Mathematical Biology

Control with uncertain data of socially structured compartmental epidemic models

Giacomo Albi[1] · Lorenzo Pareschi[2] · Mattia Zanella[3]

Abstract

The adoption of containment measures to reduce the amplitude of the epidemic peak is a key aspect in tackling the rapid spread of an epidemic. Classical compartmental models must be modified and studied to correctly describe the effects of forced external actions to reduce the impact of the disease. The importance of social structure, such as the age dependence that proved essential in the recent COVID-19 pandemic, must be considered, and in addition, the available data are often incomplete and heterogeneous, so a high degree of uncertainty must be incorporated into the model from the beginning. In this work we address these aspects, through an optimal control formulation of a socially structured epidemic model in presence of uncertain data. After the introduction of the optimal control problem, we formulate an instantaneous approximation of the control that allows us to derive new feedback controlled compartmental models capable of describing the epidemic peak reduction. The need for long-term interventions shows that alternative actions based on the social structure of the system can be as effective as the more expensive global strategy. The timing and intensity of interventions, however, is particularly relevant in the case of uncertain parameters on the actual number of infected people. Simulations related to data from the first wave of the recent COVID-19 outbreak in Italy are presented and discussed.

Keywords Epidemic modelling · Uncertainty quantification · Social structure · Optimal control · Non-pharmaceutical interventions · COVID-19

✉ Mattia Zanella
  mattia.zanella@unipv.it

  Giacomo Albi
  giacomo.albi@univr.it

  Lorenzo Pareschi
  lorenzo.pareschi@unife.it

[1] Department of Computer Science, University of Verona, Verona, Italy

[2] Department of Mathematics and Computer Science, University of Ferrara, Ferrara, Italy

[3] Department of Mathematics "F. Casorati", University of Pavia, Pavia, Italy

Mathematics Subject Classification 92D30 · 34H05 · 92D25 · 49N90 · 60-08

1 Introduction

From the digital tracking systems of the Koreans to UK's initial choice of not wanting to do anything to counter the spread of the virus, passing through the militarized quarantines of the Chinese and those less authoritarian and more involved of the Italians, the reaction of different countries to the COVID-19 outbreak has shown a series of very different approaches that can also be explained considering the different cultural and political attitudes of the countries concerned.

In all cases, after an initial phase, even those governments that were less restrictive in the face of the pandemic's inexorable progress had to take strong containment measures. There's a graph that has become the symbol of the COVID-19 pandemic most of all. It shows in a simple and intuitive way the importance of slowing down the spread of an epidemic as much as possible ("flattening the curve"), so that the healthcare system can take care of all the sick without collapsing. Its success has helped to save many lives, raising awareness of good practices to slow down an epidemic: stay as much as possible at home, reduce social interactions and wash your hands often and well.

These "non-pharmaceutical" intervention measures, however, entail significant social and economic costs and thus policy makers may not be able to maintain them for more than a short period of time. Therefore, a modeling approach based on a limited time horizon that takes into account the social structure of the population is necessary in order to optimize containment strategies. Most current research, however, has focused on control procedures aimed at optimizing the use of vaccinations and medical treatments (Barro et al. 2018; Bolzoni et al. 2017; Lee et al. 2010; Colombo and Garavello 2019) and only recently the problem has been tackled from the perspective of non-pharmaceutical interventions (Lin et al. 2010; Morris et al. 2020). In addition, data collected by governments are often incomplete and heterogeneous, so a high degree of uncertainty needs to be incorporated into predictive models (Capaldi et al. 2012; Capistrán et al. 2012; Chowell 2017; Jagodnik et al. 2020; Mizumoto et al. 2020; Roberts 2013). This is the case of the spreading of COVID-19 worldwide, which have been often mistakenly underestimated due to a combination of factors, including deficiencies in surveillance and diagnostic capacity, and the large number of infectious but asymptomatic individuals (Jagodnik et al. 2020; Mizumoto et al. 2020; Zhang et al. 2020).

For almost a hundred years, mathematical models have been used to describe the spread of epidemics (Kermack and McKendrick 1927). The models currently used largely originate from the model proposed by Kermack and McKendrick at the beginning of last century. Even if the model contains strong simplification assumptions, the concepts introduced through this model are essential to provide a first intuition on the dynamics of epidemics, an intuition that remains confirmed in more complex models, albeit with numerous modifications (see for example Hethcote 2000; Capasso and Serio 1978). The model provides for the division of the population into compartments, the susceptible, healthy individuals who may be infected, the infectious, who

have already contracted the disease and can transmit it, and the removed, compartment that includes those who are healed and immune.

The hypothesis made by Kermack and McKendrick is that of the homogeneous "mixing"; that is, it is assumed that each individual has the same probability of contacting any other individual in the population. One understands how this hypothesis is unrealistic: we are often in contact with people from our family, our workplace, school class, group of friends and very rarely with those who live in a different place, have different ages and professions. In recent years, therefore, computational models have been developed that try to take into account additional social characteristics of individuals in order to arrive at more accurate predictions by keeping, however, the simplicity of compartmental models (Castillo-Chavez et al. 1989; Glasser et al. 2012; Lee et al. 2012; Hethcote 1996; Iannelli et al. 1992; Franceschetti and Pugliese 2008).

In this paper starting from a general compartmental model with social structure, typically the age dependence, we consider the external action of a policy maker that aims at reducing the spread of the epidemics by applying non pharmaceutical intervention measures, such as social distancing and quarantine. The mathematical problem is formulated as an optimal control problem characterized by a functional cost whose objective is to minimize the number of infectious people in a given time horizon. Through an instantaneous control strategy we compute an explicit feedback control that allows us to derive new SIR-type compartmental models capable of describing the epidemic peak reduction. Previously, this type of approach has been used successfully in the case of social models of consensus (Albi and Pareschi 2018; Albi et al. 2014, 2015a, b; Bongini et al. 2015; Caponigro et al. 2013; Düring et al. 2018).

The feedback controlled models are subsequently extended to take into account the presence of uncertain infection parameters and data. In fact, to have reliable forecasts it is of paramount importance to consider the presence of uncertain quantities as a structural feature of the epidemic dynamics. This aspect is of paramount importance in the case of pandemic COVID-19, in which undetected infectious individuals play a key role in the spread of the disease. In this regard, it is worth noting that our methodology can be easily extended to the case of more complex compartmental models. The decision to limit ourselves to a simple SIR-type compartmentalization was related on the one hand to the increased complexity given by the dependence on the social structure, which proved to be crucial in the case of the COVID-19 pandemic, and on the other hand to the introduction of a systematic uncertainty in the number of infected to avoid a complex sub-compartmentalization of the infectious population and the consequent difficulties due to parameter identification and the inability to follow a data-driven approach (Giordano et al. 2020; Flaxman et al. 2020). From a mathematical point of view, we can rely on the methods of uncertainty quantification (UQ) to obtain efficient and accurate solutions based on stochastic orthogonal polynomials for the differential model with random inputs (Xiu 2010).

Few results are actually available regarding methods of UQ in epidemic systems, we mention in this direction (Capistrán et al. 2012; Chowell 2017; Capaldi et al. 2012; Roberts 2013). The main idea is to increase the dimensionality of the problem adding the possible sources of uncertainty from the very beginning of the modeling. Hence, we extrapolate statistics by looking at the so-called quantities of interest, i.e. statistical quantities that can be obtained from the solution and that give some global information

with respect to the input parameter like expected solution of the problem or higher order moments. Several techniques can be adopted for the approximation of the quantities of interest, in this paper we adopt stochastic Galerkin methods that allow to reduce the problem to a set of deterministic equations for the numerical evaluation of the solution in presence of uncertainties. Compared to conventional Monte Carlo methods, based on stochastic sampling, these methods guarantee an exponential convergence in the case of smooth uncertainty distributions and allow a much more accurate and efficient estimation of random parameters. We refer the interested reader to recent surveys and monographs on the topic (Dimarco et al. 2017; Jin and Pareschi 2017; Pareschi 2021; Xiu 2010).

In particular, we consider the case in which the policy maker applies his control based on several possible estimators on the actual number of infected people. The need for long-term interventions shows that alternative actions based on the social structure of the system can be as effective as the more expensive optimal strategy. The importance of the timing and intensity of interventions is particularly relevant in the case of uncertain parameters on the actual number of infected people.

The rest of the manuscript is organized as follows. In Sect. 2 we introduce the structured social SIR model and formulate the mathematical approach for containment measures to reduce the spread of the disease. Next, a feedback controlled model used in the subsequent analysis is derived within a short time horizon approximation. In Sect. 3 we generalize the feedback controlled model to take into account the presence of uncertainties. Section 4 is dedicated to the presentation of some numerical examples including applications to the first wave of the COVID-19 epidemic in Italy. In separate Appendices we provide details on the generalizations of the present approach to more realistic epidemic models for COVID-19 including additional compartmentalizations, on the stochastic Galerkin method employed to efficiently address the uncertainties, and on the social interaction matrices characterizing the contact rates.

2 Control of epidemic dynamics

The starting model in our discussion is a SIR-type compartmental model with a social structure. The presence of a social structure is in fact essential in deriving appropriate sustainable control techniques from the population for a protracted period, as in the case of the recent COVID-19 epidemic. We will discuss in Sect. 3 how to modify the model through the introduction of a stochastic parameter that takes into account the dependence on uncertain data, and thus implicitly introduce the role of undetected infectious in the dynamics.

2.1 Compartmental models with social structure

The heterogeneity of the social structure, which impacts the diffusion of the infective disease, is characterized by the vector $\mathbf{a} \in \Lambda \subseteq \mathbb{R}^{d_a}$ characterizing its social state and whose components summarize, for example, the age of the individual, its number of social connections or its economic status (Hethcote 1996, 2000). We denote by

$s(\mathbf{a}, t)$, $i(\mathbf{a}, t)$ and $r(\mathbf{a}, t)$, the distributions at time $t > 0$ of susceptible, infectious and recovered individuals, respectively in relation to specific social characteristics. We assume that the rapid spread of the disease and the low mortality rate allows to ignore changes in the social structure, such as the aging process, births and deaths.

Consequently, for a given population of total number N, we have that

$$s(\mathbf{a}, t) + i(\mathbf{a}, t) + r(\mathbf{a}, t) = f(\mathbf{a}), \qquad \int_\Lambda f(\mathbf{a}) d\mathbf{a} = N,$$

where $f(\mathbf{a})$ is the total distribution of the social features defined by the vector $\mathbf{a}$. Hence, we recover the total fraction of the population which belong to the susceptible, infected and recovered as follows

$$S(t) = \int_\Lambda s(\mathbf{a}, t)\, d\mathbf{a}, \quad I(t) = \int_\Lambda i(\mathbf{a}, t)\, d\mathbf{a}, \quad R(t) = \int_\Lambda r(\mathbf{a}, t)\, d\mathbf{a}. \qquad (1)$$

In a situation where changes in the social features act on a slower scale with respect to the spread of the disease, the socially structured compartmental model follows the dynamics

$$\frac{d}{dt}s(\mathbf{a}, t) = -s(\mathbf{a}, t)\frac{1}{N}\int_\Lambda \beta(\mathbf{a}, \mathbf{a}_*)i(\mathbf{a}_*, t)\, d\mathbf{a}_*$$
$$\frac{d}{dt}i(\mathbf{a}, t) = s(\mathbf{a}, t)\frac{1}{N}\int_\Lambda \beta(\mathbf{a}, \mathbf{a}_*)i(\mathbf{a}_*, t)\, d\mathbf{a}_* - \gamma(\mathbf{a})i(\mathbf{a}, t) \qquad (2)$$
$$\frac{d}{dt}r(\mathbf{a}, t) = \gamma(\mathbf{a})i(\mathbf{a}, t),$$

where the function $\beta(\mathbf{a}, \mathbf{a}_*) \geq 0$ represents the uncertain interaction rate among individuals with different social features and $\gamma(\mathbf{a}) \geq 0$ the recovery rate which may depend on the social feature.

Often, in socially structured models the interaction rate between people is assumed to be separable, and proportionate to the activity level of the social feature (Hethcote 1996, 2000), as follows

$$\beta(\mathbf{a}, \mathbf{a}_*) = \frac{b(\mathbf{a})b(\mathbf{a}_*)}{\int_0^{+\infty} b(\mathbf{a})f(\mathbf{a})\, d\mathbf{a}} \qquad (3)$$

with $b(\mathbf{a})$ the average number of people contacted by a person with social feature $\mathbf{a}$ per unit time. Alternative approaches are based on preferential mixing (Glasser et al. 2012; Castillo-Chavez et al. 1989). Specific examples of age-dependent social interaction matrices are reported in Appendix C.

We introduce the usual normalization scaling

$$\frac{s(\mathbf{a}, t)}{N} \to s(\mathbf{a}, t), \quad \frac{i(\mathbf{a}, t)}{N} \to i(\mathbf{a}, t), \quad \frac{r(\mathbf{a}, t)}{N} \to r(\mathbf{a}, t), \quad \int_\Lambda f(\mathbf{a}) d\mathbf{a} = 1,$$

and observe that the quantities $S(t)$, $I(t)$ and $R(t)$ satisfy the conventional SIR dynamics

$$\frac{d}{dt}S(t) = -\int_\Lambda s(\mathbf{a}, t) \int_\Lambda \beta(\mathbf{a}, \mathbf{a}_*)i(\mathbf{a}_*, t)\, d\mathbf{a}_*\, d\mathbf{a}$$
$$\frac{d}{dt}I(t) = \int_\Lambda s(\mathbf{a}, t) \int_\Lambda \beta(\mathbf{a}, \mathbf{a}_*)i(\mathbf{a}_*, t)\, d\mathbf{a}_*\, d\mathbf{a} - \int_\Lambda \gamma(\mathbf{a})i(\mathbf{a}, t)\, d\mathbf{a}, \tag{4}$$

where the fraction of recovered is obtained from $R(t) = 1 - S(t) - I(t)$. We refer to Hethcote (1996), Hethcote (2000) for analytical results concerning model (2) and (4). In the following we will adopt the simple compartmental model (2) to derive our feedback controlled formulation in presence of uncertainty. The extension to more realistic compartmental models in epidemiology, designed specifically for the COVID-19 pandemic, can be carried out in a similar way (Gatto et al. 2020).

In order to simplify the description, we will consider the one-dimensional case $d_a = 1$ and set the social dependence as the age a of the individual because of its importance in epidemic dynamics. It is clear, however, that similar containment procedures can impact also on other social features, like the wealth of the individuals (Dimarco et al. 2020). We will first formulate the feedback controlled SIR model in the deterministic case and subsequently extend our approach to the presence of uncertain parameters.

2.2 Optimal control of structured compartmental model

In order to define the action of a policy maker introducing a control over the system based on social distancing and other containment measures linked to the social structure we consider an optimal control framework. The choice of an appropriate functional is problem dependent (Lee et al. 2010).

In our setting, we account the minimization of the total number of the infected population $I(t)$ through the an age dependent control action depending both on time and pairwise interactions among individuals with different ages.

Thus, we introduce the optimal control problem

$$\min_{u \in \mathcal{U}} J(u) := \int_0^T \psi(I(t))dt + \frac{1}{2}\int_0^T \int_{\Lambda \times \Lambda} v(a, a_*, t)|u(a, a_*, t)|^2\, da\, da_*\, dt, \tag{5}$$

subject to

$$\frac{d}{dt}s(a, t) = -s(a, t)\int_\Lambda (\beta(a, a_*) - u(a, a_*, t))i(a_*, t)\, da_*$$
$$\frac{d}{dt}i(a, t) = s(a, t)\int_\Lambda (\beta(a, a_*) - u(a, a_*, t))i(a_*, t)\, da_* - \gamma(a)i(a, t) \tag{6}$$
$$\frac{d}{dt}r(a, t) = \gamma(a)i(a, t)$$

with initial condition $i(a, 0) = i_0(a)$, $s(a, 0) = s_0(a)$ and $r(a, 0) = r_0(a)$.

The number of infected individual is measured by a monotone increasing function $\psi(\cdot)$ such that $\psi : [0, 1] \rightarrow \mathbb{R}_+$. This function models the policy maker's perception of the impact of the epidemic by the number of people currently infected and in the sequel will be referred to as *perception function*. For example $\psi(I) = I^q/q$, for $q > 1$ implies an underestimation of the actual number of infected corresponding to $q = 1$. The control aims to minimize this measure of the total infected population by reducing the rate of interaction between individuals. We consider a quadratic cost for its actuation.

Such control is restricted to the space of admissible controls

$$\mathcal{U} \equiv \left\{ u \mid 0 \leq u(a, a_*, t) \leq \min\{M, \beta(a, a_*)\}, \ \forall (a, a_*, t) \in \Lambda^2 \times [0, T], \ M > 0 \right\},$$

which ensure the admissibility of the solution for (6). The above restriction on admissible controls can be relaxed if we consider controls that violate the previous condition locally but preserve the inequality in integral form after integration against $i(a_*, t)$.

The solution to problem (5–6) is computed through the optimality conditions obtained from the Euler-Lagrangian as follows

$$
\begin{aligned}
\mathcal{L}(s, i, r, p_s, p_i, p_r, u) &= J(u) \\
&+ \int_0^T \int_\Lambda p_s \cdot \left(\frac{d}{dt}s + s(a, t) \int_\Lambda (\beta(a, a_*) - u(a, a_*, t)) \, i(a_*, t) da_* \right) da \, dt \\
&+ \int_0^T \int_\Lambda p_i \cdot \left(\frac{d}{dt}i - s(a, t) \int_\Lambda (\beta(a, a_*) - u(a, a_*, t)) \, i(a_*, t) da_* + \gamma(a) i(a, t) \right) da \, dt \\
&+ \int_0^T \int_\Lambda p_r \cdot \left(\frac{d}{dt}r - \gamma(a) i(a, t) \right) da \, dt
\end{aligned}
$$

where $p_s(a, t)$, $p_i(a, t)$, $p_r(a, t)$ are the associated Lagrangian multipliers. By computing the variations with respect to (s, i, r) we retrieve the adjoint system

$$
\begin{aligned}
\frac{d}{dt}p_s &= (p_s - p_i) \int_\Lambda (\beta(a, a_*) - u(a, a_*, t)) i(a_*, t) da_* \\
\frac{d}{dt}p_i &= \psi'(I(t)) + \int_\Lambda (p_s(a_*, t) - p_i(a_*, t))(\beta(a_*, a) \\
&\quad - u(a_*, a, t)) s(a_*, t) da_* + \gamma(a) p_i
\end{aligned}
\tag{7}
$$

with terminal conditions $p_s(a, T) = 0$, $p_i(a, T) = 0$ and $p_r(a, T) = 0$. Note that the contribution of $p_r(a, t)$ vanishes since the control does not act directly on population R, and the removed population is not considered in the minimization of the functional. The optimality condition reads

$$
\nu(a, a_*, t) u(a, a_*, t) = (p_s - p_i) \, s(a, t) i(a_*, t).
\tag{8}
$$

The optimality conditions (7–8) are first order necessary conditions for the optimal control $u(a, a_*, t)$. In order to be admissible then the control reads

$$u(a, a_*, t) = \max \left\{ 0, \min \left\{ \frac{p_s - p_i}{\nu(a, a_*, t)} s(a, t) i(a_*, t), \phi_{M, \beta}(a, a_*) \right\} \right\},$$

where $\phi_{M, \beta}(a, a_*) = \min\{\beta(a, a_*), M\}$.

The approach just described, however, is generally quite complicated when there are uncertainties as it involves solving simultaneously the forward problem (5–6) and the backward problem (7–8). Moreover, the assumption that the policy maker follows an optimal strategy over a long time horizon seems rather unrealistic in the case of a rapidly spreading disease such as the COVID-19 epidemic. Let us emphasize that extending the above optimal control formulation to more complex compartmental models designed specifically for COVID-19, like SEPIAR or SIDHARTE (Gatto et al. 2020; Giordano et al. 2020), can be done by generalizing the control functional (5) to include, for example, the hospitalized compartment, or other specific indicators that can be measured from the data. For all of these models, the feedback control strategy described in the next section does not change substantially. We refer the reader to Appendix A for more details.

2.3 Feedback controlled compartmental models

In this section we consider short time horizon strategies which permits to derive suitable feedback controlled models. These strategies are suboptimal with respect the original problem (5–6) but they have proved to be very successful in several social modeling problems (Albi and Pareschi 2018; Albi et al. 2015a, b, 2014; Düring et al. 2018). To this aim, we consider a short time horizon of length $h > 0$ and formulate a time discretize optimal control problem through the functional $J_h(u)$ restricted to the interval $[t, t + h]$, as follows

$$\min_{u \in \mathcal{U}} J_h(u) := \psi(I(t + h)) + \frac{1}{2} \int_{\Lambda \times \Lambda} \nu(a, a_*, t) |u(a, a_*, t)|^2 da\, da_* \qquad (9)$$

subject to

$$s(a, t + h) = s(a, t) - hs(a, t) \int_{\Lambda} (\beta(a, a_*) - u(a, a_*, t)) \, i(a_*, t) da_* \qquad (10)$$

$$i(a, t + h) = i(a, t) + hs(a, t) \int_{\Lambda} (\beta(a, a_*) - u(a, a_*, t)) \, i(a_*, t) da_* - h\gamma(a) i(a, t). \qquad (11)$$

By recalling that the macroscopic information on the infected is

$$I(t + h) = I(t) + h \int_{\Lambda} \left[s(a, t) \int_{\Lambda} (\beta(a, a_*) - u(a, a_*, t)) \, i(a_*, t) da_* - \gamma(a) i(a, t) \right] da,$$

we can derive the minimizer of J_h computing $D_u J_h(u) \equiv 0$. We retrieve the following nonlinear equation

$$v(a, a_*, t)u(a, t) = hs(a, t)i(a_*, t)\psi'(I(t+h)). \tag{12}$$

In order to pass to the limit $h \to 0$ we must rescale the penalization term as $v(a, a_*, t) = h\kappa(a, a_*, t)$ so that we can introduce the above instantaneous strategy directly in the discrete system (10)-(11).

The resulting controlled dynamic, corresponding to the feedback controlled continuous system (6), reads as follows

$$\frac{d}{dt}s(a, t) = -s(a, t)\int_\Lambda \left(\beta(a, a_*) - \frac{s(a, t)i(a_*, t)\psi'(I(t))}{\kappa(a, a_*, t)}\right)i(a_*, t)da_* \tag{13a}$$

$$\frac{d}{dt}i(a, t) = s(a, t)\int_\Lambda \left(\beta(a, a_*) - \frac{s(a, t)i(a_*, t)\psi'(I(t))}{\kappa(a, a_*, t)}\right)i(a_*, t)da_* - \gamma(a)i(a, t) \tag{13b}$$

$$\frac{d}{dt}r(a, t) = \gamma(a)i(a, t). \tag{13c}$$

In what follows we provide a sufficient conditions for the admissibility of the instantaneous control in terms of the penalization term $\kappa(a, a_*, t)$. Indeed we want to assure that the dynamics preserve the monotonicity of the number of susceptible population $s(a, t)$.

Proposition 1 *Let $\beta(a, a_*) \geq \delta > 0$, and $\psi'(\cdot)$ a monotonically non decreasing function, then for all $(a, a_*) \in \Lambda \times \Lambda$ and $t > 0$, solutions to (13) are admissible if the penalization κ satisfies the following inequality*

$$\delta\kappa(a, a_*, t) \geq s_0(a)\bar{i}(a_*)\psi'(\bar{I}), \qquad \forall(a, a_*) \in \Lambda \times \Lambda \tag{14}$$

where $\bar{i}(a)$ and $\bar{I}$ are respectively the peak reached by the infected of age a and by the total population of infected.

Proof By imposing the non-negativity of the controlled reproduction rate inside the integral we have

$$\beta(a, a_*)\kappa(a, a_*, t) \geq s(a, t)i(a_*, t)\psi'(I(t)).$$

This inequality has to be satisfied for every time $t \geq 0$. Next we observe that the number of susceptible $s(a, t)$ is decreasing in time, therefore $s_0(a) \geq s(a, t)$ for all t. Moreover $i(a, t)$ reaches a peak before decreasing to 0 (note that this peak can also be in $t = 0$), say $\bar{I}$ for the macroscopic variable and $\bar{i}(a)$. Thus, thanks to the monotonicity of $\psi'(\cdot)$, we can restrict the previous inequality as follows

$$\delta\kappa(a, a_*, t) \geq s_0(a)\bar{i}(a_*)\psi'(\bar{I}).$$
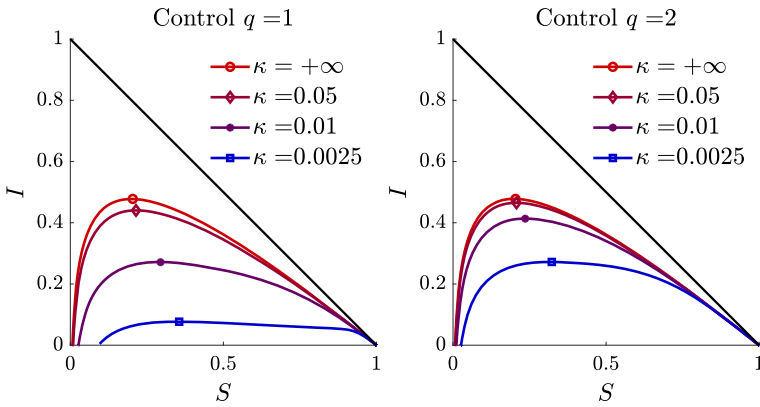
$\square$

Fig. 1 Phase diagram of susceptible-infected trajectories for the controlled SIR-type model with homogenous mixing and $\psi(I) = I^q/q$. Different choices of the penalization term κ are reported. Left plot the case $q = 1$, right plot $q = 2$. The line markers point out the peaks of the infected population for each choice of κ

In Fig. 1 we report the phase diagram of susceptible-infected trajectories for the controlled model with homogeneous mixing

$$\frac{d}{dt}S(t) = -\left(\beta - \frac{1}{\kappa}S(t)I(t)\psi'(I(t))\right)S(t)I(t)$$
$$\frac{d}{dt}I(t) = \left(\beta - \frac{1}{\kappa}S(t)I(t)\psi'(I(t))\right)S(t)I(t) - \gamma I(t),$$
(15)

with $\psi(I) = I^q/q$. The dynamics are similar to the classical SIR model but with a nonlinear contact rate. In particular, the trajectories are flattened when the value of the control is such that $\kappa\beta \approx S(t)I(t)\psi'(I(t))$. Note, however that this status is not an equilibrium point of the system.

To understand this, let us observe that an equilibrium state (S^*, I^*) for (15) satisfies the equations

$$\left(\kappa\beta - S^*I^*\psi'(I^*)\right)S^*I^* = 0$$
$$\left((\kappa\beta - S^*I^*\psi'(I^*))S^* - \kappa\gamma\right)I^* = 0.$$

An equilibrium point corresponds to the classical state in which we have the extinction of the disease $I^* = 0$ and S^* arbitrary and defined by the asymptotic state of the dynamics (Hethcote 2000). Now, let's suppose that $I^* \neq 0$, $S^* \neq 0$, we can look for solutions where control is able to perfectly balance the spread of the disease

$$\kappa\beta = S^*I^*\psi'(I^*)$$
$$\kappa\beta S^* = (S^*)^2 I^*\psi'(I^*) + \kappa\gamma,$$

consequently

$$\kappa \beta S^* = (S^*)^2 \frac{\kappa \beta}{S^*} + \kappa \gamma = 0,$$

which is satisfied only for $\gamma = 0$ when $\kappa \neq 0$. The stability analysis of this unique equilibrium point can be performed by standard arguments and we omit the details.

3 Control of epidemic dynamics with uncertainties

Since the beginning of the outbreak of new infectious diseases, the actual number of infected and recovered people is typically underestimated, causing fatal delays in the implementation of public health policies facing the propagation of epidemic fronts. This is the case of the spreading of COVID-19 worldwide, often mistakenly underestimated due to deficiencies in surveillance and diagnostic capacity (Jagodnik et al. 2020; Remuzzi and Remuzzi 2020). Health systems are struggling to adopt systematic testing to monitor actual cases. Moreover, another important epidemiological factor affecting data reliability is the large proportion of asymptomatic (Jagodnik et al. 2020; Mizumoto et al. 2020; Zhang et al. 2020).

Among the common sources of uncertainties for dynamical systems modeling epidemic outbreaks we may consider the following

- noisy and incomplete available data;
- structural uncertainty due to the possible inadequacy of the mathematical model used to describe the phenomena under consideration.

In the following we consider the effects on the dynamics of uncertain data, such as the initial conditions on the number of infected people or the interaction and recovery rates. On the numerical level we consider techniques based on stochastic Galerkin methods, for which spectral convergence on random variables is obtained under appropriate regularity assumptions (Xiu 2010). For simplicity, the details of the numerical method that allows to reduce the uncertain dynamic system to a set of deterministic equations are reported in Appendix B.

3.1 Socially structured models with uncertain inputs

We introduce the random vector $\mathbf{z} = (z_1, \ldots, z_{d_z})$ whose components are assumed to be independent real valued random variables

$$z_k : (\Omega, F) \to (\mathbb{R}, \mathcal{B}_{\mathbb{R}}), \qquad k = 1, \ldots, d_z$$

with $\mathcal{B}_{\mathbb{R}}$ the Borel set. We assume to know the probability density $p(\mathbf{z}) : \mathbb{R}^{d_z} \to \mathbb{R}_+^{d_z}$ characterizing the distribution of $\mathbf{z}$. Here, $z \in \mathbb{R}^{d_z}$ is a random vector taking into account various possible sources of uncertainty in the model.

In presence of uncertainties we generalize the initial modeling by introducing the quantities $s(\mathbf{z}, a, t), i(\mathbf{z}, a, t)$ and $r(\mathbf{z}, a, t)$ representing the distributions at time $t \geq 0$

of susceptible, infectious and recovered individuals. The total size of the population is a deterministic conserved quantity in time, i.e.

$$s(\mathbf{z}, a, t) + i(\mathbf{z}, a, t) + r(\mathbf{z}, a, t) = f(a), \qquad \int_\Lambda f(a)\,da = N.$$

Hence, the quantities

$$S(\mathbf{z}, t) = \int_\Lambda s(\mathbf{z}, a, t)\,da, \quad I(\mathbf{z}, t) = \int_\Lambda i(\mathbf{z}, a, t)\,da, \quad R(\mathbf{z}, t) = \int_\Lambda r(\mathbf{z}, a, t)\,da, \tag{16}$$

denote the uncertain fractions of the population that are susceptible, infectious and recovered respectively.

The social structured model with uncertainties reads

$$
\begin{aligned}
\frac{d}{dt}s(\mathbf{z}, a, t) &= -s(\mathbf{z}, a, t) \int_\Lambda \beta(\mathbf{z}, a, a_*) \frac{i(\mathbf{z}, a_*, t)}{N}\,da_* \\
\frac{d}{dt}i(\mathbf{z}, a, t) &= s(\mathbf{z}, a, t) \int_\Lambda \beta(\mathbf{z}, a, a_*) \frac{i(\mathbf{z}, a_*, t)}{N}\,da_* - \gamma(\mathbf{z}, a)i(\mathbf{z}, a, t) \\
\frac{d}{dt}r(\mathbf{z}, a, t) &= \gamma(\mathbf{z}, a)i(\mathbf{z}, a, t)
\end{aligned}
\tag{17}
$$

To illustrate the impact of uncertainties let us consider the simple following example. In the case of homogeneous mixing with uncertain contact rate $\beta(z) = \beta + \alpha z, \alpha > 0$, $z \in \mathbb{R}$ distributed as $p(z)$ the model reads

$$
\begin{aligned}
\frac{d}{dt}S(z, t) &= -(\beta + \alpha z)\,S(z, t)I(z, t) \\
\frac{d}{dt}I(z, t) &= (\beta + \alpha z)\,S(z, t)I(z, t) - \gamma I(z, t),
\end{aligned}
\tag{18}
$$

with deterministic initial values $I(z, 0) = I_0$ and $S(z, 0) = S_0$. The solution for the proportion of infectious during the initial exponential phase is (Roberts 2013)

$$I(z, t) = I_0 e^{(\beta + \alpha z) S_0 t - \gamma t},$$

and its expectation

$$\mathbb{E}[I(\cdot, t)] = I_0 \int_\mathbb{R} e^{(\beta + \alpha z) S_0 t - \gamma t} p(z)\,dz = I_0 e^{\beta S_0 t - \gamma t} W(t), \tag{19}$$

where the function

$$W(t) = \int_\mathbb{R} e^{\alpha z S_0 t} p(z)\,dz$$

represents the statistical correction factor to the standard deterministic exponential phase of the disease $I_0 e^{\beta S_0 t - \gamma t}$. If z is uniformly distributed in $[-1, 1]$ we can explicitly compute

$$W(t) = \frac{\sinh(\alpha S_0 t)}{\alpha S_0 t} > 1, \quad t > 0.$$

More in general, if z has zero mean then by Jensen's inequality we have $W(t) > 1$ for $t > 0$, so that the expected exponential phase is amplified by the uncertainty (see Roberts 2013).

In a similar way, keeping β constant, but introducing a source of uncertainty in the initial data $I(z, 0) = I_0 + \mu z$, $\mu > 0$ and $z \in \mathbb{R}$ distributed as $p(z)$ the solution in the exponential phase reads

$$I(z, t) = (I_0 + \mu z) e^{\beta S_0 t - \gamma t},$$

and then its expectation

$$\mathbb{E}[I(\cdot, t)] = \int_{\mathbb{R}} (I_0 + \mu z) e^{\beta S_0 t - \gamma t} p(z)\, dz = (I_0 + \mu \bar{z}) e^{\beta S_0 t - \gamma t}, \qquad (20)$$

where $\bar{z}$ is the mean of the variable z. Therefore, the expected initial exponential growth behaves as the one with deterministic initial data $I_0 + \mu \bar{z}$. Of course, if both sources of uncertainty are present the two effects just described sum up in the dynamics.

Remark 1 The presence of a large number of undetected infected is at the basis of the construction of numerous epidemiological models with an increasingly complex compartmental structure in which the original compartment of the infected is subdivided into further compartments with different roles in the propagation of the disease (Giordano et al. 2020; Gatto et al. 2020; Flaxman et al. 2020). To clarify the relationships to other deterministic compartmental models, let us consider, for simplicity, model (17) in absence of a social structure

$$\begin{aligned}
\frac{d}{dt} S(z, t) &= -\beta(z) S(z, t) I(z, t) \\
\frac{d}{dt} I(z, t) &= \beta(z) S(z, t) I(z, t) - \gamma(z) I(z, t), \\
\frac{d}{dt} R(z, t) &= \gamma(z) I(z, t),
\end{aligned} \qquad (21)$$

and with a one-dimensional random input $z \in \mathbb{R}$ distributed as $p(z)$. Furthermore, for a function $F(z, t)$ we will denote its expected value as $\bar{F}(t) = \mathbb{E}[F(\cdot, t)]$. Now, starting from a discrete probability density function

$$p_k = P\{Z = z_k\}, \quad \sum_{k=1}^{n} p_k = 1,$$

we have $\bar{F}(t) = \sum_{k=1}^{n} p_k F_k$, with $F_k = F(z_k)$. Taking the expectation in (21), we can write

$$\frac{d}{dt}\bar{S}(t) = -\bar{S}(t)\sum_{k=1}^{n}\tilde{\beta}_k p_k I_k(t)$$

$$\frac{d}{dt}\bar{I}(t) = \bar{S}(t)\sum_{k=1}^{n}\tilde{\beta}_k p_k I_k(t) - \sum_{k=1}^{n}\gamma_k p_k I_k(t), \qquad (22)$$

$$\frac{d}{dt}\bar{R}(t) = \sum_{k=1}^{n}\gamma_k p_k I_k(t),$$

with $\tilde{\beta}_k = S_k \beta_k / \bar{S}$, $k = 1, \ldots, n$. For example, in the case n=2, by identifying $I_d = p_1 I_1$ and $I_u = p_2 I_2$ with the compartments of detected and undetected infectious individuals, respectively, we can formulate the equivalent partitioning

$$\frac{d}{dt}\bar{S}(t) = -\bar{S}(t)\left(\tilde{\beta}_1 I_d(t) + \tilde{\beta}_2 I_u(t)\right)$$

$$\frac{d}{dt}I_d(t) = \bar{S}(t)\tilde{\beta}_1 I_d(t) - \gamma_1 I_d(t),$$

$$\frac{d}{dt}I_u(t) = \bar{S}(t)\tilde{\beta}_2 I_u(t) - \gamma_2 I_u(t), \qquad (23)$$

$$\frac{d}{dt}\bar{R}(t) = \gamma_1 I_d(t) + \gamma_2 I_u(t),$$

which has the same structure of a SIAR compartmental model including the undetected (or the asymptomatic) class. In the following, we will not rely on discrete probability distributions, but on continuous representations that can be associated with the overall probability of having a certain number of infectious individuals (detected or undetected). The additional dependence of the epidemiological parameters from the random variable allows us to take into account changes in the corresponding dynamics of disease transmission and recovery.

3.2 The feedback controlled model with random inputs

In presence of uncertainties the optimal control problem (5–6) is modified as follows

$$\min_{u \in \mathcal{U}} J(u) := \int_0^T \mathcal{R}[\psi(I(\cdot, t))]dt + \frac{1}{2}\int_0^T \int_{\Lambda \times \Lambda} \nu(a, a_*, t)|u(a, a_*, t)|^2 \, da da_* dt, \qquad (24)$$

being $\mathcal{R}[\psi(I(\cdot, t))]$ a suitable operator taking into account the presence of the uncertainties **z**. Examples of such operator that are of interest in epidemic modeling are the

expectation with respect to uncertainties

$$\mathcal{R}[\psi(I(\cdot,t))] = \mathbb{E}[\psi(I(\cdot,t))] = \int_{\mathbb{R}^{d_z}} \psi(I(\mathbf{z},t))\, p(\mathbf{z})d\mathbf{z} \qquad (25)$$

or relying on data which underestimate the number of infected

$$\mathcal{R}[\psi(I(\cdot,t))] = \psi(I(\mathbf{z}_0,t)), \qquad (26)$$

where $\mathbf{z}_0$ is a given value such that $\psi(I(\mathbf{z}_0,t)) \le \psi(I(\mathbf{z},t))$, $\forall\,\mathbf{z} \in \mathbb{R}^{d_z}$ and $t > 0$.

In (24) $\mathcal{U}$ the space of admissible controls is defined as

$$\mathcal{U} = \left\{ u \,|\, 0 \le u(a, a_*, t) \le \min\{M, \min_{\mathbf{z}} \beta(\mathbf{z}, a, a_*)\},\ \forall\,(a, a_*, t) \in \Lambda^2 \times [0, T],\ M > 0 \right\},$$

or in a more relaxed form if we consider the above inequality after integration against $i(a_*, t)$.

The minimization problem (24) is subject to the following dynamics

$$\frac{d}{dt}s(\mathbf{z}, a, t) = -s(\mathbf{z}, a, t)\int_\Lambda (\beta(\mathbf{z}, a, a_*) - u(a, a_*, t))i(\mathbf{z}, a_*, t)\,da_*$$

$$\frac{d}{dt}i(\mathbf{z}, a, t) = s(\mathbf{z}, a, t)\int_\Lambda (\beta(\mathbf{z}, a, a_*) - u(a, a_*, t))i(\mathbf{z}, a_*, t)\,da_* - \gamma(\mathbf{z}, a)i(\mathbf{z}, a, t) \quad (27)$$

$$\frac{d}{dt}r(\mathbf{z}, a, t) = \gamma(\mathbf{z}, a)i(\mathbf{z}, a, t)$$

with initial condition $i(\mathbf{z}, a, 0) = i_0(\mathbf{z}, a)$, $s(\mathbf{z}, a, 0) = s_0(\mathbf{z}, a)$ and $r(\mathbf{z}, a, 0) = r_0(\mathbf{z}, a)$.

The implementation of instantaneous control for dynamics in presence of uncertainties follows from the derivation presented in Sect. 2.3. We can derive the minimizer of J_h computing $D_u J_h(u) \equiv 0$ from the restriction of the minimization problem (24) in the interval $[t, t + h]$ or equivalently

$$\mathcal{R}\left[\frac{\partial \psi(I(\cdot, t + h))}{\partial u}\right] = v(a, t)u(a, a_*, t),$$

where we assumed $\partial \mathcal{R}[\psi(I(\cdot, t + h))]/\partial u = \mathcal{R}[\partial \psi(I(\cdot, t + h))/\partial u]$, to obtain the following nonlinear identity

$$v(a, t)u(a, a_*, t) = h\mathcal{R}[s(\cdot, a, t)i(\cdot, a_*, t)\psi'(I(\cdot, t))]. \qquad (28)$$

The above assumption on $\mathcal{R}[\cdot]$ is clearly satisfied by (25) and (26), where in the case of (26) we used the notation

$$\mathcal{R}[s(\cdot, a, t)i(\cdot, a_*, t)\psi'(I(\cdot, t))] = s(\mathbf{z}_0, a, t)i(\mathbf{z}_0, a_*, t)\psi'(I(\cdot, t)).$$

By introducing the scaling $v(a, t) = h\kappa(a, a_*, t)$ we can pass to the limit for $h \to 0$ to get

$$u(a, a_*, t) = \frac{1}{\kappa(a, a_*)}\mathcal{R}\left[s(\cdot, a, t)i(\cdot, a_*, t)\psi'(I(\cdot, t))\right], \tag{29}$$

which defines the feedback controlled model in presence of uncertainties.

4 Examples from the COVID-19 outbreak in Italy

In this section we present several numerical tests on the constrained compartmental model with uncertain data. Details of the numerical method used are given in Appendix B. In an attempt to analyze sufficiently realistic scenarios, in the following examples we will refer to values taken from Italian data on the COVID-19 epidemic (del Consiglio dei Ministri 2020). More precisely, in the first test case we illustrate the behavior of the model in a simplified setting in absence of uncertainty and social structure and without trying to reproduce scenarios closely related to current data. In the second test case, following a progressively more realistic approach, we consider the impact of the presence of uncertain data in the controlled model with homogeneous social mixing and calibrated on Italian data. The same setting is then considered in Test 3 taking into account the additional effects given by the social structure of the system, characterized by suitable social interaction functions and an age-dependent recovery rate. The final scenario, explored in Test 4, examines the influence on the spread of infectious disease induced by relaxed confinement measures related to the social structure of the system.

4.1 Test 1: Containment in homogeneous social mixing dynamics

To illustrate the effects of controls introduced that mimic containment procedures, let us first consider the case where the social structure is not present. Furthermore, to simplify further the modeling, in this first example we neglect any dependence on uncertain data.

We consider as initial small number of infected and recovered $i(0) = 3.68 \times 10^{-6}$, $r(0) = 8.33 \times 10^{-8}$. These normalized fractions refer specifically to the first reported values in the case of the Italian outbreak of COVID-19, even if in this simple test case we will not try to match the data in a quantitative setting but simply to illustrate the behavior of the feedback controlled model. Based on recent studies (Jagodnik et al. 2020; Zhang et al. 2020; Liu et al. 2020), the initial infection rate of COVID-19 $R_0 = \beta/\gamma$ has been estimated between 2 and 6.5. Here, to exemplify the possible evolution of the pandemic we consider a value close to the lower bound, taking $\beta = 0.25$ and $\gamma = 0.10$, namely a recovery rate of 10 days, so that $R_0 = 2.5$.

In Figs. 2 and 3 we report the infected and recovered dynamics based on the activation of the control in two different time frames. In Fig. 2 the activation for $t \in [50, 100]$, which means that after 100 days we suppose that all containment restrictions are cancelled. In Fig. 3 we consider a larger activation time frame $t \in [50, 200]$.
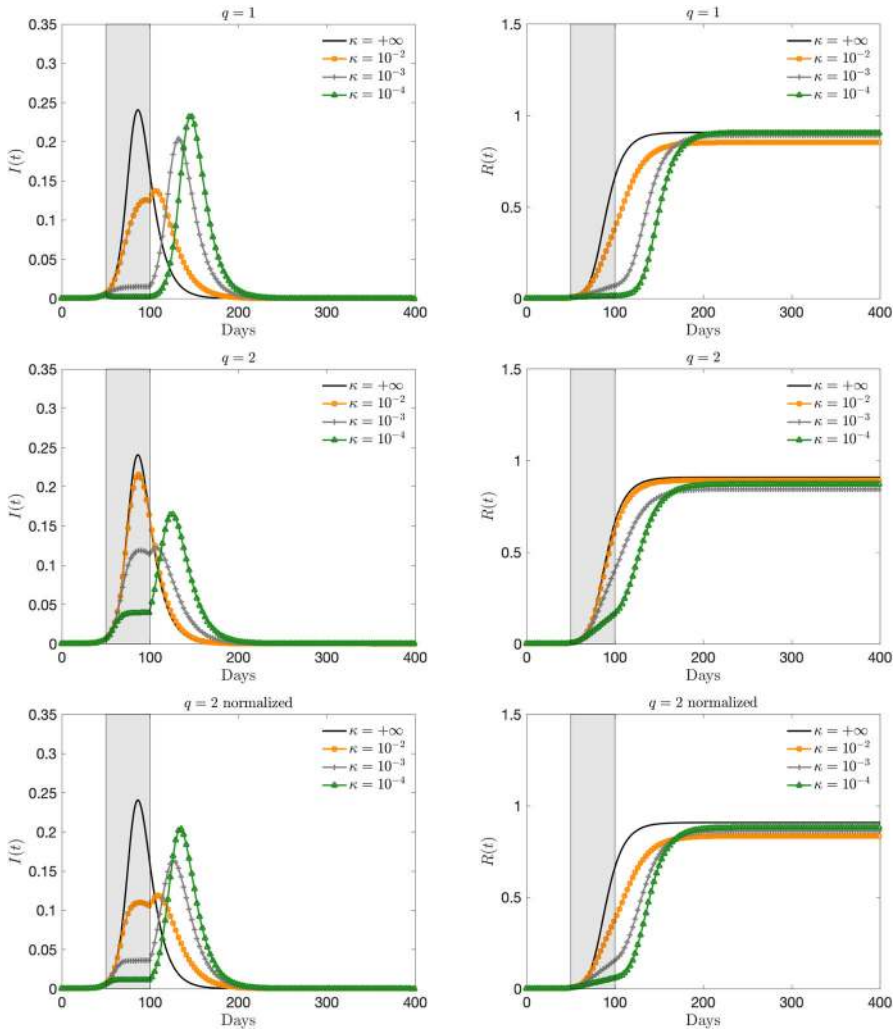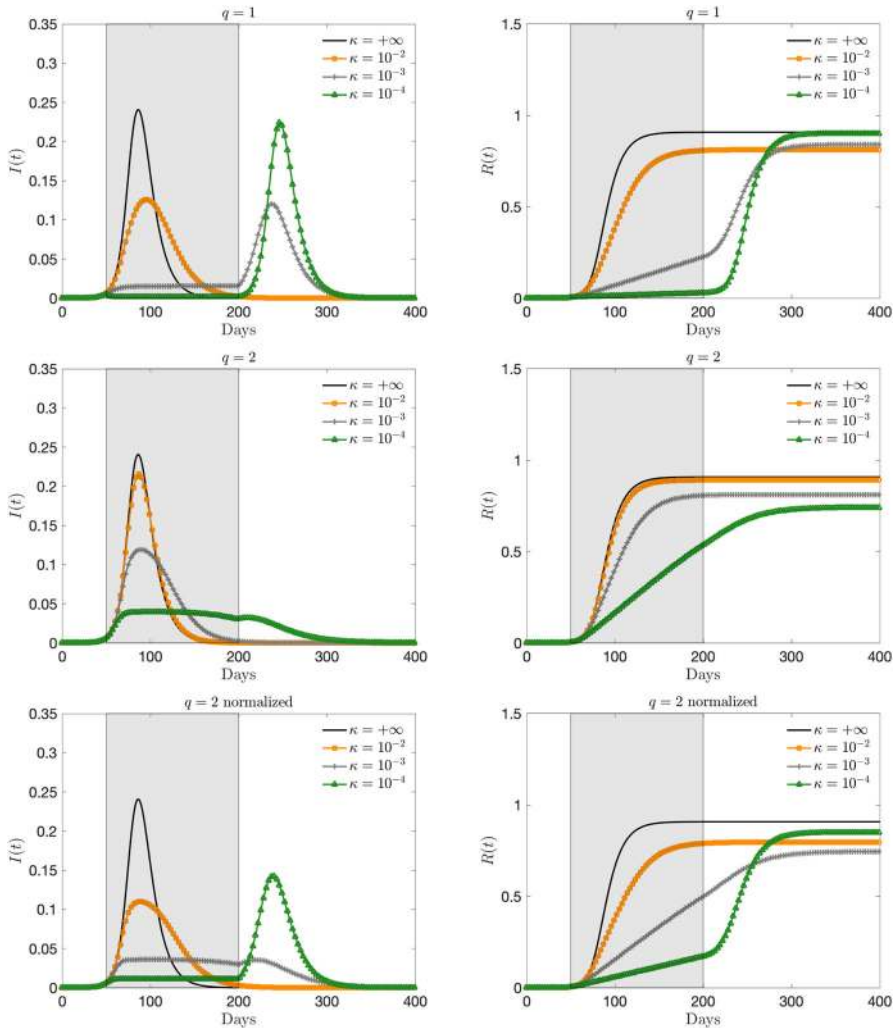
Fig. 2 Test 1. Evolution of the fraction of infected (left) and recovered (right) based on the feedback constrained model (15) for $t \in [50, 100]$, a perception function $\psi(I) = I^q/q$, $q = 1, 2$ and several penalizations $\kappa = 10^{-2}, 10^{-3}, 10^{-4}$. The choice $\kappa = +\infty$ corresponds to the unconstrained case. In last row the normalized case $\psi(I) = C_2 I^2/2$, with $C_2 = 12$

With the choice of the perception function $\psi(I) = I^q/q$, $q \geq 1$, we can observe how the control term is able to flatten the curve even if, as expected, the case $q = 2$ gives rise to a weaker control action. To make the two controls, $q = 1$ and $q = 2$, comparable for the same penalization factor κ we also consider the normalized case, where $\psi(I) = C_q I^q/q$, and the constant $C_q > 0$ is a normalization constant such

Fig. 3 Test 1. Evolution of the fraction of infected (left) and recovered (right) based on the feedback constrained model (15) for $t \in [50, 200]$, a perception function $\psi(I) = I^q/q$, $q = 1, 2$ and several penalizations $\kappa = 10^{-2}, 10^{-3}, 10^{-4}$. The choice $\kappa = +\infty$ corresponds to the unconstrained case. In last row the normalized case $\psi(I) = C_2 I^2/2$, with $C_2 = 12$

that

$$\int_0^{I_{\max}} C_1 I \, dI = \int_0^{I_{\max}} C_2 \frac{I^2}{2} \, dI,$$

with $I_{\max}$ an estimate of the maximum number of infected in absence of control. In particular, in our test case, from $I_{\max} \approx \frac{1}{4}$, assuming $C_1 = 1$ we obtain $C_2 = 12$.
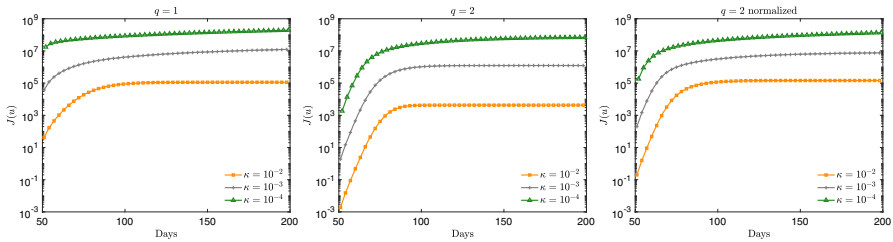
Fig. 4 Test 1. Evaluation of the cost functional $J(u)$ for the dynamics in Fig. 3

Note that, if the activation time is too short the control is not able to significantly change the total number of infected (and therefore recovered). On the other hand, by enlarging the activation time in combination with a sufficiently small penalty constant, the peak infection is not only reduced, but the total number of infected people is decreased. To achieve this, the control should be kept activated for a sufficiently long time and with the right intensity in a kind of plateau regime where there is a perfect balance between the containment effect and the spread of the disease. On the contrary, if the control is too strong, the majority of the population remains susceptible and consequently the disease will start spreading again forming a second wave after the containment policy is removed. Similar conclusions (that may appear counterintuitive) have been shown also by other authors (see for example (Britton et al. 2020; Lunelli et al. 2009)).

The cost functional depends on the value of q and can be evaluated summing up contributions in (9) with explicit form of the control given by (12). In Fig. 4 the cost of the two interventions is compared. We can see how a higher cost is associated with $q = 1$ that can be obtained with the control $q = 2$ only for smaller penalizations. As expected, the normalized case $q = 2$ essentially realigns the cost of interventions. Then, in Fig. 5 we compare the performance of the two controls in $[50, t_d]$ with a deactivation time $t_d \in [100, 300]$. We consider $\kappa = 10^{-3}$ for both $q = 1$ and $q = 2$ normalized. It can be observed that there is a minimum control horizon for both strategies, in order to avoid the onset of a second infection peak. A sufficiently long control horizon is therefore necessary to reduce the impact of the infection.

4.2 Test 2: Impact of uncertain data on the epidemic outbreak

Next we focus on the influence of uncertain quantities on the controlled system with homogeneous mixing focusing on available data for COVID-19 outbreak in Italy, see del Consiglio dei Ministri (2020). The estimation of epidemiological parameters is a very difficult problem that can be addressed with different approaches (Capaldi et al. 2012; Chowell 2017; Roberts 2013). It is worth to mention that, in the case of COVID-19, the number of infected and recovered has been largely underestimated, especially in the early phases of the epidemic, see (Jagodnik et al. 2020; Mizumoto et al. 2020). Here, we restrict ourselves to identifying the deterministic parameters of the model through a suitable data fitting procedure, considering possible deviations due to such underestimates as part of the subsequent uncertainty quantification process.
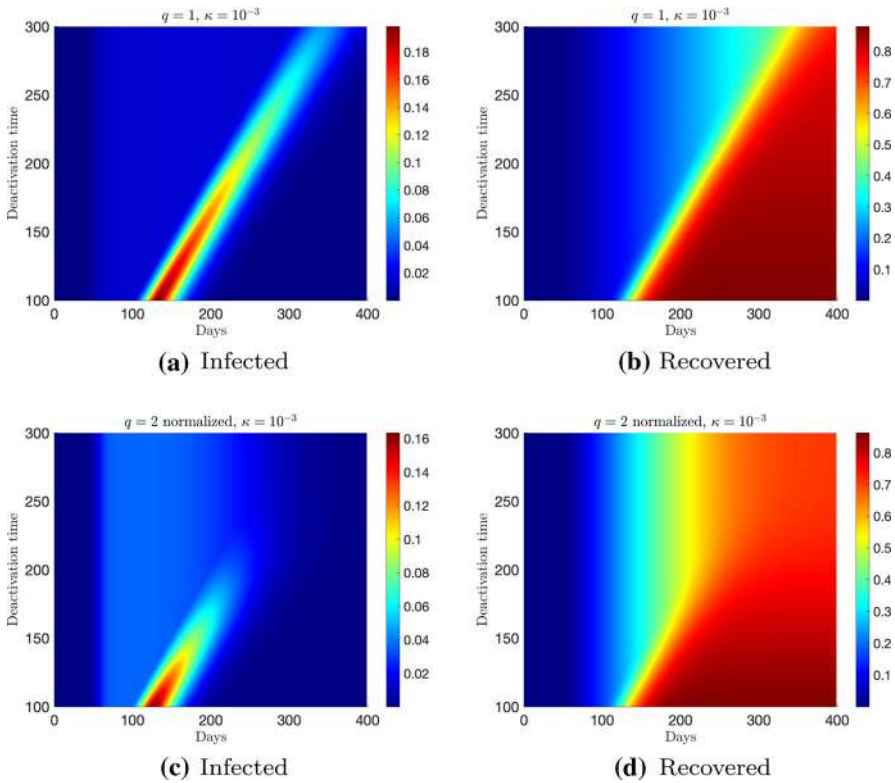
(a) Infected

(b) Recovered

(c) Infected

(d) Recovered

Fig. 5 Test 1. Evolution of the of the fraction of infected (left) and recovered (right) based on the feedback constrained model (15) for a perception function $\psi(I) = I^q/q$, $q = 1$ and $q = 2$ normalized, for different control actions in $[50, t_d]$ with a deactivation time $t_d \in [100, 300]$ and a fixed penalization $\kappa = 10^{-3}$

The data fitting process

In details, we have adopted a two-level approach in estimating the parameters in absence of uncertainties. In the phase preceding the lockdown we estimated the epidemic parameters in an unconstrained regime, where we assumed no social containment procedure was activated. This estimate was then kept int the subsequent lockdown phase where we estimated as a function of time the value of the control penalty parameter. Both these two calibration steps were analyzed under the assumption of homogeneous mixing, therefore model (15) has been used in lockdown phase.

First, we estimated the parameters $\beta, \gamma > 0$ in the time interval $t \in [t_0, t_u]$ solving a lest square problem based on the minimization of the relative L^2 norm of the difference between the reported number of infected $\hat{I}(t)$ and recovered $\hat{R}(t)$, and the theoretical evolution of the unconstrained model whose solution at time $t \geq 0$ is indicated with $I(t)$ and $R(t)$. More precisely, we considered the following minimization problem

$$\min_{\beta, \gamma \in \mathbb{R}_+} \left[(1 - \theta) \| I(t) - \hat{I}(t) \|_{L^2([t_0, t_u])} + \theta \| R(t) - \hat{R}(t) \|_{L^2([t_0, t_u])} \right], \quad (30)$$
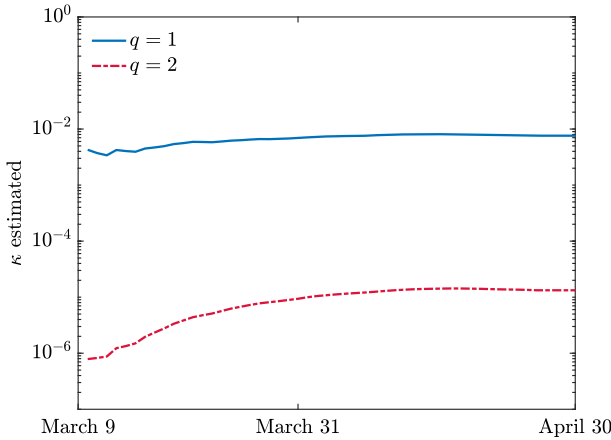
Fig. 6 **Test 2**. Estimated control penalization terms over time from reported data on number of infected and recovered in the case of COVID-19 outbreak in Italy

where $\theta \in [0, 1]$ and $\| \cdot \|_{L^2([t,s])}$ denotes the relative L^2 norm over a time horizon $[t, s]$.

Problem (30) has been solved with the constraints $\beta \in [0, 1]$ and $\gamma \in \left[\dfrac{1}{24}, \dfrac{1}{10} \right]$. Indeed, according to several studies the time to viral clearance during the early phase of the epidemic, corresponding to the time from the first positive test to the first negative test, can approximately span in average from 10 to 24 days, see (Chen 2020; Gatto et al. 2020; Lavezzo 2020).

At the end of the above optimization process we obtained the values $\beta_e \approx 0.31$, $\gamma_e \approx 0.049$ computed by averaging the optimization results with $\theta = 10^{-2}$ and $\theta = 10^{-6}$. The choice of a small value for θ is due to the low reliability on the recovered data at this early stage.

Next, we estimate the penalization $\kappa = \kappa(t)$ in time by solving (15), in the lockdown time interval $t \in (t_u, t_c]$ and for a sequence of time steps t_i, the corresponding least square problems in $[t_i - k_l, t_i + k_r]$ where $k_\ell, k_r \geq 1$ are integers, and where we fix the values β_e, γ_e estimated in the first optimization step. In details, we solve the following minimization problem

$$\min_{\kappa(t_i) \in \mathbb{R}_+} \left[(1 - \theta) \| I(t) - \hat{I}(t) \|_{L^2([(t_i - k_\ell, t_i + k_r])} + \theta \| R(t) - \hat{R}(t) \|_{L^2([t_i - k_\ell, t_i + k_r])} \right], \tag{31}$$

over a window of seven days corresponding to $k_\ell = 3$ and $k_r = 4$ for regularization along one week of available data. Both minimization problems (30–31) have been solved testing various numerical methods in combination with adaptive solves for the systems of ODEs. The results have been obtained using Matlab functions `fmincon` in combination with `ode45`. The available data start on February 24 2020, when

moderate social restrictions were enforced by the Italian government, and since the lockdown started on March 9 2020, thus we considered $t_u - t_0 = 14$ (days).

The corresponding time dependent values for the expected penalization for a perception function $\psi(I) = I^q/q$ are reported in Fig. 6. After an initial adjustment phase the penalty terms converge towards a constant value that we can assume as fixed in predictive terms for future times in a lockdown scenario. This is consistent with a situation in which society needs a certain period of time to adapt to the lockdown policy.

Remark 2 Finally, we remark that the data fitting procedure is easily generalizable to epidemic models with additional compartments (Gatto et al. 2020; Giordano et al. 2020). Let us denote with $C_i(t)$ the compartments that are related to the reported data $\hat{C}_j(t)$, $j = 1, \ldots, h$, as the number of actual cases, hospitalized, deaths, etc. In the first step one minimizes a weighted L^2 norm in the form

$$\min_{\mathbf{p} \in \mathcal{P}} \sum_{j=1}^{h} w_j \|C_j(t) - \hat{C}_j(t)\|_{L^2[t_0, t_u]}, \tag{32}$$

where $w_j \in [0, 1]$ are suitable weights such that $\sum_{j=1}^{h} w_j = 1$ and $\mathbf{p}$ is the vector of the model parameters that need to be estimated. Since the problem may admit multiple minima leading to unrealistic solutions one usually perform the above optimization process under constraints on the range of values $\mathcal{P}$ of some parameters such as recovery rate, incubation period, etc.

An important difference, compared to a simple SIR compartmentalization, is the presence of compartments that are not data-driven as exposed, pre-symptomatic, asymptomatic, etc. that prevent the realization of the data fitting since their values are unknown. A way to overcome this difficulty is to solve the differential model starting from an unknown time $t^* < t_0$ using as initial data the presence of a single individual in the first compartment promoting the infection (for example the exposed) and zero individuals in all other compartments. The idea is to simulate the early phase of the epidemic, by optimizing also the initial time t^* in problem (32).

After this, in the second step one considers the corresponding feedback controlled model (see Appendix A) and solves the minimization problem

$$\min_{\boldsymbol{\kappa}(t_i)} \sum_{j=1}^{h} w_j \|C_j(t) - \hat{C}_j(t)\|_{L^2([t_i - k_\ell, t_i + k_r])}, \tag{33}$$

where $\boldsymbol{\kappa}(t_i)$ is the vector of penalization terms that need to be estimated. Even in this case, assumptions on the range of values of the penalization terms may be necessary to avoid unrealistic solutions. Finally, we underline that most epidemic models are not data-driven, but one can always assume that the total number of reported cases underestimates the actual number of cases and perform the above data fitting to obtain a lower bound on the evolution of the epidemic. Then including a suitable data uncertainty, as in the present work, allows the recovery of realistic scenarios for the pandemic progression.
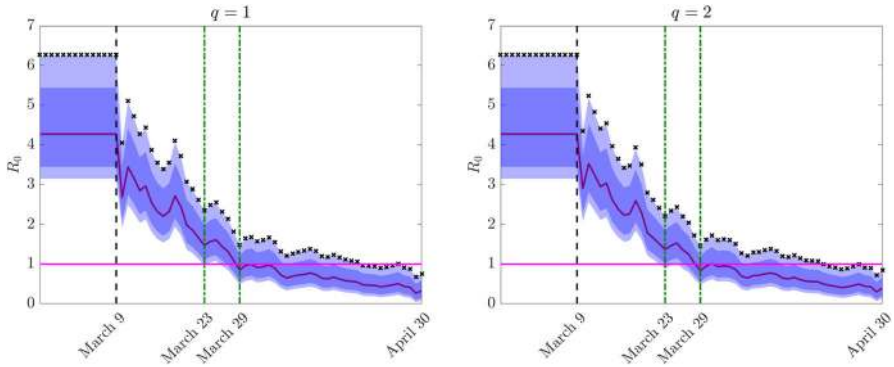
Fig. 7 Test 2. Estimated reproduction number R_0 from the feedback controlled model with uncertain data (35) for a perception function $\psi(I) = I^q$, $q = 1$ (left) and $q = 2$ (right) together with the confidence bands. We mark with dash-dotted green lines the days in which the lower 95% band and the expected R_0 fell below one, and with x-markers the estimated reproduction number relative to data fitting

Introducing data uncertainty

To account for the impact of uncertainties in the data and parameters we then consider a two-dimensional random variable $\mathbf{z} = (z_1, z_2)$ with independent components such that

$$i(\mathbf{z}, 0) = i_0(1 + \mu z_1), \qquad r(\mathbf{z}, 0) = r_0(1 + \mu z_1), \qquad \mu > 0, \tag{34}$$

and

$$\gamma(\mathbf{z}) = \gamma_e + \alpha_\gamma z_2, \qquad \beta(\mathbf{z}) = \beta_e - \alpha_\beta z_2, \qquad \alpha_\gamma, \alpha_\beta > 0, \tag{35}$$

where z_1, z_2 are chosen distributed as symmetric Beta functions in [0, 1], i_0 and r_0 are the initial number of reported cases and recovered taken from (GitHub: https://github.com/pcmdpc/COVID-19) on February 24, 2020. The choice of a Beta distribution for $p(\mathbf{z}) = p_1(z_1)p_2(z_2)$ is coherent with other authors (Roberts 2013; Xiu 2010). However, different probability distribution functions may be considered if additional information on the nature of the uncertainties are available.

It should be noted that the estimated reproduction number, computed in the first optimization step, corresponds to $R_0^e = \beta_e/\gamma_e \approx 6.3$, which is at the upper limit of the values reported in the literature for COVID-19 (Jagodnik et al. 2020; Zhang et al. 2020; Liu et al. 2020). Being aware of the limitations of the data fitting on reported data, to consider a more realistic range of values we assumed a stochastic dependence in β and γ taking into account the faster recovery of asymptomatic individuals (Gatto et al. 2020) and the fact that asymptomatic individuals might be slightly less infectious than symptomatic cases (Sayampanathan et al. 2021). In the simulations we take $\alpha_\beta = 0.03$, $\alpha_\gamma = 0.04$ and $z_2 \sim B(2, 2)$ in (35). Under this assumptions the reproduction number covers a range of values approximatively in [3.13, 6.27] with an expected value around 4.25.
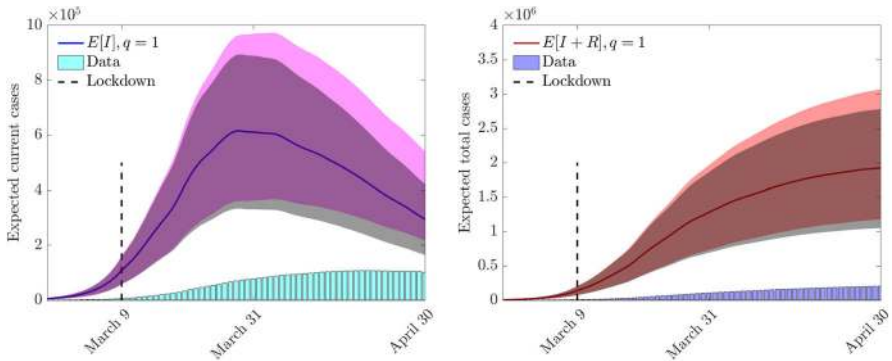
Fig. 8 Test 2. Evolution of expected current cases (left) and of the expected total cases (right) and their 95% confidence bands with respect to z_1 (shaded color) and z_2 (shaded gray) for the feedback controlled model with perception function $\psi(I) = I$, and uncertain initial data (34–35)

The reproduction number in the feedback controlled model is estimated from

$$R_0(z_2, t) = \frac{\beta(z_2) - u(t)\chi(t > \bar{t})}{\gamma(z_2)}. \tag{36}$$

In (36) the time $\bar{t}$ is the lockdown time, in the case under study March 9th, and $\chi(\cdot)$ the indicator function. The feedback control $u(t)$ is defined from (29) in the case of homogeneous mixing and assuming $\mathcal{R}[S(\cdot, t)I(\cdot, t)\psi'(I(\cdot, t))] = S(\mathbf{z}_0, t)I(\mathbf{z}_0, t)^q$ as in (26), where $I(\mathbf{z}_0, t)$ is the total number of infected reported at time t. This leads to

$$u(t) = \frac{1}{\kappa(t)} S(\mathbf{z}_0, t)I(\mathbf{z}_0, t)^q, \tag{37}$$

in agreement with the fact that the confinement restrictions have been implemented accordingly to the reported data. Note that, otherwise the action of the uncertainty translates into the control and leads to the unrealistic effect that the largest is the number of unreported infected the largest is the action of the control in the system.

In Fig. 7 we report the expected value of R_0 together with the 95% and 50% confidence levels with respect to the variable z_2. The estimated reproduction number relative to data fitting is reported with x-marked symbols and represents an upper bound for $R_0(z_2, t)$. The results show that the R_0 reproduction number, thanks to the containment actions, has been drastically reduced and its expected value fell below one between March 23rd and March 29th for both $q = 1$ and $q = 2$. After March 29th the observed R_0 is stably below unity. Note that, in both controls the results are very similar, without any need of renormalization for $q = 2$ due to the data fitting process.

Next we considered the evolution of the uncertain number of infected. In the following we assumed a strongly underestimated initial number of infected (including asymptomatic), taking $\mu = 50$ so that the reported infected along the time horizon of the simulation represent approximately a 20% portion of the total infected persons
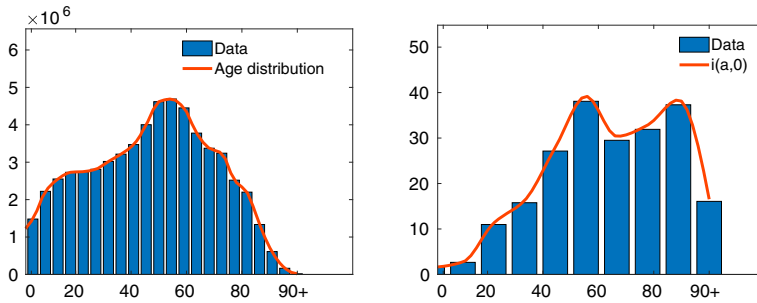
Fig. 9 Test 3. Distribution of age in Italy (left) and distribution of infected (right) together with the corresponding continuous approximations[2]

computed by the feedback controlled model. This is in accordance with the WHO findings that around 80% of infected are asymptomatic[1] and with the results of preliminary serological campaigns promoted in Italy[2].

In Fig. 8 we represent the evolution of the expected value of the number of infected obtained by the controlled model with perception function $\psi(I) = I$ in presence of uncertain contact and recovery rates (35) and initial uncertain data (34) assuming $z_1 \sim B(40, 40)$ and $z_2 \sim B(2, 2)$. We represent the expected values of the current cases (left) and of the total cases (right) along with the 95% confidence level with respect to the variables z_1 and z_2. The shaded color band is relative to the variability in z_1 whereas the shaded gray band is relative to the variability in z_2. The bars below the graph are the reported values of the number of infected on which the model has been calibrated. The results with $q = 2$ do not highlight significant differences with respect to the case $q = 1$ and therefore have been omitted.

4.3 Test 3: The effect of social contacts in the population

Subsequently, we analyze the effects of the inclusion of age dependence and social interactions in the above scenario. More precisely we consider the social interaction rate $\beta = \beta(a, a_*)$, recovery rate $\gamma = \gamma(a)$ and uncertain initial number of infected. These functions were normalized using the estimated parameters β_e and γ_e in accordance with

$$\beta_e = \int_{\Lambda \times \Lambda} \beta(a, a_*) f(a) f(a_*) \, da \, da_*, \qquad \gamma_e = \int_{\Lambda} \gamma(a) f(a) \, da, \qquad (38)$$

where $f(a)$ is the age distribution with $\Lambda = [0, a_{\max}]$, $a_{\max} = 100$.

The age dependent social interaction rate $\beta(a, a_*)$ is defined as follows,

$$\beta(a, a_*) = (1 - \xi)\beta_e + \xi\beta_{\text{social}}(a, a_*), \qquad (39)$$

---

[1] Q&A: Similarities and differences COVID-19 and influenza.
https://www.who.int/news-room/q-a-detail/q-a-similarities-and-differences-covid-19-and-influenza.

[2] Preliminary results on the seroprevalence of SARS-CoV-2 in Italy:
https://www.istat.it/it/files//2020/08/ReportPrimiRisultatiIndagineSiero.pdf
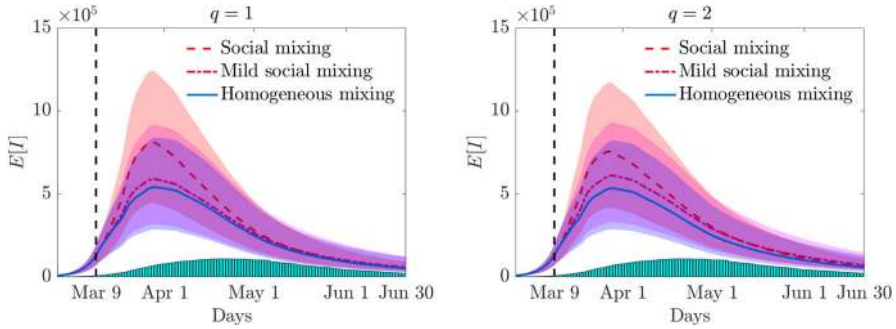
Fig. 10 Test 3. Expected number of infected in time for the perception function $\psi(I) = I^q$, $q = 1$ (left) and $q = 2$ (right) and a constant recovery rate together with the confidence bands for homogeneous mixing ($\xi = 0$), mild social mixing ($\xi = 0.75$) and full social mixing ($\xi = 1$)

where $0 \leq \xi \leq 1$, thus for $\xi = 0$ we recover the homogeneuos mixing, whereas for $\xi = 1$ we have a full social mixing behavior. The social interaction function, $\beta_{\text{social}}(a, a_*)$, accounts for the interactions due to specific activities $\mathcal{A} = \{$Family, Education, Profession$\}$ and is defined by (55) in Appendix C. However, since after the discovery of the first case (February 21), schools, and many places of aggregation were closed in most regions of Northern Italy, we assume that β_E is 0 from February 24 onwards, while β_P is reduced by a factor one-half from March 9 onwards.

The choice of age dependent recovery rate $\gamma(a)$ involves a certain degree of arbitrariness, nevertheless it is reasonable to account such heterogeneity as observed in different studies (Faes et al. 2020; Voinsky et al. 2020; Verity 2020; Wang et al. 2020; Russell et al. 2020; Paradisi and Rinaldi 2020). In order to account fast recovery rate of young people, and slow recovery of the eldest we chose $\gamma(a)$ to be constant up to a specific age a_0 and then a decreasing function of the age. We express mathematically the recovery rate as

$$\gamma(a) = C_\gamma \left(\chi(a \leq a_0) + (1 - \chi(a \leq a_o))e^{-r(a-a_o)} \right), \tag{40}$$

with $r = 4.5$, $a_o = 20$ and $C_\gamma \in \mathbb{R}_+$ such that (38) holds.

We divided the computation time frame into two zones and used different models in each zone, in accordance with the policy adopted by the Italian Government. The first time interval defines the period without any form of containment from 24 February to 9 March, the second the lockdown period from 9 March. In the first zone we adopt the uncontrolled model with homogeneous mixing. Hence, in the second zone we compute the evolution of the feedback controlled age dependent model (27–29) with matching (on average) interaction and recovery rates accordingly to (38) and with the estimated value of the control penalization $\kappa(t)$ as reported in Fig. 6 until April 30. After April 30 the computation advances in time using as penalization term the constant asymptotic
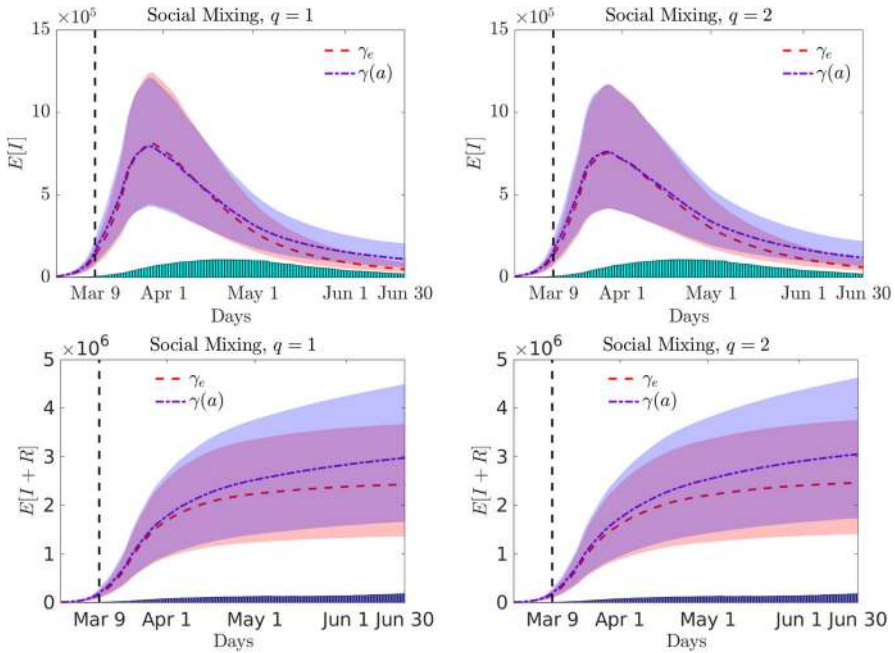
Fig. 11 Test 3. Expected number of infected and total cases of infected and recovered in time for a perception function $\psi(I) = I^q$, $q = 1$ (left) and $q = 2$ (right) together with the confidence bands for the social mixing scenario with age-independent or age-dependent recovery rate

value $\bar{\kappa}$ reached by $\kappa(t)$. The initial values for the age distributions of susceptible and infectious individuals are shown in Fig. 9 in agreement with reported data[3].

In Fig. 10 we report the results of the expected number of infected with the related confidence bands in case of homogeneous mixing and different levels of social mixing ($\xi = 0.75$, $\xi = 1$) for the constant recovery rate γ_e. The homogeneous mixing hypothesis leads to a lower estimate of the maximum number of infected and shows a slower decay over time in the case $q = 1$, whereas for $q = 2$ the decay is comparable. In Fig. 11 we compare the case of constant and age-dependent recovery rates for the social mixing scenario. The expected number of infected are shown in the top row, bottom row depicts the total number of recovered and infected people. The heterogeneity of the recovery rate makes the epidemic more persistent and causes an increase in the total number of cases with respect to the homogeneous recovery rate. Finally, in Fig. 12 we report the expected age distribution of infectious individuals in time for $q = 1$. It is evident the effect of the age dependent recovery rate in the increase of cases among the oldest part of the population. Note that, this effect is partially compensated by the strength of the social mixing which reduces interaction among the elderly.

---

[3] Source ISTAT (https://www.istat.it) and Istituto Superiore di Sanità (https://www.epicentro.iss.it)
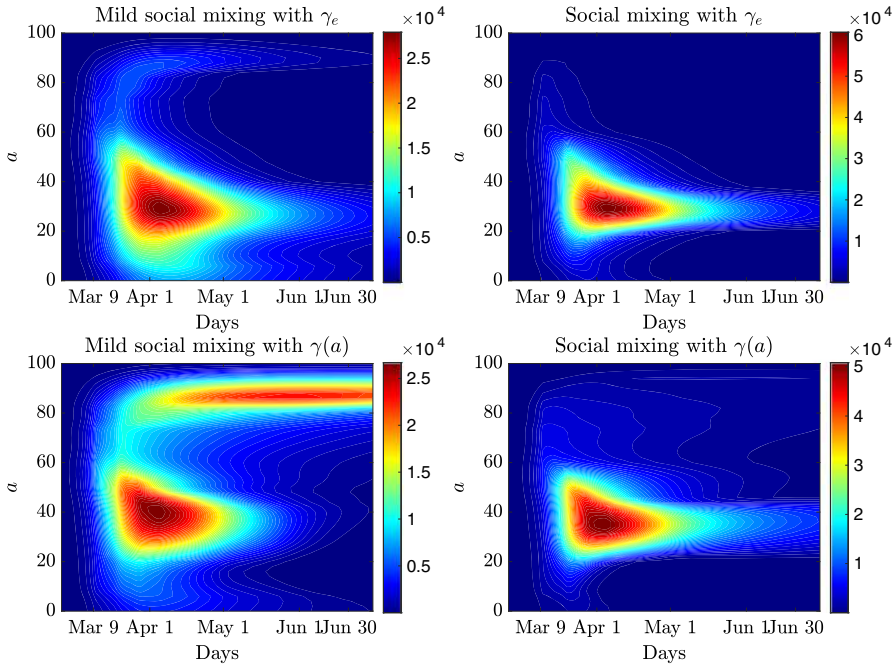
Fig. 12 **Test 3**. Expected age distribution of infectious individuals for a perception function $\psi(I) = I$ with mild (left) and full (right) social mixing. In the top row $\gamma = \gamma_e$ and in the bottom row $\gamma = \gamma(a)$ defined in (40)

4.4 Test 4: Reducing the epidemic through relaxed social containment

One of the major problems in the application of very strong containment strategies, such as the lockdown applied in Italy, is the difficulty in maintaining them over a long period, both for the economic impact and for the impact on the population from a social point of view. In order to analyze sustainable control strategies, therefore, it is necessary to resort to models with a social structure and control methods based on specific forms of social distancing that allow the economy to restart and the population to dedicate itself, albeit in a limited way, to its pre-pandemic activities.

In accordance with the interaction function introduced in the Appendix C, we considered the following age dependent relaxation function

$$\Psi(a, a_*, t) = \begin{cases} p_s & a, a_* \in \Lambda_P \\ p_w & \text{elsewhere} \end{cases} \tag{41}$$

where the interval Λ_P defines the age group related to a stronger relaxation of the restrictions (in the sequel we assume $\Lambda_P = [25, 65]$), and the parameters $0 \leq p_w < p_s \leq 1$ characterize the intensity of the heterogeneity of the relaxation process over the different age classes. Hence we relax the control parameter defined in (29) according

Table 1 Reduction of the feedback control (42) over different time periods due to the relaxation of the lockdown processes by the choice of the parameter p_s and p_w of the age dependent function Ψ defined in (41)

	Until March 9	March 9–May 3	May 4–June 2	from June 3
$p_s \times 100\%$	–	0%	15%	20%
$p_w \times 100\%$	–	0%	5%	10%

to

$$u_{\text{relax}}(a, a_*, t) = (1 - \Psi(a, a_*, t))u(a, a_*, t). \tag{42}$$

The evolution of the infection is considered within two different relaxation times, at May 4 and at June 3, as actuated by the Italian Government during the first wave of the pandemic. We report in Table 1 the specific choice of the parameter p_s and p_w associated to different periods. Note that, these values are directly related with an increase of the disease transmission rather than to a relaxation of restrictions. In fact, it is clear that under suitable safety precautions a relaxation of restrictions may not contribute to the progress of the epidemic.

In Fig. 13 (top row) we report the evolution of the age-controlled model with perception function $\psi(I) = I$, with mild social mixing and with full social mixing and homogeneous recovery rate. Each plot compare the evolution with relaxation of the containment measure and without, respectively indicated with $E[I_{\text{relax}}]$ and $E[I]$. It is easily observed how the relaxation process increases the number of infected persons. Although the expected value remains under control, the wider confidence bands highlight the risk of a resumption of the epidemic. Of course, a further relaxation of the values in Table 1 will lead to a higher risk of restarting the epidemic wave. In Fig. 13 (bottom row) we report also the evolution of the age-dependent expected number of infected individuals, both for mild social mixing and full social mixing. We remark that, as expected, the relaxation process focuses the increase of infections in the interval characterized by Λ_P.

As can be seen, a gradual strategy can keep the average number of infections under control and have an outcome comparable to the fully controlled model at a potentially lower social cost. The timing and intensity of interventions, however, are crucial to prevent the restart of the epidemic wave.

5 Conclusions

Quantifying the impact of uncertain data in the context of an epidemic emergency is essential in order to design appropriate containment measures. Such containment measures, implemented by several countries in the course of the COVID-19 epidemic, have proved effective in reducing the R_0 reproduction number to below or very close to one. These large-scale non-pharmaceutical interventions vary from country to country but include social distancing (banning large mass events, closing public places and advis-
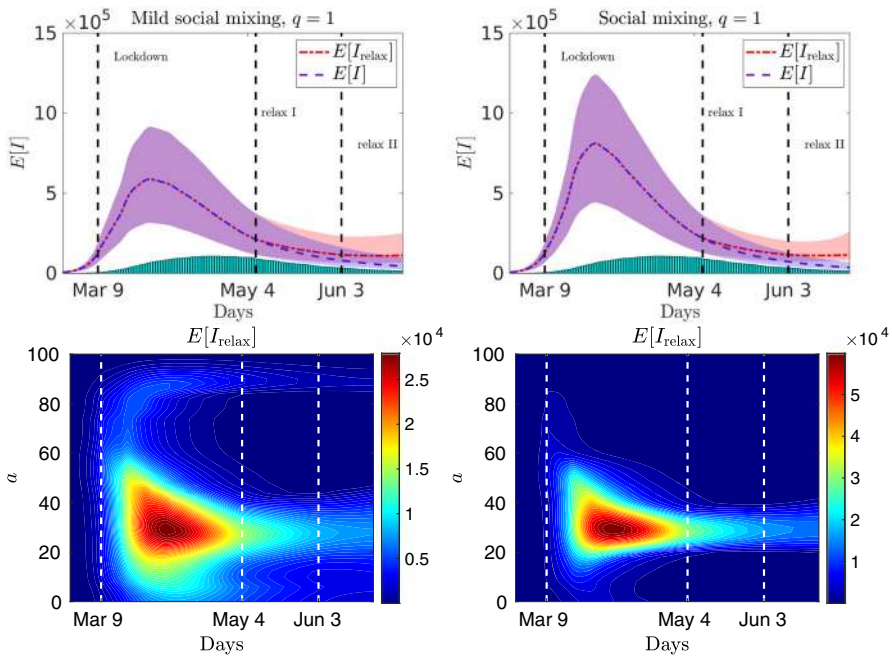
Fig. 13 Test 4. Expected number of infected with relaxed control (42) characterized by Table 1, perception function $\psi(I) = I$, and homogeneous recovery rate, using mild social mixing (left) and with full social mixing (right). In the bottom row the corresponding expected age distribution of infectious individuals is reported

ing people not to socialize outside their families), closing borders, closing schools, measures to isolate symptomatic individuals and their contacts, and the large-scale lock-down of populations with all but essential prohibited travel.

One of the main problems is the sustainability of these interventions, which until the introduction of a vaccine will have to be maintained in the field for long periods. However, estimating the reproductive numbers of SARS-CoV-2 is a major challenge due to the high proportion of infections not detected by health care systems and differences in test application, resulting in diverse proportions of infections detected over time and between countries. Most countries initially had only the capacity to test a small proportion of suspected cases and tests were reserved for severely ill patients or high risk groups. The available data therefore offer a systematic partial overview of trends.

In this article, starting from a SIR-type compartmental model with social structure, we developed new stochastic mathematical models describing the actions of a government agency to contain infections among the population in presence of an uncertain number of infectious individuals. More precisely, in the present model, the state of the infected is defined by a multi-dimensional time-dependent function characterized by the age, which proved essential in the description of the COVID-19 pandemic, and by a systematic uncertainty which permits to avoid additional sub-compartmentalization of the infectious population.

These assumptions allows to derive a socially structured model that contains the control action in feedback form based on the perception of the policy maker of the spread of the disease. Subsequently, the uncertainty in the model has been approximated by expanding the state variables into orthogonal polynomials in the random space, reducing the problem to a set of deterministic equations for the distribution of the solution through the course of the epidemic. The resulting controlled stochastic dynamical system is then solved using this deterministic formulation, which in the case of sufficiently regular uncertainty distributions allows efficient and accurate estimations of the random parameters.

The numerical simulations, performed using data from the recent COVID-19 outbreak in Italy, show, on the one hand, the model's ability to characterize the presence of the asymptomatic population trough the introduction of an uncertainty in the number of infected and in their epidemic role, and, on the other hand, in presence of control to well describe the effects of non pharmaceutical interventions aimed at flattening the infection curve. In particular, identifying some scenarios in agreement with government actions, containment measures by the population based on the resumption of certain occupational activities characterized by specific age groups and social interaction matrices were studied.

Further studies in this direction will aim to consider more comprehensive epidemic models, including the effects of other clinical compartments of interest, along with generalization of the control term to different objective functions or to the case of multiple control terms for each social activity, in order to design optimal strategies to mitigate the overall epidemic impact.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

A Feedback controlled models with additional compartmentalizations

In this Appendix, we will detail how to extend the instantaneous control approach introduced in Sect. 2.3 in the case of a a socially-structured SEPIAR-type compartmentalization (Gatto et al. 2020) where the control functional is given by

$$
\min_{\mathbf{u}\in\mathcal{U}} J(\mathbf{u}) := \int_0^T \psi(I_T(t))dt
$$
$$
+ \frac{1}{2} \sum_{j\in\{P,I,A\}} \int_0^T \int_{\Lambda\times\Lambda} \nu_j(a, a_*, t)|u_j(a, a_*, t)|^2 \, da da_* dt, \tag{43}
$$

with $\mathbf{u} = (u_P, u_I, u_A)$ and subject to

$$
\frac{d}{dt}s(a, t) = -s(a, t) \sum_{j\in\{P,I,A\}} \int_\Lambda (\beta_j(a, a_*) - u_j(a, a_*, t))i_j(a_*, t) \, da_*
$$
$$
\frac{d}{dt}e(a, t) = s(a, t) \sum_{j\in\{P,I,A\}} \int_\Lambda (\beta_j(a, a_*) - u_j(a, a_*, t))i_j(a_*, t) \, da_* - \delta_E(a)e(a, t)
$$
$$
\frac{d}{dt}i_P(a, t) = \delta_E(a)e(a, t) - \delta_P(a)i_P(a, t)
$$
$$
\frac{d}{dt}i_I(a, t) = \sigma(a)\delta_P(a)i_P(a, t) - \gamma_I(a)i_I(a, t)
$$
$$
\frac{d}{dt}i_A(a, t) = (1 - \sigma(a))\delta_P(a)i_P(a, t) - \gamma_A(a)i_A(a, t)
$$
$$
\frac{d}{dt}r(a, t) = \gamma_I(a)i(a, t) + \gamma_A(a)i_A(a, t). \tag{44}
$$

In the above compartmentalization e, i_P, i_I, and i_A represent the number of exposed, pre-asymptomatic, infected with heavy symptoms and asymptomatic/mildly symptomatic individuals, respectively (Gatto et al. 2020). The parameters β_j, $j \in \{P, I, A\}$ are the specific transmission rates of the three infectious classes. Exposed individuals, still not contagious, enter the presymptom stage at rate δ_E and only then become infectious. Pre-symptomatic individuals progress at rate δ_P to become symptomatic who develop severe symptoms with probability σ or asymptomatic individuals with probability $1 - \sigma$. Symptomatic individuals recover from infection at rate γ_I whereas asymptomatic individuals leave their compartment after having recovered from infection at rate γ_A.

In (43) we assume that the policy maker aims at minimizing the total number of infected

$$
I_T(t) = \int_\Lambda (i_P(a, t) + i_I(a, t) + i_A(a, t)) \, da.
$$

We highlight that other control strategies may be considered as well leading to different feedback control formulations.

For the socially-structured SEPIAR model the feedback controlled formulation is then obtained from the discrete approximation

$$\min_{\mathbf{u}\in\mathcal{U}} J_h(\mathbf{u}) := \psi(I_T(t+h)) + \frac{1}{2} \sum_{j\in\{P,I,A\}} \int_{\Lambda\times\Lambda} v_j(a, a_*, t)|u_j(a, a_*, t)|^2 \, da da_* \quad (45)$$

subject to

$$s(a, t+h) = s(a, t) - hs(a, t) \sum_{j\in\{P,I,A\}} \int_\Lambda (\beta_j(a, a_*) - u_j(a, a_*, t)) i_j(a_*, t) \, da_*$$

$$e(a, t+h) = e(a, t) + hs(a, t) \sum_{j\in\{P,I,A\}} \int_\Lambda (\beta_j(a, a_*) - u_j(a, a_*, t)) i_j(a_*, t) \, da_*$$

$$-h\delta_E(a)e(a, t)$$

$$i_P(a, t+h) = i_P(a, t) + h\delta_E(a)e(a, t+h) - h\delta_P(a)i_P(a, t)$$

$$i_I(a, t+h) = i_I(a, t) + h\sigma(a)\delta_P(a)i_P(a, t+h) - h\gamma_I(a)i_I(a, t)$$

$$i_A(a, t+h) = i_A(a, t) + h(1 - \sigma(a))\delta_P(a)i_P(a, t+h) - h\gamma_A(a)i_A(a, t) \quad (46)$$

where the discrete equation for $r(a, t)$ can be omitted since the control does not play any direct role in its evolution. The total number of infected evolved accordingly to

$$I_T(t+h) = I_T(t) + h \int_\Lambda (\delta_E(a)e(a, t+h) - \gamma_I(a)i_I(a, t) - \gamma(a)i_A(a, t)) \, da. \quad (47)$$

We can derive the minimizer of J_h computing $\nabla_\mathbf{u} J_h(\mathbf{u}) \equiv 0$ or equivalently

$$\frac{\partial J_h(u_i)}{\partial u_j} = 0, \quad i, j \in \{P, I, A\}.$$

Using (45) we can compute

$$\psi'(I_T(t+h))\frac{\partial I_T(a, t+h)}{\partial u_j} = -v_j(a, a_*, t)u_j(a, a_*, t), \quad j \in \{P, I, A\},$$

which by virtue of (47) leads to the discrete feedback controls

$$u_j(a, a_*, t) = \frac{h^2\delta_E(a)}{v_j(a, a_*, t)}s(a, t)i_j(a_*, t)\psi'(I_T(t+h)), \quad j \in \{P, I, A\}. \quad (48)$$

In order to pass to the limit $h \to 0$ in (46) and obtain the feedback controlled model, we must rescale the penalization parameters with respect to the short time-horizon h. With the aim of preserving the compartment of exposed, the penalization terms

should rescale as $v_j(a, a_*, t) = h^2 \kappa_j(a, a_*, t)$ to get the feedback controlled model (44) where the control terms are given by

$$u_j(a, a_*, t) = \frac{\delta_E(a)}{\kappa_j(a, a_*, t)} s(a, t) i_j(a_*, t) \psi'(I_T(t)), \quad j \in \{P, I, A\}. \quad (49)$$

It is interesting to observe that the control action in (49) is inversely proportional to the incubation time, namely shorter incubation periods requires a stronger control. On the contrary, if one assumes the same rescaling adopted for the SIR model in (12), namely $v_j(a, a_*, t) = h \kappa_j(a, a_*, t)$, and select a short time horizon equal to the incubation time $h = 1/\delta_E$, the exposed compartment in (46) reduces to

$$e(a, t + h) = h s(a, t) \sum_{j \in \{P, I, A\}} \int_\Lambda (\beta_j(a, a_*) - u_j(a, a_*, t)) i_j(a_*, t) \, da_*$$

which substituted into the equation for $i_P(a, t + h)$ in the limit $h \to 0$ leads to a feedback controlled model in the form (44) without the compartment of exposed and where the pre-symptomatic satisfy

$$\frac{d}{dt} i_P(a, t) = s(a, t) \sum_{j \in \{P, I, A\}} \int_\Lambda (\beta_j(a, a_*) - u_j(a, a_*, t)) i_j(a_*, t) \, da_* - \delta_P(a) i_P(a, t).$$

Finally, we emphasize that the generalization to the case with uncertainty contains no difficulty and follows the same methodology introduced in Sect. 3. In terms of data-fitting, similar to the SIR model, one considers the feedback controlled model in absence of age-dependence and proceeds along the lines indicated in Remark 2. We will omit the details for brevity.

B Stochastic Galerkin approximation

In this Appendix we give the details of the stochastic Galerkin (sG) method used to solve the feedback controlled system (27–29) with uncertainties. To this aim, we consider a random vector $\mathbf{z} = (z_1, \ldots, z_d)$ with independent components and whose distribution is $p(\mathbf{z}) : \mathbb{R}^{d_z} \to \mathbb{R}_+^{d_z}$. The stochastic Galerkin approximation of the differential model (27–29) is based on stochastic orthogonal polynomials and provides a spectrally accurate solution under suitable regularity assumptions, see (Xiu 2010). We consider the linear space $\mathbb{P}_M$ of polynomials of degree up to M generated by a family of polynomials $\{\Phi_\mathbf{h}(\mathbf{z})\}_{|\mathbf{h}|=0}^M$ that are orthonormal in the space $L^2(\Omega)$ such that

$$\mathbb{E}[\Phi_h(\cdot) \Phi_k(\cdot)] = \delta_{\mathbf{kh}}, \qquad 0 \le |\mathbf{k}|, |\mathbf{h}| \le M$$

being $\mathbf{k} = (k_1, \ldots, k_d)$ a multi-index, $|\mathbf{k}| = k_1 + \cdots + k_{d_z}$ with δ_{kh} the Kronecker delta function, and $\mathbb{E}[\cdot]$ the expectation with respect to $p(\mathbf{z})$. The construction of the polynomial basis $\{\Phi_\mathbf{h}(\mathbf{z})\}_{|\mathbf{h}|=0}^M$ depends on the distribution of the uncertainties and

Table 2 The different polynomial expansions connected to the probability distribution of the random component $z_k, k = 1, \ldots, d_z$

Probability law	Expansion polynomials	Support
Gaussian	Hermite	$(-\infty, +\infty)$
Uniform	Legendre	$[a, b]$
Beta	Jacobi	$[a, b]$
Gamma	Laguerre	$[0, +\infty)$
Poisson	Charlier	$\mathbb{N}$

must be chosen in agreement with the Askey scheme (Xiu 2010). We summarize in Table 2 several polynomials bases in connection with the law of a random component of **z**.

By assuming $s(\mathbf{z}, a, t)$, $i(\mathbf{z}, a, t)$ and $r(\mathbf{z}, a, t)$ in $L^2(\Omega)$ we may approximate these terms through a generalized polynomial chaos expansion in the random space as follows

$$s(\mathbf{z}, a, t) \approx s^M(\mathbf{z}, a, t) = \sum_{|\mathbf{k}|=0}^{M} \hat{s}_{\mathbf{k}}(a, t) \Phi_{\mathbf{k}}(\mathbf{z})$$

$$i(\mathbf{z}, a, t) \approx i^M(\mathbf{z}, a, t) = \sum_{|\mathbf{k}|=0}^{M} \hat{i}_{\mathbf{k}}(a, t) \Phi_{\mathbf{k}}(\mathbf{z}) \qquad (50)$$

$$r(\mathbf{z}, a, t) \approx r^M(\mathbf{z}, a, t) = \sum_{|\mathbf{k}|=0}^{M} \hat{r}_{\mathbf{k}}(a, t) \Phi_{\mathbf{k}}(\mathbf{z}),$$

where the quantities $\hat{s}_{\mathbf{k}}, \hat{i}_{\mathbf{k}}, \hat{r}_{\mathbf{k}}$ are projections in the polynomial space

$$\hat{s}_{\mathbf{k}} = \mathbb{E}\left[s(\cdot, a, t)\Phi_{\mathbf{k}}(\cdot)\right], \quad \hat{i}_{\mathbf{k}} = \mathbb{E}\left[i(\cdot, a, t)\Phi_{\mathbf{k}}(\cdot)\right], \quad \hat{r}_{\mathbf{k}} = \mathbb{E}\left[r(\cdot, a, t)\Phi_{\mathbf{k}}(\cdot)\right].$$

The sG formulation of system (27) is obtained first by replacing the solutions with their stochastic polynomial expansions

$$\frac{d}{dt} s^M(\mathbf{z}, a, t) = -s^M(\mathbf{z}, a, t) \int_{\Lambda} \left(\beta(\mathbf{z}, a, a_*) - u^M(a, a_*)\right) i^M(\mathbf{z}, a_*, t) da_*$$

$$\frac{d}{dt} i^M(\mathbf{z}, a, t) = s^M(\mathbf{z}, a, t) \int_{\Lambda} \left(\beta(\mathbf{z}, a, a_*) - u^M(a, a_*)\right) i^M(\mathbf{z}, a_*, t) da_*$$

$$\qquad\qquad - \gamma(\mathbf{z}, a) i^M(\mathbf{z}, a, t) \qquad (51)$$

$$\frac{d}{dt} r^M(\mathbf{z}, a, t) = \gamma(\mathbf{z}, a) i^M(\mathbf{z}, a, t),$$

with

$$u^M(a, a_*) = \frac{1}{\kappa(a, a_*)} \mathcal{R}\left[s^M(\mathbf{z}, a, t) i^M(\mathbf{z}, a_*, t) \psi'(I^M(\mathbf{z}, t))\right],$$

$$I^M(\mathbf{z}, t) = \sum_{|\mathbf{k}|=0}^{M} \hat{I}_{\mathbf{k}}(t)\Phi_{\mathbf{k}}(\mathbf{z}), \qquad \hat{I}_{\mathbf{k}}(t) = \int_\Lambda i^M(\mathbf{z}, a, t)da, \tag{52}$$

and where s^M, i^M, r^M are defined by (50). Then, thanks to the orthonormality of the polynomial basis of $\mathbb{P}_M$, multiplying by $\Phi_{\mathbf{m}}$, for all $|\mathbf{m}| \leq M$, and taking the expectation with respect to $p(\mathbf{z})$ we obtain the following coupled system of $M+1$ deterministic equations for the evolution of each projection

$$\begin{cases} \dfrac{d}{dt}\hat{s}_{\mathbf{k}}(a, t) = -\sum_{|\mathbf{m}|=0}^{M} \mathbf{B}_{\mathbf{km}}(a, t)\hat{s}_{\mathbf{m}}(a, t) \\ \dfrac{d}{dt}\hat{i}_{\mathbf{k}}(a, t) = \sum_{|\mathbf{m}|=0}^{M} \mathbf{B}_{\mathbf{km}}(a, t)\hat{s}_{\mathbf{m}}(a, t) - \sum_{|\mathbf{m}|=0}^{M} \mathbf{G}_{\mathbf{km}}\hat{i}_{\mathbf{m}}(a, t) \\ \dfrac{d}{dt}\hat{r}_{\mathbf{k}}(a, t) = \sum_{|\mathbf{m}|=0}^{M} \mathbf{G}_{\mathbf{km}}\hat{i}_{\mathbf{m}}(a, t) \end{cases} \tag{53}$$

where

$$\mathbf{B}_{\mathbf{km}} = \sum_{|\mathbf{l}|=0}^{M} \int_{\mathbb{R}^{d_z}} \left(\int_\Lambda \left(\beta(\mathbf{z}, a, a_*) - u^M(a, a_*) \right) \hat{i}_{\mathbf{l}}(a_*, t)da_* \right) \Phi_{\mathbf{k}}(\mathbf{z})\Phi_{\mathbf{m}}(\mathbf{z})\Phi_{\mathbf{l}}(\mathbf{z})p(\mathbf{z})d\mathbf{z} \tag{54}$$

$$\mathbf{G}_{\mathbf{km}} = \int_{\mathbb{R}^{d_z}} \gamma(\mathbf{z}, a)\Phi_{\mathbf{k}}(\mathbf{z})\Phi_{\mathbf{m}}(\mathbf{z})p(\mathbf{z})d\mathbf{z}.$$

The above system is then integrated in time directly in the space of projections. We remark that statistical quantities of interest, such as expectation and variance of infected, can be recovered as

$$\mathbb{E}[i^M(\cdot, a, t)] = \int_{\mathbb{R}^{d_z}} i^M(\mathbf{z}, a, t)p(\mathbf{z})\,d\mathbf{z} = \sum_{|\mathbf{k}|=0}^{M} \hat{i}_{\mathbf{k}}(a, t)\mathbb{E}[\Phi_{\mathbf{k}}(\cdot)] = \hat{i}_{\mathbf{0}}(a, t),$$

whereas for the variance we get

$$\begin{aligned} \mathrm{Var}(i^M(\cdot, a, t)) &= \mathbb{E}[i^M(\cdot, a, t)^2] - \mathbb{E}[i^M(\cdot, a, t)]^2 \\ &= \sum_{|\mathbf{k}|=0}^{M} \sum_{|\mathbf{h}|=0}^{M} \hat{i}_{\mathbf{k}}\hat{i}_{\mathbf{h}}\mathbb{E}[\Phi_{\mathbf{k}}(\cdot)\Phi_{\mathbf{h}}(\cdot)] - \hat{i}_{\mathbf{0}}^2(a, t) = \sum_{|\mathbf{k}|=0}^{M} \hat{i}_{\mathbf{k}}^2(a, t) - \hat{i}_{\mathbf{0}}^2(a, t). \end{aligned}$$

In all the simulations reported we used $M = 10$ and a fourth order Runge-Kutta method for the time integration.

C Social mixing functions

This last appendix is devoted to report the details of the social interaction functions that characterize the dynamics of social mixing. These characteristics are in fact crucial for a correct prediction of outcomes, especially in diseases transmitted by close contacts. Several large-scale studies have been designed in the last decade to determine relevant age-based models in social mixing. Without attempting to review the vast literature on this topic, we mention (Béraud et al. 2015; Mossong et al. 2008; Prem et al. 2017) and the references therein.

The number of contacts per person generally shows considerable variability according to age, occupation, country and even day of the week, in relation to the social habits of the population. Nevertheless, some universal behaviors can be extracted which emerge as a function of specific social activities. Social mixing is highly age-related, which means that people usually tend to interact with other people of a similar age. Young people have a high rate of contact with adults aged around 30-39 and older people over 65, i.e. their parents and grandparents respectively. Contact rates are indeed very high at home and at school. On the other hand, professional mixing is weakly assortative by age and tends to be determined by uniform interactions, approximately between people from 25 and 65 years old.

For these reasons we consider an interaction function determined by three main sub-functions that characterize the family, the school and the professional mixing. Therefore, a stylized function approximating a realistic contact matrix can be written as follows

$$\beta_{\text{social}}(a, a_*) = C_\beta \sum_{j \in \mathcal{A}} \omega_j \beta_j(a, a_*), \tag{55}$$

where the functions $\beta_j(a, a_*)$ take into account the different contact rates of people with ages a and a_* in relation to specific social activities of the type $\mathcal{A} = \{F, E, P\}$, where we identify family contacts with F, education and school contacts with E and professional contacts with P, and the associated weights ω_F, ω_E and ω_P. The particular structure of these social interaction matrices was determined empirically in Béraud et al. (2015); Prem et al. (2017). Here, according to these observations, we propose suitable mathematical functions that can be calibrated to reproduce empirical observations, in what follow we normalize $\beta_e = 1$ and $a_{\max} = 1$.

In details, familiar contacts tend to concentrate on a three-band matrix with a peak around younger ages. This can be reproduced considering the function

$$\tilde{\beta}_F(a, a_*) = \frac{\lambda_{F,1}}{(1 + (a - a_*)^2)^{\lambda_{F,2}}} + \sqrt{(a^2 + a_*^2)} \exp\left\{-\frac{1}{\sigma_F^2} \frac{a^2 + a_*^2}{(1 + (a - a_*)^2)^{\lambda_{F,2}}}\right\}.$$

Hence, we define the family interactions as

$$\beta_F(a, a_*) = \beta_0 + \sum_{\ell = \{-\mu, 0, \mu\}} \frac{C_\ell}{2} \left[\tilde{\beta}_F(a + \mu, a_*) + \tilde{\beta}_F(a, a_* + \mu)\right],$$
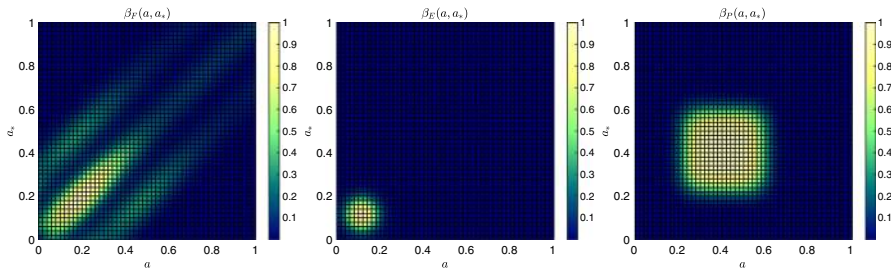
Fig. 14 From left to right, contour plot of the social interaction functions β_F, β_E and β_P taking into account the different contact rates of people with ages a and a_* in relation to specific social activities. The function β_F characterizes the family contacts, β_E the education and school contacts, and β_P the professional contacts

Fig. 15 Social interaction functions
$\beta_{\text{social}} = \beta_F + \beta_E + \beta_P$
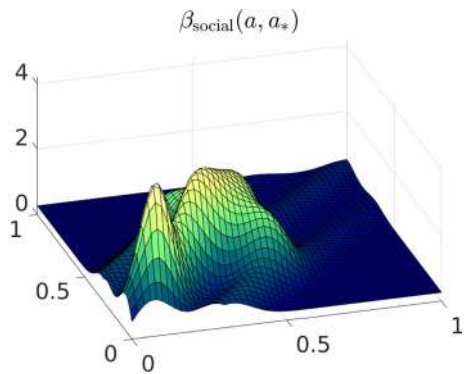
being $\mu > 0$ the age shift at which family contacts occur and $C_{\pm\mu} = 1/2$ and $C_0 = 1$. are reported in Table 3

On the other hand, school and professional interactions are more age-specific and the corresponding matrices can be reproduced as follows

$$\beta_E(a, a_*) = \beta_0 + \exp\left\{-\frac{1}{\sigma_E^2}\left((a - \lambda_E)^2 + (a_* - \lambda_E)^2\right)\right\}$$

$$\beta_P(a, a_*) = \beta_0 + \exp\left\{-\frac{1}{\sigma_P^2}\left((a - \lambda_P)^4 + (a_* - \lambda_P)^4\right)\right\},$$

being $\lambda_E > 0$ the average contact age at school, and $\lambda_P > 0$ the average professional contact age. The coefficient $C_F, C_E, C_P > 0$ measure the impact of the different contact function in the dynamics. In Figs. 14 and 15 we represent the three social interaction functions and the resulting global social interaction function $\beta_{\text{social}}(a, a_*)$ assuming $\omega_F = \omega_E = \omega_P = 1$. In Table 3 we report the choice of the parameters used in the simulations.

Table 3 The parameters defining the details of the interaction functions used in the simulations

Contact function	Parameters							
$\beta_F(a, a_*)$	β_0	σ_F^2	$\lambda_{F,1}$	$\lambda_{F,2}$	μ	$C_{\pm\mu}$	C_0	ω_F
	0.04	0.125	0.02	100	0.3	0.5	1	1
$\beta_E(a, a_*)$	β_0	σ_E^2	λ_E					ω_E
	0.04	0	0.105					0
$\beta_P(a, a_*)$	β_0	σ_P^2	λ_P					ω_P
	0.04	0.00125	0.4					0.5

References

Albi G, Pareschi L (2018) Selective model-predictive control for flocking systems. Commun Appl Ind Math 9(2):4–21

Albi G, Pareschi L, Zanella M (2014) Boltzmann-type control of opinion consensus through leaders. Philos Trans R Soc Lond Ser A Math Phys Eng Sci 372(2028):20140138

Albi G, Pareschi L, Zanella M (2015a) Uncertainty quantification in control problems for flocking models. Math Probl Eng **2015**, 850124

Albi G, Herty M, Pareschi L (2015b) Kinetic description of optimal control problems and applications to opinion consensus. Commun Math Sci 13(6):1407–1429

Barro M, Guiro A, Ouedraogo D (2018) Optimal control of a SIR epidemic model with general incidence function and a time delays. Cubo 20(2):53–66

Béraud G, Kazmercziak S, Beutels P, Levy-Bruhl D, Lenne X, Mielcarek N, Yazdanpanah Y, Boëlle PY, Hens N, Dervaux B (2015) The French connection: the first large population-based contact survey in France relevant for the spread of infectious diseases. PLoS ONE 10(7):e0133203

Bolzoni L, Bonacini E, Soresina C, Groppi M (2017) Time-optimal control strategies in SIR epidemic models. Math Biosci 292:86–96

Bongini M, Fornasier M, Kalise D (2015) (Un)conditional consensus emergence under perturbed and decentralized feedback controls. Discrete Contin Dyn Syst 35(9):4071–4094

Britton T, Ball F, Trapman P (2020) The disease-induced herd immunity level for Covid-19 is substantially lower than the classical herd immunity level. *preprint* arXiv:2005.03085

Capaldi A, Behrend S, Berman B, Smith J, Wright J, Lloyd AL (2012) Parameter estimation and uncertainty quantification for an epidemic model. Math Biosci Eng 9(3):553–576

Capasso V, Serio G (1978) A generalization of the Kermack–McKendrick deterministic epidemic model. Math Biosci 42(1):43–61

Capistrán MA, Christen JA, Velasco-Hernández JX (2012) Towards uncertainty quantification and inference in the stochastic SIR epidemic model. Math Biosci 240(2):250–259

Caponigro M, Fornasier M, Piccoli B, Trélat E (2013) Sparse stabilization and optimal control of the Cucker–Smale model. Math Control Relat Fields 3(4):447–466

Castillo-Chavez C, Hethcote HW, Andreasen V, Levin SA, Liu WM (1989) Epidemiological models with age structure, proportionate mixing, and cross-immunity. J Math Biol 27(3):233–258

Colombo RM, Garavello M (2019) Optimizing vaccination strategies in an age structured SIR model. Math Bios Eng 17(2):1074–1089

Chen J et al (2020) COVID-19 infection: the China and Italy perspectives. Cell Death Dis 11:438

Chowell G (2017) Fitting dynamic models to epidemic outbreaks with quantified uncertainty: a primer for parameter uncertainty, identifiability, and forecasts. Infect Dis Model 2(3):379–398

Dimarco G, Pareschi L, Zanella M (2017) Uncertainty quantification for kinetic models in socio-economic and life sciences. In: Jin S, Pareschi L (eds)Uncertainty quantification for hyperbolic and kinetic equations, SEMA SIMAI Springer Series, **14**, pp 151–191

Dimarco G, Pareschi L, Toscani G, Zanella M (2020) Wealth distribution under the spread of infectious diseases. Phys Rev E 102:022303

Düring B, Pareschi L, Toscani G (2018) Kinetic models for optimal control of wealth inequalities. Eur Phys J B 91:265

Faes C, Abrams S, Van Beckhoven D, Meyfroidt G, Vlieghe E, Hens N (2020) Belgian collaborative group on COVID-19 Hospital surveillance, time between symptom onset, hospitalisation and recovery or death: statistical analysis of Belgian COVID-19 patients. Int J Environ Res Public Health 17(20):7560

Flaxman S et al (2020) Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries, Report 13. Imperial College COVID-19 Response Team, 30 March 2020

Franceschetti A, Pugliese A (2008) Threshold behaviour of a SIR epidemic model with age structure and immigration. J Math Biol 57(1):1–27

Gatto M, Bertuzzo E, Mari L, Miccoli S, Carraro L, Casagrandi R, Rinaldo A (2020) Spread and dynamics of the COVID-19 epidemic in Italy: effect of emergency containment measures. PNAS 117(19):10484–10491

Giordano G, Blanchini F, Bruno R, Colaneri P, Di Filippo A, Di Matteo A, Colaneri M (2020) Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. Nat Med 26:855–860

Glasser J, Feng Z, Moylan A, Del Valle S, Castillo-Chavez C (2012) Mixing in age-structured population models of infectious diseases. Math Bios 235(1):1–7

Hethcote HW (1996) Modeling heterogeneous mixing in infectious disease dynamics. In: Isham V, Medley GFH (eds) Models for infectious human diseases. Cambridge University Press, Cambridge, pp 215–238

Hethcote HW (2000) The mathematics of infectious diseases. SIAM Rev 42(4):599–653

Sayampanathan AA, Heng CS, Pin PH, Pang J, Leong TY, Lee VJ (2021) Infectivity of asymptomatic versus symptomatic COVID-19. The Lancet 397:93–94

Iannelli M, Milner FA, Pugliese A (1992) Analytical and numerical results for the age-structured S-I-S epidemic model with mixed inter-intracohort transmission. SIAM J Math Anal 23(3):662–688

Jagodnik K, Ray F, Giorgi FM, Lachmann A. Correcting under-reported COVID-19 case numbers: estimating the true scale of the pandemic. Preprint medRvix:2020.03.14.20036178

Jin S, Pareschi L (2017) Uncertainty quantification for hyperbolic and kinetic equations. SEMA-SIMAI Springer series, 14

Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. Proc R Soc Lond A 115:700–721

Lavezzo E et al (2020) Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'. Nature 584:425–429

Lee S, Chowell G, Castillo-Chávez C (2010) Optimal control for pandemic influenza: the role of limited antiviral treatment and isolation. J Theor Biol 265(2):136–150

Lee S, Golinski M, Chowell G (2012) Modeling optimal age-specific vaccination strategies against pandemic influenza. Bull Math Biol 74(4):958–980

Lin F, Muthuraman K, Lawley M (2010) An optimal control theory approach to non-pharmaceutical interventions. BMC Infect Dis 10(1):32

Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J (2020) The reproductive number of COVID-19 is higher compared to SARS coronavirus. J Travel Med 27(2):1–4

Lunelli A, Pugliese A, Rizzo C (2009) Epidemic patch models applied to pandemic influenza: contact matrix, stochasticity, robustness of predictions. Math Biosci 220:24–33

Mizumoto K, Kagaya K, Zarebski A, Chowell G (2020) Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. Euro Surveill 25(10):2000180

Morris DH, Rossine FW, Plotkin JB, Levin SA (2020) Optimal, near-optimal, and robust epidemic control. Preprint arXiv:2004.02209

Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, Massati M, Salmaso S, Scalia Tomba G, Wallinga J, Heijne J, Sadkowska-Todys M, Rosinska M, Edmunds WJ (2008) Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med 5(3):e75

Paradisi M, Rinaldi G (2020) An empirical estimate of the infection fatality rate of COVID-19 from the First Italian Outbreak. SSRN: https://ssrn.com/abstract=3582811 or https://doi.org/10.2139/ssrn.3582811. 18 April 2020

Pareschi L (2021) An introduction to uncertainty quantification for kinetic equations and related problems. In: Albi G, Merino-Aceituno S, Nota A, Zanella M (eds) Trails in kinetic theory: foundational aspects and numerical methods, SEMA-SIMAI Springer series, vol 25

Prem K, Cook AR, Jit M (2017) Projecting social contact matrices in 152 countries using contact surveys and demographic data. PLoS ONE 13(9):e1005697

del Consiglio dei Ministri Presidenza (2020) Dipartimento della Protezione Civile. GitHub: COVID-19 Italia - Monitoraggio situazione. https://github.com/pcm-dpc/COVID-19

Remuzzi A, Remuzzi G (2020) COVID-19 and Italy: What next? Lancet 395:1225–1228

Roberts MG (2013) Epidemic models with uncertainty in the reproduction. J Math Biol 66:1463–1474

Russell TW, Hellewell J, Jarvis CI, van Zandvoort K, Abbott S, Ratnayake R, Flasche S, Eggo RM, Edmunds WJ, Kucharski AJ (2020) Estimating the infection and case fatality ratio for coronavirus disease (COVID-19) using age-adjusted data from the outbreak on the Diamond Princess cruise ship. Euro Surveill 25(12):pii=2000256

Verity R et al (2020) Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet Infect Dis 20(6):669–677

Voinsky I, Baristaite G, Gurwitz D (2020) Effects of age and sex on recovery from COVID-19: analysis of 5769 Israeli patients. J Infect 81(2):102–103

Wang S, Zhong F, Bao W, Li Y, Liu L, Wang H, He Y (2020) Age-dependent risks of Incidence and Mortality of COVID- 19 in Hubei Province and Other Parts of China Hongdou. Front Med 7:190

Xiu D (2010) Numerical methods for stochastic computations: a spectral methods approach. Princeton University Press, Princeton

Zhang S, Diao M, Yu W, Pei L, Lin Z, Chen D (2020) Estimation of the reproductive number of novel coronavirus (COVID-19) and the probable outbreak size on the Diamond Princess cruise ship: A data-driven analysis. Int J Infect Dis 93:201–204