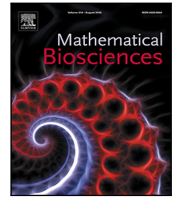




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Original Research Article

Piecewise-constant optimal control strategies for controlling the outbreak of COVID-19 in the Irish population

Lennon Ó Náraigh*, Áine Byrne

School of Mathematics and Statistics, University College Dublin, Belfield, Dublin 4, Ireland



ARTICLE INFO

Keywords:

SEIR model
Optimal control
Simulated annealing

ABSTRACT

We introduce a deterministic SEIR model and fit it to epidemiological data for the COVID-19 outbreak in Ireland. We couple the model to economic considerations — we formulate an optimal control problem in which the cost to the economy of the various non-pharmaceutical interventions is minimized, subject to hospital admissions never exceeding a threshold value corresponding to health-service capacity. Within the framework of the model, the optimal strategy of disease control is revealed to be one of disease suppression, rather than disease mitigation.

1. Introduction

Since late 2019, the COVID-19 respiratory disease originating from the SARS-CoV-2 virus [1] has caused a global pandemic, leading to over 9 million confirmed cases and over 660,000 confirmed deaths as of June 2020 [2]. As the epicentre of the pandemic has shifted from Asia to Europe and the Americas, governments across the world have taken steps to curb the spread of the disease. Due to the current lack of a vaccine for COVID-19, and the current very limited number of therapeutic agents available for treatment of COVID-19 [3], these efforts have taken the form of non-pharmaceutical interventions (NPIs).

Non-pharmaceutical interventions (NPIs) to curb the spread a pandemic like COVID-19 may take various forms, ranging from mass publicity campaigns to educate the public about proper hygiene, social distancing, testing, up to quarantining. The main purpose of such interventions is to reduce the effective reproduction number of the disease. Reducing the effective reproduction number serves either to mitigate the disease, or to suppress it entirely [4]. As per Ref. [4], we take ‘suppression’ to mean ‘driving infections to the (unstable) $I = 0$ fixed point’, that is, driving the number of infectious cases to zero. Therefore, for the present purposes, ‘suppression’ and ‘elimination’ are synonymous. In contrast, mitigation serves to make the spread of the disease in the population more manageable (e.g. preventing a breakdown in health services due to an excess of hospital admissions). Mitigation may also involve a slow build-up of the disease in the population, with a view to achieving ‘herd immunity’ [5–7]. In contrast, suppression aims to reduce the effective reproduction number of the disease so far below unity that the disease dies out in the population.

In this work, we address the tradeoffs between mitigation and suppression for a specific country case (Ireland). We use official epidemiological data to generate a standard SEIR-type model for the

outbreak of COVID-19 in Ireland. Thereafter, we carry out a counterfactual exercise and look at what would be the optimal strategy of NPIs to reduce the effective reproduction number of the disease while simultaneously minimizing the cost to the national economy. The basic methodology is a standard compartmental SEIR model with multiple additional compartments for fitting the model to the extant data; the model is deterministic and based on coupled ordinary-differential equations. Thereafter, the model is reformulated in the context of Optimal Control Theory — the model is solved in such a way as to minimize a cost function, while satisfying certain state constraints which correspond to the requirement that hospital admissions remain below capacity for the duration of the outbreak.

Optimal control theory has already found great use in mathematical epidemiology. It has been used to compute optimal vaccination strategies for various diseases lethal to humans [8–10], as well as optimal containment strategies for animal diseases [11,12]. These works use the Pontryagin Maximum Principle [13] to formulate semi-analytical conditions for the optimal control strategy. In cases with a linear cost function (including optimal stopping time problems [11,12]), the optimal control is revealed to be a so-called *bang–bang* control, where the control switches between a maximum and a minimum value during at a finite number of switching points.

Our work is a departure from these prior works, in that we recognize the importance of state constraints in a context where it may be desirable to strictly control the number of infectious cases, with a view to preserving the integrity of a health system. Including state constraints is challenging from the mathematical point of view, and often a numerical approach is pursued [14]. Our work follows in this vein.

* Corresponding author.

E-mail address: onaraigh@maths.ucd.ie (L.Ó. Náraigh).

<https://doi.org/10.1016/j.mbs.2020.108496>

Available online 16 October 2020

0025-5564/© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

In this work, we perform optimal control theory over a fixed time horizon of one year. The motivation is that this is a realistic time horizon over which to manage the spread of COVID-19 in population, before the advent of anticipated therapeutic interventions such as antiviral drugs and vaccines. This approach follows the recent work in Ref. [5]. This prior work is an important first step in the application of Optimal Control Theory to the COVID-19 outbreak. We build on this prior work by working with a linear cost function (Reference [5] uses a cost function that is quadratic in the intensity of the non-pharmaceutical interventions). A linear cost function may be more realistic from the point-of-view of economic modelling [15]. Ref. [5] involves modelling in the absence of state constraints; in the present article we include state constraints. This makes the mathematical analysis more involved, so a numerical optimization approach is favoured herein. Finally, we emphasize that although the model is calibrated for the Irish context, and the optimization performed in that context also, the methods are entirely generic, and will find use more broadly.

This article is organized as follows. In Section 2 we introduce the basic SEIR model and demonstrate how it is fit to epidemiological data for the COVID-19 outbreak in Ireland. In Section 3 we couple the model to economic considerations — we formulate an optimal control problem in which the cost to the economy of the various NPIs is minimized, subject to the satisfaction of certain key state constraints. Results are presented in Section 4, and conclusions are presented in Section 5.

2. Mathematical model

In this work, we develop an optimal control theory based on a deterministic compartmental model of epidemiology. Therefore, in this section we introduce an appropriate multi-compartment susceptible–exposed–infectious–removed (SEIR) model. The model is tailored to COVID-19, and is based on prior works [16,17]. Thus, the model includes specific compartments for susceptible individuals S , exposed (but not yet infectious) individuals E , and three separate infectious classes: infectious-pre-symptomatic (I_P), infectious-asymptomatic (I_A), and infectious-symptomatic (I_S) individuals. Removed cases (R) are considered also in a standard fashion. Finally, in order to fit the model to existing data, further compartments are introduced to the model to enable the recording of confirmed cases and deaths. In this way, we arrive at the following model. Time is measured in days.

$$\frac{dS}{dt} = -\frac{cS(q_P I_P + q_A I_A + q_S I_S)}{N}, \tag{1a}$$

$$\frac{dE}{dt} = \frac{cS(q_P I_P + q_A I_A + q_S I_S)}{N} - \frac{1}{\tau_E} E, \tag{1b}$$

$$\frac{dI_P}{dt} = \frac{1}{\tau_E} E - \frac{1}{\tau_{IP}} I_P, \tag{1c}$$

$$\frac{dI_A}{dt} = \frac{f}{\tau_{IP}} I_P - \frac{1}{\tau_I} I_A, \tag{1d}$$

$$\frac{dI_S}{dt} = \frac{1-f}{\tau_{IP}} I_P - \frac{1}{\tau_I} I_S, \tag{1e}$$

$$\frac{dR}{dt} = \frac{1}{\tau_I} I_A + \frac{1-g}{\tau_I} I_S. \tag{1f}$$

Additional, non-standard compartments are introduced to allow for comparison with recorded data:

$$\frac{d}{dt}(\text{Ailing}) = \frac{g}{\tau_I} I_S - \frac{1}{\tau_D}(\text{Ailing}), \tag{1g}$$

$$\frac{dD}{dt} = \frac{1}{\tau_D}(\text{Ailing}), \tag{1h}$$

$$\frac{d}{dt}(\text{Awaiting Test}) = \frac{1-f}{\tau_{IP}} I_P - \frac{1}{\tau_T}(\text{Awaiting Test}), \tag{1i}$$

$$\frac{dC}{dt} = \frac{1}{\tau_T}(\text{Awaiting Test}). \tag{1j}$$

Here, cases in the ‘Ailing’ compartment die and enter the D -compartment at a rate τ_D . Also, it is assumed that all individuals entering the

symptomatic compartment are tested — these cases enter the ‘Awaiting Test’ compartment, wait on average τ_T days for a test; results are subsequently reported and attributed to the C -compartment. Including deaths in the model is important, as it enables us to fit the model to a wider range of epidemiological data involving both confirmed cases and deaths, thereby improving the fit of the model. Also, it opens up the possibility of performing optimal control on the model in such a way as to control or even limit the number of deaths over the course of the epidemic.

We further elaborate on the enumerated constants. The quantity N is the total initial susceptible population. The coefficient c denotes the average number of contacts between a susceptible person and an infectious person, per day. Correspondingly, the coefficient q_P denotes the probability that an infected, pre-symptomatic person transmits the disease to a susceptible person during one of those contacts, and similarly for q_A and q_P . By allowing for different values between q_P , q_S , and q_A , we can account for the possibility that asymptomatic cases are less infectious than symptomatic cases. We can also allow for the effect of case isolation, which will reduce the value of q_S . The time τ_E (in days) denotes an incubation time when an exposed individual is not yet infectious, the time τ_{IP} (in days) denotes the time interval during which an exposed individual is infectious before symptom onset. The coefficient f denotes the fraction of pre-symptomatic cases that never show any symptoms of the disease. All infectious cases (whether asymptomatic or symptomatic) are assumed to remain infectious for a period τ_I days. It is assumed that all asymptomatic cases recover and are immune, while only $1 - g$ of the symptomatic cases recover; the remaining g become ailing and subsequently, die. Thus, g represents the long-term case fatality rate of the disease.

2.1. Model fitting

For definiteness, we fit the model (1) to epidemiological data for the COVID-19 outbreak in Ireland. The epidemiological data is obtained from official sources [18], and involves both confirmed cases of COVID-19, and deaths attributed to COVID-19. To fit the data to the model (1), we have minimized the L^2 norm,

$$\begin{aligned} \Delta = & \sum_{j=0}^n \left\{ \ln [C_{model}(t_j)] - \ln [C_{data}(t_j)] \right\}^2 \\ & + \sum_{j=0}^n \left\{ \ln \left(\frac{dC_{model}}{dt} \right)_{t_j} - \ln \left(\frac{dC_{data}}{dt} \right)_{t_j} \right\}^2 \\ & + \sum_{j=j_{D1}}^n \left\{ \ln [D_{model}(t_j)] - \ln [D_{data}(t_j)] \right\}^2 \\ & + \sum_{j=j_{D2}}^n \left\{ \ln \left(\frac{dD_{model}}{dt} \right)_{t_j} - \ln \left(\frac{dD_{data}}{dt} \right)_{t_j} \right\}^2. \end{aligned} \tag{2}$$

Here, t_j labels time in days. Here also, $n = 76$, corresponding to 77 days’ worth of data, starting with the first recorded case of COVID-19 in Ireland on February 29th 2020 ($t_j = 0$) and ending on May 17th 2020 ($t_j = 76$). The justification for choosing this time-period is that it includes the initial phase of the COVID-19 outbreak before disease-control measures were introduced (beginning 12th March 2020). The chosen time-period ends before the beginning of the phased lifting of these controls (May 18th 2020). Finally, j_{D1} denotes the day with the first recorded death and j_{D2} denotes the first day in the outbreak after which the number of new recorded deaths per day is strictly greater than zero.

We use simulated annealing to minimize Δ as a function of the model parameters. To reduce the complexity of the problem slightly, we re-write Eq. (1a) as:

$$\frac{dS}{dt} = -\beta_0 \frac{S}{N} (I_P + i_A I_A + i_S I_S); \tag{3}$$

a similar modification applies to Eq. (1b). Here, $\beta_0 = cq_P$, $i_A = q_A/q_P$, and $i_S = q_S/q_P$. As such, i_A and i_S should henceforth be understood as

Table 1

Effect of various interventions on the value of the basic reproductive number R_0 . The formula for R_0 is given in Eq. (8), with β_0 replaced by β . The given values of β correspond to β_1 and β_2 respectively in the optimization problem (6), the numerical values of which are given in Appendix A. Similarly, the value for the case-isolation parameter i_S is obtained from the optimization problem (6), with the details provided again in Appendix A.

Intervention	β	i_S	$R_0(\beta, i_S)$
Case isolation, no other interventions	1.4695	0.3405	2.6945
Case isolation, schools and universities are closed, mass gatherings are banned	1.1009	0.3405	2.0187
Case isolation, all non-essential services and industries closed, travel restrictions	0.3576	0.3405	0.6557

ratios of probabilities – e.g. i_A should be taken to mean the probability that an asymptomatic individual transmits the disease to a susceptible individual, divided by the probability that a pre-symptomatic individual transmits the disease to a susceptible individual. A similar interpretation should henceforth be applied to i_S .

At the same time, the optimization problem is rendered more complex, in several respects. First, the model is initialized at $t = -t_{offset}$, with

$$S(t = -t_{offset}) = N - 1 = (4.9 \times 10^6) - 1, \quad I_P(t = -t_{offset}) = 1, \quad (4)$$

and all other compartments equal to zero. Here, t_{offset} is an unknown parameter, In this context, the total population is assumed to be susceptible to the disease, and $I_P(t = -t_{offset}) = 1$ corresponds to ‘patient zero’.

Furthermore, β_0 changes over the course of the epidemic, corresponding to the implementation of disease controls [19]. To account for this, we replace β_0 in the model (1) by:

$$\beta_j = \begin{cases} \beta_0, & j < 13, \\ \beta_1, & 13 \leq j < 28, \\ \beta_2, & j \geq 28. \end{cases} \quad (5)$$

In the first time period ($j < 13$), no disease controls are implemented. In the second time period ($13 \leq j < 28$), all bars, schools and universities are closed, and mass gatherings are banned. Finally, in the third time period $j \geq 28$, all non-essential services and industries are closed down, and travel is restricted to within 2 km of a person’s house, with certain limited exceptions. Eq. (5) already contains some simplifications that do not correspond exactly to the actual course of the COVID-19 control efforts in Ireland – e.g. a four-day period (12th March 2020–15th March 2020, inclusive) where schools and universities were shut down and mass gatherings banned but bars remained open. Also, after May 5th 2020 ($j = 66$), the 2 km restriction was extended to 5 km. However, we ignore these details, as the main purpose of this article is not necessarily to track and predict the course of COVID-19 in Ireland, but rather to develop a simple, generic methodology to compute the optimum application of non-pharmaceutical interventions to control the spread of COVID-19. As such, we seek to solve the optimization problem for A_0 , where

$$A_0 = \min [A(t_{offset}, \beta_0, \beta_1, \beta_2, f, g, \tau_E, \tau_{IP}, \tau_I, \tau_D, i_A, i_S)], \quad (6)$$

which is a 13-dimensional optimization problem.

We solve the optimization problem (6) using simulated annealing. Bounds for the search space are estimated from the literature — these are reported in Appendix A, where we also provide the optimal parameter values corresponding to Eq. (6). Results are shown in Fig. 1. Qualitative agreement between the SEIR model (1) and the data can be seen in the figure. To demonstrate quantitatively the agreement between the model, the data, and prior characterizations of the epidemiology of COVID-19, we have computed the basic reproduction number of the model (1) using the maximum eigenvalue of the next-generation matrix [20], $R_0 = \max \text{spec}(FV^{-1})$. For the model (1), the

matrices F and V are given by:

$$F = \begin{pmatrix} 0 & \beta_0 & \beta_0 i_A & \beta_0 i_S \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad V = \begin{pmatrix} \tau_E^{-1} & 0 & 0 & 0 \\ -\tau_L^{-1} & \tau_{IP}^{-1} & 0 & 0 \\ 0 & -f\tau_{IP}^{-1} & \tau_I^{-1} & 0 \\ 0 & 0 & -(1-f)\tau_{IP}^{-1} & \tau_I^{-1} \end{pmatrix}, \quad (7)$$

hence

$$R_0(\beta_0, i_S) = \beta_0 \left(\tau_{IP} + i_A f \tau_I + f(1-f) i_S \frac{\tau_I^2}{\tau_{IP}} \right). \quad (8)$$

(the reason for including the explicit functional dependence on the parameters β_0 and i_S will become clear in what follows). We have used the parameters obtained from solving the optimization problem (6) to compute the value of R_0 . With case isolation (i.e. $i_S = 0.3405$, as per the solution of Eq. (6)), we have computed $R_0 = 2.6945$. Without case isolation (i.e. $i_S = 1$), we have computed $R_0 = 3.4495$. These results are consistent with an existing estimates of the basic reproduction number for COVID-19 [21]. Furthermore, by replacing β_0 in Eq. (8) with both β_1 and β_2 , we can estimate the effect of various measures to control the spread of the epidemic. This is done in Table 1 – the results are very close to other independent estimates of the same quantities [19]. These findings confirm the robustness of the basic mathematical model (1) and therefore provide a solid foundation on which to basic a theory of optimal control in the remaining sections.

Finally, we emphasize that the focus of this article is on developing a numerical framework for optimal control for epidemic outbreaks. Therefore, quantifying the uncertainty in the fit of the SEIR model (1) to the epidemiological data takes a back seat. However, in Appendix A, along with the optimal parameter values corresponding to Eq. (6), we also provide the confidence intervals for the same, as well as confidence intervals for the basic reproductive number R_0 and the effective reproductive numbers in Table 1.

3. Optimal control theory

In this Section we model the imposition of non-pharmaceutical interventions (NPIs) by replacing the constant value β_0 in Eq. (1) with a piecewise-constant function denoted by $\beta(t)$. We solve the equations from the initial time $t = -t_{offset}$ (initial conditions (4)) out to a fixed final time $t = T$. Specifically, we take:

$$\beta(t) = \begin{cases} \beta_0, & t \in (-t_{offset}, t_0], \\ \beta_0 [1 - u(t)], & t \in (t_0, T]. \end{cases} \quad (9a)$$

Here, $u(t)$ is a second piecewise-constant functions. In this way, the model (1) is solved with no controls for $t \leq t_0$; controls are introduced at time t_0 . The time t_0 is fixed. For consistency, the following bounds apply to $u(t)$:

$$0 \leq u(t) \leq 1, \quad t \in [t_0, T]. \quad (9b)$$

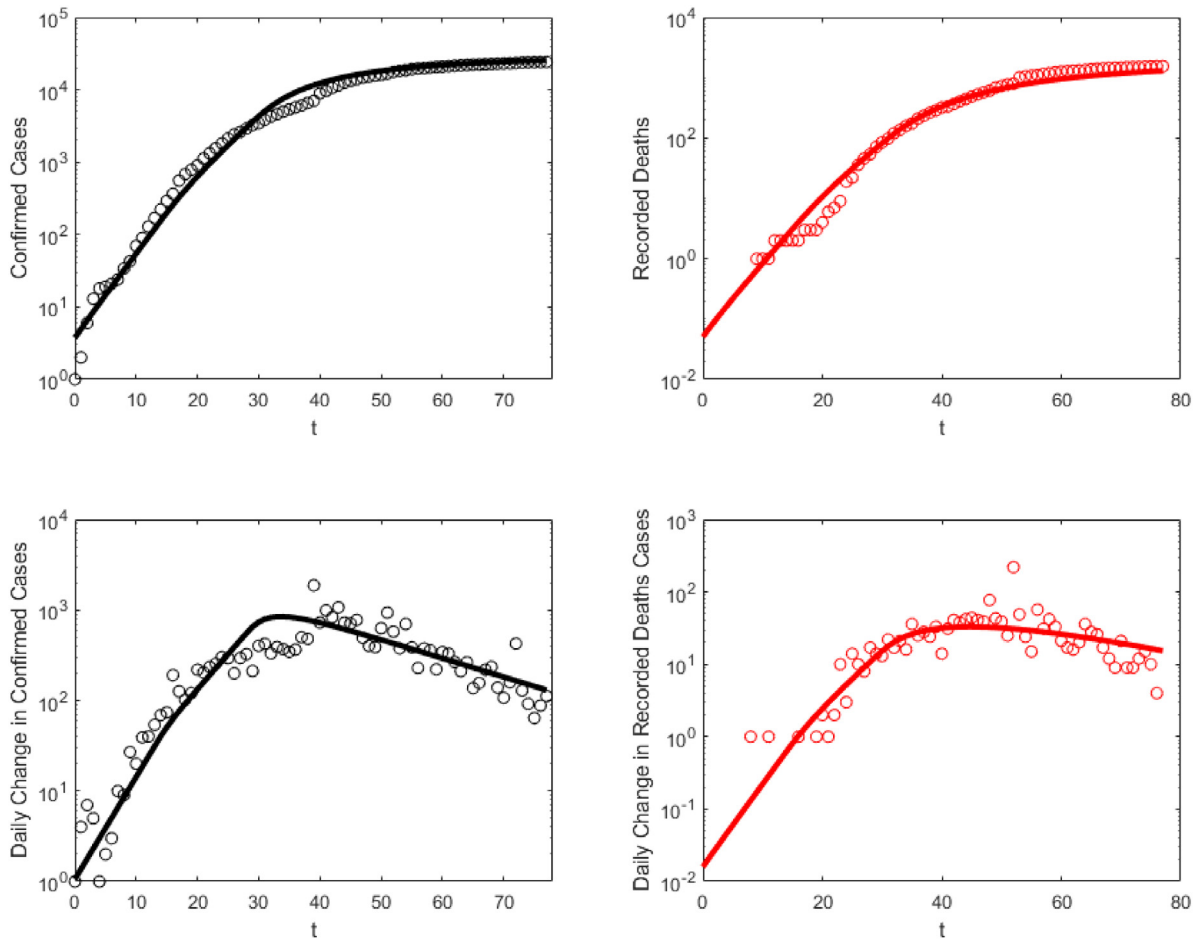


Fig. 1. Fit of the extended SEIR model (1) to the epidemiological data for the outbreak of COVID-19 in Ireland. $t = 0$ is the 29th February 2020, the day of the first recorded case. Data source: Ref. [18]. Top: Cumulative recorded cases and recorded deaths. Bottom: New daily recorded cases and daily recorded deaths.

The piecewise-constant function $u(t)$ is characterized as follows:

$$u(t) = \begin{cases} u_1, & t_0 < t \leq t_{s1}, \\ u_2, & t_{s1} < t \leq t_{s2}, \\ \vdots & \\ u_n, & t_{s,n-1} < t \leq T. \end{cases} \quad (10)$$

Here, n is an integer (one or greater), u_1, \dots, u_n are real numbers between zero and one, and $t_{s1}, \dots, t_{s,n-1}$ are switching times (real numbers between t_0 T , such that $t_{s1} < \dots < t_{s,n-1} < T$). In practice, the switching times correspond to how often the NPIs are updated.

We recall that $\beta_0 = cq_p$, where c is the average number of contacts between a susceptible person and an infections person, per day, and q_p denotes the probability that an infected, pre-symptomatic person transmits the disease to a susceptible person during one of those contacts. Therefore, the function $u(t)$ describes at a population level how to implement a strategy of NPIs — through a reduction of either c or q_p . For example, $u(t) = 0.5$ may correspond to a 50% reduction in c , resulting from a corresponding reduction in mobility for the susceptible population.

3.1. Formulation

We quantify the cost to the economy of implementing a sequence $\{u_1, \dots, u_n\}$ of controls by the cost function

$$J = \sum_{i=1}^n u_i(t_{s,i} - t_{s,i-1}), \quad (11)$$

with the convention that $t_{s,0} = t_0$ and $t_{s,n} = T$. As such, J is a function of the real variables $\{u_1, \dots, u_n, t_{s1}, \dots, t_{s,n-1}\}$. The rationale for this choice is as follows. The aim of each NPI is to reduce the transmission of the virus. Other than the ‘easy’ interventions such as public information campaigns, handwashing, etc. the interventions aim to reduce the number of daily contacts between individuals. Assuming economic activity is proportional to this number of contacts, there is a natural linear relationship between the level of the NPIs and the cost to the economy (e.g. if people’s daily contracts are reduced by 10%, then one can expect economic activity to be reduced by the same factor). However, we recognize that some of the more intense NPIs may carry a disproportionate cost to the economy and for that reason, the linear relationship in Eq. (11) may break down as $u_i \rightarrow 1$. For that reason, in what follows, we investigate the robustness of our results to the assumption of a linear cost function (Section 4).

We furthermore insist on an optimal control problem where human life is put on a very high footing. As such, we propose the following additional constraints on the optimal control problem:

$$kI_S(t) \leq B, \quad t \in (-t_{offset}, T] \quad (12)$$

In this context, the positive constant k may be thought of as the percentage of symptomatic cases who require a hospital bed (or a bed in ICU) at time t , and the positive constant B represents a corresponding capacity limit (number of hospital beds, number of ICU beds, etc.). Finally, we impose the additional constraint

$$\frac{dI_S}{dt} \leq 0, \quad t = T, \quad (13)$$

which is true if and only if $\tau_{I_p}^{-1}(1-f)I_p - \tau_I^{-1}I_S \leq 0$ at $t = T$. This rules out any ‘optimal’ strategy in which a large epidemic peak

would occur just beyond the horizon at $t = T$. The exclusion of such strategies is desirable for public-health reasons. As such, the optimal control problem to be solved is the following:

Under the epidemic modelled by Eq. (1), over a time horizon T , compute the minimizer of

$$\min_{n \geq 1} \left\{ \min_{u_1, \dots, u_n, t_{s1}, \dots, t_{s, n-1}} \left[\sum_{i=1}^n u_i(t_{s,i} - t_{s,i-1}) \right] \right\}, \quad (14)$$

subject to the constraints (12)–(13).

3.2. Discussion

We discuss the optimal control problem (14) in a more general context. We use the notation $\mathbf{x} = (S, E, I_p, I_A, I_S, R)$. We look at the optimal control problem

$$\left\{ \begin{array}{l} \text{Minimize} \quad J(\mathbf{x}, u) = \int_0^T u(t) dt, \\ \text{Subject to} \quad dx/dt \text{ satisfies Eq. (1), } 0 \leq t \leq T, \\ \quad \mathbf{x}(t=0) = \mathbf{x}_0, \quad \mathbf{x}_0 = \text{initial conditions implied by Eq. (4)}, \\ \quad 0 \leq u(t) \leq 1, \quad 0 \leq t \leq T, \\ \quad kI_S(t) \leq B, \quad 0 \leq t \leq T, \\ \quad \tau_{I_p}^{-1}(1-f)I_p - \tau_{I_S}^{-1}I_S \leq 0, \quad t = T. \end{array} \right. \quad (15)$$

The control is activated at $t = 0$; between $t = -t_{of\ fset}$ and $t = 0$, the evolution of dx/dt is given by Eq. (1) with $u = 0$. This particular optimal control problem can be put into the framework of Theorem 23.11 in Ref. [22] – this framework provides the conditions such that, if there is at least one admissible process (\mathbf{x}, u) for which J is finite, the optimal control problem admits a solution. The structure of the problem (15) is relatively simple, and to show that the conditions of Theorem 23.11 in Ref. [22] apply here, it suffices to show that solutions $\mathbf{x}(t)$ of the ordinary differential equation (1) remain in a bounded hypercube in \mathbb{R}^6 ; this is done in Appendix C. Furthermore, if the number of infectious individuals is sufficiently small at $t = 0$, then by making $u(t) = \text{Const.}$ sufficiently large, the solution $\mathbf{x}(t)$ will follow a standard epidemic curve with $kI_S(t) \leq B$ – this gives the required admissible process.

This framework gives very general necessary conditions for a solution to the problem (15) to exist. Often, standard analytical approaches can be applied to compute this solution analytically. The application of these approaches is hampered in the present context by the presence of the state constraints in Eqs. (12)–(13). For a system of ordinary differential equations $dx/dt = f(x, u)$, state constraints are generically expressed as $S(x) \leq 0$ (for this brief presentation we assume for discussion purposes that there is only one state constraint, hence S is a scalar-valued function of its arguments). The order of the constraint is the minimum value of p for which $d^p S/dt^p$ can be expressed as a linear function of u , that is, $d^p S/dt^p = A(x, t) + u(t)B(x, t)$. Here, $dS/dt = (\nabla S) \cdot f$, and A and B are continuous functions of their arguments. In the absence of such state constraints, a convenient way of solving optimal control problems is the Pontryagin Maximum Principle [13]. For cost functions that are linear in the control $u(t)$, a consequence of the Pontryagin Maximum Principle is that the optimal control is bang–bang. That is, if $0 \leq u(t) \leq u_{max}$, then the optimal control switches between $u(t) = 0$ and $u(t) = u_{max}$, with a finite number of switching times. With state constraints, the optimal control tends to be a mixture of bang–bang control sequences and ‘boundary arcs’, during which time the control $u(t)$ is such that the system maintains $d^p S/dt^p = 0$ [23].

The state constraints in Problem (14) are of high order (for instance, Eq. (12) is a third-order state constraint). The analysis of such problems is demanding, as the constraints involve the analysis of new multipliers associated with each individual constraint [14]. For this reason, the focus of this work is on computing optimal controls numerically. Clearly, the resulting absence of theoretical analysis means that our

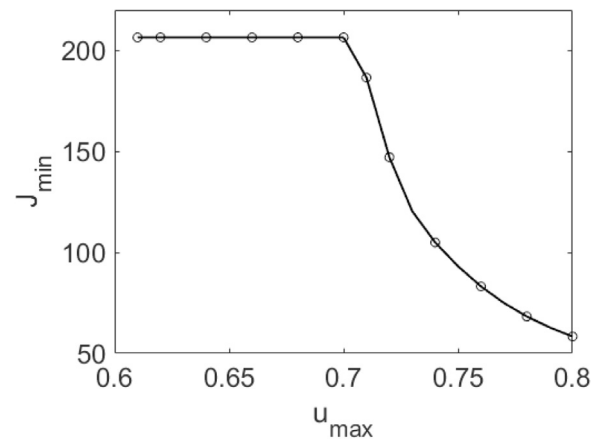


Fig. 2. Plot of J_{min} as a function of u_{max} , with $n = 2$. The plot range is $u_{max} \in [0.61, 0.8]$. The upper limit is justified in the text. The lower limit is $u_{max} \geq 0.61$; for smaller values of u_{max} the state constraints (12)–(13) fail to be satisfied.

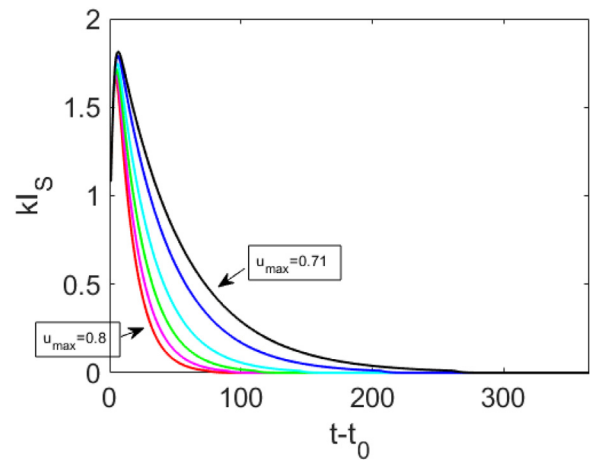


Fig. 3. Plot of $kI_S(t)$ under a bang–bang control with a single switch. Here, $k = 0.016$ and $t_0 = 13$ Days. The curves correspond in monotone order to $u_{max} = 0.8, 0.78, 0.76, 0.74, 0.72, 0.71$.

results should be treated with care since there is no guarantee that they are optimal. However, we have carefully examined the robustness of our numerical calculations – this is described in detail in Section 4 and also, in Appendix B.

There is also a second and more compelling reason for our use of numerical optimization methods in this work. From the outset we rule out the search for optimal controls that involve boundary arcs. Although this restriction may result in the ruling out of globally optimum solutions, we are guided by the intended application of the work, where the outbreak of an epidemic is to be controlled using public-health interventions. These interventions tend to be on the population level. Communication with, and collaboration from, the entire population is necessary for such measures to work. As such, piecewise-controls of the type (10) are more realistic. The formulation of the optimization problem in terms of such discrete controls lends itself naturally to a numerical approach.

3.3. Methodology

We solve the optimization problem (14) using simulated annealing. We use the built-in simulated annealing algorithms in Matlab/Octave. The source code is obtainable via an online repository [24]. During each round of simulated annealing, a random set of controls

$\{u_1, \dots, u_n, t_{s1}, \dots, t_{s,n-1}\}$ is drawn from the feasible values. The model (1) is then solved numerically using ODE45 in Matlab/Octave and the penalty function (11) is calculated. Extra terms are added to the penalty function for cases where the choice of controls causes the constraints (12)–(13) to be breached. The simulated annealing algorithm generates successive random sets of controls until a global minimum value of the penalty function is attained. Simulated annealing is guaranteed asymptotically to converge to the global minimum — more precisely, for any finite-dimensional problem, the probability that the simulated annealing algorithm terminates with a global optimal solution approaches 1 as the annealing schedule is extended [25]. In practice, the convergence can be very slow, and the Matlab/Octave codes employ built-in stopping criteria to confirm if a global minimum has been reached. These built-in stopping criteria can occasionally cause the simulated-annealing algorithm to converge before the attainment of the global minimum. We have checked this by comparing the results of the simulated annealing with other fundamentally different optimization methods with independent stopping criteria (e.g. particle-swarm optimization), and the results are the same. This inspires confidence in our numerical method, although the absence of theoretical analysis of the optimal control problem (14) at this time means that our results should be treated with some care.

4. Numerical results

We solve a variant of optimization problem (14) numerically. Specifically, for various values of the number of controls n , we compute

$$J_{min} = \min_{\substack{u_1, \dots, u_n, \\ t_{s1}, \dots, t_{s,n-1}}} \left[\sum_{i=1}^n u_i (t_{s,i} - t_{s,i-1}) \right] \quad (16)$$

subject to the constraints (12)–(13) (we also compute the corresponding values $\{u_1, \dots, u_n\}$ and $\{t_{s1}, \dots, t_{s,n-1}\}$ which achieve the minimum). We report on the trends as the integer n is increased from $n = 2$ to $n = 12$. The model (1) is solved subject to the initial conditions (4), with $t = 0$ corresponding to February 29th 2020, the day on which the first case of COVID-19 was recorded in Ireland. The model is solved without any controls until $t = t_0 = 13$ Days (cf. Eq. (5)), whereupon the control $u(t)$ is activated. In this way we look at counter-factual scenarios, to determine what would have been the optimal sequence of controls to implement, to control the outbreak of COVID-19 in Ireland. We fix the final time to be $T = t_0 + 365$ days. In this, we are motivated by similar recent work which looks at how to control an outbreak of COVID-19 over a timeframe prior to the development of pharmaceutical treatments of the disease [5]. Finally, based on the official statistics [26], we use $k = 0.016$ and $B = 300$ in the state constraints (12)–(13), this corresponds to an ICU admission rate of 1.6% among all confirmed cases of COVID-19.

To take account of strategies which reduce the spread of the disease to a very low level, we supplement Eq. (3) by introducing an indicator function,

$$I(t) = \begin{cases} 0, & \text{if } I_P + i_A I_A + i_S I_S < 1, \\ 1, & \text{otherwise.} \end{cases} \quad (17)$$

Then, Eq. (3) is further modified to read,

$$\frac{dS}{dt} = -I(t)\beta_0 \frac{S}{N} (I_P + i_A I_A + i_S I_S); \quad (18)$$

a similar modification applies to the equation for the ‘exposed’ compartment. In this way, once the disease is reduced to a low level in the population, in the sense that $I_P + i_A I_A + i_S I_S < 1$, an unrealistic resurgence in the disease by a notional ‘fraction of a person’ is ruled out.

The main purpose of this section is to characterize the solution of Eq. (16) as a function of u_{max} , where

$$0 \leq u(t) \leq u_{max}, \quad u_{max} \leq 1. \quad (19)$$

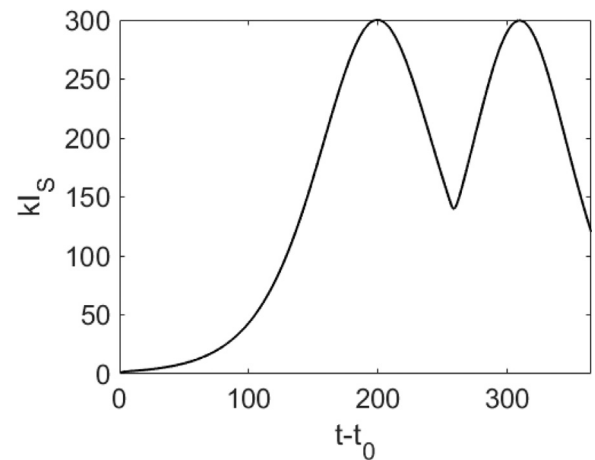


Fig. 4. Plot of $kI_S(t)$ under an optimal-control strategy in which the boundary is attained. Control parameters: $u_1 \approx 0.607$, $u_2 \approx 0.468$, with switching time at $t - t_0 \approx 256.5$ Days.

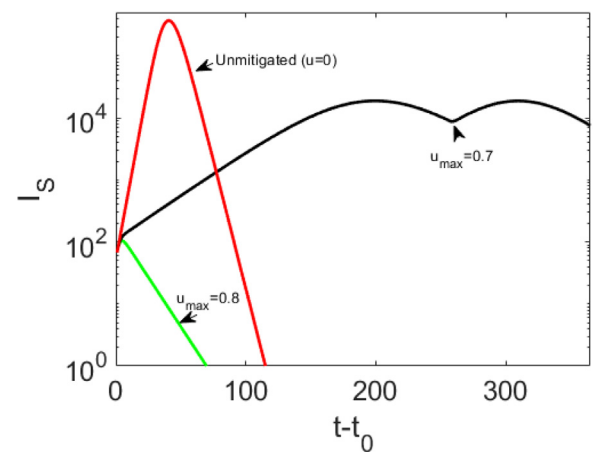


Fig. 5. $I_S(t)$ versus t for three scenarios: disease outbreak with controls $u_{max} = 0.7$ and $u_{max} = 0.8$ and also, disease outbreak with no controls.

That is, we look at the structure of the solution of the problem (16) for various values of u_{max} . We deliberately restrict ourselves to $u_{max} \leq 0.8$. When the model (1) was fitted to the data for Ireland, $u_{max} \approx 0.8$ was the maximum reduction in β achieved, under a severe set of interventions designed to curb the spread of COVID-19 (all non-essential services and industries closed down, travel restricted to within 2 km of a person’s house, with certain exceptions).

4.1. One switching time, $n = 2$

We start by looking at the structure of the solution of the problem (16), for a range of values of u_{max} , and for $n = 2$. Although switching between different sets of controls only once in the time period T might be unrealistic, this is a first step that helps with understanding (we investigate $n > 2$ in what follows). As such, we plot the minimum value J_{min} as a function of u_{max} in Fig. 2. For large values of u_{max} , $u_{max} \geq 0.7$, the optimal control is bang–bang with one switch: the control is enabled at the maximum value $u(t) = u_{max}$ for between 58 ($u_{max} = 0.8$) and 193 days ($u_{max} = 0.71$). The evolution of the infection under these controls is shown in Fig. 3 – under these bang–bang controls, the epidemic is seen to be suppressed, and the number of infectious cases is zero at the end-time T . In contrast, for $u_{max} \leq 0.7$, the optimal control is no longer bang–bang. Instead, the optimal control strategy is revealed to be independent of u_{max} : for an initial time

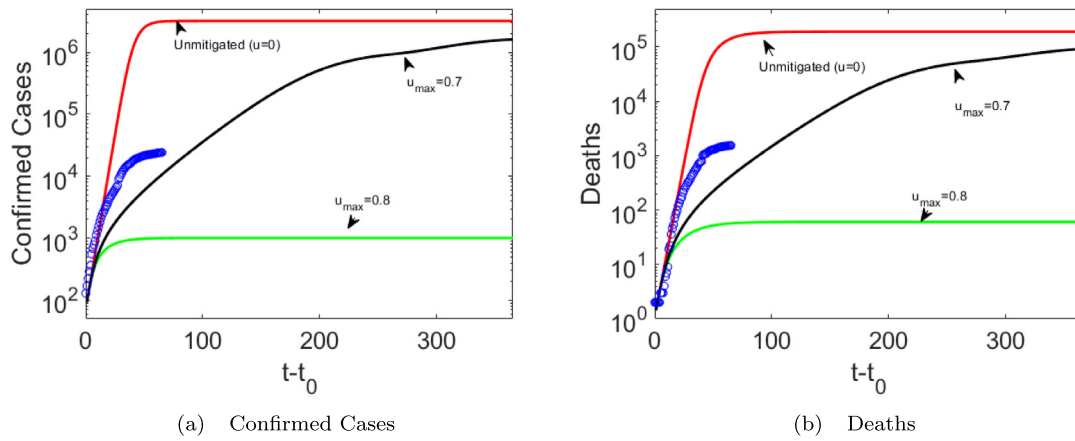


Fig. 6. Time evolution of the outbreak for three scenarios: disease outbreak with controls $u_{max} = 0.7$ and $u_{max} = 0.8$ and also, disease outbreak with no controls. The $-o$ curves show the actual course of the outbreak over the time interval for which the model is fitted to the data.

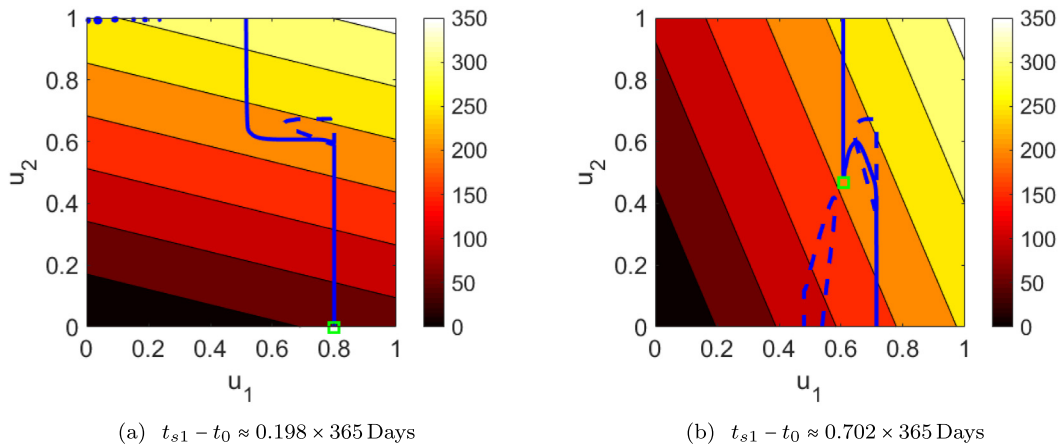


Fig. 7. Two-dimensional plots of $J(u_1, u_2, t_{s1})$ (cf. Eq. (20)) at fixed t_{s1} showing the dependence of J on the controls u_1 and u_2 . A local minimum is shown using a square marker in each panel. Solid line: the state constraint $k \max_{t \in [t_0, T]} I_S(t) = B$. Broken line: the state constraint $(dI_S/dt)_{t=T} = 0$. The closed regions demarcated by the broken lines correspond to the forbidden regions $(dI_S/dt)_{t=T} > 0$; regions to the left of the solid lines are forbidden also, and correspond to regimes where the constraint $kI_S(t) \leq B$ is breached.

interval, the value $u \approx 0.607$ is optimal, during which the boundary of the state constraint $kI_S(t) = B$ is attained, for some t . In a second time interval, the value $u \approx 0.468$ is optimal, during which time the boundary of the state constraint is again attained. The switching time occurs at $t_{s1} - t_0 \approx 256.5$ Days. The evolution of the epidemic under this scenario is shown in Fig. 4. This scenario corresponds to a ‘containment’ of the epidemic: the number of infectious cases is non-zero at the end-time T , but the maximum capacity of the hospital system is not breached. Finally, a comparison between a disease outbreak with controls, and one without any controls, is shown in Fig. 5. The imposition of controls contains or suppresses the epidemic: specifically, $\max(I_S) \approx 3.7 \times 10^5$ with no controls, $\max(I_S) = B/k = 18,750$ with $u_{max} = 0.7$, and $\max(I_S) \approx 100$ with the optimal control $u_{max} = 0.8$. Comparison between these scenarios and the actual course of the outbreak is shown in Fig. 6. The model predictions for the counterfactual scenario of an unmitigated outbreak are consistent with the initial modelling of the epidemiology of COVID-19 elsewhere, e.g. Ref. [4].

In the case with one switching time (i.e. $n = 2$), the optimization problem (16) contains only three parameters, and the optimal control strategies found in Figs. 3–4 can be investigated separately via a ‘brute-force approach’. This is desirable because it helps with our understanding and also, because it gives further confidence in the correctness of the simulated-annealing approach in isolating the global minimum. As such, we have evaluated

$$J(u_1, u_2, t_{s1}) = u_1(t_{s1} - t_0) + u_2(T - t_{s1}) \tag{20}$$

for a wide range of values, by direct computation. By fixing t_{s1} , a two-dimensional plot can be made, and a value $j(t_{s1}) = \min_{u_1, u_2} [u_1(t_{s1} - t_0) + u_2(T - t_{s1})]$ (subject to the state constraints) can be obtained graphically. By repeating this procedure over all values of t_{s1} , the optimal control strategies reported in Figs. 3–4 are recovered. Sample two-dimensional plots along these lines are shown in Fig. 7. The minimum marked in panel (a) corresponds to the global minimum with the imposed constraint $u_{max} = 0.8$ on the control. This corresponds to a bang–bang control with $u = u_{max}$ with one switching time at $t_{s1} - t_0 \approx 0.198 \times 365$ Days, and $u = 0$ thereafter. It can be asked if the minimum ($u_1 = 0.8, u_2 = 0, t_{s1} - t_0 \approx 0.198 \times 365$) is a true bang–bang control or is instead a boundary control, as this optimal control appears to occur along the curve $k \max_{t \in [t_0, T]} I_S(t) = B$ in Fig. 7(a). The answer is that the surface $\Phi(u_1, u_2) = k \max_{t \in [t_0, T]} I_S(t) - B$ possesses a jump discontinuity, which arises as a result of the extra condition (17)–(18) in the model (1). Therefore, the optimal control ($u_1 = 0.8, u_2 = 0, t_{s1} - t_0 \approx 0.198 \times 365$ Days) occurs on one side of a jump discontinuity and thus produces a solution that never attains the boundary value $k \max_{t \in [t_0, T]} I_S(t) = B$. This explanation can be verified by removing the condition (17)–(18) from the model; in that case, the discontinuity goes away, and the optimal control ($u_1 = 0.8, u_2 = 0, t_{s1} - t_0 \approx 0.198 \times 365$ Days) shifts elsewhere. However, we rule out further study of this special case as the condition (17)–(18) is deemed necessary for model realism.

We emphasize finally that although the two control strategies represented by Figs. 3–4 each represent a solution to the optimal control

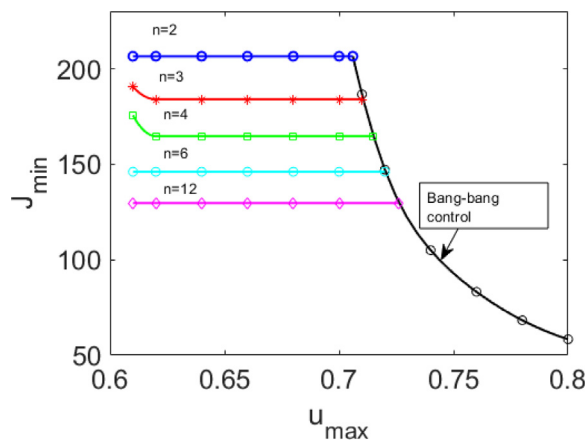


Fig. 8. Plot of J_{min} as a function of u_{max} , with $n \geq 2$.

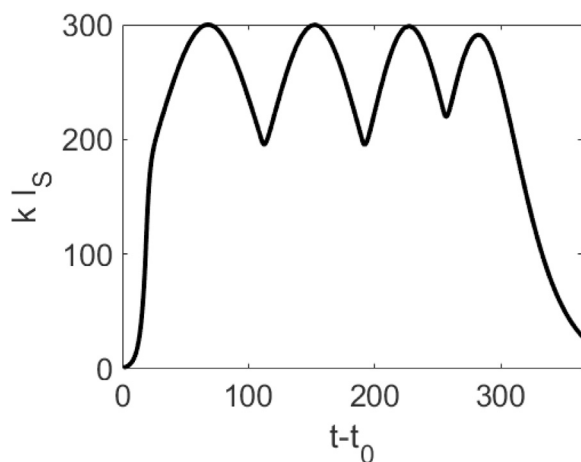


Fig. 9. Plot of $kI_S(t)$ under an optimal-control strategy in which the boundary is attained. Control parameters: $n = 6$, $u_{max} = 0.68$. Cost: $J_{min} \approx 146.5$. Similar results are found for other values of n , e.g. with $n = 12$ an eight-peak plot is obtained with $J_{min} \approx 129.6$.

problem (16), they are by no means equivalent. The ‘bang–bang’ control strategy (high u_{max}) comes with a much lower economic cost (lower J_{min}). Therefore, subject to the model assumptions and the framework of the optimal control problem (14), an eradication strategy is far preferable to a containment strategy.

4.2. Several switching times, $n > 2$

We extend the foregoing analysis to include control strategies with multiple switching times. The results are summarized succinctly in Fig. 8. The results are qualitatively similar to what is observed in the case of a single-switch ($n = 2$) control — that is, for large values of u_{max} , the bang–bang control with a single switch and the complete suppression of the disease is optimal, while for lower values of u_{max} a containment strategy is optimal. Under the mitigation strategy, the state constraint $k \max_{t \in [t_0, T]} I_S(t) = B$ is attained at several times — the attainment of the bound corresponds to local maxima in the function $I_S(t)$ of infectious cases (e.g. Fig. 9). Also, the cost to the economy of the mitigation strategy is reduced as the number of switching times is increased. Overall, however, the strategy of complete suppression is optimal: the higher the value of u_{max} permitted, the lower the economic cost of the control measures.

The results in Fig. 8 are for $n \leq 12$; we do not rule out that for $n > 12$, a mitigation strategy may become competitive with a suppression strategy. However, all mitigation strategies (no matter how many switches) have an inherent instability-like behaviour — small deviations from the optimal control lead to large deviations in the outcome of the epidemic. This is illustrated in Fig. 10, where we show 10 different simulation results corresponding to control sequences that are within 10% of the optimal control. In several cases, this small deviation away from the optimal control leads to a breach of the state constraint concerning ICU capacity by a factor of 4. Thus, the maintenance of the state constraints is highly sensitive to small deviations away from the optimal control. Now, the non-pharmaceutical interventions discussed in this article are population-based interventions, concerning such public-health initiatives as social distancing, testing suspect cases, quarantining infected individuals, up to so-called ‘lockdown’ of entire populations. It can be appreciated that these are blunt tools for epidemic control, and are dependent on a high degree of compliance from the public. Consequently, a deviation of 10% in the implementation of these controls seems reasonable. And yet such a small deviation leads to a four-fold breach in the ICU capacity. This sensitivity in the outcomes to small deviations in the controls is another reason why the mitigation strategy is inferior to the suppression strategy.

4.3. Quadratic penalty function

We have repeated the analysis of the previous sections for a quadratic cost function,

$$J = \sum_{i=1}^n u_i^2 (t_{s,i} - t_{s,i-1}). \tag{21}$$

The motivation for considering Eq. (21) is to look at the hypothetical scenario where the cost to the economy of the non-pharmaceutical interventions is nonlinear in the intensity u_i of the interventions. In this way, Eq. (21) describes a scenario where the more intense interventions ($u_i = 1$) carry a disproportionately severe cost in comparison to less intense interventions ($u_i < 1$). In spite of the differences between the linear and quadratic cost functions, the results in both cases are qualitatively the same – e.g. a plot of the minimum value of J as a function of u_{max} for the quadratic case is shown in Fig. 11, to be compared with the linear cost function in Fig. 8. The main difference between these two figures is that the crossover where the optimal strategy switches from mitigation to suppression occurs at a higher value of u_{max} for the quadratic cost function — that is, under the assumption of the quadratic cost function, mitigation is an optimal strategy for a wider range of values of u_{max} . That being said, the sensitivity of the mitigation strategy to small deviations in the controls still applies for the case of the quadratic penalty function.

4.4. Sensitivity of results to parameter variations

We briefly discuss the sensitivity of our results to parameter variations, focusing on some key parameters of interest. In the first instance, we remark that we have taken t_0 as fixed in this work — that is, the day on which NPIs are first introduced is fixed *a priori*. We have taken $t_0 = 13$ days, corresponding to the day on which NPIs were introduced in Ireland, on foot of the outbreak of COVID-19. By fixing t_0 in this way, we are asking what would be the optimal sequence of NPIs, given that the NPIs commence on a particular day. Certainly, it may be of interest to look at varying t_0 . Based on the other results in this work, it would seem that the optimal strategy in this case would still be to suppress the disease from the outset (and to make t_0 as small as possible). In this way, any initial sharp rise in case numbers (due to exponential growth) would be more strongly suppressed (compared to fixed t_0), and high-intensity NPIs could be lifted sooner, corresponding to a lower cost to the economy.

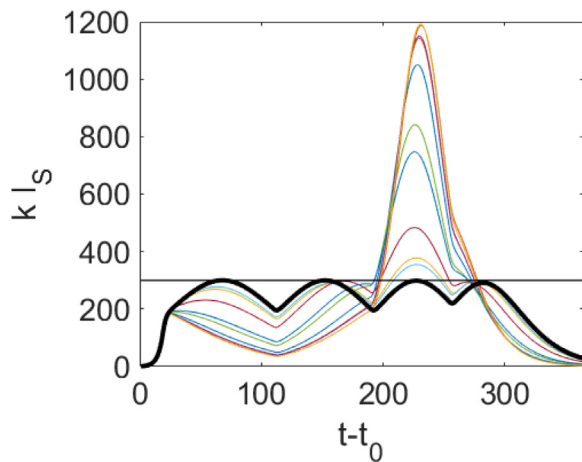


Fig. 10. Black line: Plot of $kI_S(t)$ under an optimal-control strategy in which the boundary is attained. Control parameters: $n = 6$, $u_{max} = 0.68$. Other curves: plot of $kI_S(t)$ for control strategies that are within 10% of the optimal control strategy — specifically, each u_n and t_n is varied independently by 10% with respect to the optimal values.

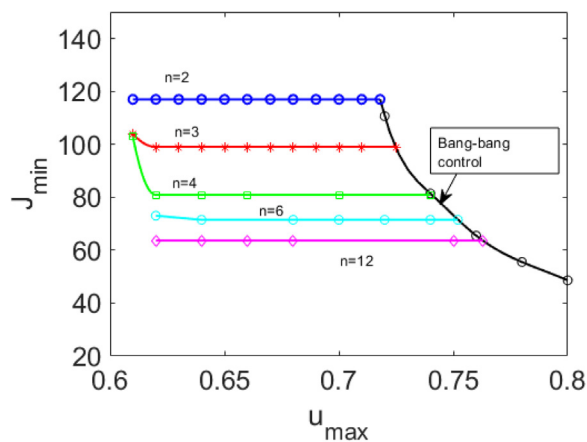


Fig. 11. Plot of J_{min} as a function of u_{max} , with $n \geq 2$, quadratic cost function (Eq. (21)).

In the same way, we emphasize again (cf. Section 1) that the time horizon of the optimal control problem is fixed, with $T = 1$ year. The motivation is that this is a realistic time horizon over which to manage the spread of COVID-19 in population, before the advent of anticipated therapeutic interventions such as anti-viral drugs and vaccines. Introducing a variable time horizon for the optimal control problem can be addressed in future work — the same observations as above will apply, namely the importance of suppressing the initial exponential growth of the epidemic.

Finally, it can be noted that optimal control solutions are well known to suffer from low robustness. The same caveat applies here — it may be anticipated that the control parameters $\{u_1, \dots, u_n, t_{s1}, \dots, t_{s,n-1}\}$ will vary under small changes in the model parameters (for instance, parameter variations within the bounds of the confidence intervals in Appendix A). As such, the optimal control parameters are not expected to be prescriptive. Instead, the general trends in the results are more important — in particular, the importance of suppressing any initial sharp rise in cases, such that high-intensity NPIs can be lifted as soon as possible, with a view to minimizing the economic cost of the suppression strategy.

5. Conclusions

Summarizing, we have introduced a mathematical model to describe the outbreak of COVID-19 in Ireland in March–May 2020. Based on this model, which we compare to actual epidemiological data, we have used numerical methods from optimal control theory to answer a hypothetical question — what would have been the optimal strategy for the control of the epidemic? The optimal control theory is based on minimizing a cost function which is related to the cost to the economy of the various control measures. At the same time, the theory introduces state constraints as a means of maintaining the number of infectious cases below a maximum threshold such that hospital capacity is not overwhelmed. By insisting on piecewise-constant controls (appropriate for modelling rather crude population-level public-health interventions) and state constraints, the optimal-control model is not amenable to theoretical analysis — hence a numerical approach is considered. As such, we have performed the necessary optimizations via numerical simulated annealing — the results show that a suppression strategy (implemented promptly) is superior to a mitigation strategy.

The optimal suppression strategy may be compared to the strategy actually implemented by the Irish Government in the early stages of the epidemic (March–May 2020). Broadly, the optimal strategy is similar to the actually implemented strategy, in the sense that both involve high-intensity NPIs at the beginning of the epidemic, with a view to suppressing the outbreak. However, the optimal strategy envisages maximum-intensity NPIs at the very outset — this saves lives (e.g. Fig. 6, with 1547 deaths at $t - t_0 = 65$ of the outbreak in real life, and 57 deaths on the same day under the optimal strategy, and an asymptotic final number of deaths of 61 under the same optimal strategy). Therefore, imposing high-intensity NPIs from the very beginning saves lives; it is also economically less costly, under the penalty function formulated in this work.

The optimal strategy predicted by the model envisages the maximum level of non-pharmaceutical interventions (NPIs) from the outset (Day 13 in our calculations, cf. Eq. (5)), and lasting 58 days at $u_{max} = 0.8$. In contrast, the strategy implemented by the Irish Government involved some controls on Day 13, maximum controls on Day 28, followed by gradual easing of controls on Day 79 (May 18th 2020). The period between the initial NPIs and the first easing of NPIs is therefore $79 - 13 = 66$ Days. Thus, the immediate imposition of maximum controls on Day 13 would have resulted in the lifting of the most intense NPIs roughly 14 days sooner. Naturally, this estimate is subject to model uncertainty, but the clear message (supported by the other findings in this paper) is that prompt implementation of maximum NPIs is best.

Finally, it can be emphasized that the coupling of the epidemiological model to the economy in the present model framework is rather crude (e.g. the linear cost function (11)). In future, it can be envisaged that epidemiological models can be coupled to economic models of production [15] or employment [27] to better understand the interplay between the different societal costs of a disease outbreak. Equally, the approach presented in this work may form a basis for the design of an optimal vaccination strategy if and when a vaccine becomes available for COVID-19.

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.

Acknowledgements

The author acknowledge technical support by Alexander Gubic in facilitating access to the Orr compute cluster in UCD. The author also acknowledges helpful discussions with Miguel Bustamante, Michelle Carey, James Herterich, and Richard Martin.

Table A.2

Table of the range of allowed parameter values for the solution of the optimization problem (6) by simulated annealing.

Parameter	Lower bound	Upper bound	Reference/Justification
t_{offset}	1	15	Optimal value is checked to make sure it is strictly within the bounds
β_0	0	5	Optimal value checked to make sure it is strictly within the bounds and checked also for consistency with known values of the basic reproduction number R_0
β_1	0	5	Optimal values are checked and the consistency criterion $\beta_1 < \beta_0$ is satisfied.
β_2	0	5	Optimal values are checked and the consistency criterion $\beta_2 < \beta_1$ is satisfied.
f	0.1	0.5	In a recent review, the proportion of positive cases that remained asymptomatic was estimated at 16%, with a range from 6 to 41% [28].
g	0.05	0.07	Case Fatality Ratio. Ireland: 5.6% [26]. World: 5.6% [29].
τ_E	3.05	4.35	Mean value 3.7 days [16]. Variation in this parameter accounted for in line with the reference.
τ_{IP}	0.85	2.15	Mean value 3.7 days [16]. Variation in this parameter accounted for in line with the reference. Mean value of $\tau_E + \tau_I = 5.2$ days.
τ_I	1.3	3.3	Mean value 2.3 days [16]. Significant variation accounted for in the optimization. Optimal value checked to make sure it is strictly within the bounds.
τ_D	10	30	Ref. [30]
τ_T	1	5	Optimal value is checked to make sure it is strictly within the bounds
i_A	0	1	$q_S \approx q_{PS}$ due to occurrence of peak viral load close to symptom onset [31]. Also, $q_A < q_S$ [28]. Hence, $i_A = q_A/q_P \leq 1$.
i_S	0	1	$i_S \leq 1$ to account for some level of case isolation.

Appendix A. Fitting the extended SEIR model to epidemiological data for the outbreak of COVID-19 in Ireland.

In this Appendix we provide the necessary background information to fit the model (1) to the epidemiological data in Ref. [18] for the outbreak of COVID-19 in Ireland. A nonlinear optimization is performed to minimize the penalty function $\Delta(t_{offset}, \beta_0, \beta_1, \beta_2, f, g, \tau_E, \tau_{IP}, \tau_I, \tau_D, i_A, i_S)$ in Eq. (6). First, bounds on the range of acceptable parameter values are obtained in Table A.2. The resulting optimal parameter values are provided in Table A.3. In this table, we also present confidence intervals for these optimal parameter values – these have been obtained by bootstrapping – the model residuals are resampled with replacement $n_{bootstrap}$ times. These residuals are then used to compute $n_{bootstrap}$ synthetic data sets, and the model (1) is fitted to each data set in turn. This then gives a range of $n_{bootstrap}$ ‘optimal’ parameter values; a histogram based on this range of values is constructed and the 2.5 and 97.5 percentiles are computed to produce the reported confidence intervals. We use $n_{bootstrap} = 100$ and $n_{bootstrap} = 1000$, with little or no difference observed between these two choices.

Furthermore, we have used the data in Table A.3 to generate estimates of the basic reproduction number and the effective reproduction number under various NPIs, with appropriate lower and upper bounds. The results are shown in Table A.4, these may be compared with Table 1 in the main paper.

According to the estimates in Table A.2, the mean value of the infectious periods $\tau_{IP} + \tau_I$ is between 2.15 and 5.45 days. This is consistent with Ref. [16]. However, some other studies indicate that the infectious period may be significantly longer (e.g. up to 10 days [31]). As the main purpose of this article is to establish a methodology for computing optimal control strategies, this discrepancy is not important in the current context. However, exploring the effect of this range of values on the results may be of interest in future work.

Appendix B. Description of simulated annealing method for optimization

To maximize the chance of finding a globally optimal control, we use the method of Simulated Annealing. We exploit the built-in simulated annealing capabilities in Matlab and Octave for these purposes. As such, we solve Eq. (1) for an arbitrary sequence of controls $\{u_1, \dots, u_n, t_{s1}, \dots, t_{s,n-1}\}$. In Matlab/Octave programming we execute the following command:

```
penalty=ode_solve_seir(u);
```

Here, `ode_solve_seir` is a Matlab/Octave ODE45 solver which takes in the input $u = \{u_1, \dots, u_n, t_{s1}, \dots, t_{s,n-1}\}$, solves Eqs. (1) until the final time $t = T$, and returns the penalty function (11). The state constraints (12)–(13) are taken into account by adjoining to the penalty function an additional term

$$\left[\tanh \left(\frac{0.016 \max(I_S) - 300}{0.1} \right) + 1 \right] [0.016 \max(I_S) - 300]^2 + \left[\tanh \left(\frac{(dI_S/dt)_{t=T}}{0.1} \right) + 1 \right] [(dI_S/dt)_{t=T}]^2, \quad (B.1)$$

this imposes an additional penalty on configurations $\{u_1, \dots, u_n, t_{s1}, \dots, t_{s,n-1}\}$ that violate the state constraints (12)–(13) but leaves other states un-penalized.

We use the default convergence criteria in the simulated annealing algorithm in Matlab. In particular, the algorithm is deemed to have converged if the average change in the objective function over the previous $500 \times$ (Problem Size) is less than 10^{-6} . Here, (Problem Size) is the size of the optimization problem which in the present case is $2n - 1$. We have systematically varied these parameters and find no significant change to the estimated global minima.

Table A.3

Optimal parameter values corresponding to the optimization problem (6). The reported confidence intervals use $n_{bootstrap} = 1000$.

	t_{offset} (days)	β_0 (days) ⁻¹	β_1 (days) ⁻¹	β_2 (days) ⁻¹	f	g	τ_E (days)
Lower bound	7.9057	1.4679	0.9524	0.3205	0.2245	0.0500	3.6257
Best estimate	9.9831	1.4695	1.1009	0.3576	0.3084	0.0599	3.7486
Upper bound	9.9936	1.6961	1.2201	0.3908	0.4445	0.0614	3.7976

	τ_{IP}	τ_I	τ_D	τ_T	i_A	i_S
Lower bound	1.1394	1.9160	14.9812	3.3146	0.3303	0.2741
Best estimate	1.2938	2.1738	15.2210	3.8911	0.4096	0.3405
Upper bound	1.3841	2.4337	16.5007	5.2055	0.4758	0.4367

Table A.4

Estimates of the basic reproduction number and the effective reproduction number under various NPIs, with lower and upper bounds generated from the confidence intervals in Table A.3.

Intervention	β (Best fit)	i_S (Best fit)	$R_0(\beta, i_S)$ (Lower bound)	$R_0(\beta, i_S)$ (Best fit)	$R_0(\beta, i_S)$ (Upper bound)
No interventions, basic reproduction number	1.4695	1	2.7664	3.4495	4.9079
Case isolation, no other interventions	1.4695	0.3405	2.0798	2.6945	3.7549
Case isolation, schools and universities are closed, mass gatherings are banned	1.1009	0.3405	1.2880	2.0187	2.7204
Case isolation, all non-essential services and industries closed, travel restrictions	0.3576	0.3405	0.4847	0.6557	0.8324

In order to carry out large-scale simulated annealing runs (e.g. over 40 runs to generate the graph in Fig. 8, some taking over 3 days), we have implemented the simulated annealing algorithms in Octave (a GNU open-source alternative to Matlab) and carried out large-scale runs on a cluster. The cluster is made up of 10-core/20-thread Intel(R) ‘Ivy Bridge’ Xeon(R) CPUs (2.50 GHz, 32 GB RAM). Each optimization job is run on a single core; however, the jobs are carried out simultaneously in batch mode. The default convergence criteria for the simulated annealing algorithm are different from those in Matlab [32] (the average change in the objective function over the 5 previous pseudo-temperature changes in the simulated annealing algorithm should be less than 10^{-10}). The results under the Matlab convergence criteria and the Octave convergence criteria are near-identical (exactly-identical results are unlikely, since simulated annealing is a non-deterministic algorithm). This inspires confidence in the robustness of the results generated in the paper, although certainly, without a theoretical analysis of the optimal control problem, convergence to a global minimum is never guaranteed.

Appendix C. Solutions of the SEIR model remain in a bounded set

We show that solutions (S, E, I_P, I_A, I_S, R) of Eq. (1) remain bounded for all time, once suitable initial conditions are provided. For clarity of the exposition, we assume that Eq. (1) is posed on a time interval $(0, \tau]$, with initial conditions provided at $t = 0$; specifically,

$$S(0) = S_0, E(0) = E_0, I_P(0) = I_{P0}, I_A(0) = I_{A0},$$

$$I_S(0) = I_{S0}, R(0) = R_0, \tag{C.1}$$

where S_0, \dots, R_0 are non-negative real numbers, and where specifically,

$$S_0 > 0, E_0 = 0, I_{P0} > 0, \tag{C.2}$$

corresponding to standard initial conditions for the outbreak of an epidemic.

Eqs. (1)(a)–(f) can be summed up to yield

$$\frac{d}{dt} (S + E + I_P + I_A + I_S + R) = 0, \tag{C.3}$$

hence

$$(S + E + I_P + I_A + I_S + R)(t) = S_0 + E_0 + I_{P0} + I_{A0} + I_{S0} + R_0 = N,$$

$$t \in [0, \tau], \tag{C.4}$$

where N is the total population (positive constant). We formally integrate Eq. (1)(a) to obtain

$$S(t) = S_0 e^{-\int_0^t c(q_A I_A + q_P I_P + q_S I_S) dt}, \quad S_0 > 0, \tag{C.5}$$

hence $S(t) \geq 0$.

We assume as per Eq. (C.2) that $E_0 = 0$, with $I_{P0} > 0$, and $I_{A0}, I_{S0} \geq 0$. By continuity, there is a time interval on which $I_P(t) > 0$. Let t_* be the first time such that $I_P(t) = 0$. We assume for contradiction that $t_* \in [0, \tau]$.

From Eq. (1), we have

$$I_A(t) = I_A(0)e^{-t/\tau_I} + \frac{f}{\tau_{IP}} \int_0^t I_P(t')e^{-(t-t')/\tau_I} dt', \tag{C.6}$$

hence $I_A(t) > 0$ for $t \in (0, t_*]$. Similarly,

$$I_S(t) = I_S(0)e^{-t/\tau_I} + \frac{1-f}{\tau_{IP}} \int_0^t I_P(t')e^{-(t-t')/\tau_I} dt', \tag{C.7}$$

hence $I_A(t) > 0$ for $t \in (0, t_*]$. Thus,

$$\sum_{j=P,A,S} q_j I_j > 0, \quad t \in [0, t_*]. \tag{C.8}$$

Now, from Eq. (1) again, we have

$$E(t) = \int_0^t e^{-(t-t')/\tau_E} \left[\sum_{j=P,A,S} q_j I_j(t') \right] c S(t') dt', \tag{C.9}$$

hence $E(t) > 0$ for $t \in (0, t_*]$. Finally, from Eq. (1) again, we have

$$\frac{dI_P}{dt} = \frac{1}{\tau_E} E - \frac{1}{\tau_{IP}} I_P > -\frac{1}{\tau_{IP}} I_P, \quad t \in (0, t_*]. \tag{C.10}$$

Hence, by Gronwall’s Inequality,

$$I_P(t) > I_{P0} e^{-t/\tau_{IP}}, \quad t \in (0, t_*]. \tag{C.11}$$

But this is a contradiction, since $I_P(t_*) = 0$. Hence, $t_* \notin [0, \tau]$, and I_P remains positive for all $t \in [0, \tau]$. From Eqs. (C.6), (C.7), and (C.9), the same conclusion follows for I_A, I_S , and E respectively, and by similar reasoning, $R(t) \geq 0$ also, for all $t \in [0, \tau]$. Thus,

$$S(t) \geq 0, E(t) \geq 0, \dots, R(t) \geq 0, \quad t \in [0, \tau].$$

Since $(S + E + I_P + I_A + I_S + R)(t) = N$, it follows that each of (S, E, I_P, I_A, I_S, R) remains between 0 and N for all $t \in [0, \tau]$ and hence, the solution of the ordinary differential equation (1) is bounded.

References

- [1] Chaolin Huang, Yeming Wang, Xingwang Li, Lili Ren, Jianping Zhao, Yi Hu, Li Zhang, Guohui Fan, Jiuyang Xu, Xiaoying Gu, et al., Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, *Lancet* 395 (10223) (2020) 497–506.
- [2] World Health Organization, Coronavirus disease (COVID-2019) situation reports, 2020, <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/>. Visited: 31st July 2020.
- [3] John H. Beigel, Kay M. Tomashek, Lori E. Dodd, Aneesh K. Mehta, Barry S. Zingman, Andre C. Kalil, Elizabeth Hohmann, Helen Y. Chu, Annie Luetkemeyer, Susan Kline, et al., Remdesivir for the treatment of Covid-19—preliminary report, *New Engl. J. Med.* (2020).
- [4] S. Eubank, I. Eckstrand, B. Lewis, S. Venkatramanan, M. Marathe, C.L. Barrett, Commentary on Ferguson, et al., “Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand”, *Bull. Math. Biol.* 82 (2020) 1–7.
- [5] Ramses Djidjou-Demasse, Yannis Michalakis, Marc Choisy, Micea T. Sofonea, Samuel Alizon, Optimal COVID-19 epidemic control until vaccine deployment, 2020, medRxiv.
- [6] Patrick G.T. Walker, Charles Whittaker, Oliver J. Watson, Marc Baguelin, Peter Winskill, Arran Hamlet, Bimandra A. Djafaara, Zulma Cucunubá, Daniela Olivera Mesa, Will Green, et al., The impact of COVID-19 and strategies for mitigation and suppression in low-and middle-income countries, *Science* (2020).
- [7] Abdel Meguid Kassem, COVID-19: Mitigation or suppression?, *Arab J. Gastroenterol.* 21 (1) (2020) 1.
- [8] Ellina V. Grigorieva, Evgenii N. Khailov, Optimal intervention strategies for a SEIR control model of Ebola epidemics, *Mathematics* 3 (4) (2015) 961–983.
- [9] E.V. Grigorieva, E.N. Khailov, Optimal vaccination, treatment, and preventive campaigns in regard to the SIR epidemic model, *Math. Model. Nat. Phenom.* 9 (4) (2014) 105–121.
- [10] A. Ishaku, A.M. Gazzali, S.A. Abdullahi, N. Hussaini, Analysis and optimal control of an HIV model based on CD4 count, *J. Math. Biol.* (2020) 1–33.
- [11] Luca Bolzoni, Elena Bonacini, Cinzia Soresina, Maria Groppi, Time-optimal control strategies in SIR epidemic models, *Math. Biosci.* 292 (2017) 86–96.
- [12] Luca Bolzoni, Elena Bonacini, Rossella Della Marca, Maria Groppi, Optimal control of epidemic size and duration with limited resources, *Math. Biosci.* 315 (2019) 108232.
- [13] Arthur Earl Bryson, *Applied Optimal Control: Optimization, Estimation and Control*, CRC Press, 1975.
- [14] Md Haider Ali Biswas, Luís Tiago Paiva, M.D.R. De Pinho, A SEIR model for control of infectious diseases with constraints, *Math. Biosci. Eng.* 11 (4) (2014) 761.
- [15] Lukasz Drozd, Marina Tavares, Responding to Covid-19: A note, 2020, FRB of Philadelphia Working Paper.
- [16] Laura Di Domenico, Giulia Pullano, Chiara E. Sabbatini, Pierre-Yves Boëlle, Vittoria Colizza, Expected impact of lockdown in Île-de-France and possible exit strategies, 2020, medRxiv.
- [17] IEMAG Epidemiology Modelling subgroup, A population-level SEIR model for COVID-19 scenarios, 2020, <https://www.gov.ie/en/publication/dc5711-irish-epidemiology-modelling-advisory-group-to-nphet-technical-notes/>. Visited: 15th June 2020.
- [18] Ordnance Survey Ireland, CovidStatisticsProfileHPSCirelandOpenData, 2020, <https://data.gov.ie/dataset/covidstatisticsprofilehpscirelandopendata>. Visited: 15th June 2020.
- [19] IEMAG Epidemiology Modelling subgroup, Time-Series Based Estimation of the Effective Reproduction Number (Rt) - 21 May 2020, 2020, <https://www.gov.ie/en/publication/dc5711-irish-epidemiology-modelling-advisory-group-to-nphet-technical-notes/>. Visited: 15th June 2020.
- [20] Odo Diekmann, J.A.P. Heesterbeek, Michael G. Roberts, The construction of next-generation matrices for compartmental epidemic models, *J. R. Soc. Interface* 7 (47) (2010) 873–885.
- [21] Ying Liu, Albert A. Gayle, Annelies Wilder-Smith, Joacim Rocklöv, The reproductive number of COVID-19 is higher compared to SARS coronavirus, *J. Travel Med.* 27 (2020).
- [22] Francis Clarke, *Functional Analysis, Calculus of Variations and Optimal Control*, vol. 264, Springer Science & Business Media, 2013.
- [23] Kazufumi Ito, Karl Kunisch, Minimal effort problems and their treatment by semismooth Newton methods, *SIAM J. Control Optim.* 49 (5) (2011) 2083–2100.
- [24] Lennon Ó. Náraigh, Áine Byrne, Source Code, https://github.com/ainebyrne/ONaraighByrne_COVID19_optimal_control.
- [25] Vincent Granville, Mirko Krivánek, J.-P. Rasson, Simulated annealing: A proof of convergence, *IEEE Trans. Pattern Anal. Mach. Intell.* 16 (6) (1994) 652–656.
- [26] Health Protection Surveillance Centre, Epidemiology of COVID-19 in Ireland, 2020, [https://www.hpsc.ie/a-z/respiratory/coronavirus/novelcoronavirus/casesinireland/epidemiologyofcovid-19inireland/COVID-19_Daily_epidemiology_report_\(NPHT\)_06062020_final_website.pdf](https://www.hpsc.ie/a-z/respiratory/coronavirus/novelcoronavirus/casesinireland/epidemiologyofcovid-19inireland/COVID-19_Daily_epidemiology_report_(NPHT)_06062020_final_website.pdf). Visited: 15th June 2020.
- [27] David Baqaee, Emmanuel Farhi, Michael Mina, James H. Stock, Reopening Scenarios, NBER Working Paper No. w27244, 2020, Available at SSRN: <https://ssrn.com/abstract=3609697>.
- [28] Oyungerel Byambasuren, Magnolia Cardona, Katy Bell, Justin Clark, Mary-Louise McLaws, Paul Glasziou, Estimating the extent of true asymptomatic COVID-19 and its potential for community transmission: systematic review and meta-analysis, 2020, Available at SSRN 3586675.
- [29] David Baud, Xiaolong Qi, Karin Nielsen-Saines, Didier Musso, Léo Pomar, Guillaume Favre, Real estimates of mortality following COVID-19 infection, *Lancet Inf. Dis.* 20 (2020).
- [30] World Health Organization, Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19), 2020, <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>. Visited: 15th June 2020.
- [31] Kelvin Kai-Wang To, Owen Tak-Yin Tsang, Wai-Shing Leung, Anthony Raymond Tam, Tak-Chiu Wu, David Christopher Lung, Cyril Chik-Yan Yip, Jian-Piao Cai, Jacky Man-Chun Chan, Thomas Shiu-Hong Chik, et al., Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study, *Lancet Inf. Dis.* (2020).
- [32] Michael Creel, Unconstrained optimization with MINTOOLKIT for GNU octave, 2004.