Optimal control of a SIR epidemic with ICU constraints and target objectives

Florin Avram, Lorenzo Freddi, Dan Goreac

Optimal Control of a SIR Epidemic With ICU Constraints and Target Objectives

Florin Avram[1], Lorenzo Freddi[2], and Dan Goreac[3,4]

[1]Laboratoire de Mathématiques Appliquées, Université de Pau, F-64012 Pau, France, avramf3@gmail.com
[2]Dipartimento di Scienze Matematiche, Informatiche e Fisiche, via delle Scienze 206, 33100 Udine, Italy, lorenzo.freddi@uniud.it
[3]School of Mathematics and Statistics, Shandong University, Weihai, Weihai 264209, PR China
[4]LAMA, Univ Gustave Eiffel, UPEM, Univ Paris Est Creteil, CNRS, F-77447 Marne-la-Valleée, France, dan.goreac@univ-eiffel.fr

Abstract

The aim of this paper is to provide a rigorous mathematical analysis of an optimal control problem with SIR dynamics. The main feature of our study is the presence of state constraints (related to intensive care units ICU capacity) and strict target objectives (related to the immunity threshold). The first class of results provides a comprehensive description of different zones of interest using viability tools. The second achievement is a thorough mathematical analysis of Pontryagin extremals for the aforementioned problem allowing to obtain an explicit closed-loop feedback optimal control. All our theoretical results are numerically illustrated for a further understanding of the geometrical features and scenarios.

Keywords: Optimal control; SIR; Pontryagin principle; State constraints; Viability; Epidemics; Feedback control

1 Introduction

The optimal control of epidemics ([AAM92, Beh00, HD11, Mar15], ...) has been awakening lots of interest recently, and even more so after the start of the COVID-19 pandemic – see for example [AAL20, KS20, Ket21]. In this paper, we focus on the optimal "contact control" of the following two-dimensional SIR model [KMW27] :

$$\begin{cases} \dfrac{ds}{dt}(t) = -b(t)\,s(t)\,i(t) \\ \dfrac{di}{dt}(t) = b(t)\,s(t)\,i(t) - \gamma i(t) \end{cases} \tag{1.1}$$

(the third "recovered class" of SIR being classically obtained by using the conservation of mass).

In our problem, the control parameters are chosen as $b \in B := [\beta_*, \beta]$ for some $0 < \beta_* < \beta$. The derivatives in (1.1) are meant in a distibutional sense and the trajectories are constructed from *admissible controls* $b \in \mathbb{L}^0(\mathbb{R}; B)$ (B-valued Borel-measurable functions). Whenever the control b and the initial conditions $s(0) = s_0,\ i(0) = i_0$ are fixed, the unique solution to (1.1) will also be denoted by $\left(s^{s_0,i_0,b}, i^{s_0,i_0,b}\right)$. The reader is invited to note that the trajectory can be computed for positive and negative times $t \in \mathbb{R}$.

Our aim is to consider the optimization problem of minimizing an effort-related *cost functional* of the form

$$J(t_f, b) = \int_0^{t_f} [\lambda_1 + \lambda_2 (\beta - b(t))] \ dt, \tag{1.2}$$

for two fixed non-negative weight parameters λ_1, λ_2, under an *intensive care unit (ICU) constraint* on the number of infected [KK20, MSW20]

$$i(t) \leq i_M,$$

for some fixed $0 < i_M < 1$, and starting from initial positions

$$(s_0, i_0) \in \mathbb{T}^{i_M} := \left\{(s, i) \in \mathbb{R}_+^2 : \ s + i \leq 1, \ i \leq i_M\right\}.$$

In this model, the prescribed level i_M represents an upper bound to the capacity of the health-care system to treat infected patients. The time t_f should be seen as a monitoring horizon. The λ_1 cost is related to the (psychological) stress of the epidemics and λ_2 is an economic cost of confinement. The second part $\lambda_2 (\beta - b)$ depends on the severity of the confinement measures, $b = \beta$ meaning no confinement, whileas $b = \beta_*$ meaning maximal confinement that avoids economic shut-down.

Even for the simplest cost (1.2), the decision maker is faced with hard fundamental choices: to aim for eradication of the epidemics [BBSG17, BBDMG19], which corresponds to a target $i(t_f) = \varepsilon$, where $\varepsilon < i_0$ is very small, or merely for a "modus vivendi", which could be modeled via the "*safe zone/no-effort zone constraint*" in the spirit of [ACMM+21]. More precisely, this no-effort condition requires that the trajectory (1.1) controlled with β satisfy the ICU constraint:

$$(s(t_f), i(t_f)) \in \mathcal{A} := \left\{(s, i) \in \mathbb{R}_+^2 : \ i^{s,i,\beta}(t) \leq i_M, \ \forall t \geq 0\right\}.$$

We will see shortly after that this constraint can be regularized and it can be completely described cf. Theorem 2.3, assertion 5, b. In our precise statement, we will ask a final qualification stricter than $(s(t_f), i(t_f)) \in \mathcal{A}$, i.e. $s(t_f) < \frac{\gamma}{\beta}$. On one hand, the method we develop here can be applied to a wide range of real-life systems and, in particular, it can be extended to more complex models of epidemiology. Roughly said, we exhibit a set $\mathcal{B}_0$ that is the largest subset of the viability kernel in which the ICU constraint can be fully saturated, while the remaining of the manageable configurations ($\mathcal{B}_1$) describes a backward-in-time "trace" of $\mathcal{B}_0$. We are currently working on a paper detailing this tool-kit in the identification of safe zones in other epidemic models.

Let us further emphasize that this kind of backward-in-time invariance is of particular relevance in control problems under state constraints (cf. [FP00]), especially when the boundary may be degenerate. It allows extensions of the control problems to discontinuous costs and the use of verification theorems coming from the Hamilton-Jacobi-Bellman theory.
On the other hand, the proof of Theorem 2.3 makes explicit how the curves describing the boundaries ($\partial\mathcal{A}$ and $\partial\mathcal{B}$) can be obtained and are generalizable to different models.

The main tool for solving the problem is Pontryagin's maximum/(minimum) principle (see for example [Pon18, BP03, BP05, SL12, SM17, DGI18, SS78]). However, unlike the usual approach for SIR problems consisting in writing down Pontryagin's conditions and using a numerical method to find the/an optimal control, we provide a thorough analysis of extremals and optimal control(s). The main features of the work are the following.

1. It provides a comprehensive, self-contained description of different zones of interest (in Section 2) via classical viability and invariance tools.
2. It provides, in Section 4 a work-through mathematical analysis of the features of Pontryagin extremals for the considered problems. As opposed to an important part of the literature on the subject, we look deeply into the extremals instead of using numerical methods to deal with the (dual) co-state problem. We would like to emphasize that our Problems 3.1 and 3.2 deal with both *state constraints* and *target constraints*, and, for 3.1, due to a strict constraint, existence of optimal controls is obtained *a posteriori*.

The analysis of the optimality via Pontryagin extremals allows to prove that the optimal control in such problems is unique, which is only assumed when directly applying numerical tools. Furthermore, we show that, for the problem we consider, an "act only when imperative" confinement is, indeed, optimal (see Theorem 5.6).

The paper is organized as follows. In Section 2 we define the different sets of initial configurations according to the controls one can employ. Precise characterizations of these regions (in terms both of reachable sets and of the explicit separating borders) are provided in Theorem 2.3. The proof of this theorem, relaying on viability tools is relegated to an appendix (Section 8). Section 3 provides the description of two types of control problem under investigation with fixed time horizon. The first (Problem 3.1) deals with a relaxed constraint on the infections and a strict no-effort target. The second problem (Problem 3.2) relaxes the no-effort target and approximates the previous one. Section 4 presents general considerations on Pontryagin's principle in the presence of state and target constraints. In Section 5, we provide a detailed analysis of the shape of Pontryagin extremals in each region (see Theorem 5.1 and Theorem 5.2). After having done it for Problem 3.2, we present, by extrapolation, the implications on Problem 3.1. The optimality is gathered in Theorem 5.6 with emphasis on the *a posteriori* continuity of the value function(s), on zone-related feedback forms and uniqueness considerations. Section 6 provides numerical illustrations in the different region-related scenarios. The last two sections are devoted to conclusions and appendix.

Notations: Throughout the paper, we will make use of the following notations:

1. given an interval $I \subset \mathbb{R}$ and a (subspace of a) metric space B, $\mathbb{L}^0(I;B)$ will stand for the family of Borel-measurable B-valued functions whose domain is I;
2. the usual 0/1-valued indicator function of sets will be denoted by $\mathbf{1}$, while the $0/\infty$-version is denoted by χ.

2 No-Effort, All-Control and Feasible Zones

Before getting into optimality considerations, let us take some time in order to describe regions of feasibility and the no-effort zone by means of viability tools. For our readers' sake, and in an effort of providing a self-contained material, we gather these notions of *viability, viability kernel, invariance, capture basin* in the Appendix.

Definition 2.1 *1. We call a* no-effort zone *the set* $\mathcal{A} \subseteq \mathbb{T} := \left\{(s_0, i_0) \in \mathbb{R}^2_+ : \ s_0 + i_0 \leq 1\right\}$ *of all initial configurations* $(s_0, i_0) \in \mathbb{T}$ *such that the associated trajectories controlled with* β *satisfy the state constraint* $i^{s_0,i_0,\beta}(t) \leq i_M$*, for all* $t \geq 0$.

2. We call an all-control zone *the family* $\mathcal{A}^0$ *of all initial data such that, for every* $b \in \mathbb{L}^0\left(\mathbb{R}; [\beta_*, \beta]\right)$*, one has* $i^{s_0,i_0,b}(t) \leq i_M$*, for all* $t \geq 0$.

3. We call feasible *(or* viable*, see next remark) zone the set* $\mathcal{B} \subseteq \mathbb{T}$ *of all initial configurations* $(s_0, i_0) \in \mathbb{R}^2$ *for which there exists a control* $b \in \mathbb{L}^0\left(\mathbb{R}; [\beta_*, \beta]\right)$ *keeping the associated trajectory* $i^{s_0,i_0,b} \leq i_M$.

Remark 2.2 1. The feasible or viable zone $\mathcal{B}$ is the maximal set of initial configurations on which at least one trajectory satisfies the afore-mentioned constraint. All-control zones are the sets on which constrained and unconstrained problems give the same value (independently on the cost). No-effort zones are useful in target problems.

2. The reader will easily note that $\mathcal{A}$ is the viability kernel of

$$\mathbb{T}^{i_M} := \left\{(s_0, i_0) \in \mathbb{T} : \ i^{s_0,i_0,\beta} \leq i_M\right\}, \tag{2.1}$$

when the flow is only controlled with $B = \{\beta\}$.

3. Similarly, $\mathcal{B}$ is the viability kernel of $\mathbb{T}^{i_M}$ when $B = [\beta_*, \beta]$. This is why feasible zones are viable in the sense of the preceding subsection (see also Remark 8.2).

4. The set $\mathcal{A}^0$ is the largest subset of $\mathbb{T}^{i_M}$ that is time-invariant (using $B = [\beta_*, \beta]$). It can be referred to as an invariance kernel.
5. Finally, the simple inclusion $\mathcal{A}^0 \subseteq \mathcal{A} \subseteq \mathcal{B}$ holds true.

The following result, whose proof is based solely on viability theory, provides an extensive characterization of the previously-introduced zones.

Theorem 2.3 *1. The all-control zone $\mathcal{A}^0$ contains the set*

$$\mathcal{A}_0 := \Big([0, \frac{\gamma}{\beta}] \times [0, i_M]\Big) \cap \mathbb{T},$$

as well as $[0,1] \times \{0\}$.

2. The viable zone $\mathcal{B}$ contains the set

$$\mathcal{B}_0 := \Big([0, \frac{\gamma}{\beta_*}] \times [0, i_M]\Big) \cap \mathbb{T},$$

as well as $[0,1] \times \{0\}$.

3. One has the explicit (capture basin) characterizations

(a) $\mathcal{A} = \mathcal{A}^0 \cup \mathcal{A}_1$, *where*

$$\mathcal{A}_1 := \Big\{\big(s^{\frac{\gamma}{\beta}, i_0, \beta}(-t), i^{\frac{\gamma}{\beta}, i_0, \beta}(-t)\big) : \ \big(\frac{\gamma}{\beta}, i_0\big) \in \mathcal{A}_0, \ t \geq 0\Big\} \cap \mathbb{T}.$$

(b) $\mathcal{B} = \mathcal{B}_0 \cup ([0,1] \times \{0\}) \cup \mathcal{B}_1$, *where*

$$\mathcal{B}_1 := \Big\{\big(s^{\frac{\gamma}{\beta_*}, i_0, b}(-t), i^{\frac{\gamma}{\beta_*}, i_0, b}(-t)\big) : \ \big(\frac{\gamma}{\beta_*}, i_0\big) \in \mathcal{B}_0, \ b \in \mathbb{L}^0\left(\mathbb{R}; [\beta_*, \beta]\right), \ t \geq 0\Big\} \cap \mathbb{T}.$$

4. If $(s_0, i_0) \in \mathcal{B}$ (resp. $\mathcal{A}^0$ or $\mathcal{A}$), then, for every $0 < i_1 \leq i_0$, $(s_0, i_1) \in \mathcal{B}$ (resp. $\mathcal{A}^0$ or $\mathcal{A}$). As a consequence, there exist maps $\Phi_{\mathcal{A}^0}$, $\Phi_{\mathcal{A}}$ and $\Phi_{\mathcal{B}}$ from $[0,1]$ to $[0, i_M]$ such that

$$\begin{aligned} \mathcal{A}^0 &= \big\{(s,i) \in \mathbb{T} : \ i \leq \Phi_{\mathcal{A}^0}(s)\big\}, \ \mathcal{A} = \big\{(s,i) \in \mathbb{T} : \ i \leq \Phi_{\mathcal{A}}(s)\big\}, \\ \mathcal{B} &= \big\{(s,i) \in \mathbb{T} : \ i \leq \Phi_{\mathcal{B}}(s)\big\}. \end{aligned} \tag{2.2}$$

5. The sets in (2.2) can be expressed using the following regular (piecewise C^1) functions

(a)

$$\Phi_{\mathcal{B}}(s) = \begin{cases} i_M, & \text{if } s \leq \frac{\gamma}{\beta_*}, \\ i_M - s + \frac{\gamma}{\beta_*} + \frac{\gamma}{\beta_*} \log\big(\frac{\beta_* s}{\gamma}\big), & \text{if } s \in \big(\frac{\gamma}{\beta_*}, s_M^*\big), \\ 0, & \text{otherwise,} \end{cases} \tag{2.3}$$

where

$$i_M - s_M^* + \frac{\gamma}{\beta_*} + \frac{\gamma}{\beta_*} \log\big(\frac{\beta_* s_M^*}{\gamma}\big) = 0.$$

(b) $\mathcal{A}^0 = \mathcal{A}$ *and*

$$\Phi_{\mathcal{A}}(s) = \begin{cases} i_M, & \text{if } s \leq \frac{\gamma}{\beta}, \\ i_M - s + \frac{\gamma}{\beta} + \frac{\gamma}{\beta} \log\big(\frac{\beta s}{\gamma}\big), & \text{if } s \in \big(\frac{\gamma}{\beta}, s_M\big), \\ 0, & \text{otherwise,} \end{cases} \tag{2.4}$$

where

$$i_M - s_M + \frac{\gamma}{\beta} + \frac{\gamma}{\beta} \log\big(\frac{\beta s_M}{\gamma}\big) = 0.$$

6. *If $s_M^* \le 1$ then, starting from $(s_0, i_0) \in \partial\mathcal{B}$ with $s_0 > 0$ and $i_0 > 0$, the only viable controls are identically β_* on $\big[0, \tau_0^{s_0,i_0}\big)$, where*

$$\tau_0^{s_0,i_0} := \inf\big\{t \ge 0 : \; s^{s_0,i_0,\beta_*}(t) \le \frac{\gamma}{\beta_*}\big\},$$

with the convention $\inf \varnothing = +\infty$.

The proof is relegated to Subsection 8.2, allowing the reader to get familiarized with the viability notions. The pictures show the intersection with the triangle $\mathbb{T}$ of the qualitative graphs of the viability boundaries $\Phi_{\mathcal{A}}$ and $\Phi_{\mathcal{B}}$ in the cases in which $s_M^* < 1$ and $s_M^* > 1$. From them, the reader can recognize the sets $\mathcal{A}_0$ and $\mathcal{B}_0$, $\mathcal{A}$ and $\mathcal{B}$.

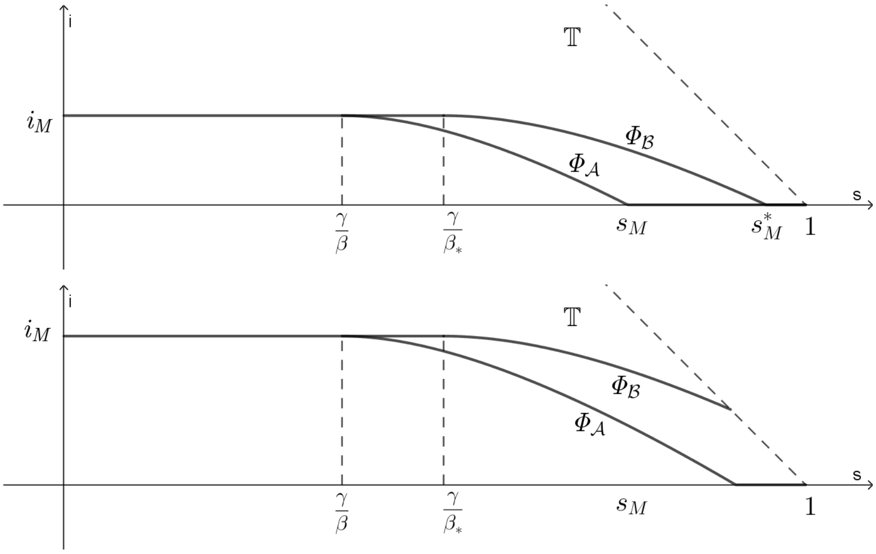


Figure 1: Viability regions

From now on, and unless otherwise stated, we assume that $s_M^* \le 1$. This implies that

$$\big\{(s,i) \in [0,1]^2 \;:\; i \le \Phi_{\mathcal{B}}\big\} \subseteq \mathbb{T}.$$

Remark 2.4 1. Assertion 3 tells us that the feasible/viable region $\mathcal{B}$ is the capture basin of $\mathcal{B}_0$ (completed by the stationary regimes $[0,1] \times \{0\}$).

2. Similarly, the no-effort region $\mathcal{A}$ is the capture basin of $\mathcal{A}_0$ (completed by the stationary regimes $[0,1] \times \{0\}$).

3 A Strict No-Effort Formulation

Let us begin with an optimal control problem on a finite and fixed time horizon $[0, t_f]$. To lighten notation, from this point on and unless explicitly stated otherwise, the states $i^{s_0,i_0,b}$ and $s^{s_0,i_0,b}$, corresponding to the specific initial conditions (s_0, i_0) and control b, will be simply denoted by i and s, respectively.

Problem 3.1 *Given a* fixed $t_f > 0$ *(large enough) and the initial data $(s_0, i_0) \in \mathcal{B}$ (i.e. $i_0 \le \Phi_{\mathcal{B}}(s_0)$) in the feasible region, minimize, over all admissible controls $b \in \mathbb{L}^0\big(\mathbb{R}; [\beta_*, \beta]\big)$,*

- *the* cost functional
$$J(t_f, b) = \int_0^{t_f} \big[\lambda_1 + \lambda_2 \left(\beta - b(t)\right)\big]\, dt; \tag{3.1}$$

- *under the* ICU constraint *on the trajectory of* (1.1)

$$\big(s(t), i(t)\big) \in \mathcal{B},\ \forall t \in [0, t_f]\,; \tag{3.2}$$

- *under a* strict no-effort constraint $s(t_f) < \frac{\gamma}{\beta}$.

In the sequel, we refer to this formulation as problem $\mathcal{P}$.

Due to the strict constraint $s(t_f) < \frac{\gamma}{\beta}$, the existence of an optimal control does not follow from the embedding in a lower semicontinuous functional. However, should such an optimal control exist, it is also the optimal solution of the following problem (for some $\varepsilon_0 > 0$ and all $\varepsilon < \varepsilon_0$).

Problem 3.2 *Given a* fixed $t_f > 0$ *(large enough) and* $\varepsilon > 0$ *(small enough) and the initial data* $(s_0, i_0) \in \mathcal{B}$ *(i.e.* $i_0 \le \Phi_{\mathcal{B}}(s_0)$*) in the feasible region, minimize, over all admissible controls* $b \in \mathbb{L}^0(\mathbb{R}; [\beta_*, \beta])$*,*

- *the* cost functional (3.1)*;*
- *under the* ICU constraint *on the trajectory* (3.2)*;*
- *under a* no-effort constraint $s(t_f) \le \frac{\gamma}{\beta} - \varepsilon$.

In the sequel, we refer to this formulation as problem $\mathcal{P}_\varepsilon$.

Remark 3.1 Form a mathematical point of view, because of the properties of indicator functions, instead of directly working with a half-open set $s(t_f) \in [0, \frac{\gamma}{\beta})$, we prefer working with closed sets $s(t_f) \in [0, \frac{\gamma}{\beta} - \varepsilon_0]$ by approximating $\frac{\gamma}{\beta}$ from the left. Let us note that

1. the ε-formulation (as opposed to merely writing an interior condition $\big(i(t_f), s(t_f)\big) \in \overset{\circ}{\mathcal{A}}$) is due to the fact that the lower-semicontinuous penalty $\chi_{s(t_f) \le \frac{\gamma}{\beta} - \varepsilon}$ is needed in order to guarantee the existence of optimal controls for every ε;
2. non-emptiness of the set of viable controls satisfying the no-effort constraint is guaranteed for t_f large enough by using, for example, (5.13);
3. furthermore, in view of the comment preceding the introduction of problem $\mathcal{P}_\varepsilon$, we can restrict our attention to controls b satisfying $s(t_f) < \frac{\gamma}{\beta} - \varepsilon$.

Whenever needed, we will specify the initial data $(s_0, i_0) \in \mathcal{B}$ by writing $J(t_f, b; s_0, i_0)$. The value functions of $\mathcal{P}$ and $\mathcal{P}_\varepsilon$ will be denoted by $V(t_f; s_0, i_0)$, respectively $V_\varepsilon(t_f; s_0, i_0)$.

Proposition 3.2
1. *For every* $\varepsilon > 0$, $(s_0, i_0) \in \mathcal{B}$ *with* $i_0 > 0$ *and every* t_f *large enough, the problem* $\mathcal{P}_\varepsilon$ *has an optimal solution.*
2. *Furthermore, for every* $\varepsilon > 0$, $(s_0, i_0) \in \mathcal{B}$*, with* $i_0 > 0$*, for every* t_f *large enough and every* $\frac{i_0}{2} > \eta > 0$ *such that* $V_\varepsilon(t_f; s_1, i_1) < \infty$ *for every* $(s_1, i_1) \in \mathcal{B} \cap \mathbb{B}_\eta(s_0, i_0)$*, the value function* $V_\varepsilon(t_f; \cdot, \cdot)$ *is lower semi-continuous on* $\mathcal{B} \cap \mathbb{B}_\eta(s_0, i_0)$ *(with* $\mathbb{B}_\eta$ *being the* η*-radius open ball);*
3. $V = \inf_{\varepsilon > 0} V_\varepsilon$.

Proof. The existence of optimal solutions to problem $\mathcal{P}_\varepsilon$ with fixed $\varepsilon > 0$ is a standard matter. It follows by the Direct Method of the Calculus of Variations as, for instance, in [Fre] or using [Cla13, Theorem 23.11] by noting that $\mathcal{B}$ and $\left\{(s, i) \in \mathcal{B} :\ s \le \frac{\gamma}{\beta} - \varepsilon\right\}$ are compact ([Cla13, Theorem 23.11 (a), (c), (e)]), $B := [\beta_*, \beta]$ is convex and compact ([Cla13, Theorem 23.11 (b), (f.i)]), the running cost $f(b) := \lambda_1 + \lambda_2(\beta - b)$ is lower-bounded (by λ_1), continuous and convex ([Cla13, Theorem 23.11 (d)]).

The final state constraints in the problem $\mathcal{P}_\varepsilon$ can classically be dropped by considering a lower semi-continuous final cost $g(s, i) := \chi_{s \le \frac{\gamma}{\beta} - \varepsilon}$, where χ stands for the usual $0/\infty$-valued indicator function. As a consequence, one gets the lower semi-continuity of V_ε. This result is standard, and further references for state-constrained dynamics can be found in [FP00, Proposition 4] (see also [BFZ11, Remark 2.2]).

The last assertion is straightforward. For every $\varepsilon > 0$, V does not exceed V_ε, as it minimizes over a larger set, implying the inequality $V \leq \inf_{\varepsilon>0} V_\varepsilon$. On the other hand, if (s_0, i_0) is fixed and b is an admissible control for $\mathcal{P}$, it is admissible for $\mathcal{P}_{\varepsilon_0}$ for some $\varepsilon_0 > 0$. As such, $J(t_f, b; s_0, i_0) \geq V_{\varepsilon_0}(s_0, i_0) \geq \inf_{\varepsilon>0} V_\varepsilon(s_0, i_0)$. The conclusion follows by taking the infimum over any admissible b. □

4 Pontryagin Approach. General Considerations

To write necessary conditions of optimality let us introduce the adjoint variables $p_0 \geq 0$, $p_s, p_i \in \mathbb{R}$, and the pre-Hamiltonian

$$H(t, b, s, i, p_0, p_s, p_i) = p_0 f_0(s, i, b) + p_s f_s + p_i f_i$$

where $f_0(s, i, b) = \lambda_1 + \lambda_2 (\beta - b)$ is the running cost function and $f_s = -sbi$, $f_i = sbi - \gamma i$ are the dynamics of the state equations. After some manipulations, the pre-Hamiltonian turns out to be

$$H(t, b, s, i, p_0, p_s, p_i) = p_0 (\lambda_1 + \lambda_2 (\beta - b)) + \eta sbi - \gamma p_i i$$

where $\eta := p_i - p_s$. The usage of η is quite natural. Nevertheless, the idea that two adjoint variables can be summarized into a single new variable is already in [Beh00] and used also in [KS20] and [Fre].

In the sequel we use a constrained version of Pontryagin's theorem developed in [BdlV10, BDLVD13]. In particular, we refer to [BDLVD13] for the definition of the space of functions with bounded variation $BV([0, t_f])$ which is given by extending functions in a constant way on an open interval containing $[0, T]$. We adopt here also the notation used in [BdlV10, BDLVD13] of denoting the distributional derivative of a BV function f (which is a measure) by df, instead than $\dot{f}$ that we reserve to measures which are absolutely continuous with respect to Lebesgue as, for instance, in the state equations.

4.1 Optimality Conditions for Problem $\mathcal{P}_\varepsilon$

By Pontryagin's theorem, given an optimal solution (s, i, b), there exist a constant $p_0 \in \{0, 1\}$, adjoint state real functions p_s, $p_i \in BV([0, t_f])$, a multiplier for the state constraint $\mu \in BV([0, t_f])$ with a nondecreasing representative (hence with measure distributional derivative $d\mu \geq 0$) such that $\mu(t_f^+) = 0$ (recall that the functions are extended outside $[0, t_f]$), and a multiplier for the final state constraint $p_1 \geq 0$ that satisfy

(P1) the non-degeneration property

$$p_0 + d\mu([0, t_f]) + p_1 > 0; \tag{4.1}$$

indeed, our problem $\mathcal{P}_\varepsilon$ corresponds to problem (P) in [BdlV10] with a final condition $\Phi(y_T) \in K$ where $y_T = \big(s(t_f), i(t_f)\big)$, $\Phi(s, i) = s - \frac{\gamma}{\beta} + \varepsilon$ and $K = (-\infty, 0]$; the normal cone to K is given by $N_K(0) = [0, \infty)$ implying $p_1 \geq 0$; furthermore, $N_K(x) = \{0\}$, for all $x < 0$, implying $p_1 = 0$ whenever $s(t_f) < \frac{\gamma}{\beta} - \varepsilon$;

(P2) the complementarity condition

$$\int_{[0,t_f]} \big(i(t) - i_M\big)\, d\mu(t) = 0; \tag{4.2}$$

(P3) the conjugate equations with transversality conditions

$$\begin{cases} dp_s = -\eta bi, \\ dp_i = -\big(\eta bs - \gamma p_i\big) - d\mu, \\ p_s(t_f) = p_1,\ p_i(t_f^+) = 0, \end{cases}$$

which hold as equalities between measures on $[0, t_f]$; we observe that the boundary condition for the costate p_i is given on the right limit in t_f, since p_i could be discontinuous in t_f if the measure $d\mu$ charges this point; on the contrary, p_s is continuous in $[0, t_f]$ since the derivative is absolutely continuous with respect to the Lebesgue measure;

(P4) the minimality property

$$H\big(t, b(t), s(t), i(t), p_0, p_s(t), p_i(t)\big) = \inf_{b \in B} H\big(t, b, s(t), i(t), p_0, p_s(t), p_i(t)\big),$$

for almost every $t \in [0, t_f]$;

(P5) the conservation property

$$H\big(t, b(t), s(t), i(t), p_0, p_s(t), p_i(t)\big) = k,$$

with k constant, for a.e. $t \in [0, t_f]$ (see [Bon20, Lemma 7.7]).

Definition 4.1 *We recall that a* Pontryagin extremal *for problem P_ε is any control $b \in \mathbb{L}^0(\mathbb{R}; [\beta_*, \beta])$ that satisfies the constraints of problem P_ε and conditions (P1)-(P5).*

4.2 Remarks and consequences

Throughout all this section we assume that $(s = s^{s_0,i_0,b}, i = i^{s_0,i_0,b}, b)$ be a local solution of the control Problem 3.2 (with initial data $(s_0, i_0) \in \mathcal{B}$ and write consequences of Pontryagin's necessary conditions.

We have the following consequences.

(C1) By (P3) the jump condition $[\eta(t)] = [p_i(t)] = -[\mu(t)] \leq 0$ holds for every $t \in [0, t_f]$ (where the inequality follows by the fact that μ is non-decreasing); The functions μ, p_i and η have the same set of discontinuity points D; since such functions have bounded variation, the set D is at most countable; in particular it is a Lebesgue-null set;

(C2) By the conservation of the Hamiltonian

$$p_0\big(\lambda_1 + \lambda_2(\beta - b(t))\big) + \eta(t)b(t)s(t)i(t) - \gamma p_i(t)i(t) = k \quad \text{for a.e. } t \in [0, t_f]. \tag{4.3}$$

(C3) (P3) implies that

$$d\eta = dp_i - dp_s = \eta b(i - s) + \gamma p_i - d\mu, \tag{4.4}$$

or, owing to (4.3),

$$d\eta = \eta b i + \frac{p_0\big(\lambda_1 + \lambda_2(\beta - b(t))\big) - k}{i} - d\mu. \tag{4.5}$$

Remark 4.2 We remark here that, differently from the case in which state constraints are not considered, the computation of the constant k is not a straightforward consequence of the transversality conditions, because, as already observed, p_i is allowed to jump in the final point t_f.

Since the cost is linear in the control b, the minimum value of the Hamiltonian on $K = [\beta_*, \beta]$ is achieved when $b \in \{\beta_*, \beta\}$. Hence, setting the *switching function*

$$\psi := \eta s i$$

the optimal control has to satisfy

$$b(t) = \begin{cases} \beta, & \text{if } \psi(t) < p_0\lambda_2, \\ \beta_*, & \text{if } \psi(t) > p_0\lambda_2. \end{cases} \tag{4.6}$$

for almost every $t \in [0, t_f]$, where ψ denotes any pointwise representative of the switching function.

Proposition 4.3 *If $k - p_0\lambda_1 \geq 0$ in (4.3) and $p_1 = 0$ then we have $\eta(t) \geq 0$ for almost every $t \in [0, t_f]$.*

Proof. We begin with noting that, since μ is nondecreasing, $d\mu \geq 0$ and, therefore, (4.5) implies

$$d\eta \leq \eta bi + \frac{p_0\big(\lambda_1 + \lambda_2\,(\beta - b(t))\big) - k}{i}. \tag{4.7}$$

Without loss of generality, we assume η, and thus ψ, to be right-continuous.

Arguing by contradiction, let us assume that there exists $t \in (0, t_f)$ such that $\eta(t) = \eta(t^+) < 0$. Since the switching function $\psi = \eta si$ has the same sign as η, and since $p_0 \geq 0$, we have $\psi(t^+) < p_0\lambda_2$. On the other hand, since ψ is right-continuous, one is able to find some $\varepsilon > 0$ such that $\psi < p_0\lambda_2$, and hence $b = \beta$, a.e. in $J := (t, t+\varepsilon)$. Owing to (4.7) with $b = \beta$, in J we have

$$d\eta \leq \eta bi - \frac{k - p_0\lambda_1}{i}.$$

Under the standing assumption of our assertion i.e. $k \geq p_0\lambda_1$, we have that

$$d\eta \leq \eta bi$$

and $\eta bi < 0$ in J, and therefore η is negative and decreasing in J.

As we shall see, this implies that $\eta < 0$, and decreasing, in (t, t_f), which implies $\eta(t_f^-) < 0$ and hence $\eta(t_f^+) < 0$ since (by (C1)) it cannot have increasing jumps, thus contradicting the fact that $\eta(t_f^+) = -p_1 = 0$ by the transversality conditions and the assumption $p_1 = 0$.

To prove the claim that $\eta(s) < 0$ in (t, t_f), let us set

$$t_0 := \sup\{s \in (t, t_f] \ : \ \eta(s) < 0\}.$$

Assume, again by contradiction, that $t_0 < t_f$. Since $\eta \in BV$, then there exists the left limit $\eta(t_0^-) \leq 0$. Since η has no positive jumps in $[0, t_f]$, it follows that $\eta(t_0^+) \leq 0$. On the other hand, $\eta(t_0^+) < 0$ would contradict the choice of t_0. It follows that $\eta(t_0) = 0$. On the other hand, in the interval (t, t_0) we have that $\eta < 0$, hence $\psi < 0$. Then, in (t, t_0) we have that $\psi < p_0\lambda_2$, hence $b = \beta$ a.e., hence $d\eta \leq \eta bi < 0$, hence η is decreasing, which contradicts $\eta(t_0) = 0$. Our assumption on t_0 is wrong such that $t_0 = t_f$ and $\eta(s) < 0$ in (t, t_f). □

Proposition 4.4 *If $s(t_f) < \frac{\gamma}{\beta} - \varepsilon$ then*

1. *$k = p_0\lambda_1$ in (4.3),*
2. *$\eta(t) \geq 0$ for almost every $t \in [0, t_f]$,*
3. *p_s is (continuous) nonincreasing and nonnegative in $[0, t_f]$.*

Proof. The assumption $s(t_f) < \frac{\gamma}{\beta} - \varepsilon$ implies $p_1 = 0$, as observed in (P1). Since s is continuous, then we have that $s < \frac{\gamma}{\beta}$ in an interval $(t_I, t_f]$ with $t_I < t_f$. This implies that $di = (sb - \gamma)i < \frac{\gamma}{\beta}(b - \beta)i \leq 0$, hence i is decreasing and therefore $i < i_M$ in $(t_I, t_f]$. Then, by complementarity, we have $d\mu((t_I, t_f]) = 0$ which implies that p_i is continuous in $(t_I, t_f]$.

- If $p_0 = 0$ then (4.3) implies

  $$k = \eta(t)b(t)s(t)i(t) - \gamma p_i(t)i(t) \text{ a.e. } t \in (t_I, t_f].$$

  By taking a sequence $t_n \to t_f$ on which the equality holds and taking the limit an $n \to \infty$ then we get $k = 0$, which proves *1.* in this case.
- If $p_0 = 1$ then $\psi(t_f) = 0 < p_0\lambda_2$ and, by continuity, we have $\psi(t) < p_0\lambda_2$ and hence $b(t) = \beta$ in a left neighborhood of t_f which we still call $(t_I, t_f]$; (4.3) implies

  $$k = \eta(t)\beta s(t)i(t) - \gamma p_i(t)i(t) + p_0\lambda_1 \text{ a.e. } t \in (t_I, t_f].$$

  By arguing as before, we get $k = p_0\lambda_1$, which proves *1.* also in this case.

Point *2.* follows by Proposition 4.3. Point *3.* comes from the previous point *2.*, the first adjoint equation and the final conditions.

□

Using (4.4), one easily proves the following result.

Proposition 4.5 *The distributional derivative of ψ is the measure given by*

$$d\psi = si\gamma p_s - si\, d\mu.$$

Theorem 4.6 *Let us assume that $\lambda_2 > 0$. Let $(t_1, t_2) \subset [0, t_f]$ be an interval in which $\psi = p_0\lambda_2$. Then the following hold true*

1. *if $p_0 = 1$, then $p_s(t) > 0$ and $i(t) = i_M$ for every $t \in (t_1, t_2)$;*
2. *if $p_0 = 0$, then we have that, either*
   (a) *i is constant in (t_1, t_2), or*
   (b) *$k = p_0\lambda_1$, $\eta \geq 0$ in $[0, t_f]$, $\eta = p_i = p_s = 0$ in $(t_1, t_f]$ and $d\mu([0, t_1]) > 0$.*

Proof. Let us consider the case when $p_0 = 1$ and prove the first assertion. First of all, the reader is invited to note that $\eta > 0$ in (t_1, t_2). Since ψ is constant, on this interval, it follows that $d\psi = 0$ in (t_1, t_2). Since s and i are strictly positive, by Proposition (4.5), we have

$$d\mu = \gamma p_s \text{ in } (t_1, t_2). \tag{4.8}$$

Then, using the complementarity condition (P2) and $\gamma > 0$, we have that

$$\int_{(t_1,t_2)} p_s(i_M - i(t))\, dt = 0. \tag{4.9}$$

Since i is continuous, in order to prove that $i = i_M$, it suffices to prove that $p_s > 0$.

Since $d\mu \geq 0$, from (4.8) we have

$$p_s \geq 0 \text{ in } (t_1, t_2). \tag{4.10}$$

Now we prove that the strict inequality holds. Suppose now, by contradiction, that there exists $t_0 \in (t_1, t_2)$ such that $p_s(t_0) = 0$. By the adjoint equation $dp_s = -\eta bi$ we have that $dp_s \leq 0$ in (t_1, t_2), hence $p_s(t) \leq 0$ for every $t \in (t_0, t_2)$ and, by (4.10), the equality holds in this interval. Then in (t_0, t_2) we would have $dp_s = 0$, hence $\eta = 0$ by the adjoint equation, and finally $\psi = 0 < p_0\lambda_2$, thus giving a contradiction. The first assertion is now completely proved.

Let us assume $p_0 = 0$ and prove *2*. In the interval (t_1, t_2), we have that $\psi = 0$, hence $\eta = 0$; by the adjoint equations then $dp_s = 0$, hence there exists a constant c such that $p_s = c$ on this interval. Since $0 = \eta = p_i - p_s$, we also get $p_i = c$. By conservation of the Hamiltonian we have

$$p_0\lambda_1 - \gamma ci(t) = k, \tag{4.11}$$

hence i is constant in (t_1, t_2) whenever $c \neq 0$.

In the case in which $c = 0$, by (4.11), we have $k = p_0\lambda_1 = 0$. By Proposition (4.3), it follows that $\eta \geq 0$ in $[0, t_f]$. Then $dp_s \leq 0$ and p_s is nonincreasing. Since p_s is zero on (t_1, t_2) and at the end point t_f, we must have $p_s = 0$ in $(t_1, t_f]$, hence $p_1 = 0$. By the first adjoint equation, it follows that $\eta = 0$ on $(t_1, t_f]$ which implies that $p_i = 0$ on $(t_1, t_f]$. By (4.5), we have $d\mu((t_1, t_f]) = -d\eta((t_1, t_f]) = 0$. Finally, the non degeneration condition (4.1) requires $d\mu\,([0, t_f]) > 0$, which implies $d\mu([0, t_1]) > 0$.

□

Proposition 4.7 *If $i = i_0 \in (0, i_M]$ in an interval (t_1, t_2) then there exists a positive constant k_s such that*

$$b(t) = \frac{\gamma}{s(t)} = \frac{\gamma}{k_s - \gamma i_0 t} \tag{4.12}$$

a.e. in the interval. The constant is given by

$$k_s = s(t_1) + \gamma i_0 t_1 = s(t_2) + \gamma i_0 t_2.$$

Moreover, since $b \in [\beta_*, \beta]$, we have

$$s(t_1) \leq \frac{\gamma}{\beta_*} \quad \text{and} \quad s(t_2) \geq \frac{\gamma}{\beta}.$$

Proof. By the second state equation with $i = i_0$ we immediately have $sb = \gamma$. Then the first becomes $ds = -\gamma i_0$. Integrating we obtain that there exists a constant k_s such that $s(t) = k_s - \gamma i_0 t$. The moreover part of the statement follows by imposing $\beta_* \leq \frac{\gamma}{s} \leq \beta$ and using the fact that s does not increase.

□

5 Pontryagin Extremals and Optimal Controls

Let us set, for $(s_0, i_0) \in \mathcal{A}$, the *reaching time*[*]

$$\bar{t}_f^{s_0,i_0} := \inf\big\{t \geq 0: \; s^{s_0,i_0,\beta}(t) = \frac{\gamma}{\beta}\big\}. \tag{5.1}$$

It is easy to see that this quantity is well defined, that is the set on the right hand side is non-empty. By continuity we have $s^{s_0,i_0,\beta}(\bar{t}_f^{s_0,i_0}) = \frac{\gamma}{\beta}$. If $i_0 > 0$, the state s is strictly decreasing, so that, $t_f > \bar{t}_f^{s_0,i_0}$ implies $\frac{\gamma}{\beta} - s^{s_0,i_0,\beta}(t_f) > 0$. Hence, for every $\varepsilon > 0$ small enough (i.e. $\varepsilon < \frac{\gamma}{\beta} - s^{s_0,i_0,\beta}(t_f)$) we have

$$s^{s_0,i_0,\beta}(t_f) < \frac{\gamma}{\beta} - \varepsilon.$$

As a consequence, the control $b \equiv \beta$ is admissible for problem $\mathcal{P}_\varepsilon$ for this choice of ε (and actually for every $\varepsilon' < \varepsilon$). The reader is invited to keep in mind these elements of reasoning whenever encountering the expression "$\varepsilon > 0$ small enough".

Theorem 5.1 *1. If $(s_0, i_0) \in \mathcal{A} \setminus \partial\mathcal{A}$, b is a Pontryagin extremal of $\mathcal{P}_\varepsilon$ and $s(t_f) < \frac{\gamma}{\beta} - \varepsilon$, then $b = \beta$ almost everywhere.*

2. If $(s_0, i_0) \in \mathcal{A}$ with $i_0 > 0$, and $t_f > \bar{t}_f^{s_0,i_0}$, then,

   (a) for every $\varepsilon > 0$ small enough, the optimal solution of $\mathcal{P}_\varepsilon$ is $b = \beta$, almost everywhere;

   (b) for every $\varepsilon > 0$ small enough, the restriction of the value function $V_\varepsilon(t_f; \cdot, \cdot)$ to $\mathcal{A}$ is continuous at (s_0, i_0);

   (c) The problem $\mathcal{P}$ admits the unique optimal control $b = \beta$ almost everywhere. The value is the constant $\lambda_1 t_f$.

Proof. Let us recall that the states $i^{s_0,i_0,b}$ and $s^{s_0,i_0,b}$, corresponding to the specific initial conditions (s_0, i_0) and control b, will be simply denoted by i and s, respectively.

First, let us note that if $(s_0, i_0) \in \mathcal{A} \setminus \partial\mathcal{A}$, then we have $i < i_M$ on $[0, t_f]$ for every choice of b among the admissible controls. Indeed, in the sense of distributions,

$$\frac{d}{dt}\big(i - \Phi_{\mathcal{A}}(s)\big) = bsi\left(\frac{\gamma}{\beta s} - 1\right)^{-} + (bs - \gamma)\, i.$$

The later quantity is non-positive. As a consequence, $i(t) \leq \Phi_{\mathcal{A}}(s(t)) + i_0 - \Phi_{\mathcal{A}}(s_0) < \Phi_{\mathcal{A}}(s(t))$ which implies the desired inequality. This implies a series of consequences. Indeed, we get

1. $d\mu([0, t_f]) = 0$, by complementarity and the fact that $i < i_M$ on $[0, t_f]$;
2. $p_1 = 0$, as observed in (P1), since $s(t_f) < \frac{\gamma}{\beta} - \varepsilon$;
3. $p_0 = 1$, by the previous items and the non-degeneration condition (4.1);
4. $\psi = p_0\lambda_2$ in a subinterval $(t_1, t_2) \subset [0, t_f]$ is impossible, due to Theorem 4.6 (otherwise, we would have $i = i_M$, which is a contradiction);
5. by the first item and *3.* of Proposition (4.4), it holds $d\psi = si\gamma p_s \geq 0$ and hence ψ is continuous and non-decreasing;

[*]Note that in the statement of Theorem 5.6, equation (5.14), the reaching time has been redefined using the same notation. When restricted to $\mathcal{A}$ the two definitions turn out to be equivalent.

6. by the second item, $\psi(t_f) = 0$.

By assumption, there exists some interval $[t_I, t_f]$ on which $s < \frac{\gamma}{\beta}$. Owing to Proposition 4.7, it is clear that i cannot be constant on any interval $(t_1, t_2) \subset [t_I, t_f]$. Combined with $\psi(t_f) = 0$ and the monotonicity of ψ, one gets $\psi(t) < 0 \leq p_0\lambda_2$ for all $t \in (0, t_f)$ and the first assertion follows by (4.6).

For the second assertion point (a), let us fix $t_f > \bar{t}_f^{s_0,i_0}$ and $0 < \varepsilon < \frac{\gamma}{\beta} - s^{s_0,i_0,\beta}(t_f)$.

- If $(s_0, i_0) \notin \partial\mathcal{A}$, and b is a Pontryagin extremal for $\mathcal{P}_\varepsilon$ (with horizon t_f), it is also a Pontryagin extremal for $\mathcal{P}_{\frac{\varepsilon}{2}}$ with the same horizon. Moreover, $s(t_f) \leq \frac{\gamma}{\beta} - \varepsilon < \frac{\gamma}{\beta} - \frac{\varepsilon}{2}$. By the first assertion, it follows that $b = \beta$ (which is admissible for $\mathcal{P}_\varepsilon$ by the choice of ε). Since this is the only extremal, it follows that it is the optimal control for $\mathcal{P}_\varepsilon$.
- Let us fix $(s_0, i_0) \in \partial\mathcal{A}$. We define $\delta := \inf\{t \geq 0 : (s(t), i(t)) \notin \partial\mathcal{A}\}$. Clearly, δ is well defined and $\delta < t_f$; indeed, since $s(t_f) < \frac{\gamma}{\beta} - \varepsilon$ then i is strictly decreasing in the last part of the time horizon, hence $i(t_f) < i_M$, which implies $(s(t_f), i(t_f)) \notin \partial\mathcal{A}$. It is also clear by the previous point that $b = \beta$ a. e. on (δ, t_f). Furthermore, the only control keeping the solution on ∂A is again β. It follows that $b = \beta$ (almost surely).

By (a) and the continuity of $(s, i) \mapsto \frac{\gamma}{\beta} - s^{s,i,\beta}(t_f)$ (for $i > 0$) there exists a neighbourhood U of (s_0, i_0) such that the restriction of $V_\varepsilon(t_f; \cdot, \cdot)$ to $U \cap \mathcal{A}$ is the constant $\lambda_1 t_f$, hence proving (b).

Let us prove (c). Since $b \leq \beta$, one has $J(t_f, b) \geq \lambda_1 t_f$. On the other hand, we have $J(t_f, b) = \lambda_1 t_f$ if and only if $b = \beta$ a.e.. On the other hand, the assumption $t_f > \bar{t}_f^{s_0,i_0}$ implies $s^{s_0,i_0,\beta}(t_f) < \gamma/\beta$ which means that $b \equiv \beta$ is admissible, hence optimal and unique.

□

Theorem 5.2 *Let $\lambda_2 > 0$. Let us assume that $0 < i_0$, and $0 < s_0$ are such that $(s_0, i_0) \in \mathcal{B} \setminus \mathcal{A}$. Assume moreover that b is a Pontryagin extremal for Problem $\mathcal{P}_\varepsilon$ such that $s(t_f) < \frac{\gamma}{\beta} - \varepsilon$.*

1. *The ICU* saturation time

$$\tau_1 := \sup\{t \in (0, t_f] \ : \ i < i_M \text{ in } [0, t)\}$$

if the set on the right hand side is nonempty and $\tau_1 := 0$ if it is empty, satisfies $\tau_1 < t_f$.

2. *If $(s_0, i_0) \in \mathcal{B}_0 \setminus \mathcal{A}$ and $i_0 < i_M$, then the control b must have the following structure:*

$$b(t) = \begin{cases} \beta & \text{if } t \in (0, \tau_1), \\ \dfrac{\gamma}{s(\tau_2) + \gamma i_M(\tau_2 - t)} & \text{if } t \in (\tau_1, \tau_2), \end{cases} \tag{5.2}$$

where τ_1 has been defined above, and

$$\tau_2 := \sup\{t \in [\tau_1, t_f] \ : \ i = i_M \text{ in } [\tau_1, t]\}. \tag{5.3}$$

Furthermore,

$$s(\tau_2) = \frac{\gamma}{\beta}.$$

3. *If $(s_0, i_0) \in \mathcal{B} \setminus \mathcal{B}_0$, then, prior to reaching $\mathcal{B}_0$, the control b must have the following structure:*

$$b(t) = \begin{cases} \beta & \text{if } t \in (0, \tau_1 \wedge \tau_0), \\ \beta_* & \text{if } \tau_0 < \tau_1, \ t \in (\tau_0, \tau_1), \end{cases} \tag{5.4}$$

where τ_1 has been defined above and

$$\tau_0 := \inf\{t \geq 0 : \ i(t) = \Phi_{\mathcal{B}}(s(t))\}. \tag{5.5}$$

Proof. Let us start with some structural remarks.

i. Since, by assumption, $s(t_f) < \frac{\gamma}{\beta} - \varepsilon$, Proposition 4.4 implies that $\eta \geq 0$ in $[0, t_f]$, and p_s is non-increasing and non-negative; moreover, ψ is non-negative and, by (P1), we have $p_1 = 0$.

ii. In the interval $[0, \tau_1)$, whenever nonempty, we have $i < i_M$ and hence

   (a) $d\mu([0, \tau_1)) = 0$ (by complementarity);

   (b) if ψ is equal to $p_0\lambda_2 = 0$ in a subinterval $(t_1, t_2) \subset [0, \tau_1)$, then $p_0 = 0$ and, owing to Theorem 4.6, we have that i is constant in (t_1, t_2) (the alternative $d\mu([0, t_1]) > 0$ being impossible since $[0, t_1] \subset [0, \tau_1)$ and $d\mu([0, \tau_1)) = 0$ by (a));

   (c) $\psi = p_0\lambda_2 > 0$ in a subinterval $(t_1, t_2) \subset [0, \tau_1)$ is impossible, again due to Theorem 4.6 (otherwise, we would have $i = i_M$, which is a contradiction);

   (d) by the first assertion, on this same interval $[0, \tau_1)$, it holds $d\psi = si\gamma p_s\, dt \geq 0$ and hence ψ is continuous and non-decreasing (and non-negative as already observed in i.).

Assertion 1. To prove the first assertion, we assume by contradiction that $\tau_1 = t_f$, that is $i < i_M$ on $[0, t_f)$. Then the same applies on $[0, t_f]$ since, by assumption, there exists some interval $[t_I, t_f]$ on which $s < \frac{\gamma}{\beta}$ and i is strictly decreasing on $(t_I, t_f]$. We claim that $(s_0, i_0) \in \mathcal{A}$, against one of the hypotheses of the theorem. This will be achieved by showing that among the three exhaustive cases

1.1. $\psi(0) > p_0\lambda_2$,

1.2. $p_0\lambda_2 > 0$ and $\psi(0) \leq p_0\lambda_2$,

1.3. $\psi(0) = p_0\lambda_2 = 0$,

only the second can happen. In that case, $\psi(s) = p_0\lambda_2$ on some interval is excluded as discussed in item (c) at the beginning of the proof. Similarly, $\psi(s) > p_0\lambda_2$ at some point $s < t_f$ will be excluded as in 1.1. It follows that $\psi < p_0\lambda_2$ and $b = \beta$ on $(0, t_f)$ (by Definition 2.1, this corresponds to $(s_0, i_0) \in \mathcal{A}$), hence showing that the extremal policy can only keep $i < i_M$ if $(s_0, i_0) \in \mathcal{A}$.

It remains then to exclude the other cases. Since ψ is non-decreasing, in case *1.1* we have $\psi > p_0\lambda_2$ and the optimal control is β_* on $(0, t_f)$, which contradicts the nature of our problem for which β is optimal as soon as one reaches $\mathcal{A}$ (thus, at least on (t_I, t_f)). In case *1.3* we have $p_0 = 0$. Since, as observed in point (i) of our preliminar discussion, also $p_1 = 0$, by the non-degeneration condition (4.1) we get $d\mu([0, t_f]) > 0$. $\psi(s) > 0$ at some point s is excluded as in 1.1, thus $\psi = 0$ on $(0, t_f)$. By Theorem 4.6, using $d\mu([0, t_f]) > 0$, we have $i = i(0)$ on $[0, t_f)$; this is in contradiction with the fact that i is strictly decreasing on $[t_I, t_f)$ as observed in 1.

The remaining assertions concern $(s_0, i_0) \in \mathcal{B} \setminus \mathcal{A}$. Due to the previous argument, $\tau_1 < t_f$, and, by continuity, we have that $i(\tau_1) = i_M$. Moreover, our assumptions imply $\tau_1 > 0$.

Assertion 3. We now turn our attention to the last assertion of the theorem. By definition of the sets $\mathcal{B}$ and $\mathcal{B}_0$, in this case we have $i_0 < i_M$ and hence $\tau_1 > 0$. It is convenient to split the proof according to the following subcases:

3.1. $(s_0, i_0) \in \partial\mathcal{B} \setminus \mathcal{B}_0$,

3.2. $(s_0, i_0) \in \mathcal{B} \setminus (\partial\mathcal{B} \cup \mathcal{B}_0)$.

Subcase 3.1. When $(s_0, i_0) \in \partial\mathcal{B} \setminus \mathcal{B}_0$, the claim directly follows from Theorem 2.3, assertion 6. Indeed, in this case we have $\tau_0 = 0$ and $b = \beta_*$ until the time $\tau_0^{s_0, i_0}$, and to conclude it suffices to prove that $\tau_0^{s_0, i_0} = \tau_1$. In fact, $s_0 > \frac{\gamma}{\beta_*}$ (by assumption) and, as long as $s \geq \frac{\gamma}{\beta_*}$, we have

$$\frac{d}{dt}\big(i - \Phi_{\mathcal{B}}(s)\big) = (\beta_* s - \gamma)i - (1 - \frac{\gamma}{\beta_* s})\beta_* si = 0. \tag{5.6}$$

This implies that $i = \Phi_{\mathcal{B}}(s)$ on $(0, \tau_0^{s_0,i_0})$. Furthermore, along the associated trajectory, we have $\left(s^{s_0,i_0,\beta_*}(\tau_0^{s_0,i_0}), i^{s_0,i_0,\beta_*}(\tau_0^{s_0,i_0})\right) \in \partial\mathcal{B}$. Since $s^{s_0,i_0,\beta_*}(\tau_0^{s_0,i_0}) = \frac{\gamma}{\beta_*}$, it follows that $i^{s_0,i_0,\beta_*}(\tau_0^{s_0,i_0}) = i_M$ which implies $\tau_1 \leq \tau_0^{s_0,i_0}$. On the other hand, if $\tau_1 < \tau_0^{s_0,i_0}$ then there would exist $\bar{t} \in (\tau_1, \tau_0^{s_0,i_0})$ such that $i^{s_0,i_0,\beta_*}(\bar{t}) = i_M$ and $s^{s_0,i_0,\beta_*}(\bar{t}) > \gamma/\beta_*$ so exiting the viable zone $\mathcal{B}$, which is a contradiction.

At $\tau_0^{s_0,i_0} = \tau_1$, the trajectory enters in the set $\mathcal{B}_0$.

Subcase 3.2. When $(s_0, i_0) \in \mathcal{B} \setminus (\partial\mathcal{B} \cup \mathcal{B}_0)$ we have $i_0 < \Phi_{\mathcal{B}}(s_0)$ and $s_0 > \frac{\gamma}{\beta_*}$. The proof of the assertion in this subcase is divided in three steps.

Step 3.2.1. We claim that $\psi(0) < p_0\lambda_2$ (hence $p_0\lambda_2 > 0$, since η and ψ are non-negative).

Indeed, otherwise, since ψ is non-decreasing in $[0, \tau_1)$ and does not stay equal to $p_0\lambda_2$ on an interval, we must have $\psi > p_0\lambda_2$ in $(0, \tau_1)$. According to (4.6), we would have $b = \beta_*$ in this interval. Then the associated trajectory $\left(s^{s_0,i_0,\beta_*}, i^{s_0,i_0,\beta_*}\right)$ is kept in the interior of $\mathcal{B}$. To see this, it suffices to note that, as in (5.6), as long as $s \geq \frac{\gamma}{\beta_*}$, we have $\frac{d}{dt}\left(i - \Phi_{\mathcal{B}}(s)\right) = 0$. On the other hand, since $i^{s_0,i_0,\beta_*}(\tau_1) = i_M$, the trajectory has to reach the boundary by entering in the zone in which $s < \frac{\gamma}{\beta_*}$. In particular, we have $s^{s_0,i_0,\beta_*}(\tau_1) < \frac{\gamma}{\beta_*}$. But, then, there exists some $t < \tau_1$ such that $s^{s_0,i_0,\beta_*}(t) = \frac{\gamma}{\beta_*}$ (and $i^{s_0,i_0,\beta_*}(t) < i_M$). Since i is decreasing on (t, τ_1), we get a contradiction.

Step 3.2.2. Definition of σ_1. Since $\psi(0) < p_0\lambda_2$ and it is continuous, there exists a right neighbourhood of 0 in which $\psi < p_0\lambda_2$ and where, according to (4.6), we have $b = \beta$. Let us define

$$\sigma_1 := \sup\left\{t \in [0, t_f] \ : \ \psi < p_0\lambda_2 \text{ in } [0, t)\right\}.$$

The reader is invited to note that $\sigma_1 < t_f$. Otherwise, β would be always admissible for (s_0, i_0) which contradicts the choice $(s_0, i_0) \notin \mathcal{A}$.

Step 3.2.3. The time τ_0 is well defined since, due to $i(\tau_1) = i_M$, the set on the right hand side in (5.5) is nonempty. We claim that $\sigma_1 = \tau_0 \wedge \tau_1$.

We focus first on the case in which $\tau_0 < \tau_1$. Then, by definition of τ_1, $i(t) < i_M$ for every $t \leq \tau_0$. Moreover, $s(\tau_0) > \frac{\gamma}{\beta_*}$ because $(s(\tau_0), i(\tau_0)) \in \partial\mathcal{B}$ and $i(\tau_0) < i_M$. Since s is decreasing then we have $s(t) > \frac{\gamma}{\beta_*}$ for every $t \leq \tau_0$ and we can apply an argument similar to that of Step 3.2.1 to $\left(\tilde{s}_0, \tilde{i}_0\right) = (s(\sigma_1), i(\sigma_1))$ to get a contradiction with $\sigma_1 < \tau_0$. On the other hand, the only control admissible at $(s(\tau_0), i(\tau_0))$ is (locally in time) β_* (see Subcase 3.1), and this implies $\sigma_1 \leq \tau_0$. The conclusion follows.

If $\tau_0 \geq \tau_1 > 0$, we aim at proving that $\sigma_1 = \tau_1$. Assume, by contradiction, that $\sigma_1 > \tau_1$. The control β is admissible at $(s(\tau_1), i_M) \in \partial\mathcal{B}$ only if $s(\tau_1) \leq \frac{\gamma}{\beta}$ (since, otherwise, i would be increasing in a neighbourhood of τ_1 and the ICU state constraint would be violated) and, in this case, we would have $\sigma_1 = t_f$ (excluded before). Thus $\sigma_1 \leq \tau_1$.

If, by contradiction, $\sigma_1 < \tau_1$, then, in the interval (σ_1, τ_1) we would have $\psi > p_0\lambda_2$ (since ψ is non-decreasing in $[0, \tau_1)$ and cannot stay equal to $p_0\lambda_2$ which is positive, by Step 3.2.1.). Then we would have $b = \beta_*$ in (σ_1, τ_1) and i would be decreasing so contradicting $i(\tau_1) = i_M$.

The structure (5.4) follows now by definition of σ_1, the fact that $\sigma_1 = \tau_0 \wedge \tau_1$ and (4.6). Actually, assertion *3* is completely proved.

Assertion 2. Let us now prove the second assertion. The reader is invited to note that the assumption $i_0 < i_M$ implies $\tau_1 > 0$. The proof is subdivided in five steps.

Step 2.1. We claim that $\psi(0) < p_0\lambda_2$ (and, thus, $p_0 = 1$, since η and ψ are non-negative).

Again, the case $\psi(0) > p_0\lambda_2$ leads to the control β_* in $(0, \tau_1)$ for which i is non-increasing thus cannot reach i_M at time τ_1. By remark ii. (c) at the beginning of the proof, the case $\psi(0) = p_0\lambda_2 > 0$ cannot hold on some non-empty sub-interval. By remark ii. (b), the case $\psi(0) = 0$ yields i constant on some initial (possibly empty) interval followed by i decreasing and it is also excluded.

Step 2.2. The argument on $(0, \tau_1)$ is identical to the one in Step 3.2.2, with the notable exception that, starting from $\mathcal{B}_0$, we get $\tau_0 = \tau_1$.

Step 2.3. Since $i(\tau_1) = i_M$, it follows that the time τ_2 is well defined, that is, the set on the right hand side in (5.3) is nonempty. By the assumption $s(t_f) < \frac{\gamma}{\beta}$, one has $\tau_2 < t_f$. We focus on the case when $\tau_1 < \tau_2$ for which, in (τ_1, τ_2), we have $i = i_M$. Then the optimal control in (τ_1, τ_2) is given by (4.12), that is

$$b(t) = \frac{\gamma}{s(\tau_2) + \gamma i_M(\tau_2 - t)}. \tag{5.7}$$

Since $b \le \beta$, we have $s(\tau_2) \ge \frac{\gamma}{\beta}$. Since s is strictly decreasing, it follows that

$$s(t) > s(\tau_2) \ge \frac{\gamma}{\beta} \quad \text{for every } t \in [\tau_1, \tau_2). \tag{5.8}$$

In other terms, we have $i(t) > \Phi_{\mathcal{A}}(s(t))$ for every $t \in [\tau_1, \tau_2)$. Let us set

$$\overline{\tau}_2 := \sup\left\{t \in (\tau_1, t_f] \ : \ i > \Phi_{\mathcal{A}}(s) \text{ in } (\tau_1, t)\right\}.$$

We note that the right-hand side set is nonempty and $\tau_2 \le \overline{\tau}_2 < t_f$.

The remaining part of the proof aims to show that $\tau_2 = \overline{\tau}_2$. By continuity, this yields $\partial A \ni (s(\tau_2), i_M)$, which implies $s(\tau_2) \le \frac{\gamma}{\beta}$ and, together with (5.8), provides the "furthermore" part of the assertion.

Step 2.4. We claim that if $\tau_2 < \overline{\tau}_2$, then $i < i_M$ in $(\tau_2, \overline{\tau}_2]$.

First of all, we prove that

$$i < i_M \text{ in a right neighborhood of } \tau_2. \tag{5.9}$$

We assume, by contradiction, the existence of a strictly decreasing sequence of points $t_n \in (\tau_2, \overline{\tau}_2)$ with $i(t_n) = i_M$ and such that $t_n \to \tau_2$. We claim that this implies that

$$i = i_M \text{ in } [t_n, t_{n+1}] \text{ for every } n. \tag{5.10}$$

Again, by contradiction, should there exist $\hat{t}_n \in (t_{n+1}, t_n)$ such that $i(\hat{t}_n) < i_M$, the continuity of i implies the existence of some open interval $(\underline{a}, \overline{a}) \subseteq [t_{n+1}, t_n]$ such that $i < i_M$ inside and $i = i_M$ at the boundary[†]. By complementarity, we have that $d\mu((\underline{a}, \overline{a})) = 0$ and, therefore, $d\psi = si\gamma p_s \ge 0$ such that ψ is continuous and non-decreasing in $(\underline{a}, \overline{a})$. Moreover ψ cannot stay equal to $p_0\lambda_2$ in $(\underline{a}, \overline{a})$: otherwise, recalling that $p_0 = 1$, by Theorem 4.6, we would have $i = i_M$, which is a contradiction with the choice of $\hat{t}_n$. Then, we have all the ingredients to repeat in $(\underline{a}, \overline{a})$ the argument that we have used in the interval $(0, \tau_1)$ (albeit we are now in the interior of $\mathcal{B}$ if $\tau_2 - \tau_1 > 0$), obtaining that $b = \beta$ in $(\underline{a}, \overline{a})$.

Since we are inside the interval $(\tau_2, \overline{\tau}_2)$, it holds that $s(t) > \frac{\gamma}{\beta}$ and, therefore, $di = (s\beta - \gamma)i > 0$ in $(\underline{a}, \overline{a})$, that is i is increasing in $(\underline{a}, \overline{a})$ which is a contradiction with $i(\underline{a}) = i(\overline{a}) = i_M$.

[†]One may take, for instance, $\underline{a} = \inf\{t \in [t_{n+1}, \hat{t}_n] \ : \ i(t) < i_M\}$ and $\overline{a} = \sup\{t \in [\hat{t}_n, t_n] \ : \ i(t) < i_M\}$.

Then, we have proved that (5.10) holds true. Since the sequence t_n converges to τ_2 then, by continuity, this implies that $i = i_M$ on (τ_1, t_1) with $t_1 > \tau_2$, against the definition of τ_2. This proves (5.9) and there exists a right neighborhood of τ_2 in which we have $i < i_M$.

We can set

$$\hat{\tau}_2 := \sup \{ t \in (\tau_2, \overline{\tau}_2] \ : \ i < i_M \text{ in } (\tau_2, t) \}$$

By the same argument as before we must have $\hat{\tau}_2 = \overline{\tau}_2$. Indeed, otherwise, in the interval $(\tau_2, \hat{\tau}_2)$, we would have $i < i_M$ inside and $i = i_M$ at the boundary and we repeat the same argument used in the interval $(\underline{a}, \overline{a})$, leading to a contradiction. By the same argument we get that $i(\overline{\tau}_2) < i_M$. Then we have proved that $i < i_M$ in $(\tau_2, \overline{\tau}_2]$.

Step 2.5. We claim that if $\tau_2 < \overline{\tau}_2$ then there exists $\sigma > \overline{\tau}_2$ such that $\psi > p_0\lambda_2$; hence $b = \beta_*$, in (τ_2, σ).

Owing to the assertion in Step 2.4 and to the continuity of i, there exists $\sigma > \overline{\tau}_2$ such that $i < i_M$ on (τ_2, σ). Due to the complementarity conditions, we have that ψ is non-decreasing and continuous in (τ_2, σ). We can exclude that $\psi(\tau_2^+) < p_0\lambda_2$. Indeed, in such case, we would have $\psi < p_0\lambda_2$ in a right neighborhood of τ_2 in which we have also $i < i_M$ and $i(\tau_2) = i_M$. However, since in such a case $b = \beta$ and $s > \frac{\gamma}{\beta}$, we would have that $di = (s\beta - \gamma)i > 0$ and this provides a contradiction. Moreover it does not stay equal to $p_0\lambda_2$ in a subinterval (as constancy would imply $i = i_M$). Then we must have $\psi(t) > p_0\lambda_2$, on some interval (τ_2, σ'). Since ψ is non-decreasing and does not stay equal to $p_0\lambda_2$ in any sub-interval of (τ_2, σ), we have $\psi > p_0\lambda_2$ in (τ_2, σ). Therefore, $\psi > p_0\lambda_2$ and $b = \beta_*$ in (τ_2, σ), as claimed.

By definition, $(s(\overline{\tau}_2), i(\overline{\tau}_2)) \in \mathcal{A}$ and, as we have seen, $i(\overline{\tau}_2) < i_M$. Then, for every $t \in (\overline{\tau}_2, \sigma)$, $(s(t), i(t)) = \left(s^{s(\overline{\tau}_2), i(\overline{\tau}_2), \beta_*}(t - \overline{\tau}_2), i^{s(\overline{\tau}_2), i(\overline{\tau}_2), \beta_*}(t - \overline{\tau}_2)\right) \in \mathcal{A} \setminus \partial\mathcal{A}$ and $b = \beta_*$ on $[t, \sigma)$. This comes in contradiction with Theorem 5.1, *1.*

It follows that $\tau_2 = \overline{\tau}_2$ and the theorem is completely proved. □

Remark 5.3 Substituting $s(\tau_2) = \frac{\gamma}{\beta}$ in the expression (5.2) of $b(t)$, in assertion 2 we have

$$b(t) = \begin{cases} \beta & \text{if } t \in (0, \tau_1), \\ \dfrac{\beta}{1 + \beta i_M(\tau_2 - t)} & \text{if } t \in (\tau_1, \tau_2). \end{cases} \tag{5.11}$$

Remark 5.4

1. Assertion 2 of the previous theorem yields that the only optimal control for $(s_0, i_0) \in \mathcal{B}_0$, with $i_0 < i_M$ is of the form (5.2). This optimality statement can easily be extended to $(s_0, i_M) \in \mathcal{B}_0$ if $\frac{\gamma}{\beta} < s_0 < \frac{\gamma}{\beta_*}$. Indeed, each point of this type can be written as $(s_0, i_M) = \left(s^{s_1, i_1, \beta}(t), i^{s_1, i_1, \beta}(t)\right)$ starting from $(s_1, i_1) = \left(s^{s_0, i_M, \beta}(-t), i^{s_0, i_M, \beta}(-t)\right)$. Since $s_0 < \frac{\gamma}{\beta_*}$, it follows that $s^{s_0, i_M, \beta}(-t) < \frac{\gamma}{\beta_*}$ for t small enough. Furthermore, $s_0 > \frac{\gamma}{\beta}$ implies that $r \mapsto i^{s_0, i_M, \beta}(-r)$ is decreasing (locally in time). This shows that, for small $t > 0$, $(s_1, i_1) \in \mathcal{B}_0$ with $i_1 < i_M$. Whenever b is an admissible control for t_f, s_0, i_0 and the problem $\mathcal{P}_\varepsilon$, one sets $\tilde{b}(r) := \beta \mathbf{1}_{r \le t} + b(r - t)\mathbf{1}_{r > t}$. The optimality of b^{opt} given by (5.2) for the problem $\mathcal{P}_\varepsilon$, the initial data (s_1, i_1) and for the time horizon t_f (note also that $t = \tau_1$!) yields $J(t_f, b^{opt}(\cdot + t); s_0, i_M) = J(t_f + t, b^{opt}; s_1, i_1) = V(t_f + t; s_1, i_1) \le J(t_f + t, \tilde{b}; s_1, i_1) = J(t_f, b; s_0, i_M)$.
2. In the same spirit, for the upper-right corner of $\mathcal{B}_0$ i.e. $(s_0, i_0) = \left(\frac{\gamma}{\beta_*}, i_M\right)$, every (locally-in-time) admissible control $b\mathbf{1}_{[0,\delta]}(t)$ should be followed (on (δ, t_f)), by (5.2). ‡ As such, in this case too, an optimal control can be chosen of form (5.2).

‡This can be made rigorous using the dynamic programming principle.

Remark 5.5 1. By Theorem 5.2, we have $i = i_M$ in the interval (τ_1, τ_2). By integrating the state equations (see Proposition 4.7), in this interval the susceptible population s turns out to be

$$s(t) = s(\tau_2) + \gamma i_M(\tau_2 - t).$$

Then it decreases linearly from $s(\tau_1)$ to $s(\tau_2) = \gamma/\beta$. In particular, we have

$$s(\tau_1) = \frac{\gamma}{\beta} + \gamma i_M(\tau_2 - \tau_1)$$

from which we obtain that the time amplitude of this control regime is given by

$$\tau_2 - \tau_1 = \frac{s(\tau_1) - \frac{\gamma}{\beta}}{\gamma i_M}.$$

The later provides also an alternative definition of the switching time τ_2 in terms of τ_1.

2. One can give an estimate of $\tau_1 - \tau_0$ solely based on the initial data and the parameters $\beta,\ \beta_*,\ \gamma,\ i_M$. Using the fact that between τ_0 and τ_1 one employs the constant control β_* and that $\left(s^{s_0,i_0,\beta_*}, i^{s_0,i_0,\beta_*}\right) \in \partial\mathcal{B}$ with $s^{s_0,i_0,\beta_*} \geq \frac{\gamma}{\beta_*}$, one can easily show that

$$\frac{\beta_*\vartheta_1 - \gamma}{\beta_*\beta s_0 i_M} \leq \tau_1 - \tau_0 \leq \frac{\beta_*\vartheta_1 - \gamma}{\beta\gamma\vartheta_2}, \tag{5.12}$$

where

$$\begin{aligned}
\vartheta_0 &:= i_M + \frac{\gamma}{\beta_*} - \frac{\gamma}{\beta_*}\log\frac{\gamma}{\beta_*};\\
\vartheta_1 &:= \exp\left(\frac{s_0 + i_0 - \frac{\gamma}{\beta}\log s_0 - \vartheta_0}{\frac{\gamma}{\beta_*} - \frac{\gamma}{\beta}}\right);\\
\vartheta_2 &:= \frac{\beta_*}{\beta - \beta_*}\vartheta_0 + \frac{s_0 + i_0 - \frac{\gamma}{\beta}\log s_0}{1 - \frac{\beta_*}{\beta}} - \vartheta_1.
\end{aligned}$$

5.1 The Optimal Control

Gathering these pieces of information provided by Theorem 5.1, Theorem 5.2, Remark 5.3, Remark 5.4 and Remark 5.5, we get the following complete characterizations of an optimal control for problems $\mathcal{P}_\varepsilon$ and $\mathcal{P}$.

Theorem 5.6 *Let* $(s_0, i_0) \in \mathcal{B}$ *with* $0 < i_0,\ 0 < s_0$. *Let*

$$b^{opt}(t) = \begin{cases} \beta & \text{if } t \in (0, \tau_1 \wedge \tau_0),\\ \beta_* & \text{if } \tau_0 < \tau_1,\ t \in (\tau_0, \tau_1),\\ \dfrac{\beta}{1 + \beta i_M(\tau_2 - t)} & \text{if } t \in (\tau_1, \tau_2),\\ \beta & \text{if } t > \tau_2, \end{cases} \tag{5.13}$$

where

$$\begin{aligned}
\tau_0 &:= \inf\left\{t \geq 0 :\ i^{s_0,i_0,\beta}(t) = \Phi_{\mathcal{B}}\left(i^{s_0,i_0,\beta}(t)\right)\right\},\\
\tau_1^\beta &:= \inf\left\{t \in (0, t_f]\ :\ i^{s_0,i_0,\beta}(t) = i_M\right\},\\
\tau_1 &:= \inf\left\{t \in (0, t_f]\ :\ i^{s_0,i_0,\beta 1_{(0,\tau_1^\beta \wedge \tau_0)} + \beta_* 1_{(\tau_1^\beta \wedge \tau_0, t_f)}}(t) = i_M\right\},\\
\tau_2 &:= \tau_1 + \frac{s^{s_0,i_0,\beta 1_{(0,\tau_1 \wedge \tau_0)} + \beta_* 1_{(\tau_1 \wedge \tau_0, t_f)}}(\tau_1) - \frac{\gamma}{\beta}}{\gamma i_M}.
\end{aligned}$$

Furthermore, let us set

1. the reaching time

$$\bar{t}_f^{s_0,i_0} := \inf\{t \geq 0 : \; s^{s_0,i_0,b^{opt}}(t) = \frac{\gamma}{\beta}\}; \tag{5.14}$$

2. ω^{s_0,i_0} *to be the function*

$$\tau \mapsto \omega^{s_0,i_0}(\tau) := \frac{\gamma}{\beta} - s^{\frac{\gamma}{\beta}, i^{s_0,i_0,b^{opt}}(\bar{t}_f^{s_0,i_0}), b^{opt}}(\tau). \tag{5.15}$$

Then

1. *For every* $t_f > \bar{t}_f^{s_0,i_0}$ *and every* $\varepsilon < \omega^{s_0,i_0}(t_f - \bar{t}_f^{s_0,i_0})$*, the problem* $\mathcal{P}_\varepsilon$ *has a continuous value and* b^{opt} *is an optimal control.*
2. *For every* $t_f > \bar{t}_f^{s_0,i_0}$*, the problem* $\mathcal{P}$ *admits* b^{opt} *as optimal control.*
3. *The optimal control can be given in a feed-back form*

$$\tilde{b}^{opt}(s,i) = \begin{cases} \beta & \text{if } (s,i) \in (\mathcal{B} \setminus \partial\mathcal{B}) \cup \partial\mathcal{A}, \\ \beta_* & \text{if } (s,i) \in \partial\mathcal{B} \setminus \mathcal{B}_0, \\ \dfrac{\gamma}{s} & \text{if } (s,i) \in \left(\frac{\gamma}{\beta}, \frac{\gamma}{\beta_*}\right) \times \{i_M\}, \end{cases}$$

and is unique up to the Lebesgue-null set $\partial\mathcal{A} \cup \left(\frac{\gamma}{\beta_*}, i_M\right)$.

Remark 5.7 1. The function ω^{s_0,i_0} represents the deviation of the susceptible population from the value γ/β taken at the reaching time, when the latter is taken as new origin of times and the control is b^{opt}; accordingly, we have $\omega^{s_0,i_0}(0) = 0$. By the monotonicity of s, we have that ω^{s_0,i_0} is increasing and strictly positive (if $i_0 > 0$) on the interval $(0, +\infty)$. Moreover, if $t_f > \bar{t}_f^{s_0,i_0}$, then

$$\omega^{s_0,i_0}(t_f - \bar{t}_f^{s_0,i_0}) > \varepsilon \iff s^{s_0,i_0,b^{opt}}(t_f) < \frac{\gamma}{\beta} - \varepsilon.$$

2. The only control keeping the solution on ∂A is β (and only as long as $s > \frac{\gamma}{\beta}$). Furthermore, an optimal trajectory starting from $\mathcal{B} \setminus \mathcal{A}$ enters $\mathcal{A}$ at $\left(\frac{\gamma}{\beta}, i_M\right)$ (see assertion 2 in Theorem 5.2). It follows that the occupation time of ∂A is non-singular with respect to the Lebesgue measure on $\mathbb{R}_+$ only if $(s_0, i_0) \in \partial\mathcal{A}$, $s_0 > \frac{\gamma}{\beta}$ and only if, on the set $\{(s,i) \in \partial\mathcal{A}\}$, the control coincides with $\tilde{b}^{opt}$. This shows that b^{opt} is unique Lebesque-almost surely (in time).
3. If one does not already start in the interior of the no-effort zone, i.e. $(s_0, i_0) \notin \mathcal{A} \setminus \partial\mathcal{A}$, the modulus ω is independent of (s_0, i_0) and $\omega(t) = \frac{\gamma}{\beta} - s^{\frac{\gamma}{\beta}, i_M, \beta}(t)$.
4. The optimal controls need to be extremals and they are unique in almost all the settings (excepting $i = i_M$ and $s \in \left(\frac{\gamma}{\beta}, \frac{\gamma}{\beta_*}\right)$ which is dealt with in Remark 5.4. Assertion 2 follows from assertion 3 in Proposition 3.2. Finally, assertion 3 is just a way to say that we actually have a "positional strategy" but this is merely a way of rewriting the first assertion.
5. The value function in assertion *1* only depends (in an integral formulation) on the position $s^{s_0,i_0,\beta}(\tau)$ where τ is the hitting time of $\mathcal{B}$ (with β control). The continuity of $(s_0, i_0) \mapsto \tau$ implies that of the value function V itself. Note that the (relevant part of the) boundary $\partial\mathcal{B}$ can be divided into two sets $\Gamma_1 := \left(\frac{\gamma}{\beta}, \frac{\gamma}{\beta_*}\right) \times \{i_M\}$ and $\Gamma_2 := \left\{(s, \Phi_{\mathcal{B}}(s)) : s \geq \frac{\gamma}{\beta_*}\right\}$ with different "qualification" behaviour (see, for instance, the introduction in [FP00]). Indeed, Γ_1 satisfies the "inward pointing" condition, while Γ_2 is "outward pointing" with all but the viable control. This hybrid qualification does not allow direct application of classical results of continuity of the value function.

6. The control b^{opt} is continuous at τ_2; moreover, starting from $\mathcal{B}_0 \setminus \mathcal{A}$ its expression coincides with the one obtained in [MSW20, Theorem 1].

We end this section with a few elements of comparison with [ACMM$^+$21]. While it is true that the optimal control is the greedy one in both cases, we wish to emphasize that our problem is fundamentally different than the one in [ACMM$^+$21]. Indeed, [ACMM$^+$21] focuses on the *minimal time* to reach $\mathcal{A}$ (cf. line 4 after Eq. (2.1) in [ACMM$^+$21]). Roughly speaking,

1. their cost is $\lambda_1 = 1$ and $\lambda_2 = 0$, prior to reaching the safe zone; this has a strong impact on the proof as their comparison in S1.4 strongly uses the independence of control in the cost;
2. since they deal with a minimal time problem, the "final" time varies with the controls used; for us, there is a surveillance until a *fixed* time t_f.

6 Bocop Simulations

To conclude, we present some numeric simulations done by using the Bocop package, [TC17, BGG$^+$17]. For simplicity, the simulations are made on a fixed time interval $[0, t_f]$ where t_f is taken to be large enough to ensure that in the last part of the epidemic horizon the optimal control is β (no effort condition). Consequently, we can consider a cost functional in which $\lambda_1 = 0$ and $\lambda_2 = 1$ (indeed, since t_f is fixed, the choice $\lambda_1 > 0$ would simply result in a constant additive contribution to the cost functional), that is

$$J(b) = \int_0^{t_f} \big(\beta - b(t)\big)\, dt.$$

The cost $J(b)$ is minimized under the SIR state equations

$$\begin{cases} s' = -sbi \\ i' = sbi - \gamma i \\ s(0) = s_0, \; i(0) = i_0 \end{cases}$$

with the ICU contraint

$$i(t) \le i_M \;\forall\, t \in [0, t_f].$$

The Bocop package implements a local optimization method. The optimal control problem is approximated by a finite dimensional optimization problem (NLP) using a time discretization (the direct transcription approach). The NLP problem is solved by the well known software Ipopt, using sparse exact derivatives computed by CppAD. From the list of discretization formulas proposed by the package (Euler, Midpoint, Gauss II and Lobatto III C), we have chosen to use the Gauss II implicit method since it appears to be stable enough for the problem under consideration.

In our simulations we consider only viable initial conditions outside the no-effort zone (in which the optimal control would be identically equal to β), in other words we take $(s_0, i_0) \in \mathcal{B} \setminus \mathcal{A}$. Once having choosen s_0, this condition writes

$$\Phi_{\mathcal{A}}(s_0) < i_0 < \Phi_{\mathcal{B}}(s_0).$$

By the theory developed in the previous sections we expect qualitatively different optimal controls according to the following two scenarios:

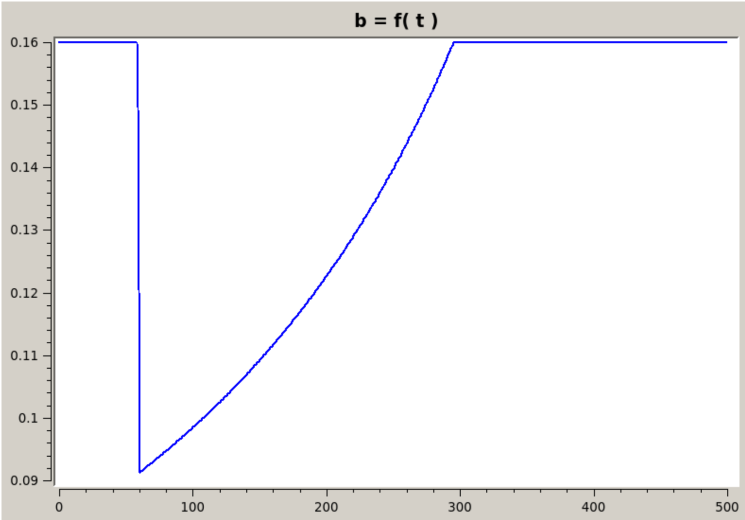
1. $s_0 < \frac{\gamma}{\beta_*}$, in which we expect an optimal control b with a bang-boundary-bang structure,
2. $s_0 > \frac{\gamma}{\beta_*}$, in which we expect an optimal control b with a bang-bang-boundary-bang structure.

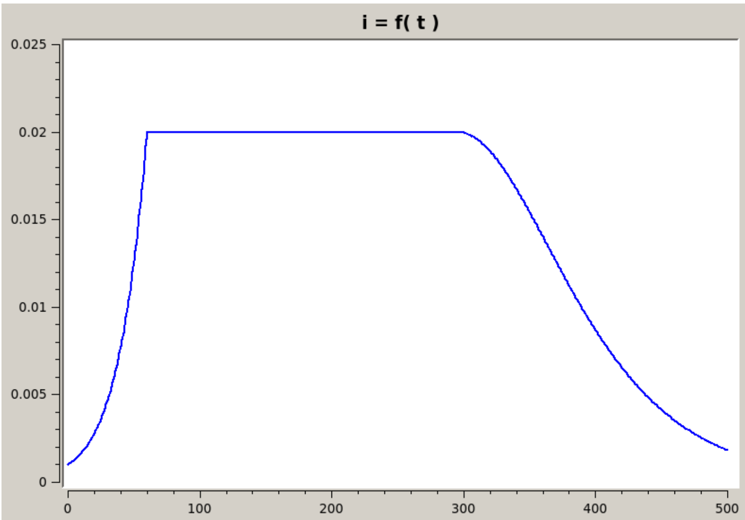
In both scenarios we consider a time horizon t_f of 500 days and choose the coefficients $\beta = 0.16$, $\gamma = 0.06$, $\beta_* = 0.08$ (so that $\frac{\gamma}{\beta_*} = 0.75$), the initial conditions $i_0 = 0.001$ and the ICU upper bound $i_M = 0.02$, but different values of s_0; precisely

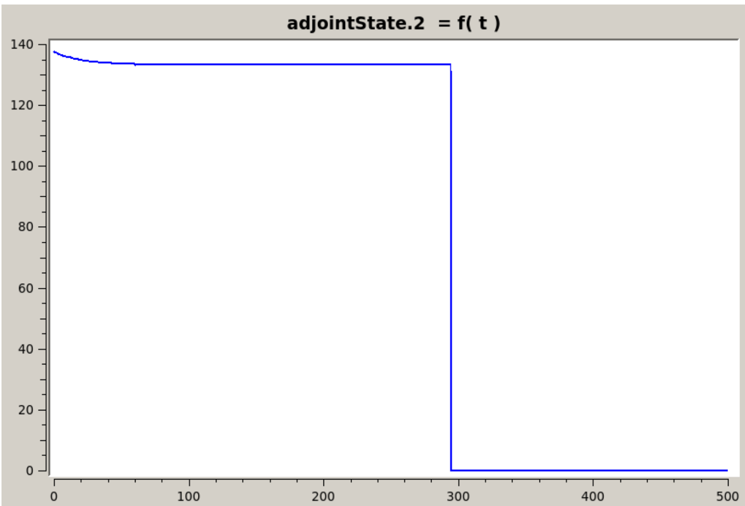
1. $s_0 = 0.7 < \frac{\gamma}{\beta_*}$ in the first scenario, so that $\Phi_{\mathcal{A}}(s_0) \simeq -0.071 < i_0 < \Phi_{\mathcal{B}}(s_0) \simeq 0.018$;
2. $s_0 = 0.85 > \frac{\gamma}{\beta_*}$ in the second, so that $\Phi_{\mathcal{A}}(s_0) \simeq -0.148 < i_0 < \Phi_{\mathcal{B}}(s_0) \simeq 0.014$.

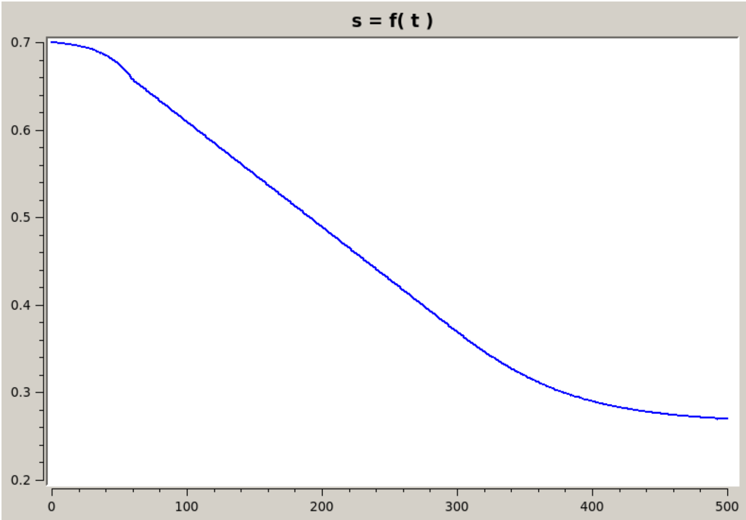
The figures below show the graph of the optimal control, the states and the adjoint states (adjointState.2 stands for p_i and adjointState.1 is p_s). In both scenarios the expected structure of the optimal control is confirmed by the numerical solutions. Moreover, in the second scenario the adjoint variable p_i has two jumps while in the first there is just one discontinuity. As expected, p_s is always continuous and decreasing.

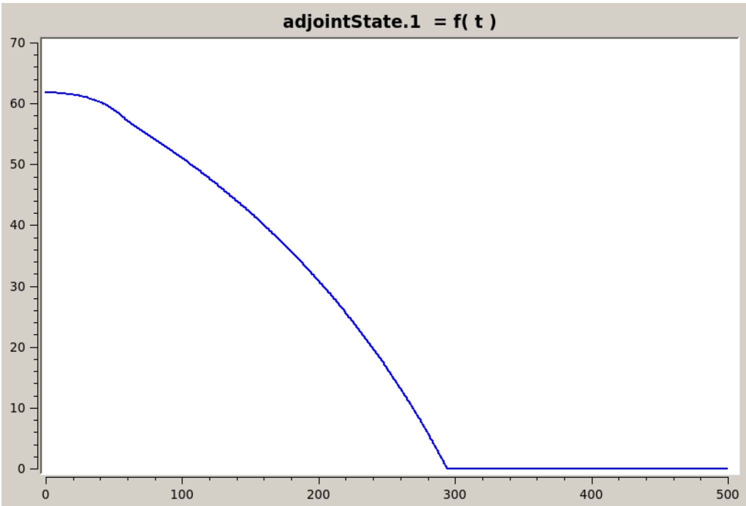
Scenario 1


b = f(t)
0.16
0.15
0.14
0.13
0.12
0.11
0.1
0.09
0
100
200
300
400
500

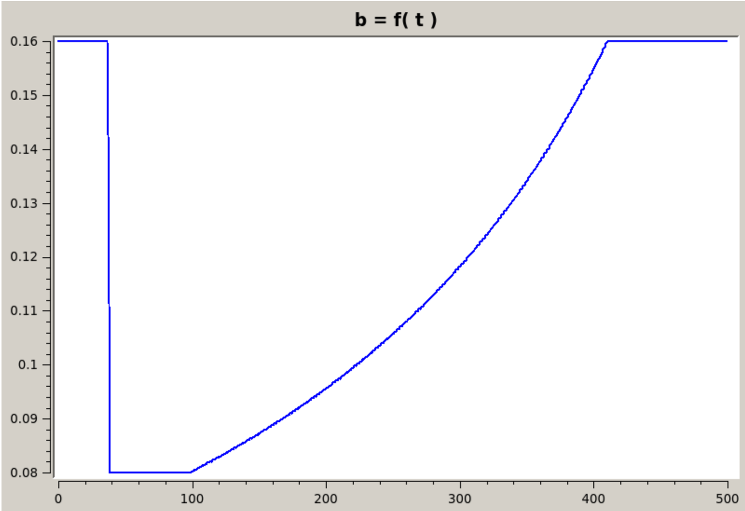

i = f(t)
0.025
0.02
0.015
0.01
0.005
0
0
100
200
300
400
500

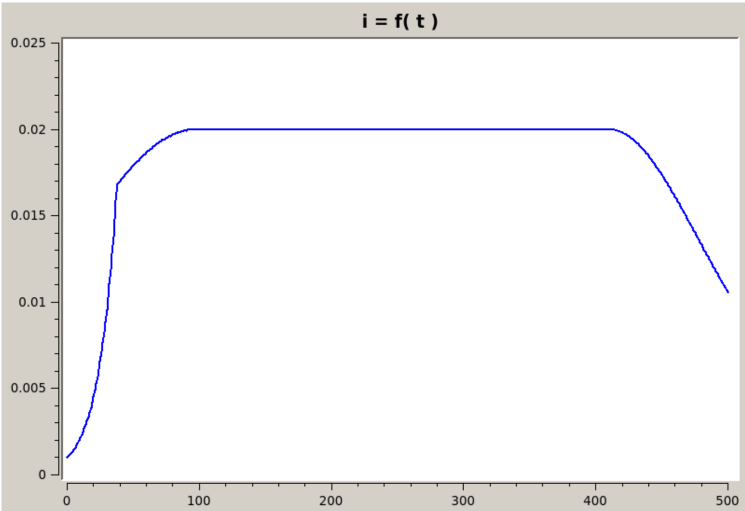

adjointState.2 = f(t)
140
120
100
80
60
40
20
0
0
100
200
300
400
500

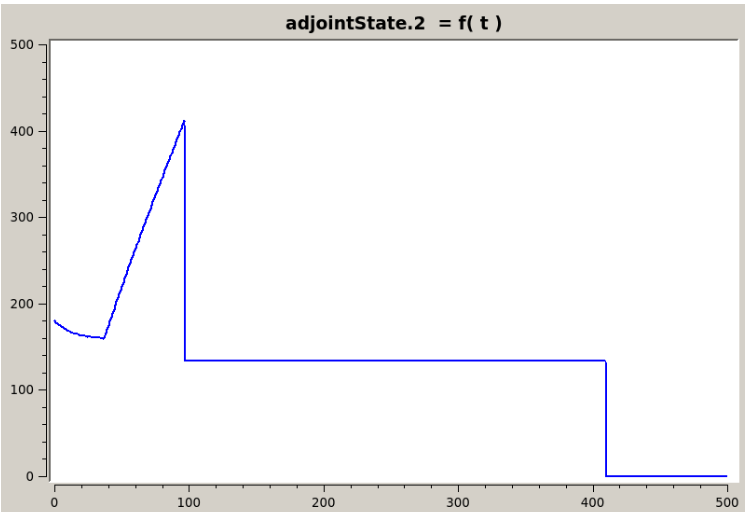

s = f(t)
0.7
0.6
0.5
0.4
0.3
0.2
0
100
200
300
400
500

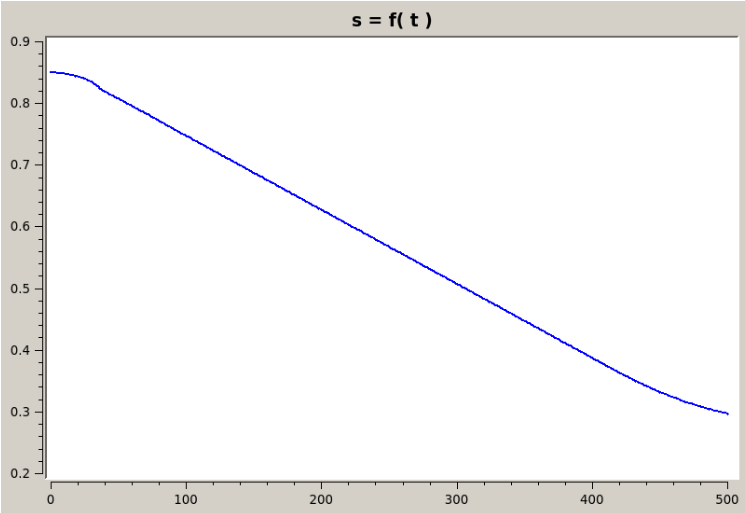

adjointState.1 = f(t)
70
60
50
40
30
20
10
0
0
100
200
300
400
500

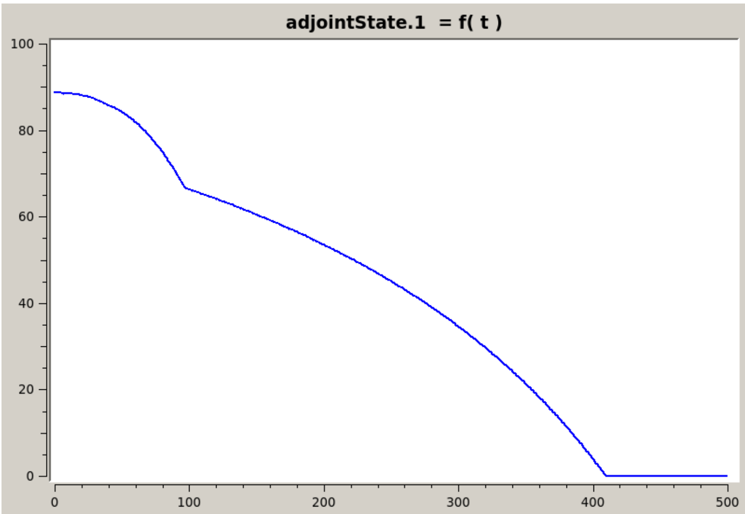
Scenario 2


b = f(t)
0.16
0.15
0.14
0.13
0.12
0.11
0.1
0.09
0.08
0
100
200
300
400
500


i = f(t)
0.025
0.02
0.015
0.01
0.005
0
0
100
200
300
400
500


adjointState.2 = f(t)
500
400
300
200
100
0
0
100
200
300
400
500





Let us end this section with some further insights on the optimal decisions to be taken depending on the state-of-epidemics. For better understanding, we illustrate this on the COVID-19 situation in France and explain the optimal decisions in connection with the associated zones (see Figure 1).

We deal here with a population $N = 67,000,000$ individuals and an alert level of 400 cases per $100,000$ inhabitants (as announced in May 2021 for basis of de-confinement; a very large level if compared to 40 in Japan or 100 in several other countries). The incubation-to-recovery average time is set to $\frac{1}{\gamma} = 14$ (days). [§] A simple computation yields a normalized $i_M := \frac{1}{\gamma} \times \frac{400}{100000} \simeq 0.056$. We consider $\beta = 0.5$ (this parameter is computed from the Delta-variant $R_0 \simeq 7$ and γ using the Cori method with $c = 1$ contact per time unit) and $\beta_* = 0.25$ (i.e. roughly 50% of the population is not directly involved in the essential economy or can work remotely and can be confined if needed), we get

$$\frac{\gamma}{\beta} \simeq 0.1428571, \ \frac{\gamma}{\beta_*} \simeq 0.2857143, \ s_M \simeq 0.3091318, \ s_M^* \simeq 0.5037254.$$

Given a state-of-epidemics (S, I) (under the curve $\Phi_{\mathcal{B}}$!):

- If $I \leq I_M = i_M \times N = 3,752,000$ and $S \leq 0.1428571 \times N = 9,571,425$ (meaning that $53,676,571$ individuals have recovered or are vaccinated corresponding to a percentage of roughly 80%), then do nothing. The herd immunity is achieved.

[§] Indeed, with an exponential law, the average is the inverse of the exponential parameter.

- If $I = I_M = i_M \times N = 3,752,000$ and $S \in (9,571,425,\ 19,142,857]$, then apply progressive de-confinement $b = \frac{\gamma \times N}{S}$ (high b means less confinement) until herd immunity (see previous case).
- If $I < 3,752,000$, compute $S_- := N \times \Phi_{\mathcal{A}}\left(\frac{I}{N}\right)$ and $S^+ := N \times \Phi_{\mathcal{B}}\left(\frac{I}{N}\right)$.
  - If $S \leq S_-$, then do nothing (the herd immunity will be achieved).
  - If $S > S_-$, monitor the situation without intervention until $S = S_+$. If $S_+ > 19,142,857$, then a strict confinement is in order until the number of susceptibles reaches $19,142,857$. Otherwise, we are in the second case (progressive de-confinement).
- The reader is invited to note that $\mathcal{A}_0$ represents the first case (herd immunity already attained), $\mathcal{A}$ is the zone where no measure is necessary, $\mathcal{B}_0$ is a part of the region in which a progressive de-confinement is possible, while the part of $\mathcal{B}$ for which $S_+ > 19,142,857$ corresponds to configurations for which a strict confinement cannot be avoided.

7 Conclusions

We considered an optimal control problem for a SIR epidemic where the control is on the transmission rate coefficient b, under an ICU state constraint which prescribes that the infected population i has to stay always below a critical threshold i_M (representing the estimated maximum capacity of the health-care system) and a final condition which requires that at the final time t_f the susceptible population should be under the immunity threshold $\frac{\gamma}{\beta}$. The cost functional is assumed to be affine and depending only on the control variable.

We proved that, under viable initial conditions, the optimal strategy is as follows

- *do nothing* if the initial conditions allow for an evolution in which $i(t) \leq i_M$ for every $y \in [0, t_f]$ (meaning that the capacity of the health-care system is never exceeded, i.e. $(s_0, i_0) \in \mathcal{A}$);
- *do nothing until time τ_1 at which $i(\tau_1) = i_M$, then preserve the saturation $i = i_M$ until reaching the immunity threshold $\frac{\gamma}{\beta}$* if the initial conditions allow for this kind of control (i.e. if $(s_0, i_0) \in \mathcal{B}$ are below the curve $\mathbb{R}_+ \ni t \mapsto \left(s^{\frac{\gamma}{\beta_*}, i_M, \beta}(-t), i^{\frac{\gamma}{\beta_*}, i_M, \beta}(-t)\right)$);
- otherwise, *actuate the maximum level of lock-down before reaching $i = i_M$ and then preserve saturation;* the time length of the lock-down regime $(\tau_1 - \tau_0)$ depends on the initial conditions and on the coefficient β_* which gives the strength of the lock-down (see the inequality 5.12).

8 Appendix

8.1 Viability, Invariance and Other Tools

We recall here some tools from the theory of viability and see their qualitative implications on the system under study.

Definition 8.1 *Let $x^{x_0,b}$ be a solution to a controlled system governed by a regular (Lipschitz-continuous) field $f : \mathbb{R}^n \times B \to \mathbb{R}^n$ i.e.*

$$dx^{x_0,b}(t) = f\left(x^{x_0,b}(t), b(t)\right)dt,\ \ x^{x_0,b}(0) = x_0, \tag{8.1}$$

where B is some compact (subset of a) metric space and the admissible controls are Borel measurable functions $b \in \mathbb{L}^0(\mathbb{R}; B)$.

1. *A closed set $K \subset \mathbb{R}^n$ is said to be (forward-in-time)* viable *w.r.t.* (8.1) *if for every initial datum $x_0 \in K$ there exists an admissible control b such that the associated trajectory satisfies $x^{x_0,b}(t) \in K$ for all $t \geq 0$.*

2. *Given a closed set $K^+ \subset \mathbb{R}^n$, the largest subset $K \subset K^+$ such that trajectories starting at $x \in K$ can be maintained in K^+ is called a* viability kernel.
3. *A closed set $K \subset \mathbb{R}^n$ is said to be (forward-in-time)* invariant *w.r.t.* (8.1) *if for every initial datum $x_0 \in K$ and every admissible control b, the associated trajectory satisfies $x^{x_0,b}(t) \in K$ for all $t \geq 0$.*
4. *Given a viable constraint set $K \subset \mathbb{R}^n$ and a target set $K_0 \subseteq K$, the family $L \subset K$ of initial data that can be steered to K_0 in finite time (such that the associated trajectories do not leave K) is called a* capture basin *of K_0 from K.*

Remark 8.2 It is easy to note that a viability kernel is viable in time (i.e. trajectories starting in K remain not only in K^+ but actually in K). Indeed, if $x_0 \in K$, then $x^{x_0,b}(t) \in K^+$ for some control and all t (by definition of the viability kernel). Taking now $y_0 := x^{x_0,b}(t_0)$ (for some $t_0 > 0$), then $x^{y_0,b(\cdot+t_0)}(s) = x^{x_0,b}(s+t_0) \in K^+$. This proves that $y \in K$ (instead of the weaker condition $y \in K^+$). Thus the viability kernel is viable. It follows that a viability kernel is the largest viable set contained in K^+.

These notions are related to the Bouligand tangent (or contingent) and the (negative polar) normal cones. For our readers' sake, we recall these notions hereafter.

Definition 8.3 *Given a closed set $K \subset \mathbb{R}^n$, the* tangent cone *to K at a point $x \in K$ is the set*

$$T_K(x) = \left\{ d \in \mathbb{R}^n : \liminf_{\varepsilon \to 0+} \frac{d_K(x + \varepsilon d)}{\varepsilon} = 0 \right\},$$

where $d_K(y)$ denotes the distance of y from the set K. The normal cone *to K at $x \in K$ is the negative polar cone to $T_K(x)$, i.e.,*

$$N_K(x) = \{ p \in \mathbb{R}^n : \langle p, d \rangle \leq 0, \ \forall d \in T_K(x) \}.$$

The following result(s) gather some tools when dealing with viability and invariance cf. Propositions 3.4.1, 3.4.2, Theorem 3.2.4, Theorem 5.2.1 in [Aub09]¶.

Theorem 8.4 *With the notation of Definition 8.1, let us assume B to be convex and compact, $f(x,\cdot)$ to be convex and f to be Lipschitz in x and globally uniformly continuous.*

1. *A closed set $K \subset \mathbb{R}^n$ is viable with respect to* (8.1) *if and only if for every $x \in \partial K$ and every $p \in N_K(x)$,*
$$\inf_{b \in B} \langle p, f(x,b) \rangle \leq 0;$$
2. *A closed set $K \subset \mathbb{R}^n$ is invariant with respect to* (8.1) *if and only if for every $x \in \partial K$ and every $p \in N_K(x)$,*
$$\sup_{b \in B} \langle p, f(x,b) \rangle \leq 0;$$

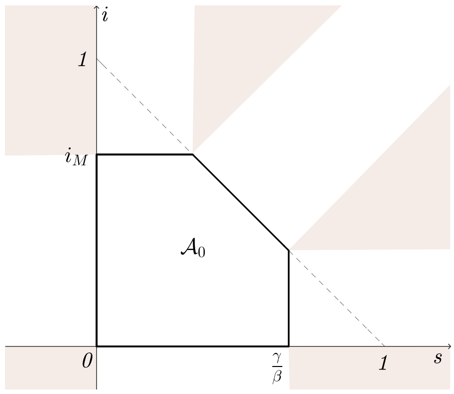
Remark 8.5 In practice, the domain K is often described by a regular frontier ϕ (C^1-diffeomorphism) i.e. $K = \{x \in \mathbb{R}^n : \phi(x) \leq 0\}$, $\partial K = \{x \in \mathbb{R}^n : \phi(x) = 0\}$ and the computation of normal sets is made to direct and inverse images roughly leading to the semi-line spanned by $\nabla\phi(x)$ (at $x \in \partial K$ with positive multiplicative constants) or intersections of such domains.

8.2 Proof of Theorem 2.3

This section is devoted to the proof of Theorem 2.3.

¶The cited reference considers differential inclusions but the properties can easily be obtained in these simple cases via the regularity assumptions we state in the theorem.

1. We use *2.* of Theorem 8.4 with the vector field
$$f(s,i,b) = \big(-sbi, (bs-\gamma)i\big).$$
The picture displays the set $\mathcal{A}_0$ with the normal cones in the singular points of the boundary. The set degenerates into a rectangle if $i_M \le 1 - \frac{\gamma}{\beta}$.



To prove that $\mathcal{A}_0$ is invariant (viable with any measurable control that takes the values in $[\beta_*, \beta]$) by checking that along the boundary condition *2.* of Theorem 8.4 is always satisfied is an easy task. Let us do it in the less trivial cases and left the others to the reader:

- along the upper side of $\partial\mathcal{A}_0$ we have
$$N_{\mathcal{A}_0}(s, i_M) = \{r(0,1) \ : \ r \ge 0\}, \ \forall s \in \big(0, \min\{\frac{\gamma}{\beta}, 1 - i_M\}\big);$$
$$\langle f(s, i_M, b), (0,1)\rangle = (bs-\gamma)i_M \le 0, \ \forall s \le \frac{\gamma}{\beta}, \ \forall b \le \beta.$$
- the diagonal side of $\partial\mathcal{A}_0$ appears only if $\frac{\gamma}{\beta} > 1 - i_M$; assuming to be in such case:
  - on the vertex between the upper and the diagonal sides we have
$$N_{\mathcal{A}_0}(1-i_M, i_M) = \{r(\lambda,1) \ : \ \lambda \in [0,1], \, r \ge 0\};$$
$$\langle f(1-i_M, i_M, b), (\lambda, 1)\rangle = \big(b(1-i_M)(1-\lambda) - \gamma\big)i_M \le 0, \ \forall b \le \beta, \ \forall \lambda \in [0,1].$$
  - along the diagonal side of $\partial\mathcal{A}_0$ we have
$$N_{\mathcal{A}_0}(s, 1-s) = \{r(1,1) \ : \ r \ge 0\}, \ \forall s \in \big(1 - i_M, \frac{\gamma}{\beta}\big);$$
$$\langle f(s, 1-s, b), (1,1)\rangle = -\gamma(1-s) \le 0, \ \forall s.$$
  - on the vertex between the diagonal and the vertical sides we have
$$N_{\mathcal{A}_0}\left(\frac{\gamma}{\beta}, 1 - \frac{\gamma}{\beta}\right) = \{r(1,\lambda) \ : \ \lambda \in [0,1], \, r \ge 0\};$$
$$\left\langle f(\frac{\gamma}{\beta}, 1 - \frac{\gamma}{\beta}, b), (1,\lambda)\right\rangle = \big(1 - \frac{\gamma}{\beta}\big)\big(b\frac{\gamma}{\beta}(1-\lambda) - \gamma\big) \le 0, \ \forall b \le \beta, \ \forall \lambda \in [0,1].$$

The remaining cases are almost trivial since the s component is non-increasing on $\mathcal{A}_0$ and when the initial $i = 0$ the system is stationary. The case in which $\frac{\gamma}{\beta} \le 1 - i_M$ and the set $\mathcal{A}_0$ is a rectangle is very similar.

2. The argument is quite similar to the previous one but one employs $b = \beta_*$ for every $s \le \frac{\gamma}{\beta_*}$.

3. We present the argument for $\mathcal{B}$, the argument for $\mathcal{A}$ being quite similar. The right-hand set in the definition of $\mathcal{B}_1$ is included in $\mathcal{B}$. To see this, one notes that, given $\left(\frac{\gamma}{\beta_*}, i_0\right) \in \mathcal{B}_0$, and an admissible control b, the function $t \mapsto s^{\frac{\gamma}{\beta_*}, i_0, b}(-t)$ is non-decreasing, hence $s^{\frac{\gamma}{\beta_*}, i_0, b}(-t) \geq \frac{\gamma}{\beta_*}$ for every $t \geq 0$. It follows that, in the same interval, the function $t \mapsto i^{\frac{\gamma}{\beta_*}, i_0, b}(-t)$ is non-increasing and, thus, $i^{\frac{\gamma}{\beta_*}, i_0, b}(-t) \leq i_M$ for all $t \geq 0$, hence satisfying the state-constraint. As a consequence, starting at some position $\left(s^{\frac{\gamma}{\beta_*}, i_0, b}(-t_0), i^{\frac{\gamma}{\beta_*}, i_0, b}(-t_0)\right) \in \mathcal{B}_1$, one reverses the time in b up till t_0 (thus reaching $\mathcal{B}_0$), then uses β_*.
Conversely, let $(s, i) \in \mathcal{B} \setminus \mathcal{B}_0$. We only need to consider the case when $i > 0$. In this case, $s > \frac{\gamma}{\beta_*}$ and, as long as $s^{s,i,b}(t) > \frac{\gamma}{\beta_*}$, the map $t \mapsto i^{s,i,b}(t)$ is increasing (for every $b \in \mathbb{L}^0(\mathbb{R}_+; [\beta_*, \beta])$), hence $i^{s,i,b}(t) \geq i$, for all t as before and $s^{s,i,b}(t) \leq s e^{-\beta_* i t}$. It follows that there exists (a unique) $t_1 > 0$ such that $s^{s,i,b}(t_1) = \frac{\gamma}{\beta_*}$. Since $(s, i) \in \mathcal{B}$, one has (for the viable control b), $i^{s,i,b}(t_1) \leq i_M$. It follows that $\left(s^{s,i,b}(t_1), i^{s,i,b}(t_1)\right) \in \mathcal{B}_0$ and, hence, by reversing the time, and setting $b^-(t) := b(t_1 - t)$, one gets $(s, i) = \left(s^{s^{s,i,b}(t_1), i^{s,i,b}(t_1), b^-}(-t_1), i^{s^{s,i,b}(t_1), i^{s,i,b}(t_1), b^-}(-t_1)\right)$, thus concluding our argument.

4. We only prove the assertion for $\mathcal{B}$ (the remaining relations being quite similar). We take $(s_0, i_0) \in \mathcal{B}$ and a viable control b. If $t_1 = t_0$ the conclusion is trivial. Let us then consider the case in which $i_1 < i_0$. Any trajectory (say with control b') starting from (s_0, i_1) has a non-increasing s and, thus, satisfies

$$\left\{\left(s^{s_0, i_1, b'}(t), i^{s_0, i_1, b'}(t)\right) : \ t \geq 0\right\} \subset R := [0, s_0] \times [0, 1].$$

The associated trajectory $Reach(s_0, i_0) := \left\{\left(s^{s_0, i_0, b}(t), i^{s_0, i_0, b}(t)\right) : \ t \geq 0\right\} \subset R$ is asymptotically (as t$\to \infty$) directed towards $i = 0$ and, thus, separates the line $i = i_M$ and the point (s_0, i_1) (in the rectangle R). We consider the hitting time

$$0 < t_0 := \inf\left\{t > 0 : \left(s^{s_0, i_1, b'}(t), s^{s_0, i_1, b'}(t)\right) \in Reach(s_0, i_0)\right\}.$$

If $t_0 = \infty$, then b' is obviously a viable control for (s_0, i_1). Otherwise, for some $t > 0$, $\left(s^{s_0, i_1, b'}(t_0), s^{s_0, i_1, b'}(t_0)\right) = \left(s^{s_0, i_0, b}(t), s^{s_0, i_0, b}(t)\right)$. We modify b' by setting $\tilde{b}'(r) := b'(r)\mathbf{1}_{[0,t_0]}(r) + b(r - t_0 + t)\mathbf{1}_{r > t_0}$ to get a viable control for (s_0, i_1).
Our second assertion follows by setting

$$\begin{aligned}
\Phi_{\mathcal{A}^0}(s) &:= \sup\left\{i : \ (s, i) \in \mathbb{T}, i^{s,i,b}(t) \leq i_M, \ \forall t \geq 0, \ \forall b \in \mathbb{L}^0(\mathbb{R}; [\beta_*, \beta])\right\}, \\
\Phi_{\mathcal{A}}(s) &:= \sup\left\{i : \ (s, i) \in \mathbb{T}, i^{s,i,\beta}(t) \leq i_M, \ \forall t \geq 0\right\}, \\
\Phi_{\mathcal{B}}(s) &:= \sup\left\{i : \ (s, i) \in \mathbb{T}, \exists\, b \in \mathbb{L}^0(\mathbb{R}; [\beta_*, \beta]) \ s.t. \ i^{s,i,b}(t) \leq i_M, \ \forall t \geq 0\right\}.
\end{aligned}$$

Let us prove that

$$\mathcal{B} = \{(s, i) \in \mathbb{T} : \ i \leq \Phi_{\mathcal{B}}(s)\},$$

the other cases being similar and easier.
Let us prove first the inclusion $\subseteq$. Let $(s_0, i_0) \in \mathcal{B}$. Then, there exists b such that $i^{s_0, i_0, b}(t) \leq i_M$ for any $t \geq 0$. Then

$$i_0 \in \left\{i : \ (s_0, i) \in \mathbb{T}, \ \exists\, b \in \mathbb{L}^0(\mathbb{R}; [\beta_*, \beta]) \text{ s.t. } i^{s_0, i, b}(t) \leq i_M \ \forall t \geq 0,\right\}$$

and therefore

$$i_0 \leq \sup\left\{i : \ (s_0, i) \in \mathbb{T}, \ \exists\, b \in \mathbb{L}^0(\mathbb{R}; [\beta_*, \beta]) \text{ s.t. } i^{s_0, i, b}(t) \leq i_M \ \forall t \geq 0,\right\} = \Phi_{\mathcal{B}}(s_0),$$

that is $(s_0, i_0) \in \{(s, i) \in \mathbb{T} : \ i \leq \Phi_{\mathcal{B}}(s)\}$, which proves the claimed inclusion.

Let us prove the opposite inclusion $\supseteq$. Let $(s_0, i_0) \in \{(s,i) \in \mathbb{T} : \ i \le \Phi_{\mathcal{B}}(s)\}$, that is

$$i_0 \le \Phi_{\mathcal{B}}(s_0) = \sup\left\{ i : \ (s_0, i) \in \mathbb{T}, \ \exists\, b \in \mathbb{L}^0\left(\mathbb{R}; [\beta_*, \beta]\right) \text{ s.t. } i^{s_0,i,b}(t) \le i_M \ \forall t \ge 0, \right\}. \tag{8.2}$$

If the strict inequality holds, then there exists $i_1 > i_0$ and $b \in \mathbb{L}^0\left(\mathbb{R}; [\beta_*, \beta]\right)$ such that $(s_0, i_1) \in \mathbb{T}$, $i^{s_0,i,b}(t) \le i_M$, $\forall\, t \ge 0$. Then, using the first part of this item, we have the existence of some admissible b' such that $i^{s_0,i_0,b'}(t) \le i_M$, i.e. $(s_0, i_0) \in \mathcal{B}$ and the proof is concluded in this case.

If, otherwise, in (8.2) the equality holds, then we have that for any $\varepsilon > 0$ there exists $i_\varepsilon > i_0 - \varepsilon$ and $b_\varepsilon \in \mathbb{L}^0\left(\mathbb{R}; [\beta_*, \beta]\right)$ such that $(s_0, i_\varepsilon) \in \mathbb{T}$, $i^{s_0,i_\varepsilon,b_\varepsilon}(t) \le i_M$, $\forall\, t \ge 0$. Then we have (for some modified controls b'_ε,

$$i^{s_0,i_0-\varepsilon,b'_\varepsilon}(t) \le i_M \tag{8.3}$$

By extracting from (b'_ε) a weakly-* converging sequence and sending $\varepsilon \to 0$, by using continuous dependence on the data and the uniqueness of the solution of the initial value problem for the system of state equations, we can deduce that $i^{s_0,i_0,b}(t) \le i_M$.

5. If ϕ is regular enough, then the explicit computation of contingent cones and normal cones to $K := \{(s,i) : \ i \le \phi(s)\}$ yields

$$N_K(s, \phi(s)) = \{r(\partial_s, \partial_i)\,(i - \phi(s)) \ : \ r \ge 0\} = \{r\,(-\phi'(s), 1) \ : \ r \ge 0\}.$$

The viability condition *1.* of Theorem 8.4 for $\mathcal{B}$ yields, for frontier points $(s,i) = (s, \phi(s))$,

$$\exists\, b \in [\beta_*, \beta] \ \text{ s.t. } \phi'(s) bs\phi(s) + (bs - \gamma)\,\phi(s) \le 0, \tag{8.4}$$

which implies

$$\phi'(s) \le -1 + \frac{\gamma}{\beta_* s}.$$

It follows that a necessary condition for ϕ to be constant is $s \le \frac{\gamma}{\beta_*}$. On the other hand, the maximal admissible constant is $\phi = i_M$. It is immediately seen that, for this constant, the condition is also sufficient, that is $s_0 \le \frac{\gamma}{\beta_*}$, $i_0 \le i_M$ implies $(s_0, i_0) \in \mathcal{B}$. Indeed, since s is nonincreasing, with β_* control we have $i' = (s\beta_* - \gamma) i \le 0$. Then we have proven that $\Phi_{\mathcal{B}}(s) = i_M$, $\forall s \le \frac{\gamma}{\beta_*}$ and decreases afterwards. Since $\mathbb{T}$ is invariant, it follows that the intersection with $\mathbb{T}$ is viable. For $s \ge \frac{\gamma}{\beta_*}$, the (maximal) solution (realizing the equality in the previous inequality) is given by

$$\Phi_{\mathcal{B}}(s) = i_M - s + \frac{\gamma}{\beta_*} + \frac{\gamma}{\beta_*} \log\left(\frac{\beta_* s}{\gamma}\right).$$

Since $\Phi_{\mathcal{B}}$ is non-negative and $i = 0$ is stationary, the conclusion follows. The argument for $\mathcal{A}$ is similar but one reasons for $b = \beta$. For $\mathcal{A}_0$ one asks that (8.4) be satisfied for all $b \in [\beta_*, \beta]$, thus arriving on the same set as for $\mathcal{A}$.

6. One writes

$$\frac{d}{dt}\left(i - \Phi_{\mathcal{B}}(s)\right) = bsi - \gamma i + \left(-1 + \frac{\gamma}{\beta_* s}\right) bsi = \gamma i \left(\frac{b}{\beta_*} - 1\right) \ge 0,$$

whenever b is admissible, and as long as $s > \frac{\gamma}{\beta_*}$. As a consequence, starting from $(s_0, i_0) \in \partial\mathcal{B}$ with $s_0 > \frac{\gamma}{\beta_*}$, one either exits this region (thus violating viability) or at most stays on the boundary and, in this case, the viable control is β_* as described in the statement.

The theorem is completely proved.

□

References

[AAL20] Fernando E Alvarez, David Argente, and Francesco Lippi, *A simple planning problem for covid-19 lockdown*, Tech. report, National Bureau of Economic Research, 2020.

[AAM92] Roy M Anderson, B Anderson, and Robert M May, *Infectious diseases of humans: dynamics and control*, Oxford university press, 1992.

[ACMM+21] Marco Tulio Angulo, Fernando Castaños, Rodrigo Moreno-Morton, Jorge X. Velasco-Hernández, and Jaime A. Moreno, *A simple criterion to design optimal non-pharmaceutical interventions for mitigating epidemic outbreaks*, Journal of The Royal Society Interface **18** (2021), no. 178, 20200803.

[Aub09] Jean-Pierre Aubin, *Viability theory*, Birkhäuser Boston, Boston, MA, 2009.

[BBDMG19] Luca Bolzoni, Elena Bonacini, Rossella Della Marca, and Maria Groppi, *Optimal control of epidemic size and duration with limited resources*, Mathematical biosciences **315** (2019), 108232.

[BBSG17] Luca Bolzoni, Elena Bonacini, Cinzia Soresina, and Maria Groppi, *Time-optimal control strategies in sir epidemic models*, Mathematical biosciences **292** (2017), 86–96.

[BdlV10] Joseph Frédéric Bonnans and Constanza Sánchez Fernández de la Vega, *Optimal control of state constrained integral equations*, Set-Valued and Variational Analysis **18** (2010), no. 3-4, 307–326.

[BDLVD13] J Frédéric Bonnans, Constanza De La Vega, and Xavier Dupuis, *First-and second-order optimality conditions for optimal control problems of state constrained integral equations*, Journal of Optimization Theory and Applications **159** (2013), no. 1, 1–40.

[Beh00] Horst Behncke, *Optimal control of deterministic epidemics*, Optimal control applications and methods **21** (2000), no. 6, 269–285.

[BFZ11] Bokanowski, Olivier, Forcadel, Nicolas, and Zidani, Hasnaa, *Deterministic state-constrained optimal control problems without controllability assumptions*, ESAIM: COCV **17** (2011), no. 4, 995–1015.

[BGG+17] J. Bonnans, Frederic, Daphne Giorgi, Vincent Grelard, Benjamin Heymann, Stephan Maindrault, Pierre Martinon, Olivier Tissot, and Jinyan Liu, *Bocop – A collection of examples*, Tech. report, INRIA, 2017.

[Bon20] J. F. Bonnans, *Course on optimal control. part i: the Pontryagin approach.*, 2020.

[BP03] Ugo Boscain and Benedetto Piccoli, *Optimal syntheses for control systems on 2-d manifolds*, vol. 43, Springer Science & Business Media, 2003.

[BP05] Ugo Boscain and Benetto Piccoli, *An introduction to optimal control*, Contrôle non linéaire et applications, Herman, Paris (2005), 19–66.

[Cla13] Francis Clarke, *Functional analysis, calculus of variations and optimal control*, Springer London, London, 2013.

[DGI18] Paolo Di Giamberardino and Daniela Iacoviello, *Direct integrability for state feedback optimal control with singular solutions*, International Conference on Informatics in Control, Automation and Robotics, Springer, 2018, pp. 482–502.

[FP00] Hélène Frankowska and Slawomir Plaskacz, *Semicontinuous solutions of hamilton–jacobi–bellman equations with degenerate state constraints*, Journal of Mathematical Analysis and Applications **251** (2000), no. 2, 818–838.

[Fre] Lorenzo Freddi, *Optimal control of the transmission rate in compartmental epidemics*, Mathematical Control and Related Fields, doi: 10.3934/mcrf.2021007.

[HD11] Elsa Hansen and Troy Day, *Optimal control of epidemics with limited resources*, Journal of mathematical biology **62** (2011), no. 3, 423–451.

[Ket21] David I. Ketcheson, *Optimal control of an sir epidemic through finite-time non-pharmaceutical intervention*, Journal of Mathematical Biology **83** (2021), no. 1, 7.

[KK20] Markus Kantner and Thomas Koprucki, *Beyond just "flattening the curve": Optimal control of epidemics with purely non-pharmaceutical interventions*, Journal of Mathematics in Industry **10** (2020), no. 1, 1–23.

[KMW27] William Ogilvy Kermack, A. G. McKendrick, and Gilbert Thomas 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** (1927), no. 772, 700–721.

[KS20] Thomas Kruse and Philipp Strack, *Optimal control of an epidemic through social distancing.*

[Mar15] Maia Martcheva, *An introduction to mathematical epidemiology*, vol. 61, Springer, 2015.

[MSW20] Laurent Miclo, Daniel Spiro, and Jörgen Weibull, *Optimal epidemic suppression under an icu constraint*, arXiv preprint arXiv:2005.01327 (2020).

[Pon18] Lev Semenovich Pontryagin, *Mathematical theory of optimal processes*, Routledge, 2018.

[SL12] Heinz Schättler and Urszula Ledzewicz, *Geometric optimal control: theory, methods and examples*, vol. 38, Springer Science & Business Media, 2012.

[SM17] Oluwaseun Sharomi and Tufail Malik, *Optimal control in epidemiology*, Annals of Operations Research **251** (2017), no. 1-2, 55–71.

[SS78] Suresh P Sethi and Preston W Staats, *Optimal control of some simple deterministic epidemic models*, Journal of the Operational Research Society **29** (1978), no. 2, 129–136.

[TC17] Inria Saclay Team Commands, *Bocop: an open source toolbox for optimal control*, `http://bocop.org`, 2017.