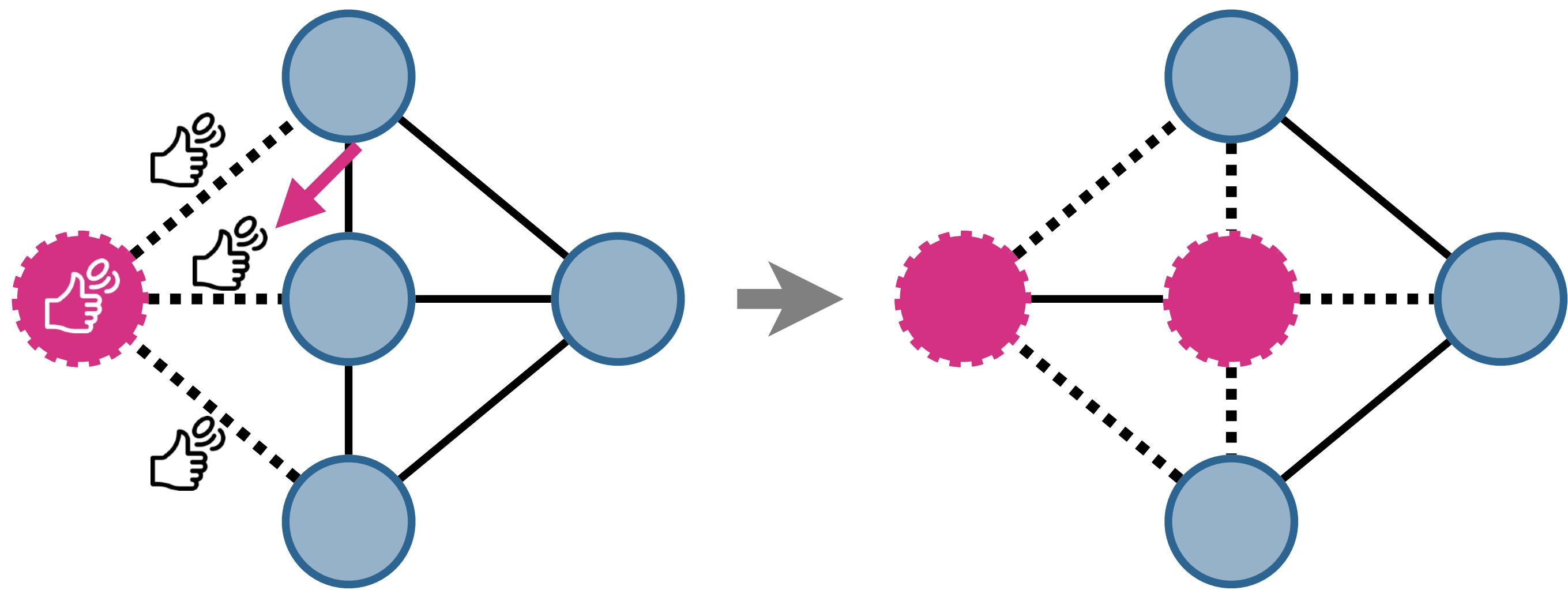
SIS Simulation

naïve:



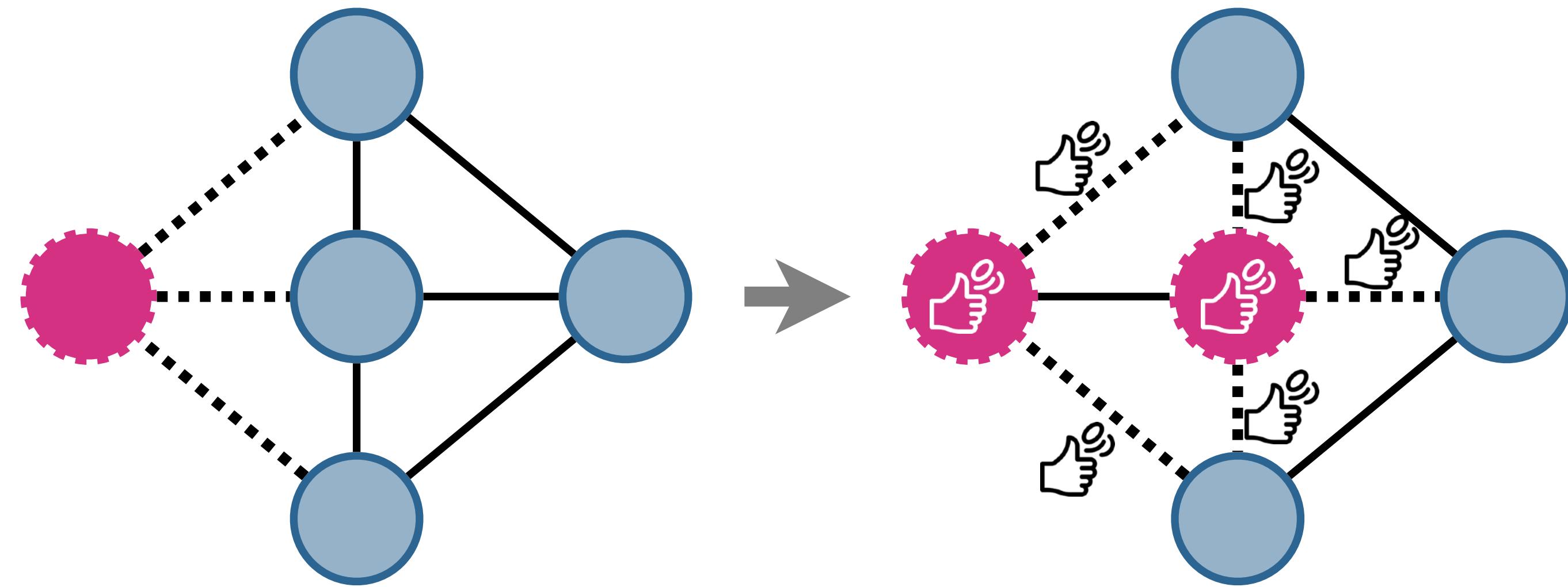
SIS Simulation

naïve:



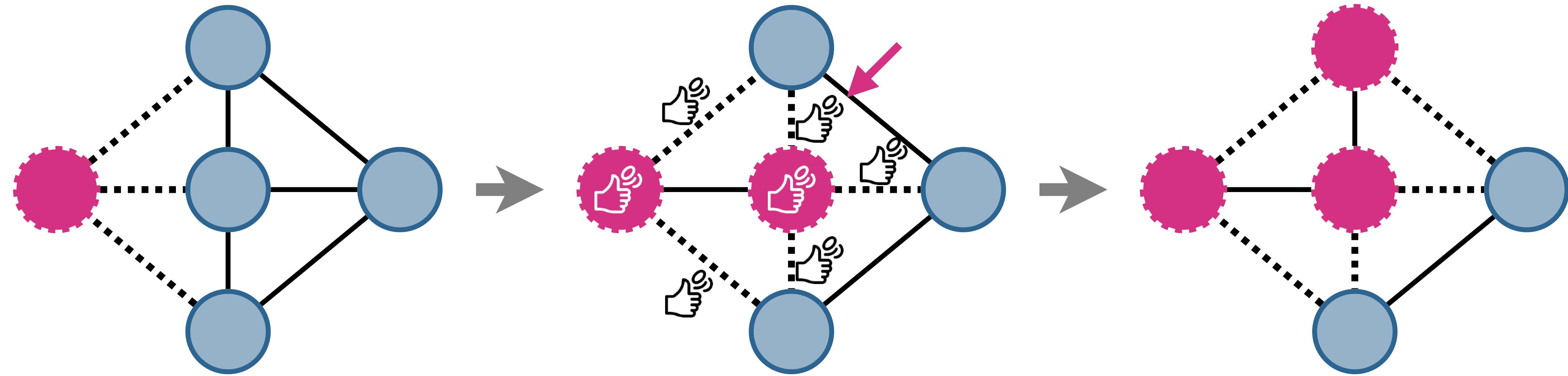
SIS Simulation

naïve:



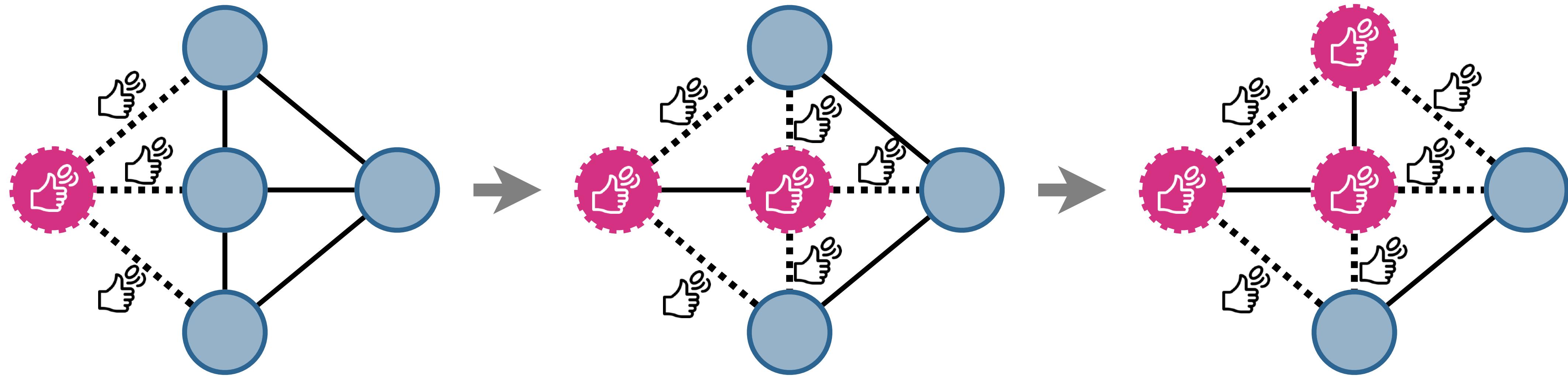
SIS Simulation

naïve:



SIS Simulation

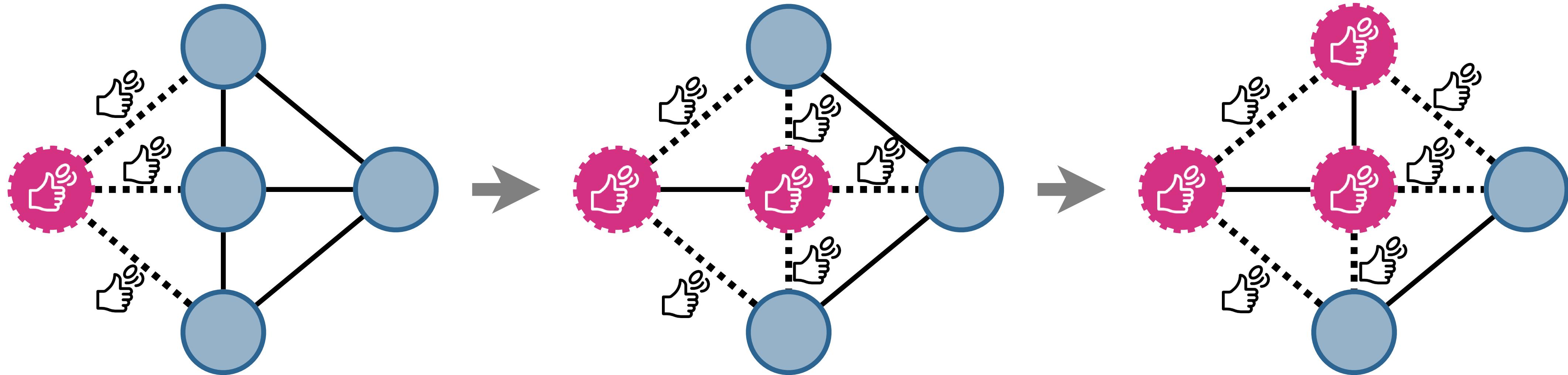
naïve:



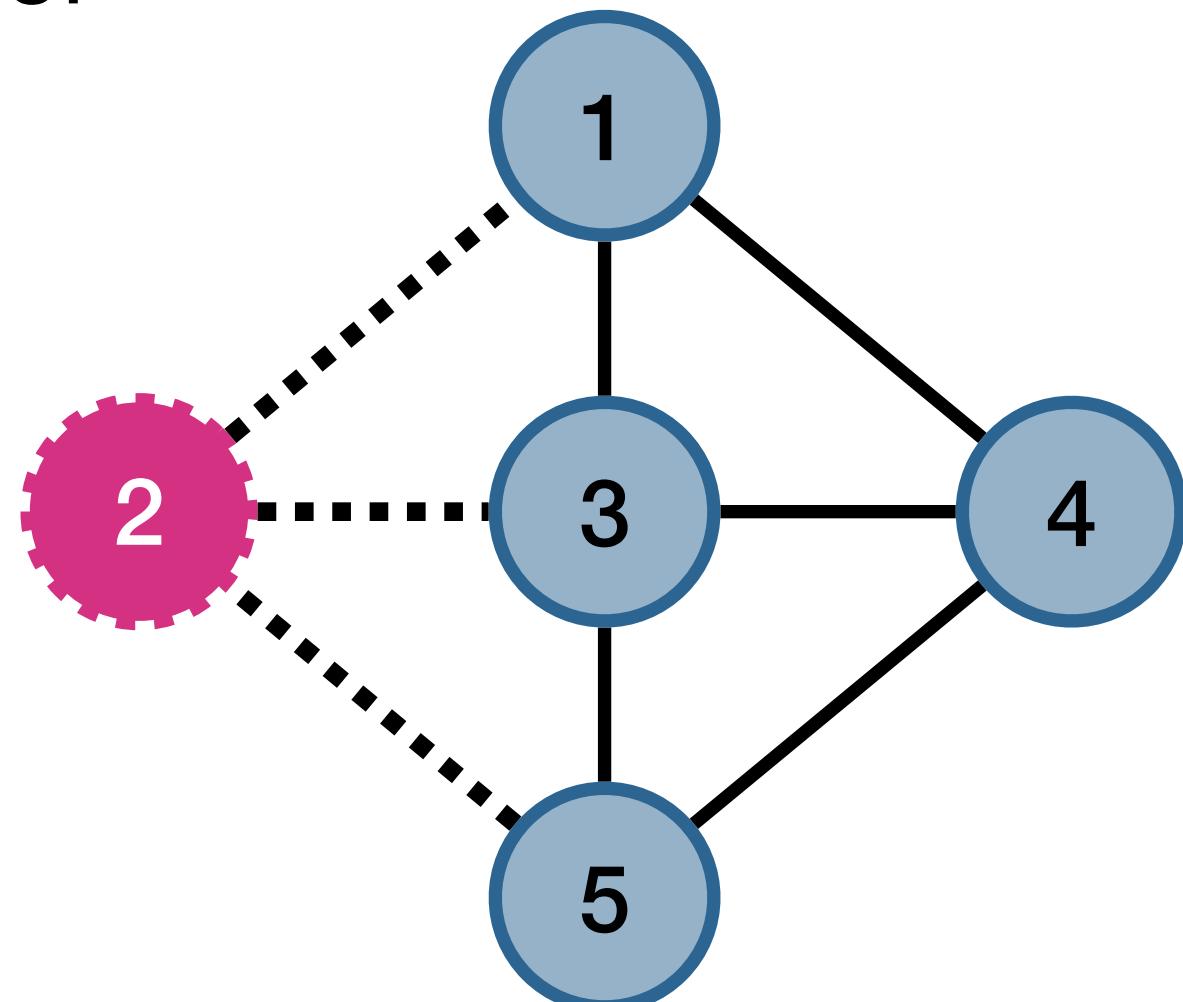
- ▶ **Expensive:** Amount of random numbers
- ▶ **Also Expensive:** Iterate over whole network in each step

SIS Simulation

naïve:



Gillespie:



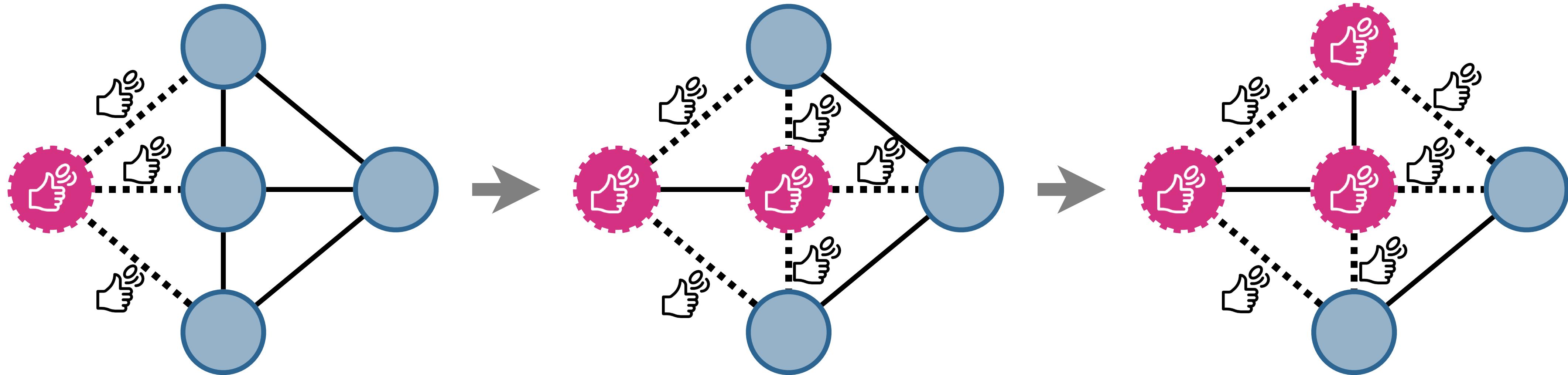
I-nodes: [2]

SI-edges [2-1, 2-3, 2-5]

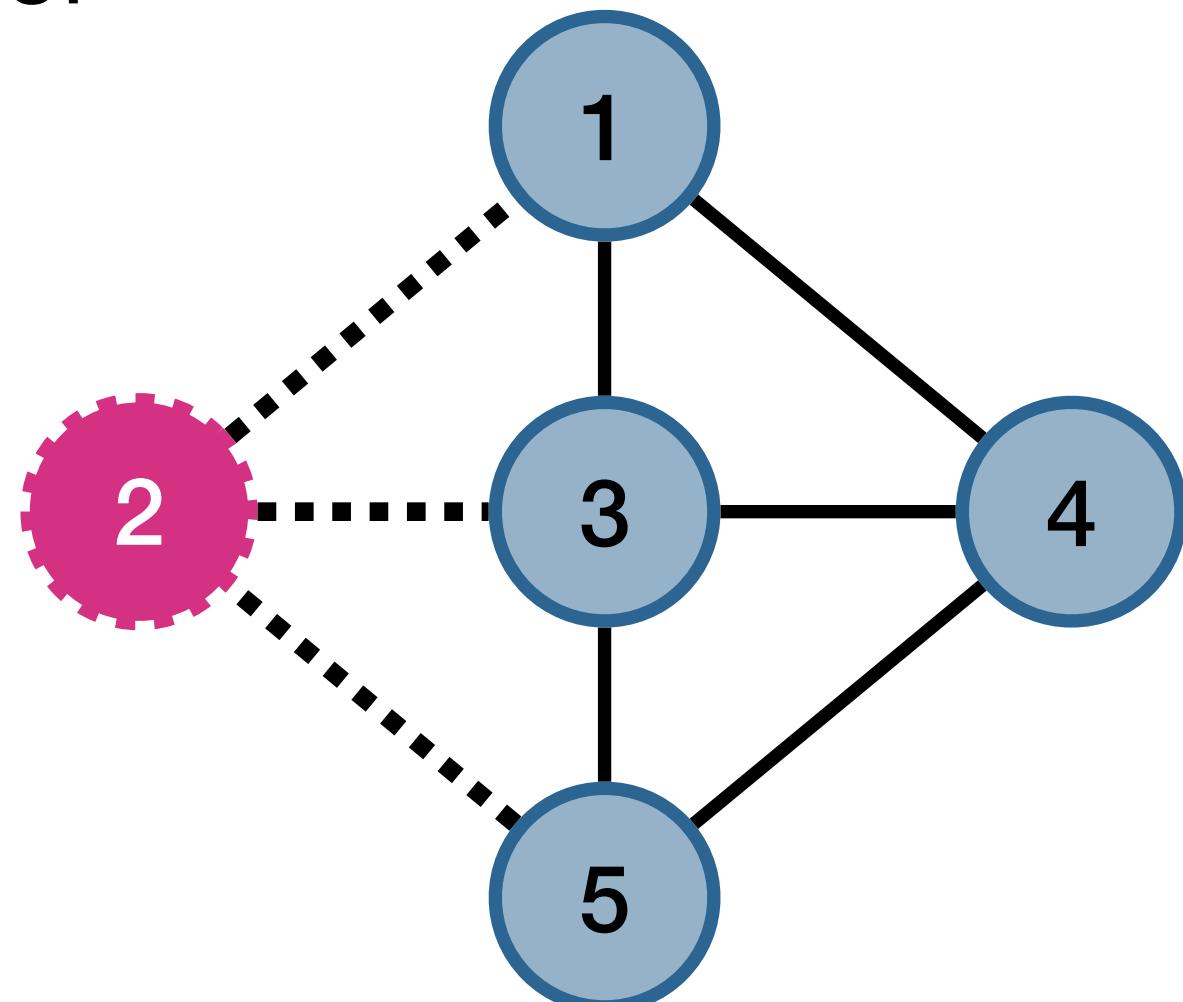


SIS Simulation

naïve:



Gillespie:

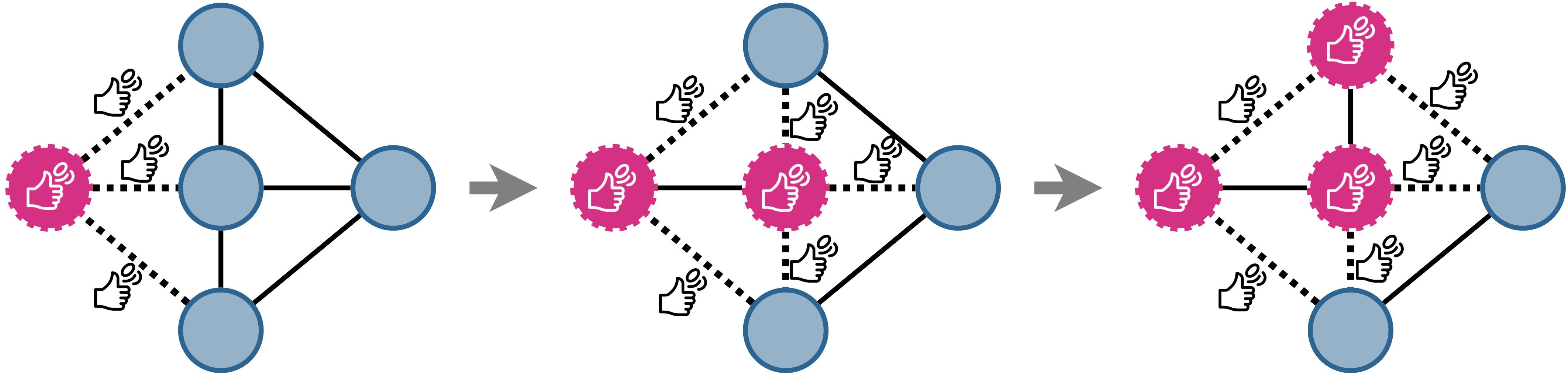


I-nodes: [2]

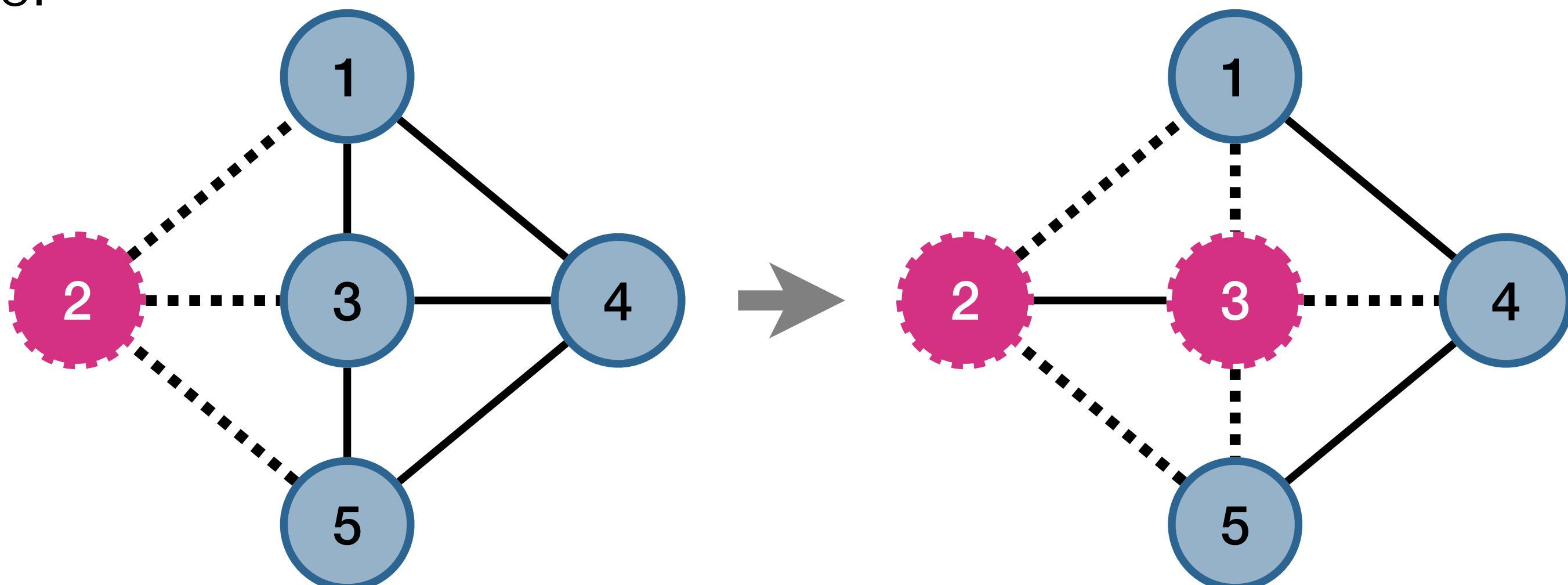
SIR-edges [2-1, 2-3, 2-5]

SIS Simulation

naïve:



Gillespie:



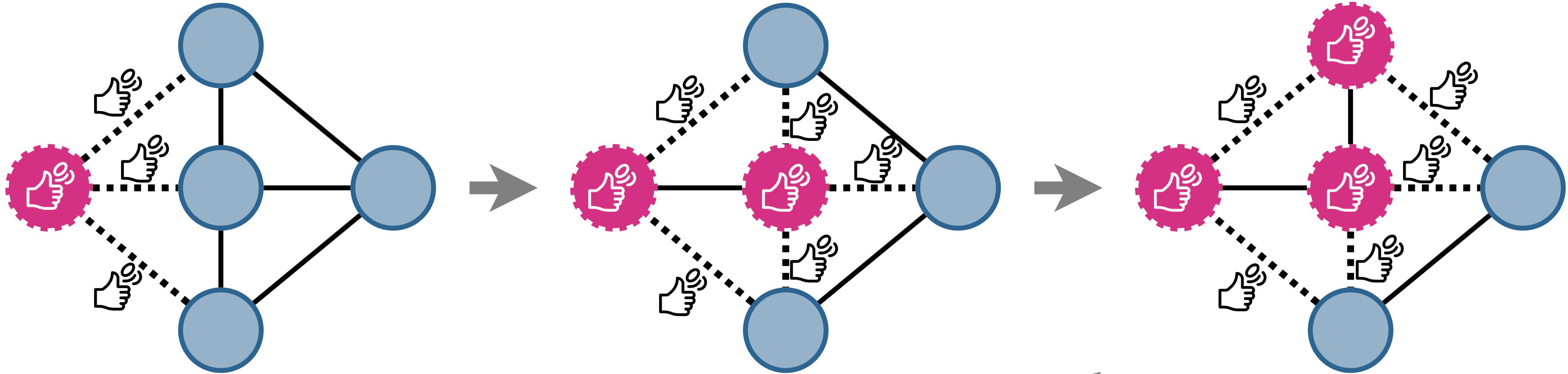
I-nodes: [2, 3]

SI-edges [2-1, 2-5, 3-1, 3-4, 3-5]

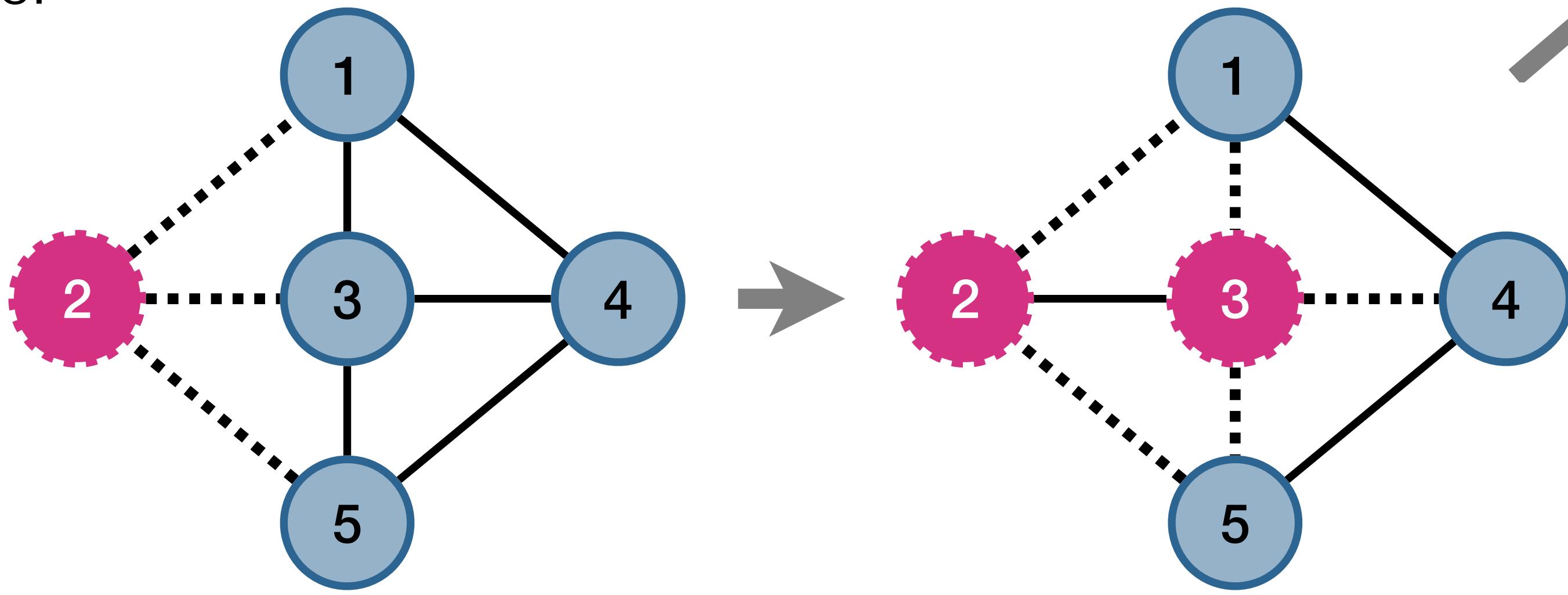


SIS Simulation

naïve:



Gillespie:

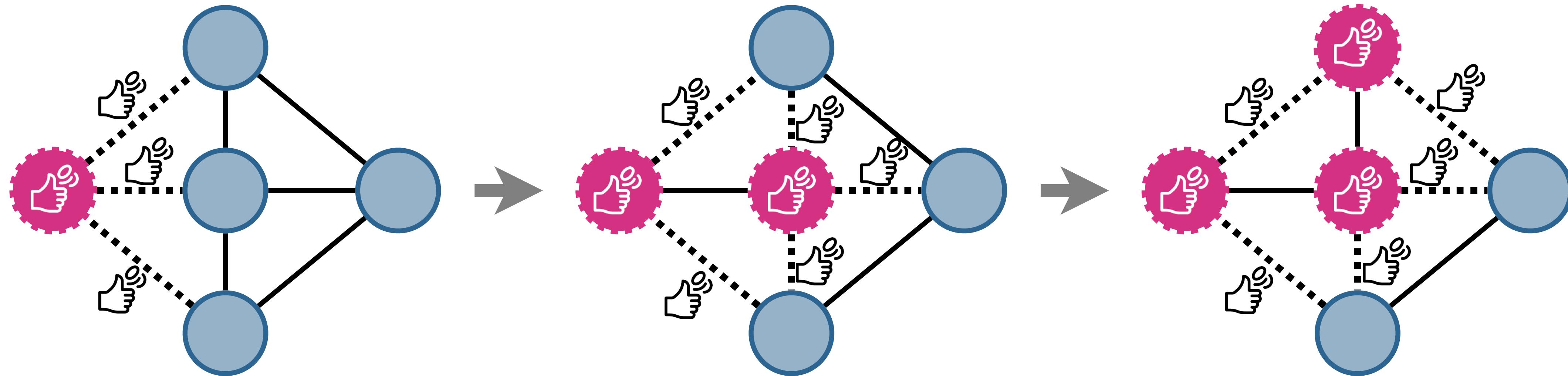


I-nodes: [2, 3]

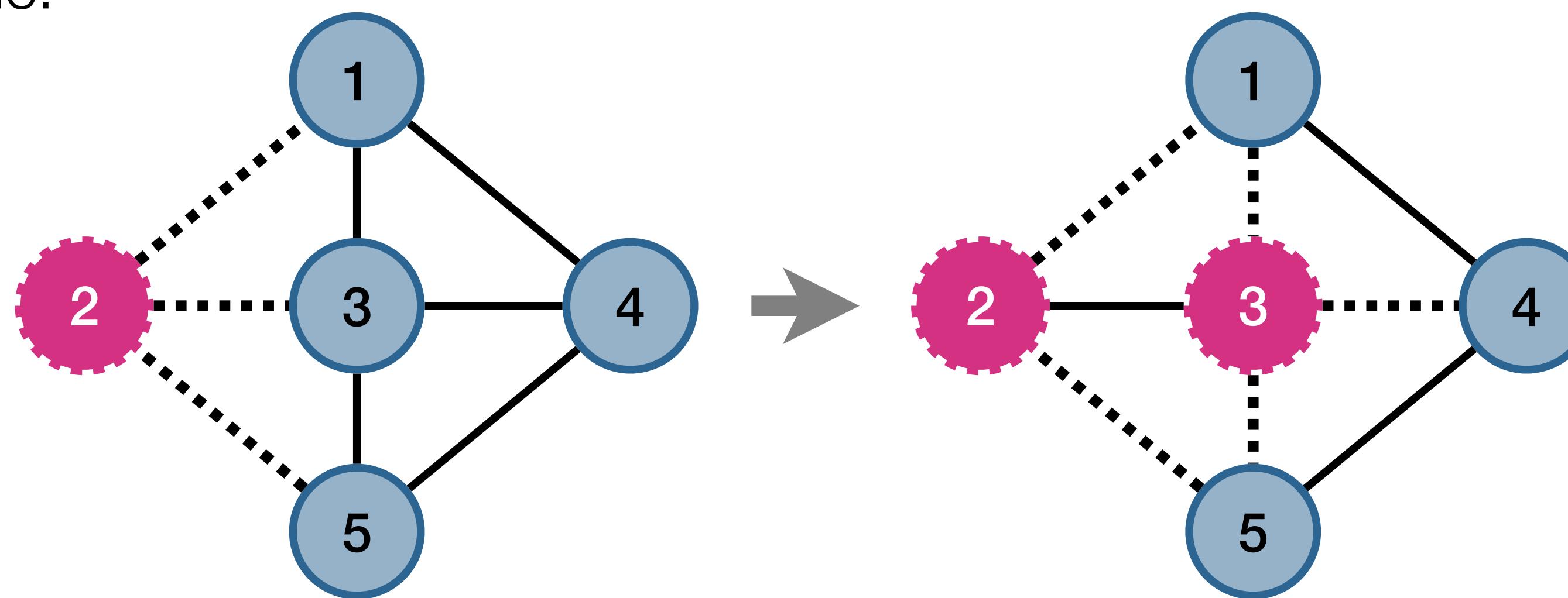
SI-edges [2-1, 2-5, 3-1, 3-4, 3-5]

SIS Simulation

naïve:



Gillespie:



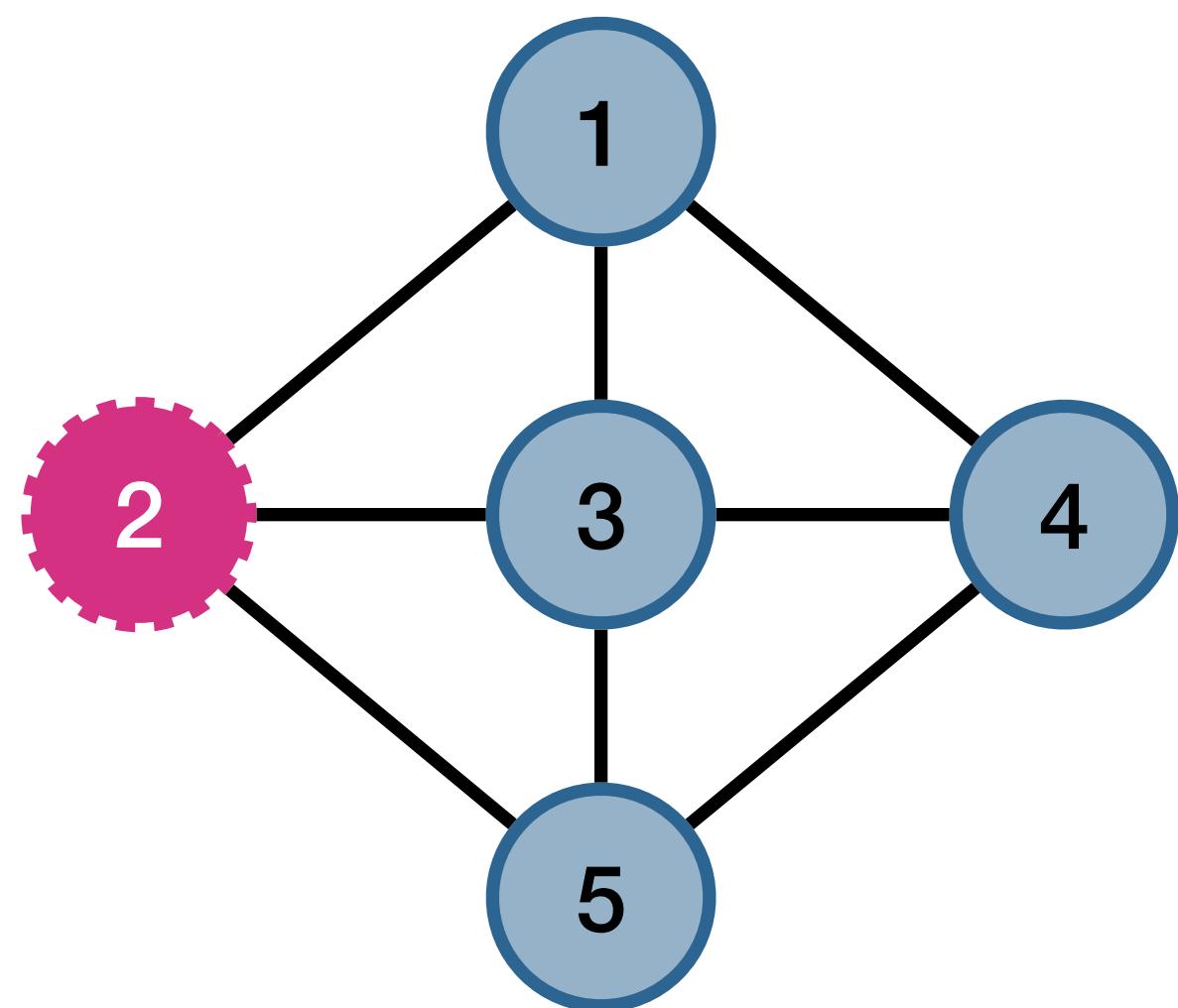
I-nodes: [2, 3]

SI-edges [2-1, 2-5, 3-1, 3-4, 3-5]

- ▶ **Cheap:** Sampling
- ▶ **Expensive:** Update lists

SIS Simulation

rejection-based:



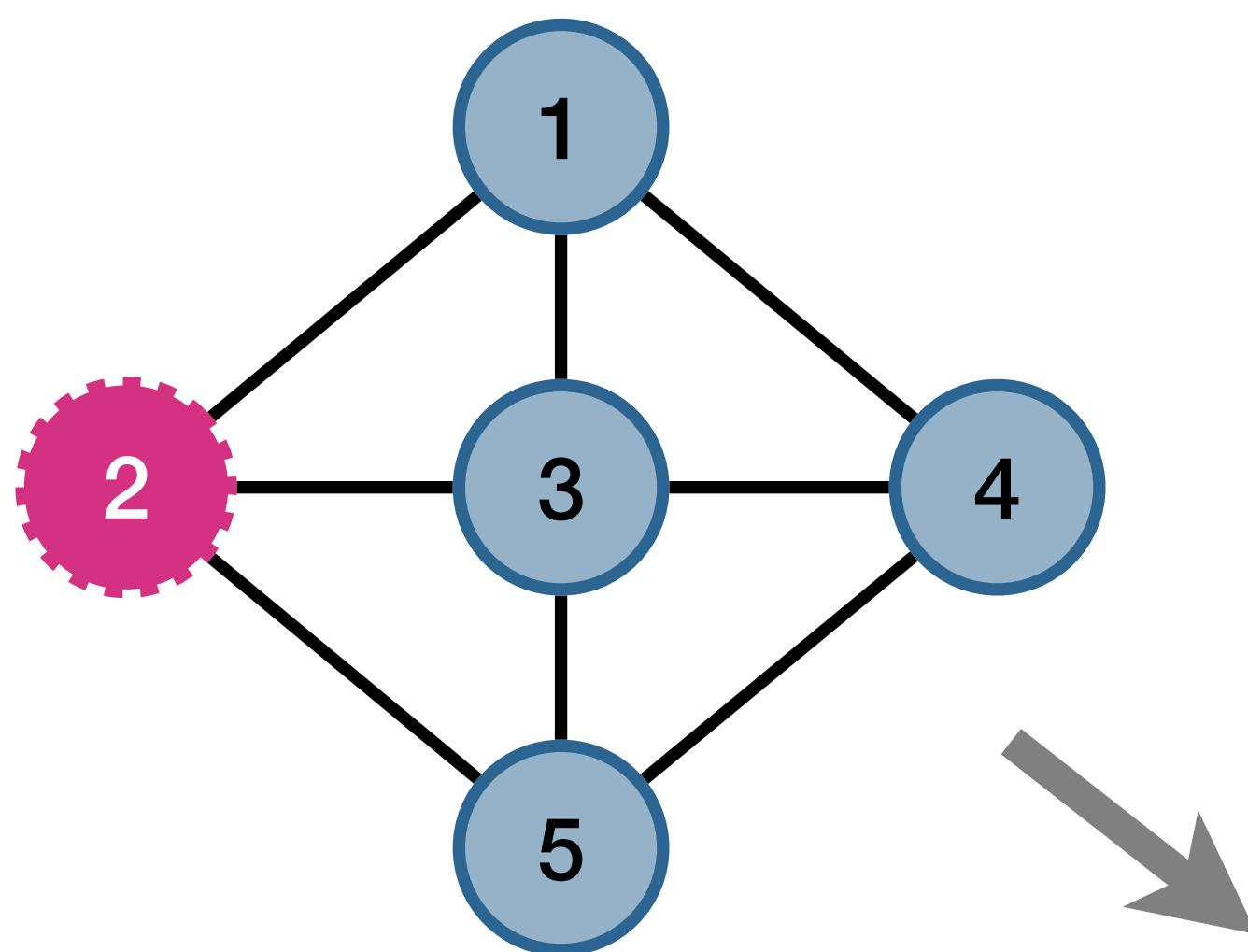
I-nodes: [2]



Sample infection source
assuming **4 susceptible**
neighbors

SIS Simulation

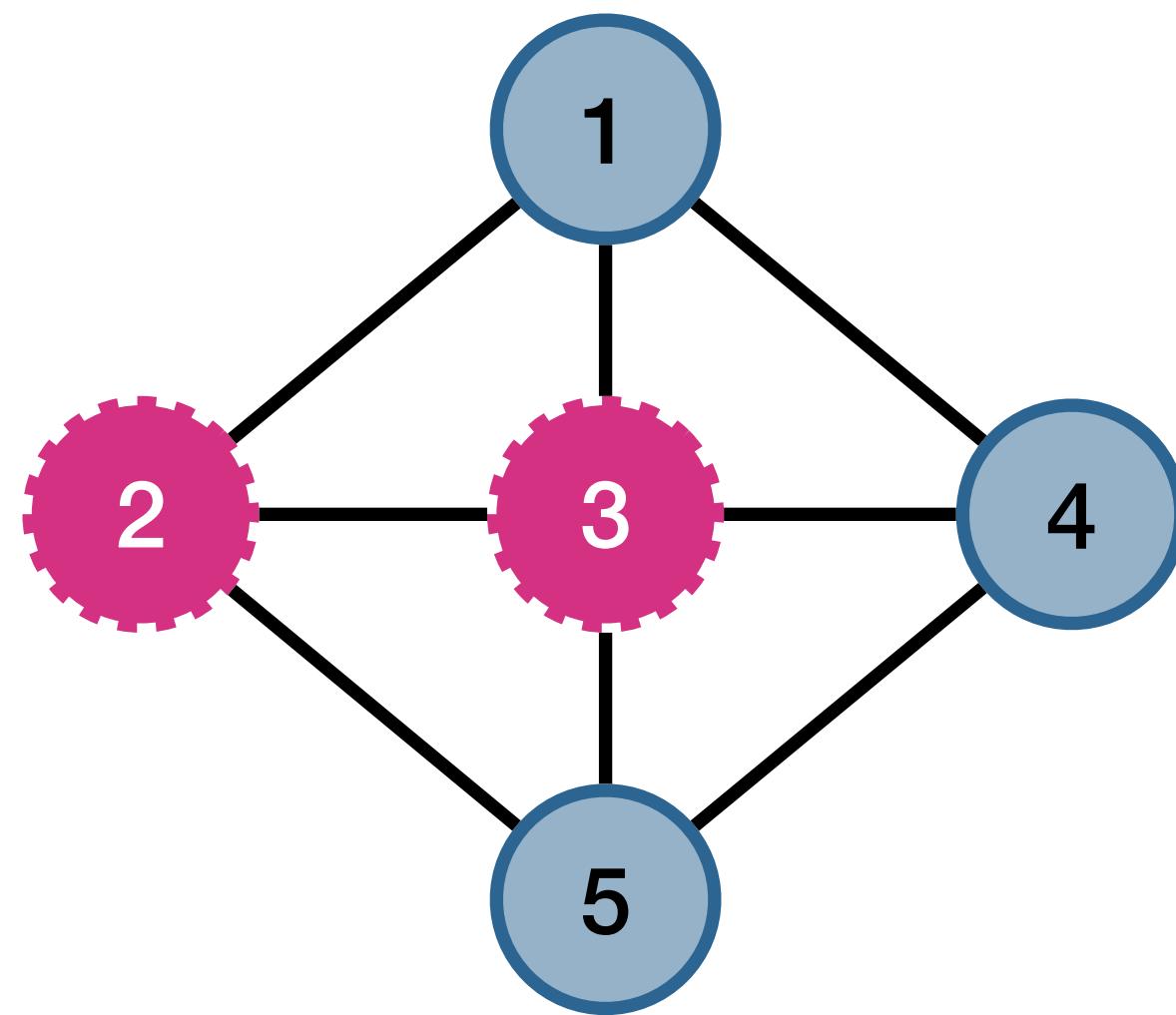
rejection-based:



I-nodes: [2]



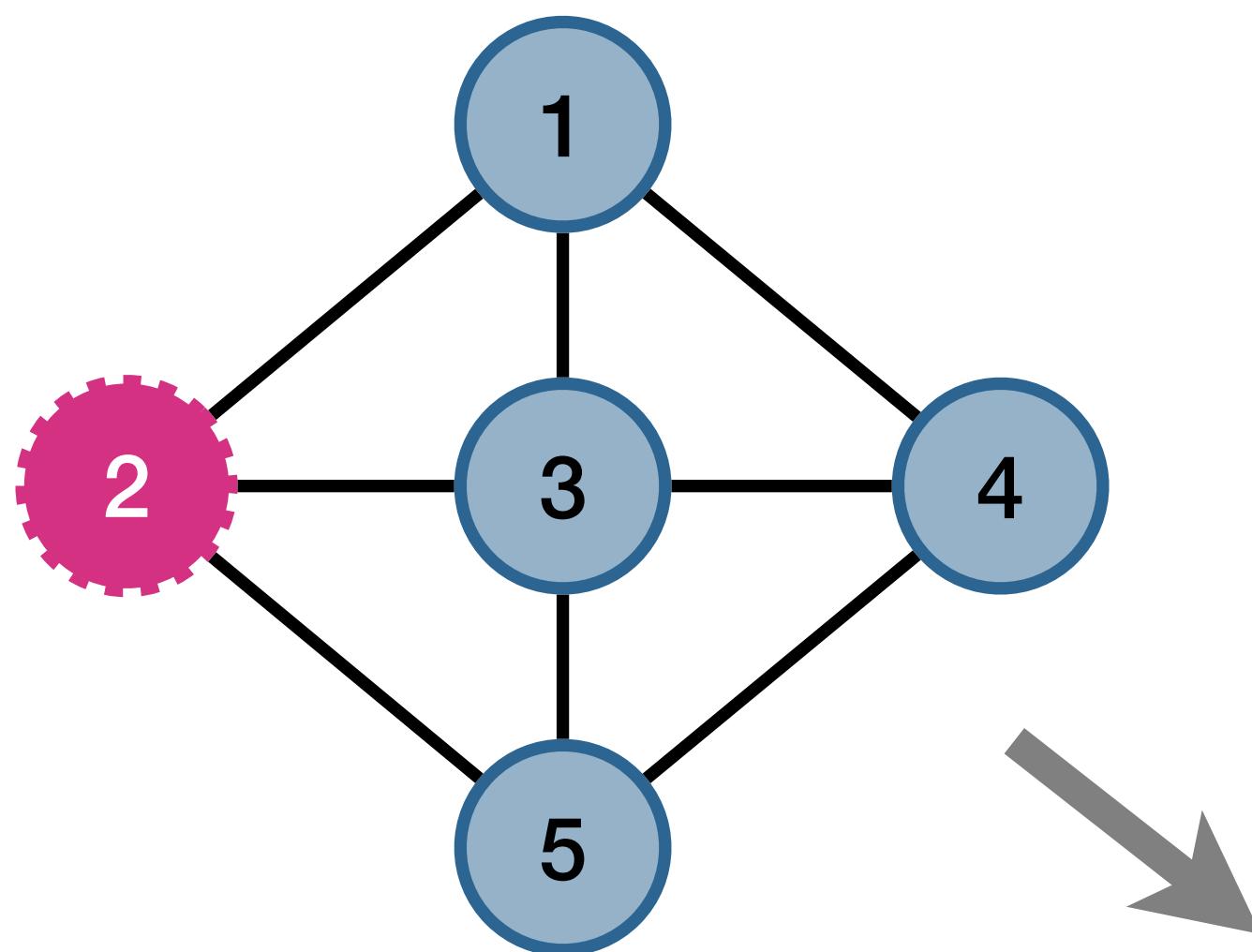
Sample infection source
assuming **4 susceptible**
neighbors



rejection probability
= 25%

SIS Simulation

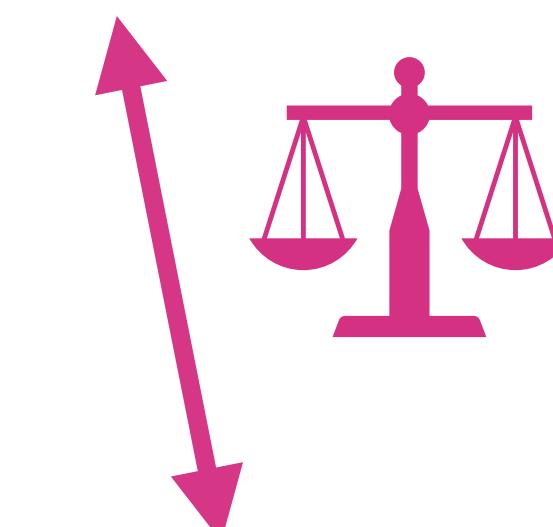
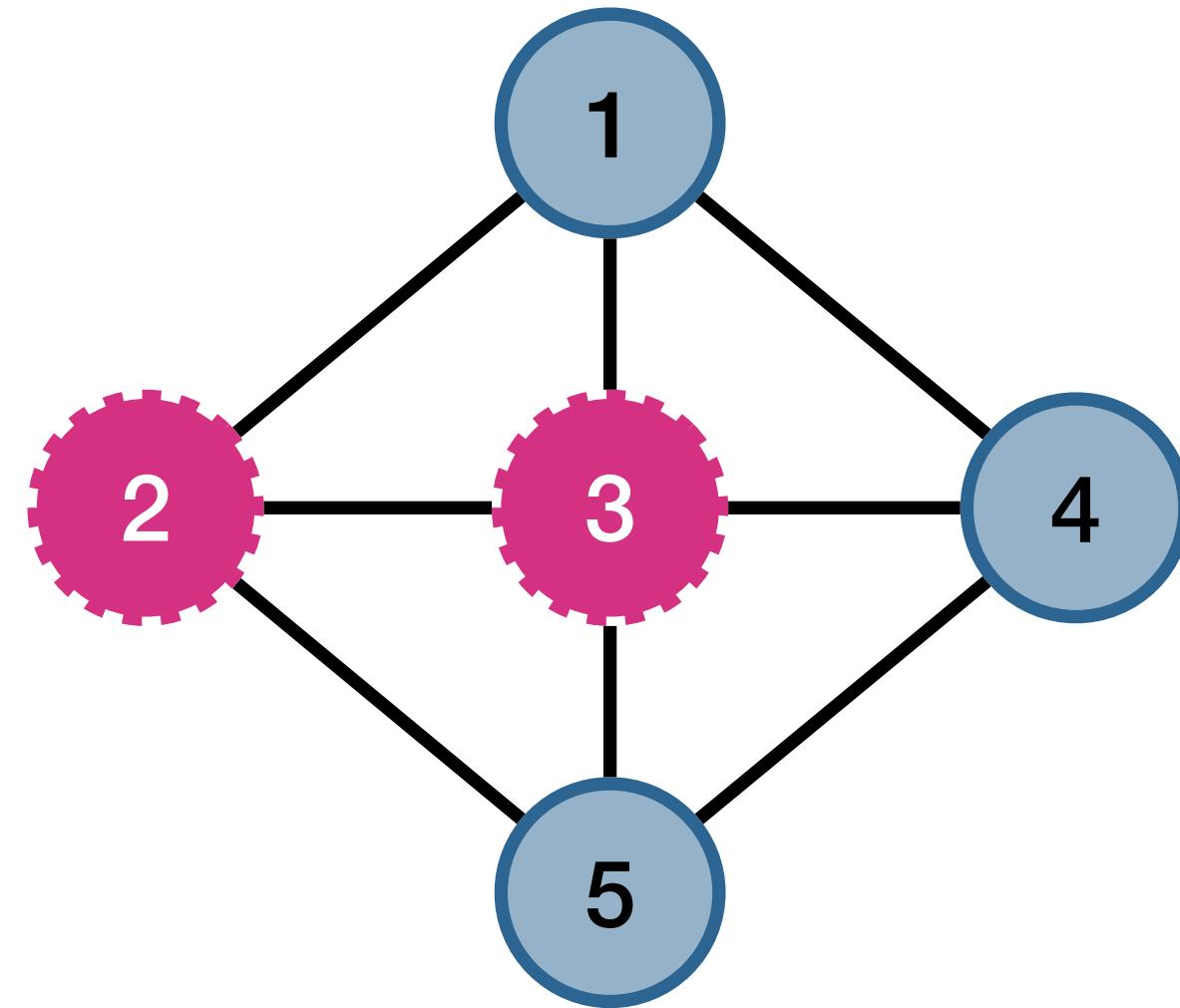
rejection-based:



I-nodes: [2, 3]



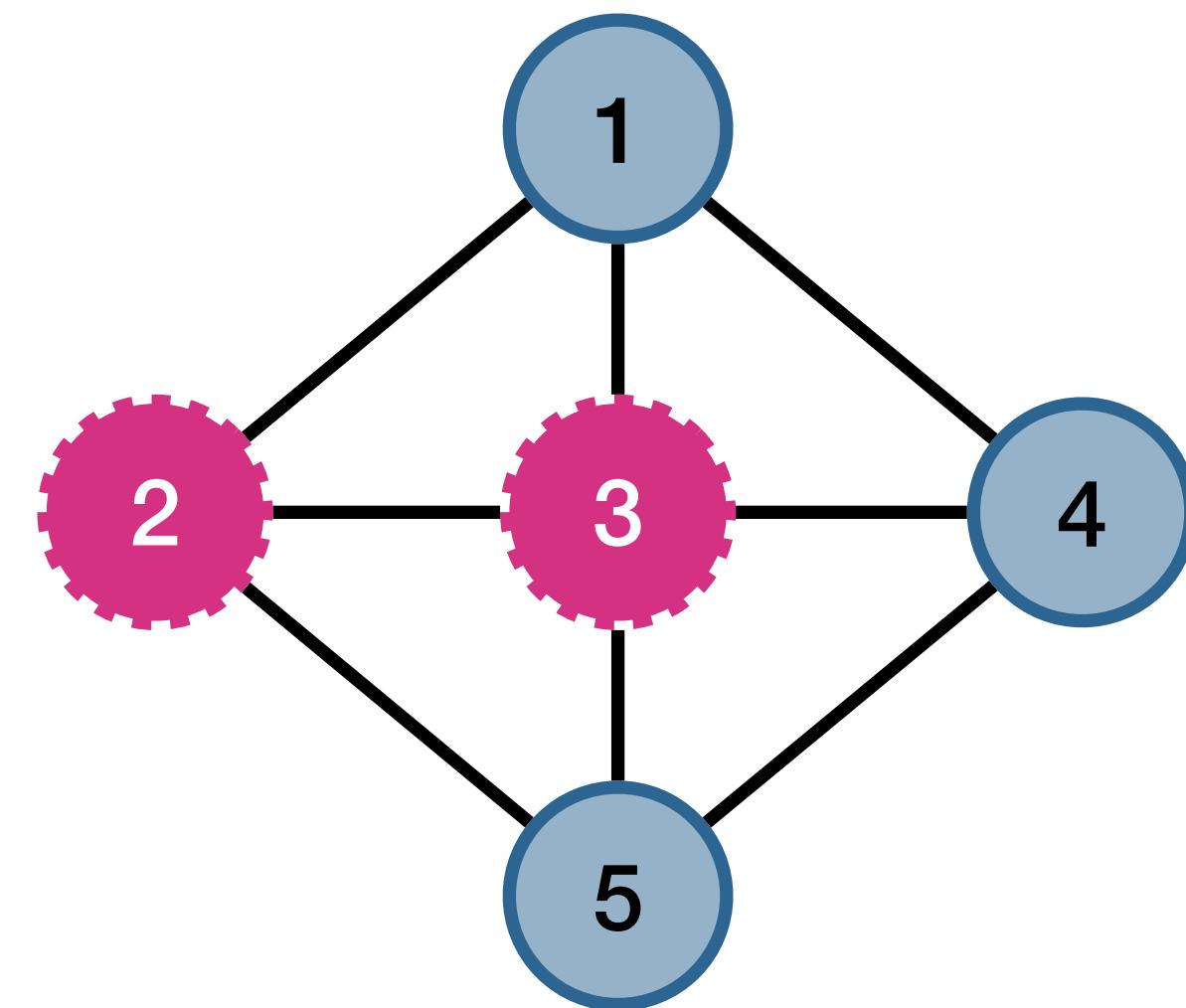
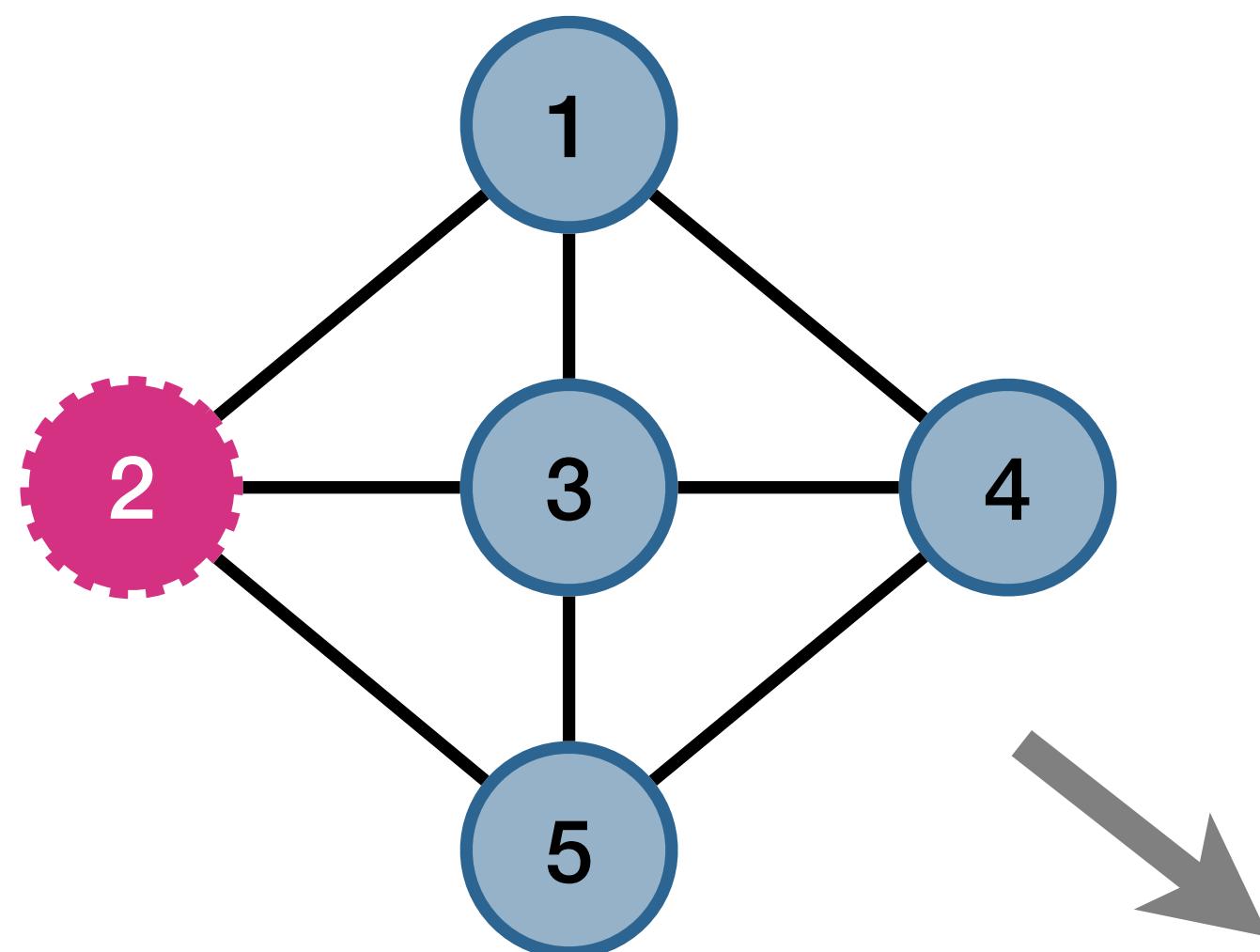
Sample infection source
assuming **4 susceptible**
neighbors



rejection probability
= 25%

SIS Simulation

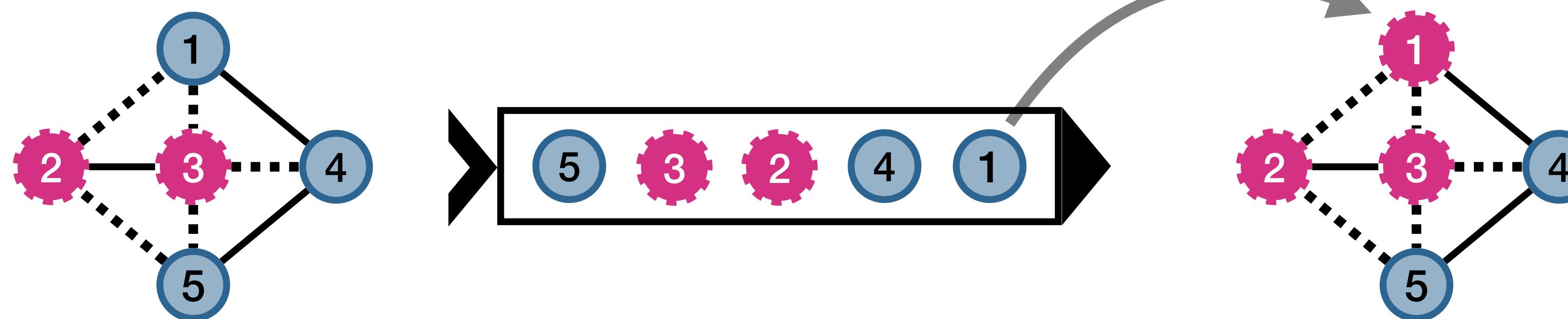
rejection-based:



- ▶ **Cheap:** Sampling
- ▶ **Also cheap:** Update lists
- ▶ **Expensive:** Amount of rejections

Event-Driven Simulation

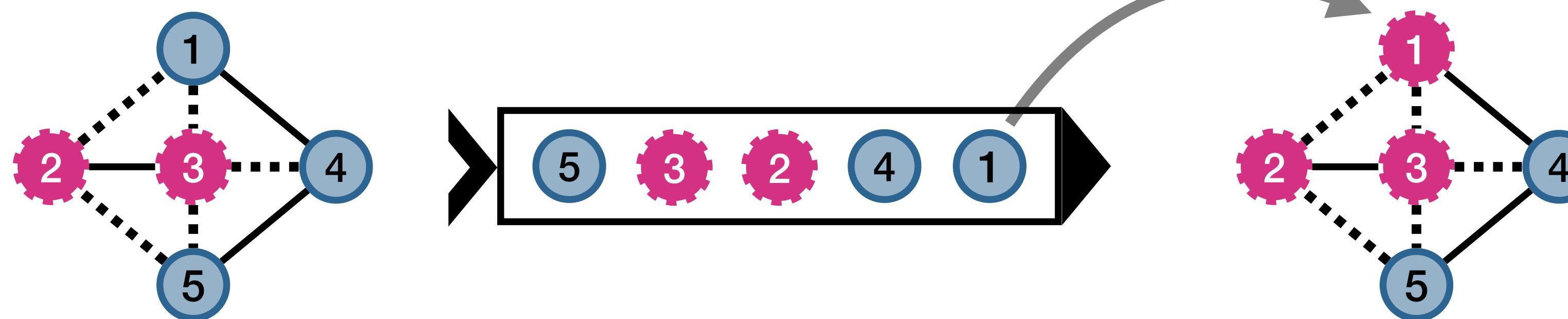
event-driven:



1. Pop event.
2. Update node.
3. Create new event.
4. Update neighborhood.

Event-Driven Simulation

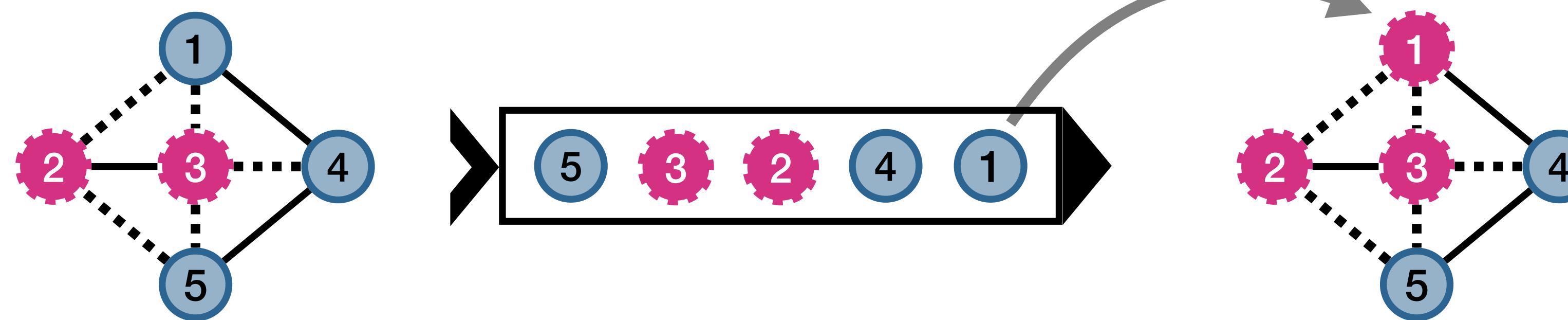
event-driven:



1. Pop event.
2. Update node.
3. Create new event.
4. Update neighborhood **(expensive)**.

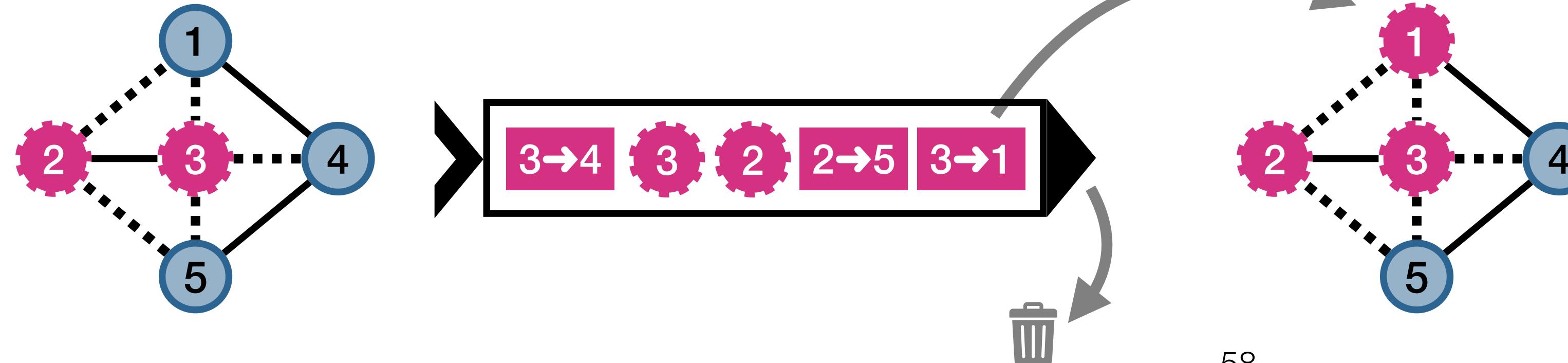
Our Method

event-driven:



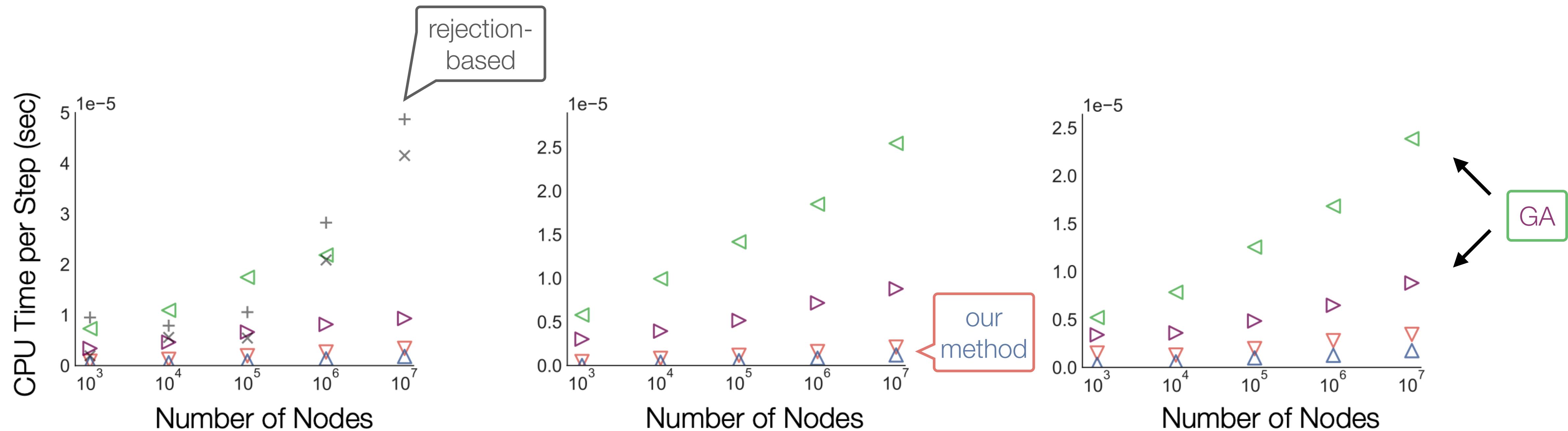
1. Pop event.
2. Update node.
3. Create new event.
4. Update neighborhood **(expensive)**.

rejection-based + event-driven (our method):



1. Pop event **(possible rejection)**.
2. Update node.
3. Create new event **(possible rejection)**.
4. **Update neighborhood**.

SIS Simulation: Results

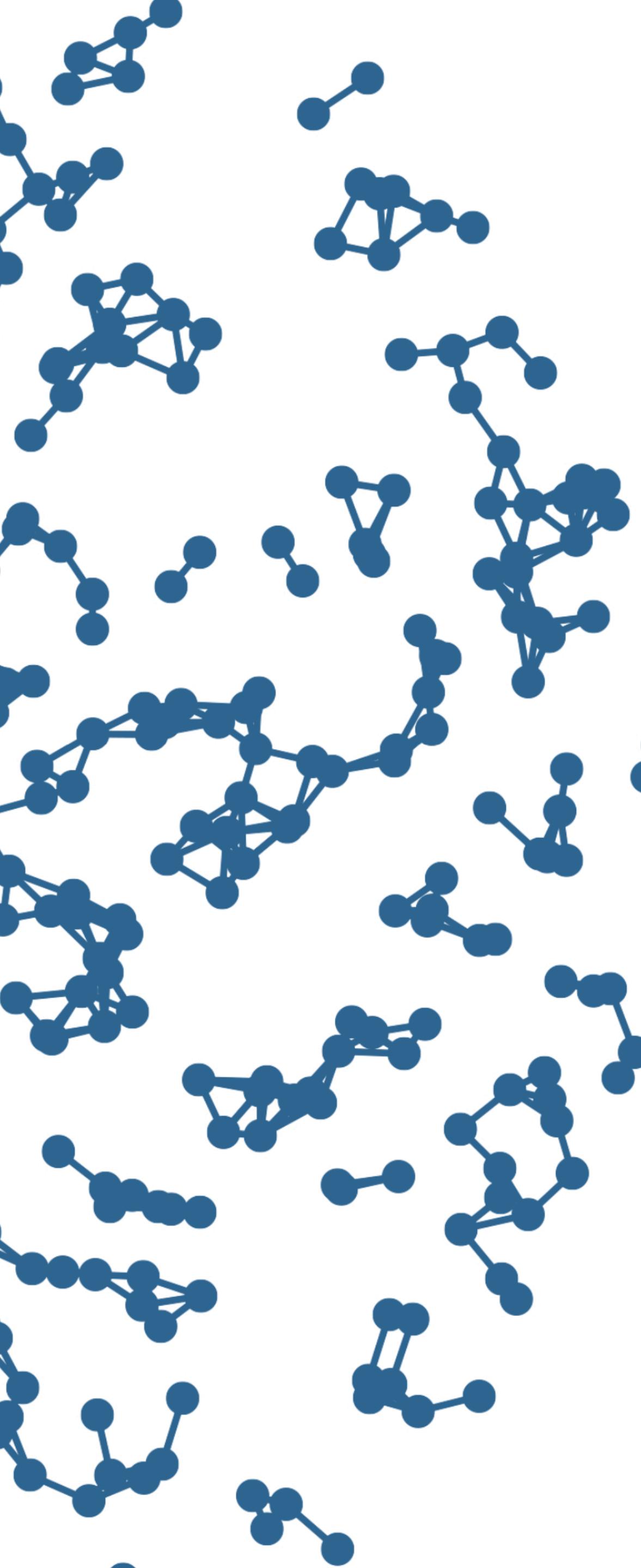


SIS

SIRS

SIJS

(competing pathogens)



Agenda

METHODS

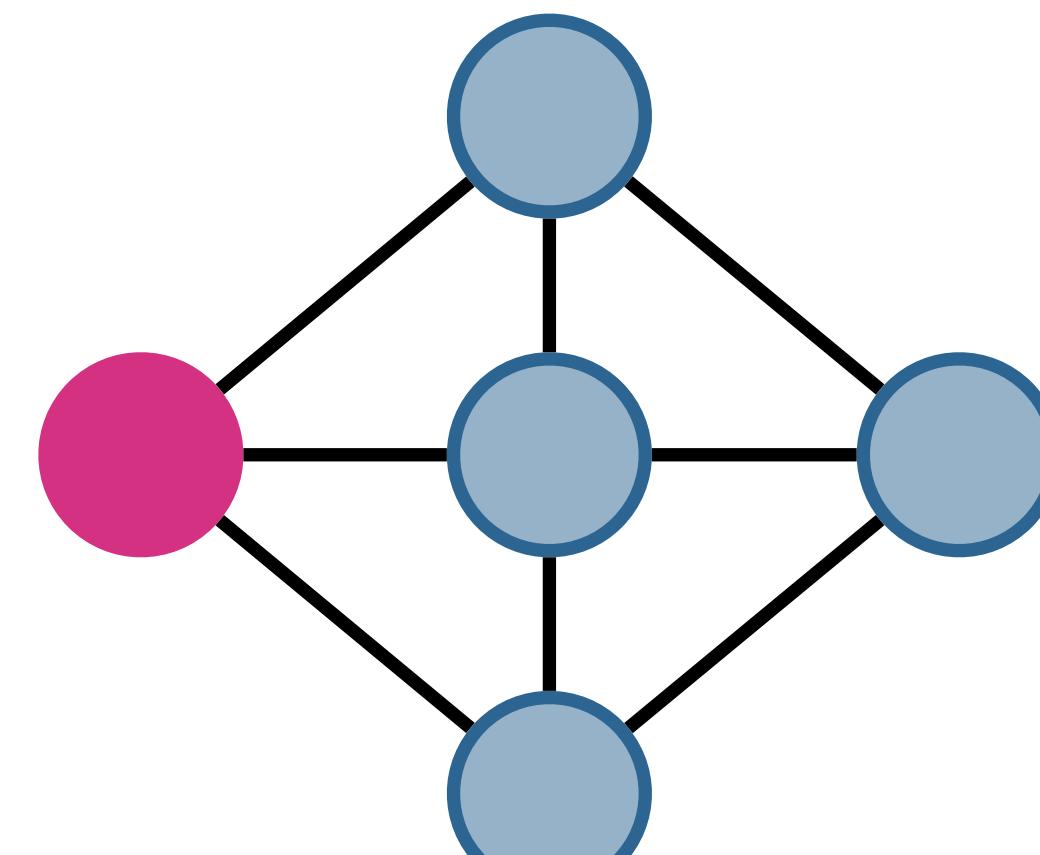
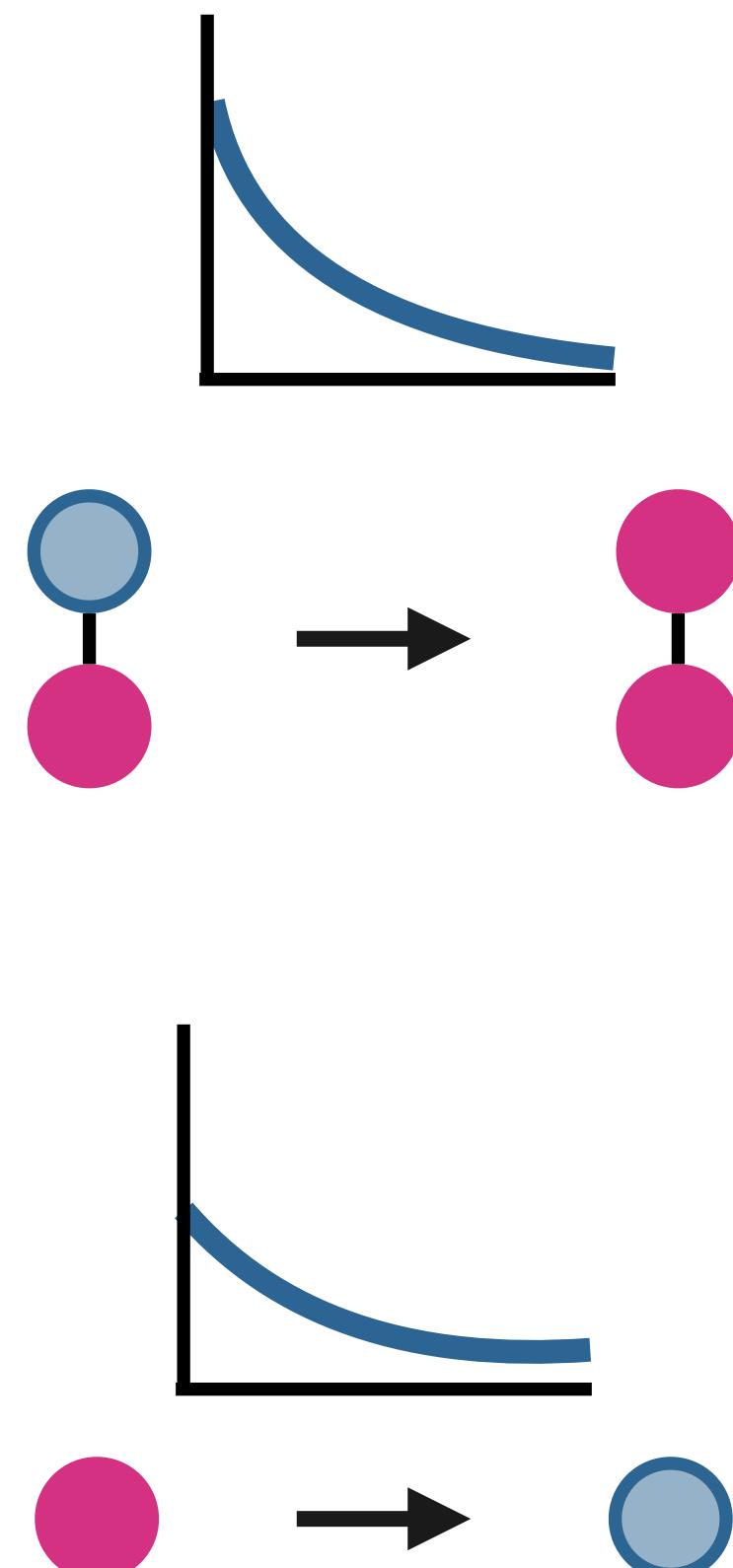
COVID

(NON-)MARKOVIAN SIMULATION

CONTROL
INFERENCE

Non-Markovian Models

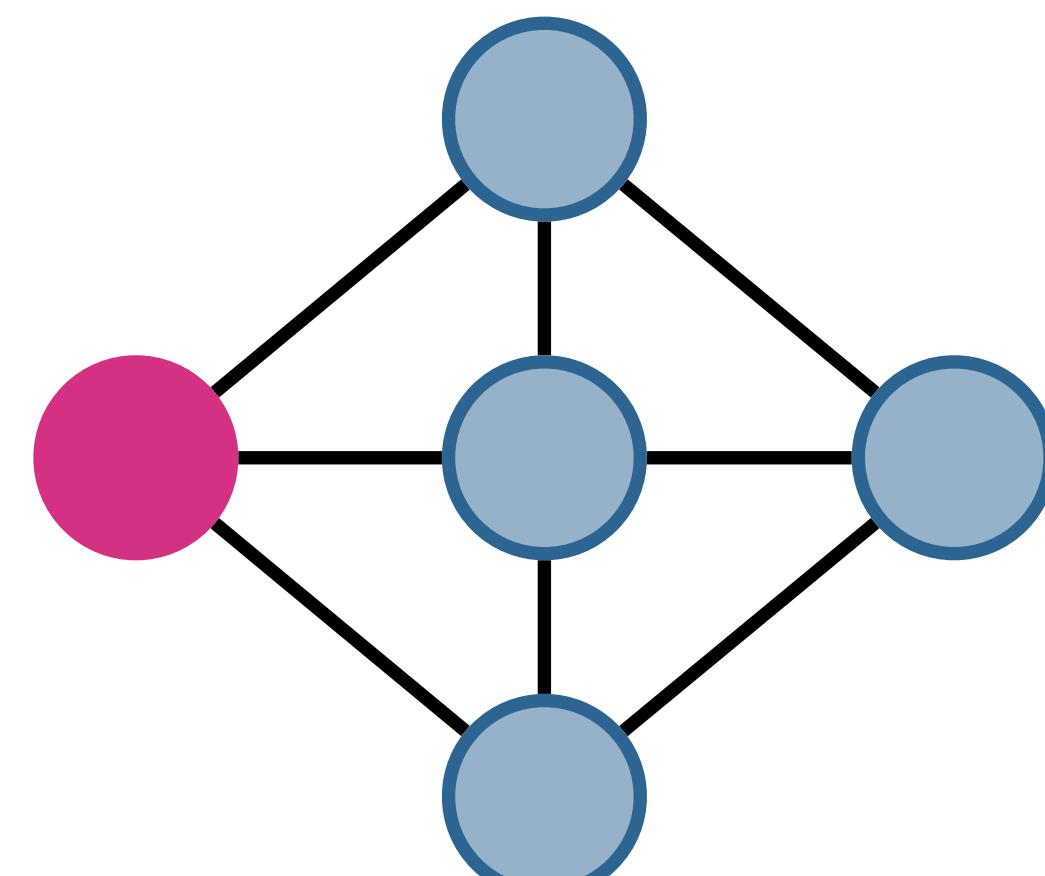
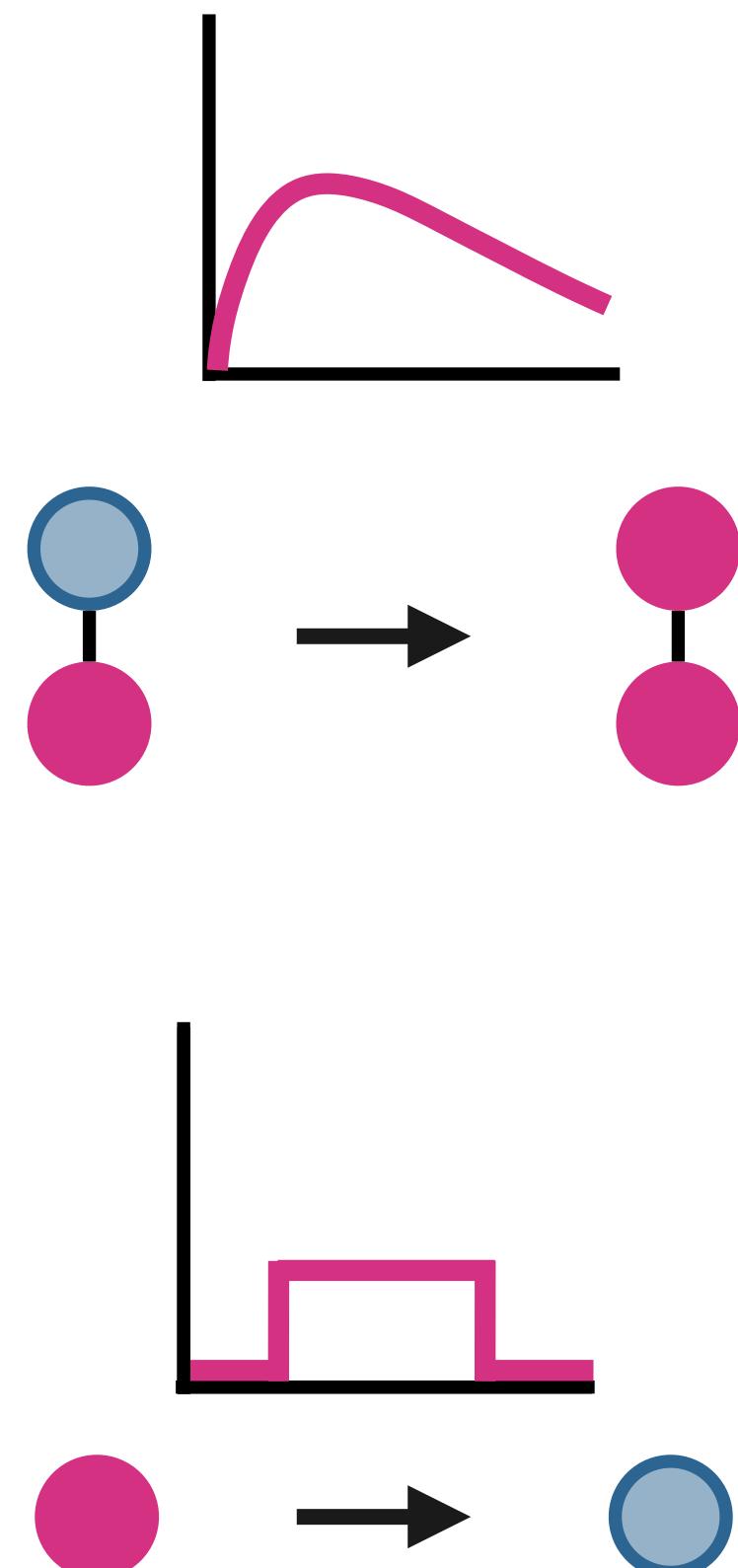
Markovian:



- ▶ **Trade-off:** simplicity vs accuracy
- ▶ Vastly different properties
- ▶ Difficult to analyse

Non-Markovian Models

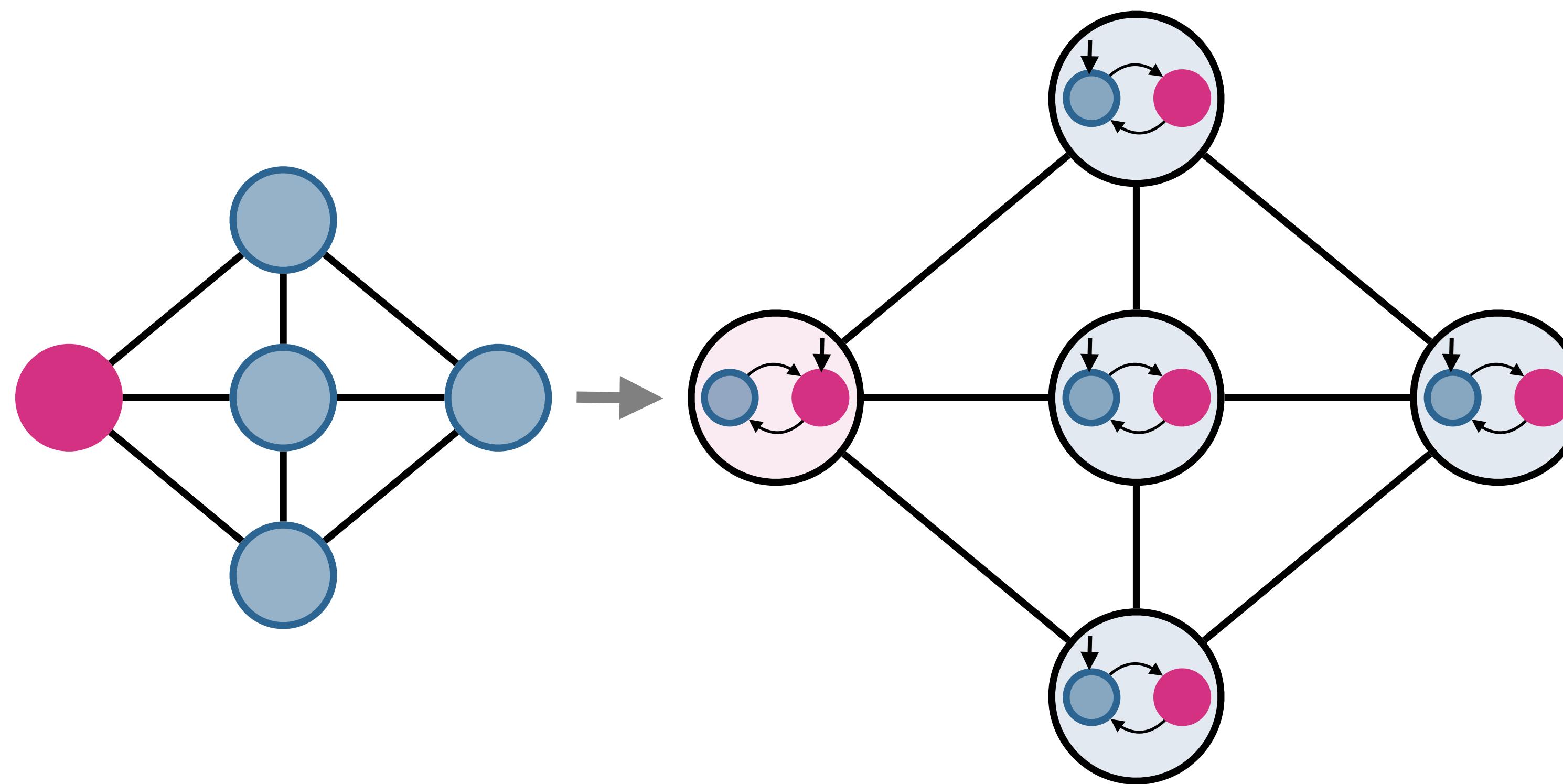
Non-Markovian:



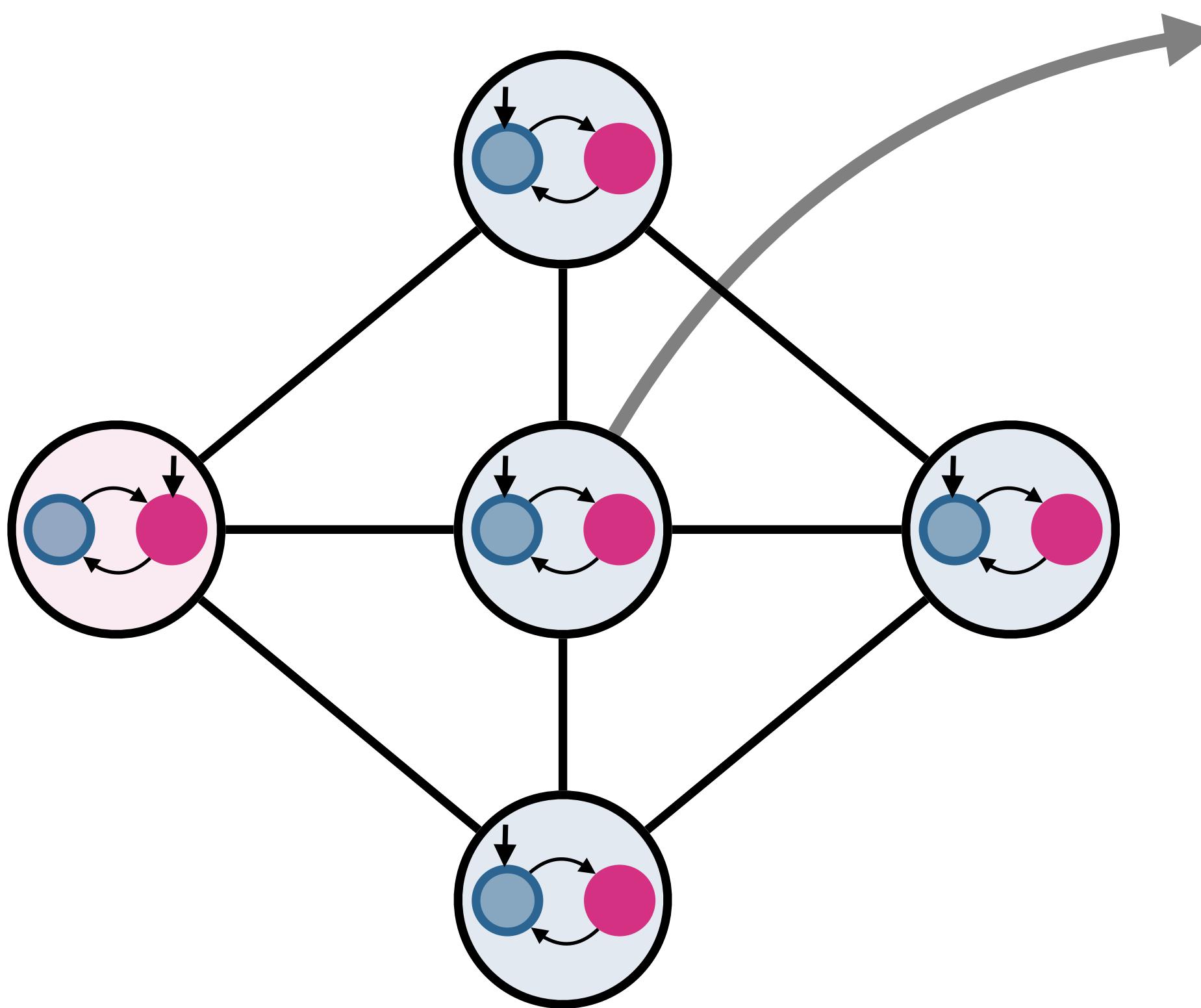
- ▶ **Trade-off:** simplicity vs accuracy
- ▶ Vastly different properties
- ▶ Difficult to analyse

Non-Markovian Models

Goal: Universal framework for non-Markovian dynamical systems

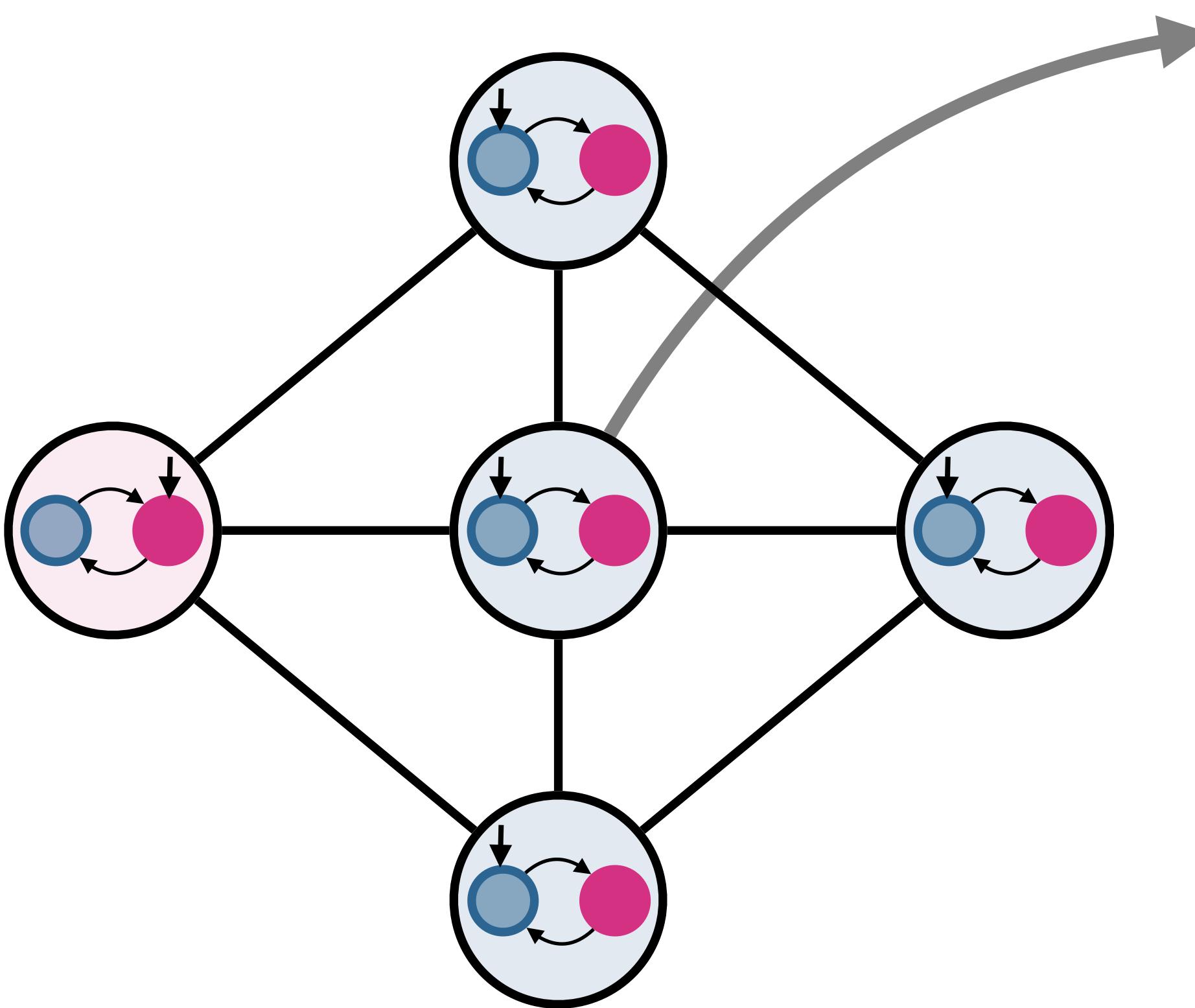


Non-Markovian Models



- ▶ Agent-specific rate function
- ▶ Changes continuously in time.
- ▶ Depends on Markov blanket.

Non-Markovian Models

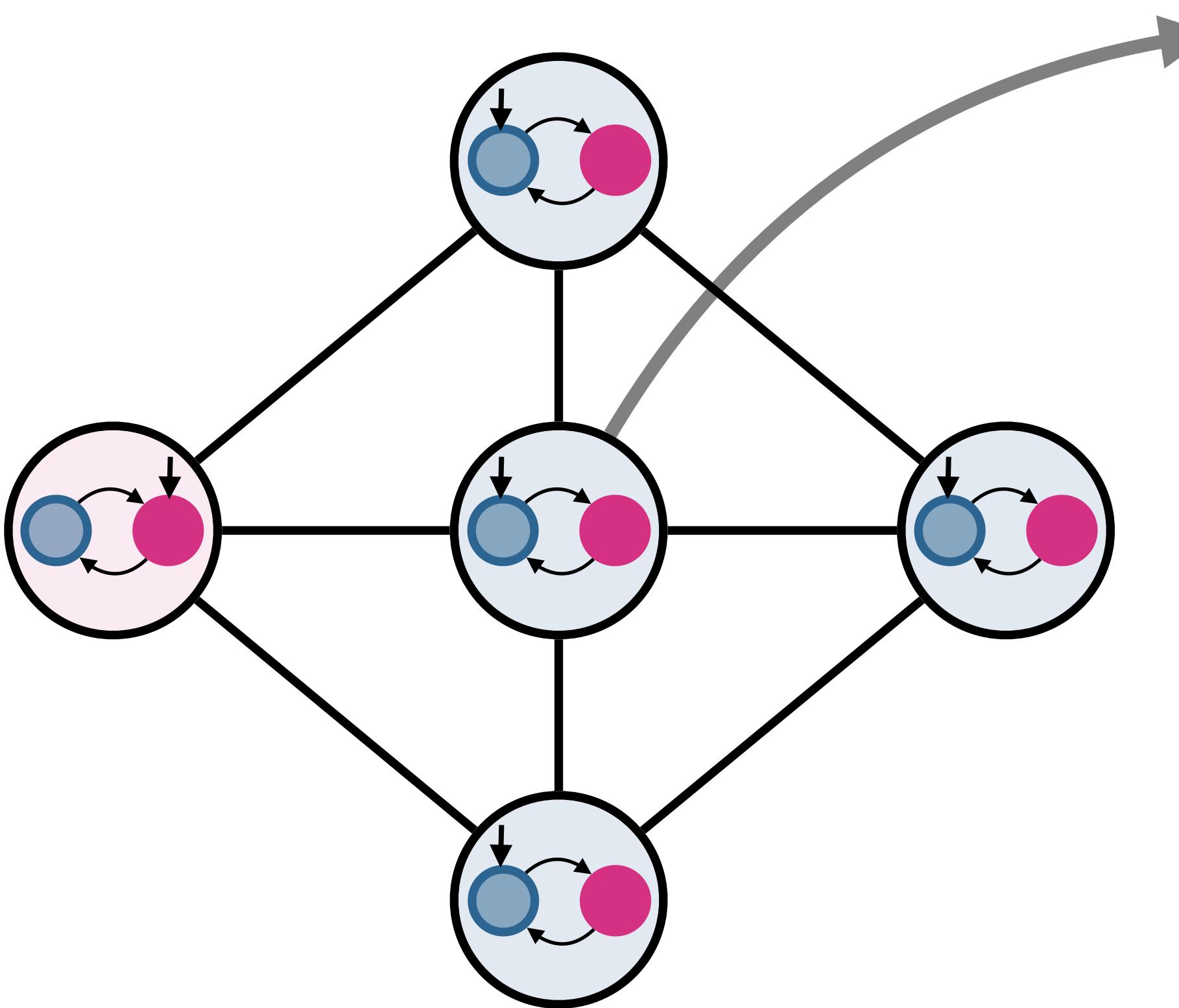


Input:

- ▶ Current state:
- ▶ Residence time: 0.42
- ▶ State and residence time of all neighbors:

$\{(\text{pink}, 0.1), (\text{blue}, 2.7), (\text{blue}, 1.5), (\text{blue}, 0.2)\}$

Non-Markovian Models



Input:

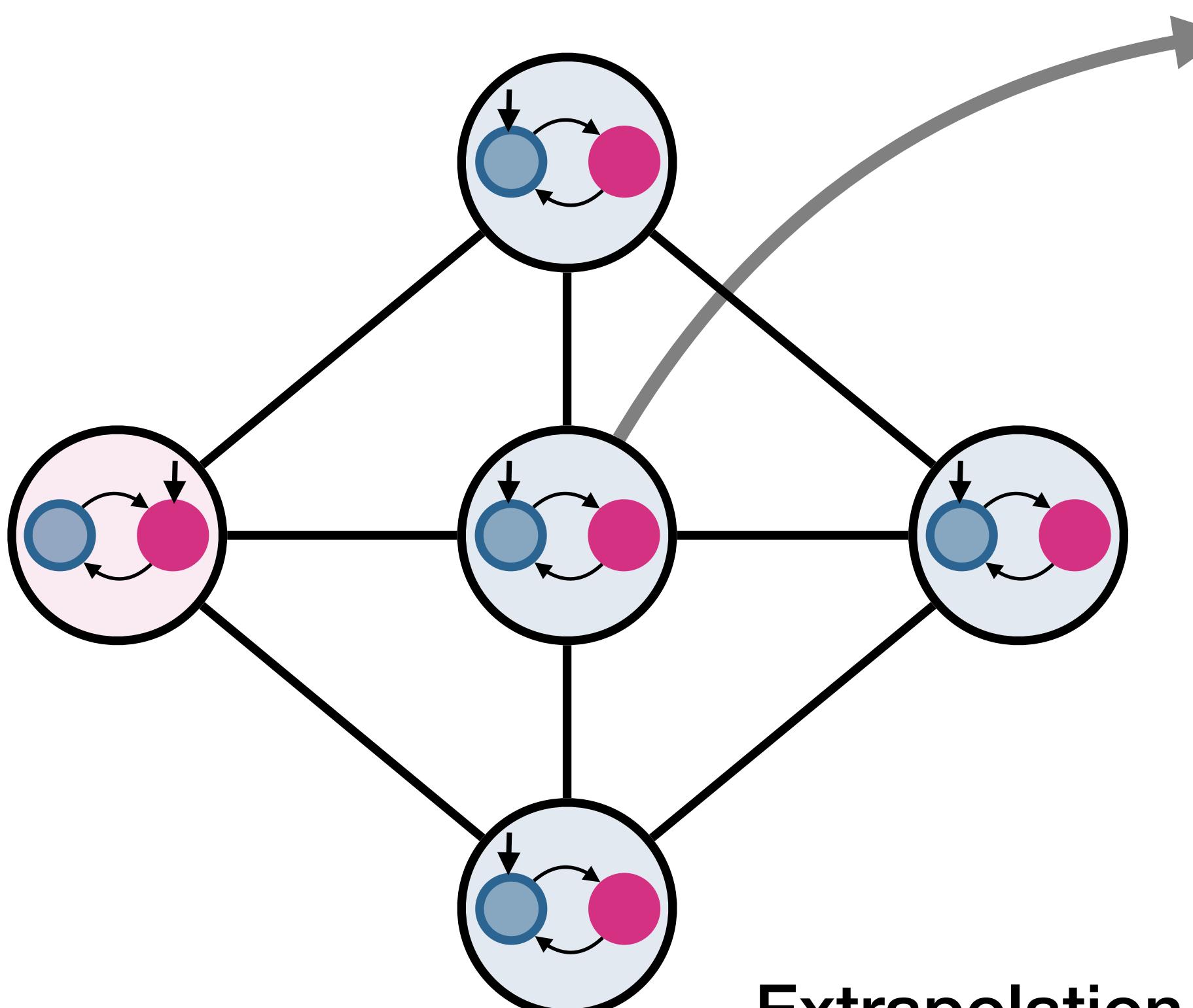
- ▶ Current state:
- ▶ Residence time: 0.42
- ▶ State and residence time of all neighbors:

$\{((\text{pink}), 0.1), ((\text{blue}), 2.7), ((\text{blue}), 1.5), ((\text{blue}), 0.2)\}$

Output:

- ▶ Next state
- ▶ Instantaneous rate: 1.5

Non-Markovian Models

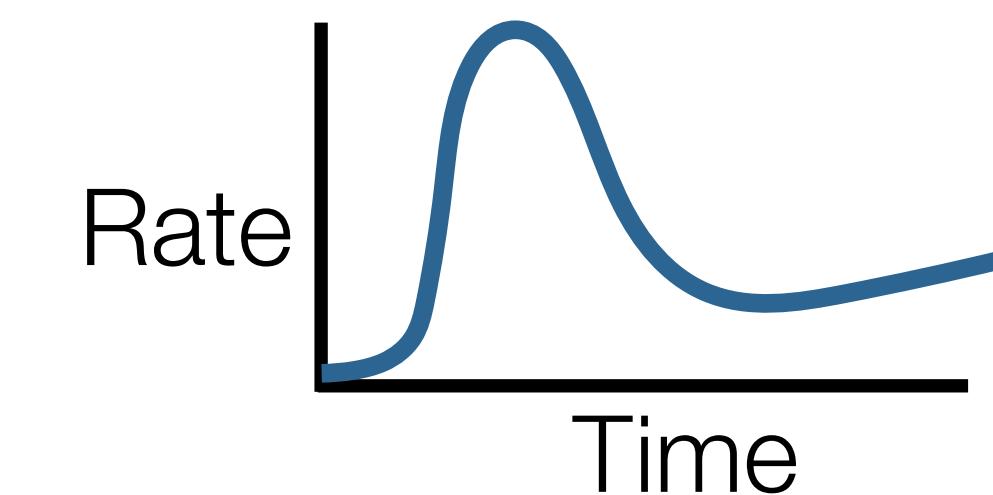


Extrapolation assuming fixed neighborhood:

Input:

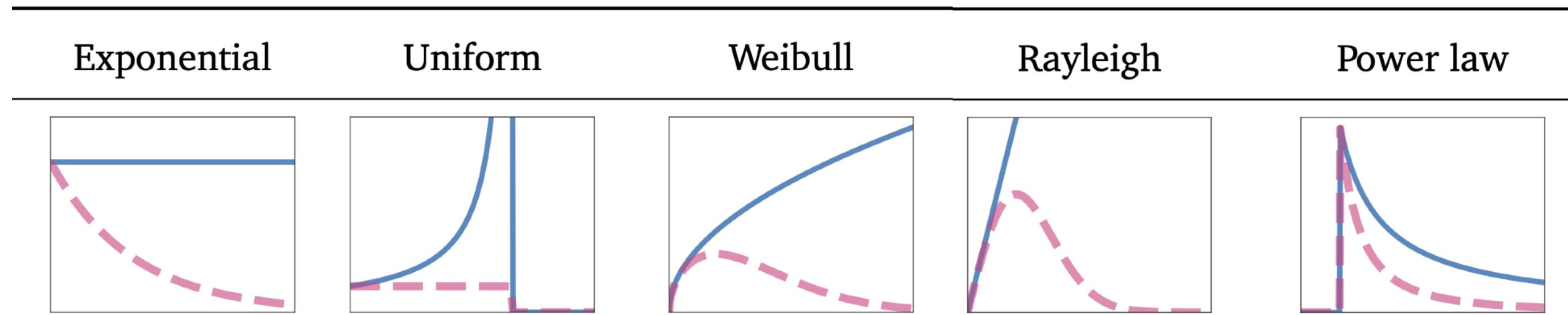
- ▶ Current state:
- ▶ Residence time: 0.42
- ▶ State and residence time of all neighbors:

$\{(\text{pink}, 0.1), (\text{blue}, 2.7), (\text{blue}, 1.5), (\text{blue}, 0.2)\}$



Non-Markovian Simulation

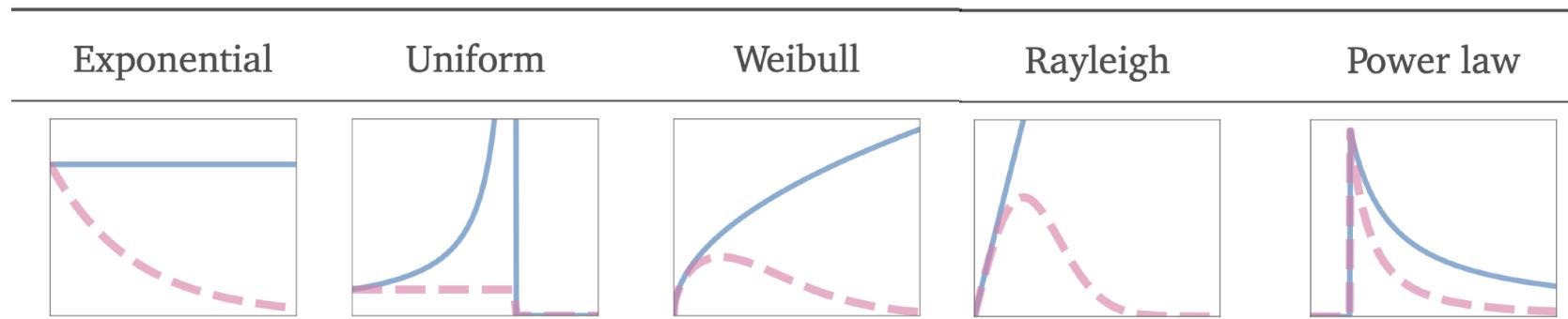
Rate functions are universal:



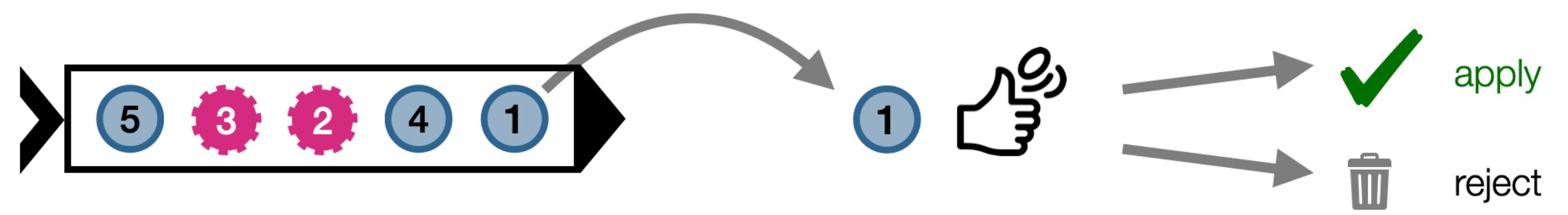
- ▶ Time-dependent rate function
- ▶ Jump-time probability density

Non-Markovian Simulation

Rate functions are universal:



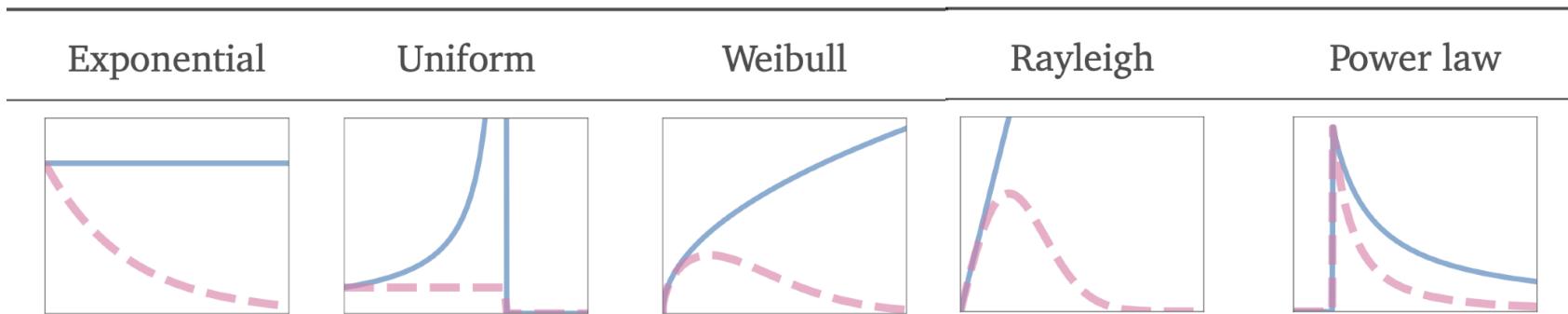
Over-approximations can be balanced out with rejections.



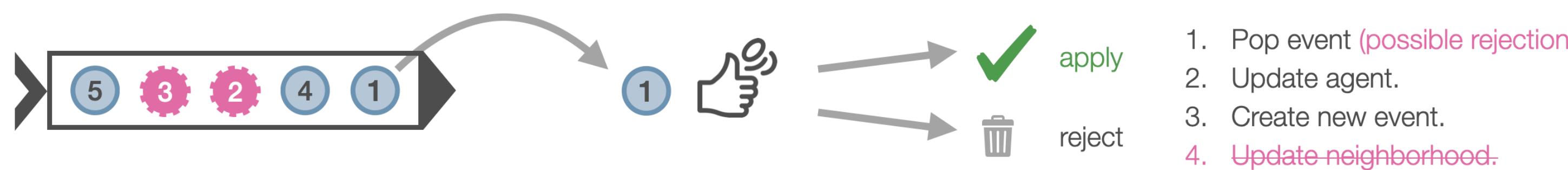
1. Pop event (possible rejection).
2. Update agent.
3. Create new event.
4. ~~Update neighborhood.~~

Non-Markovian Results

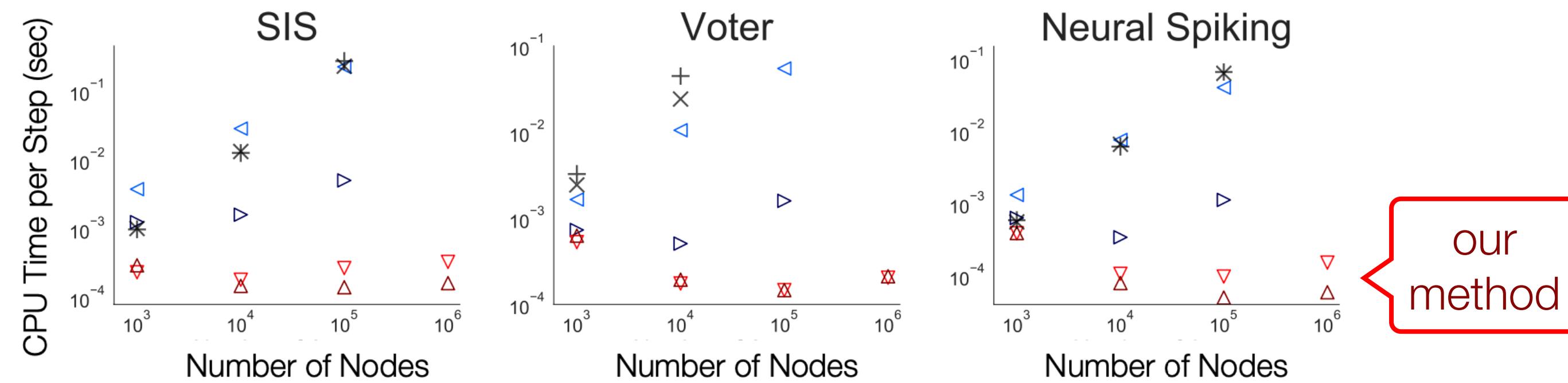
Rate functions are universal:



Over-approximations can be balanced out with rejections.



Results:



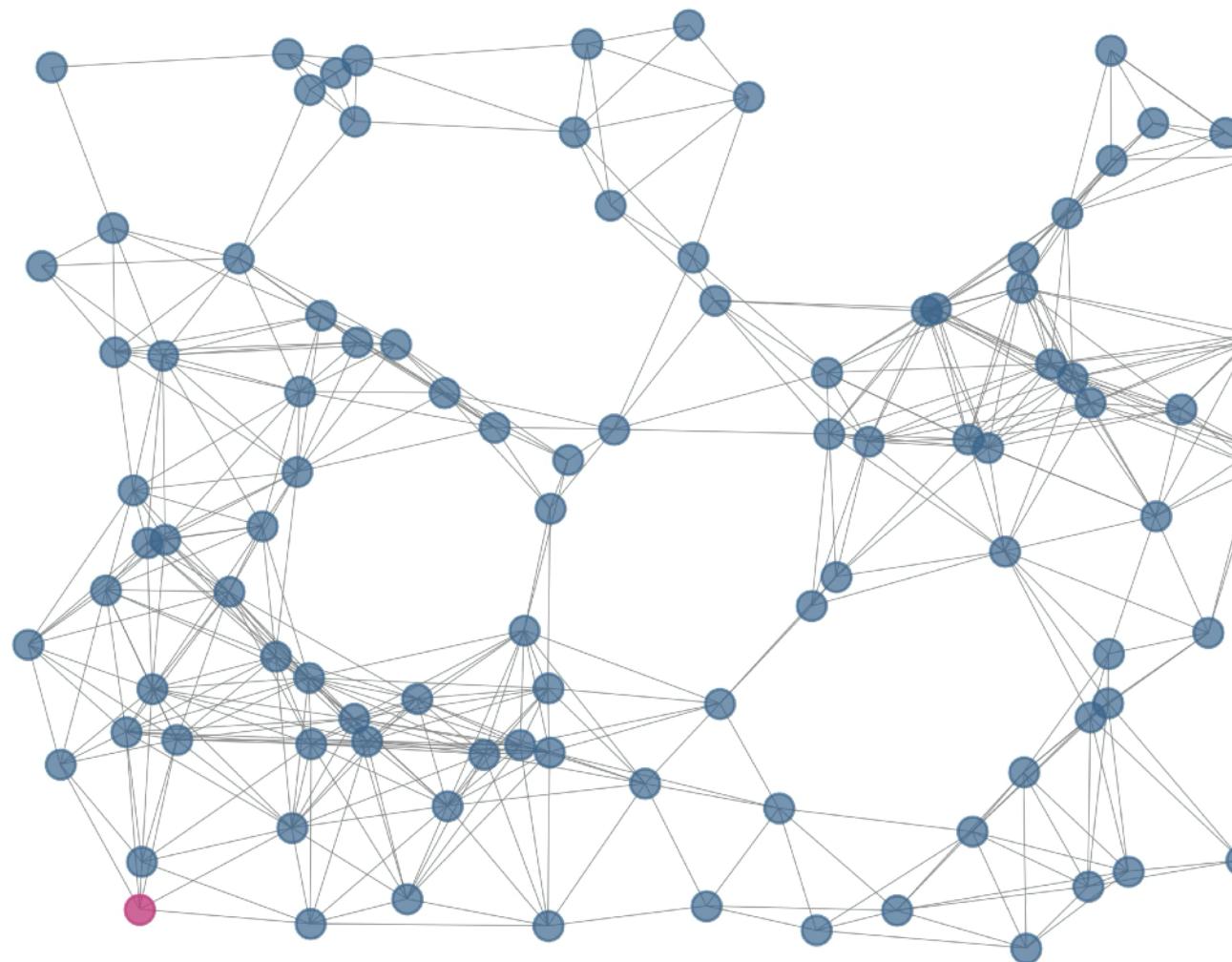


Agenda

METHODS

COVID
(NON-)MARKOVIAN SIMULATION
CONTROL
INFERENCE

Problem Setting

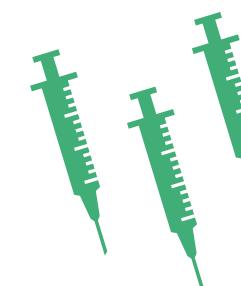


Input

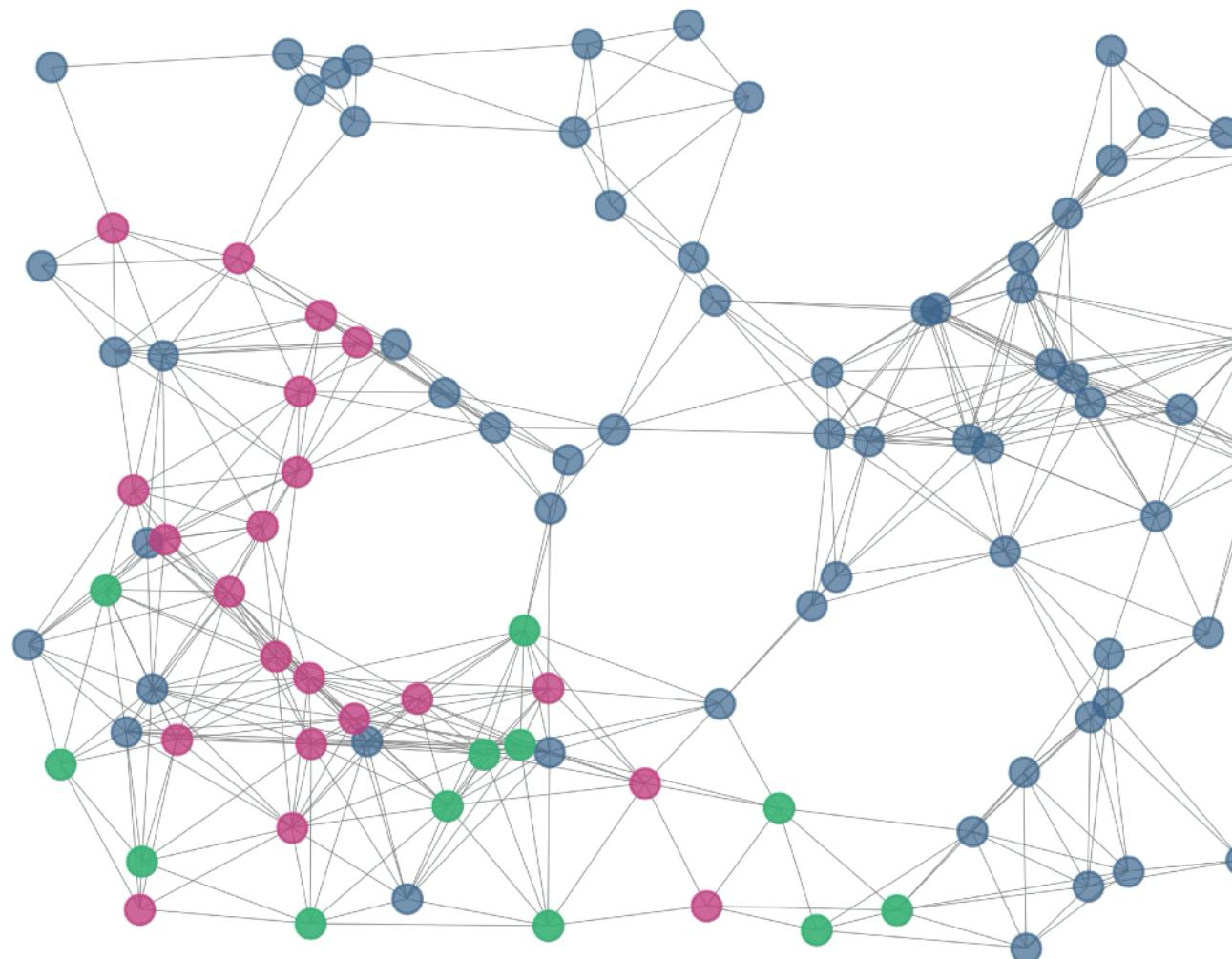
- ▶ Contact graph
- ▶ **SIR**-type spreading model
- ▶ Number of vaccine doses
- ▶ Patient zero

Question

- ▶ Optimal distribution of vaccine doses?



Problem Setting

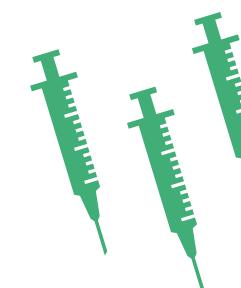


Input

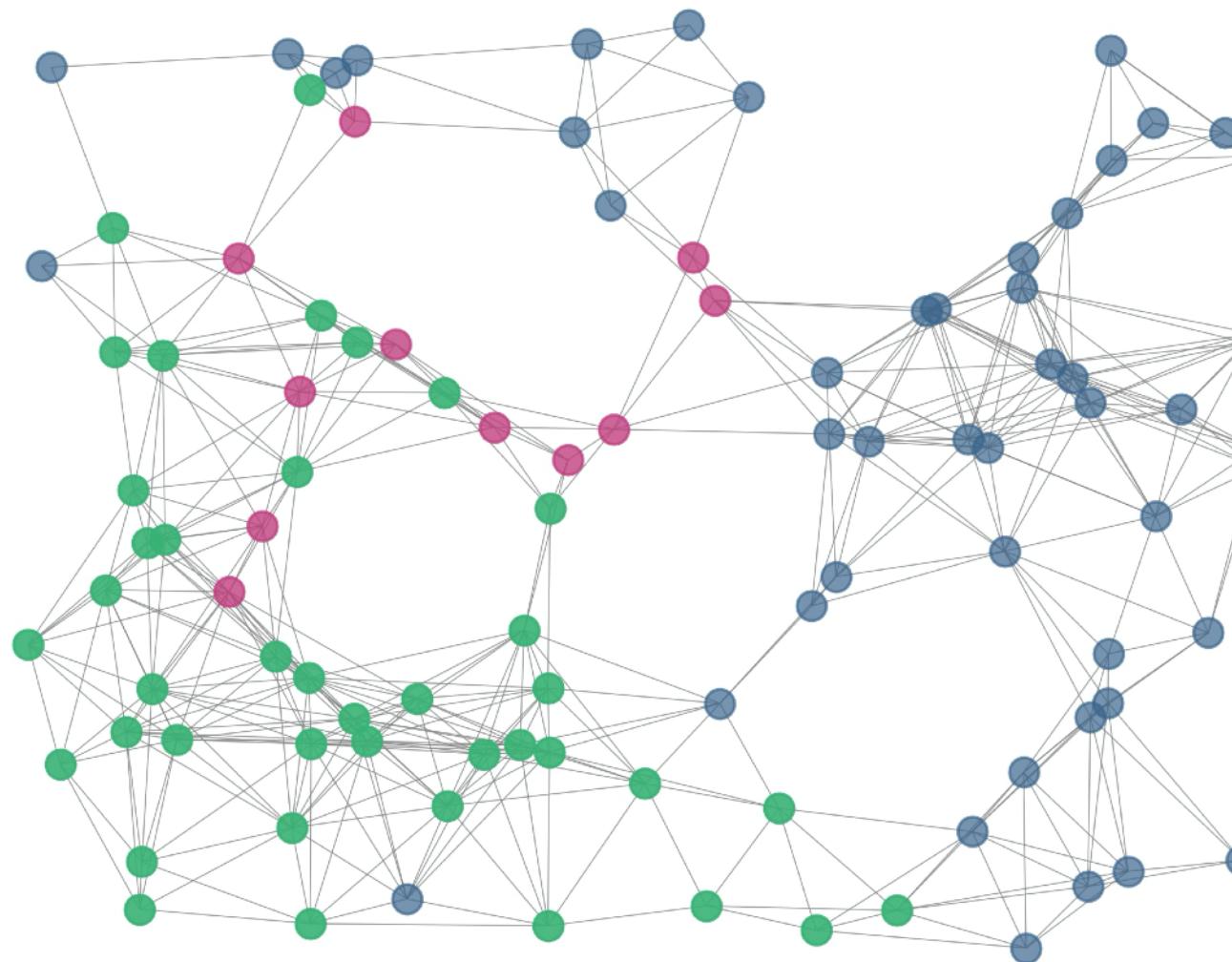
- ▶ Contact graph
- ▶ **SIR**-type spreading model
- ▶ Number of vaccine doses
- ▶ Patient zero

Question

- ▶ Optimal distribution of vaccine doses?



Problem Setting

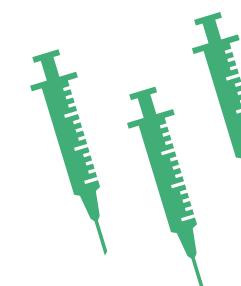


Input

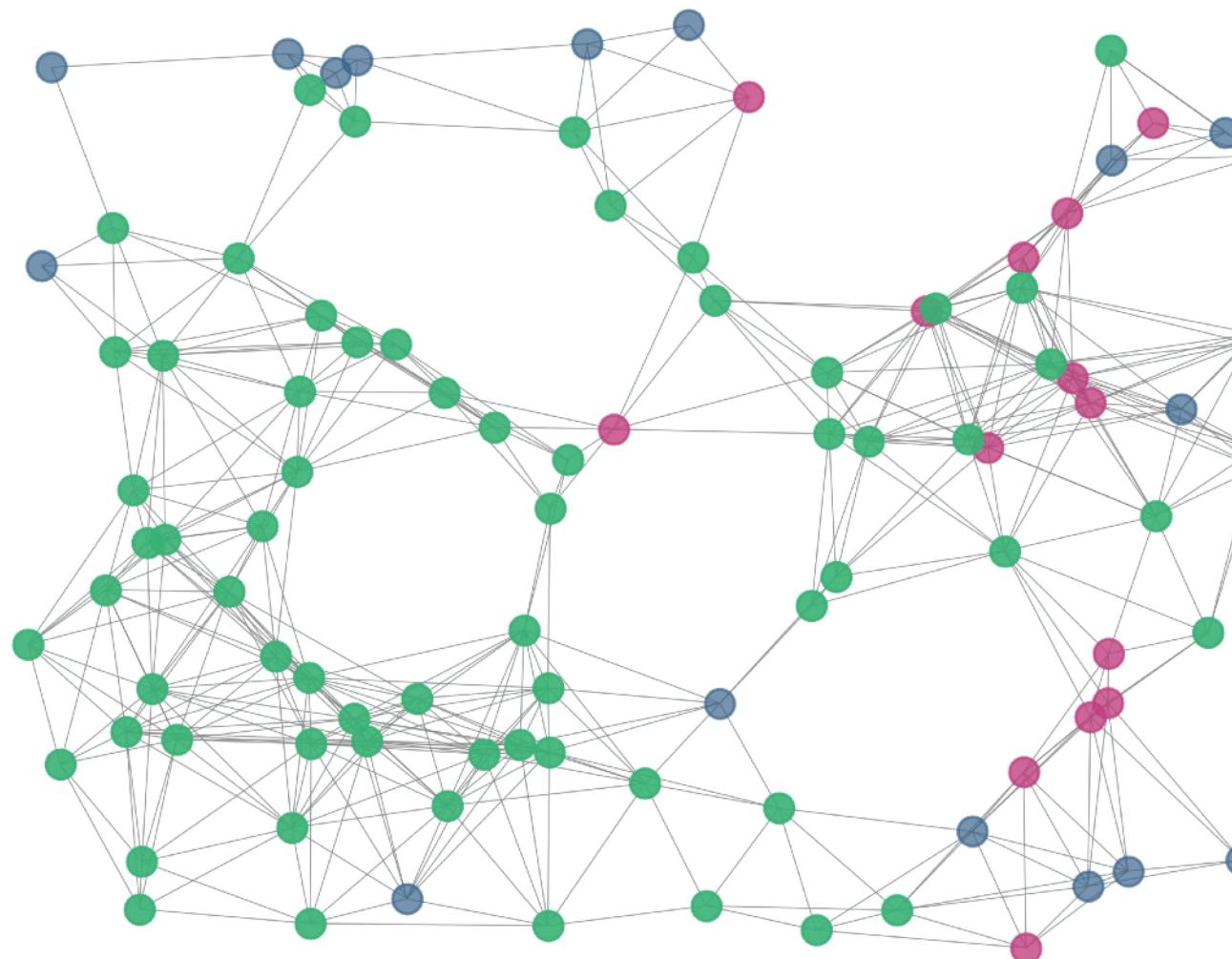
- ▶ Contact graph
- ▶ **SIR**-type spreading model
- ▶ Number of vaccine doses
- ▶ Patient zero

Question

- ▶ Optimal distribution of vaccine doses?



Problem Setting

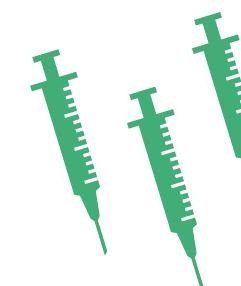


Input

- ▶ Contact graph
- ▶ **SIR**-type spreading model
- ▶ Number of vaccine doses
- ▶ Patient zero

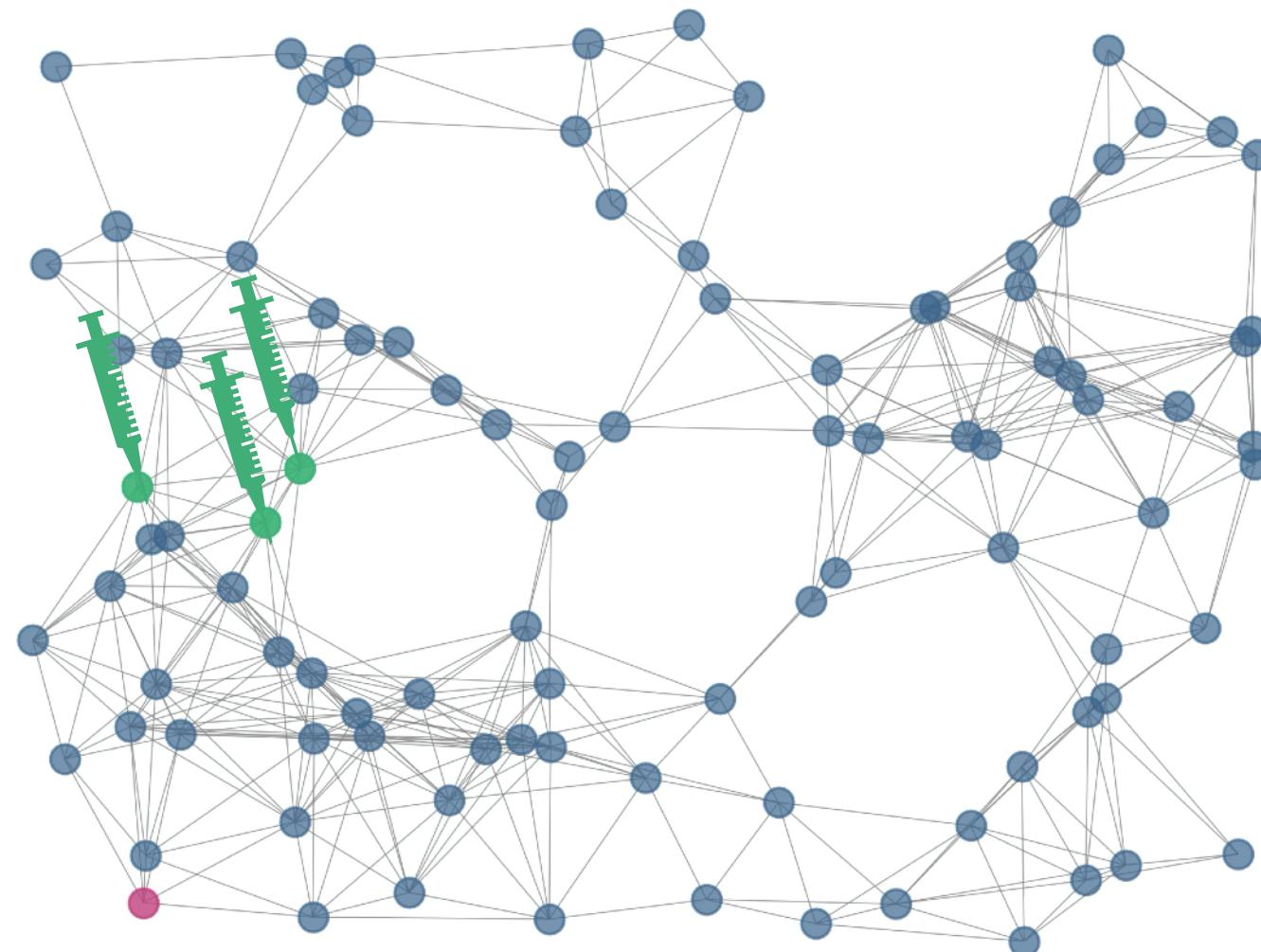
Question

- ▶ Optimal distribution of vaccine doses?

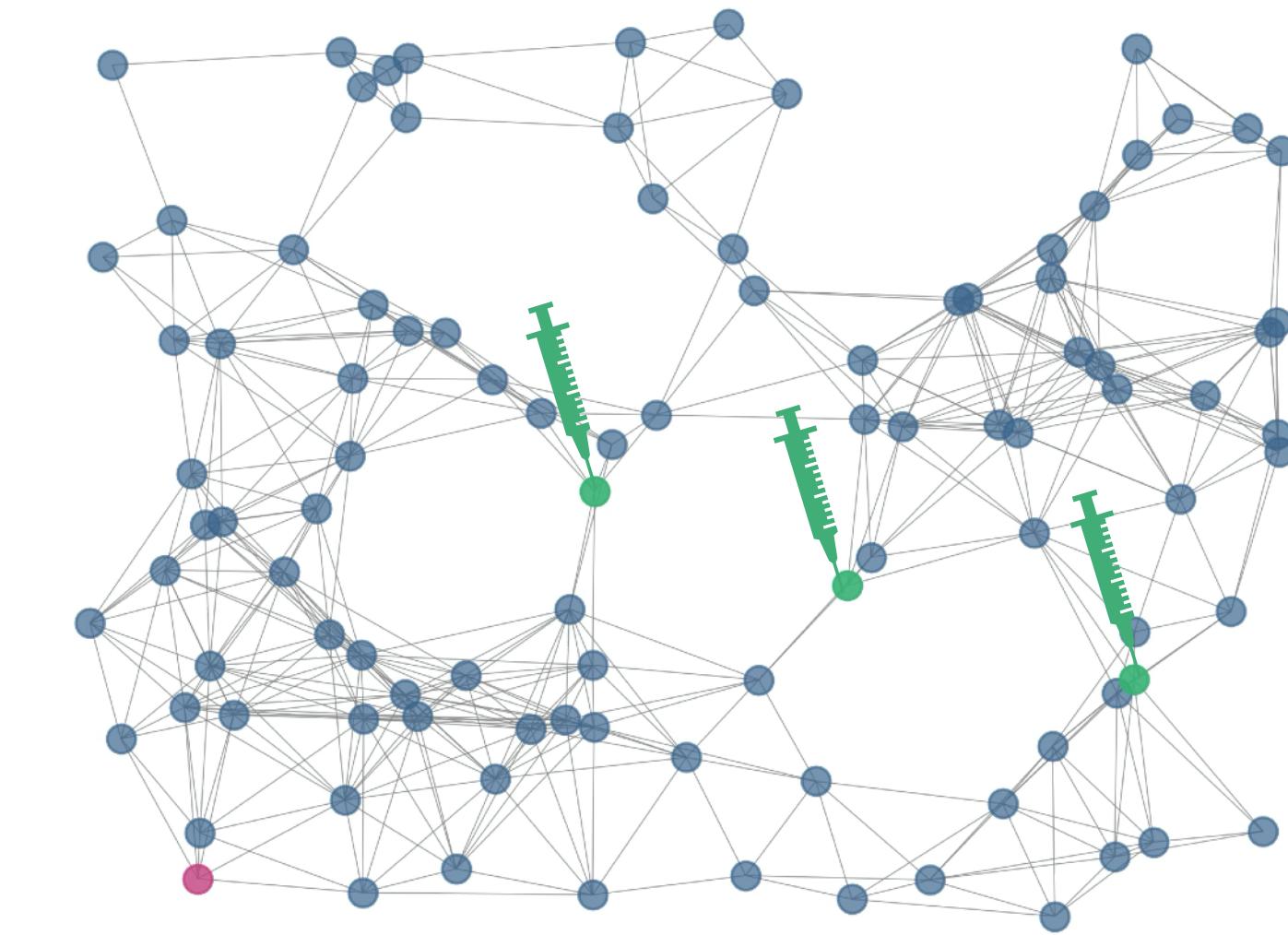
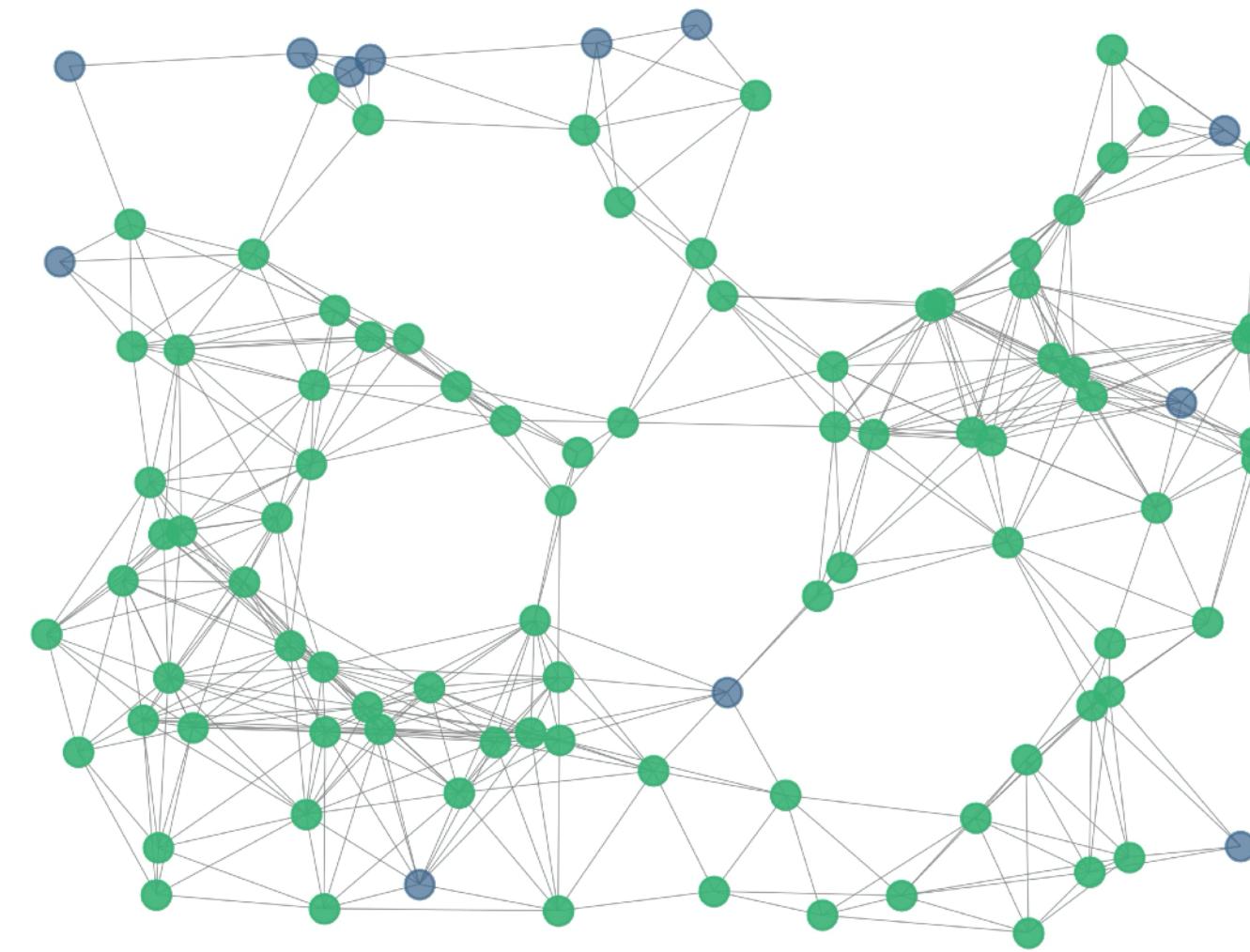


Problem Setting

Strategy 1



Strategy 2



Input

- ▶ Contact graph
- ▶ **SIR**-type spreading model
- ▶ Number of vaccine doses
- ▶ Patient zero

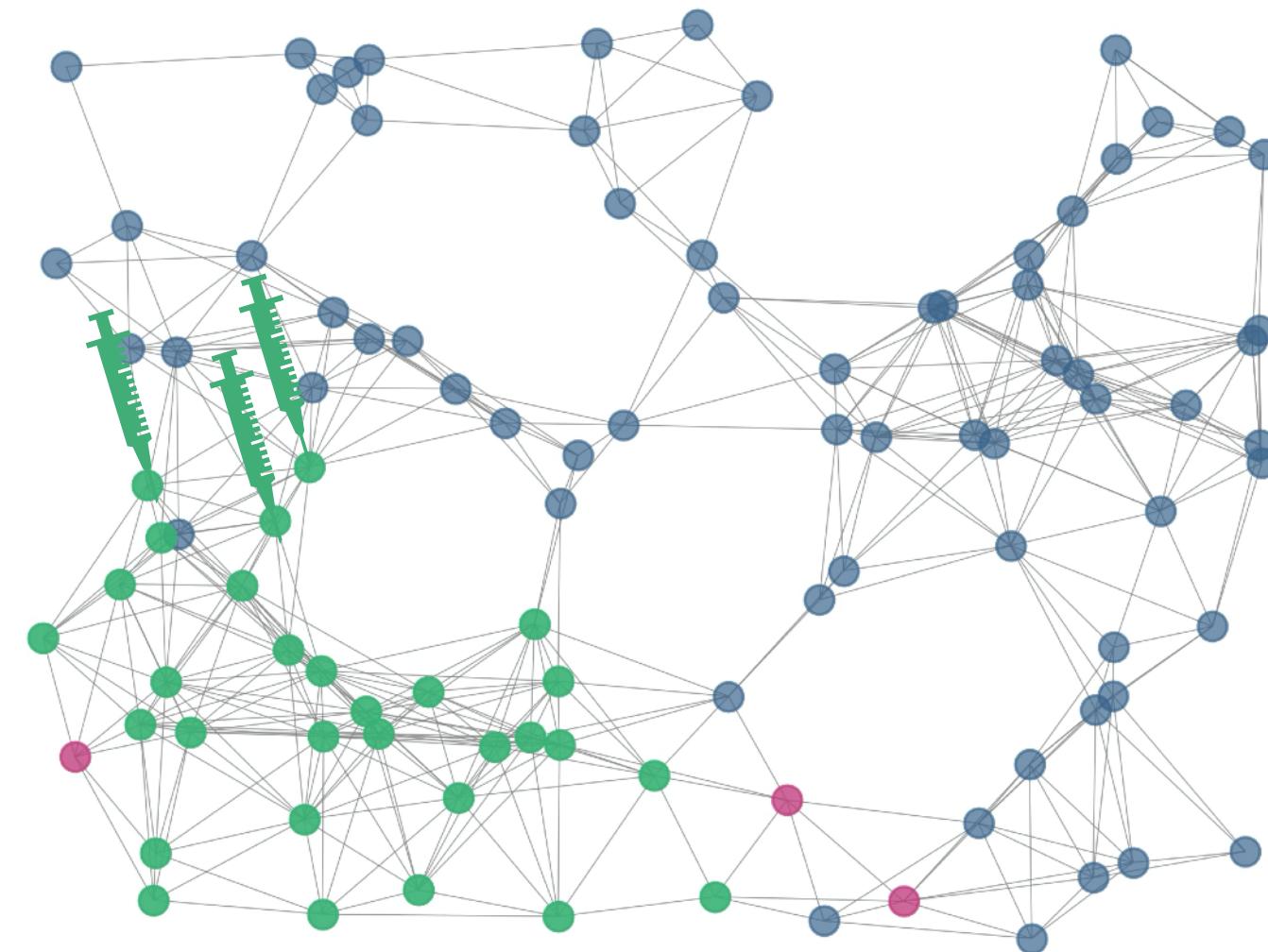
Question

- ▶ Optimal distribution of vaccine doses?

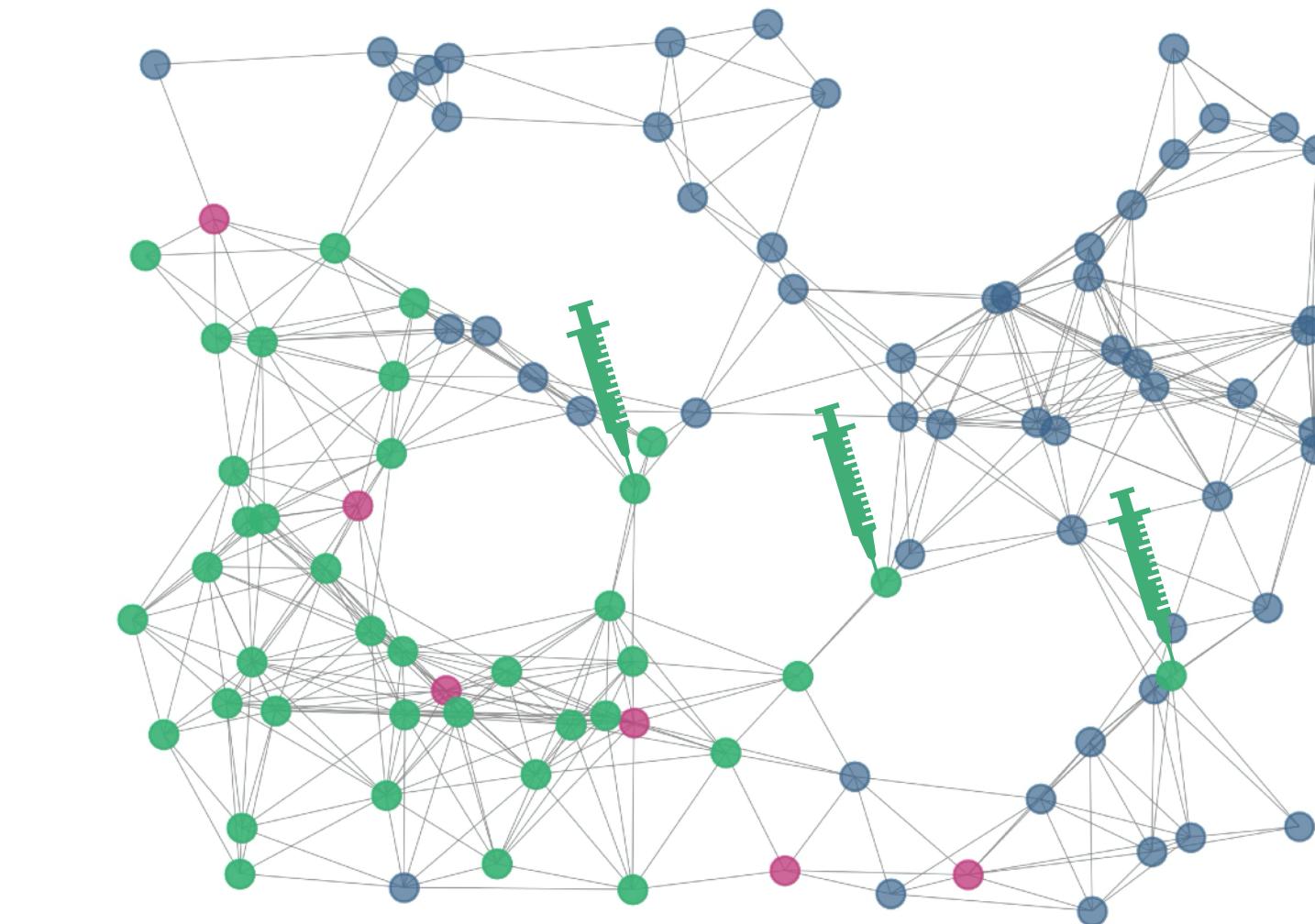
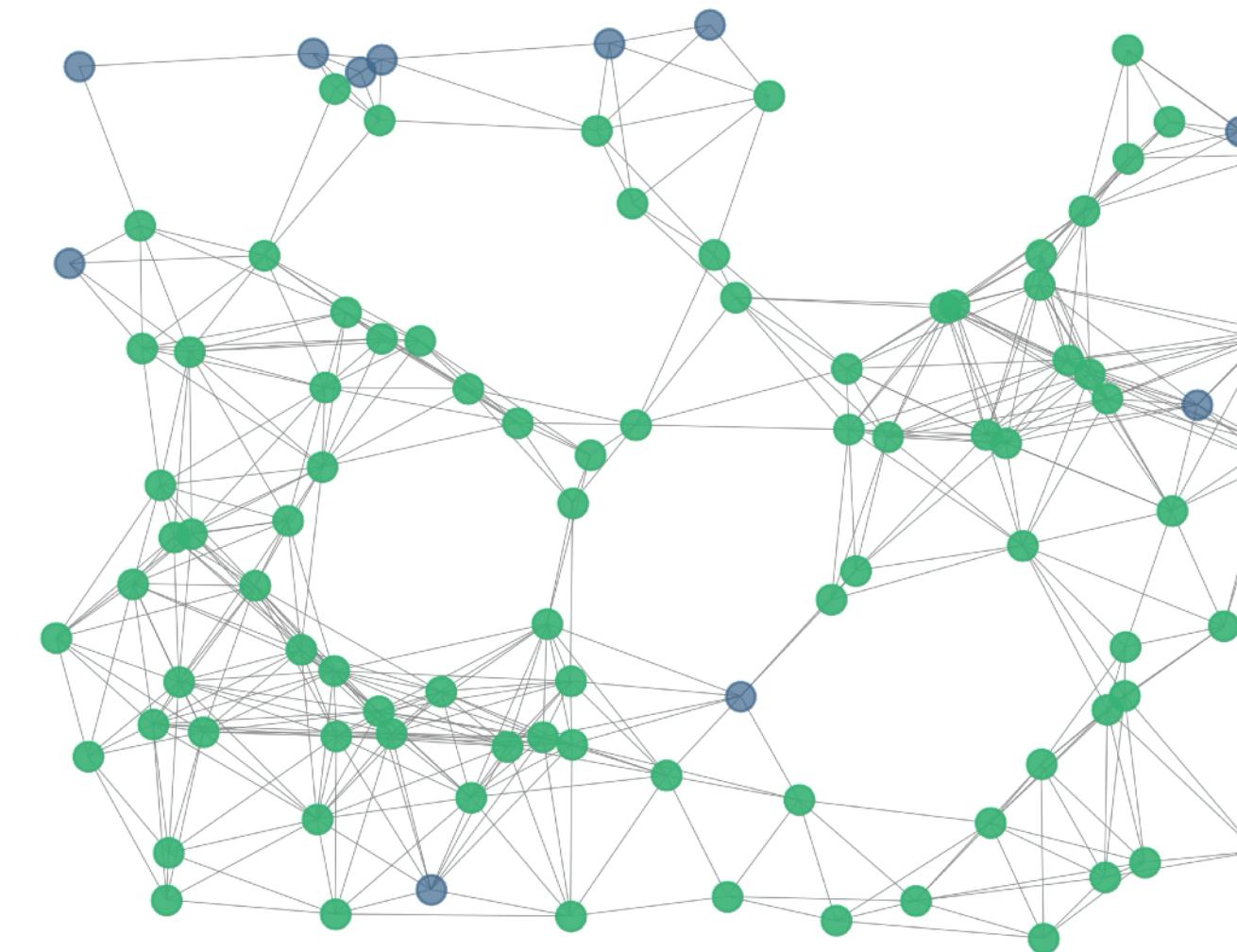


Problem Setting

Strategy 1



Strategy 2

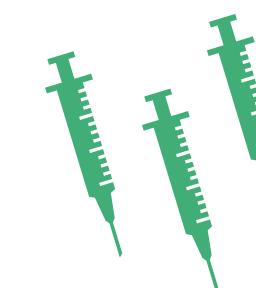


Input

- ▶ Contact graph
- ▶ **SIR**-type spreading model
- ▶ Number of vaccine doses
- ▶ Patient zero

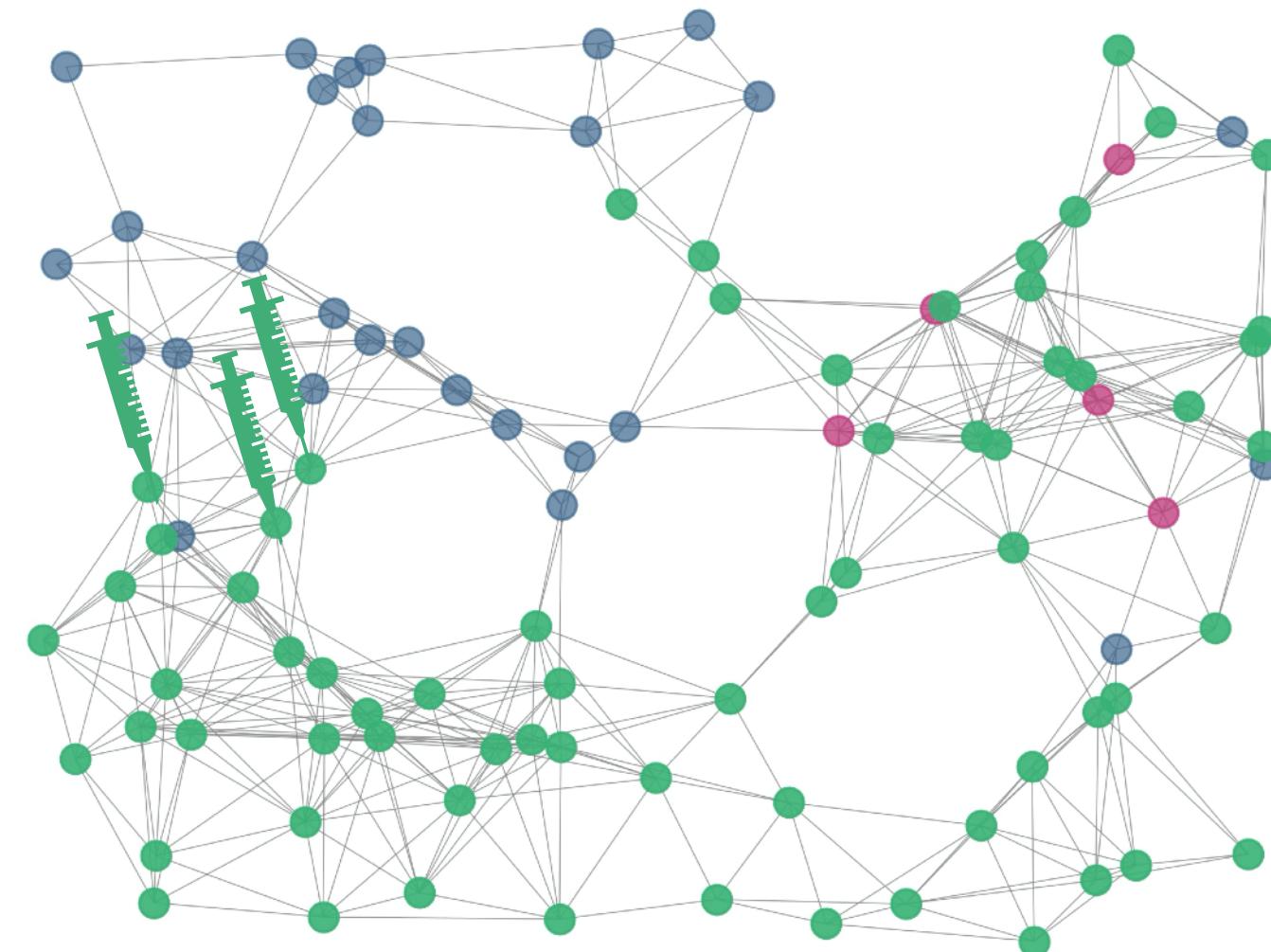
Question

- ▶ Optimal distribution of vaccine doses?

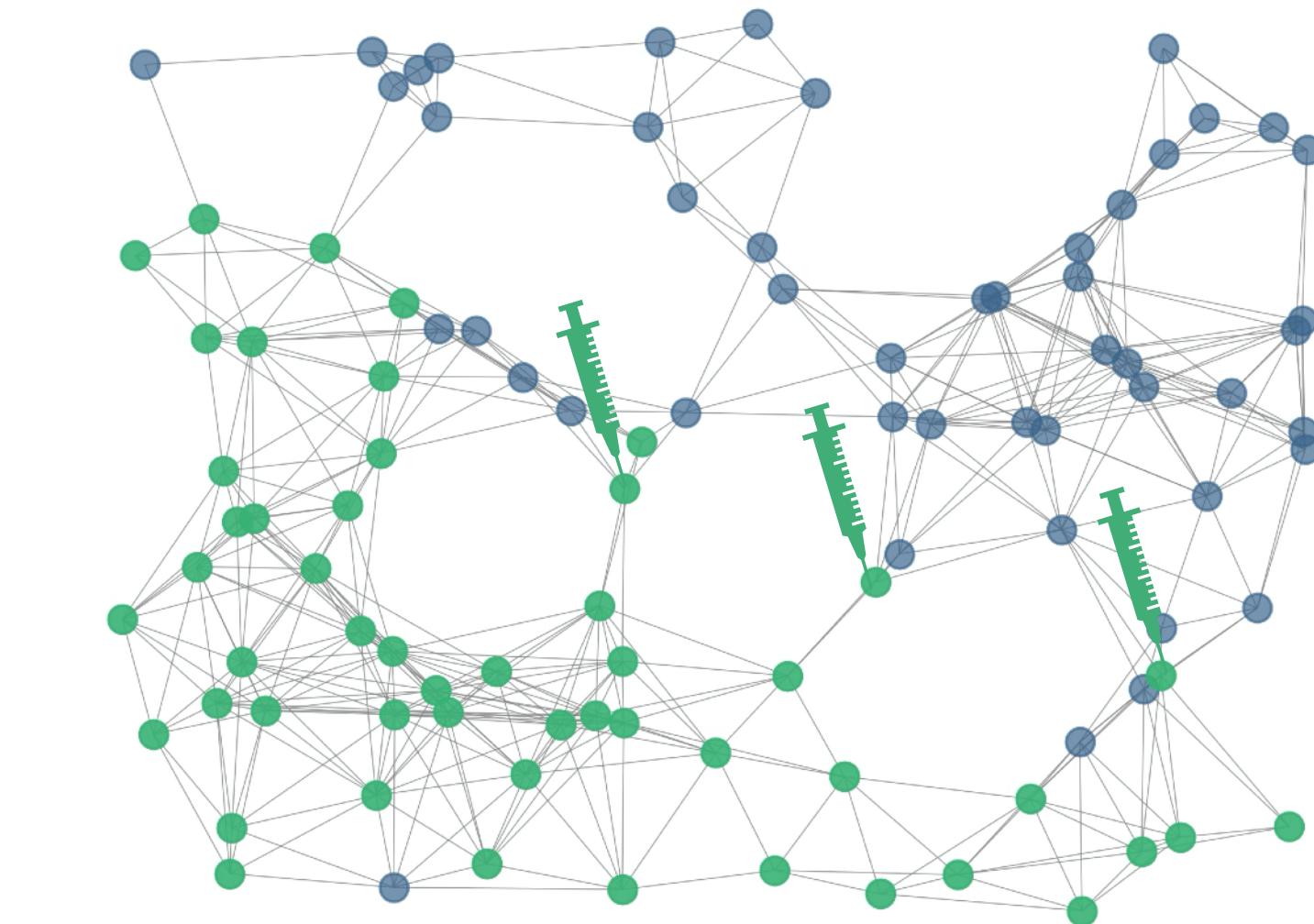
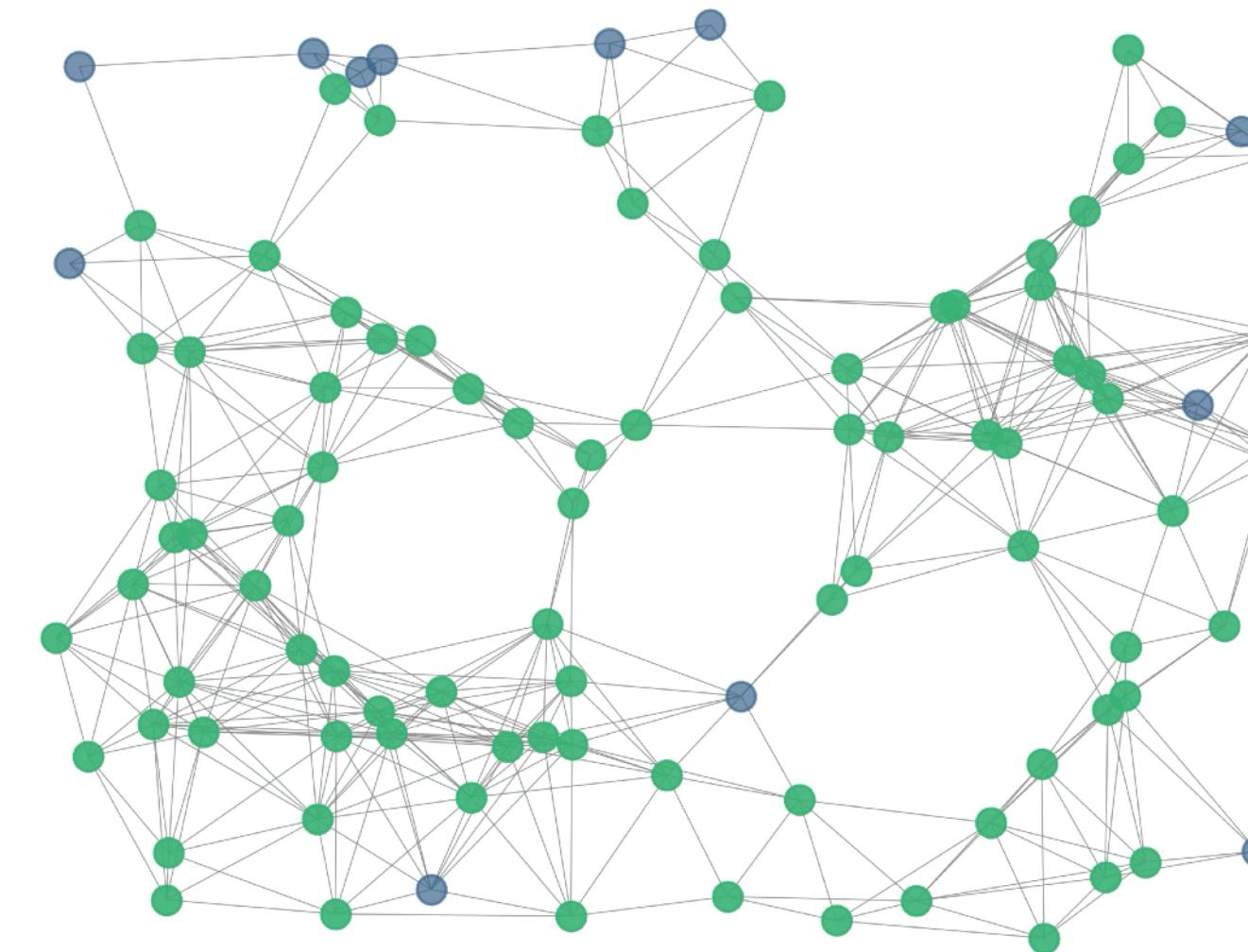


Problem Setting

Strategy 1



Strategy 2

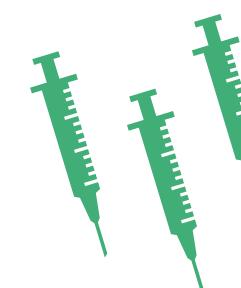


Input

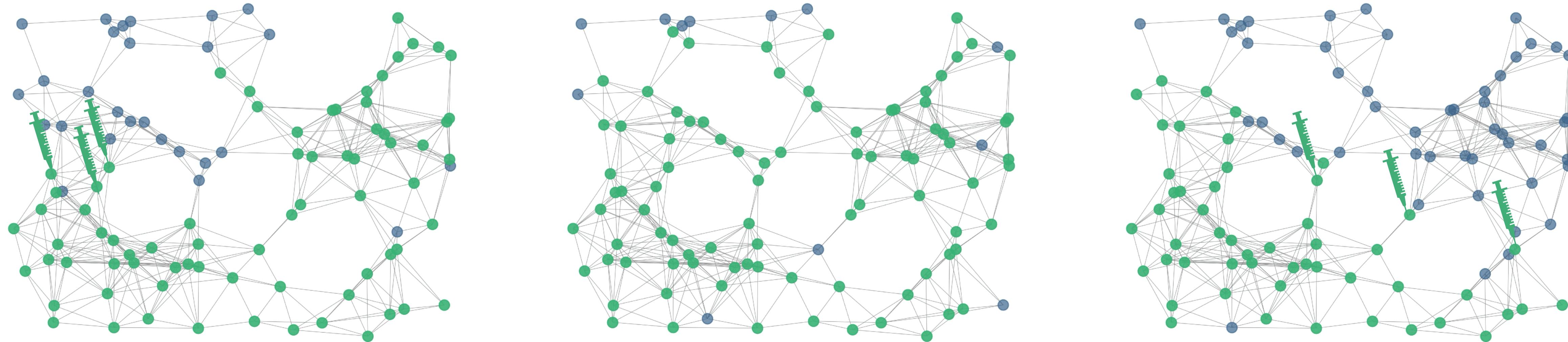
- ▶ Contact graph
- ▶ **SIR**-type spreading model
- ▶ Number of vaccine doses
- ▶ Patient zero

Question

- ▶ Optimal distribution of vaccine doses?



Problem Setting

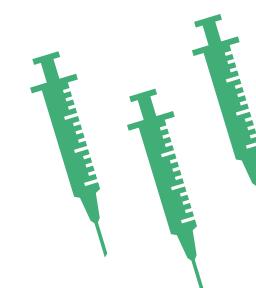


Input

- ▶ Contact graph
- ▶ **SIR**-type spreading model
- ▶ Number of vaccine doses
- ▶ Patient zero

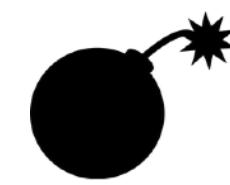
Question

- ▶ Optimal distribution of vaccine doses?

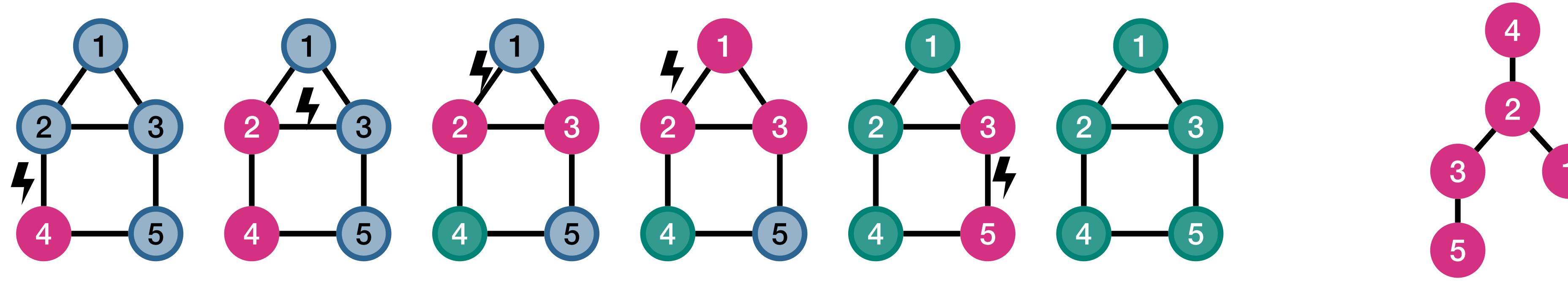


Problem

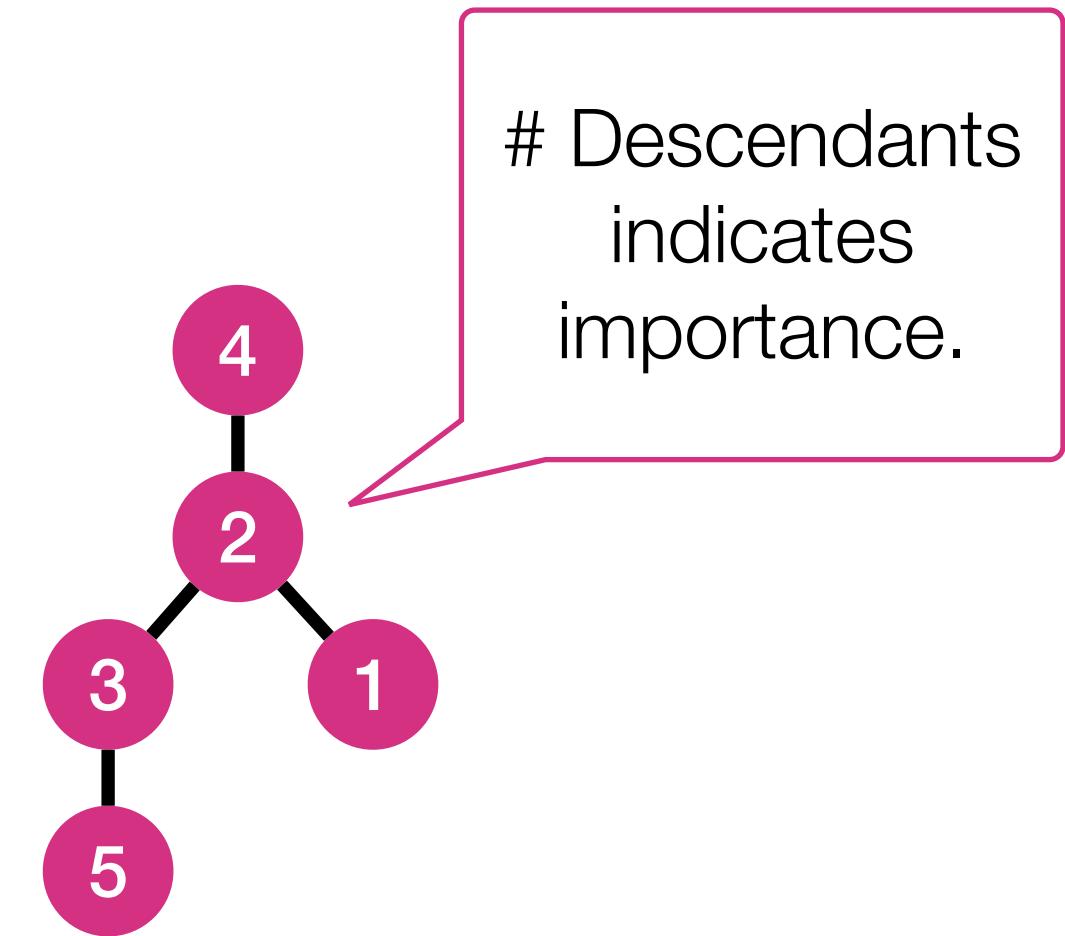
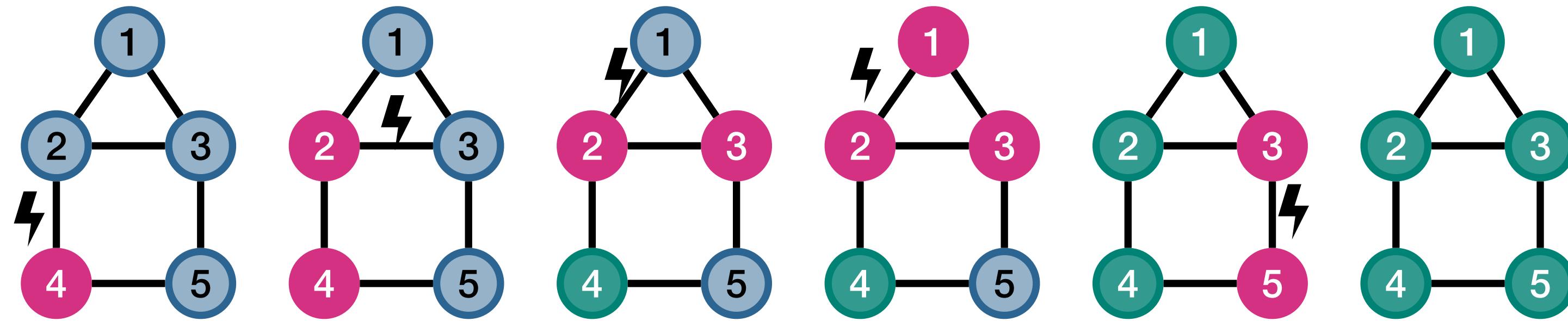
- ▶ Combinatorial explosion



Transmission Trees

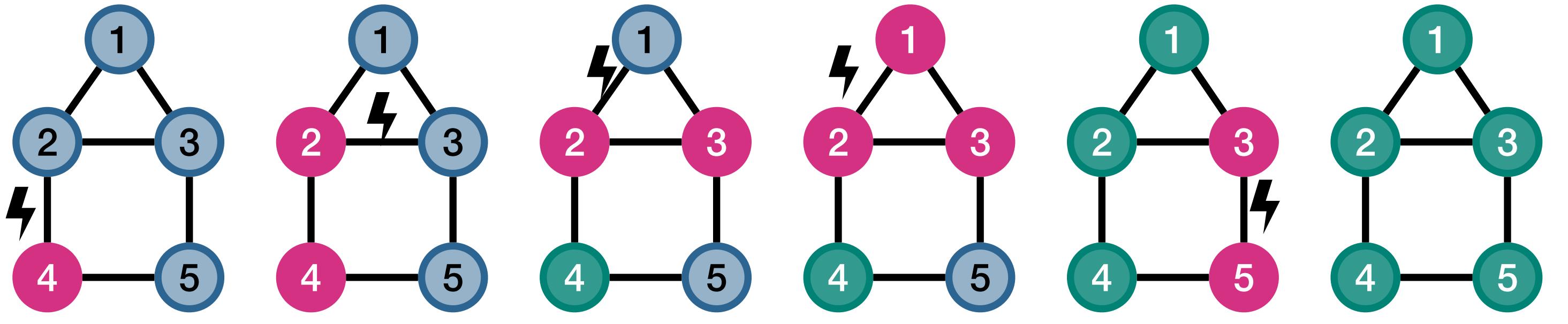


Transmission Trees

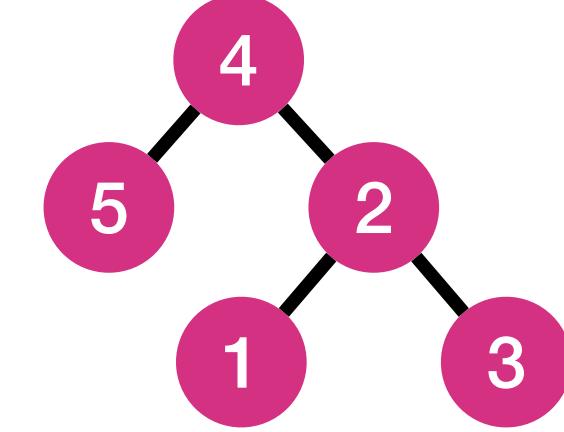
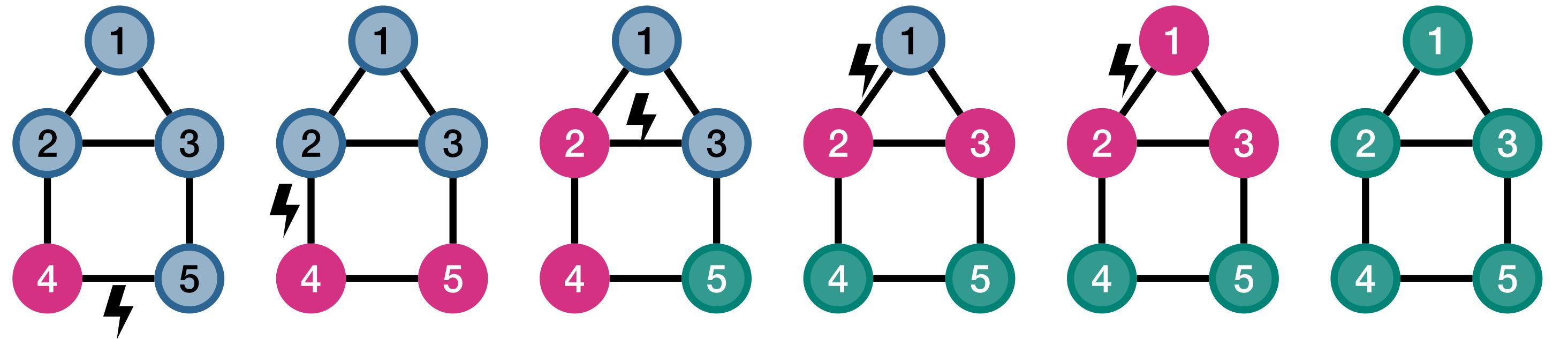
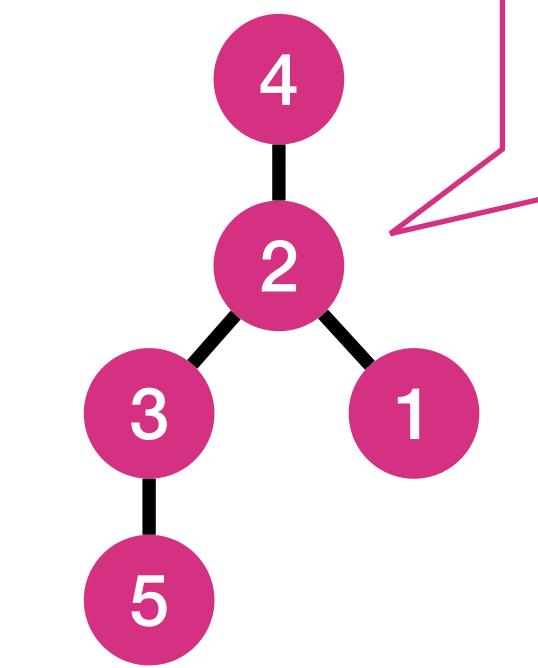


Descendants
indicates
importance.

Transmission Trees

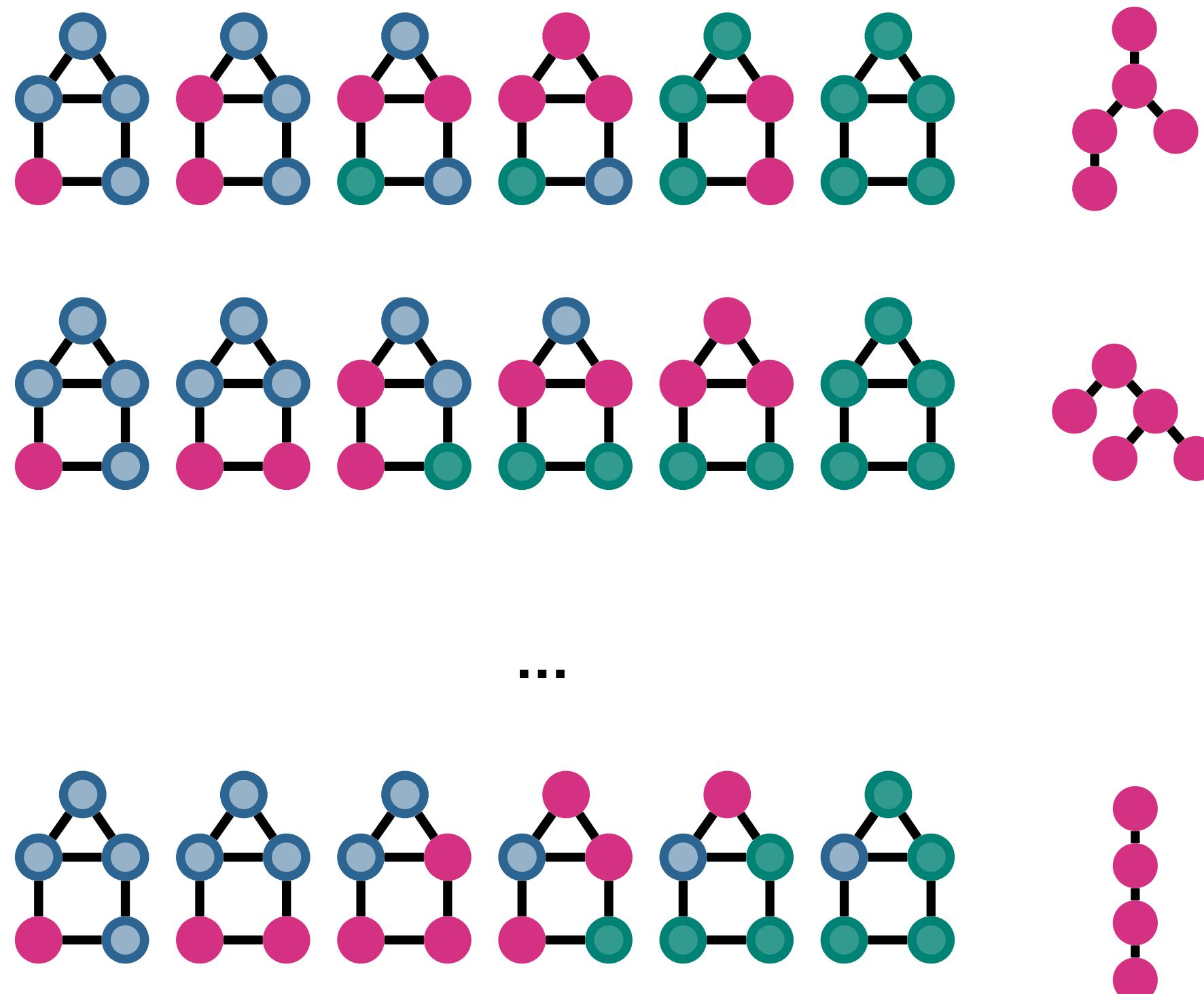


Descendants
indicates
importance.

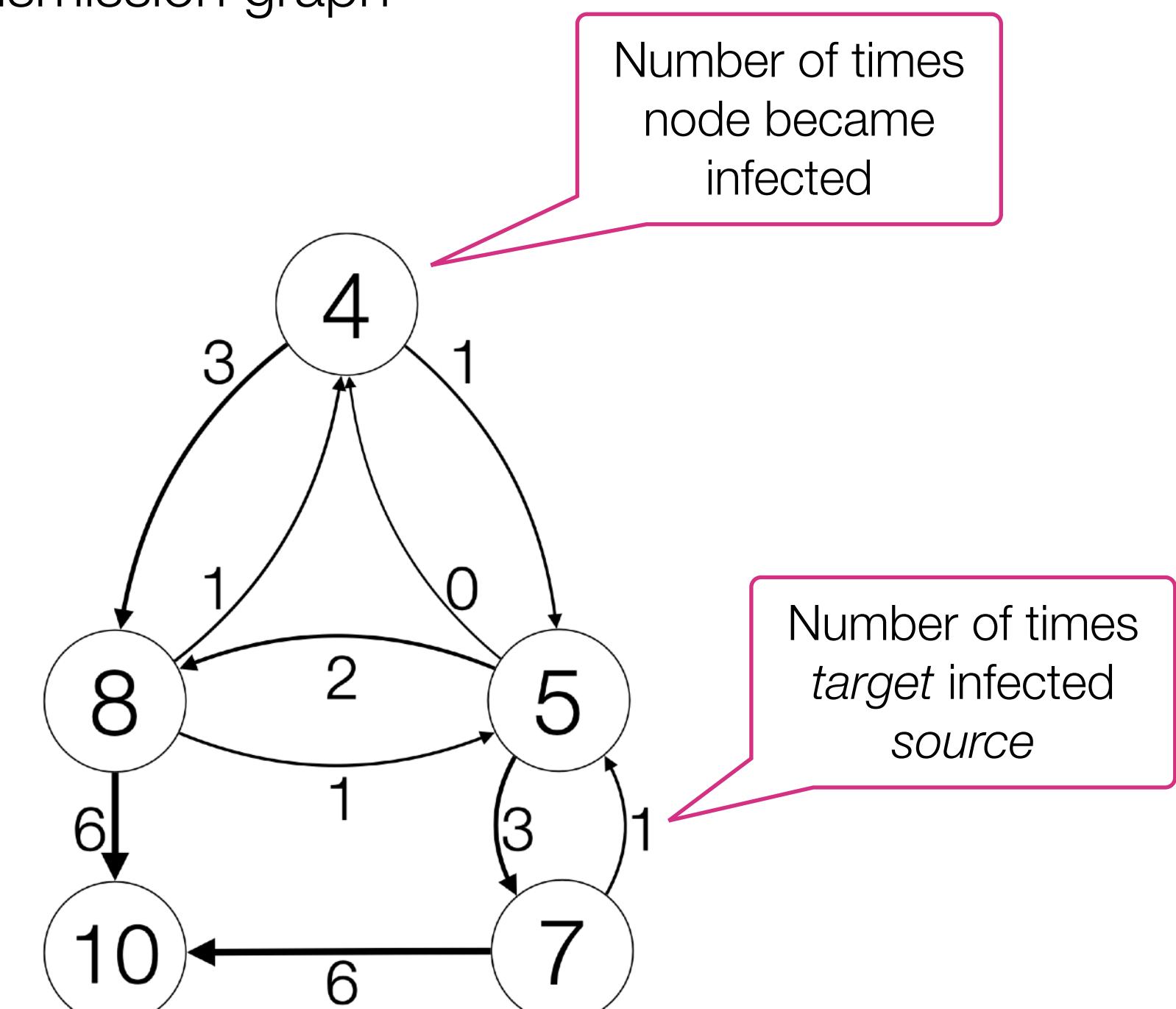


From Trees to Graphs

Simulation runs
(10 samples)

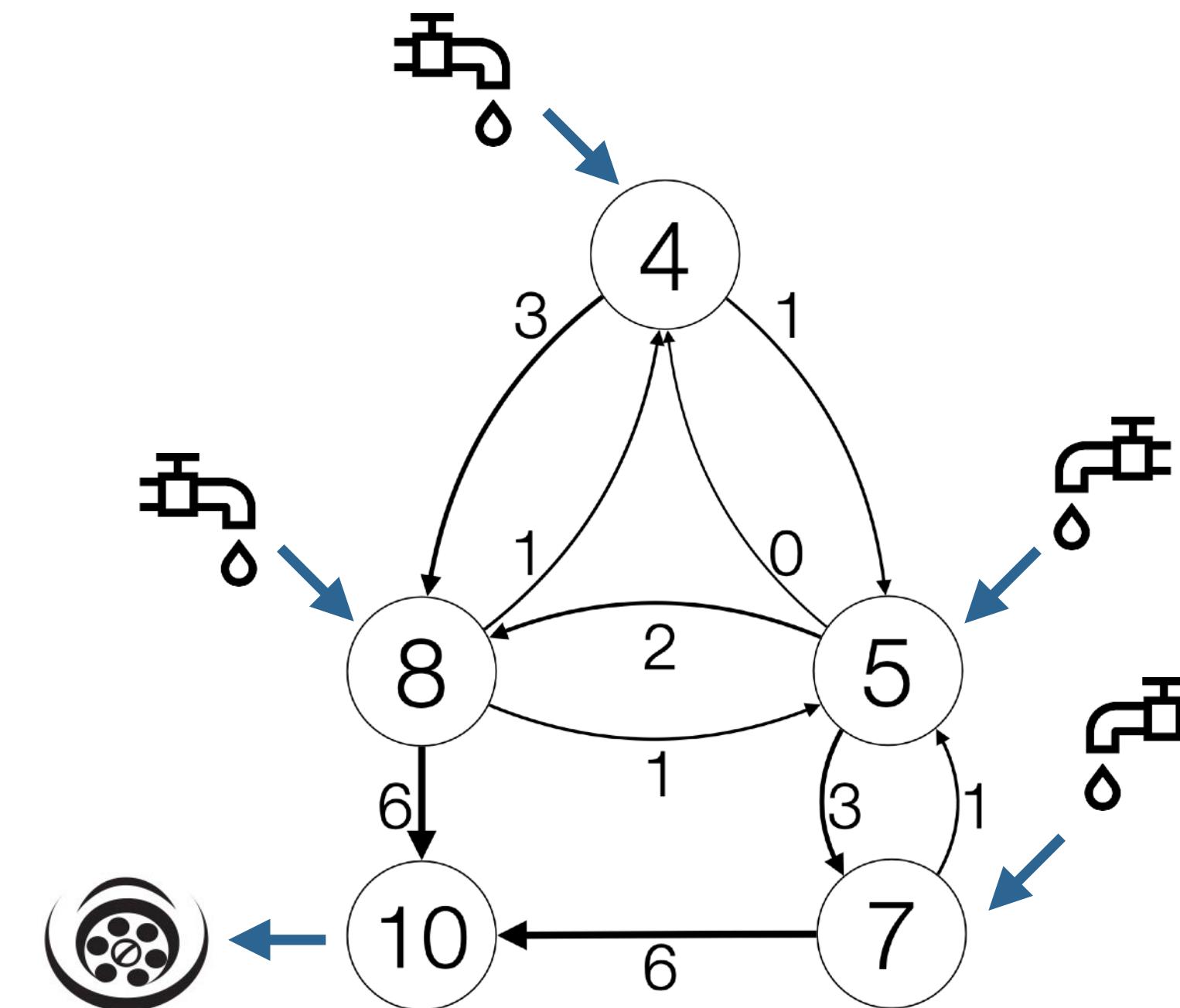


Transmission graph



Flow Analysis

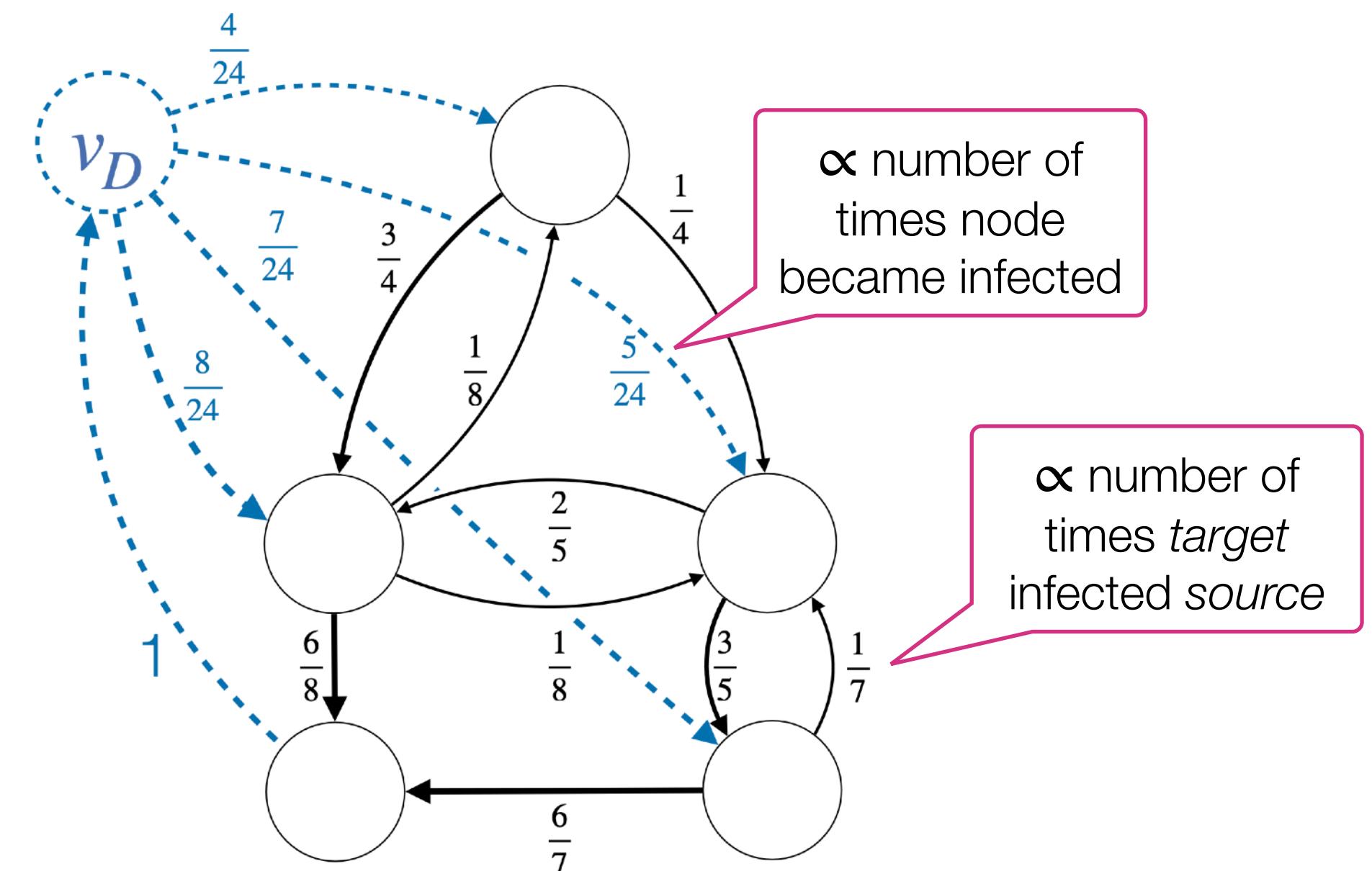
Transmission graph



- ▶ **Source:** Each infected node
- ▶ **Sink:** Patient zero
- ▶ Reversed direction of propagation
- ▶ Flow strength = proportional to transmission impact.

Flows as Random Walks

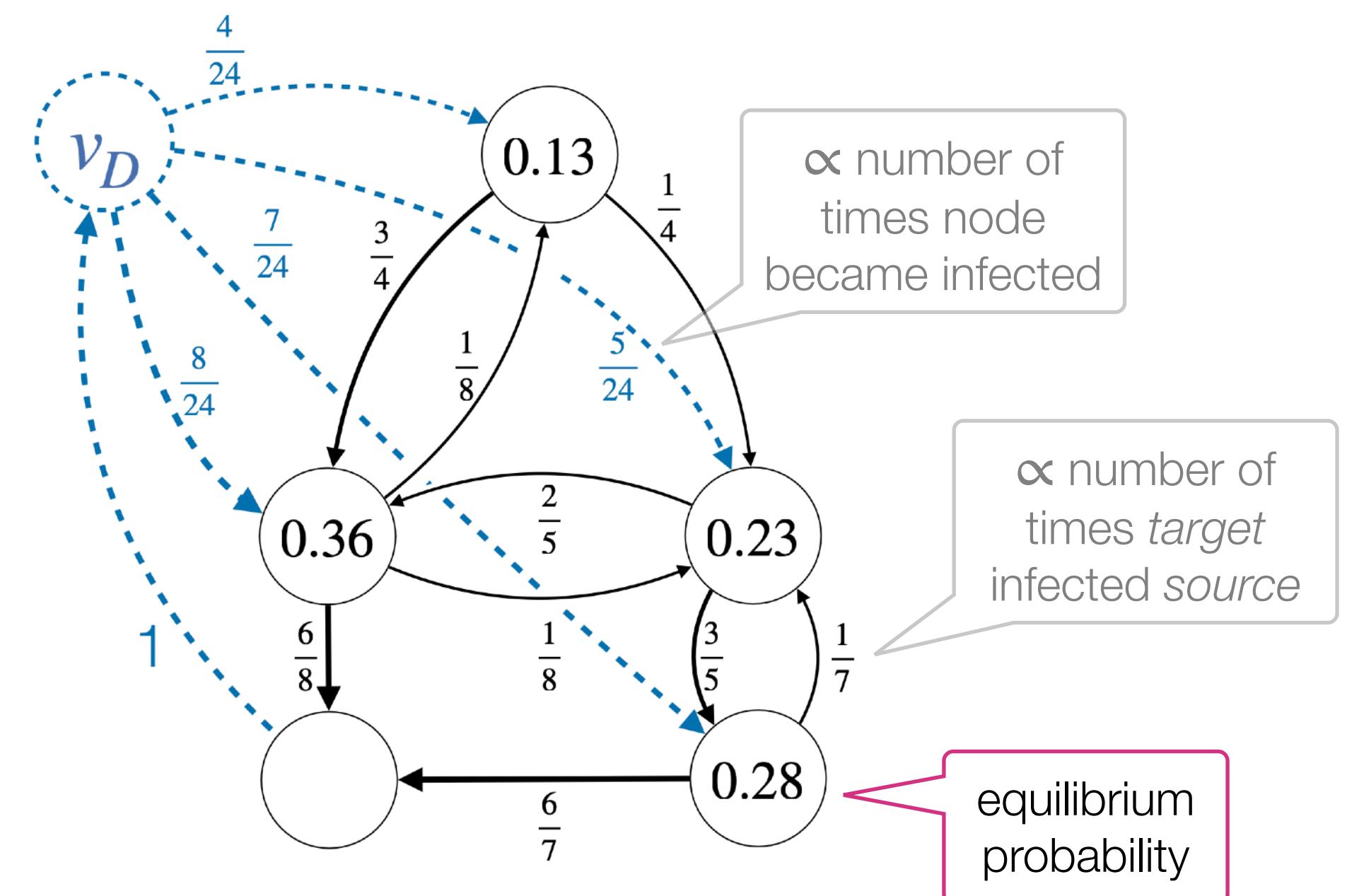
Transmission graph



- ▶ **Source:** Each infected node
- ▶ **Sink:** Patient zero
- ▶ Reversed direction of propagation
- ▶ Flow strength = proportional to transmission impact.

Flows as Random Walks

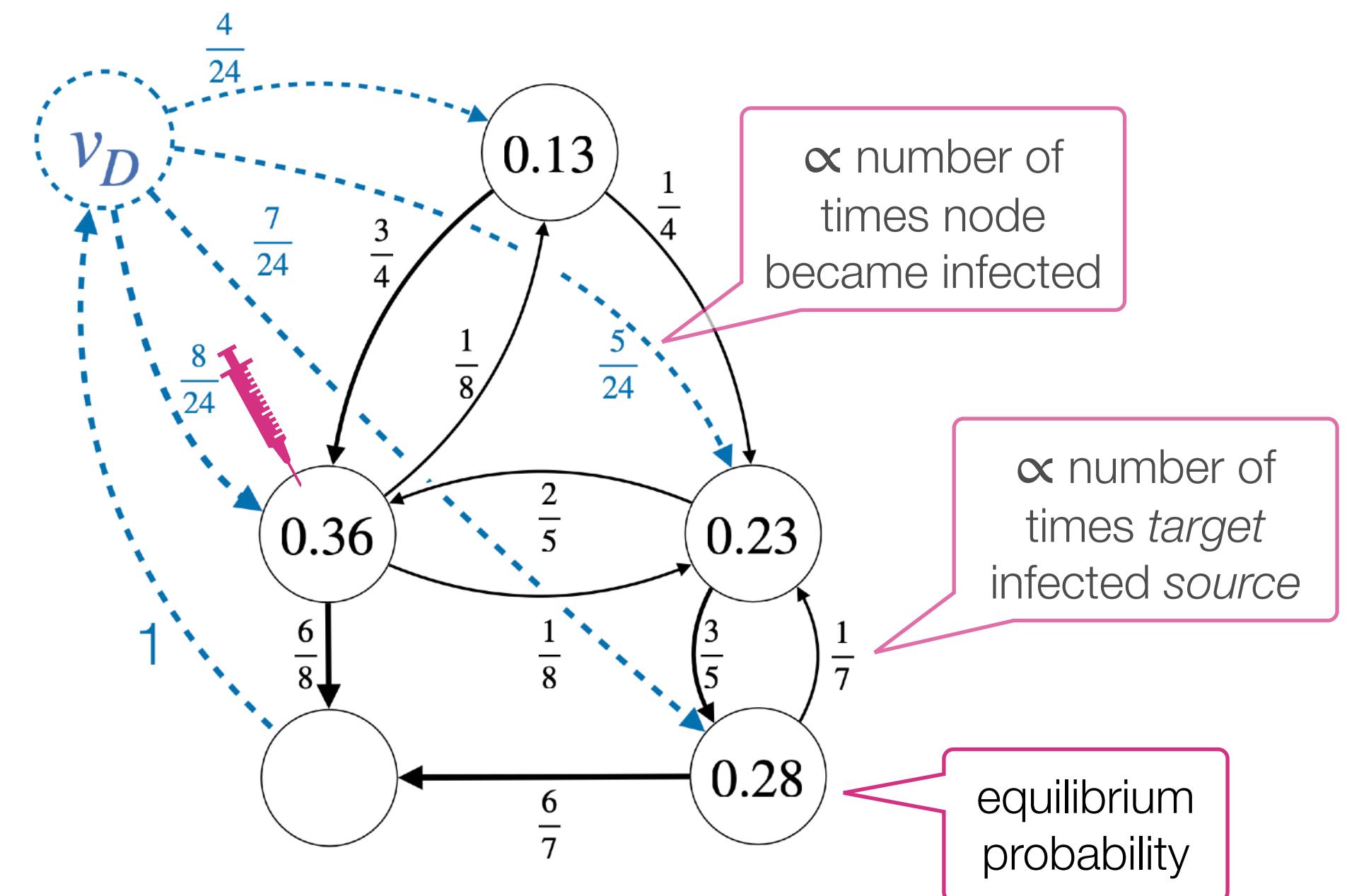
Transmission graph



- **Source:** Each infected node
- **Sink:** Patient zero
- Reversed direction of propagation
- Flow strength = proportional to transmission impact.

Flows as Random Walks

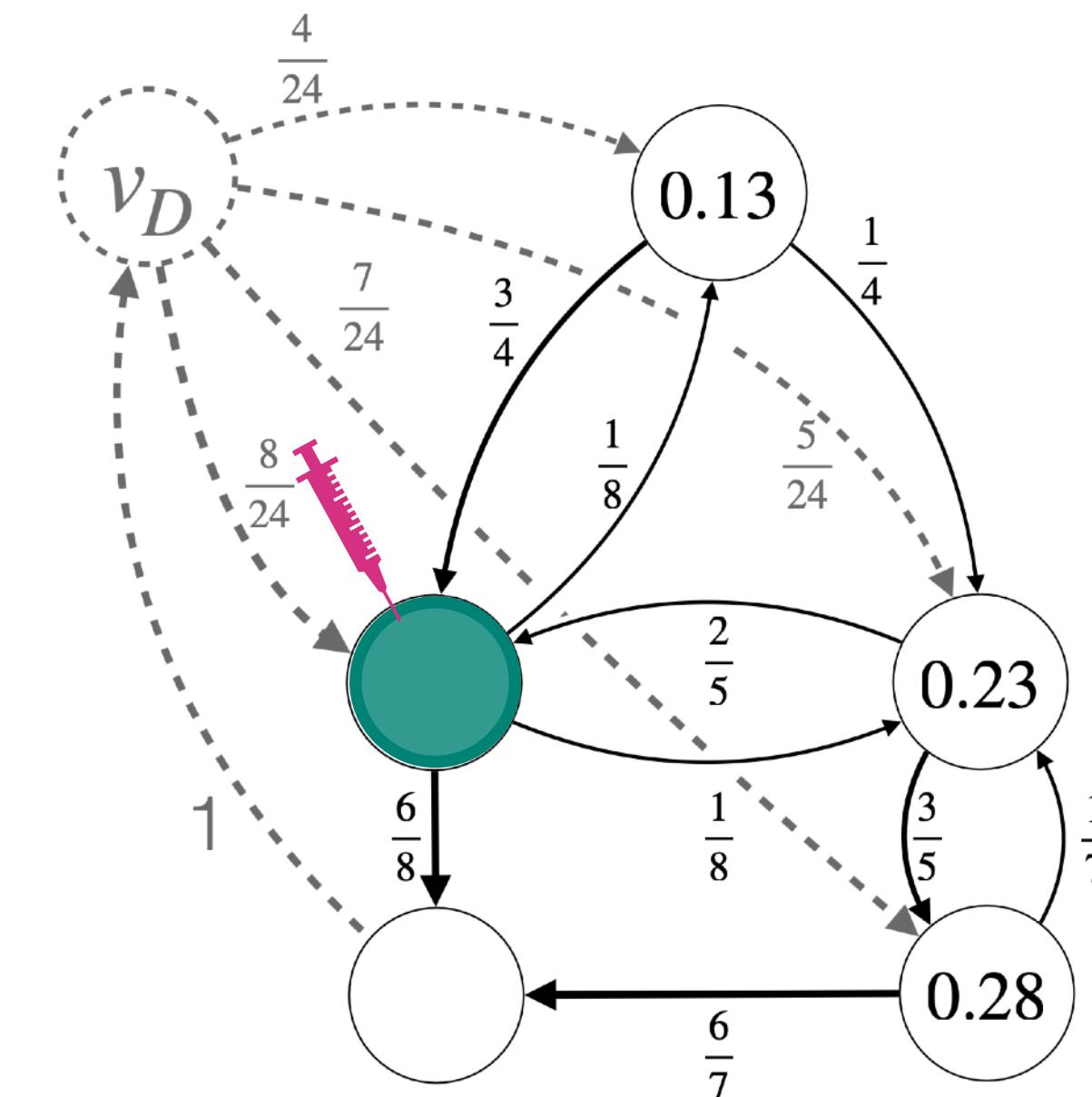
Transmission graph



- **Source:** Each infected node
- **Sink:** Patient zero
- Reversed direction of propagation
- Flow strength = proportional to transmission impact.

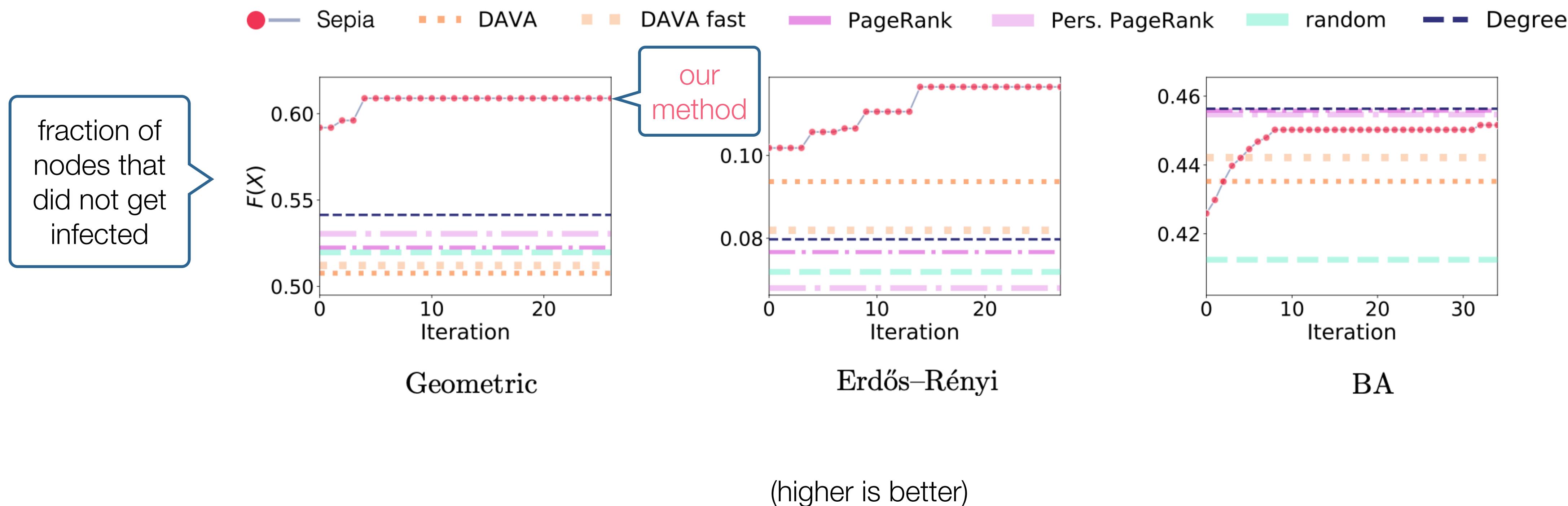
Optimization

Transmission graph



Vaccination
+
Recompute impact score
+
Iterative optimization

Control: Results



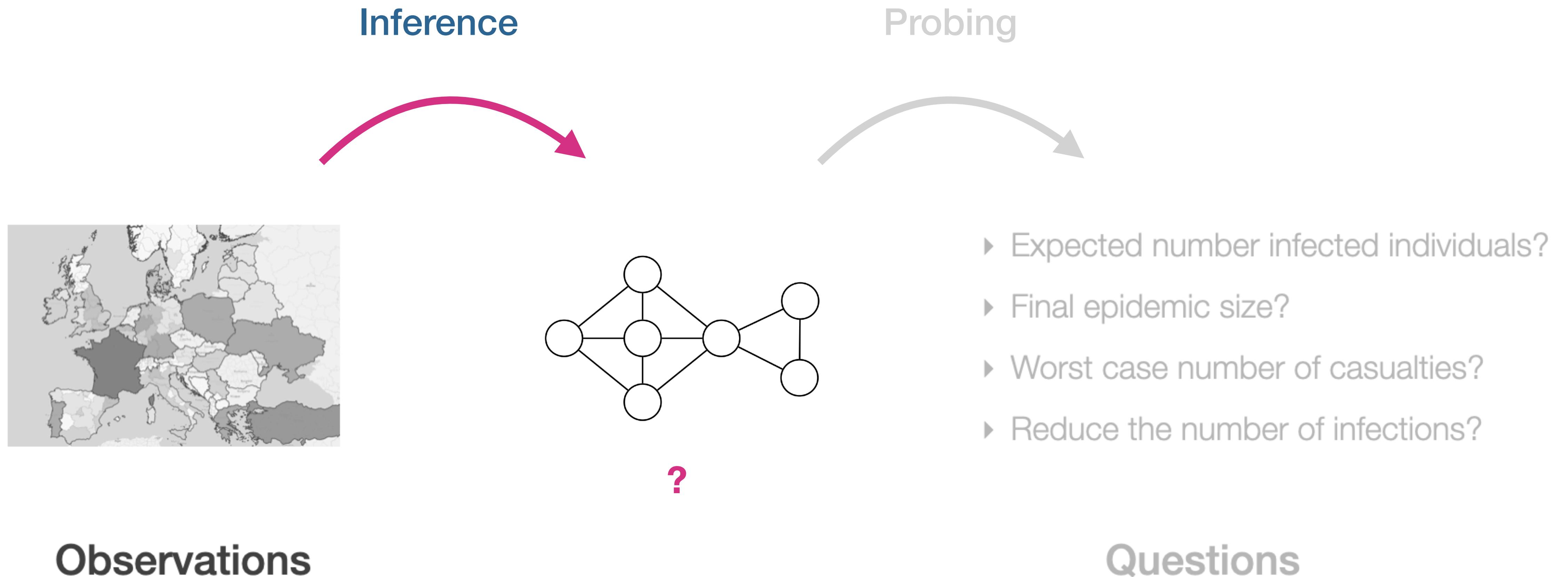


Agenda

METHODS

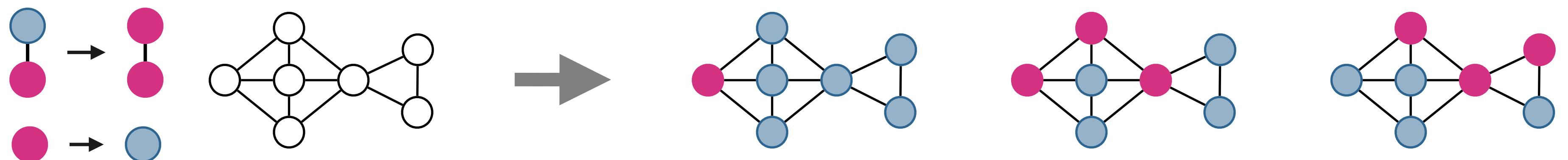
COVID
(NON-)MARKOVIAN SIMULATION
CONTROL
INFERENCE

Pipeline

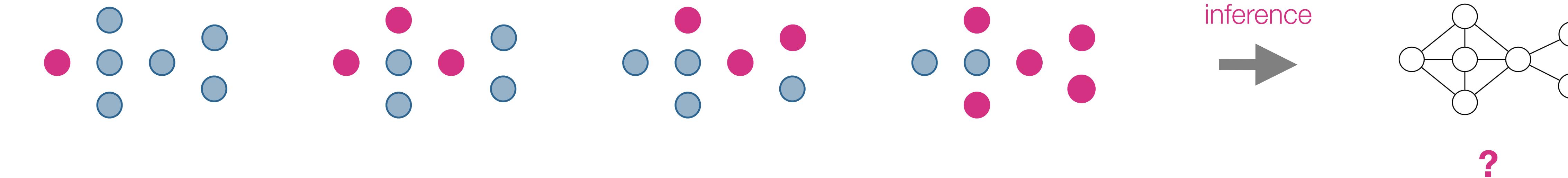


Problem Setting

Generation:

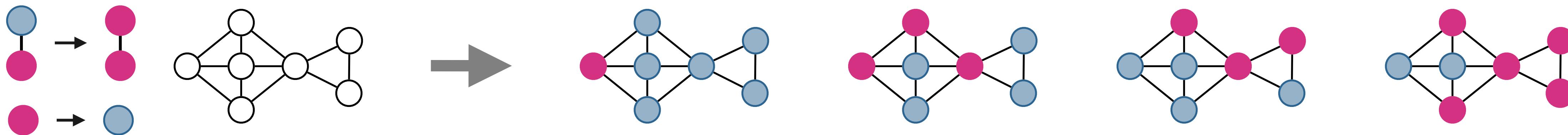


Our task:

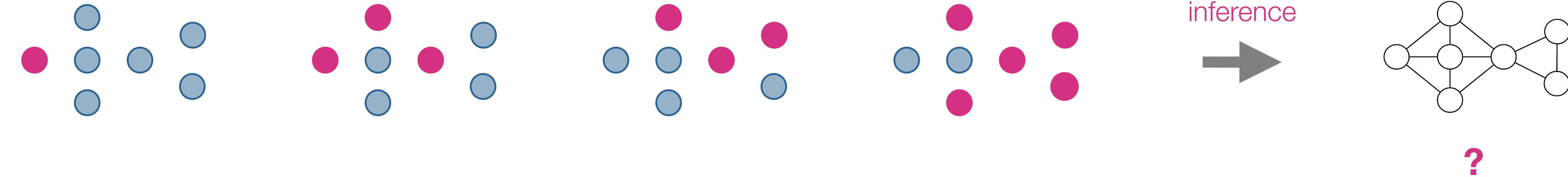


Problem Setting

Generation:



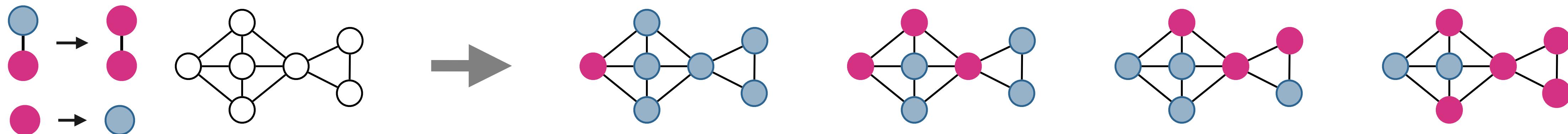
Our task:



- ▶ Node-wise observations
- ▶ **Typically:** from time-series data
- ▶ **Here:** from independent snapshots

Problem Setting

Generation:

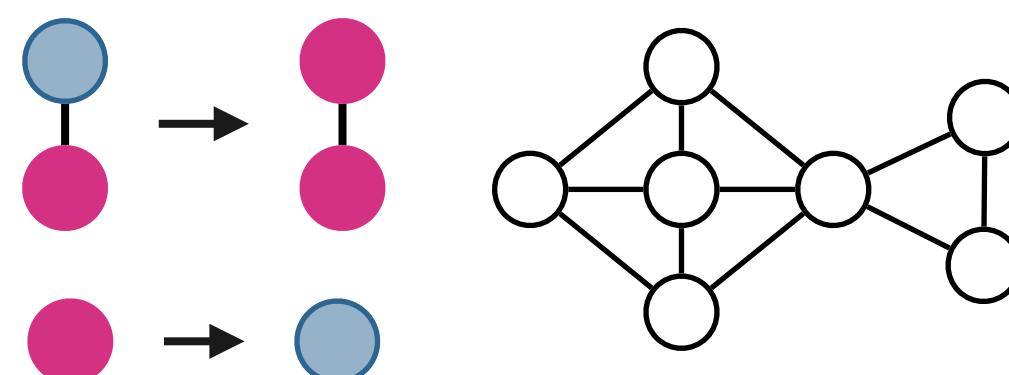


Our task:

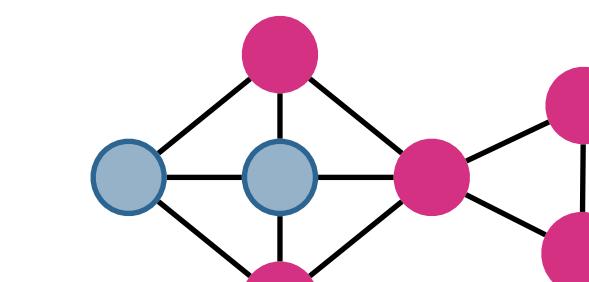
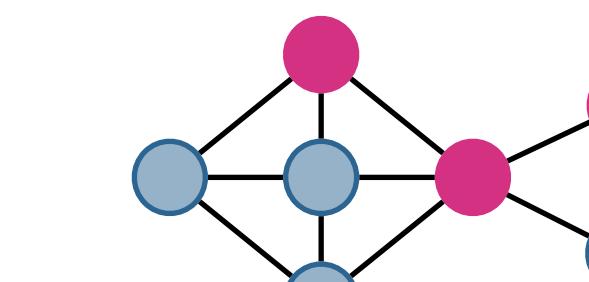
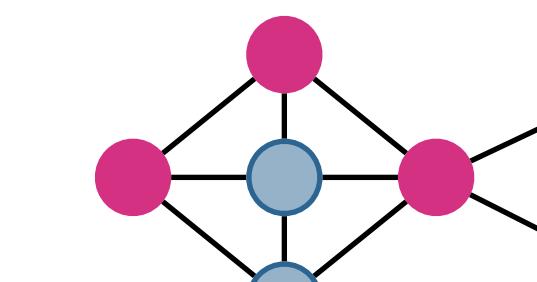
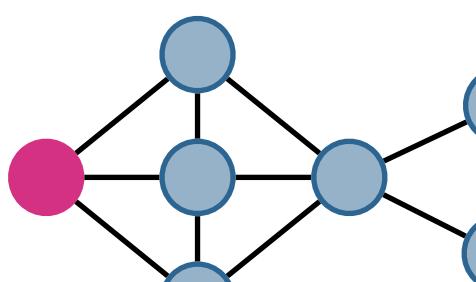


Problem Setting

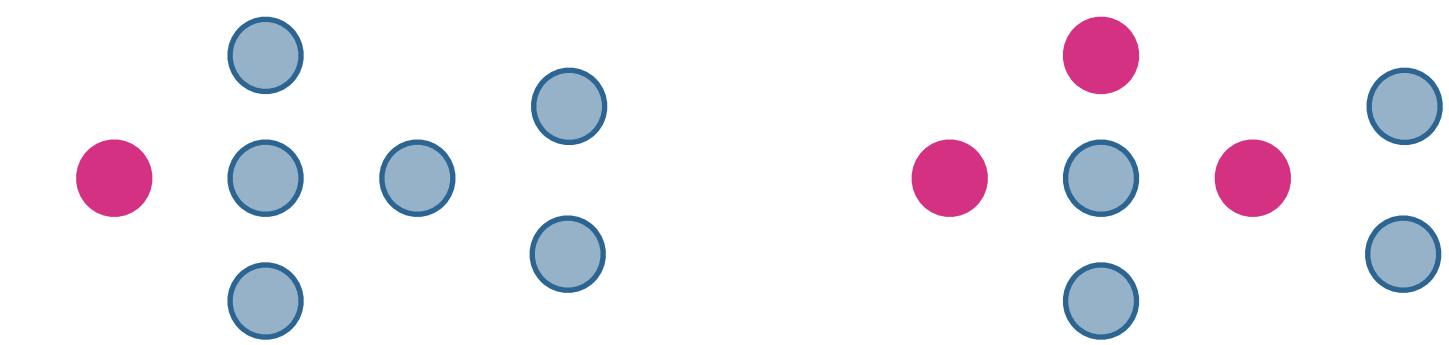
Generation:



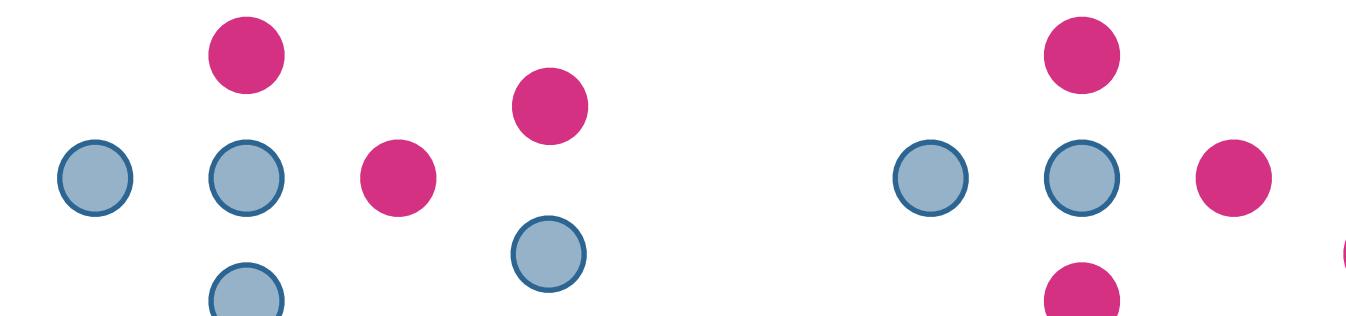
Unknown
dynamical
model.



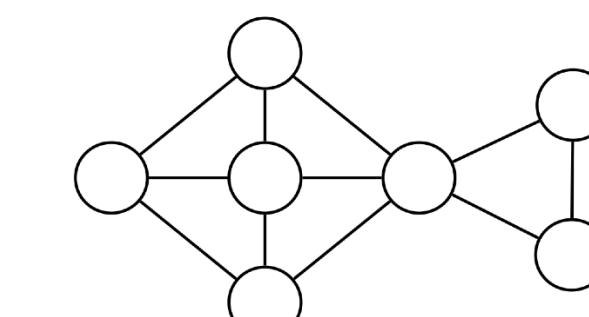
Our task:



Do not assume any
correlation.

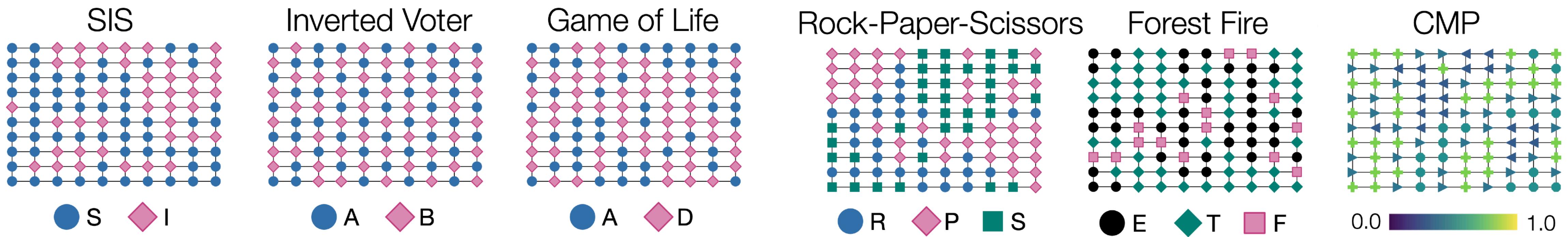


network
inference
→



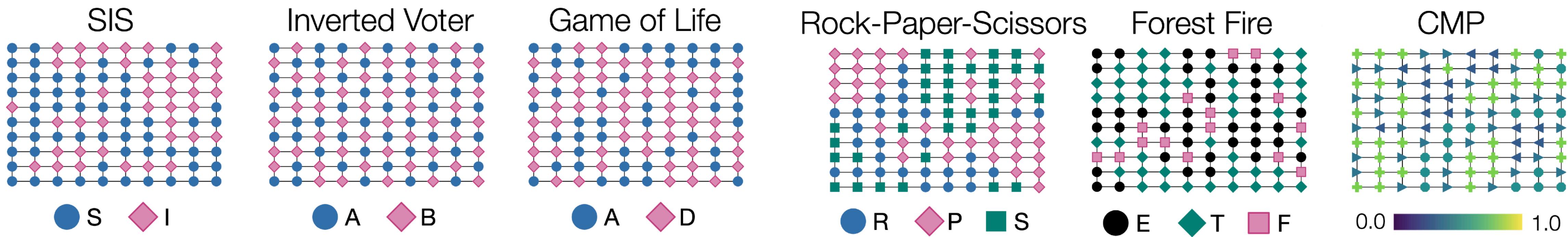
- ▶ Node-wise observations
- ▶ **Typically:** from time-series data
- ▶ **Here:** from independent snapshots

Snapshots



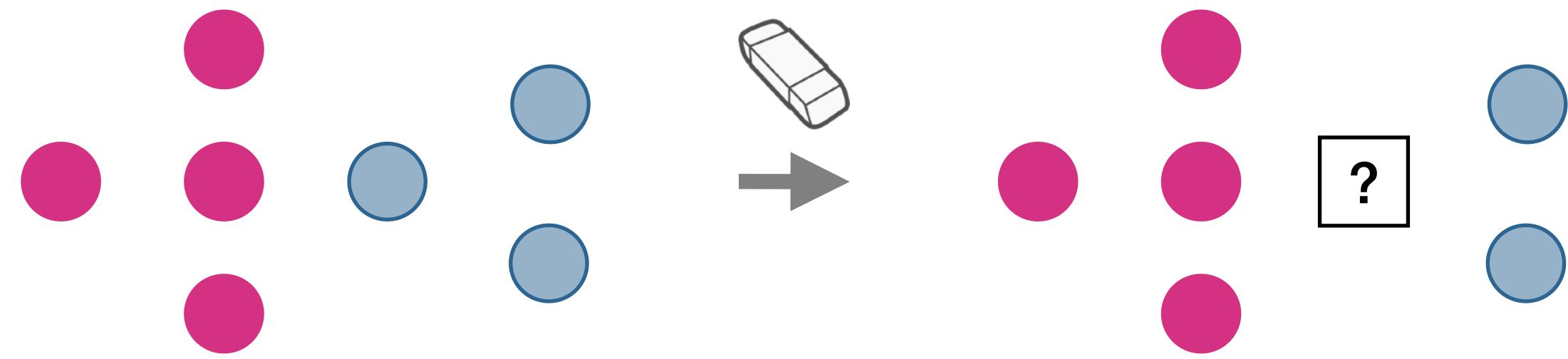
- ▶ **Given:** Set of independent snapshots
- ▶ **Unknown:** Interaction graph and dynamical laws
- ▶ **Task:** Find interaction graph “*best fitting*” the observed data
- ▶ **Assumption:** Patterns result from local interactions

Snapshots



- ▶ **Given:** Set of independent snapshots
- ▶ **Unknown:** Interaction graph and dynamical laws
- ▶ **Task:** Find interaction graph “*best fitting*” the observed data
- ▶ **Assumption:** Patterns result from local interactions

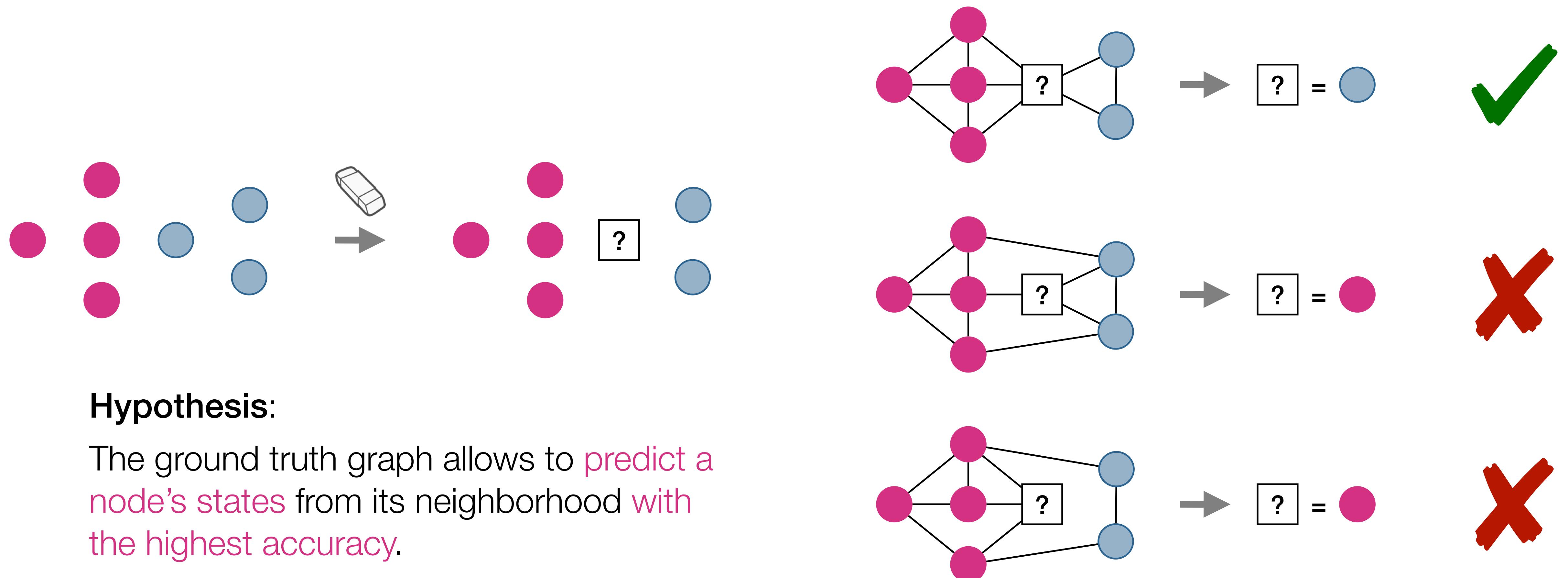
Idea



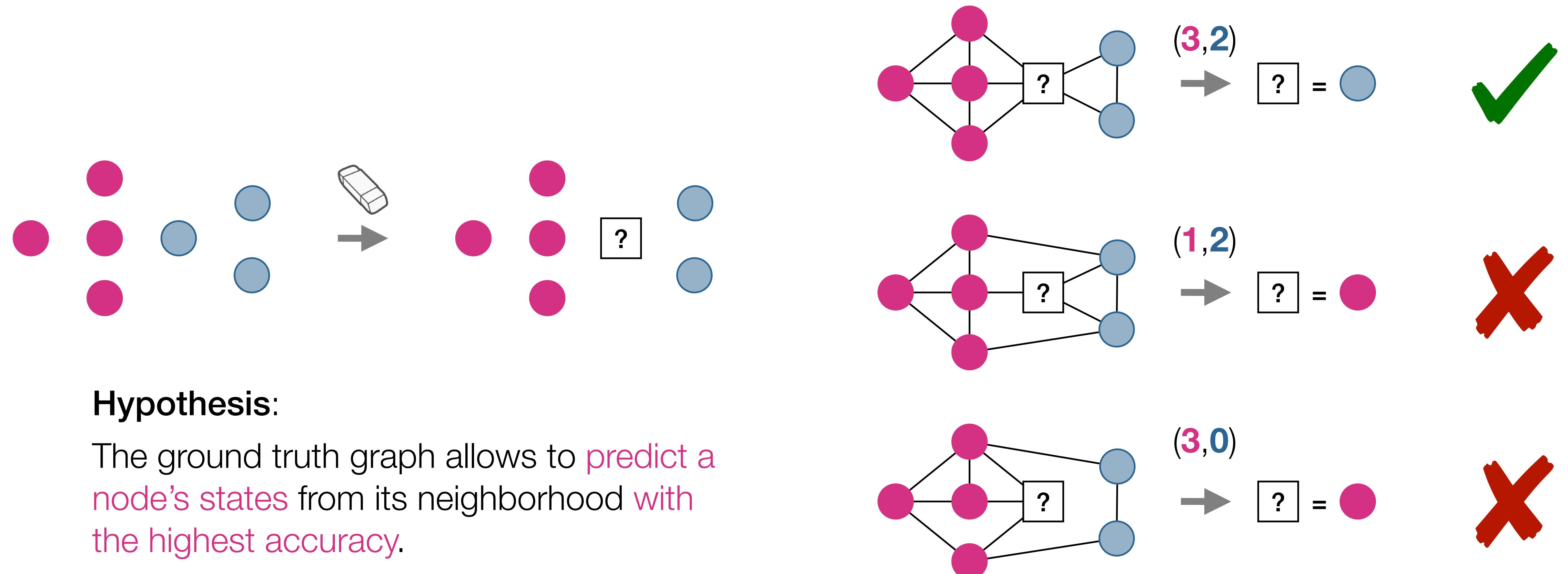
Hypothesis:

The ground truth graph allows to predict a node's states from its neighborhood with the highest accuracy.

Idea

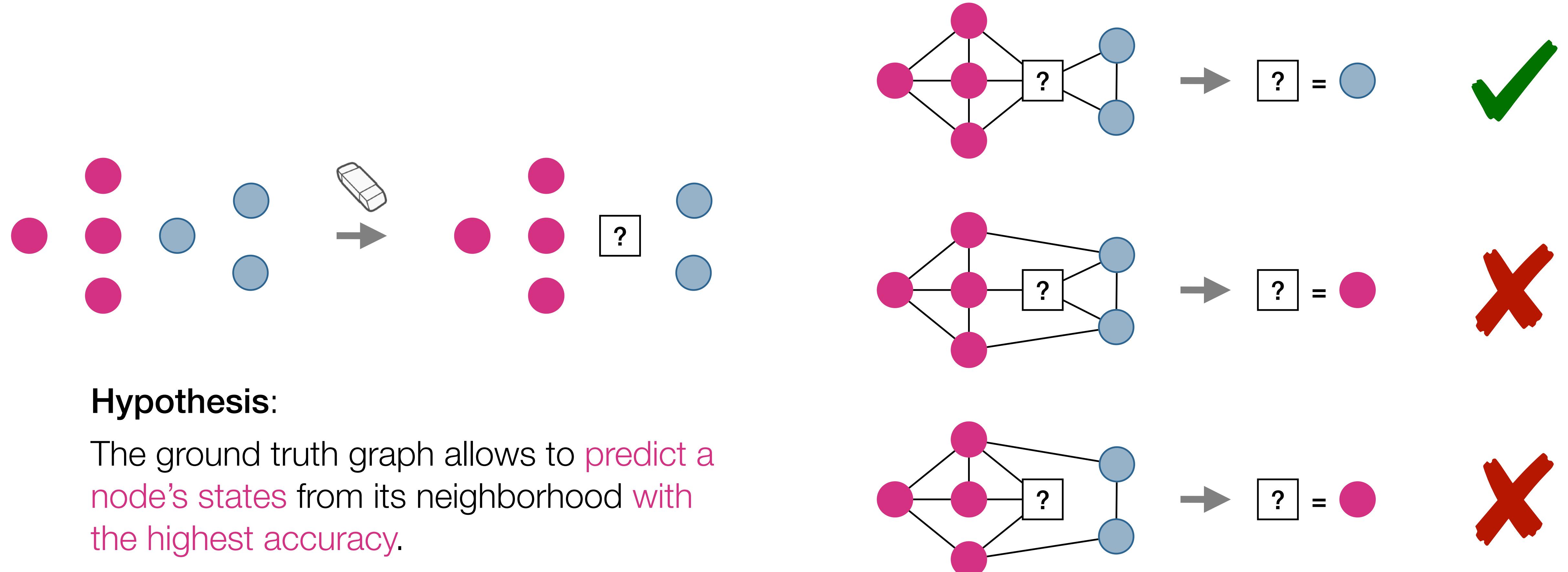


Idea

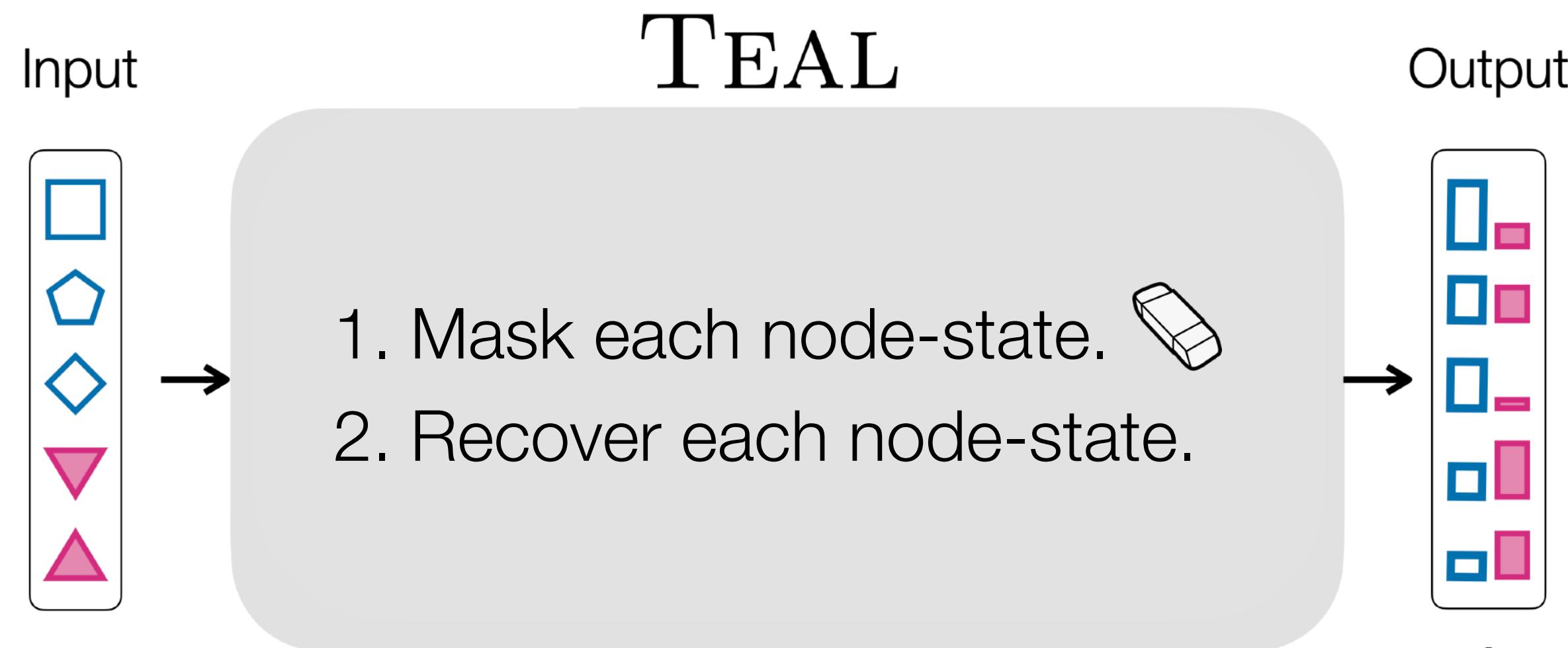


Idea

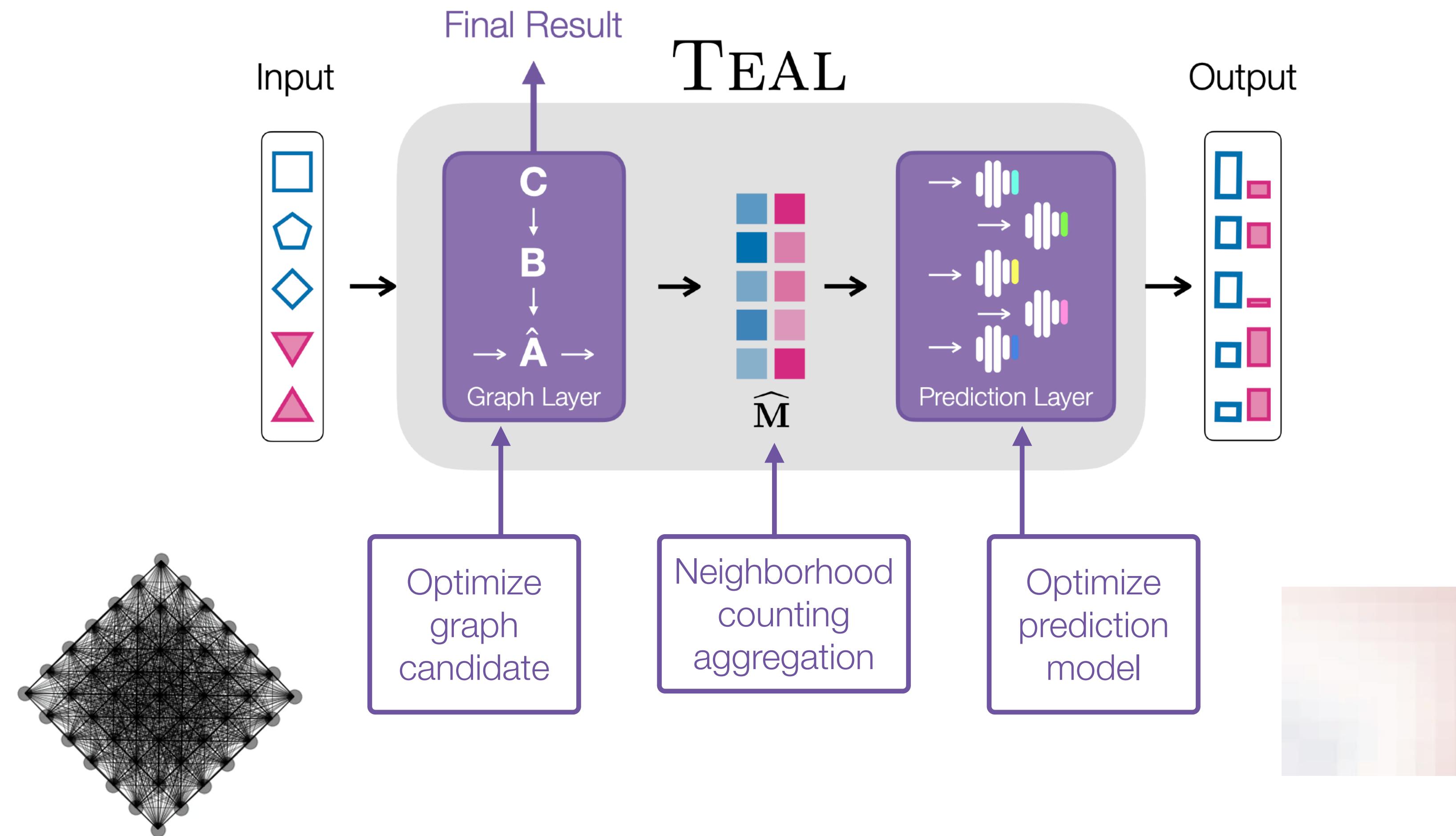
Problem: search space size



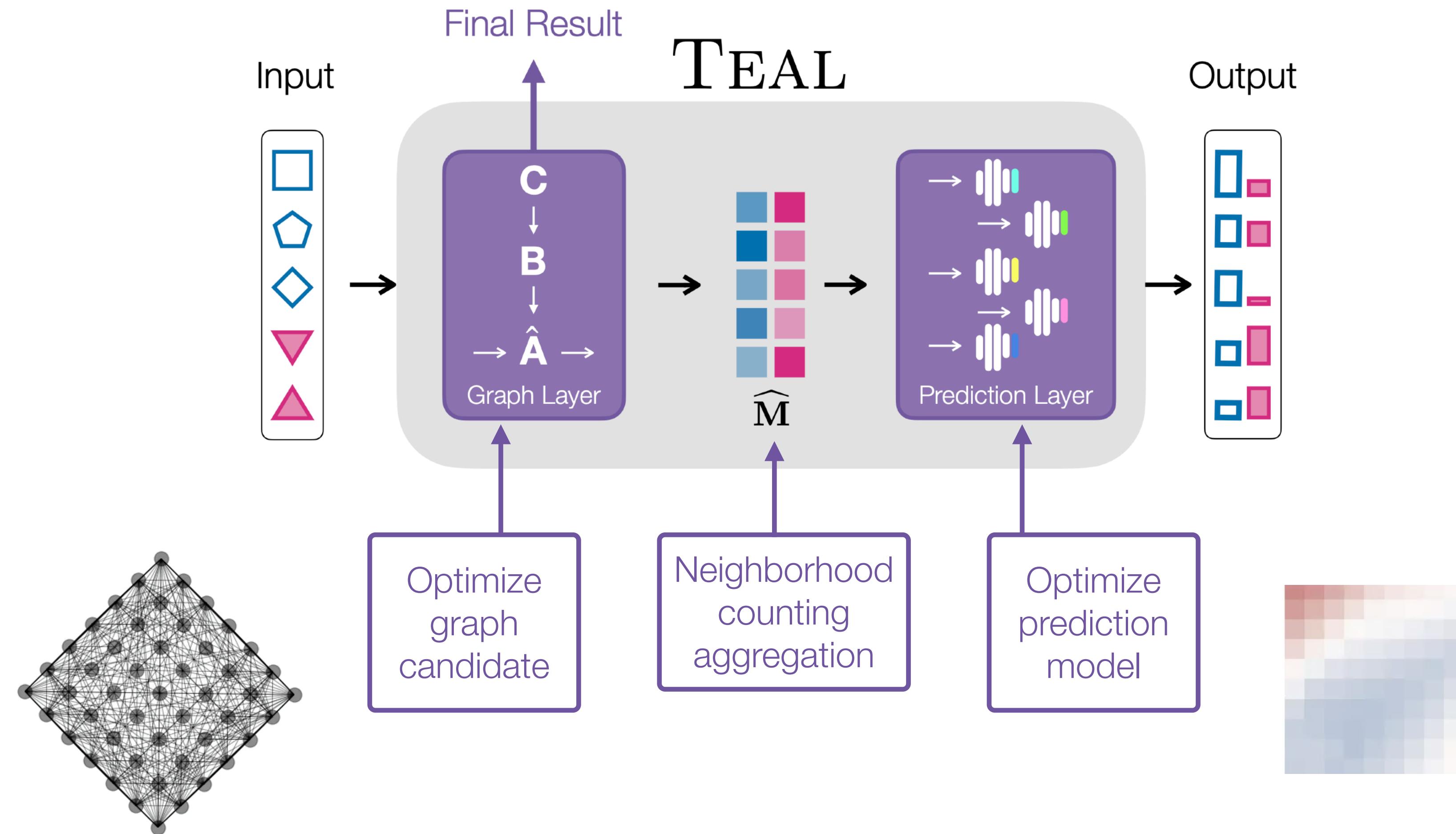
Our Method



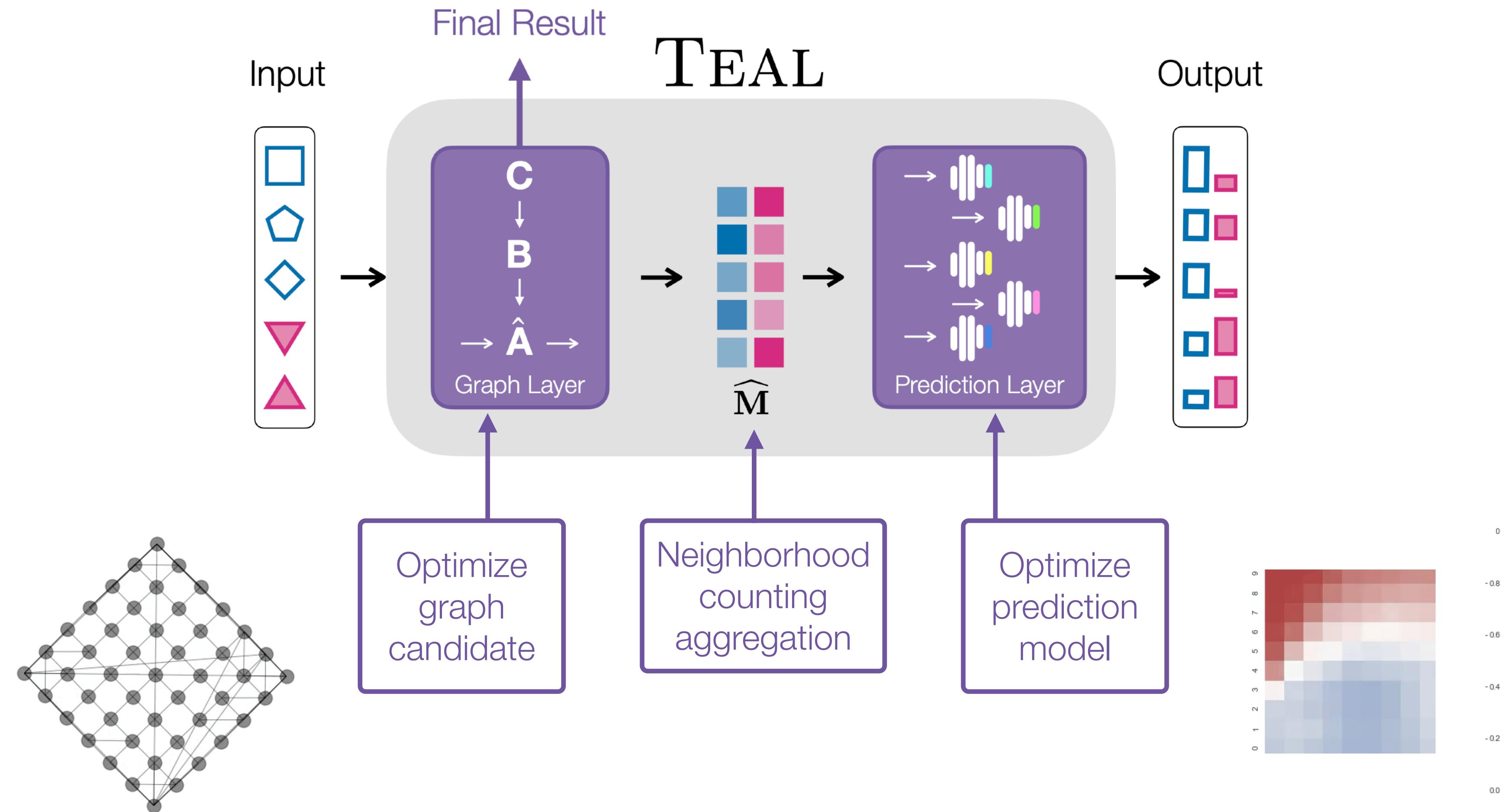
Our Method



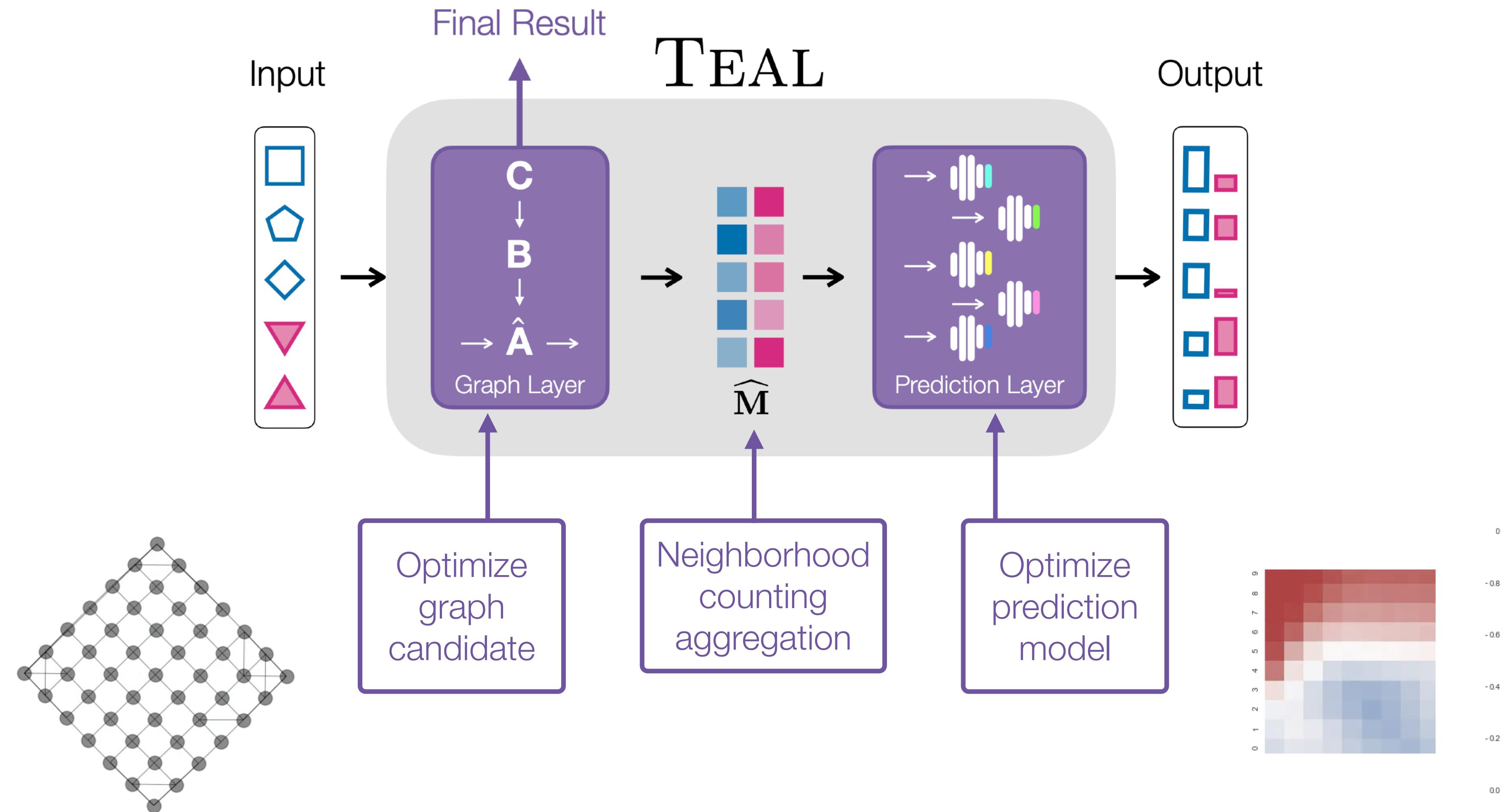
Our Method



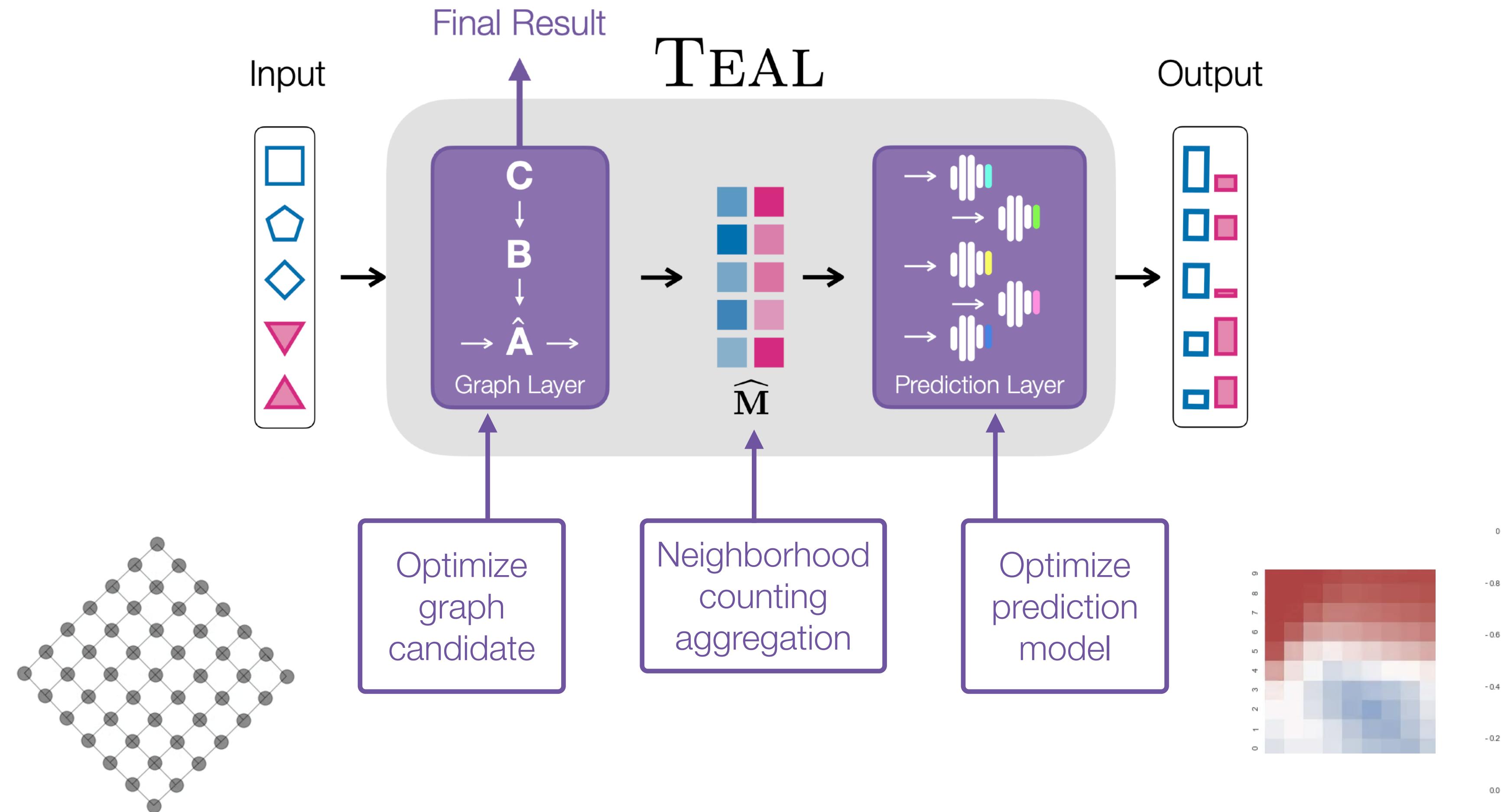
Our Method



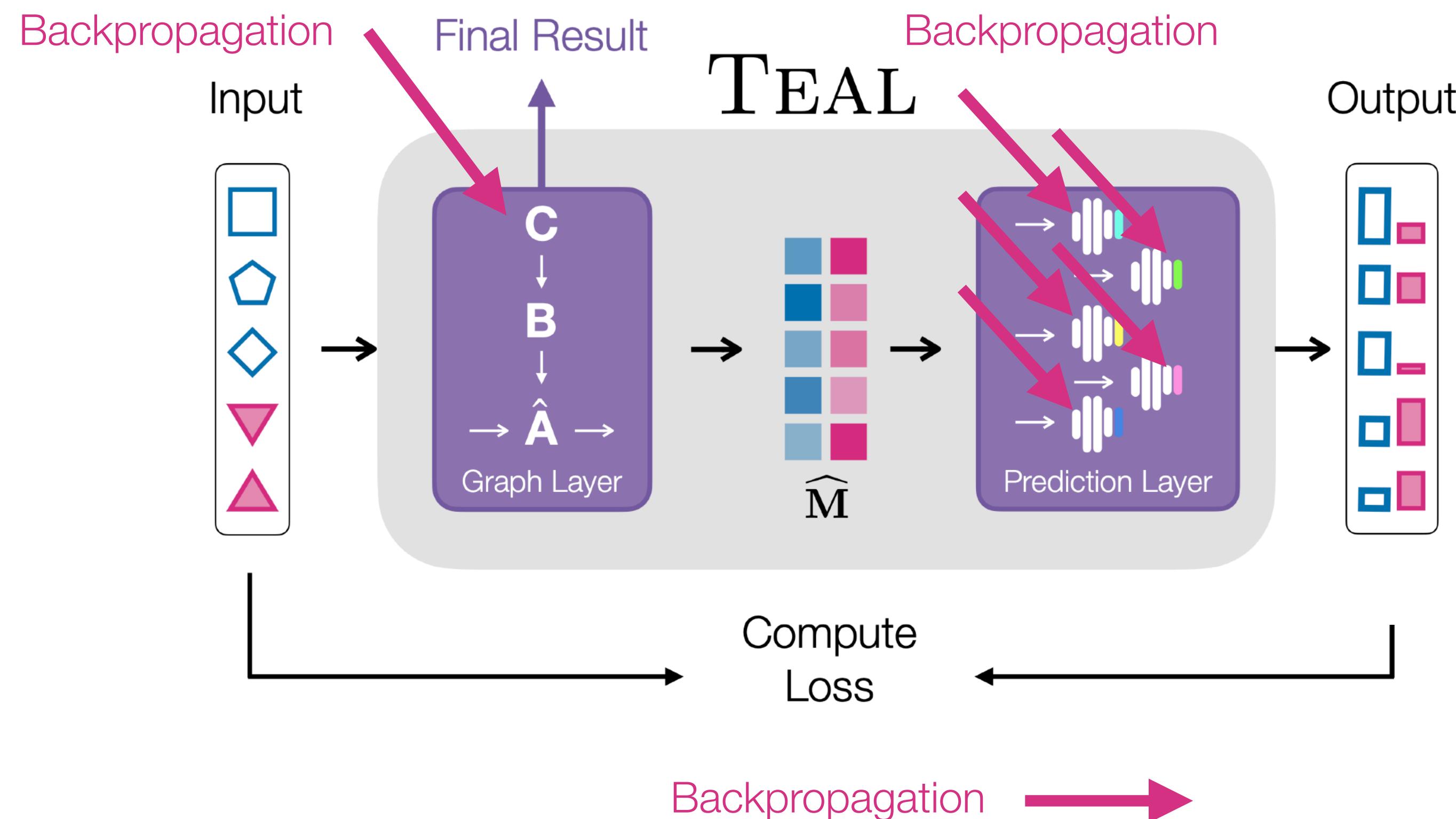
Our Method



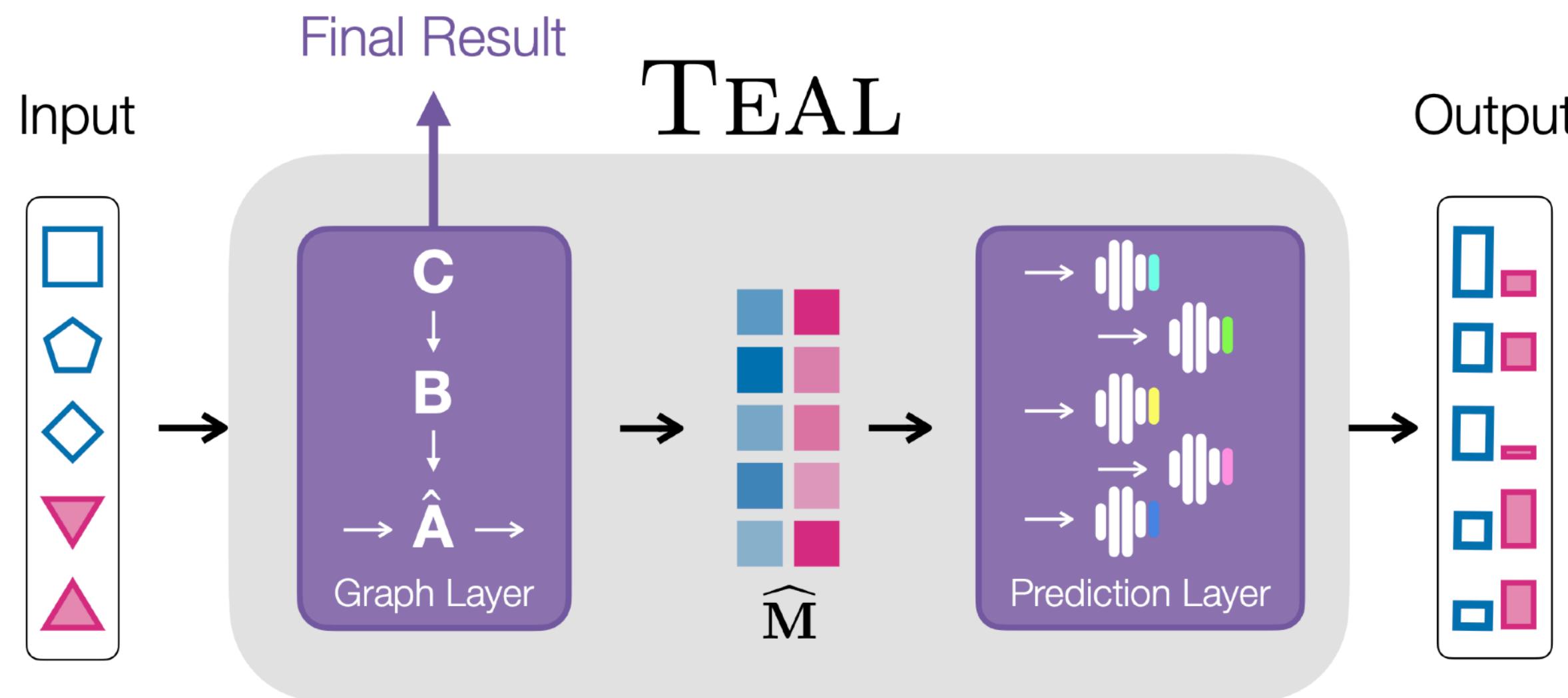
Our Method



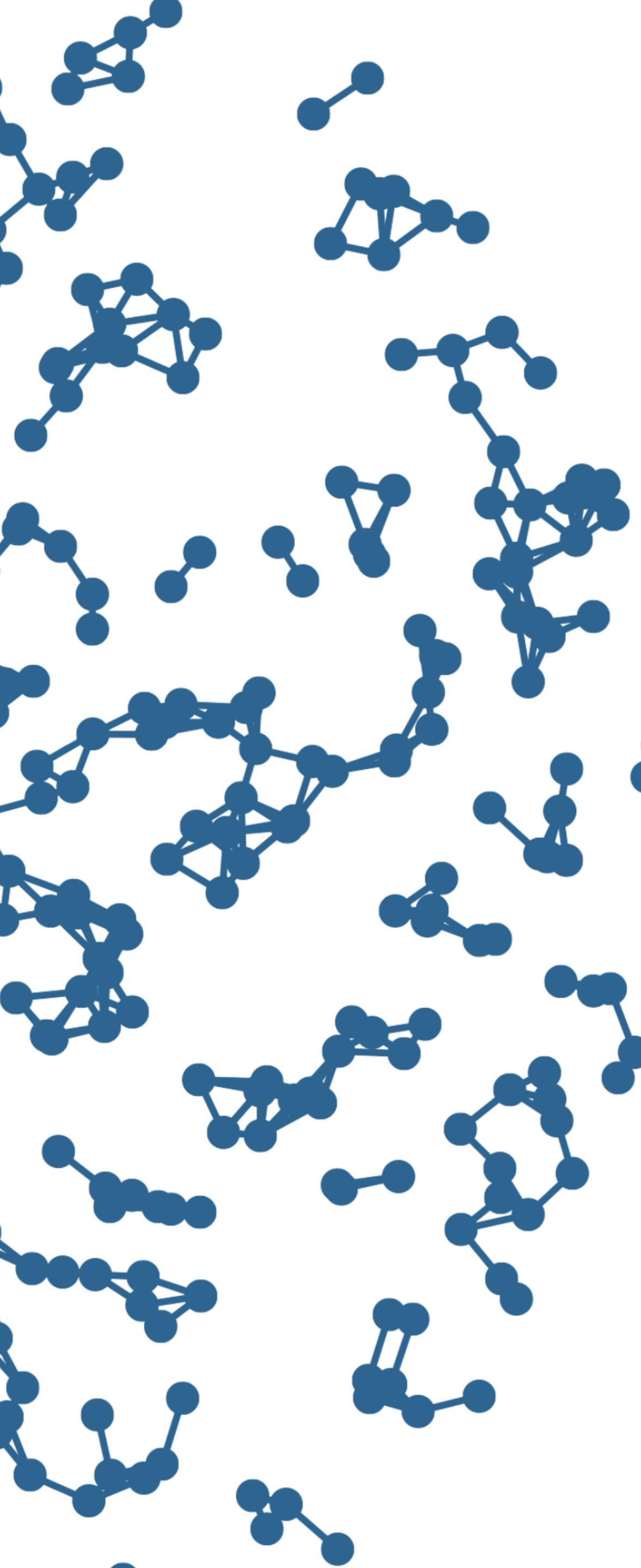
Our Method



Our Method



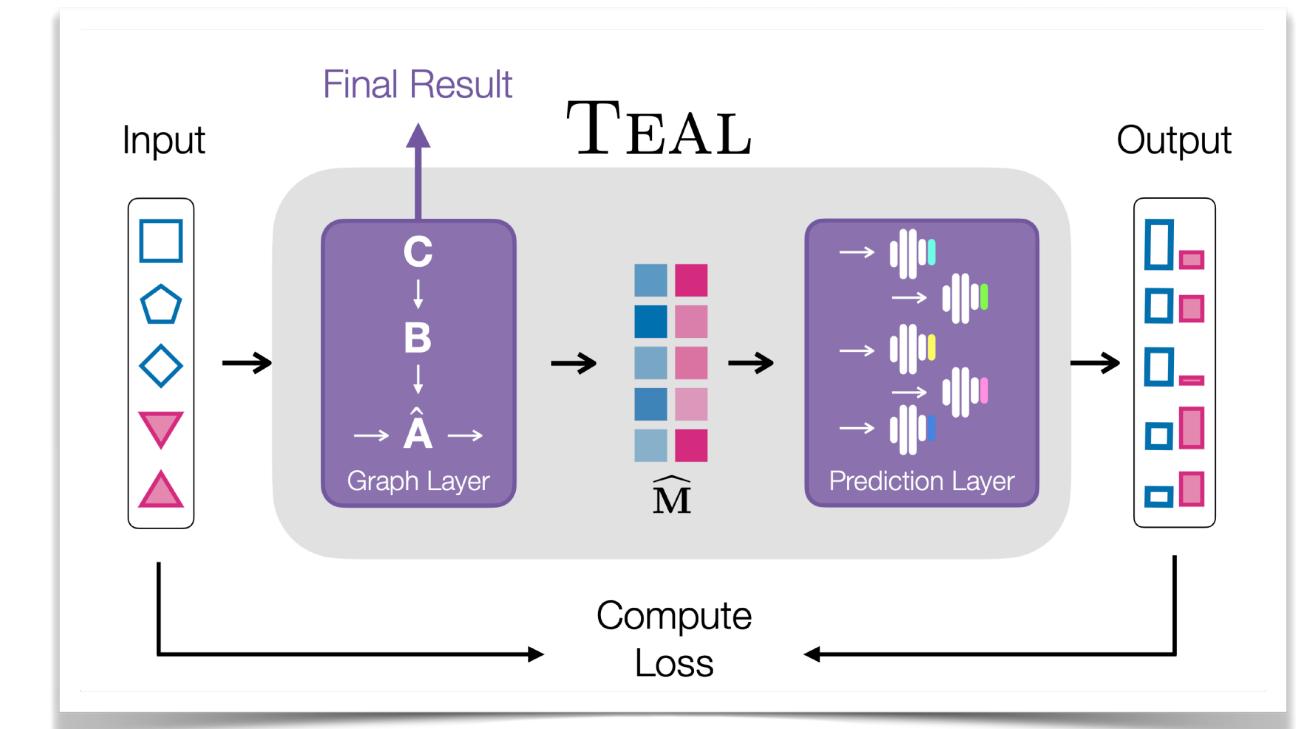
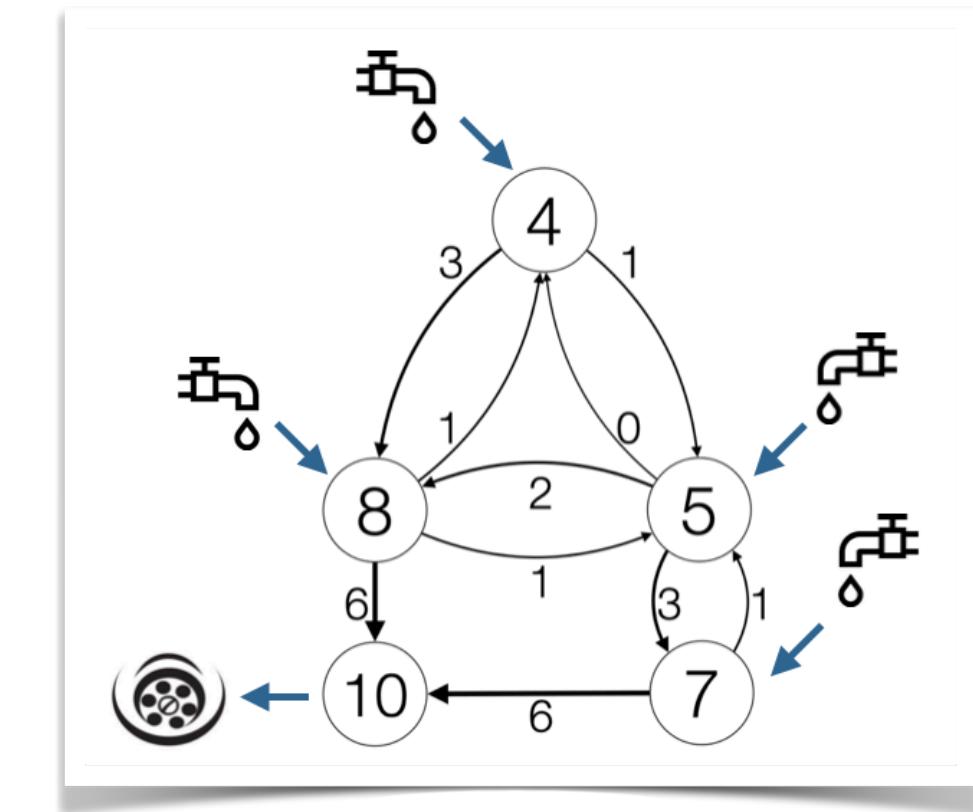
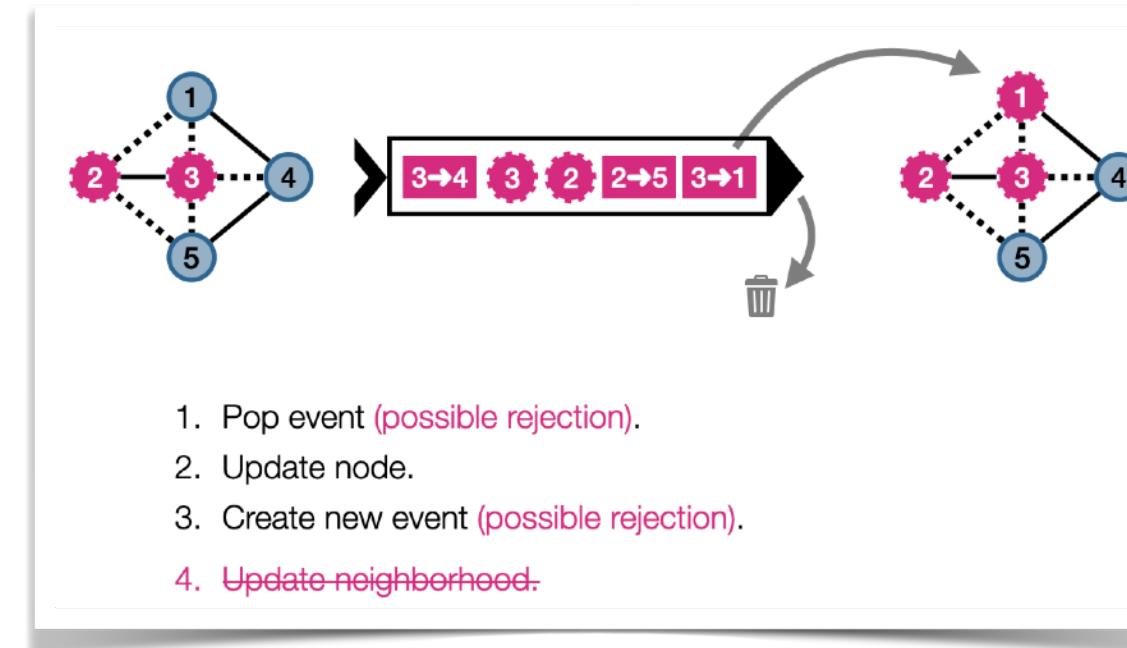
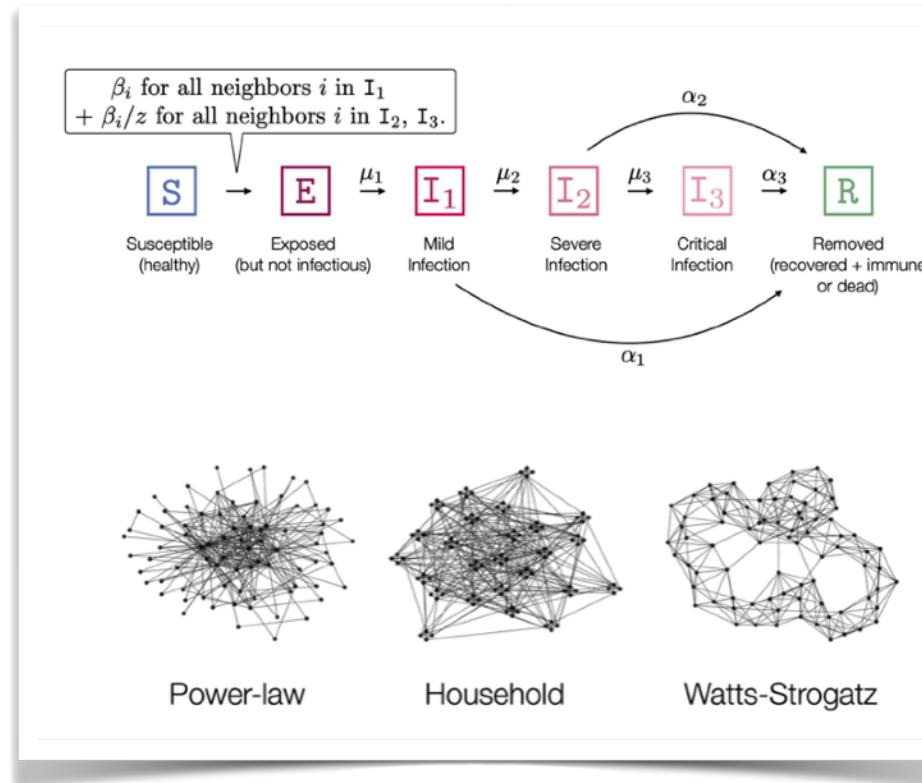
- ▶ **Masked reconstruction quantifies** which graph “best describes” observed data.
- ▶ **Tame vast search space** using SGD.
- ▶ **Model-free** and threshold-free network inference;
- ▶ **Outperforms** statistical and machine learning baselines.



Epilogue

CONCLUDING REMARKS

Lessons Learned



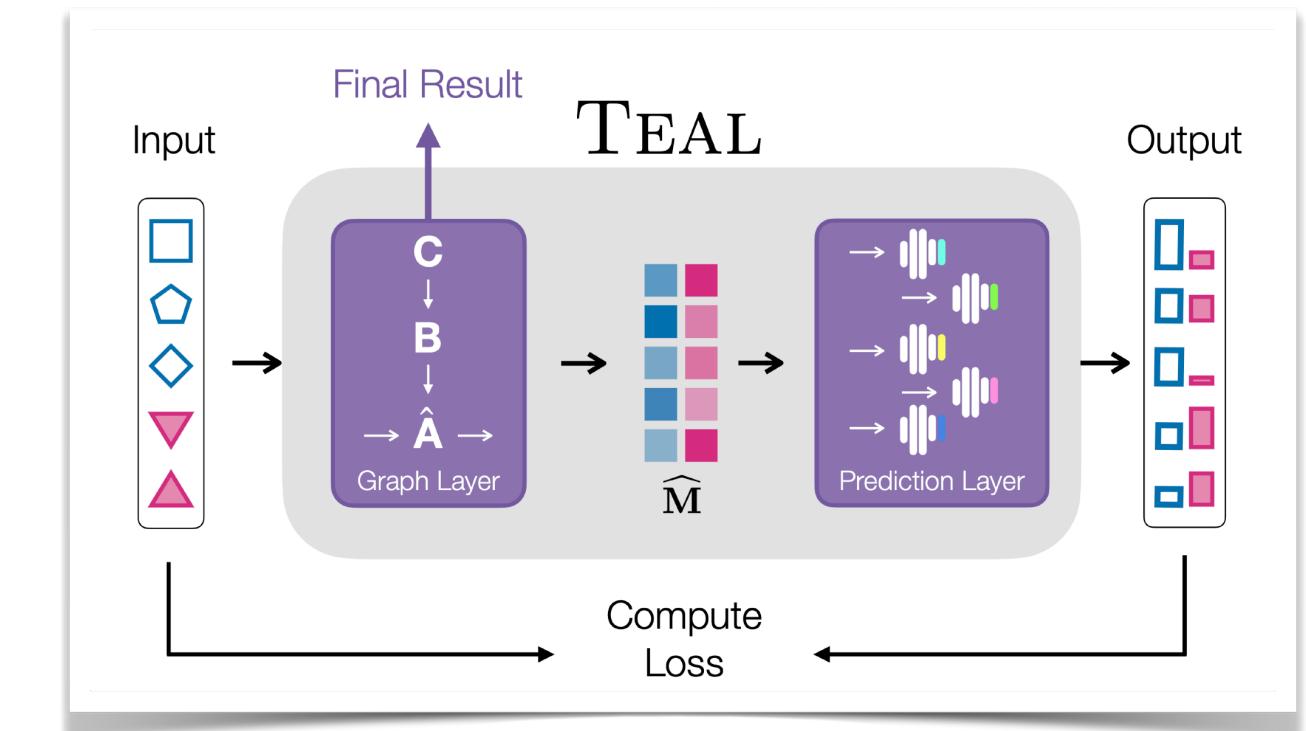
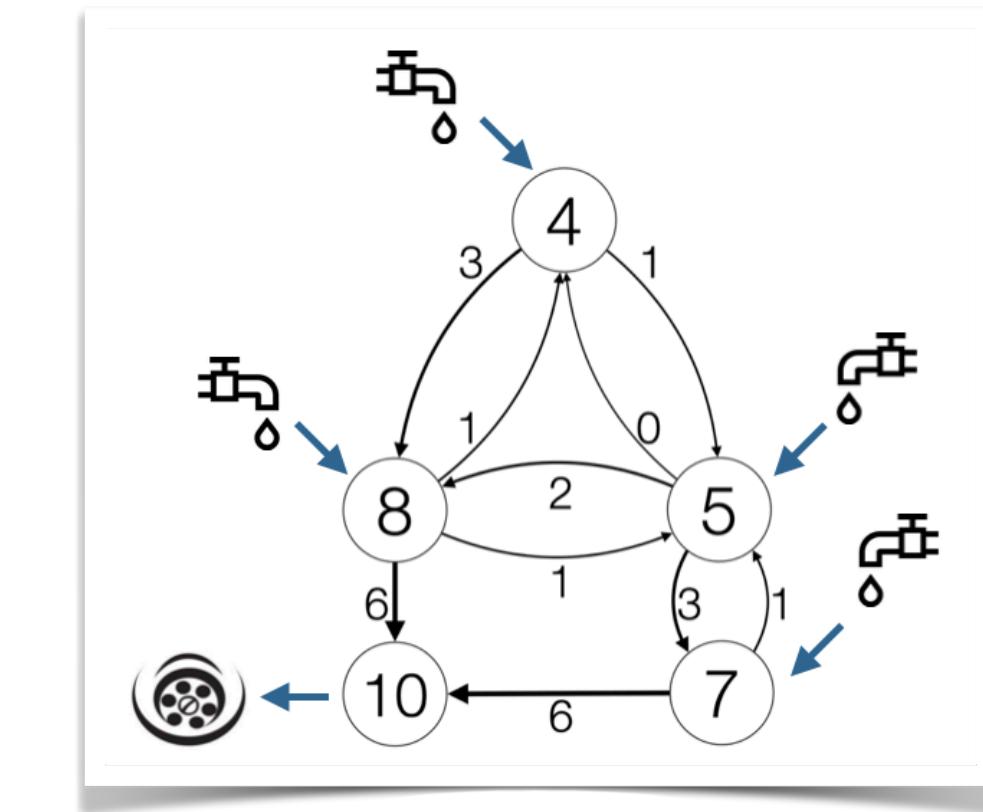
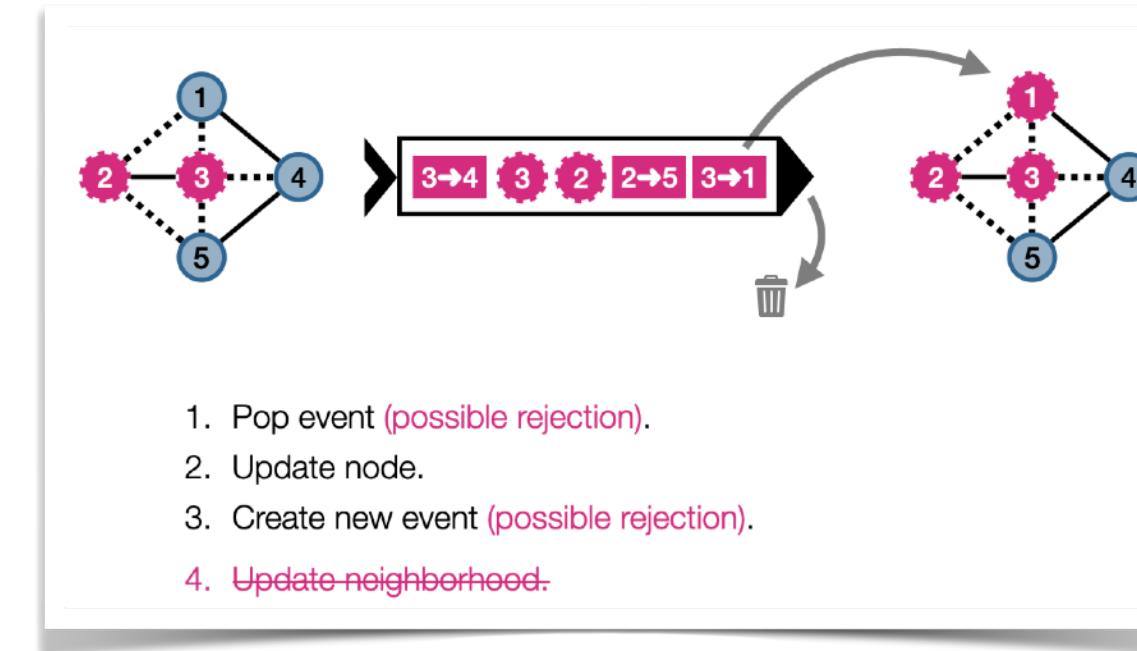
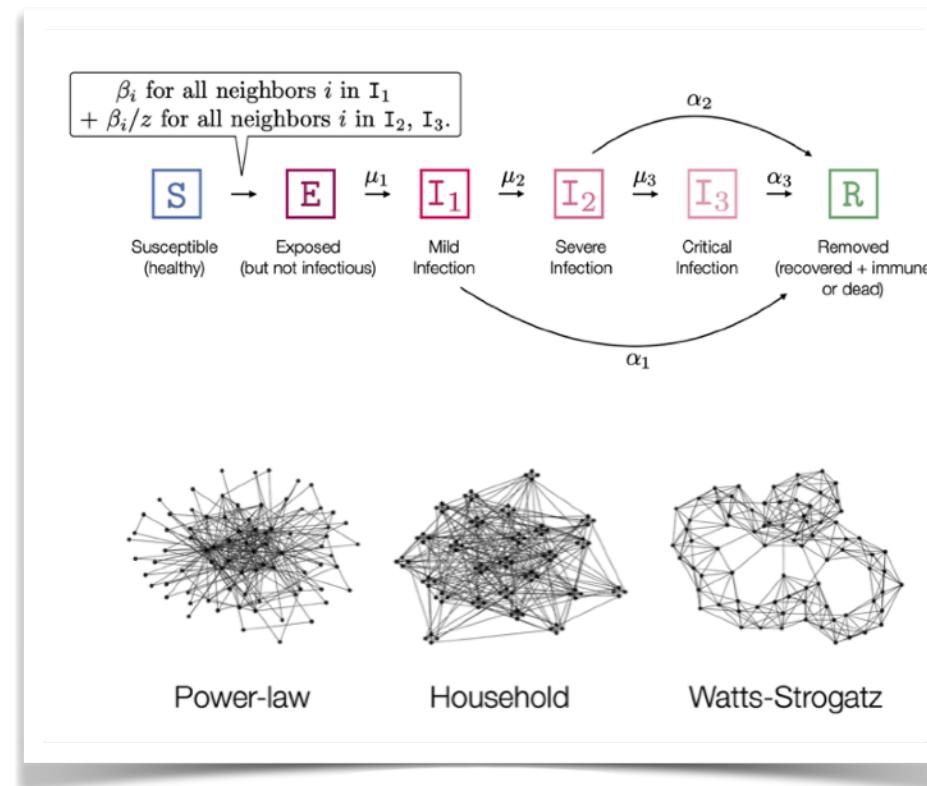
Networks are necessary.

Rejections create chances.

Flow uncovers importance.

Masking enables inference.

Lessons Learned



Networks are necessary.

Rejections create chances.

Flow uncovers importance.

Masking enables inference.

Future Challenges

- ▶ **Obvious extensions:** temporal + adaptive + multilayer + higher order networks
- ▶ **Summary statistics:** Something better than prevalence plots?
- ▶ **Theoretical challenges:** Mean-field approximation for non-Markov. models
- ▶ **Uniform view on emergence:** SIS + Ising model + spin glass + point processes + percolation theory