# Lumping the Approximate Master Equation for Multistate Processes on Complex Networks

**Gerrit Großmann**
Charalampos Kyriakopoulos
Luca Bortolussi
Verena Wolf

# Lumping the Approximate Master Equation for Multistate **Processes on Complex Networks**

**Gerrit Großmann**
Charalampos Kyriakopoulos
Luca Bortolussi
Verena Wolf

# Complex Networks

Networks are everywhere

- Friendship networks
- Online social networks
- Telecommunication networks
- Infrastructure networks
- …

# Motivation

Understand spreading phenomenon of

- Infectious diseases
- Computer viruses
- Rumours/opinions/emotions
- Blackouts
- …

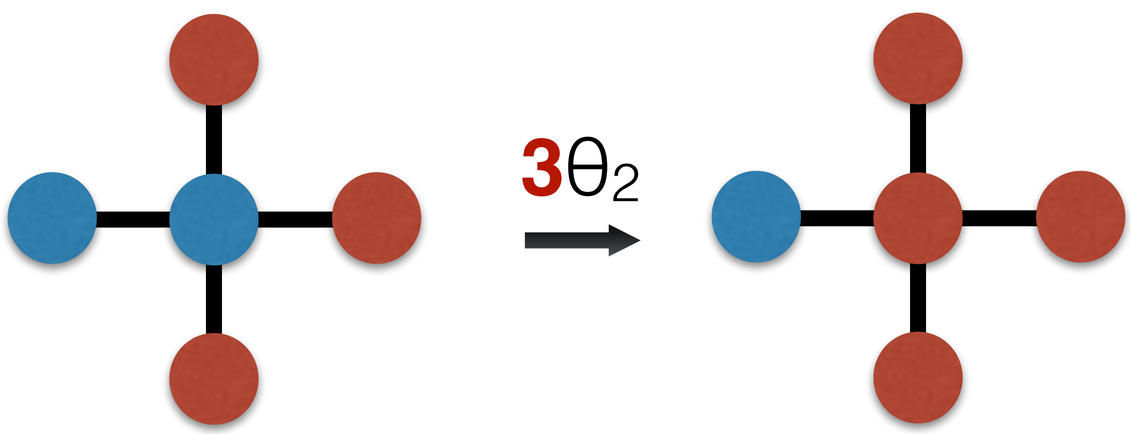
# Spreading Process

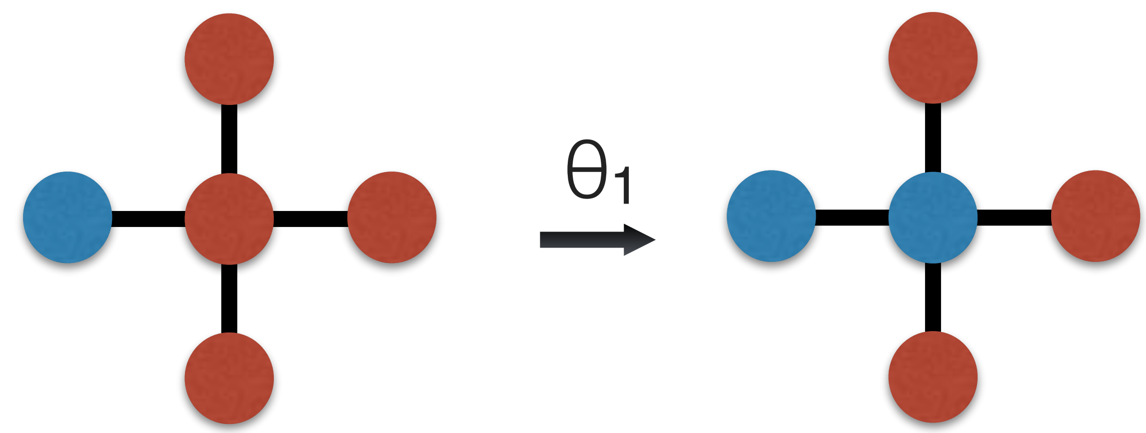
- Fixed graph topology
- Continuous time dynamics
- Nodes occupy node-states
- Node-states change randomly w.r.t. rules

Classical example: SIS model

- 2 node-states (infected, susceptible)
- 2 rules (infection, recovery)

# SIS Model



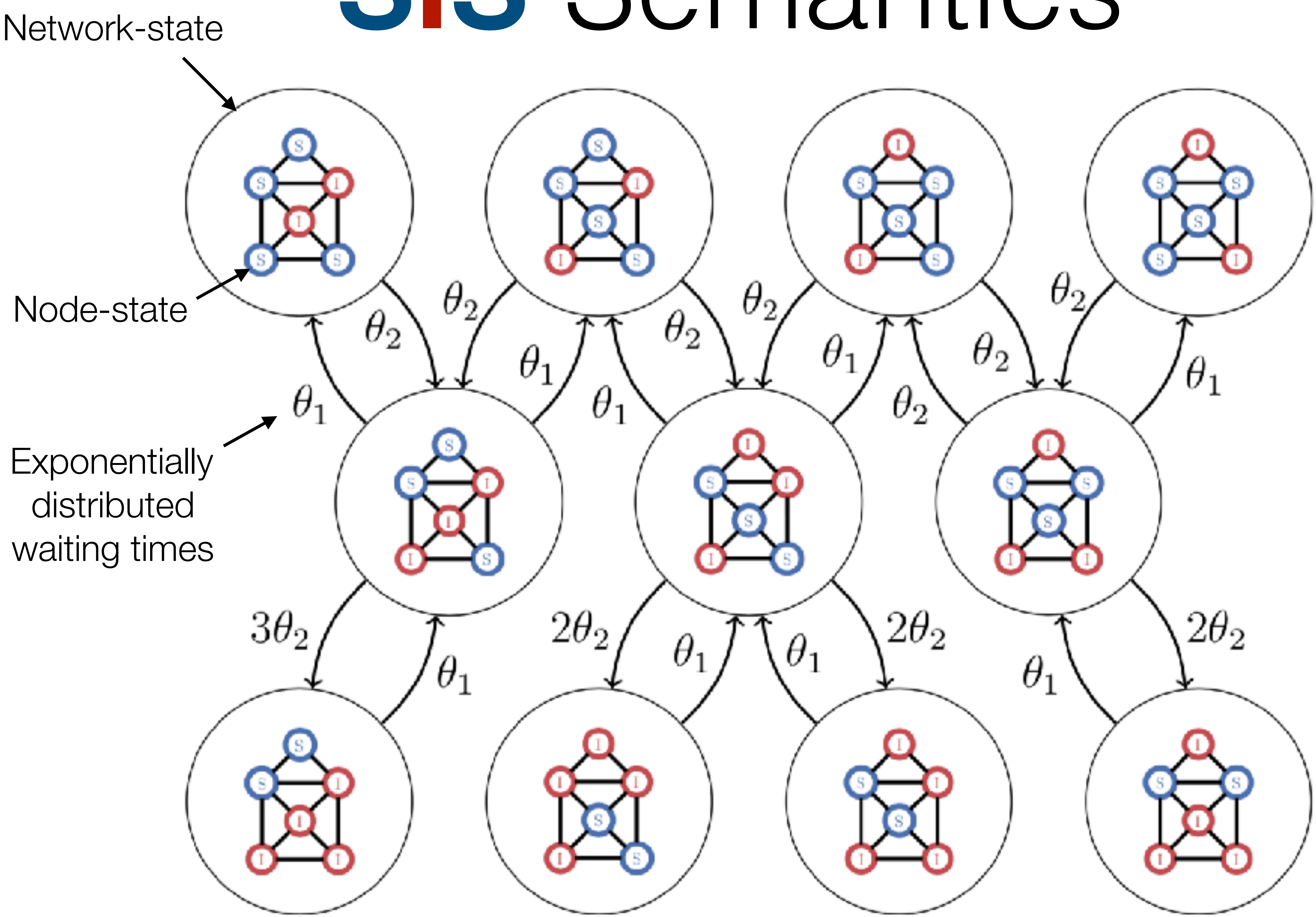




Infection (edge-based)

Recovery (node-based)

- 2 node-states (**i**nfected, **s**usceptible)
- 2 rules (infection, recovery)

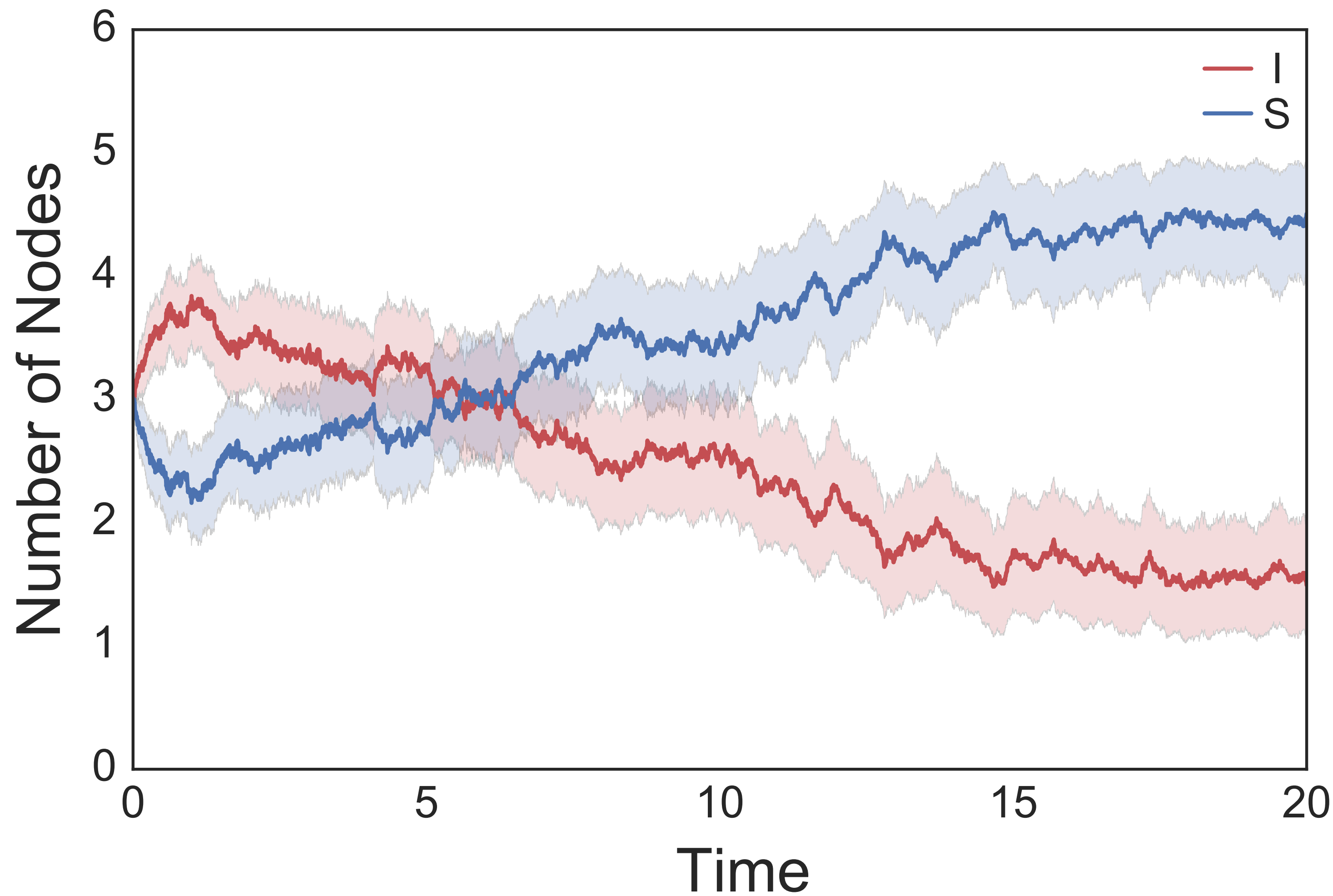
# SIS Semantics



# Fraction of Infected Nodes

Measure of interest:

- Fractions of nodes in each node-state
- Can be estimated with Monte-Carlo simulation (expensive)

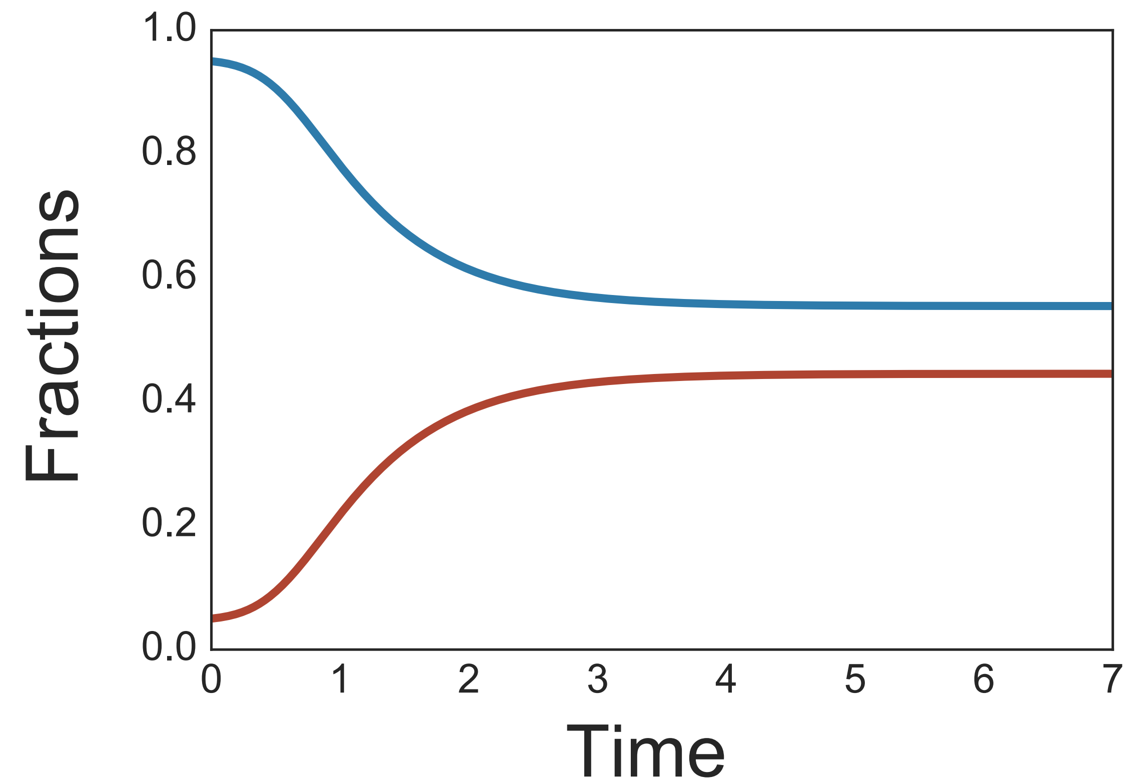


# Lumping the **Approximate Master Equation** for Multistate Processes on Complex Networks

*High-Accuracy Approximation of Binary-State Dynamics on Networks* (James P. Gleeson, 2011)

# Approximate Master Equation (AME)

- ODEs to approximate expected fractions over time
- One ODE for each node-state and possible neighbourhood

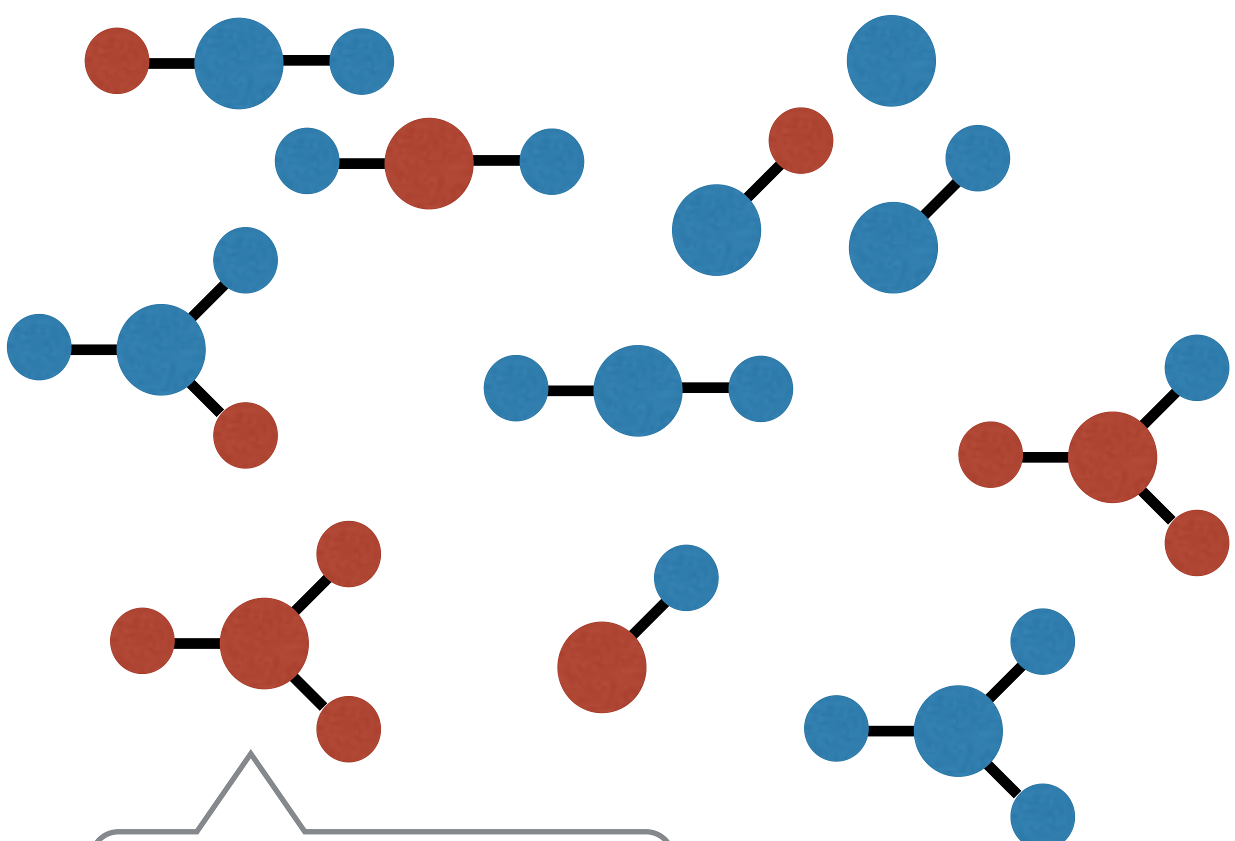



*High-Accuracy Approximation of Binary-State Dynamics on Networks* (James P. Gleeson, 2011)

# AME

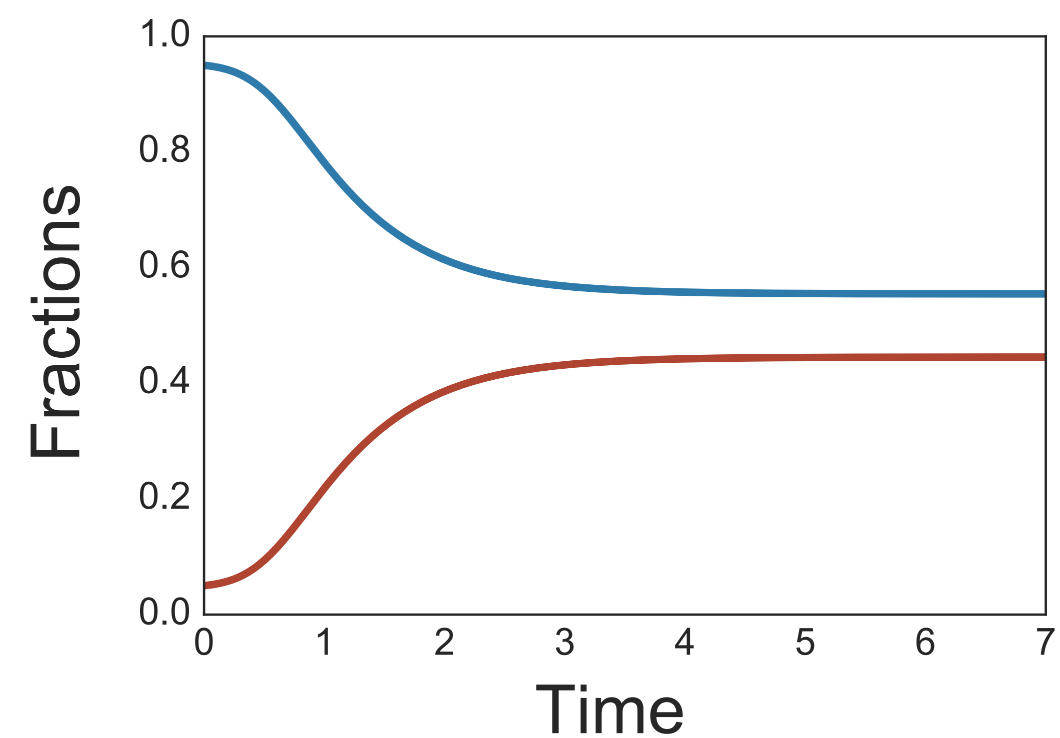
## one ODE for each node-state and neighbourhood

Fraction of susceptible nodes with 1 susceptible and 1 infected neighbour

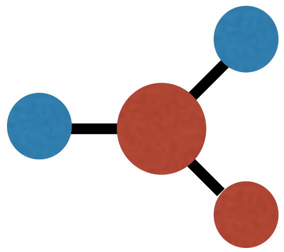


Fraction of infected nodes with 3 infected neighbours

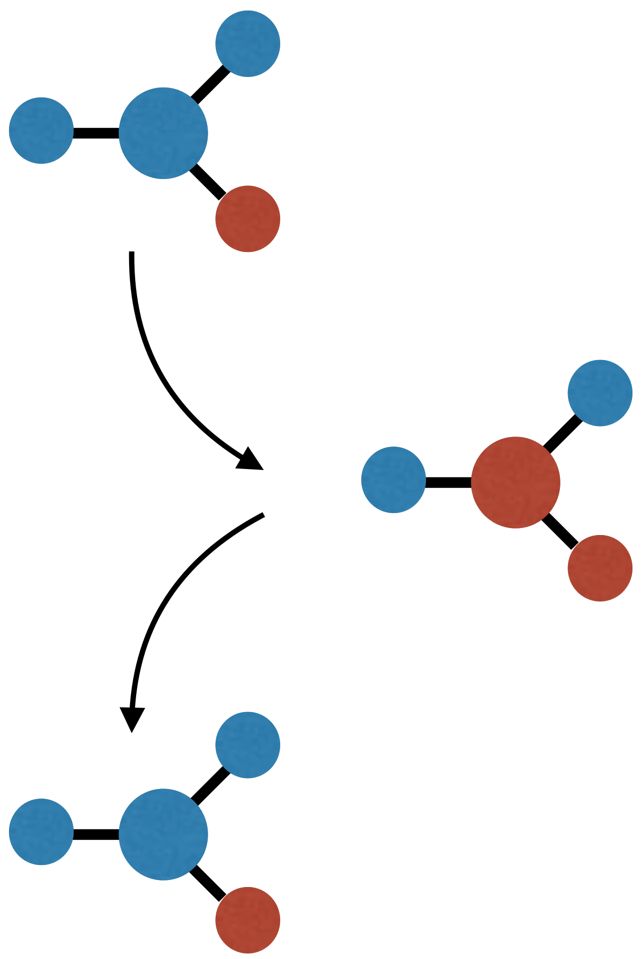
Numerical integration



# AME

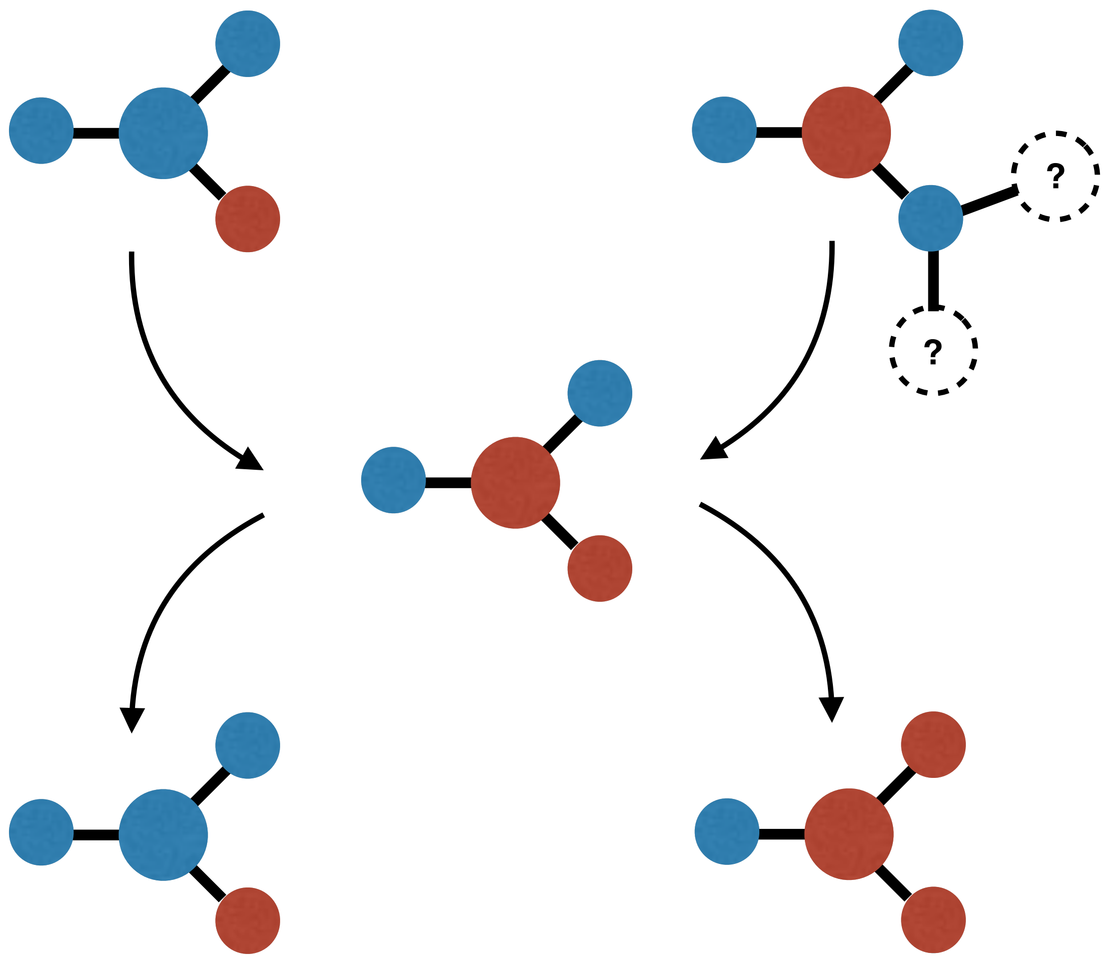


# AME



node
changes

# AME



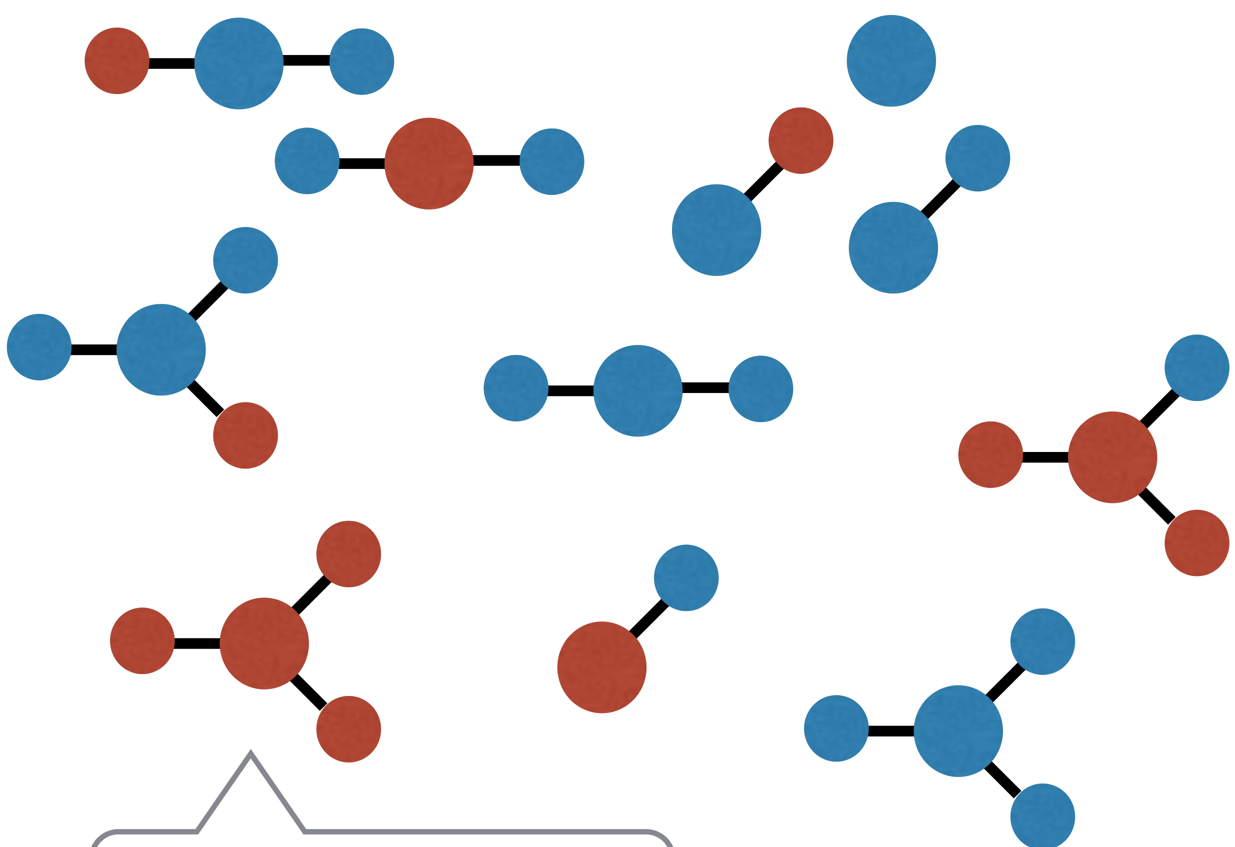

node changes

neighbourhood changes

# AME

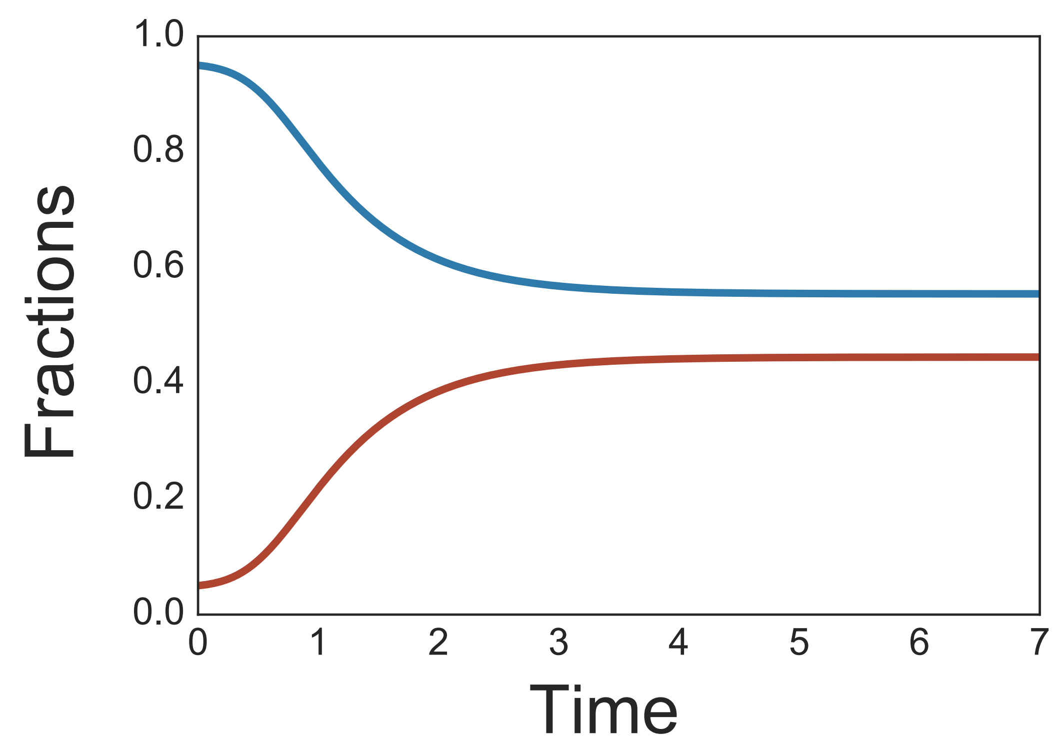
## one ODE for each node-state and neighbourhood

Fraction of susceptible nodes with 1 susceptible and 1 infected neighbour

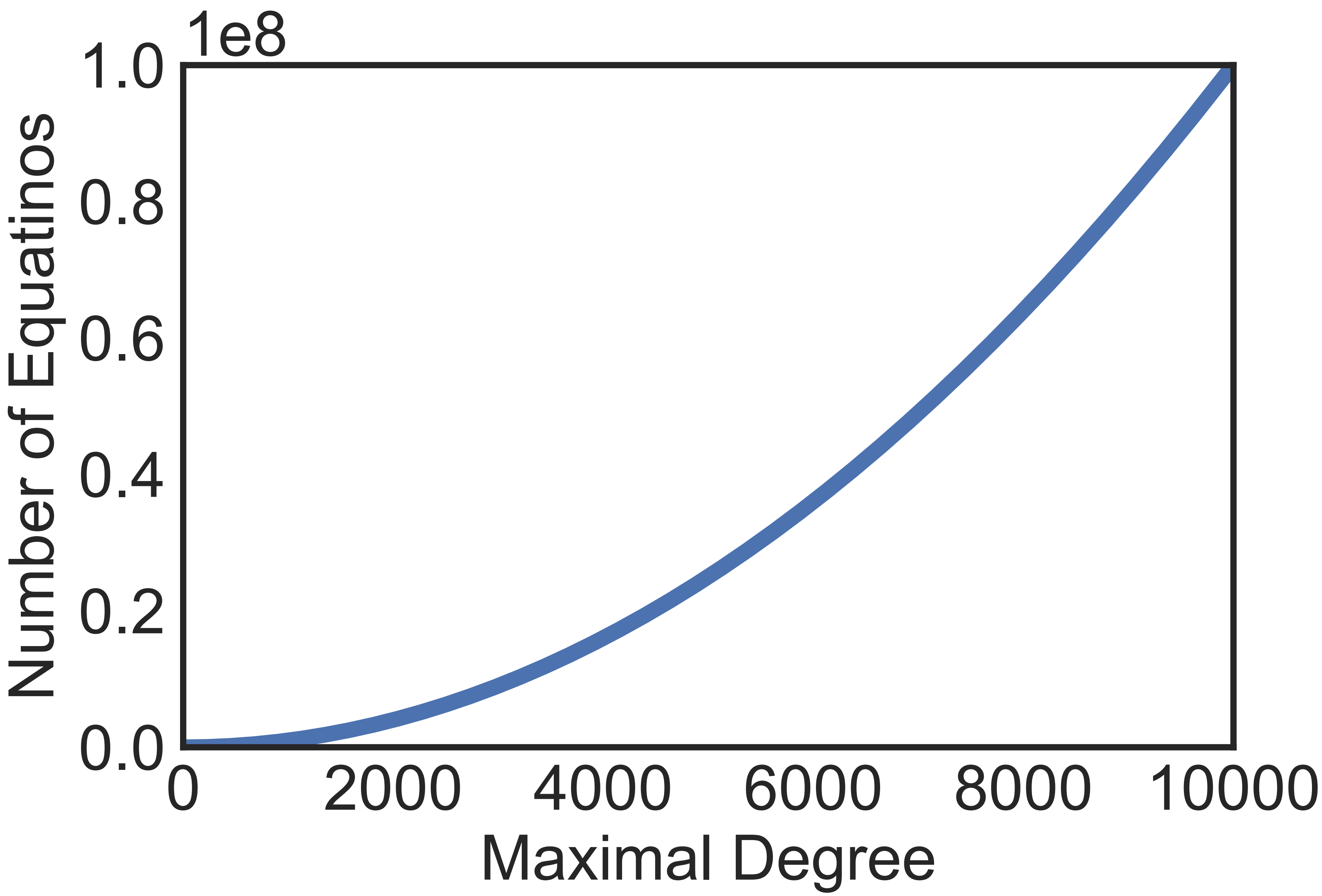


Fraction of infected nodes with 3 infected neighbours

Numerical integration

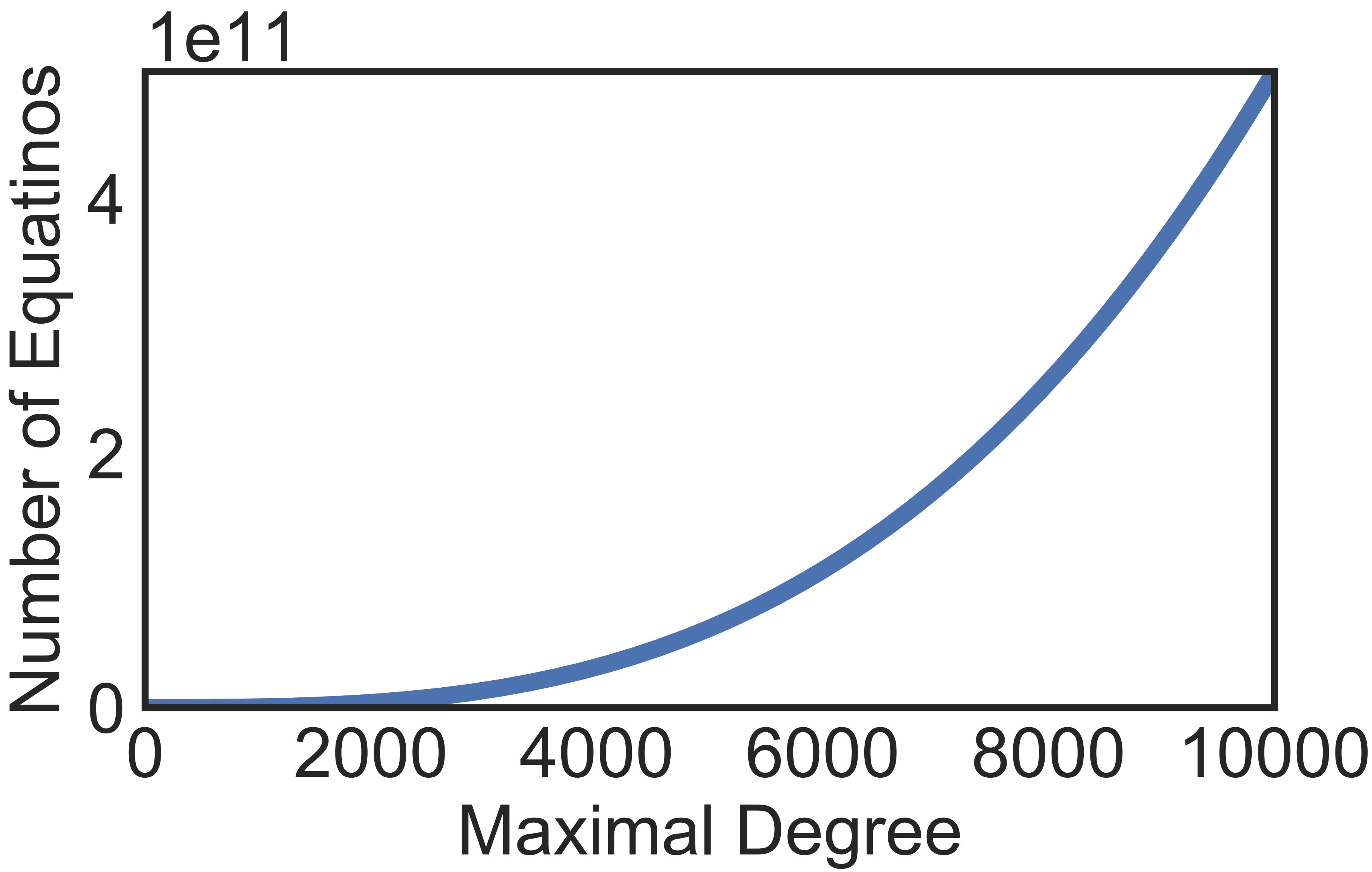


# Number of Equations




**2** node-states (SIS)

# Number of Equations



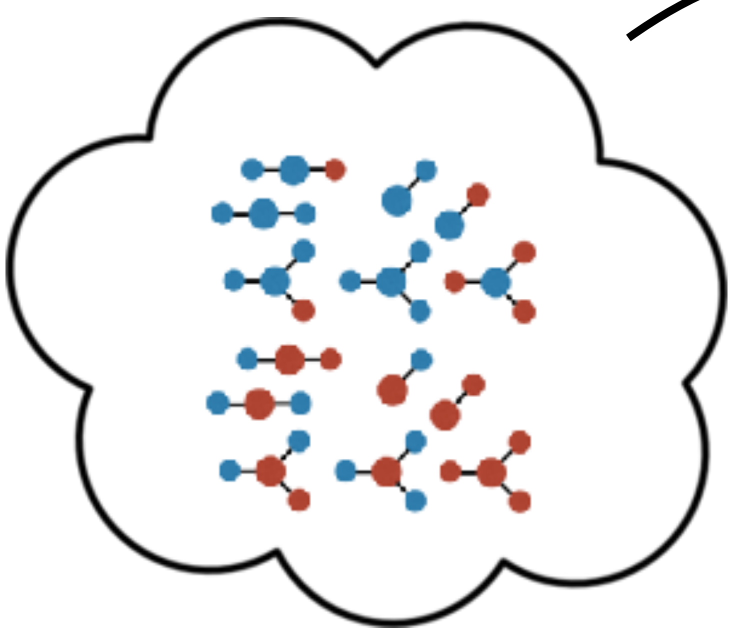

**3** node-states (SIR)

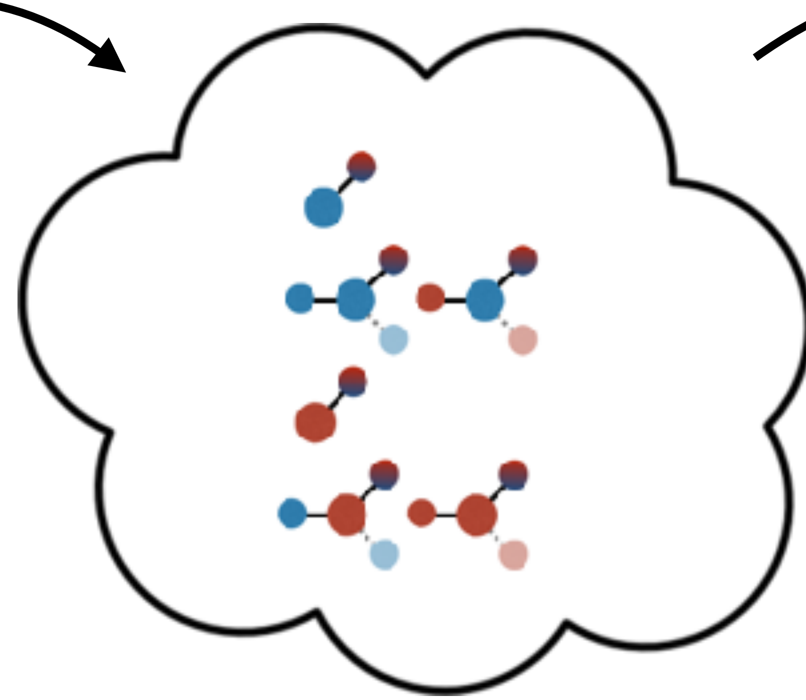
**Lumping** the Approximate Master Equation for Multistate Processes on Complex Networks

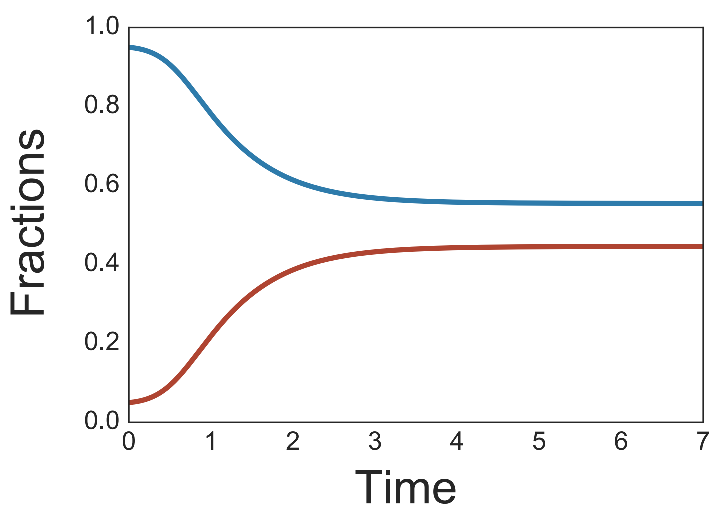
# Lumping

Aggregate ODEs

Solve ODEs





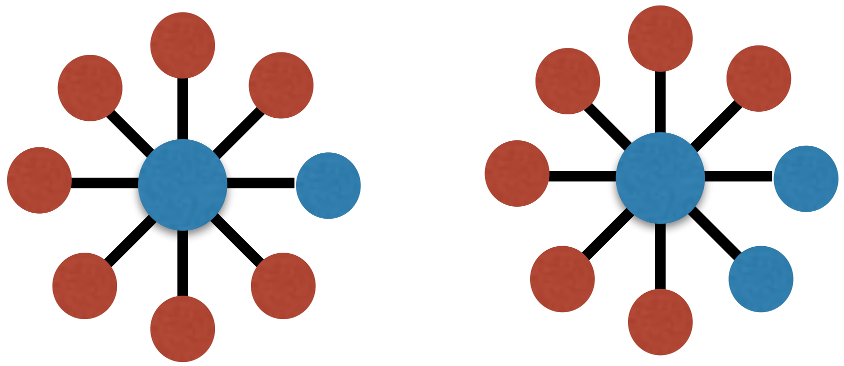



Approximate Master Equation

lumped equations

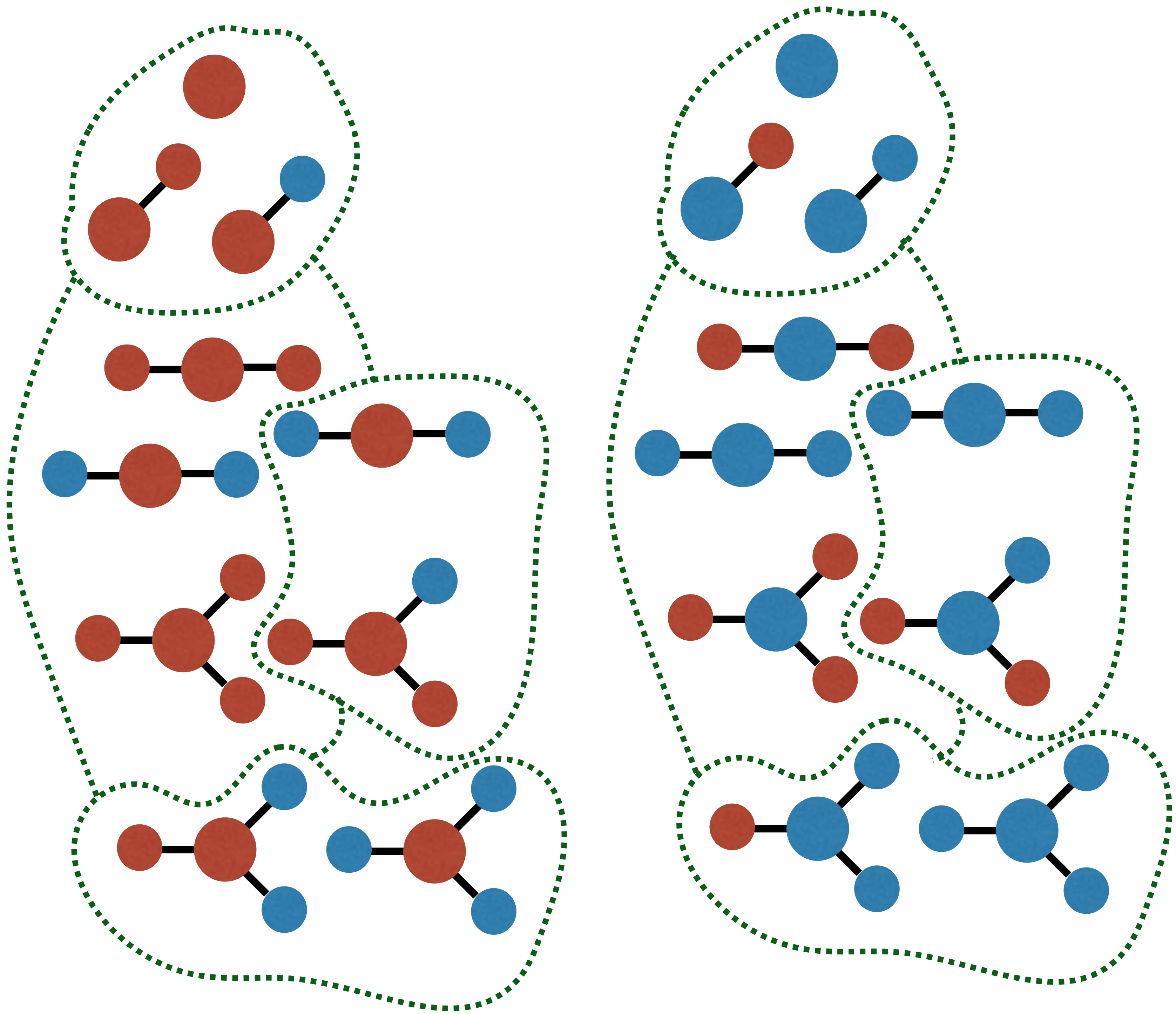
solution

# Aggregate similar ODEs together



E.g.
Store fraction of susceptible nodes with 1-2 susceptible neighbours and 6-8 infected neighbours

# Partitioning



# Lumping

$$\frac{\partial x_1}{\partial t} = x_1 + x_2$$

$$\frac{\partial x_2}{\partial t} = x_1 \cdot x_2$$

average

$$\frac{\partial x_{12}}{\partial t} = \frac{1}{2}(x_1 + x_2) + \frac{1}{2}(x_1 \cdot x_2)$$

# Lumping

$$\begin{aligned}\frac{\partial x_1}{\partial t} &= x_1 + x_2 \\ \frac{\partial x_2}{\partial t} &= x_1 \cdot x_2\end{aligned}$$

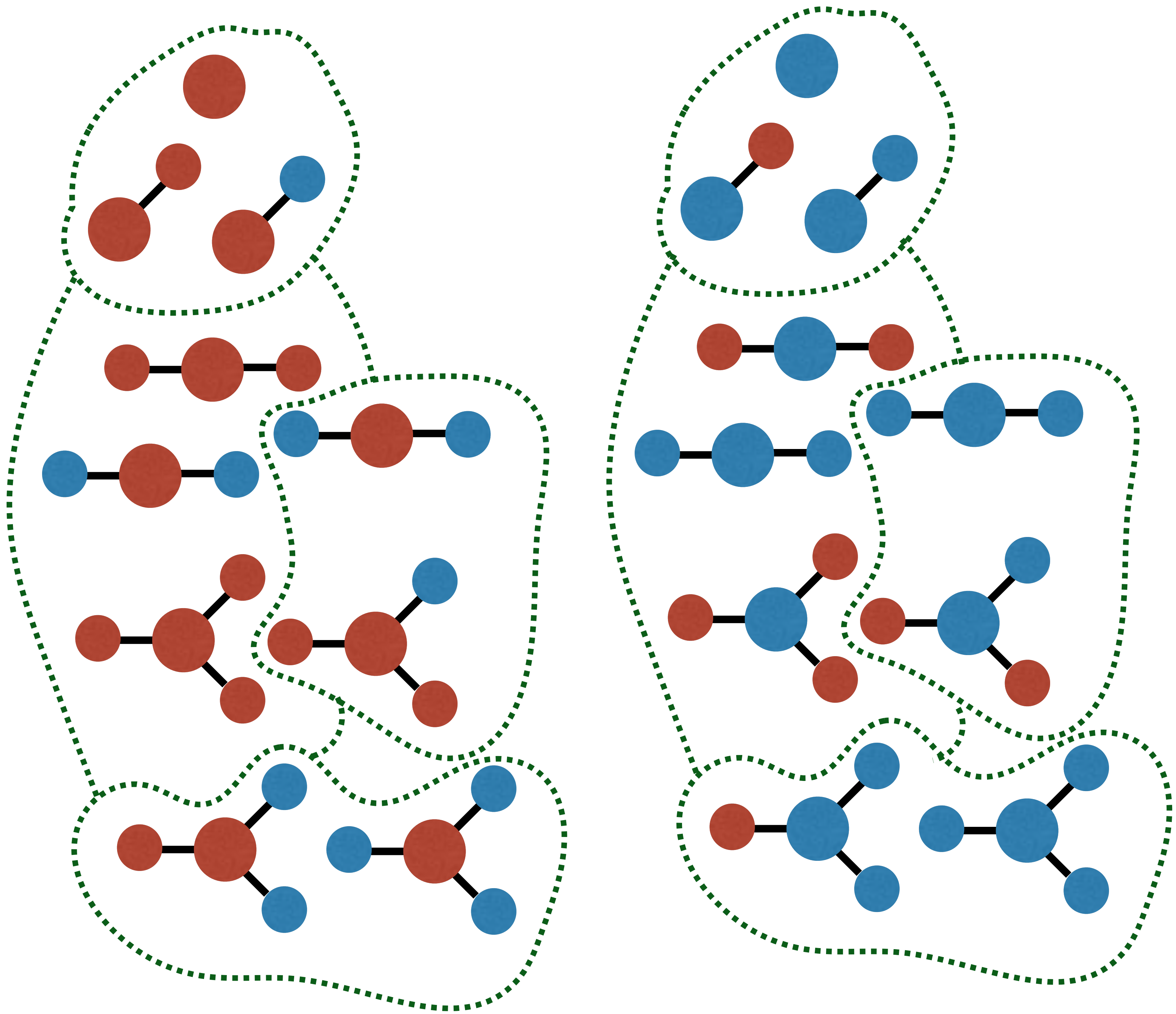
average

$$\frac{\partial x_{12}}{\partial t} = \frac{1}{2}(x_1 + x_2) + \frac{1}{2}(x_1 \cdot x_2)$$

substitute

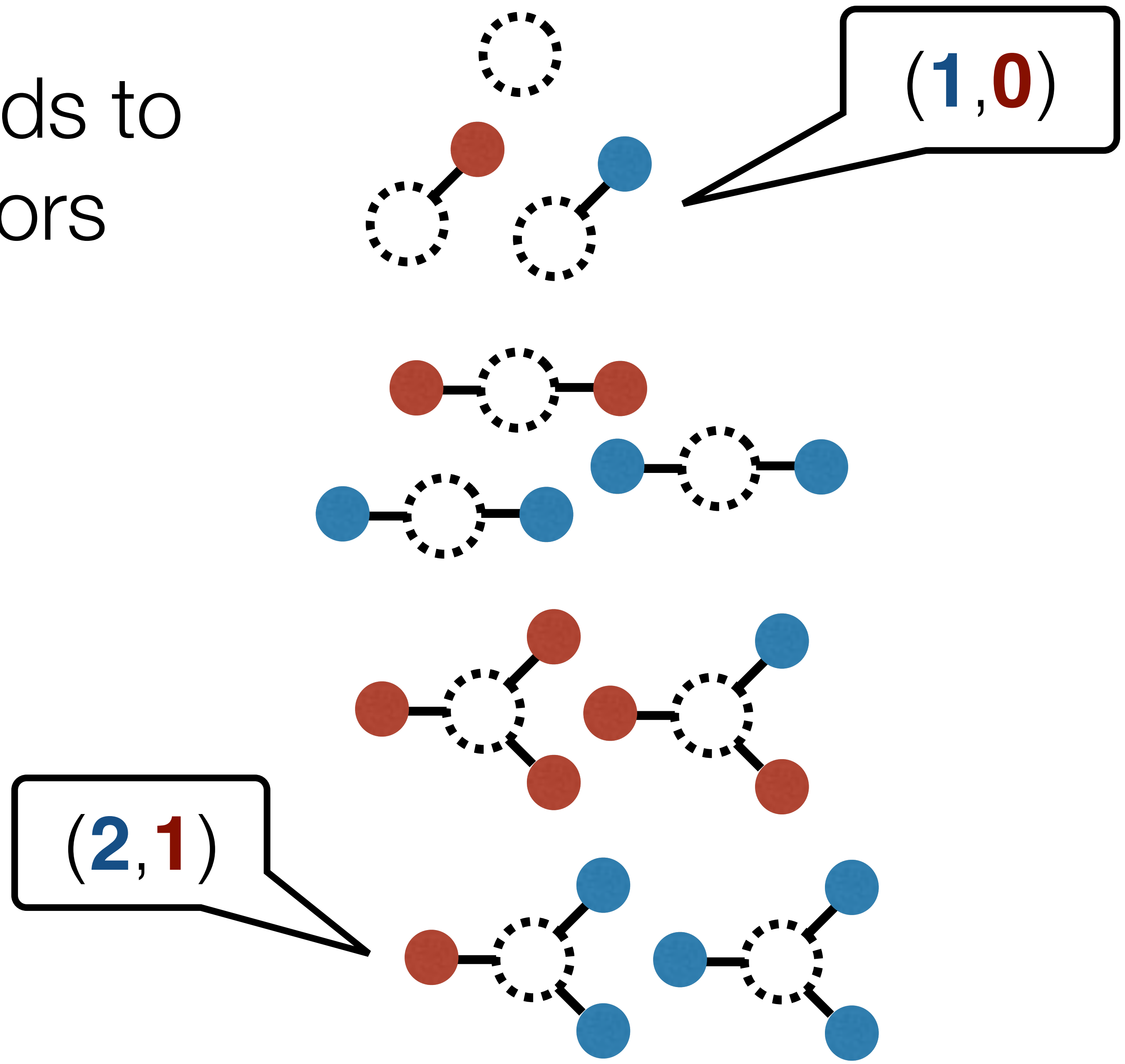
$$\frac{\partial x_{12}}{\partial t} = \frac{1}{2}(x_{12} + x_{12}) + \frac{1}{2}(x_{12} \cdot x_{12})$$

# Partitioning



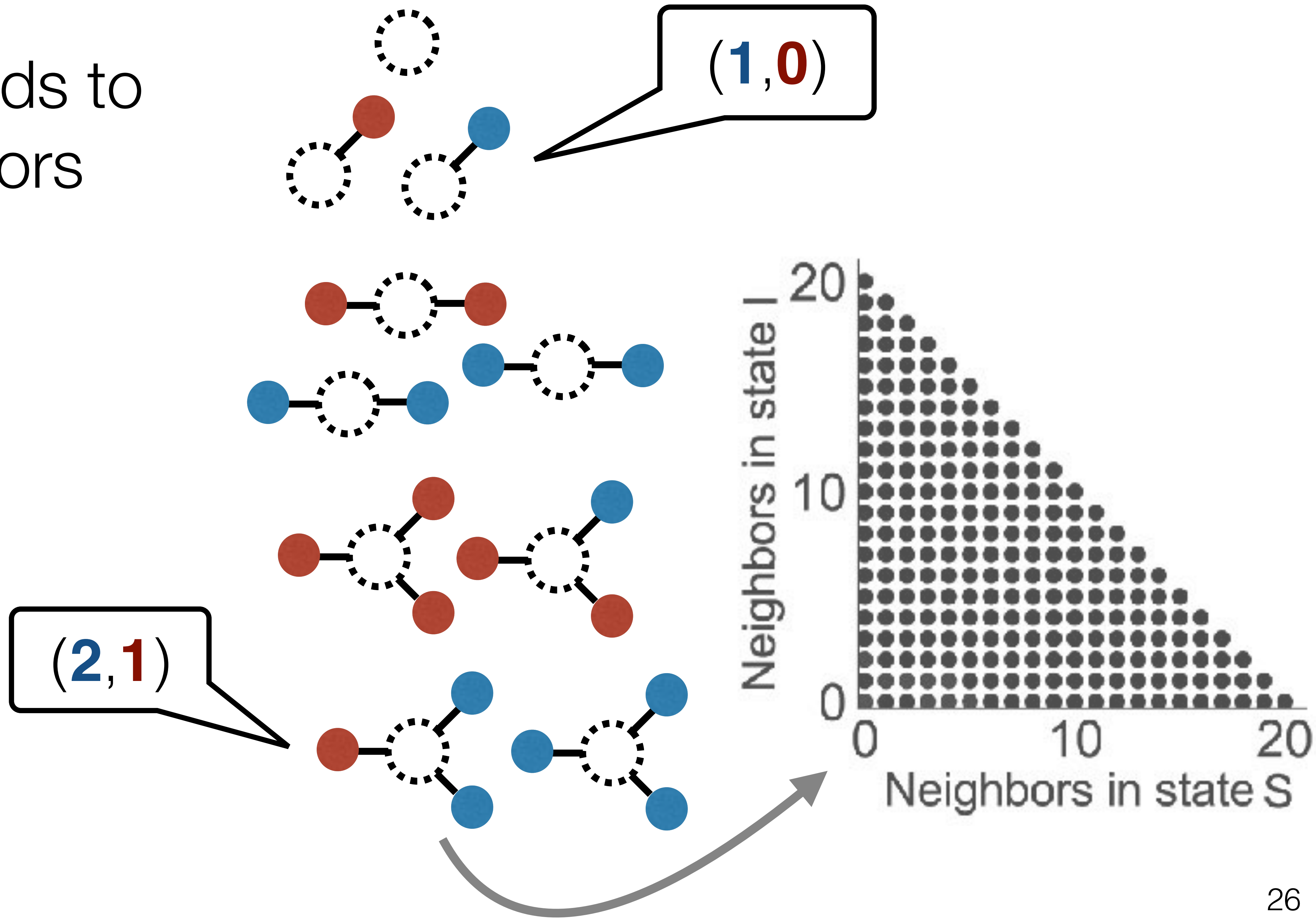
# Partitioning

Map neighbourhoods to integer vectors

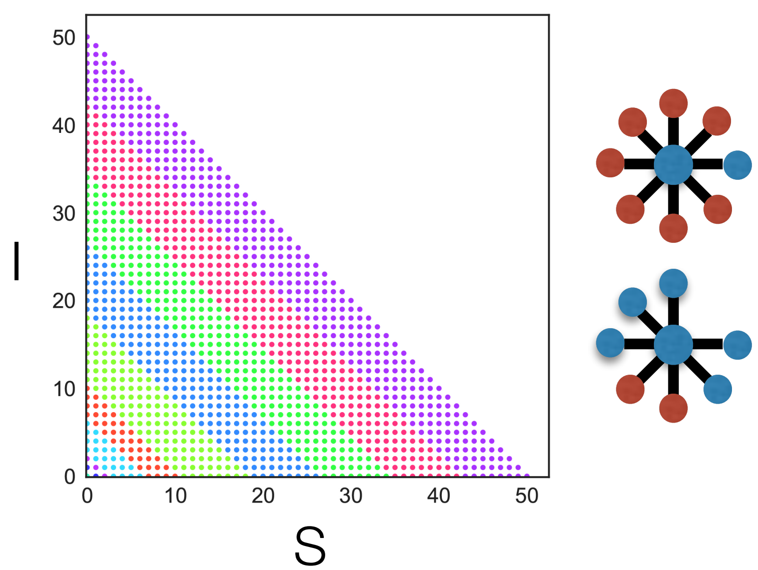


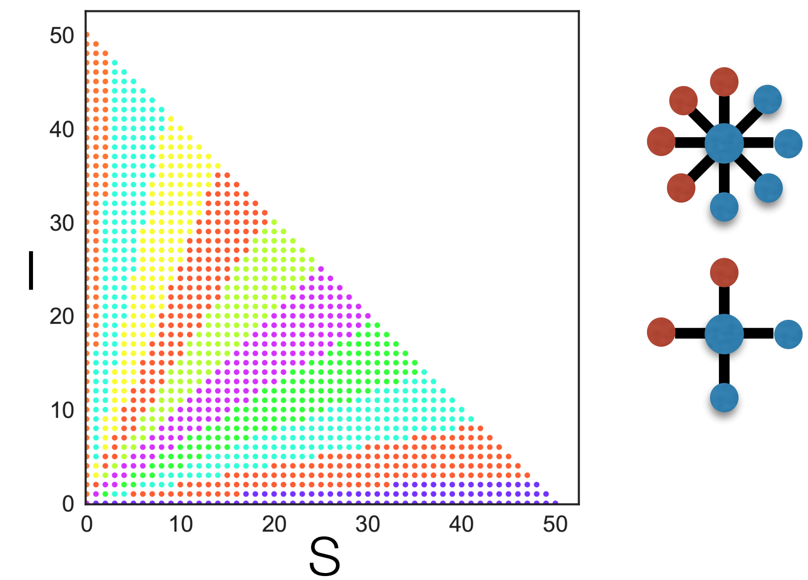
# Partitioning

Map neighbourhoods to integer vectors



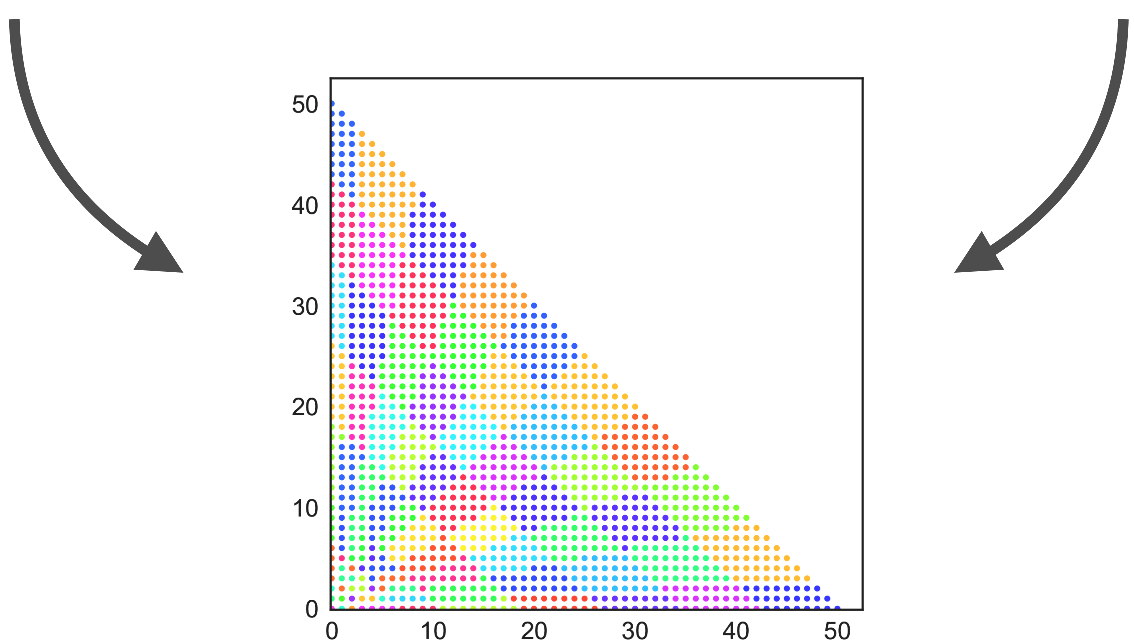
# Partitioning



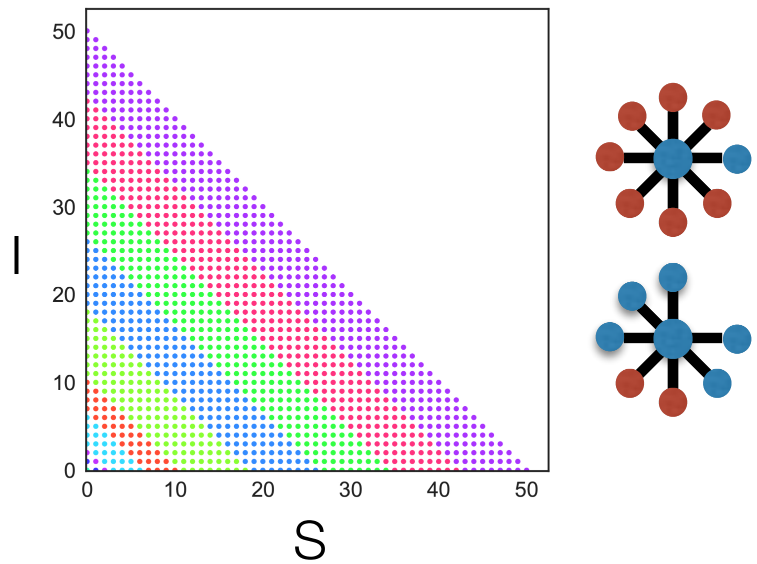




Cluster degrees $k$ based on degree-distribution $P(k)$

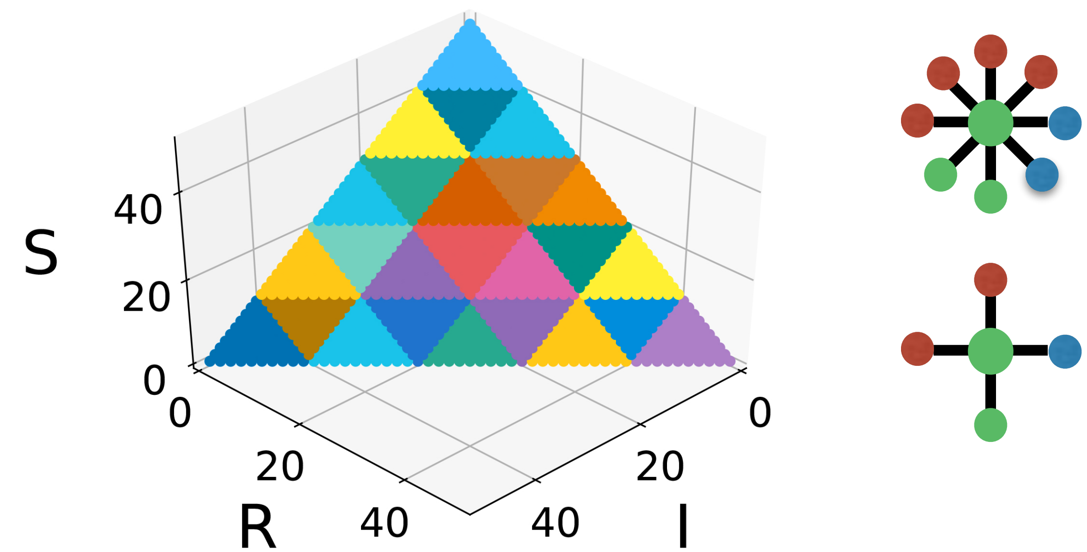
Cluster neighbourhood vectors based on proportionality



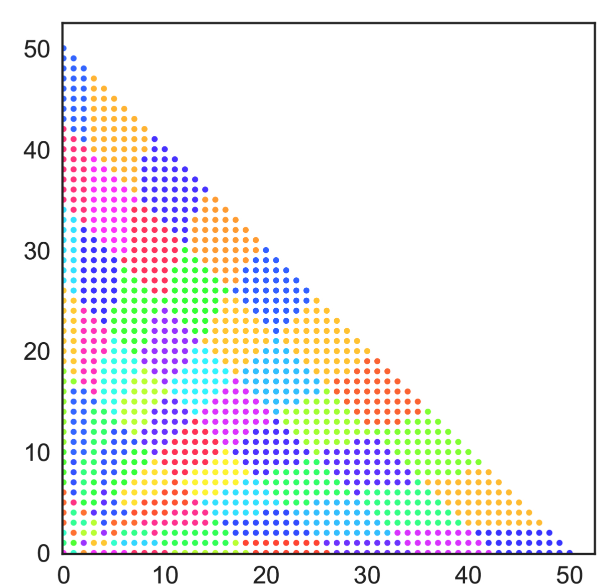
# Partitioning 3D




Cluster degrees $k$ based on degree-distribution $P(k)$




Cluster neighbourhood vectors based on proportionality



# How Many Partitions?

## Simple Stopping Heuristic

- Start with small number of partitions
- Increase partition number and solve model as long as solution changes

# How Many Partitions?

Simple Stopping Heuristic

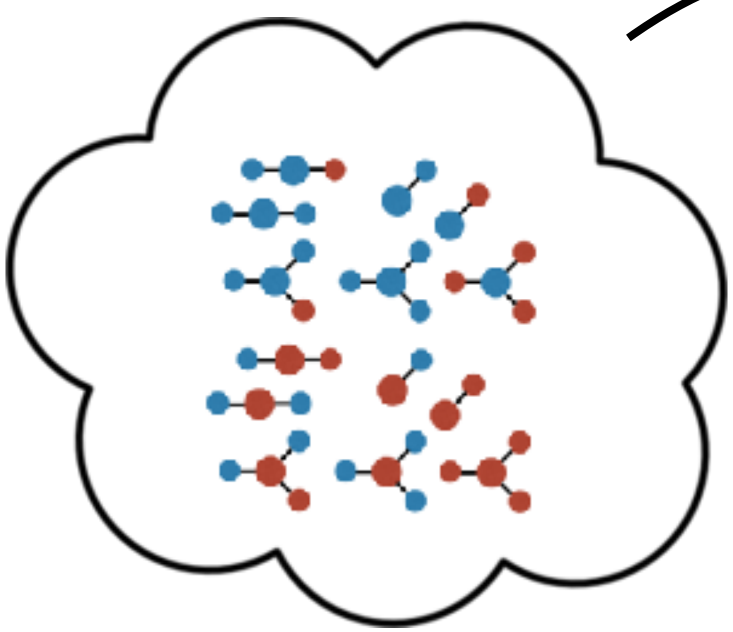
- Start with small number of partitions
- Increase partition number and solve model as long as solution changes

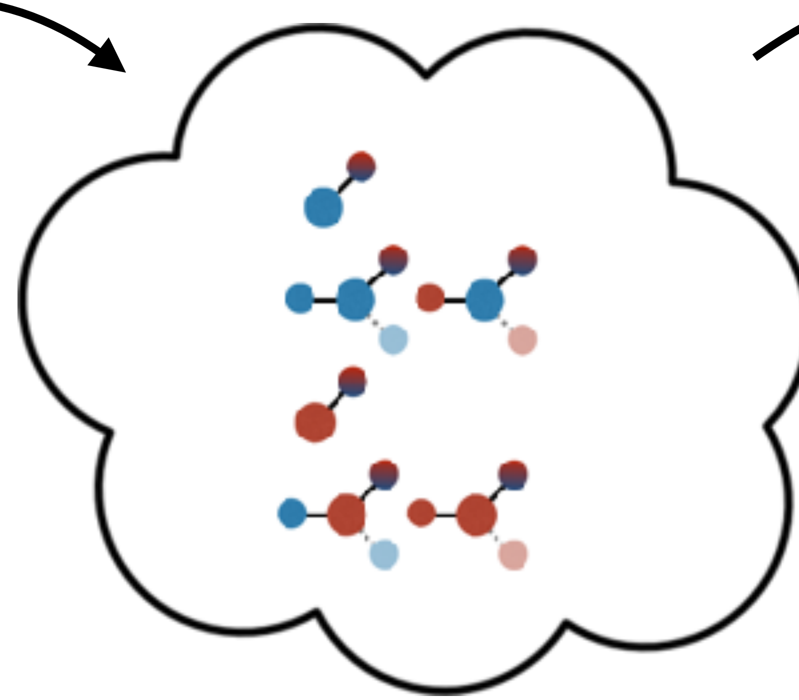
Difference between consecutive solutions can be used to predict lumping error

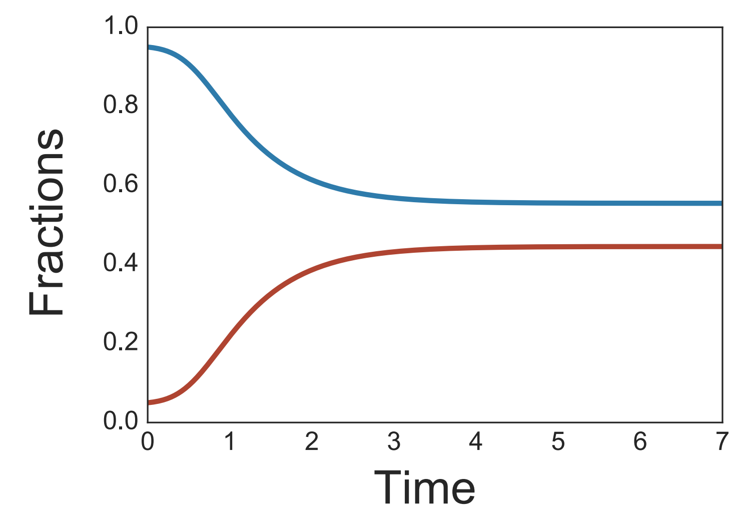
# Lumping

Aggregate ODEs

Solve ODEs





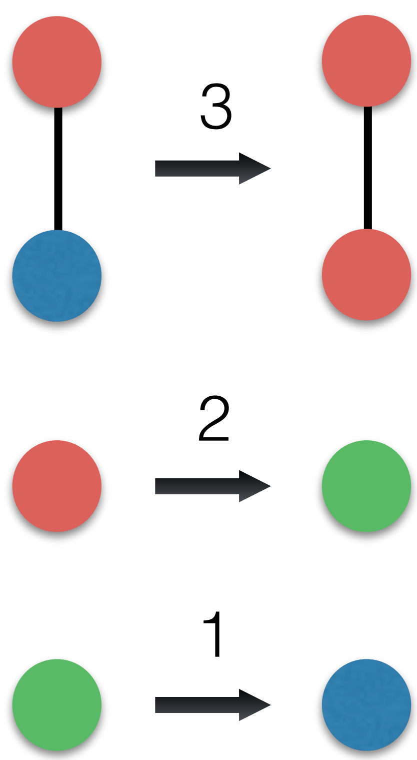


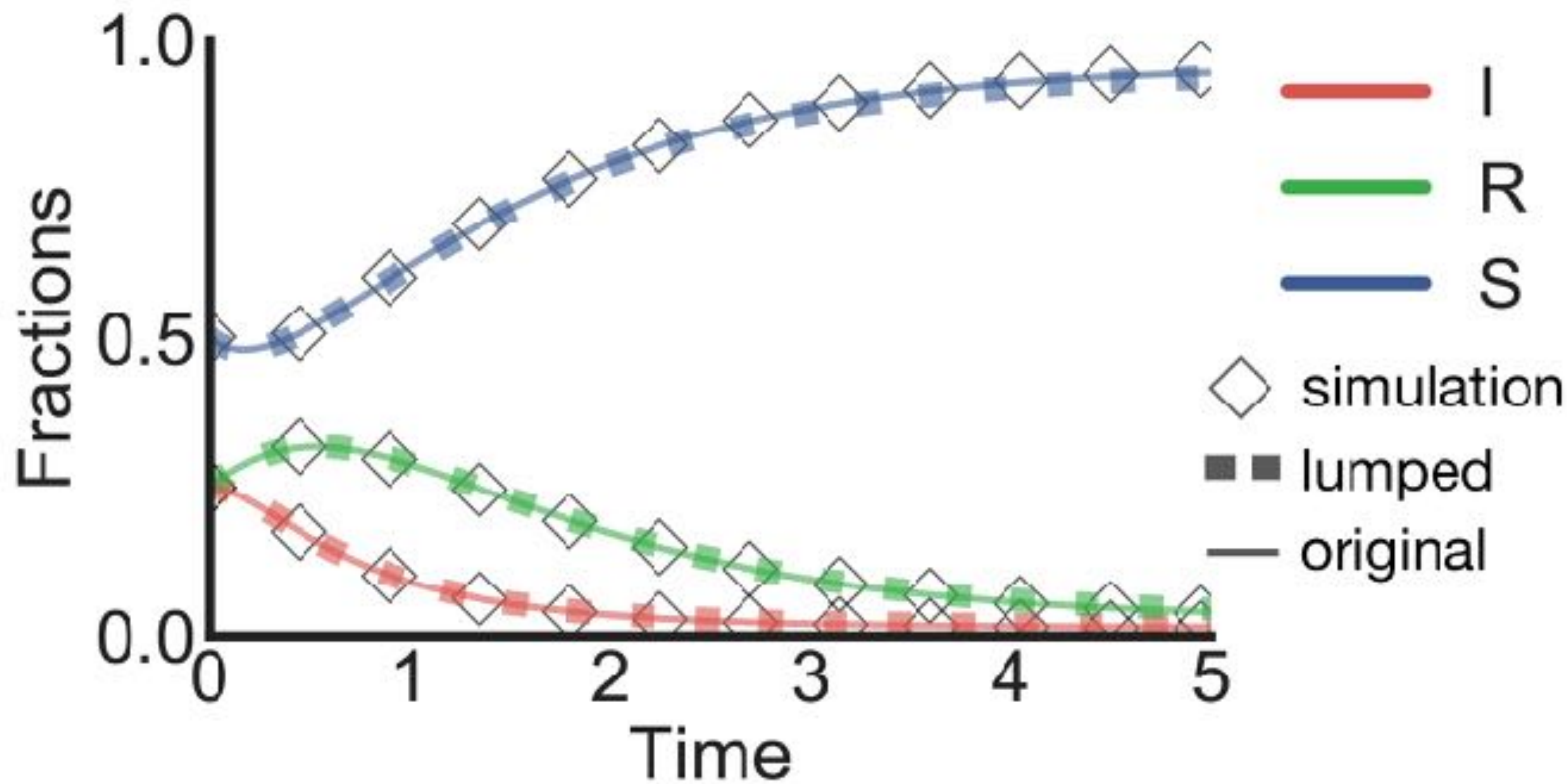

Approximate Master Equation

lumped equations

solution

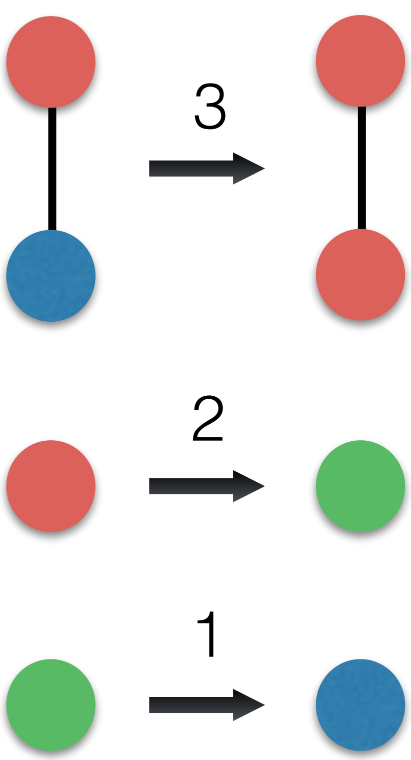
# Results SIR



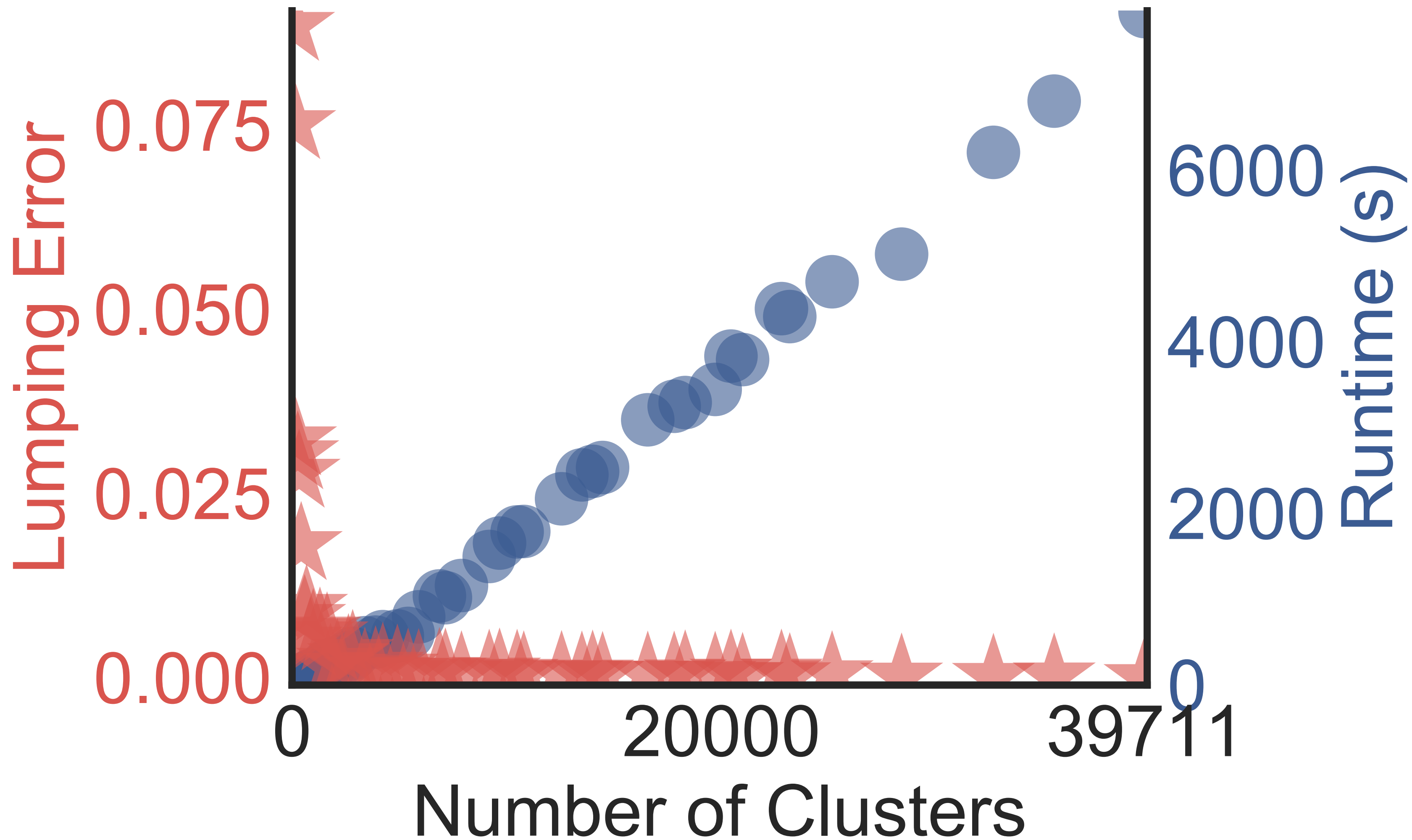




5 373 ODEs vs 119 133 ODEs    Truncated power-law with $\gamma = 2.5$ and $k_{\max} = 60$

# Results SIR

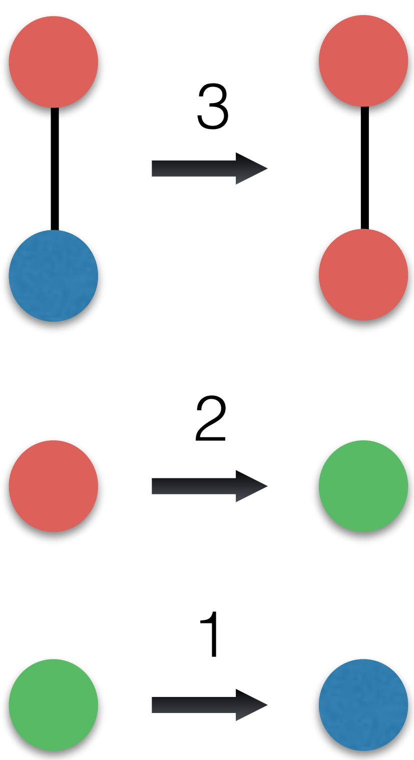


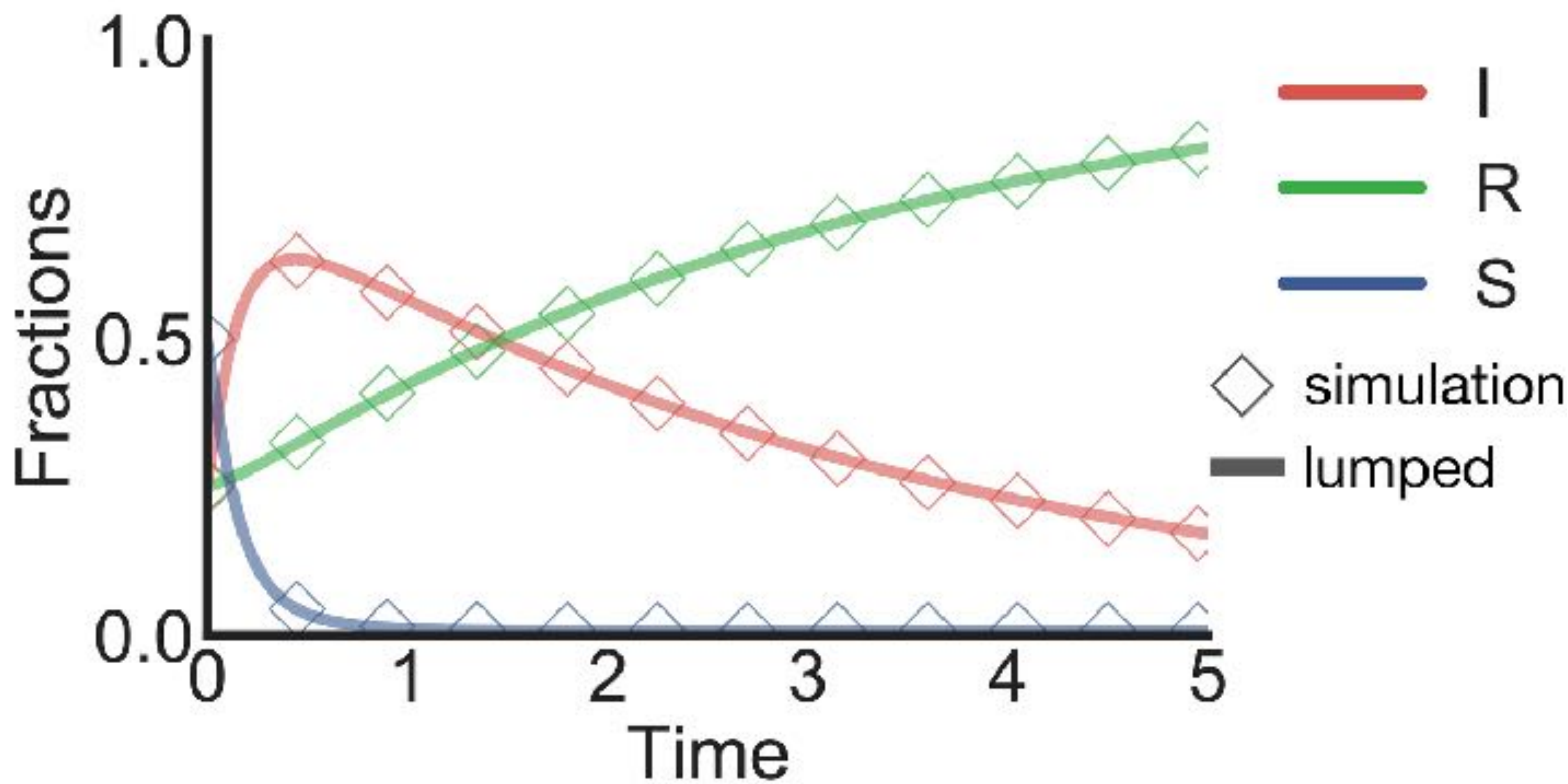
3
2
1

Lumping Error
0.075
0.050
0.025
0.000
6000
4000
2000
0
Runtime (s)
0
20000
39711
Number of Clusters

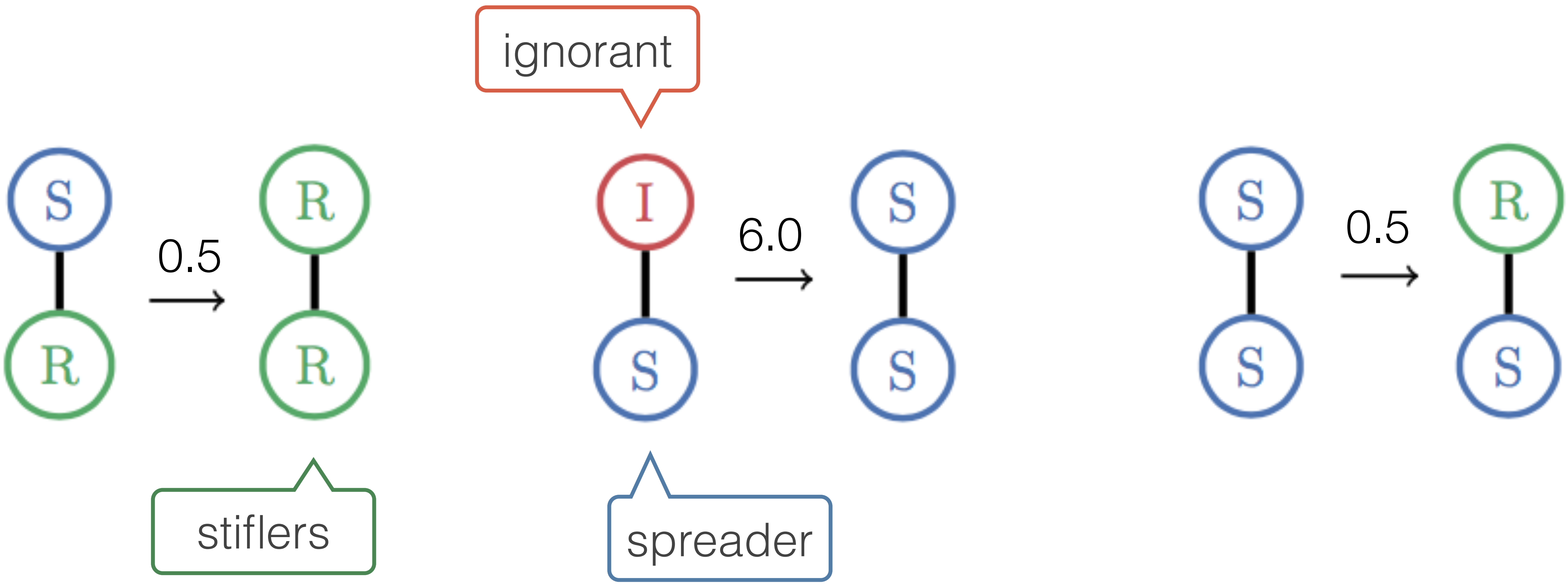
# Results SIR



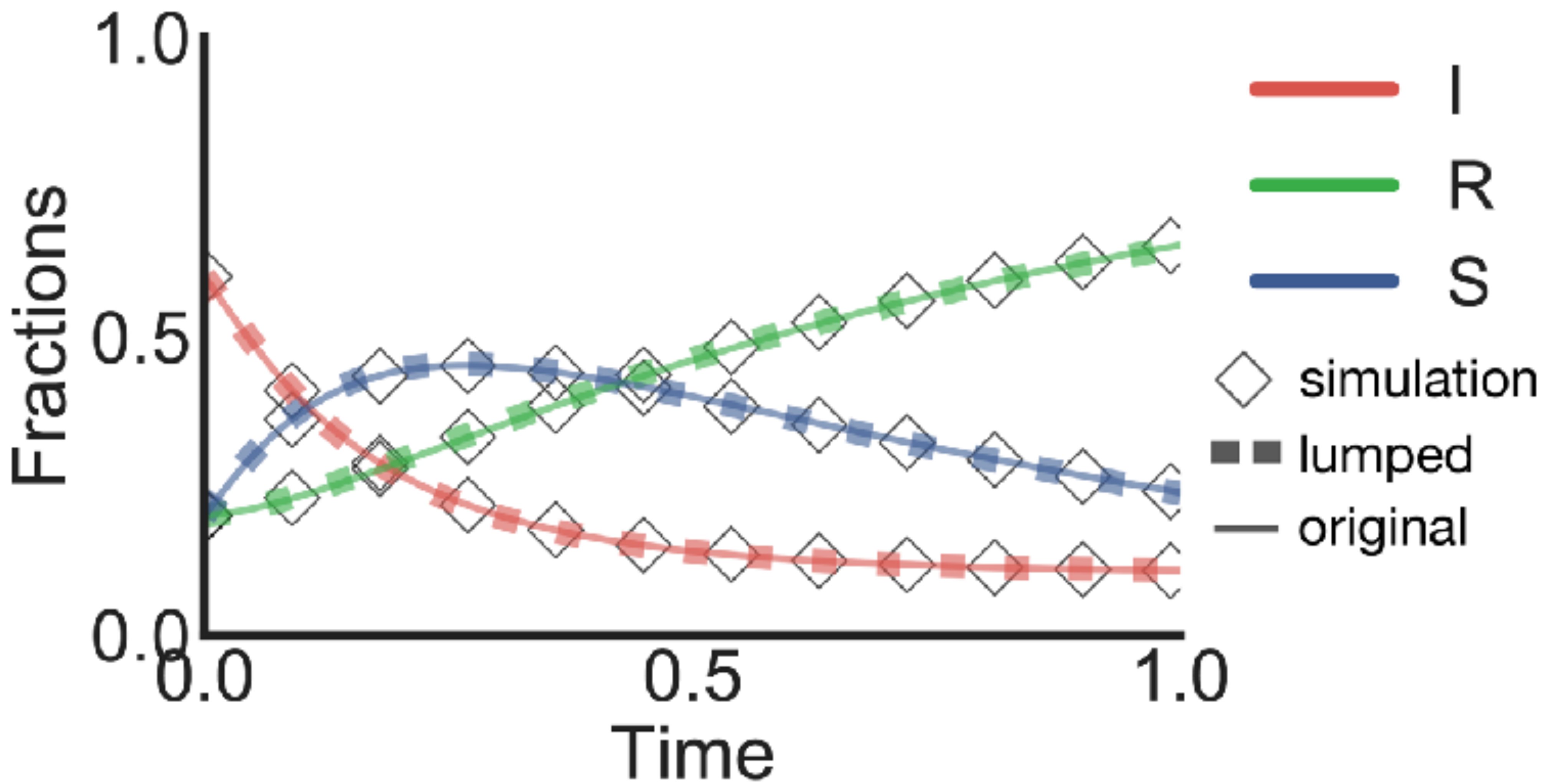




25K ODEs vs 63M ODEs     Truncated power-law with $\gamma = 2.5$ and $k_{\max} = 600$

# Rumor Spreading

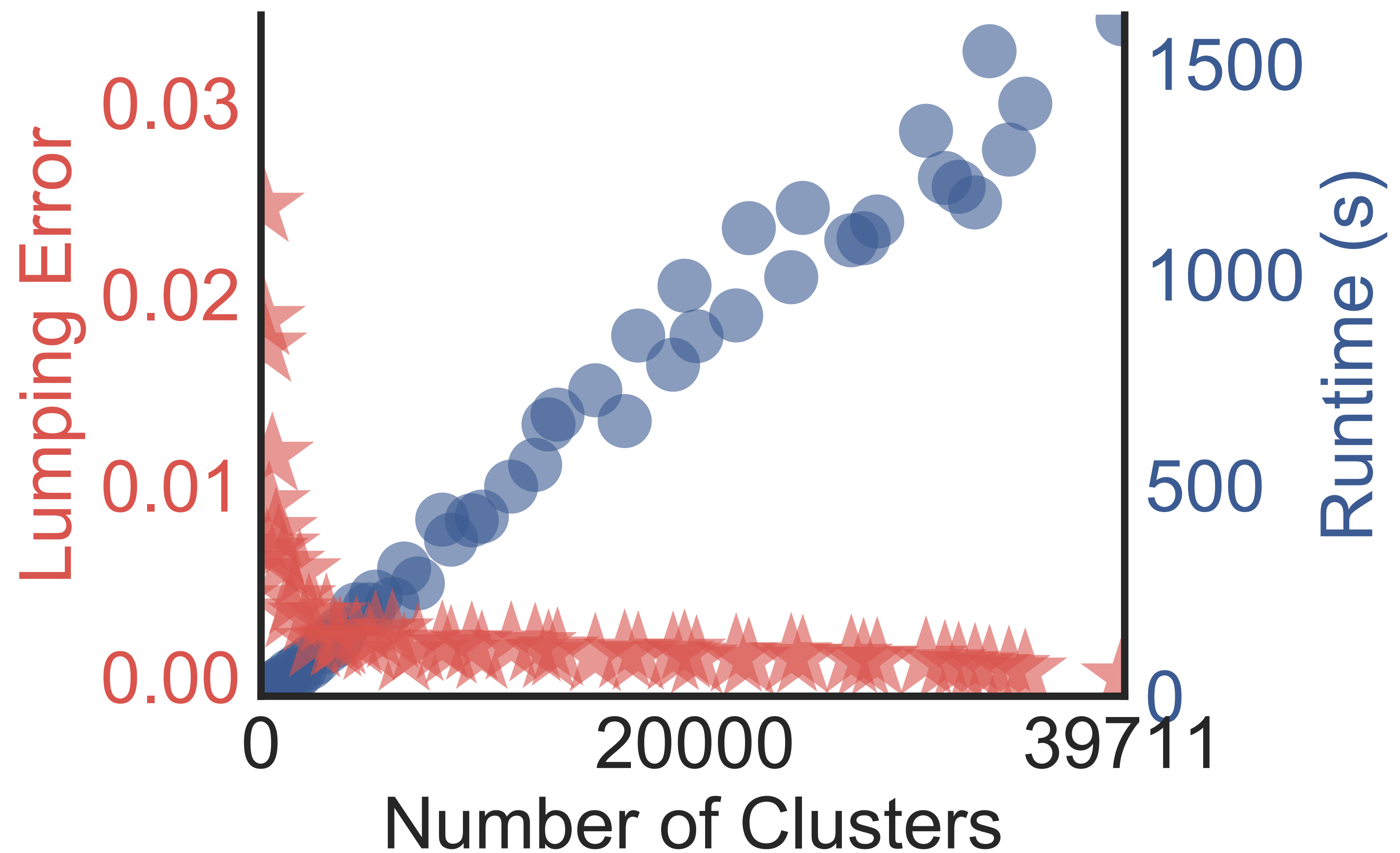


# Rumor Spreading




3 096 ODEs vs 119 133 ODEs  Truncated power-law with $\gamma = 3.0$ and $k_{\max} = 60$

# Rumor Spreading



# Conclusion

- Lumping exploits redundancies in AME
- Massive reduction of computational time
- Lumping error of similar order as AME error

Future Work:

- Take spatial information into account

*Thank you*

Slides and Code: github.com/gerritgr/LumPyQest

# AME

Fraction of label *s* with neighbourhood ***m***

ODEs

$$\frac{\partial x_{s,\mathbf{m}}}{\partial t} = \sum_{(s',f,s)\in R^{s+}} f(\mathbf{m})x_{s',\mathbf{m}} - \sum_{(s,f,s')\in R^{s-}} f(\mathbf{m})x_{s,\mathbf{m}}$$
$$+ \sum_{\substack{(s_1,s_2)\in\mathbb{S}^2 \\ s_1\neq s_2}} \beta^{ss_1\to ss_2} x_{s,\mathbf{m}^{\{s_1^+,s_2^-\}}}\, \mathbf{m}^{\{s_1^+,s_2^-\}}[s_1]$$
$$- \sum_{\substack{(s_1,s_2)\in\mathbb{S}^2 \\ s_1\neq s_2}} \beta^{ss_1\to ss_2} x_{s,\mathbf{m}}\mathbf{m}[s_1]$$

local

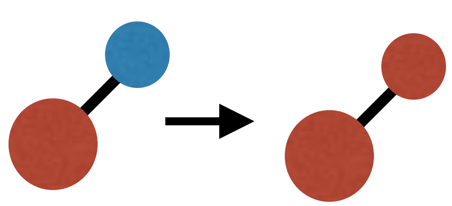
global

weighted average over all degrees

$$x_s(t) = \sum_{\mathbf{m}\in\mathbb{M}} x_{s,\mathbf{m}}(t)\cdot P(k_{\mathbf{m}})$$

$$\beta^{ss_1\to ss_2} = \frac{\sum_{\mathbf{m}\in\mathbb{M}} P(k_{\mathbf{m}}) \sum_{(s_1,f,s_2)\in R^{s_1\to s_2}} f(\mathbf{m})x_{s_1,\mathbf{m}}\mathbf{m}[s]}{\sum_{\mathbf{m}\in\mathbb{M}} P(k_{\mathbf{m}})x_{s_1,\mathbf{m}}\mathbf{m}[s]}$$

average rates of pairs



# Lumping the AME

Original

$$\frac{\partial x_{s,\mathbf{m}}}{\partial t} = \sum_{(s',f,s)\in R^{s+}} f(\mathbf{m}) x_{s',\mathbf{m}} - \sum_{(s,f,s')\in R^{s-}} f(\mathbf{m}) x_{s,\mathbf{m}} + \sum_{\substack{(s_1,s_2)\in\mathbb{S}^2\\ s_1\neq s_2}} \beta^{ss_1\to ss_2} x_{s,\mathbf{m}^{\{s_1^+,s_2^-\}}} \mathbf{m}^{\{s_1^+,s_2^-\}}[s_1] - \sum_{\substack{(s_1,s_2)\in\mathbb{S}^2\\ s_1\neq s_2}} \beta^{ss_1\to ss_2} x_{s,\mathbf{m}} \mathbf{m}[s_1]$$

Lumped

$$\frac{\partial z_{s,C}}{\partial t} = \underbrace{\sum_{\mathbf{m}\in C} P(k_{\mathbf{m}}|C)}_{\text{Average}} \cdot \Bigg( \sum_{(s',f,s)\in R^{s+}} f(\mathbf{m}) \underline{r_{C,k_{\mathbf{m}}} z_{s',C}} - \sum_{(s,f,s')\in R^{s-}} f(\mathbf{m}) \underline{r_{C,k_{\mathbf{m}}} z_{s,C}} + \sum_{\substack{(s_1,s_2)\in\mathbb{S}^2\\ s_1\neq s_2}} \beta_{\mathcal{L}}^{ss_1\to ss_2} \underline{r_{C,k_{\mathbf{m}}} z_{s,C_{\mathbf{m}}^{\{s_1^+,s_2^-\}}}} \mathbf{m}^{\{s_1^+,s_2^-\}}[s_1] - \sum_{\substack{(s_1,s_2)\in\mathbb{S}^2\\ s_1\neq s_2}} \beta_{\mathcal{L}}^{ss_1\to ss_2} \underline{r_{C,k_{\mathbf{m}}} z_{s,C}} \mathbf{m}[s_1] \Bigg)$$

Substitution