Network epidemic models with two levels of mixing

Frank Ball and David Sirl

Frank.Ball@nottingham.ac.uk, David.Sirl@nottingham.ac.uk

University of Nottingham

29th October 2008

Research supported by EPSRC and The Leverhulme Trust

Overview

This talk considers two extensions of the standard network SIR epidemic model to include important real-life features yet remain susceptible to analysis

- Casual contacts — Frank Ball and Peter Neal
- Households — Frank Ball, David Sirl and Pieter Trapman

NETWORK — Configuration model

- Population $\mathcal{N} = \{1, 2, \ldots, N\}$
- D= degree of typical individual

$$p_k = \mathrm{P}(D = k) \quad (k = 0, 1, \ldots) \quad \text{specified} \qquad \mu_{\mathrm{D}} = \mathrm{E}[D]$$

- $D_1, D_2, \ldots, D_N$ iid copies of D, conditioned on $S_N = D_1 + D_2 + \cdots + D_N$ being even
- Attach D_i half-edges to individual i $(i = 1, 2, \ldots, N)$
- Pair up the S_N half-edges uniformly at random to form the network

IMPERFECTIONS — sparse if $\sigma^2_{\mathrm{D}} = \mathrm{var}(D) < \infty$

(Bollobás (2001))

Epidemic model

- SIR (susceptible $\longrightarrow$ infective $\longrightarrow$ removed)
- Infectious periods $I_1, I_2, \ldots, I_N$ iid $\sim I$ (arbitrary but specified)
- If infected, individual i makes
  - (i) local contacts along each of the D_i edges emanating from him/her independently at rate λ_{L} for each edge
  - (ii) global contacts at rate λ_{G}, with individuals chosen independently and uniformly from $\mathcal{N}$ (CASUAL CONTACTS)
- Latent period
- SPECIAL CASES
  - (i) $\lambda_{\mathrm{L}} = 0$ Standard homogeneously mixing SIR epidemic
  - (ii) $\lambda_{\mathrm{G}} = 0$ Standard network SIR epidemic

Diekmann et al. (1998), Ball and Neal (2002), (2008a), Kiss et al (2006))

Approximate deterministic model

- For $t \geq 0$, let $x_i(t)$ and $y_i(t)$ be the proportion of the population that have degree i and are susceptible and infective, respectively, at time t and $y(t) = \sum_{i=1}^{\infty} y_i(t)$.

- $$\frac{dx_i}{dt} = -\lambda_{\mathrm{G}} y x_i - \frac{\lambda_{\mathrm{L}} i x_i}{\mu_{\mathrm{D}}} \sum_{j=1}^{\infty} (j-1) y_j$$

  $$\frac{dy_i}{dt} = \lambda_{\mathrm{G}} y x_i + \frac{\lambda_{\mathrm{L}} i x_i}{\mu_{\mathrm{D}}} \sum_{j=1}^{\infty} (j-1) y_j - \frac{1}{\mu_{\mathrm{I}}} y_i \quad (i = 0, 1, \ldots)$$

- Model makes three approximations

  (i) Globally contacted individuals lose one neighbour on infection

  (ii) Allows repeated local transmission down same edge

  (iii) Effective degrees of individuals do not decrease as epidemic progresses

- Exact deterministic model given in Ball and Neal (2008a)

(Kiss et al. (2006); cf. May and Lloyd(2001), Moreno et al.(2002))

Exact deterministic model

- Form network as epidemic evolves.
- For $t \geq 0$, let $x_i(t)$ and $y_i(t)$ be the proportion of the population that have effective degree i and are susceptible and infective, respectively, at time t.
- $$\frac{dy_i}{dt} = \lambda_{\mathrm{G}} y x_i + \lambda_{\mathrm{L}}[(i+1)y_{i+1} - i y_i] - \gamma y_i$$
$$+\rho_E(t)\{(i+1)[\lambda_{\mathrm{L}}(x_{i+1}+y_{i+1}) + \gamma y_{i+1}] - i(\lambda_{\mathrm{L}}+\gamma)y_i(t)\},$$
$$\frac{dx_i}{dt} = -\lambda_{\mathrm{G}} y x_i - \rho_E(t)[(\lambda_{\mathrm{L}}+\gamma)i x_i - \gamma(i+1)x_{i+1}], \quad (i = 0, 1, \ldots),$$

  where $\gamma = \frac{1}{\mu_{\mathrm{I}}}$, $\rho_E(t) = \left(\sum_{i=1}^{\infty} i y_i(t)\right) / \sum_{i=1}^{\infty} i(y_i(t) + x_i(t))$ and $y(t) = \sum_{i=1}^{\infty} y_i(t)$.

(Ball and Neal (2008a))

Basic reproduction number R_0

- Suppose N is large and there are few initial infectives
- Early stages of epidemic can be approximated by a two-type branching process

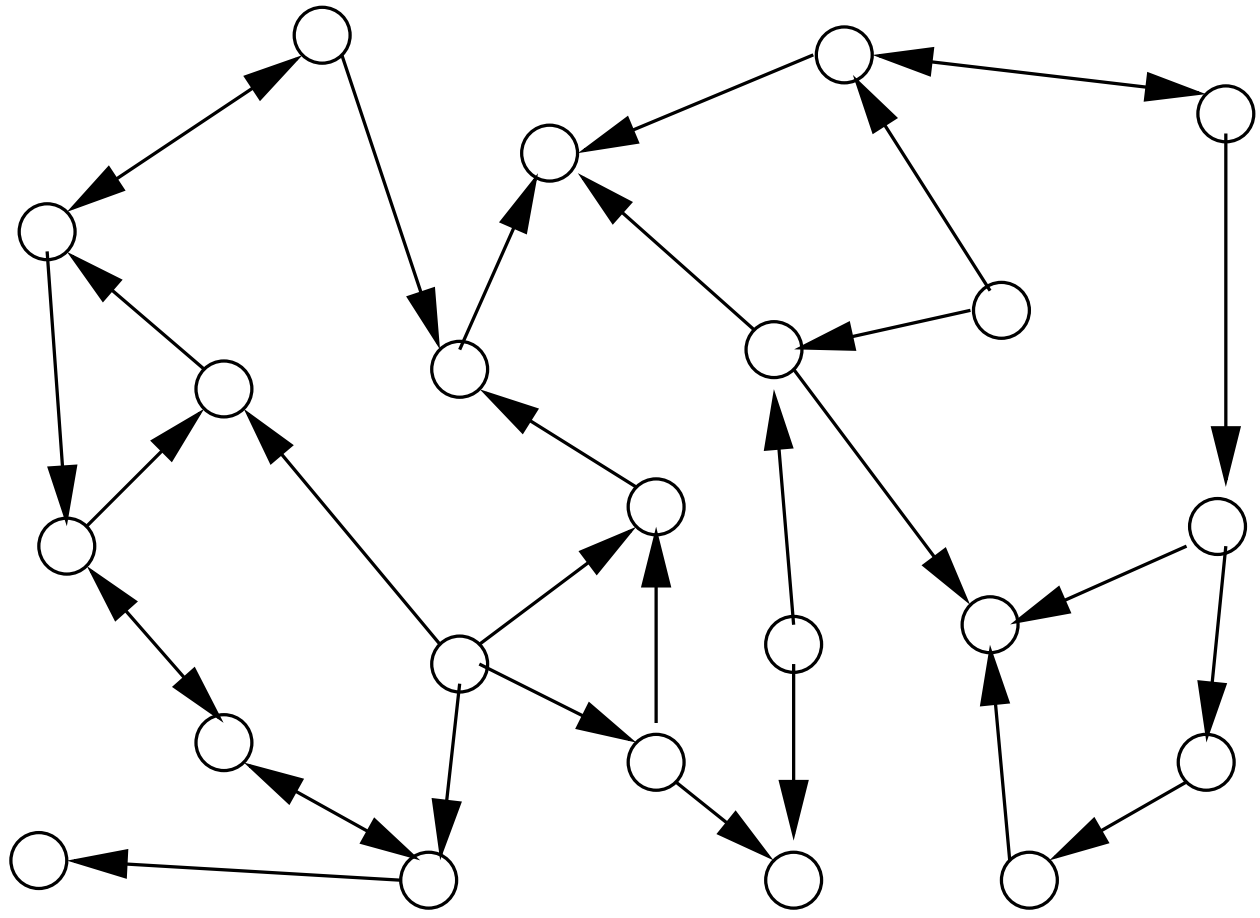
  Type-L infectives — infected locally (i.e. through the network)

  Type-G infectives — infected globally (i.e. by a casual contact)
- Mean offspring matrix

$$M = \begin{bmatrix} m_{\mathrm{GG}} & m_{\mathrm{GL}} \\ m_{\mathrm{LG}} & m_{\mathrm{LL}} \end{bmatrix} = \begin{bmatrix} \lambda_{\mathrm{G}}\mu_{\mathrm{I}} & p_{\mathrm{L}}\mu_{\mathrm{D}} \\ \lambda_{\mathrm{G}}\mu_{\mathrm{I}} & p_{\mathrm{L}}(\mu_{\tilde{\mathrm{D}}} - 1) \end{bmatrix},$$

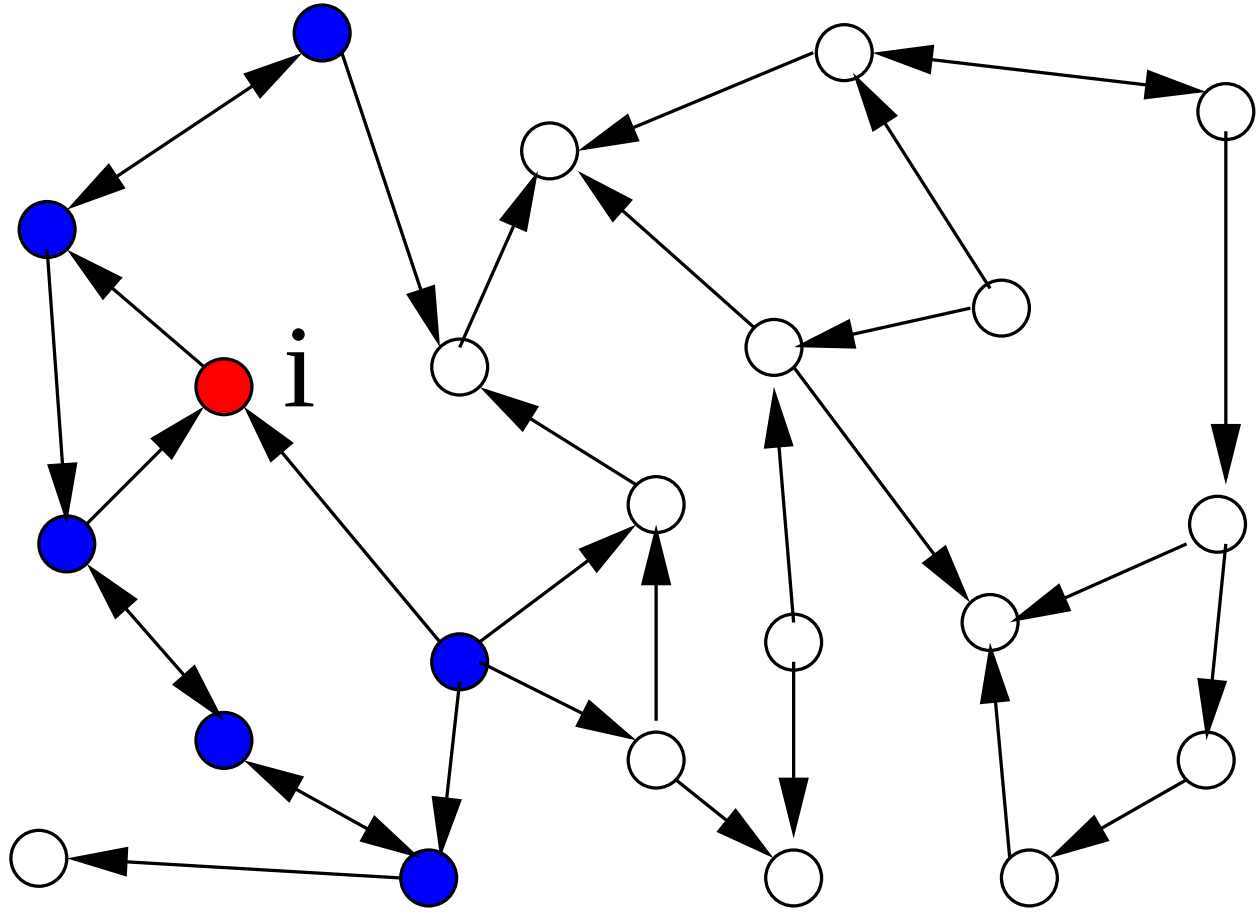
  where $p_{\mathrm{L}} = \mathrm{P}(\text{individual infects given neighbour locally}) = 1 - \mathrm{E}[\mathrm{e}^{-\lambda_{\mathrm{L}} I}]$ and $\mu_{\tilde{\mathrm{D}}} = \mathrm{E}[\tilde{D}]$. Here $\tilde{D}$= degree of typical locally contacted individual [$\mathrm{P}(\tilde{D} = k) = kp_k/\mu_{\mathrm{D}} \quad k = 1, 2, \ldots$]
- R_0 is the maximal eigenvalue of M
- $\mathrm{P}(\text{global epidemic}) > 0 \iff R_0 > 1$

Directed graph of potential local contacts



$i \to j$ if and only if i, if infected, contacts j locally (i.e. through the network).

Local susceptibility set $\mathcal{S}_i^N$



$\mathcal{S}_i^N = \{j \in \mathcal{N} : j \rightsquigarrow i\}$, where $j \rightsquigarrow i$ if and only if there exists a chain of directed arcs from j to i, and $S_i^N = |\mathcal{S}_i^N|$.

Final outcome of global epidemic

- Suppose N is large and there are few initial infectives. Let z be the expected proportion of the population who are infected by the epidemic. Then

$$\pi = \mathrm{P}(\text{typical susceptible avoids global infection}) = \exp\left(-\frac{\lambda_{\mathrm{G}}}{N} N z \mu_{\mathrm{I}}\right) = \exp(-\lambda_{\mathrm{G}} z \mu_{\mathrm{I}})$$

and

$$\begin{aligned} 1 - z &= \mathrm{P}(\text{typical susceptible avoids infection by epidemic}) \\ &= \mathrm{P}(\text{typical local susceptibility set avoids global infection}) \\ &= \sum_{k=1}^{\infty} \mathrm{P}(S = k)\pi^k = f_S(\pi) = f_S(\mathrm{e}^{-\lambda_{\mathrm{G}} z \mu_{\mathrm{I}}}) \end{aligned} \tag{1}$$

- $R_0 \leq 1$ $z = 0$ is the only solution of (1) in $[0, 1]$

  $R_0 > 1$ unique second solution $\hat{z} \in (0, 1)$, giving mean 'size' of global epidemic

- Fully rigorous proof and central limit theorem for final size of global epidemic is available using Scalia-Tomba (1985) embedding technique (Ball and Neal(2008b))

Size S^N of typical local susceptibility set

- $S^N \xrightarrow{a.s.} S$ as $N \to \infty$, where S is the total size of a branching process having offspring law $Y_0 \sim \mathrm{Bin}(D, p_{\mathrm{L}})$ for the initial individual and $Y_1 \sim \mathrm{Bin}(\tilde{D}-1, p_{\mathrm{L}})$ for all subsequent individuals
- $S = 1 + \sum_{k=1}^{Y_0} \tilde{S}_k$, where $\tilde{S}_1, \tilde{S}_2, \ldots$ are iid $\sim \tilde{S}$ and $\tilde{S}$ is the total size of the branching process in which all individuals have law Y_1. Thus,

$$\tilde{S} = 1 + \sum_{k=1}^{Y_1} \tilde{S}'_k, \quad \text{where } \tilde{S}'_1, \tilde{S}'_2, \ldots \text{ are iid} \sim \tilde{S}$$

- Let $f_{\mathrm{D}}(x) = \sum_{k=0}^{\infty} \mathrm{P}(D=k)x^k$ be the pgf of D. Then,

$$f_S(x) = \mathrm{E}[x^S] = \mathrm{E}[\mathrm{E}[x^S|D]] = \mathrm{E}[x(1-p_{\mathrm{L}}+p_{\mathrm{L}}f_{\tilde{S}}(x))^D] = x f_{\mathrm{D}}(1-p_{\mathrm{L}}+p_{\mathrm{L}}f_{\tilde{S}}(x))$$

and

$$f_{\tilde{S}}(x) = \mathrm{E}[x^{\tilde{S}}] = \mathrm{E}[x(1-p_{\mathrm{L}}+p_{\mathrm{L}}f_{\tilde{S}}(x))^{\tilde{D}-1}] = \frac{x}{\mu_{\mathrm{D}}} f^{(1)}_{\mathrm{D}}(1-p_{\mathrm{L}}+p_{\mathrm{L}}f_{\tilde{S}}(x))$$

- Proportion z of population infected by a global epidemic is given by the largest root of

$$1 - z = f_S(\mathrm{e}^{-\lambda_{\mathrm{G}} z \mu_{\mathrm{I}}})$$

No global infection ($\lambda_{\mathrm{G}} = 0$)

- Proportion z of population infected by a global epidemic is given by the largest root of

$$1 - z = f_S(\mathrm{e}^{-\lambda_{\mathrm{G}} z \mu_{\mathrm{I}}}), \tag{2}$$

where $f_S(x) = x f_{\mathrm{D}}(1 - p_{\mathrm{L}} + p_{\mathrm{L}} f_{\tilde{S}}(x))$ and $f_{\tilde{S}}(x) = \frac{x}{\mu_{\mathrm{D}}} f_{\mathrm{D}}^{(1)}(1 - p_{\mathrm{L}} + p_{\mathrm{L}} f_{\tilde{S}}(x))$

- Let $\lambda_{\mathrm{G}} \downarrow 0$. Then model becomes standard network SIR epidemic and (2) yields

$$1 - z = f_S(1-) = \mathrm{P}(S < \infty) \Longrightarrow z = \mathrm{P}(S = \infty),$$

so $z > 0 \iff m_{LL} > 1$ (i.e. local epidemic above threshold)

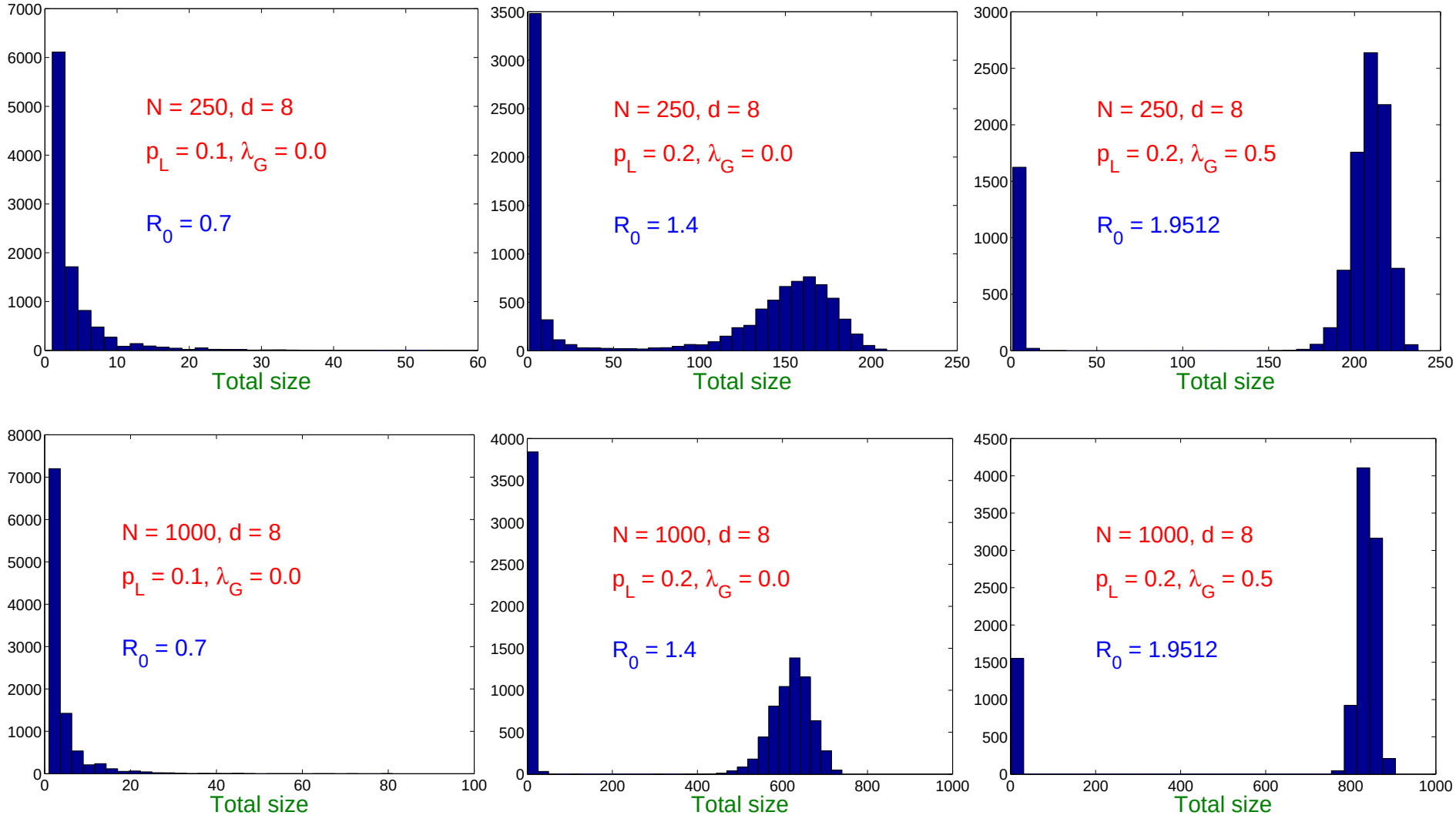
- If $m_{LL} > 1$ then

$$z = 1 - f_{\mathrm{D}}(1 - p_{\mathrm{L}} + p_{\mathrm{L}} u), \text{ where } u = f_{\tilde{S}}(1-) = \frac{1}{\mu_{\mathrm{D}}} f_{\mathrm{D}}^{(1)}(1 - p_{\mathrm{L}} + p_{\mathrm{L}} u),$$

agreeing with previous results (e.g. Andersson (1999) and Newman (2002)).

- Letting $\lambda_{\mathrm{G}} \downarrow 0$ in CLT gives heuristically a CLT for (i) the final size of the standard network SIR epidemic and (ii) (by then letting $\lambda_L \uparrow \infty$) the size of the giant component in the configuration model.

Illustration of threshold



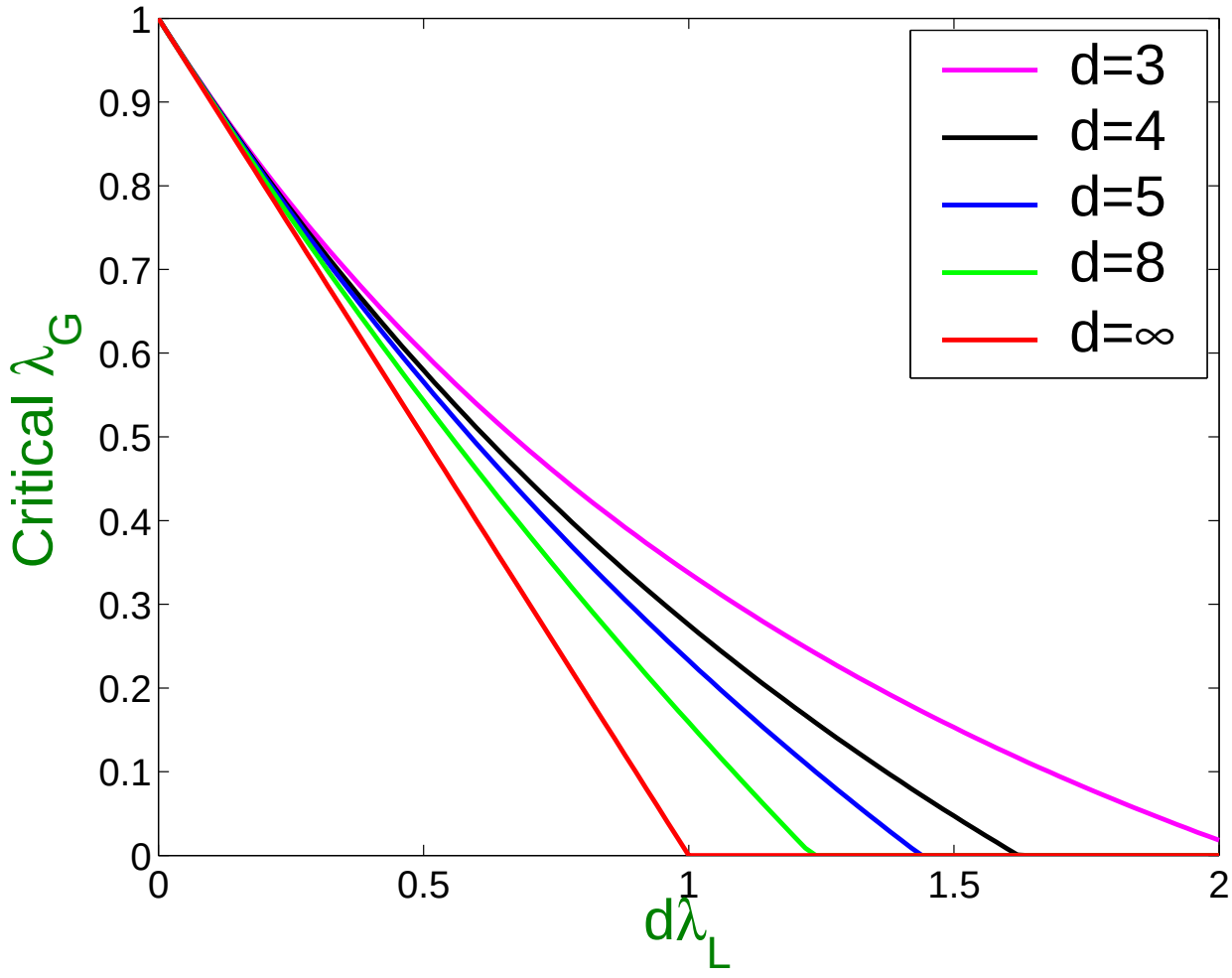
Histograms of size of 10,000 simulated epidemics per parameter combination, for a constant-degree network with $D \equiv d, I \equiv 1$ and other parameters as shown.

Comparisons of simulations and theory

Parameters		Theoretical	$N = 250$	$N = 1000$	$N = 5000$	$N = 10000$
$d = 4$	$\hat{p}_{\mathrm{EXT}}$	0.3137	0.3372	0.3199	0.3119	0.3185
$\lambda_{\mathrm{G}} = 0.5$	Mean	0.6864	0.6758	0.6842	0.6863	0.6859
$p_{\mathrm{L}} = 0.3$	Variance	1.3864	1.7852	1.4423	1.3566	1.3971
$d = 8$	$\hat{p}_{\mathrm{EXT}}$	0.1601	0.1662	0.1618	0.1645	0.1610
$\lambda_{\mathrm{G}} = 0.5$	Mean	0.8399	0.8361	0.8393	0.8395	0.8399
$p_{\mathrm{L}} = 0.2$	Variance	0.3743	0.3792	0.3765	0.3771	0.3671
$d = 8$	$\hat{p}_{\mathrm{EXT}}$	0.3767	0.4118	0.3786	0.3740	0.3685
$\lambda_{\mathrm{G}} = 0$	Mean	0.6233	0.6099	0.6196	0.6226	0.6230
$p_{\mathrm{L}} = 0.2$	Variance	2.0351	2.6698	2.1677	2.0334	2.1171

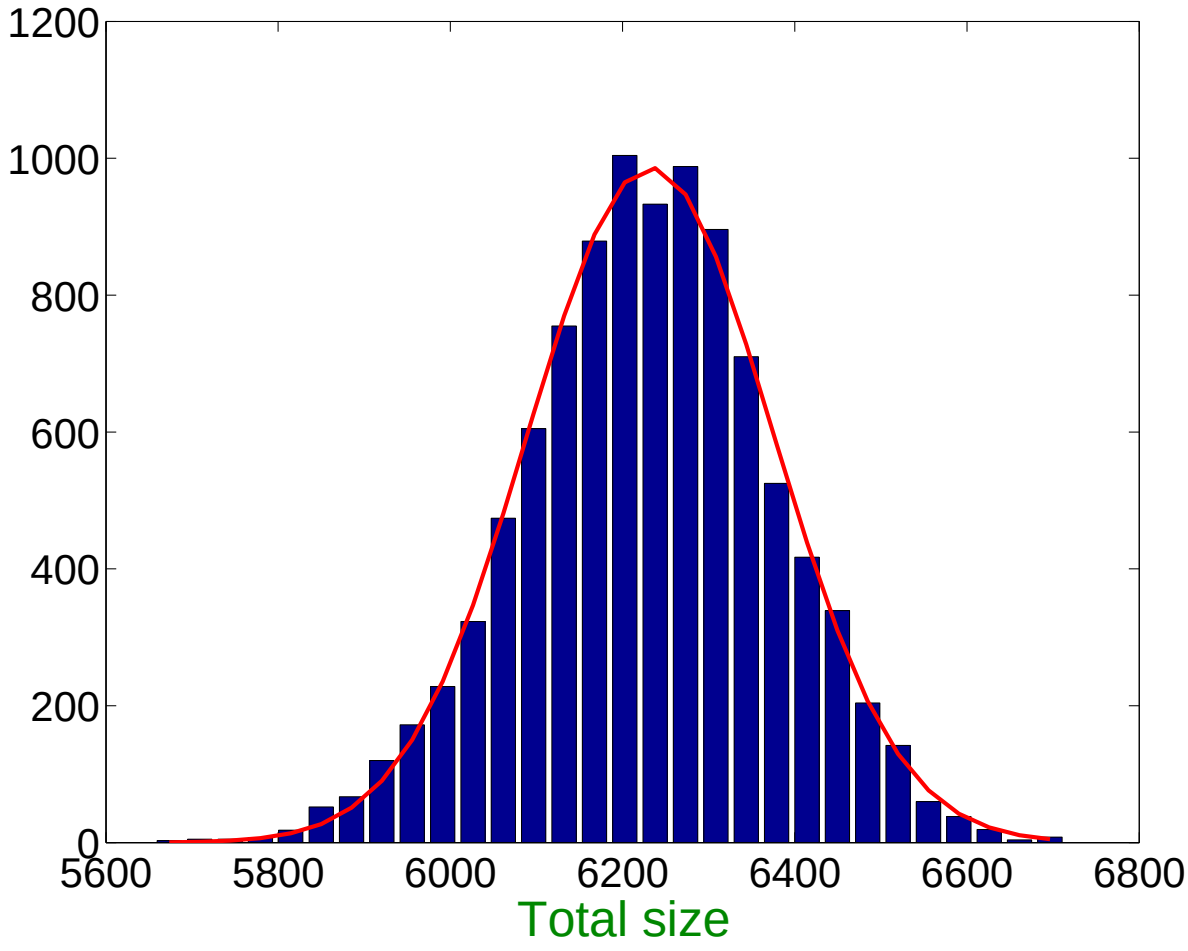
Comparisons of the (scaled) mean, variance and probability of global epidemics for finite N and their asymptotic limit as $N \to \infty$, when $I \equiv 1$ and $D \equiv d$, based on 10,000 simulations per parameter combination.

Critical values of $(\lambda_{\mathrm{G}}, d\lambda_L)$



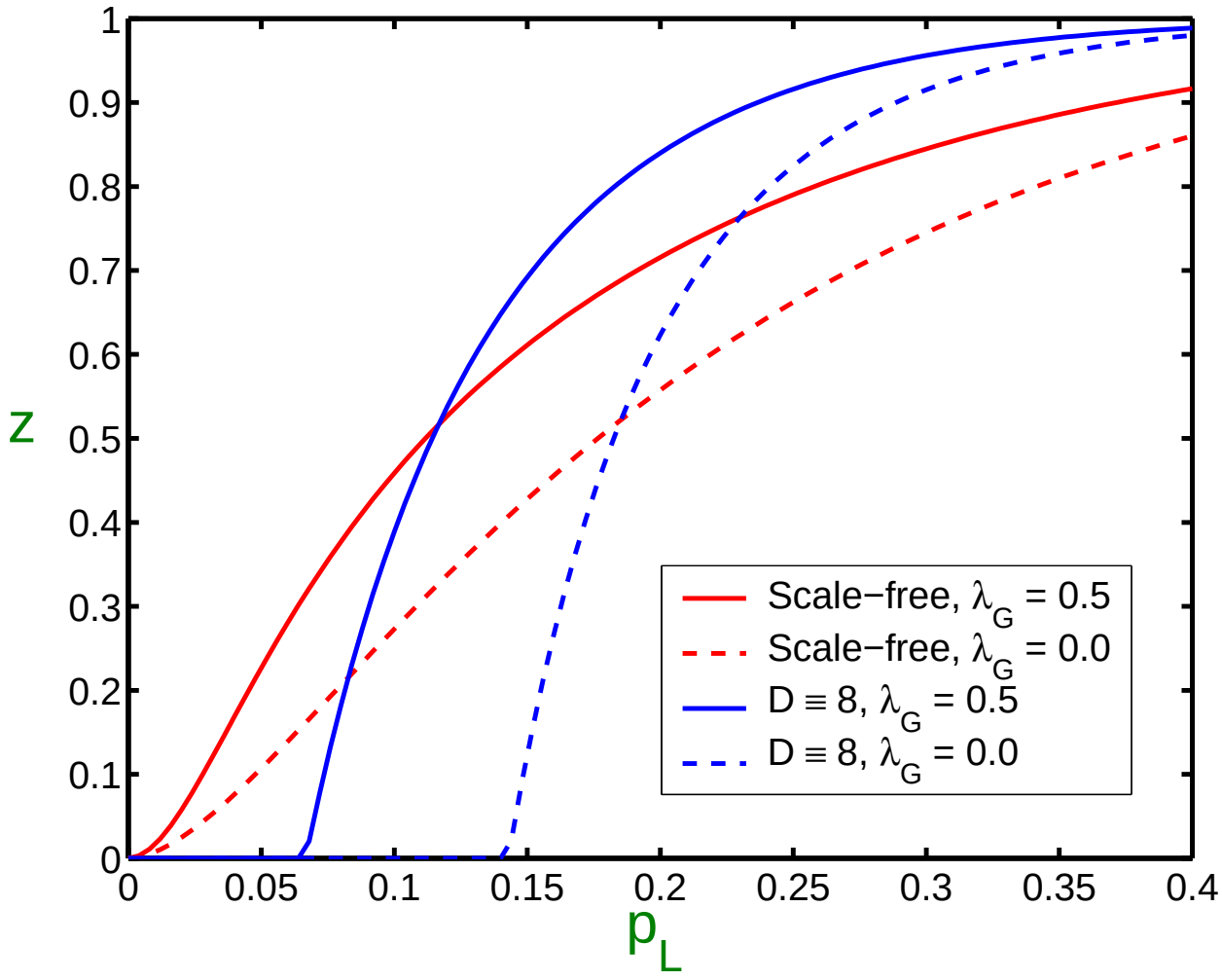
Critical values of $(\lambda_{\mathrm{L}}, \lambda_{\mathrm{G}})$ so that $R_0 = 1$ when $I \equiv 1$. [Expected number of potentially infectious contacts made by an infective is $\lambda_{\mathrm{G}} + d\lambda_{\mathrm{L}}$.]

Illustration of CLT



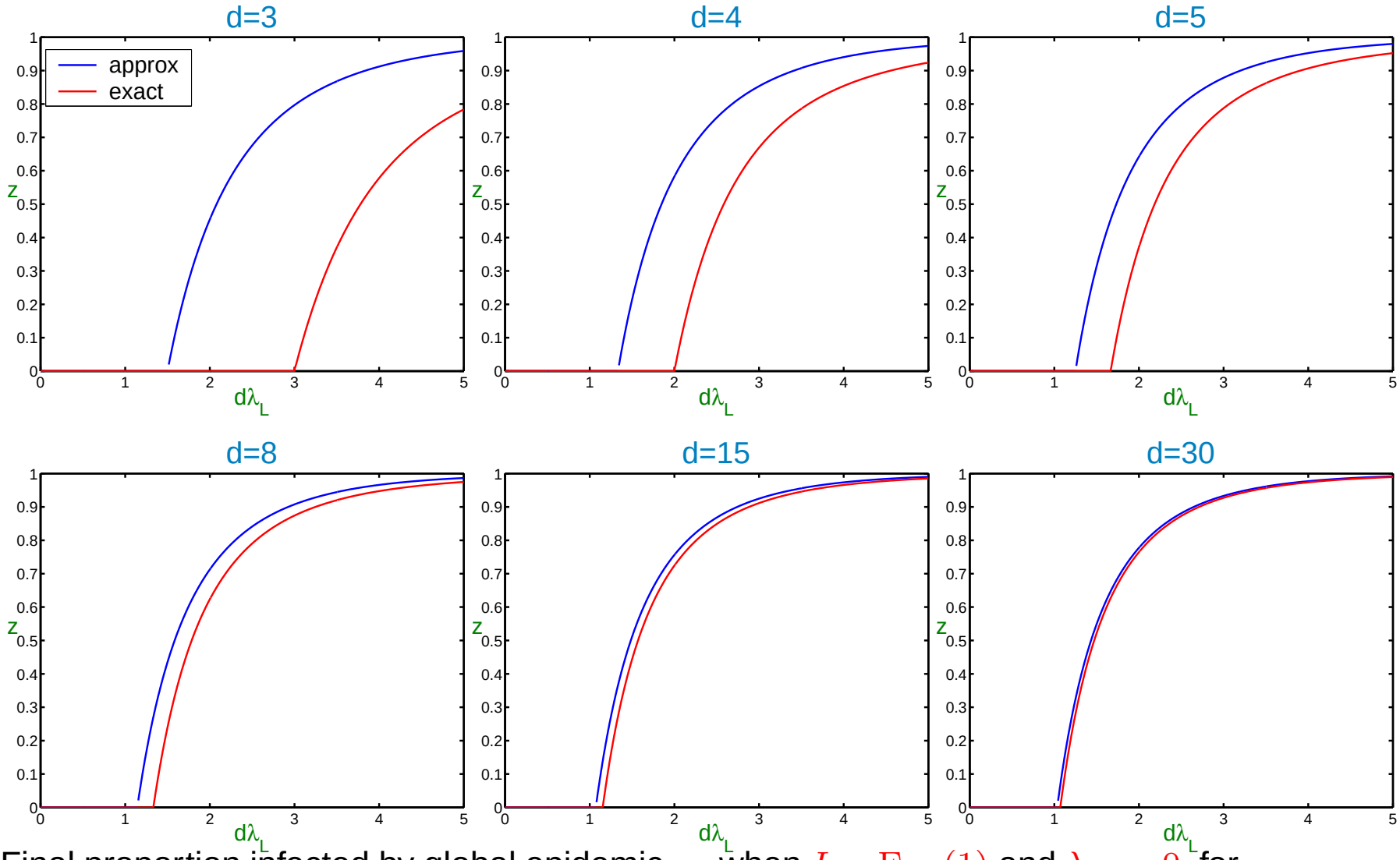
Histogram of size of 10,000 simulated global epidemics in a population of size $N = 10,000$ when $D \equiv 8, \lambda_{\mathrm{G}} = 0$ and $p_{\mathrm{L}} = 0.2$ ($I \equiv 1$ and $\lambda_{\mathrm{L}} = -\log 0.8$), with asymptotic normal approximation superimposed.

Mean final outcome



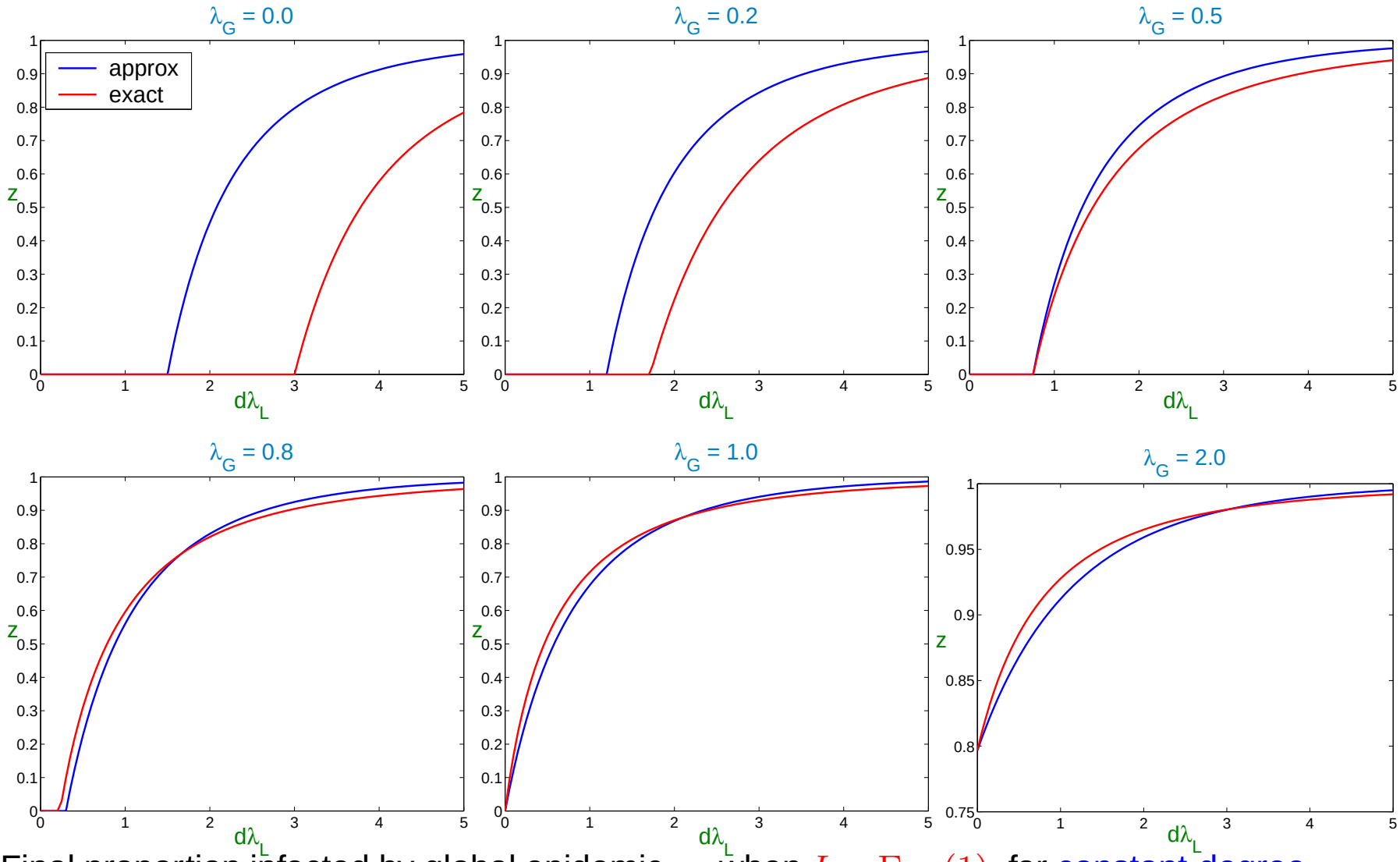
Asymptotic proportion of population infected by global epidemic, z, for constant-degree and scale-free ($\mathrm{P}(D=k) \propto k^{-2.466956}\ (k=3,4,\ldots)$) networks with $\mu_{\mathrm{D}} = 8$ when $I \equiv 1$.

Accuracy of Kiss et al. approximation



Final proportion infected by global epidemic, z, when $I \sim \mathrm{Exp}(1)$ and $\lambda_{\mathrm{G}} = 0$, for constant-degree network (i.e. $D \equiv d$).

Accuracy of Kiss et al. approximation

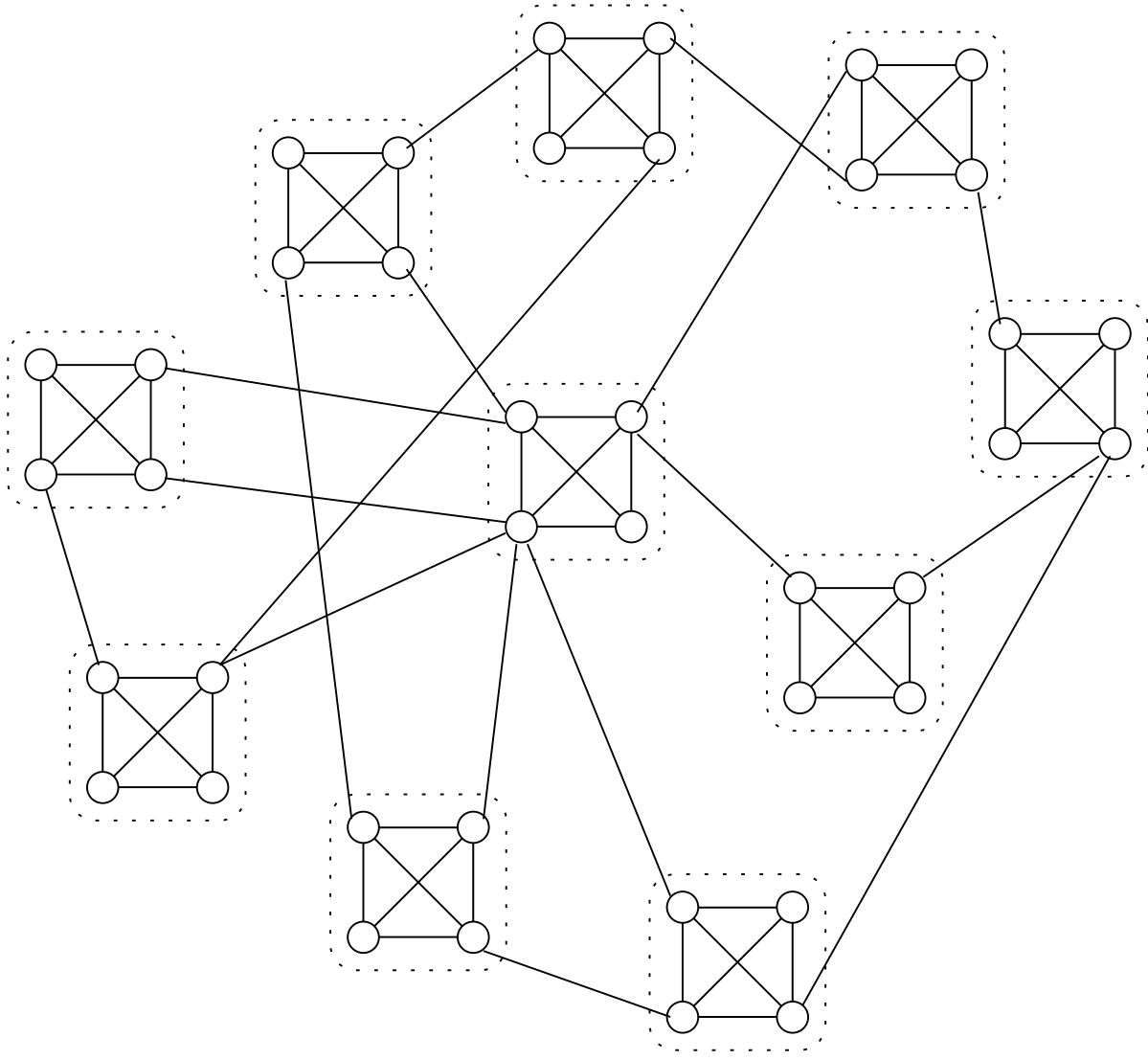


Final proportion infected by global epidemic, z, when $I \sim \mathrm{Exp}(1)$, for constant-degree network with $d = 3$.

Interlude

- We have seen a model which extends the standard network SIR epidemic model by incorporating casual, homogeneously mixing contacts.
  - Casual contacts can have major impact on outcome of epidemic.
  - Letting $\lambda_G \downarrow 0$ yields useful heuristic approach for studying final outcome of epidemics without casual contacts.
  - Kiss et al. (2006) and related deterministic approximations may be poor, particularly when μ_D is small and epidemic is close to threshold.
- Now we examine another extension where we treat the network as a global structure and allow for local structure in the form of households.

An example



Model

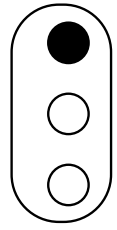
- SIR dynamics, closed population.
- The population structure:
  - A fixed number, m, of households of size n.
  - A distribution, D, with $\mathbb{P}(D = k) = p_k$, $k \geq 0$, $\mu_D < \infty$. Recall that $\tilde{D}$ satisfies $\mathbb{P}(\tilde{D} = k) = kp_k/\mu_D$.
- The epidemic:
  - A single initial infective.
  - Infectious period distribution I, with Laplace transform $\phi(\theta) = \mathbb{E}[e^{-\theta I}]$, $\theta \geq 0$.
  - Whilst infective, individuals contact each of their local (global) neighbours at the points of Poisson processes with rate λ_L (λ_G).

Construction of network

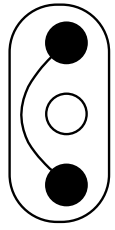
- Put down the $N = mn$ individuals (in their households).
- Construct the network of global contacts according to the configuration model described earlier: Assign each individual a degree from D and pair up half-edges uniformly at random

Recall that the density of imperfections in the graph (parallel edges, self-loops) tends to 0 as m increases. This is also true of the probability that two edges emanating from the same household lead to individuals who also share a household.

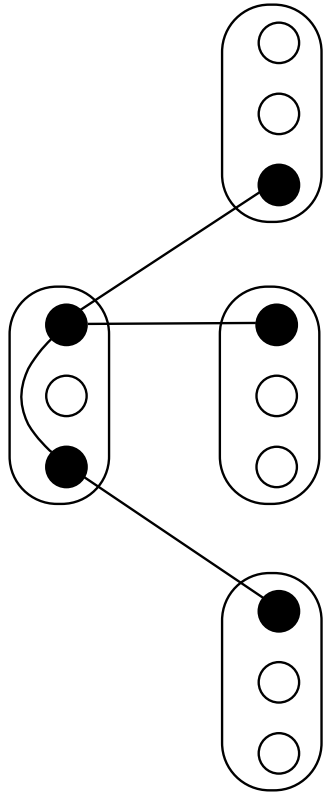
Infection spread



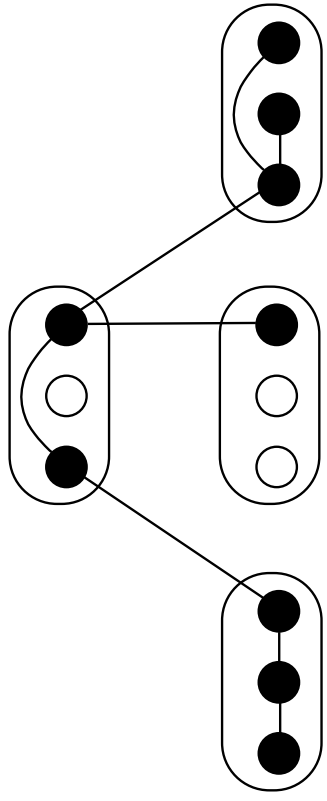
Infection spread



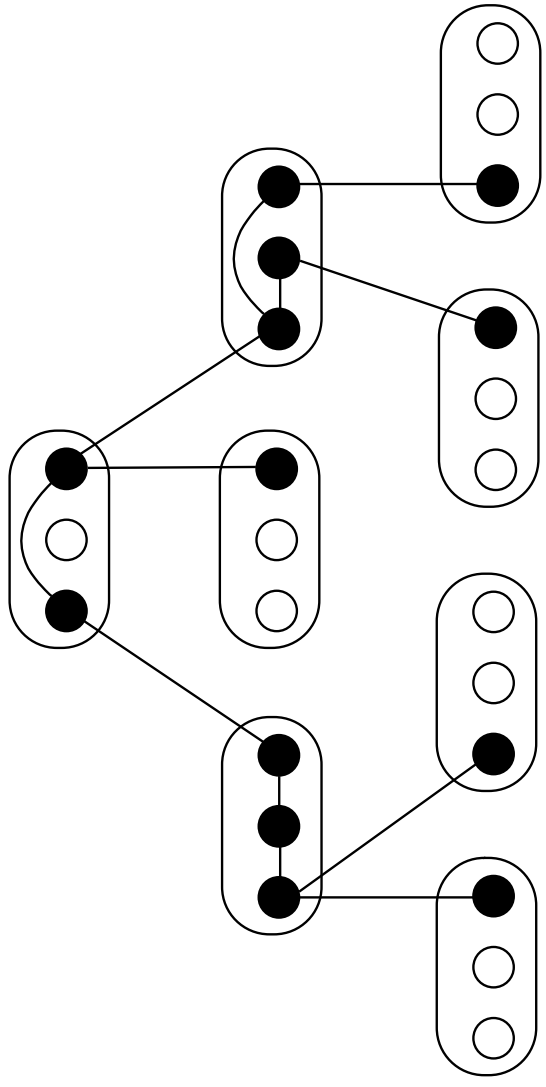
Infection spread



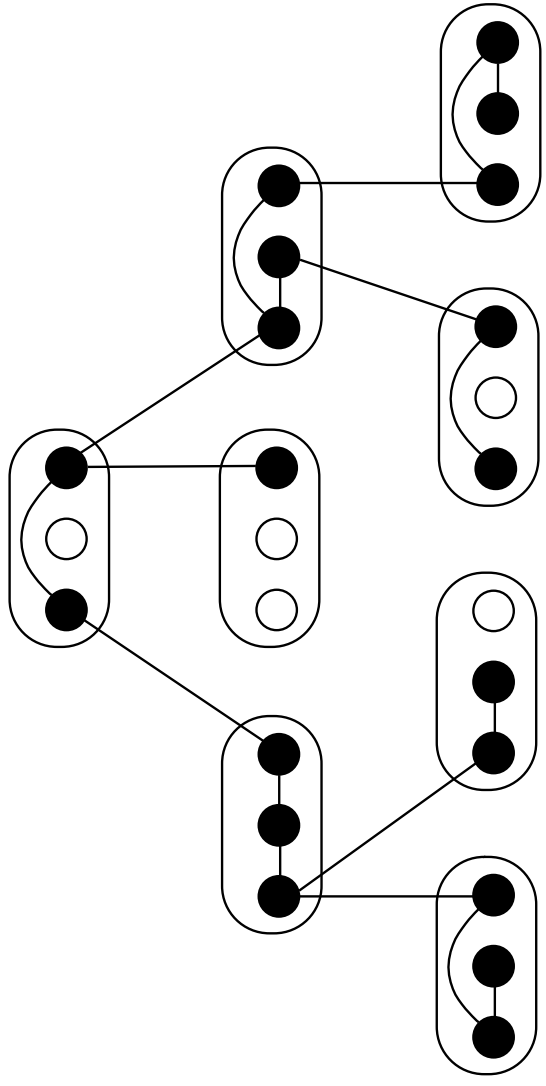
Infection spread



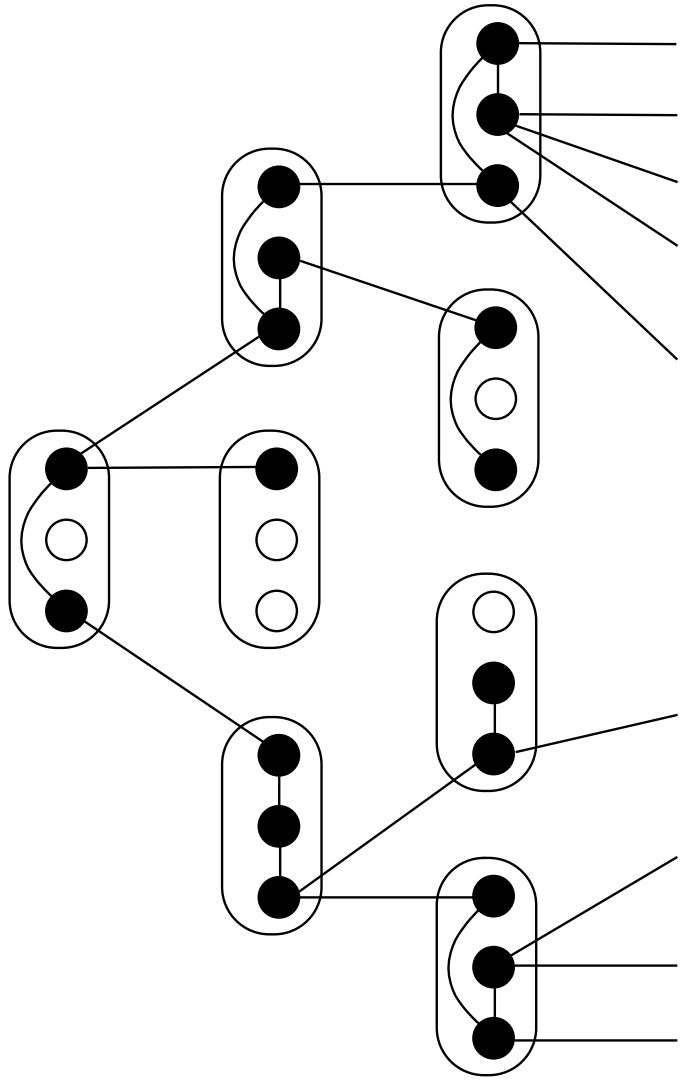
Infection spread



Infection spread



Infection spread



Early stages

- Initially the number of infected households behaves like a branching process (exact as $m \to \infty$).
- The offspring of a household are the households it infects.
- We must determine the distribution of the number of successful global infectious contacts emanating from a household with a single initial infective who was infected globally. Denote by C a random variable with this distribution.
- A major outbreak is possible iff $R_* = \mathbb{E}\, C > 1$.
- The probability of a major outbreak occuring can be determined from the PGF of C.

Early stages—threshold parameter

- Letting T be the final size of a local epidemic amongst $n-1$ initial susceptibles, we find that

$$\begin{aligned} R_* &= \mathbb{E}\,C_0 + \mathbb{E}\,T\,\mathbb{E}\,C_1 \\ &= \left(\mu_{\tilde{D}-1} + \mu_T\mu_D\right)(1-\phi(\lambda_G)) \\ &= \left(\mu_D(\mu_T+1) + \frac{\sigma_D^2}{\mu_D} - 1\right)(1-\phi(\lambda_G)), \end{aligned}$$

  since $\mathbb{E}\,\tilde{D} = \mathbb{E}\,D + \mathrm{Var}\,D/\,\mathbb{E}\,D$.

- Unless n is very small any analytical formula for μ_T is very complicated, so we evaluate this quantity numerically.

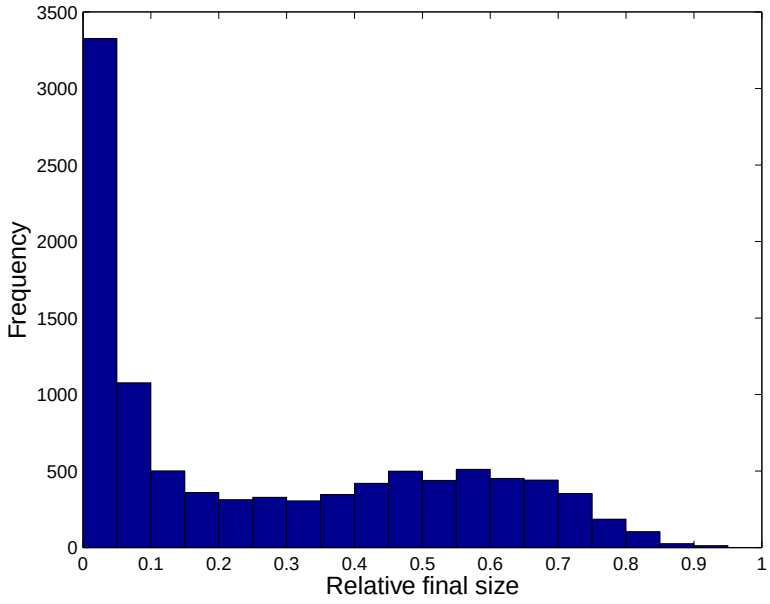
Major outbreak probability

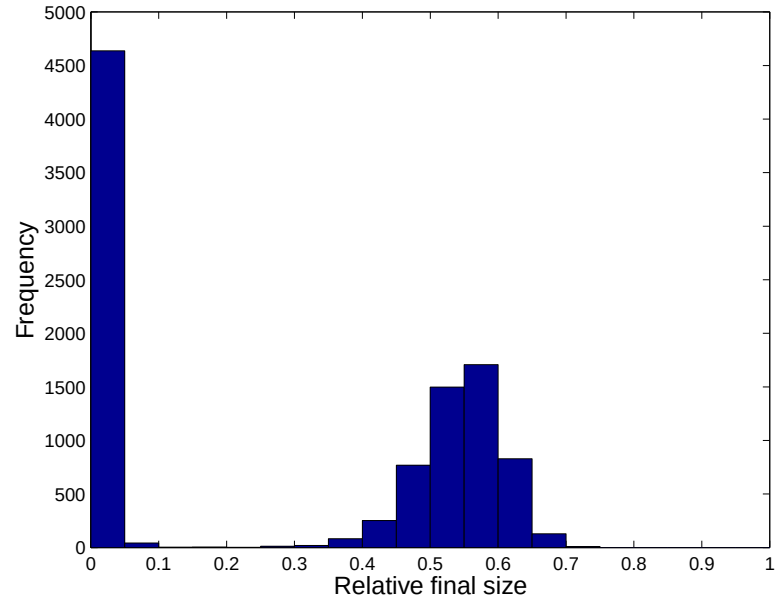
- The probability of a major outbreak is approximated by the probability that the branching process avoids extinction, determined from the PGF of the offspring distribution.
- Dependencies between the number of global infections made by individuals within a household complicates matters somewhat.
- The ‘final state random variable’ framework of Ball and O’Neill (1999) provides a methodology for us to calculate the desired PGF numerically.
- Note that due to the size-biasing effect the first generation has a different offspring distribution.

Susceptibility sets and final size

- Susceptibility set size is again important in determining the mean final size of a major outbreak.
- We can construct the susceptibility set of an individual by ‘generations’ in a manner similar to our analysis of the early stages of the epidemic.
- This leads to a branching process approximation for the size of an individual’s susceptibility set in the limit as $m \to \infty$.
- The offspring distribution for this BP is the same as the distribution of the number of individuals that make global contact with the members of a given individual’s local susceptibility set.

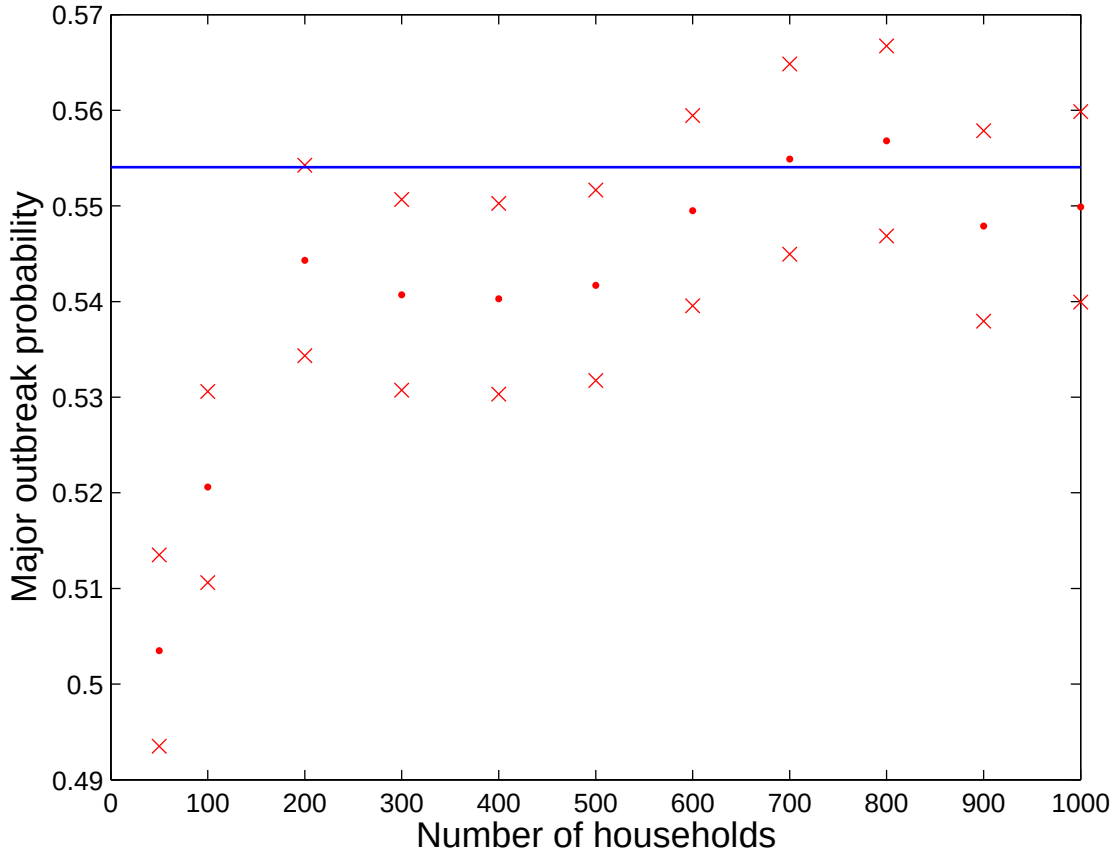
Numerical results





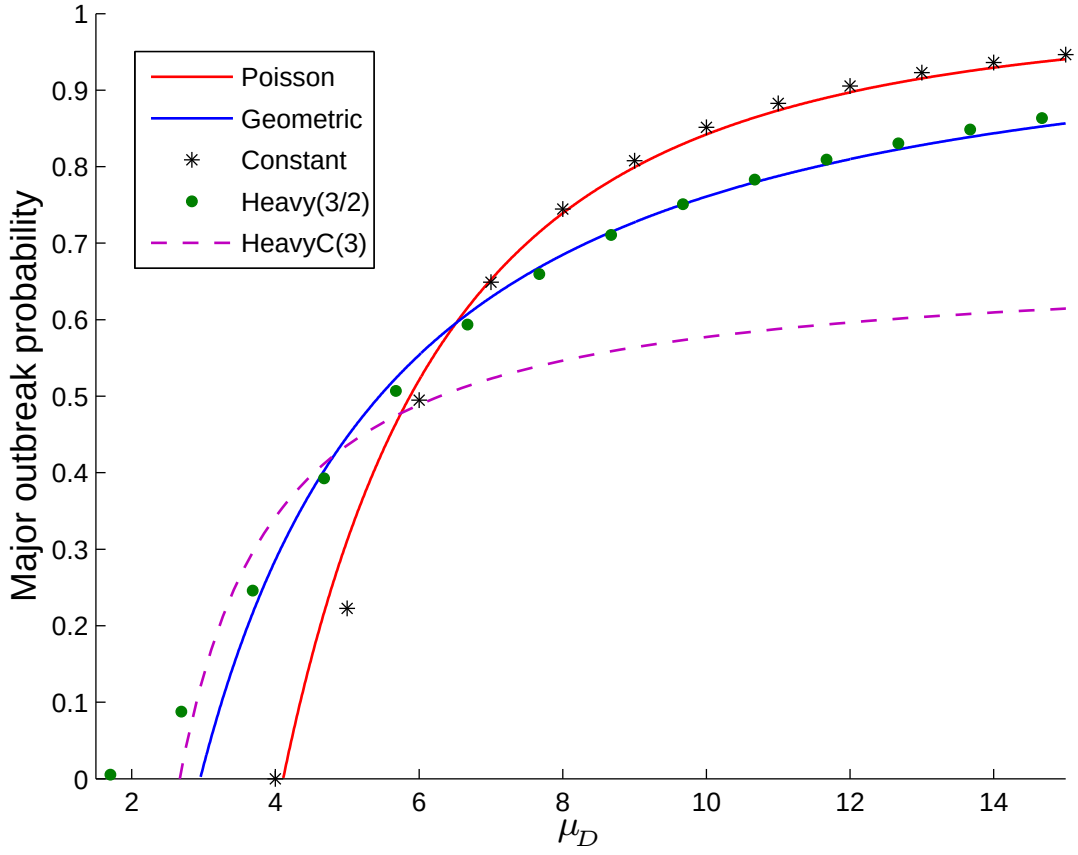
Histograms of relative final sizes for 10,000 simulations of the model with $H \sim \mathrm{U}(\{1,2,3\})$, $D \sim \mathrm{Geom}(3/4)$, $I \equiv 1$, $\lambda_L = 2$, $\lambda_G = 1/4$, on networks of 20 and 200 households.

Numerical results



Simulation-based estimates of major outbreak probability against number of households, together with asymptotic value, for the model with $H \sim \mathrm{U}(\{1,2,3\})$, $D \sim \mathrm{Geom}(3/4)$, $I \equiv 1$, $\lambda_L = 2$, $\lambda_G = 1/4$. Each estimate is based on 10,000 simulations and the plot shows the sample proportions $\pm$ 2SE.

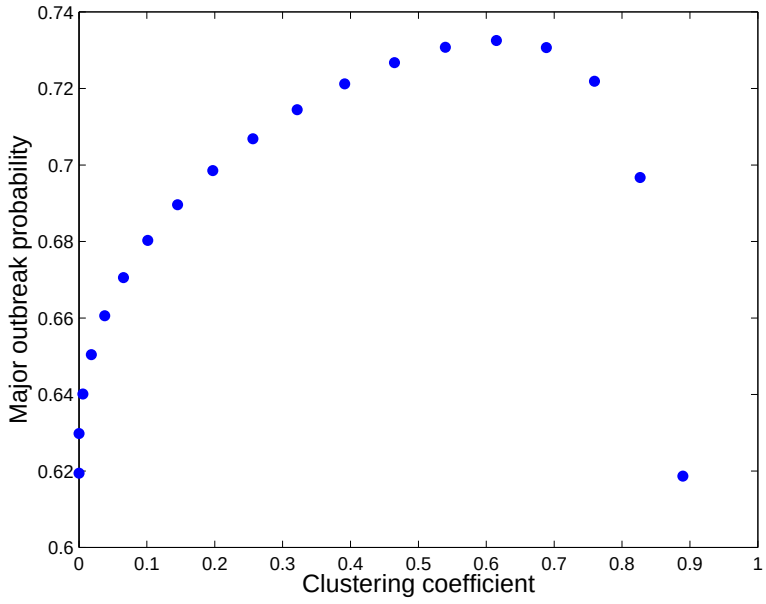
Numerical results

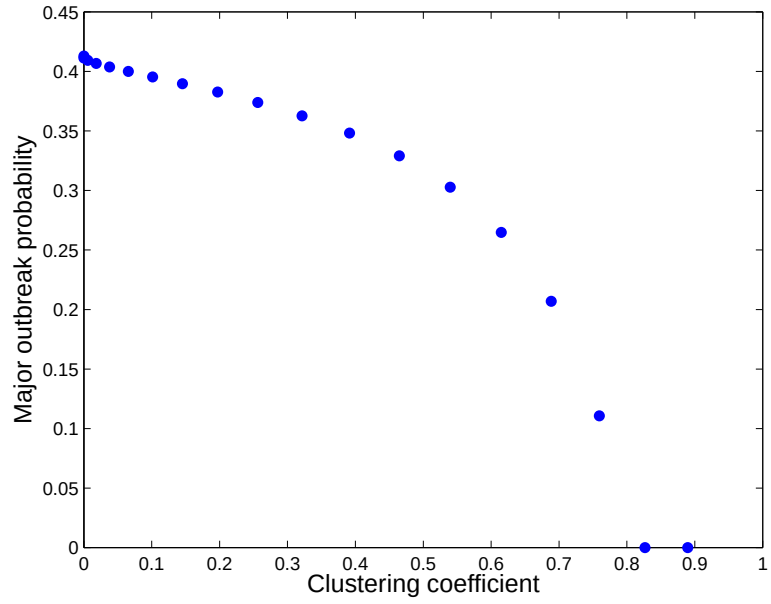


Major outbreak probability dependence on D.

Other parameters are $H \equiv 3$, $I \equiv 1$, $\lambda_L = 1$, $\lambda_G = 1/10$.

Numerical results





Plots showing the effect of clustering on major outbreak probability, with expected number of neighbours of an individual remaining constant.

(a) $I \equiv 1$, $\lambda_L = \lambda_G = 1/10$, average 20 neighbours, D geometric.

(b) $I \equiv 1$, $\lambda_L = \lambda_G = 1/15$, average 20 neighbours, D Poisson.

Vaccination

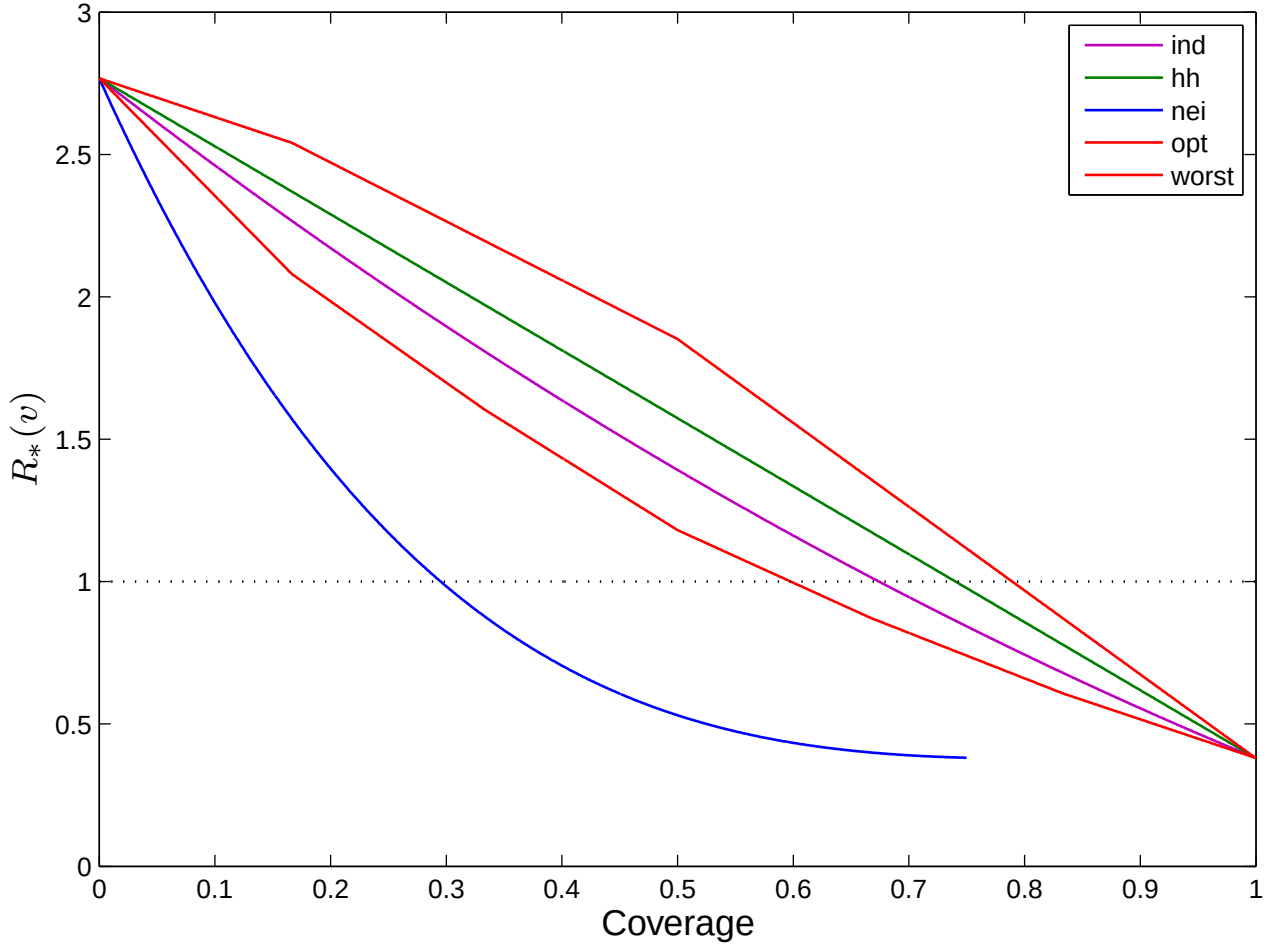
We also have results which take account of the effect of an all-or-nothing vaccine on disease spread for this model with several vaccine allocation strategies:

- vaccinate individuals uniformly at random,
- vaccinate whole households uniformly at random,
- optimal/worst household size based allocation and
- acquaintance vaccination.

The analysis is very similar to that for the original model and can in principle be extended to deal with more complex vaccine action models.

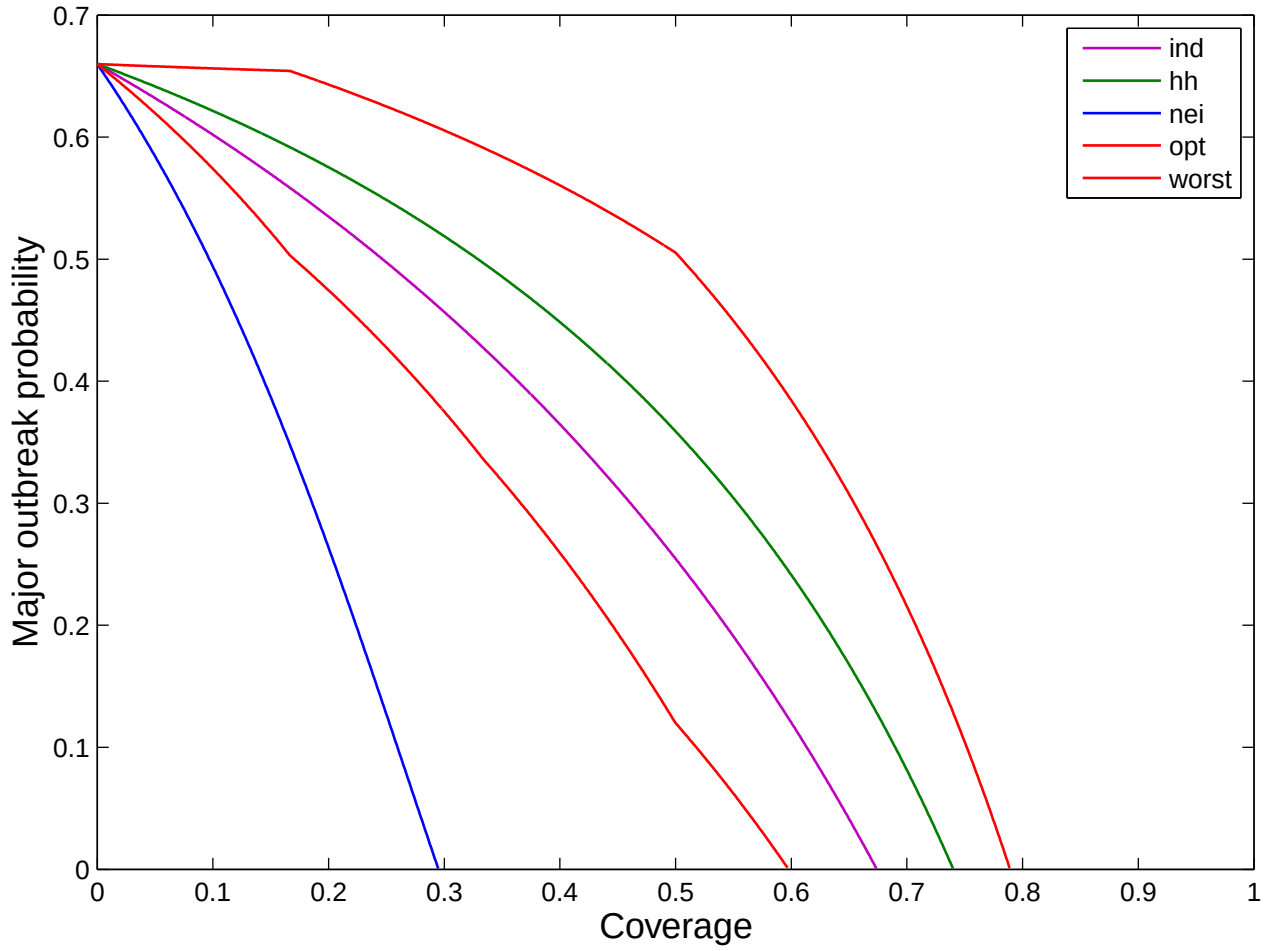
cf. Ball and Lyne (2006); Britton, Janson and Martin-Löf (2007)

Numerical results



Potential effects of an all-or-nothing vaccine with efficacy $\varepsilon = 0.8$ on the threshold parameter. Other parameters are $H \sim \mathrm{U}(\{1,2,3\})$, $D \sim \mathrm{Geom}(3/4)$, $I \equiv 1$, $\lambda_L = 2$, $\lambda_G = 1/3$.

Numerical results



Potential effects of an all-or-nothing vaccine with efficacy $\varepsilon = 0.8$ on major outbreak probability. Other parameters are $H \sim \mathrm{U}(\{1,2,3\})$, $D \sim \mathrm{Geom}(3/4)$, $I \equiv 1$, $\lambda_L = 2$, $\lambda_G = 1/3$.

Summary / discussion

- We have seen two extensions of the standard SIR network epidemic model—incorporating casual contacts and household structure.
- These models capture important real-life departures from homogeneous mixing, both of which have a significant impact on model behaviour and performance of vaccination strategies, whilst retaining mathematical tractability.
- What further departures from homogeneous mixing can be considered that are susceptible to analysis?

References

Andersson (1999), Epidemic models and social networks. *Math. Sci.* 24:128–147.

Ball and Lyne (2006), Optimal vaccination schemes for epidemics among a population of households, with application to variola minor in Brazil, *Stat. Methods Med. Res.* 15:481–497.

Ball and Neal (2002), A general model for stochastic SIR epidemics with two levels of mixing, *Math. Biosci.* 180:73–102.

Ball and Neal (2008a), Network epidemic models with two levels of mixing, *Math. Biosci.* 212:69–87.

Ball and Neal (2008b), Epidemics upon random graphs with two levels of mixing, in preparation.

Ball and O'Neill (1999), The distribution of general final state random variables for stochastic epidemic models, *J. Appl. Probab.* 36:473âĂŞ491.

Bollobás, B. (2001), Random graphs, Cambridge University Press.

References

Britton, Janson and Martin-Löf (2007), Graphs with specified degree distributions, simple epidemics and local vaccination strategies, *Adv. Appl. Probab.* 39:922–948.

Diekmann, De Jong and Metz (1998), A deterministic epidemic model taking account of repeated contacts between the same individuals, *J. Appl. Probab.* 35:448–462.

Kiss, Green and Kao (2006), The effect of contact heterogeniety and multiple routes of transmission on final epidemic size, *Math. Biosci.* 203:124–136.

May and Lloyd (2001), Infection dynamics on scale-free networks, *Phys. Rev. E* 64:066112.

Moreno, Pastor-Satorras and Vespignani (2002), Epidemic outbreaks in complex heterogeneous networks, *Eur. Phys. J. (B)* 26:521–529.

Newman (2002), Spread of epidemic disease on networks, *Phys. Rev. E* 66:016128.

Scalia-Tomba (1985), Asymptotic final-size distribution for some chain-binomial processes, *Adv. in Appl. Probab.* 17:477–495.