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
- …

img © Michael Kreil

Motivation

Understand spreading phenomenon of

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

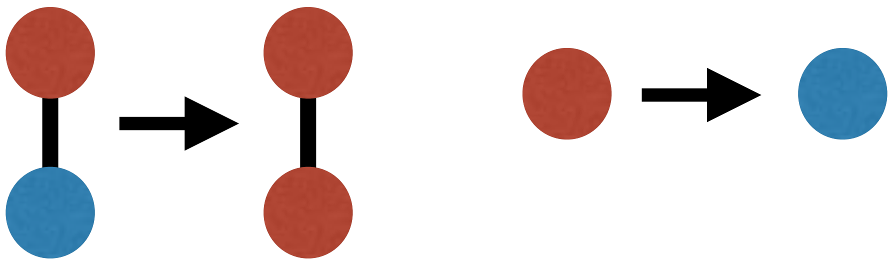
img © Michael Kreil

Spreading Process

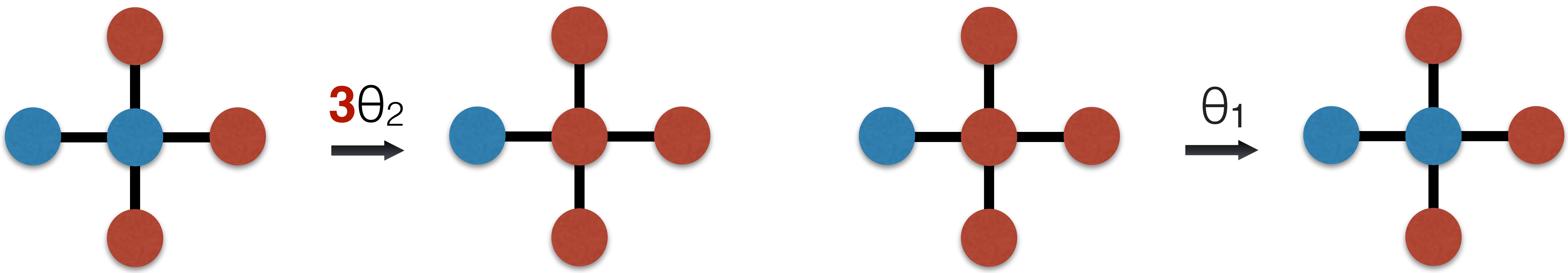
- Fixed graph topology
- Continuous time dynamics
- Nodes have local states
- Nodes' states change randomly w.r.t. rules

Classical example: SIS model

- 2 local states (**infected**, **susceptible**)
- 2 rules (infection,   recovery)



SIS Model

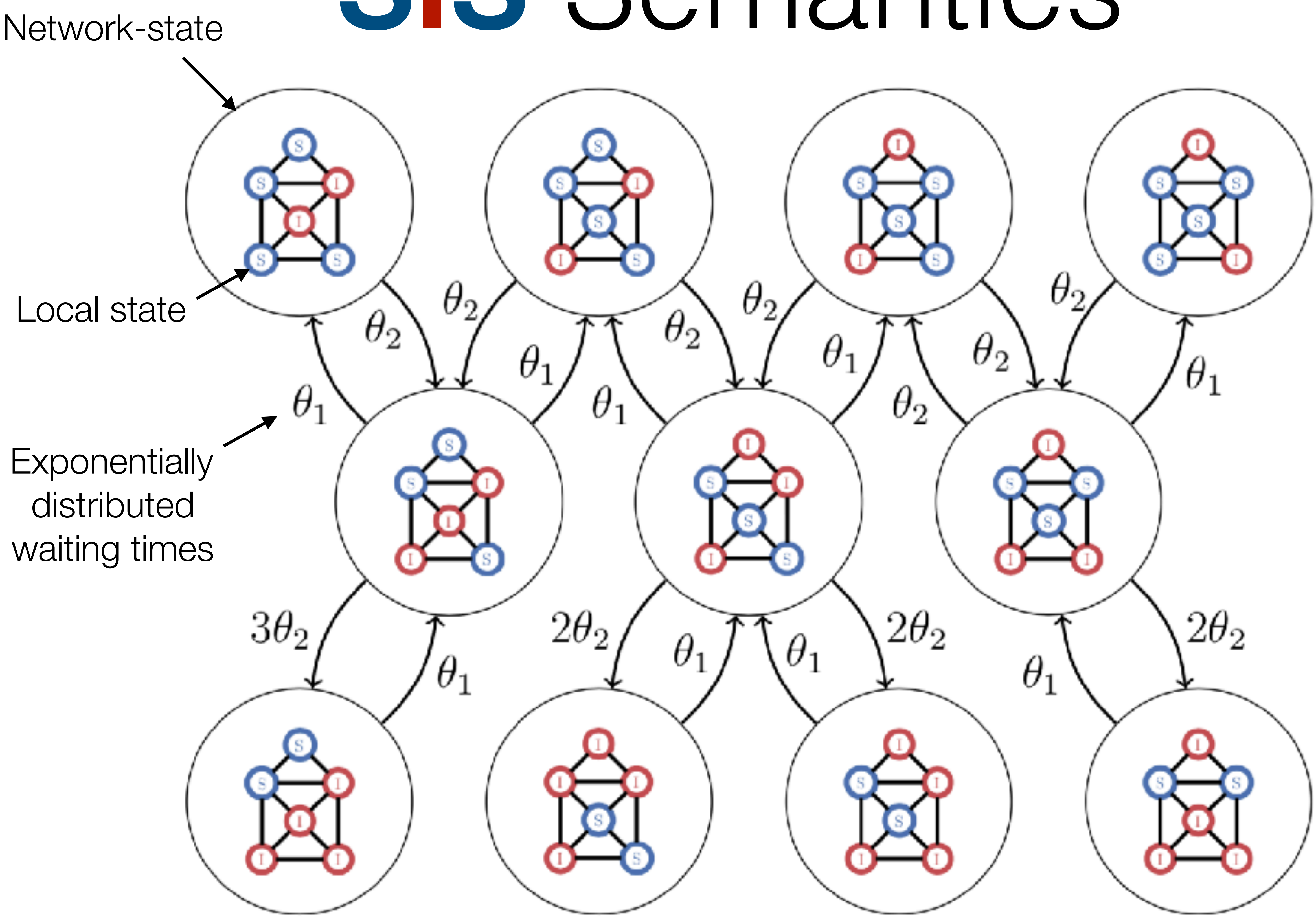


Infection (edge-based)

Recovery (node-based)

- 2 local states (**infected**, **susceptible**)
- 2 rules (infection, recovery)

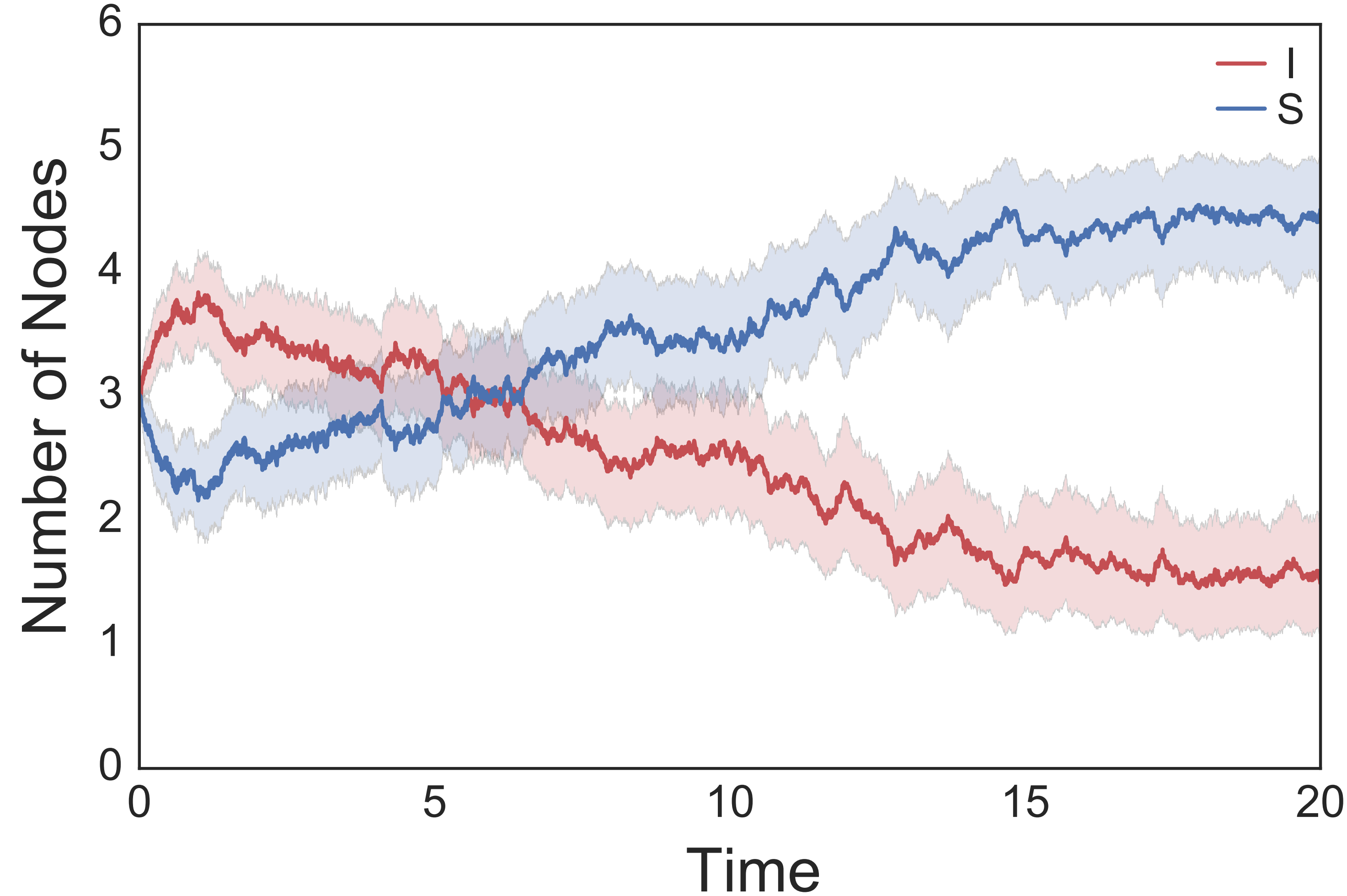
SIS Semantics



Fraction of Infected Nodes

Measure of interest:

- Fractions of nodes in each local state
- Can be estimated with Monte-Carlo simulation (expensive)
- Simple mean field approximations are faster but not always accurate

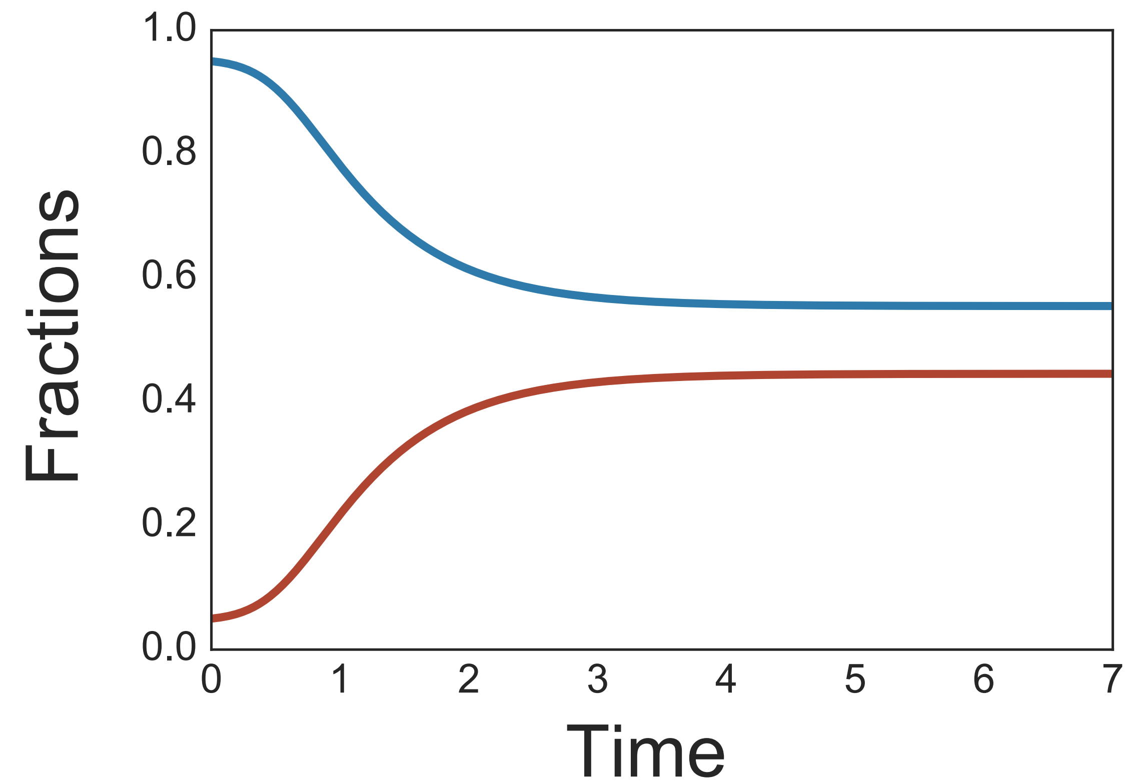


Lumping the **Approximate Master Equation** for Multistate Processes on Complex Networks

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 local state and possible neighbourhood

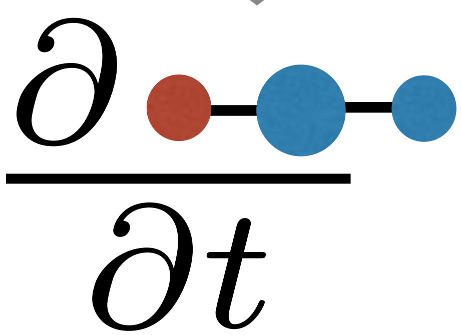


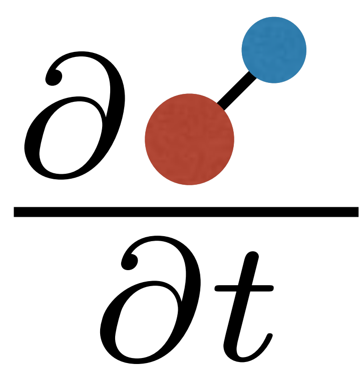
High-Accuracy Approximation of Binary-State Dynamics on Networks (James P. Gleeson, 2011)

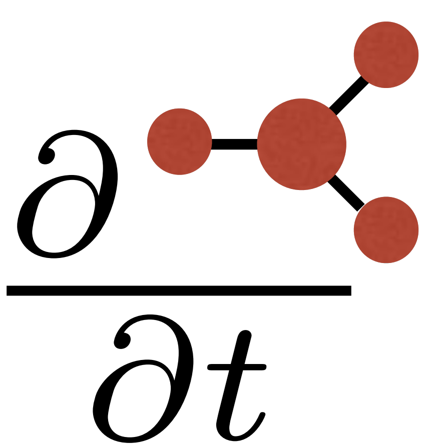
AME

one ODE for each local state and neighbourhood

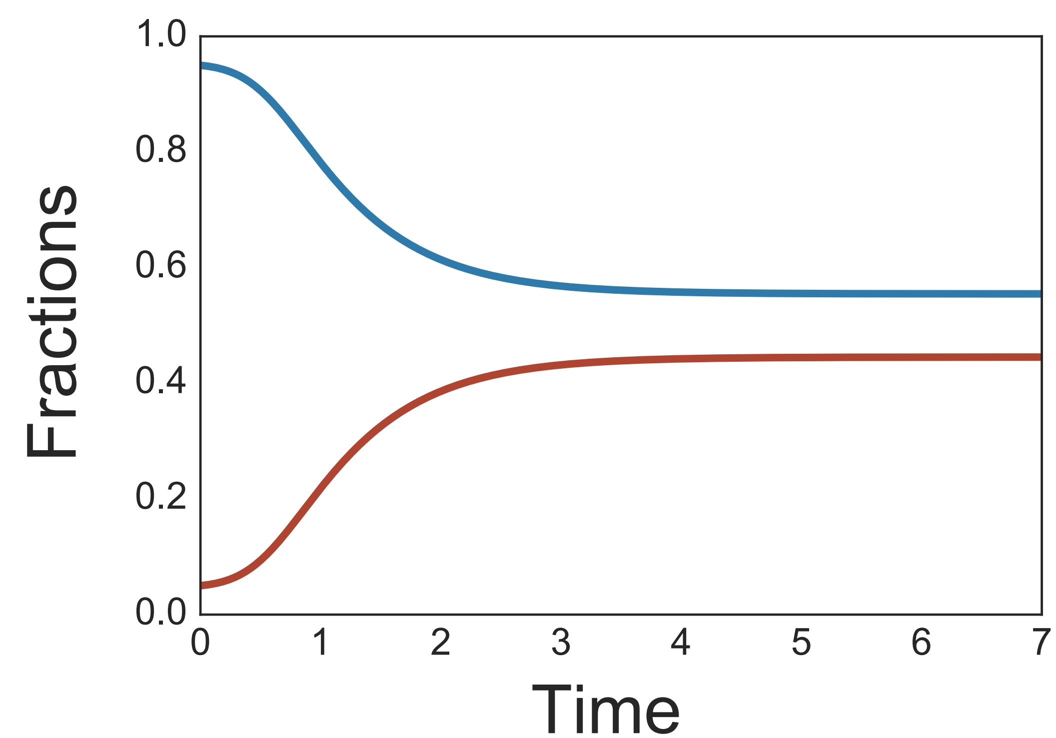
Fraction of susceptible nodes with 1 susceptible and 1 infected neighbour



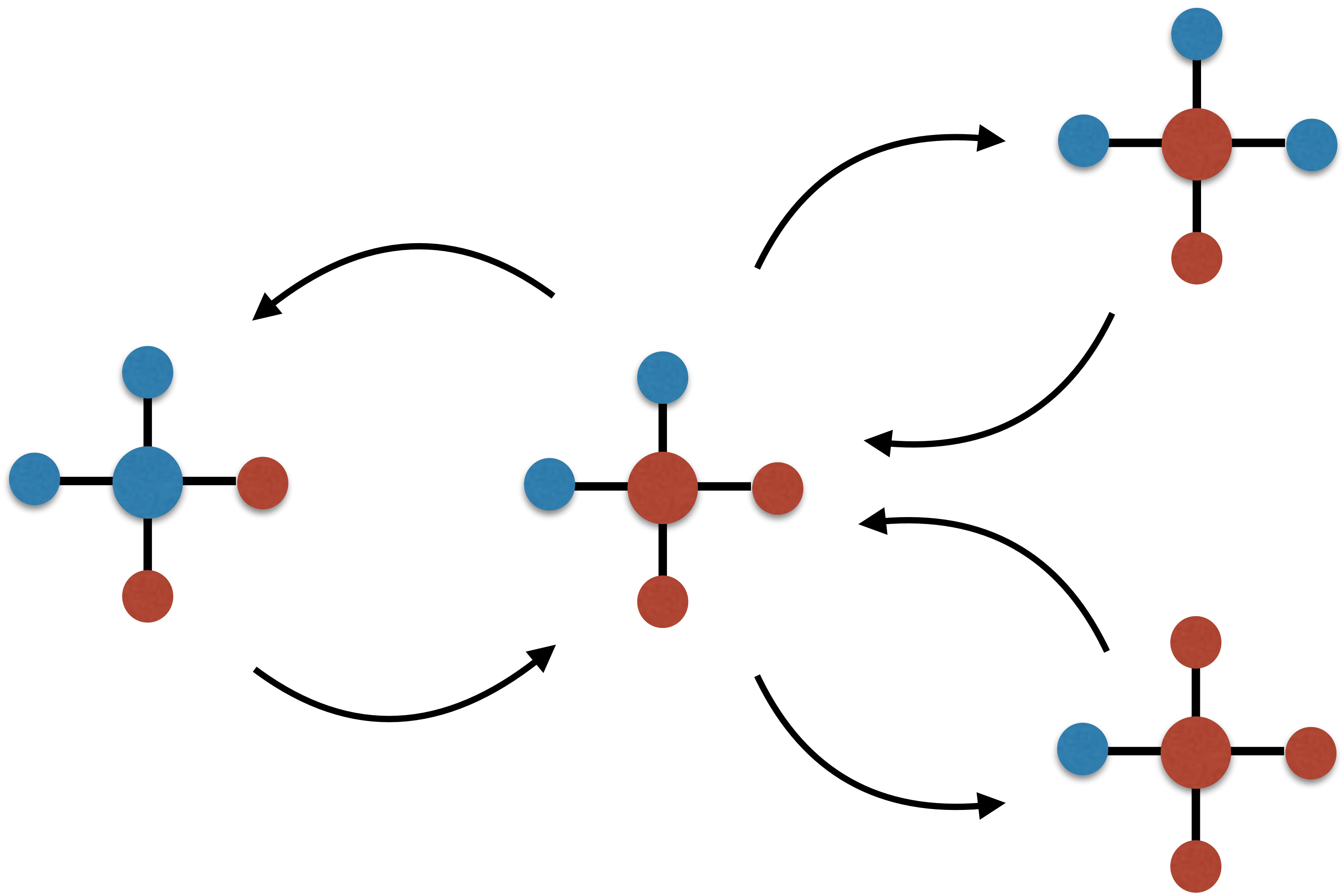


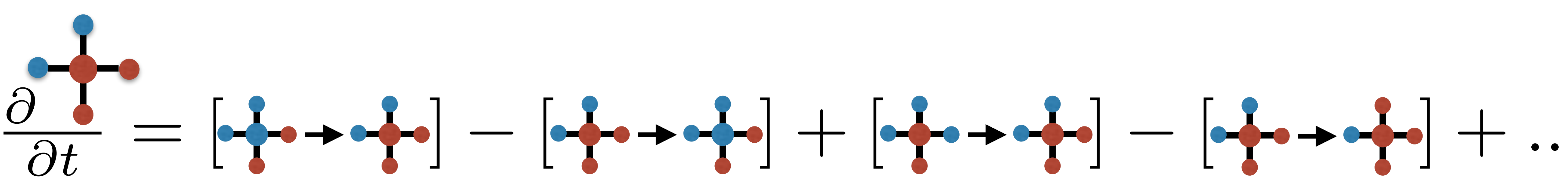


Numerical integration



AME





AME

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

ODEs

local

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

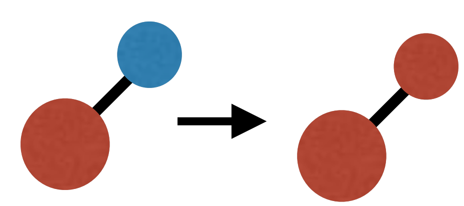
global

weighted average over all degrees

$$x_s(t) = \sum_{\mathbf{m}\in\mathcal{M}} x_{s,\mathbf{m}}(t)$$

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

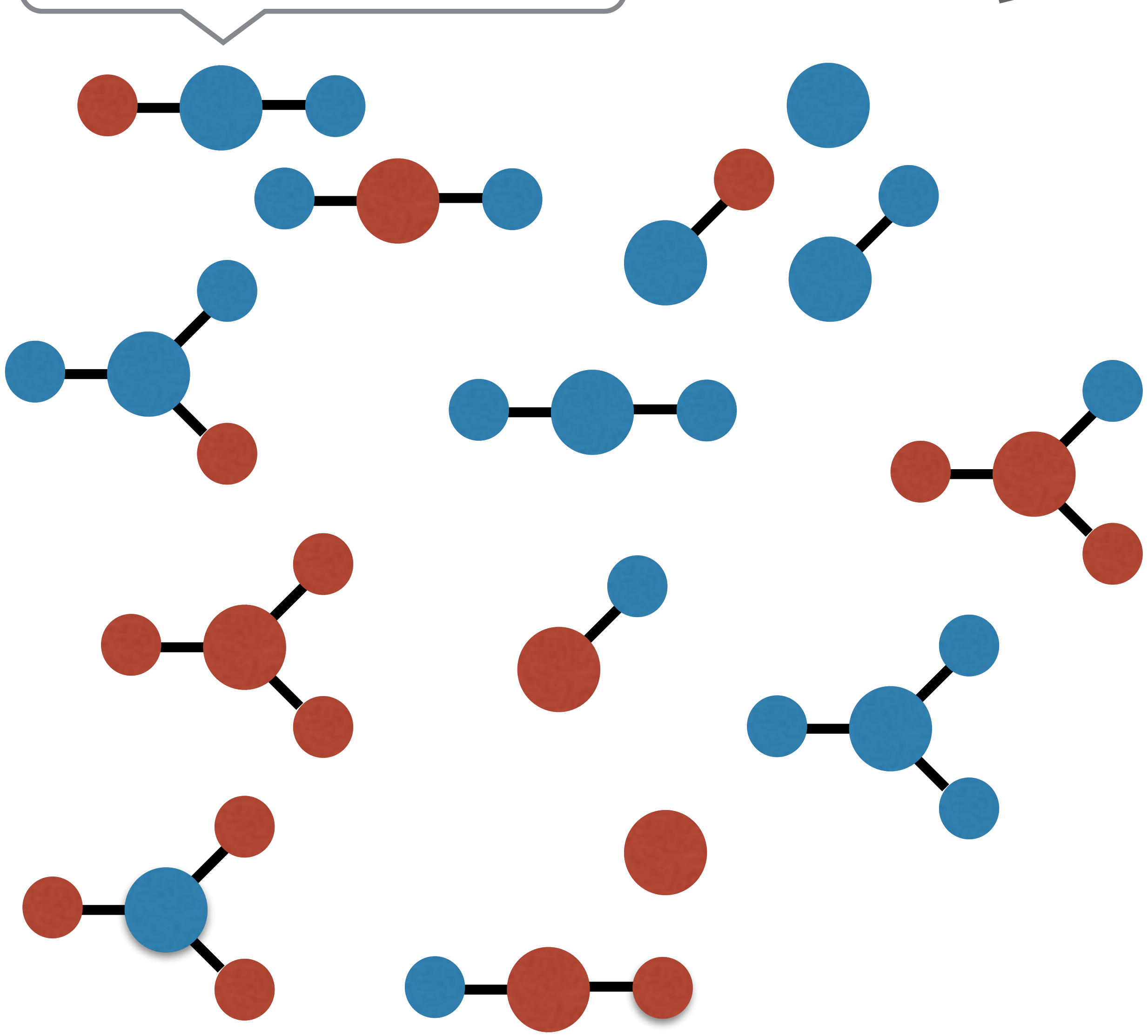
average rates of pairs



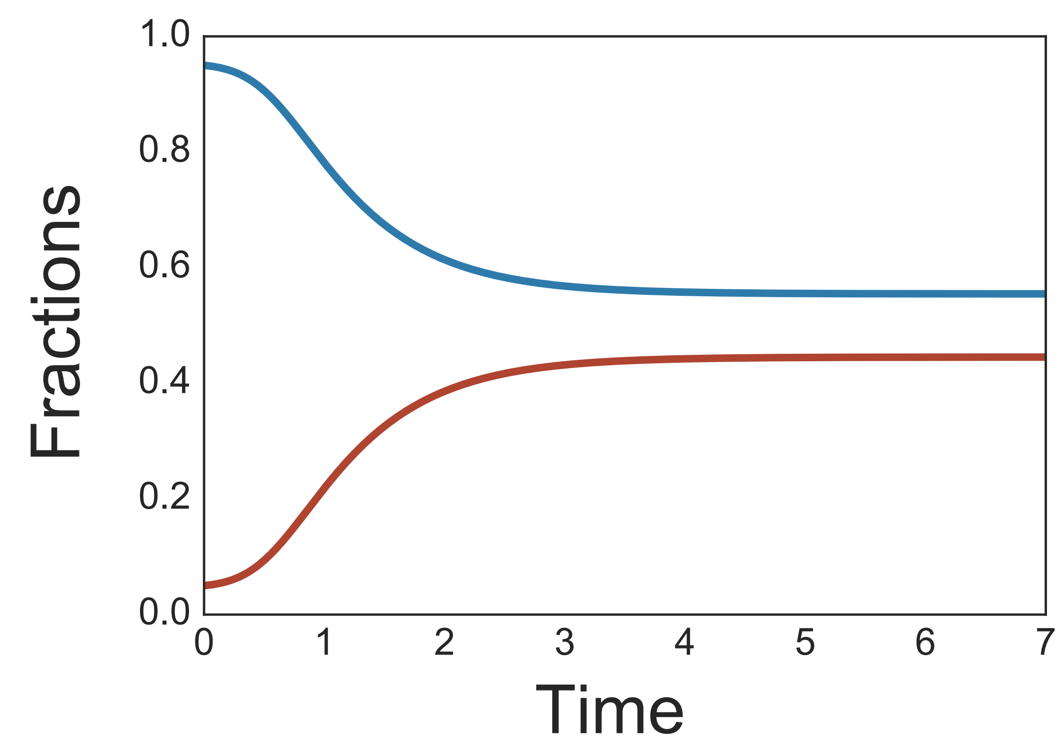
AME

one ODE for each local state and neighbourhood

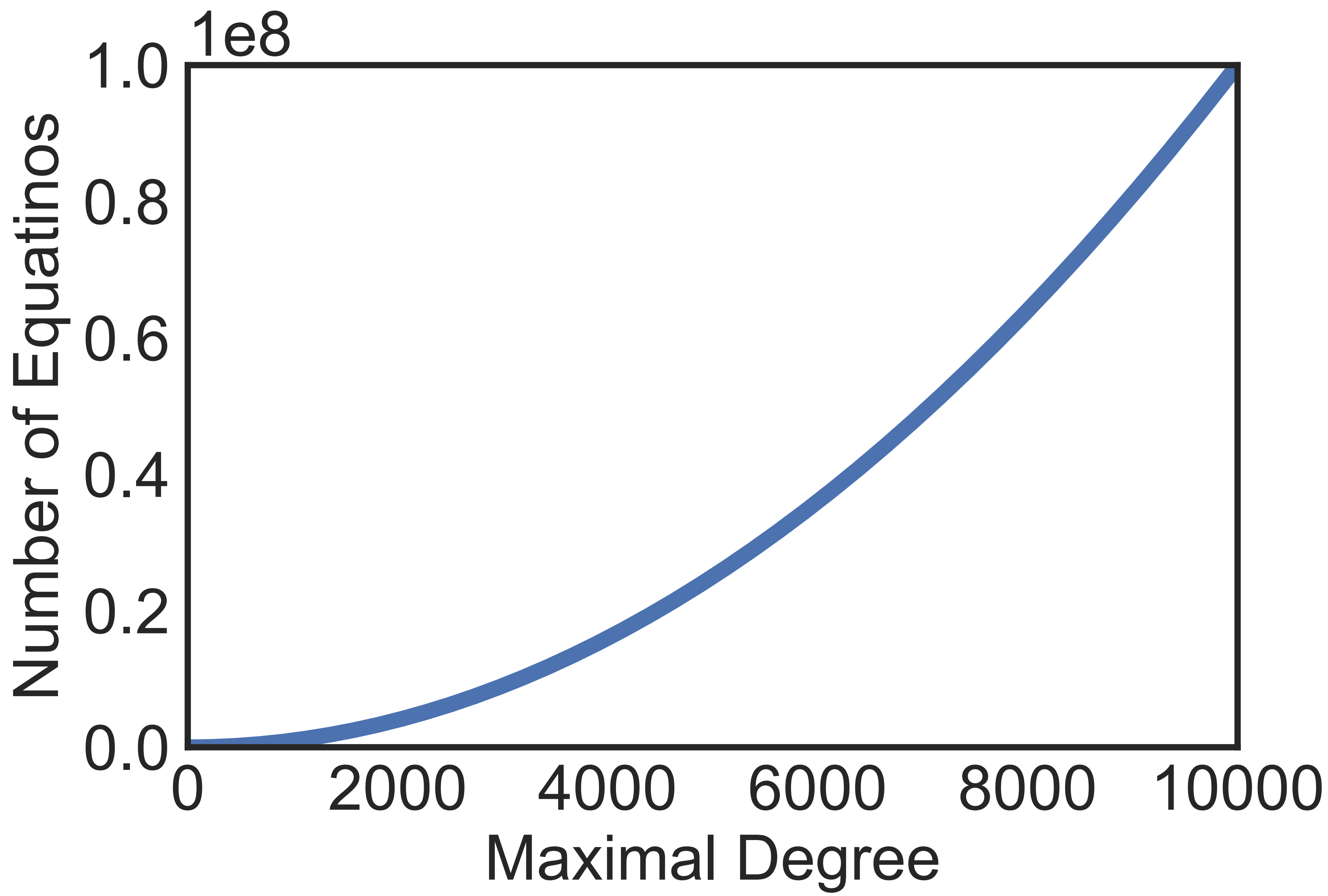
Fraction of susceptible nodes with 1 susceptible and 1 infected neighbour



Numerical integration

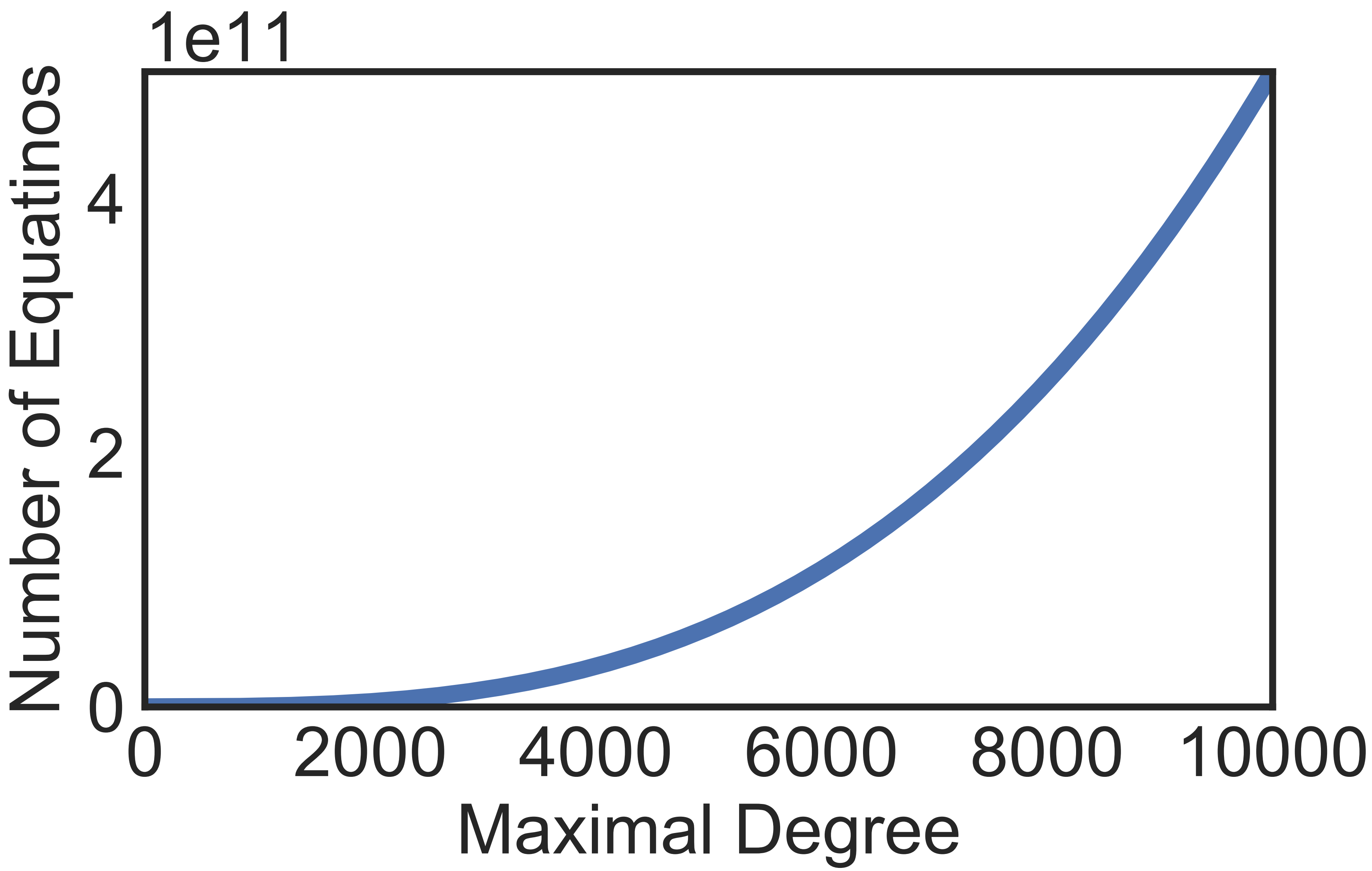


Number of Equations



2 node-states (SIS)

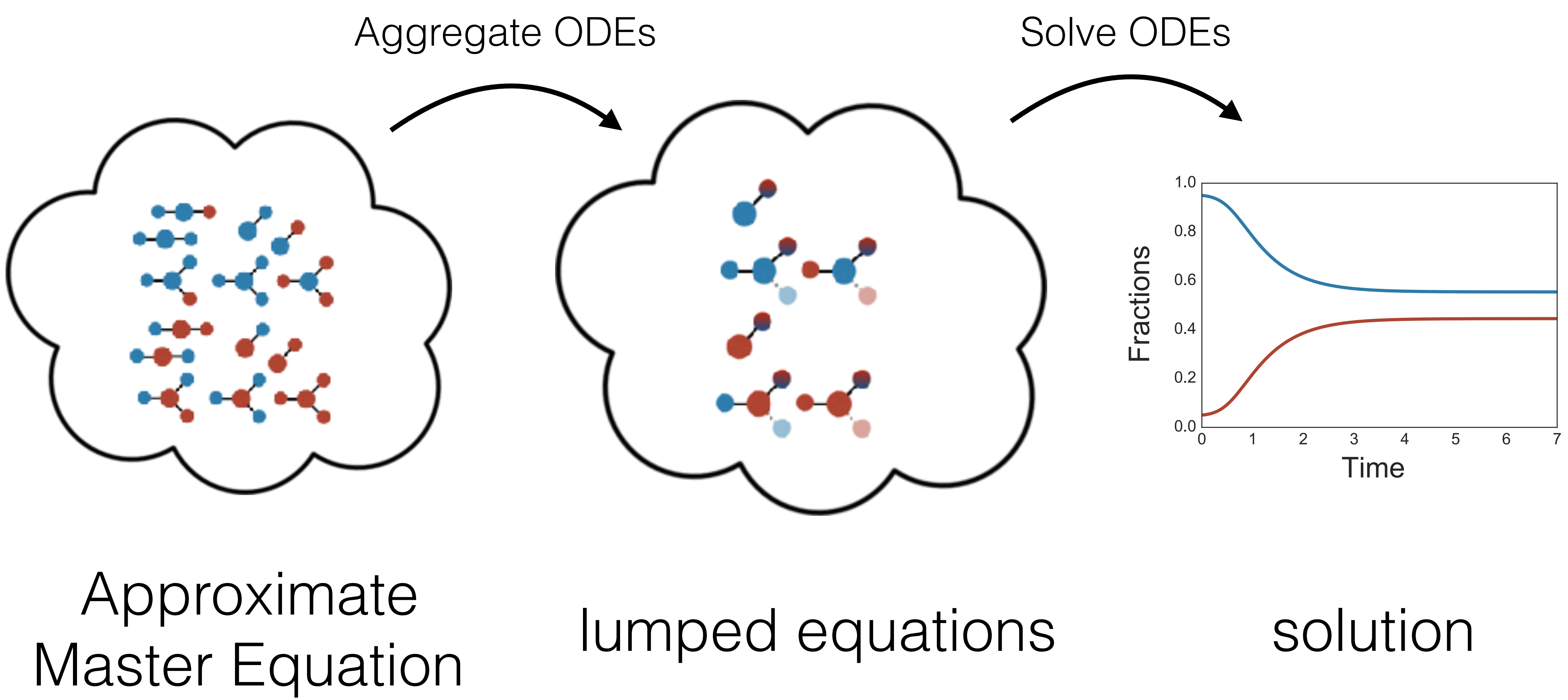
Number of Equations



3 node-states (SIR)

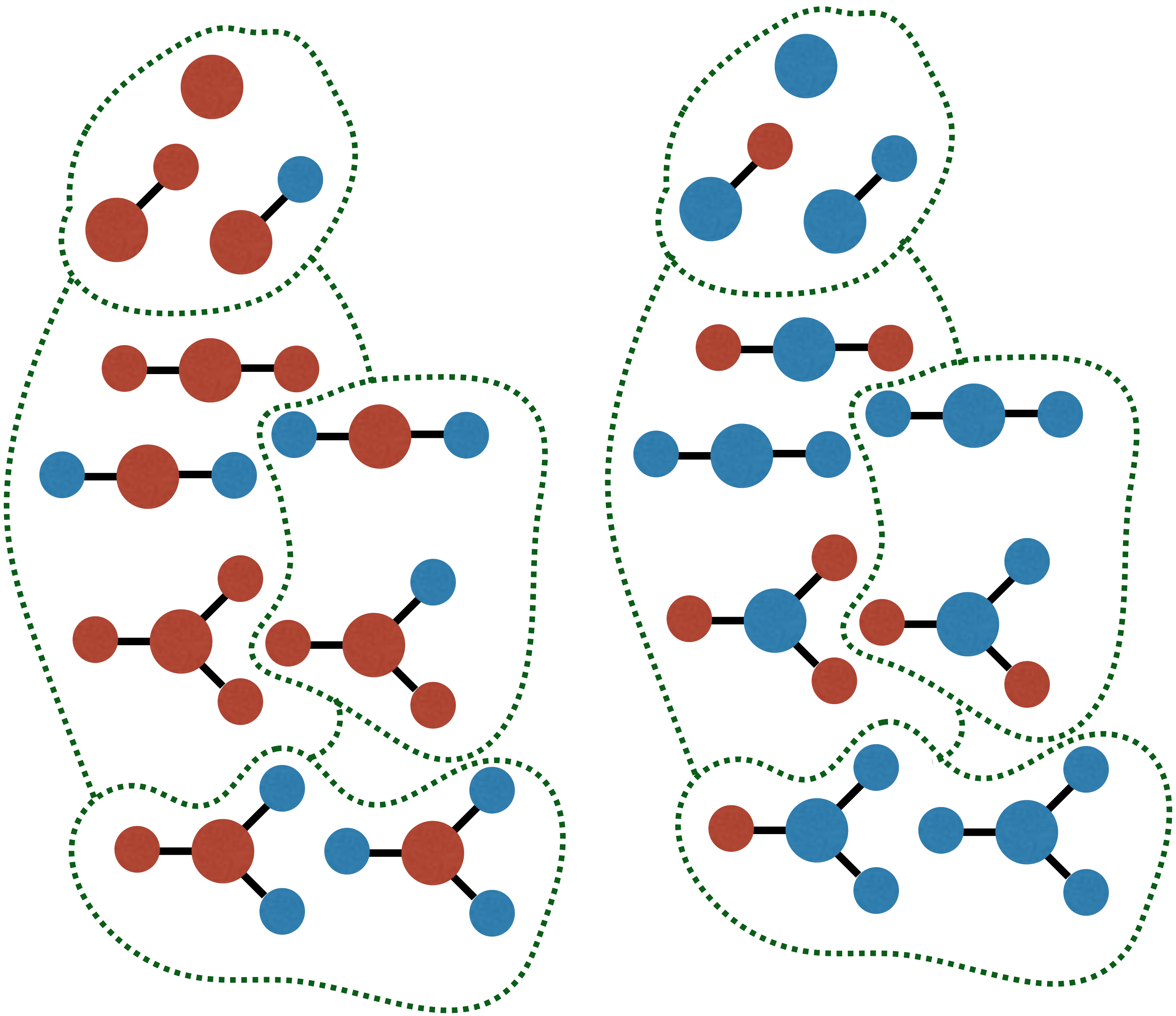
Lumping the Approximate Master Equation for Multistate Processes on Complex Networks

Lumping


Aggregate ODEs
Solve ODEs
Fractions
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
4
5
6
7
Time
Approximate
Master Equation
lumped equations
solution

Partitioning

Solve one equation for each partition



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

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

average

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

substitute

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

AME

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

Node changes

ODEs

$$
\begin{aligned}
\frac{\partial z_{s,C}}{\partial t} = & \sum_{(s',f,s)\in R^{s+}} z_{s',C}\Big(\sum_{\mathbf{m}\in C} w_{C,k_{\mathbf{m}}} f(\mathbf{m})\Big) \\
& - \sum_{(s,f,s')\in R^{s-}} z_{s,C}\Big(\sum_{\mathbf{m}\in C} w_{C,k_{\mathbf{m}}} f(\mathbf{m})\Big) \\
& + \sum_{\substack{(s_1,s_2)\in \mathcal{S}^2 \\ s_1\neq s_2}} \beta_{\mathcal{L}}^{ss_1\to ss_2}\Big(\sum_{\mathbf{m}\in C} w_{C(\mathbf{m}^{\{s_1^+,s_2^-\}}),k_{\mathbf{m}}} z_{s,C(\mathbf{m}^{\{s_1^+,s_2^-\}})} \mathbf{m}^{\{s_1^+,s_2^-\}}[s_1]\Big) \\
& - \sum_{\substack{(s_1,s_2)\in \mathcal{S}^2 \\ s_1\neq s_2}} \beta_{\mathcal{L}}^{ss_1\to ss_2} z_{s,C}\Big(\sum_{\mathbf{m}\in C} w_{C,k_{\mathbf{m}}} \mathbf{m}[s_1]\Big) ,
\end{aligned}
$$

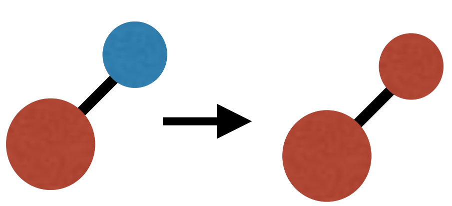
Neighbourhood changes

weighted average over all degrees

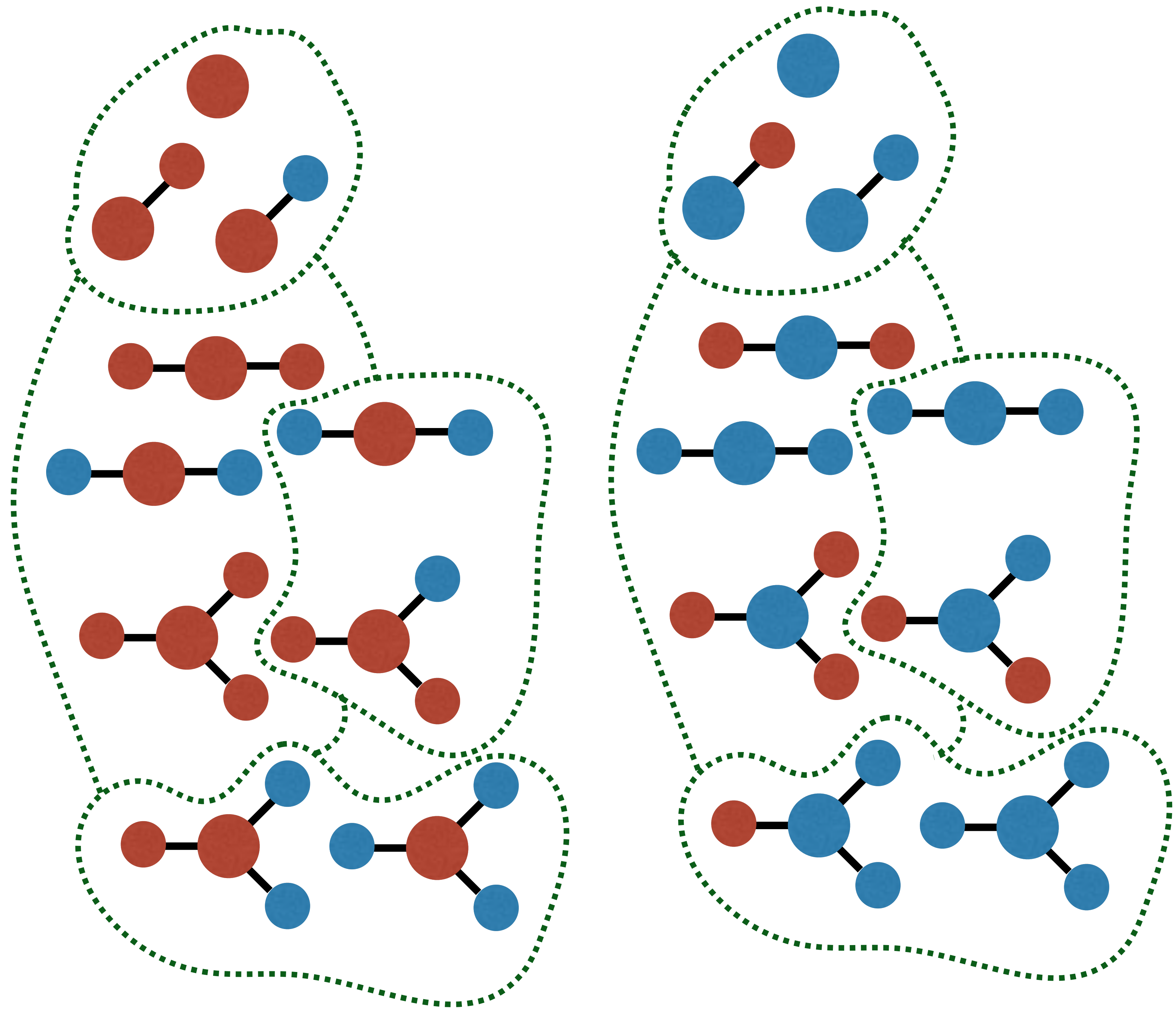
$$
x_s(t) \approx \sum_{C\in\mathcal{C}} z_{s,C}(t)
$$

$$
\beta_{\mathcal{L}}^{ss_1\to ss_2} = \frac{\sum_{C\in\mathcal{C}} z_{s_1,C} \sum_{(s_1,f,s_2)\in R^{s_1\to s_2}} \sum_{\mathbf{m}\in C} f(\mathbf{m}) w_{C,k_{\mathbf{m}}} \mathbf{m}[s]}{\sum_{C\in\mathcal{C}} z_{s_1,C} \sum_{\mathbf{m}\in C} w_{C,k_{\mathbf{m}}} \mathbf{m}[s]}
$$

average rates of pairs

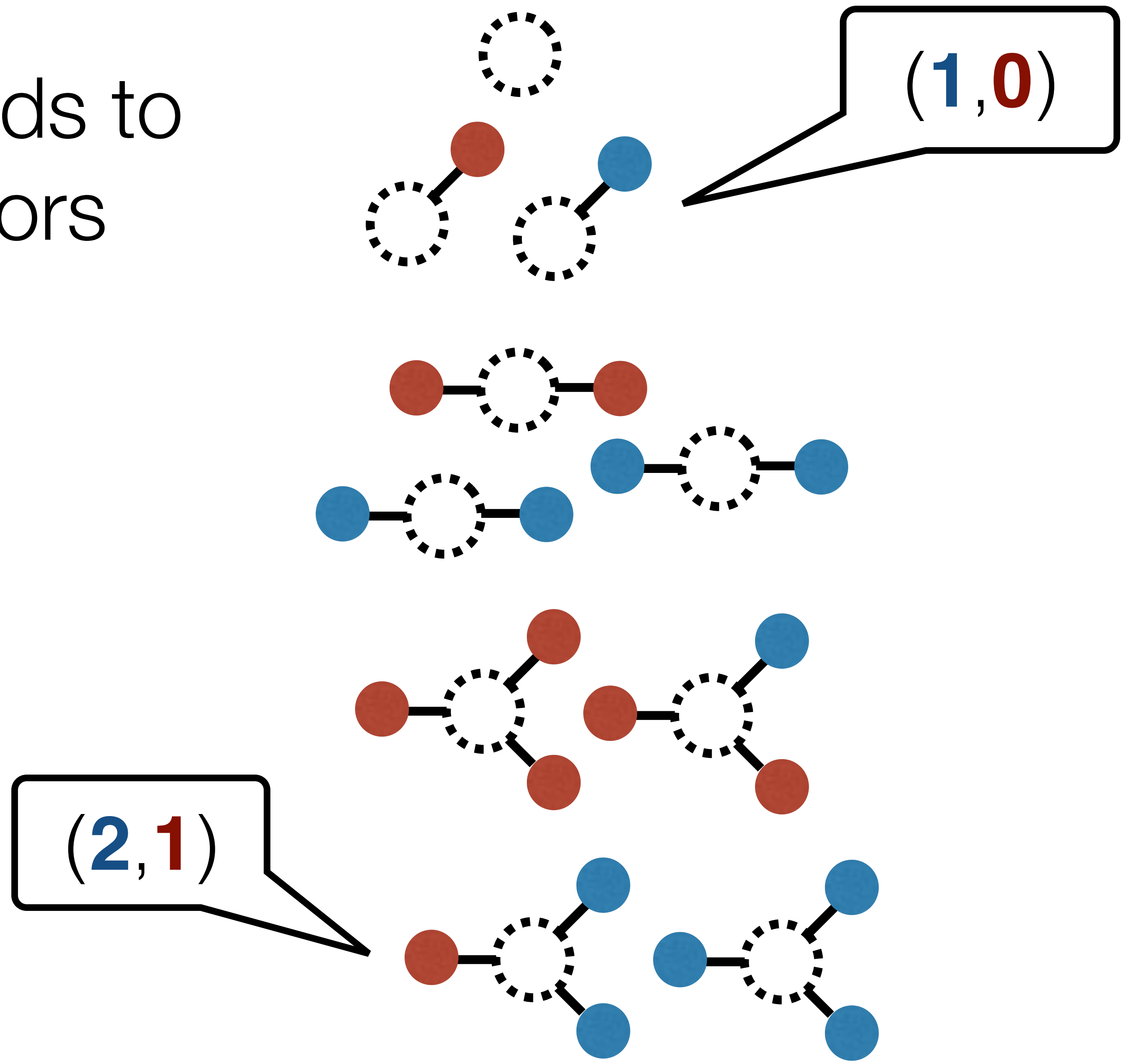


Partitioning



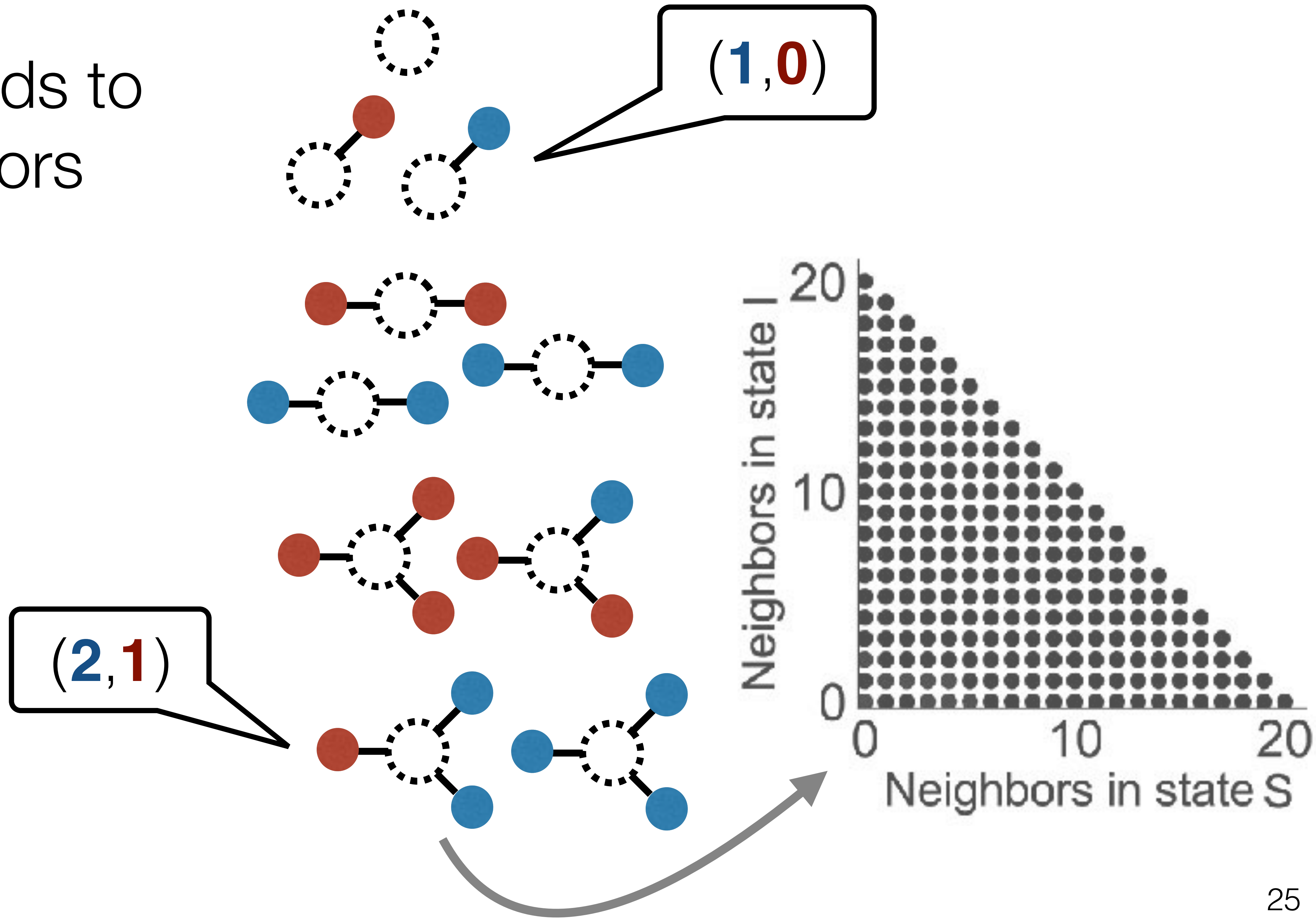
Partitioning

Map neighbourhoods to integer vectors

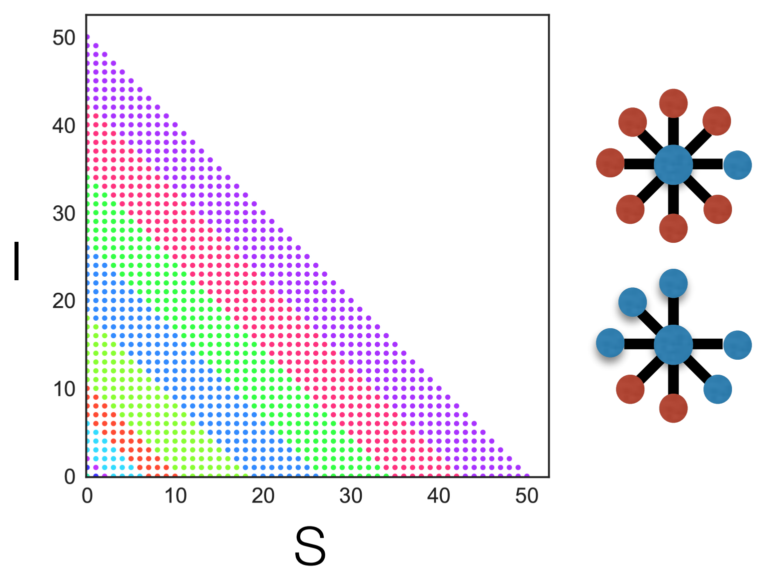


Partitioning

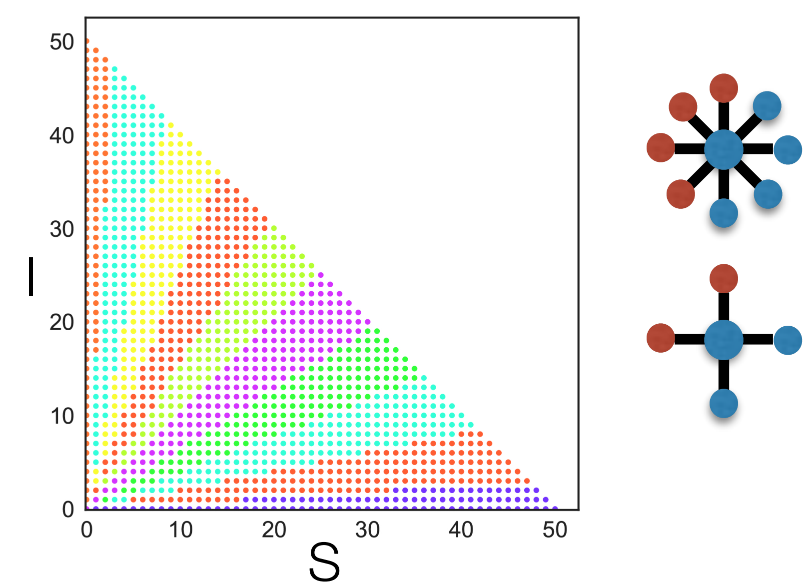
Map neighbourhoods to integer vectors



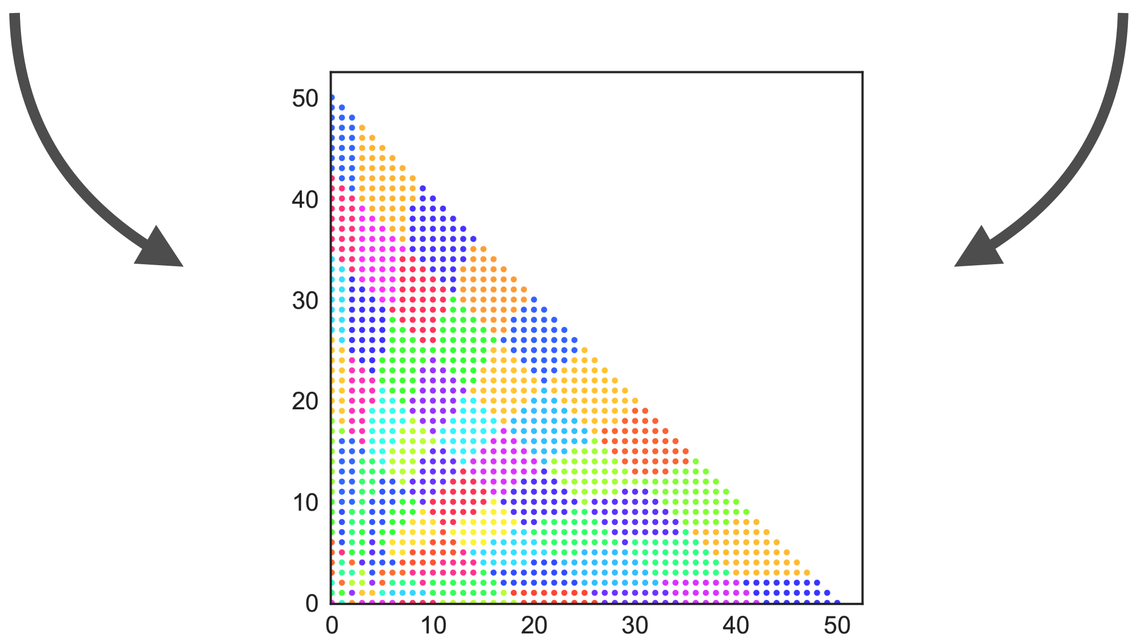
Partitioning



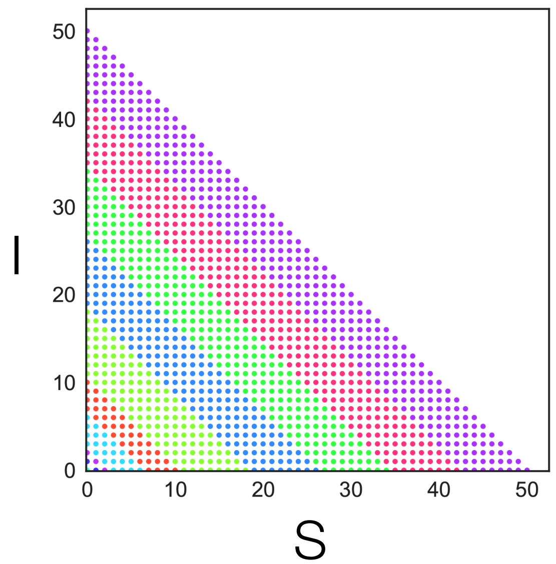
Cluster degrees k based on degree-distribution $P(k)$

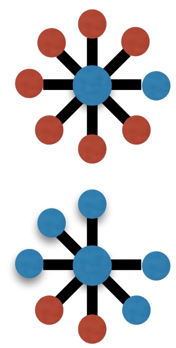


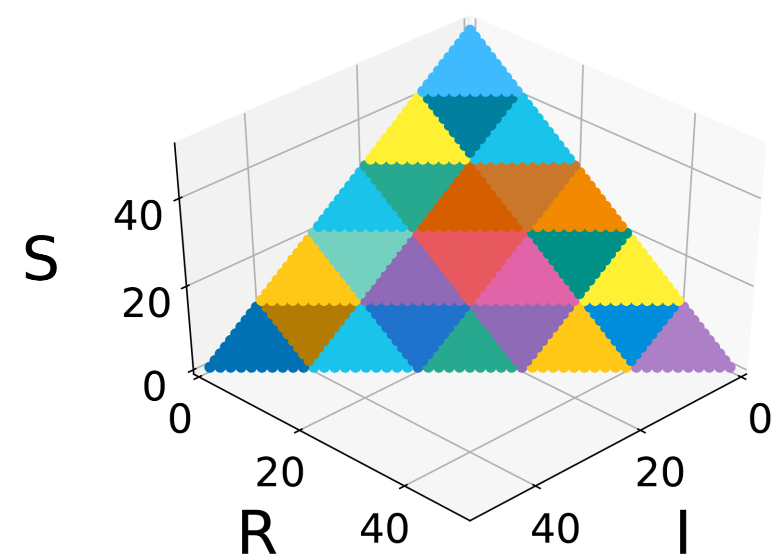
Cluster neighbourhood vectors based on proportionality

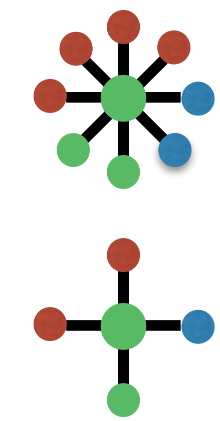


Partitioning 3D



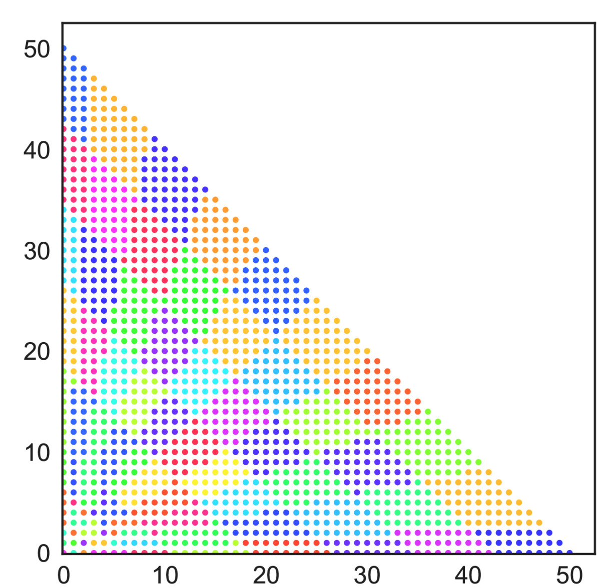






Cluster degrees k based on degree-distribution $P(k)$

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

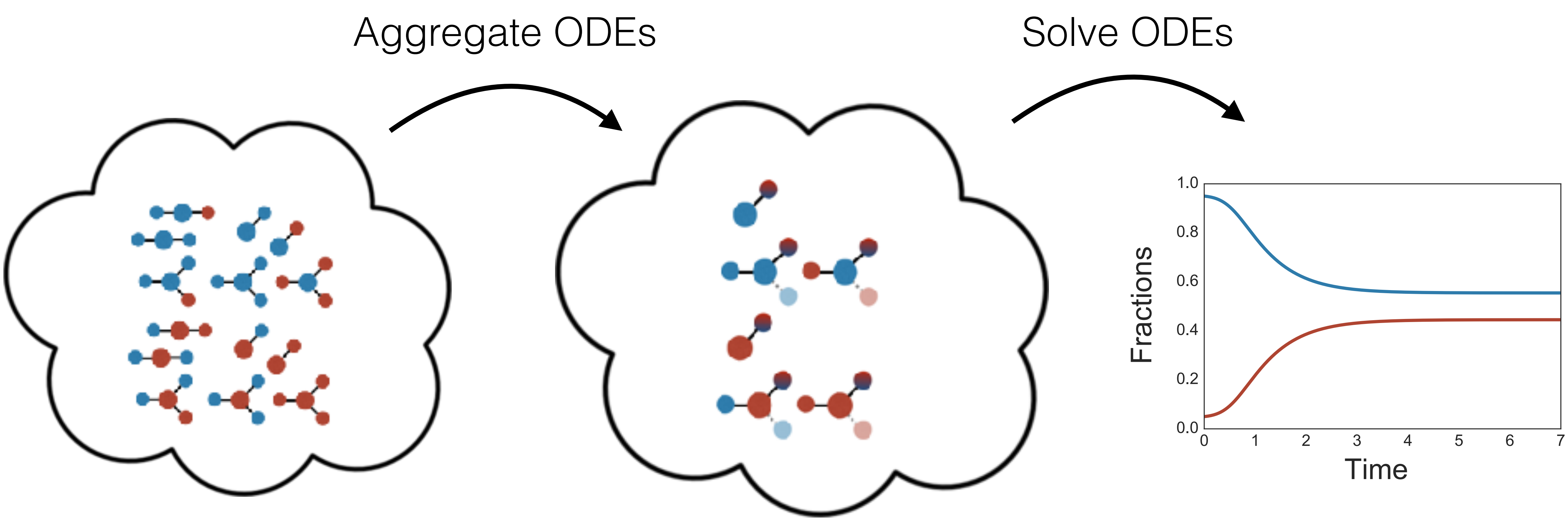
How Many Partitions?

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 overall lumping error

Lumping

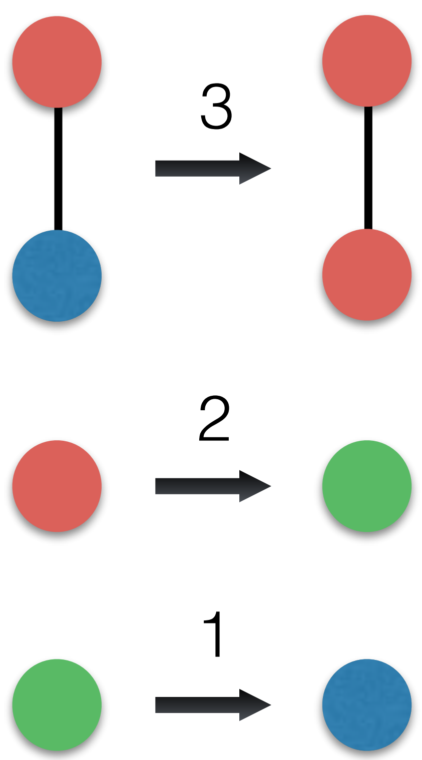


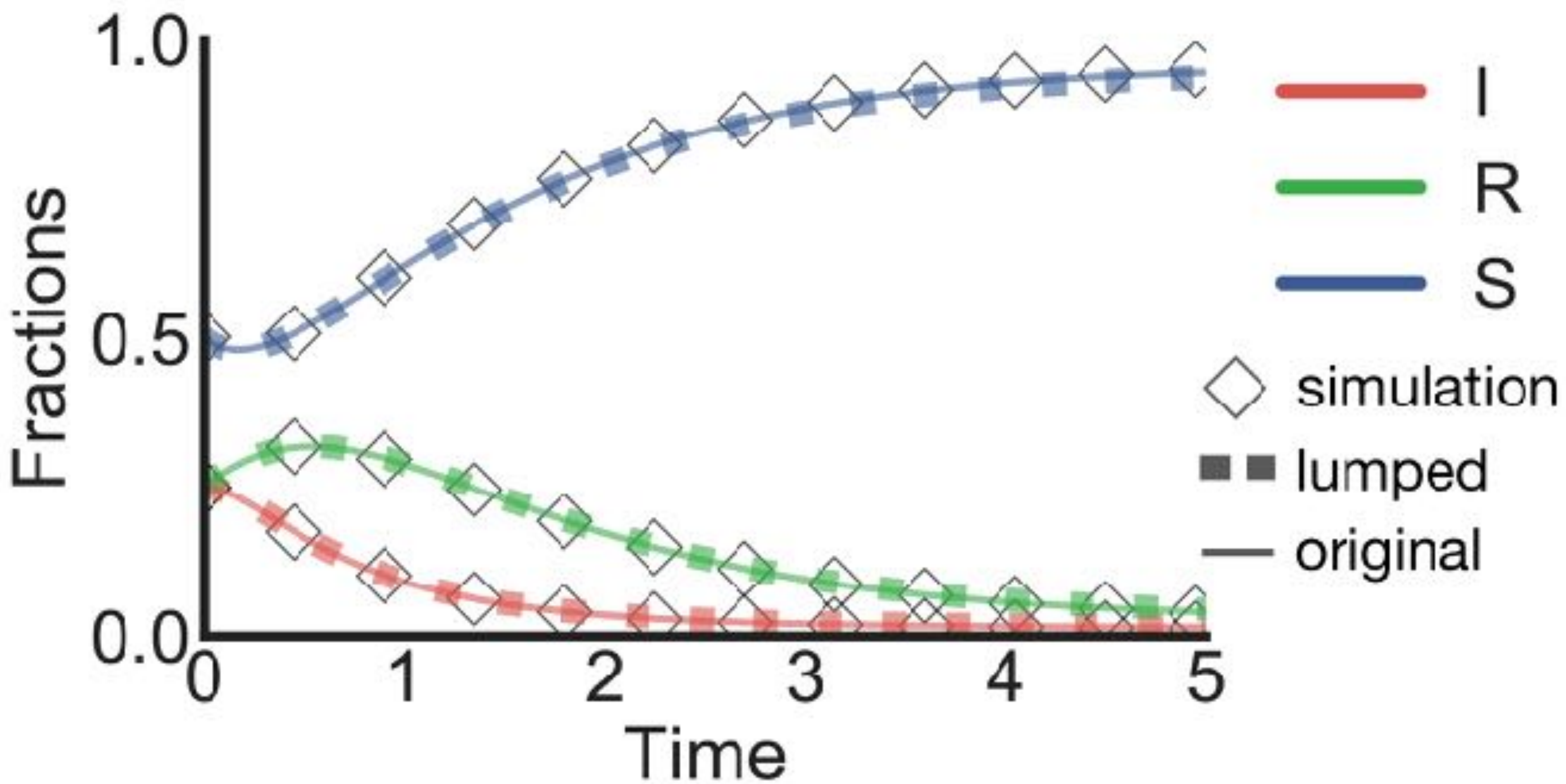
Approximate Master Equation

lumped equations

solution

Results SIR

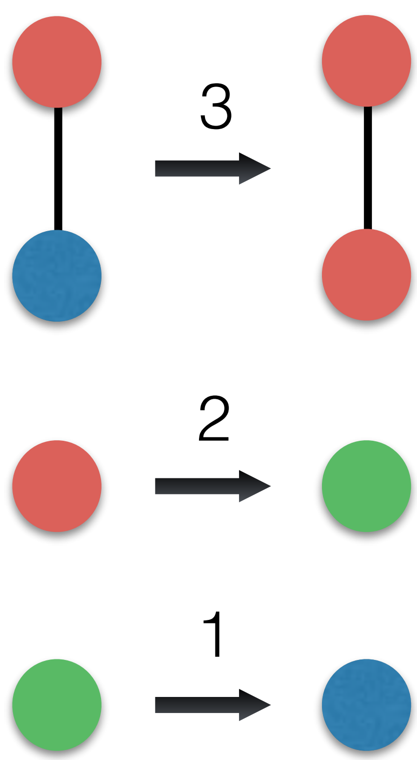


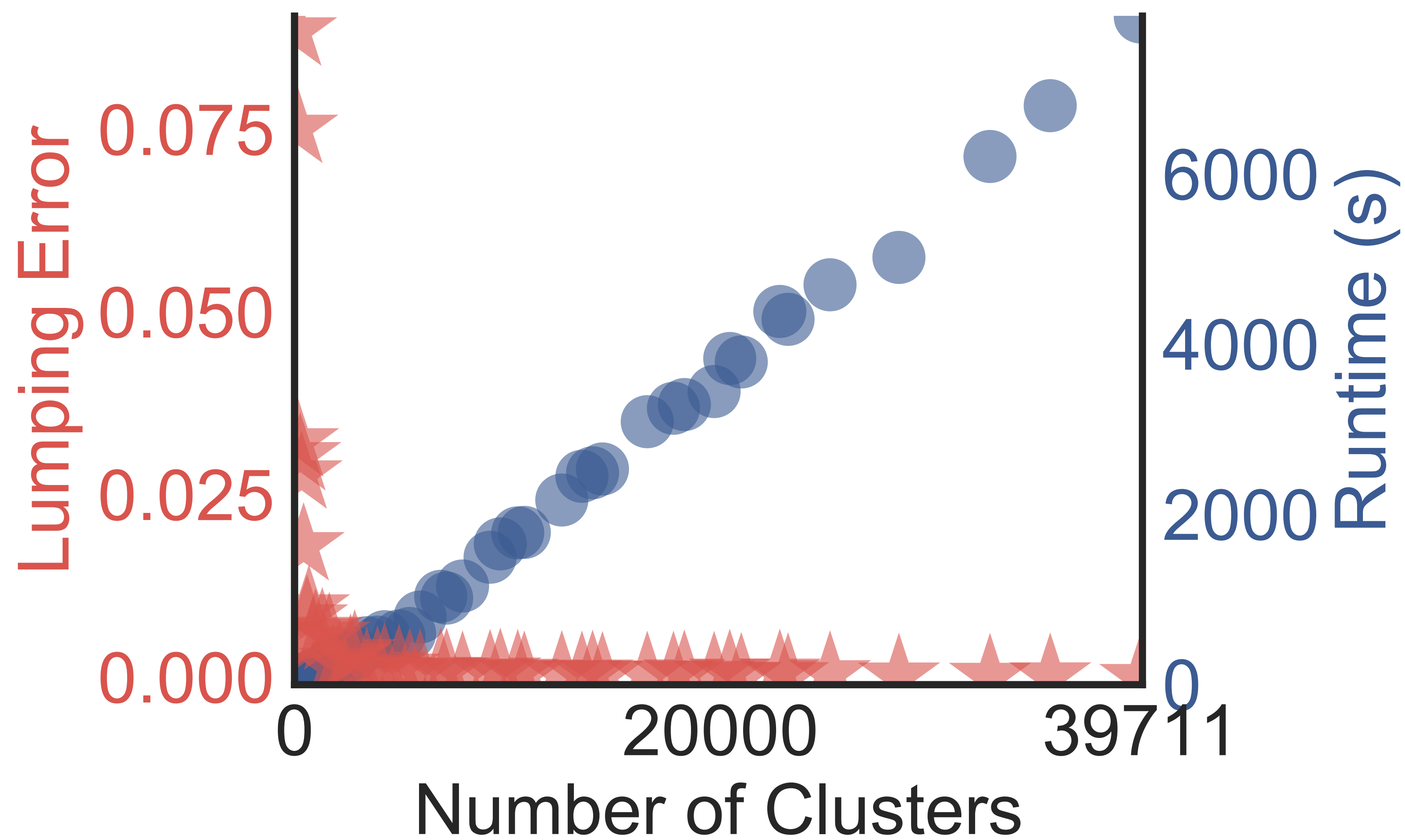


5 373 ODEs vs 119 133 ODEs

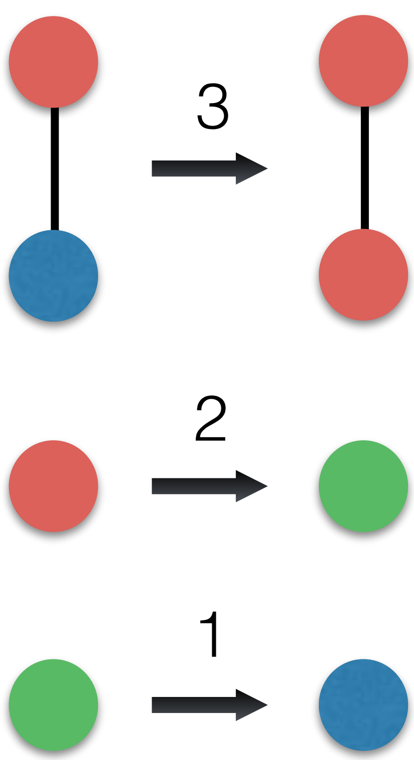
Truncated power-law with $\gamma = 2.5$ and $k_{\max} = 60$

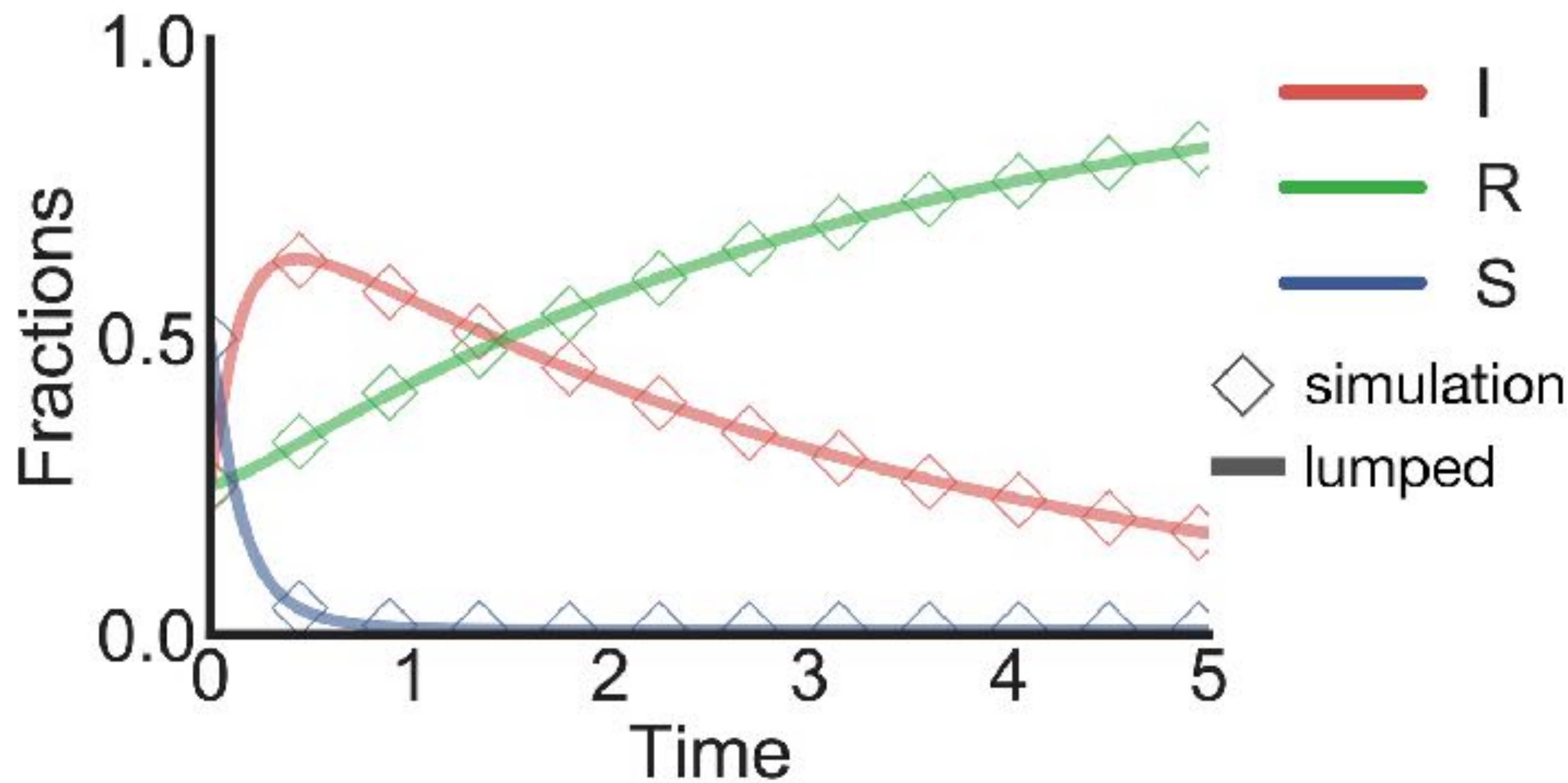
Results SIR


3
2
1


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

Results SIR

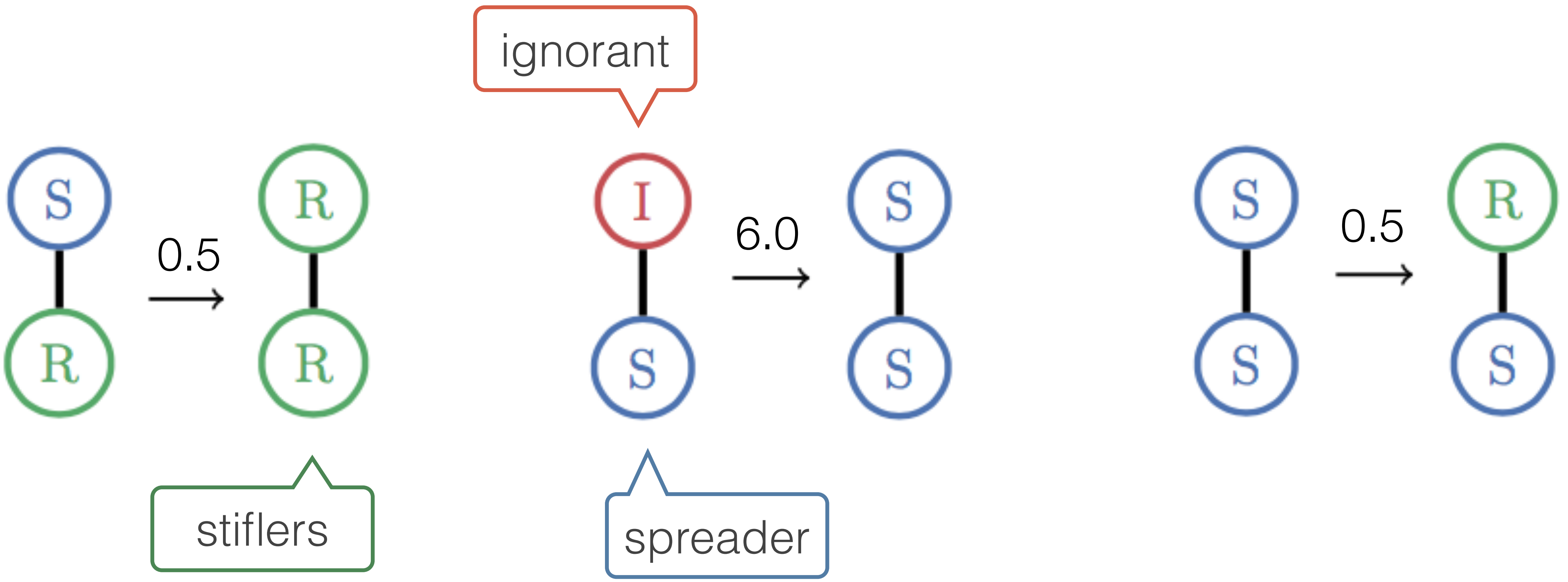




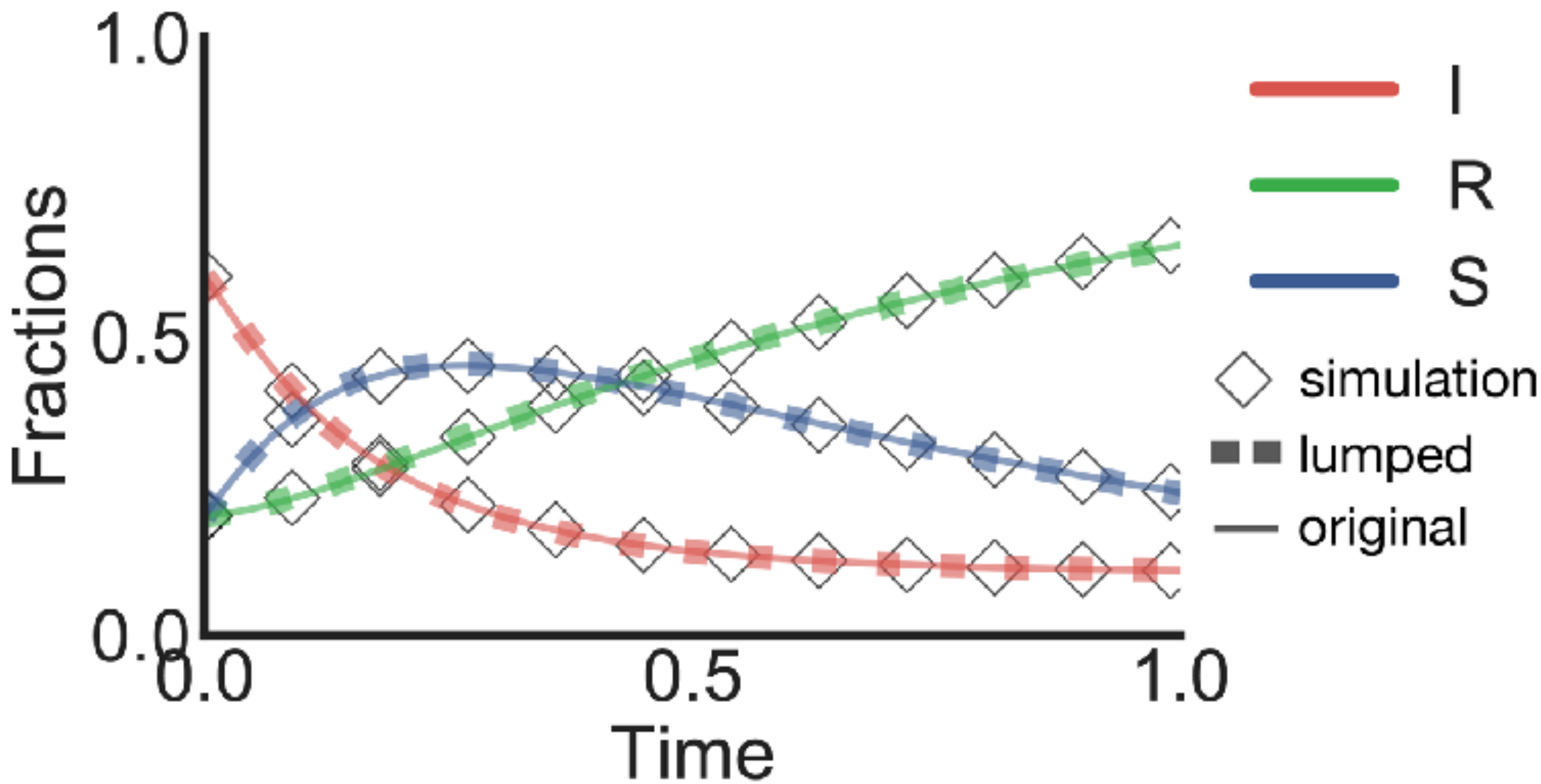
25K ODEs vs 63M ODEs

Truncated power-law with $\gamma = 2.5$ and $k_{\max} = 600$

Rumor Spreading



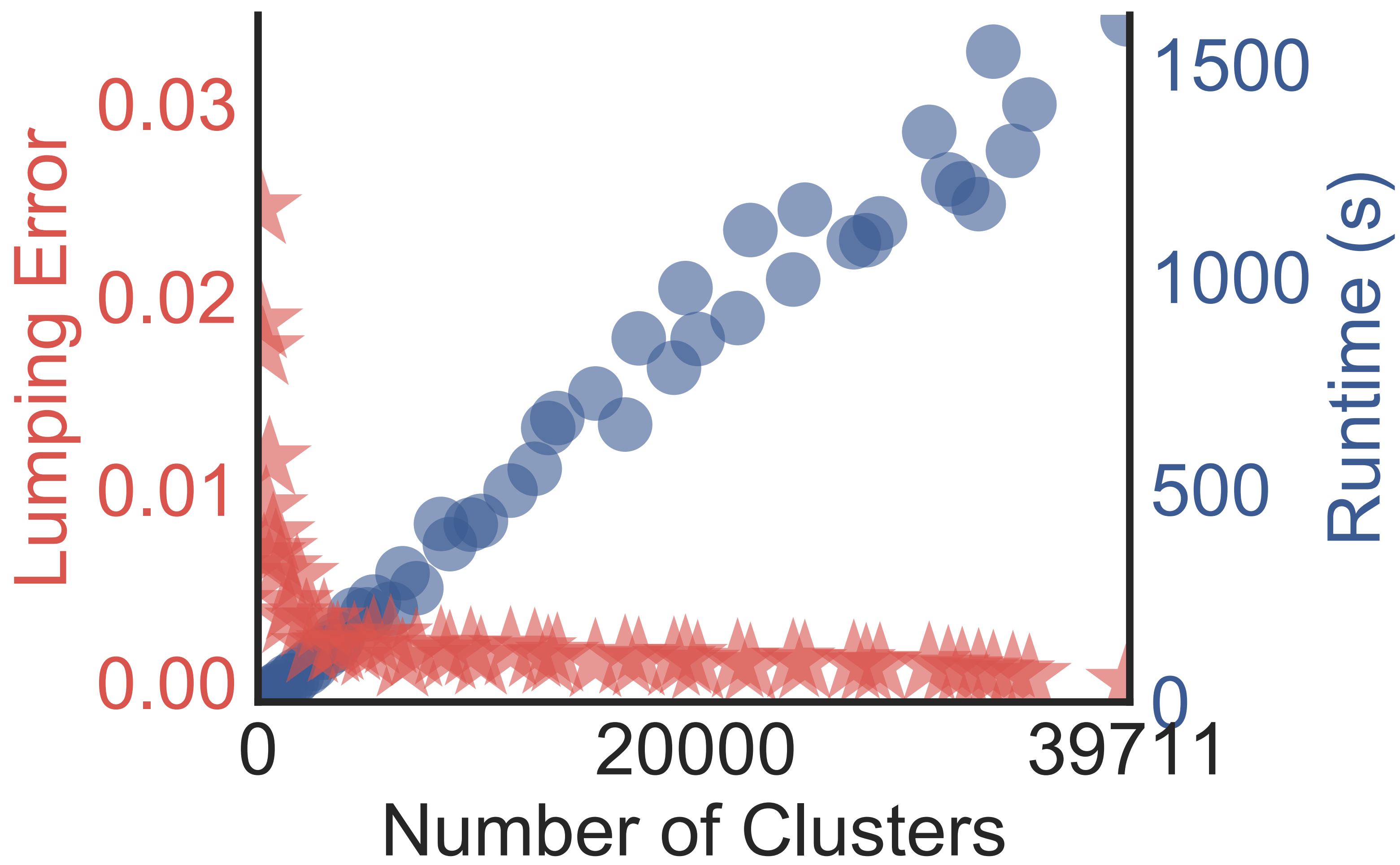
Rumor Spreading



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

local

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

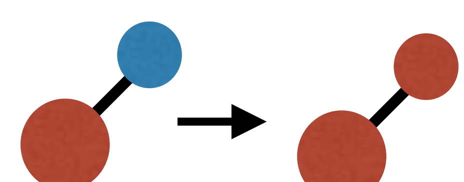
global

weighted average over all degrees

$$x_s(t) = \sum_{\mathbf{m}\in\mathcal{M}} x_{s,\mathbf{m}}(t)$$

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

average rates of pairs



Lumping the AME

Original

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

Lumped

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

Substitution