Optimal control techniques based on infection age for the study of the COVID-19 epidemic

Joseph Frédéric Bonnans, Justina Gianatti

OPTIMAL CONTROL TECHNIQUES BASED ON INFECTION AGE FOR THE STUDY OF THE COVID-19 EPIDEMIC [*]

J. FRÉDÉRIC BONNANS[†] AND JUSTINA GIANATTI[‡]

September 11, 2020

Abstract. We propose a model for the COVID-19 epidemic where the population is partitioned into classes corresponding to ages (that remain constant during the epidemic). The main feature is to take into account the infection age of the infected population. This allows to better simulate the infection propagation that crucially depend on the infection age. We discuss how to estimate the coefficients from data available in the future, and introduce a confinement variable as control. The cost function is a compromise between a confinement term, the hospitalization peak and the death toll. Our numerical experiments allow to evaluate the interest of confinement varying with age classes.

Key words. COVID-19, epidemic modelling, confinement, age-structured systems, optimal control.

AMS subject classifications. 49M25, 92D25, 92D30, 35Q93

1. Introduction. The COVID-19 epidemic appears for the first time in Wuhan, China at the end of 2019. It quickly spread throughout the world, and on March 12, 2020 it was declared a pandemic by the World Health Organization (WHO). So far there are no vaccines and it is not clear if there is any effective treatment to combat it. That is why the governments of the different countries began to implement containment measures in the population, trying to control the disease. Since these actions have a great impact on the daily life of the population, one of the big questions now is knowing how long to carry out these measures.

The recent COVID-19 epidemic has already been modelized in e.g. [8, 10, 9]. These papers use variants and extensions of the SIR (Susceptible-Infected-Recovered) model based on ODEs (ordinary differential equations). SIR is the standard approach for the modelization of epidemics and takes its origin in [7]. Of course it is difficult to model confinement. The simplest way is to assume that it boils down to a multiplicative factor coefficient with value in $[0, 1]$ in the formula for new infections. That is, the dynamics in the SIR setting would be

$$
\begin{aligned}
\dot{S}(t) &= -\delta(1 - u(t))S(t)I(t) \\
\dot{I}(t) &= \delta(1 - u(t))S(t)I(t) - (\eta + \mu)I(t) \\
\dot{R}(t) &= \mu I(t)
\end{aligned}
\tag{1.1}
$$

with $\delta > 0$ infection coefficient, $\eta > 0$ death rate, and $\mu > 0$ recovery rate. Observe that, here as in the sequel, our models do not take into account birth, as well as death for other reasons than the epidemic itself. We recover the classical SIR dynamics in the case of an uncontrolled system, i.e., when $u(t) = 0$ for all t. Total confinement is out of reach and therefore we may assume that we have the control constraint $u(t) \leq u^M$, with $u^M \in (0, 1)$. Of course it is not easy to specify a confinement cost, but we might take

$$
c(u) := p_u \int_0^T \varphi(u(t))\mathrm{d}t
\tag{1.2}
$$

with $p_u \geq 0$ penalty factor, and for example $\varphi(s) = s$ (as we did in our numerical tests) or φ quadratic.

Our proposal is to distinguish, among infected people, those who are hospitalized and those who are not, and to take into account the limited duration of the infection. The latter can be achieved by associating with the infected population an "infection age" variable. That also allows us to differentiate between the incubation and infectious phases of the disease. Furthermore, since the

severity of this COVID-19 seems to depend on the age of the infected population, it may be useful to take into account the population age, in addition to the infection age. The cost function will minimize a combination of confinement cost, value of hospitalization peak, and death toll.

Concerning the optimal control of age-structured population, see [4] (without diffusion), and [1] (with diffusion). The optimal control of age models with state constraints is discussed in [3]. About optimal control application to epidemiological models, see e.g. [11]. About the related question of models involving delays, see e.g. [13]. On optimal control approaches for COVID-19, see [5]. During the revision process, we noticed the appearance of [12] that also propose an optimal control model with infection age. The main difference with respect to our model is that we consider the peak value in the objective function, which leads to a problem with state constraints. Let us also mention a mean field game approach to the COVID-19 epidemic [6].

The paper is organized as follows. In Section 2 we present a continuous time (and infection age) optimal control problem. The corresponding discrete time problem is discussed in Section 3. The data analysis (how to estimate the coefficients of the problem) as well as the initial conditions are the subject of Section 4. Finally in Section 5 we discuss various numerical results. The paper ends with a conclusion giving a perspective of possible extensions.

Our open access code, based on the optimal control toolbox Bocop http://www.bocop.org, can be downloaded from http://www.cmap.polytechnique.fr/∼bonnans/covid-19/covid-19.html or https://www.fceia.unr.edu.ar/∼justina/covid-19/.

2. Continuous time optimal control problem.

2.1. An infection age structured model. We have two kinds of patients, the exposed ones, who are asymptomatic or with mild symptoms, and those who need hospitalization. Let $z(t, a, b)$ (resp. $h(t, a, b)$) denote the number of exposed (resp. hospitalized) patients of age $a \in [0, A]$ at time t, that have been infected for duration $b \in [0, B]$. Here A and B are positive constants. Let $y(t, a)$ (resp. $\bar{y}(t, a)$) denote the number of non immunized (resp. immunized) people of age a at time t. Since in practice the age variable a does not significantly vary during the epidemic, we assume that individuals keep the same age over time. The control variable $u(t, a) \in [0, 1]$ is the confinement policy (no confinement if $u(t, a) = 0$, total confinement if $u(t, a) = 1$). We consider the following dynamics

$$(2.1) \quad \begin{cases} \dot{y}(t, a) = -\delta(a)(1 - u(t, a))Z(t)y(t, a) \\ y(0, a) = y_0(a), \\ (z_t + z_b)(t, a, b) = -\nu(a, b)z(t, a, b) \\ z(0, a, b) = z_0(a, b) > 0, \quad z(t, a, 0) = \delta(a)(1 - u(t, a))Z(t)y(t, a) \\ (h_t + h_b)(t, a, b) = \nu(a, b)z(t, a, b) - (\eta(a, b) + \gamma(a, b)E(t)) h(t, a, b) \\ h(0, a, b) = h_0(a, b) \geq 0, \quad h(t, a, 0) = 0 \\ \dot{\bar{y}}(t, a) = z(t, a, B) + h(t, a, B) \\ \bar{y}(0, a) = 0. \end{cases}$$

Here $\delta(a) \geq 0$ is the "transmission coefficient", $Z(t)$ is the "global transmission factor", modeled as

$$(2.2) \quad Z(t) = \int_0^A \int_0^B e(a, b)z(t, a, b)\mathrm{d}a\mathrm{d}b,$$

with $e(a, b) \geq 0$ "transmission factor" (we assume that hospitalized patients do not contribute to the transmission of the disease). Also, $\nu(a, b) \geq 0$ is the hospitalization coefficient, and $E(t) \in [0, 1]$ is the hospital saturation estimate, given by

$$(2.3) \quad E(t) := \frac{(H(t) - C)_+}{H(t) + C}; \quad H(t) := \int_0^A \int_0^B h(t, a, b)\mathrm{d}a\mathrm{d}b,$$

where $C > 0$ is the nominal capacity. We assume that only hospitalized patients die, with death rate

$$(2.4) \qquad d(t, a, b) := \eta(a, b) + \gamma(a, b) E(t),$$

where $\eta(a, b) \geq 0$ is the minimal death rate and $\gamma(a, b)$ sensitivity of the death rate w.r.t. the hospital saturation estimate.

Remark 2.1. In our numerical tests we will assume that there exists an *incubation phase* $[0, n_0)$ during which transmission and hospitalization do not occur. Additionally, in the *infectious phase* for which $b \in [n_0, B]$, we assume that $e(a, b) = 1$ and that $\nu(a, b)$, $\eta(a, b)$ and $\gamma(a, b)$ have constant values denoted by $\bar{\nu}(a)$, $\bar{\eta}(a)$ and $\bar{\gamma}(a)$ resp. Then, the coefficients take the form:

$$(2.5)$$
$$e(a, b) = \mathbb{1}_{[n_0, B]}(b), \quad \nu(a, b) = \mathbb{1}_{[n_0, B]}(b) \bar{\nu}(a), \quad \eta(a, b) = \mathbb{1}_{[n_0, B]}(b) \bar{\eta}(a), \quad \text{and} \quad \gamma(a, b) = \mathbb{1}_{[n_0, B]}(b) \bar{\gamma}(a).$$

Consequently, $h(t, a, b) = 0$ whenever $b \in [0, n_0]$ and

$$(2.6) \qquad Z(t) = \int_0^A \int_{n_0}^B z(t, a, b) \mathrm{d}b \mathrm{d}a.$$

So, in this case, the model has four parameters for each value of a: $\delta(a)$, $\bar{\nu}(a)$, $\bar{\eta}(a)$ and $\bar{\gamma}(a)$.

2.2. Optimal control problem. We next introduce the cost function to be minimized. Given the horizon T, the death toll for age a is equal to the difference between the initial and final population for this age:

$$(2.7) \qquad \begin{aligned} D(a) := \quad & y(0, a) + \bar{y}(0, a) + \int_0^B z(0, a, b) + h(0, a, b) \mathrm{d}b \\ & - \left(y(T, a) + \bar{y}(T, a) + \int_0^B (z(T, a, b) + h(T, a, b)) \, \mathrm{d}b \right). \end{aligned}$$

The total death toll is $D_T := \int_0^A D(a) \mathrm{d}a$. Denote by

$$(2.8) \qquad c(u) := \int_0^T \int_0^A c_e(a) u(t, a) \mathrm{d}a \mathrm{d}t$$

the weighted confinement cost, with weights $c_e(a) \geq 0$. The cost to be minimized is a combination of several costs. Its expression is

$$(2.9) \qquad J(M, u, D_T) := p_M M + p_u c(u) + p_D D_T,$$

where the penalty coefficients p_M, p_u and p_D are nonnegative. Here M is an additional variable, called the peak value, subject to

$$(2.10) \qquad H(t) \leq M, \quad \forall t \in [0, T].$$

Indeed, if p_M is nonzero, the optimal M will be the minimal hospital capacity needed to take care in an optimal way of all patients, which justifies the peak value terminology. The ability to minimize this amount has been recognized of paramount importance.

The second term in the cost function involves the weighted confinement effort. The associated economic cost is difficult to evaluate; this question being out of the scope of the present paper, we simply take it linear with somewhat arbitrary weights, having in mind the obvious fact that the confinement of retired people has a smaller economic impact than for the rest of the population, although psychological aspects should also be taken into account.

Finally remember that the death rate depends on the saturation of hospital capacities; so the confinement has an impact on the death toll too (assuming the coefficient γ to be nonzero and the hospital capacity to be saturated).

In addition, in our model we include confinement duration constraints for each age, more precisely:

$$(2.11) \qquad \int_0^T u(t, a) \mathrm{d}t \leq M(a), \quad \text{for a.a. } a \in (0, A).$$

3. Discretization of the control problem. In this section we present the discrete model associated to the previous formulation. The main objective of this section is to explain how to write the discrete problem so that an optimal control solver (based on an ODE model) can address it (See appendix A for more details).

For the age $a \in [0, A]$ we take discrete values $\{1, \ldots, n_a\}$. As both time unit, time step and discretization of the infection age we choose one day.

The horizon of the problem will be $N \in \mathbb{N}$. Let n_0, n_b, with $n_0 < n_b$ be the duration of the incubation phase and of the infection, respectively.

The time index is $k \in \mathbb{N}$, starting from value 0. The infection age is $j \in \{1, \ldots, n_b\}$. The initial conditions are

$$
\begin{aligned}
y_a^0 &= y_0(a) & a = 1, \ldots, n_a, \\
z_{a,j}^0 &= z_0(a, j) & a = 1, \ldots, n_a, \ j = 1, \ldots, n_b, \\
h_{a,j}^0 &= h_0(a, j) & a = 1, \ldots, n_a, \ j = 1, \ldots, n_b.
\end{aligned}
\tag{3.1}
$$

The global transmission factor and the size of the hospitalized population are respectively

$$
Z^k = \sum_{a=1}^{n_a} \sum_{j=1}^{n_b} e_{a,j} z_{a,j}^k, \qquad H^k = \sum_{a=1}^{n_a} \sum_{j=1}^{n_b} h_{a,j}^k.
\tag{3.2}
$$

Also $E^k \in [0, 1]$ is a discrete counterpart to the hospital saturation estimate $E(t)$, given by

$$
E^k := \frac{(H^k - C)_+}{H^k + C}.
\tag{3.3}
$$

Willing to keep conservation laws, we may write the dynamics as follows: for all $a = 1, \ldots, n_a$,

$$
\begin{aligned}
y_a^{k+1} &= \left(1 - \delta_a(1 - u_a^k)Z^k\right) y_a^k \\
z_{a,1}^{k+1} &= \delta_a(1 - u_a^k)Z^k y_a^k \\
z_{a,j}^{k+1} &= (1 - \nu_{a,j-1})z_{a,j-1}^k & j = 2, \ldots, n_b \\
h_{a,1}^{k+1} &= 0 \\
h_{a,j}^{k+1} &= \nu_{a,j-1}z_{a,j-1}^k + (1 - \eta_{a,j-1} - \gamma_{a,j-1}E^k)h_{a,j-1}^k & j = 2, \ldots, n_b \\
\bar{y}_a^{k+1} &= \bar{y}_a^k + z_{a,n_b}^k + h_{a,n_b}^k.
\end{aligned}
\tag{3.4}
$$

The death toll per age and day, and per day, are resp.

$$
D_a^k := \sum_{j=1}^{n_b - 1} (\eta_{a,j} + \gamma_{a,j}E^k)h_{a,j}^k; \quad D^k := \sum_{a=1}^{n_a} D_a^k.
\tag{3.5}
$$

The parameters of the dynamics are:

$$
\begin{aligned}
\delta_a&: & \text{transmission rate} \\
\nu_{a,j}&: & \text{hospitalization rate} \\
\eta_{a,j}&: & \text{minimal death rate} \\
\gamma_{a,j}&: & \text{sensitivity of the death rate w.r.t. hospital saturation.}
\end{aligned}
$$

Remark 3.1. The state variable y_a^k remains nonnegative iff

$$
\delta_a(1 - u_a^k)Z^k \le 1.
\tag{3.6}
$$

In the case of normalized (i.e., summing to 1) initial population, we have that $Z^k \le 1$. So, the previous nonnegativity condition holds whenever $\delta_a \le 1$. In our numerical test, it may happen that $\delta_a \ge 1$, but y_a^k always remained positive.

On the other hand, in our tests $\eta_{a,j} + \gamma_{a,j} \le 1$. Since $E^k \le 1$ by definition, it follows that $h_{a,j}^k$ always remains nonnegative.

We have the following constraints: maximal achievable confinement, peak value for the hospital capacity, and cumulated confinement effort. Namely, for some $u_a^M \in [0,1]$, $M \geq 0$, and $M_a \geq 0$:

$$
\begin{aligned}
& 0 \leq u_a^k \leq u_a^M, \quad k = 0, \ldots, N, \ a \in \{1, n_a\} \\
& H^k \leq M, \quad k = 0, \ldots, N, \\
& \textstyle\sum_{k=0}^{N-1} u_a^k \leq M_a, \ a \in \{1, n_a\}.
\end{aligned}
$$

(3.7)

The expression of the cost is

(3.8) $$ J_d(M, u, D_T) := p_M M + p_u c_u + p_D D_T, $$

with c_u weighted confinement cost, D_T total death toll:

(3.9) $$ c_u := \sum_{k=0}^{N-1} \sum_{a=1}^{n_a} c_{e,a} u_a^k; \quad D_T = \sum_{k=0}^{N-1} D^k, $$

and p_u, p_M and p_D are nonnegative penalty coefficients.

Due to the incubation phase, the value of control for the last n_0 days has no influence on the hospitalizations and therefore, neither on the death toll nor on the hospital capacity constraint.

4. Estimation of coefficients and initial state.

4.1. Incubation and infectious phases. As we mentioned in Remark 2.1, for our numerical tests we assume that there exists an incubation phase (days 1 to $n_0 - 1$ of infection) where no transmission and hospitalizations occur. So,

(4.1) $e_{a,j} = 0$ for $1 \leq j \leq n_0 - 1$, $e_{a,j} = 1$ for $n_0 \leq j \leq n_b$, and $\nu_{a,j} = 0$ for $1 \leq j \leq n_0 - 1$.

The above assumptions together with the initial condition $h_{a,1}^k = 0$ implies that

(4.2) $$ Z_a^k = \sum_{a=1}^{n_a} \sum_{j=n_0}^{n_b} z_{a,j}^k, \quad \text{and} \quad h_{a,j}^k = 0, \quad \text{whenever} \ 1 \leq j \leq n_0. $$

During the infectious phase (j varying from $n_0 + 1$ to n_b) we may assume that the coefficients $\nu_{a,j}$, for $j \geq n_0$ $\eta_{a,j}$ and $\gamma_{a,j}$, for $j \geq n_0 + 1$, do not depend on j and therefore are equal to some positive constants $\bar{\nu}_a$, $\bar{\eta}_a$, $\bar{\gamma}_a$.

We assume that we know for each age the total proportion $\hat{\nu}$ of transfers from z to h, that satisfies

(4.3) $\hat{\nu}_a = \bar{\nu}_a + \bar{\nu}_a(1 - \bar{\nu}_a) + \cdots + \bar{\nu}_a(1 - \bar{\nu}_a)^{n_b - n_0 - 1} = \bar{\nu}_a \sum_{j=0}^{n_b - n_0 - 1} (1 - \bar{\nu}_a)^j = 1 - (1 - \bar{\nu}_a)^{n_b - n_0}$,

so that we can estimate

(4.4) $$ \bar{\nu}_a = 1 - (1 - \hat{\nu}_a)^{1/(n_b - n_0)}. $$

Similarly, we may assume that we know for each age the proportion $\hat{\eta}_a$ of hospitalized people that will die, when the hospital capacity is not saturated. Since $h_{a,j}^k = 0$, for $j = 1, \ldots, n_0$, we have by similar arguments

(4.5) $$ \bar{\eta}_a = 1 - (1 - \hat{\eta}_a)^{1/(n_b - n_0 - 1)}. $$

As expected, $\bar{\nu}$ and $\bar{\eta}$ take values in $[0,1]$.

When the total hospital capacity is exceeded we assume that it can be measured so that $\bar{\gamma}_a$ itself can be evaluated.

4.2. Initial condition. Assuming that we know a growth factor g for the infected population given a time period of p time steps, then $e^{p\lambda} = g$, so that

$$(4.6) \qquad \lambda = (\log g)/p.$$

We assume that the initial time corresponds to an early phase of the epidemic, and therefore that the h and $\bar{y}$ variables are equal to zero, and (since the order of magnitude of y is almost constant) search for the z variable a solution of the form

$$(4.7) \qquad z_{a,j}^k = c_{a,j} e^{\lambda(k-j)}.$$

We can express $c_{a,j}$ as function of $c_{a,1}$:

$$(4.8) \qquad c_{a,j} = \begin{cases} c_{a,1} & \text{if } j \le n_0 \\ c_{a,1}(1 - \bar{\nu}_a)^{j-n_0} & \text{otherwise.} \end{cases}$$

So the total and infectious population of z for each age are resp.

$$(4.9) \qquad \bar{Z}_a^k := \sum_{j=1}^{n_b} z_{a,j}^k = c_{a,1}\bar{\chi}_a e^{\lambda k}; \quad \bar{\chi}_a := \left(\sum_{j=1}^{n_0-1} e^{-\lambda j} + \sum_{j=n_0}^{n_b} (1 - \bar{\nu}_a)^{b-n_0} e^{-\lambda j} \right)$$

$$(4.10) \qquad Z_a^k := \sum_{j=n_0}^{n_b} z_{a,j}^k = c_{a,1}\chi_a e^{\lambda k}; \quad \chi_a := \sum_{j=n_0}^{n_b} (1 - \bar{\nu}_a)^{j-n_0} e^{-\lambda j}.$$

In the absence of confinement at initial time, estimating $c_{a,1}$ with the previous relation, we deduce the value of δ from:

$$(4.11) \qquad c_{a,1} = z_{a,1}^1 = \delta_a Z^0 y_a^0.$$

As noted in remark 3.1, in order that y^k remains nonnegative, we need that

$$(4.12) \qquad \delta_a(1 - u_a^k)Z^k \le 1, \quad \text{for all } k = 0, \cdots, N.$$

Remark 4.1. We have seen that, since the initial time corresponds to an early stage of the epidemic, we can assume that $h^0 = 0$ and $\bar{y}^0 = 0$, and additionally, we can assume that there is no saturation of the hospital capacity. We could consider that we have statistical estimates of the "initial parameters" λ, $\hat{\nu}$, $\hat{\eta}$, y^0, Z^0.

Then we can deduce $\bar{\nu}$ and $\bar{\eta}$ from (4.4) and (4.5), respectively. The value of χ and $\bar{\chi}$ follow. We can deduce $c_{a,1}$ from (4.10) and the values of z^0 follow with (4.7). Finally, we deduce δ from (4.11).

For the estimate of the γ parameters, specific methods are needed, which are beyond the scope of this paper.

5. Numerical tests.

5.1. Choice of parameters and initial conditions. We consider the available data for France up to April 6, 2020. From Santé publique France and data.gouv.fr we obtain the following information:

Total confirmed cases:	74390
Total hospitalized cases:	54036
Total confirmed deaths:	8911

From these data we can estimate the values of $\hat{\nu}$ and $\hat{\eta}$ that represent the proportion of infected population in needs of hospitalization and the proportion of deaths with respect to the hospitalizations, respectively.

$$(5.1) \qquad \hat{\nu} = 0.726, \quad \text{and} \quad \hat{\eta} = 0.165.$$

Furthermore, from the above mentioned sources, we can observe that in the beginning of the epidemic, the growth factor per week is between 2.5 and 3, we can consider that

$$(5.2) \qquad \lambda \in [0.13, 0.15].$$

In the French demographic institute website INED we find detailed information on deaths in hospitals, taking into account the age of the population. We deduce that the strong population y_1 can be considered as the population up to 59 years and the weak population y_2 as the population over 60 years. In this case we have

	Total population	%	Total confirmed deaths	%
strong	49213491	73.4	416	6.5
weak	17850240	26.6	6021	93.5
total	67063703	100	6037	100

Since they only consider deaths in hospitals, the number of deaths in the two previous tables is not the same. From the last table we can calculate the relative mortality rate as

$$(5.3) \qquad \eta_1^r = \frac{416}{49213491} = 8.45e - 6, \quad \text{and} \quad \eta_2^r = \frac{6021}{17850240} = 3.4e - 4,$$

which implies that the ratio of this rate is

$$(5.4) \qquad \frac{\eta_1^r}{\eta_2^r} = 0.0248.$$

In order to obtain the values of $\hat{\eta}_1$ and $\hat{\eta}_2$, we assume that ratio is the same, i.e. $\hat{\eta}_1/\hat{\eta}_2 = 0.0248$. Normalizing the population, the global death rate $\hat{\eta}$ is equal to

$$(5.5) \qquad \hat{\eta} = \hat{\eta}_1 y_1 + \hat{\eta}_2 y_2 = (0.0248\hat{\eta}_2)0.734 + \hat{\eta}_2 0.266.$$

By (5.1) we obtain,

$$(5.6) \qquad \hat{\eta}_1 = 0.014, \quad \hat{\eta}_2 = 0.58.$$

Since we do not have detailed information about the age of hospitalized people, we assume:

$$(5.7) \qquad \hat{\nu} = \hat{\nu}_2 = \hat{\nu}_1.$$

Finally, from data.gouv.fr we know the total of confirmed cases per day, but again we do not have the information by age. For this reason, and since in our formulation, the discount term in the first dynamics equation is depending on the sum of infected population (Z^k), we consider that if $\bar{Z}^0$ is the initial confirmed cases, then we can define the (normalized) initial confirmed cases by age, as

$$(5.8) \qquad \bar{Z}_1^0 = \bar{Z}^0 \frac{0.734}{y^0}, \qquad \bar{Z}_2^0 = \bar{Z}^0 \frac{0.266}{y^0},$$

where $y^0 = 67063703$ is the total population of France.

Since we have taken the same values of $\bar{\nu}$ for the two populations, and the initial infected population proportional to the total population, the values for δ are identical. The initial infected population is computed as it was explained, taking the data from data.gouv.fr.

In our numerical tests the first day is March 15, 2020. For France this was actually the first day of confinement, with 6633 confirmed cases. We consider the following parameters:

y_1^0	initial strong population	0.734
y_2^0	initial weak population	0.266
Z_1^0	initial infected strong population	$7.26e-5$
Z_2^0	initial infected weak population	$2.63e-5$
λ	growth coefficient given by (4.6)	0.13
$\hat{\nu}_1$	total proportion of transfers from z_1 to h_1	0.726
$\hat{\nu}_2$	total proportion of transfers from z_2 to h_2	0.726
$\hat{\eta}_1$	strong minimal death rate	0.014
$\hat{\eta}_2$	weak minimal death rate	0.58

We choose an incubation duration of 6 days, and a total infection duration as 14 days. From the above table, following the methodology described in Remark 4.1, we obtain:

$$
\begin{array}{lll}
\delta_1 = 1.656, & \bar{\nu}_1 = 0.149412, & \bar{\eta}_1 = 0.002012 \\
\delta_2 = 1.656, & \bar{\nu}_2 = 0.149412, & \bar{\eta}_2 = 0.116557
\end{array}
\tag{5.9}
$$

In addition, we assume:

$$
\bar{\gamma}_{a,j} = \bar{\eta}_{a,j}, \quad \forall a = 1, 2, \ 1 \leq j \leq n_b; \ \ C = 0.005.
\tag{5.10}
$$

5.2. Numerical results in the case of an age independent control. In these first four tests, we assume the control variables to be independent of the age a, i.e. $u_1^k = u_2^k$ for all k, that the upper bound for the control is $u_M = 0.75$, and that the confinement cost, given by (3.9), satisfies $c_{e,1} + c_{e,2} = 1$. We present different scenarios, taking different penalty terms in the objective function. We do not consider any cumulated control constraints.

In each of the figures below, we show four graphics. The first one (upper left), is the optimal confinement, the second one (upper right) is the cumulative number of deaths for the strong and weak population, the third one (bottom left) is the number of people in need of hospitalization per day and the last one (bottom right) shows the evolution of the susceptible strong and weak population.

Test 1: $p_M = 1$, $p_u = 1$, $p_D = 1$, Figure 1.
Test 2: $p_M = 1$, $p_u = 0.000001$, $p_D = 0$, Figure 2
Test 3: $p_M = 1e - 5$, $p_u = 0$, $p_D = 1$, Figure 3
Test 4: $p_M = 1$, $p_u = 0.0005$, $p_D = 1$, Figure 4

In the following table we show the death toll for strong, weak and total population, follow by $I(u) := \sum_{k=0}^{N} u^k$ and the peak value M.

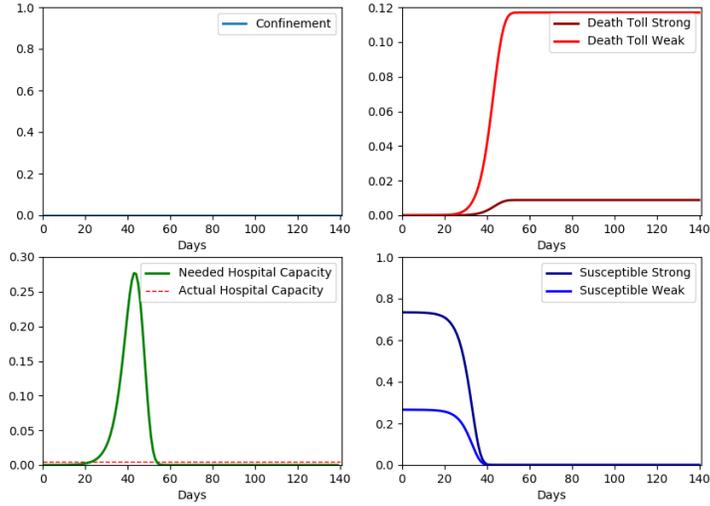| Test	$D_{1,T}$	$D_{2,T}$	D_T	$I(u)$	M
1	0.0088192	0.116966	0.1257852	0	0.27665
2	0.00816993	0.111278	0.11944793	54.49	0.0698543
3	0.00660044	0.0906906	0.09729104	102.375	0.0693109
4	0.00688871	0.094924	0.10181271	65.9702	0.0715698

Fig. 1. *Test 1:* $p_M = 1$, $p_u = 1$, $p_D = 1$. *The optimal control is 0. Without confinement, the death toll is greater than* 10% *and the peak value is very large, as expected.*
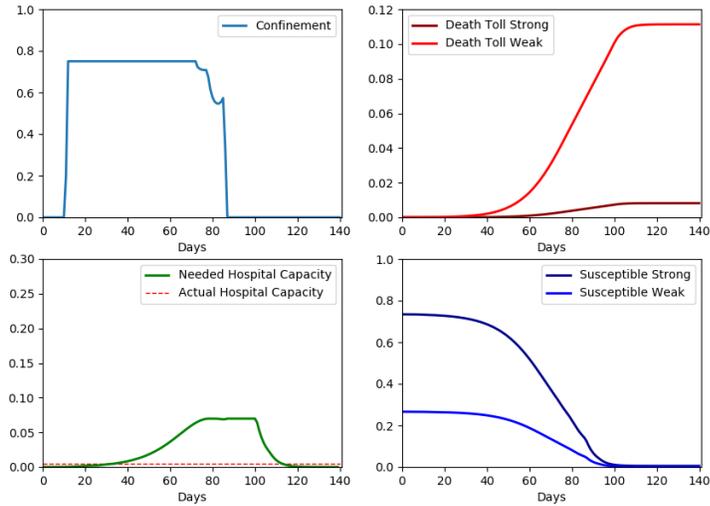

Fig. 2. *Test 2:* $p_M = 1$, $p_u = 0.000001$, $p_D = 0$. *In this case the main objective is to reduce the peak value. We can observe the delay, between the confinement decision and the response in the hospitalized population. Indeed, when the confinement stops, the size of hospitalized population remains constant for several days. The death toll is greater than* 10%.
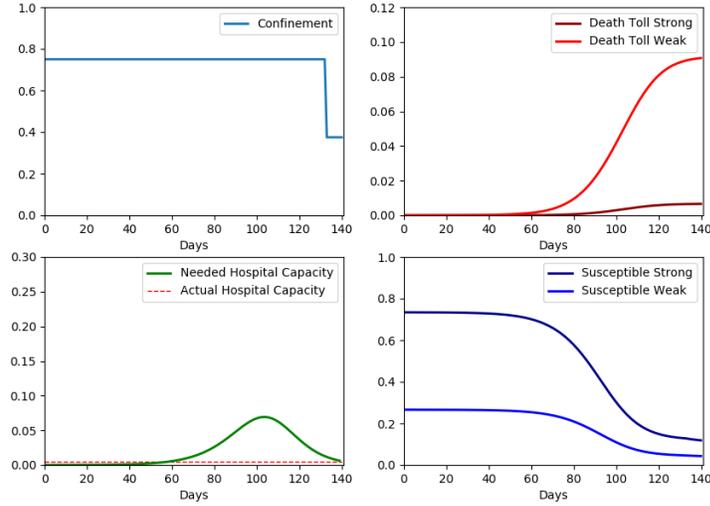
FIG. 3. *Test 3:* $p_M = 1e - 5$, $p_u = 0$, $p_D = 1$. *The main objective is to reduce the death toll, without considering the confinement cost. Therefore, the optimal policy assumes the maximum confinement value until the last days. Since the saturation of hospital capacity impacts in the death rate, as a consequence the peak value is reduced.*
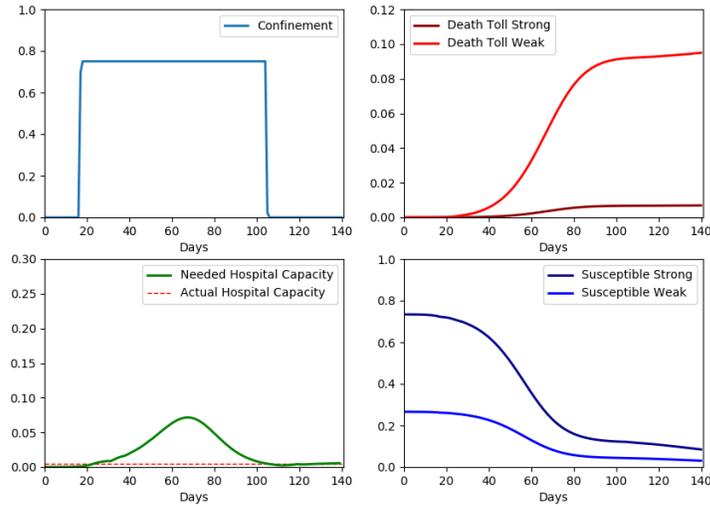

FIG. 4. *Test 4:* $p_M = 1$, $p_u = 0.0005$, $p_D = 1$. *In this case we want to reduce the peak value and the death toll, taking into account the economical impact.*

Discussion of the numerical results for an age independent control:

To analyze and compare the complete behavior of the epidemic when different confinement policies are applied, we assume that the maximum for this control is 0.75. Of course, this is the reason why in the tests we obtain that almost the entire population is immunized in the period considered. On the contrary, at the moment in France and in many countries around the world, containment measures are stronger and after more than a month of confinement we are far from the end of the epidemic. If these policies continue, the time needed to reach this level of immunization will be even longer.

From the above tests we can observe that if there is no confinement, which is the case of Test 1, then the health system will collapse and the death toll will be very large, as expected.

It is interesting to compare Test 2 and Test 3. In Test 3, the main objective is to reduce the death toll, and we observe that in comparison with Test 2, we obtain more or less the same peak,

some days after, but with a small death toll. The difference between them, is that in the first case, when the hospitalized population attains its maximum, the confinement stops, whereas in Test 3, the confinement is maximal until the end. That means that reducing the peak is not enough to reduce the final number of deaths.

Finally, as expected, in Test 4 since the control penalty is greater than those previously considered, the confinement is reduced, obtaining a greater peak value and death toll, but even though the death toll is smaller than in Test 2. We can also observe that in Test 4 since the optimal policy is 0 during the first 20 days, the peak value is obtained several weeks before than in Test 2 and Test 3.

As we mentioned in the introduction, it is very difficult to weigh economic losses against the number of lives lost.

5.3. Numerical results with age dependent controls. Taking into account these results, it is clear that most deaths occur in the weak population, so it is natural to address the issue of considering two different levels of confinement, depending on the age of the population. We take 0.75 as upper bound for the two confinement controls. In this case the cost function is given by (3.9). For the numerical tests, we assume that $c_{e,1} = 0.734$ and $c_{e,2} = 0.133$. With this coefficients we take into account the proportion of the population that is weak and strong and we also assume that confinement for the elderly has less impact on the economy than confinement for the young. We use the same initial conditions and parameters as in the previous tests. For Test 5 and Test 6 we do not consider any cumulated control constraints.

Test 5: $p_M = 1$, $p_u = 0.000001$, $p_D = 0$, Figure 5
Test 6: $p_M = 1$, $p_u = 0.0005$, $p_D = 1$, Figure 6

In the last test, we take the same values of Test 6, but we include cumulated control constraints (3.7) with $M_1 = 25$ and $M_2 = 45$.

Test 7: $p_M = 1$, $p_u = 0.0005$, $p_D = 1$, Figure 7

In the following table we summarized the results of the tests with differentiated confinement:

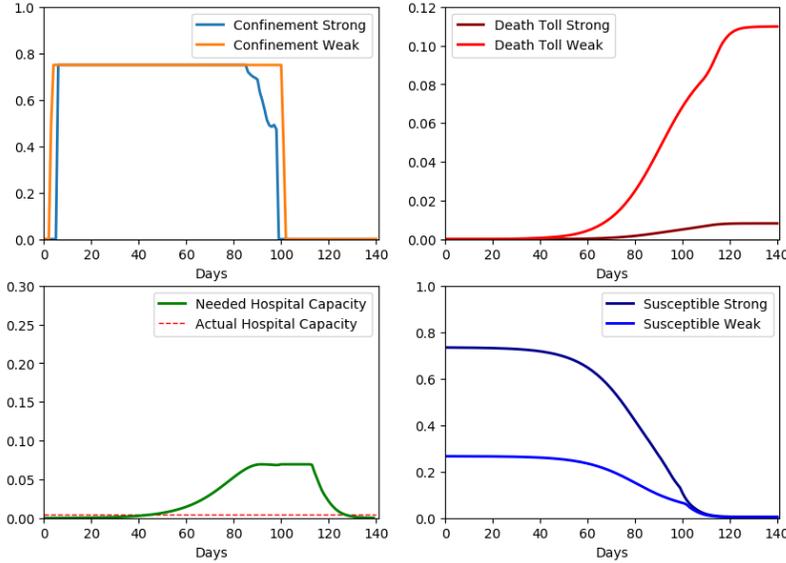| Test	$D_{1,T}$	$D_{2,T}$	D_T	$I(u_1)$	$I(u_2)$	M
5	0.00817212	0.109833	0.11800512	67.7524	73.6568	0.0693917
6	0.00824508	0.0983062	0.10655128	35.8425	62.25	0.0729996
7	0.00837618	0.10037	0.10874618	25	45	0.0805676

FIG. 5. *Test 5:* $p_M = 1$, $p_u = 0.000001$, $p_D = 0$. *Since hospitalizations rates $\bar{\nu}_1$ and $\bar{\nu}_2$ are assumed equals, to reduce the amount of hospitalizations that occur at the same time, it makes sense that both populations are confined for almost the same period. In this case, once the peak is reached, the confinement stops. We still observe a large number of deaths, greater than 10%.*
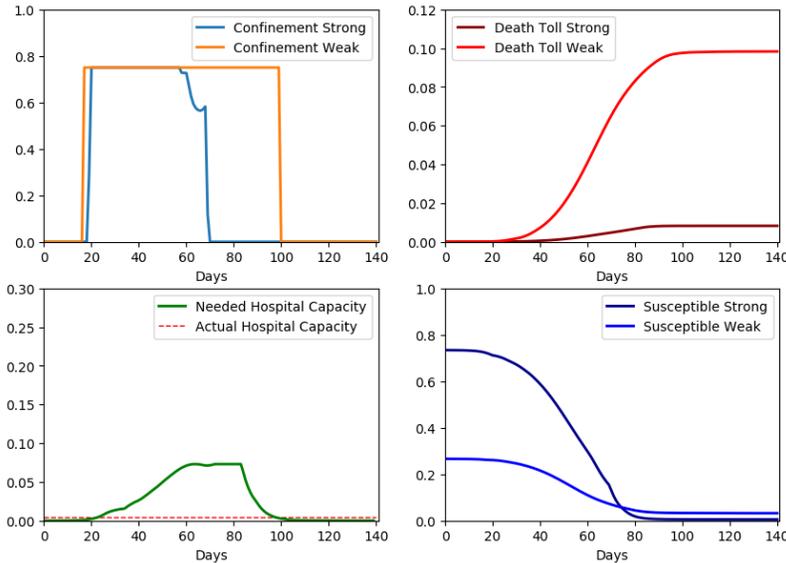

FIG. 6. *Test 6:* $p_M = 1$, $p_u = 0.0005$, $p_D = 1$. *As in Test 4, the aim is to balance between reducing peak value, death toll and confinement cost. We observe an important difference between optimal confinement for strong and weak populations. Of course, this occurs due to the difference between the mortality rates for both populations.*

Discussion of the numerical results for age dependent controls:

We analyze the difference between taking a confinement policy dependent or independent of the age. Test 5 is similar to Test 2, in both cases we achieve to reduce the peak value, but we obtain a very high death toll. In Test 5, there is no significant difference between confinement for the strong and weak population, and the behavior is similar to Test 2, once the peak is reached, the confinement stops (taking into account the delay effects).

The comparison between Test 4 and Test 6 is maybe the most interesting one. With similar penalty terms, we obtain almost the same death toll and peak value. The main difference is that the
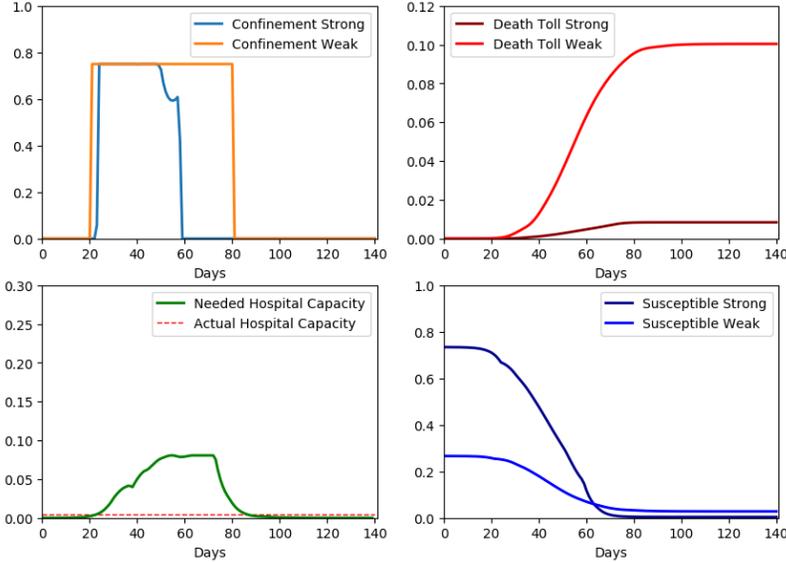
FIG. 7. *Test 7:* $p_M = 1$, $p_u = 0.0005$, $p_D = 1$, *with the confinement lengths $M_1 = 25$ and $M_2 = 45$. The objective of this test is to show how the optimal policy changes when we add constraints in the confinement lengths. Taking the delay into account, we see that the end of confinement of the strong population corresponds to the end of the hospitalization peak.*

strong population has a considerably shorter confinement in Test 6, while the weak population has almost the same as in Test 4. This confirms that performing differentiated confinement is perhaps the best response when seeking to minimize the economic impact of confinement, while at the same time reducing the number of deaths.

From the available data, it is evident that the majority of deaths belong to the weakest group, so it is natural to think that if the objective is to reduce the number of final deaths, the confinement for this group should be larger than for the other. When the sole objective is to reduce the peak, we do not observe this behavior because it is more important to reduce the number of infected people than the total number of deaths. Since in our model we take into account the saturation of the hospital system in the mortality rate of hospitalized people, when we minimize the number of deaths, we also obtain a reduction in the peak.

5.4. Two cases of specific interest. As we mentioned in the previous discussions, the objective of the previous tests was to analyze different scenarios that show that our numerical methods work well and that all the results obtained are consistent with our model. Now we will try to analyze cases closer to the current data.

5.4.1. A quasi-periodic control. In order to show an example which may be closer to the actual available data so far in France, we consider the case when after 15 March, the transmission rate δ is reduced to the half. This can be interpreted as a result of a change in the population behavior, in addition to the confinement policy itself, such as washing hands, social distancing, and the use of masks. We allow the confinement policy to take values between 0 and 0.9. We choose $p_M = 12$, $p_u = 0.005$, $p_D = 1$ and assume that the hospital capacity is $C = 0.01$. See figure 8.

After several weeks of confinement, hospitalizations per day begin to decrease. As expected, since the total number of immunized people is still too small, the epidemic is far from the end and when the confinement policy is relaxed, a rebound can be observed. This quasi-periodic feature of confinement could be a good strategy not to overwhelm the health system.

Given that the containment policy is very strong, after 140 days we did not observe great changes in the number of immunized people. However, the death toll is close to 0.0011, which is equivalent to approximately 73,000 people. Again, the reason for this large number of deaths is due to the hospitalization and mortality rates considered.
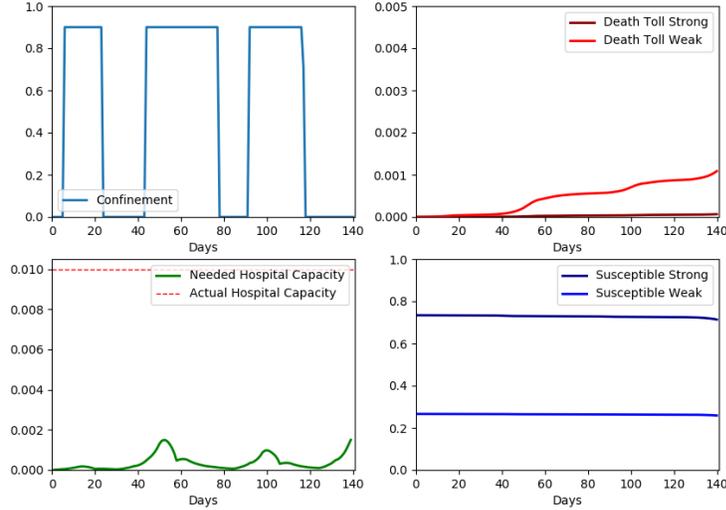
FIG. 8. *Quasi-periodic case:* $p_M = 12$, $p_u = 0.005$, $p_D = 1$, $C = 0.01$.

5.4.2. Long term behavior. In this case we take the parameters values of the above test, in particular the transmission rate $\delta = 0.825$ and the control takes values in $[0, 0.9]$. The penalty terms are $p_M = 12$, $p_u = 0.005$, $p_D = 1$, which implies that the main objective is to reduce the peak value, taking into account the death toll, but the costs associated with confinement are quite high. In this test we increase the horizon to 260 days.

We observe that some days after the peak of hospitalizations is achieved, the control is very high, and then it begins to slowly decrease to zero, which is when the epidemic is close to ending, see figure 9. When the situation of the country do not allow to take a long and strong confinement, this may be the more accurate strategy. Of course, due to penalties considered, the death toll is still too high.
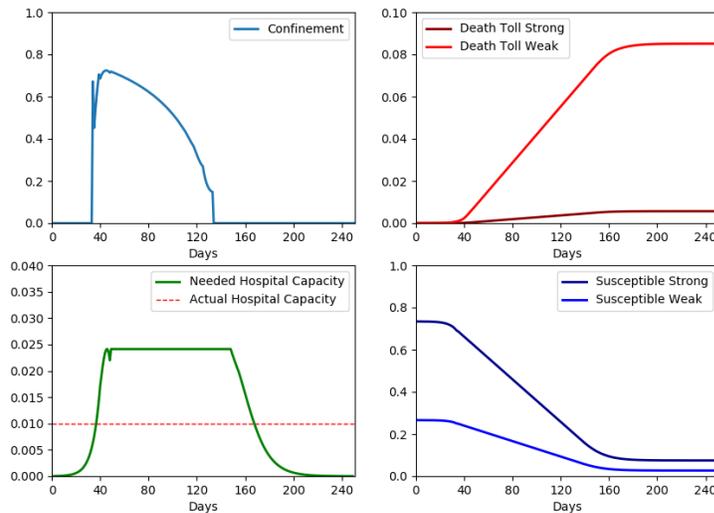

FIG. 9. *Long-term behavior:* $p_M = 12$, $p_u = 0.005$, $p_D = 1$, $C = 0.01$, $T = 260$.

5.5. Comments on the numerical results. A critical question is to check if data are available for evaluating the coefficients. Exchanges on the model with epidemic experts would be of

course very useful.

We want to point out that all the results obtained in the tests are illustrative. It is well known that the available data on confirmed cases could be far from the actual number, due to asymptomatic cases. A higher number of infected people would imply that hospitalization and mortality rates are lower and, therefore, the death toll as well as the maximum number of hospitalizations will be lower. In fact, we took the available data on March 15, 2020, when the death toll was almost 16%. This is why, in our tests, when the epidemic is almost over, the number of deaths is so high.

In addition, the other parameter that influences the number of deaths and the peak value, is the hospitalization rate $\bar{\nu}$, which at the beginning was close to 70%. Given that there are apparently many cases that are not serious, at a later stage of the epidemic when there is more demand in hospitals, some of these cases could be treated at home, obtaining a lower rate. Furthermore, we have no information on the length of hospitalization.

6. Conclusion.

6.1. Possible extensions.
We want to highlight several possible extensions or modifications of our model. Of course, such changes require further information, but we may have more accurate data in the near future. For instance, if a vaccine is available we can introduce a new control $v(t,a)$ which is the per capita rate of vaccination. The resulting modification in the dynamics would be:

$$(6.1) \qquad \begin{aligned} \dot{y}(t,a) &= -\delta(a)(1-u(t,a))Z(t)y(t,a) - v(t,a)y(t,a), \\ \dot{\bar{y}}(t,a) &= z(t,a,B) + h(t,a,B) + v(t,a)y(t,a). \end{aligned}$$

It is natural to think that v depends on the age of the population, and the limitation of M_V vaccines per day would read as

$$(6.2) \qquad \int_0^A v(t,a)y(t,a)\mathrm{d}a \le M_V, \quad \text{for a.a. } t \in (0,T).$$

Here we neglect any delay for the vaccination to become effective. Also, if data was available, it may be useful to split the infected population between the confirmed cases (taking advantage of a testing policy) and non-confirmed (among then some asymptomatic cases). Keeping the variable z for the non-confirmed cases, we could introduce a new state $w(t,a,b)$ representing the confirmed cases. Assuming the latter to be isolated, only the z population contributes to the global transmission factor $Z(t)$. We obtain the following dynamics:

$$(6.3) \qquad \begin{cases} \dot{y}(t,a) = -\delta(a)(1-u(t,a))Z(t)y(t,a) \\ (z_t + z_b)(t,a,b) = -\mu(t,a,b)z(t,a,b) - \nu(a,b)z(t,a,b) \\ (w_t + w_b)(t,a,b) = \mu(t,a,b)z(t,a,b) - \nu(a,b)w(t,a,b) \\ (h_t + h_b)(t,a,b) = \nu(a,b)(z(t,a,b) + w(t,a,b)) - (\eta(a,b) + \gamma(a,b)E(t))\,h(t,a,b) \\ \dot{\bar{y}}(t,a) = z(t,a,B) + w(t,a,B) + h(t,a,B). \end{cases}$$

Here $\mu(t,a,b) \ge 0$ is another control, which represents the detection rate. A possible associated cost for the testing policy would be, for some positive coefficients c_1 and c_2:

$$(6.4) \qquad \int_0^T \int_0^A \int_0^B \mu(t,a,b)(c_1 z(t,a,b) + c_2 y(t,a,b))\mathrm{d}b\mathrm{d}a\mathrm{d}t.$$

In addition, we can consider the hospitalization age to be a significant parameter. Denoting it by $c \in [0,C]$, and assuming that h depends on (t,a,c), we get the following dynamics for the hospitalized population:

$$(6.5) \qquad \begin{cases} h(t,a,0) = \int_0^b \nu(a,b)(z(t,a,b) + w(t,a,b))\mathrm{d}b \\ (h_t + h_c)(t,a,c) = - (\eta(a,c) + \gamma(a,c)E(t))\,h(t,a,c), \quad c \in [0,C]. \end{cases}$$

It is also possible to introduce a recovery rate, that could depend on the hospitalization age, and would contribute to the dynamics of $\bar{y}$.

As we can see there are many possible adaptations of our model. On the other hand, one should avoid to excessively increase the number of states and parameters. A model with too many states is difficult to optimize; one with too many parameters makes identifiability uneasy.

6.2. Final remarks. Population models with infection age need (after infection age discretization) many states. Nevertheless, we could obtain a numerical convergence of our solvers in various cases. Sometimes we needed to solve the problem first with a small horizon (e.g. 3 months) and then to increase progressively the horizon, in order to obtain the convergence.

The main advantage of our model is the age-structure. Despite the limited accuracy of the parameters and the obtained data, in our tests we recovered the particular behavior of the epidemic with respect to the incubation phase and almost fixed duration, at least for non-severe cases. We can see the delay in the response when a containment policy is taken, which is crucial in this type of disease where the infected population grows exponentially.

Our model is quite flexible and allows to take for instance variable coefficients during or after the incubation phase. In future studies one should of course take advantages of clinical studies in order to take into account a more accurate dynamical model, including the response to drugs, as well as the extensions just previously discussed.

It is of interest to observe how is the containment variable in the 'long term' case of subsection 5.4.2. The containment is initially absent, then close to its maximal value at a time close to the saturation of the hospital capacity, then smoothly decreases and goes back to 0. This (happily) seems to be in accordance with the expected social behavior (to accept more easily strict measures at the beginning).

Besides, this study raises interesting questions related to the analysis of the optimality conditions, especially since the model involves state constraints. For this purpose the techniques of the recent paper [3] could be useful.

Appendix A. General time steps.

We used the open-source optimal control solver Bocop ([2], http://www.bocop.org). Since our software solves optimal control problems described by an ODE, the present problem does not enter in its format. However, note that we can rewrite the discrete system (3.4) in this way:

$$
\begin{aligned}
y_a^{k+1} - y_a^k &= -\delta_a(1-u_a^k)Z^k y_a^k \\
z_{a,1}^{k+1} - z_{a,1}^k &= -z_{a,1}^k + \delta_a(1-u_a^k)Z^k y_a^k \\
\text{(A.1)} \qquad z_{a,j}^{k+1} - z_{a,j}^k &= -z_{a,j}^k + (1-\nu_{a,j-1})z_{a,j-1}^k \\
h_{a,j}^{k+1} - h_{a,j}^k &= -h_{a,j}^k + \nu_{a,j-1}z_{a,j-1}^k + (1-\eta_{a,j-1}-\gamma_{a,j-1}E^k)h_{a,j-1}^k \\
\bar{y}_a^{k+1} - \bar{y}_a^k &= z_{a,n_b}^k + h_{a,n_b}^k.
\end{aligned}
$$

Taking advantage of the fact that the time step is 1, we see that the discrete system (3.4) can be interpreted as the explicit Euler discretization scheme, for dynamics corresponding to the r.h.s. of the above array.

While a time step of one day is natural, in the case of large horizons it may be useful to conduct experiments with time steps of several days. So, consider a time step (equal to the step for the infection age) $\Delta > 0$. We need $n_b' := n_b/\Delta$ to be integer. Clearly the following system is a consistent approximation of (2.1):

$$
\begin{aligned}
\frac{y_a^{k+1} - y_a^k}{\Delta} &= -\delta_a(1-u_a^k)Z^k y_a^k \\
\frac{z_{a,1}^{k+1} - z_{a,1}^k}{\Delta} &= \frac{-z_{a,1}^k}{\Delta} + \delta_a(1-u_a^k)Z^k y_a^k \\
\text{(A.2)} \qquad \frac{z_{a,j+1}^{k+1} - z_{a,j+1}^k}{\Delta} &= \frac{z_{a,j}^k - z_{a,j+1}^k}{\Delta} - \nu_{a,j-1}z_{a,j}^k \\
\frac{h_{a,j+1}^{k+1} - h_{a,j+1}^k}{\Delta} &= \frac{h_{a,j}^k - h_{a,j+1}^k}{\Delta} + \nu_{a,j}z_{a,j}^k - (\eta_{a,j}+\gamma_{a,j}E^k)h_{a,j}^k \\
\frac{\bar{y}_a^{k+1} - \bar{y}_a^k}{\Delta} &= z_{a,n_b'}^k + h_{a,n_b'}^k.
\end{aligned}
$$

In addition the above scheme is conservative, and can be interpreted as the explicit Euler discretization scheme, for dynamics corresponding to the r.h.s.

REFERENCES

[1] B. Ainseba, S. Aniţa, and M. Langlais. Optimal control for a nonlinear age-structured population dynamics model. *Electron. J. Differential Equations*, pages No. 28, 9, 2003.

[2] J. Bonnans, D. Giorgi, V. Grélard, B. Heymann, S. Maindrault, P. Martinon, O. Tissot, and J. Liu. Bocop - A collection of examples. Technical report, INRIA, 2017.

[3] J. F. Bonnans and J. Gianatti. Optimal control of state constrained age-structured problems. *SIAM Journal on Control and Optimization*, 58(4):2206–2235, 2020. https://doi.org/10.1137/19M1271257.

[4] M. Brokate. Pontryagin's principle for control problems in age-dependent population dynamics. *Journal of Mathematical Biology*, 23(1):75–101, Dec 1985.

[5] R. Djidjou-Demasse, Y. Michalakis, M. Choisy, M. T. Sofonea, and S. Alizon. Optimal COVID-19 epidemic control until vaccine deployment. https://www.medrxiv.org/content/10.1101/2020.04.02.20049189v2, 2020.

[6] R. Elie, E. Hubert, and G. Turinici. Contact rate epidemic control of covid-19: an equilibrium view. https://hal.archives-ouvertes.fr/hal-02545930, 2020.

[7] W. O. Kermack, A. G. McKendrick, and G. T. Walker. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 115(772):700–721, 1927. https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1927.0118.

[8] Q. Lin, S. Zhao, D. Gao, Y. Lou, S. Yang, S. S. Musa, M. H. Wang, Y. Cai, W. Wang, L. Yang, and D. He. A conceptual model for the coronavirus disease 2019 (covid-19) outbreak in Wuhan, China with individual reaction and governmental action. *International Journal of Infectious Diseases*, 93:211–216, April 2020. https://www.sciencedirect.com/science/article/pii/S120197122030117X.

[9] Z. Liu, P. Magal, O. Seydi, and G. Webb. Predicting the cumulative number of cases for the COVID-19 epidemic in China from early data. https://www.medrxiv.org/content/10.1101/2020.03.11.20034314v1, March 2020.

[10] Z. Liu, P. Magal, O. Seydi, and G. Webb. Understanding unreported cases in the COVID-19 epidemic outbreak in Wuhan, China, and the importance of major public health interventions. *Biology*, 9(3), March 2020. https://www.mdpi.com/2079-7737/9/3/50/.

[11] H. Maurer and M. do Rosario de Pinho. Optimal control of epidemiological SEIR models with L^1-objectives and control-state constraints. *Pac. J. Optim.*, 12(2):415–436, 2016.

[12] Q. Richard, S. Alizon, M. Choisy, M. T. Sofonea, and R. Djidjou-Demasse. Age-structured non-pharmaceutical interventions for optimal control of covid-19 epidemic. *medRxiv*, 2020. https://www.medrxiv.org/content/early/2020/08/18/2020.06.23.20138099.

[13] C. Silva, H. Maurer, and D. Torres. Optimal control of a tuberculosis model with state and control delays. *Math. Biosci. Eng.*, 14(1):321–337, 2017.