A feedback SIR (fSIR) model highlights advantages and limitations of infection-based social distancing

Elisa Franco,
Department of Mechanical and Aerospace Engineering,
Department of Bioengineering,
University of California, Los Angeles

***Abstract*— Transmission rates in epidemic outbreaks vary over time depending on the societal and government response to infections and mortality data, as evidenced in the course of the COVID-19 pandemic. Following a mean field approach that models individuals like particles in a well-mixed solution, I derive a modified SIR model in which the contacts between susceptible and infected population are reduced based on the known infection levels. This approach yields a time-varying reproduction number that is continuously adjusted, based on infection information through a negative-feedback term that is similar to Michaelis-Menten and Hill functions in chemistry and molecular biology. This feedback-adjustment of the transmission rate causes a structural reduction in infection peak, and simulations indicate that such reduction persists even in the presence of information delays. Simulations also show that a distancing policy based on infection data may substantially extend the duration of an epidemic. The main advantage of this model is that it adds a *single* parameter to the original SIR, making it useful to illustrate the effects of social distancing enforced based on awareness of infections.**

I. Introduction

Compartment models are the simplest and most established approach to modeling epidemic outbreaks. Like mean-field models in physics and chemical reaction networks governed by the law of mass action, compartment models assume a well-mixed population in which individuals have average interaction and recovery patterns. The population is binned in categories, and at a minimum include those susceptible to disease (S), those who become infected (I), and those who recover and/or die (R), which are included in the well-known SIR model by Kermack and McKendrick [12].

Taking inspiration from models widely adopted in molecular biology, I examine a variant of the SIR model in which the transmission rate is not constant, rather it changes based on information of current infections through a Hill-type function: the number of reported infections causes a reduction of the transmission rate, thereby reducing the likelihood of additional infections. Because infection information is available to policy-makers and to a large number of individuals through apps or websites that track contagion information [5], [18], [17], it is reasonable (not to mention desirable) to assume that individuals react to knowledge of infections by self-isolating, using personal protective equipment, or increasing hygiene standards [4]. Because this model includes an explicit feedback loop, I will refer to it as feedback SIR (fSIR).

Similar models have been adopted to model how *awareness* of infections modifies social habits and transmission rates. For instance, Bootsma and Ferguson included a Hill-type term to capture the effects of awareness of *deaths* (rather than infections) on social interactions within the 1918 influenza epidemic in the United States [2]. Previous models in the literature have examined how infection awareness reduces the susceptible fraction of the population [14], [7], in particular by increasing vaccination rates [3]. The reduction of transmission and contact rates achieved by awareness programs has also been evaluated within SIS models [9], [19]. I am not aware of studies focusing on the direct effects of infection information on epidemic transmission rates, a scenario that has become plausible in the COVID-19 pandemic thanks to the widespread social connectivity and rapid access to media and news reports of most individuals [5], [18]. The question of whether testing information can be used successfully to regulate distancing policies is important in the COVID-19 pandemic, and it has been recently investigated considering a time-varying control function of transmission rate that depends on the number of daily tests [4].

Analysis and simulations of the fSIR model show that a reduction of the transmission rate that depends on the level of detected infections has the following effects: 1) the peak size of the infection is structurally reduced, and this peak mitigation persists even in the presence of infection information delay of several days; 2) the peak of infections is moderately delayed; 3) the duration of the epidemic, measured as the time for which infections persist, may significantly increase. Further, the fSIR model can capture historical infection data for the COVID-19 pandemic.

The main contribution of this report is that it repurposes a well-known epidemic model to capture soci-

etal responses to epidemic information with only one parameter in addition to the reproduction number. The fSIR model may be helpful as an illustrative yet rigorous framework to examine alternative scenarios for the epidemic spread depending on the strictness of social distancing. Further, this model supports social distancing guidelines as it clearly shows that the infection curve can be flattened without postponing the peak, a misleading (and demotivating) scenario suggested by models that use a constant transmission rate. At the same time, the model highlights that policies relying exclusively on infection data to regulate social distancing can majorly extend the time required to reach a disease-free equilibrium [4].

II. Background: qualitative analysis of the non-dimensional SIR model

The SIR model is the simplest compartment model that can qualitatively capture the evolution of an epidemic at the population level. The model assumes well-mixed population that remains constant (without birth and death processes) and has two key parameters: 1) the transmission coefficient β, which depends on the social interactions among individuals (average daily contacts) and on the viral infection characteristics; the transmission rate is generally thought as the product of the average frequency of contacts between infected and susceptible and the likelihood that infection occurs given a contact; 2) the recovery coefficient γ, which captures the average time for recovery (or death) of infected individuals. The inverse $1/\gamma$ is also known as duration of infectiousness. Assuming the total population is N, the original SIR model is:

$$\frac{dS}{dt} = -\beta \frac{S}{N} I, \tag{1}$$

$$\frac{dI}{dt} = \beta \frac{S}{N} I - \gamma I, \tag{2}$$

$$\frac{dR}{dt} = \gamma I. \tag{3}$$

Because $r = N - i - s$, the model can be reduced to two ODEs. Further, the variables can be normalized by the total population setting $s = S/N$, $i = I/N$ (and $r = R/N$); by rescaling time as $\tau = t\gamma$, the SIR model becomes non-dimensional, with a single coefficient $\mathcal{R}_0 = \beta/\gamma$, the well known reproduction ratio or reproduction number [1].

$$\frac{ds}{d\tau} = -\mathcal{R}_0 s i, \tag{4}$$

$$\frac{di}{d\tau} = (\mathcal{R}_0 s - 1) i. \tag{5}$$

It is well-known that the solutions are positive and satisfy the conservation law $s + i + r = 1$ [11], and exact expressions have been computed [10]. If there are no infected individuals ($i_0 = 0$), the system remains in the disease-free equilibrium $E_0 = (s_0, 0, 0)$ because all derivatives are identically zero. For any initial value of infections $i_0 > 0$, the solutions $s(t)$ and $i(t)$ are bounded and evolve in the invariant set $\mathcal{P} = \{s \leq s_0,\, i \leq 1,\, r \leq 1\}$. This follows from the fact that $ds/d\tau \leq 0$, so $s(\tau) \leq s_0$, $\forall \tau \geq \tau_0$. The solutions and the admissible equilibria depend on the value of $\mathcal{R}_0$ and on the initial value of the susceptible population s_0.

If $\mathcal{R}_0 s_0 < 1$, the infected population is non-increasing because $di/d\tau \leq 0$, thus the epidemic does not start (the system reaches a disease-free equilibrium $\tilde{E} = (\tilde{s}, 0, \tilde{r})$).

If $\mathcal{R}_0 s_0 > 1$, then $di/d\tau$ initially increases, reaches a peak when $s = s_{crit} = 1/\mathcal{R}_0 \leq s_0$, and finally decreases to zero. The equilibrium in this case is $E = (\bar{s}, 0, \bar{r})$. Because $s_0 \leq 1$, $\mathcal{R}_0 s_0 > 1 \Rightarrow \mathcal{R}_0 > 1$. For any positive i_0 and $\mathcal{R}_0 s_0 > 1$, the relation between susceptible and infected can be computed exactly from the ratio of $di/d\tau$ and $ds/d\tau$ (directional derivative) [11]:

$$\frac{di}{ds} = \frac{\mathcal{R}_0 s - 1}{-\mathcal{R}_0 s} = -1 + \frac{1}{\mathcal{R}_0 s} \;\Rightarrow\; di = -ds + \frac{ds}{\mathcal{R}_0 s}.$$

Integrating we obtain the relation between $i(\tau)$ and $s(\tau)$:

$$i(\tau) = i_0 + s_0 - s(\tau) - \frac{1}{\mathcal{R}_0} \log \frac{s_0}{s(\tau)}. \tag{6}$$

The peak of infections occurs when $s = s^* = 1/\mathcal{R}_0$ ($s = s^*$ yields $di/d\tau = 0$). Substituting s^* we find:

$$i_{max} = i_0 + s_0 - \frac{1}{\mathcal{R}_0} \left(1 + \log(s_0 \mathcal{R}_0)\right), \tag{7}$$

with $\log(s_0 \mathcal{R}_0) > 0$ because $s_0 \mathcal{R}_0 > 1$. From expression (6), by setting $\bar{i} = 0$, we can also derive an implicit equation to find the equilibrium value of the recovered population:

$$\log \frac{s_0}{\bar{s}} = \mathcal{R}_0 (1 - \bar{s}),$$

which has one positive root (because $\bar{s} < s_0 \leq 1$ and $\mathcal{R}_0 > 1$). In other words, the equilibrium susceptible population is positive (not all the population has become infected), unless $\mathcal{R}_0$ is extremely large.

A. Flattening the curve: a low reproduction number reduces and delays the infection peak

The SIR model has been used to illustrate how a low reproduction number $\mathcal{R}_0$ (or a low transmission rate β) has the effect of "flattening the (infection) curve", *i.e.* reducing the infection peak while lengthening the duration of the epidemic. The simulations in Fig. 1 compare the SIR solutions for values of $\mathcal{R}_0 = 2.5$, which is close to recent estimates for the COVID-19 outbreak [16], and $\mathcal{R}_0 = 1.5$. The infection peak is clearly reduced when $\mathcal{R}_0 = 1.5$, however the infection peak is also significantly delayed.

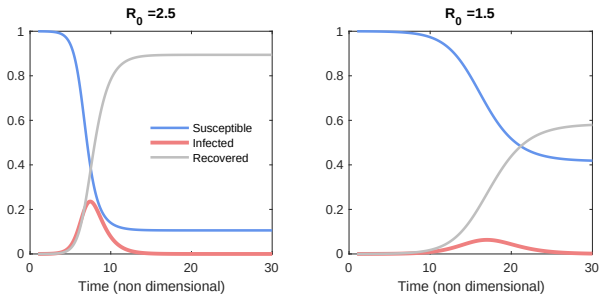

Fig. 1. Illustrative numerical simulations showing the SIR dynamics for different values of (constatn) transmission rate β. The right plot illustrates how a lower value of β "flattens the curve" while also significantly delaying the infection peak. This illustration may be misleading to the public because the introduction of social distancing causes the transmission rate β to vary in time. Further, this picture suggests that social distancing measures may have to be imposed for a very long time to be effective, but with a time-varying β this may not be necessary. This manuscript describes an SIR model in which distancing depends on the (known) infection levels, introducing feedback that changes the reproduction number as a function of time.

The problem with assuming a low, constant reproduction number $\mathcal{R}_0$ is that it is misleading to the public and to policymakers. A time-dependent reproduction number $\mathcal{R}(\tau)$ captures better a societal response in which disease awareness, social habits, and government policies may fluctuate over time. During the COVID-19 epidemic, enormous research efforts are dedicated to a continuous estimation and forecasting of the reproduction number as a function of social distancing measures [1], [15]. Here I report the derivation of a simple candidate model for $\mathcal{R}(\tau)$ that is based on a mean field model of social distancing, and may be helpful for illustrative purposes, to compare alternative scenarios of collective response, or to model and compare regional epidemic data.

III. The feedback SIR (fSIR) model

Social distancing policies suggested or officially imposed by government agencies typically depend on the reported infections or deaths. With fast spread of information about testing results [5], [18], [17], knowledge of *infections* may be more helpful in containing epidemics (the average time to death for COVID-19 patients, for example, is 17 days [20], which means average lethality information of an epidemic may be available with a significant delay).

We can model the average effects of social distancing by thinking about individuals and their interactions as molecules in a well mixed solution, and using the law of mass action in chemistry. A contagion may occur when a susceptible individual (S) and an infected individual (I) are in spatial proximity for some time (associated or contact state C); this encounter may then result in two infected individuals. This can be modeled using the equivalent reactions:

$$S + I \underset{\rho^-}{\overset{\rho^+}{\rightleftharpoons}} C \xrightarrow{\phi} 2I,$$

where ρ^+ and ρ^- are the rates of association and dissociation of a susceptible and an infected individual, and we can associate ϕ with the probability that infection occurs. The law of mass action yields an ODE for the density of individuals in the associated state C (the ODE is written in terms of non-dimensional variables normalized to the total population, all in lowercase letters):

$$\frac{dc}{dt} = \rho^+ s \cdot i - (\rho^- + \phi)c.$$

Because contacts occur on an hourly or daily basis, which is much faster than timescale of the epidemic, it is sensible to assume $dc/dt = 0$ and derive an expression for the equilibrium level of associated individuals:

$$\bar{c} = \frac{\rho^+}{\rho^- + \phi} s \cdot i.$$

This value of $\bar{c}$ is intended to represent a dynamic equilibrium at the population level, so it indicates the average number of contacts per day. With this definition, the transmission rate β introduced in model (1) is:

$$\beta = \phi \frac{\rho^+}{\rho^- + \phi},$$

where ϕ is the probability of infection per contact, and $\rho^+/(\rho^- + \phi)$ is the average number of contacts per day, a definition consistent with the literature. [a] The corresponding (non-dimensional) reproduction coefficient can be computed as earlier $\mathcal{R}_0 = \beta/\gamma$.

In the presence of infection awareness or policies that discourage or prevent association of individuals, the level of individuals in associated state C should decrease. A sensible "continuum" approximation of this phenomenon is to model an additional dissociation process that depends on the known infection levels:

$$C \xrightarrow{\psi I} S + I$$

With this model for dissociation, individuals in state c evolve according to the ODE:

$$\frac{dc}{dt} = \rho^+ s \cdot i - (\rho^- + \phi)c - \psi i \cdot c,$$

which equilibrates to:

$$\bar{c} = \left(\frac{\rho^+}{\rho^- + \phi} \right) \frac{1}{1 + \kappa i} s \cdot i, \quad \kappa = \frac{\psi}{\rho^- + \phi}.$$

[a] This definition of β can be verified by using the law of mass action to write the ODEs of s and i. For example

$$\frac{ds}{dt} = -\rho^+ s \cdot i - \rho^- c,$$

in which c has to be replaced by its equilibrium value $\bar{c}$.

With this equilibrium value for the average contacts, we derive a time-varying expression for the reproduction number that depends on the infection levels:

$$\mathcal{R}(i) = \mathcal{R}_0 \frac{1}{1+\kappa i}. \tag{8}$$

The coefficient κ is in units of /time/individual (or fraction of individuals, the equivalent of "copy number" or molar in chemical reaction networks). Thus $\mathcal{R}(i)$ is non-dimensional like $\mathcal{R}_0$.

Expression (8) resembles Michaelis-Menten/Hill functions in chemical kinetics, and indicates that under a policy in which social distancing depends on average on the infection levels, the reproduction number $\mathcal{R}(i)$ decreases as the infection numbers raise. One can think about the feedback term $1/(1+\kappa i)$ as a reduction of either the duration or frequency of infectious contacts (social distancing) or of the likelihood of infection through the use of personal protective equipment, which effectively reduces the level of infectiousness of a contact.

$\mathcal{R}(i)$ decreases monotonically as a function of i, and it decreases more steeply for large values of κ, as illustrated in Fig. 2. The larger κ, the smaller the value of i that induces a significant reduction in $\mathcal{R}_0$ (i.e. "social distancing" occurs in response to a very small known infection level). For example, a value of $\kappa = 2$ results in $\mathcal{R}(i) = \mathcal{R}_0/2$ when $i = 0.5$; a value of $\kappa = 10$ cuts in half $\mathcal{R}_0$ much sooner, when $i = 0.1$.

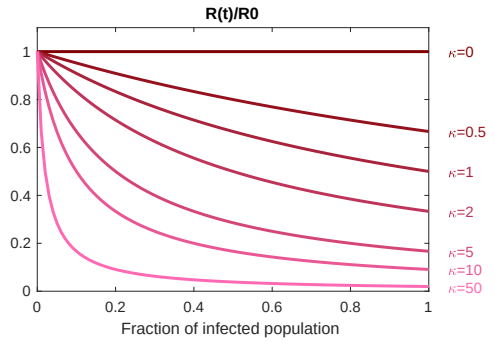

Fig. 2. The infection-dependent coefficient $\lambda(i) = 1/(1+\kappa i)$ reduces the reproduction number (8) for any choice of $\kappa \geq 0$.

With this definition, the non-dimensional fSIR model is:

$$\frac{ds}{d\tau} = -\mathcal{R}_0 \frac{1}{1+\kappa i} si = -\mathcal{R}(i) si \tag{9}$$

$$\frac{di}{d\tau} = \left(\mathcal{R}_0 \frac{1}{1+\kappa i} s - 1 \right) i = (\mathcal{R}(i)s - 1) i \tag{10}$$

The coefficient κ models the average population response to knowledge of current infection numbers, in relation to typical interaction patterns; this coefficient could also be used to model the collective "trust" in infection information. For $\kappa = 0$, i.e. there is no reaction/policy, nor trust on infection data, then $\mathcal{R}(i(\tau)) = \mathcal{R}_0$. (Similarly, if there are no infections and $i(\tau) = 0$, then for any value of κ we have no change in $\mathcal{R}(i(\tau)) = \mathcal{R}_0$).

The time varying reproduction number $\mathcal{R}(i)$ introduces a negative feedback loop in the epidemic model, because captures the fact that society decreases interactions in response to an increase of infections, thereby reducing the reproduction number. The model captures also the fact that a society is expected to return to typical interaction patterns when infections are not present.

IV. Structural properties of the fSIR model

A. Analysis of equilibria

The fSIR model (9) inherits the equilibrium properties of the original SIR model. If $i_0 = 0$ ($r_0 = 0$), the system remain at the disease-free equilibrium $E_0 = (s_0, 0, 0)$ because all derivatives are identically zero. For any $i_0 > 0$, the solutions are bounded and evolve in the invariant set $\mathcal{P} = \{s \leq s_0, i \leq 1, r \leq 1\}$. If $\mathcal{R}_0 s_0 \leq 1 + \kappa i_0$, the infected population is non-increasing because $di/d\tau \leq 0$, the epidemic does not start and the system reaches the disease-free equilibrium $\tilde{E} = (\tilde{s}, 0, \tilde{r})$. Like in the SIR model, because $s_0 \leq 1$, for the epidemic to start it is necessary that $\mathcal{R}_0 > 1 + \kappa i_0$.

If $\mathcal{R}_0 s_0 > 1 + \kappa i_0$, $di/d\tau > 0$ until the susceptible population decreases to the value $s = s_{crit} = (1 + \kappa i_{max})/\mathcal{R}_0 > \mathcal{R}_0$ at which $i(\tau) = i_{max}$. As the susceptible population continues to decrease, so does the infected population and the system reaches the disease-free equilibrium $E = (\bar{s}, 0, \bar{r})$.

Proposition 1 *Assume $\mathcal{R}_0 s_0 > 1 + \kappa i_0$. The disease-free equilibrium $E = (\bar{s}, 0, \bar{r})$ is structurally stable.*

Proof The Jacobian of the fSIR model is:

$$J = \mathcal{R}_0 \begin{bmatrix} -\frac{\bar{i}}{1+\kappa\bar{i}} & -\frac{\bar{s}}{(1+\kappa\bar{i})} \\ \frac{\bar{i}}{1+\kappa\bar{i}} & \frac{\bar{s}}{(1+\kappa\bar{i})^2} - \frac{1}{\mathcal{R}_0} \end{bmatrix}. \tag{11}$$

At the disease-free equilibrium $E = (\bar{s}, 0, \bar{r})$, J is identical to the Jacobian of the SIR model:

$$J_0 = \begin{bmatrix} 0 & -\bar{s} \\ 0 & \mathcal{R}_0\bar{s} - 1 \end{bmatrix},$$

which is a structurally stable matrix (at equilibrium $\mathcal{R}_0\bar{s} < 1$). □

B. Analysis of the solutions

The peak of infections in the fSIR model is *structurally* smaller than the infection peak for the SIR model, for any $\kappa > 0$.

Problem 1 *The fSIR model (9) with initial conditions s_0, i_0, r_0 and $s_0\mathcal{R}_0 > 1 + \kappa i_0$ defines an initial value*

problem (IVP) with non-negative solutions. We look for properties of the solutions of this IVP when $\kappa > 0$ that hold for any $\mathcal{R}_0$ and are therefore structural. These properties will be contrasted to the limit case $\kappa = 0$ that corresponds to the IVP defined by the SIR model (4)*. The solution for $\kappa = 0$ as well as its features will be denoted with the superscript ^{0} (*e.g. *$i(\tau) = i^0(\tau)$ if $\kappa = 0$).*

Proposition 2 *In Problem 1, for any $\mathcal{R}_0$ and for any $\kappa > 0$, we have:*

$$i_{max} < i^0_{max}.$$

Proof Following the same approach used to derive (7), the peak of infection for the fSIR model can be estimated from the directional derivative:

$$\frac{di}{ds} = -1 + \frac{1+\kappa i}{\mathcal{R}_0 s}.$$

We then obtain the infinitesimal expression:

$$di = -ds + \frac{ds}{\mathcal{R}_0 s} + \frac{\kappa}{\mathcal{R}_0}\frac{i\,ds}{s}, \tag{12}$$

in which the last term cannot be easily integrated, but it can be replaced by a simpler expression. We note that $dr = id\tau$, therefore $di = -ds - id\tau$. At the same time we just showed that:

$$di = -ds + \frac{(1+\kappa i)}{\mathcal{R}_0}\frac{ds}{s}, \tag{13}$$

which means

$$id\tau = \frac{(1+\kappa i)}{\mathcal{R}_0}\frac{ds}{s}.$$

Rearranging terms, we find that

$$\frac{1}{\mathcal{R}_0}\frac{ds}{s} = -\frac{i}{1+\kappa i}d\tau,$$

which can be substituted in equation (12):

$$di = -ds + \frac{ds}{\mathcal{R}_0 s} - i\frac{\kappa i}{1+\kappa i}d\tau,$$

thus we obtain the expression:

$$i(\tau) = s_0 + i_0 - s + \frac{1}{\mathcal{R}_0}\log\frac{s}{s_0} - \int_0^\tau i\frac{\kappa i}{1+\kappa i}d\sigma, \tag{14}$$

The infection peak occurs at $s_{crit} = (1+\kappa i_{max})/\mathcal{R}_0$, which can be substituted in (14):

$$i_{max}(\tau_{max}) = s_0 + i_0 - \frac{1+\kappa i_{max}}{\mathcal{R}_0} + \tag{15}$$
$$+ \frac{1}{\mathcal{R}_0}\log\left(\frac{1+\kappa i_{max}}{\mathcal{R}_0 s_0}\right) - \int_0^{\tau_{max}} i\frac{\kappa i}{1+\kappa i}d\sigma.$$

When $\kappa = 0$ we recover the original SIR infection peak expression (7). The difference between the peak value (15) and the peak when $\kappa = 0$ (i^0_{max}) is:

$$i_{max} - i^0_{max} = -\frac{1}{\mathcal{R}_0}\left(\kappa i_{max} - \log(1+\kappa i_{max})\right) -$$
$$- \int_0^{\tau_{max}} i\frac{\kappa i}{1+\kappa i}d\sigma.$$

Because $\log(1+x) < x$ for any $x > 0$, and because the last integral is strictly positive, we conclude that $i_{max} < i^0_{max}$ for any $\kappa > 0$. □

Corollary 1 *In Problem 1, the time τ_{max} at which the peak of infection occurs is always larger than the peak time τ^0_{max} corresponding to $\kappa = 0$:*

$$\tau_{max} > \tau^0_{max}. \tag{16}$$

Proof We show that $d\tau_{max}/d\kappa > 0$ for any $\kappa > 0$, which means that the peak time can only increase as κ increases. First, recall that $s(\tau_{max}) = s_{crit} = (1+\kappa i_{max})/\mathcal{R}_0$. Then:

$$\frac{ds(\tau)}{d\kappa}_{|\tau=\tau_{max}} = \frac{\kappa}{\mathcal{R}_0}\frac{di_{max}}{d\kappa}$$

Using the chain rule $ds/d\kappa = (ds/d\tau)(d\tau/d\kappa)$. At $\tau = \tau_{max}$, $ds/d\tau = -i_{max}$, therefore:

$$\frac{d\tau_{max}}{d\kappa} = -\frac{\kappa}{i_{max}\mathcal{R}_0}\frac{di_{max}}{d\kappa}$$

From Proposition 2, we know that $di_{max}/d\kappa < 0$ and we conclude that $d\tau_{max}/d\kappa > 0$. □

Corollary 2 *In Problem 1, the solution $s(\tau)$ is always lower bounded by the solution $s^0(\tau)$. In particular the equilibrium satisfies $\bar{s} > \bar{s}^0$.*

Proof First note that $ds/d\tau$ is always negative. Then we have that:

$$\frac{ds}{d\tau} = -\mathcal{R}_0\frac{1}{1+\kappa i}si > -\mathcal{R}_0 s^0 i^0,$$

since the function $\frac{1}{1+\kappa i}$ is strictly less than one for any positive value of κ and i. Because this inequality holds for arbitrary values of i, we can invoke the comparison principle [13]:

$$\frac{ds^0}{d\tau} < \frac{ds}{d\tau} \Rightarrow s^0(\tau) < s(\tau).$$

In the limit for $\tau \to \infty$, the steady state values satisfy $\bar{s} > \bar{s}^0$. □

This proposition shows that, relative to an epidemic that lacks negative feedback, the fSIR model structurally settles to a larger susceptible population in the disease-free equilibrium. As a consequence, the equilibrium recovered population satisfies $\bar{r} < \bar{r}^0$.

Corollary 3 *The solution $i(\tau)$ in Problem 1 is upper bounded by the solution $i^0(\tau)$ for $0 < \tau < \tau^0_{max}$, where τ^0_{max} is the time at which the maximum i^0_{max} occurs.*

Proof For $0 < \tau < \tau_{max}$ we know that $di/d\tau > 0$, so $i(\tau)$ is monotonically increasing for any value of κ including $\kappa = 0$. Due to Proposition 2, $i_{max} < i^0_{max}$ for any value of $\kappa > 0$. Due to monotonicity of the solutions, it must be that $i(\tau) < i^0(\tau)$ for all $0 < \tau < \min(\tau_{max}, \tau^0_{max})$. Because Proposition 1 shows that $\tau^0_{max} < \tau_{max}$ for all $\kappa > 0$, we conclude that $i(\tau) < i^0(\tau)$ for all $0 < \tau < \tau^0_{max}$. □

We conclude with a qualitative observation on the fSIR solutions (Problem 1) when κ is very large, thus $\kappa i \gg 1$. In this case, the fSIR can be approximated by the linear system:

$$\frac{d\hat{s}}{d\tau} \approx -\frac{\mathcal{R}_0}{\kappa}\hat{s}, \qquad \frac{d\hat{i}}{d\tau} \approx \frac{\mathcal{R}_0}{\kappa}\hat{s} - \hat{i}. \tag{17}$$

The solution $\hat{i}(\tau)$ can be found exactly:

$$\hat{i}(\tau) = i_0 e^{-\tau} + s_0 \frac{\mathcal{R}_0}{\kappa - \mathcal{R}_0}\left(e^{-\tau} - e^{-\frac{\mathcal{R}_0}{\kappa}\tau}\right).$$

This approximation shows that if $\mathcal{R}_0/\kappa \ll 1$ the infection dynamics converge very slowly to the disease free equilibrium $\hat{i} = 0$ (convergence is dominated by the constant $\mathcal{R}_0/\kappa$).

V. Numerical analysis

A. Infection-aware distancing reduces the peak of infection and does not postpone the peak significantly

Fig. 3, top, shows the numerically integrated solution of the fSIR model (9) with $\mathcal{R}_0 = 2.5$ as the parameter κ is varied. ($\mathcal{R}_0 = 2.5$ corresponds to a choice of $\beta = 0.25$ and $\gamma = 1/10$, i.e. the average time to recovery or death assumed to be 10 days; for comparison, the estimated average time to recovery in the COVID-19 epidemic is about 17 days for hospitalized patients [20]). These simulations confirm that the peak of infections decreases with a large κ, relative to the case $\kappa = 0$ (SIR without feedback). Fig. 3, bottom, shows the temporal evolution of the reproduction number in each simulation in the top panel: when infections increase, $\mathcal{R}(\tau)$ decreases and adjusts to the nominal level ($\mathcal{R}_0 = 2.5$) when infections are not present. These simulations also confirm the results of Propositions 2, and 1, as the infection peak is always reduced and delayed: Fig. 4 shows that a feedback parameter $\kappa = 2$ (taken as an illustrative example) the infection peak size can be reduced by about 30%, but this also causes a 30% extension of the time during which more than 2.5% of the population is infected.

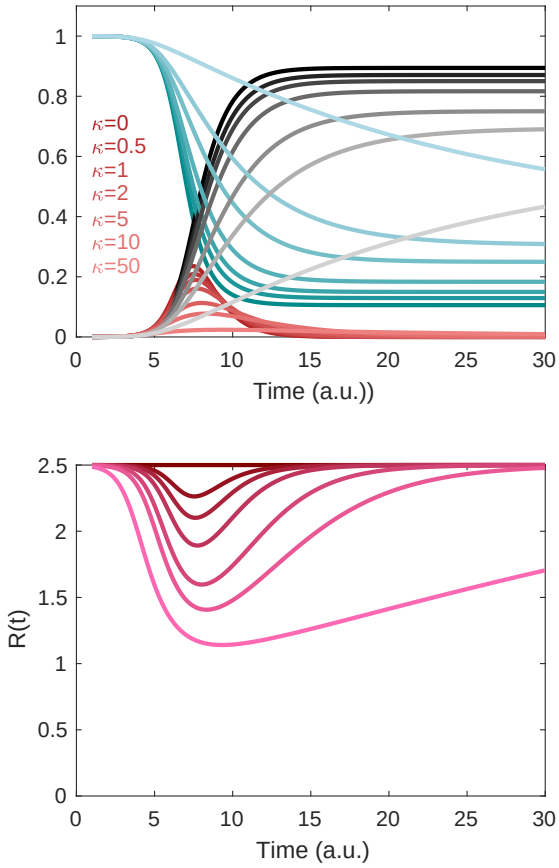

Fig. 3. Numerically integrated solutions of the fSIR model. Top: Susceptible (green), infected (red), and recovered (gray) individuals when the parameter κ is varied (low to high, color shades from dark to light). Bottom: Evolution of the reproduction number in time computed from the simulations above; this can be interpreted as a qualitative measure of the implemented social distancing policies.

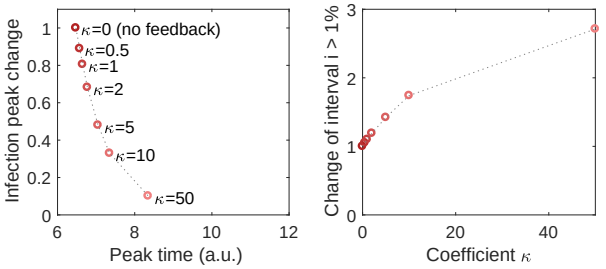

Fig. 4. Left: Peak time versus peak value of infections for different values of the feedback parameter κ. This plot evidences that the peak is not delayed as in models where the transmission rate is constant and low. Right: The duration of infections is longer in the presence of feedback; here it is measured as the time interval for which the fraction of infected individuals is larger than 2.5% of the population.

B. Effects of delayed infection awareness

Here I examine whether a delay Δ in obtaining infection information can compromise the effects of feedback. A delay is included in the transmission rate expression:

$$\frac{ds}{d\tau} = -\mathcal{R}(i)si, \qquad \mathcal{R}(i) = \mathcal{R}_0 \frac{1}{1+\kappa i(\tau-\Delta)}, \quad (18)$$

$$\frac{di}{d\tau} = (\mathcal{R}(i)s - 1)i. \quad (19)$$

For illustrative purposes, I choose a feedback parameter $\kappa = 2$ that remains fixed in these simulations, with $\mathcal{R}_0 = 2.5$ ($\beta = 0.25$ and $\gamma = 1/10$). Fig. 5 shows that a delay of up to 7 days increases the peak by less than 10%, but a 14 day delay causes a 25% increase in the peak, offsetting the peak reduction obtained by introducing feedback (the simulated non-dimensional delay is divided by the rescaling constant $\gamma = 1/10$).

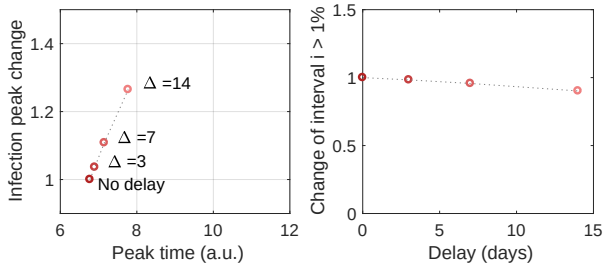

Fig. 5. Effects of delays on the peak size and duration. Left: Change in peak size in the presence of delays, relative to the case in which feedback is present without delay and $\kappa = 2$. Right: The amount of time for which the fraction of infected population exceeds 2.5% is slightly reduced when delays are between 0 and 14 days.

1) The fSIR model captures current COVID-19 infection trends: The fSIR model was fitted to COVID-19 temporal series data for infections, recoveries, and deaths available from the Johns Hopkins Github repository [6]. I selected data from six western democracies (Italy, France, UK, Spain, Germany and the US) with comparable infection and recovery reporting patterns. In addition, in these countries strict social distancing measures were not *immediately* enforced during the early stages of the epidemic, thus it is reasonable to use a compartment model. Further, the reported infections and deaths in these countries have a similar doubling time of 2-4 days in the early (exponential) stages [1]. Finally infection, recovery, and deaths reports are unlikely to have been systematically manipulated, although there are substantial differences in the protocols for infection and post-mortem testing. Because all these countries enforced social distancing measures at different times during the local evolution of the pandemic, the fSIR model should capture these differences in the fitted parameter κ. Yet, we should keep in mind that the population of infected individuals is largely underestimated (worldwide) due to lack of testing resources.

Fitting results are in Fig. 6. The data were processed to compute active infections in a given day, and recoveries and deaths were summed and consolidated into the "recovered" compartment. All data were normalized by country population and thresholded to include only data collected after infections exceed two per million. Parameters were fitted with constraints $\beta \in [0 \quad 0.6]$, $\beta \in [1/17 \quad 1/10]$, and $\kappa \in [0 \quad 10 \cdot 10^3]$; in the fitting score function, the infection prediction error was assigned a 100-fold penalty relative to the recovery data, with the expectation that recoveries may not be accurately reported for non-hospitalized patients; as a consequence, the fitted model reproduces infection data much more closely than recovery data. (An SIR model fits the same data very poorly, as it unrealistically assumes a constant transmission rate β.)

If the fSIR model is fitted to data scaled by X-fold (*i.e.* infections and recoveries are believed to be X-times larger than reported), the fitted κ qualitatively scales by a factor $1/X$, while changes in fitted β and γ are negligible.

The fitted value of κ can be interpreted as the inverse of the infections observed when the reproduction number is half its initial value. Thus, low fitted values of κ indicate that a country enacted infection-dependent social distancing late in the epidemic (higher infection numbers). Unsurprisingly, among the countries considered here the highest κ is fitted for Germany, and the lowest for the UK and the US. The fact that overall the model fits poorly infection data from Germany points to the fact that their management of $\mathcal{R}(t)$ may not be based on mere infection numbers, and may have been substantially different from other countries from the very initial stages, resulting in lower infection levels and lower mortality.

The data fitting reported here has largely an illustrative purpose, and is not meant to make predictions. However, from this exercise it is clear that a protocol of social distancing based on infection awareness (enforce distancing when infections are high, relax distancing when infections decrease) does reduce the peak of infections, but it can also considerably lengthen the duration of the epidemic as shown by the linear approximation (17) and by simulations in Fig. 5. With this in mind, in Fig. 7 we show a projection of infections for Spain and Italy based on the fitted fSIR parameters. Although the peak of infections appears near, infection numbers may remain roughly constant for years to come, consistently with the observations made on the fSIR linear approximation (17). The computed reproduction number remains very close or slightly above

one throughout the projected simulation. This may be or may not be a reasonable societal target, depending on considerations that balance the sustainability of a long-term lockdown with the emergence of therapeutic advances and vaccines that may reduce mortality rates and improve recovery time, making it possible for a healthcare system to absorb a bounded number of infections. Similar predictions highlighting the pitfalls of relaxing social distancing are put forward by many models that are more complex and accurate than the one presented here [15], [8].

VI. Conclusion and Discussion

I have presented the derivation and structural analysis of a modified SIR model, here named feedback SIR (fSIR), in which the reproduction number is reduced as a continuous function of infection levels generating a negative feedback loop. The model is derived from first principles assuming a well-mixed population and assuming individuals reduce their contacts the more infections are reported, and using a standard time-scale separation argument the transmission rate function takes the form of a Michaelis-Menten or Hill function that are widely adopted in chemistry and biology. This model requires only one additional parameter to capture the effects of social distancing, and illustrates the tradeoffs between reducing the infection peak while maintaining a short epidemic course.

The reduction of infection peak and other structural properties of the fSIR model presented here can be extended to the case of feedback based on death information, rather than infections [2]. This is trivial if we assume a constant mortality rate, thus deaths are linearly proportional to infections.

A Michaelis-Menten expression modifying transmission rates in a SEIR model was previously adopted by Bootsma and Ferguson to model the 1918 influenza epidemic in the United States [2], in which social interactions were thought to be decreased due to awareness of *death numbers* $M(t)$ rather than infections:

$$\lambda(t) = \beta(t)\frac{\kappa_m}{\kappa_m + M(t)}$$

The model by Bootsma and Ferguson and others [14], [3], [9] introduce coefficient κ_m as a *threshold* before societal reaction (reduction of contact rates), as opposed to a scaling factor, which I believe is more convenient to describe a top-down feedback policy. The limit $\kappa_m \to \infty$ corresponds to $\kappa = 0$ (no feedback) in the model presented here.

The fSIR model could be modified to assume a time-varying κ, even with discrete variations, to better capture enforced lockdown or distancing protocols that drastically change in time, as recently proposed by Casella [4]; the structural analysis presented here serves however to clarify that *any* positive κ would reduce and postpone the infection peak, as well as extend the epidemic.

Finally, it must be noted that the SIR model is not suited to capture epidemics with a long incubation time, a large population of asymptomatic individuals, and high lethality. For these reasons, many SIR variants with additional compartments have been developed and tailored to model specific epidemic outbreaks [2], [8]. We conjecture that structural results similar to those presented here may hold for other SIR variants modified to include negative feedback on infection or death information.

VII. Methods

Differential equations were integrated with a forward Euler method in MATLAB using custom scripts, or using MATLAB's `ode45`. Data fitting was done using MATLAB's `fmincon`.

References

[1] Andrea L Bertozzi, Elisa Franco, George Mohler, Martin B Short, and Daniel Sledge. The challenges of modeling and forecasting the spread of COVID-19. *arXiv preprint arXiv:2004.04741*, 2020.

[2] Martin CJ Bootsma and Neil M Ferguson. The effect of public health measures on the 1918 influenza pandemic in US cities. *Proceedings of the National Academy of Sciences*, 104(18):7588–7593, 2007.

[3] Bruno Buonomo, Alberto dOnofrio, and Deborah Lacitignola. Global stability of an SIR epidemic model with information dependent vaccination. *Mathematical biosciences*, 216(1):9–16, 2008.

[4] Francesco Casella. Can the COVID-19 epidemic be managed on the basis of daily data? *arXiv preprint arXiv:2003.06967*, 2020.

[5] Ensheng Dong, Hongru Du, and Lauren Gardner. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet*, 2020. https://plague.com/.

[6] Ensheng Dong, Hongru Du, and Lauren Gardner. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet infectious diseases*, 2020.

[7] Sebastian Funk, Erez Gilad, Chris Watkins, and Vincent AA Jansen. The spread of awareness and its impact on epidemic outbreaks. *Proceedings of the National Academy of Sciences*, 106(16):6872–6877, 2009.

[8] Giulia Giordano, Franco Blanchini, Raffaele Bruno, Patrizio Colaneri, Alessandro Di Filippo, Angela Di Matteo, and Marta Colaneri. Modelling the COVID-19 epidemic and implementation of population-wide interventions in italy. *Nature Medicine*, pages 1–6, 2020.

[9] David Greenhalgh, Sourav Rana, Sudip Samanta, Tridip Sardar, Sabyasachi Bhattacharya, and Joydev Chattopadhyay. Awareness programs control infectious disease–multiple delay induced mathematical model. *Applied Mathematics and Computation*, 251:539–563, 2015.

[10] Tiberiu Harko, Francisco S. N. Lobo, and M. K. Mak. Exact analytical solutions of the susceptible-infected-recovered (SIR) epidemic model and of the SIR model with equal death and birth rates. *Applied Mathematics and Computation*, 236:184194, 2014.

[11] Herbert W Hethcote. Qualitative analyses of communicable disease models. *Mathematical Biosciences*, 28(3-4):335–356, 1976.

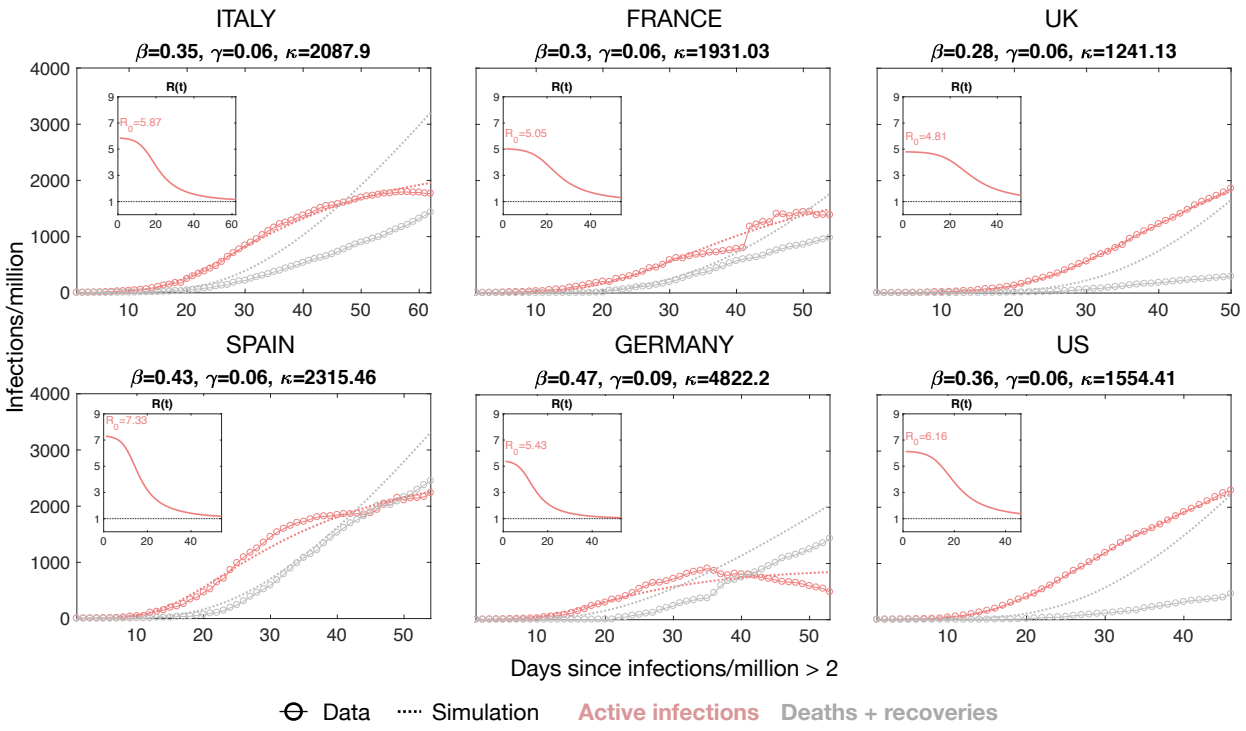

Fig. 6. The fSIR model parameters β, γ, and κ were fitted to COVID-19 data of active infections as well as recoveries and deaths for six different countries up to April 25, 2020. Data were obtained from the JHU Github repository [6]. A 100-fold fitting penalty was assigned to infection data relative to recoveries. The model captures well most infection data with the exception of Germany, which presents the highest fitted κ indicating a stronger distancing in relation to infection information. Results are discussed in Section V-B.1.

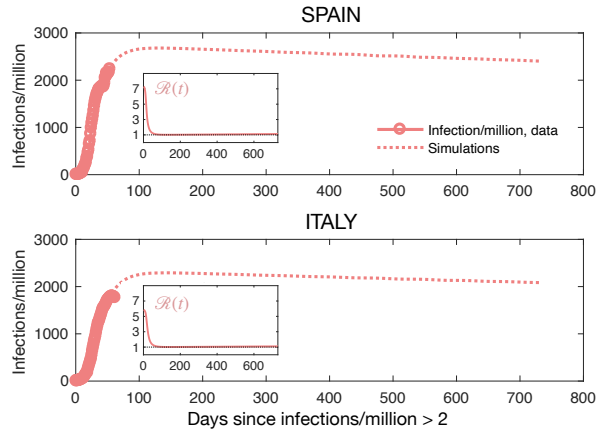

Fig. 7. Theoretical projection of infection data using the fitted fSIR model for Spain (top) and Italy (bottom); inlet shows predicted $\mathcal{R}(t)$. This simulation is not meant to provide accurate predictions, rather to illustrate the concept that feedback based purely on infection-aware distancing may significantly extend the epidemic according to the very simple model presented here. Adjustments of the coefficient κ may overcome this problem.

[12] William Ogilvy Kermack and Anderson 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(772):700–721, 1927.

[13] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, 2002.

[14] Istvan Z Kiss, Jackie Cassell, Mario Recker, and Péter L Simon. The impact of information transmission on epidemic outbreaks. *Mathematical biosciences*, 225(1):1–10, 2010.

[15] Stephen M Kissler, Christine Tedijanto, Edward Goldstein, Yonatan H Grad, and Marc Lipsitch. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science*, 2020.

[16] Adam J Kucharski, Timothy W Russell, Charlie Diamond, Yang Liu, John Edmunds, Sebastian Funk, and Rosalind M Eggo. Early dynamics of transmission and control of covid-19: a mathematical modelling study. *The Lancet, Infectious Diseases*, 2020. March 11, 2020.

[17] Bastian Prasse, Massimo A. Achterberg, Long Ma, and Piet Van Mieghem. Network-based prediction of the 2019-ncov epidemic outbreak in the chinese province hubei, 2020. https://arxiv.org/pdf/2002.04482.pdf.

[18] Umberto Rosini. COVID-19 Italia - Monitoraggio situazione, 2020. GitHub repository of data from Italy COVID-19 epidemic https://github.com/pcm-dpc/COVID-19.

[19] Sudip Samanta and Joydev Chattopadhyay. Effect of awareness program in disease outbreak–a slow–fast dynamics. *Applied Mathematics and Computation*, 237:98–109, 2014.

[20] Fei Zhou, Ting Yu, Ronghui Du, Guohui Fan, Ying Liu, Zhibo Liu, Jie Xiang, Yeming Wang, Bin Song, Xiaoying Gu, et al. Clinical course and risk factors for mortality of adult inpatients with covid-19 in wuhan, china: a retrospective cohort study. *The Lancet*, 2020.