Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Mathematical Biosciences 340 (2021) 108671

Contents lists available at ScienceDirect

Mathematical Biosciences

journal homepage: www.elsevier.com/locate/mbs

Original Research Article

Dynamics of partially mitigated multi-phasic epidemics at low susceptible depletion: phases of COVID-19 control in Italy as case study

Alberto d'Onofrio [a], Piero Manfredi [b], Mimmo Iannelli [c,*]

[a] 37 Quai du Docteur Gailleton, Lyon, France
[b] Department of Economics and Management, University of Pisa, Italy
[c] Mathematics Department, University of Trento, Italy

ARTICLE INFO

MSC:
92D25
92D30

Keywords:
Multi-phasic age-structured epidemics
Non-pharmaceutical interventions
Social distancing
Stable age distributions
Linear-chain trick
COVID-19

ABSTRACT

To mitigate the harmful effects of the COVID-19 pandemic, world countries have resorted – though with different timing and intensities – to a range of interventions. These interventions and their relaxation have shaped the epidemic into a multi-phase form, namely an early invasion phase often followed by a lockdown phase, whose unlocking triggered a second epidemic wave, and so on. In this article, we provide a kinematic description of an epidemic whose time course is subdivided by mitigation interventions into a sequence of phases, on the assumption that interventions are effective enough to prevent the susceptible proportion to largely depart from 100% (or from any other relevant level). By applying this hypothesis to a general SIR epidemic model with age-since-infection and piece-wise constant contact and recovery rates, we supply a unified treatment of this multi-phase epidemic showing how the different phases unfold over time. Subsequently, by exploiting a wide class of infectiousness and recovery kernels allowing reducibility (either to ordinary or delayed differential equations), we investigate in depth a low-dimensional case allowing a non-trivial full analytical treatment also of the transient dynamics connecting the different phases of the epidemic. Finally, we illustrate our theoretical results by a fit to the overall Italian COVID-19 epidemic since March 2020 till February 2021 i.e., before the mass vaccination campaign. This show the abilities of the proposed model in effectively describing the entire course of an observed multi-phasic epidemic with a minimal set of data and parameters, and in providing useful insight on a number of aspects including e.g., the inertial phenomena surrounding the switch between different phases.

1. Introduction

Since its detection in the city of Wuhan, China, by end December 2019, the novel SARS-CoV-2 coronavirus has rapidly spread worldwide with a dramatic direct disease burden (WHO situation reports 1–209 and Weekly epidemiological updates, [1]), and a dramatic impact on the economy, health and the society as a whole. During the first COVID-19 pandemic wave, an increasing number of countries worldwide have opted, following China, for massive social distancing measures — what is now universally identified as *lockdown*, integrated with a range of supplementary measures such as tracing and isolation of confirmed cases. The aim of lockdown is essentially that of abruptly bringing the basic reproduction number of the infection, $\mathcal{R}_0$, below threshold, therefore stopping sustained transmission. If this actually occurs, and the measures are not prematurely interrupted, the epidemic will be brought on a *suppression* path (Ferguson et al. [2]). The China initial experience, especially the one of the Hubei province, has shown that a combination of severity and appropriate duration of intervention can achieve high degrees of suppression even for epidemics that have reached a non-negligible scale. A somewhat different story occurred in Western Europe, where in many cases the lockdown was declared with substantial delays, with the aggravating circumstance that its implementation was done through several steps instead than abruptly, therefore yielding much larger epidemics than in China (see [1]). For example in Italy, whose Northern regions (especially Lombardia) experienced one of the most dramatic local epidemic worldwide, full lockdown was achieved by a long number of subsequent decrees of the Prime Minister covering a span of more than twenty days. These delays not only caused the epidemic to overwhelm available public health resources, with huge human costs, but also made eventually unavoidable to unlock long before the achievement of adequate suppression levels, due to the dramatically growing economic loss. Common to several European countries, the subsequent unlocking phase have shown prolonged *honey-moon* epochs of low epidemic activity, possibly resulting from a combination of reasons such as e.g., an enduring the maintenance of prudent behavior, temperature effects and school holidays.

* Corresponding author.
E-mail addresses: adonofrio1967@gmail.com (A. d'Onofrio), manfredi@ec.unipi.it (P. Manfredi), iannelli.mimmo@gmail.com (M. Iannelli).

https://doi.org/10.1016/j.mbs.2021.108671

Available online 21 July 2021

However, the unavoidable restart of epidemic activity again found the country unprepared, generating a massive second wave [1]. The ensuing mitigation interventions, milder than the generalized lockdown adopted during the spring wave possibly on trust of the announced start of the vaccination campaigns, were partly successful i.e., the epidemic curve was just brought to stationarity with the current reproduction index $\mathcal{R}_t$ in the region of one at the national level, until the new variants of concern initiated their replacement of the original ones [1], giving rise to a third wave [3], and forcing new restrictions, including regional lockdown. A number of key questions were continuously debated by the public during each pandemic wave about the effects of the enacted interventions such as e.g., (i) when the effects of social distancing measures will be detectable in the data, or when the epidemic peak of that particular phase will be achieved, (ii) which the further epidemic growth will be, in terms of hospitalizations, and deaths, after the measures implementation, etc.

Previous questions have represented major concerns throughout the COVID-19 pandemic and were responsible of widespread anxiety in the public opinion. They therefore call for general answers, even more so in the current moment when a third wave has started and it is clear that the benefits of vaccination – at current rates of administration – will take time to manifest. And, obviously, in view of future pandemics as they are not anymore classifiable as *rare events*. The COVID-19 pandemic has called for an unprecedented interest of modelers for epidemic dynamics both *per-se*, to investigate the direct epidemiological consequences of the pandemic and related mitigation measures, and also to investigate its broader implications on society. This has led to an endless number of papers and preprints of which we quote here only a few focusing on COVID-19 epidemiological modeling [2,4–21]. Most of the cited investigations have resorted to simulation. However, even under fairly general conditions, many of these phenomena can be investigated in a fully analytical manner, therefore providing a range of further robust insights about epidemic control.

Therefore, motivated by the main feature of the COVID-19 epidemic i.e., being characterized by a sequence of phases driven by the intensity and duration of the undertaken mitigation measures, in this paper we attempt to analyze, in an analytical manner, the dynamics of a generic multi-phasic epidemic, much in the same way as we observed for COVID-19 prior the current mass vaccination campaign started to deploy its main effects. These phases include an early invasion phase, the subsequent mitigation interventions (e.g., generalized lockdown), the unlocking/releasing phase, the second wave phase, the further mitigating interventions etc. To make the problem tractable, we assumed that the interventions considered are sufficiently long-lasting and effective so to prevent that the susceptible proportion of the population excessively departed from its initial value (presumably around 100% for COVID-19) or from any other reference value. This hypothesis can be always made valid by considering suitable observation windows. Besides the rapid control achieved in China already by February 2020 at very low attack rates, in Italy a nation-wide large-scale serological survey conducted during after the first wave (June–July 2020) showed that — with the exception of Lombardy which had entire areas ravaged by the epidemic (eventually bringing attack rates around 7.5%), the national average was only slightly in excess of 2%, with most regions only mildly attacked [22]. And still at the onset of the third wave (end February 2021) the overall nation-wide proportions of confirmed infections was in the range of 5%. This is strongly suggestive of the fact that the potential impact on the course of the epidemics, due to the depletion of the susceptible population, would always be secondary with respect to the impact due to the undertaken mitigation measures. More in general we expect this to be largely true in all countries and subnational settings where social distancing arrived at a sufficiently earlier stage of the epidemic so to halt it before it depleted the susceptible proportion. Note additionally that this approach is valid for any combination of epidemic phases where interventions are able to approximately "freeze" the susceptible fraction around some given level.

In these circumstances, the analysis of a multi-phasic epidemic yields to a linear problem which can be analyzed by the tools of age-structured populations mathematics ([23–25] and references therein). Our principal aim here will be to provide a unified framework to multi-phase epidemics while keeping our results as more analytical as possible. Our approach will be *kinematic* in the sense that will look at how the different epidemic phases unfold over time reshaping the epidemic, without minding at whether, for example, a certain value of the current reproduction number R_t achieved during e.g., a lockdown phase, was the mere outcome of government-enforced restrictions or was also reflecting the beneficial (or harmful) effect of virtuous (resistant) behavior spread in the population, as would instead be the approach of behavioral epidemiologists [26,27]. Consistently, we will simply consider a SIR epidemic model including *class age* i.e., age since infection, in the infective compartment and piece-wise constant transmission and recovery rates over time reflecting different *intervention epochs*. The reason why we use such a simple model (e.g., without including more compartments as symptomatic cases, hospitalization, deaths etc as a faithful description of COVID-19 would require), is that this simple model provides the backbone description of a general multi-phasic epidemic. This backbone can be easily refined by adding further compartments representing delayed events following infection [28].

Despite its simplicity, our SIR model is general in that we consider both a latency delay and general forms for the infectiousness and recovery processes. By the proposed model, we first provide a general characterization of a multi-phasic epidemic. Second, by adopting a wide class of generation time and recovery kernels allowing reducibility to ordinary or delayed differential equations [29], we were able to provide a detailed investigation of a non-trivial low-dimensional case allowing a full analytical treatment of the transient dynamics of a multi-phase epidemic. Finally, we complemented our theoretical results by a number of illustrations, using realistic parameter constellations drawn from available estimates of COVID parameters. In particular, we fitted the model to the overall Italian COVID-19 epidemic since March 2020 until late February 2021. This illustrates the abilities of the proposed model in effectively explaining the entire course of the epidemic, including the effects of interventions, using a minimal set of data and parameters, and in providing useful insight on a number of aspects including e.g., the demographic inertia occurring during the switch between different epidemic phases.

The article is organized as follows. Section 2 develops our model for a multi-phasic epidemic. Section 3 reports the main theoretical results. Section 4 presents the proposed class of reducible kernels and reports a number of sub-cases and illustrations. Section 5 reports the application to Italian data. Concluding remarks follow.

2. A backbone model for a multi-phase epidemic

The model we present here is a SIR epidemic model structured by age since infection or *class age*. The evolution of the epidemics over time is subdivided into phases, each one characterized by a certain intervention intensity.

2.1. Model specification

Our model is based on the state variable

$$Y(x,t), \quad x \in [0, x_+],$$

where x denotes class age, representing the (improper) age-density of the number of infected individuals. In particular, x_+ denotes the maximal age since infection and may be finite or infinite. Our main assumption is that interventions are early and effective enough to prevent a large depletion of the susceptible population during the overall epidemic course i.e., the susceptible fraction does not significantly departs from 1. Note that this hypothesis can actually be applied at any level

S of the susceptible proportion, provided subsequent interventions are effective enough to "freeze" the susceptible population thereafter.

In this manner, the model equations read

$$\begin{cases} \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial x}\right) Y(x,t) = -\gamma(x,t)Y(x,t), \\ Y(0,t) = \int_0^{x_+} \beta(x,t)Y(x,t)dx, \\ Y(x,0) = Y^0(x). \end{cases} \tag{1}$$

where

- $\gamma(x,t)$ = *the removal rate of infected individuals aged* x *at time* t, that is, the overall rate at which infected individuals aged x are removed by any cause as, e.g., recovery or death (during an uncontrolled epidemic) or by screening, tracing, isolation, hospitalization etc (in the presence of interventions).
- $\beta(x,t)$ = *the transmission rate of infected individuals aged* x *at time* t, representing the number of secondary infections caused by a single infected aged x per unit of time in the hypothetical situation of no removal.

Note that, thanks to the stated hypotheses, the incidence of new infections

$$U(t) = \int_0^{x_+} \beta(x,t)Y(x,t)dx$$

depends linearly on the infected density $Y(x,t)$. Both rates $\beta(t,x), \gamma(x,t)$ are taken as functions of time as well, to reflect the possibility of interventions (e.g., social distancing, testing, isolating), or the removal of such measures. To mimic the various phases of the epidemic experienced by most European countries we will represent these rates as piecewise function of time. Just for illustration, to represent a 3-phase epidemic, as it might be the case for an epidemic with an initial outbreak phase, a lockdown phase, and a subsequent release phase, we will take

$$\gamma(x,t) = \mathcal{I}_{[0,t_L)}(t)\ \gamma_1(x) + \mathcal{I}_{[t_L,t_R)}(t)\ \gamma_2(x) + \mathcal{I}_{[t_R,+\infty)}(t)\ \gamma_3(x) \tag{2}$$

and

$$\beta(x,t) = c(t)\beta_0(x) = \left[\mathcal{I}_{[0,t_L)}(t)\ c_1 + \mathcal{I}_{[t_L,t_R)}(t)\ c_2 + \mathcal{I}_{[t_R,+\infty)}(t)\ c_3\right]\beta_0(x), \tag{3}$$

where $\mathcal{I}_{[a,b)}(t)$ represents the characteristic function of the interval $[a,b)$, while t_L and t_R, $0 < t_L < t_R$, are the lockdown and unlocking times respectively. Of course, it is possible to consider any number of epidemic phases. Specification (2) models interventions aiming to affect the age-density of removal in the broad sense defined above. In (3), factor $c(t)$ denotes the number of adequate contacts per person and per unit time, while $\beta_0(x)$ tunes the (average) intrinsic infectiousness of an infected individual aged x (i.e, the probability that an infected individual aged x infects a susceptible during an adequate contact), possibly related to her/his viral load. Our formulation therefore assumes that intervention strategies may act either on the number of adequate contacts (as is the case for social distancing) or the scale of transmission (e.g., by using masks), but not on the shape of the distribution of infectiousness by age since infection. Note moreover that the previous form describes abrupt switches between phases, reflecting an ideal world where compliant agents suddenly adjust their behavior upon governmental intervention. However, extension to smooth transitions over time is possible. Overall, the proposed model represents a non-autonomous, infinite-dimensional, dynamical system with piecewise constant *switching* parameters (see for instance [30,31]).

2.2. Stepwise solution of the model

Within the previous framework, the model can be solved in a stepwise manner, phase by phase. For the sake of simplicity we limit the discussion to three "representative" phases, namely an invasion (or outbreak) phase, a lockdown phase and a release one (respectively labeled by $i = 1$, $i = 2$, $i = 3$), but the analysis can continue to include successive phases.

We first note that under our hypotheses each phase of the epidemics ($i = 1,2,3$) is characterized by a phase-specific *reproduction number* given by

$$\mathcal{R}_0^i = c_i \int_0^{x_+} \beta_0(x)\Gamma_i(x)dx, \quad (i = 1,2,3)$$

where

$$\Gamma_i(x) = e^{-\int_0^x \gamma_i(s)ds}, \quad x \in [0, x_+],$$

is the phase-specific *survival-to-removal probabilities*, defining the probability that a newly infected individual is still infective by age x conditional on the interventions undertaken during any given phase. Concerning the reproduction numbers, we will make the assumptions (consistent with the experience of countries which resorted to lockdown) $\mathcal{R}_0^1 > 1$, $\mathcal{R}_0^2 < 1$, $\mathcal{R}_0^3 > 1$ (but other hypotheses can obviously be made). For each phase ($i = 1,2,3$) we have the Lotka–Von Foerster problem

$$\begin{cases} \left(\frac{\partial}{\partial t}+\frac{\partial}{\partial x}\right) Y_i(x,t) = -\gamma_i(x)Y_i(x,t), \\ Y_i(0,t) = c_i \int_0^{x_+} \beta_0(x)Y_i(x,t)dx, \\ Y_i(x,0) = Y_i^0(x), \end{cases} \tag{4}$$

where the initial age distribution of the next phase is the terminal distribution of the previous one

$$\begin{aligned} Y_1^0(x) &= Y^0(x), \\ Y_2^0(x) &= Y_1(t_L,x), \\ Y_3^0(x) &= Y_2(t_R - t_L, x), \end{aligned}$$

and the solution to (1) reads

$$Y(t,x) = \begin{cases} Y_1(x,t) & \text{for } t \in [0,t_L], \\ Y_2(x,t-t_L) & \text{for } t \in [t_L,t_R], \\ Y_3(x,t-t_R) & \text{for } t \in [t_R,\infty]. \end{cases}$$

By integration along characteristic lines (see for instance [24]), each system (4) can be transformed into an integral equation in the phase-specific incidence $U_i(t)$ ($i = 1,2,3$)

$$U_i(t) = c_i \int_0^{x_+} \beta_0(x)Y_i(x,t)dx.$$

Namely we obtain the phase-specific renewal equation

$$U_i(t) = \mathcal{R}_0^i \left(G_i(t) + \int_0^t K_i(x)U_i(t-x)dx\right), \tag{5}$$

where

$$K_i(x) = \frac{\beta_0(x)\Gamma_i(x)}{\int_0^\infty \beta_0(x)\Gamma_i(x)dx}$$

is a normalized kernel representing the *generation time* density, and

$$G_i(t) = \int_t^\infty K_i(x)\frac{Y_i^0(x-t)}{\Gamma_i(x-t)}dx,$$

where all the function involved are extended as zero over the interval of definition $[0, x_+]$.

Once Eq. (5) is solved, the solution to (4) is then given by

$$Y_i(x,t) = \begin{cases} Y_i^0(x-t)\frac{\Gamma_i(x)}{\Gamma_i(x-t)} & \text{for } t \le x \\ U_i(t-x)\Gamma_i(x) & \text{for } t > x. \end{cases} \tag{6}$$

Note that in the case of a finite maximum age $x_+ < +\infty$ we eventually have (for $t > x_+$)

$$Y_i(x,t) = U_i(t-x)\Gamma_i(x).$$

On the other hand, if the maximum age is not finite, which is an hypothesis common in epidemiological modeling where generation time densities are typically represented by infinitely-supported parametric distributions (as e.g., Gamma or Weibull), the age-density of infectives $Y_i(x,t)$ will continue to include, at any time, its initial distribution even if at a steadily declining "weight".

2.3. Analysis of single phases

Our analysis of the epidemic dynamics during each phase will be based on the renewal equation (5) that is well known in demography and epidemiology. The basic theory (briefly outlined in Appendix A) states that the behavior of $U_i(t)$ is related to the roots of the corresponding characteristic equation

$$\mathcal{R}_0^i \hat{K}_i(\lambda) = 1. \tag{7}$$

Under generic hypotheses on kernel $K_i(x)$ that we assume are satisfied, this equation has one dominant real root α_i^* and the other roots can be ordered in a sequence $\{\alpha_i^j\}$ such that

$$\Re\alpha_i^{j+1} \le \Re\alpha_i^j < \alpha_i^* \qquad (j = 1,\ 2,\ \dots), \tag{8}$$

and the solution for the incidence during the phase i has the asymptotic expansion

$$U_i(t) = U_i^* e^{\alpha_i^* t} + \sum_{j=1}^{\infty} U_i^j e^{\alpha_i^j t}.$$

where the constants of the expansion respectively obey

$$U_i^* = \frac{\hat{G}_i(\alpha_i^*)}{\int_0^\infty x K_i(x) e^{-\alpha_i^* x} dx}. \tag{9}$$

and

$$U_i^j = \frac{\hat{G}_i(\alpha_i^j)}{\int_0^\infty x K_i(x) e^{-\alpha_i^j x} dx}.$$

Since (8) holds, we have

$$\lim_{t\to\infty} e^{-\alpha_i^* t}\ U_i(t) = U_i^*. \tag{10}$$

On the further assumption that the temporal duration of each phase is sufficient to allow the emergence of long-term behavior, the ultimate behavior of the incidence during the ith phase is therefore exponential

$$U_i(t) \approx U_i^* e^{\alpha_i^* t},$$

at a Lotka's *intrinsic* rate α_i^* and scale constant U_i^*, also termed the *stable equivalent* in demographic jargon (see [23]). The requirement that intervention phases are sufficiently long is by no means unrealistic: during invasion phases convergence to the exponential epidemic path is usually very fast, while for e.g., a lockdown phase aiming to bring reproduction below threshold, it would purely be irrational to stop the intervention before the epidemic has entered its exponentially declining path.

The *asymptotic* behavior of incidence, as stated in (10), in turn implies the emergence of the *stable age distribution* (i.e., a stable age profile) of the infected population:

$$Y_i^*(x) = \frac{e^{-\alpha_i^* x}\Gamma_i(x)}{\int_0^\infty e^{-\alpha_i^* x}\Gamma_i(x)dx}, \quad x \in [0, x_+], \tag{11}$$

where, if $x_+ < +\infty$, in the denominator we intend that the function Γ_i is extended as 0 for $x > x_+$ (indeed, the stable distribution arises if the function $\Gamma_i(x)$ is Laplace transformable at α_i^*).

Actually, based on (10), the solution of (4) [see (6)] will approach the asymptotic form

$$Y_i^\infty(x,t) \approx U_i^*\ \hat{\Gamma}_i(\alpha_i^*)\ e^{\alpha_i^* t}\ Y_i^*(x).$$

Precisely, we have [see (6)]

$$\lim_{t\to\infty} e^{-\alpha_i^* t}\ Y_i(x,t) = U_i^*\ \hat{\Gamma}_i(\alpha_i^*)\ Y_i^*(x) \tag{12}$$

where the convergence occurs pointwise in x.

2.4. Components of the stable equivalents

We briefly comment about the components of the stable equivalent U_1^* of COVID-19 dynamics during the exponential phase. The denominator in (9) reads

$$\int_0^\infty x K_i(x) e^{-\alpha_i^* x} dx = \frac{\mu_i}{\mathcal{R}_0^i}.$$

where

$$\begin{aligned}\mu_i &= \mathcal{R}_0^i \int_0^\infty x K_i(x) e^{-\alpha_i^* x} dx \\ &= \int_0^\infty x \frac{\beta_0(x)\Gamma_i(x)e^{-\alpha_i^* x}}{\int_0^\infty \beta_0(s)\Gamma_i(s)e^{-\alpha_i^* s}ds} dx = \int_0^\infty x \frac{\beta_0(x)Y_i^*(x)}{\int_0^\infty \beta_0(s)Y_i^*(s)ds} dx\end{aligned}$$

is the *average age at which infected individual produce their secondary cases*, at the stable age distribution prevailing during invasion. Therefore, the stable equivalent can be represented as

$$U_i^* = \frac{\mathcal{R}_0^i \hat{G}_i(\alpha_i^*)}{\mu_i}, \tag{13}$$

Notably, at the numerator of the expression we have the product

$$\mathcal{R}_0^i \hat{G}_i(\alpha_i^*) = \mathcal{R}_0^i \int_0^\infty e^{-\alpha_i^* t} \int_t^\infty K_i(x) \frac{Y_i^0(x-t)}{\Gamma_i(x-t)} dx dt = \int_0^\infty \Pi_i(x) Y_i^0(x) dx. \tag{14}$$

where

$$\Pi_i(x) = c_i \int_x^\infty e^{-\alpha_i^*(s-x)} \beta_0(s) \frac{\Gamma_i(s)}{\Gamma_i(x)} ds \tag{15}$$

represents, still resorting to demographic jargon, Fisher's *reproductive value of an infected aged* x [23]. Therefore, the numerator in (13), which scales the magnitude of the epidemic, also has a noteworthy demographic interpretation, i.e., it represents the *total reproductive value* i.e., the infection potential embedded in the initial age distribution of infected individuals.

3. The dynamics of the different phases

This section summarizes the main model results in distinct subsections focusing on the distinct phases ($i = 1, 2, 3$) of the epidemic using Eqs. (5). During each phase, the infected distribution evolves toward its asymptotic *stable age distribution* and provides the new initial age distribution for the subsequent phase. Thus each phase is ultimately represented by the dominant real root α_i^* of the characteristic equation (7) and by the *stable equivalent* U_i^*.

We discuss in greater detail the invasion phase ($i = 1$), where no intervention is in place, and the subsequent social distancing phase ($i = 2$). Finally we briefly extend the results to the release phase ($i = 3$).

3.1. The outbreak phase

We first consider the initial phase of the epidemic, that is the solution of (1) in the interval $[0, t_L]$ where no control measures are in place yet. This is governed by the solution of (4) for $i = 1$ with $\mathcal{R}_0^1 > 1$, so that the dominant real root α_1^* is strictly positive. Thus, whatever be the initial distribution $Y^0(x)$, the infected population will reach its *stable age distribution* [(11)]

$$Y_1^*(x) = \frac{e^{-\alpha_1^* x}\Gamma_1(x)}{\hat{\Gamma}_1(\alpha_1^*)}, \quad x \in [0, x_+],$$

Note that, since $\alpha_1^* > 0$, the function $\Gamma_1(x)$, being bounded, is Laplace transformable at α_1^*.

As known from classical mathematical demography, the stable age distribution (SAD), which here represents the *relative* distribution of the infected by (class) age, is time invariant due to the fact that the absolute numbers of infective individuals in each age group grow at the same rate α_1^* over time. In particular, the normalizing term of the SAD, given by the reciprocal of the Laplace transform of the *survival to removal* function $\widehat{\Gamma}_1(\alpha_1^*)$ has a clear epidemiological meaning. It indeed represents the long-term per-capita birth rate of new infective individuals – generically defined as the ratio between the rate of new births, i.e., the incidence $U_i(t)$, and the overall size of the infected population $I(t)$ – in the asymptotic regime of exponential growth of incidence defined in (10).

In case the age profile of the initial datum corresponds to the SAD (11), namely if

$$Y^0(x) = I(0)\, Y_1^*(x), \tag{16}$$

where $I(0)$ is the number of infected individuals at the beginning of the epidemic, then

$$G_1(t) = \frac{I(0)}{\widehat{\Gamma}_1(\alpha_1^*)}\, e^{\alpha_1^* t} \int_t^\infty K_1(x) e^{-\alpha_1^* x} dx$$

and the solution to (5) is exactly (see Appendix A)

$$U_1(t) = \frac{I(0)}{\widehat{\Gamma}_1(\alpha_1^*)}\, e^{\alpha_1^* t}. \tag{17}$$

From (17), the corresponding solution to (4) reads

$$Y_1(x,t) = I(0)\, e^{\alpha_1^* t}\, Y_1^*(x),$$

i.e., the age-distribution is stable at all times. The latter developments state that if the infective population lies on its SAD already at the initial phase of the epidemic, then the dynamics of incidence is simply given by the exponential propagation at the intrinsic rate α_1^* of the total initial infective population $I(0)$ through its stable (per-capita) birth rate $1/\widehat{\Gamma}_1(\alpha_1^*)$.

Concerning the behavior of the total number of infected during the outbreak phase

$$I_1(t) = \int_0^{x_+} Y_1(x,t) dx,$$

from (6) and (12) we see that it splits into two terms

$$\begin{aligned} I_1(t) &= \int_t^\infty Y^0(x-t) \frac{\Gamma_1(x)}{\Gamma_1(x-t)} dx + \int_0^t U_1(t-x)\Gamma_1(x) dx \\ &= I_1^0(t) + I_1^+(t) \end{aligned}$$

with

$$\lim_{t\to\infty} I_1^0(t) = 0, \quad \lim_{t\to\infty} e^{-\alpha_1^* t}\, I_1^+(t) = U_1^*\, \widehat{\Gamma}_1(\alpha_1^*). \tag{18}$$

Thus, the first term is a transient one accounting for the initial group of infected people, while the second term accounts for the epidemic dynamics resulting from new cases generated for $t > 0$ which for large times converges to its stable form. Precisely, the initial group $I_1^0(t)$ decays at a rate related to the removal probability $\Gamma_1(x)$. In fact, if $x_+ < +\infty$ we simply have

$$I_1^0(t) = 0 \quad \text{for} \quad t > x_+,$$

otherwise, if $x_+ = +\infty$, for any $0 < \omega < -\sigma_{\Gamma_1}$, where σ_{Γ_1} is the abscissa of convergence of the Laplace transform of $\Gamma_1(x)$, we have

$$\lim_{t\to\infty} e^{\omega t} I_1^0(t) = \lim_{t\to\infty} \int_0^\infty Y^0(x) \frac{e^{\omega(x+t)}\Gamma_1(x+t)}{e^{\omega x}\Gamma_1(x)} dx = 0. \tag{19}$$

Note that the ultimate behavior of the total number of infected during the first phase

$$I_1^+(t) \approx U_1^*\, \widehat{\Gamma}_1(\alpha_1^*)\, e^{\alpha_1^* t} \tag{20}$$

differs from incidence only in the scale constant, which is now given by $U_1^*\, \widehat{\Gamma}_1(\alpha_1^*)$. The latter represents the stable equivalent of the total infected population, which differs from that of incidence by a multiplication for the reciprocal of the asymptotic birth rate of new infectives, $\widehat{\Gamma}_1(\alpha_1^*)$.

Finally, as regards the lapse of time necessary for observing the emergence of the asymptotic behavior of the solution, we can estimate the approximate minimum time for a tolerance ϵ as [see (A.4)]

$$T_\epsilon = \frac{\ln(|U_1^*|) - \ln(\epsilon)}{\alpha_1^* - \Re\alpha_1^1}.$$

3.2. The lockdown phase

Let us now consider the implications of a sufficiently long lasting mitigation phase starting at the time t_L, when the reproduction number is abruptly set at $\mathcal{R}_0^2 < 1$. Since in this case we have $\alpha_2^* < 0$, if social distancing is maintained for a sufficiently long time, the epidemic will eventually set on an exponentially declining path that we will term a *suppression path*. The suppression path would eventually bring to epidemic extinction if the lockdown phase would continue rather than being halted by unlocking. However, before setting on the suppression path, the lockdown dynamics will be somewhat articulated as it will have to connect the pattern of fast exponential epidemic growth inherited from the first phase with its regime phase of exponential decline. This suggests a transient phase resulting from the abrupt switch between dynamic regimes, which will be dominated by the inertia inherited from the fast growth of the first phase, and might result in an epidemic peak, after which epidemic decline will start. During this phase the age distribution of infective individuals will continuously adjust, and will eventually converge to its new stable limiting form $Y_2^*(x)$ that will promote the exponential decline of the suppression phase. In what follows, we discuss the second-phase dynamics by distinguishing between the limit case of epidemic *annihilation* ($\mathcal{R}_0^2 = 0$) from the standard case of (eventual) elimination $0 < \mathcal{R}_0^2 < 1$.

3.2.1. Epidemic annihilation

In this case, setting $\mathcal{R}_0^2 = 0$, we have $U_2(t) \equiv 0$ and

$$Y_2(x,t) = \begin{cases} Y_2^0(x-t)\dfrac{\Gamma_2(x)}{\Gamma_2(x-t)} & \text{for} \quad t \le x \\ 0 & \text{for} \quad t > x. \end{cases} \tag{21}$$

Thus, for the prevalence of infected we have only the transient component $I_2^0(t)$ that either is identically 0 for $t > x_+$ (if $x_+ < +\infty$) or goes extinct at any rate ω such that $0 < \omega < -\sigma_{\Gamma_2}$ [see the analogous (19)].

3.2.2. The case $0 < \mathcal{R}_0^2 < 1$

If $0 < \mathcal{R}_0^2 < 1$, then there is also the component $I_2^+(t)$ and we have

$$\lim_{t\to\infty} e^{-\alpha_2^* t} Y_2(t,x) = U_2^* e^{-\alpha_2^* x} \Gamma_2(x) \tag{22}$$

pointwise in x. Note that, in the case $x_+ < +\infty$, we certainly have

$$I_2^+(t) \approx U_2^* \widehat{\Gamma}_2(\alpha_2^*) e^{\alpha_2^* t}, \tag{23}$$

but, if $x_+ = +\infty$, since $\alpha_2^* < 0$, the limit on the right hand side of (22) is integrable if and only if $\alpha_2^* > \sigma_{\Gamma_2}$. In this case (23) holds true [compare with (20)]. However if $\alpha_2^* < \sigma_{\Gamma_2}$ we instead have

$$\begin{aligned} &\lim_{t\to\infty} e^{\omega t} I_2^+(t) = \lim_{t\to\infty} \int_0^t U_2(t-x) e^{\omega(t-x)} e^{\omega x} \Gamma_2(x) dx \\ &\le \lim_{t\to\infty} \int_0^t U_2(t-x) e^{-\alpha_2^*(t-x)} e^{\omega x} \Gamma_2(x) dx = U_2^* \widehat{\Gamma}_2(-\omega) \end{aligned}$$

for any ω such that $0 < \omega < -\alpha_2^* < -\sigma_{\Gamma_2}$, i.e., $I_2^+(t)$ decays at the rate $-\omega$.

3.2.3. Stable equivalents of a lockdown phase

Much information on epidemic trend during the lockdown phase is provided by the two parameters governing the resulting long-term declining exponential path, namely the intrinsic rate α_2^* and the stable equivalent, where the latter tunes the scale of the epidemic curve over the suppression path. We assume that at the lockdown time t_L, the epidemic has reached its stable form (12), so that the initial datum at t_L is

$$Y_2^0(x) = U_1^*\ \widehat{\Gamma}_1(\alpha_1^*)\ e^{\alpha_1^* t_L}\ Y_1^*(x) = U_1^*\ e^{\alpha_1^*(t_L - x)}\Gamma_1(x),$$

implying a total number of infected individuals

$$I_2(0) = U_1^*\ \widehat{\Gamma}_1(\alpha_1^*)\ e^{\alpha_1^* t_L}.$$

Here U_1^* is in general given by (13) while, if the initial population of infectious individuals at the beginning of the epidemic was distributed according to the stable distribution of the first phase, then [see (17)]

$$U_1^* = \frac{I(0)}{\widehat{\Gamma}_1(\alpha_1^*)}.$$

The stable equivalent of the incidence function for the lockdown phase is [see (13) and (14)]

$$U_2^* = \frac{I(0)\ e^{\alpha_1^* t_L}}{\mu_2}\int_0^\infty \Pi_2(x)\ Y_1^*(x)dx,$$

In the particular case of $\Gamma_1(x) \equiv \Gamma_2(x)$, i.e. if the lockdown measures only involve the lowering of the contact rate from its free-epidemic level c_1 to $c_2 < c_1$, the latter quantity reads

$$U_2^* = \frac{U_1^* e^{\alpha_1^* t_L}(\mathcal{R}_0^1 - \mathcal{R}_0^2)}{\mu_2(\alpha_1^* - \alpha_2^*)\mathcal{R}_0^1}.$$

The previous quantity encapsulates the inertial effects arising from the transition between two regimes of stable growth, namely an initial regime of exponential epidemic growth and the subsequent suppression regime characterized by asymptotic exponential decline.

3.3. The release of social distancing

For simplicity we assume that the release of social distancing measures starts when the stable age distribution of the suppression phase was fully established so that we have to consider the incidence equation

$$U_3(t) = \mathcal{R}_0^3\left(G_3(t) + \int_0^t K_3(x)U_3(t-x)dx\right)$$

with the initial datum

$$Y_3^0(x) = U_2^*\ e^{\alpha_2^*(t_R - t_L)}e^{-\alpha_2^* x}\Gamma_2(x).$$

The analysis proceeds in the same way as the first two phases ending in a long-term regime of exponential growth ($\mathcal{R}_0^3 > 1$) or decay ($\mathcal{R}_0^3 < 1$). In particular the corresponding stable equivalent

$$U_3^* = \frac{\mathcal{R}_0^3 \widehat{G}_3(\alpha_3^*)}{\mu_3}.$$

will embed the inertial effects due to the aging experienced by the age distribution of infection inherited from the suppression dynamics of the lockdown phase, on the unlocking dynamics.

4. Illustrations and simulations

To better appreciate the general discussion above, it is worth to consider specific parametrizations allowing a more detailed description of the epidemic phases as well as comparisons with data. Moreover, by a parametrized description of the process, we can exploit *Erlang* kernels allowing the reduction of infinite-dimension systems to systems of ordinary differential equations (ODEs) or to mixed systems of delayed and ordinary differential equations (DDE–ODE since now on). This is a classical reduction procedure for integral equations of Volterra type, known as the *linear chain trick* [29], that can be directly applied to our systems (4). We recall that the compartmental approximation through sequences of ODEs with constant transition rates represent the easiest way to reliably simulate the PDE systems studied here.

4.1. Reducible kernels

Our general model depends on two key age-specific epidemiological functions i.e., the (phase-specific) survival-to-infection function $\Gamma_i(x)$, which tunes the age distribution of removal, and the infection reproduction kernel $K_i(x)$, which is the normalized product of the infectivity kernel $\beta_0(x)$ and $\Gamma_i(x)$, tuning the generation time (age-) distribution. A wide class of reducible problems appears if both distributions are translated Erlang densities, that is Gamma densities with integer index, i.e.,

(i) the infectivity kernel $\beta_0(x)$ is a (non-proper) translated Erlang density of order n and rate φ:

$$\beta_0(x) = \beta_0\frac{\varphi^n(x-\tau)^{n-1}e^{-\varphi(x-\tau)}}{(n-1)!}\ \mathcal{I}_{[\tau,+\infty)}(x); \tag{24}$$

(ii) the removal rate $\gamma_i(x)$ is taken as

$$\gamma_i(x) = \frac{\gamma_i^m(x-\tau)^{m-1}}{(m-1)!\sum_{j=0}^{m-1}\frac{\gamma_i^j(x-\tau)^j}{j!}}\ \mathcal{I}_{[\tau,+\infty)}(x), \tag{25}$$

corresponding to the survival function of a translated Erlang density of order m and rate γ_i, that is:

$$\Gamma_i(x) = \mathcal{I}_{[0,\tau)}(x) + \sum_{j=0}^{m-1}\frac{\gamma_i^j(x-\tau)^j e^{-\gamma_i(x-\tau)}}{j!}\ \mathcal{I}_{[\tau,+\infty)}(x), \tag{26}$$

Under previous specifications, the phase-specific reproduction numbers are given by

$$\mathcal{R}_0^i = \frac{c_i\beta_0\varphi^n}{\theta_i^n}\sum_{j=0}^{m-1}\binom{n+j-1}{j}\left(\frac{\gamma_i}{\theta_i}\right)^j$$

where

$$\theta_i = \varphi + \gamma_i.$$

Typical shapes of both $\beta_0(x)$ and $\gamma_i(x)$ functions, for different values of the Erlang index parameter, are reported in Fig. 1

Formulations (24) and (25) assume a latently infective phase having fixed duration $\tau, \tau \geq 0$) during which infected individuals are not yet infective and cannot be removed. In relation to COVID-19 data, we remark that although the Gamma kernels best fitting observed data of serial intervals (taken as surrogates of generation times) were obviously characterized by non-integer indices [2,32], therefore implying non-reducible kernels, nonetheless integer-indexed Gamma (i.e., Erlang-type) distributions can usefully bound observed distributions.

As we now show by means of examples, if duration τ is strictly positive, the linear chain trick will lead to a mixed DDE–ODE system which, in the limit case $\tau = 0$, will collapse into a pure ODE system.

4.2. Reduction to a 2-dimensional DDE-ODE system

To illustrate how the chosen kernels allow reduction (for each phase), we discuss the case $n = 2$, $m = 1$ yielding

$$\begin{aligned} K_i(x) &= \mathcal{I}_{[\tau,+\infty)}(x)\ \theta_i^2\ (x-\tau)\ e^{-\theta_i(x-\tau)},\\ \gamma_i(x) &= \gamma_i\ \mathcal{I}_{[\tau,+\infty)}(x),\\ \mathcal{R}_0^i &= \frac{c_i\beta_0\varphi^2}{\theta_i^2}, \end{aligned} \tag{27}$$

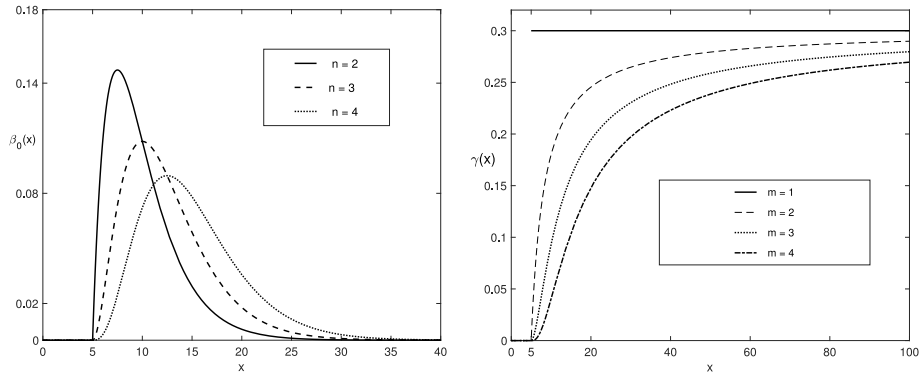

Fig. 1. Instances of the basic age-phase specific parameters. Left panel: specific infectivity rate $\beta_0(x)$, parametrized according to (24) for $\tau = 5$ day, $\beta_0 = 1$ day, $\varphi = 0.3$ day^{-1} and $n = 2, 3, 4$; note that increasing the order parameter n delays the growth of the function. Right panel: removal rate $\gamma_i(x)$, parametrized according to (25), for $\tau = 5$ day, $\gamma_i = 0.3$ day^{-1} and $m = 1, 2, 3, 4$; the shape of the rate is increasing after the end of the latency phase and converges for large ages to the constant level γ_i; also in this case, increasing the index smooths the function and delays its growth.

By considering the basic variables

$$\begin{aligned} U_i(t) &= \mathcal{R}_0^i \int_0^\infty K_i(x) y_i(x,t) dx, \\ J_i(t) &= \int_\tau^\infty e^{-\theta_i(x-\tau)} y_i(x,t) dx, \end{aligned} \tag{28}$$

where $y_i(x,t)$ is the scaled variable

$$y_i(x,t) = \frac{Y_i(x,t)}{\Gamma_i(x)}, \tag{29}$$

we obtain (see details in Appendix B) the following DDE–ODE system having as state variables the epidemic incidence $U_i(t)$ and the auxiliary variable $J_i(t)$:

$$\begin{aligned} \frac{d}{dt} U_i(t) &= -\theta_i\ U_i(t) + \mathcal{R}_0^i \theta_i^2 J_i(t) \\ \frac{d}{dt} J_i(t) &= U_i(t-\tau) - \theta_i J_i(t) \end{aligned} \tag{30}$$

Note that the variable $J_i(t)$, that appears upon differentiation of $U_i(t)$, contributes to determine the behavior of $U_i(t)$ that in turn contains all the relevant information. The variable $J_i(t)$ is an auxiliary quantity representing the hypothetical incidence of new cases of infection that would appear if the normalized infection kernel K_i would have the *diminished* index $n - 1$ instead of n. This confers to the resulting DDE–ODE system, the classical structure of *reduced* problems [29].

Once $U_i(t)$ is known, using $Y_i(x,t)$ from formula (6), other relevant epidemiological quantities can be computed such as

$$\begin{aligned} &prevalence\ of\ infected\ individuals: && I_i(t) = \int_0^\infty Y_i(x,t) dx, \\ &prevalence\ of\ infective\ individuals: && I_i^\#(t) = \int_\tau^\infty Y_i(x,t) dx, \\ &prevalence\ of\ exposed\ individuals && E_i(t) = \int_0^\tau Y_i(x,t) dx. \end{aligned}$$

For these variables we can also state additional equations that can be added to system (30). Indeed, proceeding as before we obtain

$$\begin{aligned} \frac{d}{dt} I_i^\#(t) &= U_i(t-\tau) - \gamma_i I_i^\#(t), \\ \frac{d}{dt} E_i(t) &= U_i(t) - U_i(t-\tau), \\ \frac{d}{dt} I_i(t) &= U_i(t) - \gamma_i I_i^\#(t). \end{aligned}$$

The previous equations explicitly show the underlying SEIR structure implicit in the present reduced model, which can indeed be described as a SEIR model with *fixed-duration* latency time τ.

If in particular $\tau = 0$, the delayed system (30) reduces to the 2-dimensional ODE system

$$\begin{aligned} \frac{d}{dt} U_i(t) &= -\theta_i\ U_i(t) + \mathcal{R}_0^i \theta_i^2 J_i(t) \\ \frac{d}{dt} J_i(t) &= U_i(t) - \theta_i J_i(t). \end{aligned} \tag{31}$$

and $I_i(t)$ (note that, since $\tau = 0$, it holds $I_i(t) = I_i^\#(t)$, landing on an SIR structure) satisfies the additional equation

$$\frac{d}{dt} I_i(t) = U_i(t) - \gamma_i I_i(t). \tag{32}$$

In this case, the model solution can be calculated explicitly through the roots of the characteristic equation [compare with (7)]

$$\left(\lambda + \theta_i\right)^2 = \mathcal{R}_0^i \theta_i^2,$$

yielding

$$\alpha_i^* = \theta_i \left(\sqrt{\mathcal{R}_0^i} - 1\right), \quad \alpha_i^1 = -\theta_i \left(\sqrt{\mathcal{R}_0^i} + 1\right)$$

and, for each phase

$$\begin{aligned} U_i(t) &= A_i^+ e^{\alpha_i^* t} + A_i^- e^{\alpha_i^1 t} \\ J_i(t) &= \frac{1}{\sqrt{\mathcal{R}_0^i}\theta_i} \left(A_i^+ e^{\alpha_i^* t} - A_i^- e^{\alpha_i^1 t}\right) \end{aligned}$$

where

$$A_i^\pm = \frac{1}{2} \left(U_i(0) \pm J_i(0) \sqrt{\mathcal{R}_0^i}\ \theta_i\right).$$

From these, we finally calculate $I_i(t)$ through (32). Note that if the initial distribution is assumed to be the stable one (16), then for the first phase we have

$$\begin{aligned} U_1(t) &= I(0)(\alpha_1^* + \gamma_1) e^{\alpha_1^* t}, \\ J_1(t) &= \frac{I(0)(\alpha_1^* + \gamma_1)}{\alpha_1^* + \theta_1} e^{\alpha_1^* t}, \\ I_1(t) &= I(0) e^{\alpha_1^* t}. \end{aligned}$$

Next, for the second phase (the "'lockdown"') we have (note this holds for any subsequent phase)

$$I_2(t) = I_2^0(t) + I_2^+(t) \tag{33}$$

where

$$\begin{aligned} I_2^0(t) &= I(0) e^{\alpha_1^* t_L} e^{-\gamma_2 t} \\ I_2^+(t) &= \left(B^0 e^{-\gamma_2 t} + B^+ e^{\alpha_2^* t} + B^- e^{\alpha_2^1 t}\right) \end{aligned}$$

with

$$\begin{aligned} B^0 &= \frac{\mathcal{R}_0^2 \theta_2^2}{\mathcal{R}_0^1 \theta_1^2} \frac{(\mathcal{R}_0^1 \theta_1^2 - \phi^2)}{(\phi^2 - \mathcal{R}_0^2 \theta_2^2)} I(0) e^{\alpha_1^* t_L}, \\ B^\pm &= \frac{1}{2} \frac{\sqrt{\mathcal{R}_0^2}\theta_2}{\mathcal{R}_0^1 \theta_1^2} \frac{(\sqrt{\mathcal{R}_0^1}\theta_1 - \phi)(\sqrt{\mathcal{R}_0^1}\theta_1 \pm \sqrt{\mathcal{R}_0^2}\theta_2)}{(\sqrt{\mathcal{R}_0^2}\theta_2 \mp \phi)} I(0) e^{\alpha_1^* t_L}, \end{aligned}$$

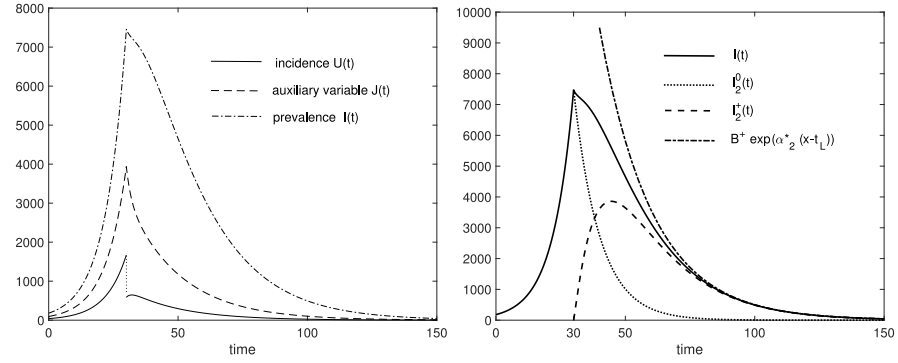

Fig. 2. Illustration of the transition (initiated at the lockdown time $t_L = 30$ day) from the outbreak phase to the lockdown phase, through the solution of the ODE model (31). Parameter values are: $\tau = 0$ day, $\mathcal{R}_0^1 = 2$, $t_L = 30$ day, $\mathcal{R}_0^2 = 0.7$, $\phi = 0.2$ day^{-1}, $\gamma_1 = 0.1$ day^{-1}, $\gamma_2 = 0.1$ day^{-1}. Left panel: plot of the two components $U(t), J(t)$ of the solution to system (31), together with the prevalence function $I(t)$; note that the incidence is discontinuous at t_L. Right panel: plot of the prevalence $I(t)$; for the second phase $I_2(t)$ is split into its two components $I_2^0(t)$ and $I_2^+(t)$ as in (33); the parameter choice implies $\alpha_2^* + \gamma_2 > 0$ so that $I_2^+(t)$ eventually behaves as $B^+ e^{\alpha_2^*(t-t_L)}$; the two components $I_2^0(t)$ and $I_2^+(t)$ are plotted together with $I_2(t)$ and $B^+ e^{\alpha_2^*(t-t_L)}$.

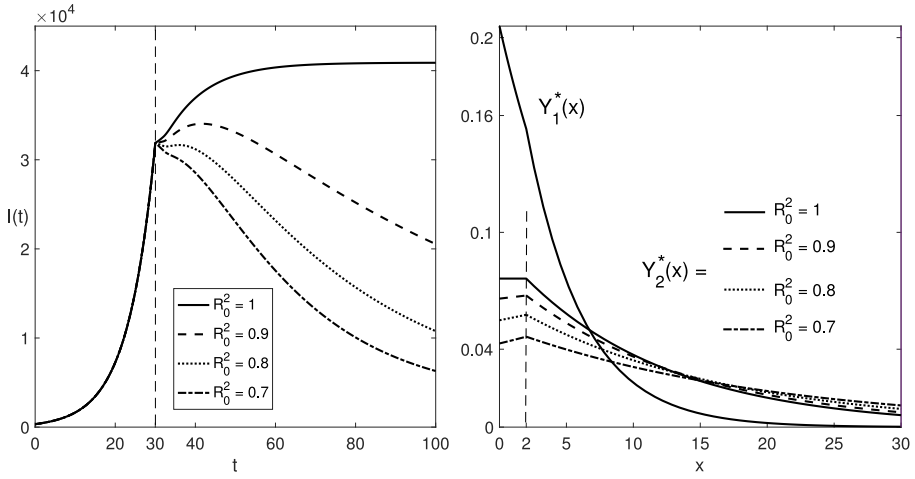

Fig. 3. Further analysis of the transition of the prevalence function $I(t)$ from the outbreak phase to the lockdown phase based on the DDE system (30). Baseline parameter values are: $\tau = 2$ day and $\mathcal{R}_0^1 = 3$ day^{-1}, $\gamma_1 = 0.09$ day^{-1}, $\gamma_2 = 0.09$ day^{-1}. Left panel: comparison of trends of prevalence $I(t)$ for different values of $\mathcal{R}_0^2$. Right panel: comparison between the stable distribution $Y_1^*(x)$ emerging in the first phase and the corresponding distributions $Y_2^*(x)$ emerging in the lockdown phase for the different values of $\mathcal{R}_0^2$.

and, in particular,

$$B^0 + B^+ + B^- = 0.$$

The first term in (33) is due to the prevalence of infected individuals at the beginning of the second phase, which decays at a rate γ_2 (compare with (18)). The second term accounts for transmission events occurred during the lockdown phase. In this case, since $\alpha_2^* < 0$, the asymptotic behavior of $I_2^+(t)$ depends on $\alpha_2^* + \gamma_2$. Indeed we have

$$\text{If } \alpha_2^* + \gamma_2 > 0 \text{ then } I_2^+(t) \approx B^+ e^{\alpha_2^* t}$$
$$\text{If } \alpha_2^* + \gamma_2 < 0 \text{ then } I_2^+(t) \approx B^0 e^{-\gamma_2 t}$$

Some interesting demographic remarks can be made about the dynamics of the infective population $I_2(t)$ during the lockdown phase (see (33)) depending on the sign of $\alpha_2^* + \gamma_2$. On the one hand, for $\alpha_2^* + \gamma_2 > 0$ it is immediate to see that $I_2(t)$ asymptotically behaves as $U_2(t)/(\alpha_2^* + \gamma_2)$ where $(\alpha_2^* + \gamma_2)$ has a clear demographic meaning i.e., it represents the instantaneous per-capita birth rate of the infective population. That is to say, the infective population asymptotically evolves in a stable regime of exponential decline (given that $\alpha_2^* < 0$ as resulting from the hypothesis $\mathcal{R}_0^2 < 1$) where each infective individual produces on average $(\alpha_2^* + \gamma_2)$ new births (i.e., secondary cases) per unit of time.

On the other hand, for $\alpha_2^* + \gamma_2 < 0$, $I_2(t)$ asymptotically behaves as $e^{-\gamma_2 t}(I_2^0 - U_2^0/(\alpha_2^* + \gamma_2))$. The interpretation for this case is that for $\alpha_2^* + \gamma_2 < 0$ the infection reproductivity $\mathcal{R}_0^2$ is brought by the lockdown to such a low level that – essentially – the infective population generated at any time $t > t_L$ does not differ significantly from the sum of the initial infective population present at the beginning of the lockdown phase I_2^0, plus *their* new infective cases (given by $-U_2^0/(\alpha_2^* + \gamma_2)$ as in this case the per-capita birth rate is exactly given by $-(\alpha_2^* + \gamma_2)$) that decay over time at the removal/mortality rate γ_2. In other words, in this case subsequent new infections generated at any time $t > t_L$ are so few that they are hardly distinguishable from the case of a population lacking reproductive power (i.e., having $\mathcal{R}_0 = 0$), despite we specifically assumed $\mathcal{R}_0^2 > 0$. The separation between these two cases occur for $\alpha_2^* + \gamma_2 = 0$, to which it corresponds a reproduction number given by

$$\mathcal{R}_0^{2,*} = \frac{\varphi^2}{(\varphi + \gamma_2)^2} \tag{34}$$

The latter expression is informative about the role played by the two functions shaping the generation time kernel during the second phase, $K_2(x)$. If $\varphi = 0$ i.e., if the transmission rate $\beta_0(x)$ is just an increasing power function, then $\mathcal{R}_0^{2,*} = 0$. On the other hand, if $\varphi > 0$ then $\mathcal{R}_0^{2,*} > 0$ i.e., it becomes a non trivial threshold. This peculiar situation is due to the fact that in our formulation declining infectivity (occurring at the rate $\varphi > 0$) and removal (occurring at the rate $\gamma_2 > 0$) act as additive and independent causes of *infection removal*.

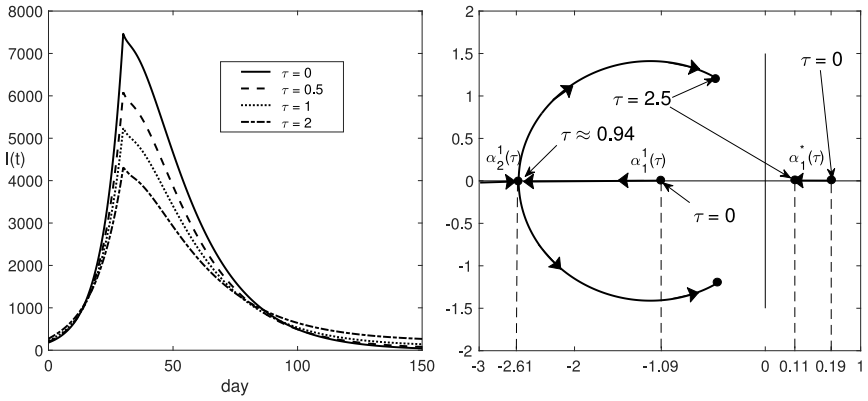

Fig. 4. Dependence of the model solution on the latency delay τ. Left panel: demonstrative solution curves (outbreak and lockdown phases), for different values of τ; baseline parameters are: $\mathcal{R}_0^1 = 2$, $\mathcal{R}_0^2 = 0.7$, $\phi = 0.2$ day^{-1}, $\gamma_1 = 0.1$ day^{-1}, $\gamma_2 = 0.1$ day^{-1}, $t_L = 30$ day. Right panel: characteristic roots of Eq. (35) for $\tau \in [0, 2.5]$; baseline parameters as in left panel; the leading root $\alpha_1^*(\tau)$ is positive and decreases as τ varies from 0 to 2.5 day; the root $\alpha_1^1(\tau)$ is real and decreasing for $\tau \in [0, 0.94]$, then it splits into two complex conjugate roots; a third real root $\alpha_1^2(\tau)$ comes from the left, meets $\alpha_1^1(\tau)$ at $\tau \approx 0.94$ and disappears.

4.3. Numerical illustrations

Fig. 2 illustrates, for the ODE system (31), the behavior of the main variables $U(t)$, $J(t)$, $I(t)$ when the change from phase 1 (invasion) to phase 2 (lockdown) occurs. The left panel shows the discontinuity of the incidence function $U(t)$ at the switching time t_L, resulting from the massive abrupt change in infection reproduction ratios (from $\mathcal{R}_0^1 = 2.0$ to $\mathcal{R}_0^2 = 0.7$). On the right panel we analyze the prevalence $I(t)$ and the two separated terms $I_2^0(t)$ and $I_2^+(t)$ for the lockdown phase. We reported only a case with $\alpha_2^* + \gamma_2 > 0$, the other case with $\alpha_2^* + \gamma_2 < 0$ being very similar.

The left panel of Fig. 3 illustrates the transition pattern of the prevalence function $I(t)$ for the DDE system (30) considering different values of the reproduction number $\mathcal{R}_0^2$ relative to the lockdown phase. This nicely shows the *intrinsic demographic inertia* of the infective population. We termed this type of inertia as "demographic inertia" because it exactly represents the type of inertia that arises in an age-structured population as a consequence of the switch between two stable dynamic regimes characterized by their SAD's. This notion was popularized in classical mathematical demography by Nathan Keyfitz's through the parameter he called *population momentum* [23]. In our model this inertia is marked at relatively high values of $\mathcal{R}_0^2$ (say, in the range $0.85 < \mathcal{R}_0^2 < 1.0$), where prevalence may continue to grow for weeks after the lockdown implementation. This demographic inertia represents the key component of COVID-19 inertia following mitigation interventions, which can be further increased by the characteristic delays arising for COVID-19 (e.g., those due to diagnosis and cases confirmation). At lower values of $\mathcal{R}_0^2$ this inertia is obviously mitigated. The right panel of the same figure compares the stable age distributions $Y_1^*(x)$ emerging during the invasion phase to those relative to the second phase (still considering different values of $\mathcal{R}_0^2$). The difference between these stable age distributions are the ultimate responsible of the transient phase leading from the stable asymptotic behavior during phase 1 to that of phase 2. Notably, during the first phase the stable age distribution is a piecewise-declining exponential function reflecting only epidemic growth (at rate $\alpha_1^* > 0$) at ages $x < \tau$ (where no removal is in place) and the additional effect of removal at ages larger than τ, when also removal starts occurring. On the other hand, during the second phase, the stable distribution is increasing at ages $x < \tau$, reflecting the aging of the infective population, which is asymptotically declining at the rate $\alpha_2^* < 0$.

Concerning the dependence of the DDE system (30) on the latency delay τ, the left panel of Fig. 4 compares the model solution for infected prevalence $I(t)$ for $\tau = 0$ (the unlagged case) with those resulting from different positive values of $\tau > 0$. In the latter case, the characteristic equation is transcendental and reads

$$\left(\lambda + \theta_i\right)^2 = \mathcal{R}_0^i \theta_i^2 e^{-\lambda\tau}. \tag{35}$$

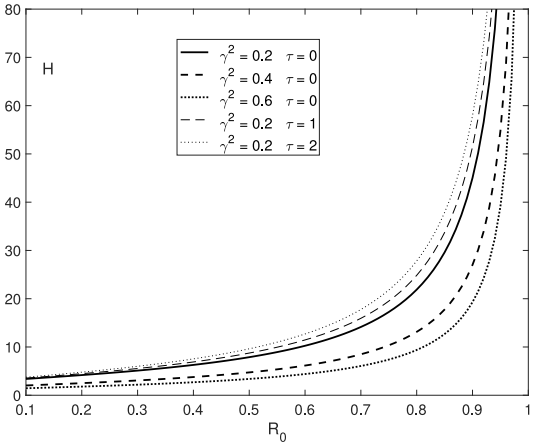

Fig. 5. Halving time H for the lockdown phase versus the reproduction number $\mathcal{R}_0^2$, with $\phi = 0.1$ day^{-1} and for different combinations of τ and γ_2.

Thus, we may have infinitely many roots, though we are mainly interested in the leading root which is real. In the right panel of Fig. 4 we plot the first roots of (35) as a function of τ.

Finally, Fig. 5 reports the halving time of the various epidemic curves in the stable regime of the lockdown phase versus the reproduction number $\mathcal{R}_0^2$, for different combinations of values of the latency delay τ and of the removal rate γ_2. Note that, given the value of $\mathcal{R}_0^2$, an increase of τ produces an increase of the halving time while an increase of the removal rate γ_2 has the opposite effect. Interestingly, the shape of the relationship in Fig. 5 indicates that the greatest part of the decline in the halving time of the suppression path – which is the most straightforward measure of the speed at which the community controls the epidemic – is achieved for $\mathcal{R}_0^2$ values up to say 0.7. Beyond that point, further substantial achievements in decreasing the halving time would require further substantial decreases in $\mathcal{R}_0^2$. This would in turn require a dramatic strengthening of the social distancing conditions, possibly causing an explosion in the epidemic social and economic costs. This effect should primarily be a consequence of the specific steepness of the involved curve whose shape depends on the generation time kernel adopted. Indeed, in the case $\tau = 0$ the halving time is given by

$$H = \frac{\log 2}{\theta_2\left(1 - \sqrt{\mathcal{R}_0^2}\right)}.$$

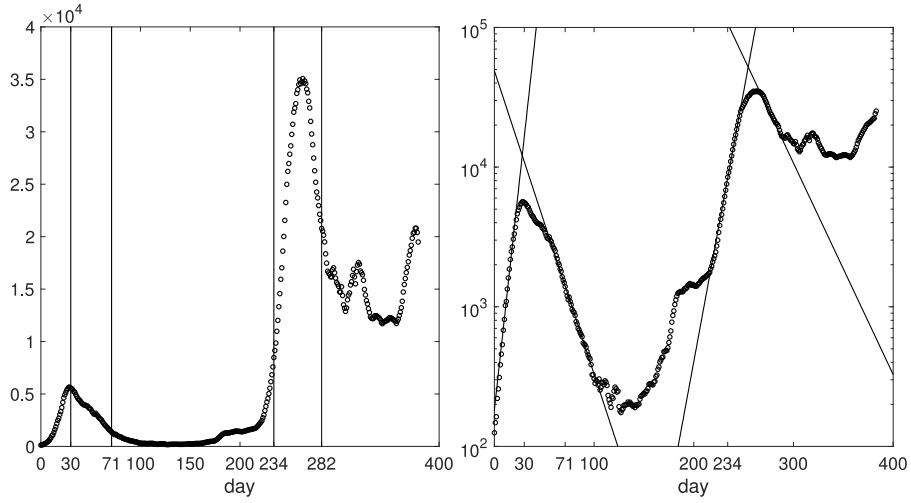

Fig. 6. Data set for daily reported cases, smoothed by a 7 days moving average. Left panel: the number of daily reported cases showing the onset of the epidemics starting with February 24 (day n=0) and the subsequent phase of lockdown, officially starting on March 25 (n=30) and ending on May 5 (n=71), followed by new distancing measures on October 15 (day n=234). Right panel: same data in semi-log scale showing linear traits corresponding to the exponential behavior of the epidemics.

From the latter standpoint, Fig. 5, even if not including social costs, is illustrative of the main societal tradeoff of a generalized lockdown. On one hand, brute-force epidemic suppression requires to bring $\mathcal{R}_0^2$ strongly below one (perhaps below the threshold $\mathcal{R}_0^{2,*}$ stated in (34)). This would require a dramatic haltening of socio-economic activity, with large temporary costs, but at the advantage of getting rid of the epidemic much faster. On the other hand, the attempt to keep the maximum feasible of socio-economic activities opened i.e., bringing $\mathcal{R}_0^2$ just slightly below (or about) 1, allows the epidemic to continue for long time, with a persistently large and slowly declining incidence and prevalence.

5. Application to COVID-19 in Italy

Though our model was mainly theoretical and aiming to investigate the basic mechanisms of a multi-phasic epidemic, it is nonetheless interesting to try to fit the model to COVID-19 data in order to highlight the potential and limits of such a simple tool with real data. Before passing to results we report a short summary of the Italian epidemic as observed so far.

The Italian COVID-19 epidemics has been officially documented since February 24, 2020, with the first 229 cases reported by the devoted Governmental unit *Protezione civile*. Since then, data have been updated daily on a dedicated site (https://github.com/pcm-dpc/COVID-19) providing information on some main stock/flows (confirmed cases, hospitalized cases, ICU occupancy, deaths and tests) at a regional level. However, more detailed data such as those on contact tracing and symptom onset used to compute the weekly figures of the reproduction number R_t by more sophisticated methods were not made available until very late. This makes it of interest an analysis as the present one also for comparing with Governmental estimates.

After a number of purely local measures, such as the creation of *hotspots* in most affected Northern Italy areas since February 20, the full national lockdown was completed on March 25, when only essential activities and services were allowed to be continued. This was preceded by a number of earlier Government measures including the national closure of all schools grades and universities since March 4, 2020, and continued on March 11 by the first nation-wide closure of non-fundamental economic activities. Pairwise, though the lockdown was declared officially over by May 5th, some measures were continued till July 3 in the light of the still critical situation of hospitals, thereby maintaining some degrees of social distancing and individual protection.

However, by the start of holidays (August 2020), the Government publicly encouraged people to avoid international travels and to spend vacations in Italy, contributing to spread optimism and confidence that the epidemic was over, possibly bringing a decline in individuals' attentions. This decline, documented by the press in an endless list of episodes and partly captured by google mobility data, led to a growing pool of new infections clusters that were the premise for the *second wave* once workplaces and schools re-opened since the beginning of September. The subsequent collapse of the tracing system and the ensuing acceleration of the epidemic, brought to further social distancing measures since the second half of October 2020. These measures were, as a rule, aimed to avoid further generalized global lockdown by local-level targeted interventions informed by local figures such as the current reproduction number $\mathcal{R}_t$, the incidence of confirmed cases, and the hospital pressure. Such interventions were able to downturn the second wave and to bring the $\mathcal{R}_t$ in the region of 1 during January and February 2021, but proved insufficient when the alpha "UK" variant became predominant initiating a third epidemic wave [3].

In order to fit reported data with the simplest versions of our model, namely those analyzed in Section 4.2 using parametrization (27), we made a few simplifying assumptions. First, we considered data at the National level. This is not inappropriate for the phases after the first one where epidemic circulation involved the entire country, possibly it is not for the first phase where the epidemic was localized in a few provinces of a few regions. Moreover, being our model highly stylized with a minimal parametrization, it does not include further epidemiological compartments (e.g., isolation, hospitalizations, ICU, and deaths) which are necessary for realistic epidemic modeling. Therefore, we simply used daily reported new cases to compute daily prevalences. The time series D_n of daily reported data was smoothed by a 7 days moving average, in order to remove weekly oscillations due to the sudden fall in the number of tests observed every week-end (Fig. 6). We deliberately did not detrend data by using the time series of tests actually carried out. Assuming, quite crudely, that newly confirmed cases are isolated and therefore removed from the active infective population, newly reported cases are considered as removed individuals $R_i(t)$, satisfying the equation

$$\frac{d}{dt}R_i(t) = \int_0^\infty \gamma_i(x)Y_i(x,t)dx = \gamma_i \int_\tau^\infty Y_i(x,t)dx = \gamma_i I_i^{\#}(t).$$

We therefore can use the formula:

$$\text{infective prevalence} = \frac{1}{\gamma_i} \times \text{daily new (confirmed) cases},$$

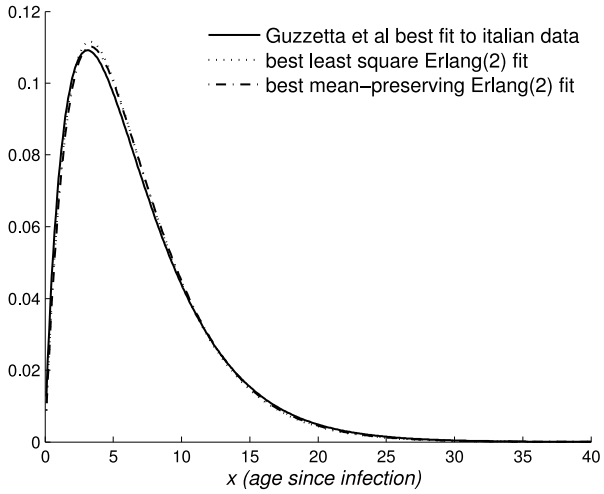

Fig. 7. Graphic comparison between the generation Gamma kernel estimated for Italy in [32] vs two possible Erlang approximation with index 2 namely (i) the mean preserving Erlang distribution used in our application in this section, (ii) the best fitting Erlang distribution computed by a least squares criterion.

where γ_i is the removal rate of the corresponding phase, getting a sequence I_n to be fitted against the theoretical prevalence $I^{\#}(t)$ provided by the model.

The generation time kernel $K_i(x)$ (27) was chosen based on available estimates of the generation time distribution for COVID-19 in Italy [32]. In that study, data on observed serial intervals (i.e., the distance between symptoms in primary infectors and symptoms in their secondary cases), were fitted by a *Gamma* kernel with shape parameter equal to 1.87 and rate 0.28 day^{-1}, yielding a mean generation time of 6.6 day. By sticking on their estimated mean generation time, our stylized kernel $K_1(x)$ in (27) provided an extremely good approximation of it (see Fig. 7), by simply taking (an analogous approach was used in [33])

$$\theta_1 = 0.30 \text{ day}^{-1}.$$

To fit our model throughout the various epidemic phases (Fig. 6, left panel), we checked in the data for evidence of the presence of the key model feature, namely the emergence of epochs of stable exponential behavior eventually resulting from the implementation (or the relaxation) – for a sufficiently long period of time – of mitigation interventions. This is clear from the log-scale diagram in the right panel of Fig. 6 where we could well identify at least four phases characterized by an initial transient dynamics which eventually approaches an exponential trend. These four phases are represented by (i) the initial phase of epidemic growth (roughly days 0–19), (ii) the declining phase after the peak induced buy the lockdown (roughly days 80–100), (iii) the phase of exponential increase during the second wave (days 210–250), (iv) the phase of the exponential decline resulting from the measures adopted to mitigate the second wave (days 275–290).

This shape in turn suggests a 2-stage fitting procedure, to be repeated for each phase transition, where: (i) in the first stage the characteristic root α_i^* for the particular phase considered is estimated, (ii) in the second stage the remaining parameters (τ, γ_i, t_i.) are adjusted conditional on α_i^* (actually τ is fitted only once because it is a biological constant pertaining to the infection and it will not plausibly change through the different phases). In what follows, we detail the adopted approach for the first two phases and briefly summarize it for the remaining ones.

5.1. The outbreak phase

As discussed in Section 4.2, we assume that the epidemic was initiated by an initial cohort of infective individuals distributed according to the corresponding stable age distribution of the invasion phase. Consequently, the initial growth of the epidemic curve is expected to be exponential since initial time with rate α_1^*. This parameter was estimated by a (least squares) exponential fit choosing the data window yielding the highest value of the determination coefficient R^2. This yielded

$$I_0 = 181 \qquad \alpha_1^* = 0.15 \text{ day}^{-1},$$

over the period from day $n = 0$ to $n = 19$, with $R^2 \approx 0.99$. The fact that including further data points beyond $n = 19$ day always worsened the fit, may be interpreted as due to the onset of social distancing gradually showing its effect soon after time $n = 19$ day, that is since March 15.

5.2. The lockdown phase

At a crude glance, the lockdown phase might seem to display its effects on the epidemic curve from day $n = 27$ to $n = 170$. However, we eventually selected the period from $n = 80$ to $n = 105$, showing a markedly exponential pattern, to identify α_2^*, obtaining

$$\alpha_2^* = -0.05 \text{ day}^{-1} \quad (R^2 \approx 0.99)$$

Adjusting the remaining model parameters by a conditional fit we obtained

$$\tau = 2 \text{ day}, \quad \gamma_1 = \gamma_2 = 0.09 \text{ day}^{-1}, \quad t_L = 22 \text{ day},$$

yielding in particular a latency delay of about two days. Thus we could compute

$$\phi = \theta_1 - \gamma_1 = 0.21 \text{ day}^{-1}, \quad \theta_2 = 0.30 \text{ day}^{-1}.$$

This in turn allowed to also estimate the reproduction number $\mathcal{R}_0^1$ of the first phase

$$\mathcal{R}_0^1 = e^{\tau\alpha_1^*}\left(1 + \frac{\alpha_1^*}{\theta_1}\right)^2 \approx 3.06,$$

as well as that of the second phase:

$$\mathcal{R}_0^2 = e^{\tau\alpha_2^*}\left(1 + \frac{\alpha_2^*}{\theta_2}\right)^2 \approx 0.61.$$

The overall fit of the model (30) over the first two phases is shown in the left panel of Fig. 8. Though the overall quality might seem acceptable, there is a clear loss of fit around the transition time t_L. This is the consequence of the abrupt change of the reproduction number from $\mathcal{R}_0^1$ to $\mathcal{R}_0^2$ causing a sharp discontinuity in the incidence $U(t)$ function. A correction of the model to remove the lack of realism of the hypothesis of abrupt transition, obtained through a continuous exponential decline of R_0^2 distributed over a time span of two weeks is shown (8, right panel) to remarkably correct the lack of fit previously commented. In this experiment, we considered

$$R_0^2(t) = R_0^1 e^{-\delta(t-t_L)} \quad \text{where} \quad \delta = (1/\Delta)\ln(\mathcal{R}_0^1/\mathcal{R}_0^2)$$

and run the modified model obtaining $t_L = 17.3$ and $\Delta = 10.5$, suggesting that the decline in the reproduction number R_0^2 of the second phase realistically occurs through a time span of about 10 days.

5.3. Further phases

Concerning the subsequent phases, namely (i) the second wave phase which started due to the relaxation of most measures adopted during the lockdown, (ii) the subsequent mitigation phase with the so-called "multicolor strategy" [3] based on locally targeted (typically: at the regional level) measures in view of the documented seriousness of the local situation, up to (iii) the current emergence of more transmissible variants of the virus, we used the same procedure detailed above. The results are reported in Table 1 and Fig. 9.

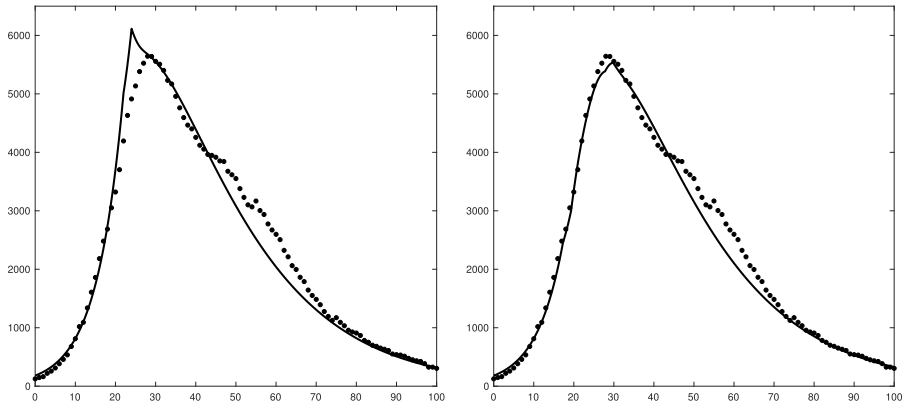

Fig. 8. The transition from the outbreak phase to the lockdown phase fitted by the proposed model (30) (parameter estimates as described in Sections 5.1 and 5.2). Left panel: observed vs fitted prevalence $I(t)$ under our baseline assumption that transition between phases occurs instantaneously; in this case the predicted curve shows a cuspid at the transition time t_L, due to the discontinuity of the incidence function $U(t)$. Right panel: fit obtained through a slight modification of the model allowing a smooth transition.

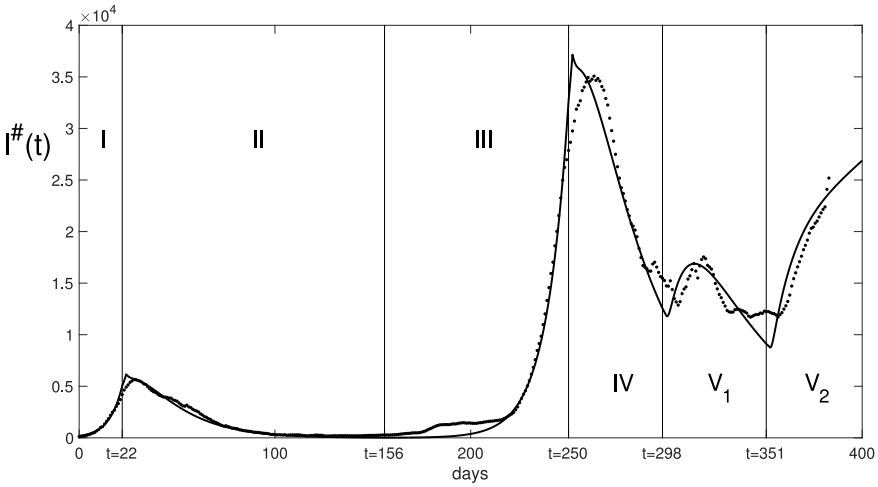

Fig. 9. Overall fit by the proposed model to the entire course of the COVID-19 epidemic in Italy (parameter estimates as described in the present Section 5).

The estimates of the phase-specific reproduction numbers $\mathcal{R}_0^i$ are systematically close to the nation-wide figures of the current reproduction number $\mathcal{R}_t$ estimated weekly by the Governmental unit [3].

Fig. 9, provides a summary view of the explanatory power of the model (30), by comparing its solution with the data over the entire course of the epidemic observed in Italy so far. Note that we deliberately avoided to include generalized smooth transitions between phase specific reproduction numbers just for purposed of obtaining better-looking fits. Note also that the poorer fit over the last two phases is the consequence of their short duration compared to phases I–IV, preventing the emergence of any meaningful stable exponential phase.

Table 1
Summary of model parameters resulting from the fitting procedure. Common parameters to all phases are $\phi = 0.21$, $\tau = 2$, $\gamma = 0.09$.

Phase	α^*	$\mathcal{R}_0$	start	end
I (outbreak)	0.15	3.06	t=0	t=22
II (lockdown)	−0.05	0.63	t=23	t=156
III (release)	0.09	2.01	t=157	t=250
IV (re-lock)	−0.035	0.73	t=251	t=289
V_1 (multi-color)	–	0.8	t=290	t=351
V_2 (multi-color)	–	1.05	t=352	–

6. Discussion

To sum up, the overall fit in Fig. 9 was very good almost everywhere but in the most recent phases, despite the little information used, namely the data from the sole exponential windows in each phase. Essentially, prior to the recent period made complicated by the emergence of the new variants, there are only two periods where some lack of fit emerges, namely a portion of the declining trend during the lockdown phase and the first epoch of the re-opening phase. This mostly lies in (i) the minimalistic structure of our model which does not explicitly include any compartment directly comparable with real data, (ii) the stylized hypothesis by which switches from one phase to the successive one occur abruptly. Refined data explanations would require to add a number of further compartments to account for the delayed processes of the disease pathways, testing, isolation, hospitalization etc, and to amend our homogeneous mixing framework to include heterogeneities in transmission and disease (over age and space) and testing.

This said, the fitted model is the simplest setup illustrating the naturally embedded *demographic inertia* of COVID-19 dynamics despite the lack of any delayed dimensions (such as ICU and deaths). Here, by "demographic inertia" we mean the natural tendency of a multi-phasic epidemic to display its true phase-specific trends with a temporal delay compared to the moment when that particular phase was initiated, for example due to the activation of a mitigation intervention. This demographic inertia is to be distinguished from other components of COVID-19 observed inertia that are due to the fact that its main public health outcomes e.g., cases confirmations (not to say of hospitalizations

and deaths) arise with substantial time-delays with respect to the time when transmission occurred. This natural demographic inertia can be macroscopic or negligible, other parameters being equal, depending on the relative magnitude of the reproduction numbers $\mathcal{R}_0^i, \mathcal{R}_0^{i+1}$ that are characteristic of two successive phases. This was illustrated in the left panel of Fig. 3.

Therefore, looking at Fig. 9 we can say that: (i) the continued growth in prevalence after the first lockdown declaration is largely attributable to delayed patterns rather than to demographic inertia (as can be understood from the model lack of fit at the switch between the first two phases). The reason is that the lockdown brought the reproduction number representative of the second phase down to a quite low level. On the other hand, the continued decline in prevalence for a very long time beyond the lockdown official end-date at May 5 (about time $n = 80$ day) can be partly explained by demographic inertia exactly due to the very low level of $\mathcal{R}_0^i$ eventually achieved during the lockdown phase. Borrowing from a language typical of vaccination programs, the latter prolonged *honey-moon* period likely surely contributed to the generation of an optimism wave in the public opinion that eventually was responsible for the second pandemic wave.

7. Concluding remarks

Based on a simple hypothesis, namely that the depletion of the susceptible population remains contained over time, in this manuscript we have proposed a minimal SIR model for the overall dynamics of an epidemic evolving in a multi-phasic form due to a sequence of epochs of mitigation intervention and their relaxations, as has been the case for the COVID-19 epidemic. Our model provides, first of all, a unified representation of such multi-phasic epidemics, describing how the different epidemic phases unfold over time as a consequence of the interplay between the transmission process and intervention responses. Second, by using a wide class of generation time kernels allowing reducibility (either to ordinary or delayed differential equations), we investigated in depth a low-dimensional case allowing a non-trivial full analytical treatment also of the transient dynamics connecting the different epidemic phases. The latter model provided an excellent fit to the entire course of the epidemic observed in Italy since February 2020, despite its parsimonious parametric structure, not including any disease compartments and using minimal data. Last, the proposed model represents the simplest setup illustrating the intrinsic demographic component of the *inertia* of COVID-19 dynamics. This demographic inertia adds to the one due to the various delayed phenomena characteristic of COVID-19 to determine the overall COVID-19 inertia. The limitations of the proposed model lie in its parsimonious structure and in its kinematic nature. As for the first drawback, this can be easily overcome by simply adding disease-related and other relevant compartments, as well as by introducing relevant heterogeneities, as those due to chronological age such as, e.g., in the onset of symptoms and serious disease. This however represents only a part of the story of COVID-19 which is clearly more complicated than described by simple models. For example, we described the incoming story of the second COVID-19 wave, debuted in Italy since the start of September 2020, in a short report (Iannelli et al. [34]) where we correctly predicted that the observed epidemic was dramatically accelerating since the beginning of October, with $\mathcal{R}_t$ shifting upward, and would continue to do so for a while, due to the overwhelming of the testing–tracing system by epidemic growth. We also suggested that the subsequent epidemic increase would have continued up to a maximum corresponding to the reproduction number generated in the preceding weeks by the behavior of risky groups, and would have halted or slowed down once the more prudent population groups would be reached. This is suggestive of the fact that, at a fine scale, the epidemic description requires to account for a number of further dimensions. These include saturation effects, as the finiteness of tracing resources, capable – if alone – to slow-down the epidemic but only for a while, as well as to go beyond the kinematic representation to include a number of behavioral aspects. Behavioral effects are numerous but – just for illustrative purposes – one of these is surely understandable as a consequence of the inertia phenomenon that we termed the post-lockdown *honey-moon* period. This prolonged period of low epidemic activity surely contributed to the generation of the optimism wave in the Italian public that contributed to the drop of prudent behaviors, and eventually accelerated the second pandemic wave. This calls for behavioral epidemiology explanations [26,27]. Learning from such aspects is of dramatic importance also in prospective terms, namely the control of future pandemic events.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the anonymous referees whose suggestions improved this manuscript. Mimmo Iannelli wishes to thank the Mathematical Department of Trento University for the support received.

Appendix A. Survey of results from the renewal equation theory

Here we recall some results from the theory of the renewal equation

$$u(t) = F(t) + \int_0^t K(x)u(t-x)dx, \tag{A.1}$$

and nomenclature coming from the demographic theory of the stable population. Under rather general assumption such as

- $F(t) \geq 0 \quad (t \in [0,\infty))$ is continuous and absolutely Laplace transformable,
- $K(x) \geq 0 \quad (x \in [0,\infty))$ is absolutely Laplace transformable,

Eq. (A.1) has a unique continuous non-negative solution $u(t) \quad (t \in [0,\infty))$ whose asymptotic behavior has been extensively discussed by Feller [35] and is related (under some general assumption that we suppose fulfilled by the equation coming from our model) to the roots of the characteristic equation

$$\widehat{K}(\lambda) = 1.$$

where we use symbol $\widehat{f}$ to denote the Laplace transform of the function f.

Following Feller [35], it is known that this equation has one and only one leading root α^*, which is real, simple and strictly positive (negative) if the condition $\widehat{K}(0) > 1$ (respectively $\widehat{K}(0) < 1$) is satisfied. Precisely we have the sequence $\alpha^*, \{\alpha^j\}_{j=1}^{j=\infty}$ such that

$$\Re\alpha^{j+1} \leq \Re\alpha^j < \alpha^* \qquad (j = 1,\ 2,\ \ldots)$$

and the solution has the asymptotic expansion

$$u(t) = u^* e^{\alpha^* t} + \sum_{j=1}^{\infty} u^j e^{\alpha^j t}. \tag{A.2}$$

This means that, for any integer N, $u(t)$ can be written as

$$u(t) = \left(u^* e^{\alpha^* t} + \sum_{j=1}^{N} u^j e^{\alpha^j t}\right)(1 + \Omega(t)),$$

with

$$\lim_{t\to\infty} \Omega(t) = 0.$$

As for the coefficients u^j we have

$$u^j = \frac{\widehat{F}(\alpha^j)}{\displaystyle\int_0^\infty xK(x)e^{-\alpha^j x}dx}.$$

In the particular case of

$$F(t) = ce^{\alpha^* t}\int_t^\infty K(x)e^{-\alpha^* x}dx,$$

then the expansion (A.2) reduces to the first term

$$u(t) = u^* e^{\alpha^* t}, \qquad u^* = c.$$

In any other case such a solution is reached asymptotically as we have

$$\lim_{t\to\infty} e^{-\alpha^* t}u(t) = u^*. \tag{A.3}$$

We can estimate the speed of this convergence by truncating the asymptotic expansion (A.2) at the second term

$$u(t) \approx u^* e^{\alpha^* t} + u^1\ e^{\alpha^1 t},$$

to have

$$\left|e^{-\alpha^* t}u(t) - u^*\right| \le \left|u^1\right|\ e^{(\Re\alpha^1 - \alpha^*)t}$$

so that

$$T_\epsilon = \frac{\ln(|u^1|) - \ln(\epsilon)}{\alpha^* - \Re\alpha^1}. \tag{A.4}$$

is an estimate of the minimum time required to get the limit (A.3) reached within a tolerance ϵ.

Appendix B. The linear chain-trick

Reduction of the main system (4) to a cascade of ODE–DDE equation is allowed by the special form of the basic parameters (24), (25). Here we illustrate the method through the special case (27).

The starting point are the variables (28) and (29). From (4) we see that the latter satisfy the reduced system

$$\begin{cases} \left(\dfrac{\partial}{\partial t} + \dfrac{\partial}{\partial x}\right) y_i(x,t) = 0, \\ y_i(0,t) = R_0^i \displaystyle\int_0^{x_+} K_i(x)y_i(x,t)dx, \\ y_i(x,0) = \dfrac{Y_i^0(x)}{\Gamma_i(x)}. \end{cases} \tag{B.1}$$

then

$$\frac{d}{dt}U_i(t) = \mathcal{R}_0^i\int_\tau^\infty K_i(x)\frac{\partial y_i}{\partial t}(x,t)dx = -\mathcal{R}_0^i\int_\tau^\infty K_i(x)\frac{\partial y_i}{\partial x}(x,t)dx$$

and, integrating by parts, we have

$$\frac{d}{dt}U_i(t) = \mathcal{R}_0^i\theta_i^2\int_\tau^\infty \big(1-\theta_i(x-\tau)\big)\,e^{-\theta_i(x-\tau)}y_i(x,t)dx = \mathcal{R}_0^i\theta_i^2 J_i(t) - \theta_i U_i(t).$$

Analogously

$$\frac{d}{dt}J_i(t) = y_i(\tau,t) - \theta_i\int_\tau^\infty e^{-\theta_i(x-\tau)}y_i(x,t)dx = U_i(t-\tau) - \theta_i J_i(t),$$

where we have used (6) to express $y_i(\tau,t)$, and $U_i(t)$ is extended to the interval $[-\tau,0]$ by

$$U_i(t) = y_i^0(-t), \quad t\in[-\tau,0].$$

Thus we obtain system (30), endowed with the initial data

$$\begin{aligned} &U_1(t) = y_1^0(-t),\ t\in[-\tau,0],\\ &J_1(0) = \int_\tau^\infty e^{-\theta_1(x-\tau)}y_1^0(x)dx,\\ &U_2(t) = U_1(t_L+t)\ t\in[-\tau,0],\\ &J_2(0) = J_1(t_L),\\ &U_3(t) = U_2(t_R - t_L + t)\ t\in[-\tau,0],\\ &J_3(0) = J_2(t_R). \end{aligned}$$

When the initial distribution $Y^0(x)$ is assumed to be the stable one (11) then

$$U_1(t) = \frac{\alpha_1^*(\alpha_1^*+\gamma_1)}{\alpha_1^* + \gamma_1(1-e^{-\alpha_1^*\tau})}I_0 e^{\alpha_1^* t}, \quad t\in[-\tau,0]$$

$$J_1(0) = \frac{e^{-\alpha_1^*\tau}}{\alpha_1^*+\theta_1}U_1(0).$$

The previous reduction procedure, can be applied to any choice of the functions (24), (25), and leads to a cascade of equations where several auxiliary variables are necessary in order to complete the chain.

References

[1] WHO, COVID-10 situation reports 1-209 and weekly epidemiological updates, https://www.who.int/publications/m/item/weekly-epidemiological-update.

[2] N.M. Ferguson, D. Laydon, G. Nedjati-Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunubá, G. Cuomo-Dannenburg, et al., Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand. Imperial College COVID-19 Response Team, Report No. 9, 2020, Imperial College COVID-19 Response Team, 2020, p. 20, http://dx.doi.org/10.25561/77482.

[3] ISS, COVID-19 Epidemic in Italy. National and Regional Upgrade, Bollettino-Sorveglianza-Integrata-COVID-19-10-Marzo-2021 (Weekly Report on COVID-19 Epidemic), 2020, (in Italian).

[4] A. Aleta, D. Martín-Corral, A. Pastore y Piontti, M. Ajelli, M. Litvinova, M. Chinazzi, N.E. Dean, M.E. Halloran, I.M. Longini Jr., S. Merler, et al., What Does and Does Not Correlate with COVID-19 Death Rates, Springer Science and Business Media LLC, 2020.

[5] A. Aleta, D. Martin-Corral, A.P. y Piontti, M. Ajelli, M. Litvinova, M. Chinazzi, N.E. Dean, M.E. Halloran, I.M. Longini Jr., S. Merler, et al., Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19, Nat. Hum. Behav. 4 (9) (2020) 964–971.

[6] M. Biggerstaff, B.J. Cowling, Z.M. Cucunubá, L. Dinh, N.M. Ferguson, H. Gao, V. Hill, N. Imai, M.A. Johansson, S. Kada, et al., Early insights from statistical and mathematical modeling of key epidemiologic parameters of COVID-19, Emerg. Infect. Diseases 26 (11) (2020) e201074, http://dx.doi.org/10.3201/eid2611.201074.

[7] M. Chinazzi, J.T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, A.P. y Piontti, K. Mu, L. Rossi, K. Sun, et al., The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak, Science 368 (6489) (2020) 395–400, http://dx.doi.org/10.1126/science.aba9757.

[8] E. Estrada, COVID-19 and SARS-CoV-2. Modeling the present, looking at the future, Phys. Rep. (2020).

[9] L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, L. Abeler-Dörner, M. Parker, D. Bonsall, C. Fraser, Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing, Science 368 (6491) (2020) http://dx.doi.org/10.1126/science.abb6936.

[10] S. Flaxman, S. Mishra, A. Gandy, H.J.T. Unwin, T.A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J.W. Eaton, et al., Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe, Nature 584 (7820) (2020) 257–261, http://dx.doi.org/10.1038/s41586-020-2405-7.

[11] M. Gatto, E. Bertuzzo, L. Mari, S. Miccoli, L. Carraro, R. Casagrandi, A. Rinaldo, Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures, Proc. Natl. Acad. Sci. 117 (19) (2020) 10484–10491.

[12] G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. Di Filippo, A. Di Matteo, M. Colaneri, Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy, Nature Med. 26 (6) (2020) 855–860.

[13] M.U. Kraemer, C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, D.M. Pigott, L. Du Plessis, N.R. Faria, R. Li, W.P. Hanage, et al., The effect of human mobility and control measures on the COVID-19 epidemic in China, Science 368 (6490) (2020) 493–497.

[14] A.J. Kucharski, T.W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, R.M. Eggo, F. Sun, M. Jit, J.D. Munday, et al., Early dynamics of transmission and control of COVID-19: a mathematical modelling study, Lancet Infect. Dis. 20 (5) (2020) 553–558.

[15] T. Kuniya, H. Inaba, Possible effects of mixed prevention strategy for COVID-19 epidemic: massive testing, quarantine and social distancing, AIMS Public Health 7 (3) (2020) 490.

[16] Z. Liu, P. Magal, O. Seydi, G. Webb, A COVID-19 epidemic model with latency period, Infect. Dis. Modell. 5 (2020) 323–337.

[17] Z. Liu, P. Magal, G. Webb, Predicting the number of reported and unreported cases for the COVID-19 epidemics in China, South Korea, Italy, France, Germany and United Kingdom, J. Theoret. Biol. 509 (2021) 110501.

[18] C. Poletto, S. Scarpino, E. Volz, Applications of predictive modelling early in the COVID-19 epidemic, Lancet Digit. Health 2 (10) (2020) E498–E499.

[19] A. Vespignani, H. Tian, C. Dye, J.O. Lloyd-Smith, R.M. Eggo, M. Shrestha, S.V. Scarpino, B. Gutierrez, M.U. Kraemer, J. Wu, et al., Modelling covid-19, Nature Rev. Phys. 2 (6) (2020) 279–281.

[20] P.G.T. Walker, C. Whittaker, O.J. Watson, M. Baguelin, P. Winskill, A. Hamlet, B.A. Djafaara, Z. Cucunubá, D. Olivera Mesa, W. Green, H. Thompson, S. Nayagam, K.E.C. Ainslie, S. Bhatia, S. Bhatt, A. Boonyasiri, O. Boyd, N.F. Brazeau, L. Cattarino, G. Cuomo-Dannenburg, A. Dighe, C.A. Donnelly, I. Dorigatti, S.L. van Elsland, R. FitzJohn, H. Fu, K.A.M. Gaythorpe, L. Geidelberg, N. Grassly, D. Haw, S. Hayes, W. Hinsley, N. Imai, D. Jorgensen, E. Knock, D. Laydon, S. Mishra, G. Nedjati-Gilani, L.C. Okell, H.J. Unwin, R. Verity, M. Vollmer, C.E. Walters, H. Wang, Y. Wang, X. Xi, D.G. Lalloo, N.M. Ferguson, A.C. Ghani, The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries, Science 369 (6502) (2020) 413–422, http://dx.doi.org/10.1126/science.abc0035.

[21] J. Zhang, M. Litvinova, Y. Liang, Y. Wang, W. Wang, S. Zhao, Q. Wu, S. Merler, C. Viboud, A. Vespignani, et al., Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China, Science 368 (6498) (2020) 1481–1486.

[22] ISTAT, First results of the national serosurvey on SARS-COV-2, 2020, https://www.istat.it/it/files//2020/08/ReportPrimiRisultatiIndagineSiero.pdf.

[23] N. Keyfitz, Applied Mathematical Demography, Springer, New York, 1985.

[24] M. Iannelli, F. Milner, The Basic Approach to Age-Structured Population Dynamics, Springer, New York, 2017.

[25] H. Inaba, Age-Structured Population Dynamics in Demography and Epidemiology, Springer, 2017.

[26] P. Manfredi, A. D'Onofrio, Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases, Springer Science & Business Media, 2013.

[27] Z. Wang, C.T. Bauch, S. Bhattacharyya, A. d'Onofrio, P. Manfredi, M. Perc, N. Perra, M. Salathé, D. Zhao, Statistical physics of vaccination, Phys. Rep. 664 (2016) 1–113.

[28] T. Britton, G. Scalia Tomba, Estimation in emerging epidemics: Biases and remedies, J. R. Soc. Interface 16 (150) (2019) 20180670.

[29] N. MacDonald, Biological Delay Systems: Linear Stability Theory, Cambridge University Press, Cambridge, 1989.

[30] X. Liu, P. Stechlinski, Pulse and constant control schemes for epidemic models with seasonality, Nonlinear Anal. RWA 12 (2) (2011) 931–946.

[31] X. Liu, P. Stechlinski, Infectious Disease Modeling, Springer, 2017.

[32] G. Guzzetta, F. Riccardo, V. Marziano, P. Poletti, F. Trentini, A. Bella, X. Andrianou, M. Del Manso, M. Fabiani, S. Bellino, et al., The impact of a nation-wide lockdown on COVID-19 transmissibility in Italy, 2020, arXiv preprint arXiv:2004.12338.

[33] K.B. Blyuss, Y.N. Kyrychko, Effects of latency and age structure on the dynamics and containment of COVID-19, J. Theoret. Biol. 513 (2021) 110587.

[34] M. Iannelli, P. Manfredi, S. Salmaso, G. ScaliaTomba, Se l'epidemia accelera cosa resta da fare? 2020, Scienza in rete, https://www.scienzainrete.it/articolo/se-lepidemia-accelera-cosa-resta-da-fare/mimmo-iannelli-piero-manfredi-stefania-salmaso.

[35] W. Feller, On the integral equation of renewal theory, Ann. Math. Stat. 12 (1941) 243–267.