J. Math. Biol. (2015) 70:1065–1092
DOI 10.1007/s00285-014-0791-x

Mathematical Biology

Disease invasion on community networks with environmental pathogen movement

Joseph H. Tien · Zhisheng Shuai · Marisa C. Eisenberg · P. van den Driessche

Received: 11 July 2013 / Revised: 14 April 2014 / Published online: 5 May 2014

Abstract The ability of disease to invade a community network that is connected by environmental pathogen movement is examined. Each community is modeled by a susceptible–infectious–recovered (SIR) framework that includes an environmental pathogen reservoir, and the communities are connected by pathogen movement on a strongly connected, weighted, directed graph. Disease invasibility is determined by the basic reproduction number $\mathcal{R}_0$ for the domain. The domain $\mathcal{R}_0$ is computed through a Laurent series expansion, with perturbation parameter corresponding to the ratio of the pathogen decay rate to the rate of water movement. When movement is fast relative to decay, $\mathcal{R}_0$ is determined by the product of two weighted averages of the community characteristics. The weights in these averages correspond to the network structure through the rooted spanning trees of the weighted, directed graph. Clustering of disease "hot spots" influences disease invasibility. In particular, clustering hot spots together according to a generalization of the group inverse of the Laplacian matrix facilitates disease invasion.

J. H. Tien (✉)
Department of Mathematics, Ohio State University, Columbus, OH 43210, USA
e-mail: jtien@math.ohio-state.edu

Z. Shuai
Department of Mathematics, University of Central Florida, Orlando, FL 32816, USA
e-mail: shuai@ucf.edu

M. C. Eisenberg
Departments of Epidemiology and Mathematics, University of Michigan, Ann Arbor, MI 48109, USA
e-mail: marisae@umich.edu

P. van den Driessche
Department of Mathematics and Statistics, University of Victoria, Victoria, BC V8W 2Y2, Canada
e-mail: pvdd@math.uvic.ca

Keywords Cholera · Waterborne disease · Basic reproduction number · Spanning trees · Group inverse

Mathematics Subject Classification 05C20 · 15A09 · 92D30

1 Introduction

Consider a network of communities (patches), connected to one another according to a weighted, directed graph (digraph). When can a disease invade this network? Intuitively, the answer to this question must depend on both the network structure as well as on the properties of the individual patches. The purpose of this paper is to make this relationship explicit for a large class of networks and patch types, by examining the dependence of the basic reproduction number $\mathcal{R}_0$ for the domain as a function of the network structure and patch characteristics.

We consider community networks where the disease dynamics in each patch are governed by an SIR (susceptible–infectious–recovered) framework together with an environmental (free-living) pathogen state, through which the disease is transmitted. Our motivating disease is cholera, with free-living pathogen residing in water. As such, we subsequently refer to the environmental pathogen reservoir as water, but the modeling framework considered here can be applied to other waterborne diseases, or to pathogens with a free-living, non-aquatic state (for example, in soil or fomites). Infectious individuals shed pathogen into the water compartment within their corresponding patch. We make minimal assumptions on the form of the disease transmission and shedding functions. Regarding network structure, there are two types of movement networks to consider: movement of individuals, and pathogen movement through the environment (water). The role these networks play in the spread of cholera is an important public health consideration (Dowell and Braden 2011; Piarroux et al. 2011), and recent modeling work has begun to examine this question using a variety of approaches (Bertuzzo et al. 2010; Chao et al. 2011; Eisenberg et al. 2013). We make two assumptions on these networks: the water network is strongly connected, and infectious individuals are too sick to move. Other properties of the network are arbitrary.

Our analysis exploits the presence of two time scales in the system: the movement rate in the water network, and the pathogen decay rate in the water. In particular, our computation of the basic reproduction number $\mathcal{R}_0$ for the domain hinges upon a Laurent series expansion of a perturbation of the Laplacian matrix for the water network, with the perturbation parameter ε corresponding to the ratio of the pathogen decay rate to the rate of water movement. Small ε values thus correspond to rapid movement relative to decay. Our main mathematical results are the following:

(1) To lowest order in ε, invasibility is determined by two weighted averages of patch characteristics: the first a measure of transmission within patch, and second the pathogen decay rates in each patch. The weights in both these averages explicitly reflect the network structure through the rooted spanning trees of the weighted, directed graph.

(2) Higher order corrections correspond to a generalization of the group inverse of the Laplacian matrix for the water network. This generalized group inverse is related to how the vertices (communities) in the network can be grouped together into meta-communities.

These mathematical results in turn give the following biological insights on how the placement of disease "hot spots" in a network affects the ability of a disease to invade:

(i) Patches that are the roots of rooted spanning trees with the largest weight (loosely speaking, patches with the most net "inflow") in the network have the greatest impact on invasibility.
(ii) Clustering disease hot spots together with respect to meta-communities increases the ability of the disease to invade. The strength of this clustering effect increases as the disparity in patch risk between hot spots and non-hot spots increases.

The remainder of this paper is organized as follows. Precise description of the modeling framework and accompanying assumptions are given in Sect. 2. In Sect. 3 the Laurent series for the Laplacian matrix is computed, based upon the elegant work of Langenhop (Langenhop 1971) on nearly singular matrices. In Sect. 4 the domain $\mathcal{R}_0$ is examined in terms of this series. These two sections are the analytical heart of the paper. Examples of specific networks are given in Sect. 5. These are cartoon networks, chosen to illustrate points (i)–(ii) listed above. The paper concludes with a discussion in Sect. 6. Two appendices are included: Appendix A contains basic terminology and standard graph-theoretic results used throughout the paper, and Appendix B presents mathematical details on the group inverse and its generalization.

2 Modeling framework

2.1 Network structure and patch dynamics

Patch dynamics are assumed to correspond to a "SIWR" (susceptible–infectious–water–recovered) type ordinary differential equation (ODE) model, a generalization of the model for cholera formulated by Codeço (2001). The human population at each patch is divided into susceptible, infectious, and recovered compartments. Infection in patch i occurs through contact with an environmental water compartment at rate $g_i(S_i, W_i)$; direct (human–human) transmission is ignored. Humans in patch i recover at rate γ_i. Loss of infection-derived immunity (σ_i) and disease-induced mortality (α_i) are both included in the model, as well as natural mortality (d_i) and susceptible recruitment (A_i). The water compartment in patch i is in turn contaminated by pathogen shedding from infectious individuals at rate $h_i(I_i)$, and pathogen decays at rate δ_i. All parameters are allowed to vary between patches. A list of the model parameters is given in Table 1.

The individual patches are connected through three different networks: the movement of susceptible and recovered individuals, and the movement of water. Infectious individuals are assumed to be too sick to move. Let $M_S = [m_{ij}^S]$, $M_R = [m_{ij}^R]$, and $M_W = [m_{ij}^W]$ denote the corresponding movement matrices, with $m_{ij}^{\#}$ giving the move-

Table 1 Model parameters

$A_i > 0$	Susceptible recruitment in patch i
$d_i > 0$	Natural death rate in patch i
$\alpha_i \geq 0$	Disease-induced death rate in patch i
$\gamma_i > 0$	Recovery rate of infectious individuals in patch i
$\sigma_i \geq 0$	Rate of loss of immunity of recovered individuals in patch i
$\delta_i > 0$	Pathogen decay rate in patch i
$m_{ij}^{\#} \geq 0$	Movement rate from patch j to patch i of susceptible / recovered individuals or pathogen
$g_i(S_i, W_i) \geq 0$	Incidence function for waterborne disease transmission in patch i
$h_i(I_i) \geq 0$	Pathogen shedding rate function in patch i

ment rate from patch j to i and $m_{ii}^{\#} = 0$ in compartment $\# \in \{S, R, W\}$. We then have the following system:

$$\begin{aligned}
\dot{S}_i &= A_i - g_i(S_i, W_i) - d_i S_i + \sigma_i R_i + \sum_{j=1}^{n} \left(m_{ij}^{S} S_j - m_{ji}^{S} S_i \right) \\
\dot{I}_i &= g_i(S_i, W_i) - (d_i + \alpha_i + \gamma_i) I_i \\
\dot{R}_i &= \gamma_i I_i - (d_i + \sigma_i) R_i + \sum_{j=1}^{n} \left(m_{ij}^{R} R_j - m_{ji}^{R} R_i \right) \\
\dot{W}_i &= h_i(I_i) - \delta_i W_i + \sum_{j=1}^{n} \left(m_{ij}^{W} W_j - m_{ji}^{W} W_i \right),
\end{aligned} \tag{1}$$

for $i = 1, \ldots, n$, where n is the number of patches, and $\dot{\cdot} = \frac{d}{dt}$. The parameters A_i, d_i, γ_i and δ_i are all assumed positive, and the α_i, σ_i nonnegative. A schematic is shown in Fig. 1.

We make the following assumptions throughout:

A1: M_W is irreducible.
A2: Infectious individuals do not move between patches.
A3: The incidence functions g_i are differentiable, and satisfy $g_i(S_i, W_i) \geq 0$, $g_i(S_i, 0) = g_i(0, W_i) = 0$ for each i.
A4: The shedding functions h_i are differentiable, with $h_i(I_i) \geq 0$ and $h_i(0) = 0$ for all i.

Assumption **A1** states that the water network is strongly connected. This is a reasonable assumption for many situations, including wetlands, tidal rivers, and settings where both diffusion and advection play significant roles. The assumption that infectious individuals do not move (**A2**) is reasonable for individuals with symptomatic cholera, as the severe diarrhea associated with the disease likely restricts movement. Mild and asymptomatic cholera cases, however, are common (Kaper et al. 1995), and the extent that asymptomatic infection contributes to spatial spread of the disease is not well known. We note that the techniques used here can be applied to settings where pathogen movement occurs solely through human movement, i.e. where infectious

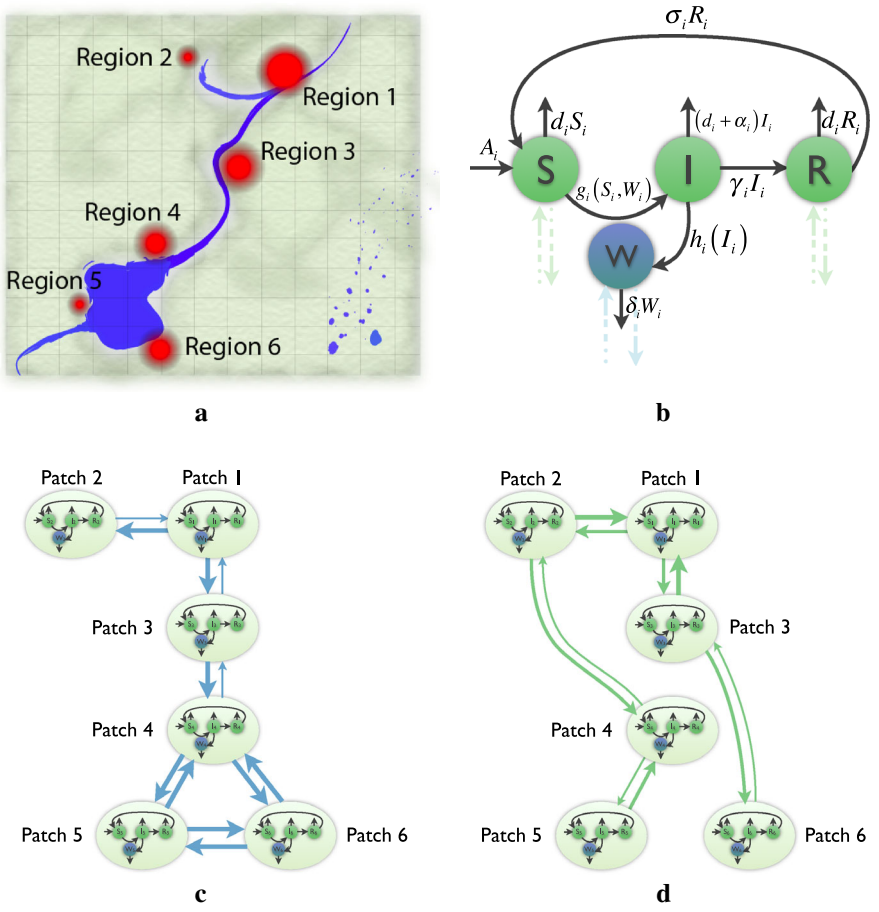

Fig. 1 Modeling framework schematic. **a** Communities in a landscape, connected by a water network. Size of the *red circles* indicates disease transmission within each patch, with *larger circles* corresponding to disease hot spots. **b** Flow diagram showing within patch dynamics. *Dashed lines* indicate movement of water (*blue*) and healthy individuals (*green*) between patches. **c** A sample water network. **d** A sample movement network for healthy individuals (infectious individuals are assumed not to move). Thickness of the *arrows* in **c** and **d** correspond to the weight of the arc between patches (color figure online)

individuals move and the environmental network plays a negligible role in pathogen spatial spread. Assumptions **A3** and **A4** are satisfied by most biologically reasonable incidence and shedding functions, such as mass action and simple saturating incidence, and linear shedding.

Under these assumptions, a solution of system (1) with nonnegative initial conditions is unique, nonnegative, and bounded. System (1) admits a unique disease free equilibrium (DFE), given by $S_i = S_i^0 = A_i/d_i$, $I_i = R_i = W_i = 0$ (Eisenberg et al. 2013). The ability of disease to invade the DFE is determined by the basic reproduction number, which in turn involves the linearized incidence and shedding rates at the DFE. We make the following additional assumption on the incidence functions:

A5: When evaluated at the disease free equilibrium, $\frac{\partial g_i}{\partial W_i} > 0$ for all i.

Assumption **A5** is a technical assumption that allows a similarity transform useful for our analysis (see Sect. 2.2). This assumption will be satisfied if the reproduction number of each patch in isolation is nonzero.

The basic reproduction number for the domain can then be computed as the spectral radius of the "next generation" matrix FV^{-1} (Diekmann and Heesterbeek 2000; van den Driessche and Watmough 2002). For system (1), this gives the following form for the domain $\mathcal{R}_0$:

$$\begin{aligned}\mathcal{R}_0 &= \rho(FV^{-1}) \\ &= \rho(D_q G_W^{-1} D_r G_I^{-1}),\end{aligned} \tag{2}$$

where

$$G_W = \begin{pmatrix} \delta_1 + \Sigma_{j=1}^n m_{j1}^W & -m_{12}^W & \dots & -m_{1n}^W \\ -m_{21}^W & \delta_2 + \Sigma_{j=1}^n m_{j2}^W & \dots & -m_{2n}^W \\ \vdots & \vdots & \ddots & \vdots \\ -m_{n1}^W & -m_{n2}^W & \dots & \delta_n + \Sigma_{j=1}^n m_{jn}^W \end{pmatrix}, \tag{3}$$

with $m_{jj}^W = 0$ for $j = 1, \dots, n$, and

$$G_I = \text{diag}\{d_i + \alpha_i + \gamma_i\}. \tag{4}$$

Note that G_W can be written as $\text{diag}\{\delta_i\} + L_W$, where L_W is the Laplacian matrix associated with M_W, i.e. the off-diagonal entries of L_W are defined as the negative of the corresponding off-diagonal entries of M_W, while the diagonal entries of L_W are chosen such that the sum of the entries in each column of L_W equals zero. Further details on the Laplacian matrix are given in Appendix A. As all off-diagonal entries of G_W are nonpositive (i.e. the Z sign pattern) and the sum of the entries of each column is positive, G_W is a non-singular M-matrix and $G_W^{-1} \geq 0$ (Berman and Plemmons 1979, p. 137). The linearized incidence and shedding at the DFE are given by the diagonal matrices D_q and D_r:

$$D_q = \text{diag}\{q_i\}, \text{ with } q_i = \frac{\partial g_i(S_i^0, 0)}{\partial W_i}, \tag{5}$$

$$D_r = \text{diag}\{r_i\}, \text{ with } r_i = h_i'(0), \tag{6}$$

where S^0 denotes the vector of susceptibles at the disease free equilibrium. Detailed analysis of equilibria and $\mathcal{R}_0$ for a version of (1) that includes direct transmission within patches is given in Eisenberg et al. (2013).

The next generation matrix can also be used to compute the basic reproduction number $\mathcal{R}_0^{(i)}$ of patch i in isolation, i.e. system (1) with $M_S = M_R = M_W = 0$:

$$\mathcal{R}_0^{(i)} = \frac{1}{\delta_i} \frac{q_i r_i}{d_i + \alpha_i + \gamma_i}. \tag{7}$$

This patch reproduction number is the product of the expected pathogen lifetime in the environment $1/\delta_i$ times a "transmission rate" $q_i r_i/(d_i+\alpha_i+\gamma_i)$. These terms will be relevant when further considering $\mathcal{R}_0$ for the entire domain.

2.2 Time scales, perturbation parameter ε, and notation

Let d_W denote a characteristic rate of water movement across the network, such that $M_W = d_W \bar{M}_W$. For example, d_W can be $\|M_W\|$ for some matrix norm. Let L denote the Laplacian matrix corresponding to the scaled water movement matrix $\bar{M}_W$, $\tau = d_W t$, and $\varepsilon = \frac{\delta}{d_W}$, where $\delta = \max_i \delta_i$. Let G denote the transfer rate matrix G_W in scaled time:

$$G = L + \varepsilon D, \text{ with} \tag{8}$$

$$D = \operatorname{diag}\left\{\frac{\delta_i}{\delta}\right\}. \tag{9}$$

From (2), the next generation matrix is then given by

$$\begin{aligned} FV^{-1} &= D_q G^{-1} D_r G_I^{-1} \\ &= \operatorname{diag}\left\{\frac{q_i}{d_W}\right\} (L+\varepsilon D)^{-1} \operatorname{diag}\left\{\frac{r_i}{d_i+\alpha_i+\gamma_i}\right\}. \end{aligned} \tag{10}$$

The Laplacian matrix L is a singular M-matrix, and by assumption **A1** has a one dimensional nullspace (Guo et al. 2006; Moon 1970). Thus for ε small, $G = L + \varepsilon D$ can be regarded as a perturbation of the singular matrix L. Laurent series expansions for nearly singular matrices have been considered by several authors (e.g. Avrachenkov et al. 2001; Langenhop 1971; Rothblum 1981; Schweitzer and Stewart 1993). Using the approach of Langenhop (1971), we show in Sect. 3 that the Laurent series for G^{-1} has a simple pole:

$$\begin{aligned} G^{-1} &= (L+\varepsilon D)^{-1} \\ &= \frac{1}{\varepsilon} X_{-1} + X_0 + \varepsilon X_1 + \ldots \end{aligned} \tag{11}$$

Then

$$FV^{-1} = \operatorname{diag}\{q_i\}\left(\frac{1}{\delta} X_{-1} + \frac{1}{d_W} X_0 + \frac{\delta}{d_W^2} X_1 + \ldots\right) \operatorname{diag}\left\{\frac{r_i}{d_i+\alpha_i+\gamma_i}\right\}. \tag{12}$$

At times, it will be convenient to work with the matrix $\widetilde{FV^{-1}} = G^{-1} D_r G_I^{-1} D_q$. Note that $\widetilde{FV^{-1}}$ and FV^{-1} are similar matrices, with $\widetilde{FV^{-1}} = D_q^{-1} FV^{-1} D_q$. The corresponding expression for $\widetilde{FV^{-1}}$ in terms of the Laurent series expansion is

$$\widetilde{FV^{-1}} = \left(\frac{1}{\delta} X_{-1} + \frac{1}{d_W} X_0 + \frac{\delta}{d_W^2} X_1 + \ldots\right) \operatorname{diag}\left\{\frac{q_i r_i}{d_i+\alpha_i+\gamma_i}\right\}. \tag{13}$$

Table 2 Summary of notation

G_W	Transfer matrix for the water compartments
G	Transfer matrix for the water compartments in scaled time
L	Laplacian matrix of G in scaled time
D	Diagonal matrix of scaled pathogen decay rates
δ	Maximum pathogen decay rate
d_W	Characteristic rate of water movement
ε	Ratio of rates of decay to movement, δ/d_W
X_i	Terms of the Laurent series expansion for G^{-1}
u	Basis for ker L, with u positive and $\sum_{i=1}^{n} u_i = 1$
$\hat{\delta}$	Average pathogen decay rate with respect to $\hat{\delta} = \sum_{i=1}^{n} \delta_i u_i$, u

The terms of the Laurent series depend heavily upon the nullspace of L. We use $u \in \mathbb{R}^n$ to denote a basis for ker L, normalized so that $\sum_{i=1}^{n} u_i = 1$. It follows from **A1** and the matrix tree theorem (Moon 1970; see also Theorem 2 in Appendix A) that all components of u can be taken to be positive, so we assume this throughout.

Table 2 summarizes some of the notation used repeatedly in the remainder of the paper.

3 Terms of the Laurent series expansion

We follow the approach of Langenhop (1971) to compute the terms of the Laurent series for $(L + \varepsilon D)^{-1}$ in (11). In addition to range L and ker L, the following spaces are fundamental to Langenhop's construction:

$$\begin{aligned} N_{1,0} &= \{x \in \mathbb{R}^n : Dx \in \text{range } L\} \\ R_{1,0} &= \{Dx : x \in \ker L\} \\ N_{0,1} &= \{x \in \mathbb{R}^n : Lx \in R_{1,0}\} \\ R_{0,1} &= \{Lx : x \in N_{1,0}\} \end{aligned} \tag{14}$$

We assume throughout that the network consists of a single strongly connected component (assumption **A1**), and thus the Laplacian matrix L has a one dimensional nullspace. The perturbation to L is of the form εD, where D has full rank. Langenhop's results in this setting gives the following result.

Lemma 1 *There exists a convergent Laurent series for* $(L + \varepsilon D)^{-1}$ *in a punctured neighborhood of 0. This Laurent series has a simple pole.*

Proof The pole is simple if $N_{0,1}$ and ker L coincide (this criterion determines $\mu = 0$ in equation (3.1) of Langenhop (1971), corresponding to a simple pole). Consider any $x \in N_{0,1}$. Then there exists a scalar α such that

$$Lx = \alpha \begin{pmatrix} \delta_1 u_1 \\ \vdots \\ \delta_n u_n \end{pmatrix}, \tag{15}$$

where $u = (u_1, \ldots, u_n)^T$ is a positive basis for $\ker L$ with $\sum_{i=1}^n u_i = 1$. Multiplying both sides of (15) by the row vector $(1, \ldots, 1)$ and noting that each column of L sums to zero gives

$$0 = \alpha \hat{\delta}, \tag{16}$$

where $\hat{\delta} = \sum_{i=1}^n \delta_i u_i$. Since $\hat{\delta} > 0$, it follows that $\alpha = 0$, and thus $x \in \ker L$ and $N_{0,1} \subseteq \ker L$. On the other hand, let $x \in \ker L$. Then $Lx = 0 \in R_{1,0}$, giving $\ker L \subseteq N_{0,1}$. Thus $N_{0,1}$ and $\ker L$ coincide and the pole is simple. Finally, as the matrix D has full rank by assumption, $\ker D \cap \ker L = \{0\}$. Theorem 4.1 of Langenhop (1971) then gives the convergence of the Laurent series in a punctured neighborhood of 0. □

Our computation of the terms in the Laurent series is based on the following facts, due to Langenhop:

Theorem 1 (Langenhop 1971, Lemma 3.11, Definition 4.1 and Lemma 4.1)

(a) $\mathbb{R}^n = N_{1,0} \oplus \ker L$.
(b) $\mathbb{R}^n = R_{1,0} \oplus \text{range } L$.
(c) *L is a bijection from $N_{1,0}$ onto* range L, *and D is a bijection from* $\ker L$ *onto* $R_{1,0}$.
(d) *$X_{-1}D$ is a projection onto* $\ker L$, *which sends $N_{1,0}$ to 0.*
(e) *$X_0 L$ is a projection on $N_{1,0}$, and X_0 sends $R_{1,0}$ to 0.*

Multiplying $G = L + \varepsilon D$ by the expansion (11) for G^{-1}, collecting powers of ε, and setting equal to the identity matrix Id gives the following equations:

$$\begin{array}{ll} (\varepsilon^{-1}) & LX_{-1} = 0 \\ (\varepsilon^{0}) & LX_0 + DX_{-1} = Id \\ (\varepsilon^{1}) & LX_1 + DX_0 = 0 \\ (\varepsilon^{2}) & LX_2 + DX_1 = 0 \\ & \vdots \end{array} \tag{17}$$

The (ε^{-1}) equation in (17) will be used for computing X_{-1}.

We also use the concept of the *index* of a square matrix, defined as the smallest integer k such that $\text{rank } L^k = \text{rank } L^{k+1}$. For networks with a single component, zero is a simple eigenvalue of L. This implies that L has index 1 (e.g. this can be seen by considering the Jordan form of L). For a matrix M of index ℓ, $\text{range } M^\ell \cap \ker M^\ell = \{0\}$ (e.g. Ben-Israel and Greville 2003, p. 155). Thus range L and $\ker L$ are complementary spaces, and $\mathbb{R}^n = \ker L \oplus \text{range } L$. With these preliminaries out of the way, we are ready to compute the terms of the Laurent series for $(L + \varepsilon D)^{-1}$.

3.1 Singular term X_{-1}

From the (ε^{-1}) equation in (17), each column of X_{-1} must be a multiple of u. Let $c_j u$ denote the jth column of X_{-1}. To determine the c_j, note that $X_{-1}D$ is a projection onto ker L, which sends $N_{1,0}$ to 0 (Theorem 1d). Setting $X_{-1}Du = u$ and multiplying both sides by the row vector $(1, \ldots, 1)$ gives

$$\delta = \sum_{j=1}^{n} c_j \delta_j u_j. \tag{18}$$

As D has full rank, each column of L is in the image of D, i.e. there exists x_j such that Dx_j is equal to the jth column of L for any j. As the columns of L are in range L, it follows that $x_j \in N_{1,0}$, and thus $0 = X_{-1}Dx_j = X_{-1}Le_j$ for $j = 1, \ldots, n$, giving $X_{-1}L = 0$. Multiplying both sides by the row vector $(1, \ldots, 1)$ and using $\sum_{j=1}^{n} u_j = 1$ gives

$$(c_1, \ldots, c_n)L = 0. \tag{19}$$

Thus $(c_1, \ldots, c_n)$ is a left nullvector of L. However, L has a one dimensional left nullspace, and $(1, \ldots, 1)$ is a left nullvector of L, giving $c_i = c_j$ for all i, j. Thus the columns of X_{-1} are identically equal to cu, for some constant c. Setting $c_j = c$ in (18) and solving for c, we have

$$c = \frac{\delta}{\hat{\delta}}, \tag{20}$$

where $\hat{\delta} = \sum_{i=1}^{n} \delta_i u_i$ as in the proof of Lemma 1. Thus we have shown the following result.

Proposition 1 *The singular term in the Laurent expansion for $G^{-1} = (L + \varepsilon D)^{-1}$ is given by $X_{-1} = \frac{\delta}{\hat{\delta}} U$, where U has identical columns corresponding to u.*

Note that a restricted case of this result is proved by using cofactors in Eisenberg et al. (2013, Lemma 4.3).

3.2 Zeroth order term X_0.

The two criteria defining X_0 in Theorem 1(e) are that X_0 is a left inverse of L restricted to $N_{1,0}$, and that X_0 send $R_{1,0}$ to 0. We first construct a map satisfying the first criterion, and then modify it to have the appropriate zero set.

The following three lemmas construct a left inverse of L restricted to $N_{1,0}$.

Lemma 2 *$Id - X_{-1}D$ is a projection onto $N_{1,0}$, which maps* ker L *to 0.*

Proof Theorem 1(d) gives that $X_{-1}D$ is a projection onto ker L, which sends $N_{1,0}$ to 0. The result then follows. □

Lemma 3 *$L - X_{-1}D$ is nonsingular.*

Proof Let x satisfy $(L - X_{-1}D)x = 0$. Then $Lx = X_{-1}Dx$. As $X_{-1}D$ is a projection onto $\ker L$, $X_{-1}Dx \in \ker L$ and thus $Lx \in \ker L$. But Lx also belongs to the range of L. As L has index 1, range L and $\ker L$ are complementary spaces, so $Lx = 0$. Thus x is in the kernel of L. Hence $0 = Lx = X_{-1}Dx = x$, showing that $(L - X_{-1}D)$ is nonsingular. □

Lemma 4 *For all $x \in N_{1,0}$, $(Id - X_{-1}D)(L - X_{-1}D)^{-1}Lx = x$.*

Proof Let $x \in N_{1,0}$. As $X_{-1}D$ sends $N_{1,0}$ to 0 by Theorem 1(d), $(L - X_{-1}D)x = Lx$ and thus $(L - X_{-1}D)^{-1}Lx = x$ for all $x \in N_{1,0}$. The result then follows from $Id - X_{-1}D$ being a projection onto $N_{1,0}$ (Lemma 2). □

Note that $L - X_{-1}D$ maps $\ker L$ one-to-one onto itself. The kernel of $(Id - X_{-1}D)(L - X_{-1}D)^{-1}$ is thus $\ker L$, rather than the desired $R_{1,0}$ specified in Theorem 1(e). In general, $\ker L \neq R_{1,0}$. The following proposition modifies the map to give the desired zero set.

Proposition 2 *Let $P : \mathbb{R}^n \to \mathbb{R}^n$ be the linear map defined by*

$$Px = \begin{cases} x, & x \in \text{range } L \\ X_{-1}Dx, & x \in R_{1,0}. \end{cases}$$

Then $X_0 = (Id - X_{-1}D)(L - X_{-1}D)^{-1}P$.

Proof According to the above definition, P acts as the identity on range L, and projects $R_{1,0}$ onto $\ker L$. As $\mathbb{R}^n = \text{range } L \oplus R_{1,0}$ (Theorem 1(b)), P is well defined.

Let $x \in R_{1,0}$. Then $Px \in \ker L$, and thus $(Id - X_{-1}D)(L - X_{-1}D)^{-1}Px = 0$ as desired.

Now let $x \in N_{1,0}$. As $Lx \in \text{range } L$, the definition of P gives

$$\begin{aligned}(Id - X_{-1}D)(L - X_{-1}D)^{-1}PLx &= (Id - X_{-1}D)(L - X_{-1}D)^{-1}Lx \\ &= x,\end{aligned}$$

by Lemma 4. Thus $(Id - X_{-1}D)(L - X_{-1}D)^{-1}P$ satisfies property (e) in Theorem 1 defining X_0. □

Propositions 1 and 2 establish the first two terms X_{-1}, X_0 of the Laurent series expansion for G^{-1}. Langenhop (1971, Definition 4.3 and Lemma 4.5) shows that the remaining terms in the Laurent series can be expressed in terms of X_0 and D, with

$$X_k = (-X_0D)^k X_0, \quad \text{for } k \geq 1. \tag{21}$$

From (21), the order zero and greater terms of the Laurent expansion for G^{-1} can be written as:

$$\begin{aligned} X_0 + \varepsilon X_1 + \varepsilon^2 X_2 + \ldots &= X_0 + \varepsilon(-X_0 D)X_0 + \varepsilon^2(-X_0 D)^2 X_0 + \ldots \\ &= \left(Id + \varepsilon(-X_0 D) + [\varepsilon(-X_0 D)]^2 + \ldots\right) X_0 \\ &= \left(Id + \sum_{k=1}^{\infty} [\varepsilon(-X_0 D)]^k\right) X_0. \end{aligned} \tag{22}$$

Equation (22) involves a Neumann series that converges if and only if $\rho(\varepsilon X_0 D) < 1$. This gives the following convergence criterion for the Laurent series in terms of ε:

$$0 < \varepsilon < \frac{1}{\rho(X_0 D)}. \tag{23}$$

3.2.1 X_0 as a generalized inverse

The zeroth order term X_0 is related to the group inverse of L (definition of the group inverse given in Appendix B). When the pathogen decay rates δ_i are equal for all patches, X_0 and the group inverse coincide. As pointed out by Schweitzer and Stewart (1993) in the general case of unequal δ_i, X_0 can be viewed as a generalization of the group inverse. This will be relevant for biological interpretation of X_0 and its impact on $\mathcal{R}_0$ (Sect. 4.2). Detailed comparison of X_0 with the group inverse is given in Appendix B.

4 $\mathcal{R}_0$ and the Laurent series

The terms of the Laurent series expansion for G^{-1} in (11) give an associated series of approximations for $\mathcal{R}_0 = \rho(FV^{-1})$ with FV^{-1} from (12), or $\rho(\widetilde{FV^{-1}})$ with $\widetilde{FV^{-1}}$ from (13). Here we consider the approximations for $\mathcal{R}_0$ corresponding to the first two terms X_{-1} and X_0 of the Laurent series, as these reveal how network structure and patch characteristics combine to determine the ability of disease to invade the network. In particular, let the *network risk* of patch i be defined as u_i, where u as previously denotes the normalized nullspace of the Laplacian matrix L, and let the *patch transmissibility* of i be defined as $q_i r_i/(d_i + \alpha_i + \gamma_i)$. As implied by the names, network risk is determined by the network structure, while patch transmissibility reflects intrinsic patch qualities. Specifically, the network risk of i is related by the matrix tree theorem to the spanning trees rooted at vertex i (Appendix A), while the patch transmissibility is the "transmission" term that appears in the basic reproduction number for patch i in isolation (7). Note that the network risk is non-dimensional, whereas patch transmissibility has units of inverse time. In this section, we show that $\mathcal{R}_0$ is closely tied to both these quantities.

4.1 $\mathcal{R}_0$ and averaging on networks

To lowest order, $G^{-1} \approx \frac{1}{\varepsilon} X_{-1}$ from (11), giving $\widetilde{FV^{-1}} \approx \frac{1}{\delta} X_{-1}\,\mathrm{diag}\{\frac{q_i r_i}{d_i+\alpha_i+\gamma_i}\}$ from (13). Proposition 1 then gives

$$\mathcal{R}_0 \approx \frac{1}{\hat{\delta}} \sum_{i=1}^{n} \frac{q_i r_i}{d_i + \alpha_i + \gamma_i} u_i$$
$$\equiv \hat{\mathcal{R}}_0. \tag{24}$$

We have used the fact that FV^{-1} and $\widetilde{FV^{-1}}$ are similar matrices, and thus have the same eigenvalues. The dominant (i.e. corresponding to $\hat{\mathcal{R}}_0$) right and left eigenvectors for $\widetilde{FV^{-1}}$ are u and

$$w = \left(\frac{q_1 r_1}{d_1 + \alpha_1 + \gamma_1}, \ldots, \frac{q_n r_n}{d_n + \alpha_n + \gamma_n} \right)^T, \tag{25}$$

respectively.

Intuitively, when movement between patches is fast (small ε) we expect the domain $\mathcal{R}_0$ to be some sort of average involving the patch parameters. Equation (24) shows that this is indeed the case, and specifies how the average should be taken to respect the network structure. The Laplacian matrix generates a regular Markov process, and to lowest order, averaging occurs according to the stationary distribution u of this process. The network structure thus appears in (24) through the weights u_i associated with each node. Indeed, this relationship between u_i and the network can be made explicit: consider the set of all spanning trees rooted at vertex i, oriented inward towards the root. Assign a weight to each in-tree, given by the product of all edge weights of the tree. The matrix tree theorem states that u_i is proportional to the sum of the weights of all in-trees rooted at i (Appendix A).

There are two averages that appear in the approximation $\hat{\mathcal{R}}_0$: the mean pathogen decay rate, $\hat{\delta} = \sum_{i=1}^{n} \delta_i u_i$, and an average transmission rate $\sum_{i=1}^{n} \frac{q_i r_i}{d_i + \alpha_i + \gamma_i} u_i$. The domain $\mathcal{R}_0$ is approximately the product of the average transmission rate times the reciprocal of the mean pathogen decay rate $1/\hat{\delta}$. In the special case where the pathogen decay rates are equal for each patch, $\hat{\mathcal{R}}_0$ becomes the weighted average of the patch reproduction numbers $\mathcal{R}_0^{(i)}$, with weights according to the stationary distribution u_i. This result in the restricted case of equal δ_i is proven using a different approach involving cofactor expansion in Eisenberg et al. (2013).

When the pathogen decay rates vary between patches, $\hat{\mathcal{R}}_0$ can still be expressed as a weighted average of the patch reproduction numbers $\mathcal{R}_0^{(i)}$, with respect to a different probability measure. From Eqs. (24) and (7),

$$\hat{\mathcal{R}}_0 = \frac{1}{\hat{\delta}} \sum_{i=1}^{n} \frac{1}{\delta_i} \frac{q_i r_i}{d_i + \alpha_i + \gamma_i} \delta_i u_i$$
$$= \frac{1}{\hat{\delta}} \sum_{i=1}^{n} \mathcal{R}_0^{(i)} \delta_i u_i$$
$$= \mathrm{E}[\mathcal{R}_0^{(i)}], \tag{26}$$

where the expectation is taken with respect to the probability measure $\delta_i u_i / \hat{\delta}$. Note that this measure involves both the network structure (u_i) and patch characteristics (δ_i).

4.2 $\mathcal{R}_0$, the generalized group inverse, and residence times

The matrix G corresponds to a singular perturbation from the regular Markov process generated by L, as including pathogen decay converts all of the states from recurrent to transient. The preceding discussion shows that to lowest order, $\mathcal{R}_0$ involves averaging according to the stationary distribution u of the unperturbed, regular process. Here we show that the next order correction to $\mathcal{R}_0$ involves deviations from this stationary distribution, which can be described by the generalized group inverse X_0 of L.

The entries $[G^{-1}]_{ij}$ give the expected time spent in patch i starting from patch j under the transient Markov process generated by G. The first term X_{-1} in the Laurent series for G^{-1} has identical columns proportional to u. To lowest order, then, the expected residence times under the transient process are proportional to the stationary distribution of the unperturbed, regular Markov process. As the columns of X_{-1} are identical, to lowest order the time spent in each patch is independent of the starting patch. The second term in the Laurent series X_0 then provides a measure of the deviation in expected residence times between the regular (unperturbed) and transient (perturbed) Markov processes. These deviations depend upon the initial conditions.

The entries of X_0 are related both to the network structure, as well as to the pathogen decay rates δ_i. To see the connection with network structure, consider the special case where the pathogen decay rates are equal for all patches. In this case, X_0 corresponds to the group inverse of L (see Sect. 3.2.1 and Appendix B). Meyer (1975) shows that the group inverse contains all the information associated with the classical fundamental matrix of a Markov process. In particular, for a regular Markov process the entries of the group inverse can be interpreted as the deviation from the expected time spent in patch i, due to starting from j:

$$L^{\#}_{ij} = \lim_{T\to\infty} \left[(\text{Expected time spent in } i \text{ starting from } j) - Tu_i\right]. \tag{27}$$

In the general situation where the decay rates differ between patches, the group inverse and X_0 no longer coincide, and the entries of X_0 depend upon the values of δ_i. The entries of X_0 can be interpreted analogously to (27), taking into account different decay rates:

$$[X_0]_{ij} = (\text{Expected time spent in } i \text{ starting from } j) - \frac{1}{\hat{\delta}}u_i + \mathcal{O}(\varepsilon). \tag{28}$$

The term $(1/\hat{\delta})u$ corresponds to the expected residence times in each patch to lowest order, and reflects averaging according to the unperturbed regular Markov process. There is no dependence of this term on the initial conditions, and the reciprocal of the mean decay rate $1/\hat{\delta}$ corresponds to the expected time to absorption. The generalized group inverse X_0 then gives a next order correction to the residence times that takes the initial conditions into account.

The effect of these residence time fluctuations on $\mathcal{R}_0$ can be seen by taking the first two terms of the Laurent series (13), viewing the X_0 term as a perturbation, and computing the corresponding perturbation in the dominant eigenvalue $\hat{\mathcal{R}}_0$ of the one term approximation:

$$\begin{aligned}\mathcal{R}_0 &\approx \rho\left(\left(\frac{1}{\delta}X_{-1}+\frac{1}{d_W}X_0\right)\operatorname{diag}\left\{\frac{q_k r_k}{d_k+\alpha_k+\gamma_k}\right\}\right)\\ &\approx \rho\left(\frac{1}{\delta}X_{-1}\operatorname{diag}\left\{\frac{q_k r_k}{d_k+\alpha_k+\gamma_k}\right\}\right)+\frac{1}{d_W}\sum_{i,j=1}^{n}s_{ij}\left[X_0\operatorname{diag}\left\{\frac{q_k r_k}{d_k+\alpha_k+\gamma_k}\right\}\right]_{ij}\\ &= \hat{\mathcal{R}}_0+\frac{1}{d_W}\sum_{i,j=1}^{n}s_{ij}[X_0]_{ij}\frac{q_j r_j}{d_j+\alpha_j+\gamma_j},\end{aligned} \tag{29}$$

where s_{ij} is the sensitivity of $\hat{\mathcal{R}}_0$ to the i,j entry of $X_{-1}\operatorname{diag}\{\frac{q_i r_i}{d_i+\alpha_i+\gamma_i}\}$. Using the dominant eigenvectors u and w (25) to compute these sensitivities (e.g. Horn and Johnson 1985, p. 372) gives

$$\begin{aligned}s_{ij} &= \frac{w_i u_j}{\langle w,u\rangle}\\ &= \frac{1}{\langle w,u\rangle}\frac{q_i r_i}{d_i+\alpha_i+\gamma_i}u_j.\end{aligned} \tag{30}$$

Combining (29) and (30) gives

$$\mathcal{R}_0-\hat{\mathcal{R}}_0 \approx \frac{1}{d_W}\sum_{i,j=1}^{n}\frac{1}{\langle w,u\rangle}\frac{q_i r_i}{d_i+\alpha_i+\gamma_i}\frac{q_j r_j}{d_j+\alpha_j+\gamma_j}[X_0]_{ij}u_j. \tag{31}$$

It is instructive to consider the sum in (31) in more detail. Note that $\langle w,u\rangle = \sum_{j=1}^{n}\frac{q_j r_j}{d_j+\alpha_j+\gamma_j}u_j$. Let

$$z_j = \frac{q_j r_j}{d_j+\alpha_j+\gamma_j}u_j, \tag{32}$$

$$\mathrm{E}_i[X_0] = \frac{1}{\langle w,u\rangle}\sum_{j=1}^{n}[X_0]_{ij}z_j. \tag{33}$$

For fixed i, $\mathrm{E}_i[X_0]$ is a weighted average of the entries in row i of X_0. The weights in this average are given by z_j, the product of the patch transmissibility of j times the expected time spent in j according to the unperturbed Markov process. The sum in (31) then becomes

$$\begin{aligned}&\sum_{i,j=1}^{n}\frac{1}{\langle w,u\rangle}\frac{q_i r_i}{d_i+\alpha_i+\gamma_i}\frac{q_j r_j}{d_j+\alpha_j+\gamma_j}[X_0]_{ij}u_j\\ &\quad=\sum_{i=1}^{n}\frac{q_i r_i}{d_i+\alpha_i+\gamma_i}\sum_{j=1}^{n}\frac{1}{\langle w,u\rangle}[X_0]_{ij}z_j\\ &\quad=\sum_{i=1}^{n}\frac{q_i r_i}{d_i+\alpha_i+\gamma_i}\mathrm{E}_i[X_0]\\ &\quad\equiv\sum_{i=1}^{n}\Delta\mathcal{R}^{(i)},\end{aligned} \tag{34}$$

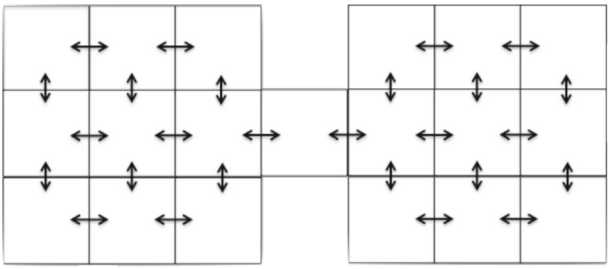

Fig. 2 Network schematic of two 3×3 grids, connected by a single "bridge" patch

giving from (31)

$$\mathcal{R}_0 \approx \hat{\mathcal{R}}_0 + \frac{1}{d_W} \sum_{i=1}^{n} \Delta\mathcal{R}^{(i)}. \tag{35}$$

Each patch i contributes a correction term $\Delta\mathcal{R}^{(i)}$ given by the product of the patch transmissibility of i with $\mathrm{E}_i[X_0]$. Using (28), $\mathrm{E}_i[X_0]$ is a weighted average of the deviations from the expected time spent in i due to starting from j. These latter weights, given in (32), also involve the patch transmissibilities. There are thus three parts to each $\Delta\mathcal{R}^{(i)}$: (1) the importance of patch i, as measured by the patch transmissibility; (2) fluctuations from the expected time spent at i, due to transients associated with the initial conditions; (3) the relative importance of the different possible starting points (in terms of patch transmissibility, together with the likelihood of occurrence of each starting point).

4.3 Meta-communities and the generalized group inverse

The interpretation of X_0 in terms of deviations of expected residence times from u (28) suggests that X_0 can be used to identify sets of vertices that form natural groupings ('meta-communities'). Let $\mathcal{S}$ be a set of vertices such that $[X_0]_{ij}$ is large for all $i, j \in \mathcal{S}$. Then flows starting in $\mathcal{S}$ tend to stay in $\mathcal{S}$, and thus we say that $\mathcal{S}$ forms a *meta-community* within the network. Meta-communities arise when there are bottlenecks to mixing between $\mathcal{S}$ and the remaining vertices in the network. For example, consider the network shown in Fig. 2 consisting of two nearly disconnected components, connected through a narrow bridge. The bridge serves as a bottleneck to mixing, and thus $[X_0]_{ij}$ is large and positive for i, j on the same side of the bridge, compared with large and negative for vertices on opposite sides of the bridge (see Sect. 5.3 for an illustration).

The placement of disease "hot spots" with respect to these meta-communities affects $\mathcal{R}_0$. From the preceding discussion of (34)–(35), placing disease hot spots together within a meta-community tends to facilitate disease invasion. This is considered further in Sect. 5.3.

5 Examples

Let us make the results from the preceding sections concrete by considering some examples.

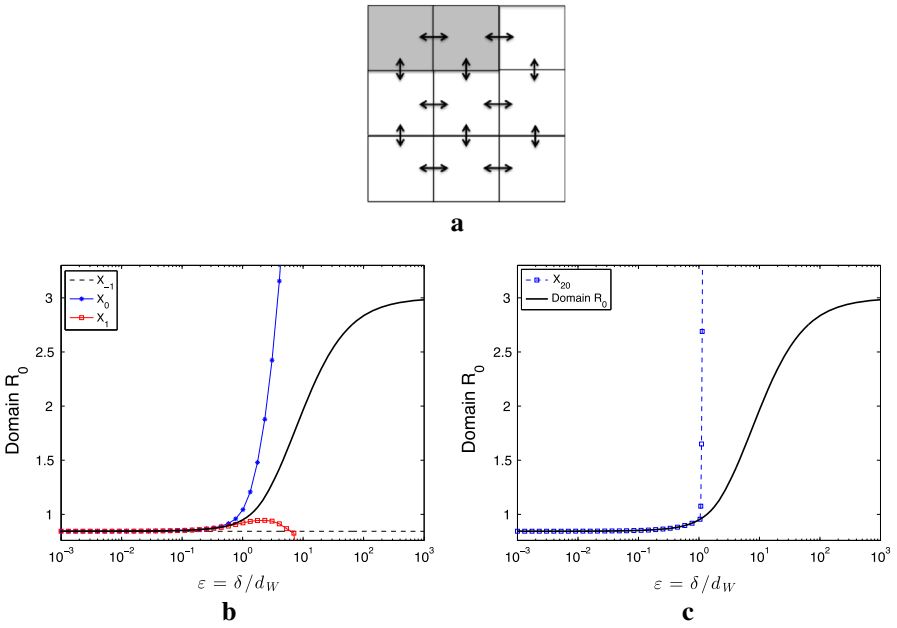

Fig. 3 **a** Network schematic of a 3×3 grid with nearest neighbor connections. *Shaded* patches have patch $\mathcal{R}_0$ values of 3, while *non-shaded* patches have patch $\mathcal{R}_0 = 0.7$. *Shaded* patches have $1/\delta_i = 30$ days, while *non-shaded* patches have $1/\delta_i = 7$ days. All patches have a 3 day infectious period, case fatality rate of 0.01, a 50 year expected lifespan, and linearized incidence $q_i = 0.1835$ per day and shedding $r_i = 0.1835$ per milliliter per day. **b** Approximations for $\mathcal{R}_0$ for the domain, by including only the X_{-1} term (*dashed line*), through X_0 (*blue asterisks*), and through X_1 (*red squares*). The true domain $\mathcal{R}_0$ corresponds to the *solid black line*. **c** Quality of the approximation from using the first 22 terms of the Laurent series (through X_{20}). The approximation is excellent up to the radius of convergence (color figure online)

5.1 Convergence of the Laurent series, quality of the approximation, and balanced networks

Theorem 4.1 of Langenhop (1971) gives convergence of the Laurent series in a punctured neighborhood of zero. The size of this neighborhood, and the quality of the approximation to the domain $\mathcal{R}_0$ from using a finite number of terms of the series (13), depends upon the specific network considered. Figure 3a shows one sample network, consisting of a 3×3 grid with equal connections between non-diagonal neighbors. This is an example of a *balanced network*, defined as a network where the net outflow is equal to the net inflow for each patch (Moon 1970). Balanced networks can be viewed as a generalization of symmetric networks: symmetric networks such as Fig. 3a are necessarily balanced, but asymmetric networks can be balanced as well. For balanced networks, the network risks u_i are equal for all i (i.e. $u_i = 1/n$; see Appendix A), and from (24) the lowest order approximation to $\mathcal{R}_0$ is simply the arithmetic mean of the patch transmissibilities $q_i r_i/(d_i + \alpha_i + \gamma_i)$ times the reciprocal of the mean of the pathogen decay rates δ_i.

Figure 3b shows the true value of the domain $\mathcal{R}_0$ computed directly from (2) over a range of ε values, compared with the approximate $\mathcal{R}_0$ values obtained from using one, two, and three terms in the Laurent series (computed from Propositions 1, 2 and equation (21)). The lowest order approximation $\hat{\mathcal{R}}_0$ (dashed line) is reasonable for ε values through 10^{-1}. Including additional terms X_0 (blue) and X_1 (red) improves the approximation, extending the range of ε where the approximation is reasonable to $\mathcal{O}(1)$. The radius of convergence here can be computed from (23) as $\varepsilon = 1.04$. For this example, then, accurate approximations to $\mathcal{R}_0$ can be obtained up to the radius of convergence from using only the first few terms of the Laurent series. Note, however, that the true value of $\mathcal{R}_0$ varies relatively little for ε less than the radius of convergence.

5.2 Spanning trees and network risk: inflow versus outflow

For networks that are not balanced, the stationary distribution u is not uniform, and thus the network risk differs between vertices. It is of interest to identify vertices with high network risk, as these vertices may contribute the most weight in the expression for $\hat{\mathcal{R}}_0$ (24). We use two toy examples here to illustrate that high network risk is associated with vertices with large net inflow.

5.2.1 River

Consider the network shown in Fig. 4: the four patches are arrayed in a line, with flow occurring only between neighboring patches, at rate a to neighbors on the right, and rate b to those on the left. The flow rate is biased, with $a > b$. The network is thus a caricature of flow on a river, with the downstream direction to the right.

For the simple network as in Fig. 4 with n patches arrayed in a line, the rooted spanning trees (and thus the network risks) can be computed explicitly. When rooted at vertex k, each vertex has a single spanning tree oriented inwards towards the root, with $k-1$ downstream edges (weight a) and $n-k$ upstream edges (weight b). The weight of the in-tree rooted at k is then $a^{k-1}b^{n-k}$, giving $u_k = \frac{a^{k-1}b^{n-k}}{\sum_{i=1}^{n} a^{i-1}b^{n-i}}$. Comparing the network risks of adjacent patches gives $\frac{u_{k+1}}{u_k} = \frac{a}{b}$, with network risk increasing downstream.

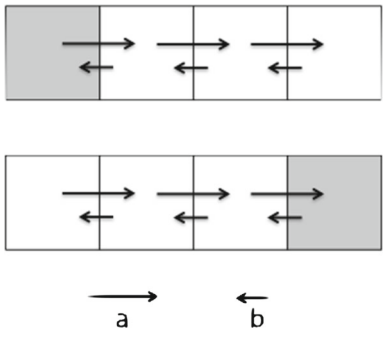

Fig. 4 Cartoon one dimensional directed flow. Flow rate of a to the *right* ("downstream"), and rate b to the *left* ("upstream"). *Shaded* patches denote disease hot spots. *Top* figure shows a network with a disease hot spot located upstream, and *bottom* figure shows a network with disease hot spot located downstream

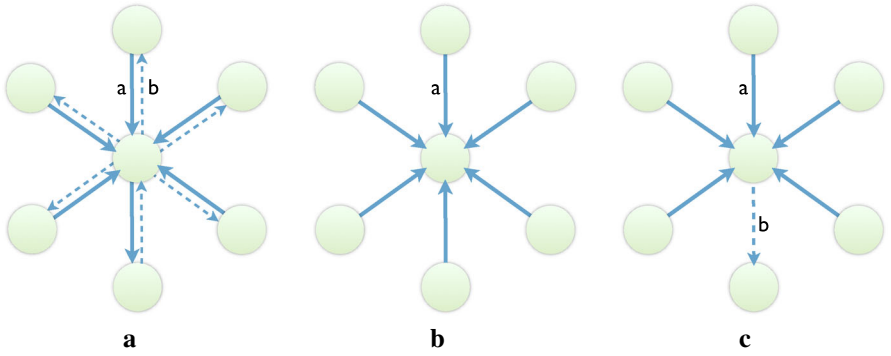

Fig. 5 Star arrangement. **a** Flow diagram. Flow rate from periphery to center at rate a, and from center to periphery at rate b. **b** Spanning tree rooted at the center. **c** Spanning tree rooted at a periphery patch. In **b** and **c**, *arcs* point to the root vertex

Suppose now that there are two types of patches: disease "hot spots" with high patch transmissibilities $q_i r_i/(d_i+\alpha_i+\gamma_i)$ or long pathogen lifetimes $1/\delta_i$, and non-hot spots with small values for these quantities. How does the placement of hot spots in the network affect the domain $\mathcal{R}_0$? From the preceding calculation, the further downstream a hot spot is located, the greater the value of $\hat{\mathcal{R}}_0$. From (29), for sufficiently large d_W, the X_{-1} term in the Laurent series dominates, $\mathcal{R}_0$ approaches $\hat{\mathcal{R}}_0$, and thus $\mathcal{R}_0$ is maximized with furthest downstream placement of a hot spot. In the situation where there is a single disease hot spot, the best case scenario (i.e. lowest $\mathcal{R}_0$) corresponds to putting the hot spot in the furthest upstream position (Fig. 4, top), and the worst case scenario (highest $\mathcal{R}_0$) corresponds to the hot spot located in the furthest downstream position (Fig. 4, bottom). It is thus the vertex with the largest inflow that contributes most greatly to disease spread, as opposed to the vertex with the greatest outflow.

5.2.2 Star

Figure 5 shows a star arrangement for the network: n peripheral patches surround a single, central (hub) patch. Flow from the periphery into the center occurs at rate a, and flow in the opposite direction from center to the periphery at rate b. The rooted spanning trees for center and periphery are shown in Fig. 5b, c, respectively, giving weight $a^{n-1}b$ for trees rooted at the periphery, and weight a^n when rooted at the center. The network risk for the center patch is then

$$u_{center} = \frac{a}{a+nb}. \tag{36}$$

In the case where $a \gg nb$, inflow to the center dominates, the center network risk is nearly one, and the ability of disease to invade is largely determined by the center patch parameters. On the other hand, when $nb \gg a$, outflow dominates and the center network risk is small. In this case it is the parameters of the periphery that determine invasibility.

5.3 Bottlenecks to mixing, meta-communities, and disease hot spots

As mentioned in Sect. 4, the generalized group inverse X_0 can be used to identify groups of vertices in which flow tends to be trapped. These groups thus form meta-communities reflecting network substructure, and the placement of disease hot spots with respect to these meta-communities affects $\mathcal{R}_0$. Here we illustrate this with an example corresponding to the balanced network shown in Fig. 2, consisting of two 3×3 grids, each with symmetric nearest neighbor connections as in the preceding paragraph, but connected to one another by a single patch acting as a bridge. For any balanced network, the network risks u_i are equal, and the lowest order approximation $\hat{\mathcal{R}}_0$ is independent of hot spot placement. Any dependence of $\mathcal{R}_0$ on hot spot location thus reflects higher order terms in the Laurent series.

The single bridge patch serves as a bottleneck to mixing between the two 3×3 blocks. This bottleneck is reflected in the entries of X_0, with large positive $[X_0]_{ij}$ values for i, j within the same block, and large negative values for i, j in different blocks, as shown in Fig. 6a, b. Each 3×3 block thus comprises a meta-community. To see how hot spot placement with respect to these meta-communities affects $\mathcal{R}_0$, consider $\Delta\mathcal{R}^{(i)}$ from (34, 35) for a disease hot spot. The contribution of a given patch $\Delta\mathcal{R}^{(i)}$ to $\mathcal{R}_0$ is proportional to the patch transmissibility, and thus hot spots (patches with high patch transmissibility) are especially important in determining the ability of disease to invade the network. The expression for $\Delta\mathcal{R}^{(i)}$ involves a weighted average of row i of X_0, with weights corresponding to z_j (32, 33). For a balanced network, the z_j are proportional to the patch transmissibilities. Whether disease hot spots are clustered together or not is therefore important for $\mathcal{R}_0$: placing two hot spots in the same meta-community increases $\Delta\mathcal{R}^{(i)}$ and thus $\mathcal{R}_0$, whereas placing hot spots in different meta-communities decreases $\mathcal{R}_0$.

Three different configurations for disease hot spots are shown in Fig. 7a–c: (a) four hot spots clustered together in a corner, on the same side of the bridge, (b) three hot spots clustered together, with the fourth hot spot located on the other side of the bridge, and (c) two hot spots apiece on either side of the bridge. The domain $\mathcal{R}_0$ values for these configurations are shown in the bottom of Fig. 7 as a function of ε. As expected, clustering the hot spots together on the same side of the bridge (as in configuration (a)) gives higher domain $\mathcal{R}_0$ values than splitting the hot spots on either side of the bridge (configuration (c)), with configuration (b) intermediate between the two. Note that this ordering holds for all ε values. The relative effect of clustering on $\mathcal{R}_0$ depends upon the value of ε, as can be seen in Fig. 8a. The largest relative change in $\mathcal{R}_0$ when comparing clustered vs. non-clustered hot spot locations occurs for intermediate ε values, with peak change in $\mathcal{R}_0$ of over 40 % between configurations Fig. 7a, c.

Both the extent of the bottleneck as well as the disparity in patch transmissibilities between hot spots and non-hot spots affect the difference in $\mathcal{R}_0$ between configurations Fig. 7a and Fig. 7c. Decreasing the flow to the bridge patch between the two 3×3 blocks results in a more severe bottleneck to mixing, more clearly defined meta-communities according to X_0, and thus a more marked difference in $\mathcal{R}_0$ between Fig. 7a and Fig. 7c, as shown in Fig. 8b. The effect of clustering hot spots together within a meta-community on $\mathcal{R}_0$ likewise becomes more marked as the difference in patch transmissibilities between hot spots to non-hot spots increases. Increasing the

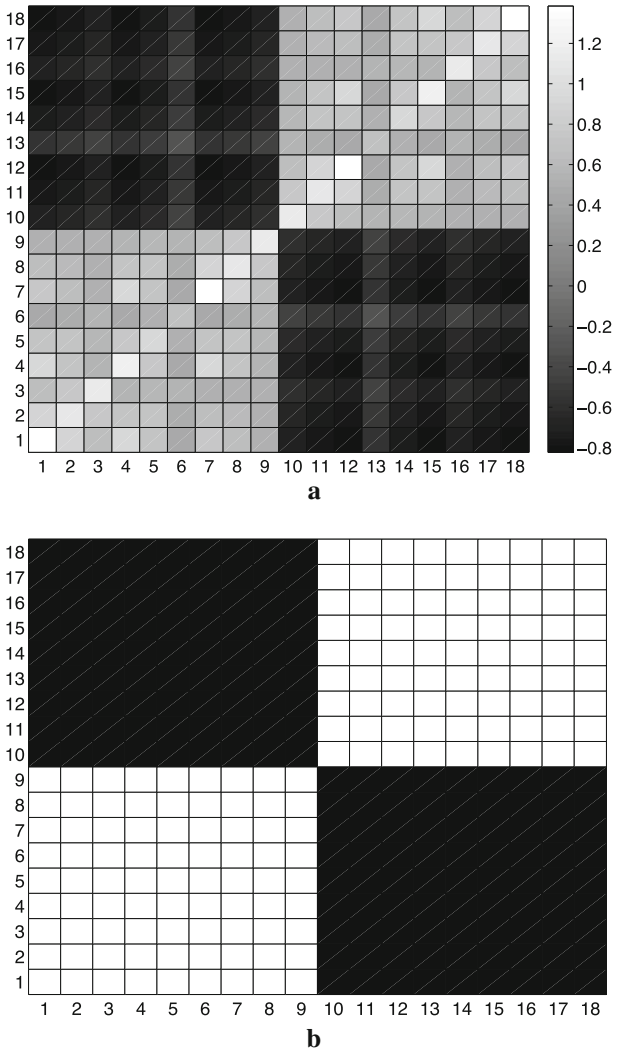

Fig. 6 The zeroth order term X_0 in the Laurent series corresponding to the configuration shown in Fig. 2, with pathogen decay rate $\delta_i = 1/30\ \text{day}^{-1}$ for all patches. Patches *left* of the bridge are labeled 1 through 9, and to the *right* of the bridge 10 through 18. **a** Entries of X_0. *Grayscale* corresponds to value of the entry, with lighter shades corresponding to higher values. **b** Plot of the sign of $[X_0]_{ij}$. Positive values shown in *white*, negative values in *black*

ratio of hot spot to non-hot spot patch transmissibilities both increases the domain $\mathcal{R}_0$ (Fig. 8c) as well as the relative effect of clustering of the hot spots (Fig. 8d).

6 Discussion

The analysis in this paper highlights the fundamental role of rooted spanning trees and the generalized group inverse of the Laplacian in determining the ability of disease to invade a network of communities.

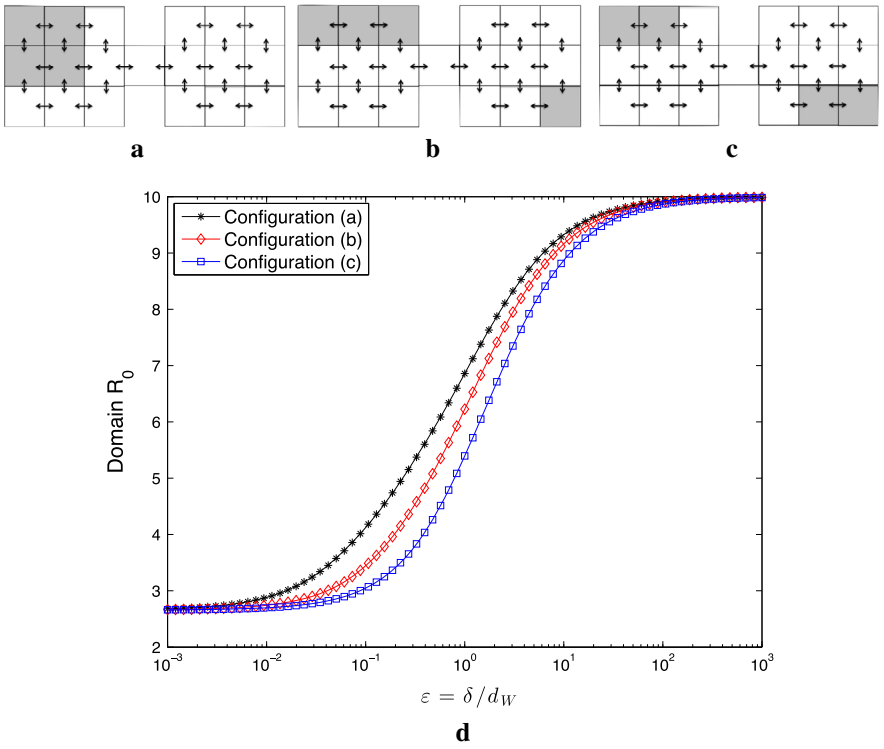

Fig. 7 Comparison of domain $\mathcal{R}_0$ for different hot spot arrangements, shown in **a**–**c**. Hot spots (*shaded*) have patch reproduction values of $\mathcal{R}_0^{high} = 10$, non-hot spots have $\mathcal{R}_0^{low} = 0.7$. All patches have pathogen decay rate $\delta_i = 1/30$ day^{-1}, infectious period $1/\gamma_i = 3$ days, case fatality ratio = 0.01, life expectancy = 50 years. Patch transmissibility is adjusted to give the desired patch reproduction numbers for hot spots/non-hot spots. The domain reproduction number approaches $\hat{\mathcal{R}}_0 = 2.6579$ as $\varepsilon \to 0$

The rooted spanning trees inform how to average across the network in a manner that respects the network structure, resulting in a lowest order approximation to $\mathcal{R}_0$; see (24). As these spanning trees are oriented inward towards the root vertex, there is an important distinction between inflow and outflow in the community network setting considered here. Due to the inward edge orientation, vertices with large amounts of inflow relative to outflow tend to have large network risks, and thus contribute disproportionately to $\mathcal{R}_0$. This is illustrated in the river and star examples from Sect. 5. In the language of Kleinberg (1999), "hubs" and "authorities" thus contribute very differently to $\mathcal{R}_0$. The situation here is also different than previous work on sexually transmitted diseases and highly sexually active core groups (Hethcote and Yorke 1984). A common assumption in modeling core groups is to assume that high activity rates of the core translate into both high transmission (outflow) as well as high contraction (inflow) of the disease (e.g. discussed in Galvani and May 2005), resulting in a disproportionate increase in $\mathcal{R}_0$ due to the core group. This need not hold in the community network setting: a vertex may have both large inflow as well as large out-

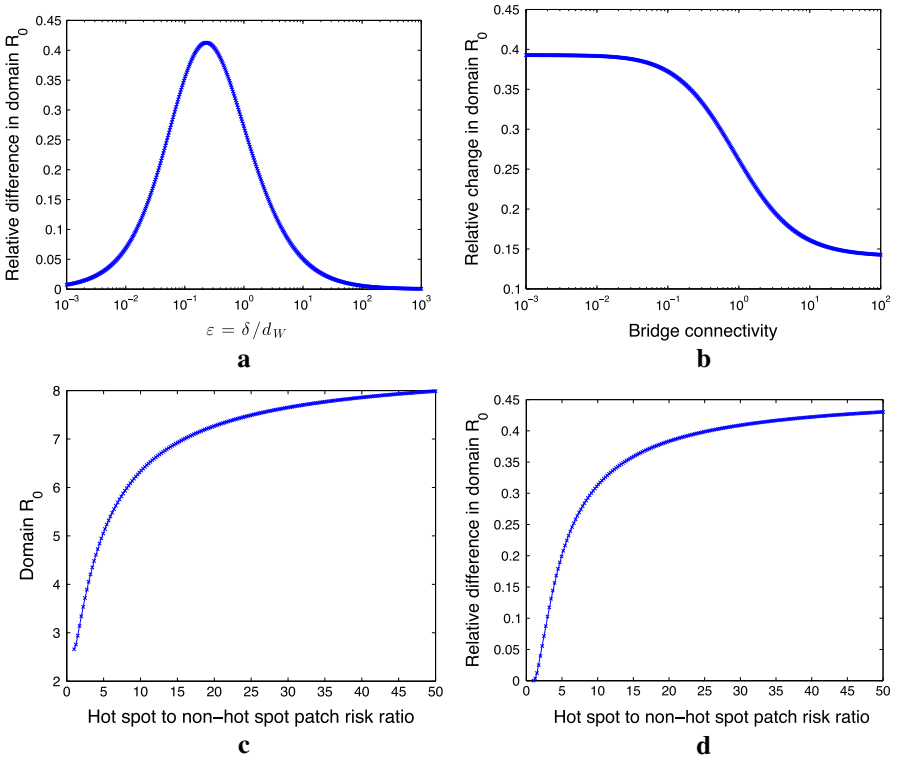

Fig. 8 **a** Relative difference in domain $\mathcal{R}_0$ values between configurations in Fig. 7a, c, as a function of ε. **b** Relative difference in $\mathcal{R}_0$ between configurations in Fig. 7a, c, as the flow to the bridge patch is varied. The parameter ε is fixed at 0.1. **c** Effect of increasing heterogeneity between hot spots and non-hot spots on $\mathcal{R}_0$. Configuration matches Fig. 7a. The ratio of patch risks $q_i r_i / (d_i + \alpha_i + \gamma_i)$ between hot spots and non-hot spots is varied (x-axis), while keeping $\hat{\mathcal{R}}_0$ fixed at 2.6579. The parameter ε is fixed at 0.1; other parameters correspond to those given in Fig. 7a. **d** Relative difference between domain $\mathcal{R}_0$ values for configurations in Fig. 7a, c. As the disparity in patch risk increases between hot spots and non-hot spots, the effect of clustering the hot spots together becomes more marked

flow, but if the network is balanced, then to lowest order all vertices contribute equally to $\mathcal{R}_0$ (e.g. consider Fig. 5 with $a = b$).

Higher order corrections to $\mathcal{R}_0$ involve the generalized group inverse X_0 of the Laplacian matrix, which can be interpreted in terms of deviations of expected residence times from the stationary distribution used in the preceding average; see (28). These deviations depend both upon the network structure as well as the pathogen decay rates, and can reveal groupings of the vertices into meta-communities. As the examples presented in Sect. 5 illustrate, the location of disease hot spots with respect to these meta-communities can be important for disease invasibility. Many different community detection algorithms exist for identifying substructure within a network (Newman 2010). The connection between $\mathcal{R}_0$ and X_0 indicates that when invasibility is a primary concern, an algorithm based on X_0 is a natural approach to take. As X_0 depends both upon the network structure as well as the pathogen decay rates δ_i, com-

munity detection algorithms based on X_0 will combine both structural and dynamic features of the network.

One of the insights from our study of the generalized group inverse in relation to $\mathcal{R}_0$ is that clustering hot spots together facilitates disease invasion in community networks. It is interesting to compare this with the extensive literature on disease spread in networks of individuals. For example, Keeling (1999) shows that increasing the clustering coefficient of a network (roughly speaking the likelihood that two of your friends are themselves friends; See, for example, Newman and Barabasi 2006; Newman et al. 2001; Watts and Strogatz 1998) decreases the ability of a disease to invade, due to spatial correlations leading to smaller numbers of susceptibles encountered by infectious individuals. Considering networks at different levels of resolution may thus provide different insights into disease dynamics.

The modeling framework considered here is reasonably flexible: apart from being strongly connected, no restrictions are placed on the network, and minimal assumptions are made on the disease incidence and shedding functions. Our original motivation was for waterborne diseases such as cholera, for which data on water networks is important. Water network data can be used to compute the network risk associated with different communities. The results presented here suggest that for disease control efforts, particular attention should be paid to communities with high network risk. In some settings, these networks may not be strongly connected (for example, in rivers with strong currents), while in others, flow in both directions between nodes may be important (e.g. wetlands, tidal rivers). Tracer studies may be useful for assessing the degree of flow between locations (Variano et al. 2009). On the other hand, the rooted spanning trees and generalized group inverse are fundamental features of networks, and we expect our results on how these quantities combine with patch characteristics to affect invasibility will extend to other settings besides waterborne disease. For example, DeVille and Peskin (2012) find that "in-hubs" play a much larger role than "out-hubs" in driving bursts in stochastic neural networks, analogous to our results for the hub and periphery example considered in Sect. 5.

Acknowledgments JHT and MCE acknowledge support from the National Science Foundation (OCE-1115881) and the Mathematical Biosciences Institute (DMS-0931642). The research of PvdD is partially supported through a Discovery Grant from the Natural Science and Engineering Research Council of Canada (NSERC). ZS acknowledges support from the University of Central Florida through a start-up grant. The authors are grateful to the anonymous reviewers for their thoughtful, constructive comments. This paper was improved by discussions at a Research in Teams meeting (13rit168) held at the Banff International Research Station.

Appendix A: Graph-theoretic background

For the convenience of the reader, we present some definitions and standard results from graph theory used throughout the paper. Further information can be found in graph theory textbooks such as West (2001). We include also a statement of the matrix tree theorem for weighted, directed graphs. A proof can be found in Moon (1970).

A *directed graph* (*digraph*) $\mathcal{G} = (V, E)$ consists of a set $V = \{1, 2, \ldots, n\}$ of vertices and a set $E = E(\mathcal{G})$ of directed arcs (i, j) from vertex i to vertex j. A directed graph $\mathcal{G}$ is *weighted* if each arc (j, i) is assigned a positive weight m_{ij};

the weighted, directed graph is denoted as $(\mathcal{G}, M)$, with nonnegative weight matrix $M = [m_{ij}]$ and $m_{ij} > 0$ if and only if $(j, i) \in E(\mathcal{G})$. For example, the weight matrix M may correspond to the movement matrix in the community network, and thus the weighted directed graph $(\mathcal{G}, M)$ corresponds to the community network. A directed graph is *strongly connected* if for any ordered pair of vertices, there exists a directed path from one to the other. A weighted directed graph $(\mathcal{G}, M)$ is strongly connected if and only if the weight matrix M is irreducible (Berman and Plemmons 1979).

A *rooted spanning tree* (in-tree) $\mathcal{T}$ is a subgraph of $\mathcal{G}$ on n vertices such that $\mathcal{T}$ is connected with no cycles, and has a root vertex such that every directed path between a non-root vertex and the root is oriented towards the root. The *weight* of a rooted spanning tree is the product of all arcs in the rooted spanning tree. To illustrate, consider the star graph shown in Fig. 5, with arc weights a from periphery to center, and b from center to periphery. Figure 5b shows the single spanning in-tree rooted at the center. This tree possesses n arcs with weight a, giving a tree weight of a^n. Each peripheral vertex roots a single spanning in-tree, shown in Fig. 5c. Trees rooted at the periphery have $n - 1$ arcs with weight a and a single arc with weight b, giving a tree weight of $a^{n-1}b$.

Let $(\mathcal{G}, M)$ be a weighted, directed graph. The *Laplacian matrix* L of the graph is then defined as:

$$L = \begin{pmatrix} \Sigma_{j\neq 1} m_{j1} & -m_{12} & \dots & -m_{1n} \\ -m_{21} & \Sigma_{j\neq 2} m_{j2} & \dots & -m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -m_{n1} & -m_{n2} & \dots & \Sigma_{j\neq n} m_{jn} \end{pmatrix}. \tag{37}$$

Note that each of the columns of L sum to zero, implying both that L is singular, and that the cofactors of L are the same within each column. The matrix tree theorem states that any cofactor corresponding to column k of L is equal to the sum of the weights of the spanning in-trees rooted at vertex k. For a proof see Moon (Section 5.5 and Theorem 5.5, Moon 1970).

Theorem 2 (Matrix Tree Theorem)
Let $(\mathcal{G}, M)$ be a weighted, directed graph, and let L be the Laplacian matrix of $(\mathcal{G}, M)$. Let c_{kk} denote the (k, k) cofactor of L. Then the cofactors of L are related to the rooted spanning trees of $\mathcal{G}$ by the following:

$$c_{kk} = \sum_{\mathcal{T}\in\mathbb{T}_k} \prod_{(j,i)\in E(\mathcal{T})} m_{ij}, \tag{38}$$

where $\mathbb{T}_k$ is the set of all spanning in-trees rooted at vertex k, $E(\mathcal{T})$ is the arc set of rooted spanning in-tree $\mathcal{T}$, and m_{ij} the weight of the arc from j to i.

A direct consequence of the matrix tree theorem is a relationship between $\ker L$ and the rooted spanning trees of $\mathcal{G}$. This is pointed out, for example, in Lemma 2.1 of Guo et al. (2006). Consider the vector of cofactors $c = (c_{11}, \dots, c_{nn})^T$. Then $(Lc)_i = \det L = 0$ for $i = 1, \dots, n$, so c belongs to $\ker L$. The dimension of

the nullspace of L is equal to the number of connected components of $\mathcal{G}$, so for strongly connected digraphs, L has a one dimensional nullspace spanned by c. For strongly connected digraphs, there is at least one in-tree rooted at each vertex, and thus the entries of c are all positive for digraphs with non-negative arc weights. Let $u_i = c_{ii}/\sum_{j=1}^{n} c_{jj}$, and let $u = (u_1, \ldots, u_n)^T$. Then u provides a basis vector for $\ker L$, with all positive entries and normalized so that $\sum_{i=1}^{n} u_i = 1$.

A weighted directed graph $(\mathcal{G}, M)$ is *balanced* if the net inflow is equal to the net outflow at each vertex, i.e. for each i, $\sum_{j\neq i} m_{ij} = \sum_{j\neq i} m_{ji}$. For balanced community networks, both the row sums and column sums of L are equal to zero, giving that all cofactors of L are equal. This in turn implies that all entries of u are equal to $1/n$ for balanced networks.

Appendix B: X_0 and the group inverse

In the equal pathogen decay case, $G = (L + \varepsilon Id)$. The resolvent expansion for matrices of this form is considered by Rothblum (1981), who shows that the zeroth order term in the series is given by the Drazin inverse of L. For matrices of index 1, the Drazin inverse is equal to the group inverse (e.g. Ben-Israel and Greville 2003, Chapter 4). As pointed out by Schweitzer and Stewart (1993), X_0 in the general case is a generalization of the Drazin (here, group) inverse. It is interesting to consider the relationship between X_0 in Proposition 2 and the group inverse in more detail.

The group inverse of a square matrix M of index 1 is the unique matrix $M^{\#}$ that satisfies the following three conditions (Ben-Israel and Greville 2003, Section 4.4):

$$MM^{\#}M = M, \tag{39}$$

$$M^{\#}MM^{\#} = M^{\#}, \tag{40}$$

$$MM^{\#} = M^{\#}M. \tag{41}$$

The matrix X_0 in Proposition 2 satisfies conditions (39) and (40) with $M = L$, but in general not (41). In the language of generalized inverses, X_0 is said to be a "{1, 2}-inverse" of L (Ben-Israel and Greville 2003).

We first show that X_0 acts as a right inverse of L on range L.

Lemma 5 *Let X_0 be given as in Proposition 2. Then*

$$LX_0x = \begin{cases} x, & x \in \text{range } L \\ 0, & x \in R_{1,0}. \end{cases}$$

Proof Let $x \in \text{range } L$. As $L : N_{1,0} \to \text{range } L$ a bijection (Theorem 1(c)), there exists $y \in N_{1,0}$ such that $Ly = x$. Then

$$\begin{aligned} LX_0x &= LX_0Ly \\ &= Ly \\ &= x. \end{aligned} \tag{42}$$

Finally, note that X_0 sends $R_{1,0}$ to 0 (Theorem 1(e)), and thus $LX_0x = 0$ for all $x \in R_{1,0}$. □

We can now show that X_0 is a $\{1, 2\}$-inverse of L, and determine where X_0 and the group inverse coincide.

Proposition 3 *Let X_0 be given as in Proposition 2. Then X_0 is a $\{1, 2\}$-inverse of L (i.e. satisfies conditions (39) and (40)), and commutes with L on $N_{1,0} \cap$ range L.*

Proof For any $x \in \mathbb{R}^n$, take $x = y + z$, where $y \in N_{1,0}$ and $z \in \ker L$. From Proposition 2,

$$\begin{aligned} X_0Lx &= X_0L(y+z) \\ &= X_0Ly \\ &= y. \end{aligned} \tag{43}$$

Multiplying on the left by L then gives

$$\begin{aligned} LX_0Lx &= Ly \\ &= L(y+z) \\ &= Lx, \end{aligned} \tag{44}$$

and thus X_0 satisfies (39).

Similarly, for any $x \in \mathbb{R}^n$ take $x = \tilde{y} + \tilde{z}$, where $\tilde{y} \in$ range L and $\tilde{z} \in R_{1,0}$. Using Lemma 5,

$$\begin{aligned} X_0LX_0x &= X_0LX_0(\tilde{y}+\tilde{z}) \\ &= X_0\tilde{y} \\ &= X_0(\tilde{y}+\tilde{z}) \\ &= X_0x, \end{aligned} \tag{45}$$

so X_0 satisfies (40).

Finally, consider $x \in N_{1,0} \cap$ range L. As $x \in$ range L, Lemma 5 gives $x = LX_0x$. But $x \in N_{1,0}$ as well, so property (e) in Theorem 1 gives $x = X_0Lx$, and thus X_0 and L commute on $N_{1,0} \cap$ range L. □

Proposition 3 shows that X_0 and the group inverse coincide on $N_{1,0} \cap$ range L. Of course, this intersection may be trivial. In the case where $D = Id$ (i.e. equal pathogen decay rates), $N_{1,0} =$ range L, and $X_0 = (Id - X_{-1}D)(L - X_{-1}D)^{-1}$ is equal to the group inverse of L.

References

Avrachenkov KE, Haviv M, Howlett PG (2001) Inversion of analytic matrix functions that are singular at the origin. SIAM J Matrix Anal Appl 22(4):1175–1189

Ben-Israel A, Greville TNE (2003) Generalized inverses: theory and applications, 2nd edn. Springer, New York
Berman A, Plemmons RJ (1979) Nonnegative matrices in the mathematical sciences. Academic Press, New York
Bertuzzo E, Casagrandi R, Gatto M, Rodriguez-Iturbe I, Rinaldo A (2010) On spatially explicit models of cholera epidemics. J R Soc Interface 7(43):321–333
Chao DL, Halloran ME, Longini IM Jr (2011) Vaccination strategies for epidemic cholera in Haiti with implications for the developing world. Proc Natl Acad Sci USA 108(17):7081–7085
Codeço C (2001) Endemic and epidemic dynamics of cholera: the role of the aquatic reservoir. BMC Infect Dis 1:1
DeVille REL, Peskin CS (2012) Synchrony and asynchrony for neuronal dynamics defined on complex networks. Bull Math Biol 74(4):769–802
Diekmann O, Heesterbeek JAP (2000) Mathematical epidemiology of infectious diseases: model building, analysis, and interpretation. Wiley, New York
Dowell SF, Braden CR (2011) Implications of the introduction of cholera to Haiti. Emerg Infect Dis 17(7):1299–1300
Eisenberg MC, Shuai Z, Tien JH, van den Driessche P (2013) A cholera model in a patchy environment with water and human movement. Math Biosci 246(1):105–112
Galvani AP, May RM (2005) Dimensions of superspreading. Nature 438:293–295
Guo H, Li MY, Shuai Z (2006) Global stability of the endemic equilibrium of multigroup SIR epidemic models. Can Appl Math Q 14:259–284
Hethcote HW, Yorke JA (1984) Gonorrhea transmission dynamics and control, Lecture Notes in Biomathematics, vol 56. Springer, New York
Horn RA, Johnson CR (1985) Matrix analysis. Cambridge University Press, Cambridge
Kaper JB, Morris JG Jr, Levine MM (1995) Cholera. Clin Microbiol Rev 8(1):48–86
Keeling MJ (1999) The effect of local spatial structure on epidemiological invasions. Proc R Soc Lond B 266:859–867
Kleinberg JM (1999) Authoritative sources in a hyperlinked environment. J Assoc Comput Mach 46:604–632
Langenhop CE (1971) The Laurent expansion for a nearly singular matrix. Linear Algebra Appl 4:329–340
Meyer CDJ (1975) The role of the group generalized inverse in the theory of finite Markov chains. SIAM Rev 17(3):443–464
Moon JW (1970) Counting labelled trees. William Clowes and Sons Limited, London
Newman M (2010) Networks: an introduction. Oxford University Press, Oxford
Newman M, Barabasi A-L, Watts DJ (2006) The structure and dynamics of networks. Princeton University Press, Princeton
Newman MEJ, Strogatz SH, Watts DJ (2001) Random graphs with arbitrary degree distributions and their applications. Phys Rev E 64:026118
Piarroux R, Barrais R, Faucher B, Haus R, Piarroux M, Gaudart J, Magloire R, Raoult D (2011) Understanding the cholera epidemic, Haiti. Emerg Infect Dis 17(7):1161–1168
Rothblum UG (1981) Resolvent expansions of matrices and applications. Linear Algebra Appl 38:33–49
Schweitzer P, Stewart GW (1993) The Laurent expansion of pencils that are singular at the origin. Linear Algebra Appl 183:237–254
van den Driessche P, Watmough J (2002) Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math Biosci 180:29–48
Variano EA, Ho DT, Engel VC, Schmeider PJ, Reid MC (2009) Flow and mixing dynamics in a patterned wetland: kilometer-scale tracer releases in the Everglades. Water Resour Res 45:W08422
Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. Nature 393(6684):440–442
West DB (2001) Introduction to graph theory, 2nd edn. Prentice Hall, Upper Saddle River