Heterogeneity in susceptible–infected–removed (SIR) epidemics on lattices

Franco M. Neri[1],*, Francisco J. Pérez-Reche[2], Sergei N. Taraskin[3] and Christopher A. Gilligan[1]

[1]*Department of Plant Sciences,* [2]*Department of Chemistry, and* [3]*St Catharine's College and Department of Chemistry, University of Cambridge, Cambridge, UK*

The percolation paradigm is widely used in spatially explicit epidemic models where disease spreads between neighbouring hosts. It has been successful in identifying epidemic thresholds for invasion, separating non-invasive regimes, where the disease never invades the system, from invasive regimes where the probability of invasion is positive. However, its power is mainly limited to homogeneous systems. When heterogeneity (environmental stochasticity) is introduced, the value of the epidemic threshold is, in general, not predictable without numerical simulations. Here, we analyse the role of heterogeneity in a stochastic susceptible–infected–removed epidemic model on a two-dimensional lattice. In the homogeneous case, equivalent to bond percolation, the probability of invasion is controlled by a single parameter, the transmissibility of the pathogen between neighbouring hosts. In the heterogeneous model, the transmissibility becomes a random variable drawn from a probability distribution. We investigate how heterogeneity in transmissibility influences the value of the invasion threshold, and find that the resilience of the system to invasion can be suitably described by two control parameters, the mean and variance of the transmissibility. We analyse a two-dimensional phase diagram, where the threshold is represented by a phase boundary separating an invasive regime in the high-mean, low-variance region from a non-invasive regime in the low-mean, high-variance region of the parameter space. We thus show that the percolation paradigm can be extended to the heterogeneous case. Our results have practical implications for the analysis of disease control strategies in realistic heterogeneous epidemic systems.

Keywords: epidemics; heterogeneity; percolation

1. INTRODUCTION

One of the main challenges in epidemiology is to explain the huge variability observed in the outcome of epidemics. Such variability affects the invasion and persistence of disease in a host population, the rate of disease progress, the final fraction of the population affected, and even whether or not the disease spreads at all [1]. Much of the observed variation is owing to the inherent stochasticity of the epidemic process, acting at different spatial scales. At the small scale, *demographic* stochasticity intervenes in the interaction between host and pathogen: pathogen transmission between two individuals is in general a random event. Epidemic spread also depends upon epidemiological parameters that vary across the population owing to stochastic *environmental* factors, whereby hosts can differ in properties such as susceptibility or infectivity. When the spatial structure of the epidemic is explicit, the interplay at different scales between environmental stochasticity (or *heterogeneity*) and demographic stochasticity (henceforth, simply stochasticity) contributes to the variability of the outcome reflected in the generation of spatio-temporal patterns of infected hosts [2–4]. Among botanical epidemics in particular, such spatio-temporal variability is evident not only between different epidemics of the same host–pathogen system, but also within the same epidemic; for example through the occurrence of patches of disease within a single field, with some areas exhibiting high and others low incidences [5].

For a broad class of host–pathogen systems, transmission only occurs between neighbouring hosts. This class includes many plant pathogens, such as soil-borne pathogens transmitted between neighbouring plants [6,7], or, at a larger scale, pathogens spreading

within a mosaic of neighbouring susceptible fields or even farms [8,9]. The class also comprises animal pathogens that spread in host populations living in a fixed habitat, as has been shown, e.g. for plague in populations of gerbils inhabiting a network of connected burrows [10]. Finally, at the microscopic scale, some diseases are known to spread among contiguous cells (cf. neural ganglion cells in [11]). Nearest-neighbour interactions lead naturally to consideration of *percolation* [12]. Percolation theory provides tools and concepts to scale up from the small-scale connectivity properties of a system (e.g. whether or not two neighbours are connected) to its large-scale properties, such as the size of clusters of connected sites. In epidemiology, percolation has been used to study the 'infection path' followed by the pathogen through the system [13,14], and the connectivity of habitable sites in the landscape [7]. The most relevant result is the existence of a *spatial threshold* for epidemic invasion. The threshold separates a non-invasive, safe regime (where the pathogen can never invade the population) from an invasive regime (where the pathogen has a non-zero probability to invade). However, including the effect of host heterogeneity in the *percolation paradigm* has proved to be not straightforward. With a few notable exceptions [15–17] and recent interest in the field of complex networks [18–21], the problem has been mostly neglected.

In this paper, we propose a significant advance towards the solution of the problem of incorporating host heterogeneity in spatial systems. We study the effects of heterogeneity on a stochastic SIR (susceptible–infected–removed) model on a lattice [22]. The model is chosen because an exact mapping onto bond percolation holds in the homogeneous case [14], where an epidemiological parameter, the *transmissibility* ψ (the probability to transmit infection along a bond), is equivalent to the bond probability in percolation. The value of ψ is thus the only parameter controlling the resilience of the system to invasion, and, in particular, the invasion threshold ψ_c coincides with the bond-percolation threshold p_c^{bond}. This also implies that in a few relevant cases the value of the invasion threshold is known analytically [12]. Thus, the model provides an ideal starting point to study the effects of inherent heterogeneity in spatially explicit systems. Our aim is to understand how heterogeneity changes the invasion threshold of the system, and whether the threshold can be still predicted without the need for numerical simulations or not. One of the main challenges in the study of a heterogeneous system is to find if it can be described by means of an equivalent homogeneous mean-field system characterized by effective parameters. It is known [16] that the effect of heterogeneity depends crucially on the presence of *correlations* in ψ along different bonds. In the uncorrelated case, the system is equivalent to an effective homogeneous system with $\psi = \langle \psi \rangle$, and the threshold is given by $\langle \psi \rangle = p_c^{\text{bond}}$ [17]. In the correlated case, no such equivalence holds, and the value of the invasion threshold cannot be inferred from $\langle \psi \rangle$. The threshold is known to shift from $\langle \psi \rangle = p_c^{\text{bond}}$ (in homogeneous and heterogeneous uncorrelated systems) to $\langle \psi \rangle = p_c^{\text{site}}$

(site percolation threshold), when the variance of the transmissibility σ_ψ^2 in correlated systems is maximal [15,16,23]. This result points to σ_ψ^2 as a good measure of heterogeneity in correlated systems, but it poses several questions. Is the invasion threshold an increasing function of the variance of ψ, and in general, how much information about the threshold is provided by σ_ψ^2? More specifically, do we need to know the whole distribution of ψ to estimate the threshold?

These questions are systematically addressed in the paper. Our main result is the existence of a phase diagram in the two-dimensional space $(\langle \psi \rangle, \sigma_\psi^2)$, with a (lattice-dependent) phase boundary separating invasive and non-invasive regimes. We also show that for all the correlated systems studied, the position of the threshold is determined with excellent approximation by the two parameters $\langle \psi \rangle$ and σ_ψ^2. Higher (and in general smaller) moments of the distribution $h(\psi)$ shift the threshold value for $\langle \psi \rangle$ within the interval $[p_c^{\text{bond}}, p_c^{\text{site}}]$ only by a small amount. For the square lattice, for example, the shift is at most 3 per cent of the interval length, which can be considered negligible for most applications. Our phase diagram extends the predictive power of the percolation paradigm for SIR epidemics to heterogeneous correlated systems. These results also have relevant implications for the implementation of control: in this respect, a heterogeneous correlated system is qualitatively different from its homogeneous counterpart, because the parameter space in which to search for an optimal control strategy is now two-dimensional. While the task of finding an optimal control strategy becomes non-trivial, we nevertheless show that the problem can be solved. This can open up new possibilities for the implementation of control in spatial epidemic systems.

2. MODEL

We give here a brief description of the model and the problems involved; a more detailed version is contained in the first section of the electronic supplementary material. We consider a stochastic SIR process on a square lattice (for concreteness) with N sites, where an epidemic spreads by the nearest-neighbour transmission of a pathogen, starting from a single infected site (primary infection). Upon infection, a host (donor) i is infectious (I) for a time τ_i (infectious period) before recovering or being removed (R). During this time, it transmits the pathogen to its susceptible neighbours (recipients) j according to a Poisson process with rate β_{ij} (infection rate). The probability that the pathogen is transmitted from i to j within the time τ_i is the *transmissibility*:

$$\psi_{ij} = \psi(\beta_{ij}, \tau_i) = 1 - e^{-\beta_{ij}\tau_i}. \tag{2.1}$$

The system is *heterogeneous* when β_{ij} and τ_i are random variables. We consider them to be independent and identically distributed (iid) random variables, drawn from the distributions $f(\beta_{ij})$ and $g(\tau_i)$, respectively. In this case, the ψ_{ij} are also random (but not necessarily independent) variables, with a probability distribution $h(\psi_{ij})$ that can be calculated from $f(\beta_{ij})$ and $g(\tau_i)$.

In a finite system, the process stops after a finite time with only S and R hosts left. The crucial question we

address in this paper is how the distributions of the infection rates $f(\beta_{ij})$ and infectious periods $g(\tau_i)$ affect the probability that a *major outbreak* or *invasion* occurs [24], when the final set of R sites spans the entire lattice. We also consider whether the resilience of the system to epidemic invasion can be described by a few effective control parameters. The evaluation of the probability of invasion for the process $P_{\mathrm{inv}}\,(f(\beta_{ij}),g(\tau_i))$ and the identification of the relevant parameters controlling invasion are the main objectives of this paper.

If the system is *homogeneous* (constant parameters: $\tau_i = \tau$, $\beta_{ij} = \beta$), the identification of the relevant parameters is straightforward. The final state of the SIR process can be mapped onto a bond percolation problem [12,14], where the probability that a bond is 'open' is equal to ψ (see the electronic supplementary material for details). This leads to the following definition of a critical (threshold) value for the transmissibility,

$$\psi(\beta, \tau) = \psi_c = p_c^{\mathrm{bond}}, \qquad (2.2)$$

where p_c^{bond} is the bond-percolation threshold ($p_c^{\mathrm{bond}} = \frac{1}{2}$ for the square lattice; [12]). In an infinite system, the critical value separates an *invasive* region in the parameter space ($P_{\mathrm{inv}}(\beta, \tau) > 0$ for $\psi_c > p_c^{\mathrm{bond}}$) from a *non-invasive*, safe region ($P_{\mathrm{inv}}(\beta, \tau) = 0$ for $\psi_c < p_c^{\mathrm{bond}}$). The same holds for a *bond-heterogeneous* process ($\tau_i = \tau$ constant, β_{ij} random), for which the ψ_{ij} are iid random variables. Since different bonds are open or closed independently, the process is still equivalent to homogeneous bond percolation, and equation (2.2) still holds, once ψ is replaced by $\langle\psi\rangle_\beta$ (the average of ψ_{ij} from equation (2.1) with respect to $f(\beta)$, [17]). Hence, there still exists a single control parameter $\langle\psi\rangle_\beta$, independent of the moments of the distribution $f(\beta)$ other than the first. The solution is more challenging in the case of a *site-heterogeneous* process ($\beta_{ij} = \beta$ constant, τ_i random). The $\psi_{ij} = 1 - e^{-\beta\tau_i} = \psi_i$ here are host-dependent and are not iid random variables). The equivalent percolation problem is *correlated*, because bonds outgoing from the same site are not open or closed independently (see the electronic supplementary material for proof and an explicit example). As a result, the relationship between the quantity $\langle\psi\rangle_\tau$ (the average of ψ_{ij} with respect to $g(\tau)$) and the resilience of the system is less straightforward.

Indeed, the average transmissibility at the invasion threshold can take any value in the range

$$p_c^{\mathrm{bond}} \leq \langle\psi\rangle_\tau \leq p_c^{\mathrm{site}}, \qquad (2.3)$$

with bounds corresponding to the bond- and site-percolation thresholds p_c^{bond} and p_c^{site}. The system is always in the non-invasive regime for $\langle\psi\rangle_\tau < p_c^{\mathrm{bond}}$, and always in the invasive regime for $\langle\psi\rangle_\tau > p_c^{\mathrm{site}}$ (whether the system is in the invasive regime or not for $\langle\psi\rangle_\tau = p_c^{\mathrm{bond}}$ and $\langle\psi\rangle_\tau = p_c^{\mathrm{site}}$ depends on the topology of the lattice and is still an open problem: e.g. [25]); we do not discuss these cases here). However, in the middle interval $p_c^{\mathrm{bond}} < \langle\psi\rangle_\tau < p_c^{\mathrm{site}}$ the system can be in either regime, depending on the particular distribution $g(\tau)$. Thus, the average transmissibility $\langle\psi\rangle_\tau$ alone is no longer a good control parameter to describe the resilience of the system to invasion. The width of the interval

changes with the topology of the system: in general, $p_c^{\mathrm{bond}} = p_c^{\mathrm{site}}$ on a tree graph, while $p_c^{\mathrm{bond}} < p_c^{\mathrm{site}}$ for lattices with dimension $d \geq 2$ [25]. For the square and triangular lattices, the interval (p_c^{bond}, p_c^{site}) is equal to ($\frac{1}{2}$, $0.593\ldots$) and ($0.347\ldots$, $\frac{1}{2}$), respectively.

The challenge is to find alternative (or additional) control parameters that allow the prediction of the value of the epidemic threshold in a site-heterogeneous system. A hint comes from the fact that the critical value p_c^{bond} is obtained when the ψ_{ij} are uncorrelated, while the value p_c^{site} is obtained when the ψ_{ij} outgoing from the same site are *maximally* correlated. In the former case, the ψ_{ij} are drawn from a distribution with null variance, in the latter case from a distribution with maximal variance (see the electronic supplementary material for details). This leads to the hypothesis that the value of the threshold increases with the variance of ψ. This hypothesis is investigated below.

In what follows, we start from the analysis of the influence of correlated heterogeneity on the invasion threshold in a model system (the *mixed* case, where both the correlated and uncorrelated kinds of heterogeneity are present, is considered in the electronic supplementary material). The model is a two-host system, where heterogeneity is owing to the existence of two classes of susceptible hosts, randomly arranged on a lattice. It is used here to test how the invasion threshold is influenced by heterogeneity, and to demonstrate to what extent a description based on a homogeneous approximation using the average transmissibility fails. The variance of the transmissibility is shown to play an important role in this case. We also demonstrate that the resilience of the system to epidemic invasion is increased when the spatial pattern of the two classes of hosts is regular. The study of the two-host system leads us to the main result of the paper that links the value of the invasion threshold to the variability of a generic heterogeneous system. We proceed in two steps. In the first step, we investigate the dependence of the threshold on σ_ψ^2 for specific families of bimodal distributions for ψ (corresponding to different families of two-host systems), and show that for a given $\langle\psi\rangle$ there is a transition from the invasive (low-variance) to the non-invasive (high-variance) regime for a critical value $\sigma_\psi^2 = \sigma_{\psi,c}^2$. This demonstrates the existence of a *phase diagram* in the plane ($\langle\psi\rangle$, σ_ψ^2), where a phase boundary separates the two regimes. The second step is to show that the influence of higher-order moments of the distribution for ψ on the location of the phase boundary is negligible. We find that the position of the phase boundary for different families of distributions (both bimodal and other than bimodal, with common values of $\langle\psi\rangle$ and σ_ψ^2 but differing in all the higher moments) is approximately the same *for all distributions* tested. This means that the pair of control parameters ($\langle\psi\rangle$, σ_ψ^2) can determine if a system is in the invasive regime or not, up to small corrections owing to higher moments.

3. A TESTING GROUND: THE TWO-HOST SYSTEM

In this model system, hosts belonging to two different classes, A and B, are arranged on the square lattice,
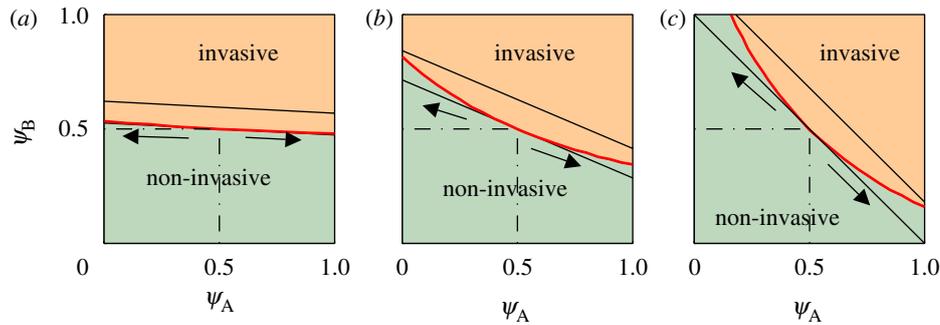
Figure 1. Effect of heterogeneity on the phase diagram of a two-host system on a square lattice. Each frame corresponds to a different value of ρ_A (concentration of A hosts; (a) $\rho_A = 0.05$; (b) $\rho_A = 0.3$; (c) $\rho_A = 0.5$). The non-invasive and invasive regimes are in green and orange, respectively. The red line marks the phase boundary, and the black solid lines are the upper and lower bounds for the phase boundary ($\langle\psi\rangle = p_c^{bond}$, p_c^{site}, respectively, from equation (2.3)). Along the phase boundary, the variance of $h(\psi)$ increases from 0 (at the homogeneous point $\psi_A = \psi_B = p_c^{bond}$) in the direction indicated by the arrows. As the level of heterogeneity ρ_A is increased (from left to right), the phase boundary moves away from the lower bound and approaches the upper bound, the minimum distance being reached at the edges ($\psi_A = 1, \psi_B = 0$) and ($\psi_A = 0, \psi_B = 1$), where the variance of $h(\psi)$ is maximal.

occupying a fraction ρ_A and $\rho_B = 1 - \rho_A$ of the lattice sites, respectively. This topology is analogous, for example, to agricultural systems where different plant species (or varieties) coexist in the same field [26,27], or at a larger scale, where some fields are treated with a pesticide and some others are not. Two different scenarios are considered below: (i) site heterogeneity with A and B hosts randomly placed on the lattice sites, and (ii) periodic arrangements of hosts. A third scenario, mixed heterogeneity, is discussed in the electronic supplementary material. In the case of site heterogeneity, each class is characterized by its own transmissibility, ψ_A and ψ_B: this occurs when two different infectious periods, τ_A and τ_B, are associated with different hosts, and the rate β is host-independent. The distribution, mean and variance of ψ are

$$h(\psi) = \rho_A \delta(\psi - \psi_A) + (1 - \rho_A)\delta(\psi - \psi_B), \quad (3.1a)$$

$$\langle\psi\rangle = \rho_A \psi_A + (1 - \rho_A)\psi_B \quad (3.1b)$$

and $\quad \sigma_\psi^2 = \rho_A(1 - \rho_A)(\psi_A - \psi_B)^2. \quad (3.1c)$

Here and in the rest of the paper, $\delta(x)$ is the Dirac delta-function with the property $\int_{-\infty}^{+\infty} f(x)\delta(x - x_0)dx = f(x_0)$. The (invasive or non-invasive) behaviour of the system depends on the values of the parameters ρ_A, ψ_A and ψ_B. In figure 1, the phase diagram is shown in the parameter plane (ψ_A, ψ_B) for different values of ρ_A. The phase boundary separating the invasive and non-invasive regimes lies between the straight lines corresponding to the lower and upper bounds, defined by the equations $\langle\psi\rangle = p_c^{bond}$ and $\langle\psi\rangle = p_c^{site}$, with $\langle\psi\rangle$ given by equation (3.1b). The point $\psi_A = \psi_B$ corresponds to a homogeneous system for which the phase boundary coincides with the lower bound. As heterogeneity is increased ($\psi_A \neq \psi_B$), the boundary approaches the upper bound. The degree of heterogeneity of the system can be described quantitatively by the variance of $h(\psi)$, given by equation (3.1c). It increases in the direction indicated by the arrows (see figure 1). The threshold for invasion increases monotonically with the variance. This behaviour is studied in

more detail in the following section. It is worth noting that in the model considered here (and in this paper in general) the transmissibility between donor and recipient depends on the donor only: recent work [19–21] has considered the more general case where the transmissibility depends both on the donor and the recipient (i.e. both host infectivity and host susceptibility are taken into account).

The results discussed so far, and the theorems giving bounds for heterogeneous SIR processes [15,16], hold for systems where the transmissibility of any host is a random variable which does not depend on the host position in space. As soon as this condition is dropped, the bounds do not hold any longer: an example is a two-host system where A and B hosts follow a regular spatial pattern. Such a system can be compared with its random counterpart with the same values of ψ_A, ψ_B and ρ_A. Comparison of the invasive regions of two periodic systems, one in the form ABAB, the other AABAAB (see figure 2), show that the periodic systems are more resilient to invasion.

4. DEPENDENCE OF THE EPIDEMIC THRESHOLD ON THE VARIANCE OF ψ

When the distribution $h(\psi)$ contains bond heterogeneity only, the resilience of a heterogeneous system to epidemic invasion is completely determined by the average transmissibility $\langle\psi\rangle$. If site heterogeneity is also present (e.g. owing to a distribution of infectious periods), the resilience of the system cannot be determined by $\langle\psi\rangle$ alone. In principle, knowledge of the whole distribution $h(\psi)$ is needed. Systems with different distributions characterized by the same $\langle\psi\rangle$ can be in different regimes (invasive or non-invasive) depending upon moments of higher order, $\langle(\psi - \langle\psi\rangle)^k\rangle$, $k = 2, 3, \ldots$ Below, we demonstrate that only the second moment plays an important role in systems with site heterogeneity.

As follows from equation (2.3), in the case of site heterogeneity the threshold $\langle\psi\rangle_c$ varies between the lower
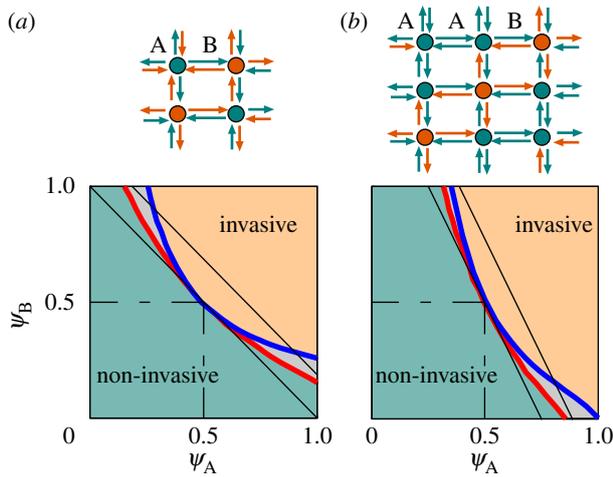
Figure 2. Comparison of the invasive regions of random and periodic two-host systems on a square lattice with given values of ψ_A, ψ_B, and ρ_A: (*a*) ABAB systems with $\rho_A = \rho_B = \frac{1}{2}$; (*b*) AABAAB systems with $\rho_A = \frac{2}{3}$. The black solid lines are the upper and lower bounds for the phase boundary of the random system (compare figure 1). The phase boundary for the random system (red line) always lies below that for the periodic system (blue line), so that the region between the two curves (shaded) is non-invasive for the periodic system and invasive for the random one, and the periodic system is in general more resilient to invasion. The orange and green regions are invasive and non-invasive, respectively, for both systems.

Figure 3. Probability of invasion P_{inv} for a process with site heterogeneity on a square 1000×1000 lattice. (*a*) P_{inv} as a function of $\langle\psi\rangle$ and σ_ψ^2. Each point in the plane $(\langle\psi\rangle, \sigma_\psi^2)$ corresponds to a choice of the parameters ψ_A and ρ_A in a family of bimodal distributions for ψ (equations (3.1)) with $\psi_B = 0$. The value $\langle\psi\rangle$ is chosen to lie in the interval (p_c^{bond}, p_c^{site}) and the variance σ_ψ^2 varies from 0 to the maximal value $\langle\psi\rangle (1 - \langle\psi\rangle)$. The phase boundary between invasive (orange) and non-invasive (green) regions is shown below the surface. (*b*) P_{inv} as a function of $\langle\psi\rangle$, with constant variance $\sigma_\psi^2 = 0.1$ (same as red curve in (*a*)). (*c*) P_{inv} as a function of σ_ψ^2, with constant average $\langle\psi\rangle = 0.55$ (same as green curve in (*a*)). Note that the horizontal axis in (*c*) is inverted.

bound p_c^{bond}, corresponding to zero variance in ψ, and the upper bound, p_c^{site}, corresponding to the maximal variance for a given value of $\langle\psi\rangle$,

$$\sigma_{\psi,max}^2 = \langle\psi\rangle(1 - \langle\psi\rangle),$$

which is achieved for a particular two-host system with $h(\psi) = \langle\psi\rangle\delta(1-\psi) + (1 - \langle\psi\rangle)\delta(\psi)$ (cf. equation (3.1*a*) and the electronic supplementary material). Therefore, when the average transmissibility falls within the window $p_c^{bond} \leq \langle\psi\rangle \leq p_c^{site}$, the system can be in different regimes (invasive or non-invasive) depending on the variance of the distribution, $0 \leq \sigma_\psi^2 \leq \sigma_{\psi,max}^2$. We carried out a numerical investigation to find the phase boundary for the SIR process in the $(\langle\psi\rangle, \sigma_\psi^2)$ plane for several different families of distributions where both $\langle\psi\rangle$ and σ_ψ^2 could be adjusted (details about the numerical approach and the distributions used are in the electronic supplementary material). The dependence of the probability of invasion P_{inv} on $\langle\psi\rangle$ and σ_ψ^2 for the family of bimodal distributions (equation (3.1*a*)) with $\psi_B = 0$ is shown in figure 3. In general, the probability of invasion P_{inv} is an increasing function of $\langle\psi\rangle$ and a decreasing function of σ_ψ^2 (see figure 3*a*). The phase transition from the invasive to the non-invasive regime can be achieved by either increasing $\langle\psi\rangle$ or decreasing σ_ψ^2 as control parameters. In the first case (red line in figure 3*a*,*b*), the transition occurs at a value $\langle\psi\rangle_c$ such that $p_c^{bond} \leq \langle\psi\rangle_c \leq p_c^{site}$. In the second case (green line in figure 3*a*,*c*), it occurs for a $\sigma_{\psi,c}^2$ such that $0 \leq \sigma_{\psi,c}^2 \leq \sigma_{\psi,max}^2$. The bound $\sigma_\psi^2 = 0$ (where $P_{inv} > 0$) corresponds to a homogeneous process equivalent to bond percolation, while $\sigma_\psi^2 = \sigma_{\psi,max}^2$ (where

$P_{inv} = 0$) corresponds to a maximally correlated process equivalent to site percolation.

The behaviour of the phase boundary between the two regimes (corresponding to the steep transition in the surface in figure 3) can be studied by projecting the surface onto the $(\langle\psi\rangle, \sigma_\psi^2)$ plane (see figure 4*a*). In the phase diagram, the phase boundary partitions the space into a non-invasive (low-$\langle\psi\rangle$, high-σ_ψ^2) and an invasive (high-$\langle\psi\rangle$, low-σ_ψ^2) region. As predicted, the phase boundary goes from the point $(p_c^{bond}, 0)$ (homogeneous system, bond percolation case) to $(p_c^{site}, p_c^{site}(1 - p_c^{site}))$; maximally heterogeneous system, site percolation case). A system with $p_c^{bond} \leq \langle\psi\rangle \leq p_c^{site}$ can be in the invasive or in the non-invasive regime depending on the value of σ_ψ^2.

By changing the average and variance of ψ in the family of bimodal distributions with $\psi_B = 0$, we are also changing the higher moments that depend on the same parameters and are not shown in figure 4*a*. The position of the phase boundary in figure 4*a* can, in general, also depend on higher moments, and thus on the particular family of distributions chosen. We investigated this effect by comparing phase diagrams for several different families of distributions. Figure 4*b* displays the phase boundaries obtained for some of the families (see the electronic supplementary material for details). The distance between the phase boundaries corresponding to different families turned out to be small for all the examples that we tested. Considering points of the boundaries with the same value of σ_ψ^2, the difference between the corresponding values of $\langle\psi\rangle$ is always $\leq 0.03\Delta_p$, where $\Delta_p = p_c^{site} - p_c^{bond}$ (cf. figure 4*b*). This shows that higher moments have a
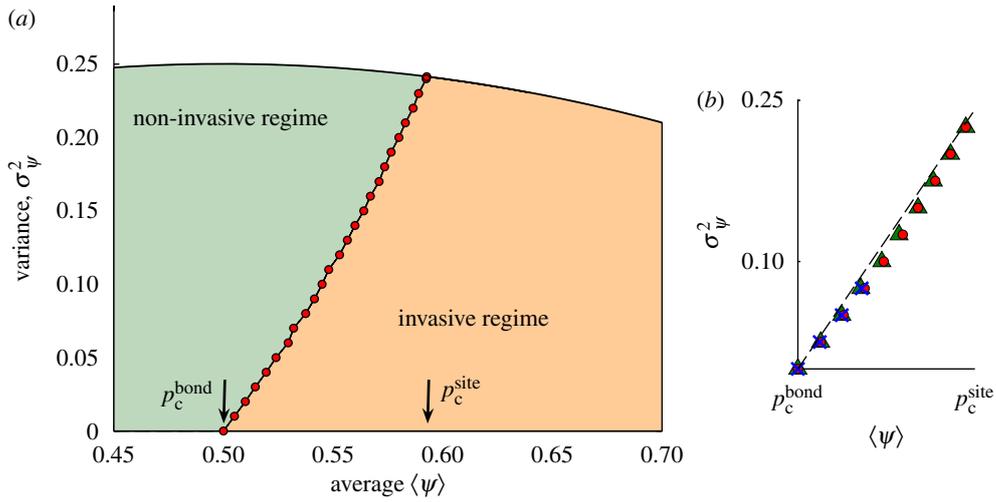
Figure 4. (*a*) Phase diagram in the ($\langle\psi\rangle$, σ_ψ^2) plane for a system with site heterogeneity. The points of the phase boundary separating the invasive and non-invasive regions (red points) were calculated numerically for the family of bimodal probability distributions for ψ given by equation (3.1*a*) with $\psi_B = 0$. The accessible points lie below the curve $\sigma_\psi^2 = \langle\psi\rangle(1 - \langle\psi\rangle)$. The values of σ_ψ^2 are exact; the error on the values of $\langle\psi\rangle$ is about $0.01\Delta_p$, $\Delta_p = p_c^{\text{site}} - p_c^{\text{bond}}$. (*b*) Points of the phase boundary on the same lattice for different families of distributions: bimodal with $\psi_B = 0$ (red circles), β distribution (green triangles), truncated Gaussian (blue crosses). The difference along the $\langle\psi\rangle$ axis between points for different distributions is less than $0.03\Delta_p$. The points lie approximately on a straight line (dashed black line).

negligible effect on the position of the phase boundary. This result was confirmed by simulations on regular lattices different from the square (see electronic supplementary material). In the case of the square lattice, the phase boundary is also well approximated by a straight line joining the two points (p_c^{bond}, 0) and (p_c^{site}, $p_c^{\text{site}}(1 - p_c^{\text{site}})$).

We conclude that the resilience of a site-heterogeneous system to epidemic invasion is mainly determined by the first two moments of the distribution for the transmissibility, $\langle\psi\rangle$ and σ_ψ^2. Changing only the higher moments of the distribution gives corrections to the position of the phase boundary that are negligible for practical applications (such as control, see §5). In other words, by going from a homogeneous to a heterogeneous system, the number of relevant control parameters increases from one to two. In the electronic supplementary material, we give a heuristic explanation of this result based on the zero-function technique introduced by Kuulasmaa [15].

Our discussion is focused on the position of the invasion threshold. We do not address here the question of the dependence of P_{inv} on heterogeneity when the system is in the invasive regime (see [15,18,19]). In the electronic supplementary material, we discuss this dependence in the case of the Bethe lattice, where an exact analytical solution can be found.

5. DISCUSSION AND CONCLUSIONS

We have developed a method to analyse the effects of heterogeneity on spread of epidemics in spatial systems. The epidemic spread for an SIR model is defined by transmissibilities between hosts. The transmissibility, in general, depends both on bond and site properties. It is convenient to distinguish between two different types of heterogeneity in transmissibility

that can occur in real systems: (i) bond heterogeneity, associated with variability of links between different hosts and (ii) site heterogeneity, determined by variability amongst hosts. The first type, bond heterogeneity, is uncorrelated when the transmissibilities between different pairs of hosts are independent of each other. Site heterogeneity is necessarily *correlated*, because all the transmissibilities from a particular donor host are dependent on the properties of the host. The analysis of SIR epidemics in systems with bond heterogeneity is relatively straightforward. Site heterogeneity is less obvious and, as it follows from our analysis, the two first moments of the transmissibility are crucial for the evaluation of the probability of invasion. One interesting finding is that in systems with correlated heterogeneity it is possible to reduce significantly the probability of invasion by increasing the variance in transmissibility and keeping the mean value unchanged. This feature is a consequence of the behaviour evident in the phase diagram for invasion in the ($\langle\psi\rangle$, σ_ψ^2) space (figure 4). This fact can be used for an effective control of epidemics, whereby both average and variance of the transmissibility can be adjusted to bring the system to the non-invasive regime.

As an example, we analyse its use for the implementation of control strategies in realistic systems with natural variability. With respect to conventional approaches for homogeneously mixed systems, the novelty lies in the fact that we are allowed to use both $\langle\psi\rangle$ and σ_ψ^2 as parameters for the optimization of the control strategy. Consider for example a homogeneous population with constant transmissibility $\psi_0 = \langle\psi\rangle > p_c^{\text{bond}}$, hence in the invasive regime (for example, a mosaic of crops, each representing a host unit). In the phase diagram (figure 5), the system is represented by a point on the $\sigma_\psi^2 = 0$ axis on the right side of the phase boundary (approximated by a straight line).
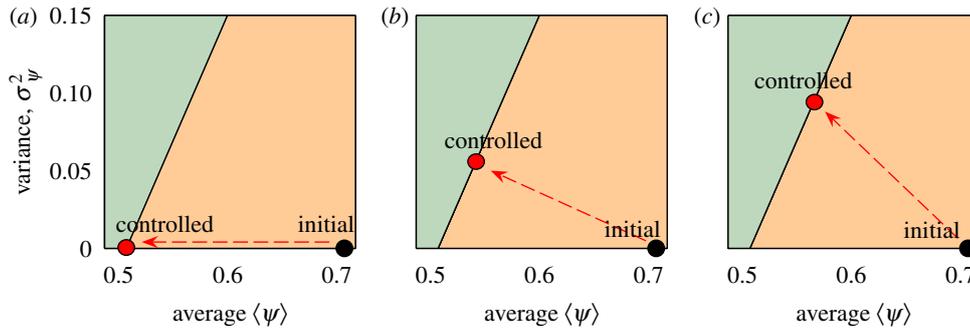
Figure 5. Effect of different control strategies on a square-lattice homogeneous system with transmissibility $\psi_0 = 0.7$, represented by a black circle on the $(\langle\psi\rangle, \sigma_\psi^2)$ plane. The red circle represents the system after treatment. The phase boundary between the invasive (orange) and non-invasive (green) regions is approximated by a straight line. The effect of treatment on a host is to decrease its transmissibility to a value $\epsilon\psi_0$, $0 \leq \epsilon < 1$. (*a*) Homogeneous strategy ($\rho = 1$, $\epsilon = \epsilon_{max} = p_c^{bond}/\psi_0$). All the hosts are treated. The final system is homogeneous with average transmissibility p_c^{bond}. (*b*) Heterogeneous strategy ($\rho = 0.4$, $\epsilon = 0.35$). A fraction ρ of the hosts is treated. The system reaches the phase boundary at a point with $\langle\psi\rangle > p_c^{bond}$, $\sigma_\psi^2 > 0$. (*c*) Maximally heterogeneous strategy ($\rho = 0.24$, $\epsilon = 0$). A smaller fraction of the hosts is treated, their transmissibility being decreased to 0. The variance of the treated system is maximal for the initial condition ψ_0.

Suppose that a control agent (a protectant) is applied before the arrival of infection in order to bring the system into the non-invasive regime. The effect of the agent on a host is to decrease the transmissibility from ψ_0 to $\epsilon\psi_0 \leq \psi_0$, where ϵ is a parameter varying from 0 (complete protection) to 1 (no treatment), depending, e.g. on the amount applied. A second adjustable parameter is the fraction ρ of hosts to be treated, giving full ($\rho = 1$) or partial ($\rho < 1$) coverage. We define a particular choice of ρ and ϵ as a *control strategy*. The minimal condition for a strategy to be successful is that the point representing the treated system lies on the phase boundary (where $P_{inv} = 0$), because further application of the control agent can only push the point further into the non-invasive regime.

We consider three strategies: (i) a 'homogeneous' treatment (only $\langle\psi\rangle$ is changed), and (ii, iii) two 'heterogeneous' treatments (both $\langle\psi\rangle$ and σ_ψ^2 are changed). In the *homogeneous* treatment (figure 5*a*), all the hosts are treated with the same amount of control agent, so that $\sigma_\psi^2 = 0$ and $\langle\psi\rangle$ is decreased to the value $\epsilon\psi_0$, reaching the boundary for $\epsilon = p_c^{bond}/\psi_0$. In a *heterogeneous* treatment (figure 5*b,c*) only a fraction $\rho < 1$ of the hosts is treated, leaving the rest unchanged. The initial homogeneous system becomes a two-host system, with average and variance of ψ given by equations (3.1*b,c*), where host A is untreated ($\psi_A = \psi_0$) and host B is treated ($\psi_B = \epsilon\psi_0$). The effect of the control agent on the system is both to decrease the average $\langle\psi\rangle$ and to increase the variance σ_ψ^2, reaching the phase boundary at a point $\langle\psi\rangle > p_c^{bond}$, $\sigma_\psi^2 > 0$. The condition that the point lies on the phase boundary given by $\sigma_\psi^2 = p_c^{site}(1 - p_c^{site})/(p_c^{site} - p_c^{bond})\langle\psi\rangle$ corresponds to a condition for ρ and ϵ to lie on a *constraint curve* in the parameter space:

$$g(\rho, \epsilon) = \frac{p_c^{site}(1 - p_c^{site})}{p_c^{site} - p_c^{bond}}((1 - \rho + \epsilon\rho)\psi_0 - p_c^{bond}) \\ - (1 - \epsilon)^2(1 - \rho)\rho\psi_0^2 = 0. \quad (5.1)$$

The constraint curve has endpoints in $(1, \epsilon_{max})$ (corresponding to the homogeneous strategy, figure 5*a*) and

$(\rho_{min}, 0)$ (maximally heterogeneous strategy, figure 5*c*), where ρ_{min} and ϵ_{max} depend on ψ_0. We remark that in a mean-field approach, where only the average transmissibility would be taken into account, there would be no difference between a heterogeneous strategy that decreases the average to $\langle\psi\rangle = \rho\epsilon\psi_0 + (1 - \rho)\psi_0 = (1 - \rho - \rho\epsilon)\psi_0$, and a homogeneous strategy decreasing the transmissibility of every host by a factor $1 - \rho - \rho\epsilon$. The difference becomes relevant in spatially explicit systems.

All the strategies are equivalent as far as the cost of control is not considered. If we introduce a cost function $c(\rho, \epsilon)$ [28], the optimal strategy can be defined as the point (ρ, ϵ) corresponding to the global minimum of $c(\rho, \epsilon)$ along the curve given by equation (5.1). The optimal strategy depends crucially on the form of the cost function. We analyse here the factorized form:

$$c(\rho, \epsilon) = \rho(1 - \epsilon)^\alpha, \quad (5.2)$$

where (i) we make the assumption that the total cost is proportional to the coverage ρ, while (ii) with the parameter $\alpha > 0$ we allow for a non-linear dependence of the cost on the *treatment efficacy*, defined as the fractional decrease of transmissibility $1 - \epsilon$ for a treated host. The non-linear dependence can be obtained, for example, by assuming that the efficacy of a given amount Q of control agent is $1 - \epsilon = Q^{1/\alpha}$, and that the cost of Q is proportional to Q. A non-linear law for ρ can also be considered, but the following results do not change qualitatively. For our choice of the cost function in equation (5.2), the minimum of $c(\rho, \epsilon)$ always coincides with one of the two endpoints of the constraint curve. However, finding which one of the two strategies is optimal is not trivial, and depends on α and the initial transmissibility ψ_0. We can eliminate one of the two parameters through $g(\rho, \epsilon) = 0$, and consider e.g. $c(\rho, \epsilon) = c(\rho, \epsilon(\rho))$ as a function of ρ only. This way, it can be shown (see figure 6*a*) that no local minima for $c(\rho, \epsilon)$ exist as a function of ρ (the same occurs for ϵ). We can then compare the values of the cost function at the two endpoints:

$$c(1, \epsilon_{max}) = (1 - \epsilon_{max})^\alpha$$
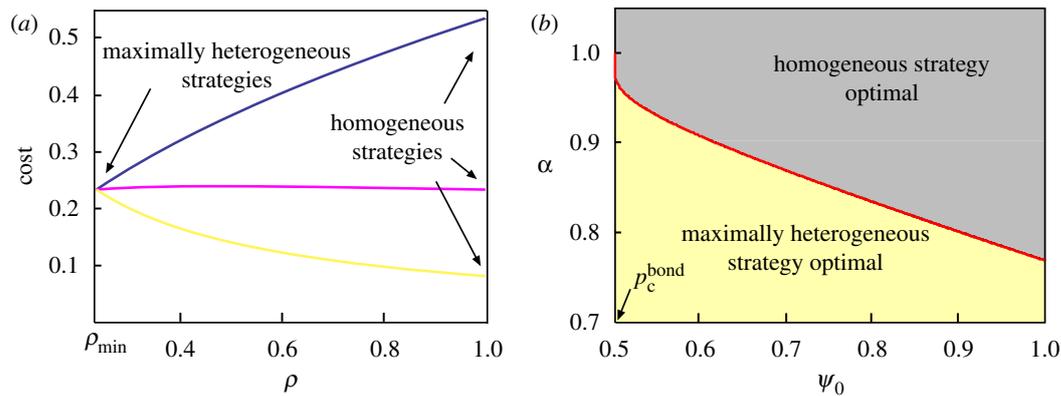
Figure 6. Optimal control strategy for a square-lattice system with initial constant transmissibility ψ_0, with treatment parametrized by ρ and ϵ, and cost function given by equation (5.2). (*a*) The cost $c(\rho, \epsilon)$ as a function of ρ for $\psi_0 = 0.7$ and $\alpha = 0.5$ (blue line), $\alpha = \kappa (\psi_0) \simeq 1.16$ (magenta), $\alpha = 2$ (yellow). The $\alpha = \kappa (\psi_0)$ line has a maximum around 0.4. (*b*) The values of ψ_0 and α determine whether the optimal strategy is the homogeneous one $(1, \epsilon_{\max})$ (grey region) or the maximally heterogeneous one $(\rho_{\min}, 0)$ (yellow region). Along the curve $\alpha = \kappa (\psi_0)$ (red line), both strategies are optimal.

and

$$c(\rho_{\min}, 1) = \rho_{\min}.$$

The homogeneous strategy $(1, \epsilon_{\max})$ (figure 5*a*) is optimal when $c(1, \epsilon_{\max}) < c(\rho_{\min}, 1)$, that is, for

$$\alpha > \frac{\log(\rho_{\min})}{\log(1 - \epsilon_{\max})} = \kappa(\psi_0). \tag{5.3}$$

Similarly, the maximally heterogeneous strategy $(\rho_{\min}, 0)$ (figure 5*c*) is optimal when $\alpha < \kappa(\psi_0)$. For $\alpha = \kappa(\psi_0)$, the two strategies are both optimal with the same cost. In general, $\kappa(\psi_0)$ is an increasing function of ψ_0 in the interval $[p_c^{bond}, 1]$ with $\kappa(p_c^{bond}) = 1$, and $\kappa(1) \simeq 0.77$ on the square lattice (see figure 6*b*). Hence, even with a simple choice for the cost function, the optimal control strategy displays a non-trivial dependence on the initial transmissibility of the system. These results are relevant for plant populations, and more generally for any system where the spatial structure of the population is known.

REFERENCES

1 Gilligan, C. A. 2002 An epidemiological framework for disease management. *Adv. Bot. Res.* **38**, 1–64. (doi:10.1016/S0065-2296(02)38027-3)

2 Durrett, R. & Levin, S. 1994 Stochastic spatial models—a user's guide to ecological applications. *Phil. Trans. R. Soc. Lond. B* **343**, 329–350. (doi:10.1098/rstb.1994.0028)

3 Shaw, M. W. 1994 Modeling stochastic processes in plant pathology. *Annu. Rev. Phytopatol.* **32**, 523–544. (doi:10.1146/annurev.py.32.090194.002515)

4 Zadoks, J. C. & van den Bosch, F. 1994 On the spread of plant disease: a theory on foci. *Annu. Rev. Phytopatol.* **32**, 503–521. (doi:10.1146/annurev.py.32.090194.002443)

5 Gibson, G. J., Gilligan, C. A. & Kleczkowski, A. 1999 Predicting variability in biological control of a plant–pathogen system using stochastic models. *Proc. R. Soc. Lond. B* **266**, 1743–1753. (doi:10.1098/rspb.1999.0841)

6 Bailey, D. J., Otten, W. & Gilligan, C. A. 2000 Saprotrophic invasion by the soil-borne fungal plant pathogen *Rhizoctonia solani* and percolation thresholds. *New Phytol.* **146**, 535–544. (doi:10.1111/j.1469-8137.2004.01086.x)

7 Otten, W., Bailey, D. J. & Gilligan, C. A. 2004 Empirical evidence of spatial thresholds to control invasion of fungal parasites and saprotrophs. *New Phytol.* **163**, 125–132. (doi:10.1111/j.1469-8137.2004.01086.x)

8 Truscott, J. E. & Gilligan, C. A. 2001 The effect of cultivation on the size, shape and persistence of disease patches in a field. *Proc. Natl Acad. Sci. USA* **98**, 7128–7133. (doi:10.1073/pnas.111548698)

9 Stacey, A. J., Truscott, J. E., Asher, M. J. C. & Gilligan, C. A 2004 A model for the invasion and spread of Rhizomania in the United Kingdom: implications for disease control strategies. *Phytopathology* **94**, 209–215. (doi:10.1094/PHYTO.2004.94.2.209)

10 Davis, S., Trapman, P., Leirs, H., Begon, M. & Heesterbeek, J. 2008 The abundance threshold for plague as a critical percolation phenomenon. *Nature* **454**, 634–637. (doi:10.1038/nature07053)

11 Pérez-Reche, F. J., Taraskin, S. N., da, F., Costa, L., Neri, F. M. & Gilligan, C. A. 2010 Complexity and anisotropy in host morphology make populations less susceptible to epidemic outbreaks. *J. R. Soc. Interface* **7**, 1083–1092. (doi:10.1098/rsif.2009.0475)

12 Stauffer, D. & Aharony, A. 1991 *Introduction to percolation theory*. London, UK: Taylor and Francis.

13 Mollison, D. 1977 Spatial contact models for ecological and epidemic spread. *J. R. Statist. Soc. B* **39**, 283–326.

14 Grassberger, P. 1983 On the critical behavior of the general epidemic process and dynamical percolation. *Math. Biosci.* **63**, 157–172. (doi:10.1016/0025-5564(82)90036-0)

15 Kuulasmaa, K. 1982 The spatial general epidemic and locally dependent random graphs. *J. Appl. Probab.* **19**, 745–758. (doi:10.2307/3213827)

16 Cox, J. T. & Durrett, R. 1988 Limit theorems for the spread of epidemics and forest fires. *Stochast. Proc. Appl.* **30**, 171–191. (doi:10.1016/0304-4149(88)90083-x)

17 Sander, L. M., Warren, C. P., Sokolov, I. M., Simon, C. & Koopman, J. 2002 Percolation on heterogeneous networks

as a model for epidemics. *Math. Biosci.* **180**, 293–305. (doi:10.1016/S0025-5564(02)00117-7)

18 Kenah, E. & Robins, J. M. 2007 Second look at the spread of epidemics on networks. *Phys. Rev. E* **76**, 036113. (doi:10.1103/PhysRevE.76.036113)

19 Miller, J. C. 2007 Epidemic size and probability in populations with heterogeneous infectivity and susceptibility. *Phys. Rev. E* **76**, 010101(R). (doi:10.1103/PhysRevE.76.010101)

20 Miller, J. C. 2008 Bounding the size and probability of epidemics on networks. *J. Appl. Probab.* **45**, 498–512. (doi:10.1239/jap/1214950363)

21 Meester, R. & Trapman, P. 2008 Bounding basic characteristics of spatial epidemics with a new percolation model. Preprint at http://arxiv.org/abs/0812.4353v2 [math.PR].

22 Murray, J. D. 2002 *Mathematical biology. I: an introduction.* Berlin, Germany: Springer.

23 Kuulasmaa, K. & Zachary, S. 1984 On spatial general epidemics and bond percolation processes. *J. Appl. Probab.* **21**, 911–914. (doi:10.2307/3213706)

24 Diekmann, O. & Heesterbeek, J. A. P. 2000 *Mathematical epidemiology of infectious diseases: model building, analysis, and interpretation.* New York, NY: Wiley.

25 Grimmett, G. R. 1999 *Percolation.* Berlin, Germany: Springer.

26 Garrett, K. A. & Mundt, C. C. 1999 Epidemiology in mixed host populations. *Phytopathology* **89**, 984–990. (doi:10.1094/PHYTO.1999.89.11.984)

27 Cook, A. R., Otten, W., Marion, G., Gibson, G. J. & Gilligan, C. A. 2007 Estimation of multiple transmission rates for epidemics in heterogeneous populations. *Proc. Natl Acad. Sci. USA* **104**, 20 392–20 397. (doi:10.1073/pnas.0706461104)

28 Pinch, E. 1993 *Optimal control and the calculus of variations.* Oxford, UK: Oxford University Press.