

Optimal Control for a Class of Infinite Dimensional Systems Involving an L^∞ -term in the Cost Functional

Sébastien Court^{1,*}, Karl Kunisch², and Laurent Pfeiffer¹

¹ Institute for Mathematics and Scientific Computing, Karl-Franzens-Universität, Heinrichstr. 36, 8010, Graz, Austria

² Institute for Mathematics and Scientific Computing, Karl-Franzens-Universität, Heinrichstr. 36, 8010, Graz, Austria and Radon Institute, Austrian Academy of Sciences

Received 30 August 2016, revised 6 July 2017, accepted 17 October 2017

Published online 13 November 2017

Key words Optimality conditions, PDE-constrained optimization, transversality conditions, hybrid optimal control problem.

MSC (2010) 49K20, 49K15, 49K21, 93C30, 90C46

An optimal control problem with a time-parameter is considered. The functional to be optimized includes the maximum over time-horizon reached by a function of the state variable, and so an L^∞ -term. In addition to the classical control function, the time at which this maximum is reached is considered as a free parameter. The problem couples the behavior of the state and the control, with this time-parameter. A change of variable is introduced to derive first and second-order optimality conditions. This allows the implementation of a Newton method. Numerical simulations are developed, for selected ordinary differential equations and a partial differential equation, which illustrate the influence of the additional parameter and the original motivation.

© 2017 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

1 Introduction

We consider optimal control problems with a cost functional involving a function of the state variable at some time τ during its time evolution. The time τ itself is free to move within the time-span $(0, T)$ of the whole experiment. As a very first thought, one might think of the prototype of a new vehicle and ask what is the highest reachable speed on the race track. Our original motivation for the class of problems under consideration stems from the domain of cardiac electro-physiology. The problem of fibrillation of a part of the heart can be mechanically tackled with the use of electric shocks acting on this muscle, leading to the re-oxygenation of the ill area, and by this means forcing this area to recover a healthy electric activity. The efficiency of the defibrillation is known to be related to the maximum reached over time by the derivative of a pressure in the heart. This quantity can be mathematically formulated as the maximum taken by a function of the state variable of the problem. We refer to [1] for more details on related optimal control problems arising in electro-cardiology. While we continue to work on this challenging problem, we focus in this article on the methodology enabling to derive optimality conditions for such a problem, for a simpler class of partial differential equations.

The optimal control problem that we shall investigate can be formulated as follows:

$$\left\{ \begin{array}{l} \max_{\tau \in [0, T], u \in L^2(0, T; U)} \int_0^T \ell(y(t), u(t)) dt + \phi_1(y(\tau)) + \phi_2(y(T)), \\ \text{subject to: } \dot{y} = f(y, u), \quad y(0) = y_0, \quad \int_0^T \|u(t)\|_U^2 dt \leq \gamma, \end{array} \right. \quad (1)$$

where the state variable y is the solution of an evolution equation controlled by u . In this problem, ℓ denotes the cost functional, ϕ_1 the functional we want to maximize at some time τ , and ϕ_2 is the terminal cost. The analytical framework will be specified later. The specificity of this kind of problem lies in the fact that a time parameter, namely τ , can be optimized. Not only do we maximize $\phi_1 \circ y$ with the help of the control u , but we also optimize the time τ for which the maximum is reached. Note that, when ϕ_1 is nonnegative, the problem can be equivalently formulated in the following

* Corresponding authors, E-mail: sebastien.court@uni-graz.at

way:

$$\left\{ \begin{array}{l} \max_{u \in L^2(0,T;U)} \int_0^T \ell(y(t), u(t)) dt + \|\phi_1(y(\cdot))\|_{L^\infty(0,T;\mathbb{R})} + \phi_2(y(T)), \\ \text{subject to: } \dot{y} = f(y, u), \quad y(0) = y_0, \quad \int_0^T \|u(t)\|_U^2 dt \leq \gamma. \end{array} \right.$$

In this fashion, the cost function incorporates the L^∞ -norm of a given function of the state variable.

With appropriate technical modifications, the results provided in this paper can be extended to the following hybrid control problem:

$$\left\{ \begin{array}{l} \max_{\substack{0=\tau_0 < \tau_1 < \dots < \tau_K=T \\ u \in L^2(0,T;U)}} \sum_{k=1}^K \int_{\tau_{k-1}}^{\tau_k} \ell_k(y(t), u(t)) dt + \phi_k(y(\tau_k)), \\ \text{subject to: } \dot{y}(t) = f_k(y(t), u(t)), \text{ for a. e. } t \in (\tau_{k-1}, \tau_k), \quad y(0) = y_0, \quad \int_0^T \|u(t)\|_U^2 dt \leq \gamma. \end{array} \right. \quad (2)$$

The above problem incorporates time parameters (or switching times) $\tau_1, \dots, \tau_{K-1}$ which can be optimized. The term *hybrid* refers here to the fact that at the switching times τ_k , the system move from a given regime (described here by the dynamics f_k) to another one (described by the dynamics f_{k+1}). At the switching times, the integral cost changes and a function of the state is also incorporated into the cost function. There are many ways to generalize problem (2) (see e.g. [2]). For example, one can consider a formulation of the problem for which the dynamics f_k (used during the interval (τ_{k-1}, τ_k)) can be itself chosen into a finite set of functions. Such generalizations are beyond the scope of the paper.

In the optimal control literature, many problems include time parameters that have to be optimized. Among them, time-optimal control problems have probably attracted the most attention. These problems basically consist in minimizing the time needed to reach a given target. We refer to the early reference [3] on this topic. Time-optimal control problems have been studied for various models: See for example [4] for the wave equation, [5, 6] for the monodomain system, [7] for the Navier-Stokes equations, [8] for the heat equation. We also mention the time-crisis management problem studied in [9]; For such a problem, the time spent out of a certain closed domain must be minimized.

For our problem, the first-order optimality conditions consist of a weak maximum principle for the control variable, and a transversality condition for the optimality of the time parameter. The derivation of the transversality condition is difficult, in so far as the Lagrangian of the problem is not differentiable with respect to the time-parameters. In e.g. [10–12], a change of variables (in time) is performed to derive the transversality condition in the case of a time-optimal control problem. This is the approach that we adopt here. It also enables us to derive second-order optimality conditions, as in [6] and [13] for finite-dimensional hybrid control problems. In this last reference, a numerical test based on a Riccati equation is provided to check the sufficient second-order conditions. In [2], specific needle variations are designed, for hybrid systems. Other approaches can also be considered. In [9], the time-crisis management problem is regularized and optimality conditions are derived by Γ -convergence. One of the contributions of the present paper is that the concept of reparameterization for the class of hybrid systems specified in (1) are systematically investigated in the infinite-dimensional setting, thus allowing for applications in the context of PDEs.

The paper is organized as follows: In Sect. 2 the functional framework is specified, and the problem is transformed with a change of variables. Sect. 3 is devoted to the derivation of first and second-order optimality conditions. The abstract framework proposed here is shown to be satisfied, as an example, by the Navier-Stokes equations in dimension 2. The issue of numerical resolution is addressed in Sect. 4: We use the theoretical expressions of the optimality condition in order to design an algorithm which solves the problem, mixing Barzilai-Borwein gradient steps with Newton steps. As illustrations, we consider two examples of ordinary differential systems and one example dealing with the Burgers equation.

Notation. When there is no ambiguity, the time variable is sometimes omitted, or written only once (e.g. $H(y, u, p)(s)$, instead of $H(y(s), u(s), p(s))$). First and second-order partial derivatives are denoted with the use of indexes. The first and second-order derivatives with respect to all variables (except the adjoint state p and the Lagrange multiplier λ , for the Hamiltonian and the Lagrangians) are denoted by D and D^2 , respectively. The partial derivative of a function φ with respect to a variable y in a direction z is denoted as $\varphi_y(y).z$. When a function h is left- and right-continuous at a given time t , the left- and right-limits are denoted by $h(t^-)$ and $h(t^+)$, respectively, and the jump is denoted by $[h]_t$.

2 Formulation of the problem

2.1 Setting

Let X be a real Hilbert space, and Y be a reflexive Banach space forming with X a Gelfand triple $Y \subseteq X \equiv X' \subseteq Y'$, with Y densely contained in X . We denote $Z = Y'$, so that Z' is isomorphic to Y . Let U be the Hilbert space of controls. Further let $\ell : X \times U \rightarrow \mathbb{R}$ and $f : Y \times U \rightarrow Y'$ be two twice continuously Fréchet-differentiable mappings. For $\tilde{u} \in L^2(0, T; U)$, the state is governed by the autonomous control system

$$\begin{cases} \dot{\tilde{y}} = f(\tilde{y}, \tilde{u}) & \text{on } (0, T), \\ \tilde{y}(0) = y_0, \end{cases} \tag{3}$$

where the initial condition $y_0 \in X$ is given. We assume that for all $y_0 \in X$ and $\tilde{u} \in L^2(0, T; U)$ this system has a unique solution \tilde{y} in the space:

$$W(0, T; Y) := L^2(0, T; Y) \cap W^{1,2}(0, T; Y').$$

Recall the continuous embedding $W(0, T; Y) \hookrightarrow C([0, T]; X)$, so that in particular the initial condition makes sense in X . Weak solutions for systems of type (3) will be understood in $W(0, T; Y)$ in the variational sense, that is to say for all $\varphi \in L^2(0, T; Y)$ they must satisfy

$$\langle \dot{\tilde{y}}(t) - f(\tilde{y}(t), \tilde{u}(t)); \varphi(t) \rangle_{Y', Y} = 0$$

for almost every $t \in (0, T)$, and $\tilde{y}(0) = y_0$ in X . Let ϕ_1 and $\phi_2 : X \rightarrow \mathbb{R}$ be two twice continuously Fréchet-differentiable mappings. We consider the following problem:

$$\begin{cases} \max_{\tau \in [0, T], \tilde{u} \in L^2(0, T; U)} \int_0^T \ell(\tilde{y}(t), \tilde{u}(t)) dt + \phi_1(\tilde{y}(\tau)) + \phi_2(\tilde{y}(T)) \\ \text{subject to: } \dot{\tilde{y}} = f(\tilde{y}, \tilde{u}) \text{ in } Y', \quad \tilde{y}(0) = y_0 \text{ in } X, \quad \int_0^T \|\tilde{u}(t)\|_U^2 dt \leq \gamma, \end{cases} \tag{4}$$

where $\gamma \in (0, \infty)$. The operators f and ℓ are assumed to generate Nemytskii operators on $W(0, T; Y) \times L^2(0, T; U)$, whose classical properties can be found in [14, Sect. 4.3], for instance. For a sake of compact notation, we keep the same notation for these Nemytskii operators so defined, and will consider in particular the relations $f(\tilde{y}, \tilde{u})(t) = f(\tilde{y}(t), \tilde{u}(t))$ and $\ell(\tilde{y}, \tilde{u})(t) = \ell(\tilde{y}(t), \tilde{u}(t))$. Throughout the paper, we consider the following set of assumptions on f and ℓ .

- A0** The mapping $\ell : W(0, T; Y) \times L^2(0, T; U) \rightarrow L^1(0, T; \mathbb{R})$ is twice continuously Fréchet-differentiable.
- A1** The mapping $f : W(0, T; Y) \times L^2(0, T; U) \rightarrow L^2(0, T; Y')$ is twice continuously Fréchet-differentiable.
- A2** For all $\tilde{u} \in L^2(0, T; U)$, $y_0 \in X$, system (3) admits a unique weak solution in $W(0, T; Y)$.
- A3** For all $(\tilde{y}, \tilde{u}) \in W(0, T; Y) \times L^2(0, T; U)$, $\xi \in L^2(0, T; Y')$, $z_0 \in X$, there exists a unique $z \in W(0, T; Y)$ solution to the following system

$$\begin{cases} \dot{z} = f_{\tilde{y}}(\tilde{y}, \tilde{u}) \cdot z + \xi & \text{on } (0, T), \\ z(0) = z_0. \end{cases}$$

We included the constraint $\|\tilde{u}\|_{L^2(0, T; U)}^2 \leq \gamma$ on the energy of the control in the problem formulation (4). While existence of solutions is not the focus of this paper, this constraint allows to argue existence in concrete cases, as illustrated in Sect. 2.2.1. In a follow-up work it can be of interest to include different constraints on the controls and/or the state variable, and to analyze their effect on the switching time τ .

2.2 Comments on the functional framework

2.2.1 Example

The variational framework $W(0, T; Y)$ is applicable, for instance, to the Burgers system (see Sect. 4.2.3), or to the unsteady Navier-Stokes system (see [15, 16]). Assumptions **A1–A3** can be easily verified for the Burgers system, if we refer to the analysis of e.g. [17].

Let us detail the verification for the Navier-Stokes system and also formulate the problem in such a manner that existence of a solution can be guaranteed.

$$\begin{aligned} \dot{y} + (y \cdot \nabla)y - \nu \Delta y + \nabla p &= Bu & \text{in } \Omega \times (0, T), \\ \operatorname{div} y &= 0 & \text{in } \Omega \times (0, T), \\ y &= 0 & \text{on } \partial\Omega \times (0, T), \\ y(0) &= y_0 & \text{in } \Omega. \end{aligned}$$

In the system above, Ω is a bounded domain of \mathbb{R}^2 , with smooth boundary, and $B \in \mathcal{L}(U, X)$ is a linear operator, as for instance $B = \mathbf{1}_\omega$ where $\omega \subset \Omega$, with $U = [L^2(\omega)]^2$, and $\mathbf{1}_\omega$ denotes the extension-by-zero operator, i.e. $\mathbf{1}_\omega = u(x)$ for $x \in \omega$ and $(Bu)(x) = 0$ otherwise. We consider the solenoidal spaces

$$Y = \{y \in [H_0^1(\Omega)]^2 \mid \operatorname{div} y = 0\}, \quad X = \{y \in [L^2(\Omega)]^2 \mid \operatorname{div} y = 0, y \cdot \nu = 0 \text{ on } \partial\Omega\}, \quad Y' \supset [H^{-1}(\Omega)]^2,$$

where ν is the outer normal to $\partial\Omega$, and $y_0 \in Y$. It is well-known that X is a closed subspace of $[L^2(\Omega)]^2$, [18]. Now we are prepared to state the variational formulation of the state equation, given $u \in L^2(0, T; U)$, in which the pressure p disappears:

Find $y \in W(0, T; Y)$ such that for all $\varphi \in L^2(0, T; Y)$ we have almost everywhere in $(0, T)$:

$$\begin{cases} \langle \dot{y} + (y \cdot \nabla)y; \varphi \rangle_{Y'; Y} + \nu \langle \nabla y; \nabla \varphi \rangle_{[L^2(\Omega)]^2} = \langle B(u); \varphi \rangle_{Y'; Y}, \\ \langle y(0); \varphi \rangle_X = \langle y_0; \varphi \rangle_X. \end{cases} \quad (5)$$

For details, we refer to [18, Chap. III.3.5]. For our example, in the context of the general framework above, we have $f(y, u) = \nu \Delta y - (y \cdot \nabla)y + P_X Bu$, where $P_X \in \mathcal{L}([L^2(\Omega)]^2, X)$ denotes the orthogonal projection onto X . To be specific, we further set for $(z, v) \in X \times U$

$$\ell(z, v) = -\|v\|_U^2, \quad \phi_1(z) = \phi_2(z) = \|z\|_{L^2(\Omega)}^2. \quad (6)$$

Let us verify **A0–A4**. Assumption **A0** is clearly satisfied. Concerning **A1**, short considerations show that first- and second-order directional differentiability and continuity of the derivatives (and hence Fréchet differentiability) of f will follow from the continuity of the Oseen-type term $(z_1, z_2, z_3) \mapsto (z_1 \cdot \nabla)z_2 + (z_3 \cdot \nabla)z_1$ from $[W(0, T; Y)]^3$ to $L^2(0, T; Y')$, which will be given below. Since $\operatorname{div} z_1 = \operatorname{div} z_2 = \operatorname{div} z_3 = 0$, we have

$$\begin{aligned} (z_1 \cdot \nabla)z_2 + (z_3 \cdot \nabla)z_1 &= \operatorname{div}(z_2 \otimes z_1 + z_1 \otimes z_3), \\ \|(z_1 \cdot \nabla)z_2 + (z_3 \cdot \nabla)z_1\|_{Y'} &\leq \|(z_1 \cdot \nabla)z_2 + (z_3 \cdot \nabla)z_1\|_{[H^{-1}(\Omega)]^2} \\ &\leq \|z_1 \otimes z_2 + z_3 \otimes z_1\|_{[L^2(\Omega)]^2} \leq \|z_1 \otimes z_2\|_{[L^2(\Omega)]^4} + \|z_3 \otimes z_1\|_{[L^2(\Omega)]^4}. \end{aligned}$$

In view of the symmetric role played by z_1 , z_2 , and z_3 , we only estimate the first term. From [19, (Appendix B, Proposition B.1)], in dimension 2 there exists a constant $C > 0$ independent of z_1 and z_2 such that

$$\|z_1 \otimes z_2\|_{[L^2(\Omega)]^4} \leq C \|z_1\|_{[H^{1/2}(\Omega)]^2} \|z_2\|_{[H^{1/2}(\Omega)]^2}.$$

By interpolation we deduce

$$\|z_1 \otimes z_2\|_{[L^2(\Omega)]^4} \leq C \|z_1\|_{[L^2(\Omega)]^2}^{1/2} \|z_1\|_{[H^1(\Omega)]^2}^{1/2} \|z_2\|_{[L^2(\Omega)]^2}^{1/2} \|z_2\|_{[H^1(\Omega)]^2}^{1/2}.$$

By integrating in time, this yields

$$\begin{aligned} \|z_1 \otimes z_2\|_{L^2(0, T; [L^2(\Omega)]^4)} &\leq C \|z_1\|_{L^\infty(0, T; X)}^{1/2} \|z_2\|_{L^\infty(0, T; X)}^{1/2} \|z_1\|_{L^2(0, T; Y)}^{1/2} \|z_2\|_{L^2(0, T; Y)}^{1/2} \\ &\leq \tilde{C} \|z_1\|_{W(0, T; Y)} \|z_2\|_{W(0, T; Y)}, \end{aligned}$$

for some constant $\tilde{C} > 0$. Thus **A1** is satisfied. Assumptions **A2** and **A3** are due to [15, (Proposition 2.1 page 928, and Proposition 2.4 page 929, respectively)].

Now we turn to the question of existence for problem (4). For this purpose we consider problem (4) with the specifications (5) and (6). Let $\{(\tau_n, u_n)\} \in [0, \tau] \times L^2(0, T; U)$ denote a maximizing sequence. Since $\|u_n\|_{L^2(0, T; U)}^2 \leq \gamma$ for all n , there exists a subsequence of $\{(\tau_n, u_n)\}$, that for simplicity we denote by the same symbol, and $(\bar{\tau}, \bar{u}) \in [0, \tau] \times L^2(0, T; U)$ such that $\{(\tau_n, u_n)\}$ converges weakly to $(\bar{\tau}, \bar{u})$. Since $B \in \mathcal{L}(U, X)$, it follows that $\{Bu_n\}$ is bounded in X . By assuming that $y_0 \in Y$, it follows that $\{y(u_n)\}$ is a bounded sequence in $W(0, T; Y) \cap L^2(0, T; [H^2(\Omega)]^2) \cap H^1(0, T; X)$, see [18, Theorem III. 3.10] and [15, Proposition 2.3]. It is now standard to pass to the limit in (5) with u replaced by u_n and

to find that $y(u_n)$ converges at least weakly in $W(0, T; Y)$ to the solution $y(\bar{u})$ associated with the control \bar{u} . The extra boundedness of $\{y(u_n)\}$ in $L^2(0, T; [H^2(\Omega)]^2) \cap H^1(0, T; X)$ is needed to pass to the lim sup in the functional

$$-\int_0^T \|u_n(t)\|_U^2 dt + \|y(\tau_n; u_n)\|_{[L^2(\Omega)]^2}^2 + \|y(T; u_n)\|_{[L^2(\Omega)]^2}^2. \tag{7}$$

From [20, Corollary 8], the embedding of $L^2(0, T; [H^2(\Omega)]^2) \cap H^1(0, T; X)$ in $C([0, T]; [[H^2(\Omega)]^2, X]_{1-\epsilon})$ is relatively compact, where $\epsilon \in (0, \frac{1}{2})$, and hence relatively compactly into $C([0, T]; [L^2(\Omega)]^2)$. Here $[[H^2(\Omega)]^2, X]_{1-\epsilon}$ denotes the interpolation space with index $1 - \epsilon$, which is well-defined since the Stokes operator associated with (5) in the space X generates an analytic semi-group, see e.g. [21, Theorem 2]. In particular, possibly after taking another subsequence, we have that $y(u_n) \rightarrow y(\bar{u})$ in $C([0, T]; [L^2(\Omega)]^2)$, see [20, Lemma 1]. Now we have the estimate

$$\begin{aligned} \|y(\tau_n; u_n) - y(\bar{\tau}; \bar{u})\|_{[L^2(\Omega)]^2} &\leq \|y(\tau_n; u_n) - y(\tau_n; \bar{u})\|_{[L^2(\Omega)]^2} + \|y(\tau_n; \bar{u}) - y(\bar{\tau}; \bar{u})\|_{[L^2(\Omega)]^2} \\ &\leq \|y(\cdot; u_n) - y(\cdot; \bar{u})\|_{C(0, T; [L^2(\Omega)]^2)} + \|y(\tau_n; \bar{u}) - y(\bar{\tau}; \bar{u})\|_{[L^2(\Omega)]^2}. \end{aligned}$$

Analogously we obtain that $y(T; u_n) \rightarrow y(T; \bar{u})$ in $[L^2(\Omega)]^2$. Since (τ_n, u_n) was taken as a maximizing sequence, we can take the lim sup in (7) and obtain

$$\begin{aligned} &-\int_0^T \|u(t)\|_U^2 dt + \|y(\tau; u)\|_{[L^2(\Omega)]^2}^2 + \|y(T; u)\|_{[L^2(\Omega)]^2}^2 \\ &\leq -\int_0^T \|\bar{u}(t)\|_U^2 dt + \|y(\bar{\tau}; \bar{u})\|_{[L^2(\Omega)]^2}^2 + \|y(T; \bar{u})\|_{[L^2(\Omega)]^2}^2 \end{aligned}$$

for all $u \in L^2(0, T; U)$ with $\|u\|_{L^2(0, T; U)}^2 \leq \gamma$ and $\tau \in [0, T]$, and thus existence of an optimal control is established.

2.2.2 Additional regularity

In other partial differential equations examples, the function space framework $W(0, T; Y)$ can be too restrictive. This is due to the fact that, in the linearized state equation given in assumption **A3**, the term $f_y(\bar{y}, \bar{u})$ appears as a coefficient. Hence the solution of this linearized system may not be well-defined, unless f satisfies appropriate growth bounds. For this purpose an alternative functional framework can be appropriate. For instance, let $\hat{Y} \subset Y \subset X$ be three Hilbert spaces endowed with a chain of continuous embeddings. Duality is understood with respect to $X \equiv X'$. We set

$$\hat{W}(0, T; \hat{Y}) = \{y \in L^2(0, T; \hat{Y}) \mid \dot{y} \in L^2(0, T; X)\}.$$

We still have $\hat{W}(0, T; \hat{Y}) \hookrightarrow C([0, T]; X)$, but in general we do not have $\hat{W}(0, T; \hat{Y}) \hookrightarrow C([0, T]; Y)$, except for instance when we have the interpolation $Y = [X; \hat{Y}]_{1/2}$. When we do not have this embedding, we can add $y \in C([0, T]; Y)$ into the definition of $\hat{W}(0, T; \hat{Y})$ above. In applications, for a given smooth domain Ω , we may think of $X = L^2(\Omega)$, and Y, \hat{Y} subspaces of $H^1(\Omega), H^2(\Omega)$, respectively. The mapping e introduced in the proof of Lemma 1 is then modified to be

$$e : \hat{W}(0, 2; \hat{Y}) \times L^2(0, 2; U) \times \mathbb{R} \rightarrow L^2(0, 2; X) \times Y.$$

The assumptions utilized in this proof are implied by **A1–A3**. In order to transpose the assumptions **A1–A3** and to derive differentiability in the framework considered here, we introduce the mapping $\hat{f} : \hat{Y} \times U \rightarrow X$ associated with f as a restriction. We would assume:

- B1** The mapping $\hat{f} : W(0, T; \hat{Y}) \times L^2(0, T; U) \times \mathbb{R} \rightarrow L^2(0, T; X)$ is twice continuously Fréchet-differentiable.
- B2** For all $(u, \tau) \in L^2(0, T; U) \times \mathbb{R}$, system (9) admits a unique solution in $\hat{W}(0, T; \hat{Y})$.
- B3** For all $\xi \in L^2(0, T; X), z_0 \in Y$, there exists a unique $z \in \hat{W}(0, T; \hat{Y})$ solution to the following system

$$\begin{cases} \dot{z} = \hat{f}_y(\bar{y}, \bar{u}) \cdot z + \xi & \text{on } (0, T), \\ z(0) = z_0. \end{cases}$$

This framework, where strong regularity is considered, is needed when the nonlinearity of the mapping f cannot be handled in the context of weak solutions. For instance, polynomial nonlinearities, as those occurring in the Schlögel and FitzHugh-Nagumo systems (see [22]), require the notion of strong solutions, in Hilbert spaces of type H^s (see [23]). These

are weak solutions which in addition lie in $\hat{W}(0, T; \hat{Y})$. In the context of parabolic equations, the time-dependent operator $\hat{f}_y(\bar{y}, \bar{u})$ can be studied with the approaches of [24] or [14] for instance.

For a sake of being specific, in the rest of the paper we will keep the framework of the function space $W(0, T; Y)$, with assumptions **A1**–**A3**. Analogous results as those obtained for $W(0, T; Y)$ also hold for $\hat{W}(0, T; \hat{Y})$.

2.3 Transformation of the problem

The formulation of problem (4) does not enable us to derive optimality conditions. Indeed, the control is not continuous at the optimal time τ , in general. Therefore, the trajectory is not differentiable at τ and the cost is not differentiable with respect to τ . This difficulty can be circumvented by introducing the following change of variables, $\pi(\cdot, \tau) : [0, 2] \rightarrow [0, T]$, for all $\tau \in (0, T)$,

$$\pi(s, \tau) = \begin{cases} \tau s & \text{if } s \in [0, 1], \\ (T - \tau)s + 2\tau - T & \text{if } s \in [1, 2]. \end{cases}$$

Observe that $\pi(1, \tau) = \tau$. For future reference we introduce the time-derivative $\dot{\pi}$ of π (with respect to s), as well its the partial derivative $\dot{\pi}_\tau$ with respect to τ :

$$\dot{\pi}(s, \tau) = \begin{cases} \tau & \text{if } s \in [0, 1), \\ (T - \tau) & \text{if } s \in (1, 2], \end{cases} \quad \dot{\pi}_\tau(s, \tau) = \begin{cases} 1 & \text{if } s \in [0, 1), \\ -1 & \text{if } s \in (1, 2]. \end{cases}$$

Observe that $\dot{\pi}_\tau$ is actually independent of τ . To simplify the notation, we will simply write $\dot{\pi}_\tau(s)$. Given $\pi(\cdot, \tau)$ we introduce the following change of unknowns

$$y : s \mapsto \tilde{y} \circ \pi(s, \tau), \quad u : s \mapsto \tilde{u} \circ \pi(s, \tau), \quad s \in [0, 2]. \quad (8)$$

Then, for $(u, \tau) \in L^2(0, 2; U) \times (0, T)$, we are lead to consider the following system

$$\begin{cases} \dot{y} = \dot{\pi}(\cdot, \tau)f(y, u) & \text{on } (0, 2), \\ y(0) = y_0, \end{cases} \quad (9)$$

and the following reformulated optimal control problem:

$$\begin{cases} \max_{\tau \in (0, T), u \in L^2(0, 2; U)} J(u, \tau) := \int_0^2 \dot{\pi}(s, \tau)\ell(y(s), u(s))ds + \phi_1(y(1)) + \phi_2(y(2)) \\ \text{subject to: } \dot{y} = \dot{\pi}(\cdot, \tau)f(y, u) \text{ in } Y', \quad y(0) = y_0 \text{ in } X, \\ G(u, \tau) := \int_0^T \dot{\pi}(s, \tau)\|u(s)\|_U^2 ds - \gamma \leq 0. \end{cases} \quad (10)$$

Using the fact that $\pi(\cdot, \tau) \in W^{1,\infty}(0, 2; \mathbb{R})$, we can verify that assumptions **A0**–**A3** imply the following ones.

A0' The mapping $\ell : W(0, 2; Y) \times L^2(0, 2; U) \rightarrow L^1(0, 2; \mathbb{R})$ is twice continuously Fréchet-differentiable.

A1' The mapping $f : W(0, 2; Y) \times L^2(0, 2; U) \rightarrow L^2(0, 2; Y')$ is twice continuously Fréchet-differentiable.

A2' For all $(u, \tau) \in L^2(0, 2; U) \times (0, T)$, $y_0 \in X$, system (9) admits a unique weak solution in $W(0, 2; Y)$.

A3' For all $(y, u) \in W(0, 2; Y) \times L^2(0, 2; U)$, $\xi \in L^2(0, 2; Y')$, $z_0 \in X$, there exists a unique $z \in W(0, 2; Y)$ solution to the following system

$$\begin{cases} \dot{z} = \dot{\pi}(\cdot, \tau)f_y(y, u).z + \xi & \text{on } (0, 2), \\ z(0) = z_0. \end{cases}$$

Observe that the mapping $G : L^2(0, 2; U) \times (0, T) \rightarrow \mathbb{R}$ modeling the constraint is twice continuously Fréchet-differentiable.

The change of unknown (8) modifies the nature of the optimal control problem under investigation: While (4) looks like a shape optimization problem with variable domains $(0, \tau)$ and (τ, T) , problem (10) has the nature of a parametric optimization problem.

The equivalence of problems (4) and (10) is straightforward because, on one hand, the time-derivative of $\tilde{y} \circ \pi(\cdot, \tau)$ is expressed with the chain rule as $\dot{\pi}(\cdot, \tau)\dot{\tilde{y}} \circ \pi(\cdot, \tau) = \dot{\pi}(\cdot, \tau)f(\tilde{y} \circ \pi(\cdot, \tau), \tilde{u} \circ \pi(\cdot, \tau))$. On the other hand, the integral on

$(0, T)$ can be split on $(0, \tau) \cup (\tau, T)$, and the change of variables $\pi(\cdot, \tau)$ is used for transforming the integrals on $(0, \tau)$ and (τ, T) into the integrals on $(0, 1)$ and $(1, 2)$, respectively. To sum up, if $\tilde{u} \in L^2(0, T; U)$, $\tau \in (0, T)$ and $u = \tilde{u} \circ \pi(\cdot, \tau)$, the pair (\tilde{u}, τ) is an optimal solution of the original problem (4) if and only if the pair (u, τ) is an optimal solution of the reformulated problem (10).

3 First and second-order optimality conditions

In this section, we provide a first and second-order sensitivity analysis for the reformulated problem (10). Due to the presence of the parameter τ in addition to the control variable u , it is non-standard to obtain elegant representations for the first- and second-order derivatives. Once they are given, their expressions are used for obtaining necessary and sufficient optimality conditions, as well as for iterative numerical methods.

Throughout this section, $\bar{u} \in L^2(0, 2; U)$ is a fixed value of the control and $\bar{\tau} \in (0, T)$ a fixed value of the variable τ . Note that we do not follow up the special cases $\bar{\tau} = 0$ and $\bar{\tau} = T$. We denote by $\bar{y} \in W(0, 2; Y)$ the corresponding state variable, which is the solution of (9) for $(u, \tau) = (\bar{u}, \bar{\tau})$.

The approach we use for deriving optimality conditions is classical, as described in [15, 25], for instance. However, due to the time transformation and the additional optimization variable, special attention is required. Considering the state equation as a constraint of the optimization problem (10), our approach mainly consists in computing the first- and second-order derivatives of the associated Lagrangian.

3.1 Linearization of the system

We introduce the control-to-state mapping $(u, \tau) \in L^2(0, 2; U) \times (0, T) \mapsto S(u, \tau) \in W(0, 2; Y)$, where $S(u, \tau)$ is the solution of (9). Remember that $\bar{y} = S(\bar{u}, \bar{\tau})$. The differentiability properties of the mapping S derive from assumptions **A1'**–**A3'**.

Lemma 1. *The mapping S is twice continuously Fréchet-differentiable on $L^2(0, 2; U) \times (0, T)$. For $v \in L^2(0, 2; U)$, the derivatives $z = S_u(\bar{u}, \bar{\tau}) \cdot v$ and $w = S_\tau(\bar{u}, \bar{\tau})$ are the respective solutions – in the weak sense – of the following systems:*

$$\begin{cases} \dot{z} = \dot{\pi}(\cdot, \bar{\tau}) f_y(\bar{y}, \bar{u}) \cdot z + \dot{\pi} f_u(\bar{y}, \bar{u}) \cdot v & \text{on } (0, 2), \\ z(0) = 0, \end{cases} \quad \begin{cases} \dot{w} = \dot{\pi}(\cdot, \bar{\tau}) f_y(\bar{y}, \bar{u}) \cdot w + \dot{\pi}_\tau f(\bar{y}, \bar{u}) & \text{on } (0, 2), \\ w(0) = 0. \end{cases}$$

Proof. Consider the mapping

$$e : W(0, 2; Y) \times L^2(0, 2; U) \times (0, T) \rightarrow L^2(0, 2; Y') \times X, \\ (y, u, \tau) \mapsto (y - \dot{\pi}(\cdot, \tau) f(y, u), y(0) - y_0).$$

Since we have the identity $e(S(u, \tau), u, \tau) = 0$, assumptions **A1'**–**A3'** enable us to apply the implicit function theorem, in the same fashion as [25], Sect. 1.6, pages 57-58. In fact, invertibility of the mapping $e_y(y, u, \tau)$ is a consequence of **A3'**, and the required smoothness conditions follow from **A1'**. The result then follows. \square

Corollary 1. *The cost function J is twice continuously Fréchet differentiable.*

Proof. The following mapping is twice continuously Fréchet differentiable

$$(y, u, \tau) \in W(0, 2; Y) \times L^2(0, 2; U) \times (0, T) \mapsto \int_0^2 \dot{\pi}(s, \tau) \ell(y(s), u(s)) ds + \phi_1(y(1)) + \phi_2(y(2)),$$

by **A0'**, and because ϕ_1 and ϕ_2 are twice continuously Fréchet differentiable. Therefore, J is also twice continuously Fréchet differentiable, by composition by S . \square

3.2 Vectorial formalism

We define the functional space \mathcal{Y} and its dual space as follows:

$$\mathcal{Y} = X \times X \times L^2(0, 2; Y), \quad \mathcal{Y}' = X \times X \times L^2(0, 2; Y').$$

Next we introduce the mapping **S** as follows:

$$\mathbf{S} : (u, \tau) \in L^2(0, 2; U) \times (0, T) \mapsto (S(u, \tau)(1), S(u, \tau)(2), S(u, \tau)) \in \mathcal{Y}.$$

As a consequence of Lemma 1, the mapping \mathbf{S} is twice continuously differentiable. Its first-order derivatives are given by:

$$\mathbf{S}_u(u, \tau) = (S_u(u, \tau)(1), S_u(u, \tau)(2), S_u(u, \tau)), \quad \mathbf{S}_\tau(u, \tau) = (S_\tau(u, \tau)(1), S_\tau(u, \tau)(2), S_\tau(u, \tau)). \quad (11)$$

Let us define the operator $\mathcal{K} \in \mathcal{L}(L^2(0, 2; Z); \mathcal{Y})$ by

$$\mathcal{K} : \xi \mapsto (z(1), z(2), z)$$

where $z \in W(0, 2; Y)$ is defined – in virtue of assumption **A3** – as the solution of

$$\begin{cases} \dot{z} = \dot{\pi}(\cdot, \bar{\tau}) f_y(\bar{y}, \bar{u}) \cdot z + \xi & \text{on } (0, 2), \\ z(0) = 0. \end{cases} \quad (12)$$

Lemma 2. *The adjoint $\mathcal{K}^* \in \mathcal{L}(\mathcal{Y}'; L^2(0, 2; Z'))$ of \mathcal{K} is given by $\mathcal{K}^*(a, b, w) = q$, where q is the solution of*

$$\begin{cases} -\dot{q} = \dot{\pi}(\cdot, \bar{\tau}) f_y(\bar{y}, \bar{u})^* \cdot q + w & \text{on } (0, 1) \cup (1, 2), \\ q(2) = b, \\ q(1^+) - q(1^-) + a = 0. \end{cases} \quad (13)$$

Proof. Let $\xi \in L^2(0, 2; Z)$ and let z be the solution of system (12) corresponding to ξ . Let $(a, b, w) \in \mathcal{Y}'$ and denote by q the solution of system (13) corresponding to (a, b, w) . We calculate by integration by parts

$$\begin{aligned} \langle (a, b, w); \mathcal{K}(\xi) \rangle_{\mathcal{Y}'; \mathcal{Y}} &= \langle w; z \rangle_{L^2(0, 2; Y); L^2(0, 2; Y)} + \langle a; z(1) \rangle_X + \langle b; z(2) \rangle_X \\ &= \int_0^1 \langle -\dot{q} - \dot{\pi} f_y^*(\bar{y}, \bar{u}) \cdot q; z \rangle_{Y'; Y} ds + \int_1^2 \langle -\dot{q} - \dot{\pi} f_y^*(\bar{y}, \bar{u}) \cdot q; z \rangle_{Y'; Y} ds \\ &\quad + \langle a; z(1) \rangle_X + \langle b; z(2) \rangle_X \\ &= \int_0^1 \langle q; \dot{z} - \dot{\pi} f_y(\bar{y}, \bar{u}) \cdot z \rangle_{Z'; Z} ds - \langle q(1^-); z(1) \rangle_X + \int_1^2 \langle q; \dot{z} - \dot{\pi} f_y(\bar{y}, \bar{u}) \cdot z \rangle_{Z'; Z} ds \\ &\quad + \langle q(1^+); z(1) \rangle_X - \langle q(2); z(2) \rangle_X + \langle a; z(1) \rangle_X + \langle b; z(2) \rangle_X \\ &= \int_0^1 \langle q; \xi \rangle_{Z'; Z} ds + \int_1^2 \langle q; \xi \rangle_{Z'; Z} ds, \end{aligned}$$

which leads to $\langle (a, b, w); \mathcal{K}(\xi) \rangle_{\mathcal{Y}'; \mathcal{Y}} = \langle q; \xi \rangle_{L^2(0, 2; Z'); L^2(0, 2; Z)}$ and thus completes the proof. \square

Lemma 2 enables us to conveniently express the adjoint operators of $\mathbf{S}_u(\bar{u}, \bar{\tau})$ and $\mathbf{S}_\tau(\bar{u}, \bar{\tau})$.

Corollary 2. *The adjoint operators $\mathbf{S}_u(\bar{u}, \bar{\tau})^* \in \mathcal{L}(\mathcal{Y}'; L^2(0, 2; U'))$ and $\mathbf{S}_\tau(\bar{u}, \bar{\tau})^* \in \mathcal{L}(\mathcal{Y}'; \mathbb{R})$ are given by*

$$\begin{aligned} \mathbf{S}_u(\bar{u}, \bar{\tau})^* \cdot (a, b, w) &= \dot{\pi}(\cdot, \bar{\tau}) f_u^*(\bar{y}, \bar{u}) \mathcal{K}^*(a, b, w), \\ \mathbf{S}_\tau(\bar{u}, \bar{\tau})^* \cdot (a, b, w) &= \int_0^2 \dot{\pi}_\tau \langle f(\bar{y}, \bar{u}); \mathcal{K}^*(a, b, w) \rangle_{Y'; Y} ds. \end{aligned}$$

Proof. As a consequence of Lemma 1 and (11), we have $\mathbf{S}_u(\bar{u}, \bar{\tau}) = \dot{\pi}(\cdot, \bar{\tau}) \mathcal{K} \circ f_u(\bar{y}, \bar{u})$ and $\mathbf{S}_\tau(\bar{u}, \bar{\tau}) = \dot{\pi}_\tau \mathcal{K} \circ f(\bar{y}, \bar{u})$. The result now follows from Lemma 2. \square

3.3 Lagrangian formulation and computation of derivatives

We recall that $\bar{u} \in L^2(0, 2; U)$ and $\bar{\tau}$ are fixed values of the control and the time variable, respectively. We also fix $\bar{y} = S(\bar{u}, \bar{\tau})$ and $\bar{\mathbf{y}} = \mathbf{S}(\bar{u}, \bar{\tau}) = (\bar{y}(1), \bar{y}(2), \bar{y})$. We introduce the Hamiltonian:

$$\begin{aligned} H : Y \times U \times Z' \times \mathbb{R} &\mapsto \mathbb{R} \\ (y, u, p, \lambda) &\rightarrow \ell(y, u) + \langle p; f(y, u) \rangle_{Z'; Z} - \lambda \|u\|_U^2. \end{aligned}$$

The adjoint state is defined as the solution of the following linear system:

$$\begin{cases} -\dot{p} = \dot{\pi}(\cdot, \bar{\tau})H_y(\bar{y}, \bar{u}, p) & \text{on } (0, 1) \cup (1, 2), \\ p(2) = D\phi_2(\bar{y}(2)), \\ p(1^+) - p(1^-) + D\phi_1(\bar{y}(1)) = 0. \end{cases} \tag{14}$$

It satisfies:

$$\bar{p}|_{(0,1)} \in W(0, 1; Z'), \quad \bar{p}|_{(1,2)} \in W(1, 2; Z').$$

In the system above, the variable λ does not appear, since H_y is independent of λ . Recall the continuous embeddings $W(I; Z') \hookrightarrow C(\bar{I}; X)$, for $I = (0, 1)$ and $I = (1, 2)$. In order to solve this backward system, we first consider $\bar{p}(2) = D\phi_2(\bar{y}(2))$ as the initial condition in X , next compute \bar{p} on $(1, 2)$ according to the first equation of (14), deduce $\bar{p}(1^-)$ from $\bar{p}(1^+)$ with the transmission condition in X , and finally compute \bar{p} on $(0, 1)$ as previously. From a more abstract point of view, the affine mapping $p \mapsto H_y(\bar{y}, \bar{u}, p) = f_y(\bar{y}, \bar{u})^* \cdot p + \ell_y(\bar{y}, \bar{u})$ is in the form of the right-hand-side in system (13), and thus Lemma 2 allows the existence and uniqueness of a solution to system (14).

We work with two Lagrangian functionals. The first one is defined as follows:

$$\begin{aligned} L : L^2(0, 2; U) \times (0, T) \times \mathbb{R} &\rightarrow \mathbb{R} \\ (u, \tau, \lambda) &\mapsto J(u, \tau) - \lambda G(u, \tau). \end{aligned}$$

We also consider the following extended Lagrangian:

$$\begin{aligned} \mathbf{L} : (X \times X \times W(0, 2; Y)) \times L^2(0, 2; U) \times (0, T) \times W(0, 2; Y) \times \mathbb{R} &\rightarrow \mathbb{R} \\ (\mathbf{y} = (a_1, a_2, y), u, \tau, p, \lambda) &\mapsto \phi_1(a_1) + \phi_2(a_2) + \int_0^2 (\dot{\pi}(s, \tau)H(y, u, p, \lambda)(s) - \langle p(s); \dot{y}(s) \rangle_{Z';Z}) ds \\ &\quad - \langle p(0); y(0) - y_0 \rangle_X + \langle p(2); y(2) - a_2 \rangle_X - \langle [p]_1; y(1) - a_1 \rangle_X, \end{aligned}$$

where we denote $[p]_1 = p(1^+) - p(1^-)$. Note that the functionals L and \mathbf{L} are both twice continuously differentiable. The following lemma enables us to derive the first and second-order derivatives of the L in a convenient way.

Lemma 3. *The following identity holds:*

$$L(u, \tau, \lambda) = \mathbf{L}(\mathbf{S}(u, \tau), u, \tau, p, \lambda), \quad \forall (u, \tau, p, \lambda) \in L^2(0, 2; U) \times (0, T) \times W(0, 2; Z') \times \mathbb{R}. \tag{15}$$

Moreover, for all $\lambda \in \mathbb{R}$ we have

$$\mathbf{L}_y(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}, \lambda) = 0 \quad \text{in } \mathcal{Y}'.$$

Proof. Identity (15) follows directly from the definitions of J, \mathbf{S}, L , and \mathbf{L} . We decompose $\mathbf{L}_y(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}, \lambda) \in \mathcal{Y}'$ into $\mathbf{L}_{a_1}(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}, \lambda) \in X, \mathbf{L}_{a_2}(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}, \lambda) \in X$ and $\mathbf{L}_y(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}, \lambda) \in Y'$. From the definition of the adjoint state, we obtain the following identities in X :

$$\mathbf{L}_{a_1}(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}, \lambda) = D\phi_1(\bar{y}(1)) + [\bar{p}]_1 = 0, \quad \mathbf{L}_{a_2}(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}, \lambda) = D\phi_2(\bar{y}(2)) - \bar{p}(2) = 0.$$

Moreover, for all $\delta y \in W(0, 2; Y)$, we obtain by integration by parts

$$\begin{aligned} \langle \mathbf{L}_y(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}, \lambda); \delta y \rangle_{Y';Y} &= \int_0^2 \dot{\pi}(s, \bar{\tau}) \langle H_y(\bar{y}, \bar{u}, \bar{p})(s); \delta y(s) \rangle_{Y';Y} ds - \int_0^2 \langle \bar{p}(s); \delta \dot{y}(s) \rangle_{Y';Y} ds \\ &\quad - \langle \bar{p}(0); \delta y(0) \rangle_X + \langle \bar{p}(2); \delta y(2) \rangle_X - \langle [\bar{p}]_1; \delta y(1) \rangle_X \\ &= 0, \end{aligned}$$

which concludes the proof. □

Proposition 1. *The first-order derivatives of L are given by:*

$$L_u(\bar{u}, \bar{\tau}, \lambda).v = \mathbf{L}_u(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}, \lambda).v = \int_0^2 \dot{\pi}(s, \bar{\tau})H_u(\bar{y}, \bar{u}, \bar{p}, \lambda)(s).v(s)ds, \tag{16}$$

$$L_\tau(\bar{u}, \bar{\tau}, \lambda) = \mathbf{L}_\tau(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}) = \int_0^2 \dot{\pi}_\tau H(\bar{y}, \bar{u}, \bar{p}, \lambda)(s) ds. \quad (17)$$

Proof. The result follows directly from Lemma 3. Applying the chain rule to (15), we obtain:

$$DL(\bar{u}, \bar{\tau}, \lambda) = \begin{pmatrix} \mathbf{L}_y(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}, \lambda) \mathbf{S}_u(\bar{u}, \bar{\tau}) + \mathbf{L}_u(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}, \lambda) \\ \mathbf{L}_y(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}, \lambda) \mathbf{S}_\tau(\bar{u}, \bar{\tau}) + \mathbf{L}_\tau(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}, \lambda) \end{pmatrix}.$$

Since from Lemma 3 we have $\mathbf{L}_y(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}, \lambda) = 0$, formulas (16) and (17) hold. \square

In the following proposition, we calculate the Hessian of L easily, thanks to the Lagrangian formalism described above. The fact that $\mathbf{L}_y(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}, \lambda) = 0$ is a key property here. We denote by \mathbf{I} the linear identity mapping in $L^2(0, 2; U')$.

Proposition 2. *The second-order derivative of L (with respect to (u, τ)) is given by:*

$$D^2L(\bar{u}, \bar{\tau}, \lambda) = \begin{pmatrix} \mathbf{S}_u^*(\bar{u}, \bar{\tau}) & \mathbf{I} & 0 \\ \mathbf{S}_\tau^*(\bar{u}, \bar{\tau}) & 0 & 1 \end{pmatrix} D^2\mathbf{L}(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}, \lambda) \begin{pmatrix} \mathbf{S}_u(\bar{u}, \bar{\tau}) & \mathbf{S}_\tau(\bar{u}, \bar{\tau}) \\ \mathbf{I} & 0 \\ 0 & 1 \end{pmatrix}.$$

The second-order derivatives read, in a more explicit form, as

$$\begin{aligned} D^2L(\bar{u}, \bar{\tau}, \lambda).((v, \theta), (\hat{v}, \hat{\theta})) &= D^2\phi_1(\bar{y}(1)).(z(1), \hat{z}(1)) + D^2\phi_2(\bar{y}(2)).(z(2), \hat{z}(2)) \\ &+ \int_0^2 \dot{\pi}(s, \bar{\tau}) D^2H(\bar{y}, \bar{u}, \bar{p}, \lambda).((z, v), (\hat{z}, \hat{v}))(s) ds \\ &+ \theta \int_0^2 \dot{\pi}_\tau(s) DH(\bar{y}, \bar{u}, \bar{p}, \lambda)(\hat{z}, \hat{v})(s) ds \\ &+ \hat{\theta} \int_0^2 \dot{\pi}_\tau(s) DH(\bar{y}, \bar{u}, \bar{p}, \lambda)(z, v)(s) ds, \end{aligned}$$

where $z = S_u(\bar{u}, \bar{\tau}).v + S_\tau(\bar{u}, \bar{\tau}).\theta$ and $\hat{z} = S_u(\bar{u}, \bar{\tau}).\hat{v} + S_\tau(\bar{u}, \bar{\tau}).\hat{\theta}$.

Proof. Once again, the proposition follows directly from Lemma 3. Applying the chain rule to (15), we obtain:

$$D^2L(\bar{u}, \bar{\tau}, \lambda) = \begin{pmatrix} \mathbf{S}_u^*(\bar{u}, \bar{\tau}) & \mathbf{I} & 0 \\ \mathbf{S}_\tau^*(\bar{u}, \bar{\tau}) & 0 & 1 \end{pmatrix} D^2\mathbf{L}(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}) \begin{pmatrix} \mathbf{S}_u(\bar{u}, \bar{\tau}) & \mathbf{S}_\tau(\bar{u}, \bar{\tau}) \\ \mathbf{I} & 0 \\ 0 & 1 \end{pmatrix} + \mathbf{L}_y(\bar{y}, \bar{u}, \bar{\tau}) D^2\mathbf{S}(\bar{u}, \bar{\tau}).$$

The term involving $D^2\mathbf{S}(\bar{u}, \bar{\tau})$ vanishes, since $\mathbf{L}_y(\bar{y}, \bar{u}, \bar{\tau}, \bar{p}) = 0$. The explicit form follows from the compact relation below, where the notation of the variable $(\bar{y}, \bar{u}, \bar{\tau}, \bar{p})$ has been omitted, for a sake of clarity:

$$D^2\mathbf{L} = \begin{pmatrix} D^2\phi_1(\bar{y}(1)) & 0 & 0 & 0 & 0 \\ 0 & D^2\phi_2(\bar{y}(2)) & 0 & 0 & 0 \\ 0 & 0 & \mathbf{L}_{yy} & \mathbf{L}_{yu} & \mathbf{L}_{y\tau} \\ 0 & 0 & \mathbf{L}_{uy} & \mathbf{L}_{uu} & \mathbf{L}_{u\tau} \\ 0 & 0 & \mathbf{L}_{\tau y} & \mathbf{L}_{\tau u} & 0 \end{pmatrix}.$$

Denoting $\mathbf{z} = (z(1), z(2), z)$ and $\hat{\mathbf{z}} = (\hat{z}(1), \hat{z}(2), \hat{z})$, the partial derivatives above are formally given by

$$\left\langle \begin{pmatrix} \mathbf{L}_{yy} & \mathbf{L}_{yu} \\ \mathbf{L}_{uy} & \mathbf{L}_{uu} \end{pmatrix} \cdot \begin{pmatrix} z \\ v \end{pmatrix}; \begin{pmatrix} \hat{z} \\ \hat{v} \end{pmatrix} \right\rangle_{L^2(0,2;Y);L^2(0,2;Y)} = \int_0^2 \dot{\pi}(s, \bar{\tau}) D^2H(\bar{y}, \bar{u}, \bar{p}, \lambda)(s).((z, v), (\hat{z}, \hat{v}))(s) ds,$$

$$\left\langle \begin{pmatrix} \mathbf{L}_{\tau y} \\ \mathbf{L}_{\tau u} \end{pmatrix}; \begin{pmatrix} z \\ v \end{pmatrix} \right\rangle_{L^2(0,2;Y);L^2(0,2;Y)} = \int_0^2 \dot{\pi}_\tau DH(\bar{y}, \bar{u}, \bar{p}, \lambda)(s) \cdot (z, v)(s) ds.$$

So the proof is complete. □

3.4 Optimality conditions

We give in this subsection necessary and sufficient optimality conditions. They involve the first and second-order derivatives of L , which have been calculated previously. First we introduce some notation. For $(v, \theta) \in L^2(0, 2; U) \times \mathbb{R}$, we denote

$$\|(v, \theta)\| = \left(\|v\|_{L^2(0,2;U)}^2 + \theta^2 \right)^{1/2}.$$

Given $\varepsilon > 0$, we denote

$$B_\varepsilon(\bar{u}, \bar{\tau}) = \{(u, \tau) \in L^2(0, 2; U) \times (0, T) : \|(u - \bar{u}, \tau - \bar{\tau})\| < \varepsilon\}.$$

3.4.1 First-order optimality conditions

Proposition 3. *If the pair $(\bar{u}, \bar{\tau})$ is locally optimal, then there exists a unique $\bar{\lambda} \geq 0$ such that*

$$L_u(\bar{u}, \bar{\tau}, \bar{\lambda}) = 0, \quad L_\tau(\bar{u}, \bar{\tau}, \bar{\lambda}) = 0 \quad \text{and} \quad \bar{\lambda} G(\bar{u}, \bar{\tau}) = 0. \tag{18}$$

Proof. If the constraint is not active (i.e. $G(\bar{u}, \bar{\tau}) < 0$), then the result is clearly satisfied with $\bar{\lambda} = 0$. If the constraint is active (i.e. $G(\bar{u}, \bar{\tau}) = 0$), then $\bar{u} \neq 0$ and therefore $D_u G(\bar{u}, \bar{\tau})$ is non-zero. As a consequence, the linear independence condition of qualification holds, which ensures the existence and uniqueness of a Lagrange multiplier $\bar{\lambda}$ satisfying the Karush-Kuhn-Tucker conditions (18). □

Remark 1. When the Hamiltonian is strongly uniformly convex with respect to u , the optimal control \bar{u} inherits some regularity properties of \bar{y} and \bar{p} . In the current framework, the optimal control is not continuous at time 1 (in general), because of the jump of the adjoint state.

3.4.2 Second-order: Case of an inactive constraint

Proposition 4. *If $(\bar{u}, \bar{\tau})$ is a solution to problem (10) such that $G(\bar{u}, \bar{\tau}) < 0$, then for all $(v, \theta) \in L^2(0, 2; U) \times \mathbb{R}$,*

$$D^2 J(\bar{u}, \bar{\tau}) \cdot (v, \theta)^2 \leq 0.$$

Conversely, if $(\bar{u}, \bar{\tau})$ is feasible, if $DJ(\bar{u}, \bar{\tau}) = 0$, and if there exists $\alpha > 0$ such that for all $(v, \theta) \in L^2(0, 2; U) \times \mathbb{R}$

$$D^2 J(\bar{u}, \bar{\tau}) \cdot (v, \theta)^2 \leq -\alpha \|(v, \theta)\|^2, \tag{19}$$

then $(\bar{u}, \bar{\tau})$ is a local solution to problem (10). More precisely, for all $\beta \in (0, \alpha)$, there exists $\varepsilon > 0$ such that for all $(u, \tau) \in B_\varepsilon(\bar{u}, \bar{\tau})$,

$$J(u, \tau) \leq J(\bar{u}, \bar{\tau}) - \frac{\beta}{2} \|(u - \bar{u}, \tau - \bar{\tau})\|^2. \tag{20}$$

Proof. If $(\bar{u}, \bar{\tau})$ is a solution to (10), then for $\varepsilon > 0$ small enough the pair $(\bar{u} + \varepsilon v, \bar{\tau} + \varepsilon \theta)$ is feasible and therefore

$$0 \geq \lim_{\varepsilon \downarrow 0} \frac{J(\bar{u} + \varepsilon v, \bar{\tau} + \varepsilon \theta) - J(\bar{u}, \bar{\tau})}{\varepsilon^2} = \frac{1}{2} D^2 J(\bar{u}, \bar{\tau}) \cdot (v, \theta)^2,$$

because $DJ(\bar{u}, \bar{\tau}) = 0$. Conversely, assume that $DJ(\bar{u}, \bar{\tau}) = 0$ and that (19) holds. Let $\varepsilon > 0$ be such that for all $(u, \tau) \in B_\varepsilon(\bar{u}, \bar{\tau})$, we have

$$|(D^2 J(u, \tau) - D^2 J(\bar{u}, \bar{\tau})) \cdot (v, \theta)^2| \leq (\alpha - \beta) \|(v, \theta)\|^2. \tag{21}$$

By Taylor's Theorem, for all $(u, \tau) \in B_\varepsilon(\bar{u}, \bar{\tau})$, there exists $\mu \in [0, 1]$ such that

$$J(u, \tau) - J(\bar{u}, \bar{\tau}) = \frac{1}{2} D^2 J(\bar{u} + \mu v, \bar{\tau} + \mu \theta) \cdot (v, \theta)^2,$$

where $v = u - \bar{u}$ and $\theta = \tau - \bar{\tau}$.

We obtain

$$\begin{aligned} J(u, \tau) - J(\bar{u}, \bar{\tau}) &\leq \frac{1}{2} D^2 J(\bar{u}, \bar{\tau}) \cdot (v, \theta)^2 + \frac{1}{2} \left| (D^2 J(\bar{u}, \bar{\tau}) - D^2 J(\bar{u} + \mu v, \bar{\tau} + \mu \theta)) \cdot (v, \theta)^2 \right| \\ &\leq -\frac{1}{2} \alpha \| (v, \theta) \|^2 + \frac{1}{2} (\alpha - \beta) \| (v, \theta) \|^2 = -\frac{1}{2} \beta \| (v, \theta) \|^2, \end{aligned}$$

which completes the proof. \square

3.4.3 Second-order analysis: Case of an active constraint with strict complementarity.

All along the paragraph, the pair $(\bar{u}, \bar{\tau})$ is assumed to be such that the first-order necessary optimality condition of Proposition 3 holds with $\bar{\lambda} > 0$ and $G(\bar{u}, \bar{\tau}) = 0$. We do not consider the case of a null Lagrange multiplier with an active constraint. We define

$$w(\cdot) = 2\bar{\pi}(\cdot, \bar{\tau})\bar{u}(\cdot) \in L^2(0, 2; U) \text{ and } r = \int_0^2 \bar{\pi}_\tau(s) \|\bar{u}(s)\|_{\bar{U}}^2 ds,$$

so that

$$DG(\bar{u}, \bar{\tau}) \cdot (v, \theta) = \langle w; v \rangle_{L^2(0, 2; U)} + r\theta.$$

Observe that $\bar{u} \neq 0$, because $G(\bar{u}, \bar{\tau}) = 0$, and thus $w \neq 0$. We also set

$$(\bar{w}, \bar{r}) = \frac{(w, r)}{\|(w, r)\|} \in L^2(0, 2; U) \times \mathbb{R}.$$

and define a linear form $\zeta : L^2(0, 2; U) \times \mathbb{R} \rightarrow \mathbb{R}$ as follows:

$$\zeta(v, \theta) = \langle \bar{w}; v \rangle_{L^2(0, 2; U)} + \bar{r}\theta.$$

The first-order optimality condition reads

$$DJ(\bar{u}, \bar{\tau}) - \bar{\lambda} \|(w, r)\| \zeta = 0. \quad (22)$$

For all $\varepsilon \geq 0$, we define the cone \mathcal{C}_ε as follows:

$$\mathcal{C}_\varepsilon = \{ (v, \theta) \in L^2(0, 2; U) \times \mathbb{R} : |\zeta(v, \theta)| \leq \varepsilon \| (v, \theta) \| \}.$$

The cone \mathcal{C}_0 is called critical cone. The following lemma is a metric regularity property. In the proofs of the following results, C is used as a generic constant.

Lemma 4. *There exist $\varepsilon > 0$, a constant $\bar{C} > 0$, and a twice continuously differentiable mapping $\varphi : (u, \tau) \in B_\varepsilon(\bar{u}, \bar{\tau}) \rightarrow L^2(0, 2; U) \times (0, T)$ such that, for all $(u, \tau) \in B_\varepsilon(\bar{u}, \bar{\tau})$,*

$$G(\varphi(u, \tau)) = 0 \text{ and } \|\varphi(u, \tau) - (u, \tau) + \zeta(u - \bar{u}, \tau - \bar{\tau})(\bar{w}, \bar{r})\| \leq \bar{C} \|(u - \bar{u}, \tau - \bar{\tau})\|^2. \quad (23)$$

Proof. Consider the following twice continuously differentiable mapping:

$$\chi : (u, \tau, \pi) \in L^2(0, 2; U) \times (0, T) \times \mathbb{R} \mapsto G((u, \tau) + \pi(\bar{w}, \bar{r})) \in \mathbb{R}. \quad (24)$$

We have: $\chi(\bar{u}, \bar{\tau}, 0) = 0$, $\chi_\pi(\bar{u}, \bar{\tau}, 0) = \langle w; \bar{w} \rangle_{L^2(0, 2; U)} + r\bar{r} = \|(w, r)\| \neq 0$. Therefore, by the implicit function theorem, the non-linear equation $\chi(u, \tau, \pi) = 0$ with unknown $\pi \in \mathbb{R}$ possesses a solution for (u, τ) close enough to $(\bar{u}, \bar{\tau})$. More precisely, there exist $\varepsilon > 0$ and a twice continuously differentiable mapping $\Pi : B_\varepsilon(\bar{u}, \bar{\tau}) \rightarrow \mathbb{R}$ such that $\Pi(\bar{u}, \bar{\tau}) = 0$ and such that

$$G((u, \tau) + \Pi(u, \tau)(\bar{w}, \bar{r})) = 0, \quad \forall (u, \tau) \in B_\varepsilon(\bar{u}, \bar{\tau}). \quad (25)$$

Differentiating (25) with respect to (u, τ) at $(\bar{u}, \bar{\tau}, 0)$, we obtain

$$D\Pi(\bar{u}, \bar{\tau}) \cdot (v, \theta) = -\zeta(v, \theta), \quad \forall (v, \theta) \in L^2(0, 2; U) \times \mathbb{R}.$$

Since Π is twice continuously differentiable, there exists a constant C such that for all $(u, \tau) \in B_\varepsilon(\bar{u}, \bar{\tau})$,

$$|\Pi(u, \tau) - \underbrace{\Pi(\bar{u}, \bar{\tau})}_{=0} - D\Pi(\bar{u}, \bar{\tau})(u - \bar{u}, \tau - \bar{\tau})| \leq C \|(u - \bar{u}, \tau - \bar{\tau})\|^2.$$

Therefore, there exists a constant C such that for all $(u, \tau) \in B_\varepsilon(\bar{u}, \bar{\tau})$,

$$\|\Pi(u, \tau)(\bar{w}, \bar{r}) + \zeta(u - \bar{u}, \tau - \bar{\tau})\| \leq C\|(u - \bar{u}, \tau - \bar{\tau})\|^2. \tag{26}$$

We finally define $\varphi : B_\varepsilon(\bar{u}, \bar{\tau}) \rightarrow L^2(0, 2; U) \times (0, T)$ by

$$\varphi(u, \tau) = (u, \tau) + \Pi(u, \tau)(\bar{w}, \bar{r}).$$

The two properties specified in (24) follow directly from (25) and (26). □

Proposition 5. *Let $(\bar{u}, \bar{\tau})$ be locally optimal. Furthermore, assume that $G(\bar{u}, \bar{\tau}) = 0$. Then,*

$$D^2L(\bar{u}, \bar{\tau}, \bar{\lambda}).((v, \theta), (v, \theta)) \leq 0, \quad \forall (v, \theta) \in \mathcal{C}_0.$$

Proof. Let $(v, \theta) \in \mathcal{C}_0$. Let $(\varepsilon_k)_{k \in \mathbb{N}} \downarrow 0$. Set $(u_k, \tau_k) = (\bar{u}, \bar{\tau}) + \varepsilon_k(v, \theta)$. For k large enough, Lemma 4 applies and therefore,

$$\|\varphi(u_k, \tau_k) - (u_k, \tau_k)\| \leq C\varepsilon_k^2, \tag{27}$$

since $\zeta(u_k - \bar{u}, \tau_k - \bar{\tau}) = \varepsilon_k \zeta(v, \theta) = 0$. We have $\varphi(u_k, \tau_k) \rightarrow (\bar{u}, \bar{\tau})$, thus together with the feasibility of $\varphi(u_k, \tau_k)$, for k large enough,

$$0 \leq J(\varphi(u_k, \tau_k)) - J(\bar{u}, \bar{\tau}).$$

Since $G(\bar{u}, \bar{\tau}) = 0$ and since $G(\varphi(u_k, \tau_k)) = 0$ for all $k \in \mathbb{N}$,

$$0 \leq J(\varphi(u_k, \tau_k)) - J(\bar{u}, \bar{\tau}) = L(\varphi(u_k, \tau_k), \bar{\lambda}) - L(\bar{u}, \bar{\tau}, \bar{\lambda}),$$

for k large enough. The Lagrangian L is twice continuously differentiable and $DL(\bar{u}, \bar{\tau}) = 0$, and so there exists $\mu_k \in [0, 1]$ such that

$$0 \leq \frac{1}{\varepsilon_k^2} (L(\varphi(u_k, \tau_k), \bar{\lambda}) - L(\bar{u}, \bar{\tau}, \bar{\lambda})) = D^2L(\tilde{u}_k, \tilde{\tau}_k, \bar{\lambda}). \left(\frac{\varphi(u_k, \tau_k) - (\bar{u}, \bar{\tau})}{\varepsilon_k}, \frac{\varphi(u_k, \tau_k) - (\bar{u}, \bar{\tau})}{\varepsilon_k} \right), \tag{28}$$

where $(\tilde{u}_k, \tilde{\tau}_k) = \mu_k(\bar{u}, \bar{\tau}) + (1 - \mu_k)(u_k, \tau_k)$. Using (27), we obtain that:

$$\left\| \frac{\varphi(u_k, \tau_k) - (\bar{u}, \bar{\tau})}{\varepsilon_k} - (v, \theta) \right\| = \frac{1}{\varepsilon_k} \|\varphi(u_k, \tau_k) - (u_k, \tau_k)\| \leq C\varepsilon_k \xrightarrow{k \rightarrow \infty} 0.$$

Moreover, $(\tilde{u}_k, \tilde{\tau}_k) \xrightarrow{k \rightarrow \infty} (\bar{u}, \bar{\tau})$, therefore, we can pass to the limit in (28). We obtain

$$0 \leq D^2L(\bar{u}, \bar{\tau}, \bar{\lambda}).(v, \theta)^2,$$

which concludes the proof. □

Consider now the following sufficient second-order optimality condition: There exists $\alpha > 0$ such that

$$D^2L(\bar{u}, \bar{\tau}, \bar{\lambda}).(v, \theta)^2 \leq -\alpha\|(v, \theta)\|^2, \quad \forall (v, \theta) \in \mathcal{C}_0. \tag{29}$$

Lemma 5. *Assume that the sufficient second-order optimality condition (29) holds. Then, for all $0 < \beta < \alpha$, there exists $\varepsilon > 0$ such that*

$$D^2L(\bar{u}, \bar{\tau}, \bar{\lambda})(v, \theta)^2 \leq -\beta\|(v, \theta)\|^2, \quad \forall (v, \theta) \in \mathcal{C}_\varepsilon.$$

Proof. To simplify the notation, we write D^2L instead of $D^2L(\bar{u}, \bar{\tau}, \bar{\lambda})$. Let $\varepsilon > 0$ and let $(v, \theta) \in \mathcal{C}_\varepsilon$. Let us set $(v', \theta') = \zeta(v, \theta)(\bar{w}, \bar{r})$. We have

$$(v, \theta) - (v', \theta') \in \mathcal{C}_0,$$

since $\zeta(\bar{w}, \bar{r}) = 1$. Therefore,

$$\begin{aligned} D^2L.(v - v', \theta - \theta')^2 &\leq -\alpha\|(v - v', \theta - \theta')\|^2 \\ &= -\alpha(\|(v, \theta)\|^2 + \|(v', \theta')\|^2 - 2\langle v; v' \rangle_{L^2(0,2;U)} - \theta\theta') \\ &= -\alpha(\|(v, \theta)\|^2 - \zeta(v, \theta)^2) \\ &\leq (-\alpha + \alpha\varepsilon^2)\|(v, \theta)\|^2. \end{aligned} \tag{30}$$

We also have

$$\begin{aligned} D^2L.(v - v', \theta - \theta')^2 &= D^2L.(v, \theta)^2 - 2D^2L.((v, \theta), (v', \theta')) + D^2L.(v', \theta')^2 \\ &= D^2L.(v, \theta)^2 - 2\zeta(v, \theta)D^2L.((v, \theta), (\bar{w}, \bar{r})) + \zeta(v, \theta)^2 D^2L.(\bar{w}, \bar{r})^2 \\ &\geq D^2L.(v, \theta)^2 - C(\varepsilon + \varepsilon^2)\|(v, \theta)\|^2. \end{aligned} \quad (31)$$

Combining (30) and (31), we obtain

$$D^2L.((v, \theta), (v, \theta)) \leq (-\alpha + C\varepsilon + C\varepsilon^2)\|(v, \theta)\|^2.$$

For any $\beta \in (0, \alpha)$, there exists $\varepsilon > 0$ small enough so that $-\alpha + C\varepsilon + C\varepsilon^2 \leq -\beta$. The lemma is proved. \square

Proposition 6. *Assume that the sufficient second-order optimality condition (29) holds and that $\bar{\lambda} > 0$. Then, for all $\beta \in (0, \alpha)$, there exists $\varepsilon > 0$ such that for all $(u, \tau) \in B_\varepsilon(\bar{u}, \bar{\tau})$, if $G(u, \tau) \leq 0$, then*

$$J(u, \tau) - J(\bar{u}, \bar{\tau}) \leq -\frac{1}{2}\beta\|(u - \bar{u}, \tau - \bar{\tau})\|^2.$$

Proof. We prove the result by contradiction. If such an $\varepsilon > 0$ does not exist, then there exists a convergent sequence $(u_k, \tau_k)_{k \in \mathbb{N}}$ with limit $(\bar{u}, \bar{\tau})$ such that for all $k \in \mathbb{N}$, $G(u_k, \tau_k) \leq 0$ and such that

$$J(u_k, \tau_k) - J(\bar{u}, \bar{\tau}) > -\frac{1}{2}\beta\|(u_k - \bar{u}, \tau_k - \bar{\tau})\|^2. \quad (32)$$

Inequality (32) implies that $(u_k, \tau_k) \neq (\bar{u}, \bar{\tau})$. We set

$$(v_k, \theta_k) = \frac{(u_k - \bar{u}, \tau_k - \bar{\tau})}{\|(u_k - \bar{u}, \tau_k - \bar{\tau})\|}.$$

Since J is twice continuously differentiable and since $(u_k, \tau_k)_{k \in \mathbb{N}}$ is a bounded sequence, there exists a constant $C > 0$ such that

$$\begin{aligned} J(u_k, \tau_k) - J(\bar{u}, \bar{\tau}) &\leq DJ(\bar{u}, \bar{\tau})(u_k - \bar{u}, \tau_k - \bar{\tau}) + C\|(u_k - \bar{u}, \tau_k - \bar{\tau})\|^2 \\ &= \bar{\lambda}\|(w, r)\|\zeta(v_k, \theta_k)\|(u_k - \bar{u}, \tau_k - \bar{\tau})\| + C\|(u_k - \bar{u}, \tau_k - \bar{\tau})\|^2, \end{aligned} \quad (33)$$

where we used (22). Since $\bar{\lambda} > 0$, we obtain by combining (32) and (33) that

$$\zeta(v_k, \theta_k) \geq -C\|(u_k - \bar{u}, \tau_k - \bar{\tau})\| \xrightarrow[k \rightarrow \infty]{} 0.$$

Thus, $\liminf_{k \rightarrow \infty} \zeta(v_k, \theta_k) \geq 0$. Since (u_k, τ_k) is bounded and feasible and since G is twice continuously differentiable, there exists a constant $C > 0$ such that for all $k \in \mathbb{N}$,

$$0 \geq G(u_k, \tau_k) - G(\bar{u}, \bar{\tau}) = \|(w, r)\|\zeta(v_k, \theta_k)\|(u_k - \bar{u}, \tau_k - \bar{\tau})\| - C\|(u_k - \bar{u}, \tau_k - \bar{\tau})\|^2.$$

Therefore,

$$\zeta(v_k, \theta_k) \leq C\|(u_k - \bar{u}, \tau_k - \bar{\tau})\| \xrightarrow[k \rightarrow \infty]{} 0.$$

Thus, $\limsup_{k \rightarrow \infty} \zeta(v_k, \theta_k) \leq 0$ and finally, $\zeta(v_k, \theta_k) \xrightarrow[k \rightarrow \infty]{} 0$.

We are ready to obtain a contradiction to (32). We have $G(\bar{u}, \bar{\tau}) = 0$. Moreover, for all $k \in \mathbb{N}$, $G(u_k, \tau_k) \leq 0$. Thus

$$J(u_k, \tau_k) - J(\bar{u}, \bar{\tau}) \leq L(u_k, \tau_k, \bar{\lambda}) - L(\bar{u}, \bar{\tau}, \bar{\lambda}). \quad (34)$$

Since L is twice continuously differentiable and since $DL(\bar{u}, \bar{\tau}, \bar{\lambda}) = 0$, there exists $\mu_k \in [0, 1]$ such that

$$L(u_k, \tau_k, \bar{\lambda}) - L(\bar{u}, \bar{\tau}, \bar{\lambda}) = \frac{1}{2}\|(u_k - \bar{u}, \tau_k - \bar{\tau})\|^2 D^2L(\tilde{u}_k, \tilde{\tau}_k, \bar{\lambda}).(v_k, \theta_k)^2, \quad (35)$$

where $(\tilde{u}_k, \tilde{\tau}_k) = (1 - \mu_k)(\bar{u}, \bar{\tau}) + \mu_k(u_k, \tau_k)$. Observe that $(\tilde{u}_k, \tilde{\tau}_k) \xrightarrow[k \rightarrow \infty]{} (\bar{u}, \bar{\tau})$. We introduce now a number $\eta \in (\beta, \alpha)$.

For k large enough, $\|D^2L(\tilde{u}_k, \tilde{\tau}_k, \bar{\lambda}) - D^2L(\bar{u}, \bar{\tau}, \bar{\lambda})\| \leq \eta - \beta$. Thus, combining (34) and (35), we obtain

$$J(u_k, \tau_k) - J(\bar{u}, \bar{\tau}) \leq \frac{1}{2}\|(u_k - \bar{u}, \tau_k - \bar{\tau})\|^2 (D^2L(\bar{u}, \bar{\tau}, \bar{\lambda})(v_k, \theta_k)^2 + \eta - \beta). \quad (36)$$

By Lemma 5, there exists $\varepsilon > 0$ such that

$$D^2L(\bar{u}, \bar{\tau}, \bar{\lambda}).(v, \theta)^2 \leq -\eta\|(v, \theta)\|^2, \quad \forall (v, \theta) \in \mathcal{C}_\varepsilon.$$

For k large enough, $\zeta(v_k, \theta_k) \in \mathcal{C}_\varepsilon$ since $\zeta(v_k, \theta_k) \rightarrow 0$. Therefore, by (36),

$$J(u_k, \tau_k) - J(\bar{u}, \bar{\tau}) \leq -\frac{1}{2}\beta\|(u_k - \bar{u}, \tau_k - \bar{\tau})\|^2,$$

for k large enough, which contradicts (32). □

4 Numerical realization

4.1 Method

The numerical realization is based on a optimize then discretize approach using Newton’s method for the reduced formulation. For this purpose we need the first and second order sensitivity information which was obtained in Propositions 1 and 2, as well as the optimality conditions of Proposition 3. For ease of computations we did not realize numerically the norm constraint, and thus the expressions of the derivatives in Propositions 1 and 2 hold for J , as well as the optimality conditions of Proposition 3 (with $\lambda = 0$). In order to reach the region of attraction for Newton’s method, first some Barzilai-Borwein gradient steps are performed (see [26] for instance), and these steps are initialized with the Armijo rule. In these gradient steps, only two inner products are computed for each step. Next, when the norm of the gradient is small enough, we switch to the full-step Newton algorithm, which is faster, but demands one linear system solve at each iteration. Switching to the Newton step is monitored by the norm of the gradient given by

$$|||(J_u, J_\tau)||| := \|(J_u, J_\tau)\|_{L^2(0,2;\mathbb{R}^m) \times \mathbb{R}}^2 = \int_0^2 |J_u|_{\mathbb{R}^m}^2(s) ds + |J_\tau|_{\mathbb{R}}^2.$$

The integral above is approximated by the trapezoidal rule. The algorithm then performed is the following:

Algorithm 1 Solving the first-order optimality conditions

Initialization: $u = 0, \tau_0 = T/2$, for $1 \leq i \leq N, s(i) = 2i/N$.

Initial gradient: Compute (J_u, J_τ) corresponding to the initial data above.

Steps for vanishing the gradient:

Gradient steps: Compute another (J_u, J_τ) with the Armijo rule.

Barzilai-Borwein steps: While $|||(J_u, J_\tau)||| > 10^{-4}$, perform this gradient method.

Newton steps: While $|||(J_u, J_\tau)||| > 10^{-12}$, do:

- Compute $(\delta u, \delta \tau)$ by solving system (37).
- Update the unknowns: $u_{k+1} = u_k + \delta u, \tau_{k+1} = \tau_k + \delta \tau$.
- Update the gradient (J_u, J_τ) .

Post-processing: for $1 \leq i \leq N, t(i) := \pi(s(i), \tau)$.

Recall that J is defined in (10). The derivatives of J are provided by Proposition (1) and Proposition (2). For solving one Newton step, we use the Gmres algorithm [27], calling only the evaluation of the mapping

$$\begin{pmatrix} v \\ \theta \end{pmatrix} \mapsto \begin{pmatrix} J_{uu} & J_{u\tau} \\ J_{\tau u} & J_{\tau\tau} \end{pmatrix} \cdot \begin{pmatrix} v \\ \theta \end{pmatrix},$$

in order to solve the system

$$\begin{pmatrix} J_{uu} & J_{u\tau} \\ J_{\tau u} & J_{\tau\tau} \end{pmatrix} \cdot \begin{pmatrix} \delta u \\ \delta \tau \end{pmatrix} = - \begin{pmatrix} J_u \\ J_\tau \end{pmatrix}. \tag{37}$$

This leads in particular, for (u, τ) and $y = S(u, \tau)$ given, to the evaluation of the following quantities:

$$\begin{aligned} J_{uu} \cdot (\delta u, \cdot) &= \dot{\pi}(\cdot, \tau) (S_u^* H_{yy} S_u \delta u + H_{uy} S_u \delta u + S_u^* H_{yu} \delta u + H_{uu} \delta u) \\ &\quad + S_u^* D^2 \phi_1(y(1)) S_u \delta u + S_u^* D^2 \phi_2(y(2)) S_u \delta u, \\ J_{\tau u} &= \dot{\pi}(\cdot, \tau) (S_u^* H_y + H_u + H_{uy} S_\tau + S_u^* H_{yy} S_\tau) + S_u^* D^2 \phi_1(y(1)) S_\tau + S_u^* D^2 \phi_2(y(2)) S_\tau. \end{aligned}$$

Note, in particular, that we do not need any evaluation of the adjoint operator S_τ^* .

4.2 Illustrations

In all the tests below, for the cost function we choose

$$\ell(y, u) = -\frac{\alpha}{2} |u|_{\mathbb{R}^m}^2,$$

for a cost parameter $\alpha > 0$. The examples dealing with ordinary differential systems, considered below, are inspired by [28]. The time discretizations, for solving the state equations as well as for solving the adjoint states, are made with the Crank-Nicolson scheme. This is a second-order scheme, which is important for getting a good accuracy for the time evolution, in particular for coupled systems whose the dynamics is complex.

4.2.1 The Lotka-Volterra prey-predator system

Consider the following differential system:

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = \begin{pmatrix} (y_1(a - by_2) + u_1 y_1)(1 - c_1 y_1) \\ (y_2(qy_1 - r) + u_2 y_2)(1 - c_2 y_2) \end{pmatrix}.$$

In this system the variables y_1 and y_2 represent the densities of population of preys and predators, respectively. The multiplicative terms of type $(1 - c_i y_i)$ are considered in order to limit the values of these densities to $1/c_i$. We assume that we can control the birth rates and death rates of both species, through the bilinear control made of u_1 and u_2 . The numerical method given in Sect. 4.1 is applied to this system, with

$$T = 30.0, \quad y_0 = (1.0, 2.0)^T, \quad \alpha = 10.0, \quad a = 0.3, \quad b = 0.1, \quad r = 0.2, \quad q = 0.1, \quad c_1 = c_2 = 0.05.$$

We do not consider any terminal cost, namely $\phi_2 \equiv 0$, and the functional we maximize at some time τ is $\phi_1(y) = y_2$. It means that we want to maximize the density of population of predators. The tests presented in Figs. 1 and 2 are obtained with a Crank-Nicolson time-discretization, with $N = 3000$ time steps.

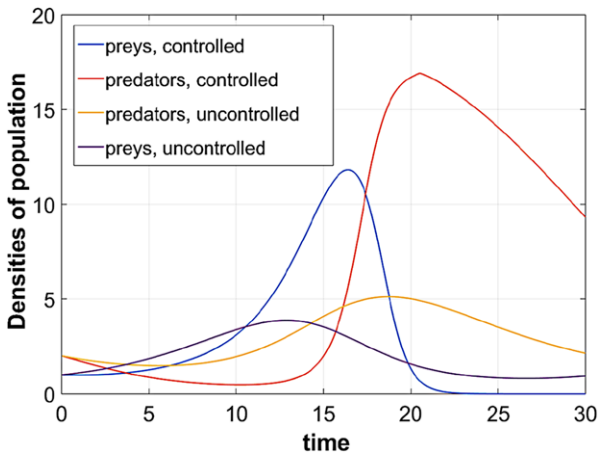


Fig. 1 Comparison of the evolutions of states for the Lotka-Volterra system, with and without control.

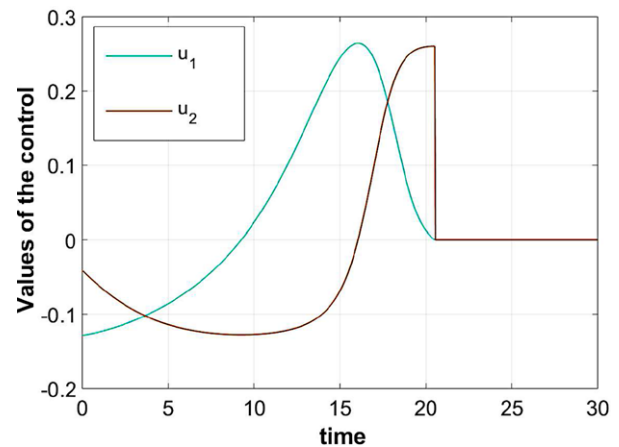


Fig. 2 Values of the bilinear control through the time.

At the beginning, the action of the control leads to diminution of the populations of both preys and predators, then leads to introduction of preys in order to feed the predators, and then favor their reproduction. The maximum of predators is reached at $\tau \approx 20.57$. As expected, the time-derivative of the second-component of the state has a jump. Note that in this example the first-order optimality condition gives in particular the equalities

$$\alpha u_1 = y_1(1 - c_1 y_1) p_1, \quad \alpha u_2 = y_2(1 - c_2 y_2) p_2.$$

Without any terminal cost, the control is indeed null for $t > \tau$, as expected in system (14) defining the adjoint-state $(p_1, p_2)^T$.

With terminal cost.

In order to avoid the extinction of population of the preys after having maximized the population of predators, we consider the terminal cost functional

$$\phi_2(y) = -\beta \log \left(\left| \frac{y_1}{y_{\text{des}}} \right| \right)^2,$$

where $\beta > 0$ is a coefficient chosen large enough and y_{des} is the desired value for the density of preys at time $t = T$. The idea is to penalize the extinction of this population (the case $y_1(T) = 0$ is forbidden), and to force to reach the desired value by choosing β large enough. With the same coefficients chosen as previously, the same time-discretization, and with $\beta = 25.0$ and $y_{\text{des}} = 1.0$, we obtain the results presented in Figs. 3 and 4.

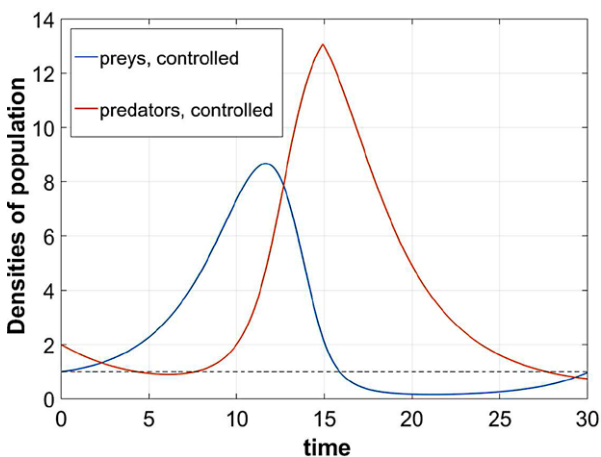


Fig. 3 Evolutions of the state for the Lotka-Volterra system, with control and terminal cost.

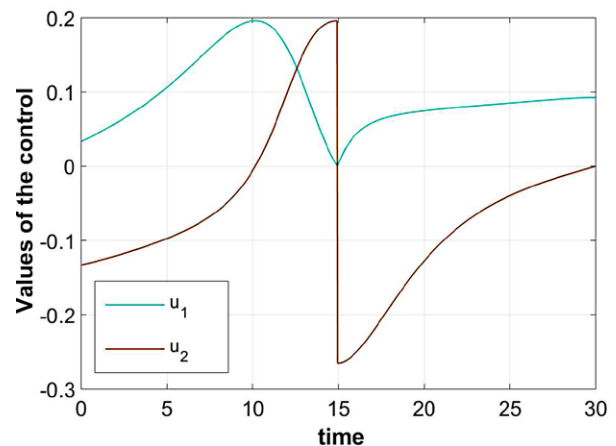


Fig. 4 Values of the bilinear control through the time, with terminal cost.

The optimal time for the maximum of predators is $\tau \approx 14.87$. As expected, the maximum is smaller than the one reached without terminal cost, because the control has to be activated for $t > \tau$ in order to take into account the functional ϕ_2 .

4.2.2 The simple damped pendulum

Consider the model of a simple pendulum, as described in Fig. 5 below.

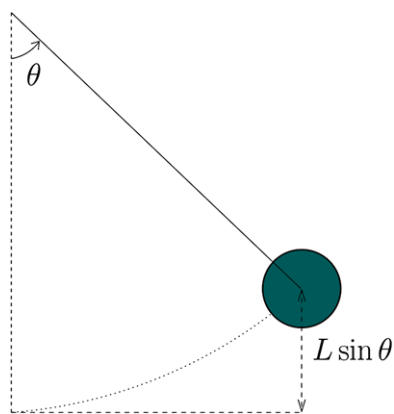


Fig. 5 Simple pendulum.

The weight is assumed to satisfy the Newton's law, and so in particular the angle θ has to satisfy the equation $\ddot{\theta} + \lambda \dot{\theta} + \mu \sin \theta = 0$, where $\lambda > 0$ is a damping term, and μ is a coefficient depending on the gravity field and the length of the taut rope. the corresponding differential system is

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = \begin{pmatrix} y_2 \\ -\lambda y_2 - \mu \sin y_1 \end{pmatrix} + \begin{pmatrix} 0 \\ u \end{pmatrix},$$

where $y = (y_1, y_2)^T := (\theta, \dot{\theta})^T$, and where the control u represents some additional horizontal force in the Newton's law. The numerical method developed in Sect. 4.1 is performed, with

$$T = 25.0, \quad y_0 = (-1.0, 0.0)^T, \quad \alpha = 10.0, \quad \lambda = 0.03, \quad \mu = 1.0.$$

Here again, we do not consider any terminal cost, namely $\phi_2 \equiv 0$. The functional we maximize at some time τ is $\phi_1(y) = y_1$, namely the angle and thus the height of the weight. The results presented in Figs. 6 and 7 are obtained with a Crank-Nicolson time-discretization, with $N = 2500$ time steps.

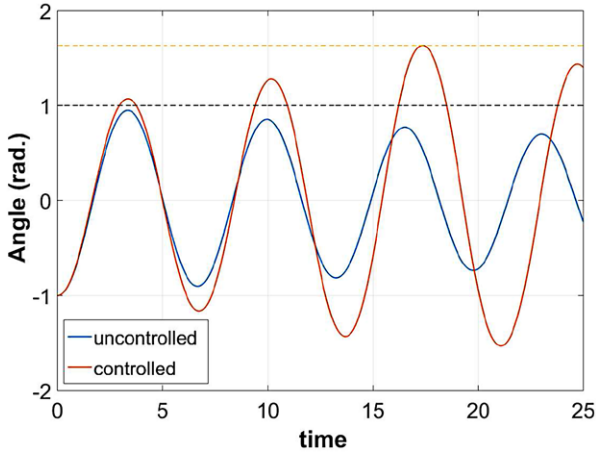


Fig. 6 Comparison of the evolutions of angles for the damped simple pendulum, with and without control.

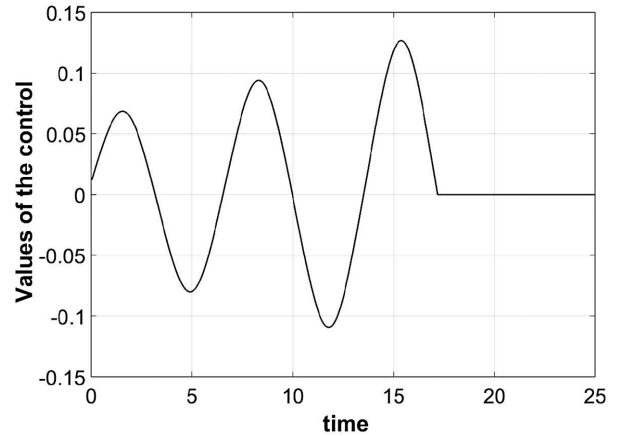


Fig. 7 Values of the control through the time.

The maximum is reached at $\tau \approx 17.22$. As we can see, the control can be kept activated on several pseudo-periods, and it can compensate the damping effect. But numerically, the difficulty lies in considering an appropriate initialization for τ (which is not $T/2$ in that case). Indeed, our numerical approach enables us only to get a critical point, and thus without a good initialization, one might find some local extremum, instead of the desired maximum. A solution consists in modifying this initialization *a posteriori*, if the critical point first found is undesired. An *a priori* knowledge of a time interval in which τ would lie may be required.

4.2.3 A partial differential equation

Consider the following Burgers-type system

$$\begin{cases} \dot{y} = \nu y_{xx} - \beta y y_x + \mathbf{1}_\omega u, & (x, t) \in (0, 1) \times (0, T), \\ y(0, t) = y(1, t) = 0, & t \in (0, T), \\ y(x, 0) = 10 \cdot (1 - e^{-(1-x)}) (e^{-(1-x)} - e^{-1}), & x \in (0, 1), \end{cases}$$

where ν and β are positive constant. Here the unknown y is considered in the space $W(0, T; H_0^1(0, 1))$ (defined in section 2.1), corresponding to the Gelfand triplet $H_0^1(0, 1) \hookrightarrow L^2(0, 1) \hookrightarrow H^{-1}(0, 1)$. Recall that $H^1(0, 1) \hookrightarrow C([0, 1])$. The question of well-posedness for such a system is addressed in [17]. The control u is considered in $L^2(\omega)$, with $\omega = [0.00; 0.25]$, and the cost term is given by $\ell(y, u) = -\frac{\alpha}{2} \|u\|_{L^2(\omega)}^2$. We want to maximize, at some optimal time $\tau \in (0, T)$, the following quantity

$$\phi_1(y(\cdot, \tau)) = \frac{1}{2} \int_D |y(x, \tau)|^2 dx$$

with $D = [0.25; 0.30]$, and without considering any terminal cost: $\phi_2 \equiv 0$. Recall that the Burgers equation is the one-dimensional version of the Navier-Stokes equations. The state y plays the role of a velocity, and thus the quantity we maximize in ϕ_1 corresponds to a kinetic energy. The space discretization is done with finite P1-elements, with $n = 101$ degrees of freedom for the state variable, and hence $m = 26$ unknowns for the control variable. The time discretization is done with a Crank-Nicolson scheme, with $N = 1000$ time steps. We consider the following parameters

$$T = 10.0, \quad \alpha = 2 \cdot 10^{-9}, \quad \nu = 2 \cdot 10^{-4}, \quad \beta = 0.05.$$

The evolutions of the state and the control through the time are represented in Fig. 8 below.

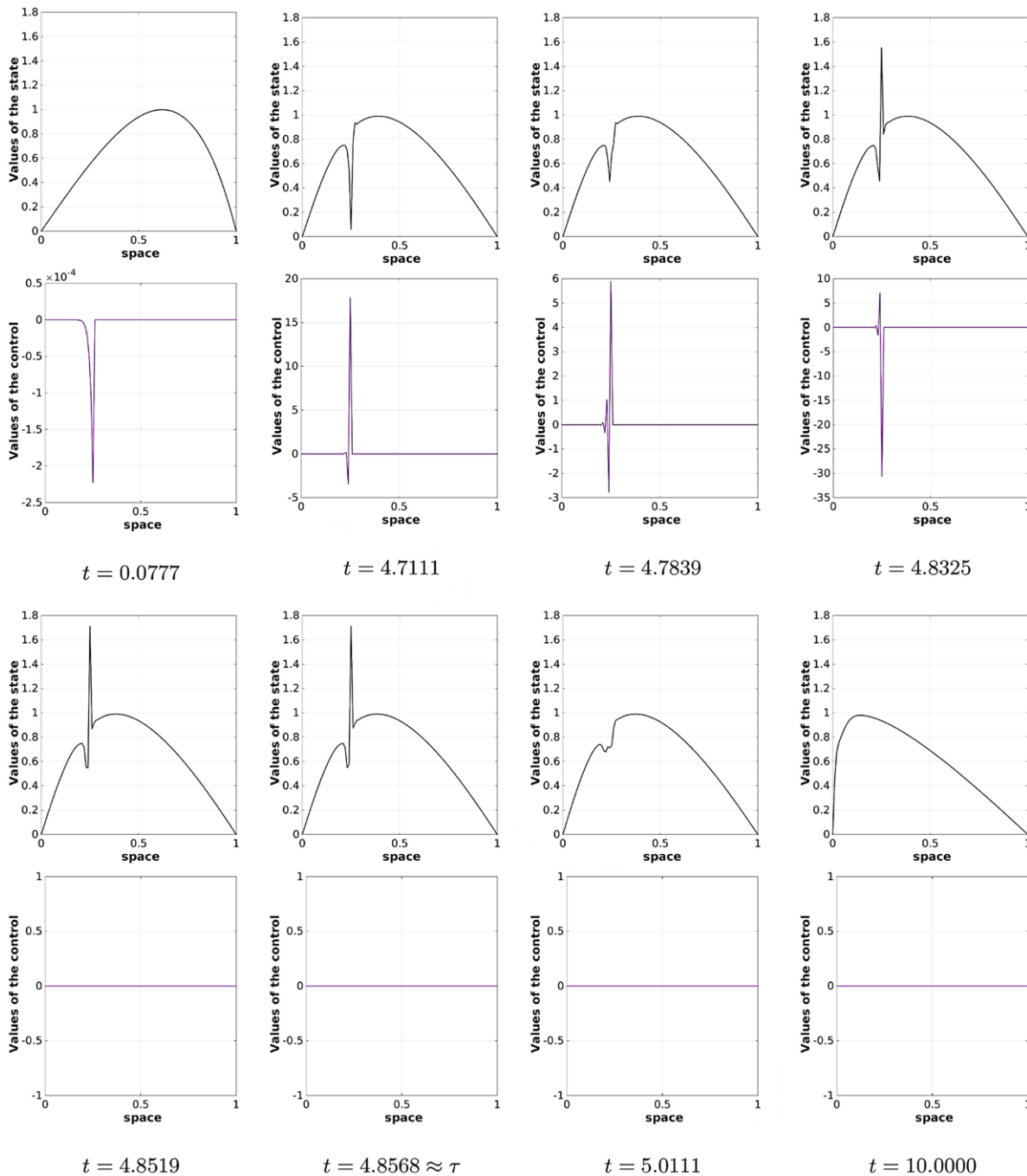


Fig. 8 Values of the state and the control, for different values of the time.

The maximum is reached for $\tau \approx 4.86$. In view of the profile of the initial condition, without control the solution is transported to the left part of the domain. The simulation shows that the control seems to wait for the so transported energy, before being mainly active during a small period before $t = \tau$, operating a bumping effect. Its influence on the state leads to transport some energy from ω to D . For $t > \tau$, the energy obtained on D is diffused into the profile (the viscosity ν is chosen here very small, so that the diffusion due to the term νy_{xx} is almost not noticed). Note that some delay is encoded into the model, that is to say when the maximum is reached for the state, at time τ , and even a little bit before this moment, the control is no longer active.

Acknowledgments The authors gratefully acknowledge support by the Austrian Science Fund (FWF) special research grant SFB-F32 “Mathematical Optimization and Applications in Biomedical Sciences”, and by the ERC advanced grant 668998 (OCLOC) under the EU’s H2020 research program.

References

- [1] N. Chamakuri, K. Kunisch, and G. Plank, *International Journal for Numerical Methods in Biomedical Engineering* **32**(4), (2016) n/a–n/a, e02742 cnm.2742.
- [2] M. Garavello and B. Piccoli, *SIAM J. Control Optim.* **43**(5), 1867–1887 (2005).
- [3] H. Hermes and J. P. LaSalle, *Functional Analysis and Time Optimal Control* (Academic Press, New York-London, Mathematics in Science and Engineering, Vol. 56. 1969).
- [4] K. Kunisch and D. Wachsmuth, *SIAM J. Control Optim.* **51**(2), 1232–1262 (2013).
- [5] K. Kunisch and A. Rund, *IMA J. Appl. Math.* **80**(6), 1664–1683 (2015).
- [6] K. Kunisch, K. Pieper, and A. Rund, *ESAIM: M2AN* **50**(2), 381–414 (2016).
- [7] V. Barbu, *Syst. Control Lett.* **30**(2), 93–100 (1997).
- [8] S. Micu, I. Roventa, and M. Tucsnak, *J. Funct. Anal.* **263**(1), 25–49 (2012).
- [9] T. Bayen and A. Rapaport, *J. Convex Anal.* **23**(1), 263–290 (2016).
- [10] K. Ito and K. Kunisch, *SIAM J. Control Optim.* **48**(6), 3997–4013 (2010).
- [11] J. P. Raymond and H. Zidani, *J. Optim. Theory Appl.* **101**(2), 375–402 (1999).
- [12] J. P. Raymond and H. Zidani, *Differential Integral Equations* **13**(7–9), 1039–1072 (2000).
- [13] L. Li, Y. Gao, and H. Wang, *J. Ind. Manag. Optim.* **11**(1), 329–343 (2015).
- [14] F. Tröltzsch, *Optimal Control of Partial Differential Equations, Vol 112 of Graduate Studies in Mathematics, Theory, Methods and Applications, Translated from the 2005 German original by Jürgen Sprekels* (American Mathematical Society, Providence, RI, 2010).
- [15] M. Hinze and K. Kunisch, *SIAM J. Control Optim.* **40**(3), 925–946 (2001) (electronic).
- [16] M. Hinze and K. Kunisch, *ZAMM Z. Angew. Math. Mech.* **84**(3), 171–187 (2004).
- [17] S. Volkwein, *Control Cybern.* **30**, 249–278 (2001).
- [18] R. Temam, *Navier-Stokes Equations, revised edn, Vol 2 of Studies in Mathematics and its Applications, Theory and Numerical Analysis, With an appendix by F. Thomasset* (North-Holland Publishing Co., Amsterdam-New York, 1979).
- [19] G. Grubb and V. A. Solonnikov, *Math. Scand.* **1991** **69**(2), 217–290 (1992).
- [20] J. Simon, *Ann. Mat. Pura Appl. (4)* **146**, 65–96 (1987).
- [21] Y. Giga, *Math. Z.* **178**(3), 297–329 (1981).
- [22] E. Casas, C. Ryll, and F. Tröltzsch, *Comput. Methods Appl. Math.* **13**(4), 415–442 (2013).
- [23] T. Breiten and K. Kunisch, *SIAM J. Control Optim.* **52**(6), 4057–4081 (2014).
- [24] C. Bardos, *J. Functional Analysis* **7**, 311–322 (1971).
- [25] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE Constraints, Vol 23 of Mathematical Modelling: Theory and Applications*, (Springer, New York, 2009).
- [26] M. Raydan, *SIAM J. Optim.* **7**(1), 26–33 (1997).
- [27] Y. Saad, M. H. Schultz, *SIAM J. Sci. Statist. Comput.* **7**(3), 856–869 (1986).
- [28] E. Trélat, *Contrôle optimal, of Mathématiques Concrètes. [Concrete Mathematics]*, Vuibert, Paris, **2005**, Théorie & Applications. [Theory and Applications].