Haven't I seen you before? Accounting for partnership duration in epidemic models

Joel C. Miller

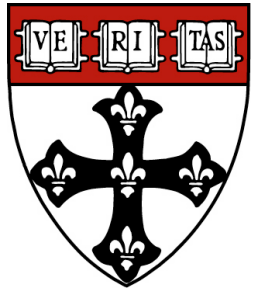
RAPIDD; Harvard School of Public Health

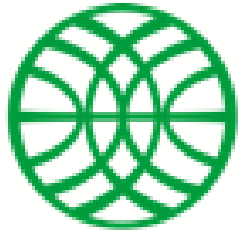
16 Aug 2011

Acknowledgments

Erik Volz (U Michigan)

CENTER *for* COMMUNICABLE DISEASE DYNAMICS





FOGARTY

Standard Mass Action SIR model

Population divided into Susceptible, Infected, and Recovered individuals, $S + I + R = 1$.

- "Mass Action" mixing
  - Behavior is homogeneous: Everyone has exactly k randomly chosen contacts.
  - Contacts are fleeting: At each moment in time, those k contacts are with new people.
- Infection is transmitted at rate β per contact.
- Recovery occurs at rate γ.

This has a simple graphical interpretation as a flow diagram.

The equations can be directly written down as

$$\dot{S} = -\hat{\beta}IS$$
$$\dot{I} = \hat{\beta}IS - \gamma I$$
$$\dot{R} = \gamma I$$

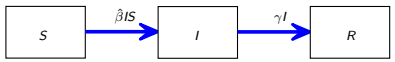
where $\hat{\beta} = \beta k$

The '$\cdot$' means: "rate of change in time".

Why use the Mass Action model?

When there is variation in contact rates or partnerships have duration, mass action model assumptions are false. Why use them?

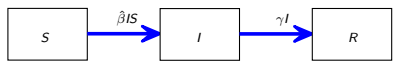
- Simple equations
- $+$ Simple graphical description
- $=$ Simple interpretation



Why use the Mass Action model?

When there is variation in contact rates or partnerships have duration, mass action model assumptions are false. Why use them?

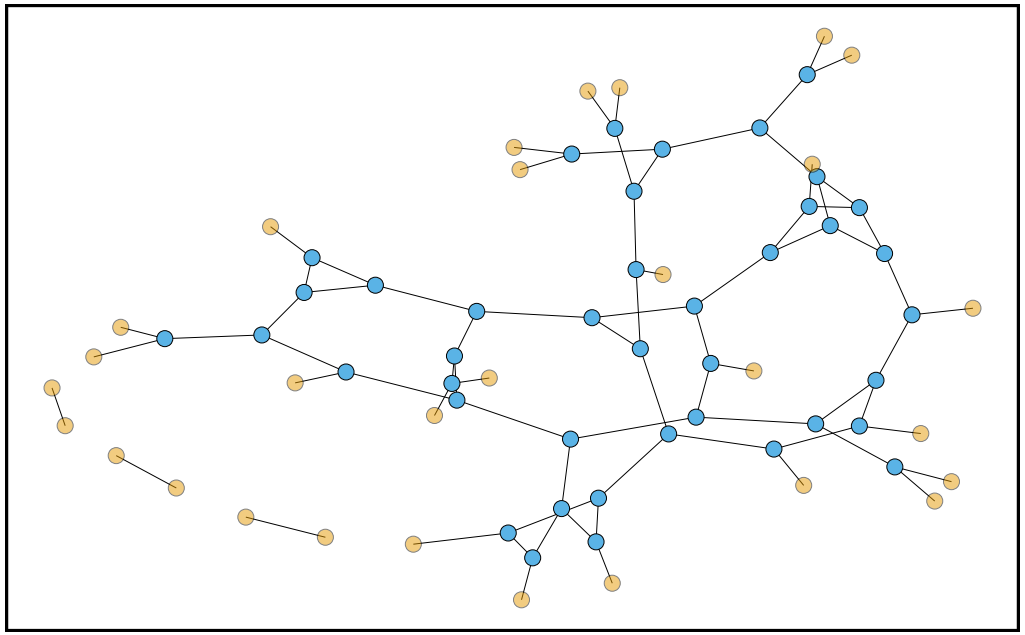
- Simple equations
- $+$ Simple graphical description
- $=$ Simple interpretation



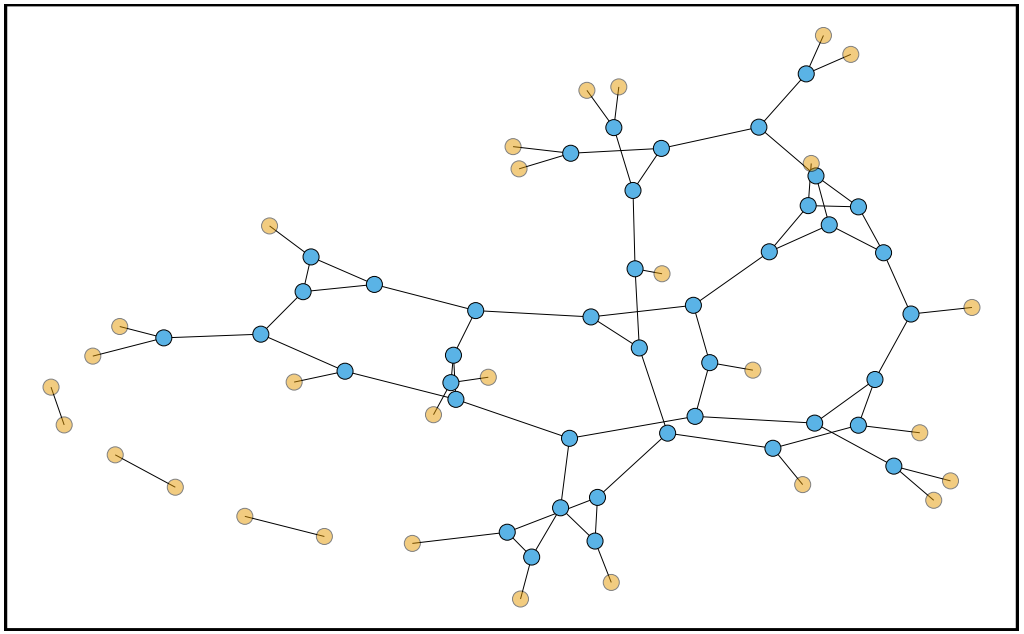
Does cost/effort of more accurate model give improved policy recommendation?

Network epidemic model

- Assume (for now) that population contacts are static. Contacts are referred to as "edges".
- Degree distribution $P(k)$ gives probability an individual has degree k, *i.e.*, k contacts. Partners chosen "randomly".
- Infection spreads along each edge at rate β. An infected individual recovers at rate γ.



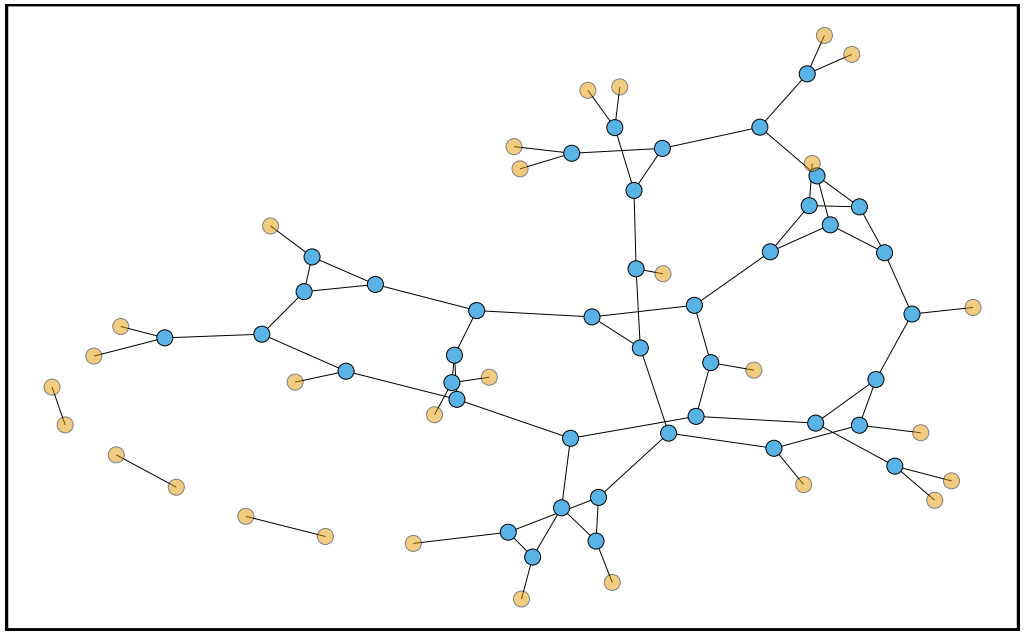
Size Bias



The probability the neighbor of a given node has degree k is

$$P_n(k) = \frac{kP(k)}{\langle K \rangle}$$

Size Bias



The probability the neighbor of a given node has degree k is

$$P_n(k) = \frac{kP(k)}{\langle K \rangle}$$

[*c.f.*, Why your friends have more friends than you do. S. Feld. American Journal of Sociology, 1991]

Previous approaches for final size

- A number of researchers have looked at the question of final size for epidemics in this (and similar) populations.
- The fundamental approach may be summarized as calculating the probability that a random individual wasn't infected during an epidemic.
- This is done by calculating the probability a neighbor was infected, and, if infected, the probability the neighbor transmitted.

Previous approaches for final size

- A number of researchers have looked at the question of final size for epidemics in this (and similar) populations.
- The fundamental approach may be summarized as calculating the probability that a random individual wasn't infected during an epidemic.
- This is done by calculating the probability a neighbor was infected, and, if infected, the probability the neighbor transmitted.
- We change our focus to calculating the probability that an individual hasn't been infected yet.

Previous approaches for dynamics

We have three pre-existing options

Previous approaches for dynamics

We have three pre-existing options

A system with $\mathcal{O}(M)$ equations where M is the maximum degree:

$$\frac{\mathrm{d}x_k}{\mathrm{d}t} = -\rho(t)\left[(\beta+\gamma)kx_k - \gamma(k+1)x_{k+1}\right]$$

$$\frac{\mathrm{d}y_k}{\mathrm{d}t} = \beta[(k+1)y_{k+1} - ky_k] - \gamma y_k + \rho(t)\left[(k+1)[\beta(x_{k+1}+y_{k+1}) + \gamma y_{k+1}] - k(\beta+\gamma)y_k\right]$$

where

$$\rho(t) = \frac{\sum ky_k}{\sum k(y_k + x_k)}$$

x_k is number of susceptible individuals u with k susceptible or infected neighbors, and y_k is the number of infected individuals with k susceptible or infected neighbors (along edges which have not transmitted to/from u).

Previous approaches for dynamics

We have three pre-existing options

A system with $\mathcal{O}(M^2)$ equations where M is the maximum degree.

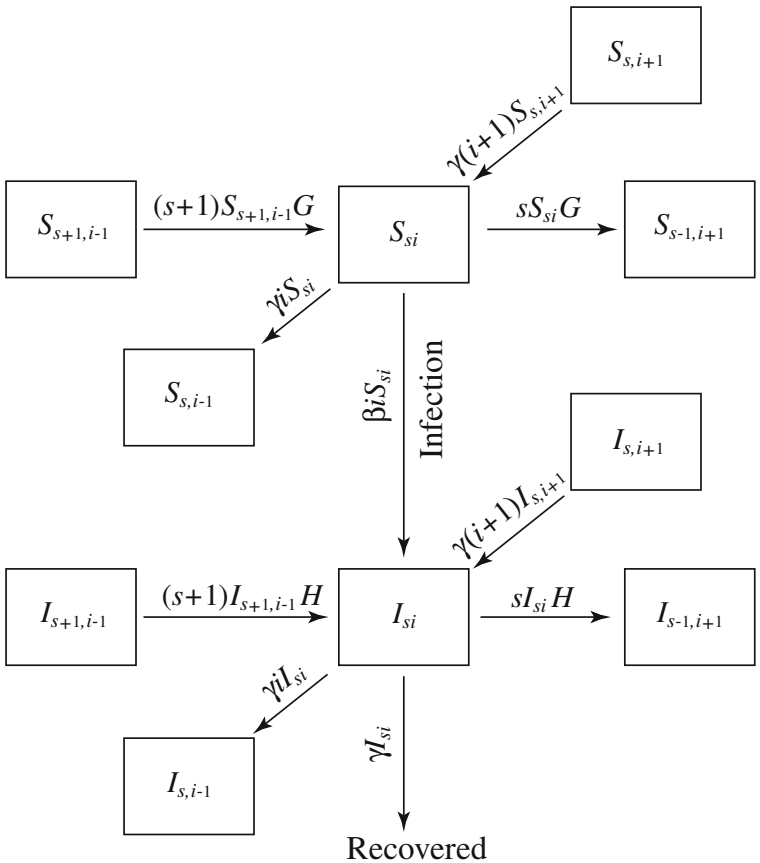


Fig. 5 Flow chart for the nodes S_{si} and I_{si} of the SIR effective degree model, where $G = \frac{\sum_{k=1}^{M}\sum_{j+l=k} j\beta l S_{jl}}{\sum_{k=1}^{M}\sum_{j+l=k} jS_{jl}}$, and $H = \frac{\sum_{k=1}^{M}\sum_{j+l=k} \beta l^2 S_{jl}}{\sum_{k=1}^{M}\sum_{j+l=k} jI_{jl}}$

The flow diagram is straightforward to understand.

Previous approaches for dynamics

We have three pre-existing options

A system with a bounded number of equations.

Table 3 A summary of the nonlinear differential equations used to the describe the spread of a simple SIR type epidemic through a random network. The degree distribution of the network is generated by $g(x)$

$$\dot{\theta} = -rp_I\theta$$

$$\dot{p}_I = rp_Sp_I\theta\frac{g''(\theta)}{g'(\theta)} - rp_I(1-p_I) - p_I\mu$$

$$\dot{p}_S = rp_Sp_I\left(1-\theta\frac{g''(\theta)}{g'(\theta)}\right)$$

$$S = g(\theta)$$

$$\dot{I} = rp_I\theta g'(\theta) - \mu I$$

There is no flow diagram, so the derivation is less intuitive.

Note $g(x) = \sum_k P(k)x^k$.

Deriving the Dynamics

The following are (almost) equivalent:

1. The proportion of the population that is susceptible, infected, or recovered at a given time.
2. The probability a random individual is susceptible, infected, or recovered given the initial conditions.
3. The probability that a random individual is susceptible, infected, or recovered given the initial conditions and that the individual is prevented from causing infection.

Deriving the Dynamics

The following are (almost) equivalent:

1. The proportion of the population that is susceptible, infected, or recovered at a given time.
2. The probability a random individual is susceptible, infected, or recovered given the initial conditions.
3. The probability that a random individual is susceptible, infected, or recovered given the initial conditions and that the individual is prevented from causing infection.

(1) $\rightarrow$ (2) is where the hard rigorous work is hiding, but if there is a deterministic limiting behavior, then it "must" be true.
(2) $\rightarrow$ (3) allows us to simplify the mathematics.

Assumptions:

- Very small initial proportion infected. (this can be weakened)
- Very large initial number of infections.

The test node

- Consider a randomly chosen test node u in the population.
- Disallow infection from the test node to its neighbors (allows independence assumption for neighbors).
- The probability the node is Susceptible, Infected, or Recovered is affected by the status of its neighbors.
- The fraction of the population that is still susceptible $S(t)$ is equal to the probability u is still susceptible.

The test node

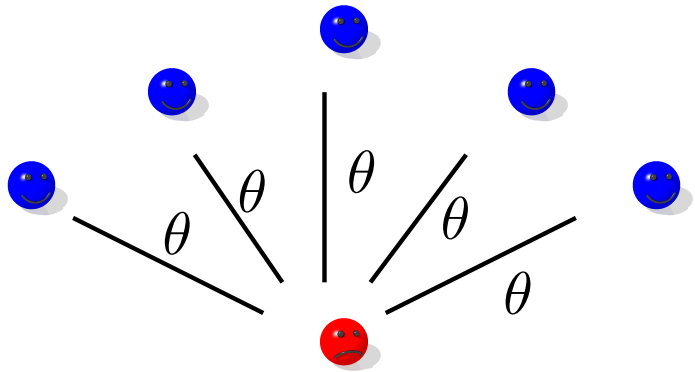
- Consider a randomly chosen test node u in the population.
- Disallow infection from the test node to its neighbors (allows independence assumption for neighbors).
- The probability the node is Susceptible, Infected, or Recovered is affected by the status of its neighbors.
- The fraction of the population that is still susceptible $S(t)$ is equal to the probability u is still susceptible.
- The fractions of the populations that are infected $I(t)$ or recovered $R(t)$ satisfy

$$I = 1 - S - R$$
$$\dot{R} = \gamma I$$

- All that remains is to determine $S(t)$, the probability u is susceptible.

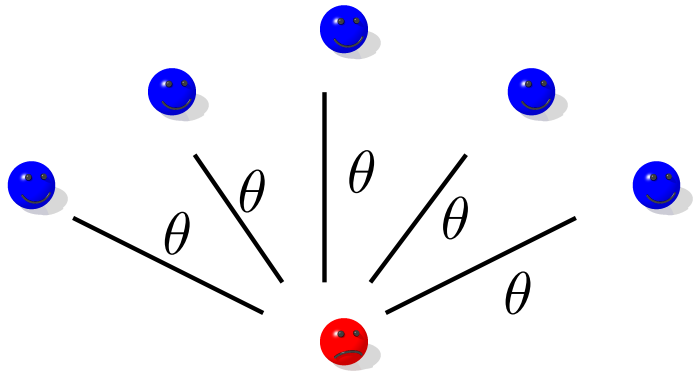
Finding $S(t)$

- Consider a random partner v of the randomly chosen test individual u.
- v has degree k with probability $P_n(k) = kP(k)/\langle K \rangle$.
- Let θ be the probability v has not yet transmitted infection to u.



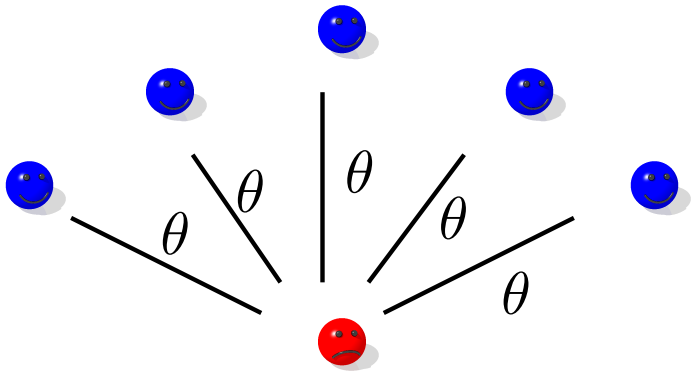
Probability a random degree k test individual still susceptible is

$$\theta(t)^k$$



Probability a random ~~degree k~~ test individual still susceptible is

$$S(t) = \sum_k P(k)\theta(t)^k$$



Probability a random ~~degree k~~ test individual still susceptible is

$$S(t) = \sum_k P(k)\theta(t)^k = \psi(\theta)$$

where

$$\psi(x) = \sum_k P(k)x^k$$

is the Probability Generating Function (pgf) of the distribution $P(k)$.

A new (partial) flow diagram

A new (partial) flow diagram

We conclude

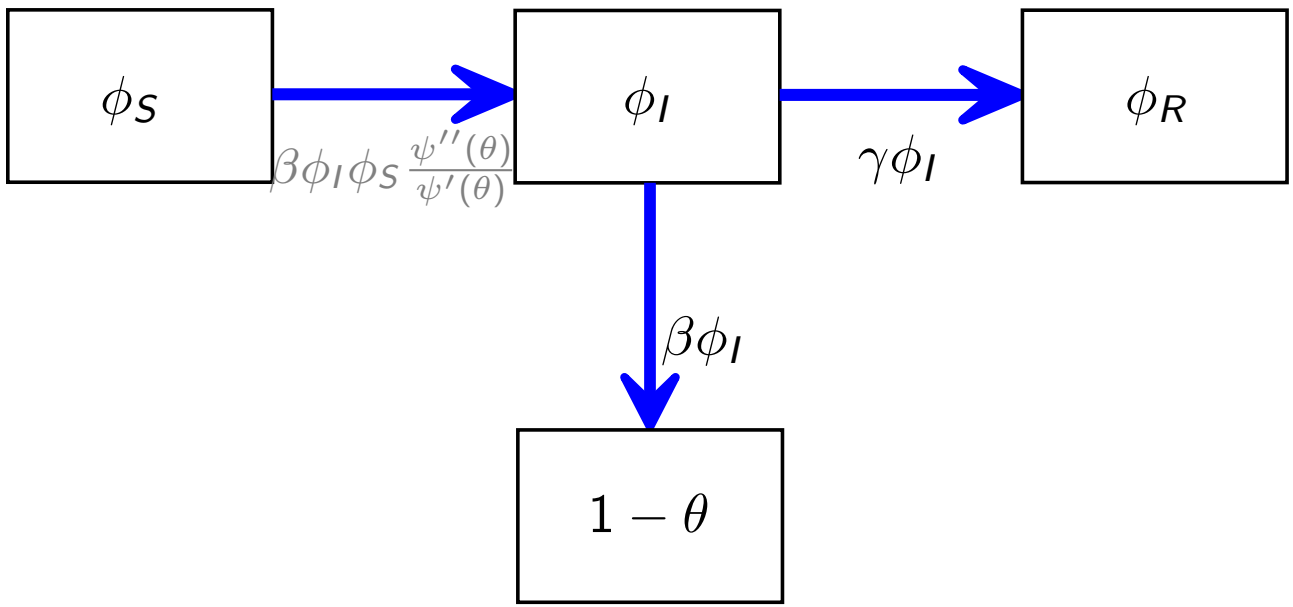
$$\dot{R} = \gamma I \qquad S = \psi(\theta) \qquad I = 1 - S - R$$

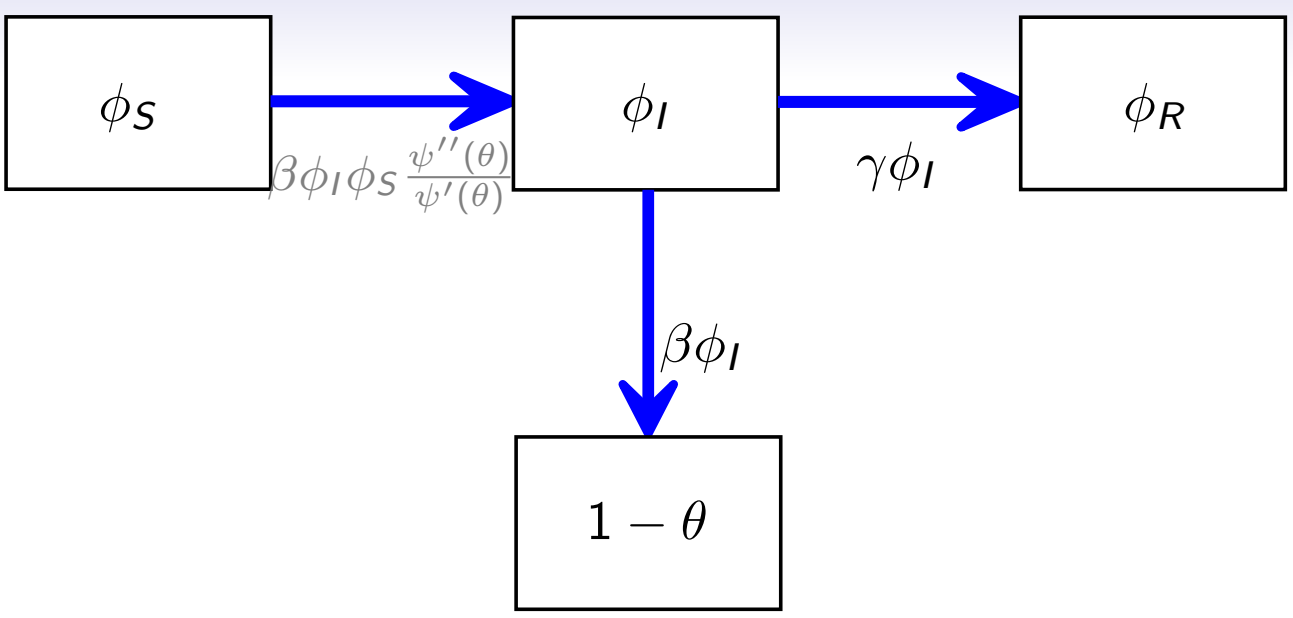
We just need θ.

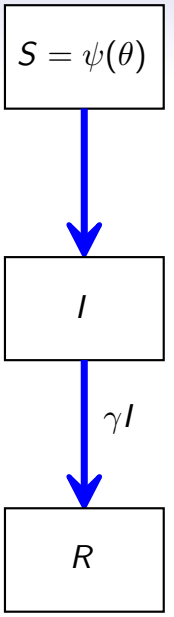
How does θ evolve?

- Let ϕ_S be the probability v (the partner) is susceptible.
- Let ϕ_I be the probability v is infected and has not infected u.
- Let ϕ_R be the probability v is recovered and did not infect u.

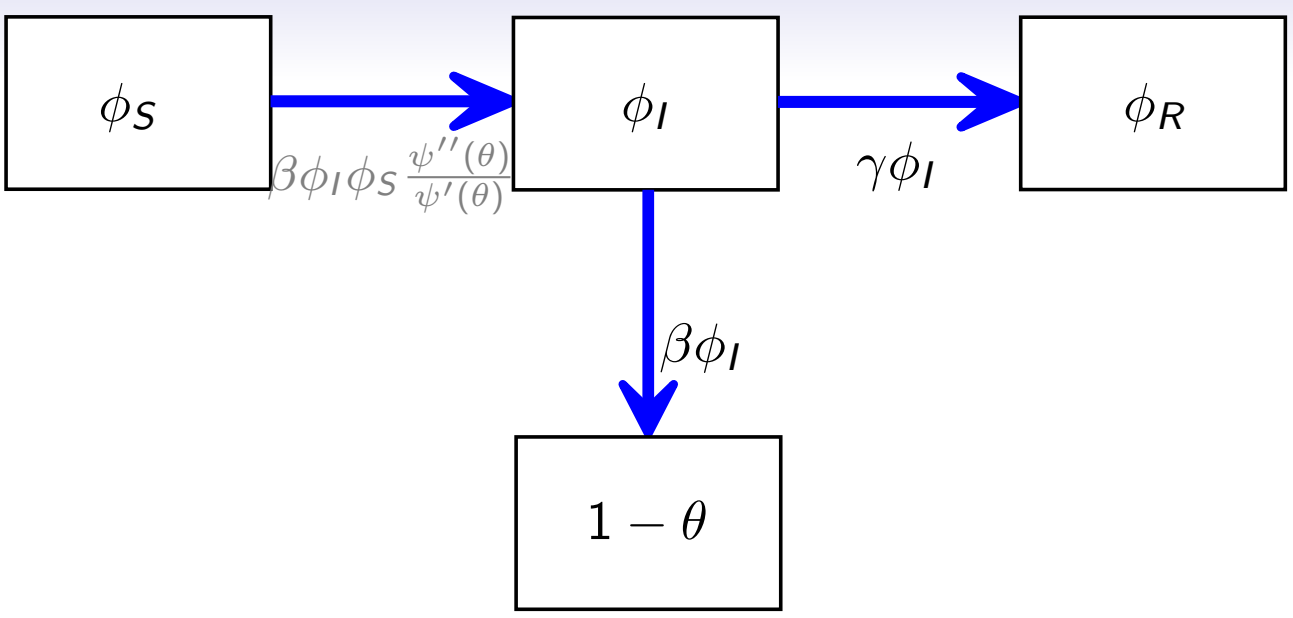
Then $\theta = \phi_S + \phi_I + \phi_R$.

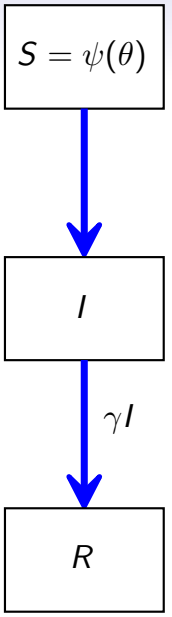




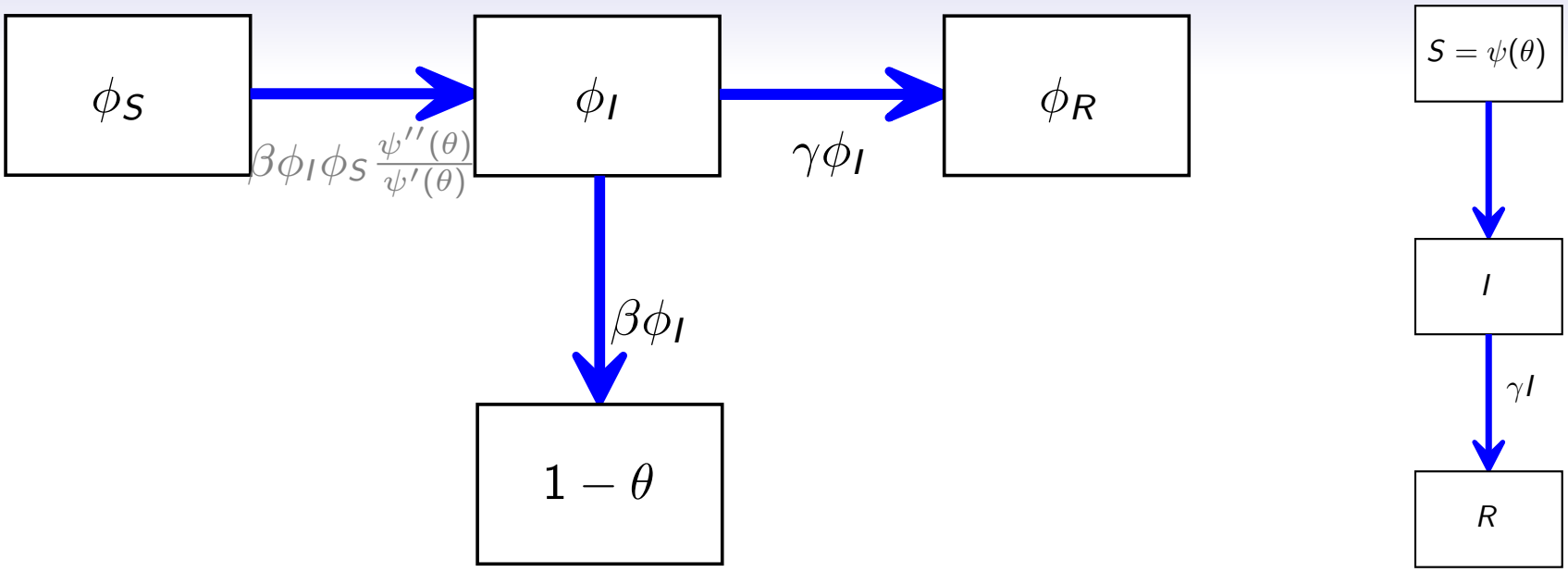


$$\phi_R = \frac{\gamma}{\beta}(1-\theta)$$

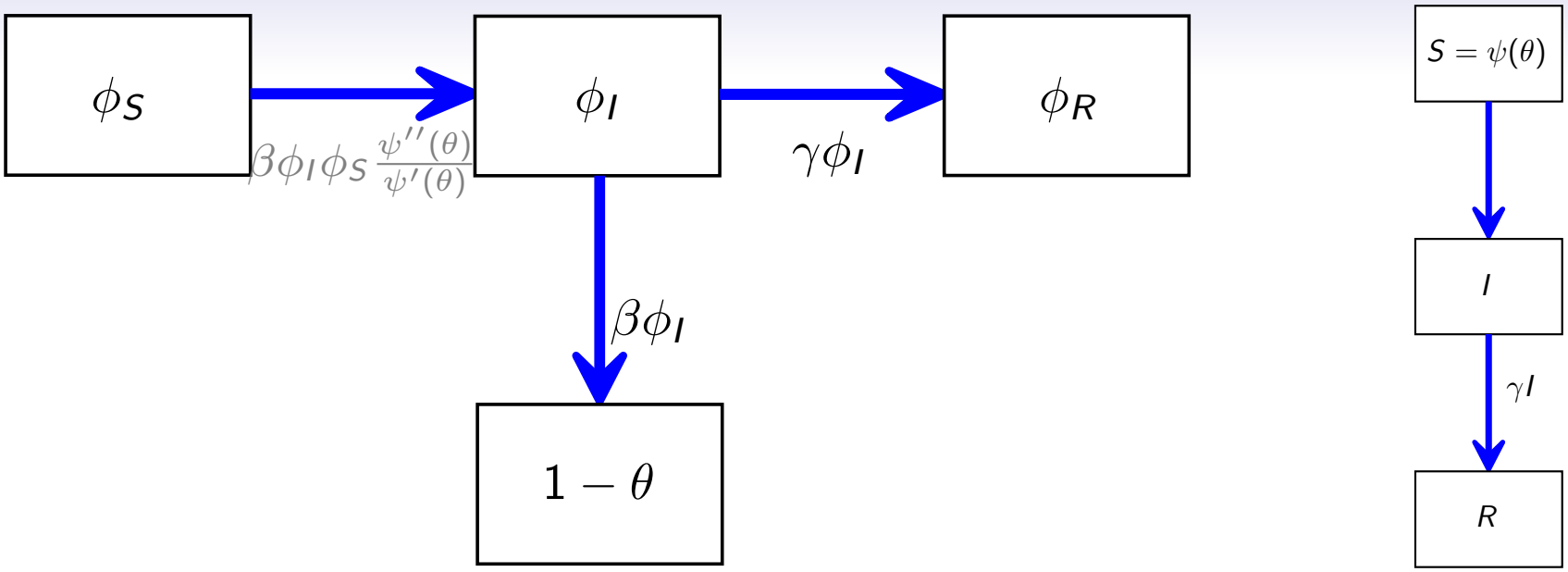




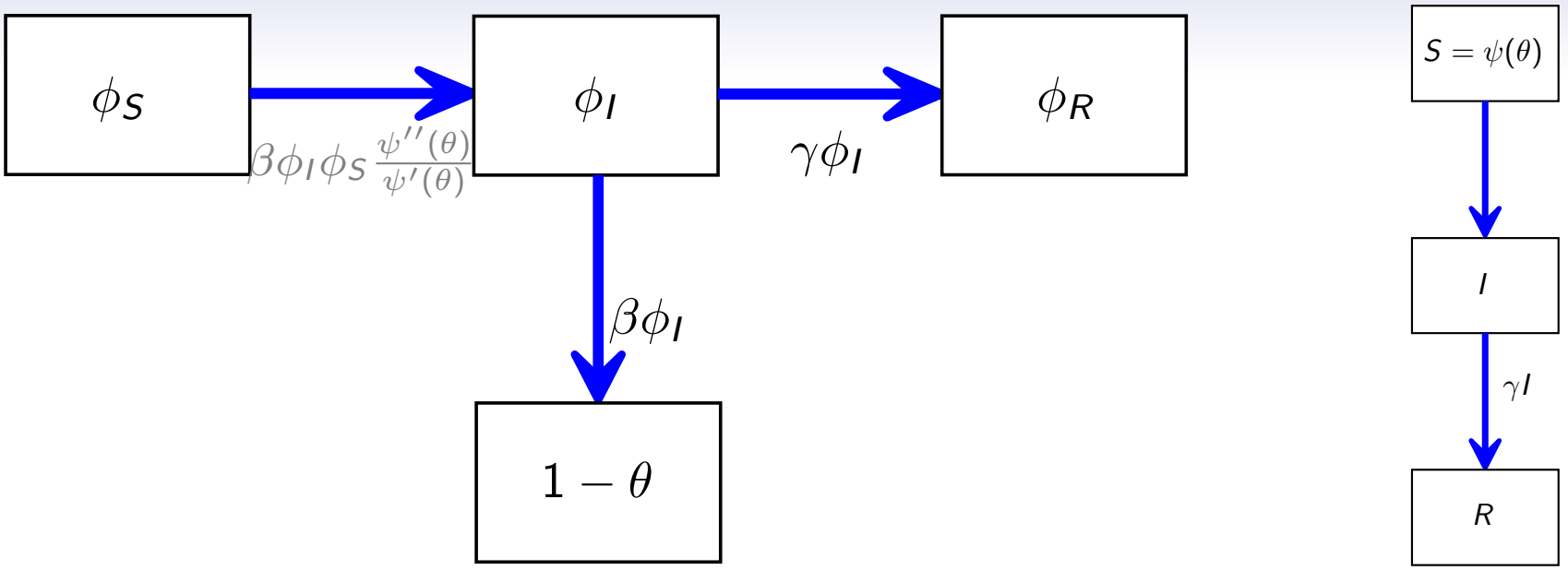
$$\phi_R = \frac{\gamma}{\beta}(1-\theta), \qquad \phi_S = \sum_k P_n(k)\theta^{k-1}$$



$$\phi_R = \frac{\gamma}{\beta}(1-\theta), \qquad \phi_S = \sum_k P_n(k)\theta^{k-1} = \sum_k \frac{kP(k)}{\langle K\rangle}\theta^{k-1}$$



$$\phi_R = \frac{\gamma}{\beta}(1-\theta), \qquad \phi_S = \sum_k P_n(k)\theta^{k-1} = \sum_k \frac{kP(k)}{\langle K\rangle}\theta^{k-1} = \frac{\psi'(\theta)}{\psi'(1)}$$



$$\phi_R = \frac{\gamma}{\beta}(1-\theta), \qquad \phi_S = \sum_k P_n(k)\theta^{k-1} = \sum_k \frac{kP(k)}{\langle K\rangle}\theta^{k-1} = \frac{\psi'(\theta)}{\psi'(1)}$$

So $\phi_I = \theta - \phi_S - \phi_R$ can be expressed in terms of θ:

$$\dot{\theta} = -\beta\phi_I = -\beta\theta + \beta\frac{\psi'(\theta)}{\psi'(1)} + \gamma(1-\theta)$$

Final System

We finally have

$$\dot{\theta} = -\beta\theta + \beta\frac{\psi'(\theta)}{\psi'(1)} + \gamma(1-\theta)$$

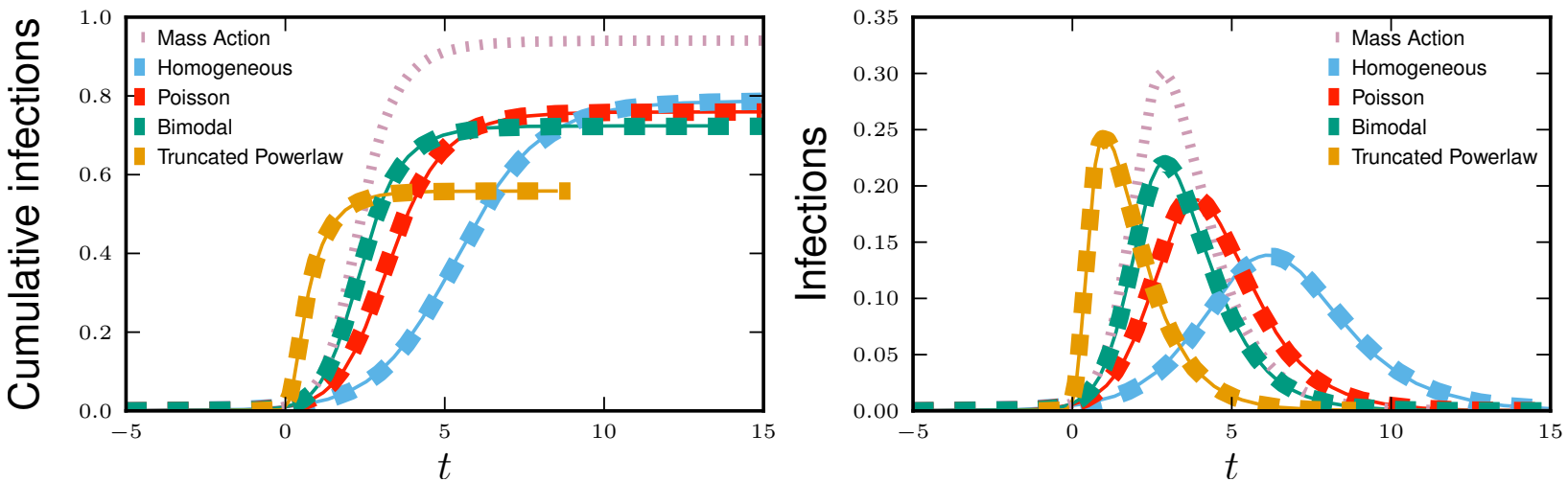
$$\dot{R} = \gamma I \qquad S = \psi(\theta) \qquad I = 1 - S - R$$

Final System

We finally have

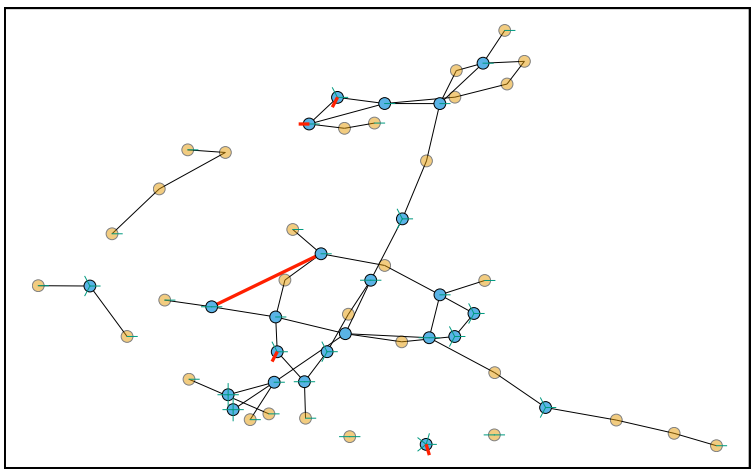
$$\dot{\theta} = -\beta\theta + \beta\frac{\psi'(\theta)}{\psi'(1)} + \gamma(1-\theta)$$

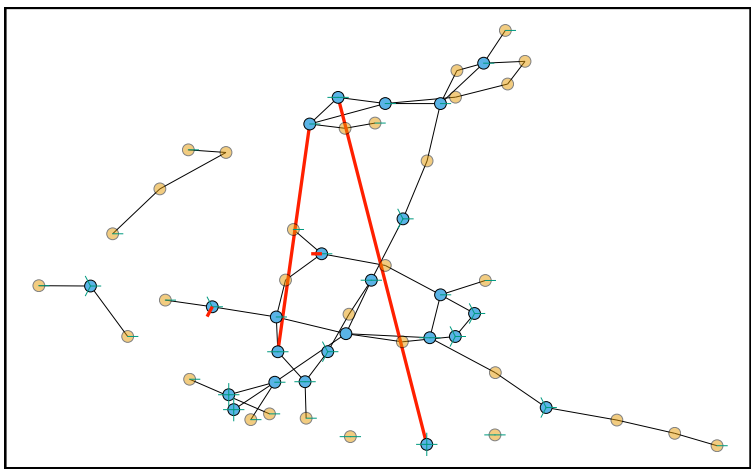
$$\dot{R} = \gamma I \qquad S = \psi(\theta) \qquad I = 1 - S - R$$



Dynamic Networks

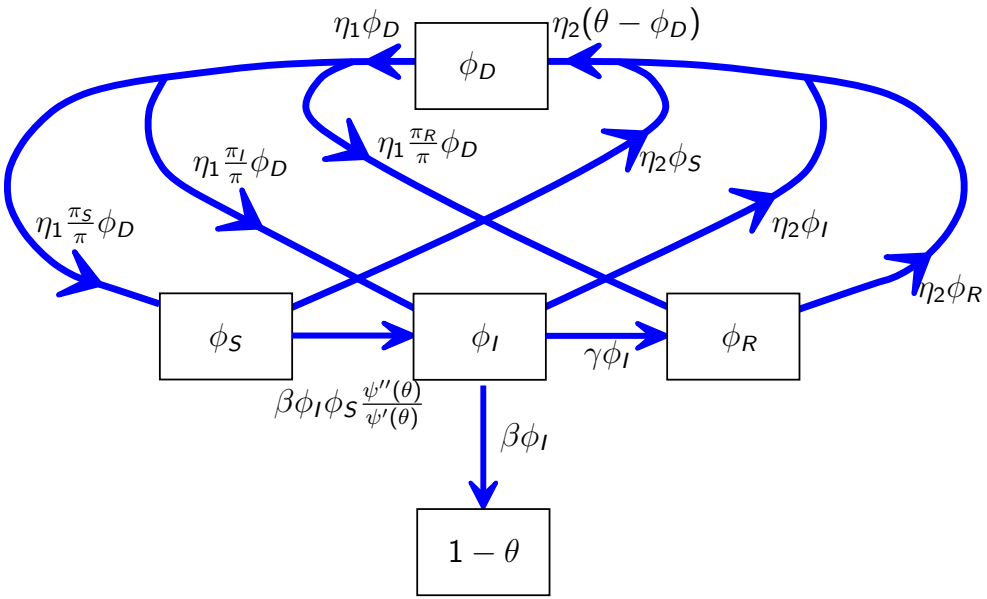
- There are several ways to generalize this to dynamic networks.
- Consider a Configuration-Model like network with nodes assigned stubs (half-edges) from a distribution.
- Stubs may be active (part of an edge) or dormant (not part of an edge). At rate η_1, dormant stubs become active and at rate η_2 active stubs become dormant.
- A node with k stubs has $k\eta_1/(\eta_1 + \eta_2)$ edges on average.



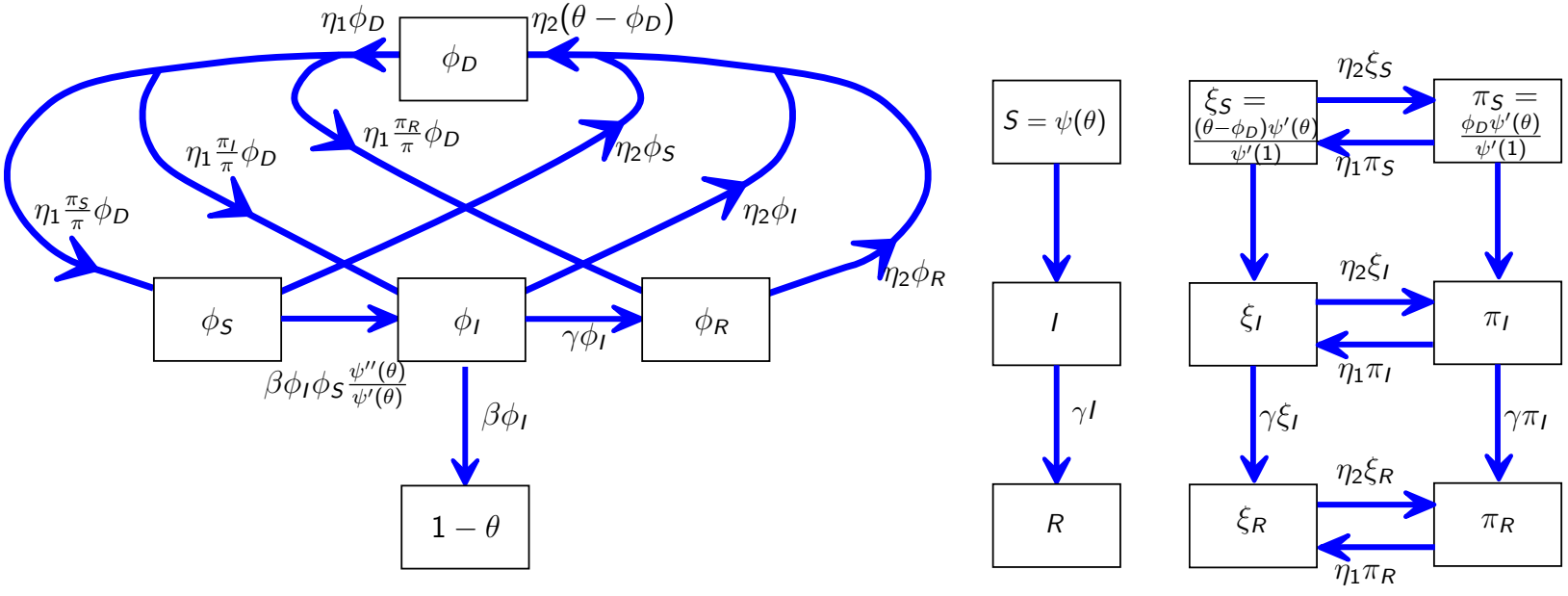


Calculating S

- Consider a randomly chosen test node u.
- The probability u is susceptible is $S(t) = \sum_k P(k)\theta^k$ where θ is the probability a stub hasn't received infection from a neighbor.
- Divide θ into ϕ_S, ϕ_I, ϕ_R, and ϕ_D where these are the probabilities that the stub has never received infection and is currently connected to a susceptible, infected, or recovered node, or dormant.



Calculating θ



The variables π_S, π_I, and π_R give the probability a random stub in the population is dormant and belongs to a susceptible, infected, or recovered node.

We have $\pi = \pi_S + \pi_I + \pi_R = \eta_2/(\eta_1 + \eta_2)$ is the probability a stub is dormant.

Similar variables exist for ξ representing active stubs.

Dormant Contact equations

We arrive at

$$\dot{\theta} = -\beta\phi_I\,,$$

$$\dot{\phi}_S = -\beta\phi_I\phi_S\frac{\psi''(\theta)}{\psi'(\theta)} + \eta_1\frac{\pi_S}{\pi}\phi_D - \eta_2\phi_S\,,$$

$$\dot{\phi}_I = \beta\phi_I\phi_S\frac{\psi''(\theta)}{\psi'(\theta)} + \eta_1\frac{\pi_I}{\pi}\phi_D - (\eta_2+\beta+\gamma)\phi_I\,,$$

$$\dot{\phi}_D = \eta_2(\theta-\phi_D) - \eta_1\phi_D\,,$$

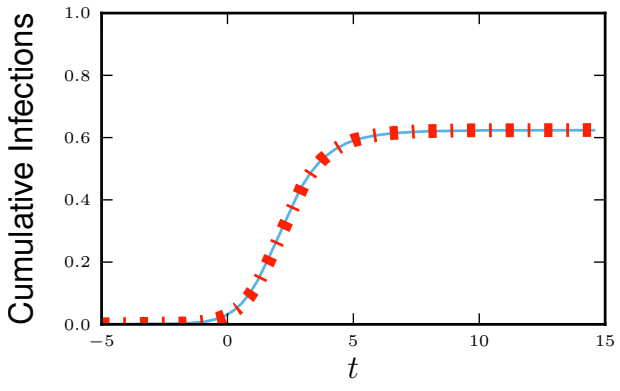
$$\dot{\xi}_R = -\eta_2\xi_R + \eta_1\pi_R + \gamma\xi_I\,, \qquad \xi_S = (\theta-\phi_D)\frac{\psi'(\theta)}{\psi'(1)}\,, \qquad \xi_I = \xi - \xi_S -$$

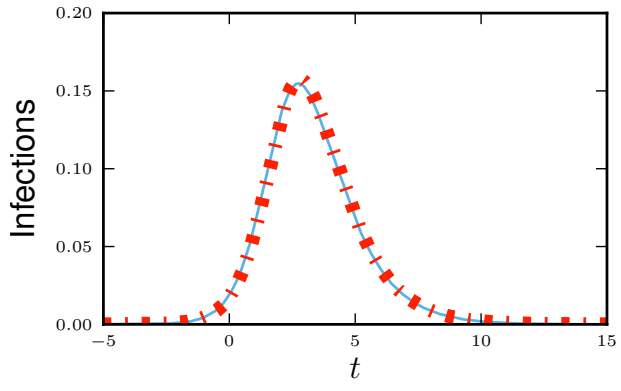
$$\dot{\pi}_R = \eta_2\xi_R - \eta_1\pi_R + \gamma\pi_I\,, \qquad \pi_S = \phi_D\frac{\psi'(\theta)}{\psi'(1)}\,, \qquad \pi_I = \pi - \pi_S - \pi_R\,,$$

$$\xi = \frac{\eta_1}{\eta_1+\eta_2}\,, \qquad \pi = \frac{\eta_2}{\eta_1+\eta_2}\,,$$

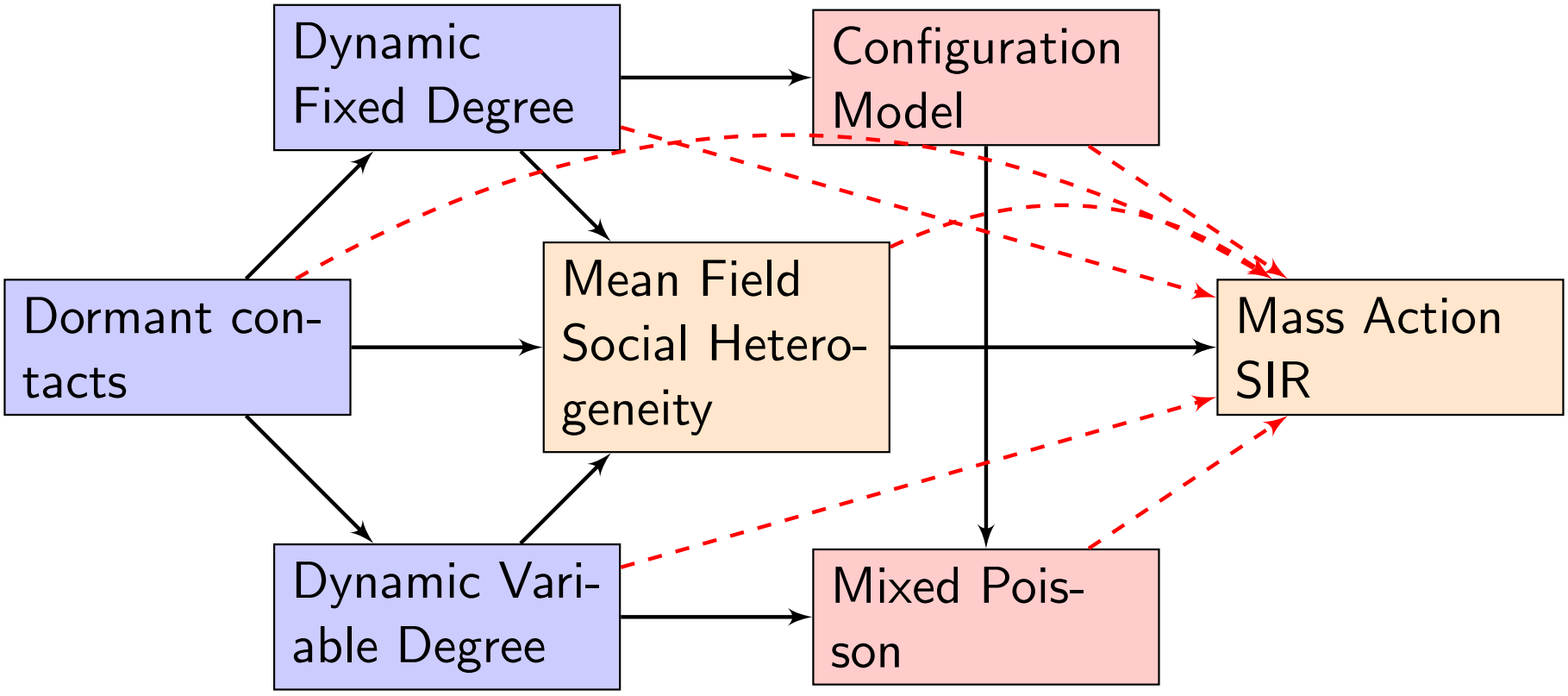
$$\dot{R} = \gamma I\,, \qquad S = \psi(\theta)\,, \qquad I = 1-S-R\,.$$

Comparison with simulation





Hierarchy



Note: Mass action almost results if $\langle K^2 \rangle / \langle K \rangle^2 \to 1$ as $\langle K \rangle \to \infty$ with $\beta \langle K \rangle$ fixed. It does result if $\langle K^4 \rangle / \langle K \rangle^4 \to 1$.

Other things we can do

- Non-constant rates
- Serosorting
- Household models
- Multitype networks
- Random Intersection Graphs
- Just about any network with configuration-model-like properties for which analytic final-size results exist (and some for which they don't).

Questions?

arXiv:1106.6320 arXiv:1106.6319 arXiv:1106.6344