

Goal

To model **stochastic processes** taking place on **complex networks**.

Analytical methods are sought. Although this work focuses on epidemiological dynamics, different phenomena may also be considered.

Recipe

- Specify a way to represent the characteristics of the system (which includes the network topology) with a **state vector** $\mathbf{x}(t)$. This mapping does not need to be a bijection (*partial representation*).
- Specify all the **events** that may occur with **shift vectors** \mathbf{r}^j . An event of type j takes the system from state \mathbf{x} to state $\mathbf{x} + \mathbf{r}^j$.
- Obtain **inference terms** $P(\mathbf{y}|\mathbf{x})$, when required. Some events may be affected by unknown quantities $\mathbf{y}(t)$ that have to be inferred from the available information $\mathbf{x}(t)$. (Recall that $\mathbf{x}(t)$ may not completely encode the status of the system.)
- Specify the **rates of occurrence** $q_j^+(\mathbf{x})$ for all events. An event of type j has probability $q_j^+(\mathbf{x})dt$ to occur during $[t, t+dt)$. When events may occur in the backward direction (*i.e.* \mathbf{x} becomes $\mathbf{x} - \mathbf{r}^j$), also specify $q_j^-(\mathbf{x})$. Use inference terms when required.

Birth-death Markov process

The **probability distribution** $P(\mathbf{x}|t)$ is governed by the **master equation**

$$\frac{dP(\mathbf{x}|t)}{dt} = \sum_j \left[q_j^+(\mathbf{x} - \mathbf{r}^j)P(\mathbf{x} - \mathbf{r}^j|t) - q_j^+(\mathbf{x})P(\mathbf{x}|t) + q_j^-(\mathbf{x} + \mathbf{r}^j)P(\mathbf{x} + \mathbf{r}^j|t) - q_j^-(\mathbf{x})P(\mathbf{x}|t) \right]$$

For large systems, using a **continuous approximation** for the state vector leads to the **stochastic differential equation** (Fokker-Planck equation)

$$\frac{\partial P(\mathbf{x}|t)}{\partial t} = - \sum_i \frac{\partial}{\partial x_i} [a_i(\mathbf{x})P(\mathbf{x}|t)] + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} [B_{ij}(\mathbf{x})P(\mathbf{x}|t)]$$

$$a_i(\mathbf{x}) = \sum_j r_i^j [q_j^+(\mathbf{x}) - q_j^-(\mathbf{x})] \quad B_{ij}(\mathbf{x}) = \sum_j r_i^j r_j^i [q_j^+(\mathbf{x}) + q_j^-(\mathbf{x})]$$

When the main contribution of $P(\mathbf{x}|t)$ is located around its mean value, using a **mean-field approximation** provides the **ordinary differential equation**

$$\frac{d}{dt} \langle \mathbf{x}(t) \rangle = \mathbf{a}(\langle \mathbf{x}(t) \rangle)$$

A **Gaussian approximation** determines how $P(\mathbf{x}|t)$ behaves around this mean value. Assuming deterministic initial conditions, this behaviour is given by

$$P(\mathbf{x}|t) = \frac{1}{(2\pi)^{d/2} |\hat{C}(t)|^{1/2}} \exp \left(-\frac{1}{2} (\mathbf{x}(t) - \langle \mathbf{x}(t) \rangle)^T \hat{C}(t)^{-1} (\mathbf{x}(t) - \langle \mathbf{x}(t) \rangle) \right)$$

$$\hat{C}(t) = \int_0^t \hat{A}(t, t') \cdot \hat{B}(\langle \mathbf{x}(t'') \rangle) \cdot \hat{A}(t, t')^T dt' \quad \hat{A}(t, t') = \exp \left[\int_{t'}^t \hat{J}_a(\langle \mathbf{x}(t'') \rangle) dt'' \right]$$

where d is the dimensionality of \mathbf{x} , $\hat{C}(t)$ is the covariance matrix of $P(\mathbf{x}|t)$, and $\hat{J}_a(\langle \mathbf{x}(t'') \rangle)$ is the Jacobian matrix of \mathbf{a} evaluated at $\langle \mathbf{x}(t'') \rangle$.

See C. W. Gardiner, *Handbook of Stochastic Methods*, Springer (2004).

Epidemiological dynamics

In simple epidemiological models, there are three possible intrinsic node states.

- **Susceptible** nodes do not have the infection (but could acquire it).
- **Infectious** nodes have the infection and can transmit it.
- **Removed** nodes neither acquire nor transmit the infection.

Changes in node states during $[t, t+dt)$ depend on the specificity of the infection.

SI: S nodes become I with probability proportional to number of I neighbours.

SIS: As in SI, except that I nodes become S with constant probability.

SIR: As in SI, except that I nodes become R with constant probability.

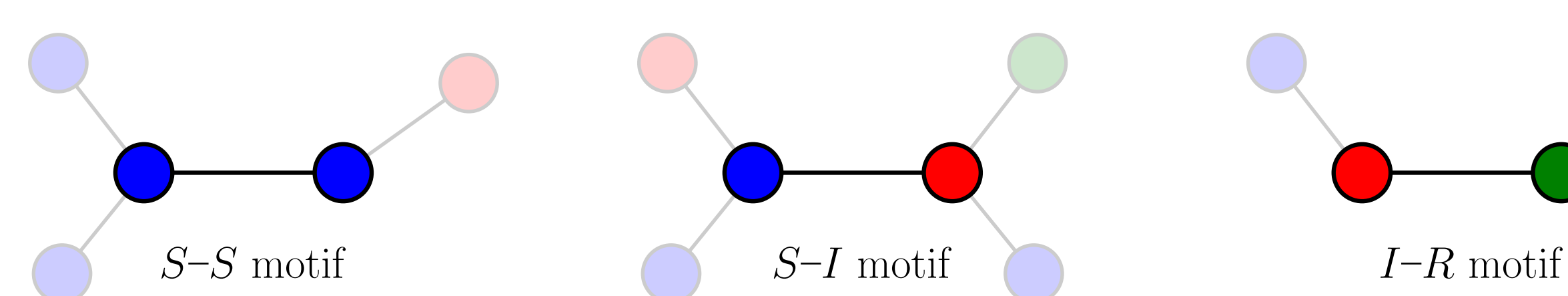
SIRS: As in SIR, except that R nodes become S with constant probability.

These changes correspond to the events of the birth-death process.

Motifs

A vast class of state vectors amounts to an **enumeration** of the network **motifs**.

Pair motifs Perhaps the simplest description of a network system.



A **pair motif** informs about the intrinsic state of the two (linked) nodes forming it; nothing is said about the presence (or absence) of other neighbours.

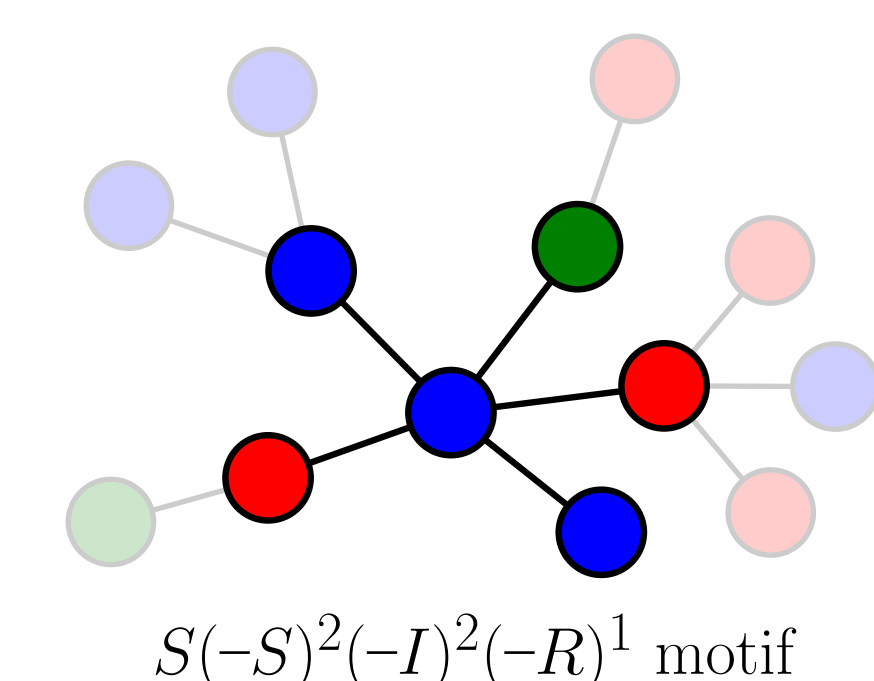
Pair models are usually supplemented by **node motifs** (intrinsic node states).

The state vector may then enumerate the total number of each types of motifs.

Example: $\mathbf{x} = (x_S, x_I, x_R, x_{S-S}, x_{S-I}, x_{S-R}, x_{I-I}, x_{I-R}, x_{R-R})$.

First neighbourhood motifs

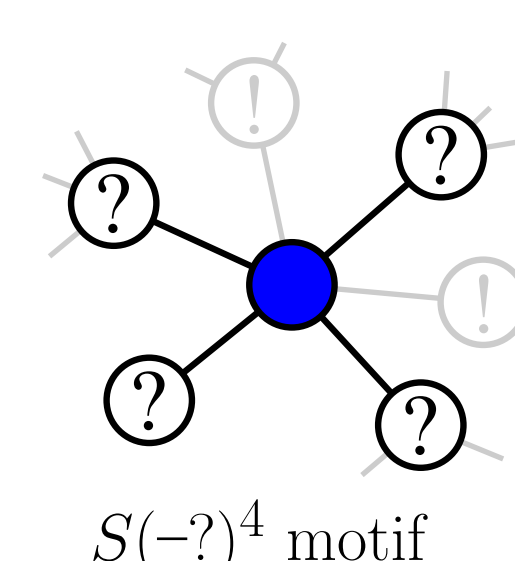
Simple yet powerful.



The intrinsic state of the central node and of each of its neighbours is known. Second neighbours are unspecified, although motifs based on higher order neighbourhood are also possible.

On-the-fly assignment

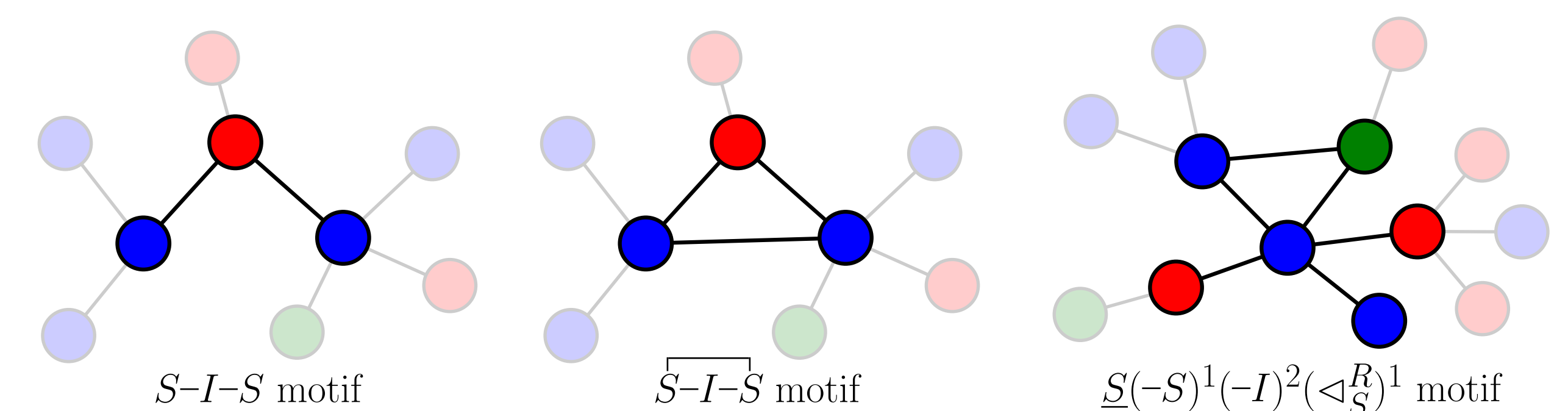
A special case of first neighbourhood.



The central node may have some neighbours **known to be irrelevant** for future evolution (Ⓛ). The total number of other neighbours, of unknown intrinsic state, is specified (Ⓢ). In some cases, SI and SIR dynamics may exactly be represented with state vectors based on such simple motifs.

Short cycles and other motifs

Endless possibilities.



Triple, triangle and **first neighbourhood with triangles** motifs are shown. Motifs for intrinsic link states and directed networks are also possible.

Notation inspired by T. House *et al.*, *Bull. Math. Biol.* **71**, 1693 (2009).

Example: SIS pair model

- **State vector** based on node and pair motifs: $\mathbf{x} = (x_S, x_I, x_{S-S}, x_{S-I}, x_{I-I})$.
- **Events** $\underline{S}(-S)^k(-I)^l \xrightarrow{kl} \underline{I}(-S)^k(-I)^l$ depend on the first neighbourhood of the changing node. The **shift vectors** are $\mathbf{r}^{kl} = (-1, 1, -k, k - l)$.
- **Inference terms** are required since \mathbf{x} only specifies node and pair motifs.
- **Rates of occurrence** depend on these inference terms.

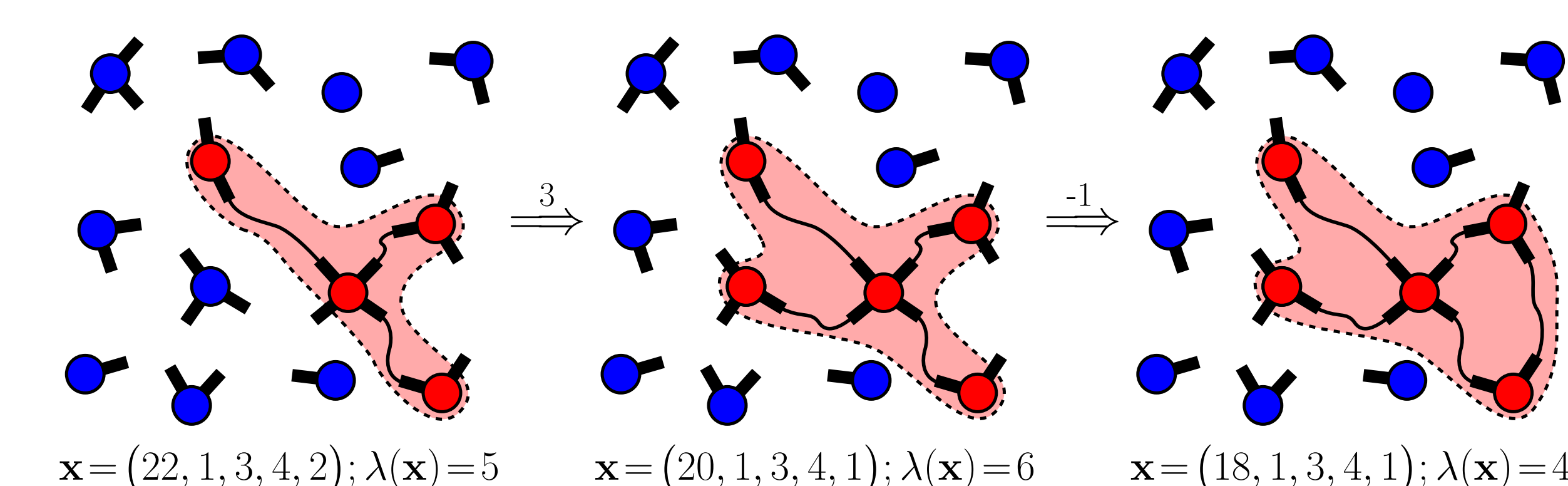
$$q_{kl}^+(\mathbf{x}) = \beta x_{S-I} P(\underline{S}(-S)^k(-I)^l | S \text{ reached through random } S-I \text{ link}, \mathbf{x})$$

$$q_{kl}^-(\mathbf{x}) = \alpha x_I P(\underline{I}(-S)^k(-I)^l | \text{random } I \text{ node}, \mathbf{x})$$

Although presented differently, the Gaussian approximation of this model for regular random graphs is equivalent to C. E. Dangerfield *et al.*, *J. R. Soc. Interface* **6**, 761 (2009).

Example: SI on-the-fly assignment

For a configuration model network allowing for repeated links and self-loops, the following representation **exactly** corresponds to the full network model.

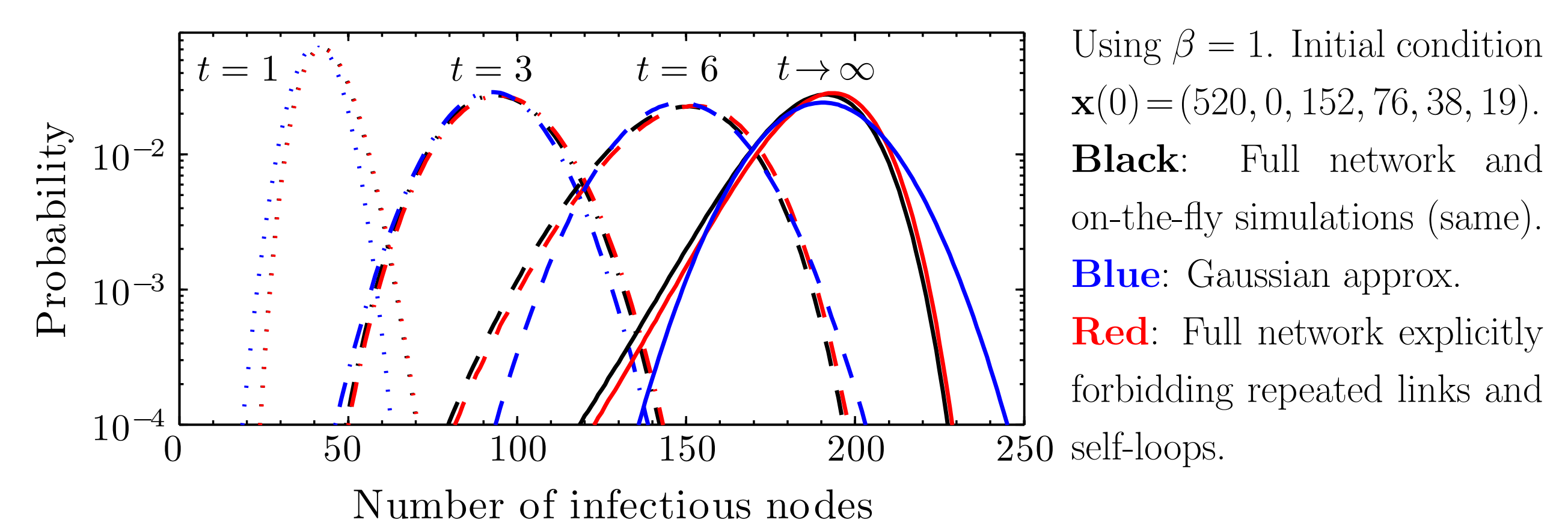


• **State vector** $\mathbf{x} = (x_{-1}, x_0, x_1, x_2, x_3, \dots)$ where x_{-1} is the total number of unassigned stubs and the other x_k are the number of $S(-?)^k$ motifs. The total number of stubs belonging to infectious nodes is $\lambda(\mathbf{x}) = x_{-1} - \sum_{k=0}^{\infty} kx_k$.

• **Events** occur when an I node's stub form a link with another stub. The event is of type $j = -1$ if the other stub belongs to any I node and of type $j \geq 0$ if it belongs to a S node with j unassigned stubs ($r_i^j = \begin{cases} -2 & \text{if } i = -1 \\ -\delta_{ij} & \text{if } i \geq 0 \end{cases}$).

• **Inference terms** are trivial in this case.

• **Rates of occurrence** $q_{-1}^+(\mathbf{x}) = \beta \lambda(\mathbf{x}) \frac{\lambda(\mathbf{x}) - 1}{x_{-1} - 1}$ and $q_j^+(\mathbf{x}) = \beta \lambda(\mathbf{x}) \frac{j x_j}{x_{-1} - 1}$ ($j \geq 0$).



Using $\beta = 1$. Initial condition $\mathbf{x}(0) = (520, 0, 152, 76, 38, 19)$.

Black: Full network and on-the-fly simulations (same).

Blue: Gaussian approx.

Red: Full network explicitly forbidding repeated links and self-loops.

Further examples

The following mean-field models are compatible with the philosophy presented here. As such, Gaussian approximations could be obtained for these models.

Adaptive networks

V. Marceau *et al.* *PRE* **82**, 036116 (2010).

Based on first neighbourhood motifs. A special SIS model where S nodes may disconnect from their I neighbours and reconnect to another S node instead.

Community structure

L. Hébert-Dufresne *et al.* *PRE* **82**, 036115 (2010).

SIS model tracking both first neighbourhood and clique motifs in the state vector.

Interacting epidemics

V. Marceau *et al.* *arXiv*:1103.4059.

On-the-fly model for two interacting SIR dynamics (total 9 intrinsic node states).

Two overlaying networks are considered using intrinsic link states.

High complexity
Usually non-solvable
Better correspondence with reality

Lower complexity
Easier to solve
Possible discrepancies

Real-world system	Full network model	Representation approx.	Continuity approx.	Gaussian approx.	Mean-field approx.
	Master equation (huge state vector)	Master equation (manageable state vector)	Stochastic differential eq.	Ordinary differential eq. + two integrals	Ordinary differential eq.