11-4-2022

Alternative SIAR models for infectious diseases and applications in the study of non-compliance

Marcelo Bongarti

Luke Diego Galvan

Lawford Hatcher

Michael R. Lindstrom
The University of Texas Rio Grande Valley, mike.lindstrom@utrgv.edu

Christian Parkinson

See next page for additional authors

Recommended Citation

Bongarti, Marcelo, Luke Diego Galvan, Lawford Hatcher, Michael R. Lindstrom, Christian Parkinson, Chuntian Wang, and Andrea L. Bertozzi. "Alternative SIAR models for infectious diseases and applications in the study of non-compliance." Mathematical Models and Methods in Applied Sciences 32, no. 10 (2022): 1987-2015. https://doi.org/10.1142/S0218202522500464

Mathematical Models and Methods in Applied Sciences

Alternative SIAR models for infectious diseases and applications in the study of non-compliance

Marcelo Bongarti

Weierstrass Institute for Applied Analysis and Stochastics
Mohrenstraße 39, 10117, Berlin, Germany
bongarti@wias-berlin.de

Luke Diego Galvan

Department of Mathematics, University of Texas, Austin
2515 Speedway, Austin, TX, USA, 78712
lukediego33@gmail.com

Lawford Hatcher

Department of Mathematics, Indiana University Bloomington
831 East 3rd St., Bloomington, IN, USA, 47405
lhhatche@iu.edu

Michael R. Lindstrom

Department of Mathematics, University of California, Los Angeles
520 Portola Plaza, Los Angeles, CA, USA, 90095
mikel@math.ucla.edu

Christian Parkinson*

Department of Mathematics, University of Arizona
617 N. Santa Rita Ave., Tucson, AZ USA, 85721
chparkin@math.arizona.edu

Chuntian Wang

Department of Mathematics, University of Alabama
505 Hackberry Lane, Tuscaloosa, AL 35401
cwang27@ua.edu

Andrea L. Bertozzi

Department of Mathematics, University of California, Los Angeles
520 Portola Plaza, Los Angeles, CA, USA, 90095
bertozzi@math.ucla.edu

Received (Day Month Year)
Revised (Day Month Year)

*Corresponding Author

Communicated by (xxxxxxxxxx)

In this paper we use modified versions of the SIAR–model for epidemics to propose two ways of understanding and quantifying the effect of non-compliance to non-pharmaceutical intervention measures on the spread of an infectious disease. The SIAR model distinguishes between symptomatic infected (I) and asymptomatic infected (A) populations. One modification, which is simpler, assumes a known proportion of the population does not comply with government mandates such as quaranting and social-distancing. In a more sophisticated approach, the modifed model treats non-compliant behavior as a social contagion. We theoretically explore different scenarios such as the occurrence of multiple waves of infections. Local and asymptotic analyses for both models are also provided.

Keywords: Epidemic models; compartmental models; stability analysis; non-compliant behavior

AMS Subject Classification: 92D30, 92B05, 34D05

1. Introduction

The COVID-19 pandemic led to the implementation of many non–pharmaceutical interventions (NPIs) to delay the rapid spread of the novel coronavirus (SARS-CoV-2). Some of the most common NPIs are mandated social distancing and self-quarantine, but these inevitably lead to non-compliant populations[3,46,55] and have societal and economic impacts. Social scientiests refer to deleterious effects of NPIs on mental health (and subsequent wane in compliance with NPIs) as *quarantine fatigue*[9] and hypothesize that social distancing mandates and all–or–nothing stay–at–home orders are similar to abstinence–only messages for issues like substance abuse and teen pregnancy prevention[36], which have mixed results [49]. The dynaics of non-compliance may spread by way of social contagion [14] in which a given behavior spreads through a population like an epidemic. This approach has been used to model illicit substance use[2] and adolescent sexual behavior[48]. In this paper, we utilize compartmental models as the primary tool for developing between-host epidemic models with non-compliant populations.

We note that between-host compartmental modeling is one of many methods for epidemic modeling being used to understand the COVID-19 pandemic. Disease models can broadly be divided into between-host analysis—those which analyze spread of the disease from one person to another—and in-host analysis—those which analyze the dynamics of the disease and immune response within a person[41,44,51,15,27,47]. Even in the realm of between-host analysis, there are a variety of approaches one may take including compartmental modeling (as in this study and many of the references below), network models [20,43,17,39,52], agent-based models [31,32,34,42], and self-exciting point process models [13,30,24]. Modeling a pandemic with high fidelity would likely require multiscale considerations, so some effort has been devoted to developing combined between-and-within-host models[5,45,12]. However, this study will be focused on between-host compartmental modeling.

1.1. *The SIR Model and $\mathscr{R}_0$*

There is an abundance of literature on epidemic models dating back to the early 20th century. A cornerstone of classical mass–action models is the Susceptible–Infected–Recovered (SIR) model. The model supposes that a disease spreads among a homogeneous and well-mixed population according to [28]:

$$\begin{cases} \dot{S}(t) = -\beta I(t)S(t) \\ \dot{I}(t) = \beta I(t)S(t) - \gamma I(t) \\ \dot{R}(t) = \gamma I(t) \\ \beta > 0 \quad \gamma > 0, \end{cases} \tag{1.1}$$

where $S(t), I(t), R(t)$ represent the susceptible, infected, and removed or recovered populations respectively at time t. The positive parameters β and γ represent the transmission and recovery rates of the disease, respectively. In this study we consider the normalized version with $S(t)+I(t)+R(t) = 1$ for all $t \geqslant 0$. Modifications of the model (1.1) have been formulated to incorporate more information about the underlying disease[1,7,29,40,54]. Additional compartments have been introduced to consider the group of people that are asymptomatic infected[23,38] and quarantined infected individuals[35]. Some researchers have also introduced and studied time–dependent transmission rates and its dynamical impact on the spread of infection[11,26].

An important quantity is the basic reproduction number $\mathscr{R}_0$, *the expected number of secondary infections generated by a typical individual in a completely susceptible population*[19]. For general compartmental models, this quantity can be derived using the *next generation matrix* method[18,19,53]. For the SIR model, $\mathscr{R}_0 = \frac{\beta}{\gamma}$. If $\mathscr{R}_0 > 1$, a small outbreak of infections will cause an epidemic (that is, cause $I(t)$ to increase for small t) whereas if $\mathscr{R}_0 < 1$, a small outbreak of infections will die off without causing an epidemic (that is, $I(t)$ will decrease). Due to its simplicity of both modeling and simulation, compartmental models are widely used in the study of infectious diseases, particularly because they reveal the worst–case scenarios.

1.2. *Herd Immunity and Final Size*

Another important quantity is the *herd immunity* threshold value $S(t_h)$ with corresponding herd immunity time t_h. This is the minimum threshold required for the recovered population to avoid continued growth of infections. In the case of the SIR model, this threshold occurs at any time t_h such that $S(t_h) < \mathscr{R}_0^{-1}$. Observe, however, that for an epidemic following (1.1), it can be shown that $S(t)$ never vanishes and in fact, $S(t) \geqslant S(0)e^{-\mathscr{R}_0}$ for all $t \geqslant 0$.

One may also wish to determine the final size of the susceptible population at the end of an epidemic period[37]. The final susceptible population size S_∞ is given implicitly by

$$S_\infty = S(0)e^{-\mathscr{R}_0(1-S_\infty)}. \tag{1.2}$$

Making the reasonable assertion of $\lim_{t\to\infty} I(t) = 0$, one can equivalently see $R_\infty = 1 - S_\infty$.

Therefore, for the basic epidemiological compartmental model, knowledge of $\mathscr{R}_0$ provides answers to three important questions:

Q1. Will an outbreak of infections cause an epidemic?
Q2. What is the herd immunity threshold $S(t_h)$ with corresponding time t_h?
Q3. What will be the final size of the outbreak?

During the outbreak of COVID-19, there has been interest in incorporating human behavior into epidemiological models, as recent research from public health experts suggests that a nontrivial portion of individuals will not comply with government mandated NPIs and that this has a significant effect on the spread of the disease[16,25,50].

In this paper we introduce two new models that take distinct routes to incorporate non-compliant populations that affect disease spread. The first model assumes that individuals have a given – potentially time–dependent – probability of being compliant to lockdown measures. In our study, the probability is taken to be piecewise constant. From a modeling standpoint, this may be unsatisfying because of the difficulties to accurately measure overall compliance with social distancing measures. Our second model is inspired by social contagion theory. The strategy of social contagion modeling is to treat a given behavior–as an epidemic, with those partaking in the behavior "infecting" others and causing them to behave accordingly. As such, this model essentially has two epidemics running in parallel: one representing the actual disease and another representing non-compliant behavior. We analyze the interaction between these two with special emphasis devoted to the effect of non-compliance on the spread of the disease. The remainder of this paper is organized as follows: in section 2, we introduce the new models. In section 3, we study stability and equilibria with a focus on showing that knowledge of each model's basic reproduction number $\mathscr{R}_0$ provides good local and long–time behavioral properties. In section 4 and section 5, we investigate herd immunity thresholds for both of the models introduced in section 2.

2. The New Models, Simulations and Discussion

In this section we present two new models for SIAR type dynamics that incorporate non-pharmaceutical intervention measures and non-compliant behavior.

2.1. *Constant or piecewise constant probability of compliance: the p–model*

The first model incorporates two important aspects of the COVID-19 pandemic. First, inspired by previous work[23], the model includes a new class of infectious individuals. Here, we split the typical infected compartment I into two compartments:

symptomatic infected I and asymptomatic infected A. We assume that new infections are symptomatic with probability ζ and asymptomatic with probability $1-\zeta$. Keeping the notation of (1.1), we take the natural transmission rate to be β. We denote by γ_1 the recovery rate for symptomatic infected individuals and by γ_2 the recovery rate of asymptomatic infected individuals. We model NPIs as a reduction of interactions quantified via a new parameter $\alpha \in (0,1)$. Specifically, we assume that a portion $(1-\alpha)$ of the compliant population is actively mixing and transmitting the disease. We also assume that a portion $\xi \in (0,1)$ of the symptomatic infected individuals isolate independently of NPIs. Finally, we assume that a given individual (who is not symptomatic infected) complies with NPIs with a probability $p(t)$ at a given instant $t > 0$.

The susceptible population loses members at a rate proportional to the product of the actively mixing susceptible and infected populations. Per the above discussion, the portion of susceptible population that is actively mixing is given by

$$(1-\alpha)pS + (1-p)S = (1-\alpha p)S,$$

and similarly the portion of the infected population that is actively mixing is given by

$$(1-\alpha p)I_M := (1-\alpha p)((1-\xi)I + A).$$

Note that I_M is the portion of the infected population that is actively mixing in the absence of any NPIs.

Thus the system of differential equations for this model, henceforth referred to as the p–model, is given by

$$\begin{cases} \dot{S} &= -\beta(1-\alpha p)^2 I_M S \\ \dot{I} &= \zeta\beta(1-\alpha p)^2 I_M S - \gamma_1 I \\ \dot{A} &= (1-\zeta)\beta(1-\alpha p)^2 I_M S - \gamma_2 A \\ \dot{R} &= \gamma_1 I + \gamma_2 A \end{cases} \tag{2.1}$$

paired with initial conditions $S(0) = S_0 = 1-\varepsilon$, $I(0) = I_0 = \zeta\varepsilon$, $A(0) = A_0 = (1-\zeta)\varepsilon$, for $\varepsilon \in (0,1)$ representing the overall initial infected population, and $R(0) = R_0 = 0$.

2.2. *Simulations and discussion for the p-model*

In the following simulations, we use the parameter values

$$\gamma_1 = 1/14, \quad \gamma_2 = 1/18, \quad \zeta = 0.75, \quad \xi = 0.8, \quad T = 100 \quad \text{and}$$
$$p(t) = \begin{cases} 0.9 & t \leqslant T/2, \\ 0.45 & T/2 < t \leqslant T. \end{cases}$$

Other parameters will vary from one simulation to another. We emphasize that these are synthetic parameter values. While faithfully estimating parameters within

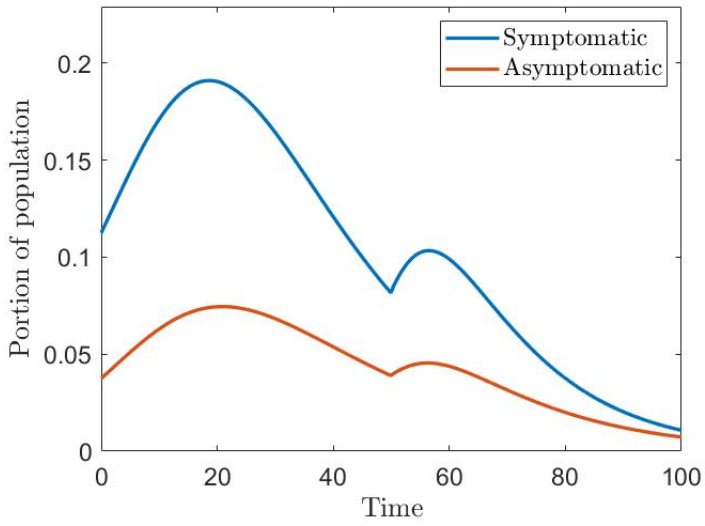

Fig. 1. A sudden decrease in compliance *might* result in a second wave of infections for the p-Model.

epidemic models is an active area of research [56,6,22,10], it is outside the scope of this qualitative study.

An interesting aspect of the p-model is the possibility of several waves of infections with different peaks correlated with a change in the probability of compliance p. Intuitively, it seems that when a large portion of the population is compliant, the implementation of very stringent lockdown measures (e.g., $\alpha = 0.9$) could eradicate the disease before quarantine fatigue becomes significant. The p-model indicates that this may not be the case, as illustrated in Figure 1, where the intensity of the NPIs is $\alpha = 0.9$, the transmission parameter is $\beta = 4$, $\varepsilon = 0.15$ so that 15% of the population is initially infected.

The structure of this model allow us to explore a few scenarios when multiple waves of infections do or do not occur. First, we note that controlling the size of the infected population is only possible if the natural uncontrolled transmission rate is small enough. If β is too large, the initial epidemic will infect a large enough portion of the susceptible population to prevent future epidemics. We illustrate this in Figure 2 where we increase the transmission rate to $\beta = 6$, but hold other parameters constant. Intuitively, if the first wave of infections is large enough, there can be no second wave.

By contrast, if the initial infection is very small and the compliance decreases, a drastically larger second wave of infections can occur, despite the application of severe NPIs. For instance, setting the transmission rate $\beta = 5$, if we assume that initial infection represents only 1% of the entire population ($\varepsilon = 0.01$), the second peak is bigger than the first, as illustrated in Figure 3. This idea is also discussed

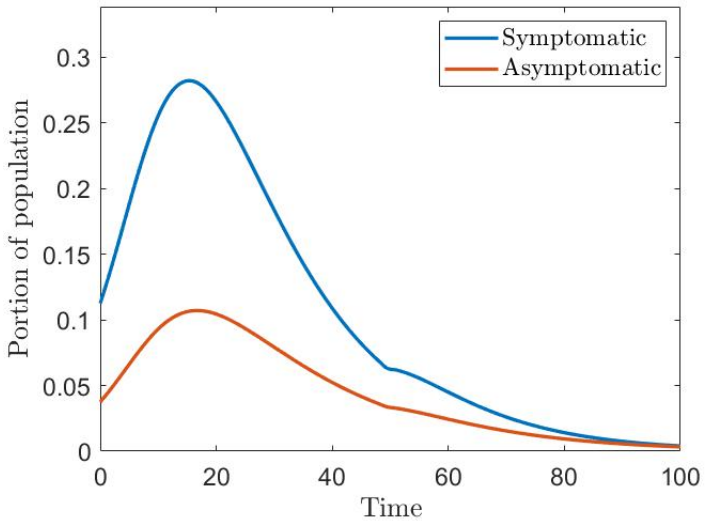

Fig. 2. A high transmission rate can negate the effect of strict lockdowns with a high compliance rate, preventing a smaller second wave even with a drop in compliance.

in some detail in [8] for the standard SIR model with controls.

The possibility of multiple waves of infection will impact the stability analysis

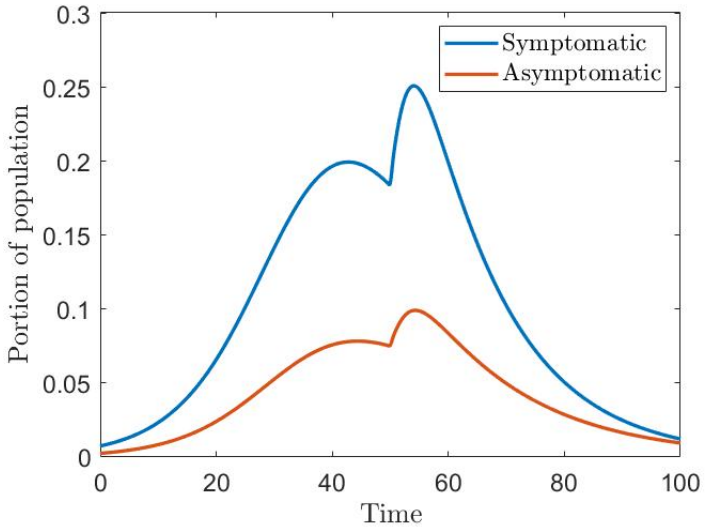

Fig. 3. A decrease in compliance before the susceptible class is sufficiently small has the potential to cause a larger second wave.

of the model. In contrast with the basic SIR-model, the knowledge of the basic reproduction number is unlikely to imply local asymptotic stability of the disease-free equilibrium points, although it is possible to prove that the infected class will eventually be driven to zero monotonically. Questions regarding herd immunity and final sizes also depend critically on the existence of multiple peaks (and therefore on compliant or non-compliant behavior). These dynamics can be further explored by introducing multiple susceptible compartments and considering the underlying fluxes among these classes. By doing so, one observes large perturbations of the basic reproduction number even if the infectious subsystem is near equilibrium. Our second model accounts for two classes of susceptible individuals and allows a parallel (social) contagion corresponding to noncompliance with NPIs that triggers the flux between them. Consequently, a more specific and detailed stability analysis is not only possible, but necessary.

2.3. *Non-compliance as a disease: the NCD–model*

As mentioned earlier, the idea of modeling certain behaviors as a social contagion has precedent in the social sciences[2,14,21]. To include non-compliance as a social contagion, we let superscripts c and nc indicate compliant and non-compliant subpopulations respectively. The variables and compartment names are otherwise the same as in the p–model. Thus we have $1 = N^c + N^{nc}$ where $N^c = S^c + I^c + A^c + R^c$ and $N^{nc} = S^{nc} + I^{nc} + A^{nc} + R^{nc}$. Again assume that a portion $\xi \in (0, 1)$ of the symptomatic infected individuals isolate independently of NPIs and that only a portion $1 - \xi$ will be affected by noncompliance and NPIs. The additional feature is to model non-compliance using mass-action terms with transmission rate $\delta > 0$. In practice, this parameter is very difficult to determine. We do not include "recovery" for the social contagion: once an individual has become non-compliant, we assume they are permanently non-compliant.

We use the same mass-action mixing argument from the previous model. The first thing to consider is the amount of infected and susceptible individuals who are actively mixing. The compliant susceptible population will isolate in compliance with lockdown mandates. Therefore, only a portion $(1 - \alpha)S^c$ will be mixing. On the other hand, non-compliant susceptible individuals do not isolate, so the compartment S^{nc} is not multiplied by the factor $(1 - \alpha)$. Applying the same reasoning for the compartment of asymptomatic infected population we arrive at

$$I_M = (1 - \xi)((1 - \alpha)I^c + I^{nc}) + (1 - \alpha)A^c + A^{nc}. \tag{2.2}$$

Here I_M again indicates the amount of infected individuals that are mixing. Hence, modeling compliance as a social contagion with transmission parameter δ means that the compliant susceptible class will lose members at a rate proportional to the the mass-action term $N^{nc}S^c$. These members transfer to the non-compliant

susceptible compartment. Thus the equations for the susceptible populations are

$$\dot{S^c} = -\beta(1-\alpha)I_M S^c - \delta N^{nc} S^c, \tag{2.3}$$

$$\dot{S^{nc}} = -\beta I_M S^{nc} + \delta N^{nc} S^c \tag{2.4}$$

Once removed from the susceptible categories, an infected individual will be added to the symptomatic subgroup with probability ζ and will recover with rate γ_1, or added to the asymptomatic subgroup and recover with rate γ_2. A similar reasoning as before regarding compliant/non-complaint populations and the flow among them yields the following equations for the infected compartments

$$\dot{I^c} = \zeta\beta(1-\alpha)I_M S^c - \gamma_1 I^c - \delta N^{nc} I^c \tag{2.5}$$

$$\dot{I^{nc}} = \zeta\beta I_M S^{nc} - \gamma_1 I^{nc} + \delta N^{nc} I^c \tag{2.6}$$

$$\dot{A^c} = (1-\zeta)(1-\alpha)\beta I_M S^c - \gamma_2 A^c - \delta N^{nc} A^c \tag{2.7}$$

$$\dot{A^{nc}} = (1-\zeta)\beta I_M S^{nc} - \gamma_2 A^{nc} + \delta N^{nc} A^c \tag{2.8}$$

The outwards flux from the infected compartments goes to the recovered/removed compartments according to their compliance status. Therefore the differential equations for the R-compartments are

$$\dot{R^c} = \gamma_1 I^c + \gamma_2 A^c - \delta N^{nc} R^c \tag{2.9}$$

$$\dot{R^{nc}} = \gamma_1 I^{nc} + \gamma_2 A^{nc} + \delta N^{nc} R^c. \tag{2.10}$$

We pair equations (2.3)-(2.10) with initial conditions

$$\begin{array}{ll} S^c(0) = S_0^c = (1-\varepsilon)(1-\sigma), & S^{nc}(0) = S_0^{nc} = (1-\varepsilon)\sigma, \\ I^c(0) = I_0^c = \zeta(1-\sigma)\varepsilon, & I^{nc}(0) = I_0^{nc} = \zeta\sigma\varepsilon, \\ A^c(0) = A_0^c = (1-\zeta)(1-\sigma)\varepsilon, & A^{nc}(0) = A_0^{nc} = (1-\zeta)\sigma\varepsilon, \\ R^c(0) = R_0 = 0, & R^{nc}(0) = R_0^{nc} = 0, \end{array}$$

where $\sigma, \varepsilon \in (0,1)$ represent the initial portion of non-compliant individuals and the initial portion of infected individuals, respectively. Henceforth, we refer to this as the Non-Compliance as a Disease (NCD) model.

The idea of splitting the non-compliant population into separate compartments is relatively novel. Peng et al.[43] propose a similar model rooted in network theory, and Barbarossa and Fuhrmann[4] simulate and discuss a similar compartmental model, but do not consider flow between compliant and non-compliant populations. To the authors' knowledge, the analysis of second waves of infections, herd immunity, and final infection size—and specifically their dependence on the human behavior parameters δ and σ—is novel.

2.4. *Basic simulations and discussion for the NCD–model*

Again, we set

$$\gamma_1 = 1/14,\ \gamma_2 = 1/18,\ \zeta = 0.75,\ \xi = 0.8, \text{and } \alpha = 0.9$$

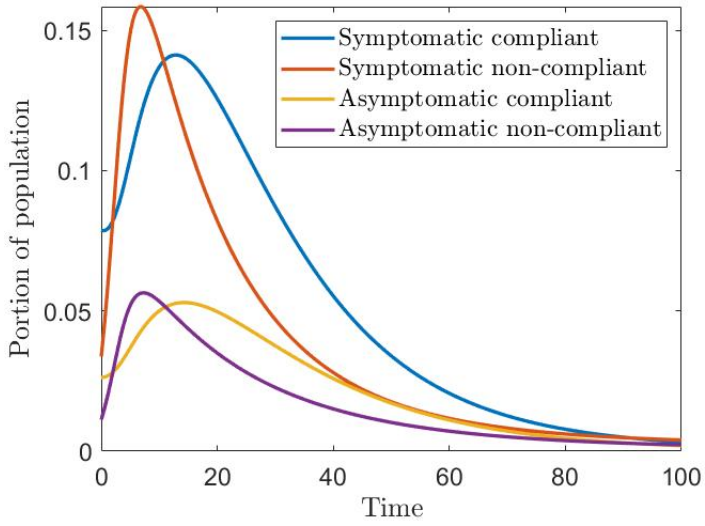

Fig. 4. Sample simulation for the NCD model taking $\delta = 0.01$, $\sigma = 0.3$, and $\beta = 5$.

and vary some other parameters as discussed below. We mention once again that these are hypothetical parameter values used to carry out a qualitative study of the behavior of the model.

Our first sample simulation includes a significant portion of non-compliant individuals. In Figure 4, we take the non-compliance transmission rate to be $\delta = 0.01$ and the initial non-compliance rate $\sigma = 0.3$, so that 30% of the population is non-compliant. We use a disease transmission rate of $\beta = 5$ and the initial total infected population of 15% ($\varepsilon = 0.15$). As the non-compliant portion of the population mixes at a higher rate than the compliant population, the non-compliant populations see earlier and larger infection peaks than the compliant population.

The simulation provided in Figure 4 has a relatively small non-compliance transmission rate, leaving the total non-compliant population nearly constant for the duration of the epidemic. If we take change the non-compliance transmission rate to $\delta = 0.1$ while all other parameters remain as before, we have the results shown in Figure 5. In particular, the non-compliant compartments see even larger epidemic peaks since they comprise a larger portion of the population during the epidemic. Furthermore, the peaks in the non-compliant compartments occur later since the influx terms in equations 2.6 and 2.8 are large in comparison to the out-flux due to recovery. It is also interesting to note initial decrease in the symptomatic compliant compartment followed by an epidemic. This phenomenon indicates that the initial lockdown measures and low non-compliance rate were strong enough to prevent an epidemic, but the swift increase in non-compliance and resulting increase in total infections quickly outweighed the effects of the lockdown.

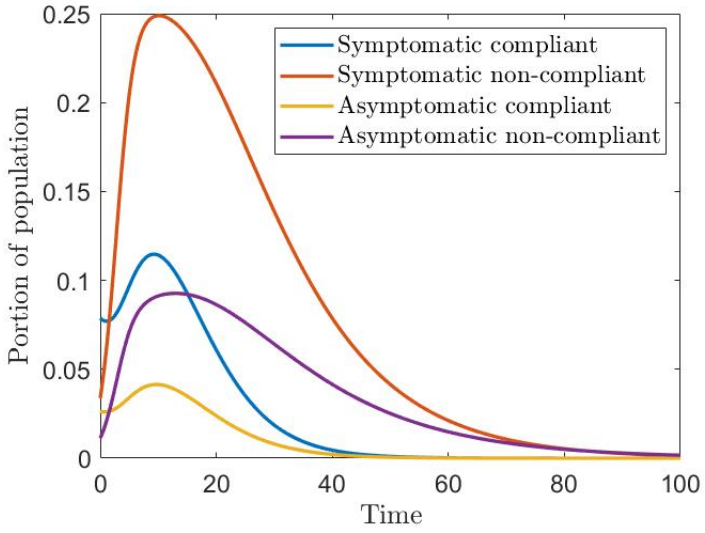

Fig. 5. Simulation for the NCD model taking $\delta = 0.1$, $\sigma = 0.3$, and $\beta = 5$.

Allowing the total non-compliant population to vary smoothly induces some interesting behavior under more extreme parameter regimes. If we take the initial non-compliant population and the non-compliance transmission rate to both be small, we see multiple epidemic waves occur in each infected compartment. As the time between these waves grows smaller, they merge into a single broad, bumpy wave, see Figure 6 for an example of this behavior. These dynamics can be explained by analyzing the susceptible non-compliant population over time, which is also displayed in Figure 6. As the initial lockdown is strong and non-compliance is nearly non-existent, there is initially no epidemic. Once non-compliant susceptible population grows large enough, an epidemic occurs within this compartment, creating a large enough total infected population to induce epidemics in the compliant population. However, since the epidemic spreads quickly through the non-compliant population, the entire epidemic dies out before the total susceptible population is small enough to prevent another epidemic. There is then again a lull in infections until the non-compliant susceptible population becomes sufficiently large, and the cycle repeats.

3. Equilibria and Stability Analysis

As discussed above, a central quantity governing the behavior of epidemic models is the basic reproduction number $\mathscr{R}_0$. For each of the models introduced in section 2, the *next generation matrix method*[18,19,53] allows us to compute the associated reproduction number $\mathscr{R}_0$. However, stability properties do not follow readily because of the structure of the models.

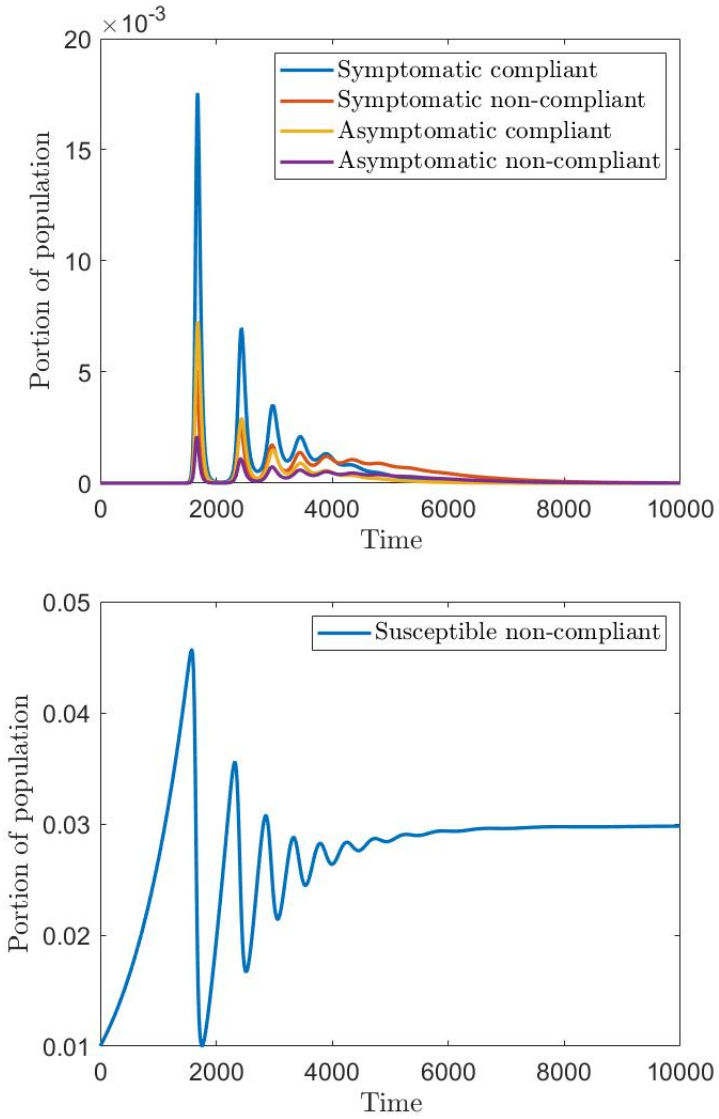

Fig. 6. Simulation for the NCD model taking $\delta = 1 \times 10^{-3}$, $\sigma = 0.01$, and $\beta = 5.0$. The initial total infected population is $\epsilon = 1 \times 10^{-7}$. The top plot shows the infected populations, and the bottom figure shows the susceptible non-compliant population.

To facilitate exposition, we employ the ensuing terminology and notation. An n-dimensional dynamical system is represented by an autonomous differential equation

$$\dot{x} = f(x), \tag{3.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ is a given nonlinear function. Given an equilibrium point $\overline{x}$ (that is, a point such that $f(\overline{x}) = 0$), a Lyapunov criteria for local asymptotic stability is

that

$$s(Df(\overline{x})) := \sup\{\mathrm{Re}(\lambda); \lambda \in \sigma(Df(\overline{x}))\} < 0 \tag{3.2}$$

where $\sigma(\cdot)$ denotes the spectrum set. We note that for basic epidemiological models such as (1.1) local stability analysis is essentially equivalent to the analysis of $\mathscr{R}_0$ discussed in section 1. On one hand, subdividing the susceptible population allows us to consider cases wherein the disease spreads differently among different demographics. On the other hand, such subdivisions complicate the analysis because they introduce the possibility of more than one disease–free equilibrium point.

We are interested in determining stability properties of the infected subsystem in the vicinity of disease–free equilibria. Let n_I be the number of infected compartments in a given system, and let $\mathbf{w}(t)$ be the vector representing them. We then write the infected subsystem as

$$\dot{\mathbf{w}} = (T + \Sigma)\mathbf{w}, \tag{3.3}$$

where T, Σ are the $n_I \times n_I$ *transmission* and *transition* matrices respectively. The transmission matrix captures the introduction of new infections, while the transmission matrix captures flow between (and out of) the infected compartments. With this representation, the basic reproduction number is defined as

$$\mathscr{R}_0 := \rho(-T\Sigma^{-1})$$

where $\rho(A) := \max\{|\lambda|; \lambda \in \sigma(A)\}$. Here $-T\Sigma$ is called the *next generation matrix*. A full exposition of the next generation matrix method is given in the references[18,19,53]

Proposition 3.1. *Let T, Σ be $n \times n$ matrices with real entries such that T is positive*[a] *and Σ is invertible and positive–off–diagonal*[b] *with $s(\Sigma) < 0$. Then*

$$\mathrm{sign}(\rho(-T\Sigma^{-1}) - 1) = \mathrm{sign}(s(T + \Sigma)). \tag{3.4}$$

Remark 3.1. We say that an equilibrium point $\overline{x}$ is locally asymptotically stable with respect to the subsystem of infected compartments if the infectious subsystem (3.3) is asymptotically stable while the system as a whole is in the vicinity of an equilibrium state. Thus combining proposition 3.1 and the Lyapunov criteria (3.2), we see that an equilibrium point $\overline{x}$ is locally asymptotically stable with respect to the infected compartments if and only if $\rho(-T\Sigma^{-1}) < 1$.

[a]Here, an $n \times n$ matrix is called *positive* if all entries are nonnegative. This does not denote positive-definiteness in the sense of operators.
[b]An $n \times n$ matrix is called positive–off–diagonal if all the entries not on the diagonal are nonnegative.

3.1. *Equilibrium and Stability analysis for the p-model*

For the analysis of the p–model we assume that $p(t) = p$ is constant. In essence, this assumes that the population is divided into static compliant and non-compliant populations, with no flow between the two.

With reference to (3.1) we denote $x = (S, I, A, R)^\top$ and let $f(x)$ denote the right hand side of (2.1). The (pre–outbreak) disease-free equilibrium point for this model is $\overline{x}_p := (\overline{S}, \overline{I}, \overline{A}, \overline{R}) = (1, 0, 0, 0)$, and no endemic equilibrium is observed.The transmission and transition matrices for the p–model are given by

$$\begin{aligned} T_p = \beta(1-\alpha p)^2 u_p v_p^\top \text{ and } \Sigma_p = \begin{pmatrix} -\gamma_1 & 0 \\ 0 & -\gamma_2 \end{pmatrix} \\ \text{where } u_p^\top = (\zeta, 1-\zeta), \quad \text{and } v_p^\top = (1-\xi, 1). \end{aligned} \tag{3.5}$$

The next generation matrix is then $-T_p {\Sigma_p}^{-1} = -\beta(1-\alpha p)^2 u_p v_p^\top \Sigma_p^{-1}$, =and the basic reproduction number is

$$\begin{aligned} \mathscr{R}_0^p &= \rho(-T_p {\Sigma_p}^{-1}) = \beta(1-\alpha p)^2 v_p^\top (-\Sigma_p)^{-1} u_p \\ &= \beta(1-\alpha p)^2 \left[\frac{\zeta(1-\xi)}{\gamma_1} + \frac{(1-\zeta)}{\gamma_2} \right]. \end{aligned} \tag{3.6}$$

Notice that in the particular cases ($\alpha = \xi = 0$ and $\zeta = 1$) or ($\alpha = \zeta = 0$), $\mathscr{R}_0^p$ reduces to $\frac{\beta}{\gamma_i}$ where $i = 1$ or $i = 2$. In these cases, the dynamics are identical to the basic SIR model (1.1) with either I or A from the p-model playing the role of I from the SIR model.

It follows from Proposition 3.1 and the Lyapunov criteria (3.2) that the disease free equilibrium $\overline{x}_p = (1, 0, 0, 0)$ is locally asymptotically stable if and only if $\mathscr{R}_0^p < 1$.

3.2. *Equilibrium and Stability analysis for the NCD-model*

The stability analysis for the NCD–model is more intricate than that of the p–model. As in the p-model, the only non-zero populations for pre-outbreak equilibrium states are the susceptible individuals. However, since we allow for two types of susceptibility, we have two disease-free equilibria:

$$\overline{x}_c := (\overline{S_c^c}, \overline{S_c^{nc}}, \overline{I_c^c}, \overline{I_c^{nc}}, \overline{A_c^c}, \overline{A_c^{nc}}, \overline{R_c^c}, \overline{R_c^{nc}}) = (1, 0, 0, 0, 0, 0, 0, 0)$$

representing the equilibrium state at which the entire population is assumed to be compliant to NPIs, and

$$\overline{x}_{nc} := (\overline{S_{nc}^c}, \overline{S_{nc}^{nc}}, \overline{I_{nc}^c}, \overline{I_{nc}^{nc}}, \overline{A_{nc}^c}, \overline{A_{nc}^{nc}}, \overline{R_{nc}^c}, \overline{R_{nc}^{nc}}) = (0, 1, 0, 0, 0, 0, 0, 0)$$

representing the equilibrium state at which the entire population is assumed to be non-compliant to NPIs.

In this case the transmission and transition matrices for the compliant equilibrium are given by

$$T_c = \beta(1-\alpha)u_c v_c^\top \text{ and } \Sigma_c = \begin{pmatrix} -\gamma_1 & 0 & 0 & 0 \\ 0 & -\gamma_1 & 0 & 0 \\ 0 & 0 & -\gamma_2 & 0 \\ 0 & 0 & 0 & -\gamma_2 \end{pmatrix}$$

$$\text{where } u_c^\top = (\zeta, 0, 1-\zeta, 0) \text{ and } v_c^\top = ((1-\xi)(1-\alpha), 1-\xi, 1-\alpha, 1).$$

For the non-compliant equilibrium, they are given by

$$T_{nc} = \beta u_{nc} v_{nc}^\top \text{ and } \Sigma_{nc} = \begin{pmatrix} -(\delta+\gamma_1) & 0 & 0 & 0 \\ \delta & -\gamma_1 & 0 & 0 \\ 0 & 0 & -(\delta+\gamma_2) & 0 \\ 0 & 0 & \delta & -\gamma_2 \end{pmatrix}$$

$$\text{where } u_{nc}^\top = (0, \zeta, 0, 1-\zeta) \text{ and } v_{nc}^\top = ((1-\xi)(1-\alpha), 1-\xi, 1-\alpha, 1).$$

Therefore the basic reproduction numbers for compliant and non-compliant equilibria are given (respectively) by

$$\mathscr{R}_0^c = \rho(-T_c\Sigma_c^{-1}) = \beta(1-\alpha)v_c^\top(-\Sigma_c)^{-1}u_c = \beta(1-\alpha)^2\left[\frac{\zeta(1-\xi)}{\gamma_1} + \frac{(1-\zeta)}{\gamma_2}\right] \tag{3.7}$$

and

$$\mathscr{R}_0^{nc} = \rho(-T_{nc}\Sigma_{nc}^{-1}) = \beta v_{nc}^\top(-\Sigma_{nc})^{-1}u_{nc} = \beta\left[\frac{\zeta(1-\xi)}{\gamma_1} + \frac{(1-\zeta)}{\gamma_2}\right] \tag{3.8}$$

For the case $\delta = 0$, it again follows from Proposition 3.1 and the Lyapunov criteria (3.2) that the disease free equilibrium $\overline{x}_c$ is locally asymptotically if and only if $\mathscr{R}_0^c < 1$. Likewise, the $\overline{x}_{nc}$ is locally asymptotically stable if and only if $\mathscr{R}_0^{nc} < 1$ (in this case, we do not need the caveat that $\delta = 0$, since there will never be a nonzero compliant population).

The basic reproduction numbers $\mathscr{R}_0^c$ and $\mathscr{R}_0^{nc}$ validate the modeling strategy employed in this work. Notice that in case the population is at $\overline{x}_{nc}$, NPI strategies – represented by $\alpha \in (0,1)$ – play no role in the stability criteria, which is expected. One may also note that $\mathscr{R}_0^c = (1-\alpha)^2\mathscr{R}_0^{nc}$ which is also as expected since perfect adherence to NPIs results in a square reduction of mixing since the reduction is applied to both the susceptible and infected classes.

Another important note is that $\overline{x}_c$ and $\overline{x}_{nc}$ represent two very extreme situations that are very unlikely to happen in practice. Moreover, in both cases the transmission of *non-compliance* plays no role: in the former because there are no non-compliant individuals to infect others and in the latter because the entire population is already non-compliant.

These give threshold values for stability provided one knows which equilibrium point the population is in. However, given that our susceptible populations were

divided based on large–scale human behavior, it is very unlikely that (1) the population will ever be at equilibrium and (2) it will ever be possible to quantify exactly how far from equilibrium the population is. Therefore, to offer a more practical threshold quantity, we provide stability results on general disease-free states.

We consider a family of possible *pre-outbreak* disease–free states of the form

$$x_\sigma = (1-\sigma, \sigma, 0, 0, 0, 0, 0, 0)$$

where $\sigma \in (0,1)$ represents the proportion of population that is non-compliant to NPIs. Note that at states x_σ, the system as a whole is not in equilibrium, but the infected compartments are all zero and thus the infectious subsystem is in equilibrium.

Supposing the initial state of the system is x_σ, we seek a threshold parameter that determines whether or not the introduction of infected individuals could lead to an epidemic. This question is deeper than simply asking whether or not an outbreak will occur *given* $\sigma \in (0,1)$. The issue is that policy makers do not have access to precise measurements of σ and therefore cannot draw conclusions from it.

Linearizing (2.3)-(2.10) around $x_\sigma = (1-\sigma, \sigma, 0, 0, 0, 0, 0, 0)$, we find that the transmission and transition matrices for the infected sub–system are given by

$$T_\sigma = \beta u_\sigma v_\sigma^\top \text{ and } \Sigma_\sigma = \begin{pmatrix} -(\sigma\delta+\gamma_1) & 0 & 0 & 0 \\ \sigma\delta & -\gamma_1 & 0 & 0 \\ 0 & 0 & -(\sigma\delta+\gamma_2) & 0 \\ 0 & 0 & \sigma\delta & -\gamma_2 \end{pmatrix}$$

with

$$\begin{aligned} u_\sigma^\top &= \Big((1-\alpha)(1-\sigma)\zeta, \sigma\zeta, (1-\alpha)(1-\sigma)(1-\zeta), \sigma(1-\zeta)\Big) \\ v_\sigma^\top &= \Big((1-\xi)(1-\alpha), (1-\xi), (1-\alpha), 1\Big), \end{aligned} \tag{3.9}$$

Therefore,

$$\begin{aligned} \mathscr{R}_0^\sigma :=& \rho(-T_\sigma \Sigma_\sigma^{-1}) = \beta v_\sigma^\top (-\Sigma_\sigma^{-1}) u_\sigma \\ =& \beta \left[\left(\frac{1}{\gamma_1} - \frac{\alpha}{\sigma\delta+\gamma_1} \right) (1-\xi)(1-\alpha)(1-\sigma)\zeta + \frac{(1-\xi)\sigma\zeta}{\gamma_1} \right] \\ &+ \beta \left[\left(\frac{1}{\gamma_2} - \frac{\alpha}{\sigma\delta+\gamma_2} \right) (1-\alpha)(1-\sigma)(1-\zeta) + \frac{(1-\zeta)\sigma}{\gamma_2} \right]. \end{aligned} \tag{3.10}$$

We observe that $\mathscr{R}_0^c = \mathscr{R}_0^0$ and $\mathscr{R}_0^{nc} = \mathscr{R}_0^1$. Moreover, assuming $\delta = 0$ we can see that the disease free state $\tilde{x}_\sigma = (1-\sigma, \sigma, 0, 0, 0, 0, 0, 0)$ is locally asymptotically stable with respect to the infectious sub–system if and only if $\mathscr{R}_0^\sigma < 1$. The assumption that $\delta = 0$ is pivotal for the validity of this stability property, but is undesirable because it significantly simplifies the model.

We now introduce a series of lemmas that will be useful for achieving our final stability result. First, we notice that the basic reproduction numbers $\mathscr{R}_0^c$ and $\mathscr{R}_0^{nc}$ give upper and lower bounds for the threshold values that imply disease-free local

stability of $\tilde{x}_\sigma$ regardless of σ. Obviously it does not fully describe the asymptotic profile in all cases. For example, if $\mathscr{R}_0^c < 1$ and $\mathscr{R}_0^{nc} > 1$, then neither of them dictates the behavior of the system thus one needs to look specifically at R_0^σ. In the case where either $R_0^c > 1$ or $R_0^{nc} < 1$, the long–time behavior is well established. The first lemma provides a monotonicity property of the basic reproduction number as a function of $\sigma \in (0,1)$.

Lemma 3.1. *The map $\sigma \mapsto \mathscr{R}_0^\sigma$ is strictly increasing on $(0,1)$. In particular, this means that*

$$\mathscr{R}_0^c < \mathscr{R}_0^\sigma < \mathscr{R}_0^{nc}. \tag{3.11}$$

for every $\sigma \in (0,1)$.

Proof. Differentiating (3.10) with respect to σ we have

$$\begin{aligned}\frac{d\mathscr{R}_0^\sigma}{d\sigma} &= \beta\left[\frac{\alpha\delta}{(\sigma\delta+\gamma_1)^2}(1-\xi)(1-\alpha)(1-\sigma)\zeta\right.\\ &\qquad\left.-\left(\frac{1}{\gamma_1}-\frac{\alpha}{\sigma\delta+\gamma_1}\right)(1-\xi)(1-\alpha)\zeta+\frac{(1-\xi)\zeta}{\gamma_1}\right]\\ &\quad+\beta\left[\frac{\alpha\delta}{(\sigma\delta+\gamma_2)^2}(1-\alpha)(1-\sigma)(1-\zeta)\right.\\ &\qquad\left.-\left(\frac{1}{\gamma_2}-\frac{\alpha}{\sigma\delta+\gamma_2}\right)(1-\alpha)(1-\zeta)+\frac{(1-\zeta)}{\gamma_2}\right]\\ &> \beta\left[-\frac{(1-\xi)(1-\alpha)\zeta}{\gamma_1}+\frac{(1-\xi)\zeta}{\gamma_1}-\frac{(1-\alpha)(1-\zeta)}{\gamma_2}+\frac{1-\zeta}{\gamma_2}\right]\geqslant 0,\end{aligned}$$

since $\alpha \geqslant 0$. □

An immediate corollary of Lemma 3.1 is the local stability of disease-free states in special cases.

Proposition 3.2. *The values $\mathscr{R}_0^c$ and $\mathscr{R}_0^{nc}$ provide stability properties of the disease-free state x_σ in the following way:*

(i) If $\mathscr{R}_0^c \geqslant 1$, then x_σ is unstable for all $\sigma \in [0,1]$.
(ii) If $\mathscr{R}_0^{nc} < 1$, then x_σ is locally asymptotically stable for all $\sigma \in [0,1]$

Proposition 3.2 eschews the assumption $\delta = 0$. This comes at the cost of making a very strong assumption that either $\mathscr{R}_0^c \geq 1$ or $\mathscr{R}_0^{nc} < 1$. Going back to (3.8) we can see that all parameters except ξ involved in $\mathscr{R}_0^{nc}$ are biological. Therefore, in the occasion of it being small, it is natural to conclude that the force of the infectious disease is already under control and, therefore, an epidemic can no longer occur, regardless of the compliant/non-compliant profile of the population.

For the next stability result, we use the notation and terminology of Laukó[33]. Let X be the subset of the positive cone $\mathbb{R}_+^8$ defined

$$X = \left\{x = (x_i)_{i=1}^8 \,:\, x_i \geq 0 \;\text{ and }\; \sum x_i = 1\right\}$$

and let $X_f \subset X$ be the set of disease-free states:

$$X_f := \{x \in X \ : \ x_3 = x_4 = x_5 = x_6 = 0\}. \tag{3.12}$$

Identifying x_i with the i^{th} compartment of our NCD-model listed in the order $(S^c, S^{nc}, I^c, I^{nc}, A^c, A^{nc}, R^c, R^{nc})$, we can rewrite the model as

$$\dot{x} = f(x), \quad x(0) = x_0 \in X. \tag{3.13}$$

Before applying the main theorem of [33], we point out some key properties of the NCD-model. First, denote $f = (f_i)_{i=1}^8$ and write each f_i as

$$f_i(x) = F_i(x) - V_i^-(x) + V_i^+(x),$$

where F_i, V_i^-, and V_i^+ denote the rate of appearance of new infections, the rate of the outwards flow, and the rate of inwards flow in the i-th compartment, respectively. In our case, conditions (A_1)-(A_5) from [33, p. 1358] are equivalent to the following remarks:

- The set X is positive invariant. This is to say that given an initial condition $x_0 \in X$, we have $x(t) \in X$ for all $t > 0$.
- The only compartments that receive new infected individuals are I^c, I^{nc}, A^c, and A^{nc}, which means that $F_i \equiv 0$ for $i = 1, 2, 7, 8$.
- If $x \in X_f$, then no infectious compartment receives new infections since the rate of inception depends on the prevalence of infected individuals. In addition, no inward flow from other infected categories occurs since all transition rates are proportional to some infected category. That is, $F_i = V_i^+ = 0$ for $i = 3, 4, 5, 6$.

Let $P : X \to \mathbb{R}_+^4$ denote the projection of X onto the compartments of infected individuals. That is,

$$P(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) = (x_3, x_4, x_5, x_6).$$

We use this projection to define stability of the disease-free stable sets.

Definition 3.1. We say that a set $E \subset X_f$ is a disease-free stable set if there exists a neighborhood $U \subset \mathbb{R}_+^8$ of E such that, for all $\tilde{x} \in U$, and $t > 0$, $\|Px(t)\| \leqslant Ke^{-\omega t}\|P\tilde{x}\|$. Here K and ω are positive constants independent of $\tilde{x}$ and $x(t)$ denotes the solution for (3.13) with $x(0) = \tilde{x}$.

Finally, to verify assumption (H) from [33, p. 1359], we define

$$E := \{(1 - \sigma, \sigma, 0, 0, 0, 0, 0, 0 \ : \ \sigma \in (0, 1)\} \subset X_f \tag{3.14}$$

and notice that for any set U such that $E \subseteq U \subseteq X_f$, the set

$$T_U := \{x(t) \ x(0) \in U \text{ and } \dot{x}(t) = f(x(t)) \text{ for } t \geqslant 0\} \subset X_f$$

is bounded. Under these conditions, we have the following Theorem, the main stability result of this paper, whose proof is a corollary of Theorem 3.1 in [33, p. 1360].

Theorem 3.1. *Assume $\mathscr{R}_0^{nc} < 1$. Then E is a disease-free stable set.*

Proof. The proof of this theorem reduces to verifying that, provided $\mathscr{R}_0^{nc} < 1$, the hypotheses of Theorem 3.1 of [33] are satisfied. This entails showing that there exists a neighborhood U of E such that the family of linearizations

$$\dot{y} = Df(x)y, \quad y(0) = y_0 \in \mathbb{R}^8_+, \tag{3.15}$$

where $x \in X_f$ is a solution of (3.13) with $x(0) \in (I-P)T_U$ is such that the operator $y_0 \mapsto Py$ is bounded with norm dominated by $Ke^{-\omega t}$ for K and ω independent of both y_0 and $x(0)$. For this, we notice that any $w \in T_U$ (here U is any set satisfying $E \subseteq U \subseteq X_f$) has the form

$$w = (w_1, w_2, 0, 0, 0, 0, w_7, w_8).$$

Thus $Pw = \vec{0}$ and so $(I-P)w = w$. Hence $(I-P)T_U = T_U$. We now need to pick a specific set U such that $E \subseteq U \subseteq X_f$,

- $T_U \subset U$, and
- $s(D\tilde{f}(x)) := \sup\left\{\mathrm{Re}(\lambda); \lambda \in \sigma(D\tilde{f}(x))\right\} < 0$ for all $x \in U$ where $\tilde{f} = P \circ f$.

The simplest choice (perhaps not unique) is $U = \overline{E}$. We then have $T_U \subset U$ from the very structure of the problem; that is, if the infected and recovered compartments are 0 at the outset, the whole evolution will occur only within the susceptible compartments, and the sum of these compartments must be 1 due to the definition of X. The second requirement above follows from $\mathscr{R}_0^{nc} < 1$ along with Proposition 3.2. □

We include a version of Theorem 3.1 that highlights its interpretation for our model.

Corollary 3.1. *Fix $x_\sigma = (1-\sigma, \sigma, 0, 0, 0, 0, 0, 0)$, $\sigma \in (0,1)$. Let $\varepsilon > 0$ and assume $\mathscr{R}_0^{nc} < 1$. Then there exists $\eta > 0$ such that if $\|x - x_\sigma\|_{\mathbb{R}^8} < \eta$ then $\|(I^c(t), I^{nc}(t), A^c(t), A^{nc}(t))\|_{\mathbb{R}^4} < \varepsilon$ for all $t \geqslant 0$ and, in particular $\|(I^c(t), I^{nc}(t), A^c(t), A^{nc}(t))\|_{\mathbb{R}^4} \to 0$ exponentially as $t \to +\infty$ with the exponential rate independent of $\sigma \in (0,1)$.*

To better formulate the next result, we use the explicit characterization of the evolution of the compliant population.

Proposition 3.3. *Let $\sigma \in (0,1)$ be the proportion of non-compliant individuals at time $t = 0$. Then, at each time $t > 0$, the proportion of non-compliant individuals is given by*

$$N^{nc}(t) = \frac{\sigma e^{\delta t}}{1 - \sigma + \sigma e^{\delta t}}. \tag{3.16}$$

Proof. This follows since the non-compliant population solves the logistic equation

$$\dot{N^{nc}}(t) = \delta N^{nc}(t) N^c(t) = \delta(1 - N^{nc}(t)) N^{nc}(t)$$

with initial condition $N^{nc}(0) = \sigma$. □

Theorem 3.2. *Suppose $x(0) = x_\sigma$ and assume that $\delta > 0$ and $\mathscr{R}_0^\sigma < 1$. Then the sum of infected compartments is locally decreasing. That is, there exists $t^* = t^*(\sigma)$ such that the function $t \mapsto I^c(t) + I^{nc}(t) + A^c(t) + A^{nc}(t)$ is decreasing for $t \in (0, t^*)$.*

Recall that $\sigma \in (0, 1)$ denotes the initial amount of non-compliant susceptible individuals. Let $\sigma^*(t) \equiv S^{nc}(t)$ denote the amount of non-compliant individuals at time $t > 0$. The proof of this Theorem 3.2 is based on the following idea. If $\mathscr{R}_0^{nc} < 1$, an outbreak could only be caused by a large insertion of infected individuals. However, with a σ–dependent reproduction number, a new possibility arises: even with a small insertion of infected individuals, the proportion of non-compliant people may grow enough so that at an instant $t^* > 0$ one has $\mathscr{R}_0^{\sigma^*(t^*)} > 1$ even when $\mathscr{R}_0^\sigma < 1$. Since we cannot solve our system explicitly, it is likely impossible to determine explicitly how long the number of infected individuals will decrease. However, thanks to formula (3.16), we have a rough estimate.

Proof of Theorem 3.2. Since the map $\sigma \mapsto \mathscr{R}_0^\sigma$ is continuous and $\mathscr{R}_0^\sigma < 1$, there exists $a \in (\sigma, 1)$ such that $\mathscr{R}_0^{\overline{\sigma}} < 1$ for $\overline{\sigma} \in [\sigma, a)$. The theorem then follows from (3.16) by taking

$$t^* = \frac{1}{\delta} \ln\left(\frac{a(1-\sigma)}{\sigma(1-a)}\right)$$

since $\sigma^*(t) \equiv S^{nc}(t) \leqslant N^{nc}(t)$ for all $t \geqslant 0$, and $N^{nc}(t^*) = a$. □

The previous Theorem guarantees that if $\mathscr{R}_0^\sigma < 1$, then the sum of the infected compartments decays, at least for a short period of time. In the last theorem of this section, we show that if the data is such that $\mathscr{R}_0^{nc} > 1$, then strong enough transmissibility of non-compliance will cause the number of infected individuals to eventually increase.

Theorem 3.3. *Suppose $x(0) = x_\sigma$ and $\mathscr{R}_0^\sigma < 1 < \mathscr{R}_0^{nc}$. Then, if $\delta > 0$ is large enough, the sum of infected compartments is **only** locally decreasing. That is, although there exists $t^* = t^*(\sigma)$ such that the function $t \mapsto I^c(t) + I^{nc}(t) + A^c(t) + A^{nc}(t)$ is decreasing in $(0, t^*)$, there exists $t^{**} > t^*$ such that the sum of infected compartments is locally increasing at t^{**}.*

Proof. Since $\mathscr{R}_0^\sigma$ is strictly increasing as a function of σ and $\mathscr{R}_0^\sigma < 1 < \mathscr{R}_0^{nc}$, it follows from the intermediate value theorem that there exists $\overline{\sigma}$ such that $\mathscr{R}_0^{\overline{\sigma}} = 1$. We claim that for the conclusion of the theorem, it is enough to have an instant t^{**} such that $\sigma(t^{**}) \equiv S^{nc}(t^{**}) > \overline{\sigma}$. Indeed, since we are excluding the case of re–infection, we can restart the evolution of the epidemic with initial time t^{**}, with the new total population equal the original population minus the already recovered population, and the new initial non-compliant population $\sigma^*(t^{**})$, for

which $\mathscr{R}_0^{\sigma^*(t^{**})} > \mathscr{R}_0^{\overline{\sigma}} = 1$. Therefore the total infected population will locally increase. It now remains to prove that for δ large enough, one has $\sigma^*(t) > \overline{\sigma}$ for some t.

Let $\lambda \in (0, 1-\sigma)$ and let $t_\lambda > 0$ the first time such that $S^c(t_\lambda) < 1-\sigma-\lambda$. Notice that this time exists since S^c is strictly decreasing and tends to 0 as $t \to \infty$. This means that for all $0 \leqslant t \leqslant t_\lambda$ we have $S^c(t) \geqslant 1-\sigma-\lambda$. Now by (2.2), we have $I_M(t) \leqslant 2-\xi$. From this, we see that

$$\int_0^t I_M(s)ds \leqslant (2-\xi)t.$$

Likewise, by (3.3), $N^{nc}(t) \geqslant \sigma$ for all t. Plugging these bounds into (2.4), we can use an integrating factor to see

$$\begin{aligned}\sigma^*(t) \equiv S^{nc}(t) &\geqslant e^{-\beta(2-\xi)t}S^{nc}(0) + \delta e^{-\beta(2-\xi)t}(1-\sigma-\lambda)\sigma t \\ &= e^{-\beta(2-\xi)t}\left[\sigma + \delta(1-\sigma-\lambda)\sigma t\right],\end{aligned}$$

for all $0 < t < t_\lambda$. Now, let $\lambda > 0$ be such that $t_\lambda > t^*$ and notice that for any $t \in (t^*, t_\lambda)$ fixed, $\delta > 0$ large enough will make $\sigma^*(t) > \overline{\sigma}$. □

The conclusion of the last theorem can be summarized as follows: information on $\mathscr{R}_0^\sigma$ can only provide insight on the local behavior of the evolution. This is because, in the presence of non-compliance, one cannot guarantee that S^{nc} is monotone, and thus cannot guarantee that the total infection decreases after the first peak. The non-monotonicity of S^{nc} and resulting multiple waves of infection are demonstrated in fig. 6 above.

4. Herd Immunity and Epidemic Final Sizes for the *p*–model

The goal of this section is to derive results for the p–model regarding herd immunity level and final epidemic, similar to those listed in section 1.2 for the basic SIR model. To do so, we again assume that $p(t) = p$ is constant.

In analogy to the basic SIR model, if $\mathscr{R}_0^p < 1$, the function $t \to I(t)+A(t)$ reaches its maximum at $t = 0$. The case where $\mathscr{R}_0^p > 1$ is more complicated. To arrive at results analogous to the basic SIR model, we introduce a weighted measurement of the infected class. We define

$$E_p(t) := \frac{\gamma_2(1-\xi)I(t) + \gamma_1 A(t)}{(1-\alpha p)^2(\gamma_2(1-\xi)\zeta + \gamma_2(1-\zeta))} \tag{4.1}$$

and prove the following herd immunity theorem.

Theorem 4.1. *Let $\mathscr{R}_0^p > 1$. Then E_p has a global maximum at $t_h > 0$ such that $S(t_h) = (\mathscr{R}_0^p)^{-1}$. Moreover, the maximum value is given by*

$$E_p(t_h) = E_p(0) + \frac{1}{(1-\alpha p)^2}\left(\frac{1}{\mathscr{R}_0^p} - S_0 - \frac{1}{\mathscr{R}_0^p}\ln\left(\frac{1}{S_0\mathscr{R}_0^p}\right)\right). \tag{4.2}$$

Proof. We introduce the following notation:

$$\mathbf{i}(t) := \begin{pmatrix} I(t) \\ A(t) \end{pmatrix}, \; v_p := \begin{pmatrix} 1-\xi \\ 1 \end{pmatrix}, \; u_p := \begin{pmatrix} \zeta \\ 1-\zeta \end{pmatrix}, \; \mathbf{v} := \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \tag{4.3}$$

and let Σ_p be the transition matrix introduced in (3.5).Then the p-model can be written

$$\begin{cases} \dfrac{dS}{dt} = -\beta(1-\alpha p)^2 v_p^\top \mathbf{i} S, \\ \dfrac{d\mathbf{i}}{dt} = [\beta(1-\alpha p)^2 v_p^\top \mathbf{i}] u_p S + \Sigma_p \mathbf{i} \\ \dfrac{dR}{dt} = \mathbf{v}^\top \mathbf{i} \end{cases} \tag{4.4}$$

In this notation, the basic reproduction number for the p-model is given by

$$\mathscr{R}_0^p = -\beta(1-\alpha p)^2 v_p^\top \Sigma^{-1} u_p$$

and $E_p(t) = -\beta(\mathscr{R}_0^p)^{-1} v_p^\top \Sigma^{-1} \mathbf{i}(t)$. Thus

$$\begin{aligned} \frac{dE_p(t)}{dt} &= -\beta(\mathscr{R}_0^p)^{-1} v_p^\top \Sigma^{-1} \frac{d\mathbf{i}(t)}{dt} \\ &= -\beta(\mathscr{R}_0^p)^{-1} v_p^\top \Sigma^{-1} ([\beta(1-\alpha p)^2 v_p^\top \mathbf{i}] u_p S + \Sigma \mathbf{i}) \\ &= -\beta v_p^\top \mathbf{i} (\mathscr{R}_0^p)^{-1} \beta(1-\alpha p)^2 v_p^\top \Sigma^{-1} u_p S - \beta(\mathscr{R}_0^p)^{-1} v_p^\top \mathbf{i} \\ &= \beta v_p^\top \mathbf{i} S - \beta(\mathscr{R}_0^p)^{-1} v_p^\top \mathbf{i} \qquad (4.5) \\ &= \beta v_p^\top \mathbf{i} \left(S - (\mathscr{R}_0^p)^{-1} \right), \qquad (4.6) \end{aligned}$$

which implies that the global maximum of $E_p(t)$ is reached at a time t_h such that $S(t_h) = (\mathscr{R}_0^p)^{-1}$.

Next notice that

$$\frac{dE_p}{dS} = \frac{1}{(1-\alpha p)^2} \left(\frac{1}{\mathscr{R}_0^p S} - 1 \right), \tag{4.7}$$

which implies that

$$E_p(t) - E_p(0) = \frac{1}{(1-\alpha p)^2} \left(\frac{1}{\mathscr{R}_0^p} \ln \left(\frac{S(t)}{S_0} \right) - S(t) + S_0 \right),$$

so at $t = t_h$ we obtain (4.2). □

From the last formula, we arrive at the following conserved quantity.

Corollary 4.1 (Conservation). *The function*

$$t \mapsto E_p(t) + \frac{S(t)}{(1-\alpha p)^2} - \frac{\ln S(t)}{(1-\alpha p)^2 \mathscr{R}_0^p}$$

is constant.

Next, we derive a final size transcendental equation for the p-model.

Theorem 4.2. *The functions $t \mapsto I(t)$ and $t \mapsto A(t)$ vanish as $t \to \infty$. Moreover, the final size of the susceptible population, denoted S_∞, satisfies the transcendental equation*

$$S_\infty = S_0 e^{-\mathscr{R}_0^p(1-S_\infty)}. \tag{4.8}$$

Proof. First notice that the boundedness and monotonicity of S guarantees its limit at infinity exists. An algebraic manipulation shows that $\zeta\dot{S}(\cdot) + \dot{I}(\cdot)$ and $(1-\zeta)\dot{S}(\cdot)+\dot{A}(\cdot)$ are both non-positive. This, along with boundedness of I and A, shows $I(\cdot)$ and $A(\cdot)$ vanish as $t \to \infty$. The result now follows by considering

$$\lim_{t\to\infty}\left(E_p(t) + \frac{S(t)}{(1-\alpha p)^2} - \frac{\ln S(t)}{(1-\alpha p)^2\mathscr{R}_0^p}\right)$$

and using Corollary (4.1). □

5. Herd Immunity and Epidemic Final Sizes for the NCD–model

In this section, we derive results analogous to those in section 4 for the NCD model. In the same manner as the above, we introduce a weighted sum of infected populations for measuring the size of the infected compartments. Since the NCD-model is more complicated, we first introduce some simplifying notation.

We define

$$\mathbf{i}_\alpha(t) := \begin{pmatrix} (1-\alpha)I^c(t) + I^{nc}(t) \\ (1-\alpha)A^c(t) + A^{nc}(t) \end{pmatrix},\ \mathbf{i}(t) := \begin{pmatrix} I^c(t) + I^{nc}(t) \\ A^c(t) + A^{nc}(t) \end{pmatrix}$$

$$v_p := \begin{pmatrix} 1-\xi \\ 1 \end{pmatrix},\ u_p := \begin{pmatrix} \zeta \\ 1-\zeta \end{pmatrix},\ \mathbf{v} := \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}$$

and let Σ_p be the transition matrix for the p-model. Finally, let Υ be the matrix

$$\Upsilon := \begin{pmatrix} 1 - \dfrac{\alpha\gamma_1}{\sigma\delta+\gamma_1} & 0 \\ 0 & 1 - \dfrac{\alpha\gamma_1}{\sigma\delta+\gamma_2} \end{pmatrix}.$$

We then measure the infected compartment sizes by defining the function

$$E_\sigma(t) = -\beta(\mathscr{R}_0^\sigma)^{-1}v_p^\top\left[\sigma + (1-\alpha)(1-\sigma)\Upsilon\right]\Sigma_p^{-1}\mathbf{i}(t). \tag{5.1}$$

We begin our analysis with a lemma from which our herd immunity results will follow.

Lemma 5.1. *The function $t \mapsto E_\sigma(t)$ satisfies the following identity:*

$$\frac{dE_\sigma(t)}{dt} = \beta v_p^\top \mathbf{i}_\alpha S_M - \beta(\mathscr{R}_0^\sigma)^{-1}v_p^\top\left[\sigma I + (1-\alpha)(1-\sigma)\Upsilon\right]\mathbf{i} \tag{5.2}$$

where I denotes the identity matrix.

Proof. Denoting $S_T = S^c + S^{nc}$, $R_T = R^c + R^{nc}$ and defining the mixing susceptible population $S_M = (1-\alpha)S^c + S^{nc}$, the system (2.3)-(2.10) can be expressed

$$\begin{cases} \dot{S_T} = -\beta v_p^\top \mathbf{i}_\alpha S_M, \\ \dot{\mathbf{i}} = [\beta v_p^\top \mathbf{i}_\alpha] u_p S_M + \Sigma_p \mathbf{i} \\ \dot{R_T} = \mathbf{v}^\top \mathbf{i} \end{cases} . \tag{5.3}$$

We then notice that

$$\mathscr{R}_0^\sigma = -\beta v_p^\top \left[\sigma I + (1-\alpha)(1-\sigma)\Upsilon\right] \Sigma_p^{-1} u_p,$$

so it follows that

$$\begin{aligned} \frac{dE_\sigma(t)}{dt} &= -\beta (\mathscr{R}_0^\sigma)^{-1} v_p^\top \left[\sigma I + (1-\alpha)(1-\sigma)\Upsilon\right] \Sigma_p^{-1} \frac{d\mathbf{i}(t)}{dt} \\ &= -\beta (\mathscr{R}_0^\sigma)^{-1} v_p^\top \left[\sigma I + (1-\alpha)(1-\sigma)\Upsilon\right] \Sigma_p^{-1} \left\{[\beta v_p^\top \mathbf{i}_\alpha] u_p S_M + \Sigma_p \mathbf{i}\right\} \\ &= \beta v_p^\top \mathbf{i}_\alpha S_M - \beta (\mathscr{R}_0^\sigma)^{-1} v_p^\top \left[\sigma I + (1-\alpha)(1-\sigma)\Upsilon\right] \mathbf{i}. \end{aligned}$$

□

Corollary 5.1 (Herd Immunity for the unstable compliant equilibrium). *Assume that $\mathscr{R}_0^c > 1$. Then the function $t \mapsto E_{\sigma=0}(t)$ attains its global maximum at time $t_h > 0$ satisfying*

$$\frac{(1-\alpha)^2}{\mathscr{R}_0^c} \leqslant S_M(t_h) \leqslant \frac{1-\alpha}{\mathscr{R}_0^c}.$$

Moreover, the function $t \mapsto E_0(t)$ increases while $S_M(t) \geqslant \dfrac{1-\alpha}{\mathscr{R}_0^c}$ and decreases while $S_M(t) \leqslant \dfrac{(1-\alpha)^2}{\mathscr{R}_0^c}$.

Proof. Taking $\sigma = 0$ in (5.2), we have

$$\frac{dE_0(t)}{dt} = \beta v_p^\top \mathbf{i}_\alpha S_M - \beta (\mathscr{R}_0^c)^{-1} v_p^\top (1-\alpha)^2 \mathbf{i},$$

so

$$\beta v_p^\top \mathbf{i}_\alpha (S_M - (\mathscr{R}_0^c)^{-1}(1-\alpha)) \leqslant \frac{dE_0(t)}{dt} \leqslant \beta v_p^\top \mathbf{i}_\alpha (S_M - (\mathscr{R}_0^c)^{-1}(1-\alpha)^2),$$

from which the claim follows. □

Corollary 5.2 (Herd Immunity for the unstable non-compliant equilibrium). *Assume that $\mathscr{R}_0^{nc} > 1$. Then the function $t \mapsto E_{\sigma=1}(t)$ attains its global maximum at a time $t_h > 0$ satisfying*

$$\frac{1}{\mathscr{R}_0^{nc}} \leqslant S_M(t_h) \leqslant \frac{1}{(1-\alpha)\mathscr{R}_0^{nc}}.$$

Moreover, $t \mapsto E_1(t)$ increases while $S_M(t) \geqslant \dfrac{1}{(1-\alpha)\mathscr{R}_0^{nc}}$ and decreases while $S_M(t) \leqslant \dfrac{1}{\mathscr{R}_0^{nc}}$.

Proof. Taking $\sigma = 1$ in (5.2) we have

$$\frac{dE_1(t)}{dt} = \beta v_p^\top \mathbf{i}_\alpha S_M - \beta(\mathscr{R}_0^{nc})^{-1} v_p^\top \mathbf{i}$$

then

$$\beta v_p^\top \mathbf{i}_\alpha (S_M - ((1-\alpha)\mathscr{R}_0^{nc})^{-1}) \leqslant \frac{dE_0(t)}{dt} \leqslant \beta v_p^\top \mathbf{i}_\alpha (S_M - (\mathscr{R}_0^{nc})^{-1}),$$

from which the claim follows. □

Corollary 5.3 (Herd Immunity for the unstable disease free states). *Assume $\mathscr{R}_0^\sigma > 1$. Then the function $t \mapsto E_\sigma(t)$ attains its global maximum at an instant t_h satisfying*

$$\frac{(\sigma + (1-\alpha)^2(1-\sigma))}{\mathscr{R}_0^{nc}} \leqslant S_M(t_h) \leqslant \frac{(\sigma + (1-\alpha)(1-\sigma))}{(1-\alpha)\mathscr{R}_0^{nc}}.$$

Moreover, $t \mapsto E_1(t)$ increases when

$$S_M(t) \geqslant \frac{(\sigma + (1-\alpha)(1-\sigma))}{(1-\alpha)\mathscr{R}_0^{nc}} \tag{5.4}$$

and decreases when

$$S_M(t) \leqslant \frac{(\sigma + (1-\alpha)^2(1-\sigma))}{\mathscr{R}_0^{nc}}. \tag{5.5}$$

Proof. From (5.2) we have

$$\beta v_p^\top \mathbf{i}_\alpha (S_M - (\sigma + (1-\alpha)(1-\sigma))((1-\alpha)\mathscr{R}_0^c)^{-1}) \leqslant \frac{dE_\sigma(t)}{dt}$$

and

$$\frac{dE_\sigma(t)}{dt} \leqslant \beta v_p^\top \mathbf{i}_\alpha (S_M - (\sigma + (1-\alpha)^2(1-\sigma))(\mathscr{R}_0^c)^{-1}),$$

from which the claim follows. □

Remark 5.1. Notice that Corollary 5.3 generalizes Corollaries 5.1 and 5.2. However, Lemma 5.1 is stronger in the sense that it also provides insights on how the *non-compliance* transmission rate $\delta > 0$ affects the dynamics of E_σ. Since $\mathscr{R}_0^\sigma$ is an increasing function of δ, equation (5.2) shows that $\frac{dE_\sigma}{dt}$ is an increasing function of δ as well. This is to be expected: if the population becomes non-compliant more rapidly, then NPIs will become ineffective more rapidly.

As a final note, we point out some limitation of this $\mathscr{R}_0$-analysis, which represents the largest difference in the analysis of the NCD-model as opposed to simpler models such as the basic SIR Model or the p-model with constant non-compliance rate p. To summarize, for these simpler models, $\mathscr{R}_0$-analysis more fully characterizes the behavior of the models by supplying, among other physically relevant quantities, exact descriptions—though sometimes in the form of implicit equations—of the

stability threshold for disease-free equilibrium, herd immunity time and final epidemic size (see, for example, equations (1.2), (4.8) which determine final epidemic size). These models satisfy a type of monotonicity: the total infected population rises to a single peak and then decays to zero. By contrast, for the NCD-model, the behavior is non-monotone: there is a possibility of multiple peaks of infections, as we see in fig. 6 and in theorem 3.3. Because of this, $\mathscr{R}_0$-analysis can only provide local stability and estimates of key quantities for the NCD-model. Explicitly, for the basic SIR model described in (1.1), the herd immunity time t_h is given exactly as the solution to the equation $S(t_h) = \mathscr{R}_0^{-1}$: the total number of infections will increase until this time and decrease thereafter. For the NCD-model, Corollaries 5.1-5.3 give estimates for the herd immunity time and establish similar qualitative results in equations (5.4) and (5.5): the total number of infections will increase so long as the total susceptible population is large enough and decrease when the total susceptible populations becomes small enough. Likewise, Theorem 3.2 establishes a local stability result for the NCD-model but Theorem 3.3 ensures that it is *only* a local result. The medium-time dynamics of the NCD-model are more intricate, and because of this $\mathscr{R}_0$-analysis cannot perfectly describe the behavior.

6. Conclusion

Inspired by the spread of SARS-CoV-2 throughout the world and the lockdown mechanisms governments originally implemented to control the spread of the virus, we developed two models to understand the interplay between the stringency of lockdown, non-compliant behavior, and the spread of the disease. In the first model (p-model), we assumed a fraction p of the population would always be compliant whereas in the second model (NCD-model), we treated non-compliance as a social contagion spreading through the general population in parallel with the disease. We have investigated our models through numerical simulations, proven a number of important qualitative results, and established estimates for key epidemiological parameters. Of particular interest, we found that even without cycles of closing and reopening, multiple waves of disease can emerge naturally in the NCD model; we have proved that if the social contagion is strong enough, it is possible for the number of active infections to initially decrease (giving the initial impression the disease is dying out) and then later increase; and we have obtained (respectively) precise formulas and bounds for the herd immunity in the p-model and NCD model. These models are quite general and not specific to the SARS-CoV-2 virus. Therefore these models and results could be useful in modeling and understanding future epidemics.

With regards to extending this work, we note that the analysis in this paper has been entirely theoretical and qualitative, without focusing on specific disease parameters or data from any specific city. Armed with the understanding thus obtained, it could be of interest to study case data in specific cities and correlate trends in such data with our models when provided parameter estimates. Likewise, one could pair these models with in-host viral models to analyze the disease spread

on multiple scales in an approach similar to some previous models[5,45,12].

Acknowledgments

This work was supported by The Simons Foundation through Math + X Investigator Award 510776 and NSF grants DMS-2027438 and DMS-1737770. This research was supported in part by an appointment with the National Science Foundation (NSF) Mathematical Sciences Graduate Internship (MSGI) Program sponsored by the NSF Division of Mathematical Sciences. This program is administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the U.S. Department of Energy (DOE) and NSF. ORISE is managed for DOE by ORAU. All opinions expressed in this paper are the author's and do not necessarily reflect the policies and views of NSF, ORAU/ORISE, or DOE.

References

1. G. G. Alcaraz and C. Vargas-De-León, Modeling control strategies for influenza A H1N1 epidemics: SIR models, *Revista Mexicana de Física* **58** (2012) 37–43.
2. M. M. Ali, A. Amialchuk and D. S. Dwyer, The social contagion effect of marijuana use among adolescents, *PloS one* **6** (2011) e16183.
3. J. Bahr, Police will respond to complaints about noncompliance with mask mandate, https://theindependent.com/news/local/police-will-respond-to-complaints-about-noncompliance-with-mask-mandate/article_0249e142-351e-11eb-8958-0f23e18314ff.html, 2020, accessed: 2021-05-09.
4. M. V. Barbarossa and J. Fuhrmann, Compliance with NPIs and possible deleterious effects on mitigation of an epidemic outbreak, *Infectious Disease Modelling* **6** (2021) 859–874.
5. N. Bellomo, D. Burini and N. Outada, Multiscale models of COVID-19 with mutations and variants, *Networks and Heterogeneous Media* **17** (2022) 293.
6. S. Bentout, A. Chekroun and T. Kuniya, Parameter estimation and prediction for coronavirus disease outbreak 2019 (COVID-19) in Algeria, *AIMS Public Health* **7** (2020) 306.
7. T. Berge, J.-S. Lubuma, G. Moremedi, N. Morris and R. Kondera-Shava, A simple mathematical model for Ebola in Africa, *Journal of biological dynamics* **11** (2017) 42–74.
8. M. C. Bootsma and N. M. Ferguson, The effect of public health measures on the 1918 influenza pandemic in us cities, *Proceedings of the National Academy of Sciences* **104** (2007) 7588–7593.
9. S. K. Brooks, R. K. Webster, L. E. Smith, L. Woodland, S. Wessely, N. Greenberg and G. J. Rubin, The psychological impact of quarantine and how to reduce it: rapid review of the evidence, *The lancet* **395** (2020) 912–920.
10. X. Chen, H. Huang, J. Ju, R. Sun and J. Zhang, Impact of vaccination on the COVID-19 pandemic in US states, *Scientific reports* **12** (2022) 1–10.
11. Y.-C. Chen, P.-E. Lu, C.-S. Chang and T.-H. Liu, A time-dependent SIR model for COVID-19 with undetectable infected persons, *arXiv preprint arXiv:2003.00122* .
12. B. Chhetri, D. Vamsi and C. Sanjeevi, A nested multi-scale model for COVID-19 viral infection, *arXiv preprint arXiv:2108.12150* .
13. W.-H. Chiang, X. Liu and G. Mohler, Hawkes process modeling of COVID-19 with

mobility leading indicators and spatial covariates, *International journal of forecasting* **38** (2022) 505–520.
14. N. A. Christakis and J. H. Fowler, Social contagion theory: examining dynamic social networks and human behavior, *Statistics in medicine* **32** (2013) 556–577.
15. C. Cockrell and G. An, Comparative computational modeling of the bat and human immune response to viral infection with the comparative biology immune agent based model, *Viruses* **13** (2021) 1620.
16. V. d'Andrea, R. Gallotti, N. Castaldo and M. De Domenico, Individual risk perception and empirical social structures shape the dynamics of infectious disease outbreaks, *PLoS Computational Biology* **18** (2022) e1009760.
17. L. Danon, A. P. Ford, T. House, C. P. Jewell, M. J. Keeling, G. O. Roberts, J. V. Ross and M. C. Vernon, Networks and the epidemiology of infectious disease, *Interdisciplinary perspectives on infectious diseases* **2011**.
18. O. Diekmann, J. Heesterbeek and M. G. Roberts, The construction of next-generation matrices for compartmental epidemic models, *Journal of the Royal Society Interface* **7** (2010) 873–885.
19. O. Diekmann, J. A. P. Heesterbeek and J. A. Metz, On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations, *Journal of mathematical biology* **28** (1990) 365–382.
20. C. Durón and A. Farrell, A mean-field approximation of SIR epidemics on an Erdös–Rényi network model, *Bulletin of Mathematical Biology* **84** (2022) 1–19.
21. R. C. Engels, R. H. Scholte, C. F. van Lieshout, R. de Kemp and G. Overbeek, Peer group reputation and smoking and alcohol consumption in early adolescence, *Addictive behaviors* **31** (2006) 440–449.
22. P. I. Frazier, J. M. Cashore, N. Duan, S. G. Henderson, A. Janmohamed, B. Liu, D. B. Shmoys, J. Wan and Y. Zhang, Modeling for COVID-19 college reopening decisions: Cornell, a case study, *Proceedings of the National Academy of Sciences* **119** (2022) e2112532119.
23. G. Gaeta, A simple sir model with a large set of asymptomatic infectives, *Mathematics in Engineering* **3** (2021) 1–39.
24. M. Garetto, E. Leonardi and G. L. Torrisi, A time-modulated Hawkes process to model the spread of COVID-19 and the impact of countermeasures, *Annual reviews in control* **51** (2021) 551–563.
25. L. Gori, P. Manfredi, S. Marsiglio and M. Sodini, COVID-19 epidemic and mitigation policies: Positive and normative analyses in a neoclassical growth model, *Journal of Public Economic Theory* .
26. K. Hadeler, Parameter estimation in epidemic models: simplified formulas, *Canadian Applied Mathematics Quarterly* **19** (2011) 343–356.
27. E. A. Hernandez-Vargas and J. X. Velasco-Hernandez, In-host mathematical modelling of COVID-19 in humans, *Annual reviews in control* **50** (2020) 448–456.
28. W. O. Kermack and A. G. McKendrick, A contribution to the mathematical theory of epidemics, *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character* **115** (1927) 700–721.
29. I. E. Kibona and C. Yang, SIR model of spread of Zika virus infections: ZIKV linked to microcephaly simulations, *Health* **9** (2017) 1190–1210.
30. C. Kresin, F. Schoenberg and G. Mohler, Comparison of Hawkes and SEIR models for the spread of COVID-19, *Statistical Science* .
31. O. Krivorotko, M. Sosnovskaia, I. Vashchenko, C. Kerr and D. Lesnic, Agent-based modeling of COVID-19 outbreaks for new york state and uk: Parameter identification algorithm, *Infectious Disease Modelling* **7** (2022) 30–44.

32. J. Lasser, J. Sorger, L. Richter, S. Thurner, D. Schmid and P. Klimek, Assessing the impact of SARS-CoV-2 prevention measures in Austrian schools using agent-based simulations and cluster tracing data, *Nature communications* **13** (2022) 1–17.
33. I. G. Laukó, Stability of disease free sets in epidemic models, *Mathematical and computer modelling* **43** (2006) 1357–1366.
34. C. M. Macal and M. J. North, Agent-based modeling and simulation, in *Proceedings of the 2009 winter simulation conference (WSC)* (IEEE, 2009), pp. 86–98.
35. M. Mandal, S. Jana, S. K. Nandi, A. Khatua, S. Adak and T. Kar, A model based study on the dynamics of COVID-19: Prediction and control, *Chaos, Solitons, and Fractals* **136**.
36. J. Marcus, Quarantine fatigue is real, https://www.theatlantic.com/ideas/archive/2020/05/quarantine-fatigue-real-and-shaming-people-wont-help/611482/, 2020, accessed: 2021-05-09.
37. J. C. Miller, A note on the derivation of epidemic final sizes, *Bulletin of mathematical biology* **74** (2012) 2125–2141.
38. M. M. Morato, S. B. Bastos, D. O. Cajueiro and J. E. Normey-Rico, An optimal predictive control strategy for COVID-19 (SARS-CoV-2) social distancing policies in Brazil, *Annual reviews in control* **50** (2020) 417–431.
39. M. Morris, *Network epidemiology: a handbook for survey design and data collection* (OUP Oxford, 2004).
40. D. Osthus, K. S. Hickmann, P. C. Caragea, D. Higdon and S. Y. Del Valle, Forecasting seasonal influenza with a state-space SIR model, *The annals of applied statistics* **11** (2017) 202.
41. S. Pankavich and C. Parkinson, Mathematical analysis of an in-host model of viral dynamics with spatial heterogeneity, *Discrete & Continuous Dynamical Systems-B* **21** (2016) 1237.
42. J. Parker, A flexible, large-scale, distributed agent based epidemic model, in *2007 Winter Simulation Conference* (IEEE, 2007), pp. 1543–1547.
43. K. Peng, Z. Lu, V. Lin, M. R. Lindstrom, C. Parkinson, C. Wang, A. L. Bertozzi and M. A. Porter, A multilayer network model of the coevolution of the spread of a disease and competing opinions, *Mathematical Models and Methods in Applied Sciences* **31** (2021) 2455–2494.
44. A. S. Perelson and R. Ke, Mechanistic modeling of SARS-CoV-2 and other infectious diseases and the effects of therapeutics, *Clinical Pharmacology & Therapeutics* **109** (2021) 829–840.
45. D. B. Prakash, D. Vamsi, D. B. Rajesh and C. B. Sanjeevi, Control intervention strategies for within-host, between-host and their efficacy in the treatment, spread of COVID-19: A multi scale modeling approach, *Computational and Mathematical Biophysics* **8** (2020) 198–210.
46. A. Press, 4 bars cited for noncompliance with coronavirus regulations, https://www.usnews.com/news/best-states/rhode-island/articles/2020-10-13/4-bars-cited-for-nomcompliance-with-coronavirus-regulations, 2020, accessed: 2022-05-09.
47. T. Sego, J. O. Aponte-Serrano, J. Ferrari Gianlupi, S. R. Heaps, K. Breithaupt, L. Brusch, J. Crawshaw, J. M. Osborne, E. M. Quardokus, R. K. Plemper et al., A modular framework for multiscale, multicellular, spatiotemporal modeling of acute primary viral infection and immune response in epithelial tissues and its application to drug therapy timing and effectiveness, *PLoS computational biology* **16** (2020) e1008451.
48. J. Sooknanan and D. M. Comissiong, When behaviour turns contagious: the use of deterministic epidemiological models in modeling social contagion phenomena, *International Journal of Dynamics and Control* **5** (2017) 1046–1050.

49. K. F. Stanger-Hall and D. W. Hall, Abstinence-only education and teen pregnancy rates: Why we need comprehensive sex education in the us, *PloS one* **6**.
50. B. K. Tay, C. A. Roby, J. W. Wu and D. Y. Tan, Dynamical analysis of universal masking on the pandemic, *International journal of environmental research and public health* **18** (2021) 9027.
51. N. K. Vaidya, A. Bloomquist and A. S. Perelson, Modeling within-host dynamics of SARS-CoV-2 infection: A case study in ferrets, *Viruses* **13** (2021) 1635.
52. T. E. Valles, H. Shoenhard, J. Zinski, S. Trick, M. A. Porter and M. R. Lindstrom, Networks of necessity: Simulating COVID-19 mitigation strategies for disabled people and their caregivers, *PLOS Computational Biology* **18** (2022) e1010042.
53. P. van den Driessche, Reproduction numbers of infectious disease models, *Infectious Disease Modelling* **2** (2017) 288–303.
54. X. Wang, Z. Wang, X. Huang and Y. Li, Dynamic analysis of a delayed fractional-order SIR model with saturated incidence and treatment functions, *International Journal of Bifurcation and Chaos* **28** (2018) 1850180.
55. M. D. Witte, Why people didn't social distance, https://news.stanford.edu/2020/04/14/people-didnt-social-distance/, 2020, accessed: 2022-05-09.
56. M.-Z. Yin, Q.-W. Zhu and X. Lü, Parameter estimation of the incubation period of COVID-19 based on the doubly interval-censored data model, *Nonlinear Dynamics* **106** (2021) 1347–1358.