Annual Reviews in Control 52 (2021) 508–522

Contents lists available at ScienceDirect

Annual Reviews in Control

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

Full Length Article

Hysteresis-based supervisory control with application to non-pharmaceutical containment of COVID-19☆

Michelangelo Bin [a,*], Emanuele Crisostomi [b], Pietro Ferraro [c], Roderick Murray-Smith [d], Thomas Parisini [a,e,f], Robert Shorten [c,g], Sebastian Stein [d]

[a] Department of Electrical and Electronic Engineering, Imperial College London, London, UK
[b] Department of Energy, Systems, Territory and Constructions Engineering, University of Pisa, Pisa, Italy
[c] Dyson School of Design Engineering, Imperial College London, London, UK
[d] School of Computing Science, University of Glasgow, Glasgow, Scotland
[e] KIOS Research and Innovation Center of Excellence, University of Cyprus, Aglantzia, Cyprus
[f] Department of Engineering and Architecture, University of Trieste, Trieste, Italy
[g] Department of Electrical and Electronic Engineering, University College Dublin, Dublin, Ireland

ARTICLE INFO

Keywords:
COVID-19
Supervisory control
Hysteresis control

ABSTRACT

The recent COVID-19 outbreak has motivated an extensive development of non-pharmaceutical intervention policies for epidemics containment. While a total lockdown is a viable solution, interesting policies are those allowing some degree of normal functioning of the society, as this allows a continued, albeit reduced, economic activity and lessens the many societal problems associated with a prolonged lockdown. Recent studies have provided evidence that fast periodic alternation of lockdown and normal-functioning days may effectively lead to a good trade-off between outbreak abatement and economic activity. Nevertheless, the correct number of normal days to allocate within each period in such a way to guarantee the desired trade-off is a highly uncertain quantity that cannot be fixed a priori and that must rather be adapted online from measured data. This adaptation task, in turn, is still a largely open problem, and it is the subject of this work. In particular, we study a class of solutions based on hysteresis logic. First, in a rather general setting, we provide general convergence and performance guarantees on the evolution of the decision variable. Then, in a more specific context relevant for epidemic control, we derive a set of results characterizing robustness with respect to uncertainty and giving insight about how a priori knowledge about the controlled process may be used for fine-tuning the control parameters. Finally, we validate the results through numerical simulations tailored on the COVID-19 outbreak.

1. Introduction

1.1. Problem description

Different recent studies Bin, Cheung et al. (2021), Della Rossa et al. (2020), Ferguson et al. (2020), Giordano et al. (2021), Karin et al. (2020), Kennedy, Zambrano, Wang, and Neto (2020), Morato, Bastos, Cajueiro, and Normey-Rico (2020), Sadeghi, Greene, and Sontag (2021) and Sontag (2021) have provided evidence that the *fast* alternation of lockdown and "normal" days may effectively hinder the COVID-19 outbreak while permitting a reduced but sustained economic activity. While full lockdown maximally abates the outbreak, leaving some normal days lessens the social and economical stress associated with a continued lockdown. Hence, a compromise between lockdown and normal days has to be found.

In particular, the main result of Bin, Cheung et al. (2021) states that if one has to allocate a given number of lockdown and normal days, the best way, from the standpoint of minimizing the number of infections, is to switch between them at the largest possible frequency.

☆ M.B., P.F., R.M.-S., T.P., R.S. & S.S. acknowledge support from EPSRC, UK project EP/V018450/1. M.B. and T.P. acknowledge funding support from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 739551 (KIOS CoE). T.P. acknowledges funding support from the Italian Ministry for Research in the framework of the 2017 Program for Research Projects of National Interest (PRIN), Grant no. 2017YKXYXJ. P.F. acknowledges the support of IOTA Foundation (SFI grant 16/IA/4610). R.M.-S. and S.S. acknowledge funding support from EPSRC, UK grant EP/R018634/1, "Closed-loop Data Science".

* Corresponding author.
E-mail address: m.bin@imperial.ac.uk (M. Bin).

https://doi.org/10.1016/j.arcontrol.2021.07.001
Received 6 May 2021; Received in revised form 7 July 2021; Accepted 20 July 2021
Available online 13 August 2021

Specifically, if one considers a periodic arrangement of lockdown and normal days (called in Bin, Cheung et al., 2021 a *Fast Periodic Switching Policy (FPSP)*), then the smaller is the period the closer the induced epidemic trajectories are to those of a fictitious epidemic characterized by a *reproductive number* which is a weighted average of the ones we would obtain, respectively, with full lockdown and no lockdown at all (Bin, Cheung et al., 2021, Theorem 1). The average weights are defined by the policy's *duty cycle*, i.e. the relative number of normal days in each period.

An FPSP is thus characterized by two degrees of freedom: its duty cycle, which defines the reproductive number of the fictitious epidemic, and its frequency, which determines how well the actual epidemic approximates the fictitious one. The frequency has to be taken as large as possible, consistent with societal constraints, and it thus represents an "open-loop parameter" (in particular, the numerical analysis of Bin, Cheung et al. (2021) shows that periods of 1–4 weeks are good choices in the case of COVID-like pandemics). The duty cycle, instead, cannot be fixed once for all in advance. In fact, the range of values for which the relative number of normal days is large whilst the epidemic is still taken under control as desired is highly uncertain, and depends on the unknown characteristics of the epidemic. To deal with this problem, in Bin, Cheung et al. (2021) a slow data-driven supervisory controller is proposed to adapt the FPSP's duty cycle at run-time on the basis of observations, suitably averaged and filtered over longer time periods. The supervisor realizes a basic hysteresis switching logic, as it is sufficiently simple to be implemented by a policy maker but sufficiently robust to deal with the high uncertainty characterizing the measured signals.

While the supervisor shows promising performance in the numerical analysis of Bin, Cheung et al. (2021), its claimed robustness and exactness properties have not been formally established yet, and some fundamental points remained open. Specifically:

P1. The regulation properties of the employed hysteresis controller have not been given a formal characterization. What guarantees does hysteresis-based control provide?

P2. The relation between the control parameters and the characteristics of the underlying controlled process and the uncertainty and delays affecting measurements is not clear. How can we use a priori information about the epidemic to improve performance?

As detailed below in Sections 1.2 and 1.3, the current state of the art of control theory does not satisfactorily cover P1 and P2 in a context relevant for epidemics mitigation. Moreover, P1 and P2 raise questions of crucial importance not only for Bin, Cheung et al. (2021), but also for all the other possible data-based mitigation techniques employing an hysteresis-like mechanism (these include many heuristic decision policies currently adopted worldwide by politicians). Therefore, developing a comprehensive theory addressing P1 and P2 in a context general enough to embrace many different hysteresis-based mitigation techniques is a timely open problem with a considerable potential impact on epidemics mitigation going well beyond the specific technique used in Bin, Cheung et al. (2021). The development of such theory is, in a nutshell, the main goal of this article.

1.2. Literature overview

The hysteresis decision logic follows a rather basic rule: when a measured *"monitoring signal"* exceeds a given value, action is taken to change the process behavior so as to bring the monitoring signals withing bounds again. Controllers based on this logic are simple to implement, and boast inherent robustness as only a rough knowledge about how decisions qualitatively affect the controlled process is required for the hysteresis logic to function well.

When the size of the hysteresis band is zero and the decision map is single-valued, the hysteresis logic reduces to a basic "step" nonlinearity, leading to state-dependent switching between two values. In this simplified form, it provides the solution to minimum-time optimal control problems for both ODEs (Bellman, Glicksberg, & Gross, 1956; Krener, 1974) and PDEs (Casas, Wachsmuth, & Wachsmuth, 2017; Mizel & Seidman, 1997), well-known under the name of *"bang–bang controller"* and boasting countless applications ranging from reservoir flooding (Zandvliet, Bosgra, Jansen, Van den Hof, & Kraaijevanger, 2007), vibration control of structures (Lim, Chung, & Moon, 2003), game theory (Olsder, 2002), control of quantum systems (Morton et al., 2006; Viola & Lloyd, 1998), and optimal intervention for cancer chemotherapy (Ledzewicz & Schättler, 2002). Moreover, when the two switching values are equal, the control logic takes the form of an amplified *"sign"* function, and in this form it is widely used especially in the context of *sliding mode control* (Slotine & Sastry, 1983), with many applications in mechatronics and robotics (Šabanovic, 2011; Slotine & Sastry, 1983; Zhihong, Paplinski, & Wu, 1994).

In the general form with non-zero band size, hysteresis controllers are instead standard components of industrial electronic devices (Bose, 1990; Buso, Fasolo, Malesani, & Mattavelli, 2000; Kawamura & Hoft, 1984; Li, Ruan, Zhang, & Lo, 2020; Mohseni & Islam, 2010; Šabanovic, 2011), and provide simple and robust solutions for a wide range of problems in engineering, from combustion devices (Guan et al., 2019) and temperature regulation (Cahlon, Schmidt, Shillor, & Zou, 1997; Gurevich, Jäger, & Skubachevskii, 2009), to epidemic control (Bin, Cheung et al., 2021; Ferguson et al., 2020), as mentioned earlier. Moreover, the hysteresis logic is a basic design principle for supervisory and hierarchical control, and lies at the core of many adaptive control approaches (Angeli & Mosca, 2004; Baldi, Battistelli, Mari, Mosca, & Tesi, 2012; Battistelli, Hespanha, & Tesi, 2012; Hespanha, Liberzon, & Morse, 2002, 2003; Hespanha & Morse, 1999; Jin & Safonov, 2012; Kosmatopoulos & Ioannou, 1999; Ma, 2008; Middleton, Goodwin, Hill, & Mayne, 1988; Morse, 1990, 1996, 1997; Morse, Mayne, & Goodwin, 1992; Stefanovic & Safonov, 2008; Stefanovic, Wang, & Safonov, 2004; Vu & Liberzon, 2011; Weller & Goodwin, 1994; Ye, 2008).

Outside application-specific analyses, adaptive control is indeed the field where hysteresis logic has been studied more thoroughly from a control standpoint, as it provides a more direct exploration tool with respect to the non-uniform, exhaustive searches of early "universal controllers" (Mårtensson, 1985; Minyue Fu & Barmish, 1986) along predefined paths. Specifically, in Baldi et al. (2012), Hespanha et al. (2002, 2003), Middleton et al. (1988), Morse (1990, 1996, 1997), Morse et al. (1992) and Weller and Goodwin (1994) hysteresis-based control is used to switch between candidate controllers on the basis of a measured monitoring signal in such a way to stabilize uncertain linear systems. A common feature of these approaches is the presence of multiple candidate controllers, each one associated with a monitoring signal measuring its performance. The control logic chooses at run time the controller associated with the best performance. Under the assumption that a stabilizing controller exists among the candidates, asymptotic stability can be concluded, with switching that typically stops in finite time. Extensions to nonlinear or time-varying systems in a discrete or continuous-time setting are provided, for instance, in Angeli and Mosca (2004), Battistelli et al. (2012), Bin, Bernard, and Marconi (2021), Bin and Marconi (2020), Bin, Marconi, and Teel (2019), Hespanha and Morse (1999), Jin and Safonov (2012), Kosmatopoulos and Ioannou (1999), Ma (2008), Stefanovic and Safonov (2008), Stefanovic et al. (2004), Vu and Liberzon (2011) and Ye (2008).

Finally, we underline that the convergence analysis of Morse et al. (1992) and Stefanovic et al. (2004) and related extensions, e.g. Hespanha et al. (2002, 2003), Hespanha and Morse (1999), Jin and Safonov (2012), Morse (1996, 1997) and Stefanovic and Safonov (2008), is carried out in a *modular way*. First, some convergence properties for the hysteresis logic decision variables are proved depending only on "open-loop assumptions" about the measured variables (see, e.g., the *"Hysteresis Switching Lemma"* of Morse et al. (1992)). Then, these are used to prove the desired regulation performance during closed-loop operation.

1.3. Contribution and organization

Although multi-model supervisory control is a well-established field, and the aforementioned modularity in the analysis is favorable in an uncertain setting such as virus outbreaks, we cannot directly rely on the aforementioned results for several reasons. First, we do not dispose of reliable models or "predictors" assessing the quality of a decision before it is implemented, so as at switching time we cannot choose the decision minimizing an expected cost. Indeed, as the recent COVID-19 shows, assessing the effect of non-pharmaceutical decisions is a complex task. Models are incapable of guaranteeing reliable predictions (Hutson, 2020), and typically effects can be only assessed a posteriori from data (Ferguson et al., 2020; Flaxman et al., 2020). Hence, motivated by the problem described in Section 1.1, we shall rather consider a case in which the effect of a decision may be only assessed a posteriori, after it is held for a while.

Second, even if predictors were available, the aforementioned approaches have as objective convergence to a stabilizing controller, typically the "most stabilizing," which in our case would result in enforcing a *full lockdown*. This is in sharp contrast with our objective which, we recall, is that of keeping the outbreak under control with the minimum number of lockdown days. Lastly, we cannot rely on exhaustive search, such as the universal controller of Mårtensson (1985) and related extensions, for obvious time constraints requiring uniformity and fast convergence. Hence, Points P1 and P2 (Section 1.1) remain substantially uncovered.

In this article, we study a class of hysteresis-based control schemes addressing Items P1 and P2. As we aim at a methodology that applies to a wider spectrum of techniques than just Bin, Cheung et al. (2021), in Section 2 we first develop a general theory in a broader context than epidemic outbreaks. Under many aspects, this is the heart of the paper, where we develop general results without relying on quantitative models such as SI-like differential equations. Specifically, along the lines of Morse et al. (1992) and Stefanovic et al. (2004), we provide analytical results on the dynamics of the decision variables and we establish a set of convergence and stability properties that are independent from the underlying controlled process and only depend on the measured signals. Then, we identify sufficient conditions ensuring that the generated decision variables properly regulate the underlying process as desired. Similar to Jin and Safonov (2012), Stefanovic and Safonov (2008) and Stefanovic et al. (2004), these conditions take the form of a *robust detectability assumption* guaranteeing that the information extracted from the measurements effectively matches the actual status of the underlying unmeasured process. Proceeding in this way allows us to separate the role of the control logic from that played by the assumptions we make on the underlying controlled process and on the relationship between its dynamics and our measurements, including the effect of uncertainty and delays. Moreover, it confers modularity on the analysis, making it applicable to a broader class of problems.

In Section 3, we then focus on the problem of virus outbreak mitigation and, in particular, on the specific context of Bin, Cheung et al. (2021). We provide robustness results with respect to uncertainty affecting the measurements, and we provide results and insights about how to use a priori information on the controlled process to choose the design parameters. Finally, in Section 4 we provide numerical simulation validating the developed theory on the case of COVID-19.

1.4. Notation

We denote by $\mathbb{R}$ and $\mathbb{N}$ the sets of real and natural numbers, respectively. If $\preceq$ is any order on a set S, for an $s \in S$, $S_{\preceq s} := \{z \in S : z \preceq s\}$. By $A \setminus B := \{a \in A : c \notin B\}$ we denote the set difference and by A^{c} the complement. Moreover, $\overline{S}$ and S° denote the closure and interior of S and $\partial S = \overline{S} \setminus S^{\circ}$ its boundary. The symbol $\subset$ denotes non-strict inclusion. Strict inclusion is denoted by $\subsetneq$. When clear from the context, we shall identify singletons with their element. For instance, we shall use $X \setminus x$ and $x \times X$ in place of $X \setminus \{x\}$ and $\{x\} \times X$ respectively. We denote by B^A the set of functions $A \to B$, as well as families or nets in B indexed by A. For a net $((a_k, b_k))_{k \in K}$ of elements (a_k, b_k), we use the short notation $(a_k, b_k)_{k \in K}$. We denote by $B^{\subset A}$ the set of functions $C \to A$ with $C \subset B$. Given a set $\mathcal{A}$ and a totally ordered real vector space $(\mathbb{T}, \leq)$, a *flow* on $(\mathbb{T}, \mathcal{A})$ is a function $\phi : \operatorname{dom}\phi \subset \mathbb{T} \times \mathcal{A} \to \mathcal{A}$ such that: **(i)** for each $a \in \mathcal{A}$ such that $(0, a) \in \operatorname{dom}\phi$, $\phi(0, a) = a$; **(ii)** if $t, s, \tau \in \mathbb{T}$ are such that $s \leq \tau \leq t$ and $(t, a), (s, a) \in \operatorname{dom}\phi$ for some $a \in \mathcal{A}$, then $(\tau, a) \in \operatorname{dom}\phi$; **(iii)** if $(s, a) \in \operatorname{dom}\phi$, then for all $t \in \mathbb{T}$ such that $(t + s, a) \in \operatorname{dom}\phi$, it holds that $(t, \phi(s, a)) \in \operatorname{dom}\phi$, and $\phi(t, \phi(s, a)) = \phi(t+s, a)$. The set of flows on $(\mathbb{T}, \mathcal{A})$ is denoted by $\boldsymbol{\Phi}(\mathbb{T}, \mathcal{A})$.

2. General theory

2.1. Basic definitions

We denote by $\mathcal{X}$ the *decision space*. For simplicity, we shall consider the case in which $\mathcal{X}$ is finite, although all results apply to countable decision spaces as well with uniformity of convergence over finite subset of initial conditions in place of fixed-time convergence. We endow $\mathcal{X}$ with the discrete topology, so that convergence in $\mathcal{X}$ means convergence in finite-time. We assume that $\mathcal{X}$ can be ordered by a total order $\leq$ (we write $x \geq y$ for $y \leq x$, $x < y$ if $x \leq y$ and $x \neq y$, and $x > y$ for $y < x$), and given any $x \in \mathcal{X}$, we denote by

$$x^+ = \begin{cases} \inf \mathcal{X}_{>x} & \text{if } \mathcal{X}_{>x} \neq \emptyset, \\ x & \text{otherwise,} \end{cases}$$

$$x^- = \begin{cases} \sup \mathcal{X}_{<x} & \text{if } \mathcal{X}_{<x} \neq \emptyset, \\ x & \text{otherwise,} \end{cases}$$

the (projected) successor and predecessor of x in $(\mathcal{X}, \leq)$.

With $\mathcal{Y}$ and $\mathcal{Z}$ sets, the controlled system is modeled by a pair $\Sigma = (\zeta, \psi)$, in which $\zeta : \mathcal{X} \to \boldsymbol{\Phi}(\mathbb{R}, \mathcal{Z})$ is a function mapping decisions $x \in \mathcal{X}$ into *process trajectory flows* $\zeta[x] \in \boldsymbol{\Phi}(\mathbb{R}, \mathcal{Z})$ (see Section 1.4), and ψ is an operator mapping trajectories $z : \operatorname{dom} z \subset \mathbb{R} \to \mathcal{Z}$ and decisions $x \in \mathcal{X}$ to *measurement signals* $\psi[z, x] : \operatorname{dom} z \to \mathcal{Y}$. Specifically, ζ represents the unknown family of trajectories of the uncertain underlying process, for instance an epidemic, which we would like to control but which we do not measure. Instead, the images of ψ represent the available measurements that can be used for control, and depend on the decision taken and on the underlying process. Both $\zeta[x]$ and $\psi[z, x]$ describe signals evolving in continuous time. This carries no loss of generality as it includes discrete-time signals as particular case, since they can always be extended to $\mathbb{R}$. Nevertheless, we restrict our focus only to discrete-time updates of the decision variable.

A *decision profile* is a pair $\xi = \big((x_k, t_k)_{k \in \operatorname{dom}\xi}, \bar{t}\big)$ in which $(x_k, t_k)_{k \in \operatorname{dom}\xi}$ is a sequence of decisions and decision times, with $\operatorname{dom}\xi$ of the form $\{0, 1, \dots, n\}$ for some $n \in \mathbb{N}$, and $\bar{t} \geq 0$. We assume that from every initial condition z_0 every decision profile ξ induces a *realization* of the process and the measured signal, denoted by $z(\cdot|\xi)$ and $y(\cdot|\xi)$, defined on

$$\operatorname{dom} z(\cdot|\xi) = \operatorname{dom} y(\cdot|\xi) = \bigcup_{k \in \operatorname{dom}\xi} [t_k, t_{k+1}], \tag{1}$$

with $z(t_0|\xi) = z_0$, and such that for all $k \in \operatorname{dom}\xi$

$$\begin{aligned} z(t|\xi) &= \zeta[x_k](t - t_k, z(t_k|\xi)) && \forall t \in (t_k, t_{k+1}], \\ y(t|\xi) &= \psi[z(\cdot|\xi), x_k](t) && \forall t \in (t_k, t_{k+1}], \end{aligned} \tag{2}$$

where we let $t_{k+1} = t_k + \bar{t}$ for $k = \sup \operatorname{dom}\xi$ in both (1) and (2). With slight abuse of notation we shall refer to y and z generically as "signals". Moreover, we shall omit references to ξ or x when clear from the context.

From now on, we suppose z_0 is fixed (albeit unknown) and we drop the dependency from it. Thus, given a decision profile ξ, in the following we denote without ambiguity by $z(\cdot|\xi)$ and $y(\cdot|\xi)$ the signals defined as above corresponding to z_0.

2.2. The logic workflow and the evaluation principle

The aim of the decision logic is to generate a suitable decision profile inducing the desired behavior of the process ζ. Decisions are taken recursively at some *decision times* $t_0, t_1, \ldots \in \mathbb{R}$. For simplicity, we assume that decisions are taken periodically, with period $\Delta \in \mathbb{R}_{>0}$. Hence, $t_{k+1} = t_k + \Delta$ for all $k \in \mathbb{N}$. This is not necessary in principle, yet it simplifies the forthcoming analysis.

The interaction between the controlled process and the decision logic follows the workflow described below starting at $k = 0$:

S1. The controller takes the decision x_k at time t_k on the basis of the information available up to time t_k, which is given by the realization y induced by the decision profile $\big((x_h, t_h)_{h=0,\ldots,k-1}, \Delta\big)$.
S2. The decision x_k is held constant on $(t_k, t_{k+1}]$.
S3. The process is then repeated for $k + 1$.

For every $x \in \mathcal{X}$, we let $\Xi(x)$ be the set of decision profiles of the form $\xi = ((t_k, x_k)_{k\in\{0,\ldots,n\}}, \Delta)$ with $n \in \mathbb{N}$ and satisfying $t_{k+1} - t_k = \Delta$ for all $k = 0, \ldots, n-1$ and $x_n = x$. Moreover, we let $\Xi := \Xi(\mathcal{X})$.

According to S1, every new decision is taken upon evaluation of the performance of the previously applied decisions. Ideally, one would evaluate the effect the past decisions had on the controlled process z. An *evaluation model* is a scheme serving such purpose. In this article, it formally consists of a tuple $(\mathcal{O}, O, \omega^\uparrow, \omega^\downarrow)$ in which $\mathcal{O}$ is a topological space, $O \subset \mathcal{O}$ is an open set, and $\omega^\uparrow, \omega^\downarrow : \mathcal{Z}^{\subset\mathbb{R}} \to \mathcal{O}$ are functions designed so that $\omega^\uparrow(z) \in O$ when z is characterized by an "excessively unstable" behavior while $\omega^\downarrow(z) \in O$, when z is characterized by an "overly stable" behavior (see Section 3 for specific choices in the COVID-19 case). Hence, O^{c} represents a compromise region such that if $\omega^\uparrow(z), \omega^\downarrow(z) \in O^{\mathrm{c}}$ then z is neither overly unstable nor stable. In this case, the behavior of z is considered satisfactory.

In our setting, however, we do not measure z. Hence, we cannot directly check whether $\omega^\downarrow(z)$ or $\omega^\uparrow(z)$ are in O. An *inference model* is an equivalent notion applicable to the measured signal y. In particular, it is a tuple $(\mathcal{A}, A, \alpha^\uparrow, \alpha^\downarrow)$ in which $\mathcal{A}$ is a topological space, $A \subset \mathcal{A}$ is open, and $\alpha^\uparrow, \alpha^\downarrow : \mathcal{Y}^{\subset\mathbb{R}} \to \mathcal{A}$ are designed so that, similarly to evaluation models, $\alpha^\uparrow(y) \in A$ when y is considered "excessively unstable", and $\alpha^\downarrow(y) \in A$ when y is considered "excessively stable". Thus, if $\alpha^\uparrow(y)$ and $\alpha^\downarrow(y)$ are both in A^{c}, the behavior of y is considered acceptable.

For this design principle to be well-posed, we make the following assumption, implying that the two conditions $\alpha^\uparrow(y) \in A$ and $\alpha^\downarrow(y) \in A$ are mutually exclusive.

Assumption 1 (*Consistency*)**.** For every decision profile $\xi \in \Xi$, the following implications hold

$$\alpha^\uparrow(y(\cdot|\xi)) \in \overline{A} \implies \alpha^\downarrow(y(\cdot|\xi)) \notin A,$$
$$\alpha^\downarrow(y(\cdot|\xi)) \in \overline{A} \implies \alpha^\downarrow(y(\cdot|\xi)) \notin A.$$

In general, the inference model has to be designed so that if we are able to choose x guaranteeing $\alpha^\uparrow(y), \alpha^\downarrow(y) \in A^{\mathrm{c}}$, then such a decision would be also associated with a satisfactory behavior for z in the sense mentioned earlier. This is, of course, possible only under certain conditions linking the underlying process and the available measurements. In our setting, these conditions are given by a detectability property described in the assumption below.

Assumption 2 (*Robust Detectability*)**.** For every decision profile $\xi \in \Xi$, the following implications hold

$$\begin{aligned} &\omega^\uparrow(z(\cdot|\xi)) \in O \implies \alpha^\uparrow(y(\cdot|\xi)) \in A, \\ &\omega^\downarrow(z(\cdot|\xi)) \in O \implies \alpha^\downarrow(y(\cdot|\xi)) \in A. \end{aligned} \tag{3}$$

Remark 1. In qualitative terms, the ability of designing inference models so as to satisfy Assumption 2 with respect to some desirable evaluation model depends on the available knowledge on the plant and on the quality of the measurements. In turn, it is here that the prior knowledge about delay, noise, parameters, and structural properties of the process come into play, and it is here that the praised robustness of hysteresis-based control originates: all the uncertainties/disturbances that do not destroy robust detectability do not affect the regulator performances (this notion of robustness, in turn, may be better framed within the more general notion of robustness in the broader context of output regulation, see e.g. Bin, Astolfi, Marconi, and Praly (2018)).

In the forthcoming Sections 2.3–2.7, we study a class of hysteresis-based control schemes seeking online a decision x guaranteeing that both $\alpha^\uparrow(y), \alpha^\downarrow(y) \notin A$. We carry out the analysis without any reference to z, thus extracting a set of convergence properties depending only on the measurements. Later, in Section 2.8, we show that robust detectability permits to extend some key guarantees to the unmeasured process z. In turn, robust detectability will not be assumed until there. The problem of designing inference models to fulfill robust detectability in the case of epidemics control is further studied in Section 3, and the results are applied to the COVID-19 case in Section 4.

2.3. The hysteresis control logic

Given an inference model $(\mathcal{A}, A, \alpha^\uparrow, \alpha^\downarrow)$, the class of hysteresis controllers we study in this article is given as follows. An initial decision x_0 is arbitrarily taken at time t_0. For every $k \in \mathbb{N}$, we denote by

$$\xi_{\leq k} := \big((t_h, x_h)_{h\in\operatorname{dom}\xi_{\leq k}}, \Delta\big) \in \Xi, \quad \operatorname{dom}\xi_{\leq k} := \mathbb{N}_{\leq k}$$

the decision profile collecting the past decisions and decision times. Then, the decision x_{k+1} is taken at time $t_{k+1} = t_k + \Delta$ according to the following inclusion

$$x_{k+1} \in F\big(x_k, y(\cdot|\xi_{\leq k})\big), \tag{4}$$

in which, for every $x \in \mathcal{X}$ and $y : \operatorname{dom} y \subset \mathbb{R} \to \mathcal{Y}$, we let

$$F(x, y) := \begin{cases} x^- & \text{if } \alpha^\uparrow(y) \in A, \\ x^+ & \text{if } \alpha^\downarrow(y) \in A, \\ \{x^-, x\} & \text{if } \alpha^\uparrow(y) \in \partial A \text{ and } \alpha^\downarrow(y) \notin \partial A, \\ \{x^+, x\} & \text{if } \alpha^\downarrow(y) \in \partial A \text{ and } \alpha^\uparrow(y) \notin \partial A, \\ \{x^-, x, x^+\} & \text{if } \alpha^\uparrow(y) \in \partial A \text{ and } \alpha^\downarrow(y) \in \partial A, \\ x & \text{otherwise.} \end{cases} \tag{5}$$

A decision profile $\xi = ((t_k, x_k)_{k\in\operatorname{dom}\xi}, t)$ for which (4) holds is called a *solution* to (4).

Remark 2 (*Wellposedness*)**.** For $\alpha^\uparrow(y) \in \partial A$ and/or $\alpha^\downarrow(y) \in \partial A$, $F(x, y)$ is a set. This is somewhat different from canonical hysteresis controllers in which, instead of F, one typically employs a selection $\widetilde{F}$ of F satisfying $\widetilde{F}(x, y) = x$ if $\alpha^\uparrow(y) \in \partial A$ and/or $\alpha^\downarrow(y) \in \partial A$. This modification has been introduced to make Eq. (4) *wellposed* in the same sense of Goebel, Sanfelice, and Teel (2012, Chapter 6), ensuring that limits of "converging sequences" of solutions are solutions as well.[1]

Remark 3. Nothing in (4) prevents one to always chose $x_{k+1} = x_k$ if $\alpha^\uparrow(y) \in \partial A$ and/or $\alpha^\downarrow(y) \in \partial A$. In fact, this choice corresponds to a solution of (4) and, as such, it is feasible. In other terms, (4)–(5) can be seen as a "robust version" of the control logic that uses the function $\widetilde{F}$ defined in Remark 2: multiple choices for the cases in which $\alpha^\uparrow(y)$

[1] More precisely, if we endow the set of signals $\operatorname{dom} y \subset \mathbb{R} \to \mathcal{Y}$ with a topology for which $\alpha^\downarrow$ and $\alpha^\downarrow$ are continuous, then by using $\widetilde{F}$, we may have cases in which a convergent net $(y_\ell)_\ell$ exists such that, for instance, $\alpha^\uparrow(y_\ell) \in A$ for all ℓ, but its limit y^* satisfies $\alpha^\uparrow(y^*) \in \partial A$. In this case, $\widetilde{F}(x, y_\ell) = x^-$ for all ℓ, so that $\lim \widetilde{F}(x, y_\ell) = x^- \neq x = \widetilde{F}(x, \lim y_\ell)$, which in turn implies that y^* does not produce a solution. Instead, by using F, we always have $\lim F(x, y_\ell) \subset F(x, \lim y_\ell)$.

and/or $\alpha^{\downarrow}(y)$ are in ∂A can model a decision logic that cannot determine the conditions $\alpha^{\downarrow}(y), \alpha^{\uparrow}(y) \in \partial A$ with arbitrary accuracy and, thus, may mistake decisions when $\alpha^{\uparrow}(y)$ and $\alpha^{\downarrow}(y)$ too close to ∂A.

2.4. The stationarity and monotonicity assumptions

We analyze System (4) under the following main "open-loop" assumptions.

Assumption 3 (*Stationarity*). For every $x \in \mathcal{X}$ and every two decision profiles $\xi^1, \xi^2 \in \Xi(x)$, the following hold

$$\alpha^{\uparrow}\big(y(\cdot|\xi^1)\big) \in A \iff \alpha^{\uparrow}\big(y(\cdot|\xi^2)\big) \in A,$$
$$\alpha^{\downarrow}\big(y(\cdot|\xi^1)\big) \in A \iff \alpha^{\downarrow}\big(y(\cdot|\xi^2)\big) \in A.$$

Assumption 3 is better understood when $\alpha(y(\cdot|\xi))$ depends only on the last decision, as in the application considered in Section 3. In this case, indeed, it requires that the same decision must lead to the same qualitative behavior of y (although, we stress, it does not require that $\alpha^{\uparrow}(y(\cdot|\xi^1)) = \alpha^{\uparrow}(y(\cdot|\xi^2))$ or $\alpha^{\downarrow}(y(\cdot|\xi^1)) = \alpha^{\downarrow}(y(\cdot|\xi^2))$).

Assumption 4 (*Monotonicity*). For every $x \in \mathcal{X}$ and every $\xi \in \Xi(x)$, $\xi^+ \in \Xi(x^+)$, $\xi^- \in \Xi(x^-)$, the following hold

$$\alpha^{\uparrow}\big(y(\cdot|\xi)\big) \in A \implies \alpha^{\uparrow}\big(y(\cdot|\xi^+)\big) \in A,$$
$$\alpha^{\downarrow}\big(y(\cdot|\xi)\big) \in A \implies \alpha^{\downarrow}\big(y(\cdot|\xi^-)\big) \in A.$$

Assumption 4 is the essence of the functioning of the hysteresis logic (4), and it is what ultimately motivates the choice of F in (4). With Assumption 3, it implies that there is a preferred direction in $\mathcal{X}$ to bring y out of the set in which $\alpha^{\uparrow}(y) \in A$, and another one to exit that in which $\alpha^{\downarrow}(y) \in A$. This assumption is what permits an hysteresis logic to avoid exhaustive explorations over "open-loop" or predefined paths, typical of universal controllers (Mårtensson, 1985; Minyue Fu & Barmish, 1986), and instead to tell the exploration direction from closed-loop measurements. In the context of epidemic control, and in particular in the context of Bin, Cheung et al. (2021) described in Section 1.1, Assumption 4 is justified by the result of Bin, Cheung et al. (2021, Theorem 1) which links increasing (resp. decreasing) duty cycles with increasing (resp. decreasing) values of the reproductive number.

2.5. Target sets of decisions

Under Assumption 3, we can define the following sets without ambiguity

$$\widetilde{X}^{\uparrow} := \Big\{ x \in \mathcal{X} \,:\, \alpha^{\uparrow}(y(\cdot|\xi)) \notin A,\ \forall \xi \in \Xi(x) \Big\},$$
$$\widetilde{X}^{\downarrow} := \Big\{ x \in \mathcal{X} \,:\, \alpha^{\downarrow}(y(\cdot|\xi)) \notin A,\ \forall \xi \in \Xi(x) \Big\},$$

and their intersection

$$X^{\star} := \widetilde{X}^{\uparrow} \cap \widetilde{X}^{\downarrow}.$$

The sets $\widetilde{X}^{\uparrow}$ and $\widetilde{X}^{\downarrow}$ contain, respectively, the decisions leading to a behavior of y which is not too unstable and not too stable according to the chosen inference model. Thus, the set $X^{\star}$ contains decisions for which y behaves satisfactorily, and it is called the *"target set"*. Under consistency and stationarity, at least one among $\widetilde{X}^{\uparrow}$ and $\widetilde{X}^{\downarrow}$ is always nonempty, as established by the following lemma.

Lemma 1 (*Non-emptiness of $X^{\uparrow} \cup X^{\downarrow}$*). *Suppose that Assumptions 1 and 3 hold. Then, $\widetilde{X}^{\uparrow} \cup \widetilde{X}^{\downarrow} \neq \emptyset$.*

Proof. If $\widetilde{X}^{\uparrow} \neq \emptyset$ there is nothing to prove. If, instead, $\widetilde{X}^{\uparrow} = \emptyset$, then for all $x \in \mathcal{X}$, there exists $\xi \in \Xi(x)$, such that $\alpha^{\uparrow}(y(\cdot|\xi)) \in A$. By Assumption 1, this implies $\alpha^{\downarrow}(y(\cdot|\xi)) \notin A$. Assumption 3, then implies that $\alpha^{\downarrow}(y(\cdot|\xi')) \notin A$ holds for all $\xi' \in \Xi(x)$. Hence $\widetilde{X}^{\downarrow} \neq \emptyset$. □

In general, however, the target set $X^{\star}$ may be empty even if both $\widetilde{X}^{\uparrow}$ and $\widetilde{X}^{\downarrow}$ are not. Moreover, as we shall clarify later in Proposition 1, $X^{\star}$ may fail to be invariant. This motivates us to study also a relaxation of $\widetilde{X}^{\uparrow}$ and $\widetilde{X}^{\downarrow}$, consisting of their one-point dilations

$$\widetilde{X}^{\uparrow}_{+} := \widetilde{X}^{\uparrow} \cup \Big\{ x \in \mathcal{X} \,:\, x = \bar{x}^{+},\ \bar{x} \in \widetilde{X}^{\uparrow} \Big\},$$
$$\widetilde{X}^{\downarrow}_{-} := \widetilde{X}^{\downarrow} \cup \Big\{ x \in \mathcal{X} \,:\, x = \bar{x}^{-},\ \bar{x} \in \widetilde{X}^{\downarrow} \Big\}.$$

and their intersection

$$\widetilde{X}_{\pm} := \widetilde{X}^{\uparrow}_{+} \cap \widetilde{X}^{\downarrow}_{-}$$

that always satisfies $\widetilde{X}_{\pm} \supset X^{\star}$. The set $\widetilde{X}_{\pm}$ is slightly larger than $X^{\star}$ but, as detailed below, in general it enjoys stronger properties, as it is nonempty under weaker conditions (Lemma 3), and it is always forward invariant (Proposition 3). The sets $\widetilde{X}^{\uparrow}$, $\widetilde{X}^{\downarrow}$, $\widetilde{X}^{\uparrow}_{+}$ and $\widetilde{X}^{\downarrow}_{-}$ satisfy the following closure properties with respect to the predecessor and successor operators.

Lemma 2. *Under Assumptions 1, 3 and 4:*

$$x \in \widetilde{X}^{\uparrow} \implies x^{-} \in \widetilde{X}^{\uparrow}, \qquad x \in \widetilde{X}^{\uparrow}_{+} \implies x^{-} \in \widetilde{X}^{\uparrow}_{+},$$
$$x \in \widetilde{X}^{\downarrow} \implies x^{+} \in \widetilde{X}^{\downarrow}, \qquad x \in \widetilde{X}^{\downarrow}_{-} \implies x^{+} \in \widetilde{X}^{\downarrow}_{-}.$$

The proof of Lemma 2 is in the Appendix. Clearly, Lemma 1 implies $\widetilde{X}^{\uparrow}_{+} \cup \widetilde{X}^{\downarrow}_{-} \neq \emptyset$ under consistency and stationarity. Moreover, if also monotonicity holds, we can conclude that, unlike $X^{\star}$, $\widetilde{X}_{\pm}$ is always nonempty if so are $\widetilde{X}^{\uparrow}$ and $\widetilde{X}^{\downarrow}$. This is established by the following lemma, proved in the Appendix.

Lemma 3. *Suppose that Assumptions 1, 3 and 4 hold, and that $\widetilde{X}^{\uparrow}$ and $\widetilde{X}^{\downarrow}$ are nonempty. Then, $\widetilde{X}_{\pm} \neq \emptyset$ and $\widetilde{X}^{\uparrow}_{+} \cup \widetilde{X}^{\downarrow}_{-} = \mathcal{X}$.*

2.6. Forward invariance

In this section, we study the forward invariance properties of the decision sets $X^{\star}$, $\widetilde{X}^{\uparrow}_{+}$, $\widetilde{X}^{\downarrow}_{-}$ and $\widetilde{X}_{\pm}$. In particular, a set X is said to be *weakly forward invariant* for (4) if, for every $x \in X$ and every decision profile $\xi \in \Xi(x)$, it holds that $F(x, y(\cdot|\xi)) \cap X \neq \emptyset$. Moreover, X is said to be *forward invariant* for (4) if, for every $x \in X$ and every decision profile $\xi \in \Xi(x)$, it holds that $F(x, y(\cdot|\xi)) \subset X$.

In general, even if nonempty, the target set $X^{\star}$ may fail to be forward invariant. This is due to F being set-valued for $\alpha^{\uparrow}(y) \in \partial A$ and/or $\alpha^{\downarrow}(y) \in \partial A$, and due to the lack of an "ordering" on $\mathcal{X}$ of the kind established by Assumption 4 when $\alpha^{\uparrow}(y) \in A$ or $\alpha^{\downarrow}(y) \in A$. Nevertheless, the target set $X^{\star}$ is always weakly forward invariant whenever nonempty, as established by the following proposition.

Proposition 1 (*Weak forward invariance of $X^{\star}$*). *Suppose that Assumptions 1, 3 and 4 hold and that $X^{\star} \neq \emptyset$. Then, $X^{\star}$ is weakly forward invariant.*

Proof. Pick $x \in X^{\star}$ and $\xi \in \Xi(x)$ arbitrarily. Then $\alpha^{\uparrow}(y(\cdot|\xi)) \notin A$ and $\alpha^{\downarrow}(y(\cdot|\xi)) \notin A$. In view of (5), this implies $x \in F(x, y(\cdot|\xi))$, so that $X^{\star} \cap F(x, y(\cdot|\xi)) \neq \emptyset$. □

While weaker than invariance, Proposition 1 guarantees that, if $x_k \in X^{\star}$, then we can always choose $x_{k+1} \in X^{\star}$ according to (4). Under some additional conditions on the set of decisions for which $\alpha^{\uparrow}(y) \notin A$ and $\alpha^{\downarrow}(y) \notin A$, forward invariance is recovered.

Proposition 2 (*Forward Invariance of $X^{\star}$*). *In addition to the assumptions of Proposition 1, suppose that*

(a) The cardinality of $X^{\star}$ is larger or equal than 2.
(b) The set of $x \in \mathcal{X}$ such that $\alpha^{\uparrow}(y(\cdot|\xi)) \in \partial A$ and $\alpha^{\downarrow}(y(\cdot|\xi)) \in \partial A$ hold at the same time for some $\xi \in \Xi(x)$ is empty.

(c) *If* $x \in X^\star$ *is such that* $\alpha^\uparrow(y(\cdot|\xi)) \in \partial A$ *for some* $\xi \in \Xi(x)$, *then* x^+ *satisfies* $\alpha^\uparrow(y(\cdot|\xi)) \in A$ *for all* $\xi \in \Xi(x^+)$.

(d) *If* $x \in X^\star$ *is such that* $\alpha^\downarrow(y(\cdot|\xi)) \in \partial A$ *for some* $\xi \in \Xi(x)$, *then* x^- *satisfies* $\alpha^\downarrow(y(\cdot|\xi)) \in A$ *for all* $\xi \in \Xi(x^-)$.

Then, $X^\star$ *is forward invariant.*

Proposition 2 is proved in the Appendix. Unlike $X^\star$, which in general is "only" weakly forward invariant, the sets $\widetilde{X}^\uparrow_+$, $\widetilde{X}^\downarrow_-$ and $\widetilde{X}_\pm$ are always forward invariant whenever nonempty. This is established by the following proposition, proved in the Appendix.

Proposition 3 (*Forward Invariance*). *Suppose that Assumptions 1, 3 and 4 hold. Then each of the sets* $\widetilde{X}^\uparrow_+$, $\widetilde{X}^\downarrow_-$ *and* $\widetilde{X}_\pm$ *is invariant whenever nonempty.*

2.7. Convergence analysis

In this section, we study attractiveness of the decision sets $\widetilde{X}^\uparrow_+$, $\widetilde{X}^\downarrow_-$, $\widetilde{X}_\pm$ and $X^\star$. In particular, a set $X \subset \mathcal{X}$ is said to be *uniformly attractive* for (4) from another set $X_0 \subset \mathcal{X}$ if there exists $h \in \mathbb{N}$ such that every solution ξ to (4) with $x_0 \in X_0$ and $h \in \operatorname{dom}\xi$ satisfies $x_k \in X$ for all $k \in (\operatorname{dom}\xi)_{\geq h}$ (recall that, being the topology on $\mathcal{X}$ discrete, convergence in $\mathcal{X}$ is finite-time convergence). If $X_0 = \mathcal{X}$, then X is said to be *globally uniformly attractive*.

Given any $X \subsetneq \mathcal{X}$, and any $x \in \mathcal{X}$, we define the *distance of* x *to* X as

$$d(x, X) := \min\left\{n \in \mathbb{N} \,:\, x^{n+} \in X \text{ or } x^{n-} \in X\right\} \in \mathbb{N},$$

in which the operator $x^{\cdot+}$ is defined by the recursion $x^{0+} = x$ and $x^{n+} = (x^{(n-1)+})^+$ for $n \in \mathbb{N}_{>0}$, and $x^{\cdot-}$ is defined similarly. Then, uniform attractiveness from X_0 is equivalent to the existence of $h \in \mathbb{N}$ such that for every solution ξ to (4) with $x_0 \in X_0$ and $h \in \operatorname{dom}\xi$, $d(x_k, X) = 0$ for all $k \in (\operatorname{dom}\xi)_{\geq h}$.

The following result establishes uniform global attractiveness of the decision sets $\widetilde{X}^\uparrow_+$, $\widetilde{X}^\downarrow_-$ and $\widetilde{X}_\pm$.

Theorem 1 (*Attractiveness of* $\widetilde{X}$). *Suppose that Assumptions 1, 3 and 4 hold, and consider the set*

$$\widetilde{X} := \begin{cases} \widetilde{X}^\uparrow_+ & \text{if } \widetilde{X}^\uparrow_+ \neq \emptyset \text{ and } \widetilde{X}^\downarrow_- = \emptyset, \\ \widetilde{X}^\downarrow_- & \text{if } \widetilde{X}^\downarrow_- \neq \emptyset \text{ and } \widetilde{X}^\uparrow_+ = \emptyset, \\ \widetilde{X}_\pm & \text{if } \widetilde{X}^\uparrow_+ \cap \widetilde{X}^\downarrow_- \neq \emptyset. \end{cases}$$

Then, for every solution ξ *to* (4), *and every* $k \in \operatorname{dom}\xi$ *such that* $k+1 \in \operatorname{dom}\xi$, *the following implication holds*

$$x_k \notin \widetilde{X} \implies d(x_{k+1}, \widetilde{X}) = d(x_k, \widetilde{X}) - 1. \tag{6}$$

Thus, in particular, $\widetilde{X}$ *is uniformly globally attractive for* (4).

Proof. We first prove (6) for $\widetilde{X} = \widetilde{X}^\uparrow_+$. Assume that $\widetilde{X}^\uparrow_+ \neq \emptyset$ and that $\widetilde{X}^\uparrow_+ \subsetneq \mathcal{X}$ (otherwise the claim trivially holds). Pick a solution ξ to (4) and $k \in \operatorname{dom}\xi$ such that $k+1 \in \operatorname{dom}\xi$, and suppose that $x_k \notin \widetilde{X}^\uparrow_+$. By Lemma 2, necessarily $x_k > \sup \widetilde{X}^\uparrow_+$. In view of (4), Lemma 2 and Assumptions 3 and 4, this implies $x_{k+1} = x_k^-$ and, hence, $d(x_{k+1}, \widetilde{X}^\uparrow_+) = d(x_k, \widetilde{X}^\uparrow_+) - 1$, which is (6). The proof of (6) for $\widetilde{X} = \widetilde{X}^\downarrow_-$ follows the same argument. The proof of (6) for $\widetilde{X} = \widetilde{X}_\pm$ follows by noticing that, since in view of Lemma 3, $\widetilde{X}^\uparrow_\pm \cup \widetilde{X}^\downarrow_- = \mathcal{X}$, then if $x_k \notin \widetilde{X}^\uparrow_+ \cap \widetilde{X}^\downarrow_-$ either **(i)** $x_k \notin \widetilde{X}^\uparrow_+ \cap (\widetilde{X}^\downarrow_-)^c$, or **(ii)** $x_k \notin \widetilde{X}^\downarrow_- \cap (\widetilde{X}^\uparrow_+)^c$. In case **(i)**, $d(x_k, \widetilde{X}) = d(x_k, \widetilde{X}^\uparrow_+)$. In case **(ii)** $d(x_k, \widetilde{X}) = d(x_k, \widetilde{X}^\downarrow_-)$. Hence, (6) holds on both cases.

Regarding global attractiveness of $\widetilde{X}$, as $\widetilde{X}$ is always forward invariant (Proposition 3), it suffices to show that there exists $h \in \mathbb{N}$ such that, for every solution of (4) with $h \in \operatorname{dom}\xi$, we have $x_h \in \widetilde{X}$. This, however, is a direct consequence of (6) and of the finiteness of $\mathcal{X}$. □

Remark 4. Theorem 1 implies that decisions always converge in finite time to the set $\widetilde{X}_\pm$ whenever nonempty. Moreover, since the topology of $\mathcal{X}$ is discrete, (Lyapunov) *stability* is always trivially implied by forward invariance, since each singleton is a neighborhood of its element. Nevertheless, Theorem 1 claims a stronger result, given by (6), and guaranteeing that the distance to $\widetilde{X}$ is always decreasing. This, in turn, permits to directly extend the uniform attractiveness result to the case in which $\mathcal{X}$ is countable.

When $X^\star$ is nonempty, one is most interested in solutions that reach and stay in it, rather than $\widetilde{X}_\pm$. As in general $X^\star$ is not forward invariant, we cannot conclude attractiveness. Yet, this is just a technical obstacle, and we can prove several properties of $X^\star$ which in practice have the same implications that Theorem 1 has for $\widetilde{X}_\pm$. In particular:

1. $X^\star$ is always reached in finite time from every initial condition. A solution may jump outside $X^\star$, but if it does, it enters $X^\star$ again within the next update.
2. From every initial condition there always exists a solution reaching $X^\star$ in finite time and staying in it for all successive times.

These properties are stated in the following theorem.

Theorem 2 (*Attractiveness of* $X^\star$). *Suppose that Assumptions 1, 3 and 4 hold, and assume that* $X^\star \neq \emptyset$. *Then, the following hold:*

(a) *For every solution* ξ *to* (4), *and every* $k \in \operatorname{dom}\xi$ *such that* $k+1 \in \operatorname{dom}\xi$, *if* $x_k \in \widetilde{X}_\pm \setminus X^\star$ *then* $x_{k+1} \in X^\star$.

(b) *From every initial condition* x_0, *there exists a solution* ξ *and an* $h \in \mathbb{N}$ *such that either* $\sup \operatorname{dom}\xi < h$ *or* $x_k \in X^\star$ *for all* $k \in (\operatorname{dom}\xi)_{\geq h}$.

(c) *For every solution* ξ *to* (4), *define* $D := \{k \in \operatorname{dom}\xi \,:\, x_k \in X^\star\}$ *and, for every* $k \in D$, *let* $\delta(k) = \min D_{>k} - k$. *Then, there exists* $h \in \mathbb{N}$ *such that either* $\sup \operatorname{dom}\xi < h$ *or* $\delta(k) \leq 1$ *for all* $k \in (\operatorname{dom}\xi)_{\geq h}$.

Proof. To prove (a), pick a solution ξ to (4) and a $k \in \operatorname{dom}\xi$ such that $k+1 \in \operatorname{dom}\xi$ and $x_k \in \widetilde{X}_\pm \setminus X^\star$. Then, either **(i)** $x_k \notin \widetilde{X}^\uparrow$, or **(ii)** $x_k \notin \widetilde{X}^\downarrow$. In case **(i)**, in view of Assumption 3 we have $\alpha^\uparrow(y(\cdot|_{\xi_{\leq k}})) \in A$. Hence, (4) implies $x_{k+1} = x_k^-$. Since $x_k \in \widetilde{X}_\pm$ implies $x_k \in \widetilde{X}^\uparrow_+$, then by definition $x_k^- \in \widetilde{X}^\uparrow \subset X^\star$, so as $x_{k+1} \in X^\star$. Case **(ii)** is proved in the same way.

Claim (b) follows by noticing that Theorem 1 implies the existence of $h_0 \in \mathbb{N}$ such that $h_0 \in \operatorname{dom}\xi \implies x_k \in \widetilde{X}_\pm$ for all $k \in (\operatorname{dom}\xi)_{\geq h_0}$. Then, either $x_{h_0} \in X^\star$, in which case the claim holds with $h = h_0$ by weak forward invariance of $X^\star$ (Proposition 1), or Claim (a) guarantees that $h_0 + 1 \in \operatorname{dom}\xi \implies x_{h_0+1} \in X^\star$, so as the claim holds with $h := h_0 + 1$.

Claim (c) follows directly from Claims (a) and (b). □

2.8. Connections with the process z

In all the precedent analysis, robust detectability (Assumption 2) was never assumed. All the claims, indeed, referred to the measured output y. In this section we show that, if the evaluation model $(\mathcal{O}, O, \omega^\uparrow, \omega^\downarrow)$ satisfies a stationarity property analogous to that required by Assumption 3 for $(\mathcal{A}, A, \alpha^\uparrow, \alpha^\downarrow)$, and if, in addition, robust detectability holds, then we can conclude convergence to a set of decision corresponding to a satisfactory behavior of the process variable z.

In particular, we assume the following.

Assumption 5 (*z-Stationarity*). For every $x \in \mathcal{X}$ and every two decision profiles $\xi^1, \xi^2 \in \Xi(x)$, the following hold

$$\omega^\uparrow\left(z(\cdot|\xi^1)\right) \in O \iff \omega^\uparrow\left(z(\cdot|\xi^2)\right) \in O,$$
$$\omega^\downarrow\left(z(\cdot|\xi^1)\right) \in O \iff \omega^\downarrow\left(z(\cdot|\xi^2)\right) \in O.$$

Under Assumption 5, we can define the sets

$$\widetilde{Z}^{\uparrow} := \left\{ x \in \mathcal{X} \ : \ \omega^{\uparrow}(z(\cdot|\xi)) \notin O, \forall \xi \in \Xi(x) \right\},$$
$$\widetilde{Z}^{\downarrow} := \left\{ x \in \mathcal{X} \ : \ \omega^{\downarrow}(z(\cdot|\xi)) \notin O, \forall \xi \in \Xi(x) \right\},$$
$$\widetilde{Z}^{\uparrow}_{+} := \widetilde{Z}^{\uparrow} \cup \left\{ x \in \mathcal{X} \ : \ x = \bar{x}^{+}, \ \bar{x} \in \widetilde{Z}^{\uparrow} \right\},$$
$$\widetilde{Z}^{\downarrow}_{-} := \widetilde{Z}^{\downarrow} \cup \left\{ x \in \mathcal{X} \ : \ x = \bar{x}^{-}, \ \bar{x} \in \widetilde{Z}^{\downarrow} \right\},$$
$$Z^{\star} := \widetilde{Z}^{\uparrow} \cap \widetilde{Z}^{\downarrow}, \qquad \widetilde{Z}_{\pm} := \widetilde{Z}^{\uparrow}_{+} \cap \widetilde{Z}^{\downarrow}_{-},$$

which are the analogous of $\widetilde{X}^{\uparrow}$, $\widetilde{X}^{\downarrow}$, $\widetilde{X}^{\uparrow}_{+}$, $\widetilde{X}^{\downarrow}_{-}$, $X^{\star}$ and $\widetilde{X}_{\pm}$ with z in place of y.

Under robust detectability, the following result establishes a set of relationships among the decision sets defined earlier on. The importance of this result stems from the fact that it implies that, if robust detectability holds, then the sets $X^{\star}$ and its relaxation $\widetilde{X}_{\pm}$ to which the solutions of (4) converge are subsets of the sets $Z^{\star}$ and $\widetilde{Z}_{\pm}$ respectively, which are the sets of decisions making the unmeasured process z behave satisfactorily. This, in turn, permits to infer a satisfactory behavior of the underlying process z from the decision based only on y.

Theorem 3. *Suppose that Assumptions 1, 2 and 5. Then* $\widetilde{X}^{\uparrow} \subset \widetilde{Z}^{\uparrow}$, $\widetilde{X}^{\downarrow} \subset \widetilde{Z}^{\downarrow}$, $\widetilde{Z}^{\uparrow}_{+} \subset \widetilde{Z}^{\uparrow}_{+}$, $\widetilde{X}^{\downarrow}_{-} \subset \widetilde{Z}^{\downarrow}_{-}$, $X^{\star} \subset Z^{\star}$ *and* $\widetilde{X}_{\pm} \subset \widetilde{Z}_{\pm}$.

Proof. To see that $\widetilde{X}^{\uparrow} \subset \widetilde{Z}^{\uparrow}$ holds, pick $x \in \mathcal{X} \setminus \widetilde{Z}^{\uparrow}$. Then, there exists $\xi \in \Xi(x)$ such that $\omega^{\uparrow}(z(\cdot|\xi)) \in O$. Assumption 5 implies that $\omega^{\uparrow}(z(\cdot|\xi)) \in O$ for all $\xi \in \Xi(x)$. Assumption 2, in turn, implies $\alpha^{\uparrow}(y(\cdot|\xi)) \in O$ for all $\xi \in \Xi(x)$. Namely, $x \in \mathcal{X} \setminus \widetilde{X}^{\uparrow}$. For arbitrariness of x, this shows that $\mathcal{X} \setminus \widetilde{Z}^{\uparrow} \subset \mathcal{X} \setminus \widetilde{X}^{\uparrow}$, which in turn implies $\widetilde{X}^{\uparrow} \subset \widetilde{Z}^{\uparrow}$. The inclusion $\widetilde{X}^{\downarrow} \subset \widetilde{Z}^{\downarrow}$ is proved in the same way, while all the others follow directly from these. □

3. Application to epidemics control

3.1. The setting

In the remainder of the paper, we focus on the application of the theory developed in the previous sections to the problem of non-pharmaceutical control of epidemics. As anticipated in Section 1.1, we build on the result of Bin, Cheung et al. (2021). In particular, we consider a FPSP alternating $N^{\downarrow}$ days of lockdown and $N^{\uparrow}$ normal days, where society works as normal (modulo additional policies, such as social distancing and use of masks).

We define the FPSP period P and duty cycle ϱ as

$$P := N^{\downarrow} + N^{\uparrow}, \quad \varrho := N^{\uparrow}/P.$$

According to Bin, Cheung et al. (2021), the period is decided in advance as the smallest possible period compatible with societal constraints (typically one, two or three weeks) and kept fixed. The duty cycle, instead, is the variable we control. In particular, the decision space is taken as $\mathcal{X} := \{0, \dots, P\}$ with the natural ordering. Each decision $x \in \mathcal{X}$ then determines a duty cycle $\varrho(x) = x/P$.

The unmeasured process z represents the actual dynamics of the epidemic. For example, although we stress this is not necessary, z may be thought of in terms of the solution map of an "SI-like" differential equation of the form

$$\begin{aligned} \dot{S} &= f_S(\beta, S, I, C), \\ \dot{I} &= f_I(\beta, S, I, C), \\ \dot{C} &= f_C(\beta, S, I, C), \end{aligned} \tag{7}$$

in which $S(t) \in \mathbb{R}_{\geq 0}$ represents the number/percentage of susceptible individuals, $I(t) \in \mathbb{R}_{\geq 0}$ the number/percentage of infected individuals (or a portion of them, depending on the actual model (Avram, Adenane, & Ketcheson, 2021; Giordano et al., 2020)), $C(t) \in \mathbb{R}^{n_C}_{\geq 0}$, $n_C \in \mathbb{N}$, represents additional compartments, $\beta \in [0, 1]$ is a parameter modulating the infection rate, and f_S, f_I, f_C are continuous functions whose values determine the actual model dynamics. In this case, the effect of the decision x on the epidemic dynamics is modeled by letting β depend on the FPSP duty cycle $\varrho(x)$. In particular, if $\beta^{\uparrow}, \beta^{\downarrow} \in [0, 1]$ denote, respectively, the value of β during normal days and lockdown, then β in (7) is assumed to satisfy

$$\beta(t, x) = \begin{cases} \beta^{\downarrow} & \text{if } \operatorname{mod}(t, P) < (1 - \varrho(x))P, \\ \beta^{\uparrow} & \text{otherwise,} \end{cases} \tag{8}$$

in which $\operatorname{mod}(t, P) := t - \max\{s \in \mathbb{N} \ : \ sP \leq t\}$. See Bin, Cheung et al. (2021) for further details. In this case, $\mathcal{Z} = \mathbb{R}^{2+n_C}_{\geq 0}$ and, for each $x \in \mathcal{X}$, the process trajectory flow $\zeta[x]$ (Section 2.1) equals the flow of (7). The operator ψ (Section 2.1), instead, is defined to select the (combination of) variables which are measurable, such as detected cases, deaths, or hospitalized people (we refer to Section 4 for an example in the case of the COVID-19 outbreak). For simplicity, we assume $\mathcal{Y} = \mathbb{R}$, as this does not lead to any conceptual loss of generality.

3.2. Evaluation and inference models

We suppose that the initial condition z_0 is fixed (albeit unknown) and we make reference to the definitions (1)–(2), which for each decision profile ξ produce the time signals $z(\cdot|\xi)$ and $y(\cdot|\xi)$. We let γ be an operator extracting from $z(\cdot|\xi)$ the (unmeasured) variables $\theta(\cdot|\xi) := \gamma(z(\cdot|\xi))$ of interest whose growth we aim to control (γ is chosen so as $\operatorname{dom}\theta(\cdot|\xi) = \operatorname{dom} z(\cdot|\xi)$). For example, in the case described above in which the epidemic dynamics is given by (7), θ may consist in just the I variable, or a combination of I and C. For ease of exposition, we suppose that $\theta(t|\xi) \in \mathbb{R}$ for all t and ξ.

We fix a number $T \geq P$ (typically, T is a multiple of P) and we define two operators, Π and D, acting on time signals $\eta : \operatorname{dom}\eta \subset \mathbb{R} \to \mathbb{R}$ as

$$\Pi\eta(t) := \frac{1}{T} \int_{[t-T,t] \cap \operatorname{dom}\eta} \Big(\eta(s) - \eta(t - T) \Big) ds,$$
$$\mathrm{D}\eta(t) := \eta(t - T),$$

in which $\Pi\eta$ and $\mathrm{D}\eta$ are defined for all $t \in \mathbb{R}_{\geq 0}$ such that $t - T \in \operatorname{dom}\eta$. For each $t \in \operatorname{dom}\Pi\eta$, $\Pi\eta(t)$ then equals the average growth of η in the interval $[t - T, t]$.

With $o^{\uparrow}, o^{\downarrow}, \mu^{\uparrow}, \mu^{\downarrow} \geq 0$, we then consider a class of evaluation models obtained with $\mathcal{O} = \mathbb{R}$, $O = \mathbb{R}_{>0}$, and

$$\begin{aligned} \omega^{\uparrow}(z) &:= \Big[\Pi - o^{\uparrow}\mathrm{D}\Pi \Big] \gamma(z)(\sup\operatorname{dom} z) - \mu^{\uparrow}, \\ \omega^{\downarrow}(z) &:= \Big[\mathrm{D}\Pi - o^{\downarrow}\Pi \Big] \gamma(z)(\sup\operatorname{dom} z) - \mu^{\downarrow}. \end{aligned} \tag{9}$$

Bearing in mind (1) and (2), for every decision profile of the form $\xi = \big((t_k, x_k)_{k \in \operatorname{dom}\xi}, \Delta\big)$, the condition $\omega^{\uparrow}(z(\cdot|\xi)) \in O$ is equivalent to

$$\frac{1}{T} \int_{t-T}^{t} \Big(\theta(s|\xi) - \theta(t - T|\xi) \Big) ds > o^{\uparrow} \frac{1}{T} \int_{t-2T}^{t-T} \Big(\theta(s|\xi) - \theta(t - 2T|\xi) \Big) ds + \mu^{\uparrow},$$

in which $t = t_{\sup\operatorname{dom}\xi} + \Delta$. Likewise, Condition $\omega^{\uparrow}(z(\cdot|\xi)) \in O$ reads

$$o^{\downarrow} \frac{1}{T} \int_{t-T}^{t} \Big(\theta(s|\xi) - \theta(t - T|\xi) \Big) ds < \frac{1}{T} \int_{t-2T}^{t-T} \Big(\theta(s|\xi) - \theta(t - 2T|\xi) \Big) ds - \mu^{\downarrow}.$$

Therefore, the evaluation principle underlying the evaluation model $(\mathcal{O}, O, \omega^{\uparrow}, \omega^{\downarrow})$ considers a behavior of z to be "excessively unstable" if the average growth during the interval $[t - T, t]$ exceeds that during $[t-2T, t-T]$ by a factor of $o^{\uparrow}$ plus a fixed bias $\mu^{\uparrow}$. Likewise, it considers a behavior of z to be "overly stable" if the average growth during the interval $[t - T, t]$ is lower than that during $[t - 2T, t - T]$ by a factor of $1/o^{\downarrow}$ minus a fixed bias $\mu^{\downarrow}/o^{\downarrow}$.

The inference models we consider are of the same kind. In particular, we let $\mathcal{A} = \mathbb{R}$, $A = \mathbb{R}_{>0}$ and, with $a^\uparrow, a^\downarrow, \varepsilon^\uparrow, \varepsilon^\downarrow \geq 0$ design parameters, we let

$$\begin{aligned} \alpha^\uparrow(y) &:= \Big[\Pi - a^\uparrow \mathrm{D}\Pi\Big] y(\sup \mathrm{dom}\, y) - \varepsilon^\uparrow, \\ \alpha^\downarrow(y) &:= \Big[\mathrm{D}\Pi - a^\downarrow \Pi\Big] y(\sup \mathrm{dom}\, y) - \varepsilon^\downarrow, \end{aligned} \tag{10}$$

whose interpretation is the same as that given above for the evaluation model.

In the following, we say that $(o^\uparrow, o^\downarrow, \mu^\uparrow, \mu^\downarrow)$ (resp. $(a^\uparrow, a^\downarrow, \varepsilon^\uparrow, \varepsilon^\downarrow)$) *generates* the evaluation model $(\mathbb{R}, \mathbb{R}_{>0}, \omega^\uparrow, \omega^\downarrow)$ (resp. the inference model $(\mathbb{R}, \mathbb{R}_{>0}, \alpha^\uparrow, \alpha^\downarrow)$), in which $\omega^\uparrow, \omega^\downarrow$ (resp. $\alpha^\uparrow, \alpha^\downarrow$) are defined as in (9) (resp. (10)).

Remark 5. The proposed definitions for the evaluation and inference models suggest (although this is not necessary, in principle (Bin, Cheung et al., 2021)) to choose the decision period Δ in such a way that

$$\Delta \geq 2T,$$

which in turn implies $\Delta \geq 2P$, namely that decisions must be held for at least two periods of the FPSP. This, indeed, permits to evaluate the average growth of the signals in the same conditions, and it is assumed in the forthcoming Section 3.3.

Remark 6. We observe that if $\varepsilon^\uparrow = \varepsilon^\downarrow = 0$ (resp. $\mu^\uparrow = \mu^\downarrow = 0$), then the conditions $\alpha^\uparrow(y) \in A$ and $\alpha^\downarrow(y) \in A$ (resp. $\omega^\uparrow(z) \in O$ and $\omega^\downarrow(z) \in O$) are *"scale-independent"* (Hespanha & Morse, 1999). Indeed, they are conditions only on the growth rate of y (resp. θ) not depending on its actual amplitude. This, in turn, is associated with a larger domain of validity of Assumptions 3 and 4.

Remark 7. If the acquisition of y is delayed by a fixed, known delay $\delta > 0$, then all what said in the remainder of the section still holds if (10) substituted by

$$\begin{aligned} \alpha^\uparrow(y) &:= \Big[\Pi - a^\uparrow \mathrm{D}\Pi\Big] y(\sup \mathrm{dom}\, y - \delta) - \varepsilon^\uparrow, \\ \alpha^\downarrow(y) &:= \Big[\mathrm{D}\Pi - a^\downarrow \Pi\Big] y(\sup \mathrm{dom}\, y - \delta) - \varepsilon^\downarrow. \end{aligned}$$

In this case, moreover, the decision period Δ must be taken so that $\Delta \geq 2T + \delta$ (cf. Remark 5).

3.3. Achieving robust detectability

Now, we study a few pathways to achieve robust detectability in a number of relevant cases of interest for epidemic control. Unless stated otherwise, in the remainder of the section we shall assume that the parameters $o^\uparrow, o^\downarrow, \mu^\uparrow, \mu^\downarrow$, and hence an evaluation model for z, have been fixed, and we focus on the inference model. Moreover, we assume that $\Delta \geq 2T$, so as for every decision profile $\xi \in \Xi$, the signals $\Pi y(\cdot|\xi)$, $\Pi\theta(\cdot|\xi)$, $\mathrm{D}\Pi y(\cdot|\xi)$, and $\mathrm{D}\Pi\theta(\cdot|\xi)$ are defined for all $t \in \mathrm{dom}\, z(\cdot|\xi)_{\geq t_1}$, where $t_1 = \Delta$ is the first decision time. Finally, we denote by $\mathcal{M}(o^\uparrow, o^\downarrow, \mu^\uparrow, \mu^\downarrow)$ the set of tuples $(a^\uparrow_*, a^\downarrow_*, \varepsilon^\uparrow_*, \varepsilon^\downarrow_*) \in \mathbb{R}^4_{\geq 0}$ generating an inference model satisfying Assumption 2 with respect to the evaluation model generated by $(o^\uparrow, o^\downarrow, \mu^\uparrow, \mu^\downarrow)$.

3.3.1. Robust detectability in presence of uncertainty

As a first case, we assume to have available some *ideal* parameters $(a^\uparrow_*, a^\downarrow_*, \varepsilon^\uparrow_*, \varepsilon^\downarrow_*)$ generating an inference model satisfying robust detectability for an *ideal* measurement signal y^*. For example, $y^* = \theta$ and $(a^\uparrow_*, a^\downarrow_*, \varepsilon^\uparrow_*, \varepsilon^\downarrow_*) = (o^\uparrow, o^\downarrow, \mu^\uparrow, \mu^\downarrow)$. Then, we suppose that, instead of y^*, we measure the "perturbed signal"

$$y(\cdot|\xi) := y^*(\cdot|\xi) + w(\cdot|\xi),$$

in which $w(\cdot|\xi) : \mathrm{dom}\, y^*(\cdot|\xi) \to \mathbb{R}$ models additive perturbations. We thus consider the problem of constructing a new tuple $(a^\uparrow, a^\downarrow, \varepsilon^\uparrow, \varepsilon^\downarrow)$ generating an inference model for the noisy output y still satisfying robust detectability.

We consider the following assumption restricting the class of tolerable uncertainty.

Assumption 6. There exist $\nu^\uparrow \in [0, \varepsilon^\uparrow_*]$ and $\nu^\downarrow \in [0, \varepsilon^\downarrow_*]$ such that, for every decision profile $\xi \in \Xi$ satisfying $t_{\sup \mathrm{dom}\, \xi} \geq 2T$, the following holds[2]

$$\begin{aligned} \frac{1}{T}\int_{t-T}^{t} \Big(w(s|\xi) - w(t-T|\xi)\Big) ds \in &\left[-\nu^\uparrow, \frac{\nu^\downarrow}{a^\downarrow_*}\right] \\ &+ \left[a^\uparrow_*, \frac{1}{a^\downarrow_*}\right] \frac{1}{T}\int_{t-2T}^{t-T} \Big(w(s|\xi) - w(t-2T|\xi)\Big) ds \end{aligned} \tag{11}$$

with,[3] $t = \sup \mathrm{dom}\, w(\cdot|\xi)$.

Under this assumption, the following proposition (proved in the Appendix) holds.

Proposition 4. *Suppose that* $(a^\uparrow_*, a^\downarrow_*, \varepsilon^\uparrow_*, \varepsilon^\downarrow_*) \in \mathcal{M}(o^\uparrow, o^\downarrow, \mu^\uparrow, \mu^\downarrow)$ *and that Assumption 6 holds. Then, for all* $\varepsilon^\uparrow \in [0, \varepsilon^\uparrow_* - \nu^\uparrow]$ *and* $\varepsilon^\downarrow \in [0, \varepsilon^\downarrow_* - \nu^\downarrow]$, $(a^\uparrow_*, a^\downarrow_*, \varepsilon^\uparrow, \varepsilon^\downarrow) \in \mathcal{M}(o^\uparrow, o^\downarrow, \mu^\uparrow, \mu^\downarrow)$.

3.3.2. Robust detectability for linear filters

In this section, we consider the case in which for all decision profiles $\xi \in \Xi$ and all $t \in \mathrm{dom}\, y(\cdot|\xi)$,

$$y(t|\xi) = r\mathrm{e}^{-\lambda t}\theta(0|\xi) + r\lambda \int_0^t \mathrm{e}^{-\lambda(t-s)}\theta(s|\xi) ds \tag{12}$$

for some $r, \lambda > 0$ *possibly unknown* (see Remark 10 below). For simplicity, we assume r and λ to be constant. Extensions to cases in which r and λ depend on time are possible (provided that they are uniformly lower bounded) at the price, however, of a more involved treatment.

Remark 8. Eq. (12) implies $y(0|\xi) = r\theta(0|\xi)$. This is assumed to simplify the forthcoming analysis, and can be relaxed to an arbitrary initial condition of $y(\cdot|\xi)$ provided that the initial decision time t_1 is sufficiently large to make the term $r\mathrm{e}^{-\lambda t}\theta(0|\xi)$ negligible.

Throughout the section we make the following assumption.

Assumption 7 (*Regularity*)**.** For every $\xi \in \Xi$, $\theta(\cdot|\xi)$ is absolutely continuous.

Under Assumption 7, the signal

$$\eta(t|x) := \int_0^t \mathrm{e}^{\lambda(s-t)}\dot{\theta}(s|\xi) ds,$$

is well defined on $\mathrm{dom}\, \theta(\cdot|\xi)$, where $\dot{\theta}(\cdot|\xi)$ is the weak derivative of $\theta(\cdot|\xi)$.

In the following, we show that under a growth condition on η we can find an inference model directly from the knowledge of the parameters $(o^\uparrow, o^\downarrow, \mu^\uparrow, \mu^\downarrow)$ defining the evaluation model. In particular, we make the following assumption.

Assumption 8. There exist $\nu^\uparrow \in [0, \mu^\uparrow]$ and $\nu^\downarrow \in [0, \mu^\downarrow]$ such that, for every decision profile $\xi \in \Xi$ satisfying $t_{\sup \mathrm{dom}\, \xi} \geq 2T$, the following hold

$$\begin{aligned} \Pi\eta(t|\xi) &\leq o^\uparrow \mathrm{D}\Pi\eta(t|\xi) + \nu^\uparrow, \\ o^\downarrow \Pi\eta(t|\xi) &\geq \mathrm{D}\Pi\eta(t|\xi) - \nu^\downarrow, \end{aligned}$$

with $t = \sup \mathrm{dom}\, \theta(\cdot|\xi)$.

[2] Given $a, b, x, y \in \mathbb{R}$, in (11) we use the notations $[a, b]x := [ax, xb]$ and $[a, b] + [x, y] := \{s_1 + s_2 \in \mathbb{R} : s_1 \in [a, b],\ s_2 \in [x, y]\}$.

[3] We recall that, by the definitions (1)–(2) for every $\xi \in \Xi$, the signals $z(\cdot|\xi)$, $y(\cdot|\xi)$ and thus $w(\cdot|\xi)$ are defined on $[0, t_{\sup \mathrm{dom}\, \xi} + \Delta]$. Hence, $\sup \mathrm{dom}\, z(\cdot|\xi) = \sup \mathrm{dom}\, y(\cdot|\xi) = \sup \mathrm{dom}\, w(\cdot|\xi) = t_{\sup \mathrm{dom}\, \xi} + \Delta$.

Then, the following proposition (proved in the Appendix) holds.

Proposition 5. *Suppose that Assumptions 7 and 8 hold and that y is given by (12). Let $a^\uparrow = o^\uparrow$ and $a^\downarrow = o^\downarrow$. Then, for every $\varepsilon^\uparrow \in [0, r(\mu^\uparrow - \nu^\uparrow)]$ and $\varepsilon^\downarrow \in [0, r(\mu^\downarrow - \nu^\downarrow)]$, $(a^\uparrow, a^\downarrow, \varepsilon^\uparrow, \varepsilon^\downarrow) \in \mathcal{M}(o^\uparrow, o^\downarrow, \mu^\uparrow, \mu^\downarrow)$.*

We now provide some sufficient conditions on θ and the evaluation model under which Assumption 8 always holds. In particular, we assume the following.

Assumption 9 (*Bounded Variations*)**.** There exists $\kappa \geq 0$ such that, for every decision profile $\xi \in \Xi$, $|\dot\theta(t|\xi)| \leq \kappa$ for almost all $t \in \operatorname{dom} \dot\theta(t|\xi)$.

Then, the proposition below (proved in the Appendix) holds.

Proposition 6. *Suppose that Assumption 9 holds, and that the parameters $(\omega^\uparrow, \omega^\downarrow, \mu^\uparrow, \mu^\downarrow)$ satisfy*

$$\mu^\uparrow \geq \frac{2\kappa}{\lambda}(1 + o^\uparrow), \quad \mu^\downarrow \geq \frac{2\kappa}{\lambda}(1 + o^\downarrow). \tag{13}$$

Then, Assumption 8 holds with $\nu^\uparrow = 2\kappa(1 + o^\uparrow)/\lambda$ and $\nu^\downarrow = 2\kappa(1 + o^\downarrow)/\lambda$.

Remark 9. The statement of Proposition 6 can be also read in the following way. If y is given by (12), and Assumption 9 holds, then *every* set of parameters $(a^\uparrow, a^\downarrow, \varepsilon^\uparrow, \varepsilon^\downarrow) \in \mathbb{R}^4_{\geq 0}$ generates a valid inference model satisfying robust detectability with respect to an evaluation model of the form $(o^\uparrow, o^\downarrow, \mu^\uparrow, \mu^\downarrow)$ with $o^\uparrow = a^\uparrow$, $o^\downarrow = a^\downarrow$ and with $\mu^\uparrow$ and $\mu^\downarrow$ satisfying the bounds (13). In this respect, we stress that the fact that Proposition 6 requires $\mu^\uparrow, \mu^\downarrow > 0$ does not imply that also $\varepsilon^\uparrow$ and $\varepsilon^\downarrow$ must be strictly positive. Indeed, we can always choose $\varepsilon^\uparrow = \varepsilon^\downarrow = 0$. The assumptions of Proposition 6 rather limit the class of evaluation models for which the inference model fixed as in Proposition 5 can guarantee robust detectability. The imposed limits, in turn, are directly related to the maximum rate of change of θ, given by κ, and reflect the fact that our promptness in detecting events of θ from y is necessarily affected by the delay introduced by the exponential filter (12).

Remark 10. The result of Proposition 6 holds for every T, r and λ which are independent quantities in general. Hence, no knowledge of r and λ is required in principle to choose an inference model for (12). In view of Remark 9, indeed, the values of r and λ only affect the resolution with which the inference model can detect events of the underlying process z (or, formally, the class of evaluation models for which it guarantees robust detectability).

We now consider the question whether it is possible to exploit prior knowledge on λ and r to reduce the bounds (13). In particular, Proposition 7 below (proved in the Appendix) establishes a result stating that, if we can choose T and $o^\downarrow$, $o^\uparrow$ in terms of the process parameter λ, then tighter bounds can be established. This ultimately results in a higher accuracy of inference models produced by Proposition 5 (see Remarks 9 and 10).

Proposition 7. *Suppose that Assumption 9 holds, and that $o^\uparrow$, $o^\downarrow$, and T satisfy*

$$o^\uparrow = \mathrm{e}^{-\lambda T}, \quad o^\downarrow = \mathrm{e}^{\lambda T}, \quad T = 1/\lambda. \tag{14}$$

If the parameters $(\omega^\uparrow, \omega^\downarrow, \mu^\uparrow, \mu^\downarrow)$ satisfy

$$\mu^\uparrow \geq \frac{2\kappa}{\lambda}, \quad \mu^\downarrow \geq \frac{2\kappa}{\lambda}, \tag{15}$$

then Assumption 8 holds with $\nu^\uparrow = 2\kappa/\lambda$ and $\nu^\downarrow = 2\kappa/\lambda$.

4. Numerical simulations: The case of COVID-19

Extensive numerical simulations showing the efficacy of FPSP in suppressing the COVID-19 outbreak have been already presented in Bin, Cheung et al. (2021). Hence, here we focus on evaluating the effect of the parameters $a^\uparrow$, $a^\downarrow$, $\varepsilon^\uparrow$ and $\varepsilon^\downarrow$ on the closed-loop performance. As in Bin, Cheung et al. (2021), we rely on the SIDARTHE model of Giordano et al. (2020), as it provides a specific model of the form (7) tuned on the early COVID outbreak in Italy. In particular, the model is obtained by letting in (7) $n_C = 6$, $C = (D, A, R, T, H, E)$, and

$$f_S(\beta, S, I, C) = -\beta S(\sigma_1 I + \sigma_2 D + \sigma_3 A + \sigma_4 R)$$
$$f_I(\beta, S, I, C) = \beta S(\sigma_1 I + \sigma_2 D + \sigma_3 A + \sigma_4 R) - (\sigma_5 + \sigma_6 + \sigma_7) I$$
$$f_C(\beta, S, I, C) = \begin{pmatrix} \sigma_5 I - (\sigma_8 + \sigma_9) D \\ \sigma_6 I - (\sigma_{10} + \sigma_{11} + \sigma_{12}) A \\ \sigma_8 D + \sigma_{10} A - (\sigma_{13} + \sigma_{14}) R \\ \sigma_{11} A + \sigma_{13} R - (\sigma_{15} + \sigma_{16}) T \\ \sigma_7 I + \sigma_9 D + \sigma_{12} A + \sigma_{14} R + \sigma_{15} T \\ \sigma_{16} T \end{pmatrix}.$$

The parameters σ_i, $i = 1, \ldots, 16$ are rates (in 1/days) and have been identified in Giordano et al. (2020) as $\sigma_1 = 0.57$, $\sigma_2 = \sigma_4 = 0.011$, $\sigma_3 = 0.456$, $\sigma_5 = 0.171$, $\sigma_6 = \sigma_8 = 0.125$, $\sigma_7 = \sigma_9 = 0.034$, $\sigma_{10} = 0.371$, $\sigma_{11} = 0.012$, $\sigma_{12} = \sigma_{14} = \sigma_{15} = 0.017$, $\sigma_{13} = 0.027$ and $\sigma_{16} = 0.003$. The dimensionless parameter $\beta \in [0, 1]$ modulates the rate of effective contacts, and it equals $\beta^\uparrow = 1$ during normal days and $\beta^\downarrow = 0.175$ during lockdown days, as suggested in Bin, Cheung et al. (2021) and Giordano et al. (2020). Finally, the compartment S (Susceptible) represents susceptible individuals, I (Infected) the infected asymptomatic and undetected, D (Detected) the infected asymptomatic and detected, A (Ailing) the infected symptomatic and undetected, R (Recognized) the infected symptomatic and detected, T (Threatened) the acutely symptomatic detected, H (Healed) the healed, and E (Extinct) the dead individuals. Initial conditions are restricted to $[0, 1]^8$, so as compartments represent percentages of people. We refer to Giordano et al. (2020) for further details.

The control goal is to keep the active cases under control while avoiding full lockdown. As anticipated in Section 3.1, in this case $\zeta[x]$ is the flow of the SIDARTHE model. Hence, once fixed an initial condition, for every decision profile $\xi = ((x_k, t_k)_{k \in \operatorname{dom} \xi}, \bar t)$ the signal $z(\cdot|\xi)$ equals the unique solution (S, I, D, A, R, T, H, E) of the SIDARTHE model subject to the time-varying parameter

$$\beta(t|\xi) = \beta(t, x_k), \quad \forall t \in [t_k, t_{k+1}],$$

where $t_{\sup \operatorname{dom} \xi + 1} := \bar t$ and where the function $\beta(t, x)$ is defined in (8). Moreover, according to the SIDARTHE model, among the active cases we measure D, R and T. Hence, our measurement is $y(\cdot|\xi) = D + R + T$.

From the model's equations, we obtain for every $\lambda \geq \max\{\sigma_9, \sigma_{14}, \sigma_{15} + \sigma_{16}\}$ (we omit the argument ξ)

$$\dot y = -\lambda y + r\lambda\theta$$

where

$$\theta = \gamma(z) := 5 \cdot \frac{k_I I + k_D D + k_A A + k_R R + k_T T}{k_I + k_D + k_A + k_R + k_T}, \tag{16}$$

$k_I := \sigma_5$, $k_D := \lambda - \sigma_9$, $k_A := \sigma_{10} + \sigma_{11}$, $k_R := \lambda - \sigma_{14}$, $k_T := \lambda - \sigma_{15} - \sigma_{16}$, and $r := (5\lambda)^{-1}(k_I + k_D + k_A + k_R + k_T)$. Therefore, the output y has the linear-filter property (12) with respect to the variable θ. Note that $I + D + A + R + T$ represents the total infected population. Hence, θ represents a (normalized) weighted sum of all the components of the infected population in which some components are weighted more than others. For example, with $\lambda = 1/7$, the above values of the coefficients σ_i give $k_I \approx 0.17$, $k_D \approx 0.1$, $k_A \approx 0.38$, $k_R \approx 0.12$, $k_T \approx 0.12$. Hence, the unmeasured variables I and A weight more in the sum. Fig. 1 shows the "free" evolution of the epidemic from the initial conditions $I(0) = 1/500$, $D(0) = A(0) = R(0) = T(0) = H(0) = E(0) = 0$, $S(0) = 1 - I(0)$. As shown in the figure, θ reaches its peak before $I + D + A + R + T + H + E$. This is due to the larger importance of I and A in the sum (16).

In the following simulations, the control logic (4)–(5) is applied to the SIDARTHE model defined above for $t_f := 365$ days, with $P = T = 7$ days, $\Delta = 2T$, $\lambda = 1/T$ (chosen according to Proposition 7), and with

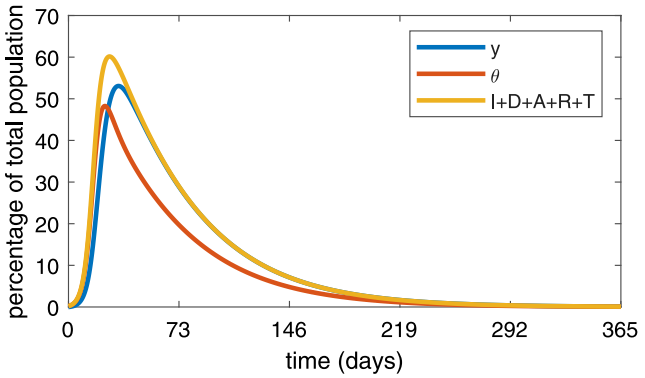

Fig. 1. Time series of the variables $y = D + R + T$, θ defined in (16), and the total infected population $I + D + A + R + T$ in the case in which no mitigation procedure is in place (namely, $\beta = \beta^{\uparrow}$). In this case, $|\theta|_{\text{RMS}} \approx 14.8$, with $|\cdot|_{\text{RMS}}$ defined in (17).

initial conditions $I(0) = 1/500$, $D(0) = A(0) = R(0) = T(0) = H(0) = E(0) = 0$, $S(0) = 1 - I(0)$, $x_0 = 0$. Figs. 2-(a) and 2-(b) summarize the closed-loop behavior obtained for different values of $a^{\uparrow}$ and $a^{\downarrow}$ and with $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 0$, while Fig. 3 shows some sample time series. In particular, Fig. 2-(a) shows the mean value of the resulting decision time series x rounded to the first decimal. The larger the value, the better it is, since x equals the number of normal days in each FPSP period. Fig. 2-(b) shows instead the RMS norm of θ, defined as

$$|\theta|_{\text{RMS}} := \sqrt{\frac{1}{t_f}\int_0^{t_f} \left(100 \cdot \theta(s|\xi)\right)^2 ds}, \tag{17}$$

and rounded to the second decimal. The smaller the value the better it is, as θ is a weighted sum of the active cases. For reference, in the uncontrolled outbreak case depicted in Fig. 1, $|\theta|_{\text{RMS}} \approx 14.8$. All the simulated pairs $(a^{\uparrow}, a^{\downarrow})$ stabilize the infection-free equilibrium, although some are better than others. In particular, the orange contours in Figs. 2-(a) and 2-(b) group pairs $(a^{\uparrow}, a^{\downarrow})$ associated with a good compromise between average x and $|\theta|_{\text{RMS}}$.

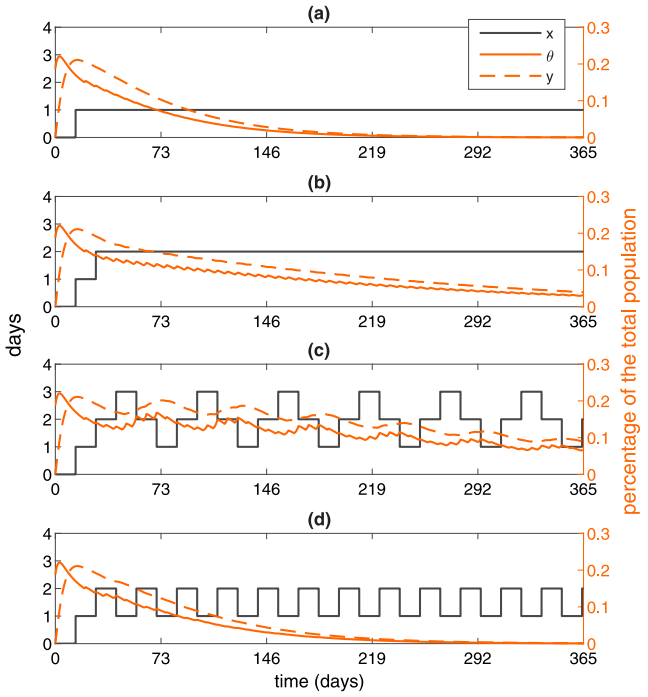

Fig. 3. Time series of the decision variable x, the observed variable y and the controlled variable θ obtained with $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 0$ and with **(a)** $(a^{\uparrow}, a^{\downarrow}) = (0.3, 0.3)$, **(b)** $(a^{\uparrow}, a^{\downarrow}) = (0.3, 0.9)$, **(c)** $(a^{\uparrow}, a^{\downarrow}) = (0.3, 2.7)$, and **(d)** $(a^{\uparrow}, a^{\downarrow}) = (0.9, 2.1)$.

Among the time series shown in Fig. 3, the second and third (shown respectively in Figs. 3-(b) and 3-(c) and obtained with $(a^{\uparrow}, a^{\downarrow}) = (0.3, 0.9)$ and $(a^{\uparrow}, a^{\downarrow}) = (0.3, 2.7)$) are inside the orange contour and, indeed, are associated with a relatively large mean value of x (about 2 days per FPSP period) and with a relatively small RMS norm of θ. Specifically, the pair $(a^{\uparrow}, a^{\downarrow}) = (0.3, 0.9)$ produces a stable behavior

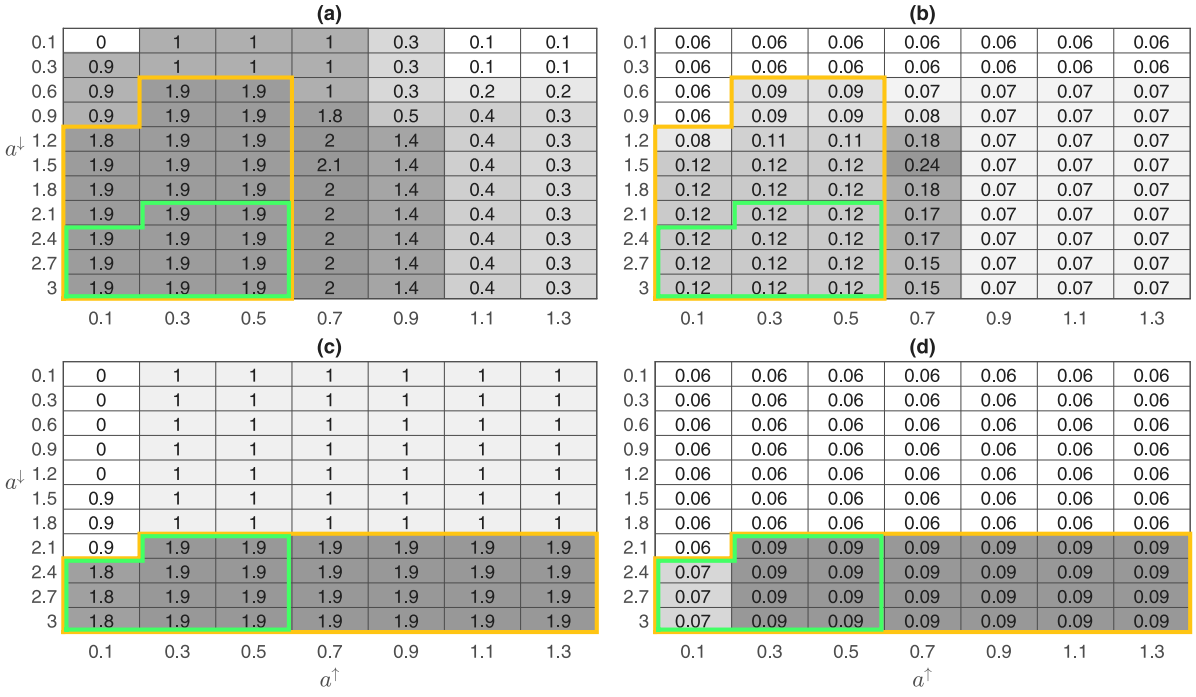

Fig. 2. Heat maps summarizing the closed-loop performance for different values of $a^{\uparrow}$ and $a^{\downarrow}$. Figures **(a)** and **(c)**: mean value of the decision variable x with $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 0$ and $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 10^{-4}$ respectively. Figures **(b)** and **(d)**: RMS norm $|\theta|_{\text{RMS}}$ of the variable θ with $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 0$ and $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 10^{-4}$ respectively. Orange contours group pairs $(a^{\uparrow}, a^{\downarrow})$ associated with a good trade-off between average of x and RMS norm of θ. Green contours group pairs $(a^{\uparrow}, a^{\downarrow})$ which represent a good trade-off for both $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 0$ and $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 10^{-4}$.

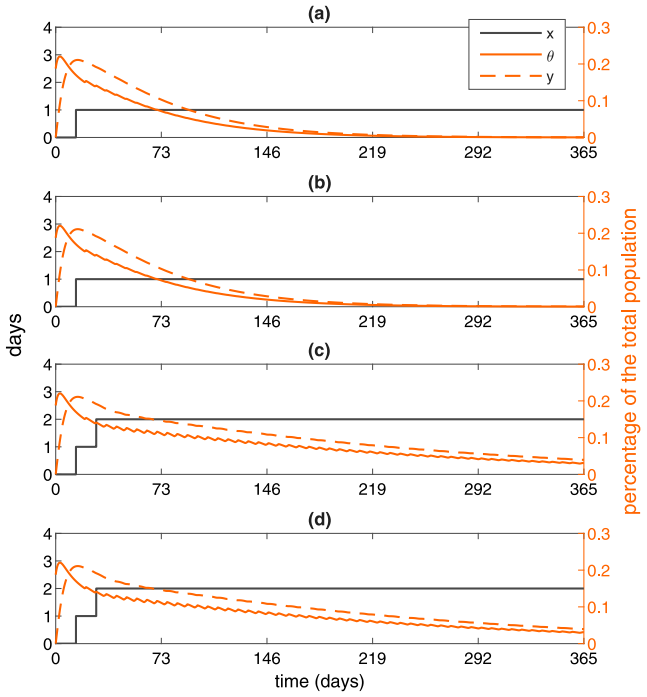

Fig. 4. Time series of the decision variable x, the observed variable y and the controlled variable θ obtained with $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 10^{-4}$ and with **(a)** $(a^{\uparrow}, a^{\downarrow}) = (0.3, 0.3)$, **(b)** $(a^{\uparrow}, a^{\downarrow}) = (0.3, 0.9)$, **(c)** $(a^{\uparrow}, a^{\downarrow}) = (0.3, 2.7)$, and **(d)** $(a^{\uparrow}, a^{\downarrow}) = (0.9, 2.1)$.

in which the decision variable converges in two steps to a stationary steady state equal to 2 days for each FPSP period (Fig. 3-(b)). The pair $(a^{\uparrow}, a^{\downarrow}) = (0.3, 2.7)$, instead, is characterized by a persistent oscillation of the decision variable, although maintaining the same average value of about 2 days per period. This persistent oscillation may be caused by the fact that stationarity (Assumption 3) does not hold in a strict sense. The time series shown respectively in e 3-(a) and 3-(d) are instead outside the orange contours, as they are too conservative in terms of the decision x, although they are associated with a faster decay of the infected population. As evident from the heat maps of Figs. 2-(a) and 2-(b), indeed, for fixed $a^{\uparrow}$ increasing $a^{\downarrow}$ leads to a less conservative control policy. Conversely, for fixed $a^{\downarrow}$ increasing $a^{\uparrow}$ leads to a more conservative control policy.

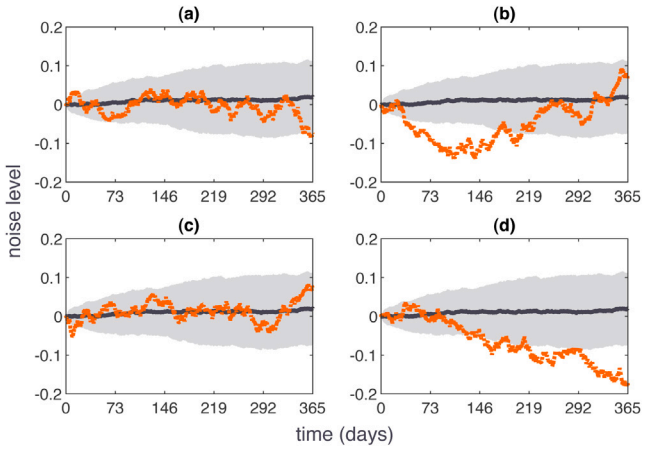

Fig. 6. Time series of the perturbation v used in the simulations of Fig. 5. Solid lines depict averages of the 100 realizations. Transparent bands depict an interval $[-\hat{\sigma}, \hat{\sigma}]$ around the average graphs, where $\hat{\sigma}$ denotes the empirical estimate of the standard deviation. Dotted lines depict the realization associated to the dotted graphs in Fig. 5.

For the same range of values for $(a^{\uparrow}, a^{\downarrow})$, Figs. 2-(c) and 2-(d) show the mean value of x and the RMS norm of θ in the case of $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 10^{-4}$. Comparing Fig. 2-(a) with 2-(c) and Fig. 2-(b) with 2-(d) shows that positive values of $\varepsilon^{\uparrow}$ and $\varepsilon^{\downarrow}$ have the effect of "shifting" in the bottom-right direction and "enlarging" the heat maps. In particular, some of the choices of $(a^{\uparrow}, a^{\downarrow})$ that with $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 0$ showed a good behavior, now with $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 10^{-4}$ are too conservative. For instance, compare the time series shown in Fig. 3, which corresponds to $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 0$ with those shown in Fig. 4, which instead are obtained with $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 10^{-4}$.

As for Figs. 2-(a) and 2-(b), the orange contours in Figs. 2-(c) and 2-(d) group choices of $(a^{\uparrow}, a^{\downarrow})$ associated with a good trade-off between average value of x and RMS norm of θ. Moreover, in all four heat maps, green contours group the "intersection" of the orange ones, thus

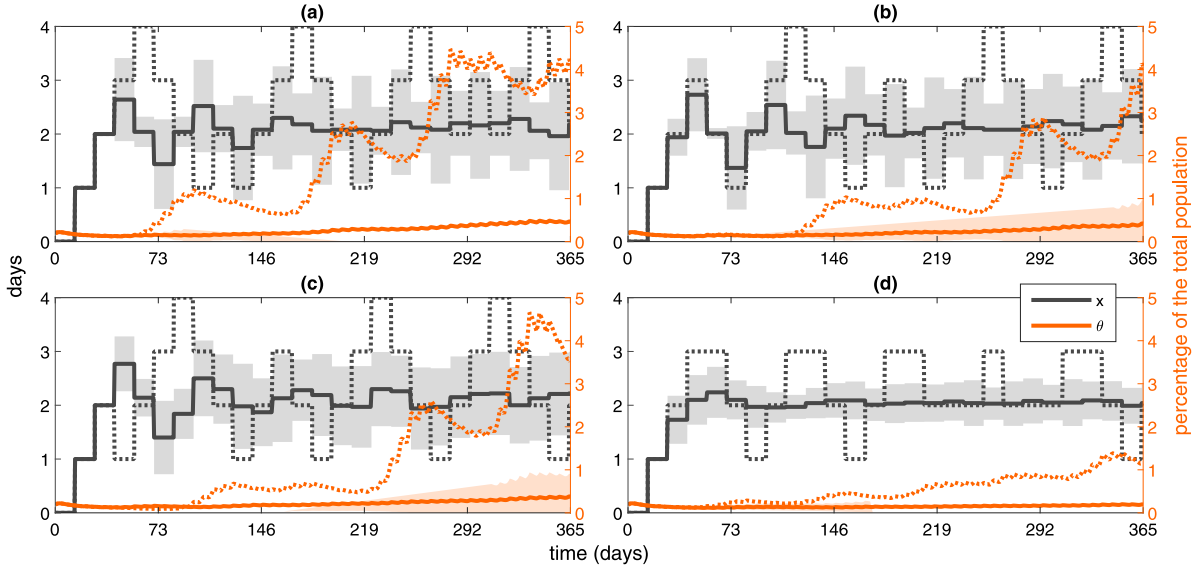

Fig. 5. Time series showing the closed-loop behavior in presence of uncertainty in the measurements. In all the simulations, $(a^{\uparrow}, a^{\downarrow}) = (e^{-1}, e^{1})$. In **(a)** $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 0$, in **(b)** $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 10^{-6}$, in **(c)** $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 10^{-5}$, and in **(d)** $\varepsilon^{\uparrow} = \varepsilon^{\downarrow} = 10^{-4}$. For each choice of parameters, the corresponding figure summarizes the result of 100 simulations, each one with a different realization of d in (20). Solid lines depict average values of the 100 realizations. Transparent bands depict an interval $[-\hat{\sigma}, \hat{\sigma}]$ around the average graphs, where $\hat{\sigma}$ denotes the empirical estimate of the standard deviation. Dotted lines depict the worst (in therms of $|\theta|_{\mathrm{RMS}}$) realizations.

individuating the set of choices of $(a^\uparrow, a^\downarrow)$ associated with a good trade-off between average value of x and RMS norm of θ for both choices of $(\varepsilon^\uparrow, \varepsilon^\downarrow)$. Those are choices which are "robust" with respect to different choices of $(\varepsilon^\uparrow, \varepsilon^\downarrow)$ inside $\{(\varepsilon^\uparrow, \varepsilon^\downarrow) \in \mathbb{R}^2_{\geq 0} \ : \ \varepsilon^\uparrow = \varepsilon^\downarrow \in [0, 10^{-4}]\}$. In this respect, we underline that, since the values y and θ are percentages of the overall population, a value of $\varepsilon^\uparrow$ and $\varepsilon^\downarrow$ of the order of 10^{-4} equals the 0.01% of the population, and thus represents a reasonable value.

We also observe that the choice suggested by Proposition 7, i.e.

$$(a^\uparrow, a^\downarrow) = \left(\mathrm{e}^{-\lambda T}, \mathrm{e}^{\lambda T}\right) = \left(\mathrm{e}^{-1}, \mathrm{e}^{1}\right) \approx (0.368, 2.718) \tag{18}$$

lies with margin inside the interior of the green contour. Thus, the values (18) suggested by Proposition 7 represent a robust choice guaranteeing good performances for different values of $\varepsilon^\uparrow$ and $\varepsilon^\downarrow$.

Regarding the choice of $\varepsilon^\uparrow$ and $\varepsilon^\downarrow$, we observe that while in an ideal case in which y is measured without uncertainty $\varepsilon^\uparrow = \varepsilon^\downarrow = 0$ is fine, $\varepsilon^\uparrow, \varepsilon^\downarrow > 0$ may instead help in presence of uncertainty, as they act as a regularizer. In particular, Fig. 5 shows some simulations obtained with $a^\uparrow$ and $a^\downarrow$ given by (18), with $\varepsilon^\uparrow = \varepsilon^\downarrow = \varepsilon$ for $\varepsilon = 0, 10^{-6}, 10^{-5}, 10^{-4}$, and with y given by

$$y = (D + R + T) \cdot (1 + \nu), \tag{19}$$

in which ν is a disturbance term produced as

$$\dot{u}_n = \ell(d - u_n), \quad \nu = A_\nu u \tag{20}$$

with $u_n(0) = 0$, $A_\nu = 10^7$, $\ell = 10^{-3}$, and where d is obtained as a linear interpolation of a uniform random noise with values in $[-10^{-5}/2, 10^{-5}/2]$ and sampled with frequency $10^5\,\mathrm{days}^{-1}$. We observe that (19) fits in the case considered in Section 3.3.1, with $w := (D + R + T)\nu$. For every value of ε, 100 simulations have been performed.

As shown in Fig. 5, increasing values of ε lead to a more robust behavior with respect to the perturbation on y. Fig. 6 shows some statistics of the perturbation signals used in the simulation. In particular, the level of perturbation is so that the measurement of y deviates also 10%–20% from its nominal value $D + R + T$.

Finally, we underline that in all the simulation the control logic considerably limits the virus growth (compare with Fig. 1) while maintaining, on average, a steady-state number of normal days of about 2 per period.

5. Conclusions

In this paper we have studied a class of hysteresis-based control schemes with the aim of providing theoretical support of the use of supervisory control for data-driven containment of epidemics. In particular, we have focused on the setting of Bin, Cheung et al. (2021) in which supervisory control is used to tune online, from measured data, the value of the FPSP duty cycle (Sections 3 and 4). Nevertheless, the theory developed in Section 2 goes far beyond such application, and can be used for a broader class of problems. Specifically, in Section 2 we have developed the main theoretical framework in a rather abstract setting, where inference and evaluation models are left generic. First, we have proved invariance and attractiveness properties of sets of decisions that lead to a good behavior of the observed variables. Then, under robust detectability, we have proved that such decisions also lead to a satisfactory behavior for the unmeasured underlying process.

In Section 3, we have restricted the focus to a case relevant for epidemic control, and we have provided additional results determining conditions under which robust detectability can be achieved in presence of uncertainty and in the relevant case in which the measured output is a filtered version of the variables whose growth must be contained. Moreover, we have shown that the knowledge of the filter time constant can be used to tune the controller parameters to improve performance (Proposition 7).

Overall, the theoretical analysis carried out in Sections 2 and 3 confirms prior intuitions on hysteresis-based control, and in particular the claims on its robustness with respect to perturbations and model uncertainties. Indeed, models enter into play in terms of evaluation models, inference models, and robust detectability (see Section 2.1 and Assumption 2). Evaluation and inference models are descriptions capturing the qualitative way in which decisions affect, respectively, the unmeasured controlled process and the measured variables. As such, they can be approximate models, and are considerably weaker hypotheses than typical models expressed in terms of differential equations. Robust detectability, on the other hand, is an assumption linking inference and evaluation models, permitting in this way to infer the response of the unmeasured variables to a decision from measurements. As discussed in Remark 1, the decisions of the hysteresis logic remain valid for all uncertainties and disturbances that do not ruin robust detectability. This, in turn, is what confers robustness on the decision logic.

The numerical simulations performed in Section 4 validate the theoretical conclusions in the context of the model of COVID-19 outbreak. The simulations confirm robustness with respect to uncertainty in the measurements and to variability of the control parameters. In particular, in all the simulated cases the outbreak is contained, and also in the worst realizations the overall closed-loop behavior is better, from the virus growth standpoint, than the uncontrolled outbreak. Remarkably, simulations provide clear evidence in favor of the theoretical values for the parameters $a^\uparrow$ and $a^\downarrow$ found in Proposition 7 in studying performance improvement. In particular, such values are associated, at the same time, with a good trade-off between virus growth and number of normal days per FPSP period, and robustness with respect to different values of $\varepsilon^\uparrow$ and $\varepsilon^\downarrow$. The latter parameters, in turn, were shown in the simulations to have a beneficial effect in presence of uncertainty in the measurements, while overall leading to a more conservative behavior for fixed $(a^\uparrow, a^\downarrow)$.

Ultimately, both theoretical analysis and simulations provide evidences supporting the use of hysteresis-based decision mechanisms in the control of epidemics. Nevertheless, many aspects of reality have been neglected in both theory and simulations, and the presented results thus only provide a starting point requiring further empirical study. Moreover, also several theoretical questions remain open. For instance, the theory of Section 2 only relies on qualitative models on how decisions affect the underlying process (evaluation and inference models), and quantitative models, such as for example SI-like differential equations, are not considered. This is inherited by Section 3, which indeed focuses on how to construct inference and evaluation models. A SIDARTHE equation is only used in Section 4 to model a plausible outbreak and, notably, to infer a relation of the form (12) linking y and θ. In turn, this is an example on how additional information or assumptions may help in building better inference models. However, there are many other ways in which such kind of prior information may be used to refine the models (e.g., how to choose $\varepsilon^\uparrow$ and $\varepsilon^\downarrow$), and here these are not discussed.

Finally, we observe that seasonality, vaccination, and even new restrictions can change the effect decisions have on the outbreak. These sources of time-variability are not considered here. But since they happen at a slower time scale than the virus dynamics, the developed theory can be still used within a limited time horizon, with models that must be possibly adapted time to time so as to reflect the new conditions. An interesting alternative, is to embed seasonality, vaccinations, and the other sources of variability directly in the inference and evaluation models, that thus become time-varying. In turn, this is a further open problem requiring additional research.

Declaration of competing interest

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

Appendix. Technical proofs

Proof of Lemma 2. Assume that $x \in \widetilde{X}^{\uparrow}$ but $x^- \notin \widetilde{X}^{\uparrow}$. This implies $\alpha^{\uparrow}(y(\cdot|\xi^-)) \in A$ for some $\xi^- \in \Xi(x^-)$. Assumption 3 then implies that $\alpha^{\uparrow}(y(\cdot|\xi^-)) \in A$ for all $\xi^- \in \Xi(x^-)$. Then, Assumption 4 implies $\alpha^{\uparrow}(y(\cdot|\xi)) \in A$ for all $\xi \in \Xi((x^-)^+) = \Xi(x)$. This, however, implies $x = (x^-)^+ \notin \widetilde{X}^{\uparrow}$ which is a contradiction. Hence, $x^- \in \widetilde{X}^{\uparrow}$ for all $x \in \widetilde{X}^{\uparrow}$, and the first implication holds. Next, notice that by definition and by the implication just proved, $x \in \widetilde{X}^{\uparrow}_{+} \implies x^- \in \widetilde{X}^{\uparrow} \subset \widetilde{X}^{\uparrow}_{+}$. Hence also $\widetilde{X}^{\uparrow}_{+}$ is closed under $(\cdot)^-$. The last two implications are proved by similar arguments. □

Proof of Lemma 3. Suppose that $\widetilde{X}^{\uparrow} = \mathcal{X}$. Then, $\widetilde{X}^{\uparrow}_{+} = \mathcal{X}$, and thus $\widetilde{X}^{\downarrow} \neq \emptyset \implies \widetilde{X}^{\downarrow}_{-} \neq \emptyset \implies \widetilde{X}_{\pm} \neq \emptyset$ and $\widetilde{X}^{\uparrow}_{+} \cup \widetilde{X}^{\downarrow}_{-} = \mathcal{X}$.

Suppose, instead, that $\widetilde{X}^{\uparrow} \subsetneq \mathcal{X}$. Then, Lemma 2 implies that $\bar{x} := (\sup \widetilde{X}^{\uparrow})^+ \in \mathcal{X} \setminus \widetilde{X}^{\uparrow}$. By definition of $\widetilde{X}^{\uparrow}_{+}$, we have $\bar{x} \in \widetilde{X}^{\uparrow}_{+}$. Moreover, in view of Assumption 3, $\bar{x} \notin \widetilde{X}^{\uparrow}$ implies $\alpha^{\uparrow}(y(\cdot|\xi)) \in A$ for all $\xi \in \Xi(\bar{x})$. In view of Assumptions 1 and 3, this implies $\alpha^{\downarrow}(y(\cdot|\xi)) \notin A$ for all $\xi \in \Xi(x)$, and hence $x \in \widetilde{X}^{\downarrow}$. Therefore, $\bar{x} \in \widetilde{X}^{\uparrow}_{+} \cap \widetilde{X}^{\downarrow} \subset \widetilde{X}_{\pm}$, which implies $\widetilde{X}_{\pm} \neq \emptyset$.

Finally, the equality $\widetilde{X}^{\uparrow}_{+} \cup \widetilde{X}^{\downarrow}_{-} = \mathcal{X}$ follows by Lemma 2 by noticing that, since $\widetilde{X}_{\pm} = \widetilde{X}^{\uparrow}_{+} \cap \widetilde{X}^{\downarrow}_{-} \neq \emptyset$, then every element of $\mathcal{X} \setminus \widetilde{X}_{\pm}$ is a predecessor of an element of $\widetilde{X}^{\uparrow}_{+}$ or a successor of an element of $\widetilde{X}^{\downarrow}_{-}$ and thus belongs to $\widetilde{X}^{\uparrow}_{+} \cup \widetilde{X}^{\downarrow}_{-}$. □

Proof of Proposition 2. Pick $x \in X^\star$ and $\xi \in \Xi(x)$ arbitrarily. In view of (5) and Condition (b), there are only two cases for which $F(x, y(\cdot|\xi)) \neq \{x\}$ (in which case we would have $F(x, y(\cdot|\xi)) \subset X^\star$): either **(i)** $\alpha^{\uparrow}(y(\cdot|\xi)) \in \partial A$ and $\alpha^{\downarrow}(y(\cdot|\xi)) \notin \partial A$, or **(ii)** $\alpha^{\downarrow}(y(\cdot|\xi)) \in \partial A$ and $\alpha^{\uparrow}(y(\cdot|\xi)) \notin \partial A$.

Consider case **(i)**. We prove that $x^- \in X^\star$ by contradiction. For, suppose that $x^- \notin X^\star$. Since $x \in X^\star \subset \widetilde{X}^{\uparrow}$ and, by Lemma 2, $\widetilde{X}^{\uparrow}$ is closed under the predecessor operator, then $\alpha^{\uparrow}(y(\cdot|\xi')) \notin A$ for all $\xi' \in \Xi(x^-)$. In view of Assumption 3, if $x^- \notin X^\star$, then necessarily $\alpha^{\downarrow}(y(\cdot|\xi')) \in A$ for all $\xi' \in \Xi(x^-)$, as otherwise x^- would be in $\widetilde{X}^{\downarrow}$, and hence in $X^\star$. But in view of Assumptions 3 and 4, this implies that all the predecessors of x lie outside $\widetilde{X}^{\downarrow}$, so that $X^\star \cap \mathcal{X}_{<x} = \emptyset$. On the other hand, Condition (c) and Assumption 4 imply that $X^\star \subset \mathcal{X}_{\leq x}$. Hence, we conclude that $X^\star = \{x\}$. However, this violates Condition (a). Thus, by contradiction, we conclude that, necessarily, $x^- \in X^\star$. Then, by (5), $F(x, y(\cdot|\xi)) = \{x, x^-\} \subset X^\star$ holds. In the case **(ii)**, $F(x, y(\cdot|\xi)) = \{x, x^+\} \subset X^\star$ is proved by a similar argument. Thus, in both cases we have $F(x, y(\cdot|\xi)) \subset X^\star$ whenever $x \in X^\star$, and the claim follows. □

Proof of Proposition 3. We first prove invariance of $\widetilde{X}^{\uparrow}_{+}$. Pick $x \in \widetilde{X}^{\uparrow}_{+}$ and $\xi \in \Xi(x)$ arbitrarily. We have two possibilities: **(i)** $x \in \widetilde{X}^{\uparrow}$, or **(ii)** $x \in \widetilde{X}^{\uparrow}_{+} \setminus \widetilde{X}^{\uparrow}$. First, assume **(i)** holds. By definition of $\widetilde{X}^{\uparrow}_{+}$, in this case $x^+ \in \widetilde{X}^{\uparrow}_{+}$. Moreover, since under Assumptions 1, 3 and 4, $\widetilde{X}^{\uparrow}$ is invariant under the predecessor operator (Lemma 2), then $\{x^-, x\} \in \widetilde{X}^{\uparrow}_{+}$ as well, implying $F(x, y(\cdot|\xi)) \subset \{x^-, x, x^+\} \subset \widetilde{X}^{\uparrow}_{+}$. Consider now case **(ii)**. As $x \notin \widetilde{X}^{\uparrow}$, then $\alpha^{\uparrow}(y(\cdot|\xi)) \in A$, so as (5) implies $F(x, y(\cdot|\xi)) = \{x^-\} \subset \widetilde{X}^{\uparrow}_{+}$ by Lemma 2. This proves that $\widetilde{X}^{\uparrow}_{+}$ is forward invariant for (4).

Forward invariance of $\widetilde{X}^{\downarrow}_{-}$ is proved by means of a symmetric argument. Finally, forward invariance of $\widetilde{X}_{\pm}$ follows by the fact that **(a)** $\widetilde{X}_{\pm} \neq \emptyset$ implies that both $\widetilde{X}^{\uparrow}_{+}$ and $\widetilde{X}^{\downarrow}_{-}$ are nonempty, and **(b)** the intersection of forward invariant sets is forward invariant. □

Proof of Proposition 4. As $(a^{\uparrow}_*, a^{\downarrow}_*, \varepsilon^{\uparrow}_*, \varepsilon^{\downarrow}_*) \in \mathcal{M}(o^{\uparrow}, o^{\downarrow}, \mu^{\uparrow}, \mu^{\downarrow})$, then for every decision profile $\xi \in \Xi$

$$\begin{aligned} &\omega^{\uparrow}(z(\cdot|\xi)) > 0 \implies \alpha^{\uparrow}_*(y^*(\cdot|\xi)) > 0, \\ &\omega^{\downarrow}(z(\cdot|\xi)) > 0 \implies \alpha^{\downarrow}_*(y^*(\cdot|\xi)) > 0, \end{aligned} \tag{A.1}$$

where $\alpha^{\uparrow}_*$ and $\alpha^{\downarrow}_*$ are given by (10) for $(a^{\uparrow}_*, a^{\downarrow}_*, \varepsilon^{\uparrow}_*, \varepsilon^{\downarrow}_*)$. Likewise, let $\alpha^{\uparrow}$ and $\alpha^{\downarrow}$ be given by (10) for $(a^{\uparrow}_*, a^{\downarrow}_*, \varepsilon^{\uparrow}, \varepsilon^{\downarrow})$ for some $\varepsilon^{\uparrow} \in [0, \varepsilon^{\uparrow}_* - \nu^{\uparrow}]$ and $\varepsilon^{\downarrow} \in [0, \varepsilon^{\uparrow}_* - \nu^{\downarrow}]$. Let $t := \sup \operatorname{dom} y^*(\cdot|\xi) = \sup \operatorname{dom} w(\cdot|\xi) = \sup \operatorname{dom} y(\cdot|\xi)$. In terms of the operators Π and D, (11) implies

$$\begin{aligned} &\Pi w(t|\xi) \geq a^{\uparrow}_* \mathrm{D}\Pi w(t|\xi) - \nu^{\uparrow}, \\ &a^{\downarrow}_* \Pi w(t|\xi) \leq \mathrm{D}\Pi w(t|\xi) + \nu^{\downarrow}. \end{aligned}$$

Hence, we have

$$\begin{aligned} \alpha^{\uparrow}_*(y^*(\cdot|\xi)) > 0 &\implies \Big[\Pi - a^{\uparrow}_* \mathrm{D}\Pi\Big] y^*(t|\xi) > \varepsilon^{\uparrow}_* \\ &\implies \Big[\Pi - a^{\uparrow}_* \mathrm{D}\Pi\Big]\Big(y(t|\xi) - w(t|\xi)\Big) > \varepsilon^{\uparrow}_* \\ &\implies \Big[\Pi - a^{\uparrow}_* \mathrm{D}\Pi\Big] y(t|\xi) > \varepsilon^{\uparrow}_* - \nu^{\uparrow} \geq \varepsilon^{\uparrow} \\ &\implies \alpha^{\uparrow}(y(\cdot|\xi)) > 0. \end{aligned}$$

Thus, $\omega^{\uparrow}(z(\cdot|\xi)) > 0 \implies \alpha^{\uparrow}(y(\cdot|\xi)) > 0$. The other implication, i.e. that $\omega^{\downarrow}(z(\cdot|\xi)) > 0 \implies \alpha^{\downarrow}(y(\cdot|\xi)) > 0$, follows by a similar argument from the second inequality of (A.1). □

Proof of Proposition 5. Pick $\xi \in \Xi$ arbitrarily. First, notice that for all t, $\mathrm{e}^{\lambda t} = \frac{d}{dt} \int_0^t \mathrm{e}^{\lambda s} ds$. Hence, integrating by parts yields (for ease of notation, we omit the argument ξ when clear)

$$\begin{aligned} y(t) &= r\mathrm{e}^{-\lambda t}\theta(0) + r\lambda \mathrm{e}^{-\lambda t}\left[\int_0^s \mathrm{e}^{\lambda\tau} d\tau\, \theta(s)\right]_0^t \\ &\quad - r\lambda \mathrm{e}^{-\lambda t} \int_0^t \left(\int_0^s \mathrm{e}^{\lambda\tau} d\tau\right) \dot{\theta}(s) ds \\ &= r\mathrm{e}^{-\lambda t}\theta(0) + r(1 - \mathrm{e}^{-\lambda t})\theta(t) \\ &\quad - r\mathrm{e}^{-\lambda t} \int_0^t (\mathrm{e}^{\lambda s} - 1)\dot{\theta}(s) ds \\ &= r\theta(t) - r\int_0^t \mathrm{e}^{\lambda(s-t)} \dot{\theta}(s) ds = r\theta(t) - r\eta(t). \end{aligned}$$

Thus, with $a^{\uparrow} = o^{\uparrow}$ and $\varepsilon^{\uparrow} \in [0, r(\mu^{\uparrow} - \nu^{\uparrow})]$, we obtain for $t = \sup \operatorname{dom} y$

$$\begin{aligned} \alpha^{\uparrow}(y) &= \Big[\Pi - a^{\uparrow}\mathrm{D}\Pi\Big] y(t) - \varepsilon^{\uparrow} \\ &= r\Big[\Pi - o^{\uparrow}\mathrm{D}\Pi\Big]\theta(t) - r\Big[\Pi - a^{\uparrow}\mathrm{D}\Pi\Big]\eta(t) - \varepsilon^{\uparrow} \\ &= r\omega^{\uparrow}(z) - r\Big[\Pi - o^{\uparrow}\mathrm{D}\Pi\Big]\eta(t) + r\mu^{\uparrow} - \varepsilon^{\uparrow} \\ &\geq r\omega^{\uparrow}(z) + r(\mu^{\uparrow} - \nu^{\uparrow}) - \varepsilon^{\uparrow} \geq r\omega^{\uparrow}(z). \end{aligned}$$

Thus, $\omega^{\uparrow}(z) > 0 \implies \alpha^{\uparrow}(y) > 0$. The implication $\omega^{\downarrow}(z) > 0 \implies \alpha^{\downarrow}(y) > 0$ is proved in a similar way. □

Proof of Proposition 6. Pick $\xi \in \Xi$ arbitrarily, and let $t = \sup \operatorname{dom} \theta(\cdot|\xi)$. Expanding the expression of $\Pi\eta(t|\xi) - o^{\uparrow}\mathrm{D}\Pi\eta(t|\xi)$ leads to (again, we omit the argument ξ)

$$\begin{aligned} \Big[\Pi - o^{\uparrow}\mathrm{D}\Pi\Big]\eta(t) &= \frac{1}{T}\int_{t-T}^{t}\int_0^s \mathrm{e}^{\lambda(\tau-s)}\dot{\theta}(\tau) d\tau ds \\ &\quad - \int_0^{t-T} \mathrm{e}^{\lambda(s+T-t)}\dot{\theta}(s) ds - \frac{o^{\uparrow}}{T}\int_{t-2T}^{t-T}\int_0^s \mathrm{e}^{\lambda(\tau-s)}\dot{\theta}(\tau) d\tau \\ &\quad + o^{\uparrow}\int_0^{t-2T} \mathrm{e}^{\lambda(s+2T-t)}\dot{\theta}(s) ds. \end{aligned}$$

Using Assumption 9 yields

$$\begin{aligned} &\Big[\Pi - o^{\uparrow}\mathrm{D}\Pi\Big]\eta(t) \\ &\leq \frac{\kappa}{T}\int_{t-T}^{t}\int_0^s \mathrm{e}^{\lambda(\tau-s)} d\tau ds + \kappa\int_0^{t-T} \mathrm{e}^{\lambda(s+T-t)} ds \\ &\quad + \frac{\kappa o^{\uparrow}}{T}\int_{t-2T}^{t-T}\int_0^s \mathrm{e}^{\lambda(\tau-s)} d\tau + \kappa o^{\uparrow}\int_0^{t-2T} \mathrm{e}^{\lambda(s+2T-t)} ds \\ &= \frac{\kappa}{\lambda}\left(2(1+o^{\uparrow}) + (1+o^{\uparrow}\mathrm{e}^{\lambda T})\mathrm{e}^{-\lambda t}\left(\frac{1-\mathrm{e}^{\lambda T}}{\lambda T} - \mathrm{e}^{\lambda T}\right)\right) \\ &\leq \frac{2\kappa}{\lambda}(1+o^{\uparrow}). \end{aligned}$$

In the same way, we find that $[o^{\downarrow}\mathrm{D}\Pi - \Pi]\eta(t) \leq 2\kappa(1+o^{\downarrow})/\lambda$. Hence, if (13) hold, then $\nu^{\uparrow} = 2\kappa(1+o^{\uparrow})/\lambda$ and $\nu^{\downarrow} = 2\kappa(1+o^{\downarrow})/\lambda$ satisfy $\nu^{\uparrow} \in [0, \mu^{\uparrow}]$ and $\nu^{\downarrow} \in [0, \mu^{\downarrow}]$ and are such that Assumption 8 holds. □

Proof of Proposition 7. Pick $\xi \in \Xi$ arbitrary, and let $t = \sup \operatorname{dom} \theta(\cdot|\xi)$. From Protter and Morrey (1985, Theorem 7, Chapter 5), we obtain (we omit the argument ξ)

$$\begin{aligned}
&\int_{t-T}^{t} \eta(s)ds = \int_{t-T}^{t}\int_{0}^{s} \mathrm{e}^{\lambda(\tau-s)}\dot\theta(\tau)d\tau ds\\
&= \int_{t-T}^{t}\int_{0}^{t-T} \mathrm{e}^{\lambda(\tau-s)}\dot\theta(\tau)d\tau ds + \int_{t-T}^{t}\int_{t-T}^{s} \mathrm{e}^{\lambda(\tau-s)}\dot\theta(\tau)d\tau ds\\
&= \int_{0}^{t-T}\int_{t-T}^{t} \mathrm{e}^{\lambda(\tau-s)}ds\dot\theta(\tau)d\tau + \int_{t-T}^{t}\int_{\tau}^{t} \mathrm{e}^{\lambda(\tau-s)}ds\dot\theta(\tau)d\tau\\
&= \frac{1}{\lambda}\mathrm{e}^{-\lambda t}(\mathrm{e}^{\lambda T}-1)\int_{0}^{t-T} \mathrm{e}^{\lambda\tau}\dot\theta(\tau)d\tau + \frac{1}{\lambda}\int_{t-T}^{t}\dot\theta(\tau)d\tau\\
&\quad - \frac{1}{\lambda}\mathrm{e}^{-\lambda t}\int_{t-T}^{t} \mathrm{e}^{\lambda\tau}\dot\theta(\tau)d\tau.
\end{aligned}$$

Therefore, using (14), we obtain

$$\begin{aligned}
&\left[\Pi - o^{\uparrow}\mathrm{D}\Pi\right]\eta(t)\\
&= -\int_{t-T}^{t} \mathrm{e}^{\lambda(s-t)}\dot\theta(s)ds + \int_{t-T}^{t}\dot\theta(s)ds - \mathrm{e}^{-1}\int_{t-2T}^{t-T}\dot\theta(s)ds\\
&\le \kappa\left(\int_{t-T}^{t} \mathrm{e}^{\lambda(s-t)}ds + T + \mathrm{e}^{-1}T\right) = \frac{2\kappa}{\lambda}.
\end{aligned}$$

A similar bound for $[\Pi - o^{\uparrow}\mathrm{D}\Pi]\eta(t)$ is obtained with the same arguments. Hence, if (15) hold, then $\nu^{\uparrow} = 2\kappa/\lambda$ and $\nu^{\downarrow} = 2\kappa/\lambda$ satisfy $\nu^{\uparrow} \in [0, \mu^{\uparrow}]$ and $\nu^{\downarrow} \in [0, \mu^{\downarrow}]$ and are such that Assumption 8 holds. □

References

Angeli, D., & Mosca, E. (2004). Adaptive switching supervisory control of nonlinear systems with no prior knowledge of noise bounds. *Automatica*, *40*(3), 449–457.

Avram, F., Adenane, R., & Ketcheson, D. I. (2021). A review of matrix SIR Arino epidemic models. arXiv:2106.00384.

Baldi, S., Battistelli, G., Mari, D., Mosca, E., & Tesi, P. (2012). Multi-model unfalsified switching control of uncertain multivariable systems. *International Journal of Adaptive Control and Signal Processing*, *26*(8), 705–722.

Battistelli, G., Hespanha, J., & Tesi, P. (2012). Supervisory control of switched nonlinear systems. *International Journal of Adaptive Control and Signal Processing*, *26*(8), 723–738.

Bellman, R., Glicksberg, I., & Gross, O. (1956). On the "bang-bang" control problem. *Quarterly of Applied Mathematics*, *14*(1), 11–18.

Bin, M., Astolfi, D., Marconi, L., & Praly, L. (2018). About robustness of internal model-based control for linear and nonlinear systems. In *IEEE conference on decision and control* (pp. 5397–5402).

Bin, M., Bernard, P., & Marconi, L. (2021). Approximate nonlinear regulation via identification-based adaptive internal models. *IEEE Transactions on Automatic Control*, *66*(8), 3534–3549.

Bin, M., Cheung, P. Y. K., Crisostomi, E., Ferraro, P., Lhachemi, H., Murray-Smith, R., et al. (2021). Post-lockdown abatement of covid-19 by fast periodic switching. *PLoS Computational Biology*, *17*(1), Article e1008604.

Bin, M., & Marconi, L. (2020). Model identification and adaptive state observation for a class of nonlinear systems. *IEEE Transactions on Automatic Control (Early Access)*, http://dx.doi.org/10.1109/TAC.2020.3041238.

Bin, M., Marconi, L., & Teel, A. R. (2019). Adaptive output regulation for linear systems via discrete-time identifiers. *Automatica*, *105*, 422–432.

Bose, B. K. (1990). An adaptive hysteresis-band current control technique of a voltage-fed PWM inverter for machine drive system. *IEEE Transactions on Industrial Electronics*, *37*(5), 402–408.

Buso, S., Fasolo, S., Malesani, L., & Mattavelli, P. (2000). A dead-beat adaptive hysteresis current control. *IEEE Transactions on Industry Applications*, *36*(4), 1174–1180.

Cahlon, B., Schmidt, D., Shillor, M., & Zou, X. (1997). Analysis of thermostat models. *European Journal of Applied Mathematics*, *8*(5), 437–455.

Casas, E., Wachsmuth, D., & Wachsmuth, G. (2017). Sufficient second-order conditions for bang-bang control problems. *SIAM Journal on Control and Optimization*, *55*(5), 3066–3090.

Della Rossa, F., Salzano, D., Di Meglio, A., De Lellis, F., Coraggio, M., Calabrese, C., et al. (2020). A network model of Italy shows that intermittent regional strategies can alleviate the COVID-19 epidemic. *Nature Communications*, *11*(1), 5106.

Ferguson, N., Laydon, D., Nedjati Gilani, G., Imai, N., Ainslie, K., Baguelin, M., et al. (2020). *Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand: Report*, Imperial College London, URL http://spiral.imperial.ac.uk/handle/10044/1/77482.

Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., et al. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, *584*(7820), 257–261.

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

Giordano, G., Colaneri, M., Di Filippo, A., Blanchini, F., Bolzern, P., De Nicolao, G., et al. (2021). Modeling vaccination rollouts, SARS-CoV-2 variants and the requirement for non-pharmaceutical interventions in Italy. *Nature Medicine*, *27*(6), 993–998.

Goebel, R., Sanfelice, R. G., & Teel, A. R. (2012). *Hybrid dynamical systems*. Princeton, N. J.: Princeton University Press.

Guan, Y., He, W., Murugesan, M., Li, Q., Liu, P., & Li, L. K. B. (2019). Control of self-excited thermoacoustic oscillations using transient forcing, hysteresis and mode switching. *Combustion and Flame*, *202*, 262–275.

Gurevich, P., Jäger, W., & Skubachevskii, A. (2009). On periodicity of solutions for thermocontrol problems with hysteresis-type switches. *SIAM Journal on Mathematical Analysis*, *41*(2), 733–752.

Hespanha, J. P., Liberzon, D., & Morse, A. S. (2002). Supervision of integral-input-to-state stabilizing controllers. *Automatica*, *38*(8), 1327–1335.

Hespanha, J. a. P., Liberzon, D., & Morse, A. S. (2003). Hysteresis-based switching algorithms for supervisory control of uncertain systems. *Automatica*, *39*(2), 263–272.

Hespanha, J. P., & Morse, A. S. (1999). Certainty equivalence implies detectability. *Systems & Control Letters*, *36*(1), 1–13.

Hutson, M. (2020). The mess behind the models: Too many of the COVID-19 models led policymakers astray. Here's how tomorrow's models will get it right. *IEEE Spectrum*, *57*(10), 30–35.

Jin, H., & Safonov, M. (2012). Unfalsified adaptive control: Controller switching algorithms for nonmonotone cost functions. *International Journal of Adaptive Control and Signal Processing*, *26*(8), 692–704.

Karin, O., Bar-On, Y. M., Milo, T., Katzir, I., Mayo, A., Korem, Y., et al. (2020). Cyclic exit strategies to suppress COVID-19 and allow economic activity. http://dx.doi.org/10.1101/2020.04.04.20053579, medRxiv arXiv:2020.04.04.20053579.

Kawamura, A., & Hoft, R. (1984). Instantaneous feedback controlled PWM inverter with adaptive hysteresis. *IEEE Transactions on Industry Applications*, *IA-20*(4), 769–775.

Kennedy, D. M., Zambrano, G. J., Wang, Y., & Neto, O. P. (2020). Modeling the effects of intervention strategies on COVID-19 transmission dynamics. *Journal of Clinical Virology*, *128*, Article 104440.

Kosmatopoulos, E. B., & Ioannou, P. A. (1999). A switching adaptive controller for feedback linearizable systems. *IEEE Transactions on Automatic Control*, *44*(4), 742–750.

Krener, A. J. (1974). A generalization of Chow's theorem and the bang-bang theorem to nonlinear control problems. *SIAM Journal on Control and Optimization*, *12*(1), 43–52.

Ledzewicz, U., & Schättler, H. (2002). Optimal bang-bang controls for a two-compartment model in cancer chemotherapy. *Journal of Optimization Theory and Applications*, *114*(3), 609–637.

Li, Y., Ruan, X., Zhang, L., & Lo, Y.-K. (2020). Multipower-level hysteresis control for the class E DC–DC converters. *IEEE Transactions on Power Electronics*, *35*(5), 5279–5289.

Lim, C. W., Chung, T. Y., & Moon, S. J. (2003). Adaptive bang-bang control for the vibration control of structures under earthquakes. *Earthquake Engineering & Structural Dynamics*, *32*(13), 1977–1994.

Ma, H. (2008). Several algorithms for finite-model adaptive control: Partial answers to finite-model adaptive control problem. *Mathematics of Control, Signals, and Systems*, *20*(3), 271–303.

Mårtensson, B. (1985). The order of any stabilizing regulator is sufficient a priori information for adaptive stabilization. *Systems & Control Letters*, *6*(2), 87–91.

Middleton, R. H., Goodwin, G. C., Hill, D. J., & Mayne, D. Q. (1988). Design issues in adaptive control. *IEEE Transactions on Automatic Control*, *33*(1), 50–58.

Minyue Fu, & Barmish, B. (1986). Adaptive stabilization of linear systems via switching control. *IEEE Transactions on Automatic Control*, *31*(12), 1097–1103.

Mizel, V. J., & Seidman, T. I. (1997). An abstract bang-bang principle and time-optimal boundary control of the heat equation. *SIAM Journal on Control and Optimization*, *35*(4), 1204–1216.

Mohseni, M., & Islam, S. M. (2010). A new vector-based hysteresis current control scheme for three-phase PWM voltage-source inverters. *IEEE Transactions on Power Electronics*, *25*(9), 2299–2309.

Morato, M. M., Bastos, S. B., Cajueiro, D. O., & Normey-Rico, J. E. (2020). An optimal predictive control strategy for COVID-19 (SARS-CoV-2) social distancing policies in Brazil. *Annual Review of Control*, *50*, 417–431.

Morse, A. S. (1990). Towards a unified theory of parameter adaptive control: tunability. *IEEE Transactions on Automatic Control*, *35*(9), 1002–1012.

Morse, A. S. (1996). Supervisory control of families of linear set-point controllers - Part I. Exact matching. *IEEE Transactions on Automatic Control*, *41*(10), 1413–1431.

Morse, A. S. (1997). Supervisory control of families of linear set-point controllers. 2. Robustness. *IEEE Transactions on Automatic Control*, *42*(11), 1500–1515.

Morse, A. S., Mayne, D. Q., & Goodwin, G. C. (1992). Applications of hysteresis switching in parameter adaptive control. *IEEE Transactions on Automatic Control*, *37*(9), 1343–1354.

Morton, J. J. L., Tyryshkin, A. M., Ardavan, A., Benjamin, S. C., Porfyrakis, K., Lyon, S. A., et al. (2006). Bang–bang control of fullerene qubits using ultrafast phase gates. *Nature Physics*, *2*(1), 40–43.

Olsder, G. J. (2002). On open-and closed-loop bang-bang control in nonzero-sum differential games. *SIAM Journal on Control and Optimization*, *40*(4), 1087–1106.

Protter, M. H., & Morrey, C. B. J. (1985). *Intermediate calculus*. New York: Springer-Verlag.

Šabanovic, A. (2011). Variable structure systems with sliding modes in motion control—A survey. *IEEE Transactions on Industrial Informatics*, *7*(2), 212–223.

Sadeghi, M., Greene, J. M., & Sontag, E. D. (2021). Universal features of epidemic models under social distancing guidelines. *Annual Review of Control*, *51*, 426–440.

Slotine, J. J., & Sastry, S. S. (1983). Tracking control of non-linear systems using sliding surfaces, with application to robot manipulators. *International Journal of Control*, *38*(2), 465–492.

Sontag, E. D. (2021). An explicit formula for minimizing the infected peak in an SIR epidemic model when using a fixed number of complete lockdowns. http://dx.doi.org/10.1101/2021.04.11.21255289, arXiv:2021.04.11.21255289.

Stefanovic, M., & Safonov, M. G. (2008). Safe adaptive switching control: Stability and convergence. *IEEE Transactions on Automatic Control*, *53*(9), 2012–2021.

Stefanovic, M., Wang, R., & Safonov, M. G. (2004). Stability and convergence in adaptive systems. In *Proc. 2004 American control conference (Vol. 2)* (pp. 1923–1928).

Viola, L., & Lloyd, S. (1998). Dynamical suppression of decoherence in two-state quantum systems. *Physical Review A*, *58*(4), 2733.

Vu, L., & Liberzon, D. (2011). Supervisory control of uncertain linear time-varying systems. *IEEE Transactions on Automatic Control*, *56*(1), 27–42.

Weller, S. R., & Goodwin, G. C. (1994). Hysteresis switching adaptive control of linear multivariable systems. *IEEE Transactions on Automatic Control*, *39*(7), 1360–1375.

Ye, X. (2008). Nonlinear adaptive control using multiple identification models. *Systems & Control Letters*, *57*(7), 578–584.

Zandvliet, M., Bosgra, O., Jansen, J., Van den Hof, P., & Kraaijevanger, J. (2007). Bang-bang control and singular arcs in reservoir flooding. *Journal for Petroleum Science and Engineering*, *58*(1-2), 186–200.

Zhihong, M., Paplinski, A. P., & Wu, H. R. (1994). A robust MIMO terminal sliding mode control scheme for rigid robotic manipulators. *IEEE Transactions on Automatic Control*, *39*(12), 2464–2469.