# ACQUAINTANCE VACCINATION IN AN EPIDEMIC ON A RANDOM GRAPH WITH SPECIFIED DEGREE DISTRIBUTION

FRANK BALL,* *University of Nottingham*

DAVID SIRL,** *Loughborough University*

## Abstract

We consider a stochastic SIR (susceptible → infective → removed) epidemic on a random graph with specified degree distribution, constructed using the configuration model, and investigate the 'acquaintance vaccination' method for targeting individuals of high degree for vaccination. Branching process approximations are developed which yield a post-vaccination threshold parameter, and the asymptotic (large population) probability and final size of a major outbreak. We find that introducing an imperfect vaccine response into the present model for acquaintance vaccination leads to sibling dependence in the approximating branching processes, which may then require infinite type spaces for their analysis and are generally not amenable to numerical calculation. Thus, we propose and analyse an alternative model for acquaintance vaccination, which avoids these difficulties. The theory is illustrated by a brief numerical study, which suggests that the two models for acquaintance vaccination yield quantitatively very similar disease properties.

*Keywords:* Branching process; epidemic process; final size; network; random graph; threshold behaviour; vaccination

2010 Mathematics Subject Classification: Primary 92D30
Secondary 05C80; 60J80

## 1. Introduction

In the past decade there has been an abundance of research published concerning the modelling of stochastic epidemics spreading on networks of 'individuals' connected by 'relationships' which in some way model the social networks that people establish with, for example, their families, friends, coworkers, and classmates. Early work on epidemic models on random graphs includes that of Diekmann *et al.* (1998), Andersson (1997), (1998), (1999), and also Newman (2002) (see also Kenah and Robins (2007)). Epidemics evolving on extensions of the basic *configuration model* (see, for example, Durrett (2007, Chapter 3)) have been studied by, for example, Kiss *et al.* (2006), Ball and Neal (2008), Ball *et al.* (2009), and Ball and Sirl (2012). Alternative random graph structures have been considered by, amongst others, Newman (2002, Section V), Britton *et al.* (2008), and Ball *et al.* (2013).

In this paper we address the question of vaccination in what we refer to as the standard network model (SNM), an SIR (susceptible → infective → removed) epidemic model on a network structure specified via the configuration model. Vaccination of individuals in these

network models seems to have first been investigated by Cohen *et al.* (2003); their work was extended and put into a more rigorous framework by Britton *et al.* (2007). However, these papers considered only perfect vaccines, where vaccination of an individual renders them completely immune to the disease, and assumed a fixed infectious period. The primary motivation for this paper is to derive analogous results for other vaccine action models where vaccination might confer only a partial resistance to infection or where there might be several possible responses to vaccination (see Becker and Starczak (1998)). In addition to allowing for individuals chosen uniformly at random to be vaccinated, the above-mentioned works of Cohen *et al.* (2003) and Britton *et al.* (2007) focused on modelling *acquaintance vaccination*, where neighbours of individuals chosen uniformly at random from the population are vaccinated. This has the effect that vaccinated individuals tend to have a higher degree and this targeting of well-connected individuals increases the effectiveness of vaccination.

We continue the work of Cohen *et al.* (2003) and Britton *et al.* (2007) by seeking to extend their results to allow for more general infectious period distributions and vaccine action models. The key tool in the analysis is branching process approximation, which enables the probability and relative final size of a major outbreak to be approximately determined for large networks. We find that, with an imperfect vaccine response, the models of Cohen *et al.* (2003) and Britton *et al.* (2007) for acquaintance vaccination lead to approximating branching processes with sibling dependence (cf. Olofsson (1996)), which may then require infinite type spaces for their analysis and are thus not generally amenable to numerical calculation. Thus, we modify the model of acquaintance vaccination slightly to avoid these difficulties. We find that the performance of the modified model is quantitatively very similar to that of the original model.

The remainder of the paper is organised as follows. In Section 2 we recall the definition of the basic SNM and its analysis as well as outlining the vaccine action models we use. In Section 3 we extend the above-mentioned analysis of Britton *et al.* (2007) to include nonconstant infectious period distributions and imperfect vaccines. In order to compare the quantitative properties of the two acquaintance vaccination models, we nevertheless analyse the original model with an imperfect vaccine. In Section 4 we propose a modification of the acquaintance vaccination model and completely analyse the basic final size properties of the SNM with this vaccination model (see Propositions 2–4). In Section 5, the performance of the old and new acquaintance vaccination models are illustrated numerically, and compared with each other and with that of uniform vaccination. We also compare these with the performance of an optimal vaccination scheme (analysed in Appendix B), in which we assume that the degrees of all individuals are known and the individuals of high degree are vaccinated.

## 1.1. Notation

We denote by $\mathbb{R}_+$ and $\mathbb{Z}_+$ the nonnegative real numbers and integers, respectively. For any $x \in \mathbb{R}_+$, we define the usual floor function $\lfloor x \rfloor = \max\{k \in \mathbb{Z}_+ : k \leq x\}$. For suitable vectors $\boldsymbol{x} = (x_1, x_2, \ldots, x_\ell)$ and $\boldsymbol{y} = (y_1, y_2, \ldots, y_\ell)$, we define $\boldsymbol{x}! = \prod_{i=1}^{\ell} x_i!$, $\boldsymbol{x}^{\boldsymbol{y}} = \prod_{i=1}^{\ell} x_i^{y_i}$, and $\binom{\boldsymbol{x}}{\boldsymbol{y}} = \prod_{i=1}^{\ell} \binom{x_i}{y_i}$, and we say that $\boldsymbol{x} \leq \boldsymbol{y}$ if the inequality holds componentwise. We also adopt the convention that summations over vector indices are rectangular, i.e. $\sum_{\boldsymbol{i}=\boldsymbol{x}}^{\boldsymbol{y}} = \sum_{i_1=x_1}^{y_1} \cdots \sum_{i_\ell=x_\ell}^{y_\ell}$. We denote the probability generating function (PGF) of a random vector $\boldsymbol{X} = (X_1, X_2, \ldots, X_\ell)$ taking values in the range $\mathcal{R} \subseteq \mathbb{Z}_+^\ell$ by $f_{\boldsymbol{X}}(\boldsymbol{s}) = \sum_{\boldsymbol{x} \in \mathcal{R}} \mathbb{P}(\boldsymbol{X} = \boldsymbol{x}) \boldsymbol{s}^{\boldsymbol{x}}$, $\boldsymbol{s} \in [0, 1]^\ell$, and, for any $\boldsymbol{k} = (k_1, k_2, \ldots, k_\ell) \in \mathbb{Z}_+^\ell$, we denote by $f^{(\boldsymbol{k})}(\boldsymbol{s})$ the mixed derivative of $f(\boldsymbol{s})$ of order $k_j$ with respect to $s_j$, $j \in \{1, 2, \ldots, \ell\}$. Of course, for a scalar random variable $X$ taking values in $\mathbb{Z}_+$, this notation simplifies to $f_X^{(k)}(s)$, $s \in [0, 1]$, $k = 0, 1, \ldots$.

For a scalar random variable $X$, we write $\mu_X$ and $\sigma_X^2$ for its mean and variance, respectively, and we denote by $\mathbf{1}$ a vector of appropriate length with all entries equal to 1.

## 2. Epidemic and vaccine models

### 2.1. Standard network model

We consider an epidemic evolving amongst a closed population of $n$ individuals. We specify a degree distribution $D$ by its mass function $p_k = \mathbb{P}(D = k)$, $k = 0, 1, 2, \ldots$, and assume that the mean ($\mu_D$) of $D$ is finite. To construct the graph describing possible infectious contacts, we follow Newman *et al.* (2001), so each individual is assigned a number of *half-edges* according to independent draws from $D$. We then pair these half-edges uniformly at random to form the edges of the random graph. (If the total number of half-edges is odd then the leftover half-edge is ignored.) The epidemic is defined as follows. Each individual is assumed to be either susceptible, infectious, or removed. We assume that initially all individuals are susceptible except for a single individual, chosen uniformly at random from the population, who is infective. (Extensions to other initial conditions, in which the number of initial infectives remains fixed as $n \to \infty$, are easily treated.) An infective individual remains so for a random time determined by a realisation of a random variable $I$, most conveniently specified by its Laplace transform $\phi(\theta) = \mathbb{E}[e^{-\theta I}]$, $\theta \geq 0$. At the end of its infectious period an infectious individual becomes removed, whence it plays no further role in the epidemic. During its infectious period an individual makes infectious contacts with any given neighbour in the random graph at the points of a Poisson process of rate $\lambda$. If an individual so contacted is susceptible then it becomes infectious; otherwise, nothing happens. We assume that the contact processes and infectious periods associated with every individual are all independent. The epidemic continues until there is no infective remaining in the population.

2.1.1. *Forward process.* We analyse the early stages of epidemic spread by approximating the number of infected individuals with a (forward) branching process. The ancestor (generation zero) of the branching process corresponds to the initially infected individual. Then generation $k$ comprises those previously uninfected individuals who are contacted by individuals in generation $k - 1$. The branching property arises because, in the early stages of the epidemic, every individual contacted by an infective is susceptible with probability tending to 1 as $n \to \infty$. The offspring distribution for the initial generation of the branching process is generally different from that of all subsequent generations. The initial infective is chosen uniformly at random from the population, and, hence, has degree distribution $D$. Subsequent infectives are chosen as the individual from which a randomly chosen half-edge emanates, and, hence, the number of uninfected neighbours of subsequent infectives is distributed as $\tilde{D} - 1$, where $\tilde{D}$ is the size-biased version of $D$ defined by the mass function $\tilde{p}_k = k p_k / \mu_D$, $k = 1, 2, \ldots$. The approximating branching process is thus defined by the distributions of $C$, the number of neighbours infected by a typical individual chosen uniformly at random from the population and of $\tilde{C}$, the number of neighbours infected by an individual that is infected by one of its neighbours. This often-used approximation is made fully rigorous, for a more general (network and households) model, in Ball *et al.* (2009).

The branching process approximation for the early stages of the epidemic justifies the use of $R_0 = \mathbb{E}[\tilde{C}]$ as a threshold parameter for the epidemic model, determining whether major outbreaks (that infect at least $\log n$ individuals) are possible. It is easy to show (see,

e.g. Andersson (1999)) that

$$R_0 = \mathbb{E}[\tilde{C}] = \mu_{\tilde{D}-1} p^I = \left( \mu_D + \frac{\sigma_D^2}{\mu_D} - 1 \right) p^I, \tag{1}$$

where $p^I = 1 - \phi(\lambda)$ is the marginal probability that an infected individual makes infectious contact with a given neighbour. As $n \to \infty$, the probability of a major outbreak converges to the survival probability of the forward branching process, so such outbreaks occur with nonzero probability if and only if $R_0 > 1$. Note that if $\sigma_D^2 = \infty$ then $R_0 = \infty$ and, as long as $p^I > 0$, the limiting probability of a major outbreak is strictly positive.

To find the probability $p_{\mathrm{maj}}$ of a major outbreak, it is necessary to determine the PGFs $f_C$ and $f_{\tilde{C}}$. Conditioning on the infectious period and degree of the individual of interest yields

$$f_C(s) = \sum_{i=0}^{\infty} \frac{(1-s)^i \phi(i\lambda)}{i!} f_D^{(i)}(s) \quad \text{and} \quad f_{\tilde{C}}(s) = \sum_{i=0}^{\infty} \frac{(1-s)^i \phi(i\lambda)}{i!} f_{\tilde{D}-1}^{(i)}(s).$$

Standard branching process theory then gives $p_{\mathrm{maj}} = 1 - f_C(\sigma)$, where $\sigma$ is the smallest solution of $f_{\tilde{C}}(s) = s$ in $[0, 1]$ (cf. Kenah and Robins (2007, Section III.B.1)).

2.1.2. *Backward process.* We can also calculate the asymptotic (as $n \to \infty$) expected relative final size of a major outbreak, that is, the proportion of individuals infected by a major outbreak, as the survival probability of a branching process which approximates an individual's *susceptibility set*, defined as follows. For every individual in the population, suppose that it is infected, choose its infectious period using the distribution of $I$, and determine, using independent Poisson processes with rate $\lambda$, which of its neighbours it would infect. Now construct a directed graph in which nodes represent individuals and an arc from individual $i$ to individual $j$ is present if and only if individual $i$, were it to be infected, would make infectious contact with individual $j$. The susceptibility set of an individual, $i^*$ say, comprises all individuals $j$ from which there is a path to $i^*$ in this (random, directed) graph, including $i^*$ itself. Observe that $i^*$ is infected by the epidemic if and only if its susceptibility set includes the initial infective.

Suppose that $i^*$ is chosen uniformly at random from the population. The (backward) branching process that approximates $i^*$'s susceptibility set is similar to the forward process describing the early stages of epidemic spread, but rather than considering how many neighbours an individual makes infectious contact with we consider how many neighbours make infectious contact with that individual. Individual $i^*$ is the ancestor (zeroth generation) of the branching process and generation $k$ consists of the individuals who join the susceptibility set by way of an infectious contact with an individual in generation $k - 1$. The offspring distributions, which define the branching process, are thus the distributions of $B$, the number of neighbours of $i^*$ that, if infected, would make contact with $i^*$, and $\tilde{B}$, the number of neighbours (excluding $i^*$) of one of these contacting individuals, $j^*$ say, that, if infected, would make contact with $j^*$. Similarly to the forward process, $B$ describes the offspring distribution for the initial generation of the branching process and $\tilde{B}$ that for subsequent generations. As each neighbour of a given individual makes infectious contact with that individual independently with probability $p^I$, the offspring PGFs are

$$f_B(s) = f_D(1 - p^I + p^I s) \quad \text{and} \quad f_{\tilde{B}}(s) = f_{\tilde{D}-1}(1 - p^I + p^I s).$$

Let $T^{(n)}$ be the number of initial susceptibles ultimately infected by a major outbreak. Then, as proved formally for a more general model in Ball *et al.* (2009), $\mathbb{E}[T^{(n)}]/n \to z$ as

$n \to \infty$, where $z$ is the survival probability of the backward branching process. Furthermore, the argument of Ball *et al.* (2013) (who consider an SIR epidemic on a random intersection graph) can be used to show that $T^{(n)}/n \to z$ in probability as $n \to \infty$. Standard branching process theory yields $z = 1 - f_B(\xi)$, where $\xi$ is the smallest solution of $f_{\tilde{B}}(s) = s$ in $[0, 1]$ (cf. Andersson (1999, Equations (11) and (12)) and Newman (2002, Equations (25) and (26))). Thus, with probability tending to 1 as $n \to \infty$, a major outbreak infects exact order $n$ (and not just $\log n$) individuals.

## 2.2. Vaccination

We work within the quite general framework of the vaccine action model proposed by Becker and Starczak (1998), in which the vaccine responses of individuals are described by independent copies of a random vector $(A, B)$ taking values in $\mathbb{R}_+^2$ (usually in $[0, 1]^2$). Here $A$ denotes the relative susceptibility (compared to a nonvaccinated individual) and $B$ denotes the relative infectivity of the vaccinated individual if he/she becomes infected. Thus, all Poisson processes concerning potential infection of the individual have their rates multiplied by $A$ and the Poisson processes governing infection by the individual have their rates multiplied by $B$. Becker and Starczak (1998) allowed for a finite number of possible outcomes of $(A, B)$, but we restrict attention to two special cases. The first of these is the *nonrandom* vaccine, where $\mathbb{P}((A, B) = (a, b)) = 1$ for some given $(a, b) \in \mathbb{R}_+^2$ and the vaccine has the same effect on every vaccinated individual; the second is the *all-or-nothing* vaccine (see Halloran *et al.* (1992)), where $\mathbb{P}((A, B) = (0, 0)) = 1 - \mathbb{P}((A, B) = (1, 1)) = \varepsilon$, so vaccinated individuals are rendered completely immune with probability $\varepsilon$ and otherwise the vaccine has no effect. If a nonrandom vaccine has $b = 1$ then it is called *leaky* (see Halloran *et al.* (1992)). A vaccine which renders all vaccinees completely immune is called *perfect*.

Vaccine allocation models are the primary focus of this paper. Perhaps the simplest strategy is to vaccinate individuals chosen uniformly at random from the population. Cohen *et al.* (2003) introduced the idea of *acquaintance* vaccination, where individuals are chosen uniformly at random from the population, but rather than vaccinating that individual, they are asked to name an acquaintance (a neighbour in the random graph) and that person is vaccinated. As we detail in Section 3, this has the effect of targeting individuals of higher degree and, therefore, has more impact on disease spread. In either situation an important quantity is the vaccine *coverage* $c$, the proportion of the population that is vaccinated.

Analysis of the case where individuals to be vaccinated are chosen uniformly at random from the population is relatively straightforward, since the vaccination status of an individual is independent of its degree. Note that, with a perfect vaccine and coverage $c$, $R_0$ is reduced to $(1 - c)R_0$; so, from (1), if $\sigma_D^2 = \infty$ then the critical vaccination coverage to eliminate the possibility of a major outbreak is 1.

## 3. Original acquaintance vaccination model

In this vaccine allocation model we sample individuals uniformly at random from the population and vaccinate a neighbour chosen uniformly at random. Thus, the degree of a vaccinated individual is distributed as $\tilde{D}$ rather than $D$. An individual that is sampled but has no neighbours is ignored and an individual is vaccinated only once, even if it is chosen to be vaccinated more than once. Adopting most of the language and notation of Britton *et al.* (2007), assume that we sample (with replacement) a random number of individuals with a Poisson distribution with mean $\kappa n$ for some $\kappa \geq 0$. The corresponding vaccine coverage $c(\kappa)$

is easily shown to be $1 - f_D(\alpha)$, where

$$\alpha = \sum_{d=1}^{\infty} e^{-\kappa/d} \tilde{p}_d \tag{2}$$

is the probability that an individual is not named by a given neighbour.

### 3.1. Perfect vaccine

The approximating forward branching process involves only unvaccinated (i.e. unnamed) individuals, since the vaccine is perfect. Consider a typical individual in a noninitial generation of this branching process and denote its degree by $\tilde{D}_U$. Unconditionally, this individual's degree is distributed as $\tilde{D}$, but we know that this individual (i) is unvaccinated and (ii) does not name its infector for vaccination. Denote these events by $U$ and $N^c$, respectively. Hence, for $d = 1, 2, \ldots,$

$$\mathbb{P}(\tilde{D}_U = d) = \mathbb{P}(\tilde{D} = d \mid U, N^c) = \frac{\mathbb{P}(\tilde{D} = d, U, N^c)}{\mathbb{P}(U, N^c)} = \frac{\tilde{p}_d \alpha^{d-1} e^{-\kappa/d}}{\sum_{k=1}^{\infty} \tilde{p}_k \alpha^{k-1} e^{-\kappa/k}}$$

and

$$\mathbb{P}(U \mid N^c) = \frac{\sum_{k=1}^{\infty} \tilde{p}_k \alpha^{k-1} e^{-\kappa/k}}{\sum_{k=1}^{\infty} \tilde{p}_k e^{-\kappa/k}} = \sum_{k=1}^{\infty} \tilde{p}_k \alpha^{k-2} e^{-\kappa/k}.$$

Then the mean of the offspring random variable $\tilde{C}_v$ in this forward branching process is given by

$$R_v = \mathbb{E}[\tilde{C}_v] = \sum_{d=1}^{\infty} \mathbb{P}(\tilde{D}_U = d)(d-1) e^{-\kappa/d} p^I \mathbb{P}(U \mid N^c) = p^I \sum_{d=2}^{\infty} \tilde{p}_d \alpha^{d-2} e^{-2\kappa/d}(d-1). \tag{3}$$

This is the same threshold parameter that appears as Equation (3.13) of Britton *et al.* (2007), though we use a different branching process approximation. (We find that our branching process approximation is easier to generalise than that given in Britton *et al.* (2007).) Note from (3) that, in contrast to uniform vaccination, the critical vaccination coverage is strictly less than 1 when $\sigma_D^2 = \infty$.

The PGFs of $\tilde{C}_v$ and the random variable $C_v$ corresponding to the initial generation may be calculated in a form that is amenable to numerical evaluation but, unless the infectious period is constant, the expressions are cumbersome and give little insight, so we do not present them here.

Turning to the backward process, the same arguments given to justify the calculation of $R_v$ (with 'an individual's infector' replaced by 'the neighbour the individual contacted to join the susceptibility set') can be used to derive the PGF $f_{\tilde{B}_v}$ of the offspring distribution $\tilde{B}_v$ in noninitial generations. Let $i^*$ denote a typical noninitial individual in the susceptibility set. Given $i^*$'s degree $d$ (distributed as $\tilde{D}_U$), each neighbour of $i^*$ (except the one it contacted to join the susceptibility set) joins the susceptibility set independently by making infectious contact with $i^*$ if it (i) is not named by $i^*$, (ii) is not named by any of its other neighbours, and (iii) makes infectious contact with $i^*$. These events are independent and occur with respective

probabilities $\mathrm{e}^{-\kappa/d}$, $\mathbb{P}(U \mid N^c)$, and $p^I$, so

$$
\begin{aligned}
f_{\tilde{B}_v}(s) &= \mathbb{E}[s^{\tilde{B}_v}] \\
&= \mathbb{E}_{\tilde{D}_U}[\mathbb{E}_{\tilde{B}_v}[s^{\tilde{B}_v} \mid \tilde{D}_U]] \\
&= \mathbb{E}_{\tilde{D}_U}[(1 - \mathrm{e}^{-\kappa/\tilde{D}_U}\mathbb{P}(U \mid N^c)p^I + s\mathrm{e}^{-\kappa/\tilde{D}_U}\mathbb{P}(U \mid N^c)p^I)^{\tilde{D}_U - 1}] \\
&= \sum_{d=1}^{\infty} \frac{\tilde{p}_d \alpha^{d-2}\mathrm{e}^{-\kappa/d}}{\mathbb{P}(U \mid N^c)}\{1 - (1-s)\mathrm{e}^{-\kappa/d}\mathbb{P}(U \mid N^c)p^I\}^{d-1}.
\end{aligned}
\tag{4}
$$

For the initial generation, writing $D_U$ for the degree distribution of an unvaccinated individual chosen uniformly at random from the population, we have, for all $d = 0, 1, \ldots, \mathbb{P}(D_U = d) = \mathbb{P}(D = d \mid U) = p_d \alpha^d / f_D(\alpha)$; and arguing as in the derivation of (4) gives

$$
f_{B_v}(s) = \sum_{d=0}^{\infty} \frac{p_d \alpha^d}{f_D(\alpha)}\{1 - (1-s)\mathrm{e}^{-\kappa/d}\mathbb{P}(U \mid N^c)p^I\}^d.
$$

The proportion of the population infected by a major outbreak is $z = f_D(\alpha)(1 - f_{B_v}(\xi))$, where $\xi$ is the smallest solution of $f_{\tilde{B}_v}(s) = s$ in $[0, 1]$. If the infectious period is constant, it is easily verified that $f_{C_v} = f_{B_v}$ and $f_{\tilde{C}_v} = f_{\tilde{B}_v}$, so $p_{\mathrm{maj}} = z$, and that the above expression for $z$ (and, hence, $p_{\mathrm{maj}}$) is the same as that obtained using Equation (3.11) of Britton *et al.* (2007).

### 3.2. Imperfect vaccine

Applying the above approach directly to deal with imperfect vaccines introduces sibling dependence (see Olofsson (1996)) into the branching processes, unless we use multitype branching processes. This is because both the forward and backward processes may now involve vaccinated individuals and, given that an individual is vaccinated, the degrees of its neighbours are not independent (see (5), below). Thus, we type infected individuals according to both their degree and their vaccination status. For every $d = 1, 2, \ldots$, define the types $N_d$, $V_d$, and $U_d$. Here $d$ denotes an individual's degree, $N$ means an individual was named by its infector (and is therefore vaccinated), $V$ means it is not named by its infector but is still vaccinated (i.e. is named by another neighbour), and $U$ means that it is unvaccinated.

We consider a forward process approximating the proliferation of infective individuals in the early stages of an epidemic. We focus on the means of the offspring distributions corresponding to a noninitial generation, with a view to calculating a threshold parameter. Define the mean matrix $M = (m_{t,t'}, t, t' \in \mathcal{T})$, where $\mathcal{T}$ is the type space $\{N_d, V_d, U_d : d = 1, 2, \ldots\}$. Also, define the marginal transmission probabilities $p_{UU}$, $p_{UV}$, $p_{VU}$, and $p_{VV}$ between unvaccinated and vaccinated (or named) individuals. If the vaccine is all-or-nothing then these are given by

$$
P^{\mathrm{AoN}} = \begin{pmatrix} p_{VV}^{\mathrm{AoN}} & p_{VU}^{\mathrm{AoN}} \\ p_{UV}^{\mathrm{AoN}} & p_{UU}^{\mathrm{AoN}} \end{pmatrix} = p^I \begin{pmatrix} 1 - \varepsilon & 1 \\ 1 - \varepsilon & 1 \end{pmatrix}
$$

and if the vaccine is nonrandom then they are given by

$$
P^{\mathrm{NR}} = \begin{pmatrix} p_{VV}^{\mathrm{NR}} & p_{VU}^{\mathrm{NR}} \\ p_{UV}^{\mathrm{NR}} & p_{UU}^{\mathrm{NR}} \end{pmatrix} = \begin{pmatrix} 1 - \phi(ab\lambda) & 1 - \phi(b\lambda) \\ 1 - \phi(a\lambda) & 1 - \phi(\lambda) \end{pmatrix}.
$$

First consider an individual of type $N_d$. The degrees of its $d - 1$ uninfected neighbours are independent copies of $\tilde{D}$. Each such neighbour is named with probability $1 - \mathrm{e}^{-\kappa/d}$.

If a neighbour is of degree $d'$ and not named then it is unvaccinated with probability $\alpha^{d'-1}$; otherwise, it is vaccinated. Therefore, for all $d, d' = 1, 2, \ldots$, it can easily be seen that $m_{N_d, N_{d'}} = (d-1)(1 - e^{-\kappa/d}) \tilde{p}_{d'} p_{VV}$, $m_{N_d, V_{d'}} = (d-1) e^{-\kappa/d} \tilde{p}_{d'} (1 - \alpha^{d'-1}) p_{VV}$, and $m_{N_d, U_{d'}} = (d-1) e^{-\kappa/d} \tilde{p}_{d'} \alpha^{d'-1} p_{VU}$.

Next consider an individual of type $U_d$. Since it is unvaccinated, all of its uninfected neighbours failed to name it; so, letting $D_1^U, D_2^U, \ldots, D_{d-1}^U$ denote the degrees of its $d-1$ uninfected neighbours, we have

$$\mathbb{P}(D_1^U = d_1', \, D_2^U = d_2', \ldots, D_{d-1}^U = d_{d-1}')$$
$$= \prod_{i=1}^{d-1} \frac{\tilde{p}_{d_i'} e^{-\kappa/d_i'}}{\alpha}, \qquad d_1', d_2', \ldots, d_{d-1}' = 1, 2, \ldots.$$

The same arguments as above show that

$$m_{U_d, N_{d'}} = (d-1)(1 - e^{-\kappa/d}) \tilde{p}_{d'} e^{-\kappa/d'} \alpha^{-1} p_{UV},$$
$$m_{U_d, V_{d'}} = (d-1) e^{-\kappa/d} \tilde{p}_{d'} e^{-\kappa/d'} \alpha^{-1} (1 - \alpha^{d'-1}) p_{UV},$$

and

$$m_{U_d, U_{d'}} = (d-1) e^{-\kappa/d} \tilde{p}_{d'} e^{-\kappa/d'} \alpha^{d'-2} p_{UU}.$$

Finally, consider a typical individual of type $V_d$ and write $D_1^V, D_2^V, \ldots, D_{d-1}^V$ for the degrees of its $d-1$ uninfected neighbours. Since the individual is of type $V$ we know that at least one of these neighbours named the individual of interest, so it follows that, for $d_1', d_2', \ldots, d_{d-1}' = 1, 2, \ldots$,

$$\mathbb{P}(D_1^V = d_1', \, D_2^V = d_2', \ldots, D_{d-1}^V = d_{d-1}')$$
$$= \frac{\tilde{p}_{d_1'} \tilde{p}_{d_2'} \cdots \tilde{p}_{d_{d-1}'} (1 - e^{-\kappa/d_1'} e^{-\kappa/d_2'} \cdots e^{-\kappa/d_{d-1}'})}{1 - \alpha^{d-1}}. \tag{5}$$

Note that $D_1^V, D_2^V, \ldots, D_{d-1}^V$ are identically distributed but *not* independent. Summing (5) over $d_2', d_3', \ldots, d_{d-1}'$ and recalling the definition of $\alpha$ given in (2) yields, for $d = 2, 3, \ldots$,

$$\mathbb{P}(D_1^V = d_1') = \frac{\tilde{p}_{d_1'}}{1 - \alpha^{d-1}} (1 - e^{-\kappa/d_1'} \alpha^{d-2}), \qquad d_1' = 1, 2, \ldots.$$

Then, by the same arguments as before,

$$m_{V_d, N_{d'}} = (d-1)(1 - e^{-\kappa/d}) \tilde{p}_{d'} (1 - e^{-\kappa/d'} \alpha^{d-2})(1 - \alpha^{d-1})^{-1} p_{VV},$$
$$m_{V_d, V_{d'}} = (d-1) e^{-\kappa/d} \tilde{p}_{d'} (1 - e^{-\kappa/d'} \alpha^{d-2})(1 - \alpha^{d-1})^{-1} (1 - \alpha^{d'-1}) p_{VV},$$

and

$$m_{V_d, U_{d'}} = (d-1) e^{-\kappa/d} \tilde{p}_{d'} (1 - e^{-\kappa/d'} \alpha^{d-2})(1 - \alpha^{d-1})^{-1} \alpha^{d'-1} p_{VU}.$$

When the support of $D$ is finite, the above branching process may be analysed using standard multitype branching process theory (see, e.g. Mode (1971)), so suppose that $D$ has countable support. For ease of exposition, we assume that $p_d > 0$ for all $d = 1, 2, \ldots$; the theory is easily modified if this condition is relaxed. Note that, except for the initial generation, individuals

with degree 1 have no offspring in the above branching process, so let $\tilde{\mathcal{B}}$ be the Galton–Watson process with type space $\tilde{\mathcal{T}} = \{N_d, V_d, U_d : d = 2, 3, \dots\}$ and offspring random variables as described implicitly above. We assume that $p_{UU}$ and $p_{VU}$ are both strictly positive. If $p_{UU} = 0$, vaccination is not required, and if $p_{VU} = 0$, the vaccine is effectively perfect.

For $d^* = 2, 3, \dots$, let $\tilde{\mathcal{B}}^{(d^*)}$ be the branching process derived from $\tilde{\mathcal{B}}$ by ignoring all individuals with degree strictly greater than $d^*$ and all descendants of such individuals. Let $\tilde{M}$ denote the mean offspring matrix for $\tilde{\mathcal{B}}$ and, for $t \in \tilde{\mathcal{T}}$, let $\tilde{\pi}_t$ be the probability that $\tilde{\mathcal{B}}$ becomes extinct given that initially there is one individual whose type is $t$. For $d^* = 2, 3, \dots$, define $\tilde{M}^{(d^*)}$ and $\tilde{\pi}_t^{(d^*)}$ ($t \in \tilde{\mathcal{T}}^{(d^*)} = \{N_d, V_d, U_d : d = 2, 3, \dots, d^*\}$) similarly for the branching process $\tilde{\mathcal{B}}^{(d^*)}$ and let $R_A^{(d^*)}$ be the dominant eigenvalue of $\tilde{M}^{(d^*)}$. Then, $R_A^{(d^*)}$ is strictly increasing in $d^*$ (see Seneta (1973, Theorem 6.8)). Let $R_A = \lim_{d^* \to \infty} R_A^{(d^*)}$; comparison with the model without vaccination shows that $R_A$ is finite if $\sigma_D^2 < \infty$. The following proposition, proved in Appendix A, shows that $R_A$ is a threshold parameter for the original acquaintance vaccination model with imperfect vaccination.

**Proposition 1.** *For any $t \in \tilde{\mathcal{T}}$, (i) $\tilde{\pi}_t = \lim_{d^* \to \infty} \tilde{\pi}_t^{(d^*)}$ and (ii) $\tilde{\pi}_t < 1$ if and only if $R_A > 1$.*

Similar results to the above hold for the corresponding backward branching process. Although it may be computationally feasible to approximate the threshold parameter $R_A$, or calculate it exactly if the support of $D$ is not too large, approximation of major outbreak probabilities and expected relative final sizes is likely to be computationally infeasible, unless the support of $D$ is small. Hence, we believe that a better approach is to modify the formulation of the acquaintance vaccination model slightly, so as to obtain numerically tractable branching process approximations.

## 4. New acquaintance vaccination model

In this model each individual in the population is sampled independently with probability $p_S$, each sampled individual then names each of its neighbours independently with probability $p_N$ and finally any individual that is named at least once is vaccinated. Thus, the probability that an individual is not named by a given neighbour is $1 - p_S p_N$, so the probability that an individual chosen uniformly at random from the population is vaccinated (i.e. the vaccine coverage) is

$$p_V = 1 - \sum_{k=0}^{\infty} p_k (1 - p_S p_N)^k = 1 - f_D(1 - p_S p_N). \tag{6}$$

The modification of the naming process so that the chance an individual names a given neighbour is independent of that individual's degree means that we do not have to type individuals by their degree in order for a branching process approximation to work and we obtain a branching process with finitely many types. However, vaccination is still targeted at higher-degree individuals since neighbours of sampled individuals are vaccinated. We now describe the typing required for the forward and backward branching processes, which we denote by $\mathcal{B}_F$ and $\mathcal{B}_B$, respectively.

In $\mathcal{B}_F$, infected individuals are typed by (i) whether they are named ($N$), vaccinated ($V$), or unvaccinated ($U$), as before, and (ii) whether or not they are sampled and, thus, might name their neighbours for vaccination ($S$ and $S^c$). Thus, there are six types of individuals involved in the noninitial generations of $\mathcal{B}_F$, labelled 1–6 as indicated in Table 1.

The typing required for the initial generation is somewhat simpler: the initial individual clearly cannot be named since it does not have a parent and we can easily condition on whether

TABLE 1: The six types of individuals involved in the noninitial generations of $\mathcal{B}_F$.

| Label | 1 | 2 | 3 | 4 | 5 | 6 |
|-------|---|---|---|---|---|---|
| Type | $(N, S)$ | $(V, S)$ | $(U, S)$ | $(N, S^c)$ | $(V, S^c)$ | $(U, S^c)$ |

or not it was sampled; thus, we need only distinguish between vaccinated and unvaccinated individuals. We denote these types by $V$ and $U$. Therefore, the forward process $\mathcal{B}_F$ is specified by the (joint) distributions of the random variables $\tilde{\boldsymbol{C}}_i = (\tilde{C}_{ij}, \, j = 1, 2, \ldots, 6)$ for $i = 1, 2, \ldots, 6$, where $\tilde{C}_{ij}$ is the number of type-$j$ individuals infected by a typical type-$i$ infective, and of the random variables $\boldsymbol{C}_A = (C_{Aj}, \, j = 1, 2, \ldots, 6)$, for $A = U, V$. The same typing is used in $\mathcal{B}_B$, so, for example, a noninitial member of a susceptibility set is named if its parent in the susceptibility set names it for vaccination. Thus, in an obvious notation, the backward process $\mathcal{B}_B$ is specified by the distributions of $\tilde{\boldsymbol{B}}_i = (\tilde{B}_{ij}, \, j = 1, 2, \ldots, 6)$ for $i = 1, 2, \ldots, 6$, and $\boldsymbol{B}_A = (B_{Aj}, \, j = 1, 2, \ldots, 6)$ for $A = U, V$. When analysing the forward branching process $\mathcal{B}_F$, it is useful to think of a generation as consisting of two phases. In the first phase we determine the types an infective's neighbours, which we call *potential offspring*, would be if they were infected, i.e. whether they are named, vaccinated, or unvaccinated and whether or not they are sampled. Then in the second phase, conditional on this information, we determine which of these neighbours actually are infected. A similar decomposition is used for the backward process $\mathcal{B}_B$. Note that, for $i = 1, 2, \ldots, 6$, the random vector giving the number of *potential* offspring of individuals of the six types from a typical type-$i$ individual has the same distribution in both $\mathcal{B}_F$ and $\mathcal{B}_B$.

Before analysing $\mathcal{B}_F$ and $\mathcal{B}_B$, we determine the degree distribution of individuals of each type. Note that this is independent of whether or not an individual is sampled, so types 1 and 4, types 2 and 5, and types 3 and 6 have the same degree distributions; we denote generic random variables with these distributions by $\tilde{D}_N$, $\tilde{D}_V$, and $\tilde{D}_U$, respectively. The degree distribution of a typical named individual follows the size-biased degree distribution, so $\tilde{D}_N$ has the same distribution as $\tilde{D}$. Next, consider a typical 'vaccinated' individual (type 2 or 5). It has unconditional degree distribution $\tilde{D}$, but we also know that it is named by at least one of its neighbours, though not by its infector. Thus, for $d = 2, 3, \ldots$,

$$\mathbb{P}(\tilde{D}_V = d) = \frac{\mathbb{P}(\tilde{D} = d)\mathbb{P}(V \mid \tilde{D} = d)}{\tilde{p}_V} = \frac{\tilde{p}_d(1 - (1 - p_N p_S)^{d-1})}{\tilde{p}_V}, \tag{7}$$

where $\tilde{p}_V = \sum_{i=2}^{\infty} \tilde{p}_i(1 - (1 - p_N p_S)^{i-1}) = 1 - f_{\tilde{D}-1}(1 - p_N p_S)$ is the *a priori* probability that the individual is vaccinated. Similarly, a typical unvaccinated (type 3 or 6) individual has unconditional degree distribution $\tilde{D}$, but we also know that it avoids vaccination by all of its neighbours, so

$$\mathbb{P}(\tilde{D}_U = d) = \frac{\tilde{p}_d(1 - p_N p_S)^{d-1}}{1 - \tilde{p}_V}, \qquad d = 1, 2, \ldots. \tag{8}$$

Turning now to initial individuals, recall that these can be of only two types, $V$ and $U$. If the initial individual is chosen uniformly at random from the population then *a priori* its degree is distributed according to $D$ and it is vaccinated with probability $p_V$ given by (6). If the initial individual is vaccinated then the distribution of its degree, say $D_V$, is given by

$$\mathbb{P}(D_V = d) = \frac{\mathbb{P}(D = d, V)}{p_V} = \frac{p_d(1 - (1 - p_S p_N)^d)}{p_V}, \qquad d = 1, 2, \ldots. \tag{9}$$

If, on the other hand, it is unvaccinated then the distribution of its degree, say $D_U$, is given by

$$\mathbb{P}(D_U = d) = \frac{p_d(1 - p_S p_N)^d}{1 - p_V}, \qquad d = 0, 1, \ldots. \tag{10}$$

We also define the marginal infection probabilities $p_{ij}^I = p_{A(i), A(j)}$ for the types $i, j = 1, 2, \ldots, 6$, where $A(i) = V$ if $i \in \{1, 2, 4, 5\}$ and $A(i) = U$ if $i \in \{3, 6\}$.

### 4.1. Threshold parameter

We now set out to calculate the mean matrix $M = (m_{ij}, i, j = 1, 2, \ldots, 6)$ of the offspring distributions for the noninitial generations of the forward branching process $\mathcal{B}_F$. The dominant eigenvalue $R_v$ of $M$ will serve as a threshold parameter determining whether or not major outbreaks infecting a significant fraction of the population are possible.

Consider a typical type-1 (i.e. $(N, S)$) individual. On average it has $\mu_{\tilde{D}-1}$ uninfected neighbours. Since a type-1 individual is sampled, each of these neighbours is named with probability $p_N$; if such a neighbour is not named then it is vaccinated with probability $\tilde{p}_V$. Independently, each such neighbour is sampled with probability $p_S$. For $i = 1, 2, \ldots, 6$, let $\hat{p}_{1i}$ be the probability that a given potential offspring of a type-1 individual is of type $i$. Then $\hat{p}_{11} = p_N p_S$, $\hat{p}_{12} = (1 - p_N)\tilde{p}_V p_S$, $\hat{p}_{13} = (1 - p_N)(1 - \tilde{p}_V)p_S$, $\hat{p}_{14} = p_N(1 - p_S)$, $\hat{p}_{15} = (1 - p_N)\tilde{p}_V(1 - p_S)$, and $\hat{p}_{16} = (1 - p_N)(1 - \tilde{p}_V)(1 - p_S)$. Furthermore, each potential type-$i$ offspring actually is an offspring with probability $p_{1i}^I$. Therefore, the first row of the matrix $M$ is specified by $m_{1j} = \mu_{\tilde{D}-1}\hat{p}_{1j}p_{1j}^I$. For a type-4 individual, the situation is very similar, the difference being that since the individual is not sampled its neighbours cannot be named by it. Thus, $m_{4j} = \mu_{\tilde{D}-1}\hat{p}_{4j}p_{4j}^I$, where $\hat{p}_{41} = 0$, $\hat{p}_{42} = \tilde{p}_V p_S$, $\hat{p}_{43} = (1 - \tilde{p}_V)p_S$, $\hat{p}_{44} = 0$, $\hat{p}_{45} = \tilde{p}_V(1 - p_S)$, and $\hat{p}_{46} = (1 - \tilde{p}_V)(1 - p_S)$.

Consideration of a typical unvaccinated (i.e. type 3 or 6) individual follows very similar arguments, the key difference being that we know that the potential offspring do not name their parent (since the parent is unvaccinated), so the probability that each potential offspring is sampled is

$$\tilde{p}_{SU} = \mathbb{P}(S \mid \text{does not name parent}) = \frac{p_S(1 - p_N)}{p_S(1 - p_N) + 1 - p_S} = \frac{p_S(1 - p_N)}{1 - p_S p_N}. \tag{11}$$

Thus, for $j = 1, 2, \ldots, 6$, $m_{3j} = \mu_{\tilde{D}_U - 1}\hat{p}_{3j}p_{3j}^I$ and $m_{6j} = \mu_{\tilde{D}_U - 1}\hat{p}_{6j}p_{6j}^I$, where, using (8),

$$\mu_{\tilde{D}_U - 1} = \sum_{k=1}^{\infty}(k - 1)\frac{\tilde{p}_k(1 - p_N p_S)^{k-1}}{1 - \tilde{p}_V} = (1 - p_N p_S)\frac{f_{\tilde{D}-1}^{(1)}(1 - p_N p_S)}{1 - \tilde{p}_V};$$

and the potential offspring type probabilities are $\hat{p}_{31} = p_N \tilde{p}_{SU}$, $\hat{p}_{32} = (1 - p_N)\tilde{p}_V \tilde{p}_{SU}$, $\hat{p}_{33} = (1 - p_N)(1 - \tilde{p}_V)\tilde{p}_{SU}$, $\hat{p}_{34} = p_N(1 - \tilde{p}_{SU})$, $\hat{p}_{35} = (1 - p_N)\tilde{p}_V(1 - \tilde{p}_{SU})$, $\hat{p}_{36} = (1 - p_N)(1 - \tilde{p}_V)(1 - \tilde{p}_{SU})$, $\hat{p}_{61} = 0$, $\hat{p}_{62} = \tilde{p}_V \tilde{p}_{SU}$, $\hat{p}_{63} = (1 - \tilde{p}_V)\tilde{p}_{SU}$, $\hat{p}_{64} = 0$, $\hat{p}_{65} = \tilde{p}_V(1 - \tilde{p}_{SU})$, and $\hat{p}_{66} = (1 - \tilde{p}_V)(1 - \tilde{p}_{SU})$.

Now consider a typical type-2 (i.e. $(V, S)$) individual, $i^*$ say. Its degree is distributed as $\tilde{D}_V$. Let $N_S$ and $N_{S^c}$ be the number of uninfected neighbours of $i^*$ that are sampled and unsampled, respectively. For $d = 1, 2, \ldots$, the conditional distribution of $N_S$, given that $i^*$'s degree $\tilde{D}_V = d$, is given by

$$\mathbb{P}(N_S = k \mid \tilde{D}_V = d) = \frac{\binom{d-1}{k}p_S^k(1 - p_S)^{d-1-k}(1 - (1 - p_N)^k)}{1 - (1 - p_N p_S)^{d-1}}, \qquad k = 0, 1, \ldots, d - 1. \tag{12}$$

Hence, using (7), $\mu_{N_S} = \tilde{p}_V^{-1} p_S (\mu_{\tilde{D}-1} - (1 - p_N) f_{\tilde{D}-1}^{(1)} (1 - p_S p_N))$ and, since $N_{S^c} = \tilde{D}_V - 1 - N_S$, we have $\mu_{N_{S^c}} = \tilde{p}_V^{-1} (1 - p_S)(\mu_{\tilde{D}-1} - f_{\tilde{D}-1}^{(1)} (1 - p_S p_N))$. Now, note that each potential offspring, whether sampled or not, is named with probability $p_N$ and otherwise vaccinated (with probability $(1 - p_N) \tilde{p}_V$) or unvaccinated (with probability $(1 - p_N)(1 - \tilde{p}_V)$). Thus, we have $m_{2j} = \mu_{N_S} \hat{p}_{2j} p_{2j}^I$ ($j = 1, 2, 3$) and $m_{2j} = \mu_{N_{S^c}} \hat{p}_{2j} p_{2j}^I$ ($j = 4, 5, 6$), where $\hat{p}_{21} = \hat{p}_{24} = p_N$, $\hat{p}_{22} = \hat{p}_{25} = (1 - p_N) \tilde{p}_V$, and $\hat{p}_{23} = \hat{p}_{26} = (1 - p_N)(1 - \tilde{p}_V)$. For a type-5 (i.e. $(V, S^c)$) individual, the situation is exactly the same except that the potential offspring cannot be named, so $m_{5j} = \mu_{N_S} \hat{p}_{5j} p_{5j}^I$ ($j = 1, 2, 3$) and $m_{5j} = \mu_{N_{S^c}} \hat{p}_{5j} p_{5j}^I$ ($j = 4, 5, 6$), where $\hat{p}_{51} = \hat{p}_{54} = 0$, $\hat{p}_{52} = \hat{p}_{55} = \tilde{p}_V$, and $\hat{p}_{53} = \hat{p}_{56} = 1 - \tilde{p}_V$.

Using similar coupling arguments to those in Ball and Sirl (2012, Section 6.2), the approximating branching process $\mathcal{B}_F$ and a sequence $(\mathcal{E}^{(n)})$ of epidemic processes with acquaintance vaccination, indexed by the population size $n$, may be constructed on the same probability space so that the number of individuals infected in $(\mathcal{E}^{(n)})$ converges almost surely to the total progeny of $\mathcal{B}_F$ as $n \to \infty$. Moreover, if $\sigma_{\tilde{D}}^2 < \infty$, similar coupling arguments to those in Ball *et al.* (2009) show that the probability that $\mathcal{E}^{(n)}$ infects at least $\log n$ individuals tends to the survival probability of $\mathcal{B}_F$ as $n \to \infty$. We then have the following proposition, in which a major outbreak is one that infects at least $\log n$ individuals. We conjecture that, as for the model without vaccination, in the limit as $n \to \infty$, a major outbreak almost surely infects exact order $n$ individuals (see Proposition 4, in Section 4.4; cf. final paragraph of Section 2.1.2).

**Proposition 2.** *The quantity $R_v$ is a threshold parameter for our epidemic model in the usual sense that, as the population size $n \to \infty$, a major outbreak occurs with strictly positive probability if and only if $R_v > 1$.*

### 4.2. PGFs of potential offspring distributions

First we consider individuals in a noninitial generation. For $i = 1, 2, \ldots, 6$, let $\tilde{X}_i = (\tilde{X}_{i1}, \tilde{X}_{i2}, \ldots, \tilde{X}_{i6})$ be the number of potential offspring of each type of a typical type-$i$ individual in a noninitial generation. (Recall that $\tilde{X}_i$ has the same distribution in both the forward and backward processes.) For types 1, 3, 4, and 6, each neighbour of the individual under consideration takes its type independently with fixed probability so, conditional on the degree of the individual, the numbers of potential offspring follow a multinomial distribution. Thus, for $s \in [0, 1]^6$, we have $f_{\tilde{X}_i}(s) = f_{\tilde{D}-1}(g_i(s))$, $i = 1, 4$, and $f_{\tilde{X}_i}(s) = f_{\tilde{D}_U - 1}(g_i(s))$, $i = 3, 6$, where $g_i(s) = \sum_{j=1}^6 \hat{p}_{ij} s_j$, $i = 1, 3, 4, 6$.

Now consider a typical type-2 individual. Conditioning on its degree (distributed according to $\tilde{D}_V$) and then on the number $N_S$ of its potential offspring that are sampled, we have

$$\mathbb{E}\left[ \prod_{j=1}^6 s_j^{\tilde{X}_{2j}} \right] = \mathbb{E}_{\tilde{D}_V}\left[ \mathbb{E}_{N_S}\left[ \mathbb{E}_{\tilde{X}_2}\left[ \prod_{j=1}^6 s_j^{\tilde{X}_{2j}} \ \middle| \ N_S, \tilde{D}_V \right] \right] \right].$$

Since each potential offspring, independently of whether or not it is sampled, is named, vaccinated, or unvaccinated with probability $\hat{p}_{21} = \hat{p}_{24}$, $\hat{p}_{22} = \hat{p}_{25}$, and $\hat{p}_{23} = \hat{p}_{26}$, respectively, we have (note that $N_S + N_{S^c} = \tilde{D}_V - 1$)

$$\mathbb{E}_{\tilde{X}_2}\left[ \prod_{j=1}^6 s_j^{\tilde{X}_{2j}} \ \middle| \ N_S, N_{S^c} \right] = (g_{21}(s))^{N_S} (g_{22}(s))^{N_{S^c}}, \tag{13}$$

where, for $s \in [0, 1]^6$, $g_{21}(s) = \sum_{j=1}^3 \hat{p}_{2j} s_j$ and $g_{22}(s) = \sum_{j=4}^6 \hat{p}_{2j} s_j$. Together with the

conditional distribution of $N_S$ given by (12), it then follows that, for $d = 2, 3, \ldots,$

$$\mathbb{E}_{\tilde{X}_2}\left[\prod_{j=1}^{6} s_j^{\tilde{X}_{2j}} \,\middle|\, \tilde{D}_V = d\right]$$

$$= \frac{1}{1 - (1 - p_S p_N)^{d-1}}$$

$$\times \sum_{k=1}^{d-1} \binom{d-1}{k}(p_S g_{21}(s))^k ((1 - p_S)g_{22}(s))^{d-1-k}(1 - (1 - p_N)^k)$$

$$= \frac{1}{1 - (1 - p_S p_N)^{d-1}}[(g_2(s, p_S, 0))^{d-1} - (g_2(s, p_S, p_N))^{d-1}], \qquad (14)$$

where $g_2(s, p_S, p_N) = p_S(1 - p_N)g_{21}(s) + (1 - p_S)g_{22}(s)$. Taking the expectation of (14) with respect to $\tilde{D}_V$, using the mass function (7), then yields

$$f_{\tilde{X}_2}(s) = \frac{1}{\tilde{p}_V}[f_{\tilde{D}-1}(g_2(s, p_S, 0)) - f_{\tilde{D}-1}(g_2(s, p_S, p_N))].$$

A similar calculation for a type-5 individual, noting that its offspring cannot be named, yields

$$f_{\tilde{X}_5}(s) = \frac{1}{\tilde{p}_V}[f_{\tilde{D}-1}(g_5(s, p_S, 0)) - f_{\tilde{D}-1}(g_5(s, p_S, p_N))],$$

where $g_5(s, p_S, p_N) = p_S(1 - p_N)g_{51}(s) + (1 - p_S)g_{52}(s)$, with $g_{51}(s) = \sum_{j=1}^{3} \hat{p}_{5j}s_j$ and $g_{52}(s) = \sum_{j=4}^{6} \hat{p}_{5j}s_j$.

For individuals in the initial generation the calculations are very similar. Recall that initial individuals are typed only by whether they are vaccinated (type $V$) or unvaccinated (type $U$). For $A = V, U$, let $X_A = (X_{Aj}, j = 1, 2, \ldots, 6)$ be the number of potential offspring of an initial individual of type $A$. A typical type-$U$ initial individual is sampled with probability $p_S$ and has degree distributed as $D_U$ defined by (10). If it is sampled then each of its neighbours (potential offspring) is of type $j$ independently with probabilities $(\hat{p}_{3j}, j = 1, 2, \ldots, 6)$ and if it is not sampled then these probabilities are $(\hat{p}_{6j}, j = 1, 2, \ldots, 6)$. Thus, for $s \in [0, 1]^6$,

$$f_{X_U}(s) = p_S f_{D_U}(g_3(s)) + (1 - p_S)f_{D_U}(g_6(s)),$$

where, using (10), $f_{D_U}(s) = f_D(s(1 - p_S p_N))/(1 - p_V)$. Now consider a typical type-$V$ individual, $i^*$ say. As before, let $N_S$ and $N_{S^c}$ be the number of $i^*$'s neighbours that are sampled and unsampled, respectively. Since $i^*$ must be named by at least one of its neighbours, it follows that, for $d = 1, 2, \ldots,$

$$\mathbb{P}(N_S = k \mid D_V = d) = \frac{\binom{d}{k}p_S^k(1 - p_S)^{d-k}(1 - (1 - p_N)^k)}{1 - (1 - p_S p_N)^d}, \qquad k = 1, 2, \ldots, d.$$

Furthermore, since $i^*$ is sampled with probability $p_S$, we have

$$\mathbb{E}_{X_V}\left[\prod_{j=1}^{6} s_j^{X_{Vj}} \,\middle|\, N_S, N_{S^c}\right] = p_S(g_{21}(s))^{N_S}(g_{22}(s))^{N_{S^c}} + (1 - p_S)(g_{51}(s))^{N_S}(g_{52}(s))^{N_{S^c}}.$$

This is an analogue of (13) in the derivation of $f_{\tilde{X}_2}$ and the corresponding equation in the derivation of $f_{\tilde{X}_5}$. Arguing as in those cases and recalling the mass function (9) yields

$$f_{X_V}(s) = \frac{1}{p_V}(p_S[f_D(g_2(s, p_S, 0)) - f_D(g_2(s, p_S, p_N))]$$
$$+ (1 - p_S)[f_D(g_5(s, p_S, 0)) - f_D(g_5(s, p_S, p_N))]).$$

### 4.3. Probability of a major outbreak

It is convenient to treat the nonrandom and all-or-nothing vaccine responses separately. Consider first the nonrandom vaccine action model and, for $i = 1, 2, \ldots, 6$, define $a_i$ to be the relative susceptibility and $b_i$ the relative infectivity of a type-$i$ individual. Thus, $a_i = b_i = 1$ for $i = 3, 6$ and $a_i = a$, $b_i = b$ for $i = 1, 2, 4, 5$. For $i = 1, 2, \ldots, 6$, conditioning on the potential offspring vector $\tilde{X}_i$ and infectious period $I$ of a typical noninitial generation type-$i$ individual yields

$$f_{\tilde{C}_i}(s) = \mathbb{E}_{I, \tilde{X}_i}\left[\mathbb{E}_{\tilde{C}_i}\left[\prod_{j=1}^{6} s_j^{\tilde{C}_{ij}} \mid \tilde{X}_i, I\right]\right] = \mathbb{E}_{I, \tilde{X}_i}\left[\prod_{j=1}^{6}(e^{-a_j b_i \lambda I} + (1 - e^{-a_j b_i \lambda I})s_j)^{\tilde{X}_{ij}}\right]. \quad (15)$$

Then, using the same manipulations as at the end of Section A.2 of Ball *et al.* (2010) and employing the vector notation defined in Section 1.1, we find that

$$f_{\tilde{C}_i}(s) = \sum_{k \in \mathbb{Z}_+^6} \frac{(1 - s)^k h_i^F(k)}{k!} f_{\tilde{X}_i}^{(k)}(s), \quad (16)$$

where $h_i^F(k) = \phi(b_i \lambda \sum_{j=1}^{6} a_j k_j)$.

Exactly the same arguments apply for the initial generation, whence, for $A = U, V$ and $s \in [0, 1]^6$,

$$f_{C_A}(s) = \sum_{k \in \mathbb{Z}_+^6} \frac{(1 - s)^k h_A^F(k)}{k!} f_{X_A}^{(k)}(s), \quad (17)$$

where $h_V^F \equiv h_1^F$ and $h_U^F \equiv h_3^F$.

If the infectious period is constant then (16) and (17) simplify appreciably, since it follows directly from (15) that, with $p_{ij}^I$ defined using $P^{\mathrm{NR}}$, for $s \in [0, 1]^6$,

$$f_{\tilde{C}_i}(s) = f_{\tilde{X}_i}(h_i^F(s)), \quad i = 1, 2, \ldots, 6, \qquad \text{and} \qquad f_{C_A}(s) = f_{X_A}(h_A^F(s)), \quad A = U, V, \quad (18)$$

where $h_i^F(s) = (1 - p_{ij}^I + p_{ij}^I s_j, \ j = 1, 2, \ldots, 6)$, $h_V^F(s) = h_1^F(s)$, and $h_U^F(s) = h_3^F(s)$.

Turning to the all-or-nothing vaccine response model, since vaccinated individuals are either rendered completely immune or are unaffected by the vaccine, in $\mathscr{B}_F$ we need only keep track of unvaccinated individuals and vaccinated individuals for whom the vaccine fails. Thus, we still have six types of individual, labelled 1 to 6 as before, but with the implicit assumption that 'named' and 'vaccinated' types are fully infectious, i.e. we ignore individuals who become fully immune. (Of course, the different types still have different degree distributions.) A similar argument to that used above shows that (15) holds but with $e^{-a_j b_i \lambda I} + (1 - e^{-a_j b_i \lambda I})s_j$ replaced by $e^{-\lambda I} + (1 - e^{-\lambda I})s_j$ for $j = 3, 6$, and by $\varepsilon + (1 - \varepsilon)e^{-\lambda I} + (1 - \varepsilon)(1 - e^{-\lambda I})s_j$ for $j = 1, 2, 4, 5$. Analogous expressions to (16) and (17) then follow. If the infectious period is constant then (18) holds with the $p_{ij}^I$ defined using $P^{\mathrm{AoN}}$.

The probability of a major outbreak in our epidemic model is approximated by the survival probability of the forward branching process $\mathcal{B}_F$. By standard branching process theory this is, assuming that the initial individual is of type $A \in \{U, V\}$, $p_{\text{maj}}^{(A)} = 1 - f_{C_A}(\boldsymbol{\sigma})$, where $\boldsymbol{\sigma}$ is the smallest solution of $\boldsymbol{f}_{\tilde{\boldsymbol{C}}}(\boldsymbol{s}) = \boldsymbol{s}$ in $[0, 1]^6$ and $\boldsymbol{f}_{\tilde{\boldsymbol{C}}}(\boldsymbol{s}) = (f_{\tilde{C}_i}(\boldsymbol{s}), i = 1, 2, \ldots, 6)$. The following proposition then follows by conditioning on the type of the initial infective.

**Proposition 3.** *Consider the epidemic taking place on the network with vaccination according to our (new) acquaintance vaccination model and a single initial infective chosen uniformly at random from the population. As the population size $n \to \infty$, the probability that a major outbreak occurs is given by*

$$p_{\text{maj}} = \begin{cases} p_V \, p_{\text{maj}}^{(V)} + (1 - p_V) \, p_{\text{maj}}^{(U)} & \text{nonrandom vaccine,} \\ p_V (1 - \varepsilon) p_{\text{maj}}^{(V)} + (1 - p_V) p_{\text{maj}}^{(U)} & \text{all-or-nothing vaccine.} \end{cases}$$

*The probability $p_{\text{maj}}$ is strictly positive if $R_v > 1$; if $R_v \leq 1$ then $p_{\text{maj}} = 0$.*

### 4.4. Final size of the major outbreak

The expected relative final size of a major outbreak is approximated by the survival probability of the backward process $\mathcal{B}_B$. Standard branching process theory tells us that if we write $\boldsymbol{f}_{\tilde{\boldsymbol{B}}}(\boldsymbol{s}) = (f_{\tilde{B}_i}(\boldsymbol{s}), i = 1, 2, \ldots, 6)$ then the expected relative final size amongst type-$A$ individuals ($A = U, V$) can be written as $z^{(A)} = 1 - f_{B_A}(\boldsymbol{\xi})$, where $\boldsymbol{\xi}$ is the smallest solution of $\boldsymbol{f}_{\tilde{\boldsymbol{B}}}(\boldsymbol{s}) = \boldsymbol{s}$ in $[0, 1]^6$.

To calculate the PGF $f_{\tilde{\boldsymbol{B}}_i}(\boldsymbol{s})$, we firstly condition on $\tilde{\boldsymbol{X}}_i = (\tilde{X}_{ij}, j = 1, 2, \ldots, 6)$, the number of potential offspring of each type. Since each potential offspring joins the susceptibility set independently, type $j$ potential offspring joining with probability $p_{ji}^I$, we find that

$$f_{\tilde{\boldsymbol{B}}_i}(\boldsymbol{s}) = f_{\tilde{\boldsymbol{X}}_i}(\boldsymbol{h}_i^B(\boldsymbol{s})),$$

where $\boldsymbol{h}_i^B(\boldsymbol{s}) = (1 - p_{ji}^I + p_{ji}^I s_j, \; j = 1, 2, \ldots, 6)$ and $p_{ji}^I \; (i, j = 1, 2, \ldots, 6)$ are the marginal infection probabilities, using $P^{\text{AoN}}$ or $P^{\text{NR}}$ as appropriate. As with $\mathcal{B}_F$, the calculations relating the potential to the actual offspring distribution are exactly the same when dealing with the initial generation; in this instance we have

$$f_{\boldsymbol{B}_A}(\boldsymbol{s}) = f_{X_A}(\boldsymbol{h}_A^B(\boldsymbol{s})), \qquad A \in \{U, V\},$$

where $\boldsymbol{h}_U^B \equiv \boldsymbol{h}_3^B$ and $\boldsymbol{h}_V^B \equiv \boldsymbol{h}_1^B$.

These formulae hold equally for all-or-nothing and nonrandom vaccine action models if we make the same implicit assumption as above, namely that with the all-or-nothing vaccine action model we count only those individuals for whom the vaccine fails in the 'named' and 'vaccinated' types. We expect that the arguments of Ball *et al.* (2009) can be adapted to prove that, as supported by extensive numerical simulations, the relative final size of a major outbreak converges in probability to the survival probability of $\mathcal{B}_B$ as $n \to \infty$ (cf. final paragraph of Section 2.1.2). A simple conditioning on the type of the randomly chosen initial susceptible thus suggests the following result.

**Proposition 4.** *Consider the epidemic taking place on the network with vaccination according to our (new) acquaintance vaccination model and a single initial infective chosen uniformly at random from the population. As the population size $n \to \infty$, the proportion of the population*

*that is ultimately infected by a major outbreak is given by*

$$z = \begin{cases} p_V z^{(V)} + (1 - p_V) z^{(U)} & \text{nonrandom vaccine,} \\ p_V(1 - \varepsilon) z^{(V)} + (1 - p_V) z^{(U)} & \text{all-or-nothing vaccine.} \end{cases}$$

## 5. Properties of the new vaccination model

In this section we briefly investigate some of the properties of our new model for acquaintance vaccination. We use the notation $c$ for the vaccine coverage; as pointed out earlier it is the same as the probability $p_V$ but the interpretation as the coverage is more pertinent here. We begin by examining the trade-off between $p_S$ and $p_N$. Clearly, the vaccine coverage $c = 1 - f_D(1 - p_S p_N)$ depends only on the product $p_S p_N$, but the precise values of $p_S$ and $p_N$ affect how effective the vaccination is. Analytical progress in this direction is possible in the case where the vaccine is perfect, we later use numerical methods to explore the situation for imperfect vaccines.

When the vaccine is perfect, only unvaccinated individuals can be infected and the approximating branching processes involve only two types of individual, $(U, S)$ and $(U, S^c)$. Moreover, in both $\mathcal{B}_F$ and $\mathcal{B}_B$ unvaccinated individuals are sampled independently, so these processes are essentially single type. The threshold parameter $R_v$ may be obtained as follows. A typical infected individual, $i^*$ say, in a noninitial generation of $\mathcal{B}_F$ is sampled with probability $\tilde{p}_{SU}$, has on average $\mu_{\tilde{D}_U - 1}$ forward neighbours (potential offspring) and fails to name a given forward neighbour with probability $1 - \tilde{p}_{SU} + \tilde{p}_{SU}(1 - p_N)$. Furthermore, each unnamed forward neighbour is vaccinated with probability $\tilde{p}_V$. Thus, $i^*$ has on average $\mu_{\tilde{D}_U - 1}(1 - \tilde{p}_{SU} + \tilde{p}_{SU}(1 - p_N))(1 - \tilde{p}_V)$ susceptible forward neighbours, so

$$R_v = p^I \mu_{\tilde{D}_U - 1}(1 - \tilde{p}_{SU} + \tilde{p}_{SU}(1 - p_N))(1 - \tilde{p}_V) = p^I(1 - 2p' + p' p_N) f^{(1)}_{\tilde{D} - 1}(1 - p'),$$

using (8) and (11) and writing $p'$ for $p_S p_N$. Thus, for fixed $p'$, i.e. fixed coverage, $R_v$ is strictly increasing in $p_N$. The best we can do with fixed coverage is to take $(p_S, p_N) = (1, p')$, and the worst we can do is to take $(p_S, p_N) = (p', 1)$; these give

$$R_v^b = p^I(1 - p')^2 f^{(1)}_{\tilde{D} - 1}(1 - p') \quad \text{and} \quad R_v^w = p^I(1 - p') f^{(1)}_{\tilde{D} - 1}(1 - p').$$

Clearly, the factor $1 - p'$ between $R_v^b$ and $R_v^w$ gives the greatest difference when $p'$ is large, i.e. the coverage $c = 1 - f_D(1 - p')$ is large. This suggests that when coverage is high the balance between $p_S$ and $p_N$ is more important than when the coverage is low. However, when the coverage is high with a perfect vaccine, the epidemic is likely to be subcritical in any event. The nature of the dependence on $p_S$ and $p_N$ is also perhaps not surprising: the effect of vaccination is greater if everyone in the population names a few friends to be vaccinated than if a few people in the population name all of their friends to be vaccinated. As is seen below, however, our numerical studies indicate that the observed differences between these 'best' and 'worst' strategies are very small. With a perfect vaccine in the old model, expression (3) for $R_v$ is explicit but cannot easily be compared to the corresponding formulae above for the new model.

Turning to the case of an imperfect vaccine, we numerically compare the performance of the old and new acquaintance vaccination models, 'uniform vaccination' (selecting individuals to be vaccinated uniformly at random), and the best possible scheme in the SNM, where the highest-degree individuals are vaccinated. The latter scheme, which we call 'SNM-optimal', is
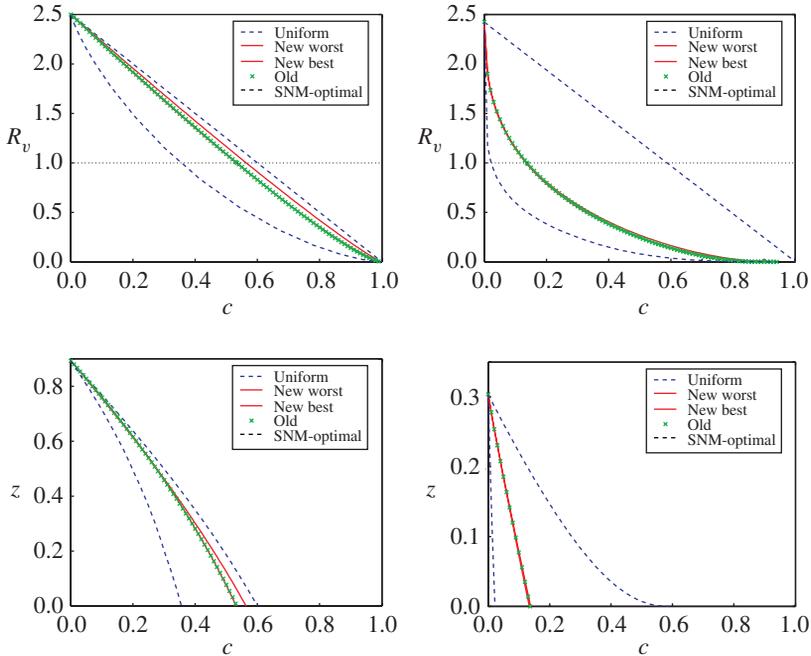
FIGURE 1: Plots of the post-vaccination threshold parameter and the expected relative final size of a major outbreak versus vaccine coverage $c$ for various vaccine allocation regimes, using a perfect vaccine and with degree distributions $D \sim \text{Poi}(10)$ (*left*) and $D \sim \text{Pow}(12, 3.4)$ (*right*), which (to 1 decimal place) has $\mu_D = 10.0$ and $\sigma_D^2 = 176.7$. Other model parameters are $\lambda = 1$ and $I \sim \exp(3)$ (*left*), $I \sim \exp(10)$ (*right*), where $\exp(\mu)$ denotes the negative exponential distribution with mean $\mu^{-1}$.

of course impossible to implement in any real-world setting, but nevertheless it gives a guide to help understand how well the various acquaintance vaccination models perform and we would hope that the performance of the acquaintance vaccination schemes comes close to that of the SNM-optimal scheme. Details of the calculations involved are given in Appendix B.

In Figure 1 we plot (using two different degree distributions) the post-vaccination threshold parameter $R_v$ and the expected relative final size of a major outbreak $z$ against vaccine coverage, using a perfect vaccine with a range of vaccine allocation regimes. The regimes we compare involve choosing individuals (i) uniformly at random, (ii) with the worst 'new' acquaintance vaccination regime, (iii) with the best 'new' acquaintance vaccination regime, (iv) with the 'old' acquaintance vaccination regime, and (v) according to the SNM-optimal regime. The degree distributions we use are the Poisson distribution with mean $\mu$, denoted by $\text{Poi}(\mu)$, and a distribution with a power law tail, for which we write $D \sim \text{Pow}(k_*, a)$ to mean

$$\mathbb{P}(D = k) = \begin{cases} k_*^{-a} & \text{for } k = 0, 1, \ldots, k_*, \\ k^{-a} & \text{for } k = k_* + 1, k_* + 2, \ldots. \end{cases}$$

In Figure 2 we present the same plots as in Figure 1 except the vaccine action model used is all-or-nothing with success probability 0.8. In the upper-right plot of Figure 2 we do not include the performance of the old acquaintance vaccination model, since the approximations $R_A^{(d^*)}$ of $R_v$ (described in Section 3.2) are very slow to converge as $d^* \to \infty$ when $D$ is heavy tailed and we have not been able to implement satisfactory numerical methods to approximate $R_v$ in
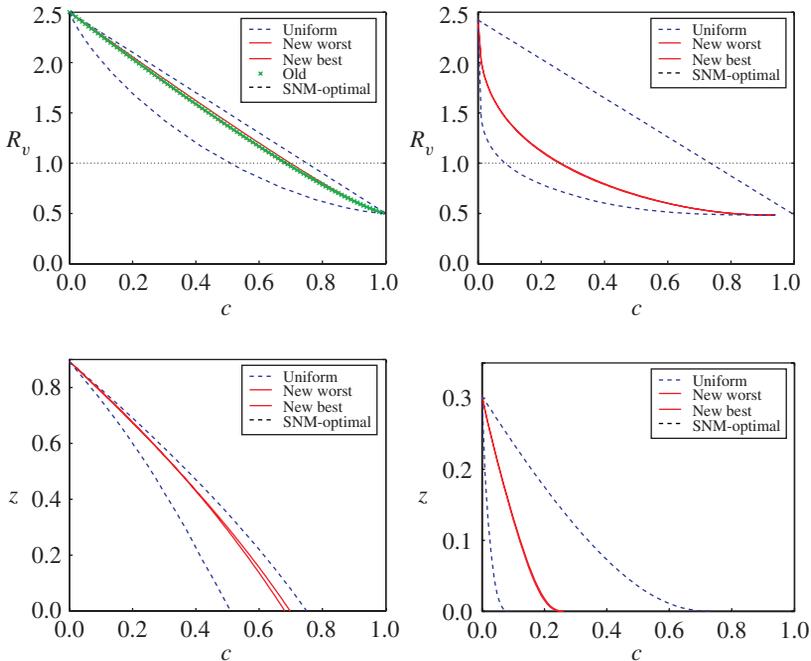
FIGURE 2: Plots of the post-vaccination threshold parameter and the expected relative final size of a major outbreak versus vaccine coverage for various vaccine allocation regimes, using an all-or-nothing vaccine with $\varepsilon = 0.8$ and with degree distributions $D \sim \text{Poi}(10)$ (*left*) and $D \sim \text{Pow}(12, 3.4)$ (*right*). Other model parameters are $\lambda = 1$ and $I \sim \exp(3)$ (*left*), $I \sim \exp(10)$ (*right*).

this situation. Very similar plots to those in Figure 2 are obtained when using a nonrandom vaccine with parameters $a = b = \sqrt{0.2}$ (chosen so that the efficacy $1 - \mathbb{E}[AB]$ of the imperfect vaccines are the same).

Unsurprisingly, we see the various acquaintance vaccination regimes perform better than vaccinating individuals chosen uniformly at random but not as well as vaccinating individuals of highest degree. It also appears that the advantage acquaintance vaccination offers over uniform vaccination is more pronounced for the more widely spread degree distribution—the effect of selecting individuals to vaccinate whose degree is distributed according to $\tilde{D}$ rather than $D$ being greater when the degree distribution has greater variability. Additionally, the extent to which the SNM-optimal allocation regime outperforms acquaintance vaccination is rather smaller for the widely spread degree distribution. This suggests that it is in realistic networks, in which individuals have vastly differing connectivities, that acquaintance vaccination both outperforms naively vaccinating uniformly at random to the greatest extent and best approximates the optimal vaccine allocation regime. We also note that the best and worst new acquaintance vaccination regimes and (when we can do the necessary calculations) also the old acquaintance-based regime all perform very similarly.

Figures 1 and 2 suggest that the difference between the best and worst performances of the new acquaintance vaccination model is very small indeed, especially for the power law degree distribution. Whilst we can analytically compare the threshold parameters of the best and worst models when the vaccine is perfect and we find that the differences are quite small, results such as those in Figure 2 and similar plots for other vaccine action models (not shown) indicate

that the discrepancy is also quite small when considering imperfect vaccines. In Figure 2 we have used the labels 'best' and 'worst' as with the perfect vaccine even though we have no real reason to do so. Extensive numerical results (not shown) indicate that for fixed $p' = p_S p_N$ the performance of an imperfect vaccine (measured by $R_v$ or $z$) is always monotonic in $p_N$ and, whilst the direction of the monotonicity is not fixed, the magnitude of the differences between best and worst is very small.

## Appendix A. Proof of Proposition 1

Let $t_0$ denote the type $U_2$, and suppose that $\tilde{\mathcal{B}}$ has a single ancestor whose type is $t_0$. Let $\tilde{\mathcal{B}}_0$ be the single-type ($t_0$) Galton–Watson process embedded in $\tilde{\mathcal{B}}$, in which, apart from the initial ancestor, for each type-$t_0$ individual in $\tilde{\mathcal{B}}$, its mother in $\tilde{\mathcal{B}}_0$ is given by its most recent type-$t_0$ ancestor when looking backwards in the family tree. Note that the offspring distribution of $\tilde{\mathcal{B}}_0$ may have a mass at $\infty$. Let $\hat{\pi}$ be the extinction probability of $\tilde{\mathcal{B}}_0$. For $d^* = 2, 3, \ldots$, let $\tilde{\mathcal{B}}_0^{(d^*)}$ be the single-type Galton–Watson process derived from $\tilde{\mathcal{B}}^{(d^*)}$ in the same fashion and let $\hat{\pi}^{(d^*)}$ denote the extinction probability of $\tilde{\mathcal{B}}_0^{(d^*)}$.

For $d = 2, 3, \ldots$, let $p_{N_d}$ be the probability that a given type-$N_d$ individual in $\tilde{\mathcal{B}}$ has at least one child of type $U_2$. Then $p_{N_d} \geq m_{N_d, U_2}/(d-1) \geq \mathrm{e}^{-\kappa/2} \tilde{p}_2 \alpha p_{VU}$. Similarly, in obvious notation, $p_{U_d} \geq \mathrm{e}^{-\kappa} \tilde{p}_2 p_{UU}$. Also, $p_{V_d} \geq \mathrm{e}^{-\kappa/d} \tilde{p}_2 (1-\mathrm{e}^{-\kappa/2} \alpha^{d-2})(1-\alpha^{d-1})^{-1} \alpha p_{VU}$ (for $d = 2, 3, \ldots$). Now $(1 - \mathrm{e}^{-\kappa/2} \alpha^{d-2})(1 - \alpha^{d-1})^{-1}$ increases with $d$ if $\alpha \leq \mathrm{e}^{-\kappa/2}$ and decreases with $d$ if $\alpha \geq \mathrm{e}^{-\kappa/2}$. Thus, $p_{V_d} \geq \mathrm{e}^{-\kappa/2} \tilde{p}_2 \min(1, (1-\mathrm{e}^{-\kappa/2})/(1-\alpha)) \alpha p_{VU}$. Hence, there exists $p_0 > 0$ such that the probability that any given individual has at least one child of type $U_2$ is at least $p_0$, so, by the strong law of large numbers, if $\tilde{\mathcal{B}}$ survives (i.e. does not become extinct) then almost surely so does $\tilde{\mathcal{B}}_0$. Clearly, $\tilde{\mathcal{B}}_0$ becomes extinct if $\tilde{\mathcal{B}}$ becomes extinct. Thus, $\tilde{\pi}_{t_0} = \hat{\pi}$, and a similar argument gives $\tilde{\pi}_{t_0}^{(d^*)} = \hat{\pi}^{(d^*)}$ ($d^* = 2, 3, \ldots$).

Let the random variable $\tilde{C}_0$ be distributed according to the offspring distribution of $\tilde{\mathcal{B}}_0$ and, for $d^* = 2, 3, \ldots$, let the random variable $\tilde{C}_0^{(d^*)}$ be distributed according to the offspring distribution of $\tilde{\mathcal{B}}_0^{(d^*)}$. Construct coupled realisations of $\tilde{C}_0$ and $\tilde{C}_0^{(d^*)}$ ($d^* = 2, 3, \ldots$) using the progeny of the ancestor in $\tilde{\mathcal{B}}$. If $\tilde{C}_0 < \infty$ then only finitely many types in $\tilde{\mathcal{B}}$ contribute to $\tilde{C}_0$, so $\tilde{C}_0 = \tilde{C}_0^{(d^*)}$ for all sufficiently large $d^*$. Alternatively, if $\tilde{C}_0 = \infty$ then, for any integer $K > 0$, let $T(K)$ be the maximum of the degrees of individuals used in $\tilde{\mathcal{B}}$ when $\tilde{C}_0$ first exceeds $K$. Then $\tilde{C}_0^{(d^*)} \geq K$ for all $d^* \geq T(K)$. Thus, $\tilde{C}_0^{(d^*)}$ converges almost surely to $\tilde{C}_0$ as $d^* \to \infty$. Hence, by the dominated convergence theorem, $\lim_{d^* \to \infty} f_{\tilde{C}_0^{(d^*)}}(s) = f_{\tilde{C}_0}(s)$, $0 \leq s \leq 1$, whence $\hat{\pi} = \lim_{d^* \to \infty} = \hat{\pi}^{(d^*)}$ (cf. Britton *et al.* (2007, Lemma 4.1)), so $\tilde{\pi}_{t_0} = \lim_{d^* \to \infty} \tilde{\pi}_{t_0}^{(d^*)}$.

Suppose now that $\tilde{\mathcal{B}}$ has a single ancestor whose type is $t \neq t_0$. Let $\tilde{\mathcal{B}}_{t,0}$ be the single-type Galton–Watson process, consisting of the ancestor and all type-$t_0$ individuals in $\tilde{\mathcal{B}}$, that is embedded in $\tilde{\mathcal{B}}$ analogously to $\tilde{\mathcal{B}}_0$ above. Note that the offspring distribution of the ancestor in $\tilde{\mathcal{B}}_{t,0}$ may be different from that of subsequent individuals. The above argument may be modified in the obvious fashion to show that $\tilde{\pi}_t = \lim_{d^* \to \infty} \tilde{\pi}_t^{(d^*)}$, thus, completing the proof of Proposition 1(i).

To prove Proposition 1(ii), note that every element of $\tilde{M}$ is strictly positive, so, for any $t \in \tilde{\mathcal{T}}$, $\tilde{\pi}_t < 1$ if and only if $\tilde{\pi}_{t_0} < 1$. Suppose that $\tilde{\pi}_{t_0} < 1$. Then, since $\hat{\pi} = \tilde{\pi}_{t_0}$ and $\hat{\pi} = \lim_{d^* \to \infty} \hat{\pi}^{(d^*)}$, there exists $d^*$ such that $\hat{\pi}^{(d^*)} < 1$, whence, since $\tilde{\pi}_{t_0}^{(d^*)} = \hat{\pi}^{(d^*)}$, Mode (1971, Theorem 1.7.1) implies that $R_A^{(d^*)} > 1$, so $R_A > 1$. Conversely, if $R_A > 1$, then there exists $d^*$ such that $R_A^{(d^*)} > 1$, whence $\tilde{\pi}_t^{(d^*)} < 1$ which implies that $\tilde{\pi}_{t_0} < 1$, since $\tilde{\pi}_{t_0} < \tilde{\pi}_t^{(d^*)}$.

## Appendix B.  SNM-optimal vaccination

Clearly, the best possible vaccination strategy in the SNM is to vaccinate all individuals of 'high' degree and no individuals of 'low' degree, the cutoff between low and high being determined by the coverage available.  Given $c$, the smallest degree of individual which we vaccinate is given by $d_c = \max\{k \in \mathbb{Z}_+ : \sum_{j=0}^{k-1} p_j < 1 - c\}$. We then vaccinate no individuals of degree $d_c - 1$ or smaller, some proportion $\delta \in (0, 1]$ of individuals of degree $d_c$, and all individuals of degree $d_c + 1$ and higher.  Setting $\delta = (c - \sum_{i=d_c+1}^{\infty} p_i)/p_{d_c}$ gives the desired coverage $c$.

Let $\tilde{p}_V$ be the probability that the individual from which a randomly chosen half-edge emanates is vaccinated.  The degree of this individual is distributed as $\tilde{D}$, so $\tilde{p}_V = \delta \tilde{p}_{d_c} + \sum_{j=d_c+1}^{\infty} \tilde{p}_j$.  The degree distribution of a typical vaccinated individual is given by $\mathbb{P}(D_V = d_c) = \delta p_{d_c} c^{-1}$ and $\mathbb{P}(D_V = k) = p_k c^{-1}$ for $k = d_c + 1, d_c + 2, \ldots$, whilst that of a typical unvaccinated individual is given by $\mathbb{P}(D_U = k) = p_k (1 - c)^{-1}$ for $k = 0, 1, \ldots, d_c - 1$ and $\mathbb{P}(D_U = d_c) = (1 - \delta) p_{d_c} (1 - c)^{-1}$.  The degree distributions of a typical contacted individual conditioned on it being vaccinated or unvaccinated are needed.  These are given by $\mathbb{P}(\tilde{D}_V = d_c) = \delta \tilde{p}_{d_c} \tilde{p}_V^{-1}$ and $\mathbb{P}(\tilde{D}_V = k) = \tilde{p}_k \tilde{p}_V^{-1}$ for $k = d_c + 1, d_c + 2, \ldots$; and $\mathbb{P}(\tilde{D}_U = k) = \tilde{p}_k (1 - \tilde{p}_V)^{-1}$ for $k = 0, 1, \ldots, d_c - 1$ and $\mathbb{P}(\tilde{D}_U = d_c) = (1 - \delta) \tilde{p}_{d_c} (1 - \tilde{p}_V)^{-1}$.  The corresponding means are

$$\mu_{\tilde{D}_V} = \tilde{p}_V^{-1} \left( \delta d_c \tilde{p}_{d_c} + \sum_{j=d_c+1}^{\infty} j \tilde{p}_j \right)$$

and

$$\mu_{\tilde{D}_U} = (1 - \tilde{p}_V)^{-1} \left( \sum_{j=1}^{d_c-1} j \tilde{p}_j + (1 - \delta) d_c \tilde{p}_{d_c} \right).$$

In the forward and backward approximating branching processes, it is sufficient to type individuals according to whether they are vaccinated (type 1) or unvaccinated (type 2).  The mean matrix for noninitial generations of the forward process is then given by

$$M = \begin{pmatrix} \mu_{\tilde{D}_V - 1} \tilde{p}_V p_{VV} & \mu_{\tilde{D}_V - 1} (1 - \tilde{p}_V) p_{VU} \\ \mu_{\tilde{D}_U - 1} \tilde{p}_V p_{UV} & \mu_{\tilde{D}_U - 1} (1 - \tilde{p}_V) p_{UU} \end{pmatrix}.$$

The dominant eigenvalue $R_v$ of this matrix is then a threshold parameter for the model, describing (asymptotically) whether or not major outbreaks affecting a positive proportion of the population are possible.  Note that if $p_{UU} p_{VV} = p_{UV} p_{VU}$ (which is satisfied when the vaccine is all-or-nothing or nonrandom with $a = 1$ or $b = 1$ (leaky)) then rank $M = 1$ and $R_v = \mu_{\tilde{D}_U - 1} (1 - \tilde{p}_V) p_{UU} + \mu_{\tilde{D}_V - 1} \tilde{p}_V p_{VV}$.  In particular, if the vaccine is perfect, $R_v = \mu_{\tilde{D}_U - 1} (1 - \tilde{p}_V) p$.

As in the analysis of our acquaintance vaccination model we consider the offspring of an individual in two phases.  First we call (in the language of the forward process) uninfected neighbours of an infected individual potential offspring and determine their types; then we determine whether they are infected.  Also, in the same way as before, with the all-or-nothing vaccine we count as vaccinated types only individuals for which the vaccine fails.

For noninitial generations the PGFs of the potential offspring distributions $\tilde{X}_i = (\tilde{X}_{i1}, \tilde{X}_{i2})$, $i = 1, 2$, follow easily from the size-biased degree distributions given above.  Since each

neighbour takes its type independently with fixed probabilities, we have, for $s \in [0, 1]^2$, $f_{\tilde{X}_1}(s) = f_{\tilde{D}_V - 1}(g(s))$ and $f_{\tilde{X}_2}(s) = f_{\tilde{D}_U - 1}(g(s))$, where $g(s) = \tilde{p}_V s_1 + (1 - \tilde{p}_V)s_2$. The same calculations for the potential offspring distributions in the initial generation (call the random variables $X_i = (X_{i1}, X_{i2})$) show that, for $s \in [0, 1]^2$, $f_{X_1}(s) = f_{D_V}(g(s))$ and $f_{X_2}(s) = f_{D_U}(g(s))$.

To calculate the PGFs of the offspring distributions for the forward process, we again follow closely the derivation in Section 4. Consider first the case of a nonrandom vaccine action model. Denoting by $\tilde{C}_i = (\tilde{C}_{i1}, \tilde{C}_{i2})$, $i = 1, 2$, the number of vaccinated and unvaccinated individuals infected by a typical type-$i$ individual, and by $a_i$ and $b_i$ the relative susceptibilities and infectivities of the types (so $a_1 = a$, $b_1 = b$, and $a_2 = b_2 = 1$), we find that

$$f_{\tilde{C}_i}(s) = \sum_{k \in \mathbb{Z}_+^2} \frac{(1 - s)^k h_i^F(k)}{k!} f_{\tilde{X}_i}^{(k)}(s),$$

where $h_i^F(k) = \phi(\lambda b_i \sum_j a_j k_j)$. Of course, exactly the same arguments hold for the initial generation and the same formula is obtained, with $C_i = (C_{i1}, C_{i2})$ and $X_i$ in place of $\tilde{C}_i$ and $\tilde{X}_i$.

Considering the case of the all-or-nothing vaccine action model, the same modifications as described at the end of Section 4.3 applied to the above argument give corresponding formulae for the offspring distributions.

The offspring PGFs for the backward process are also derived in the same way as in Section 4; writing $\tilde{B}_i = (\tilde{B}_{i1}, \tilde{B}_{i2})$, $i = 1, 2$, for the appropriate random variable we find that $f_{\tilde{B}_i}(s) = f_{\tilde{X}_i}(h_i^B(s))$, where $h_i^B(s) = (1 - p_{1i}^I + p_{1i}^I s_1, 1 - p_{2i}^I + p_{2i}^I s_2)$, with

$$p_{11}^I = p_{VV}, \qquad p_{12}^I = p_{VU}, \qquad p_{21}^I = p_{UV}, \qquad \text{and} \quad p_{22}^I = p_{UU}.$$

The same arguments apply for the initial generation, so $f_{B_i}(s) = f_{X_i}(h_i^B(s))$, with $h_i^B(s)$ as before. The differences between the nonrandom and all-or-nothing vaccine action models again arise only in the definition of the quantities $p_{ij}^I$ and the appropriate weighted average to use to determine the overall expected relative final size of a major outbreak (cf. Section 4.4).

## Acknowledgement

## References

ANDERSSON, H. (1997). Epidemics in a population with social structures. *Math. Biosci.* **140,** 79–84.

ANDERSSON, H. (1998). Limit theorems for a random graph epidemic model. *Ann. Appl. Prob.* **8,** 1331–1349.

ANDERSSON, H. (1999). Epidemic models and social networks. *Math. Scientist* **24,** 128–147.

BALL, F. AND NEAL, P. (2008). Network epidemic models with two levels of mixing. *Math. Biosci.* **212,** 69–87.

BALL, F. AND SIRL, D. (2012). An SIR epidemic model on a population with random network and household structure, and several types of individuals. *Adv. Appl. Prob.* **44,** 63–86.

BALL, F., SIRL, D. AND TRAPMAN, P. (2009). Threshold behaviour and final outcome of an epidemic on a random network with household structure. *Adv. Appl. Prob.* **41,** 765–796.

BALL, F., SIRL, D. AND TRAPMAN, P. (2010). Analysis of a stochastic SIR epidemic on a random network incorporating household structure. *Math. Biosci.* **224,** 53–73.

BALL, F., SIRL, D. AND TRAPMAN, P. (2013). Epidemics on random intersection graphs. To appear in *Ann. Appl. Prob.*

BECKER, N. G. AND STARCZAK, D. N. (1998). The effect of random vaccine response on the vaccination coverage required to prevent epidemics. *Math. Biosci.* **154,** 117–135.

BRITTON, T., JANSON, S. AND MARTIN-LÖF, A. (2007). Graphs with specified degree distributions, simple epidemics, and local vaccination strategies. *Adv. Appl. Prob.* **39,** 922–948.

BRITTON, T., DEIJFEN, M., LAGERÅS, A. N. AND LINDHOLM, M. (2008). Epidemics on random graphs with tunable clustering. *J. Appl. Prob.* **45,** 743–756.

COHEN, R., HAVLIN, S. AND BEN AVRAHAM, D. (2003). Efficient immunization strategies for computer networks and populations. *Phys. Rev. Lett.* **91,** 247901.

DIEKMANN, O., DE JONG, M. C. M. AND METZ, J. A. J. (1998). A deterministic epidemic model taking account of repeated contacts between the same individuals. *J. Appl. Prob.* **35,** 448–462.

DURRETT, R. (2007). *Random Graph Dynamics*. Cambridge University Press.

HALLORAN, M. E., HABER, M. AND LONGINI, I. M., JR. (1992). Interpretation and estimation of vaccine efficacy under heterogeneity. *Amer. J. Epidemiol.* **136,** 328–343.

KENAH, E. AND ROBINS, J. M. (2007). Second look at the spread of epidemics on networks. *Phys. Rev. E* **76,** 036113.

KISS, I. Z., GREEN, D. M. AND KAO, R. R. (2006). The effect of contact heterogeneity and multiple routes of transmission on final epidemic size. *Math. Biosci.* **203,** 124–136.

MODE, C. J. (1971). *Multitype Branching Processes. Theory and Applications* (Modern Analytic Comput. Meth. Sci. Math. **34**). Elsevier, New York.

NEWMAN, M. E. J. (2002). Spread of epidemic disease on networks. *Phys. Rev. E* **66,** 016128.

NEWMAN, M. E. J., STROGATZ, S. H. AND WATTS, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **64,** 026118.

OLOFSSON, P. (1996). Branching processes with local dependencies. *Ann. Appl. Prob.* **6,** 238–268.

SENETA, E. (1973). *Non-Negative Matrices*. Halsted Press, New York.