Automatica 147 (2023) 110680

Contents lists available at ScienceDirect

Automatica

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

Optimal control of compartmental models: The exact solution[☆]

Franco Blanchini [a,*], Paolo Bolzern [b], Patrizio Colaneri [b,c], Giuseppe De Nicolao [d], Giulia Giordano [e]

[a] Dipartimento di Scienze Matematiche, Informatiche e Fisiche, Università degli Studi di Udine, Via delle Scienze 206, 33100 Udine, Italy
[b] Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milano, Italy
[c] IEIIT, CNR, Milano, Italy
[d] Department of Computer and System Science, Università degli Studi di Pavia, Italy
[e] Department of Industrial Engineering, University of Trento, Italy

ARTICLE INFO

Article history:
Received 1 January 2022
Received in revised form 13 June 2022
Accepted 1 September 2022
Available online 4 November 2022

Keywords:
Compartmental systems
Positive systems
Optimal control
Maximum principle
HJB equations

ABSTRACT

We formulate a control problem for positive compartmental systems formed by nodes (buffers) and arcs (flows). Our main result is that, on a finite horizon, we can solve the Pontryagin equations in one shot without resorting to trial and error via shooting. As expected, the solution is bang–bang and the switching times can be easily determined. We are also able to find a cost-to-go-function, in an analytic form, by solving a simple nonlinear differential equation. On an infinite horizon, we consider the Hamilton–Jacobi–Bellman theory and we show that the HJB equation can be solved exactly. Moreover, we show that the optimal solution is constant and the cost-to-go function is linear and copositive. This function is the solution of a nonlinear equation. We propose an iterative scheme for solving this equation, which converges in finite time. We also show that an exact solution can be found if there is a positive external disturbance affecting the process and the problem is formulated in a min sup framework. We finally provide illustrative examples related to flood control and epidemiology.

1. Introduction

A compartmental system is a positive system whose state variables are associated with nodes representing *reservoirs* (compartments), where resources are stored, and whose dynamics is due to *flows* circulating on the arcs among the nodes. In the linear case, the flow from one node to another is proportional to the amount of resource in the departure node. Deeply investigated and encountered in many fields (Jacquez & Simon, 1993), compartmental models are important in the control (Lenhart & Workman, 2007) and structural analysis (Blanchini & Giordano, 2021) of biological systems, in pharmacokinetics and toxicokinetics (Yeargers, Herod, & Shonkweiler, 2012), and in the analysis and the control of infection dynamics (Brauer & Castillo-Chavez, 2012; Lee & Leitmann, 1994; Martcheva, 2015; Morton & Wickwire, 1974; Rowthorn, Laxminarayan, & Gilligan, 2009; Sharomi & Malik, 2017). The COVID-19 pandemic has recently spurred renewed interest in the control of epidemic models (Arino, Brauer, van den Driessche, Watmough, & Wu, 2007; Bloem, Alpcan, & Basar, 2009; Bussell, Dangerfield, Gilligan, & Cunniffe, 2019; Diekmann & Heesterbeek, 2000; Forster & Gilligan, 2007; Gumel et al., 2004; Hansen & Day, 2011; Hethcote, 2000; Kermack & McKendrick, 1927), leading to countless new contributions, see for instance (Bin et al., 2021; Freddi, 2022; Hayhoe, Barreras, & Preciado, 2021; Köhler et al., 2021; Mandal et al., 2020); we refer to Alamo, Reina, Millán Gata, Preciado, and Giordano (2021) for a very recent survey.

We consider here the control problem in which some of the coefficients governing the flows among the nodes are control variables. This type of situation is encountered in controlled drug delivery and treatment planning, formulated as the control of a positive switched system by Colaneri, Middleton, Chen, Caporale, and Blanchini (2014), Hernandez-Vargas, Colaneri, Middleton, and Blanchini (2011) and Hernandez Vargas, Middleton, and Colaneri (2014) for HIV mitigation and then by Devia and Giordano (2019) and Giordano, Rantzer, and Jonsson (2015) for

☆ Work supported by MURST under PRIN 2017 'Monitoring and Control Underpinning the Energy-Aware Factory of the Future: Novel Methodologies and Industrial Validation' (ID 2017YKXYXJ). G.G. acknowledges support by the Strategic Grant MOSES at the University of Trento. F.B. acknowledges support by NextGenerationEU, European Union (Grant Uniud-DM737). The material in this paper was partially presented at the 60th IEEE Conference on Decision and Control, December 13–15, 2021, Austin, Texas, USA. This paper was recommended for publication in revised form by Associate Editor Martin Monnigmann under the direction of Editor Ian R. Petersen.
* Corresponding author.
E-mail addresses: blanchini@uniud.it (F. Blanchini), paolo.bolzern@polimi.it (P. Bolzern), patrizio.colaneri@polimi.it (P. Colaneri), giuseppe.denicolao@unipv.it (G. De Nicolao), giulia.giordano@unitn.it (G. Giordano).

https://doi.org/10.1016/j.automatica.2022.110680
0005-1098/

cancer therapies. Indeed, our problem could be framed in the context of positive linear controlled switching systems. These systems have several other applications besides those mentioned, including fluid networks, thermal systems and traffic control. We refer to Blanchini, Colaneri, and Valcher (2016) for a survey and more references on the topic of positive switching systems. The optimal control of compartmental models has been successfully applied to drug administration (Kusuoka et al., 1981) and controlled cancer chemotherapic treatment (Swierniak, Ledzewicz, & Schättler, 2003).

The control problem for positive switched systems is by no means an easy one, but it can be efficiently solved under particular assumptions. The optimal control problem on a finite horizon can be solved via dynamic programming methods (Hernandez-Vargas et al., 2011) and, as long as the switching controlled coefficients are on the diagonal of the system matrix, the control reduces to a convex optimization problem (Colaneri et al., 2014; Rantzer & Bernhardsson, 2014). Unfortunately, this result does not apply in the case of switched compartmental systems with controlled arcs, because coefficients outside the diagonal are modified by the control.

We formulate the control problem for linear controlled compartmental systems with a linear cost on both finite and infinite horizon. In general, Pontryagin theory requires a shooting approach to solve the state equations, forward in time, and the co-state equations, backward in time. The convergence of the shooting scheme is a critical issue. The dynamic programming approach is even harder, since it requires the solution of the HJB equation, which, in general, involves a numerical brute-force scheme. We show that, quite surprisingly, for controlled compartmental systems, both problems can be solved efficiently via Pontryagin theory and Hamilton–Jacobi–Bellman theory. The contributions of the manuscript are summarized next.

- In the finite horizon problem, the state and co-state equations are decoupled, and the Hamiltonian minimizer control depends on the co-state only. This means that the co-state equation is integrated just once, thus producing directly the optimal control input.
- The optimal control does not depend on the initial state.
- Pontryagin equations give necessary conditions for optimality. To support the theory, we formulate the HJB equation on a finite horizon, which provides sufficient conditions. The solution is the same achieved via the Pontryagin maximum principle, which ensures the optimality of the solution. The control is switching.
- On an infinite horizon, we consider the HJB equation and we show that it reduces to an algebraic equation. The derived cost-to-go function is linear co-positive. Hence, the optimal control is constant in the long run.
- The proposed algebraic equation may have multiple solutions. Yet, we prove that there is a single stabilizing one (the only one that can be accepted). Our constructive proof provides an iterative scheme that converges to the stabilizing solution in a finite number of steps.
- When the system is affected by an additive positive external disturbance, a similar approach is proposed to solve a min sup control problem.
- The HJB approach requires stabilizability. We show how to check this assumption by solving the problem of minimizing the Frobenius eigenvalue.

We apply our results on two illustrative examples. The first is a flood control problem with 6 states. In the second, inspired by the COVID-19 pandemic, we analyze the sensitivity of the IDART model (Giordano et al., 2020), having 5 state variables, with 16 uncertain parameters treated as "control variables", and we determine minimum and maximum values of relevant variables via a minimization/maximization problem.

A very preliminary version of these results has been presented at a conference (Blanchini, Bolzern, Colaneri, De Nicolao, & Giordano, 2021).

1.1. Notation

Given a vector $v = [v_1, v_2, \dots, v_m] \in \mathbb{R}^m$, we denote

$$\mathrm{diag}(v) = \begin{bmatrix} v_1 & 0 & \dots & 0 \\ 0 & v_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & v_m \end{bmatrix} \text{ and}$$

$$\mathrm{sign}(v) = \begin{bmatrix} \mathrm{sign}(v_1) \\ \mathrm{sign}(v_2) \\ \vdots \\ \mathrm{sign}(v_m) \end{bmatrix},$$

where

$$\mathrm{sign}(v_i) = \begin{cases} 1 & \text{if} \quad v_i \geq 0 \\ -1 & \text{if} \quad v_i < 0 \end{cases}$$

The inequality $u < v$ ($u \leq v$) has to be interpreted componentwise. We write $|U| = I$ to mean that U belongs to the set

$$\{U : |U| = I\} \doteq \{\mathrm{diag}(u_1, u_2, \dots, u_m), \quad |u_k| = 1\}$$

and $|U| \leq I$ to mean that U belongs to the set

$$\{U : |U| \leq I\} \doteq \{\mathrm{diag}(u_1, u_2, \dots, u_m), \quad |u_k| \leq 1\}.$$

A square matrix M is *Metzler* if $M_{ij} \geq 0$ for $i \neq j$ and it is *irreducible* if there is no variable permutation that brings M in a block-triangular form.

2. Model description

The class of models we consider can be written as

$$\dot{x}(t) = \tilde{A}x(t) + \tilde{F}V(t)\tilde{G}x(t), \tag{1}$$

where the state $x(t) \in \mathbb{R}^n$ represents the amount of resource stored at the nodes (compartments) and the diagonal matrix $V(t)$ of control variables is associated with the flows along some of the arcs connecting the nodes. For example, in a fluid network (cf. Section 8.1), each state component represents the fluid level in a reservoir, while the arc control variables represent the opening of the valves that regulate the flow along some of the pipes connecting the reservoirs.

We consider the following assumptions.

Assumption 1. Matrix $V(t) \in \mathbb{R}^{m\times m}$ is diagonal and has elements $v_i^- \leq v_i(t) \leq v_i^+$.

Assumption 2. Matrix $\tilde{A} + \tilde{F}V\tilde{G} \in \mathbb{R}^{n\times n}$ is Metzler for all choices of $v_i \in [v_i^-, v_i^+]$, $i = 1, \dots, m$. Matrix $\tilde{G} \in \mathbb{R}^{m\times n}$ is nonnegative.

Assumption 3. The initial state $x(0) = x_0$ is nonnegative.

We consider the problem of minimizing the positive cost

$$J = h^\top x(T) + \int_0^T \tilde{\ell}^\top x(t)dt + \int_0^T r^\top V(t)\tilde{G}x(t)dt, \tag{2}$$

with nonnegative weight vectors h, $\tilde{\ell}$ and r. The first two terms of the cost are linear functions of the state, while the last term is jointly bilinear in the state and the control. This cost penalizes both the presence of fixed assets (for instance, fluid) that remain stored at the nodes over time, weighted by $\tilde{\ell}$, and the controlled flows $V(t)\tilde{G}x(t)$ representing the control effort, weighted by r.

To simplify the notation, we can parameterize matrix V, with bounds $V^- \leq V \leq V^+$, as

$$V = \frac{V^- + V^+}{2} + U\frac{V^+ - V^-}{2},$$

with

$$|U| \leq I \quad (\text{i.e., } U = \text{diag}[u],\ |u| \leq 1). \tag{3}$$

Adding the constant part to $\tilde{A}$ and scaling $\tilde{G}$ yields

$$\tilde{A} + \tilde{F}V\tilde{G} = \tilde{A} + \tilde{F}\frac{V^- + V^+}{2}\tilde{G} + \tilde{F}U\frac{V^+ - V^-}{2}\tilde{G} \doteq A + FUG,$$

by defining the new matrices $A \doteq \tilde{A} + \tilde{F}(V^- + V^+)\tilde{G}/2$, $F \doteq \tilde{F}$ and $G \doteq (V^+ - V^-)\tilde{G}/2$.

Henceforth, we thus consider a model of the form

$$\dot{x}(t) = Ax(t) + FU(t)Gx(t), \quad |U| \leq I. \tag{4}$$

Note that the cost (2) remains of the same form

$$J = h^\top x(T) + \int_0^T \ell^\top x(t)dt + \int_0^T r^\top U(t)Gx(t)dt, \tag{5}$$

with $\ell^\top = \tilde{\ell}^\top + r^\top(V^- + V^+)\tilde{G}/2$, and the new matrix G is still nonnegative.

Remark 1. System (4)–(5) encompasses (1)–(2) and Assumptions 1–3 remain valid: we just assume $-v^- = v^+ = \bar{1}$, the all-ones vector. The change of variables is useful to make the final formulas much more compact.

Positivity of the cost implies that, in the new variable U, the following assumption holds.

Assumption 4. For all $|U| \leq I$, it holds componentwise that

$$\ell^\top + r^\top UG \geq 0.$$

We will reconsider this assumption later on.

We consider the control problem with a finite horizon $T < \infty$ and with an infinite horizon $T = \infty$; in the latter case, $h = 0$.

3. Finite horizon control

The main result of this section shows that we can solve the finite horizon problem directly with a single integration of the co-state equation, without resorting to shooting.

The Hamiltonian is

$$H(x, \xi, u) = \xi^\top [A + FUG]x + r^\top UGx + \ell^\top x.$$

The minimum is achieved componentwise by solving

$$\min_{|U|\leq I} \left[\xi^\top FU + r^\top U\right] Gx. \tag{6}$$

Lemma 1. *The minimizer in (6) is given by*

$$U^* = diag\left(u^*\right) = diag\left(-sign[F^\top \xi + r]\right). \tag{7}$$

The maximizer is $U^* = diag\left(sign[F^\top \xi + r]\right)$.

Proof. The minimizer U^* is found by considering only the part depending on U: $[\xi^\top F + r^\top]UGx$. Since Gx is nonnegative by assumption, we immediately have (7). The same proof holds for the maximizer. ■

Remark 2. In view of the definition of the sign function, U^* is uniquely defined.

Since $u^*(\xi) = -\text{sign}\left[F^\top \xi + r\right]$ does not depend on x, the adjoint equation is

$$-\dot{\xi}(t)^\top = \xi^\top\left[A + FU^*(t)G\right] + r^\top U^*(t)G + \ell^\top \tag{8}$$

$$U^* = \text{diag}\left[u^*(\xi)\right] \tag{9}$$

$$\xi(T) = h \tag{10}$$

which has to be solved backward in time and can be solved *independently* of x.

Proposition 1. *Let Assumptions 1–4 be satisfied. Then the optimal control function u^* can be computed by means of a single integration of (8)–(9), backward in time, with final condition (10). The optimal control is independent of $x(0)$.*

The proof is given in the next subsection.

The same property holds if we wish to maximize the cost function: we just need to replace $u^*(\xi) = -\text{sign}\left[F^\top\xi + r\right]$ by $u^*(\xi) = \text{sign}\left[F^\top\xi + r\right]$.

Example 1. Consider a system of the form (1) with

$$\tilde{A} = 0, \quad \tilde{F}V(\alpha, \beta) = \begin{bmatrix} -\alpha & 0 \\ \alpha & -\beta \end{bmatrix}, \quad \tilde{G} = I.$$

With $T = 1$, take the parameters $\alpha \in [1, 4]$ and $\beta \in [2, 3]$ as control variables and assume the cost weights $h^\top = [0\ 1]$, $r = 0$ and $\tilde{\ell} = 0$ in (2). The differential Eq. (8) is

$$-\begin{bmatrix} \dot{\xi}_1 & \dot{\xi}_2 \end{bmatrix} = \begin{bmatrix} \xi_1 & \xi_2 \end{bmatrix} \times \left\{ \underbrace{\begin{bmatrix} -\frac{5}{2} & 0 \\ \frac{5}{2} & -\frac{5}{2} \end{bmatrix}}_{A} + \underbrace{\begin{bmatrix} -1 & 0 \\ 1 & -1 \end{bmatrix}}_{F} \underbrace{\begin{bmatrix} u_1 & 0 \\ 0 & u_2 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \frac{3}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}}_{G} \right\},$$

with $\begin{bmatrix} \xi_1(T) & \xi_2(T) \end{bmatrix}^\top = \begin{bmatrix} 0 & 1 \end{bmatrix}^\top$

If we take $u^* = -\text{sign}[\xi^\top F]$, we obtain the minimizer functions α, β in Fig. 1, upper panel. To maximize, we take $u^* = \text{sign}[\xi^\top F]$, leading to the maximizer functions α, β in Fig. 1, middle panel. While β is constant, α switches at some point. The lower panel of Fig. 1 reports the bounds on the state evolution and a set of randomly generated curves, obtained by extracting random constant values of parameters $\alpha \in [1, 4]$ and $\beta \in [2, 3]$ from uniform distributions. It is worth stressing that the bounds are valid only at the final time $t = T = 1$, as can be seen in Fig. 1 (lower panel).

To achieve bounds valid at an intermediate time $0 < T' < T$, one should re-run the procedure with the new final time T' and the curves would be different on $[0, T']$. So, to have an overall bounding function at intermediate times, we just need to iterate over different values of the final time: the computation is so fast that this goal is straightforward.

The example shows that the maximum and the minimum are not achieved for constant values of the control parameters on a finite horizon. This is not the case for infinite horizons, as we will see later.

Pontryagin's maximum principle provides a convenient theoretical framework to solve the problem. Unfortunately, in general, it provides necessary conditions only (Lewis, Vrabie, & Syrmos, 2012). To show that these conditions are also sufficient, we exploit the Hamilton–Jacobi–Bellman theory.

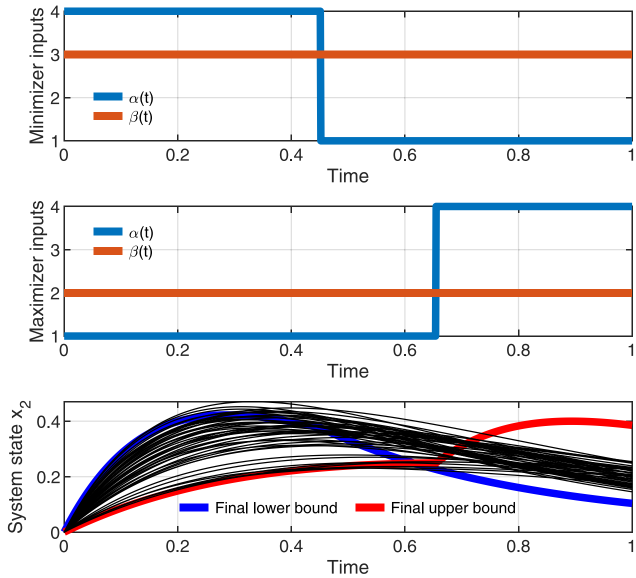

Fig. 1. Time evolution of: the minimizer functions $\alpha(t)$, blue, and $\beta(t)$, orange (upper panel); the maximizer functions $\alpha(t)$, blue, and $\beta(t)$, orange (middle panel); curves generating the lower- (blue) and upper- (red) bounding values for $x_2(T)$, along with a set of randomly generated (black) curves (lower panel). It is important to note that the bounds only hold at the final time $T = 1$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.1. The finite-time Hamilton–Jacobi-Bellman equation

Sufficient conditions can be derived by considering the Hamilton–Jacobi–Bellman equation on the finite horizon $[0, T]$, which turns out to be

$$\min_{|U|\le I}\left\{\frac{\partial\Psi(x,t)}{\partial t}+\nabla\Psi(x)\,[A+FUG]\,x+\ell^\top x+r^\top UGx\right\}$$
$$=0. \tag{11}$$

For this equation, we try a solution of the form

$$\Psi(x,t)=z^\top(t)x. \tag{12}$$

By substituting (12) into (11), we get

$$\min_{|U|\le I}\left[\dot z^\top(t)+z^\top(t)\,[A+FUG]+\ell^\top+r^\top UG\right]x=0. \tag{13}$$

From Lemma 1, the minimizer is $u^*=-\text{sign}[F^\top z(t)+r]$. Setting $U^*=\text{diag}[u^*]$ in (13) yields the following ordinary differential equation, with final condition $h^\top$:

$$\dot z^\top(t)+z^\top(t)A-|z^\top(t)F+r^\top|G+\ell^\top=0 \tag{14}$$
$$z^\top(T)=h^\top \tag{15}$$

Eq. (14) is nonlinear. However, it is Lipschitz, hence it admits a unique solution $z(t)$. We are therefore in a position to prove the following result.

Theorem 1. *Let Assumptions 1–4 be satisfied. The cost-to-go function of the finite-horizon optimal control problem is given by (12), where $z(t)$ solves (14)–(15). The optimal control is given by (7) with $\xi = z$. The optimal cost is $J_{opt} = z^\top(0)x(0)$.*

Proof. The HJB theory provides sufficient optimality conditions, while the solution of (14)–(15) exists and it is uniquely defined, yielding $\Psi = z^\top(t)x$. ■

Note that the solution $z(t)$ of (14)–(15) is the same we get with (8)–(10) if we consider the corresponding, uniquely defined, $U^*(t)=\text{diag}[\text{sign}(\xi^\top F+r^\top)]$.

3.2. Model generalization

The model can be generalized to the case in which the Metzler matrix $A(t)$ and the nonnegative matrix $G(t)$ are time-varying, also in the presence of a linear control term $Ec(t)$ and of a nonnegative drift term $d(t)\ge 0$:

$$\dot x(t)=A(t)x(t)+FU(t)G(t)x(t)+Ec(t)+d(t), \tag{16}$$

where matrix E is given and c is a control vector.

Assumption 5. Vector $c(t)$ has elements $-1\le c_i(t)\le 1$ and $Ec(t)+d(t)\ge 0$ for all $c(t)$.

Again, normalizing to unitary lower and upper bounds is not a restriction, since we can scale E and absorb any constant term into d.

Remark 3. If, in the original model (1), the bounds on $V(t)$ are time-varying, $v_i^-(t)\le v_i(t)\le v_i^+(t)$, we can still reduce the analysis to the constant bound case

$$\dot x(t)=A(t)x(t)+FU(t)G(t)x(t),\quad \text{with}\ \ |U(t)|\le I,$$

with $A(t)$ Metzler, $G(t)$ nonnegative as we did in Section 2.

We consider the problem of minimizing the positive cost

$$J=h^\top x(T)+\int_0^T[\ell^\top+r^\top U(t)G]x(t)dt+\int_0^T s^\top c(t)dt, \tag{17}$$

with the additional term $s^\top c(t)$, where s is a nonnegative vector. The new Hamiltonian is

$$H(x,\xi,u,c,d,t)=\xi^\top[A(t)x+FUG(t)x+Ec+d]$$
$$+\,r^\top U(t)G(t)x+\ell^\top x+s^\top c.$$

As long as $x(t)\ge 0$, the minimizers of the expression are:

$$u^*(\xi)=-\text{sign}\left[F^\top\xi+r\right],$$
$$c^*(\xi)=-\text{sign}\left[E^\top\xi+s\right],$$

while the adjoint equation is the same as before:

$$-\dot\xi(t)^\top=\xi^\top\left[A(t)+FU^*(t)G(t)+r^\top U^*(t)G(t)+\ell^\top\right] \tag{18}$$
$$U^*=\text{diag}\left[u^*(\xi)\right] \tag{19}$$
$$\xi(T)=h \tag{20}$$

Again, this equation does not depend on x, and hence no shooting is required. However, since variable $x(t)$ is assumed nonnegative, the control is optimal as long as $x(t)\ge 0$.

Proposition 2. *Under Assumptions 1–5, the control functions u^* and c^* can be computed by means of a single integration of (18)–(20), backward in time. The obtained control is independent of $x(0)$ and it is optimal as long as we have $x(t)\ge 0$ on $[0, T]$.*

Conditions (18)–(20) are quite convenient for computation, but they are only necessary in principle. As done before, to guarantee that the achieved solution is indeed the optimal one, we resort to the Hamilton–Jacobi–Bellman theory, which gives sufficient conditions, and leads to the same solution.

4. Infinite horizon problem

We now consider system (4) and wish to solve the infinite horizon problem of minimizing the cost

$$J=\int_0^\infty \ell^\top x(t)dt+\int_0^\infty r^\top U(t)Gx(t)dt, \tag{21}$$

under Assumptions 1–4.

For infinite horizons, we need a stabilizability assumption.

Assumption 6. There exists a matrix $\bar{U}$ such that $A + F\bar{U}G$ is Hurwitz.

We consider again the Hamilton–Jacobi–Bellman theory. Since we assume a nonnegative initial condition, we consider a cost-to-go function Ψ defined in the positive orthant. The infinite-horizon HJB equation is

$$\min_{|U|\le I} \left\{\nabla\Psi(x)\,[A+FUG]\,x+\ell^\top x+r^\top UGx\right\}=0. \tag{22}$$

With an abuse of notation, we consider $\nabla\Psi$ as the gradient of Ψ, which is not necessarily differentiable. The utmost regularity we can ensure to Ψ is its concavity.

Proposition 3. *Under Assumptions 1–4 and 6, Ψ is concave and copositively homogeneous of order 1: for $\lambda \ge 0$, $\Psi(\lambda x) = \lambda\Psi(x)$.*

Proof. It goes as in Hernandez-Vargas et al. (2011), where the proof is given for discrete-time systems. Let x_0 be a convex combination of x_A and x_B, $x_0 = \alpha x_A + \beta x_B$, $\alpha + \beta = 1$, $\alpha, \beta \ge 0$. Denote by $J(x, u)$ the cost with initial condition x and control u and denote by $J^*(x_0, u_0^*)$, $J^*(x_A, u_A^*)$, $J^*(x_B, u_B^*)$ the optimal costs, with the corresponding optimal controls u_0^*, u_A^*, u_B^*. Then

$$\begin{aligned} J^*(x_0,u^*) &= \alpha J(x_A,u^*)+\beta J(x_B,u^*)\\ &\ge \alpha J(x_A,u_A^*)+\beta J(x_B,u_B^*), \end{aligned}$$

where the first equality is due to the linearity of the cost for fixed u^*. The fact that Ψ is copositively homogeneous of order 1 is immediate. ■

Concavity ensures that the gradient $\nabla\Psi(x)$ is defined almost everywhere. Consider a point x where

$$z^\top = \nabla\Psi(x)$$

is defined and let us study it locally, relying on the next important lemma, whose proof is immediate.

Lemma 2. *The minimum in (22) is obtained on the vertices*

$$\begin{aligned} U^*(x) &= \arg\min_{|U|\le I}\left\{z^\top [A+FUG]\,x+\ell^\top x+r^\top UGx\right\}\\ &= \arg\min_{|\hat U|=I}\left\{z^\top \left[A+F\hat U G\right]x+\ell^\top x+r^\top UGx\right\} \end{aligned}$$

Letting U^* denote the optimal control, the HJB equation becomes an algebraic equation

$$z^\top\left[A+FU^*G\right]x+\ell^\top x+r^\top U^*Gx=0, \tag{23}$$

which is valid for $z^\top = \nabla\Psi(x)$, namely for the specifically chosen x.

Motivated by these considerations we wonder whether the quantity z can be constant for all x; equivalently, we look for a solution of the form

$$\Psi(x) = z^\top x,$$

with a common $z \ge 0$. By eliminating x, we get

$$z^\top\left[A+FU^*G\right]+\ell^\top+r^\top U^*G=0. \tag{24}$$

Substituting the expression (7) of U^* into Eq. (24) yields

$$z^\top A-\left|z^\top F+r^\top\right|G+\ell^\top=0. \tag{25}$$

To find the solution of (25) we proceed as follows.

Theorem 2. *Let Assumptions 1–4 and 6 be satisfied. Assume that Eq. (25) admits a single solution $z_{opt}^\top > 0$, and that $u^* = -\mathrm{sign}[F^\top z_{opt} + r]$ is stabilizing, namely, $A + FU^*G$ is Hurwitz with $U^* = diag[u^*]$. Then the function*

$$\Psi(x) = z_{opt}^\top x$$

satisfies the HJB equation and control (7) is optimal with the constant control law U^*.

The theorem has to be completed by showing the existence and uniqueness of the solution, as we discuss in the next section.

Remark 4. To solve the maximization problem, we just need to consider $U^* = \mathrm{sign}[F^\top z_{opt} + r]$, where z_{opt} now solves

$$z^\top A+\left|z^\top F+r^\top\right|G+\ell^\top=0. \tag{26}$$

Clearly, this solution has no practical significance if U^* destabilizes the system: hence, in the infinite horizon case, we need to assume that $A + FUG$ is Hurwitz for all $|U| \le I$. Note that this new assumption is limited to the claim of this remark and needed nowhere else.

5. Uniqueness and existence of a stabilizing solution

We now tackle the issue of existence and uniqueness of the solution of Eq. (25). We first consider the problem of the existence of a vector z along with a stabilizing U^*. To keep the presentation simple we strengthen Assumption 4 as follows; we will comment on this aspect later on.

Assumption 7. For all $|U| \le I$, componentwise,

$$\ell^\top + r^\top UG > 0.$$

To find a solution to (24), we propose a procedure described as pseudo-code that generates sequences $z_k > 0$, $k = 1, 2, 3\dots$ and U_k^*, $k = 1, 2\dots$ converging to the solution.

Procedure 1. *Inputs: [r, ℓ, A, F, G, U_0^*], with U_0^* stabilizing.*

Step 0. *Check the stability of $[A + FU_0^*G]$ and compute*

$$z_1^\top = -\left[\ell^\top+r^\top U_0^*G\right]\left[A+FU_0^*G\right]^{-1};$$

(where $z_1^\top$ is the solution $z^\top$ to (24) with $U^* = U_0^*$.)

Set $k := 1$.

Step 1. *Compute*

$$U_k^* = -diag\{sign[r^\top + z_k^\top F]\}; \tag{27}$$

(i.e., U_k^* is the minimizer $\arg\min_{|U|\le I}[r^\top + z_k^\top F]UGx$.)

Step 2. *Compute the solution $z_{k+1}^\top > 0$ to the linear equation*

$$z_{k+1}^\top[A+FU_k^*G]+r^\top U_k^*G+\ell^\top=0. \tag{28}$$

Step 3. *IF $z_{k+1} = z_k$ STOP and provide as Output: $z^\top = z_{k+1}^\top$ and $U^* = U_k^*$, ELSE set $k := k + 1$ and GOTO Step 1.*

Theorem 3. *Under Assumptions 1–3, 6 and 7, the proposed Procedure 1 converges to the solution $z^\top > 0$ to Eq. (24), and provides a stabilizing U^*.*

Proof. We need three steps.

Step (a) We first show that Eq. (28) can be solved for all k to find $z_{k+1}^\top$, which is positive, because the considered matrix $A + FU_k^*G$ is Hurwitz and Assumption 7 holds, and therefore $-(r^\top U_k^*G + \ell^\top)[A + FU_k^*G]^{-1} > 0$.

Matrix $A + FU_0^*G$ is Hurwitz because U_0^* is stabilizing. Hence there exists $z_1^\top > 0$ solving the equation. For $k \ge 1$, given U_k^* computed as in (27), and assuming $[A + FU_k^*G]$ Hurwitz, so that we can compute $z_{k+1}^\top > 0$ from (28), we show that $[A + FU_{k+1}^* G]$ is Hurwitz as well. Indeed

$$z_{k+1}^\top[A + FU_{k+1}^*G] = z_{k+1}^\top[A + FU_k^*G]$$
$$- z_{k+1}^\top F[U_k^* - U_{k+1}^*]G = -\ell^\top - r^\top U_k^* G$$
$$- z_{k+1}^\top F[U_k^* - U_{k+1}^*]G \pm r^\top U_{k+1}^* G$$
$$= \underbrace{-\ell^\top - r^\top U_{k+1}^* G}_{<0\ (\text{Assumption 7})} \underbrace{-[r^\top + z_{k+1}^\top F][U_k^* - U_{k+1}^*]}_{\leq 0\ (U_{k+1}^*\ \text{minimizer})}$$

This implies that $A + FU_{k+1}^* G$ is Hurwitz, hence a positive $z_{k+1}^\top$ can be found. This proves that matrices $A + FU_k^* G$ recursively generated by the procedure are Hurwitz and vectors z_k are positive.

Step (b) We now prove that the positive sequence is nonincreasing: $z_{k+1} \leq z_k$ for all k.

$$[z_k^\top - z_{k+1}^\top][A + FU_k^*G] = z_k^\top[A + FU_k^*G] + r^\top U_k^* G + \ell^\top$$
$$\pm r^\top U_{k-1}^* G \pm z_k^\top FU_{k-1}^* G = \underbrace{[z_k^\top F + r^\top][U_k^* - U_{k-1}^*]G}_{\leq 0\ (U_k^*\ \text{minimizer})}$$
$$+ \underbrace{z_k^\top[A + FU_{k-1}^*G] + \ell^\top + r^\top U_{k-1}^* G}_{=0,\ \text{because of (28)}} \leq 0$$

Since $[z_k^\top - z_{k+1}^\top][A + FU_k^*G] \leq 0$, we multiply by the nonnegative $-[A + FU_k^*G]^{-1}$, preserving the inequality, to have

$$[z_k^\top - z_{k+1}^\top][A + FU_k^*G][-(A + FU_k^*G)^{-1}]$$
$$= -[z_k^\top - z_{k+1}^\top] \leq 0$$

Step (c) Since $z_k > 0$ decreases, it has a limit $\bar{z} \geq 0$. Now we need to prove that $\bar{z} > 0$, strictly (in principle, it might have some zero components). We have to remind that U_k^* is always assumed on the vertices, therefore there are finitely many possible solutions $z_k^\top$ of (28). Therefore the sequence $z_k^\top$ takes values in a finite set.

As a consequence, convergence occurs for finite k, i.e. there exists k such that $z_k^\top = z_{k+1}^\top = \bar{z}^\top$. On the other hand $z_k > 0$ for all k hence the limit is positive, $\bar{z} > 0$. ∎

Remark 5. In principle one could consider all possible solutions to (28) for U on the vertices in a combinatorial way and find a stabilizing one. It turns out that Procedure 1 converges quickly to the right solution in a finite number of steps and is much more efficient than the combinatorial approach.

The procedure converges to some $\bar{z}$ that might depend on the initial choice of U_0^*. However, we can prove uniqueness, regardless of U_0^*.

Theorem 4. *Under Assumptions 1–3, 6 and 7, there cannot be two distinct positive solutions z_a and z_b with the property that $u_a = -sign[z_a^\top F + r^\top]$ and $u_b = -sign[z_b^\top F + r^\top]$ are both stabilizing.*

Proof. By contradiction, consider the corresponding U_a^* and U_b^* and assume both

$$z_a^\top \left[A + FU_a^*G\right] x + \ell^\top x + r^\top U_a^* G x = 0,$$
$$z_b^\top \left[A + FU_b^*G\right] x + \ell^\top x + r^\top U_b^* G x = 0.$$

Then

$$\begin{aligned} &(z_a^\top - z_b^\top)[A + FU_a^*G] \\ = &\quad z_a^\top[A + FU_a^*]G - z_b^\top[A + FU_b^*]G \\ &+ z_b^\top[A + FU_b^*]G - z_b^\top[A + FU_a^*]G \\ = &-r^\top U_a^* G + r^\top U_b^* G + z_b^\top FU_b^* G - z_b^\top FU_a^* G \\ = &-(r^\top + z_b F)[U_a^* - U_b^*]G \leq 0, \end{aligned}$$

where the last inequality holds because U_b^* is the minimizer. Again, since $(z_a^\top - z_b^\top)[A + FU_a^*G] \leq 0$, we multiply by the nonnegative $-[A + FU_a^*G]^{-1}$, to get $z_b^\top \leq z_a^\top$. With the same approach, we can show the opposite inequality, $z_a^\top \leq z_b^\top$. Hence, it must be $z_b^\top = z_a^\top$. ∎

Remark 6. Assumption 7, a stronger version of Assumption 4, is fundamental to ensure $z^\top > 0$ and the overall system stabilization. This issue is similar to the one with LQ control, when the control may be non-stabilizing if the state cost is assumed positive semi-definite. A possible relaxation is that the solutions to (28) are positive for all $|U| = I$: this can happen even if we weakly assume nonnegativity of the cost as in Assumption 4.

6. Optimality with an external disturbance

Consider the system with an additive external disturbance:

$$\dot{x}(t) = Ax(t) + FU(t)Gx(t) + Bw(t), \tag{29}$$

with B a nonnegative matrix and vector $w(t)$ a nonnegative external disturbance. We introduce the objective function

$$J = \frac{\int_0^\infty [\ell^\top x(t) + r^\top U(t)Gx(t)]dt}{\int_0^\infty \mathbf{1}^\top w(t)dt} \tag{30}$$

and we aim to solve $\min_{|U|\leq I} \sup_w J$, for some stabilizing U. The ratio in (30) is an input–output amplification measure that weighs the performance integral with the disturbance amplitude. For instance, in a flood problem, such as the one considered in Section 8.1, the numerator could be the weight of a persistent incoming flow (e.g., rain), instead of a flooding action concentrated at $t = 0$. We derive the next result.

Theorem 5. *Let Assumptions 1–3, 6 and 7 be satisfied. Assume $x(0) = 0$ and $w(t) \geq 0$, $w \in \mathcal{L}_1$. Then*

$$\min_{|U(t)|\leq I} \sup_w J = \max_k \ \hat{q}^\top B e_k, \tag{31}$$

where we denote by e_k the kth vector of the canonical basis, while vector $\hat{q} > 0$ solves the equation

$$\hat{q}^\top A - |\hat{q}^\top F + r^\top|G + \ell^\top = 0. \tag{32}$$

The optimal control is constant,

$$U^* = diag\left[-sign[\hat{q}^\top F + r^\top]\right], \tag{33}$$

and the worst-case disturbance is

$$w^* = \delta(t)e_\nu,$$

where $\delta(t)$ is the Dirac function and

$$\nu = \arg\max_k[\hat{q}^\top B e_k].$$

Proof. Consider the function $V(x) = q^\top x$, $q > 0$. We have

$$\begin{aligned} \dot{V} + (\ell^\top + r^\top UG)x = &\left(q^\top A - |q^\top F + r^\top|G + \ell^\top\right)x \\ &+ |q^\top F + r^\top|(I - U^*U)Gx + q^\top Bw \end{aligned}$$

for all $w \geq 0$ and $|U| \leq 1$, where we set

$$U^* = -\text{diag}\ [\text{sign}(q^\top F + r^\top)].$$

Take $\hat{q} > 0$ that solves (32), to annihilate the term in round brackets and get

$$\dot{V} + (\ell^\top + r^\top UG)x - \hat{q}^\top Bw = |\hat{q}^\top F + r^\top|(I - U^*U)Gx \geq 0.$$

Note that $(I - U^*U) \geq 0$ and the inequality becomes an equality for $U = U^*$. Now integrate, divide by $\int_0^\infty \mathbf{1}^\top w(t)dt$ and consider that $x(0) = 0$, so $V(x(0)) = V(x(\infty)) = 0$ and

$$\frac{\int_0^\infty (\ell^\top + r^\top U(t)G)x(t)dt}{\int_0^\infty \mathbf{1}^\top w(t)dt} \geq \frac{\int_0^\infty \hat{q}^\top Bw(t)dt}{\int_0^\infty \mathbf{1}^\top w(t)dt}, \ \forall w \in \mathscr{L}_1, \qquad (34)$$

where the equality holds if and only if $U(t) \equiv U^*$.

The fraction in the right-hand side of (34) is maximized for

$$w^*(t) = \delta(t)\, e_\nu,$$

where ν is the index corresponding to the largest component $[\hat{q}^\top B]_k$ of $\hat{q}^\top B$, formally $\nu = \arg\max_k \hat{q}^\top Be_k$. Therefore

$$\inf_{|U(t)|\leq 1} \sup_{w\in\mathscr{L}_1} \frac{\int_0^\infty (\ell^\top + r^\top U(t)G)x(t)dt}{\int_0^\infty \mathbf{1}^\top w(t)dt} = \gamma^* = \hat{q}^\top Be_\nu,$$

where $\hat{q}$ satisfies Eq. (32) and the infimum is achieved for $U(t) \equiv U^*$. ■

Remark 7. Consistently with the results of the previous sections, the optimal control is constant, $U = U^*$, and does not depend on B. Conversely, the worst case disturbance $w^* = \delta(t)e_\nu$ does depend on B.

7. Minimization of the Frobenius eigenvalue

This section considers the problem of minimizing, over $|U| \leq I$, the Frobenius dominant eigenvalue $\lambda^F(A + FUG)$ of matrix $A + FUG$, namely the eigenvalue with maximum real part, which is real because $A + FUG$ is a Metzler matrix. Besides being of interest on its own, the solution to this problem is useful to check Assumption 6 and to find a starting U_0 in Procedure 1.

To this aim, consider the functions $p^-(\lambda)$ and $p^+(\lambda)$ of a real variable λ

$$\begin{cases} p^-(\lambda) = \min_{|U|\leq I} \det[\lambda I - A - FUG], \\ p^+(\lambda) = \max_{|U|\leq I} \det[\lambda I - A - FUG], \end{cases} \qquad (35)$$

which are tight lower and upper bounds for the characteristic polynomial

$$p(\lambda, U) \doteq \det[\lambda I - A - FUG];$$

precisely,

$$p^-(\lambda) \leq p(\lambda, U) \leq p^+(\lambda).$$

We notice that $\lim_{\lambda\to\infty} p^+(\lambda) = +\infty$. So let λ^* be defined as the largest real root of function $p^+(\lambda)$:

$$\lambda^* = \max\{\lambda : p^+(\lambda) = 0\}.$$

We claim that the smallest Frobenius eigenvalue is

$$\min_{|U|\leq I} \lambda^F(A + FUG) = \lambda^*.$$

By construction, $p(\lambda^*, U) \leq p^+(\lambda^*) = 0$.

Since any characteristic polynomial diverges, $p(\lambda, U) \to +\infty$ for $\lambda \to \infty$, it must have a root larger or equal to λ^*. Hence,

$$\lambda^* \leq \min_{|U|\leq I} \lambda^F(A + FUG).$$

Now, since $|U| \leq I$ is a compact set, the maximum $p^+(\lambda)$ is achieved for some U^*, hence $p(\lambda^*, U^*) = p^+(\lambda^*) = 0$. Therefore, λ^* is the Frobenius eigenvalue of $A+FU^*G$, the minimum possible one.

We conclude with this proposition.

Proposition 4. *Under Assumptions 1 and 2, the value of U that minimizes the Frobenius eigenvalue is on the vertices:*

$$\min_{|U|\leq I} \lambda^F(A + FUG) = \min_{|U|=I} \lambda^F(A + FUG).$$

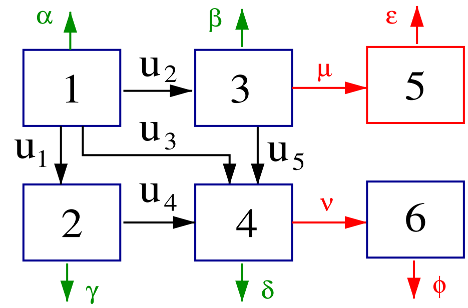

Fig. 2. Fluid network. The first compartment gets flooded: the fluid should flow outside reservoirs 1, 2, 3 and 4, while the presence of fluid in reservoirs 5 and 6 is undesired.

Proof. For fixed $\lambda = \lambda^*$, $p(\lambda^*, U)$ is a multiaffine function of the diagonal entries u_i of U. A multiaffine function defined on a hypercube (in our case $|U| \leq I$) reaches its minimum and maximum on the vertices (Giordano, Cuba Samaniego, Franco, & Blanchini, 2016), $|U| = I$. ■

The problem of minimizing the Frobenius eigenvalue is thus solved by considering all possible vertices of U, computing all the corresponding Frobenius eigenvalues $\lambda^F(A+FUG)$, and taking the minimum one. A stabilizing U exists if and only if the minimum Frobenius eigenvalue is negative.

The problem of maximizing the Frobenius eigenvalue, which is important to check stability for all $|U| \leq I$, has an analogous vertex solution: $\max_{|U|\leq I} \lambda^F(A + FUG) = \max_{|U|=I} \lambda^F(A + FUG)$.

8. Illustrative examples and applications

8.1. A flood control problem

Consider the fluid network with six compartments shown in Fig. 2, which is modeled as

$$\dot{x}(t) = Ax(t) + FV(t)Gx(t),$$

with

$$A = \begin{bmatrix} -\alpha & 0 & 0 & 0 & 0 & 0 \\ 0 & -\beta & 0 & 0 & 0 & 0 \\ 0 & 0 & -(\gamma+\mu) & 0 & 0 & 0 \\ 0 & 0 & 0 & -(\delta+\nu) & 0 & 0 \\ 0 & 0 & \mu & 0 & -\epsilon & 0 \\ 0 & 0 & 0 & \nu & 0 & -\phi \end{bmatrix},$$

$$F = \begin{bmatrix} -1 & -1 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

The state components represent the fluid level in the reservoirs (compartments, associated with the nodes), while the control $u(t)$ has components that represent the opening fractions of the valves regulating the flow along some of the pipes that connect the reservoirs: $u_i = -1$ if the valve is minimally opened and $u_i = 1$ if the valve is fully opened.

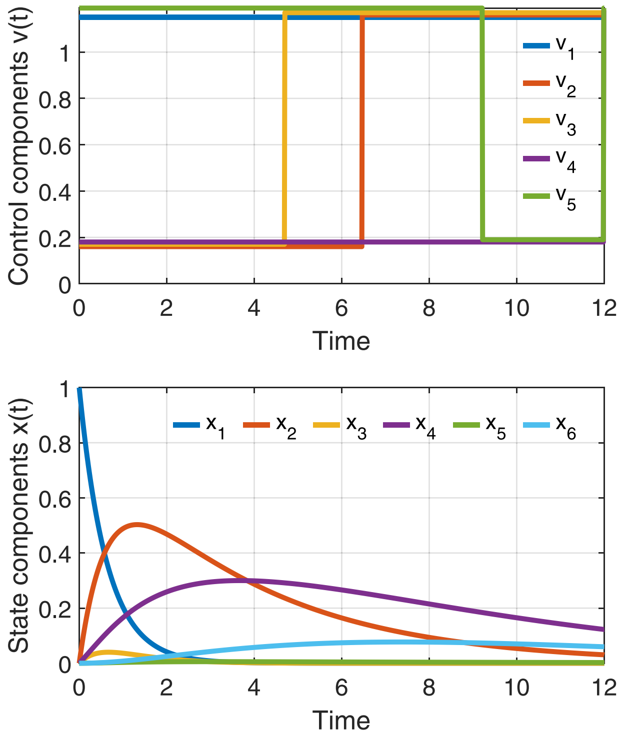

Fig. 3. The optimal (switching) control evolution on the horizon $T = 12$ (top), the corresponding optimal state evolution (bottom).

We set $\alpha = \beta = \gamma = \delta = \epsilon = \mu = \nu = 0.1$ and $\phi = 0.2$. The bounds on the control parameters v are

$$v^- = [0.150 \;\; 0.160 \;\; 0.170 \;\; 0.180 \;\; 0.190],$$
$$v^+ = [1.150 \;\; 1.160 \;\; 1.170 \;\; 1.180 \;\; 1.190].$$

We consider the scenario where an excessive fluid level is initially present in the system (flood) and we study the optimal emptying strategy. We assume that the presence of fluid in some compartments has to be avoided, yet it is temporarily necessary, in order to clear the system. Taking as a cost the integral over a finite interval of a linear functional, we penalize the weighted average of the fluid levels in the reservoirs.

We consider the cost $J = \int_0^{12} \ell^\top x(t)dt$, over [0, 12], with

$$\ell^\top = [1 \;\; 0 \;\; 0 \;\; 0 \;\; 1.2 \;\; 1],$$

which penalizes the presence of fluid in reservoirs 1, 5 and 6.

The fluid is initially present in reservoir 1, while the others are empty:

$$x(0) = [1 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0]^\top.$$

Fig. 3 shows the time evolution of the control variables (top) and of the optimal state (bottom).

If we consider the infinite horizon problem ($T = \infty$), then, solving (23), we get the cost-to-go function $V = z^\top x$ with

$$z^\top = [3.2067 \;\; 2.3571 \;\; 4.5779 \;\; 3.6667 \;\; 20.0000 \;\; 7.3333].$$

Correspondingly, the optimal constant control is

$$v = [v_1^+ \;\; v_2^- \;\; v_3^- \;\; v_4^- \;\; v_5^+],$$

meaning that, in the long run, the best strategy is to open pipes 1 and 5, and close 2, 3 and 4, as much as possible. Simulations over long horizons confirm this property.

8.2. Sensitivity analysis in epidemic evolution

Consider the SIDARTHE epidemiological model proposed by Giordano et al. (2020) and Giordano et al. (2021), rearranged as

$$\dot{x}(t) = \begin{bmatrix} -r_1 & 0 & 0 & 0 & 0 \\ \epsilon & -r_2 & 0 & 0 & 0 \\ \zeta & 0 & -r_3 & 0 & 0 \\ 0 & \eta & \theta & -r_4 & 0 \\ 0 & 0 & \mu & \nu & -r_5 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \omega(t) \tag{36}$$

$$\omega(t) = S(t)c^\top x(t) \tag{37}$$

$$\dot{S}(t) = -S(t)c^\top x(t) \tag{38}$$

$$c^\top = [\; \alpha \;\; \beta \;\; \gamma \;\; \delta \;\; 0 \;] \tag{39}$$

where $r_1 = \epsilon + \zeta + \lambda$, $r_2 = \eta + \rho$, $r_3 = \theta + \mu + \kappa$, $r_4 = \nu + \xi$, $r_5 = \sigma + \tau$. Vector $x = [I\,D\,A\,R\,T]^\top$ includes the fractions of Infected (not diagnosed), Diagnosed infected, Ailing (not diagnosed), Recognized (diagnosed with symptoms) and Threatened (diagnosed with severe illness, needing intensive care). S is the fraction of susceptible population. We wish to analyze the sensitivity of the model with respect to the parameters under either of the following assumptions:

- the susceptible population $S(t)$ is slowly varying on the interval, therefore it is assumed constant with some uncertainty;
- the susceptible population $S(t)$ is controlled, e.g. by a vaccination campaign, hence it is known with some uncertainty.

Slowly varying S. In this first case, we can absorb the uncertainty on S in the contagion parameters: $\alpha := S\alpha$, $\beta := S\beta$, $\gamma := S\gamma$, $\delta := S\delta$. Taking into account the uncertainty of S results in an additional modest uncertainty of these parameters. We then get a model of the form

$$\dot{x} = \begin{bmatrix} -r_1 + \alpha & \beta & \gamma & \delta & 0 \\ \epsilon & -r_2 & 0 & 0 & 0 \\ \zeta & 0 & -r_3 & 0 & 0 \\ 0 & \eta & \theta & -r_4 & 0 \\ 0 & 0 & \mu & \nu & -r_5 \end{bmatrix} x.$$

The adopted nominal values of the parameters (i.e., the components of vector u_{nom}) are: $\alpha_{nom} = 0.40$; $\beta_{nom} = 0.005$; $\gamma_{nom} = 0.110$; $\delta_{nom} = 0.0057$; $\epsilon_{nom} = 0.171$; $\zeta_{nom} = 0.034$; $\lambda_{nom} = 0.45$; $\eta_{nom} = 0.34$; $\rho_{nom} = 0.40$; $\theta_{nom} = 0.371$; $\mu_{nom} = 0.007$; $\kappa_{nom} = 0.017$; $\nu_{nom} = 0.007$; $\xi_{nom} = 0.017$; $\sigma_{nom} = 0.034$; $\tau_{nom} = 0.01$.

These parameters are uncertain and time-varying and we consider them as the components of u. We assume that an uncertainty is present in all parameters. Parameter α, the main infection parameter, associated with the contact between susceptible and infected asymptomatic people, is notoriously the most crucial one. It deeply affects the behavior of the disease spread. On the other hand, it is typically accurately estimated, possibly passing through the well known R_t parameter. According to Giordano et al. (2020),

$$R_t := S\frac{\alpha + \beta\epsilon/r_2 + \gamma\zeta/r_3 + \delta(\eta\epsilon/(r_2r_4) + \zeta\theta/(r_3r_4))}{r_1},$$

where parameters r_i have been defined above. According to our data, assuming the initial susceptible population fraction to be almost 1, we get $R_t = 0.6315$. This number is mostly dominated by the term $R_t = \alpha/r_1 \approx 0.6107$. From the computation of

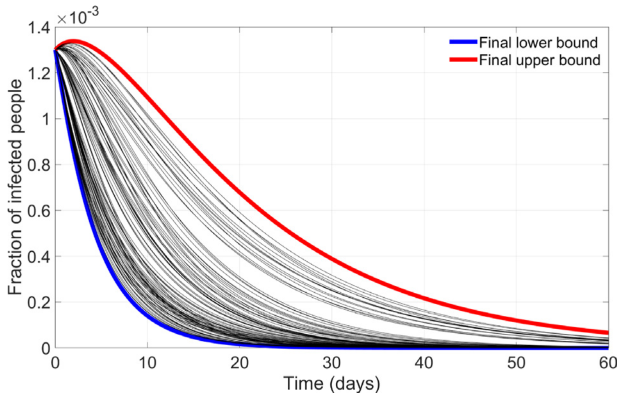

Fig. 4. Curves generating the infected upper and lower bounds (red and blue) valid at time $T = 60$ days, with $\alpha = \alpha_{nom} \pm 10\%$ and all other parameters considered with ±20% uncertainty. Randomly generated trajectories are shown in black. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

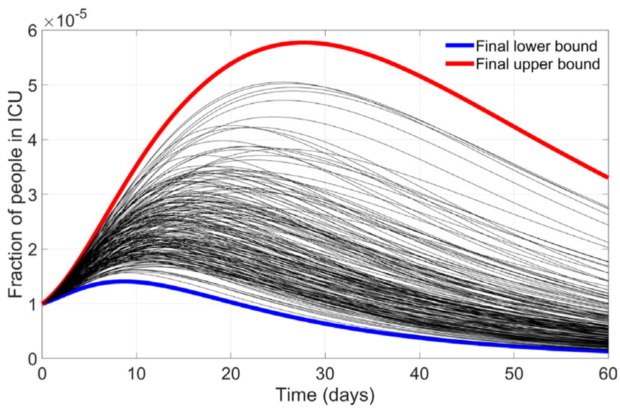

Fig. 5. Curves generating the upper and lower bounds for ICU occupancy (red and blue) valid at time $T = 60$ days, with ±10% uncertainty for α, ±20% uncertainty for all other parameters. Randomly generated trajectories are shown in black. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

R_t, we have a reasonable estimation of α. Hence, we assume an uncertainty of ±10% for α.

Conversely, the other parameters are less accurately estimated, although less crucial for the evolution. For these, we assume an uncertainty of ±20%: $\beta \in [0.8\beta_{nom}, 1.2\beta_{nom}]$, $\gamma \in [0.8\gamma_{nom}, 1.2\gamma_{nom}]$, and so on.

To find the exact bounds of the evolution, we solve an optimal control problem with $n = 5$ state variables and $m = 16$ control inputs. Note that the co-state equation introduces 5 new variables. With the proposed method, we need a single integration of the co-state equation to find the optimum.

Given a horizon of 60 days, we wish to determine bounds for the final number of infected people by selecting

$$\ell^\top = 0, \quad h^\top = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \end{bmatrix}.$$

Fig. 4 reports the bounds (red and blue) for the infected population, which are valid at time T.

Among the maximizer parameters, just one ($u_6 = \zeta$) switches from the maximum value 0.0408 to the minimum value 0.0272 at time $t \approx 58$ days: we denote this switching behavior by ± in the table below. The other parameters are constant, according to the pattern:

$\alpha\ \beta\ \gamma\ \delta$	ϵ	ζ	λ	η	ρ	θ	μ	κ	ν	ξ	σ	τ
+	−	±	−	−	−	−	−	−	−	−	−	−

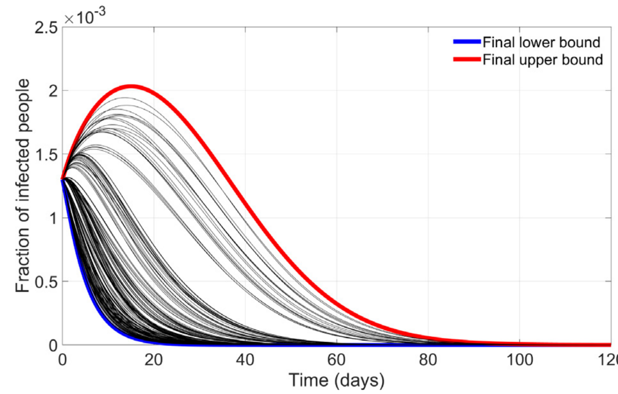

Fig. 6. Curves generating the bounds for the infected population, valid at time $T = 120$ days, with $\alpha(t) = \alpha_0 - (t/T)[\alpha_0 - \alpha_{fin}] \pm 10\%$ and all other parameters considered with ±20% parameter uncertainty.

The minimizer parameters are constant:

$\alpha\ \beta\ \gamma\ \delta$	ϵ	ζ	λ	η	ρ	θ	μ	κ	ν	ξ	σ	τ
−	+	+	+	+	+	+	+	+	+	+	+	+

We have also considered the bounds for the final values of ICU occupancy, by setting

$$\ell^\top = 0, \quad h^\top = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The results are shown in Fig. 5. The maximizer parameters are constant, with pattern:

$\alpha\ \beta\ \gamma\ \delta$	ϵ	ζ	λ	η	ρ	θ	μ	κ	ν	ξ	σ	τ
+	−	−	−	+	−	−	+	−	+	−	−	−

As for the minimizer parameters, $u_5 = \epsilon$ switches from the lower to the upper bound at time $t \approx 51$ days, while $u_6 = \zeta$ switches from the lower to the upper bound at time $t \approx 15$ days; we denote this behavior by ∓ in the table below. The remaining parameters are constant, with pattern:

$\alpha\ \beta\ \gamma\ \delta$	ϵ	ζ	λ	η	ρ	θ	μ	κ	ν	ξ	σ	τ
−	∓	∓	+	−	+	+	−	+	−	+	+	+

Controlled S: vaccination campaign. In the second case, the infection parameter α is assumed to decrease due to a vaccination campaign. This parameter is proportional to the fraction of susceptible population, $\alpha = \alpha_0 S$, and deeply subject to uncertainty. We assume the initial value $S = 1$ and the final value, after 4 months, $S_{fin} = 0.5$, so $\alpha_{fin} = \alpha_0 S_{fin}$, meaning that 50% of the population has been immunized. The term α is assumed to have a linear decreasing behavior with uncertainty of 10%:

$$\alpha(t) = \alpha_0 - \frac{t}{T}[\alpha_0 - \alpha_{fin}] \pm 10\%.$$

Fig. 6 reports upper and lower bounds for the infected population (valid at time T). The maximizer parameters do not switch and follow the pattern:

$\alpha\ \beta\ \gamma\ \delta$	ϵ	ζ	λ	η	ρ	θ	μ	κ	ν	ξ	σ	τ
+	−	−	−	+	−	−	−	−	−	−	−	−

The minimizer parameters are all constant but $u_5 = \epsilon$, which switches twice: from the upper to the lower value at time $t \approx 78$ days, and back to the upper value at time $t \approx 115$ days. We denote this behavior by ±∓:

$\alpha\ \beta\ \gamma\ \delta$	ϵ	ζ	λ	η	ρ	θ	μ	κ	ν	ξ	σ	τ
−	±∓	−	+	−	+	+	+	+	+	+	+	+

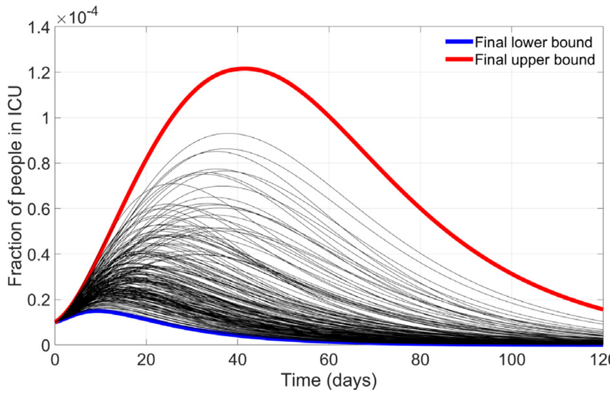

Fig. 7. Curves generating the bounds for ICU occupancy, valid at time $T = 120$ days, with $\alpha(t) = \alpha_0 - (t/T)[\alpha_0 - \alpha_{fin}] \pm 10\%$ and all other parameters considered with $\pm 20\%$ parameter uncertainty.

Finally we consider ICU occupancy, shown in Fig. 7. The maximizing parameters are all constant but $u_5 = \epsilon$ and $u_6 = \zeta$, which switch from the lower to the upper bound (a behavior denoted as $\mp$) at time $t \approx 39$ days and $t \approx 104$ days, respectively:

$\alpha\ \beta\ \gamma\ \delta$	ϵ	ζ	λ	η	ρ	θ	μ	κ	ν	ξ	σ	τ
$+$	$\mp$	$\mp$	$-$	$+$	$-$	$-$	$+$	$-$	$+$	$-$	$-$	$-$

The minimizing parameters are all constant but $u_5 = \epsilon$, which switches from the upper to the lower bound (a behavior denoted as $\pm$) at time $t \approx 4$ days:

$\alpha\ \beta\ \gamma\ \delta$	ϵ	ζ	λ	η	ρ	θ	μ	κ	ν	ξ	σ	τ
$+$	$\pm$	$-$	$+$	$-$	$+$	$+$	$-$	$+$	$-$	$+$	$+$	$+$

Infection and ICU occupancy curves initially increase and eventually decrease, once the vaccination coverage reaches a sufficient level.

9. Concluding discussion

We have solved an optimal control problem for linear compartmental systems in which some of the coefficients are control variables. This type of problem was previously addressed successfully under the assumption that the control coefficients appear on the diagonal: in this case, the problem is convex (Colaneri et al., 2014; Rantzer & Bernhardsson, 2014). In general, however, the problem is harder to solve. We show that, over both a finite and an infinite horizon, the problem can be solved by finding exact solutions to the Pontryagin equations and to the HJB equations. The key observation is that the state and co-state equations are decoupled. This is one of the few cases in which the HJB and Pontryagin equations admit a computable solution, without resorting to brute force numerical methods.

We have seen that the optimal solution in general may switch in the finite horizon problem, but it is constant in the infinite horizon case, meaning that, in the long run, the best strategy is to keep all the control parameters constant.

A limitation of the proposed theory is that the considered cost on the control action u_k associated with an arc leaving node i has been imposed on the flux $u_k x_i$, and not directly on u_k; in this latter case, the state and the co-state would no longer be decoupled. This more general problem is left for future investigation.

References

Alamo, T., Reina, D. G., Millán Gata, P., Preciado, V. M., & Giordano, G. (2021). Data-driven methods for present and future pandemics: Monitoring, modelling and managing. *Annual Reviews in Control*, *52*, 448–464.

Arino, J., Brauer, F., van den Driessche, P., Watmough, J., & Wu, J. (2007). A final size relation for epidemic models. *Mathematical Biosciences and Engineering*, *4*(2), 159–175.

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).

Blanchini, F., Bolzern, P., Colaneri, P., De Nicolao, G., & Giordano, G. (2021) Generalized epidemiological compartmental models: Guaranteed bounds via optimal control. In *Proc. 60th conf. on dec. and control*.

Blanchini, F., Colaneri, P., & Valcher, M. E. (2016). Switched linear positive systems. *Foundations and Trends in Systems and Control*, *2*(2), 101–273.

Blanchini, F., & Giordano, G. (2021). Structural analysis in biology: A control theoretic approach. *Automatica*, *126*(4), Article 109376.

Bloem, M., Alpcan, T., & Basar, T. (2009). Optimal and robust epidemic response for multiple networks. *Control Engineering Practice*, *17*, 525–533.

Brauer, F., & Castillo-Chavez, C. (2012). *Mathematical models in population biology and epidemiology* (2nd ed.). Springer.

Bussell, E. H., Dangerfield, C. E., Gilligan, C. A., & Cunniffe, N. J. (2019). Applying optimal control theory to complex epidemiological models to inform real-world disease management. *Philosophical Transactions of the Royal Society, Series B (Biological Sciences)*, *374*, Article 20180284.

Colaneri, P., Middleton, R., Chen, Z., Caporale, D., & Blanchini, F. (2014). Convexity of the cost functional in an optimal control problem for a class of positive switched systems. *Automatica*, *40*(4), 1227–1234.

Devia, C. A., & Giordano, G. (2019) Optimal duration and planning of switching treatments taking drug toxicity into account: A convex optimisation approach. In *Proc. 58th IEEE conference on decision and control* (pp. 5674–5679).

Diekmann, O., & Heesterbeek, J. A. P. (2000). *Mathematical epidemiology of infectious diseases: Model building, analysis and interpretation*. Wiley.

Forster, G. A., & Gilligan, C. A. (2007). Optimizing the control of disease infestations at the landscape scale. *Proceedings of the National Academy of Sciences*, *104*, 4984–4989.

Freddi, L. (2022). Optimal control of the transmission rate in compartmental epidemics. *Mathematical Control and Related Fields*, *12*(1), 201–223.

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*, 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*, 993–998.

Giordano, G., Cuba Samaniego, C., Franco, E., & Blanchini, F. (2016). Computing the structural influence matrix for biological systems. *Journal of Mathematical Biology*, *72*(7), 1927–1958.

Giordano, G., Rantzer, A., & Jonsson, V. (2016) A convex optimization approach to cancer treatment to address tumor heterogeneity and imperfect drug penetration in physiological compartments. In *Proc. 55th IEEE conference on decision and control* (pp. 2494–2500).

Gumel, A. B., Ruan, S., Day, T., Watmough, J., Brauer, F., van den Driessche, P., et al. (2004). Modelling strategies for controlling SARS outbreaks. *Proceedings of the Royal Society of London. Series B: Biological Sciences (London)*, *271*(1554), 2223–2232.

Hansen, E., & Day, T. (2011). Optimal control of epidemics with limited resources. *Journal of Mathematical Biology*, *62*, 423–451.

Hayhoe, M., Barreras, F., & Preciado, V. M. (2021). Multitask learning and nonlinear optimal control of the COVID-19 outbreak: A geometric programming approach. *Annual Reviews in Control*, *52*, 495–507.

Hernandez-Vargas, E., Colaneri, P., Middleton, R., & Blanchini, F. (2011). Discrete-time control for switched positive systems with application to mitigating viral escape. *International Journal of Robust and Nonlinear Control*, *21*(10), 1093–1111.

Hernandez Vargas, E., Middleton, R., & Colaneri, P. (2014). Switching strategies to mitigate HIV mutation. *IEEE Transactions on Control Systems Technology*, *22*(4), 1623–1628.

Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM Review*, *42*, 599–653.

Jacquez, J. A., & Simon, C. P. (1993). Qualitative theory of compartmental systems. *SIAM Review*, *35*(1), 43–79.

Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of The Royal Society of London*, *115*, 700–721.

Köhler, J., Schwenkel, L., Koch, A., Berberich, J., Pauli, P., & Allgöwer, F. (2021). Robust and optimal predictive control of the COVID-19 outbreak. *Annual Reviews in Control*, *51*, 525–553.

Kusuoka, H., Kodama, S., Maeda, H., Inoue, M., Hori, M., Abe, H., et al. (1981). Optimal control in compartmental systems and its application to drug administration. *Mathematical Biosciences*, *53*(1–2), 59–77.
Lee, C. S., & Leitmann, G. (1994). Control strategies for an endemic disease in the presence of uncertainty. In R. P. Agarwal (Ed.), *Recent trends in optimization theory and applications* (pp. 221–238).
Lenhart, S., & Workman, J. T. (2007). *Optimal control applied to biological models*. Chapman & Hall/CRC.
Lewis, F. L., Vrabie, D. L., & Syrmos, V. L. (2012). *Optimal control* (3rd ed.). John Wiley & Sons.
Mandal, M., Jana, S., Nandi, S. K., Khatua, A., Adak, S., & Kar, T. (2020). A model based study on the dynamics of COVID-19: Prediction and control. *Chaos, Solitons & Fractals*, *136*, Article 109889.
Martcheva, M. (2015). *An introduction to mathematical epidemiology*. Springer Science and Business Media LLC.
Morton, R., & Wickwire, K. H. (1974). On the optimal control of a deterministic epidemic. *Advances in Applied Probability*, *6*, 622–635.
Rantzer, A., & Bernhardsson, B. (2014) Control of convex-monotone systems. In *Proc. 2014 IEEE conf. decision and control* (pp. 2378–2383). Los Angeles, CA.
Rowthorn, R. E., Laxminarayan, R., & Gilligan, C. A. (2009). Optimal control of epidemics in metapopulations. *Journal of the Royal Society Interface*, *6*, 1135–1144.
Sharomi, O., & Malik, T. (2017). Optimal control in epidemiology. *Annals of Operations Research*, *251*, 55–71.
Swierniak, A., Ledzewicz, U., & Schättler, H. (2003). Optimal control for a class of compartmental models in cancer chemotherapy. *International Journal of Applied Mathematics and Computer Science*, *13*(3), 357–368.
Yeargers, E. K., Herod, J. V., & Shonkweiler, R. W. (2012). *An introduction to the mathematics of biology: With computer algebra models*. Birkhäuser.

Franco Blanchini was born on 29 December 1959, in Legnano (Italy). He is the Director of the Laboratory of System Dynamics at the University of Udine. He has been involved in the organization of several international events: in particular, he was Program Vice-Chairman of the conference Joint CDC-ECC 2005, Seville, Spain; Program Vice-Chairman of the Conference CDC 2008, Cancun, Mexico; Program Chairman of the Conference ROCOND, Aalborg, Denmark, June 2012 and Program Vice-Chairman of the Conference CDC 2013, Florence, Italy. He is co-author of the book "Set theoretic methods in control", Birkhäuser. He received the 2001 ASME Oil & Gas Application Committee Best Paper Award as a co-author of the article "Experimental evaluation of a High-Gain Control for Compressor Surge Instability", the 2002 IFAC prize Survey Paper Award as the author of the article "Set Invariance in Control - a survey", Automatica, November 1999, for which he also received the High Impact Paper Award in 2017, and the 2017 NAHS Best Paper Award as a co-author of the article "A switched system approach to dynamic race modelling", Nonlinear Analysis: Hybrid Systems, 2016. He was nominated Senior Member of the IEEE in 2003. He has been an Associate Editor for Automatica (1996–2006 and 2017–2019) and for the IEEE Transactions on Automatic Control (2012–2016), and a Senior Editor for the IEEE Control Systems Letters.

Paolo Bolzern was born in Milano, Italy in 1955. He received his Doctor's degree (Laurea) in Electronic Engineering from Politecnico di Milano in 1978. From 1983 to 1987, he was a research fellow with the Centro di Teoria dei Sistemi of the National Research Council (CNR). In 1987, he has joined Politecnico di Milano, where he is currently a Full Professor of Automatic Control in the Dipartimento di Elettronica, Informazione e Bioingegneria. In the period 2000–2002 he served as Vice-Chairman of the Department and, from 2003 to 2011, he was the coordinator of the Systems and Control area. From 2015 to 2019 he was the coordinator of the Ph.D. Program in Information Technology (area Systems and Control) at Politecnico di Milano.

He is a member of IEEE, Associate Editor of IEEE Control Systems Letters, and past Associate Editor of IEEE Transactions on Automatic Control (2015–2018). He sat in the EUCA (European Control Association) Assembly (2009–2017). He was appointed as Electronic Media Editor for IFAC (2015–2020), and served in the Organizing Committee of the IFAC World Congresses in Milano (2011) and Toulouse (2017). His current research interests include robust filtering and control, coordination of multi-agent systems, positive systems, control of deterministic and stochastic switched systems, and opinion dynamics modeling. On these and other related topics he has authored or co-authored some 90 research papers published in peer-reviewed journals and books. He is a co-author with R. Scattolini and N. Schiavoni of the textbook Fondamenti di Controlli Automatici, McGraw-Hill, 2015 (in Italian, also translated into Spanish).

Patrizio Colaneri received the Laurea degree in Electrical Engineering from the Politecnico di Milano in 1981 and the Ph.D. degree (Dottorato di Ricerca) in Automatic Control from the Italian Ministry of Education and research in 1988. He is Full Professor of Automatica at Politecnico di Milano, Italy, where he served as head of the Ph.D school on ICT (2007–2009). He held visiting positions at the University of Maryland, at the Hamilton Institute of the National University of Ireland, and at the Institute for Design and Control of Mechatronical Systems of Johannes Kepler University in Linz (Austria). He served IFAC and IEEE CSS in many capacities in journals and conferences. In particular he was Associate Editor of Automatica for six years, Senior Editor of IEEE TAC for 8 years, IPC Chair of the IFAC Symposium ROCOND in Milan (2003), IPC Vice Chair of the CDC in Florence (2013), IPC Chair of the IEEE Multi Conference on System and Control (CCA) in Buenos Aires (2016). He has been a Senior Editor of the IFAC journal Non Linear Analysis: Hybrid Systems for six years. He has been IFAC TC chair, CC chair and member of the IFAC Technical Board for 16 years, and is currently the Chair of the Italian National Member organization of IFAC. He is a Fellow of IFAC and IEEE and received the certificate of outstanding service of IFAC. His main theoretical interests are in the areas of periodic systems and control, robust filtering and control, switching control, modeling of epidemic diseases, opinion dynamics and Markov processes. His main applications are in power systems and railway automation. He authored/co-authored about 250 papers and seven books (three in Italian).

Giuseppe De Nicolao has been full professor of Systems and Control Engineering since 2000 at the University of Pavia, Italy, where, since 2020, he has been Director of the University's Doctoral School. From 1987 to 1988 he worked at the "Mario Negri" Institute for Pharmacological Researches in Milan and, since 1988, at the CNR's Center for System Theory in Milan. In 1992 he joined the University of Pavia as an associate professor. He has published more than 140 papers in refereed journals on the following topics: Systems Identification, Periodic Systems, Model Predictive Control, Machine Learning, Modeling and Control of Biomedical Systems, Advanced Process Control Applied to Semiconductor Manufacturing. He has served on the editorial boards of the IEEE Transactions on Automatic Control, IEEE Transactions on Control Systems Technology, and Automatica. He has been Team Leader of the Pavia Unit in projects funded by government agencies and the European Union, as well as in several industry-funded projects, and is co-inventor of international patents.

Giulia Giordano received the B.Sc. and M.Sc. degrees summa cum laude in electrical engineering and the Ph.D. degree (Hons.) in systems and control theory from the University of Udine, Italy, in 2010, 2012, and 2016, respectively. She visited the Control and Dynamical Systems Group, California Institute of Technology, Pasadena, CA, USA, in 2012, and the Institute of Systems Theory and Automatic Control, University of Stuttgart, Germany, in 2015. She was a Research Fellow with the LCCC Linnaeus Center and the Department of Automatic Control, Lund University, Sweden, from 2016 to 2017, and an Assistant Professor with the Delft Center for Systems and Control, Delft University of Technology, The Netherlands, from 2017 to 2019; she is currently with the Department of Industrial Engineering, University of Trento, Italy. She has been an Associate Editor for the IEEE Control Systems Letters since 2020 and for Automatica since 2022. She was recognized with the Outstanding Reviewer Letter from the IEEE Transactions on Automatic Control in 2016 and from the Annals of Internal Medicine in 2020, and chosen as the Outstanding Associate Editor of the IEEE Control Systems Letters for the year 2021. She received the EECI Ph.D. Award 2016 from the European Embedded Control Institute for her thesis "Structural Analysis and Control of Dynamical Networks", the NAHS Best Paper Prize 2017 for the article "A Switched System Approach to Dynamic Race Modelling", and the SIAM Activity Group on Control and Systems Theory Prize 2021 for "significant contributions to the development of innovative methodologies for the structural analysis of networked control systems and their applications to biological networks". Her main research interests include the analysis and the control of dynamical networks, with applications especially to biology and epidemiology.