Los Alamos NATIONAL LABORATORY EST.1943

LA-UR-18-31251 (Accepted Manuscript)

Systematic evaluation of the population-level effects of alternative treatment strategies on the basic reproduction number

Romero-Severson, Ethan
Bulla, Ingo
Gromov, Dmitry

Provided by the author(s) and the Los Alamos National Laboratory (2019-02-20).

To be published in: Journal of Theoretical Biology

DOI to publisher's version: 10.1016/j.jtbi.2018.11.029

Permalink to record: http://permalink.lanl.gov/object/view?what=info:lanl-repo/lareport/LA-UR-18-31251

Disclaimer:
Approved for public release. Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the Los Alamos National Security, LLC for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

RESEARCH LIBRARY

Systematic evaluation of the population-level effects of alternative treatment strategies on the basic reproduction number

Dmitry Gromov[a,*], Ingo Bulla[b,c], Ethan O. Romero-Severson[d]

[a] *Faculty of Applied Mathematics and Control Processes, Saint Petersburg State University, St. Petersburg, Russia*
[b] *Institut für Mathematik und Informatik, Universität Greifswald, Walther-Rathenau-Straße 47, 17487 Greifswald, Germany*
[c] *Université Perpignan Via Domitia, IHPE UMR 5244, CNRS, F-66860 Perpignan, France*
[d] *Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, New Mexico, USA*

Abstract

An approach to estimate the influence of the treatment-type controls on the basic reproduction number, R_0, is proposed and elaborated. The presented approach allows one to estimate the effect of a given treatment strategy or to compare a number of different treatment strategies on the basic reproduction number. All our results are valid for sufficiently small values of the control. However, in many cases it is possible to extend this analysis to larger values of the control as was illustrated by examples.

Keywords: Disease control, Basic reproduction number, Treatment, Compartmental epidemic model, Next generation matrix

1. Introduction

The basic reproduction number (R_0), i.e. the number of infections generated by an infected person in a fully susceptible population, is a key determinant of the dynamics of an infectious disease. Interpretation of R_0 is complex and is covered in a number of excellent papers, see, e.g., [32, 17, 19] as well as the books [13, 12]. For the purposes of this paper, we note that R_0 can be interpreted as a threshold parameter: if $R_0 < 1$ then the disease can not sustain itself in the population and will eventually die out. That is, if our goal is to push a disease towards extinction, then we can categorize a population-level intervention as 'effective' if it reduces R_0 from where it was before the intervention was enacted.

*Corresponding author
Email addresses: dv.gromov@gmail.com (Dmitry Gromov), ingobulla@gmail.com (Ingo Bulla), eoromero@lanl.gov (Ethan O. Romero-Severson)

Computation of R_0. R_0 can be computed or estimated from partially empirical considerations, for instance the "survival function"-based approach [18] that assumes knowledge of both the survival probability $F(t)$ and the infectivity $b(t)$ of an individual as functions of time. In addition to being difficult to implement, such approaches can hardly be used for computing R_0 in the case where there are more than one group of infected. This issue was addressed in [14], where an elegant approach to computing R_0 was proposed, termed the *next generation matrix* (NGM) method. This method boils down to computing R_0 as the spectral radius of a specially constructed matrix. Later, this method was detailed and substantiated in a number of papers, see, e.g., [32] and [10]. The power of the NGM approach lies in its universality. It can be applied to any population balance model as long as it satisfies a number of very natural assumptions. The result is based on certain properties of positive and inverse positive matrices, in particular M-matrices [5, 21].

Contribution. We aim at developing this approach in that we extend the notion of the basic reproduction number to a class of controlled disease propagation models. Much of applied theoretical epidemiology focuses on how treatment policies, intervention designs, and novel treatments will impact the burden of a given disease by considering alternative scenarios. There have been a large number of papers aimed at evaluating the efficacy of different treatment schemes using various methods, from numerical simulation to optimal control. We mention [9, 8, 24, 6, 4, 30, 27, 31, 16] for a short list of related research. Modeling efforts have even attempted to measure the effects of control programs in historical epidemics [7]. However, most modeling efforts do not attempt to measure intervention effects systematically, i.e. by considering the joint effects of model parameters on the relative efficacy of alternative interventions. In this paper we address treatment programs, i.e., the intervention strategies that result in moving infected individuals either into a different group of infected (with different biological or behavioral characteristics) or into a group of not-infected, which can correspond to susceptible, recovered or any other group that consists of not contagious individuals.

When applied to one or several groups of infected individuals, the action of a treatment strategy can be described by a parameter u that corresponds to a fraction of potentially eligible individuals that are administered the treatment during a unit of time. Typically, this fraction is rather small. Our first result consists in defining the notion of a controlled reproduction number $R_0(u)$ which explicitly contains u as a parameter. The next result allows to estimate the action of the particular treatment strategy on the value of the controlled reproduction number $R_0(u)$. We formulate conditions under which the application of a treatment strategy leads to the reduction in $R_0(u)$. This allows for evaluating the efficacy of a devised treatment strategy as well as for comparing the efficacy of different strategies. The obtained results are illustrated by a number of examples.

We note that the importance of the methodology, considered in the paper, extends well beyond comparing different types of treatments. This paper makes a first step towards formalizing certain problems from the field of mathematical epidemiology within the framework of control theory, thus allowing the wealth of existing control methods to be used for solving epidemiological problems (see, e.g., [3] for an elucidation of the basic control-theoretic concepts).

The paper is organized as follows: Section 2 describes a general compartmental epidemic model and briefly introduces the NGM method for computing the basic reproduction number R_0. The main result of the paper along with a discussion on further

extensions and ramifications of the developed approach are presented in Sec. 3. The obtained results are illustrated by a number of examples in Sec. 4. The paper is concluded with a discussion section. Finally, there are two appendices containing necessary technical information.

2. Epidemics dynamics

2.1. A disease propagation model

When modeling the process of disease propagation in a heterogeneous population, the standard procedure consists in dividing the total population into a number of classes (compartments) according to some criteria relevant to the disease transmission: disease status, ability to contract disease, behavioral (contact) pattern and so on. All individuals within a compartment are assumed to be identical in their evolution. Thus we can consider only the number of individuals within each compartment, denoted by x_i, where $i = 1, \dots, n$ is the number of the compartment.

General formulation. The dynamics of these groups include transitions between groups and the in- and outflows associated with these groups. In the following it is assumed that the total inflow is constant over time and does not depend on the population size while the outflow depends on both the size and the structure of the population. The evolution of the ith state is thus described by the following compartmental differential equation (DE):

$$\dot{x}_i = \Phi_i(x) = w_i + \sum_{i \neq j} \big(a_{ij}(x) - a_{ji}(x)\big) - a_i x_i, \tag{1}$$

where $x \in \mathbb{R}^n_{\geq 0}$ is the state, w_i is the constant inflow, $a_i x_i$ is the outflow from the i-th compartment, and $a_{ij}(x)$, $i \neq j$, is the flow rate from the j-th to the i-th compartment.

The flow rate functions are assumed to be C^∞ for all $x \in \mathbb{R}^n_{\geq 0}$. Furthermore, all flow rates have to satisfy the following properties:

$$w_i \geq 0, a_i \geq 0, a_{ji}(x) \geq 0, \qquad \forall x \in \mathbb{R}^n_{\geq 0}, i, j = 1, \dots, n. \tag{2a}$$

$$x_i = 0 \implies a_{ji}(x) = 0 \qquad \forall i, j = 1, \dots, n. \tag{2b}$$

The property (2b) implies that there is no outflow from an empty compartment.

Compartmental epidemic model. All state variables are divided into two groups depending on whether the respective compartment corresponds to the infected or non-infected individuals (regardless of whether they were infected earlier or not).

Let there be l state variables representing the "infected" compartments. We rearrange the states in the way that the vector of the state variables takes the form $x^\top = [x_I^\top, x_N^\top]$, where $x_I^\top = [x_1, \dots, x_l]$ and $x_N^\top = [x_{l+1}, \dots, x_n]$ are the states associated with "infected" and "non-infected" compartments. We will refer to the respective compartments as the I- and N-compartments. Further, we write $\Phi(x)$ as the sum of two vector-valued functions

$$\Phi(x) = \Phi^I(x) + \big[w + \Phi^{in}(x) - \Phi^{out}(x)\big] = \Phi^I(x) + \Phi^C(x), \tag{3}$$

where $\Phi_i^{out}(x) = \sum_{j \neq i} a_{ji}(x) + a_i x_i$, and

$$\Phi^I(x) = \begin{bmatrix} \left(\sum_{j>l} a_{ij}(x) \right)_{i=1,\dots,l} \\ \mathbf{0} \end{bmatrix}, \quad \Phi^{in}(x) = \begin{bmatrix} \left(\sum_{\substack{j \leq l \\ j \neq i}} a_{ij}(x) \right)_{i=1,\dots,l} \\ \left(\sum_{j \neq i} a_{ij}(x) \right)_{i=l+1,\dots,n} \end{bmatrix}.$$

Here, Φ_i^I is the rate at which new infections occur in the i-th compartment (new infection means that the flow is to be from an N-compartment), $\Phi^{in}(x)$ is the vector of inter-compartmental inflow rates not related to new infections, and $\Phi^{out}(x)$ are the outflow rates from the respective compartments. Thus, $\Phi^C(x)$ represents the totality of all flows not related to new infections. Note that $\Phi_i^I(x) = 0$ for all $i > l$ as there is no inflow of infected into the N-compartments.

The components of the respective vectors $\Phi^{(\cdot)}(x)$ satisfy the following properties (note that (4a) and (4b) can be derived from (2)):

$$\Phi_i^I(x), \Phi_i^{in}(x), \Phi_i^{out}(x) \geq 0, \qquad \forall i = 1, \dots, n \tag{4a}$$

$$\{x_i = 0\} \implies \{\Phi_i^{out}(x) = 0\}, \qquad \forall i = 1, \dots, n \tag{4b}$$

$$\{x_I = \mathbf{0}\} \implies \{\Phi_i^I(x) = 0, \Phi_i^{in}(x) = 0\}, \qquad \forall i = 1, \dots, l. \tag{4c}$$

The latter property implies that no new infections occur in a totally healthy population.

Assumptions. We conclude this section by making a number of epidemiologically motivated assumptions.

A1. There is no steady inflow into the I-compartments, i.e., $w_i = 0$, $i = 1, \dots, l$.

A2. There exists a positive constant μ such that $a_i \geq \mu > 0$ for all $i = 1, \dots, n$. We refer to μ as the *baseline mortality rate*.

2.2. Stability of a disease-free equilibrium

Let $\mathcal{X}_{DF} \subset \mathbb{R}^n_{\geq 0}$ be the set of disease-free states, i.e., $\mathcal{X}_{DF} = \{x \in \mathbb{R}^n_{\geq 0} | x_i = 0, i = 1, \dots, l\}$.

Definition 2.1. Let x^* be an equilibrium state, i.e., $\Phi(x^*) = 0$. Then x^* is said to be a *disease-free equilibrium* (DFE) if $x^* \in \mathcal{X}_{DF}$. Otherwise, x^* is referred to as an *endemic infection equilibrium*.

The system (1) is locally stable at a DFE x^* if the linearized model $\dot{\tilde{x}} = A\tilde{x}$ is asymptotically stable. The latter holds if the spectrum of the structure matrix A contains only the eigenvalues with negative real part, [11]. Such matrices are said to be *Hurwitz*.

Lemma 2.1 ([32])**.** *Let x^* be a DFE. The Jacobian matrices $A^I = D\Phi^I(x)\big|_{x=x^*}$ and $A^C = D\Phi^C(x)\big|_{x=x^*}$ have the following form:*

$$A^I = \begin{bmatrix} A_{11}^I & 0 \\ 0 & 0 \end{bmatrix}, \quad A^C = \begin{bmatrix} A_{11}^C & 0 \\ A_{21}^C & A_{22}^C \end{bmatrix},$$

where A_{11}^I is an $[l \times l]$ matrix with non-negative elements, $A_{11}^I \succeq 0$ and A_{11}^C is an $[l \times l]$ matrix with non-positive diagonal elements and non-negative off-diagonal elements[1].

At this point, we make one more assumption regarding the behavior of the system in the absence of the infection.

A3. A^C is *Hurwitz*.

This assumption effectively implies that the DFE is asymptotically stable provided the virus or whatever source of the infection has lost its contagiousness hence, no new infections occur.

Since the block matrix A_{22}^C is stable at a DFE, stability of the whole system is determined by the sum $A_{11}^I + A_{11}^C$. Following the common convention we write $F = A_{11}^I$ and $V = -A_{11}^C$. Hence, the stability of a DFE is determined by the stability of the matrix $F - V$.

Theorem 2.2 ([32])**.** *The matrix $F - V$ is Hurwitz if and only if $\rho(FV^{-1}) < 1$.*

In Thm. 2.2, $\rho(\cdot)$ denotes the spectral radius of a matrix. Note that for any two square matrices A and B the eigenvalues of AB and BA coincide (see, e.g., [22, Thm. 1.3.22]). Thus, we have $\rho(FV^{-1}) = \rho(V^{-1}F)$.

We conclude this section with the following two definitions:

Definition 2.2. The matrix FV^{-1} is called the *the next generation matrix*.

Definition 2.3. The parameter $R_0 = \rho(FV^{-1})$ is called the *basic reproduction number*.

To conclude this subsection we note that there is no unambiguous definition of which compartments should be considered as infected. For the details, see the discussion in [32] as well as the overview of different issues related to the computations and interpretation of R_0 in [26, Chap. 5]. While not aiming at contributing to this discussion, we mention that one can rule out improper system decomposition into infected and non-infected compartments by inspecting the structure of the linearized system. Should the linearized system have the structure described in Lemma 2.1 and such that Assumption **A3** holds, the used decomposition can be deemed appropriate for applying the NGM-based local stability analysis.

3. Evaluating the Effect of a Treatment-type Control

3.1. A controlled disease propagation model

We wish to model the effect of treatment-type controls, i.e., the controls that result in moving individuals from the infected compartments to other (both infected and healthy) compartments. In doing so, we restrict our attention to the cases where the control action enters the equations linearly and hence write the controlled population balance model in the following form:

$$\dot{x} = \Phi(x) + \Phi^u(x)u, \tag{5}$$

[1] See Appendix B for the explanation of the used notation.

where $u \in \mathbb{R}^m_{\geq 0}$ is the vector of non-negative controls and the components of the control matrix $\Phi^u(x) = [\Phi^u_{ik}(x)]$, $i = 1, \ldots, n$, $k = 1, \ldots, m$, have the following structure:

$$\Phi^u_{\cdot,k}(x) = \begin{bmatrix} \left(\sum\limits_{\substack{j \leq l \\ j \neq i}} a^u_{ij,k}(x) - \sum\limits_{j \neq i} a^u_{ji,k}(x) \right)_{i=1,\ldots,l} \\ \left(\sum\limits_{j \leq l} a^u_{ij,k}(x) \right)_{i=l+1,\ldots,n} \end{bmatrix}.$$

The control matrix describes the controlled flows from the infected compartments to both infected and healthy compartments. This may correspond to isolating infected individuals (corresponds to a transition from an I-compartment to another I-compartment) or treating an infected individual (a flow from an I-compartment to an N-compartment).

The respective controlled flow rates $a^u_{ij,k}(x)$ satisfy the conditions similar to those for the flow rates of the original (uncontrolled) model:

$$a^u_{ji,k}(x) \geq 0, \qquad \forall x \in \mathbb{R}^n_{\geq 0}, i, j = 1, \ldots, n,\ i \neq j,\ k = 1, \ldots, m. \tag{6a}$$

$$\{x_i = 0\} \implies \{a^u_{ji,k}(x) = 0\} \qquad \forall i, j = 1, \ldots, n,\ k = 1, \ldots, m. \tag{6b}$$

Remark 3.1. Note that the model (5) can be equivalently obtained from a more general nonlinear compartmental model of form

$$\dot{x} = \Phi(x, u) \tag{7}$$

by linearizing it w.r.t. u while assuming that the controls are sufficiently small. The transition from (7) to (5) is valid for most cases with a single important exception. The discussed analysis may lead to misleading conclusions if there is an interference between different treatment strategies, i.e., if the rate at which people are administered to treatment i depends on the respective rate associated with treatment j. See, for instance, [2] for an example of the model with an interference between two controls. In such a situation, one would need to fix one of the controls, say, the i-th one and study the influence of another one.

Controllability. Note that at a DFE x^*, the control matrix vanishes identically as there are no outflows from I-compartments: $\Phi^u(x^*) = 0_{[n \times m]}$. This implies that at a DFE x^*, the linearized model is uncontrollable (see, e.g., [3] for more details on controllability and related issues). Thus, one cannot use the linearized model to design a feedback control law stabilizing the DFE.

That is to say, while the stability analysis of a DFE is based upon studying a linearized model, the same approach would not work when designing the control law. Thus a different approach is needed.

3.2. Analysis of the controlled system

Due to the difficulties outlined above, we consider a somewhat simpler, but even more practically relevant situation. Namely, we wish to study the system's dynamics when a

constant control $u^* \in \mathbb{R}^m_{\geq 0}$ is applied. We first formulate the result that extends Lemma 2.1 to include the controls u.

Lemma 3.1. *Let $u(t) = u^* \in \mathbb{R}^m_{\geq 0}$ for all $t \geq 0$. The linearized model of (5) at a DFE x^* has the form*

$$\dot{x} = \left(A^I + A^C + \sum_{k=1}^{m} B_k u_k^* \right) x, \tag{8}$$

where A^I and A^C are defined in Lemma 2.1 and B_k, $k = 1, \ldots, m$ are

$$B_k = \left. \frac{D\Phi^u_{\cdot,k}(x)}{Dx} \right|_{x=x^*} = \begin{bmatrix} B_{11,k} & 0 \\ B_{21,k} & 0 \end{bmatrix}.$$

The matrices $B_{11,k}$ have non-positive diagonal and non-negative off-diagonal elements. Furthermore, $B_{11,k}$ are weakly column diagonally dominant.

Proof. See Appendix A. □

For the sake of notational simplicity and to comply with the previously accepted notation, we will write $W_i = -B_{11,i}$. Matrices W_i belong to the class of Z-matrices (see Appendix B for details). However, these are not M-matrices as matrices W_i are typically rank-deficient and hence non-invertible.

We are interested in the stability of the linearized system (8). As in the uncontrolled case, the structure matrix of (8) is a block-lower triangular matrix, whose eigenvalues coincide with the eigenvalues of the diagonal blocks. The lower right block is Hurwitz according to **A3**. Thus, the stability of (8) is determined by the eigenvalues of the matrix

$$J_u = F - V - \sum_{k=1}^{m} W_k u_k.$$

The following lemma gives an important result that will be used in the sequel.

Lemma 3.2. *Let V be a strictly column diagonally dominant Z-matrix with positive diagonal elements. Then for any $u \in \mathbb{R}^m_{\geq 0}$ the matrix $V + \sum_{k=1}^m W_k u_k$ is an M-matrix.*

Proof. According to Lemma 3.1, matrices $W_i = -B_{11,i}$ have positive diagonal elements and are weakly (non-strictly) column diagonally dominant. Thus for any positive u the matrix $V + \sum_{k=1}^m W_k u_k$ is a Z-matrix and is strictly column diagonally dominant. Then it follows from Theorem 5.2 that $V + \sum_{k=1}^m W_k u_k$ is a non-singular M-matrix. □

Now we are ready to formulate a generalized version of Theorem 2.2

Theorem 3.3. *The matrix $F - V - \sum_{k=1}^m W_k u_k$ is Hurwitz if and only if*

$$\rho\left(F \left(V + \sum_{k=1}^{m} W_k u_k \right)^{-1} \right) < 1.$$

We thus define the controlled reproduction number $R_0(u)$ as the spectral radius of the perturbed matrix $Q(u) = F\left(V + \sum_{k=1}^{m} W_k u_k\right)^{-1}$:

$$R_0(u) = \rho\left(F\left(V + \sum_{k=1}^{m} W_k u_k\right)^{-1}\right). \tag{9}$$

Obviously, we have $R_0(0) = R_0$.

Let the uncontrolled system be such that $R_0 > 1$. Theorem 3.3 allows for determining if a given constant control u^* suffices to shift the value of the basic reproduction number in order to make it less than 1. However, in many cases it is difficult to compute the perturbed reproduction number $R_0(u)$. Thus, we would like to have a result that would tell us if a given structure of the treatment allows for achieving the stated goal.

The following result provides the required estimation. We first consider the case $m = 1$, i.e., we assume that there is a scalar control $u > 0$.

Theorem 3.4. *Let R_0, x_0 and y_0 be the spectral radius of FV^{-1} as well as the right and the left eigenvectors of FV^{-1} corresponding to R_0, respectively. Let, furthermore, there be only one eigenvalue of FV^{-1} coinciding with R_0 and other eigenvalues be strictly less than R_0 in absolute value. For sufficiently small u, the sign of variation $R_0(u) - R_0$ is determined by the sign of*

$$R_1 = -y_0^\top V^{-1} W x_0 R_0 / (y_0^\top x_0) \tag{10}$$

Proof. The eigenvalues of $Q(u) = F\left(V + Wu\right)^{-1}$ change continuously with u. Hence, we can write the spectral radius $R_0(u) = R_0 + uR_1 + O(u^2)$ and the respective eigenvector as $x(u) = x_0 + ux_1 + O(u^2)$. We thus have

$$F(V + uW)^{-1}(x_0 + ux_1 + O(u^2)) = (R_0 + uR_1 + O(u^2))(x_0 + ux_1 + O(u^2)). \tag{11}$$

First, we note that $\frac{d}{du}(V + uW)^{-1} = -(V + uW)^{-1}W(V + uW)^{-1}$ and at $u = 0$ we have $\frac{d}{du}(V + uW)^{-1}\Big|_{u=0} = -V^{-1}WV^{-1}$. Differentiating the left and the right sides of (11) and setting $u = 0$ we get:

$$(FV^{-1} - R_0 I)x_1 = (FV^{-1}WV^{-1} + R_1 I)x_0. \tag{12}$$

The matrix $(FV^{-1} - R_0 I)$ has a zero eigenvalue and the respective left eigenvector is y_0, i.e., $y_0^\top(FV^{-1} - R_0 I) = 0$. Multiplying both sides of (12) with $y_0^\top$ and expressing R_1 we get:

$$R_1 = -y_0^\top W V^{-1} x_0 R_0 / (y_0^\top x_0).$$

Finally, we note that for sufficiently small (and positive) u the sign of the difference $R_0(u) - R_0$ is determined by the sign of R_1, whence the result follows. □

That is to say, we can determine if a given treatment is efficient (at least for small values of u) by checking the sign of R_1. If this sign is negative then by increasing u we decrease $R_0(u)$ and thus reduce the basic reproduction number. Otherwise (if $R_1 > 0$)

we conclude that the treatment program is inadequately formulated and does not lead to a decrease in the reproduction number. See Sec. 4 for examples.

Following the same procedure, one can write $R_0(u)$ for the case when $u \in \mathbb{R}^m$. We have

$$R_0(u^*) = R_0 + \sum_{k=1}^{m} u_k^* R_1^k + O(\|u^*\|^2).$$

If the components of u^* are sufficiently small, the contribution of each individual component is determined by the respective term R_1^k, which is defined as

$$R_1^k = -y_0^\top W_k V^{-1} x_0 R_0 / (y_0^\top x_0). \tag{13}$$

We immediately arrive at the following result.

Theorem 3.5. *For sufficiently small values of the controls u_i, the effect of each control is independent from the values of the remaining controls. The total change of $R_0(u)$ is equal to the sum of individual contributions up to the high order term: $R_0(u) - R_0 = \sum_{k=1}^{m} u_k R_1^k + O(\|u\|^2)$.*

This result indicates the second potential use of R_1. Let there be m treatment strategies that aim at decreasing the value of $R_0(u)$. Then the most efficient strategy is the one with the smallest (and necessarily negative) value of the parameter R_1^k.

Discussion. Note that the structure of the expression for R_1 does resemble that one for R_0 and it can be interpreted as follows. First, following [32], we interpret the (i, j) entry of V^{-1} as the average duration of time an individual introduced into the j-th compartment spends in the i-th one, assuming there is no reinfection and no control. The (i, i) entry of $(-W_k)$ is the rate at which infected individuals are removed from the i-th compartment, while the (i, j)-th component, $i \neq j$, corresponds to the rate at which the infected from i-th compartment are moved to the j-th compartment. The respective rates are multiplied by the control u_k.

Hence, the (i, j)-th element of $-WV^{-1}$ is the relative treatment-induced rate of flow (outflow if negative or inflow if positive) from the i-th compartment as applied to the individuals initially introduced into the j-th compartment. Note that $\sum_i (-WV^{-1})_{i,j} = 0$ if the treatment consists in redistributing the infected individuals between infected compartments and is negative if at least a part of infected individuals are removed to the susceptible compartments as a result of treatment.

To interpret the right eigenvector of FV^{-1}, corresponding to R_0, recall that $FV^{-1}x_0$ is the expected number of new infections produced by the initial distribution of infected individuals given by x_0. Hence, the right eigenvector x_0 can be interpreted as the worst case distribution of the infected. In contrast to that, the left eigenvector y_0 can be seen as a worst case transmissibility rates, i.e., the transmissibility rates that result in the maximal infection spread in the population taking into account the existing structure of the transmission routes. Note that the eigenvectors are defined up to a positive factor and hence should be seen as proportions rather than absolute values.

Finally, the expression $-y_0^\top WV^{-1}x_0(y_0^\top x_0)^{-1}$ can be interpreted as the total flow of the infection due to the treatment computed for the worst case scenario. The term $(y_0^\top x_0)^{-1}$ serves as a normalizing factor. If the total flow is negative, the treatment

leads to a decrease in infection and to an increase otherwise. Note that R_1 is negative if $\sum_i (-WV^{-1})_{i,j} < 0$ for all j.

Remark 3.2. There are two main advantages of considering R_1, resp., R_1^k in contrast to working directly with $R_0(u)$ as defined in (9):

1. The function $R_0(u)$ depends on u in a highly nonlinear way, making it difficult if ever possible to analyze the impact of the control on the controlled reproduction number $R_0(u)$. In contrast, R_1 is a well-defined quantity that does not depend on u thus making it more amenable for analysis.
2. In general, the matrix $\left(V + \sum_{k=1}^m W_k u_k\right)^{-1} F$ has a complex structure which substantially complicates the problem of finding its spectral radius in an analytic form. In contrast to that, R_1^k can be computed with much less effort. Also, we note that the vectors x_0 and y_0 need to be computed only once. The modeler can then vary the structure of Φ^u (and, respectively, W_k) to achieve the required effect in R_1^k.

Remark 3.3. Many public health measures actually involve relatively small control rates. For example PrEP was approved for use in the US in 2012, yet only about 10% of persons that met the CDC eligibility guidelines had been offered a prescription by 2015, [34]. Generally, in large populations, new public health measures will be adopted slowly as new therapies are often expensive and not widely available.

3.3. Further extensions

The described method is substantially based upon the eigenvalue perturbation theory. Namely, we analyze the behavior of the largest real eigenvalue (spectral radius) of a next generation matrix under certain structured perturbations. All the presented results are of local nature, i.e., these are valid for sufficiently small values of u. However, in many cases it is possible to extend this analysis to larger values of u. Below, we consider several possible scenarios and consider possible options and potential limitations.

The simplest case is when the rank of the matrix F is equal to 1. This corresponds to the situation when all immediately infected individuals enter the same compartment, but the infection itself can be caused by a contact with different types of infected. In this case, there is only one non-zero positive eigenvalue coinciding with the spectral radius. The only possible limitation corresponding to this situation is that the controlled reproduction number can change its behavior when the control u grows sufficiently large. One could thus wish to study the behavior of $R_0(u)$ in some more detail. To do so one can compute the second term in the series expansion which we denote by R_2. The corresponding expression is presented in Appendix A, Eq. (26). However, determination of R_2 is a rather cumbersome procedure involving computing Moore-Penrose pseudo inverse. On the other hand, it seems that most cases the sign of R_1 describes the behavior of $R_0(u)$ for arbitrary large values of u and hence can be used for making global predictions.

When the rank of F is larger than 1, the matrix $(V + uW)^{-1}F$ has in general more than one nonzero eigenvalue. A model of co-infection and two types of treatment, one for each type of infection, is a typical example of such a situation. In this case, there are two nonzero eigenvalues $\lambda_1(u_1)$ and $\lambda_2(u_2)$ controlled by the respective treatments u_1 and u_2. Suppose that for $u_1 = u_2 = 0$, $R_0(0) = \lambda_1(0) > \lambda_2(0)$. Hence, the optimal strategy (in terms of minimizing $R_0(u)$) would be to invest into u_1. However, as $\lambda_1(u_1)$ decreases, it will at some point equate with $\lambda_2(u_2)$. Further investment into u_1 will lead

to the change of the roles: $\lambda_2(u_2)$ will become greater than $\lambda_1(u_1)$ and hence will account for $R_0(u)$.

In general, the behavior of the eigenvalues can be predicted by analyzing the structure of $Q(u) = (V+uW)^{-1}F$. We assume here that this structure does not change with u, that is, $Q_{ij}(u) = 0, u > 0 \Rightarrow Q_{ij}(u) = 0, \forall u > 0$ and $Q_{ij}(u) \neq 0, u > 0 \Rightarrow Q_{ij}(u) \neq 0, \forall u > 0$. In this case there are two possible situations: the matrix $Q(u)$ can be either irreducible or reducible. The former implies that there is a simple real eigenvalue corresponding to the spectral radius and hence to R_0 (see, e.g., [28, Sec. 8.3]). If the matrix $Q(u)$ is reducible, one can find a permutation matrix P such that $\bar{Q}(u) = PQ(u)P^{-1}$ has a block-diagonal form and the blocks on the main diagonal are irreducible. Since the eigenvalues of the permuted matrix are determined by the diagonal blocks $\bar{Q}_i(u)$ it suffices to track the spectral radii of $\bar{Q}_i(u)$; denote them by $\bar{\lambda}_i(u)$. We have that $R_0(u) = \max_i \left(\bar{\lambda}_i(u)\right)$. If for some u^* it happens that $\bar{\lambda}_i(u^*) = \bar{\lambda}_j(u^*)$ we have the situation described in the preceding paragraph.

4. Examples

In this section we consider three sufficiently simple but yet non-trivial epidemiological models that are aimed at illustrating the usefulness of the proposed approach.

4.1. An SI model with acute and chronic stages and a single treatment

Model. Consider an SI-model of a disease with two stages (acute and chronic) and a treatment u. This general model structure with sequential infection compartments containing infected persons with different levels of contagiousness describes the evolution of such diseases as HIV [29] and syphilis [15] that have variable levels of contagiousness over the course of infection. For the sake of illustration we explicitly show the Φ^I, Φ^C, and Φ^u components in (14).

$$\frac{d}{dt}\begin{bmatrix} I_A \\ I_C \\ T \\ S \end{bmatrix} = \underbrace{\begin{bmatrix} \alpha(X)S \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{\Phi^I(x)} + \underbrace{\begin{bmatrix} -\beta I_A - \mu_A I_A \\ \beta I_A + \delta T - \mu_C I_C \\ -\delta T - \mu_T T \\ w - \alpha(X)S - \mu_S S \end{bmatrix}}_{\Phi^C(x)} + \underbrace{\begin{bmatrix} 0 \\ -I_C \\ I_C \\ 0 \end{bmatrix}}_{\Phi^u(x)} u \tag{14}$$

where $\alpha(X) = \dfrac{\alpha_A I_A + \alpha_C I_C + \alpha_T T}{N}$, $X = [I_A, I_C, T, S]$, and $N = I_A + I_C + T + S$. The states correspond to the number of acutely infected (I_A), chronically infected (I_C), treated (T), and susceptible (S) individuals. We assume that the first three compartments are infectious with different transmission probabilities: α_A, α_C, and α_T. Furthermore, β is the inverse duration of the acute phase, μ_S, μ_A, μ_C and μ_T are the mortality rates for susceptible, non-treated in acute and chronic phases and treated infected, with $\mu_C > \{\mu_A, \mu_T\} > \mu_S$, and δ is the rate at which the treatment fails.

Analysis. The disease-free equilibrium is unique and given by $x^* = [0, 0, 0, w/\mu_S]$. The (uncontrolled) value of R_0 is computed as the spectral radius of FV^{-1} and is equal to

$$R_0 = \frac{\alpha_C \beta + \alpha_A \mu_C}{(\beta + \mu_A)\,\mu_C}$$

Note that the DFE x^* is unstable if $\alpha_C\beta + \alpha_A\mu_C > (\beta + \mu_S)\mu_C$. We wish to evaluate if the proposed treatment strategy is efficient. To do so we compute R_1 according to (10):

$$R_1 = -\frac{\beta\left(\alpha_C\mu_T - \alpha_T\mu_C\right)}{\mu_C^2\left(\beta + \mu_A\right)\left(\delta + \mu_T\right)} \tag{15}$$

This expression is negative if $\alpha_C\mu > \alpha_T\mu_C$. This can be written as

$$\alpha_C\mu_C^{-1} > \alpha_T\mu_T^{-1}, \tag{16}$$

that is the treatment turns out to be efficient if the product (transmission probability times average residence time in the respective compartment) is lower for the treated individuals as compared with chronically infected. In the following, we will refer to this product as the *cumulative transmissibility* of the infection induced by a given compartment. The condition (16) can be interpreted as follows: If treated people live longer than untreated (but infected), but their infectivity does not reduce enough to compensate for this, such treatment will contribute to the propagation of the disease.

For the considered model it is possible to compute the controlled reproduction number $R_0(u)$:

$$R_0(u) = \frac{\left(\alpha_C\beta + \alpha_A\mu_C\right)\left(\delta + \mu_T\right) + \left(\alpha_T\beta + \alpha_A\mu_T\right)u}{(\beta + \mu_A)(\delta\mu_C + \mu_C\mu_T + \mu_T u)}.$$

One may check that $R_1 = \left.\dfrac{\partial R_0(u)}{\partial u}\right|_{u=0}$. However, we are interested in the behavior of $R_0(u)$ for (relatively) large values of u. To do so, we compute the difference

$$R_0(u) - R_0 = -\frac{\beta u\left(\alpha_C\mu_T - \alpha_T\mu_C\right)}{\mu_C\left(\beta + \mu_A\right)\left(\delta\mu_C + \mu_C\mu_T + \mu_T u\right)}. \tag{17}$$

One can readily observe that the difference is negative for any $u > 0$ when (16) holds.

Figure 1 shows the relative error in the R_1 approximation to $R_0(u)$ for a set of random parameters as a function of R_0. We obtained random parameter sets for a given value of R_0 optimizing from a random starting point for α_A, α_C, and β assuming $\mu_A = \mu_C = 120^{-1}$, $\mu_T = 360^{-1}$, and $\alpha_T = 0$. We rejected parameter sets such that there was no feasible control capable of reducing $R_0(u)$ to 1. The relative error is defined as the difference in the exact value of u^* such that $R_0(u^*) = 1$ and $u^{**} = \frac{R_0 - 1}{R_1}$ normalized by u^*. The approximate value of the control is always less than the true value and the approximation becomes worse as the value of R_0 increases, although the accuracy of the approximation is dependent on the specific parameter values. However, if R_0 is small the approximation gives a reasonable indicator of the level of control required to bring the epidemic to the threshold level.

4.2. An SEIR model with an asymptomatic stage and treatment

Model. Consider the model shown in Fig. 2. This model describes the transmission dynamics of an asymptomatic (sub-clinical) infection. Individuals with asymptomatic infection do not develop the respective symptoms, but are infectious and contribute to the distribution of the disease. Asymptomatic infection has been shown to exist for many diseases, including, e.g., herpes [33], gonorrhea [23], measles [1], and common cold.

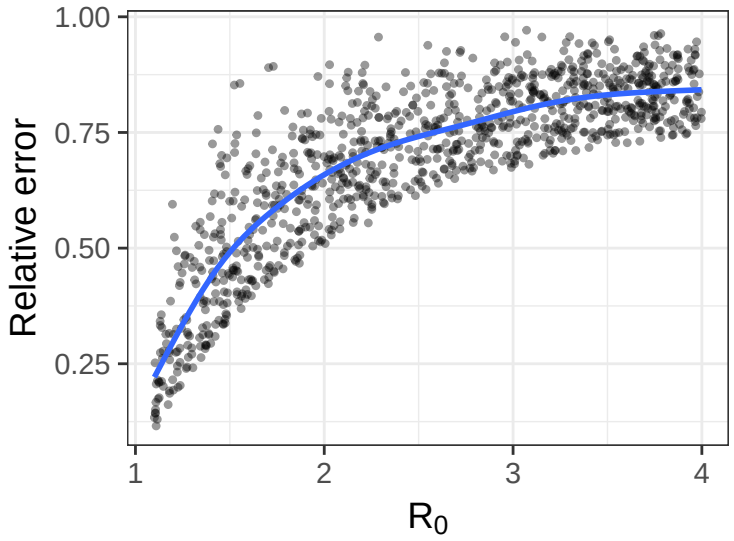

Figure 1: Relative error in approximate reduction in R_0 due to treatment. This figure shows the multiplicative-scale error defined as $\frac{u^* - u^{**}}{u^*}$ where u^* such that $R_0(u^*) = 1$ and $u^{**} = \frac{1-R_0}{R_1}$ for a set of randomly selected parameter values. Parameters with a given R_0 between 1.1 and 4 were selected at random as described in the text. The blue line is a local polynomial regression showing the general trend.

Here, we have the following compartments: E – exposed, I – infected, A – asymptomatically infected, T – treated. Λ is the inflow into the population, η is the transmission rate, p_1 is the probability that the exposition will lead to the typical clinical course and p_2 is the probability that the failed treatment will result in the typical clinical course. Furthermore, x is the rate at which the treatment is failed or canceled, and γ_I, γ_A and γ_T are the inverses of the mean residence time in respective compartments. We assume that only the symptomatic infected, i.e., I, are treated. However, the disease can be transmitted both by symptomatic and asymptomatic infected though with different transmission rates α_I and α_A. People under treatment are assumed to be non-infectious.

Its behavior is described by the following DEs:

$$\begin{cases} \dot{E} = \alpha(X)S - (\eta + \mu)E \\ \dot{A} = -(\gamma_A + \mu)A + \eta(1 - p_1)E + x(1 - p_2)T \\ \dot{I} = \eta p_1 E - (\gamma_I + \mu)\, I - uI + p_2 xT \\ \dot{T} = uI - (\gamma_T + \mu + x)T \\ \dot{S} = \Lambda - \mu S - \alpha(X)S \\ \dot{R} = \gamma_A A + \gamma_I I + \gamma_T T - \mu R, \end{cases}$$

where $\alpha(X) = \dfrac{\alpha_A A + \alpha_I I}{N}$.

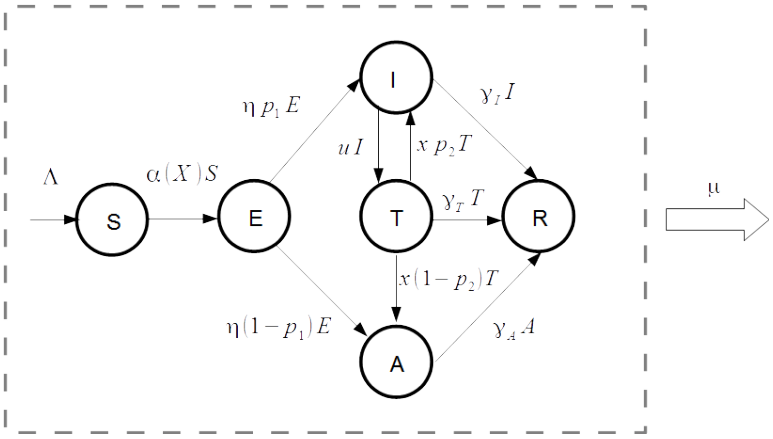

Figure 2: SEIR model.

Analysis. The basic reproduction number R_0 is found using the NGM approach to be

$$R_0 = \frac{\eta}{\eta+\mu}\left[\frac{\alpha_A}{\gamma_A+\mu}(1-p_1)+\frac{\alpha_I}{\gamma_I+\mu}p_1\right],$$

where the expression in square brackets is the expected value of the total cumulative transmissibility of the infection and the quantity in front of the brackets is the fraction of the infected that leave E toward either I or A.

We compute R_1

$$R_1 = -\frac{\eta p_1\left[\alpha_I(\gamma_A+\mu)(\gamma_T+\mu)+(\alpha_I(\gamma_A+\mu)-\alpha_A(\gamma_I+\mu))\,(1-p_2)x\right]}{(\eta+\mu)\,(\gamma_A+\mu)\,(\gamma_I+\mu)^2\,(\gamma_T+\mu+x)} \tag{18}$$

which turns out to be negative if the numerator is positive (note that the denominator is always positive).

Upon some algebraic manipulations, this condition can be written as

$$\frac{\alpha_I}{\gamma_I+\mu} > \frac{x(1-p_2)}{\gamma_T+\mu}\left[\frac{\alpha_A}{\gamma_A+\mu}-\frac{\alpha_I}{\gamma_I+\mu}\right]. \tag{19}$$

Define the cumulative transmissibilities of the A and I stages as $\zeta_A = \frac{\alpha_A}{\gamma_A+\mu}$ and $\zeta_I = \frac{\alpha_I}{\gamma_I+\mu}$. With this, the condition (19) can be rewritten as

$$\frac{\zeta_A-\zeta_I}{\zeta_A}\cdot x(1-p_2) < \gamma_T+\mu,$$

provided $\zeta_A > \zeta_I$ (otherwise the condition (19) holds trivially as the r.h.s. of (19) turns to be negative). The first term on the left side is the relative difference in cumulative transmissibility and the second one is the flow from T to A. Finally, the expression on the right side describes the flow from T to R or to outside the system. This result underscores the importance of the treatment efficiency for the successful eradication of

the disease.

We proceed by considering the total change of R_0 due to the control u

$$R_0(u) - R_0 = \\ - \frac{\eta p_1 u \left[\alpha_I(\gamma_A+\mu)(\gamma_T+\mu) + (\alpha_I(\gamma_A+\mu) - \alpha_A(\gamma_I+\mu))\,(1-p_2)x\right]}{(\eta+\mu)(\gamma_A+\mu)(\gamma_I+\mu)\left[(\gamma_I+\mu)(\gamma_T+\mu+x) + (\gamma_T+\mu+x(1-p_2))u\right]}. \quad (20)$$

The denominator of (20) is always positive so, the condition for the treatment to be efficient (that is the condition for the difference to be negative) is that the numerator is positive. But the numerator of (20) coincides with the numerator of R_1, (18), up to the control u which is positive. Hence we conclude that satisfaction of the condition (19) guarantees decrease in R_0 for any positive value of u, i.e., gives a global condition.

4.3. An SI model with high- and low-risk groups and two treatments

Model. Consider a simplified model of HIV transmission dynamics with two controls corresponding to treating infected individuals from the high- and the low-risk groups, denoted by u_T and u_L. For the details on the derivation of the model see [20]. The transmission dynamics is described by the following set of ODEs

$$\begin{cases} \dot{I}_H = -I_H\mu - I_H f_L\omega + I_L f_H\omega + \phi_H(X)S_H - I_H u_H - \mu_I I_H \\ \dot{I}_L = -I_L\mu + I_H f_L\omega - I_L f_H\omega + \phi_L(X)S_L - I_L u_L - \mu_I I_L \\ \dot{T}_H = T_L f_H\omega - T_H\mu - T_H f_L\omega + I_H u_H \\ \dot{T}_L = T_H f_L\omega - T_L\mu - T_L f_H\omega + I_L u_L \\ \dot{S}_H = \alpha_H - S_H\mu - S_H f_L\omega + S_L f_H\omega - \phi_H(X)S_H \\ \dot{S}_L = \alpha_L - S_L\mu + S_H f_L\omega - S_L f_H\omega - \phi_L(X)S_L, \end{cases} \quad (21)$$

where $S_{\{H,L\}}$, $I_{\{H,L\}}$, and $T_{\{H,L\}}$ are the susceptible, infected, and treated. The subscript denotes the behavioral pattern of the respective group: there are a (H)igh and a (L)ow risk group. Further, $\phi_H(X) = \beta\lambda_H \dfrac{\lambda_H I_H + \lambda_L I_L}{\lambda_H N_H + \lambda_L N_L}$ and $\phi_L(X) = \beta\lambda_L \dfrac{\lambda_H I_H + \lambda_L I_L}{\lambda_H N_H + \lambda_L N_L}$ are the per-capita transmission rates. μ is the mortality rate and μ_I is the disease-induced mortality; $\rho_H = f_H\omega$ and $\rho_L = (1-f_H)\omega = f_L\omega$ are the transition rates between high- and low-risk groups with ω denoting the volatility coefficient; λ_H and λ_L are the contact rates; β is the infection transmissibility. Finally, the inflow rates are $\alpha_H = f_H\mu N$ and $\alpha_L = f_L\mu N$.

For the considered model, the disease free equilibrium is $[I_H^*, I_L^*, T_H^*, T_L^*, S_H^*, S_L^*] = [0, 0, 0, 0, f_H N, f_L N]$. Note that f_H and f_L are the fractions of the respective (high- or low-risk) population at the DFE. Using the next generation matrix method we can compute R_0:

Analysis. The basic reproduction number is computed using the NGM method:

$$R_0 = \frac{\beta}{(\mu_I+\mu+\omega)} C_V(\lambda) + \frac{\beta}{(\mu_I+\mu)}[\lambda]$$

where $[\lambda] = f_L\lambda_L + f_H\lambda_H$ and $[\lambda^2] = \lambda_L^2 f_L + \lambda_H^2 f_H$ are the first and the second moments of the contact rate at the DFE, $\mathrm{Var}(\lambda)$ is the variance of the contact rate, and C_V is the

coefficient of variance defined as $C_V(\lambda) = \mathrm{Var}(\lambda) \,/\, [\lambda]$. Since there are two controls, we would like to compare their contributions in order to decide which one should be invested into. To compute the corresponding components R_1^H and R_1^L we use (13) to get

$$R_1^H = -\frac{\beta f_H(\lambda_H(\mu_I+\mu)+[\lambda]\omega)^2}{(\mu_I+\mu)^2\,[\lambda](\mu_I+\mu+\omega)^2}$$

$$R_1^L = -\frac{\beta f_L\,(\lambda_L(\mu_I+\mu)+[\lambda]\omega)^2}{(\mu_I+\mu)^2[\lambda](\mu_I+\mu+\omega)^2}$$

The first observation is that both R_1^H and R_1^L are negative thus, they contribute to reducing $R_0(u)$ for any choice of parameters. After some algebraic manipulations we find that u_H is more efficient than u_L if

$$f_H\left(\frac{\lambda_H}{[\lambda]}+\frac{\omega}{\mu}\right)^2 > f_L\left(\frac{\lambda_L}{[\lambda]}+\frac{\omega}{\mu}\right)^2$$

Figure 3 shows the value of λ_H such that the two terms in the inequality above are equal. To account for the variability in the duration of the high and low-risk periods as a function of f_H, we introduce the normalized volatility coefficient

$$\omega^* = \frac{\omega\mu^{-1}}{{f_H}^{-1}+(1-f_H)^{-1}},$$

which is the number of full high and low-risk episodes that can be contained in a typical infectious period. This plot shows that for fixed value of f_H, behavioral volatility makes high-risk intervention more plausible (i.e. the high-risk population does not have to be extremely high-risk to make a targeted intervention efficient). This simple analysis gives us a clear theoretical prescription for when to focus on high-risk group based on measurable aspects of the transmission system.

5. Discussion

The results presented in this paper can be applied to a wide class of epidemiological models as long as their dynamics can be described by a compartmental system of form (5). Our approach builds upon and further develops the next generation matrix method in that it allows one to estimate the influence of the treatment-type control(s) on the basic reproduction number R_0 which defines the ultimate condition for eventual elimination of a disease. The more a given control reduces R_0, the closer the system is to elimination and the more effective future interventions will be. Furthermore, it may turn out that in complex models an intervention could unintentionally make things worse for certain populations. The basic premise of medicine to do no harm applies to public health as well. However, the complex, non-linear dynamics of transmission limit the ability of our intuitions to predict the effects of an intervention. Likewise, measurement of the effects of interventions are often very noisy and can have long time lags. Both weak measurability of outcomes and hard to predict dynamics highlight the need for stronger theoretical guarantees that an intervention will not cause population-level harm. A

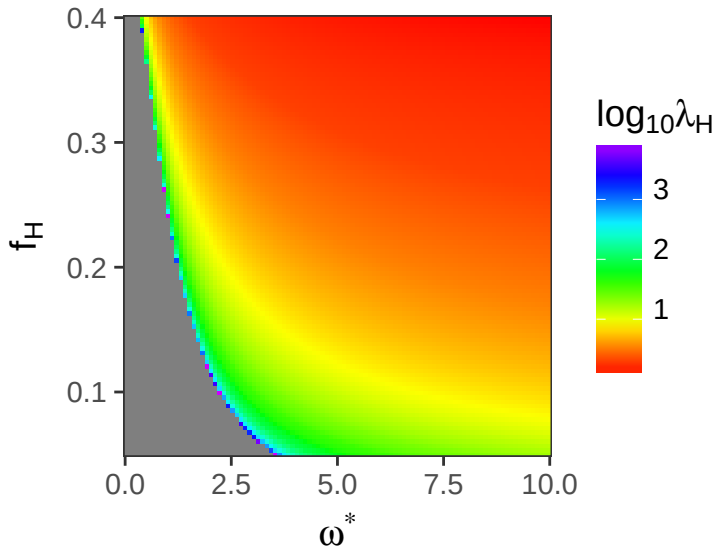

Figure 3: Minimum value of the high-risk contact rate, λ_H, for which the high-risk intervention is preferred. The color shows the value of λ_H for which the terms in the above inequality are equal. For any value of λ_H higher than the plotted value the high-risk intervention is preferred. The gray color indicates that the value is either above 10^4 or there is no value such that the terms are equal. The x-axis is the normalized volatility coefficient ω^*. The y-axis shows the fraction of the population that is high-risk in the absence of disease. The remaining parameters are $\lambda_L = 1$ and $\mu_I = 0$.

possible extension to this work could include consideration of complex models of how risk behavior changes in response to changing prevalence and incidence of disease.

All our results are of local nature, i.e., these are valid for sufficiently small values of u. However, in many cases it is possible to extend this analysis to larger values of u as was illustrated by examples It should also be noted that the proposed approach does inherit all the limitations associated with the NGM method. For instance, it provides only a local stability condition and does not allow to make a conclusion about the behavior of the system under large deviations.

Acknowledgment

Research presented in this article was supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory under project number 20180612ECR. Dmitry Gromov acknowledges the travel grant COLLAB 2018 from St. Petersburg State University.

Appendix A. Proofs and computations

Proof of Lemma 3.1. One can readily observe that $\dfrac{\partial \Phi^u_{i,k}(x)}{\partial x_j} = 0$ for all $j > l$ as $\Phi^u_{i,k}(x^*)$ turns to zero identically for any $x^* = \begin{bmatrix}\mathbf{0}_l & x^C\end{bmatrix}^\top$. We thus consider the partial derivatives

of $\Phi^u_{i,k}(x^*)$ w.r.t. x_j for $j = 1, \ldots, l$.

Let $i, j \in \{1, \ldots, l\}$. We have

$$\frac{\partial \Phi^u_{i,k}(x)}{\partial x_j} = \lim_{\delta x_j \to 0} \frac{\sum\limits_{\substack{q \le l \\ q \neq i}} a^u_{iq,k}(0, \ldots, \delta x_j, \ldots, 0) - \sum\limits_{q \neq i} a^u_{qi,k}(0, \ldots, \delta x_j, \ldots, 0)}{\delta x_j}$$

$$= \begin{cases} \lim\limits_{\delta x_j \to 0} \dfrac{-\sum\limits_{q \neq i} a^u_{qi,k}(0, \ldots, \delta x_j, \ldots, 0)}{\delta x_j} \le 0, & i = j \\ \lim\limits_{\delta x_j \to 0} \dfrac{a^u_{ij,k}(0, \ldots, \delta x_j, \ldots, 0)}{\delta x_j} \ge 0 & i \neq j \end{cases}$$

For $i \in \{l+1, \ldots, n\}$ and $j \in \{1, \ldots, l\}$, the partial derivatives are

$$\frac{\partial \Phi^u_{i,k}(x)}{\partial x_j} = \lim_{\delta x_j \to 0} \frac{a^u_{ij,k}(0, \ldots, \delta x_j, \ldots, 0)}{\delta x_j} \ge 0,$$

which yields the required sign structure. Summation over j gives $\sum_{j=1}^{n} \dfrac{\partial \Phi^u_{i,k}(x)}{\partial x_j} = 0$ for all $i = 1, \ldots, l$ and hence, we have (note that the summation is performed only for $j \le l$)

$$\left| \frac{\partial \Phi^u_{i,k}(x)}{\partial x_j} \right| \ge \sum_{\substack{j \le l \\ j \neq i}} \left| \frac{\partial \Phi^u_{i,k}(x)}{\partial x_j} \right|. \tag{22}$$

This implies that the matrix $B_{11,k} = \left[\dfrac{\partial \Phi^u_{i,k}(x)}{\partial x_j} \right]_{\substack{i=1,\ldots,l \\ j=1,\ldots,l}}$ is weakly column diagonally dominant as the inequality in (22) is not strict. □

Computation of the second order term in the expansion of $R_0(u)$. Let $A(u)$ be a matrix depending on u, r_0 be the simple eigenvalue equal to the spectral radius of $A(0)$, and w_0 and v_0 be the left and the right eigenvectors corresponding to r_0. We expand the perturbed eigenvalue $r(u)$ and the corresponding right eigenvector $x(u)$ in a Taylor series and keep the terms up to the second order: $r(u) = r_0 + r_1 u + r_2 u^2 + O(u^3)$ and $x(u) = v_0 + v_1 u + v_2 u^2 + O(u^3)$. Thus we have

$$A(u)(v_0 + v_1 u + v_2 u^2 + O(u^3)) = (v_0 + v_1 u + v_2 u^2 + O(u^3))(r_0 + r_1 u + r_2 u^2 + O(u^3)) \tag{23}$$

Differentiating (23) w.r.t. u and evaluating at $u = 0$ we get

$$(I r_0 - A(0)) v_1 = (A'(0) - I r_1) v_0, \tag{24}$$

whence the expression for r_1 can be obtained: $r_1 = w_0^\top A'(0) v_0 (w_0^\top v_0)^{-1}$ (cf. the proof of Thm. 3.4). Substituting r_1 back to (24) one gets an expression that can be used to

determine v_1 (see [25, Chap. 8] for details):

$$v_1 = (Ir_0 - A(0))^\dagger \left(I - \frac{v_0 w_0^\top}{w_0^\top v_0} \right) A'(0) v_0, \tag{25}$$

where $()^\dagger$ is the Moor-Penrose inverse operator.

To compute the second term in the expansion of $r(u)$ we differentiate (23) twice w.r.t. u and evaluate at $u = 0$ to get

$$\frac{1}{2} A''(0) v_0 + A'(0) v_1 - v_1 r_1 - v_0 r_2 = (Ir_0 - A(0)) v_2$$

Multiplying from the left by $w_0^\top$ and substituting the previously obtained expressions for r_1 and v_1 we arrive after some computations to the final expression for r_2:

$$r_2 = \left(w_0^\top v_0 \right)^{-1} w_0^\top \left[\frac{1}{2} A''(0) + A'(0) P_0 (Ir_0 - A(0))^\dagger P_0 A'(0) \right] v_0, \tag{26}$$

where $P_0 = I - \frac{v_0 w_0^\top}{w_0^\top v_0}$ is the oblique projection operator. Finally, we recall that $A(u) = (V + uW)^{-1} F$, whence $A'(0) = V^{-1} W V^{-1} F$ and $A''(0) = -2V^{-2} W V^{-2} F$.

Appendix B. Special classes of matrices and their properties.

Definition 5.1. A matrix $B \in \mathbb{R}^{n \times n}$ is *non-negative*, denoted by $B \succeq 0$, if $b_{ij} \geq 0$ for all $i, j = 1, \ldots, n$.

Definition 5.2. A non-singular matrix A is said to be *inverse positive* if it satisfies $\mathbb{R}^n_{\geq 0} \subseteq A\mathbb{R}^n_{\geq 0}$, which is equivalent to $A^{-1} \succeq 0$, i.e., $A^{-1}\mathbb{R}^n_{\geq 0} \subseteq \mathbb{R}^n_{\geq 0}$.

Theorem 5.1. *Let $A = \alpha I - B$, where $\alpha > 0$ and $B \succeq 0$. Then the following statements are equivalent:*

1. *The matrix A is inverse positive,*
2. *The spectral radius of B is strictly smaller than α,*
3. *The matrix A is positive stable, i.e., if λ is an eigenvalue of A, then $\Re(\lambda) > 0$.*

Definition 5.3. A matrix $A = \alpha I - B$ satisfying any of the properties of Thm. 5.1 is said to be an *M-matrix*.

We will occasionally write $\mathbf{M}$ to denote the class of all n-by-n non-singular M-matrices. The preceding results can be generalized in the following way. Let us define the class of *Z-matrices* as $\mathbf{Z} = \{A \in \mathbb{R}^{n \times n} | a_{ij} \leq 0, i \neq j\}$. The following theorem gives a number of conditions which guarantee that a given Z-matrix is a non-singular M-matrix. For a complete list see [5].

Theorem 5.2. *Let $A \in \mathbf{Z}$. Any of the following conditions implies $A \in \mathbf{M}$.*

1. *A is inverse-positive.*
2. *A is positive stable.*

3. *A has all positive diagonal elements and is* strictly row diagonally dominant (d.d.)*, i.e.,*

$$a_{ii} > \sum_{i \neq j} |a_{ij}|, \quad i = 1, \ldots, n.$$

4. *A has all positive diagonal elements and is* strictly column d.d.*, i.e.,*

$$a_{ii} > \sum_{i \neq j} |a_{ji}|, \quad i = 1, \ldots, n.$$

Proof. We will prove only the last item as the remaining ones are covered in [5].

Let $A \in \mathbf{Z}$ and A be strictly column d.d., then $A^\top \in \mathbf{Z}$ and is strictly row d.d. This implies $A^\top \in \mathbf{M}$. Since the spectrum of A coincides with that of $A^\top$, the positive stability property holds for A and hence $A \in \mathbf{M}$. □

References

[1] B Anlar, A Ayhan, H Hotta, M Itoh, D Engin, S Barun, and Ö Köseoglu. Measles virus RNA in tonsils of asymptomatic children. *Journal of paediatrics and child health*, 38(4):424–425, 2002.

[2] Benjamin Armbruster and Margaret L Brandeau. Cost-effective control of chronic viral diseases: finding the optimal level of screening and contact tracing. *Mathematical biosciences*, 224(1):35–42, 2010.

[3] Karl Johan Aström and Richard M. Murray. *Feedback systems: an introduction for scientists and engineers*. Princeton University Press, 2010.

[4] Rebecca F. Baggaley, Neil M. Ferguson, and Geoff P. Garnett. The epidemiological impact of antiretroviral use predicted by mathematical models: a review. *Emerging Themes in Epidemiology*, 2(9), 2005.

[5] Abraham Berman and Robert J. Plemmons. *Nonnegative matrices in the mathematical sciences*. SIAM, 1994.

[6] S. Blower, E. J. Schwartz, and J. Mills. Forecasting the future of HIV epidemics: the impact of antiretroviral therapies & imperfect vaccines. *AIDS reviews*, 5(2):113–125, 2003.

[7] M. C. J. Bootsma and N. M. Ferguson. The effect of public health measures on the 1918 influenza pandemic in U.S. cities. *Proceedings of the National Academy of Sciences*, 104(18):7588–7593, 2007.

[8] Francis J Bowden and Geoffrey P Garnett. Trichomonas vaginalis epidemiology: parameterising and analysing a model of treatment interventions. *Sexually Transmitted Infections*, 76(4):248–256, 2000.

[9] Carlos Castillo-Chavez and Zhilan Feng. To treat or not to treat: the case of tuberculosis. *Journal of mathematical biology*, 35(6):629–656, 1997.

[10] Carlos Castillo-Chavez, Zhilan Feng, and Wenzhang Huang. On the computation of R_0 and its role on global stability. In D. Bies, S. Blower, P. van den Driessche, D. Kirschner, and A.-A. Yakubu, editors, *Mathematical approaches for emerging and reemerging infectious diseases: an introduction*, volume 125 of *The IMA Volumes in Mathematics and its Applications*, pages 229–250. Springer, 2002.

[11] Earl A. Coddington and Norman Levinson. *Theory of ordinary differential equations*. Tata McGraw-Hill Education, 1955.

[12] Odo Diekmann, Hans Heesterbeek, and Tom Britton. *Mathematical tools for understanding infectious disease dynamics*. Princeton University Press, 2012.

[13] Odo Diekmann and Johan Andre Peter Heesterbeek. *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*, volume 5 of *Mathematical and computational biology*. John Wiley & Sons, 2000.

[14] Odo Diekmann, Johan Andre Peter Heesterbeek, and Johan A.J. Metz. On the definition and the computation of the basic reproduction ratio r_0 in models for infectious diseases in heterogeneous populations. *Journal of mathematical biology*, 28(4):365–382, 1990.

[15] Patrick French. Syphilis. *BMJ : British Medical Journal*, 334(7585):143–147, January 2007.

[16] Dmitry Gromov, Ingo Bulla, Oana Silvia Serea, and Ethan O. Romero-Severson. Numerical optimal control for HIV prevention with dynamic budget allocation. *Mathematical Medicine and Biology: A Journal of the IMA*, page dqx015, 2017.

[17] J.A.P. Heesterbeek. A brief history of R_0 and a recipe for its calculation. *Acta Biotheoretica*, 50(3):189–204, Sep 2002.

[18] J.A.P. Heesterbeek and K. Dietz. The concept of R_0 in epidemic theory. *Statistica Nederlandica*, 50(1):89–110, 1996.

[19] J.M. Heffernan, R.J. Smith, and L.M. Wahl. Perspectives on the basic reproductive ratio. *Journal of the Royal Society Interface*, 2(4):281–293, 2005.

[20] Christopher J Henry and James S Koopman. Strong influence of behavioral dynamics on the ability of testing and treating HIV to stop transmission. *Scientific reports*, 5:9467, 2015.

[21] Roger A. Horn and Charles R. Johnson. *Topics in matrix analysis*. Cambridge University Presss, Cambridge, 1991.

[22] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge university press, 2nd edition, 2013.

[23] Eline L Korenromp, Mondastri K Sudaryo, Sake J de Vlas, Ronald H Gray, Nelson K Sewankambo, David Serwadda, Maria J Wawer, and J Dik F Habbema. What proportion of episodes of gonorrhoea and chlamydia becomes symptomatic? *International Journal of STD & AIDS*, 13(2):91–101, 2002.

[24] Eline L Korenromp, Carina Van Vliet, Heiner Grosskurth, Awene Gavyole, Catharina PB Van der Ploeg, Lieve Fransen, Richard J Hayes, and J Dik F Habbema. Model-based evaluation of single-round mass treatment of sexually transmitted diseases for HIV control in a rural African population. *Aids*, 14(5):573–593, 2000.

[25] Jan R. Magnus and Heinz Neudecker. *Matrix differential calculus with applications in statistics and econometrics*. Wiley, 2nd rev. edition, 1999.

[26] Maia Martcheva. *An introduction to mathematical epidemiology*, volume 61 of *Texts in Applied Mathematics*. Springer, 2015.

[27] Natasha K Martin, Peter Vickerman, Jason Grebely, Margaret Hellard, Sharon J Hutchinson, Viviane D Lima, Graham R Foster, John F Dillon, David J Goldberg, Gregory J Dore, et al. Hepatitis C virus treatment for prevention among people who inject drugs: modeling treatment scale-up in the age of direct-acting antivirals. *Hepatology*, 58(5):1598–1609, 2013.

[28] Carl D Meyer. *Matrix analysis and applied linear algebra*. SIAM, 2000.

[29] World Health Organization et al. Interim WHO clinical staging of HIV/AIDS and HIV/AIDS case definitions for surveillance: African region. Technical report, Geneva: World Health Organization, 2005.

[30] Oluwaseun Sharomi, Chandra N. Podder, Abba B. Gumel, and Baojun Song. Mathematical analysis of the transmission dynamics of HIV/TB coinfection in the presence of treatment. *Mathematical Biosciences & Engineering*, 5(1):145–174, 2008.

[31] Cristiana J. Silva and Delfim F. M. Torres. A TB-HIV/AIDS coinfection model and optimal control treatment. *Discrete & Continuous Dynamical Systems - A*, 35(9):4639–4663, 2015.

[32] Pauline Van den Driessche and James Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical biosciences*, 180(1):29–48, 2002.

[33] Anna Wald, Judith Zeh, Stacy Selke, Rhoda L. Ashley, and Lawrence Corey. Virologic characteristics of subclinical and symptomatic genital herpes infections. *New England Journal of Medicine*, 333(12):770–775, 1995.

[34] M. Weinstein, O.O. Yang, and A.C. Cohen. Were we prepared for PrEP? Five years of implementation. *AIDS*, 31(16):23032305, 2017.