Bulletin of Mathematical Biology (2024) 86:41
https://doi.org/10.1007/s11538-024-01271-7

Society for Mathematical Biology

ORIGINAL ARTICLE

R_0 May Not Tell Us Everything: Transient Disease Dynamics of Some SIR Models Over Patchy Environments

Ao Li[1] · **Xingfu Zou**[1]

Received: 9 June 2023 / Accepted: 16 February 2024

Abstract
This paper examines the short-term or transient dynamics of SIR infectious disease models in patch environments. We employ reactivity of an equilibrium and amplification rates, concepts from ecology, to analyze how dispersals/travels between patches, spatial heterogeneity, and other disease-related parameters impact short-term dynamics. Our findings reveal that in certain scenarios, due to the impact of spatial heterogeneity and the dispersals, the short-term disease dynamics over a patch environment may *disagree with* the long-term disease dynamics that is typically reflected by the basic reproduction number. Such an inconsistence can mislead the public, public healthy agencies and governments when making public health policy and decisions, and hence, these findings are of practical importance.

Keywords SIR model · Patches · Dispersion · Transient dynamics · Disease amplification rate

Mathematics Subject Classification 34K20 · 92B05 · 92D25

1 Introduction

Mathematical modelling is an important and efficient tool for understanding the transmission dynamics of infectious diseases. It helps us unravel disease spread mechanisms, forecast the future course of an outbreak, and evaluate public health interventions. With respect to these topics, there are two crucial questions: (Q1) long-

Research partially supported by NSERC of Canada (No. RGPIN-2022-04744).

✉ Xingfu Zou
xzou@uwo.ca

Ao Li
ali297@uwo.ca

[1] Department of Mathematics, University of Western Ontario, London, ON N6A 5B7, Canada

Published online: 15 March 2024

term disease dynamics—will the disease eventually die out or become endemic? (Q2) short-term disease dynamics - is the epidemic escalating or being mitigated at the present time?

To better explain these two questions and our motivations for this paper, let us first look at the simplest Kermack–McKendrick model (1927):

$$\begin{cases} S'(t) = -\beta SI, \\ I'(t) = \beta SI - \gamma I, \\ R'(t) = \gamma I. \end{cases} \tag{1}$$

In this model, the host population is divided into susceptible, infected and recovered classes denoted by $S(t)$, $I(t)$ and $R(t)$. Here, $\beta > 0$ denotes the transmission rate and $\gamma > 0$ is the recovery rate. For this model, the number $\mathscr{R}_0 = \beta S_0 \cdot (1/\gamma)$ is called the basic reproduction number. Noting that $1/\gamma$ is the average infection time, it is clear that $\mathscr{R}_0$ measures the average number of new infections that an infected individual causes when the susceptible population is S_0. The value of $\mathscr{R}_0$ determines whether or not there will be an outbreak for the disease when initially $I(0) > 0$ and $S_0 > 0$: if $\mathscr{R}_0 < 1$, there will be no outbreak as $I(t)$ monotonically decreases to zero; if $\mathscr{R}_0 > 1$, the disease experiences a single outbreak before dying out as $I(t)$ first increases up to a maximum value then decreases to zero. This dichotomy can also be obtained by looking at

$$I'(0) = \beta S(0)I(0) - \gamma I(0) = [\beta S_0 - \gamma]I(0) =: \Gamma_0 I(0)$$

which indicates that there will be no outbreak if $\Gamma_0 = \beta S_0 - \gamma < 0$ (i.e., $I'(0) < 0$); and there will be an outbreak if $\Gamma_0 > 0$ (i.e., $I'(0) > 0$). Note that $\Gamma_0 = \beta S_0 - \gamma$ is the *relative change rate* of the subpopulation $I(t)$ at the initial time $t = 0$, which measures the initial amplification rate of $I(t)$. Here we have two distinct notions: the basic reproduction number $\mathscr{R}_0$ which is of *long-time nature* (during the epidemic), and the initial amplification rate Γ_0 which is of *short-time nature* (near the initial time $t = 0$). The former is supposed to predict the long-time disease dynamics while the latter is expected to predict the short-time disease dynamics. However, they amazingly agree with each other in predicting the disease dynamics described by model (1); that is, the long-term and short-term disease dynamics in terms of outbreak coincide.

The SIR model (1) has numerous variants when different transmission mechanisms and demographies are taken into account, and same or similar threshold type results have been extended accordingly. For example, when considering a birth rate $B(N)$ where $N(t) = S(t) + I(t) + R(t)$ is the total population of the host, a per capita natural death rate $d > 0$, and a per capita disease-related death rate $\epsilon > 0$, then (1) is extended to

$$\begin{cases} S'(t) = B(N) - \beta SI - dS, \\ I'(t) = \beta SI - (\gamma + d + \epsilon)I, \\ R'(t) = \gamma I - dR. \end{cases} \tag{2}$$

Assume that the demographic equation $N'(t) = B(N) - dN$ has a unique positive equilibrium $N_+ > 0$ which is globally asymptotically stable, and before the disease appears, the host population has settled at (or is close to) N_+. This implies that (2) has a unique disease-free equilibrium $E_0 = (N_+, 0, 0)$ (as opposed to (1) for which there are infinitely many disease-free equilibria). It is known that (2) also has the *long-term* threshold dynamics in terms of its basic reproduction number $\hat{\mathscr{R}}_0 = \beta N_+/(\gamma+d+\epsilon)$: if $\hat{\mathscr{R}}_0 < 1$, then E_0 is globally asymptotically stable meaning that the disease will eventually die out if $\hat{\mathscr{R}}_0 > 1$, E_0 becomes unstable and there occurs an endemic equilibrium E^+ which is globally asymptotically stable. In the meantime, from

$$I'(0) = [\beta S(0) - (\gamma + d + \epsilon)]I(0) =: \hat{\Gamma}_0 I(0),$$

one knows that if $\hat{\Gamma}_0 < 0$, then there will be no outbreak at $t = 0$; if $\hat{\Gamma}_0 > 0$, then there will be an outbreak at $t = 0$. Note that generally $R(0) = 0$ and $I(0)$ is very small in reality when a disease appears, implying that $S(0) \approx N(0) = N_+$. Hence, $\hat{\Gamma}_0 = \beta S(0) - (\gamma + d + \epsilon) \approx \beta N_+ - (\gamma + d + \epsilon)$. This implies that $\hat{\mathscr{R}}_0 - 1$ and $\hat{\Gamma}_0$ have the same sign. Therefore, $\hat{\mathscr{R}}_0$ also determines whether or not an initial *small-scale infection* $I(0)$ will lead to an outbreak in the coming short period of time after $t = 0$, the same conclusion as for model (1).

The above results for model (2) hold only for the initial time $t = 0$ when an epidemic occurs. If, at some given time $t_0 > 0$ during an epidemic, one wants to predict whether or not there will be an outbreak within a short period, one would have to look at $I'(t_0) = \hat{\Gamma}(t_0)I(t_0)$ for (2), where $\hat{\Gamma}(t_0) = \beta S(t_0) - (\gamma + d + \epsilon)$. Unfortunately, $S(t_0)$ can now be far away from $S(0)$ (hence N_+), and hence, the sign of $\hat{\mathscr{R}}_0 - 1$ (independent of t_0) may not agree with the sign of $\hat{\Gamma}(t_0)$. Thus, the value of the long-term characteristics quantity $\hat{\mathscr{R}}_0$ generally cannot predict whether or not there will be an outbreak in the coming short period of time after t_0. That is, even if $\hat{\mathscr{R}}_0 < 1$ (hence eventually $I(t) \to 0$), there can be an outbreak at some $t_0 > 0$; and even if $\hat{\mathscr{R}}_0 > 1$, there may be some time $t_0 > 0$ such that $I'(t_0) < 0$ which can be very misleading. Therefore, for a general model of infectious disease dynamics, the long-term and short-term behaviours often do not imply each other, and they both deserve careful analysis.

For the study of *long-term* disease dynamics, a substantial amount of literature has been published on various models. Typically the long-term dynamics of a disease transmission model is of threshold type in terms of the basic reproduction number $\mathscr{R}_0$. For a model that has a unique disease-free equilibrium, $\mathscr{R}_0$ is still biologically defined as the average number of secondary infections caused by a single infected individual during his/her entire period of infectiousness *in a completely susceptible population*. Mathematically, $\mathscr{R}_0$ is defined as the spectral radius of the next generation operator, which was initially introduced by Diekmann et al. (1990). For compartmental models formulated as systems of ordinary differential equations (ODEs), van den Driessche and Watmough (2002) derived an expression for the next generation matrix. The authors further demonstrated the threshold *long-term* dynamics by showing that: if $\mathscr{R}_0 < 1$ then the disease-free equilibrium of the model is locally asymptotically

stable meaning that the disease eventually dies out; if $\mathscr{R}_0 > 1$ then the disease-free equilibrium is unstable and the disease becomes endemic.

In contrast, there are only very few works in literature (see Sect. 2) that analytically investigate *short-term or transient disease dynamics* by mathematical models. This is mainly due to the lack of effective tools and methods. For simple models like (1) and (2), the equation governing the change rate of the infected subpopulation is conveniently related to its amplification rate that has an explicit formula. When there is some heterogeneity, for example, *spatial heterogeneity* as we will discuss next, analyzing short-term or transient disease dynamics becomes much more difficult, if not impossible. On the other hand, short-term disease dynamics is very important because it may affect the judgement of not only the public, but also public health agents who typically advise governments of various levels in making decisions on interventions for controlling the epidemics of an infectious disease.

Nowadays the world is highly connected, and such high connectivity has obviously enhanced the spread of infectious diseases. The ongoing pandemic of COVID-19 is such an example. Thus, when modelling the transmission dynamics of an infectious disease, we need to consider spatial structure. Typically, patch models are used with each patch representing a country, a city, a province, or a geographic area in some given context. The population dynamics of each patch is coupled by spatial dispersals or travels reflecting the movement of the host population. Such couplings bring challenges to the analysis of the resulting models. Taking the Kermack–McKendrick SIR system (1) as an example and considering $n \geq 2$ patches, the coupled system corresponding to (1) is

$$\begin{cases} S_i'(t) = \displaystyle\sum_{j\in\Omega, j\neq i} d_{ij}^S S_j - \sum_{j\in\Omega, j\neq i} d_{ji}^S S_i - \beta_i S_i I_i, \\ I_i'(t) = \displaystyle\sum_{j\in\Omega, j\neq i} d_{ij}^I I_j - \sum_{j\in\Omega, j\neq i} d_{ji}^I I_i + \beta_i S_i I_i - \gamma_i I_i, \\ R_i'(t) = \displaystyle\sum_{j\in\Omega, j\neq i} d_{ij}^R R_j - \sum_{j\in\Omega, j\neq i} d_{ji}^R R_i + \gamma_i I_i, \end{cases} \quad \text{for } i \in \Omega. \tag{3}$$

Here $\Omega = \{1, \ldots, n\}$, and the parameters $\beta_i > 0$ and $\gamma_i > 0$ have the same meanings as in (1) but for patch i. The constant $d_{ij}^X \geq 0$ is the dispersal/travel rate of individuals in class X from patch j to patch i where $X \in \{S, I, R\}, i, j \in \{1, \ldots, n\}$, and $i \neq j$. Now due to the coupling, obtaining results similar to those for the non-spatial model (1) becomes very difficult, if not impossible. This is because (A) the computation of the basic reproduction number $\mathscr{R}_0$ is more challenging; and (B) there may be some time moments at which $I_i'(t)$, $i = 1, \ldots, n$, have different signs, and hence the measurement of an outbreak should consider all patches. To our best knowledge, there are only two studies (Mari et al. 2019, 2021) that have considered the transient dynamics of disease transmission *over a patchy environment*. For this type of disease models over patches, even when a model has a unique disease-free equilibrium at which the next generation method is applicable to establish the threshold long-term dynamics, it is, in general, *very difficult* (if not impossible) to obtain an explicit expression for $\mathscr{R}_0$ as the spectral radius of a large matrix. See, e.g., Arino and Van

den Driessche (2003); Arino and Van Den Driessche (2003); Wang and Zhao (2004); Arino and Van den Driessche (2006); Allen et al. (2007); Hsieh et al. (2007); Eisenberg et al. (2013); Almarashi and McCluskey (2019); Chen et al. (2020) and the references within.

This paper is stimulated by the aforementioned need for approaches to explore short-term dynamics of infectious diseases over connected patches. To this end, we *borrow/adopt* the notion of *reactivity* used in ecology and take advantage of the developed mathematical results for reactivity in mathematical ecology (e.g. Neubert and Caswell 1997; Mari et al. 2017; Wang et al. 2019; Lutscher and Wang 2020). By applying this idea to some patch models of disease transmissions, we wish to establish a framework and develop an approach that can be used for short/transient disease dynamics for more disease models with spatial structure.

The rest of the paper is organized as below. In Sect. 2, we provide the mathematical background for some related notions in mathematical ecology, including reactivity, amplification rate, and resilience. We then move on to apply these notions and ideas behind them to some disease models over patches to examine their short-term dynamics. This will allow us to explore how the spatial dispersals/travels, spatial heterogeneity and other model parameters as well as initial values affect the short-term disease dynamics at the initial time (or can be extended to a given time during an epidemic). Two types of patch models will be examined: Sect. 3 deals with models without demography, and Sect. 4 focuses on models with demographic structure. Some numerical simulations will also be exhibited in Sect. 5 to more visually demonstrate our results. Particular attention is paid to the scenarios when the short-term disease dynamics does agree with the long-term disease dynamics, because such scenarios may mislead the public and health policymakers. We end the paper with Sect. 6 in which we summarize our main conclusions and present some discussions.

2 Amplification Rates and Reactivity

The notion of reactivity in ecology was first introduced by Neubert and Caswell (1997), as a description of the short-term response to perturbations. Specifically, it is defined as the maximum initial amplification rate over all possible perturbations to an equilibrium. An equilibrium with positive reactivity is said to be reactive, corresponding to the case when some perturbations can grow initially.

Consider the initial value problem of a linear system of ODEs:

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x}, \quad \mathbf{x}(0) = \mathbf{x}_0 \tag{4}$$

where $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{A} = \left[a_{ij}\right]_{n\times n}$ is a real matrix. Equation (4) can be the linearization of an ODE system for the population dynamics of n interacting species at an equilibrium, with $\mathbf{x}$ being the deviation from the equilibrium. Thus, the Euclidean norm of $\mathbf{x}(t)$, i.e.,

$$\|\mathbf{x}(t)\| := \sqrt{x_1^2(t) + x_2^2(t) + \cdots + x_n^2(t)},$$

measures the size of vector $\mathbf{x}(t)$, and it also measures the distance of $\mathbf{x}(t)$ to the origin, or equivalently measures how far away the population vector is from the equilibrium at time t. Denote by $\Gamma(t)$ the relative rate of change of $\|\mathbf{x}(t)\|$, that is,

$$\Gamma(t) := \frac{1}{\|\mathbf{x}\|}\frac{d\|\mathbf{x}\|}{dt}.$$

Obviously, $\Gamma(t)$ measures the amplification rate for $\|\mathbf{x}(t)\|$ at time t. Particularly, $\Gamma_0 := \Gamma(0)$ is called the initial amplification rate. If $\mathbf{x}(t)$ is a solution to (4), then direct calculation gives (see Neubert and Caswell 1997)

$$\Gamma(t) = \frac{\mathbf{x}^T(t)H(\mathbf{A})\mathbf{x}(t)}{\|\mathbf{x}(t)\|^2} \quad \text{where} \quad H(\mathbf{A}) = \frac{\mathbf{A}+\mathbf{A}^T}{2}.$$

By the definition, $\Gamma(t) > 0$ ($\Gamma(t) < 0$) means that the size of the solution $\|\mathbf{x}(t)\|$ to (4) is growing (decaying) at t. Particularly, the sign of the initial amplification rate,

$$\Gamma_0 = \frac{\mathbf{x_0}^T H(\mathbf{A})\mathbf{x_0}}{\mathbf{x_0}^T\mathbf{x_0}}, \tag{5}$$

predicts whether the solution with $\mathbf{x}_0$ will initially grow or decay. Note that $\Gamma_0 = \Gamma(\mathbf{x}_0)$ depends on the initial value $\mathbf{x}_0$ (so does $\Gamma(t)$).

Let $\lambda_1(\mathbf{A})$ denote the spectral bound of $\mathbf{A}$, i.e., the eigenvalue that has the largest real part. Then $-\lambda_1(\mathbf{A})$ is called the *resilience* of (4), which is independent of the initial value $\mathbf{x}_0$ and reflects the long-term dynamics of (4). On the other hand, $H(\mathbf{A})$ is a real symmetric matrix and hence all its eigenvalues are real. Let $\lambda_{\min}$ and $\lambda_{\max}$ denote the smallest and largest eigenvalues of $H(\mathbf{A})$ respectively. Since $\Gamma_0(\mathbf{x}_0)$ given by (5) is in the form called the Rayleigh quotient or the Rayleigh-Ritz ratio, it is known (see, e.g., Horn and Johnson 1985) that

$$\lambda_{\min} \le \Gamma_0(\mathbf{x}_0) = \frac{\mathbf{x}_0^T H(\mathbf{A})\mathbf{x}_0}{\mathbf{x}_0^T\mathbf{x}_0} \le \sup_{\mathbf{x}_0\neq 0}\frac{\mathbf{x}_0^T H(\mathbf{A})\mathbf{x}_0}{\mathbf{x}_0^T\mathbf{x}_0} = \max_{\|\mathbf{x}_0\|=1}\mathbf{x}_0^T H(\mathbf{A})\mathbf{x}_0 = \lambda_{\max}. \tag{6}$$

The reactivity of (4) is defined in Neubert and Caswell (1997) as the largest initial amplification rate over all initial values, that is,

$$\text{reactivity} = \sup_{\mathbf{x}_0\neq 0}\left(\frac{1}{\|\mathbf{x}\|}\frac{d\|\mathbf{x}\|}{dt}\bigg|_{t=0}\right) = \sup_{\mathbf{x}_0\neq 0}\frac{\mathbf{x}_0^T H(\mathbf{A})\mathbf{x}_0}{\mathbf{x}_0^T\mathbf{x}_0} = \max_{\|\mathbf{x}_0\|=1}\mathbf{x}_0^T H(\mathbf{A})\mathbf{x}_0 = \lambda_{\max}.$$

Apparently, the reactivity measures the maximal possible initial growth for (4) which is of short-term nature. Moreover,

- if the reactivity $\lambda_{\max}$ of (4) is negative, then for any initial value $\mathbf{x}_0$, the solution will initially decay in size (norm) since $\Gamma_0 = \Gamma(\mathbf{x}_0) < 0$;
- if $\lambda_{\min} > 0$, then for any initial value $\mathbf{x}_0$, the solution will initially grow in size since $\Gamma_0 = \Gamma(\mathbf{x}_0) > 0$;

- if $\lambda_{\min} < 0 < \lambda_{\max}$, then there will be initial values $\mathbf{x}_0$ for which $\Gamma_0 = \Gamma(\mathbf{x}_0) < 0$ and there will also be $\mathbf{x}_0$ for which $\Gamma_0 = \Gamma(\mathbf{x}_0) > 0$.

We remark that there have been some extensions/generalizations of the above notions of reactivity and amplification rates in ecology. For example, Mari et al. (2017) generalized *reactivity* to $\lambda_{\max}(H(\mathbf{C}^T\mathbf{C}\mathbf{A}))$ corresponding to a system output $\mathbf{y} = \mathbf{C}\mathbf{x}$ where matrix $\mathbf{C}$ reflects the interest in a subset of state variables. Wang et al. (2019) recently extended the measurements of reactivity and amplification rates to some reaction-diffusion models to explore how spatial heterogeneity affects the transient dynamics. In a more recent work, Lutscher and Wang (2020) explored the reactivity of *periodic orbits*. The results reveal some differences between the reactivity of a stable equilibrium and that of a stable periodic orbit.

In the subsequent sections, we will investigate short-term disease dynamics of the SIR model (3) without demography (Sect. 3), as well as its corresponding version with a simple demographic structure (Sect. 4). Combining the existing theory for ecology reviewed above and with some theoretical analysis, we will present explicit formulas for the initial amplification rate of the epidemics. These formulas clearly show how the transient disease dynamics is impacted by the spatial heterogeneity and the dispersals.

We point out that in the *context of disease dynamics models* (also population models with predator–prey type interactions), reactivity has been used in Hosack et al. (2008); Woodall et al. (2014); Mari et al. (2018, 2019); O'Regan et al. (2020); Mari et al. (2021) to measure the initial growth rate of infected populations for some infectious disease models. Among these works, only (Mari et al. 2019, 2021) deal with short-term disease dynamics with *discrete spatial variations*, which is the focus of this paper. However, the patch models in Mari et al. (2019, 2021) are of Lagrangian type. For such Lagrangian type patch models, mobility is implicitly presented by some parameters representing probabilities p_{ij} that a individual of home patch i will be in patch j. Moreover, the model in Mari et al. (2019) is for a water-borne disease and which assumes no host-to-host transmission and new infections can only occur from the contaminated water. Mari et al. (2021) models the transmission dynamics of COVID-19 which is a contagious disease, but the interactions between the patches are reflected by very complicated nonlocal infection force terms that have the above mentioned probabilities built in. Theoretical analysis of such model system are very challenging and hence, both (Mari et al. 2019, 2021) mainly explore the models numerically. In contrast, our models are of Eulerian type with explicit dispersal/travel rate incorporated, for which, we are able to obtain some results that can explicitly reveal the impact of the spatially heterogeneous parameters and the travel/dispersal rates on the initial amplification rate of the epidemic.

3 SIR Epidemic Patch Model

3.1 The Case of Two Patches

We start with special case $n = 2$ of (3), that is, the following two-patch SIR model without demography:

$$\begin{cases} S_1'(t) = d_{12}^S S_2 - d_{21}^S S_1 - \beta_1 S_1 I_1, \\ S_2'(t) = d_{21}^S S_1 - d_{12}^S S_2 - \beta_2 S_2 I_2, \\ I_1'(t) = d_{12}^I I_2 - d_{21}^I I_1 + \beta_1 S_1 I_1 - \gamma_1 I_1, \\ I_2'(t) = d_{21}^I I_1 - d_{12}^I I_2 + \beta_2 S_2 I_2 - \gamma_2 I_2, \\ R_1'(t) = d_{12}^R R_2 - d_{21}^R R_1 + \gamma_1 I_1, \\ R_2'(t) = d_{21}^R R_1 - d_{12}^R R_2 + \gamma_2 I_2. \end{cases} \tag{7}$$

It is easy to verify that the total population in the two patches is a constant since there are no births or natural deaths, neither there are disease caused deaths. By Proposition 1.1 in Chepyzhov and Vishik (2002), all solutions to model (7) with non-negative initial conditions remain non-negative for all $t > 0$. Let $M(t) = S_1(t) + S_2(t) + I_1(t) + I_2(t)$. Then,

$$M'(t) = -\gamma_1 I_1 - \gamma_2 I_2 \leq 0,$$

and hence, $M(t)$ is a decreasing function. In addition, $M(t)$ is non-negative, leading to the conclusion that its limit, $\lim_{t\to\infty} M(t)$, exists and thus, $\lim_{t\to\infty} M'(t) = 0$. This together with the non-negativity of $I_i(t)$ for $i \in \{1, 2\}$ implies that

$$I_1(t) \to 0 \quad \text{and} \quad I_2(t) \to 0 \quad \text{as} \quad t \to \infty.$$

Thus, the long disease dynamics is known—the disease will eventually die out in both patches, and hence, we are just concerned about whether or not there will be a short-term outbreak.

In the rest of this paper, we will use the amplification rate $\Gamma(0)$ for $x(t) = [I_1(t), I_2(t)]$ to determine whether or not there will be a short-term outbreak. To this end, we look at the linearization of (7) at the disease-free *state*

$$[S_1(0),\ S_2(0),\ I_1(0),\ I_2(0),\ R_1(0),\ R_2(0)] = [S_{10},\ S_{20},\ 0,\ 0,\ 0,\ 0] \in \mathbb{R}_+^6 \tag{8}$$

where S_{10} and S_{20} are the respective susceptible populations in the two patches before the disease was brought into the patches. We focus on the changes of the infected compartments $x(t) = [I_1(t) I_2(t)]$ when they have a perturbation $x(0) = [I_{10}, I_{20}]$. To this end, we focus on the linearization of (7) at the above disease-free *state* to obtain the linear system governing $x(t)$ near $x(0)$:

$$\frac{d\mathbf{x(t)}}{dt} = \mathbf{A}_0\mathbf{x}(t) \quad \text{with} \quad \mathbf{A}_0 = \begin{pmatrix} \beta_1 S_{10} - \gamma_1 - d_{21}^I & d_{12}^I \\ d_{21}^I & \beta_2 S_{20} - \gamma_2 - d_{12}^I \end{pmatrix}.$$

From (5) and some straightforward calculations with $\mathbf{A} = A_0$ and $\mathbf{x_0} = (I_{10}, I_{20})$, we obtain

$$\Gamma_0 = \frac{(\beta_1 S_{10} - \gamma_1 - d_{21}^I) I_{10}^2 + (\beta_2 S_{20} - \gamma_2 - d_{12}^I) I_{20}^2 + (d_{12}^I + d_{21}^I) I_{10} I_{20}}{I_{10}^2 + I_{20}^2}. \tag{9}$$

With the above formula, we can discuss the impact of the parameters and the initial values on the initial amplification rate. Firstly we observe that Γ_0 is linearly dependent on parameters d_{12}^I, d_{21}^I, β_1, β_2, γ_1, and γ_2. Moreover, Γ_0 is increasing with respect to β_1 and β_2 since

$$\frac{\partial \Gamma_0}{\partial \beta_1} = \frac{S_{10} I_{10}^2}{I_{10}^2 + I_{20}^2} > 0 \quad \text{and} \quad \frac{\partial \Gamma_0}{\partial \beta_2} = \frac{S_{20} I_{20}^2}{I_{10}^2 + I_{20}^2} > 0,$$

and it is decreasing with respect to γ_1 and γ_2 since

$$\frac{\partial \Gamma_0}{\partial \gamma_1} = -\frac{I_{10}^2}{I_{10}^2 + I_{20}^2} < 0 \quad \text{and} \quad \frac{\partial \Gamma_0}{\partial \gamma_2} = -\frac{I_{20}^2}{I_{10}^2 + I_{20}^2} < 0.$$

Suppose that some intervention measures are implemented to control the spread of an infectious disease and therefore to mitigate the outbreak. Wearing masks in public areas and practicing social distance, for instance, may result in a *lower transmission rates* (smaller β_1 and β_2) and thus a smaller initial amplification rate, helping prevent short-term outbreaks. Vaccination, as the most effective method of preventing infectious diseases, can be considered as a means of reducing the susceptible populations, yielding a smaller Γ_0 since

$$\frac{\partial \Gamma_0}{\partial S_{10}} = \frac{\beta_1 I_{10}^2}{I_{10}^2 + I_{20}^2} > 0 \quad \text{and} \quad \frac{\partial \Gamma_0}{\partial S_{20}} = \frac{\beta_2 I_{20}^2}{I_{10}^2 + I_{20}^2} > 0.$$

Note that $\Gamma_0^{(i)} = \beta_i S_{i0} - \gamma_i$ is the initial amplification rate of patch $i \in \{1, 2\}$ in isolation (i.e., $d_{ij} = 0$, meaning no dispersal/travel are allowed). Then Γ_0 given by (9) can be rewritten as

$$\begin{aligned} \Gamma_0 &= \frac{\Gamma_0^{(1)} I_{10}^2 + \Gamma_0^{(2)} I_{20}^2 + (I_{10} - I_{20})(d_{12}^I I_{20} - d_{21}^I I_{10})}{I_{10}^2 + I_{20}^2} \\ &= \left[\frac{I_{10}^2}{I_{10}^2 + I_{20}^2} \Gamma_0^{(1)} + \frac{I_{20}^2}{I_{10}^2 + I_{20}^2} \Gamma_0^{(2)} \right] + \frac{(I_{10} - I_{20})(d_{12}^I I_{20} - d_{21}^I I_{10})}{I_{10}^2 + I_{20}^2} \end{aligned} \tag{10}$$

Note that the first term on the right side of (10) is the average initial amplification rate over the two patches, and the second term thus *accounts for the deviation from that average caused by the dispersals*. When $I_{10} = I_{20}$, or $d_{12}^I I_{20} = d_{21}^I I_{10}$ (i.e., the net movement of infectives is balanced), Γ_0 is nothing but just the average and

$$\min_{i \in \{1,2\}} \Gamma_0^{(i)} \le \Gamma_0 \le \max_{i \in \{1,2\}} \Gamma_0^{(i)}.$$

If $I_{10} > I_{20}$, the second term in (10) shows that allowing an opposite *net flow* of the infected population (i.e., smaller $d_{12}^I I_{20}$ and larger $d_{21}^I I_{10}$) would help reduce Γ_0, making it possible for Γ_0 to be larger or smaller than any of the two individual initial

amplification rates $\Gamma_0^{(j)}$, $j = 1, 2$. Also if $I_{10} > I_{20}$, then Γ_0 is decreasing in d_{21}^I and increasing in d_{12}^I; if $I_{10} < I_{20}$, then Γ_0 is increasing in d_{21}^I and decreasing in d_{12}^I.

3.2 Estimate Γ_0 in Early Stage

In practice, the initial data is not always available. In the early stage of an epidemic, particularly for a newly emerging infectious disease, it takes time to diagnose or identify patients and conduct large-scale tests. Let N_{10} and N_{20} denote the total populations at $t = 0$ in the two patches. Since the number of infection cases is typically very small during the initial phase, we can approximate the initial disease free state $[S_{10}, S_{20}, 0, 0]$ in (8) by $[N_{10}, N_{20}, 0, 0]$. That is, $\Gamma_0^{(i)} = \beta_i S_{i0} - \gamma_i \approx \beta_i N_{i0} - \gamma_i$, $i = 1, 2$.

According to (6), we can predict the best and worst situations that the epidemic may develop by the upper and lower bounds of Γ_0 which are the largest and smallest eigenvalues of $H(\mathbf{A_0})$,

$$\begin{aligned} \lambda_{\max} &= \frac{1}{2}\left[m_1 + m_2 + \sqrt{(m_1+m_2)^2 + D^2 - 4m_1m_2}\right], \\ \lambda_{\min} &= \frac{1}{2}\left[m_1 + m_2 - \sqrt{(m_1+m_2)^2 + D^2 - 4m_1m_2}\right], \end{aligned} \tag{11}$$

where

$$m_1 = \Gamma_0^{(1)} - d_{21}^I, \quad m_2 = \Gamma_0^{(2)} - d_{12}^I, \quad \text{and} \quad D = d_{12}^I + d_{21}^I$$

From (11) we immediately obtain the following results:

(S1) when $4m_1m_2 - D^2 > 0$, then

(S1-a) if $m_1 + m_2 > 0$, then $\lambda_{\min} > 0$ and hence, $\Gamma_0 > 0$ for all $\mathbf{x_0} > 0$;
(S1-b) if $m_1 + m_2 < 0$, then $\lambda_{\max} < 0$ and hence, $\Gamma_0 < 0$ for all $\mathbf{x_0} > 0$;

(S2) when $4m_1m_2 - D^2 < 0$, then $\lambda_{\min} < 0 < \lambda_{\max}$ and hence, the sign of Γ_0 depends on initial conditions.

Moreover, if $\Gamma_0^{(i)} > 2\max\{d_{12}^I, d_{21}^I\}$ for $i \in \{1, 2\}$, then conditions for (S1-a) are satisfied and hence, $\|\mathbf{x}(t)\|$ will always grow at the initial time, meaning there will be an initial outbreak; if $\Gamma_0^{(i)} < \min\{d_{12}^I, d_{21}^I\} - \max\{d_{12}^I, d_{21}^I\}$ for $i \in \{1, 2\}$, then conditions for (S1-b) are satisfied and hence, $\|\mathbf{x}(t)\|$ will never grow at the initial time.

Note that the two bounds, $\lambda_{\max}$ and $\lambda_{\min}$, are independent of initial infected populations; in addition, their dependence on other parameter values is no longer linear like that of Γ_0. Analyzing the expressions given by (11), we obtain that (see Appendix A)

- $\lambda_{\max}$ and $\lambda_{\min}$ are increasing with respect to $\Gamma_0^{(i)}$ (hence, they are increasing in β_i and decreasing in γ_i) for $i \in \{1, 2\}$;
- $\lambda_{\max}$ is decreasing in d_{21}^I and increasing in d_{12}^I if $m_1 > m_2$, while it is increasing in d_{21}^I and decreasing in d_{12}^I if $m_1 < m_2$;
- $\lambda_{\min}$ is always decreasing with respect to d_{12}^I and d_{21}^I.

3.3 General Case of *n* Patches

Now we generalize the results to the case with $n \geq 2$. Given initial conditions

$$[S_i(0),\ I_i(0),\ R_i(0)] = [S_{i0},\ I_{i0},\ R_{i0}] \in \mathbb{R}^3_+, \quad i \in \Omega,$$

the long-term dynamics of this system is of the same nature as that of the two-patch case—disease will eventually die out as $t \to \infty$ (this can be shown by a similar argument as for the case $n = 2$). For short-term dynamics, we let $x(t) = [I_1(t), \cdots, I_n(t)]$. Linearize the equations governing the infected components in (3) at a disease-feee state

$$[S_i(0),\ I_i(0),\ R_i(0)] = [S_{i0},\ I_{i0},\ R_{i0}] \in \mathbb{R}^3_+, \quad i \in \Omega,$$

to obtain the linear system $\mathbf{x}(t) = \mathbf{A_0}\mathbf{x}(t)$ where

$$\mathbf{A_0} = \left[a_{ij}\right]_{n\times n}: \ a_{ii} = \beta_i S_{i0} - \gamma_i - \sum_{j\in\Omega, j\neq i} d^I_{ji}, \qquad i \in \Omega,$$
$$a_{ij} = d^I_{ij}, \qquad i, j \in \Omega,\ i \neq j.$$

The initial amplification rate, with a bit more lengthy calculations than for (9), is similarly obtained as

$$\Gamma_0 = \frac{\sum_{i\in\Omega}\left(\beta_i S_{i0} - \gamma_i - \sum_{j\in\Omega, j\neq i} d^I_{ji}\right) I^2_{i0} + \sum_{i,j\in\Omega, j>i}(d^I_{ij} + d^I_{ji}) I_{i0} I_{j0}}{\sum_{i\in\Omega} I^2_{i0}}. \tag{12}$$

This shows that the initial amplification rate is linearly increasing with respect to transmission rates β_i and initial susceptible population sizes S_{i0}, and is linearly decreasing with respect to removal rates γ_i. Its dependence on travel rate d^I_{ij} is determined by the difference in infected populations between two patches, I_{i0} and I_{j0},

$$\frac{\partial \Gamma_0}{\partial d^I_{ji}} = I_{i0}(I_{j0} - I_{i0}),$$

which clearly indicates how the impact of dispersals of infected individuals depends on the infected populations in the patches involved. Similarly, we can find the upper and lower bounds of Γ_0 by calculating the largest and smallest eigenvalues of $H(\mathbf{A_0})$. However, it is not always possible to obtain explicit expressions for the eigenvalues when the matrix is large.

Obviously, the formula (12) for the amplification rate over all patches (forming a health administration unit) can also be rewritten as a deviation from the average amplification rate over the patches (i.e., the generalization of (10)):

$$\Gamma_0 = \sum_{i\in\Omega} \frac{I^2_{i0}}{\sum_{i\in\Omega} I^2_{i0}} \Gamma^{(i)}_0 + \frac{\sum_{i,j\in\Omega, j>i}(d^I_{ij} + d^I_{ji}) I_{i0} I_{j0} - \sum_{j\in\Omega, j\neq i} d^I_{ji} I^2_{i0}}{\sum_{i\in\Omega} I^2_{i0}}. \tag{13}$$

4 SIR Endemic Patch Model

In this section, we incorporate a simple demographic structure into the SIR patch model (3). Let $N_i(t) = S_i(t) + I_i(t) + R_i(t)$ for $i \in \Omega$. Assume that there are no deaths caused by the disease, and the birth rate and the natural death rate in each patch are set to be equal. Hence, the total population size of two patches, $N = N_1(t) + N_2(t)$, remains constant. Under such a scenario, the disease dynamics of patches are now governed by the following system of ODEs:

$$\begin{cases} S_i'(t) = \sum\limits_{j\in\Omega, j\neq i} d_{ij}^S S_j - \sum\limits_{j\in\Omega, j\neq i} d_{ji}^S S_i + b_i N_i - \beta_i S_i I_i - b_i S_i, \\ I_i'(t) = \sum\limits_{j\in\Omega, j\neq i} d_{ij}^I I_j - \sum\limits_{j\in\Omega, j\neq i} d_{ji}^I I_i + \beta_i S_i I_i - \gamma_i I_i - b_i I_i, \\ R_i'(t) = \sum\limits_{j\in\Omega, j\neq i} d_{ij}^R R_j - \sum\limits_{j\in\Omega, j\neq i} d_{ji}^R R_i + \gamma_i I_i - b_i R_i, \end{cases} \quad \text{for } i \in \Omega. \tag{14}$$

Applying Proposition 1.1 in Chepyzhov and Vishik (2002) to this ODE model over patches, one can easily verify that all solutions to the initial value problem remain non-negative for all $t > 0$, provided that the initial values of the variables are non-negative. A disease-free equilibrium for model (14) is given by

$$\mathbf{E_0} = [S_1^{(0)}, \ldots, S_n^{(0)}, 0, \ldots, 0, R_1^{(0)}, \ldots, R_n^{(0)}].$$

According to the R-equation in (14), we have $R_i^{(0)} = 0$ for all $i \in \Omega$. Then, $\mathbf{S}^{(0)} = [S_1^{(0)}, \ldots, S_n^{(0)}]$ is a solution to the linear system,

$$\begin{cases} \sum\limits_{j\in\Omega, j\neq i} d_{ij}^S S_j - \sum\limits_{j\in\Omega, j\neq i} d_{ji}^S S_i = 0, \quad i \in \Omega, \\ \sum\limits_{i\in\Omega} S_i = N. \end{cases}$$

The Basic Reproduction Number. Based on the concept of next generation matrix presented in van den Driessche and Watmough (2002), we define

$$F := \begin{bmatrix} \beta_1 S_1^{(0)} & 0 & \ldots & 0 \\ 0 & \beta_2 S_2^{(0)} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \beta_n S_n^{(0)} \end{bmatrix}$$

and

$$V := \begin{bmatrix} \gamma_1 + b_1 + \sum_{j\neq 1} d_{j1}^I & -d_{12}^I & \dots & -d_{1n}^I \\ -d_{21}^I & \gamma_2 + b_2 + \sum_{j\neq 2} d_{j2}^I \dots & & -d_{2n}^I \\ \dots & \dots & \dots & \dots \\ -d_{n1}^I & -d_{n2}^I & \dots \gamma_n + b_n + \sum_{j\neq n} d_{jn}^I & \end{bmatrix}.$$

Then, the next generation matrix is FV^{-1} and the basic reproduction number is defined as its spectral radius, $\mathscr{R}_0 = \rho(FV^{-1})$. By Theorem 2 in van den Driessche and Watmough (2002), the disease-free equilibrium is locally asymptotically stable if $\mathscr{R}_0 < 1$ but is unstable if $\mathscr{R}_0 > 1$.

Reactivity of the Disease Free Equilbrium. Considering merely the infection-related variables (i.e., $I_1(t), \dots, I_n(t)$), the Jacobian matrix of model (14) is

$$\mathbf{J} = \left[J_{ij}\right]_{n\times n} : \quad J_{ii} = \beta_i S_i - \gamma_i - b_i - \sum_{j\neq i} d_{ji}^I, \qquad i \in \Omega,$$
$$J_{ij} = d_{ij}^I, \qquad i, j \in \Omega,\ i \neq j.$$

Evaluated at the disease-free equilibrium, $\mathbf{E_0}$, the Jacobian matrix becomes $\mathbf{J}_0 = F - V$. One can obtain the same linearization according to Mari et al. (2017) by letting $\mathbf{C} = [\mathbf{0}\ \mathbf{I}\ \mathbf{0}]$ (each block matrix is of the size $n \times n$). Then, the generalized reactivity of $\mathbf{E_0}$ is given by $\Lambda_0 = \lambda_{\max}(H(\mathbf{J}_0))$. The threshold index for epidemicity as defined by Hosack et al. (2008) is

$$\mathcal{E}_0 = \rho(H(F)H(V)^{-1}) = \rho(F \cdot H(V)^{-1}),$$

since F is diagonal. According to Hosack et al. (2008), if $\mathcal{E}_0 < 1$, then $\Lambda_0 < 0$ and $\mathbf{E_0}$ is non-reactive, and if $\mathcal{E}_0 > 1$, then $\Lambda_0 > 0$ and $\mathbf{E_0}$ is reactive.

The Amplification Rate at the Disease Free Equilibrium. Evaluating the Jacobian matrix, $\mathbf{J}$, at the initial point, we acquire the expression of Γ_0,

$$\Gamma_0 = \frac{\sum_{i\in\Omega}\left(\beta_i S_i^{(0)} - \gamma_i - b_i - \sum_{j\in\Omega, j\neq i} d_{ji}^I\right) I_{i0}^2 + \sum_{i,j\in\Omega, j>i}(d_{ij}^I + d_{ji}^I) I_{i0} I_{j0}}{\sum_{i\in\Omega} I_{i0}^2},$$

which is similar to that of the SIR epidemic patch model given by (12). Based on the definition of Γ_0, the size of solution, $\|[I_1(t), \dots, I_n(t)]\|$, will initially attenuate if $\Gamma_0 < 0$, while it will initially amplify if $\Gamma_0 > 0$.

It is well-known that for disease mode with demographic structure, the basic reproduction number $\mathscr{R}_0$ is an index for long-term asymptotic behaviour. Now our newly introduced initial amplification rate Γ_0 is an index for initial short-term disease dynamics. With respect to these two indices, there are four possibilities:

(I) when $\mathscr{R}_0 < 1$,

(I-a) if $\Gamma_0 > 0$, then $\|[I_1(t),\ I_2(t)]\|$ initially grows but eventually converges to zero;
(I-b) if $\Gamma_0 < 0$, then $\|[I_1(t),\ I_2(t)]\|$ initially decays and converges to zero in the long run;

(II) when $\mathscr{R}_0 > 1$,

(II-a) if $\Gamma_0 > 0$, then $\|[I_1(t),\ I_2(t)]\|$ initially grows and ultimately approaches a positive steady state;
(II-b) if $\Gamma_0 < 0$, then $\|[I_1(t),\ I_2(t)]\|$ initially decays before reaching a positive steady state.

For Γ_0, we have derived an explicit formula for Γ_0 in terms of the model parameters and $S_i^{(0)},\ i = 1, 2, \cdots, n$, together with the initial infections $I_{i0},\ i = 1, 2 \cdots, n$. But $S_i^{(0)},\ i = 1, 2, \cdots, n$ also depend on the model parameters and it is not easy to obtain explicit formula for such dependence for general n. For $\mathscr{R}_0$ however, as in most (if not all) disease model on patches, unfortunately it is also challenging to derived an explicit formula for the basic reproduction number $\mathscr{R}_0$ in terms of the model parameter. This makes analytical exploration of the above four possibilities impossible, and we need to explore them numerically. To make the numerical explorations a bit easier in the next section, we only consider simple case of $n = 2$ for which (14) reduces to

$$\begin{cases} S_1'(t) = d_{12}^S S_2 - d_{21}^S S_1 + b_1 N_1 - \beta_1 S_1 I_1 - b_1 S_1, \\ S_2'(t) = d_{21}^S S_1 - d_{12}^S S_2 + b_2 N_2 - \beta_2 S_2 I_2 - b_2 S_2, \\ I_1'(t) = d_{12}^I I_2 - d_{21}^I I_1 + \beta_1 S_1 I_1 - \gamma_1 I_1 - b_1 I_1, \\ I_2'(t) = d_{21}^I I_1 - d_{12}^I I_2 + \beta_2 S_2 I_2 - \gamma_2 I_2 - b_2 I_2, \\ R_1'(t) = d_{12}^R R_2 - d_{21}^R R_1 + \gamma_1 I_1 - b_1 R_1, \\ R_2'(t) = d_{21}^R R_1 - d_{12}^R R_2 + \gamma_2 I_2 - b_2 R_2, \end{cases} \tag{15}$$

For (15), one can easily find the unique disease-free equilibrium as

$$\mathbf{E_0} = [S_1^{(0)}, S_2^{(0)}, 0, 0, 0, 0] = \left[\frac{d_{21}^S N}{d_{12}^S + d_{21}^S},\ \frac{d_{12}^S N}{d_{12}^S + d_{21}^S},\ 0,\ 0,\ 0,\ 0 \right].$$

For this simple case of $n = 2$, our numerical computations of $\mathscr{R}_0$ and Γ_0 for (15) indicate that all four possibilities can occur.

5 Numerical Explorations

In this section, we present some numerical results for (15) and (7). The goal is to show that the four possibility can occur within appropriate parameter ranges; and also to numerically and visually demonstrate the impact of the model parameters on the two important indices $\mathscr{R}_0$ and Γ_0.

(N1). $\mathscr{R}_0 - 1$May or May Not Be Consistent with $\Gamma_0$for (15).

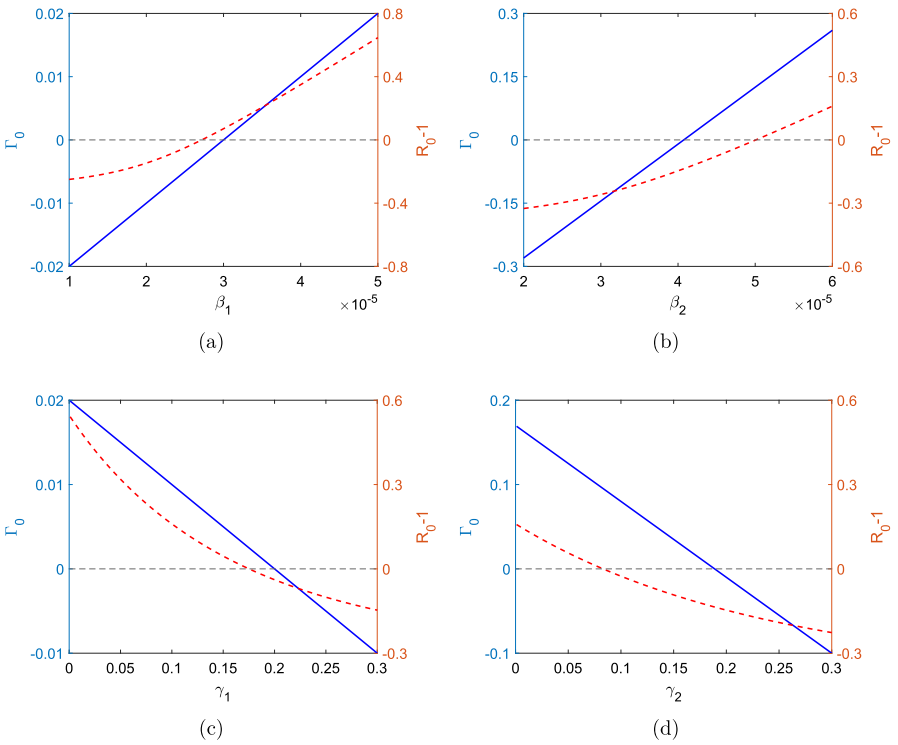


Fig. 1 Indices Γ_0 (solid line) and $\mathscr{R}_0$ (dotted line) for 4model:endemic:2-patch with the parameter values in (16). **a** $\beta_2 = 4 \times 10^{-5}$, $\gamma_1 = 0.3$, $\gamma_2 = 0.2$; **b** $\beta_1 = 2 \times 10^{-5}$, $\gamma_1 = 0.3$, $\gamma_2 = 0.2$; **c** $\beta_1 = 2 \times 10^{-5}$, $\beta_2 = 4 \times 10^{-5}$; $\gamma_2 = 0.2$; **d** $\beta_1 = 2 \times 10^{-5}$, $\beta_2 = 4 \times 10^{-5}$; $\gamma_1 = 0.3$ (Color figure online)

To demonstrate the four possibilities (I-a), (I-b), (II-a) and (II-b), we fixed the parameters:

$$\begin{cases} b_1 = b_2 = 0.2,\ d_{21}^S = 0.4,\ d_{12}^S = 0.6,\ d_{21}^I = 0,\ d_{12}^I = 0.3; \\ [S_{10}, S_{20}, I_{10}, I_{20}] = [10000, 15000, 1000, 3000]. \end{cases} \tag{16}$$

Then by varying the values of β_i and γ_i, $i = 1, 2$, we can observe the four combinations, as shown in Fig. 1. For example, from Fig. 1a, as β_1 increases, we observe switches from I-(b) to (II-b) and then to (II-a); while from Fig. 1b, we observe switches from scenario I-(b) to (I-a) and then to (II-a). Similarly, from Fig. 1c, as γ_1 increases, we observe switches from (II-a) to (I-b) and then to (I-b); while when γ_1 is fixed and γ_2 increases, Fig. 1d gives the same switch pattern as in Fig. 1c.

(N2). Impact of Spatial Heterogeneity on Initial Disease Dynamics for (7).

In Sect. 3, we have seen the initial amplification rate Γ_0 given by formula (9) is a linear increasing function of β_1, the transmission rate in patch 1. Positive/negative Γ_0 indicates an initial escalation/mitigation of the disease measured by $\|\mathbf{x}(t)\|$—a measurement of the epidemic severity over the two patches as a whole (a distance from the disease-free scenario). A decrease in β_1, yielding a lower Γ_0, leads to effective

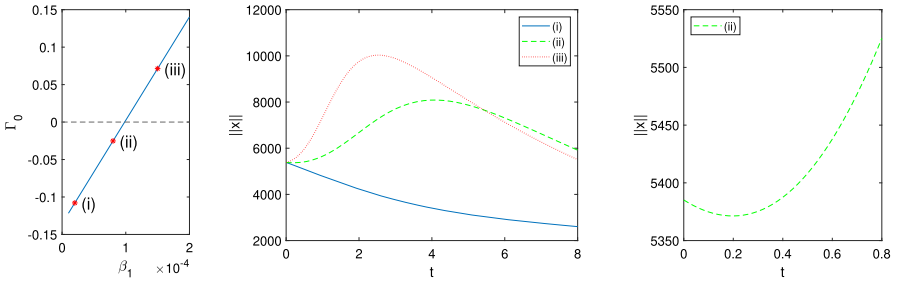


Fig. 2 Three sample points of β_1 are taken as: (i) 0.2×10^{-4}, (ii) 0.8×10^{-4}, and (iii) 1.5×10^{-4}. **a** The initial amplification rate Γ_0 is negative for (i) and (ii), while it is positive for (iii). **b** The outbreak is mitigated with β_1 decreasing from (iii) to (ii) and is completely under control when β_1 further reduces to (i). **c** For the chosen value (ii), the magnitude of the solution initially experiences a short-period reduction then grows to its maximal value. Other parameters are set to $\beta_2 = 2 \times 10^{-5}$, $\gamma_1 = 0.2$, $\gamma_2 = 0.4$, $d_{21}^S = 0.08$, $d_{12}^S = 0.1$, $d_{21}^I = 0.02$, $d_{12}^I = 0.05$ and $[S_{10}, S_{20}, I_{10}, I_{20}] = [10000, 15000, 2000, 5000]$ (Color figure online)

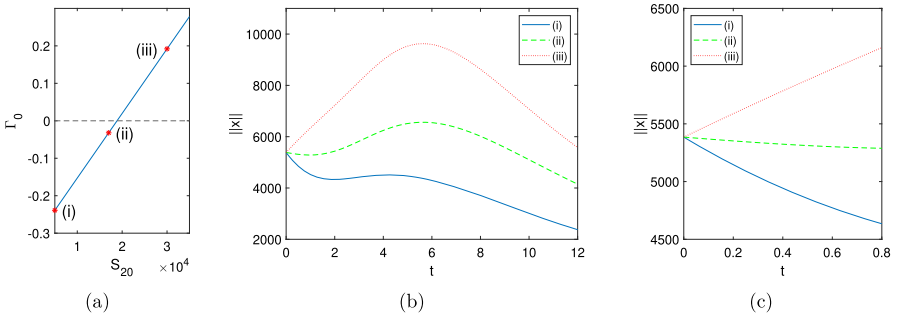


Fig. 3 Three sample points of S_{20} are taken as: (i) 5000, (ii) 17,000, and (iii) 30,000. **a** The initial amplification rate Γ_0 is a linear increasing function of S_{20}, whose value is negative for (i) and (ii) and is positive for (iii). **b** The dynamics of $\|\mathbf{x}(t)\| = \|[I_1(t), I_2(t)]\|$ for different initial conditions. **c** The linear dynamics of $\|\mathbf{x}(t)\|$ near the initial point. The other parameters are set to $\beta_1 = 5 \times 10^{-5}$, $\beta_2 = 2 \times 10^{-5}$, $\gamma_1 = 0.2$, $\gamma_2 = 0.4$, $d_{21}^S = 0.08$, $d_{12}^S = 0.1$, $d_{21}^I = 0.02$, $d_{12}^I = 0.05$ and $[S_{10}, I_{10}, I_{20}] = [10{,}000, 2000, 5000]$ (Color figure online)

control of the epidemic, as is shown in Fig. 2b. Figure 2c gives an example when $\|[I_1(t),\ I_2(t)]\|$ initially decreases but the outbreak will continue after a short period of time. In this case, the local basic reproduction numbers in the two patches without dispersal satisfy $\mathscr{R}_0^{(1)} > 1 > \mathscr{R}_0^{(2)}$, where $\mathscr{R}_0^{(i)} = \beta_i S_{i0}/\gamma_i,\ \ i = 1, 2$.

Figure 3 illustrates how vaccination impacts the initial amplification rate by reducing the initial susceptiple populations. Assume that only the number of susceptible individuals in patch 2 changes while all other initial conditions and parameter values are fixed. The numerical examples given in Fig. 3b show that the outbreak is better controlled by decreasing S_{20} (e.g., vaccinating more people). The linear dynamics of solutions near the initial point, as is displayed in Fig. 3c, are consistent with the corresponding value of Γ_0.

(N3). Impact of Dispersals on Short-term and Long-term Dynamics

The dependence of Γ_0 on the between-patch travel rates of infectious individuals is demonstrated in Fig. 4a–d for the epidemic model (7). Comparing Fig. 4a and c, 4b and d, the results imply that a higher local (when isolated) initial amplification rate produces a greater Γ_0, which is consistent with expression (10). The dynamics of $\|\mathbf{x}(t)\|$ for the sample sets of travel rates are shown in Fig. 4e, f, where the solutions (i)-1 and (iii)-1, (ii)-1 and (iv)-1, (i)-2 and (iii)-2, (ii)-2 and (iv)-2 share the same Γ_0. We observe that the system may demonstrate various behaviours when different parameter values are chosen, even if the calculated Γ_0 and initial conditions are the same. This is because such *linear approximation* is only valid in a *small neighbourhood* close to the initial point, and at later times the effect of nonlinearities becomes dominant.

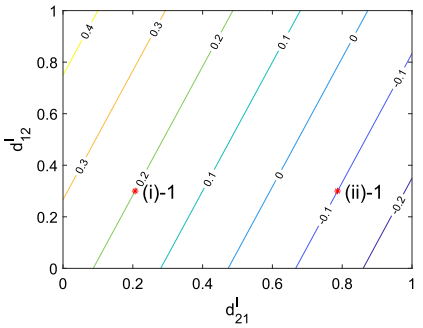
For the endemic model (14), among the four possible combinations of short-term and long-term dynamics summarized in Sect. 4, we are mostly interested in the two scenarios: (I-a) the disease dies out in the long run but there exists at least a transitory outbreak; (II-b) the disease persists for all $t > 0$ but the epidemic size drops initially. See the figures in Fig. 5 for a demonstration. These two cases are of particular importance in disease controlling, because the inconsistency of the short-term disease dynamics and the long-term disease dynamics may easily mislead not only the public but also the health policymakers. The effect of a control measure may be misunderstood without knowing its effect on both long-term and short-term dynamical behaviours. If $\mathscr{R}_0$ is the only index to be examined, the unanticipated epidemic in (I-a) may have very serious consequences. If (II-b) happens, people may be misled by the initial decrease in the size of infections when the disease is indeed an endemic.

The impact of dispersal on long-term asymptotic dynamics and transitory amplification/attenuation in infection size is demonstrated by Figs. 6 and 7. According to the definitions given in Sect. 4, both Γ_0 and $\mathscr{R}_0$ depend on the dispersal rates of infectives. An example is shown in Fig. 6 where the contours give the values of Γ_0 and $\mathscr{R}_0$ for combinations of d_{12}^I and d_{21}^I. With the chosen parameter values and initial conditions, Γ_0 and $\mathscr{R}_0$ are increasing in d_{21}^I and decreasing in d_{12}^I. Besides, there exists an area on the d_{21}^I–d_{12}^I plane within which transitory epidemic is possible before the disease dies out ($\Gamma_0 > 0$ but $\mathscr{R}_0 < 1$).

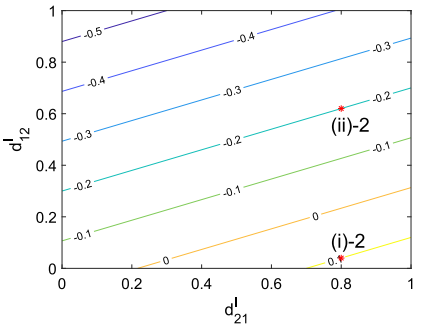
As for the dispersal rates of susceptibles, they do not affect Γ_0 but change the value of $\mathscr{R}_0$ via the population sizes in disease-free equilibrium. Figure 7 shows the contour graphs of $\mathscr{R}_0$ on the d_{21}^S–d_{12}^S plane, which includes the four possible combinations of short-term and long-term behaviours. It seems to suggest that an infectious disease is more likely to develop into an endemic when $\Gamma_0 > 0$. In addition, we point out that the dispersal of recovered individuals has no impact on Γ_0 and $\mathscr{R}_0$.

6 Conclusion and Discussion

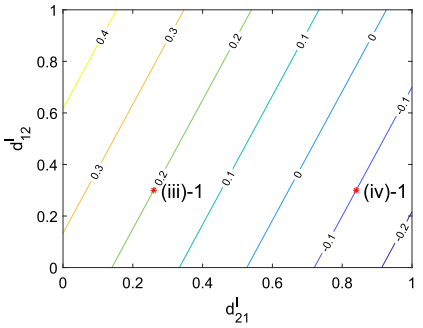
For a newly emerging infectious disease, intervention/control decisions are usually based on the initial/present short term disease dynamics which may not be reflected/implied by the long term disease dynamics. The later is typically described by the basic reproduction number and has been extensively and intensively studied for various mathematical models. The former, however, has been paid less attention, particularly in the context of mathematical modelling and analysis of disease dynam-



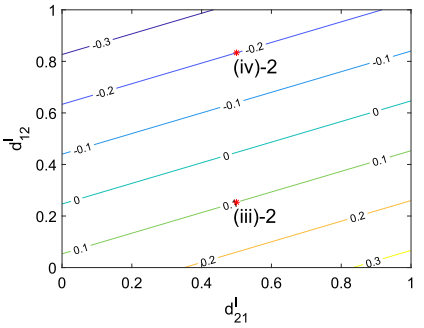
(a) $\Gamma_0^{(1)} > 0$, $\Gamma_0^{(2)} < 0$, $I_{10} > I_{20}$



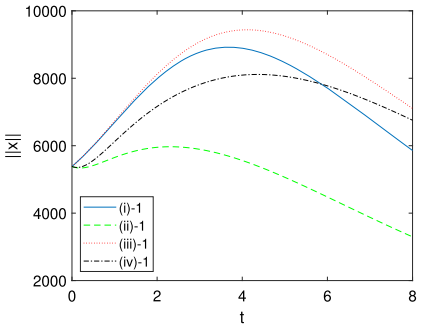
(b) $\Gamma_0^{(1)} > 0$, $\Gamma_0^{(2)} < 0$, $I_{10} < I_{20}$



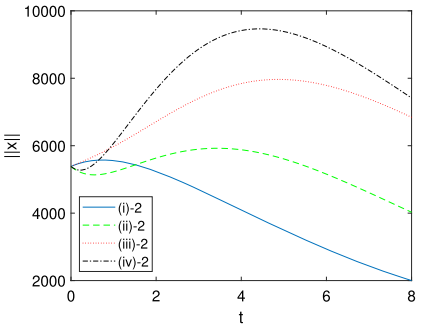
(c) $\Gamma_0^{(1)} > 0$, $\Gamma_0^{(2)} > 0$, $I_{10} > I_{20}$



(d) $\Gamma_0^{(1)} > 0$, $\Gamma_0^{(2)} > 0$, $I_{10} < I_{20}$

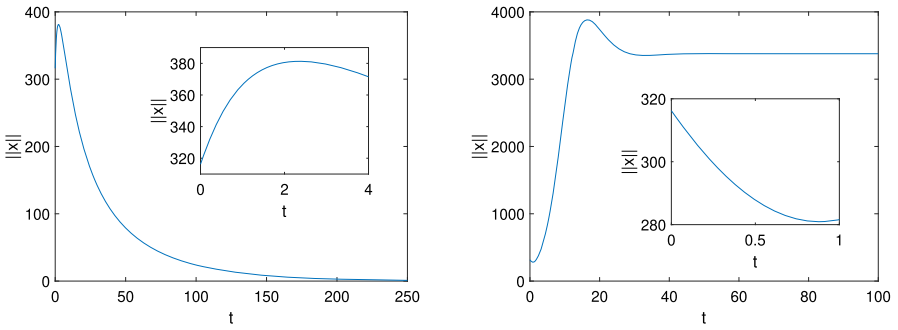


(e) $I_{10} > I_{20}$



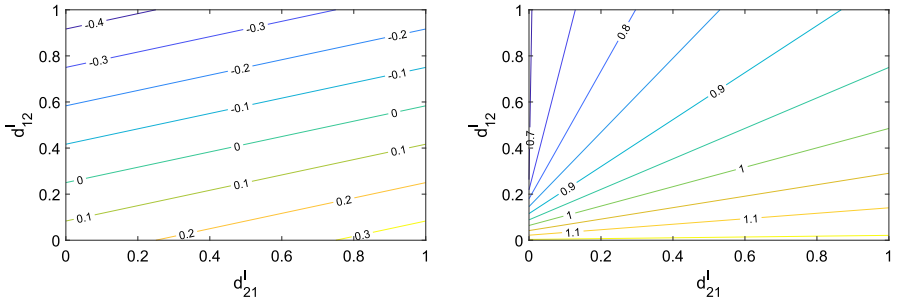
(f) $I_{10} < I_{20}$

Fig. 4 **a**–**d** are the contour plots of Γ_0 with respect to dispersal rates of infected populations. **e**–**f** illustrate the dynamics of $\|\mathbf{x}(t)\|$ for the sample sets. For the two dispersion rates d_{12}^I and d_{21}^I, the choices (i)-1 and (iii)-1, (ii)-1 and (iv)-1, (i)-2 and (iii)-2, (ii)-2 and (iv)-2, each respectively yields the same Γ_0. The other parameters are set to $\beta_1 = 5 \times 10^{-5}$, $\beta_2 = 2 \times 10^{-5}$, $\gamma_1 = 0.2$, $d_{21}^S = 0.4$, $d_{12}^S = 0.6$, and $\gamma_2 = 0.4$ in (**a**)–(**b**); $\gamma_2 = 0.2$ in (c)-(d). The initial conditions are $[S_{10}, S_{20}, I_{10}, I_{20}] = [10{,}000, 15{,}000, 5000, 2000]$ in (**a**)–(**c**)–(**e**); and $[S_{10}, S_{20}, I_{10}, I_{20}] = [10{,}000, 15{,}000, 2000, 5000]$ in (**b**)–(**d**)–(**f**) (Color figure online)



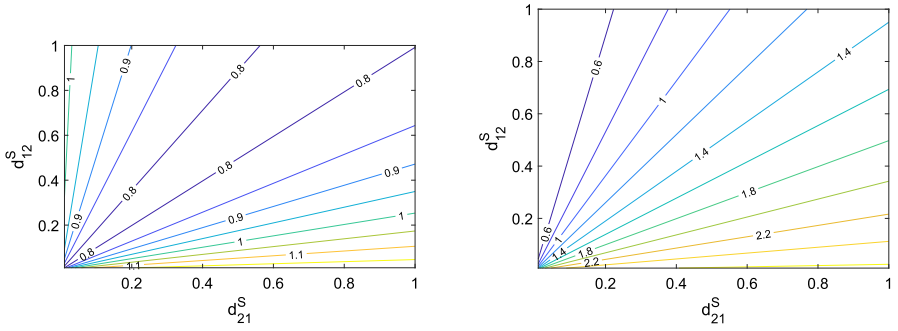
(a) $\mathscr{R}_0 < 1$ and $\Gamma_0 > 0$ (b) $\mathscr{R}_0 > 1$ and $\Gamma_0 < 0$

Fig. 5 Long-term and short-term dynamics of $\|[I_1(t),\ I_2(t)]\|$. **a** is an example of (I-a) with $\beta_1 = 10^{-5}$ and $\beta_2 = 6 \times 10^{-5}$. **b** is an example of (II-b) with $\beta_1 = 6 \times 10^{-5}$ and $\beta_2 = 2 \times 10^{-5}$. Set $d_{21}^R = 0.4$, $d_{12}^R = 0.6$ and $[S_{10}, S_{20}, I_{10}, I_{20}, R_{10}, R_{20}] = [10{,}000, 15{,}000, 100, 300, 0, 0]$. Other parameter values are the same as those used in Fig. 1 (Color figure online)



(a) (b)

Fig. 6 The contour graphs of Γ_0 and $\mathscr{R}_0$ for dispersal rates of infectives. Other parameter values and initial conditions are the same as those used in Fig. 1 (Color figure online)



(a) $\Gamma_0 < 0$ (b) $\Gamma_0 > 0$

Fig. 7 The contour graphs of $\mathscr{R}_0$ for dispersal rates of susceptibles. Set **a** $d_{21}^I = 0.2$ and $d_{12}^I = 0.8$ so that $\Gamma_0 < 0$; **b** $d_{21}^I = 0.9$ and $d_{12}^I = 0.1$ so that $\Gamma_0 > 0$. Other parameter values and initial conditions are the same as those used in Fig. 1 (Color figure online)

ics. In this work, we have studied the short-term disease dynamics, in comparison to the long-term disease dynamics, of some SIR models over a *patch environment*. After reviewing some related notions in mathematical ecology, we introduced the disease amplification rate to measure whether the epidemic of an infection disease is escalating or mitigating at the initial time or at a specified time. This amplification rate is closely related to the notion of reactivity in Neubert and Caswell (1997) and many follow-up works in ecological context (e.g. Mari et al. 2017; Wang et al. 2019; Lutscher and Wang 2020). The initial amplification rate Γ_0 is defined as the *relative rate of change* of the epidemic size at initial time ($t = 0$) measured by the Euclidean norm of the infected populations over the patches and hence, its sign can predicts escalation or mitigation of the epidemics in a near short period of time after $t = 0$. We point out that the initial time of our model ($t = 0$) is not necessary to denote the beginning of an outbreak; indeed, it can be set as any point of time during the course of a disease when disease control policies are considered.

We firstly applied Γ_0 to the SIR epidemic patch model (3). We have shown that this extended system (3) also does not allow the disease to persist since $I_i(t) \to 0$ as $t \to \infty$. The calculation of Γ_0 helps us to explore the patterns by which the disease dies out. We have obtained an expression of Γ_0 for the patch model (3) and analyzed its dependence on the involved parameters and initial conditions. Particularly, the overall amplification rate Γ_0 can be expressed as a deviation from the average amplification rate among all patches. Such an expression help us understand how the spatial heterogeneities reflected by the local amplification rates ($\Gamma_0^{(i)}$, $i = 1, \cdots, n$) and the dispersions between the patches affect the overall amplification rate Γ_0. Some numerical examples have been given for the 2-patch case and the results visually demonstrate how different interventions affect Γ_0. Based on the upper and lower bounds of the Rayleigh quotient, we have also estimated Γ_0, with $t = 0$ indicating the onset time of an epidemic when the system is at an (approximate) disease-free *state*.

We continued to study the SIR *endemic* patch model (14) with a simple demographic structure in Sect. 4. Unlike the blue epidemic model (3), this model system admits a locally asymptotically stable disease-free equilibrium. Therefore, we are able to obtain $\mathscr{R}_0$ by the next generation matrix method (van den Driessche and Watmough 2002) and calculate the initial infection amplification according to Neubert and Caswell (1997); Mari et al. (2017). The expression for Γ_0 is similar to that of the SIR *epidemic* patch model in Sect. 4. While $\mathscr{R}_0$ determines the long-term asymptotic behaviour, both reactivity and Γ_0 measure the short-term transitory dynamics. In addition, Γ_0 is the instant amplification/attenuation rate evaluated at the initial time, meaning that it depends on the initial condition. Reactivity, however, is defined as the maximal amplification rate over all possible (small) perturbations to disease-free equilibrium. We have further numerically compared Γ_0 and $\mathscr{R}_0$ as functions of infection-related parameters and dispersal rates of different compartments. The results suggest four possible combinations of transitory and asymptotic behaviours. Two of the scenarios are of particular interest in disease controlling: (I-a) the disease eventually dies out but transitory epidemics are possible; (II-b) the size of infected populations initially decreases but the disease will persist. This is because the disagreement between the short-term disease dynamics and the long-term disease dynamics in these two scenarios

may easily mislead not only the public, but also the health policymakers. Due to the influence of spatial heterogeneity and the dispersals, a disease with $\mathscr{R}_0 > 1$ may demonstrate initial decay short-term disease dynamics (harmful misleading); conversely, a disease with $\mathscr{R}_0 < 1$ may also show a short-term outbreak.

By the definition of Γ_0, we quantify the transitory behaviour for infected compartments as a whole instead of examining the specific dynamics in each patch. *Euclidean norm* has been used to measure the size of a vector. Thus, the value of Γ_0 and its sign are not always consistent with $I_1'(0)$, $I_2'(0)$ or $I_1'(0) + I_2'(0)$ (in the case of two patches for an example), and the results will be different *if other norms are adopted*. Also, our idea of solely considering the infection-related variables is a special case of the generalization proposed by Mari et al. (2017). Indeed, the measure can be evaluated based on unequally weighted state variables. We refer to that paper for a detailed discussion on amplification and reactivity of some of the state variables.

We point out that the initial amplification rate merely characterizes the linear dynamics near a given "initial time". Nonlinearities, however, can produce longer and more complex transient dynamics, as observed in Figs. 2c, 3b and 4e, f. In addition to reactivity, amplification envelope has also been proposed by Neubert and Caswell (1997) as another measure of transient dynamics, which is not included in this work in the context of infectious disease dynamics. On the other hand, Hastings and Higgins (1994) stressed the importance of transients in spatially structured ecological systems. For discrete-space models, there are some works that have explored transient behaviours by numerical simulations. See, e.g., Ruxton and Doebeli (1996); Saravia et al. (2000) for one-species models and Hastings (2001) for predator–prey systems. However, only a few studies have theoretically analyzed the effect of spatial heterogeneity on transient dynamics. The notions of reactivity and amplification envelop have been extended to advective systems by Anderson et al. (2008) and to reaction-diffusion systems by Wang et al. (2019). As for the patch model, we will leave this for future work.

To conclude this paper, we point out that this work is largely motivated by Neubert and Caswell (1997); Mari et al. (2017); Wang et al. (2019); Lutscher and Wang (2020)) in methodology. Hence, we follow these and some other follow-up works to use the Euclidian norm to measure the size of a vector. As for other norms, we have just noticed a recent publication (Harrington et al. 2022), in which other norms are discussed for the topics of reactivity, attenuation and transients in metapopulations. The authors present some examples showing that reactivity/attenuation in those discussed norms do not imply each other, but they advocate the use of l_1 norm because of the biological interpretations. Indeed, one of the referees also suggested the use of l_1 norm. Considering this manuscript is already pretty lengthy, we have to leave it as a future research project.

Acknowledgements The authors are grateful to the two anonymous referees for their careful reading and valuable comments on the first version of the paper, which have led to a significant improvement in readability of the paper.

Appendix

The Dependence of λ_{max} and λ_{min} for SIR Epidemic Patch Model

Taking partial derivatives of (11) with respect to $\Gamma_0^{(i)}$ for $i \in \{1, 2\}$, we have

$$\frac{\partial \lambda_{\max}}{\partial \Gamma_0^{(1)}} = \frac{\partial \lambda_{\min}}{\partial \Gamma_0^{(2)}} = \frac{M + \sqrt{M^2 + D^2}}{2\sqrt{M^2 + D^2}} \quad \text{and} \quad \frac{\partial \lambda_{\max}}{\partial \Gamma_0^{(2)}} = \frac{\partial \lambda_{\min}}{\partial \Gamma_0^{(1)}} = \frac{-M + \sqrt{M^2 + D^2}}{2\sqrt{M^2 + D^2}}$$

where $M := m_1 - m_2 = \Gamma_0^{(1)} - d_{21}^I - \Gamma_0^{(2)} + d_{12}^I$ and $D = d_{12}^I + d_{21}^I > 0$. All of these partial derivatives are positive since

$$\sqrt{M^2 + D^2} > |M|. \tag{17}$$

With respect to travel rates, the partial derivative of $\lambda_{\max}$,

$$\frac{\partial \lambda_{\max}}{\partial d_{21}^I} = \frac{D - M - \sqrt{M^2 + D^2}}{2\sqrt{M^2 + D^2}},$$

is positive if $M < 0$ and is negative if $M > 0$, and,

$$\frac{\partial \lambda_{\max}}{\partial d_{12}^I} = \frac{D + M - \sqrt{M^2 + D^2}}{2\sqrt{M^2 + D^2}},$$

is positive if $M > 0$ and is negative if $M < 0$. As for $\lambda_{\min}$, the partial derivatives are

$$\frac{\partial \lambda_{\min}}{\partial d_{21}^I} = \frac{-D + M - \sqrt{M^2 + D^2}}{2\sqrt{M^2 + D^2}} \quad \text{and} \quad \frac{\partial \lambda_{\min}}{\partial d_{12}^I} = \frac{-D - M - \sqrt{M^2 + D^2}}{2\sqrt{M^2 + D^2}},$$

which are both negative by the inequality (17).

References

Allen LJ, Bolker BM, Lou Y, Nevai AL (2007) Asymptotic profiles of the steady states for an SIS epidemic patch model. SIAM J Appl Math 67(5):1283–1309

Almarashi RM, McCluskey CC (2019) The effect of immigration of infectives on disease-free equilibria. J Math Biol 79:1015–1028

Anderson KE, Nisbet RM, McCauley E (2008) Transient responses to spatial perturbations in advective systems. Bull Math Biol 70:1480–1502

Andreasen V (2011) The final size of an epidemic and its relation to the basic reproduction number. Bull Math Biol 73:2305–2321

Arino J, Van den Driessche P (2003) A multi-city epidemic model. Math Popul Stud 10(3):175–193

Arino J, Van den Driessche P (2006) Disease spread in metapopulations. Fields Inst Commun 48:1–12

Arino J, Van Den Driessche P (2003) The basic reproduction number in a multi-city compartmental epidemic model. In: Positive systems. Springer, Berlin, pp 135–142

Chen S, Shi J, Shuai Z, Wu Y (2020) Asymptotic profiles of the steady states for an SIS epidemic patch model with asymmetric connectivity matrix. J Math Biol 80:2327–2361
Chepyzhov VV, Vishik MI (2002) Attractors for equations of mathematical physics. American Mathematical Society
Diekmann O, Heesterbeek JAP, Metz JA (1990) On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. J Math Biol 28:365–382
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 246:105–112
Harrington PD, Lewis MA, van den Driessche P (2022) Reactivity, attenuation, and transients in metapopulations. SIAM J Appl Dyn Syst 21:1287–1321
Hastings A (2001) Transient dynamics and persistence of ecological systems. Ecol Lett 4:215–220
Hastings A, Higgins K (1994) Persistence of transients in spatially structured ecological models. Science 263:1133–1136
Hethcote HW (1976) Qualitative analyses of communicable disease models. Math Biosci 28(3–4):335–356
Horn RA, Johnson CR (1985) Matrix analysis. Cambridge University Press, Cambridge
Hosack GR, Rossignol PA, Van Den Driessche P (2008) The control of vector-borne disease epidemics. J Theor Biol 255:16–25
Hsieh YH, Van den Driessche P, Wang L (2007) Impact of travel between patches for spatial spread of disease. Bull Math Biol 69:1355–1375
Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. Proc R Soc Lond Ser A 115(772):700–721
Lutscher F, Wang X (2020) Reactivity of communities at equilibrium and periodic orbits. J Theor Biol 493:110240
Ma J, Earn DJD (2006) Generality of the final size formula for an epidemic of a newly invading infectious disease. Bull Math Biol 68:679–702
Mari L, Casagrandi R, Rinaldo A, Gatto M (2017) A generalized definition of reactivity for ecological systems and the problem of transient species dynamics. Methods Ecol Evol 8:1574–1584
Mari L, Casagrandi R, Rinaldo A, Gatto M (2018) Epidemicity thresholds for water-borne and water-related diseases. J Theor Biol 447:126–138
Mari L, Casagrandi R, Bertuzzo E, Rinaldo A, Gatto M (2019) Conditions for transient epidemics of waterborne disease in spatially explicit systems. R Soc Open Sci 6:181517
Mari L, Casagrandi R, Bertuzzo E, Pasetto D, Miccoli S, Rinaldo A, Gatto M (2021) The epidemicity index of recurrent SARS-CoV-2 infections. Nat Commun 12:2752
Neubert MG, Caswell H (1997) Alternatives to resilience for measuring the responses of ecological systems to perturbations. Ecology 78(3):653–665
O'Regan SM, O'Dea EB, Rohani P, Drake JM (2020) Transient indicators of tipping points in infectious diseases. J R Soc Interface 17:20200094
Ruxton GD, Doebeli M (1996) Spatial self-organization and persistence of transients in a metapopulation model. Proc R Soc Lond B 263:1153–1158
Saravia LA, Ruxton GD, Coviella CE (2000) The importance of transient's dynamics in spatially extended populations. Proc R Soc Lond B 267:1781–1785
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
Wang W, Zhao XQ (2004) An epidemic model in a patchy environment. Math Biosci 190:97–112
Wang X, Efendiev M, Lutscher F (2019) How spatial heterogeneity affects transient behavior in reaction-diffusion systems for ecological interactions? Bull Math Biol 81:3889–3917
Woodall H, Bullock JM, White SM (2014) Modelling the harvest of an insect pathogen. Ecol Model 287:16–26