

MEASURE CONTROL OF A SEMILINEAR PARABOLIC EQUATION WITH A NONLOCAL TIME DELAY*

EDUARDO CASAS[†], MARIANO MATEOS[‡], AND FREDI TRÖLTZSCH[§]

Abstract. We study a control problem governed by a semilinear parabolic equation. The control is a measure that acts as the kernel of a possibly nonlocal time delay term and the functional includes a nondifferentiable term with the measure norm of the control. Existence, uniqueness, and regularity of the solution of the state equation, as well as differentiability properties of the control-to-state operator are obtained. Next, we provide first order optimality conditions for local solutions. Finally, the control space is suitably discretized and we prove convergence of the solutions of the discrete problems to the solutions of the original problem. Several numerical examples are included to illustrate the theoretical results.

Key words. optimal control, parabolic equation, nonlocal time delay, measure control

AMS subject classifications. 49K20, 35K58, 49M25

DOI. 10.1137/17M1157362

1. Introduction. We consider optimal control problems for the parabolic equation

$$(1.1) \quad \begin{cases} \frac{\partial y}{\partial t} - \Delta y + R(y) = \int_0^T y(x, t-s) \, du(s) & \text{in } Q = \Omega \times (0, T), \\ \partial_n y = 0 & \text{on } \Sigma = \Gamma \times (0, T), \\ y(x, t) = y_0(x, t) & \text{in } Q_- = \Omega \times [-T, 0] \end{cases}$$

with fixed final time $T > 0$, where the Borel measure $u \in \mathcal{M}[0, T]$ is taken as control. Depending on the particular choice of this measure and on the form of the nonlinearity R , different mathematical models of interest for theoretical physics are covered by this equation. Thanks to its generality, this equation includes the control of time delays in parabolic equations, the control of multiple time delays, and also the optimization of standard feedback operators of Pyragas type. Associated examples will be explained below.

Our paper extends the optimization of nonlocal Pyragas-type feedback operators that was investigated in [25]. The main novelty of our paper is the use of measures instead of functions. This is much more general and leads to new, partially delicate and interesting questions of analysis. The partial differential equation above includes three main difficulties: First, the equation is of semilinear type. The main ideas for the associated analysis were prepared in [11] and we are able to proceed similarly, at

*Received by the editors November 17, 2017; accepted for publication (in revised form) October 25, 2018; published electronically December 18, 2018.

<http://www.siam.org/journals/sicon/56-6/M115736.html>

Funding: The first two authors were partially supported by the Spanish Ministerio de Economía y Competitividad under projects MTM2014-57531-P and MTM2017-83185-P. The third author was supported by the collaborative research center SFB 910, TU Berlin, project B6.

[†]Departamento de Matemática Aplicada y Ciencias de la Computación, E.T.S.I. Industriales y de Telecomunicación, Universidad de Cantabria, 39005 Santander, Spain (eduardo.casas@unican.es).

[‡]Departamento de Matemáticas, Campus de Gijón, Universidad de Oviedo, 33203, Gijón, Spain (mmateos@uniovi.es).

[§]Institut für Mathematik, Technische Universität Berlin, D-10623 Berlin, Germany (troeltzsch@math.tu-berlin.de).

least partially. Second, the equation contains some kind of time delays. Finally, the integral operator includes the measure u that complicates the analysis.

The optimal control theory of ordinary or partial differential equations with time delay has a very long history. Numerous papers were contributed to this field. We mention, for instance, the papers [2, 3, 12, 18, 19], that have some relation to distributed parameter systems, or the surveys [4, 28]. More recent contributions are, e.g., [17, 23, 24].

However, to our best knowledge, the optimal control of parabolic equations with nonlocal time delay was only investigated in [25]. The case of measures as controls is new for this type of equation. However, we mention [1], where a measure-valued control function is considered in a delay equation.

Moreover, the control is not taken as a right-hand side. Here, it plays the role of a kernel in an integral operator; this is another difficulty. We should mention that the use of kernels as control functions is not new. For instance, memory kernels were taken as “controls” in identification problems in [35] and [36].

The equation above generalizes different models of Pyragas-type feedback that are very popular in theoretical physics. We mention the seminal paper by Pyragas [26], where the feedback of the form (1.2) was introduced to stabilize periodic orbits; see also [27]. We also refer to [30, 32, 33], where nonlocal Pyragas feedback operators of the type (1.5) are discussed for different kernels u .

Let us also mention [20], where the implementation of nonlocal feedback controllers is investigated. In particular, these equations have applications in laser technology; we refer to associated contributions in [30].

Let us also mention a few examples for (1.1). In the following κ is a real parameter.

Example 1.1 (Pyragas feedback control). If $\tau \in (0, T)$ is a fixed time, and δ_τ and δ_0 denote the Dirac measures concentrated at τ and 0, respectively, then the equation

$$(1.2) \quad \frac{\partial y}{\partial t} - \Delta y + R(y) = \kappa (y(x, t - \tau) - y(x, t))$$

is obtained as a particular case of (1.1) with $u = \kappa (\delta_\tau - \delta_0)$. Equations of this type with fixed time delay τ are known in the context of the so-called Pyragas-type feedback control [26, 27, 30].

Example 1.2 (Pyragas feedback with multiple time delays). A more general version of (1.2) with multiple fixed time delays is generated by

$$(1.3) \quad u = \kappa \left(\sum_{i=1}^m u_i \delta_{\tau_i} - \delta_0 \right)$$

with fixed time delays $0 < \tau_1 < \dots < \tau_m < T$. Then the equation

$$(1.4) \quad \frac{\partial y}{\partial t} - \Delta y + R(y) = \kappa \left(\sum_{i=1}^m u_i y(x, t - \tau_i) - y(x, t) \right)$$

is obtained. Here, the control $(u_1, \dots, u_m) \in \mathbb{R}^m$ is a vector of controllable weights.

Example 1.3 (nonlocal Pyragas feedback control). Finally, if the Lebesgue decomposition of u is $u = \kappa(u_r - \delta_0)$, where u_r is absolutely continuous w.r.t. the Lebesgue measure in $[0, T]$ and its Radon–Nikodym derivative is $g \in L^1(0, T)$, then equation

(1.1) takes the form

$$(1.5) \quad \frac{\partial y}{\partial t} - \Delta y + R(y) = \kappa \left(\int_0^T g(s)y(x, t-s) ds - y(x, t) \right)$$

that is used in nonlocal Pyragas-type feedback. Here, the control is the integrable kernel g of the integral operator of the partial differential equation.

2. State equation. Throughout this paper $\mathcal{M}[0, T]$ will denote the space of real and regular Borel measures in $[0, T]$. According to the Riesz representation theorem, $\mathcal{M}[0, T]$ is the dual of the space of continuous functions in $[0, T]$: $\mathcal{M}[0, T] = C[0, T]^*$, and $\mathcal{M}[0, T]$ is a Banach space endowed with the norm

$$\|u\|_{\mathcal{M}[0, T]} = |u|([0, T]) = \sup \left\{ \int_0^T \phi(s) du(s) : \phi \in C[0, T] \text{ and } \|\phi\|_{C[0, T]} \leq 1 \right\}.$$

Here, $|u|$ denotes the total variation measure of u ; see [29, pp. 130–133]. The above integrals are considered in the closed interval $[0, T]$. Notice that $u(\{0\})$ and $u(\{T\})$ could be nonzero. This notational convention will be maintained in what follows. Thus, we distinguish

$$\int_a^b \phi(s) du(s) = \int_{[a, b]} \phi(s) du(s), \int_{(a, b]} \phi(s) du(s), \int_{(a, b)} \phi(s) du(s), \int_{(a, b)} \phi(s) du(s)$$

for real numbers $a \leq b$ and continuous functions $\phi \in C[0, T]$.

We recall our general state equation (1.1),

$$\begin{cases} \frac{\partial y}{\partial t} - \Delta y + R(y) = \int_0^T y(x, t-s) du(s) & \text{in } Q, \\ \partial_n y = 0 & \text{on } \Sigma, \\ y(x, t) = y_0(x, t) & \text{in } Q_-. \end{cases}$$

In this setting, $\Omega \subset \mathbb{R}^n$, $1 \leq n \leq 3$, is a bounded Lipschitz domain with boundary Γ and, (as introduced above) $Q_- = \Omega \times [-T, 0]$. The nonlinearity $R : \mathbb{R} \rightarrow \mathbb{R}$ is a function of class C^1 such that

$$(2.1) \quad \exists C_R \in \mathbb{R} : R'(y) \geq C_R \quad \forall y \in \mathbb{R}.$$

The initial datum y_0 is taken from $C(\bar{Q}_-)$, while $u \in \mathcal{M}[0, T]$ is the control.

In particular, third order polynomials of the form

$$R(y) = \rho(y - y_1)(y - y_2)(y - y_3)$$

with $\rho > 0$ and $y_1 < y_2 < y_3$ satisfy this assumption. R has the meaning of a reaction term. The numbers y_i , $i = 1, 2, 3$, are the fixed points of the reaction; y_1 and y_3 are the stable ones, while y_2 is unstable. Such functions R play a role in bistable reactions of physical chemistry; see [22].

The assumptions are also fulfilled by higher order polynomials of odd order,

$$R(y) = \sum_{i=0}^k a_i y^i,$$

with real numbers a_i , $i = 0, \dots, k$, and $k = 2\ell + 1$, $\ell \in \mathbb{N} \cup \{0\}$, if a_k is positive. Here, the derivative R' is an even order polynomial that satisfies condition (2.1).

Remark 2.1. The theory of our paper can be extended to more general functions $R : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$, that obey the following assumptions:

- R is a Carathéodory function of class C^1 with respect to the second variable.
- There exists some $p > n/2$ such that $R(\cdot, 0) \in L^p(\Omega)$.
- For all $M > 0$ there exists a constant $C_M > 0$ such that

$$\left| \frac{\partial R}{\partial y}(x, y) \right| \leq C_M \text{ for a.a. } x \in \Omega \text{ and } \forall y \in \mathbb{R} \text{ with } |y| \leq M.$$

- The function $y \mapsto \frac{\partial R}{\partial y}(x, y)$ is bounded from below, i.e.,

$$\frac{\partial R}{\partial y}(x, y) \geq C_R \text{ for a.a. } x \in \Omega \text{ and all } y \in \mathbb{R}.$$

The theory also remains true if—in addition to these general assumptions on R —the Laplace operator $-\Delta$ is replaced by another uniformly elliptic differential operator A with L^∞ coefficients in the main part of the operator. Lipschitz regularity of these coefficients is required only in the second part of Theorem 2.5.

However, to keep the presentation simple, we concentrate on the case $A = -\Delta$, and a function $R : \mathbb{R} \rightarrow \mathbb{R}$ of class C^1 and satisfying condition (2.1).

In the following, we will denote $Y = L^2(0, T; H^1(\Omega)) \cap C(\bar{Q})$. Endowed with the norm

$$\|y\|_Y = \|y\|_{L^2(0,T;H^1(\Omega))} + \|y\|_{C(\bar{Q})},$$

Y is a Banach space.

We say that $y \in Y$ is a solution of problem (1.1) if

$$\begin{aligned} - \int_Q y \frac{\partial z}{\partial t} dx dt + \int_Q \nabla y \nabla z dx dt + \int_Q R(y)z dx dt \\ = \int_Q \int_0^T y(x, t-s) du(s)z(x, t) dx dt + \int_\Omega y_0(x, 0)z(x, 0) dx \end{aligned}$$

for all $z \in H^1(Q)$ such that $z(x, T) = 0$.

Notice that if $y \in Y$ is a solution of (1.1), then we also have that $y \in W(0, T) = L^2(0, T; H^1(\Omega)) \cap H^1(0, T; H^1(\Omega)^*)$.

We begin our analysis with the well-posedness of the state equation (1.1) that is a differential equation with time delay. Ordinary differential delay equations are well understood; we refer, for instance, to the expositions [6, 15, 13]. For parabolic partial differential equations, we only mention [5] and the references cited therein, since this book investigates oscillation effects for nonlinear partial differential equations with delay that we observe also for (1.1). Our parabolic delay equation is nonlinear and contains a nonlocal Pyragas-type feedback term defined by a measure. To our best knowledge, an associated result on existence and uniqueness of a solution is not yet known.

THEOREM 2.2. *For every $u \in \mathcal{M}[0, T]$, problem (1.1) has a unique solution $y_u \in Y$. Moreover, the estimates*

$$(2.2) \quad \|y_u\|_{L^2(0,T;H^1(\Omega))} \leq C_{1,0} (\|y_0\|_{L^2(Q_-)} \|u\|_{\mathcal{M}[0,T]} + \|y_0(\cdot, 0)\|_{L^2(\Omega)} + |R(0)|),$$

$$(2.3) \quad \|y_u\|_{C(\bar{Q})} \leq C_\infty (\|y_0\|_{C(\bar{Q}_-)} \|u\|_{\mathcal{M}[0,T]} + \|y_0(\cdot, 0)\|_{C(\bar{\Omega})} + |R(0)|),$$

are satisfied, where the constants $C_{1,0}$ and C_∞ depend on $\|u\|_{\mathcal{M}[0,T]}$, but they can be taken fixed on bounded subsets of $\mathcal{M}[0, T]$.

In order to prove this theorem, we perform the classical substitution $y^\lambda(x, t) = e^{-\lambda t}y(x, t)$ with arbitrary $\lambda > 0$. Hence (1.1) is transformed to

$$\begin{cases} \frac{\partial y^\lambda}{\partial t} - \Delta y^\lambda + e^{-\lambda t}R(e^{\lambda t}y^\lambda) + \lambda y^\lambda = \int_0^T e^{-\lambda s}y^\lambda(x, t-s) du(s) & \text{in } Q, \\ \partial_n y^\lambda = 0 & \text{on } \Sigma, \\ y^\lambda(x, t) = e^{-\lambda t}y_0(x, t) & \text{in } Q_-. \end{cases}$$

Note that the appearance of the term $e^{-\lambda s}$ under the integral is due to the fact that $y(x, t-s) = e^{\lambda(t-s)}y^\lambda(x, t-s)$ and $e^{\lambda t}$ disappears after dividing the equation by this term. To simplify the notation we introduce, for every $\lambda \geq 0$, the following family of operators $K_\lambda[u], K_\lambda^+[u] : C(\bar{Q}) \rightarrow L^\infty(Q)$, defined for $(x, t) \in \bar{Q}$ by

$$\begin{aligned} (K_\lambda[u]z)(x, t) &= \int_{[0,t)} e^{-\lambda s}z(x, t-s) du(s), \\ (K_\lambda^+[u]z)(x, t) &= \int_{(0,t)} e^{-\lambda s}z(x, t-s) du(s). \end{aligned}$$

Hence, we have $(K_\lambda[u]z)(x, t) = u(\{0\})z(x, t) + (K_\lambda^+[u]z)(x, t)$ for $t > 0$.

Moreover, we define the family of functions

$$g_{\lambda,u}(x, t) = \int_t^T e^{-\lambda s}y^\lambda(x, t-s) du(s) = e^{-\lambda t} \int_t^T y_0(x, t-s) du(s) \text{ for } (x, t) \in \bar{Q}$$

that covers the initial data y_0 .

For $\lambda = 0$ we simply write $K[u]$ and g_u instead of $K_0[u]$ and $g_{0,u}$, respectively. With this notation, the above equation and (1.1) (obtained for $\lambda = 0$) can be formulated as follows:

$$(2.4) \quad \begin{cases} \frac{\partial y^\lambda}{\partial t} - \Delta y^\lambda + e^{-\lambda t}R(e^{\lambda t}y^\lambda) + \lambda y^\lambda = K_\lambda[u]y^\lambda + g_{\lambda,u} & \text{in } Q, \\ \partial_n y^\lambda = 0 & \text{on } \Sigma, \\ y^\lambda(x, 0) = y_0(x, 0) & \text{in } \Omega \end{cases}$$

with the additional extension $y^\lambda(x, t) = e^{-\lambda t}y_0(x, t)$ for $t \in [-T, 0]$.

Remark 2.3. Notice that $K_\lambda[u]y$ and $g_{\lambda,u}$ can be discontinuous at those points t such that $u(\{t\}) \neq 0$, but the following identity holds:

$$\begin{aligned} \int_0^T e^{-\lambda s}y^\lambda(x, t-s)du(s) &= \int_{[0,t)} e^{-\lambda s}y^\lambda(x, t-s)du(s) + \int_t^T e^{-\lambda s}y_0(x, t-s)du(s) \\ &= (K_\lambda[u]y^\lambda + g_{\lambda,u})(x, t), \end{aligned}$$

and, therefore, $K_\lambda[u]y^\lambda + g_{\lambda,u} \in C(\bar{Q})$ if $y^\lambda \in C(\bar{Q})$.

LEMMA 2.4. For every $u \in \mathcal{M}[0, T]$ and $\lambda \geq 0$ we have

$$(2.5) \quad \begin{cases} \|K_\lambda[u]z\|_{L^2(Q)} & \leq \|z\|_{L^2(Q)}\|u\|_{\mathcal{M}[0,T]} & \forall z \in C(\bar{Q}), \\ \|K_\lambda[u]z\|_{L^\infty((0,T),L^2(\Omega))} & \leq \|z\|_{C([0,T],L^2(\Omega))}\|u\|_{\mathcal{M}[0,T]} & \forall z \in C(\bar{Q}), \\ \|K_\lambda[u]z\|_{L^\infty(Q)} & \leq \|z\|_{C(\bar{Q})}\|u\|_{\mathcal{M}[0,T]} & \forall z \in C(\bar{Q}), \end{cases}$$

$$(2.6) \quad \begin{cases} \|g_{\lambda,u}\|_{L^2(Q)} & \leq \|y_0\|_{L^2(Q_-)}\|u\|_{\mathcal{M}[0,T]}, \\ \|g_{\lambda,u}\|_{L^\infty(Q)} & \leq \|y_0\|_{C(\bar{Q}_-)}\|u\|_{\mathcal{M}[0,T]}. \end{cases}$$

Moreover, for every $\varepsilon > 0$ and $u \in \mathcal{M}[0, T]$ there exists $\lambda_{\varepsilon, u} > 0$ such that $\forall \lambda \geq \lambda_{\varepsilon, u}$ the following inequalities hold:

$$(2.7) \quad \begin{cases} \|K_\lambda^+[u]z\|_{L^2(Q)} & \leq \varepsilon(1 + \|u\|_{\mathcal{M}[0, T]})\|z\|_{L^2(Q)} & \forall z \in C(\bar{Q}), \\ \|K_\lambda^+[u]z\|_{L^\infty(Q)} & \leq \varepsilon(1 + \|u\|_{\mathcal{M}[0, T]})\|z\|_{C(\bar{Q})} & \forall z \in C(\bar{Q}). \end{cases}$$

Proof. By using the Schwarz inequality and the Fubini theorem, we get

$$\begin{aligned} \|K_\lambda[u]z\|_{L^2(Q)}^2 &= \int_Q \left(\int_{[0, t]} e^{-\lambda s} z(x, t - s) \, du(s) \right)^2 \, dx \, dt \\ &\leq \int_Q \left(\int_{[0, t]} z^2(x, t - s) \, d|u|(s) \right) \left(\int_{[0, t]} e^{-2\lambda s} \, d|u|(s) \right) \, dx \, dt \\ &\leq \left(\int_\Omega \int_0^T \int_s^T z^2(x, t - s) \, dt \, d|u|(s) \, dx \right) \left(\int_0^T e^{-2\lambda s} \, d|u|(s) \right) = I_1 I_2. \end{aligned}$$

Substituting $\sigma = t - s$ we get for I_1

$$\begin{aligned} I_1 &= \int_\Omega \int_0^T \int_0^{T-s} z^2(x, \sigma) \, d\sigma \, d|u|(s) \, dx \leq \int_\Omega \int_0^T \int_0^T z^2(x, \sigma) \, d\sigma \, d|u|(s) \, dx \\ &= \|z\|_{L^2(Q)}^2 \|u\|_{\mathcal{M}[0, T]}. \end{aligned}$$

To estimate I_2 we proceed as follows:

$$I_2 = \int_0^T e^{-2\lambda s} \, d|u|(s) \leq \int_0^T \, d|u|(s) = \|u\|_{\mathcal{M}[0, T]}.$$

Multiplying the estimates for I_1 and I_2 , we get the first inequality of (2.5). To prove the second estimate we proceed as follows: $\forall t \in [0, T]$

$$\begin{aligned} \|(K_\lambda[u]z)(\cdot, t)\|_{L^2(\Omega)}^2 &= \int_\Omega \left(\int_{[0, t]} e^{-\lambda s} z(x, t - s) \, du(s) \right)^2 \, dx \\ &\leq \left(\int_\Omega \int_{[0, t]} z^2(x, t - s) \, d|u|(s) \, dx \right) \left(\int_0^T e^{-2\lambda s} \, d|u|(s) \right) \\ &= \left(\int_{[0, t]} \|z(\cdot, t - s)\|_{L^2(\Omega)}^2 \, d|u|(s) \right) I_2 \leq \|z\|_{C([0, T], L^2(\Omega))}^2 \|u\|_{\mathcal{M}[0, T]}^2. \end{aligned}$$

Finally, to prove the third inequality of (2.5) we only need the following estimation:

$$\|K_\lambda[u]z\|_{L^\infty(Q)} \leq \|z\|_{C(\bar{Q})} \int_0^T e^{-\lambda s} \, d|u|(s) \leq \|z\|_{C(\bar{Q})} \|u\|_{\mathcal{M}[0, T]}.$$

In a completely analogous way we prove (2.6). To establish the first inequality of (2.7) we proceed exactly as above replacing I_2 by I_2^+ , so $\|K_\lambda^+[u]z\|_{L^2(Q)}^2 \leq I_1 I_2^+$, where

$$(2.8) \quad \begin{aligned} I_2^+ &= \int_{(0, T]} e^{-2\lambda s} \, d|u|(s) \leq \int_{(0, \frac{1}{\sqrt{\lambda}})} e^{-2\lambda s} \, d|u|(s) + \int_{\frac{1}{\sqrt{\lambda}}}^T e^{-2\lambda s} \, d|u|(s) \\ &\leq |u|(0, \frac{1}{\sqrt{\lambda}}) + e^{-2\sqrt{\lambda}} \|u\|_{\mathcal{M}[0, T]}. \end{aligned}$$

It is enough to choose $\lambda_{\varepsilon,u} > 0$ sufficiently large such that

$$|u| \left(0, \frac{1}{\sqrt{\lambda_{\varepsilon,u}}} \right) + e^{-2\sqrt{\lambda_{\varepsilon,u}}} < \varepsilon^2$$

and the first inequality in (2.7) follows from the fact that $I_1 \leq \|z\|_{L^2(Q)}^2(1 + \|u\|_{\mathcal{M}[0,T]})$ and that each of the addends in the previous inequality is smaller than ε^2 .

The second inequality in (2.7) follows trivially from (2.5) for all $\lambda > 0$ if $\varepsilon \geq 1$. For $\varepsilon < 1$, we have that if $\lambda > \lambda_{\varepsilon,u}$, chosen as before,

$$\begin{aligned} \|K_\lambda^+[u]z\|_{L^\infty(Q)} &\leq \|z\|_{C(\bar{Q})} \int_{(0,T]} e^{-\lambda s} d|u|(s) \\ &\leq \|z\|_{C(\bar{Q})} \left(|u| \left(0, \frac{1}{\sqrt{\lambda}} \right) + e^{-\sqrt{\lambda}} \|u\|_{\mathcal{M}[0,T]} \right) \\ &\leq \|z\|_{C(\bar{Q})} (\varepsilon^2 + \varepsilon \|u\|_{\mathcal{M}[0,T]}) \leq \varepsilon(1 + \|u\|_{\mathcal{M}[0,T]}) \|z\|_{C(\bar{Q})} \end{aligned}$$

and the proof is complete. \square

Proof of Theorem 2.2. We split the proof into three steps.

I: *Existence of a solution.* For every function $z \in C(\bar{Q})$, we define the problem

$$(2.9) \quad \begin{cases} \frac{\partial y}{\partial t} - \Delta y + R_\lambda^+(t, y) = K_\lambda^+[u]z + g_{\lambda,u} & \text{in } Q, \\ \partial_n y = 0 & \text{on } \Sigma, \\ y(x, 0) = y_0(x, 0) & \text{in } \Omega, \end{cases}$$

where $R_\lambda^+(t, y) = e^{-\lambda t} R(e^{\lambda t} y) + (\lambda - u(\{0\}))y$. This is a standard semilinear parabolic equation with given right-hand side. We have that $\frac{\partial R_\lambda^+}{\partial y}(t, y) \geq C_R + \lambda - u(\{0\})$ and we set $\lambda_R^+ = 1 - \min\{0, C_R - u(\{0\})\}$, hence, it follows that

$$(2.10) \quad \frac{\partial R_\lambda^+}{\partial y}(t, y) \geq 1 \quad \forall \lambda \geq \lambda_R^+.$$

Therefore, R_λ^+ is a continuous monotone increasing function with respect to y , and it is well known that the semilinear equation (2.9) has a unique solution $y \in Y$; see [7] or [34], for instance. The continuity is due to the continuity of $y_0(\cdot, 0)$ and the fact that the right-hand side of the partial differential equation in (2.9) belongs to $L^\infty(Q)$. Moreover from the above references and the inequality $|R_\lambda^+(t, 0)| = e^{-\lambda t} |R(0)| \leq |R(0)|$ we infer the estimate

$$(2.11) \quad \|y\|_{C(\bar{Q})} \leq C_0 \left(\|K_\lambda^+[u]z\|_{L^\infty(Q)} + \|g_{\lambda,u}\|_{L^\infty(Q)} + \|y_0(\cdot, 0)\|_{C(\bar{\Omega})} + |R(0)| \right).$$

Now, we define $M = C_0(1 + \|u\|_{\mathcal{M}[0,T]} \|y_0\|_{C(\bar{Q}^-)} + \|y_0(\cdot, 0)\|_{C(\bar{\Omega})} + |R(0)|)$. According to (2.7), we can select $\lambda \geq \lambda_R^+$ such that

$$(2.12) \quad \|K_\lambda^+[u]z\|_{L^\infty(Q)} \leq 1 \quad \forall z \in C(\bar{Q}) \quad \text{with } \|z\|_{C(\bar{Q})} \leq M.$$

Let B_M be the closed ball of $C(\bar{Q})$ with center at 0 and radius M . We define the

continuous mapping $F : B_M \rightarrow B_M$ that associates with every $z \in B_M$ the solution $y = F(z)$ of (2.9). The embedding $F(B_M) \subset B_M$ is an immediate consequence of (2.6), (2.11), (2.12), and the definition of M . In order to apply Schauder's fixed point theorem, we have to prove that $F(B_M)$ is relatively compact in $C(\bar{Q})$. To this end, we assume first that $y_0(\cdot, 0)$ is a Hölder function in $\bar{\Omega}$: $y_0(\cdot, 0) \in C^{0,\mu}(\bar{\Omega})$ with $\mu \in (0, 1)$. Then, there exists $\mu_0 \in (0, \mu)$ and C_μ such that $y \in C^{0,\mu_0}(\bar{Q})$ and

$$\|y\|_{C^{0,\mu_0}(\bar{Q})} \leq C_\mu \left(\|K_\lambda^+[u]z + g_{\lambda,u}\|_{L^\infty(Q)} + \|y_0(\cdot, 0)\|_{C^{0,\mu}(\bar{\Omega})} + \max_{|\rho| \leq e^{\lambda T} M} |R(\rho)| + \lambda \right);$$

see [21, section III-8]. From the compactness of the embedding $C^{0,\mu_0}(\bar{Q}) \subset C(\bar{Q})$ and the above estimate we conclude that $F(B_M)$ is relatively compact in $C(\bar{Q})$ and F has at least one fixed point y^λ . Then it is obvious that y^λ is a solution of (2.4) and $y^\lambda \in Y$.

Next we skip the assumption $y_0(\cdot, 0) \in C^{0,\mu}(\bar{\Omega})$ for every sufficiently large λ . Since $y_0 \in C(\bar{\Omega})$, then we can take a sequence $\{y_{0k}\}_{k=1}^\infty \subset C^{0,\mu}(\bar{\Omega})$ such that $y_{0k} \rightarrow y_0(\cdot, 0)$ in $C(\bar{\Omega})$. Hence, for every $k \geq 1$, (2.4) has at least one solution $y_k^\lambda \in Y$. Let us prove that $\{y_k^\lambda\}_{k=1}^\infty$ is a Cauchy sequence in Y . To this end we select two terms of the sequence y_k^λ and y_m^λ and subtract the equations satisfied by them. Then we get

$$\left\{ \begin{aligned} \frac{\partial(y_k^\lambda - y_m^\lambda)}{\partial t} - \Delta(y_k^\lambda - y_m^\lambda) + \frac{\partial R_\lambda^+}{\partial y}(t, y_{km}^\lambda)(y_k^\lambda - y_m^\lambda) &= K_\lambda^+[u](y_k^\lambda - y_m^\lambda) \text{ in } Q, \\ \partial_n(y_k^\lambda - y_m^\lambda) &= 0 \text{ on } \Sigma, \\ (y_k^\lambda - y_m^\lambda)(x, 0) &= y_{0k}(x) - y_{0m}(x) \text{ in } \Omega, \end{aligned} \right.$$

where $y_{km}^\lambda = y_k^\lambda + \theta_k(y_k^\lambda - y_m^\lambda)$ and $0 \leq \theta_{km}(x, t) \leq 1$ is a measurable function. Now, multiplying the above equation by $y_k^\lambda - y_m^\lambda$, using that $\frac{\partial R_\lambda^+}{\partial y}(t, y_{km}^\lambda) \geq 1$ due to our choice of λ (cf. (2.10)), and (2.7), we obtain

$$\begin{aligned} & \frac{1}{2} \|(y_k^\lambda - y_m^\lambda)(T)\|_{L^2(\Omega)} + \int_Q |\nabla(y_k^\lambda - y_m^\lambda)|^2 dx dt + \int_Q |y_k^\lambda - y_m^\lambda|^2 dx dt \\ & \leq \varepsilon (1 + \|u\|_{\mathcal{M}[0,T]}) \|y_k^\lambda - y_m^\lambda\|_{L^2(Q)}^2 + \frac{1}{2} \|y_{0k} - y_{0m}\|_{L^2(\Omega)}^2 \\ & \leq C\varepsilon (1 + \|u\|_{\mathcal{M}[0,T]}) \|y_k^\lambda - y_m^\lambda\|_{L^2(0,T;H^1(\Omega))}^2 + \frac{1}{2} \|y_{0k} - y_{0m}\|_{L^2(\Omega)}^2. \end{aligned}$$

For all ε sufficiently small and $\lambda \geq \lambda_{\varepsilon,u}$, this leads to

$$\|y_k^\lambda - y_m^\lambda\|_{L^2(0,T;H^1(\Omega))} \leq C_1 \|y_{0k} - y_{0m}\|_{L^2(\Omega)}.$$

Notice that the left-hand side of the above chain of inequalities absorbs the term appearing with the factor $C\varepsilon(1 + \|u\|_{\mathcal{M}[0,T]})$ in the right-hand side, if ε is small enough.

Hence $\{y_k^\lambda\}_{k=1}^\infty$ is a Cauchy sequence in $L^2(0, T; H^1(\Omega))$. To prove that it is a Cauchy sequence in $C(\bar{Q})$ as well, we handle the equation satisfied by $y_k^\lambda - y_m^\lambda$ in the

same way as we discussed (2.9). We use (2.7) to get

$$\|y_k^\lambda - y_m^\lambda\|_{C(\bar{Q})} \leq C_0(\varepsilon(1 + \|u\|_{\mathcal{M}[0,T]})\|y_k^\lambda - y_m^\lambda\|_{C(\bar{Q})} + \|y_{0k} - y_{0m}\|_{C(\bar{\Omega})}).$$

Taking again ε sufficiently small and $\lambda \geq \lambda_{\varepsilon,u}$, we deduce

$$\|y_k^\lambda - y_m^\lambda\|_{C(\bar{Q})} \leq C_2\|y_{0k} - y_{0m}\|_{C(\bar{\Omega})}.$$

Therefore, $\{y_k^\lambda - y_m^\lambda\}_{k=1}^\infty$ is a Cauchy sequence in $C(\bar{Q})$. Consequently, there exists $y^\lambda \in Y$ such that $y_k^\lambda \rightarrow y^\lambda$ in Y . It is easy to see that y^λ is a solution of (2.4). Now, we resubstitute $y_u(x, t) = e^{\lambda t}y^\lambda(x, t)$ and extend y_u to \bar{Q}_- by y_0 to get that y_u is a solution of (1.1).

II: *Uniqueness of the solution.* Let $y_1^\lambda, y_2^\lambda \in Y$ be two solutions of (2.4), and set $y^\lambda = y_2^\lambda - y_1^\lambda$. Subtracting the equations satisfied by y_2^λ and y_1^λ we obtain

$$\begin{cases} \frac{\partial y^\lambda}{\partial t} - \Delta y^\lambda + \frac{\partial R_\lambda^+}{\partial y}(t, \hat{y}^\lambda)y^\lambda = K_\lambda^+[u]y^\lambda & \text{in } Q, \\ \partial_n y^\lambda = 0 & \text{on } \Sigma, \\ y^\lambda(x, 0) = 0 & \text{in } \Omega, \end{cases}$$

where $\hat{y}^\lambda = y_1^\lambda + \theta(y_2^\lambda - y_1^\lambda)$ is some intermediate state with $0 \leq \theta(x, t) \leq 1$. Multiplying this equation by y^λ and invoking again (2.10) along with (2.7), we obtain

$$\begin{aligned} & \frac{1}{2}\|y^\lambda(T)\|_{L^2(\Omega)}^2 + \int_Q |\nabla y^\lambda|^2 dx dt + \int_Q |y^\lambda|^2 dx dt \\ & \leq \varepsilon(1 + \|u\|_{\mathcal{M}[0,T]})\|y^\lambda\|_{L^2(Q)}^2. \end{aligned}$$

Taking $\varepsilon < (1 + \|u\|_{\mathcal{M}[0,T]})^{-1}$, we conclude for $\lambda \geq \lambda_{\varepsilon,u}$ that $y^\lambda = 0$, since the last term in the left-hand side absorbs the right-hand side. Obviously the uniqueness of the solution of (2.4) is equivalent to the uniqueness of solution of (1.1).

III: *Estimates.* First we recall that $y_u(x, t) = e^{\lambda t}y^\lambda(x, t)$ is the solution of (1.1), once it has been extended to \bar{Q}_- by y_0 . Moreover, the following inequalities hold

$$\|y_u\|_{L^2(0,T;H^1(\Omega))} \leq e^{\lambda T}\|y^\lambda\|_{L^2(0,T;H^1(\Omega))} \quad \text{and} \quad \|y_u\|_{C(\bar{Q})} \leq e^{\lambda T}\|y^\lambda\|_{C(\bar{Q})}.$$

Therefore it is enough to establish the estimates for y^λ . To this end, we define this time $R_\lambda(t, y) = e^{-\lambda t}R(e^{\lambda t}y) + \lambda y$ with

$$\lambda \geq \lambda_R = 2(1 + \|u\|_{\mathcal{M}[0,T]}) - \min\{0, C_R\}.$$

Now, we multiply (2.4) by y^λ and deal with the reaction term as follows:

$$\begin{aligned} R_\lambda(t, y^\lambda)y^\lambda &= [R_\lambda(t, y^\lambda) - R_\lambda(t, 0)]y^\lambda + R_\lambda(t, 0)y^\lambda \\ &= \frac{\partial R_\lambda}{\partial y}(t, \theta y^\lambda)|y^\lambda|^2 + R_\lambda(t, 0)y^\lambda \geq 2(1 + \|u\|_{\mathcal{M}[0,T]})|y^\lambda|^2 - |R(0)||y^\lambda|, \end{aligned}$$

where $\theta, 0 < \theta(x, t) < 1$, is a measurable function in Q . Then, multiplying (2.4) by y^λ and using this inequality along with (2.5) and (2.6), we obtain for every $0 < T' < T$

and $Q_{T'} = \Omega \times (0, T')$,

$$\begin{aligned} & \frac{1}{2} \|y^\lambda(T')\|_{L^2(\Omega)}^2 + \int_{Q_{T'}} |\nabla y^\lambda|^2 \, dx \, dt + 2(1 + \|u\|_{\mathcal{M}[0,T]}) \int_{Q_{T'}} |y^\lambda|^2 \, dx \, dt \\ & \leq \int_{Q_{T'}} (K_\lambda[u]y^\lambda + g_{\lambda,u})y^\lambda \, dx \, dt + |R(0)| \int_{Q_{T'}} |y^\lambda| \, dx \, dt + \frac{1}{2} \|y_0(\cdot, 0)\|_{L^2(\Omega)}^2 \\ & \leq \|u\|_{\mathcal{M}[0,T]} (\|y^\lambda\|_{L^2(Q_{T'})} + \|y_0\|_{L^2(Q_-)}) \|y^\lambda\|_{L^2(Q_{T'})} \\ & \quad + |R(0)| |Q|^{\frac{1}{2}} \|y^\lambda\|_{L^2(Q_{T'})} + \frac{1}{2} \|y_0(\cdot, 0)\|_{L^2(\Omega)}^2 \\ & \leq \left(\frac{1}{2} + \|u\|_{\mathcal{M}[0,T]}\right) \|y^\lambda\|_{L^2(Q_{T'})}^2 + \frac{1}{2} \|u\|_{\mathcal{M}[0,T]}^2 \|y_0\|_{L^2(Q_-)}^2 \\ & \quad + \frac{|Q|}{2} |R(0)|^2 + \frac{1}{2} \|y^\lambda\|_{L^2(Q_{T'})}^2 + \frac{1}{2} \|y_0(\cdot, 0)\|_{L^2(\Omega)}^2 \\ & = (1 + \|u\|_{\mathcal{M}[0,T]}) \|y^\lambda\|_{L^2(Q_{T'})}^2 + \frac{1}{2} \|u\|_{\mathcal{M}[0,T]}^2 \|y_0\|_{L^2(Q_-)}^2 \\ & \quad + \frac{|Q|}{2} |R(0)|^2 + \frac{1}{2} \|y_0(\cdot, 0)\|_{L^2(\Omega)}^2. \end{aligned}$$

The first term of the right-hand side can be absorbed by the left-hand side. In this way, we get

$$(2.13) \quad \begin{aligned} & \|y^\lambda\|_{L^2(0,T;H^1(\Omega))} + \|y^\lambda\|_{C([0,T];L^2(\Omega))} \\ & \leq C(\|u\|_{\mathcal{M}[0,T]} \|y_0\|_{L^2(Q_-)} + |R(0)| + \|y_0(\cdot, 0)\|_{L^2(\Omega)}). \end{aligned}$$

To prove (2.3) we use the second inequality of (2.5), (2.6), and the results of [21, section III-7] applied to (2.4) to obtain

$$\begin{aligned} \|y^\lambda\|_{C(\bar{Q})} & \leq C \left(\|K_\lambda^+[u]y^\lambda\|_{L^\infty((0,T),L^2(\Omega))} + \|g_{\lambda,u}\|_{L^\infty(Q)} + \|y_0(\cdot, 0)\|_{C(\bar{\Omega})} + |R(0)| \right) \\ & \leq C \left(\|u\|_{\mathcal{M}[0,T]} \|y^\lambda\|_{C([0,T],L^2(\Omega))} + \|u\|_{\mathcal{M}[0,T]} \|y_0\|_{C(Q_-)} + \|y_0(\cdot, 0)\|_{C(\bar{\Omega})} + |R(0)| \right). \end{aligned}$$

Finally, from the equation satisfied by y^λ , the above estimates and the identity $y(x, t) = e^{\lambda t} y^\lambda(x, t)$ we conclude (2.2) and (2.3). \square

Let us prove some extra regularity of the solution of (1.1).

THEOREM 2.5. *Under the assumptions of Theorem 2.2, if $y_0(\cdot, 0) \in H^1(\Omega)$, then $y_u \in H^1(Q)$ and*

$$(2.14) \quad \|y_u\|_{H^1(Q)} \leq C_{1,1} (\|y_0\|_{L^2(Q_-)} \|u\|_{\mathcal{M}[0,T]} + \|y_0(\cdot, 0)\|_{H^1(\Omega)} + |R(0)|).$$

In addition, if either Ω is convex or Γ is of class $C^{1,1}$, then $y_u \in H^{2,1}(Q)$ and

$$(2.15) \quad \|y_u\|_{H^{2,1}(Q)} \leq C_{2,1} (\|y_0\|_{L^2(Q_-)} \|u\|_{\mathcal{M}[0,T]} + \|y_0(\cdot, 0)\|_{H^1(\Omega)} + |R(0)|).$$

The constants $C_{1,1}$ and $C_{2,1}$ depend on $\|u\|_{\mathcal{M}[0,T]}$, but they can be kept fixed on bounded subsets of $\mathcal{M}[0, T]$.

Proof. For the first part of the theorem we only have to prove that $\frac{\partial y}{\partial t}$ belongs to $L^2(Q)$ and to confirm the associated estimate. This is a simple consequence of a result that is known for linear parabolic equations; see, for instance, [31, section III.2]. Indeed, it is enough to write the equation in the form

$$\frac{\partial y}{\partial t} - \Delta y = K[u]y + g_u - R(y).$$

Thanks to [31, section III.2], the $H^1(Q)$ -norm of y_u can be estimated against the $L^2(Q)$ -norm of the right-hand side and, additionally, $y_u \in L^2(0, T; D(\Delta))$ holds. Therefore, if Ω is convex or Γ is of class $C^{1,1}$, then $D(\Delta) = H^2(\Omega)$ and the estimate (2.15) follows; see [14, Chapters 2 and 3]. \square

The next step of our analysis is the investigation of the differentiability properties of the control-to-state mapping $G : \mathcal{M}[0, T] \rightarrow Y$ that associates with $u \in \mathcal{M}[0, T]$ the solution y_u of (1.1), $G(u) = y_u$.

THEOREM 2.6. *The mapping G is of class C^1 . For every $u, v \in \mathcal{M}[0, T]$, we have that $z_v = G'(u)v$ is the solution of the problem*

$$(2.16) \quad \begin{cases} \frac{\partial z}{\partial t} - \Delta z + R'(y_u)z = K[u]z + K[v]y_u + g_v & \text{in } Q, \\ \partial_n z = 0 & \text{on } \Sigma, \\ z(x, 0) = 0 & \text{in } \Omega. \end{cases}$$

Proof. We define the space

$$\mathcal{Y} = \left\{ y \in Y : \frac{\partial y}{\partial t} - \Delta y \in L^\infty(Q) \right\},$$

endowed with the norm

$$\|y\|_{\mathcal{Y}} = \|y\|_Y + \left\| \frac{\partial y}{\partial t} - \Delta y \right\|_{L^\infty(Q)};$$

\mathcal{Y} is a Banach space. Now we consider the mapping

$$\begin{aligned} \mathcal{F} : \mathcal{Y} \times \mathcal{M}[0, T] &\longrightarrow L^\infty(Q) \times C(\bar{\Omega}), \\ \mathcal{F}(y, u) &= \left(\frac{\partial y}{\partial t} - \Delta y + R(y) - K[u]y - g_u, y(\cdot, 0) - y_0(\cdot, 0) \right). \end{aligned}$$

It is obvious that \mathcal{F} is well-defined and is of class C^1 . Moreover, we have that

$$\frac{\partial \mathcal{F}}{\partial y}(y, u)z = \left(\frac{\partial z}{\partial t} - \Delta z + R'(y)z - K[u]z, z(\cdot, 0) \right).$$

Let us confirm that $\frac{\partial \mathcal{F}}{\partial y}(y, u) : \mathcal{Y} \rightarrow L^\infty(Q) \times C(\bar{\Omega})$ is an isomorphism. Indeed, since obviously $\frac{\partial \mathcal{F}}{\partial y}(y, u)$ is a linear and continuous mapping, we only need to prove that, for every pair $(f, z_0) \in L^\infty(Q) \times C(\bar{\Omega})$, there exists a unique solution $z \in \mathcal{Y}$ of the problem

$$\begin{cases} \frac{\partial z}{\partial t} - \Delta z + R'(y)z = K[u]z + f & \text{in } Q, \\ \partial_n z = 0 & \text{on } \Sigma, \\ z(x, 0) = z_0 & \text{in } \Omega. \end{cases}$$

The existence and uniqueness of a solution z in Y is proved in the same way as for the problem (1.1). Moreover, from the partial differential equation and the fact that $z \in C(\bar{Q})$, the inclusion $\frac{\partial z}{\partial t} - \Delta z \in L^\infty(Q)$ follows and, consequently, $z \in \mathcal{Y}$ holds. Hence, an application of the implicit function theorem implies that G is of class C^1 . Equation (2.16) follows easily by differentiating the identity $\mathcal{F}(G(u), u) = 0$ with respect to u . \square

Remark 2.7. Let us mention that $z_v = G'(u)v \in H^1(Q)$ holds for every $v \in \mathcal{M}[0, T]$. This follows from (2.16) arguing similarly as in the proof of Theorem 2.5 and taking into account that $z_v(x, 0) = 0$.

3. The control problem. Now we have all the prerequisites to study our optimal control problem, namely,

$$(P) \quad \min_{u \in \mathcal{M}[0,T]} J(u) = \frac{1}{2} \int_Q |y_u - y_d|^2 \, dx \, dt + \nu \|u\|_{\mathcal{M}[0,T]},$$

where $y_d \in L^{\bar{p}}(Q)$ for some $\bar{p} > 1 + \frac{n}{2}$ and $\nu > 0$ are given.

THEOREM 3.1. *Problem (P) has at least one solution \bar{u} .*

Before proving this theorem we state the following lemma.

LEMMA 3.2. *Assume that $u_k \xrightarrow{*} \bar{u}$ in $\mathcal{M}[0, T]$ for $k \rightarrow \infty$ and let y_k and \bar{y} be the states associated with u_k and \bar{u} , respectively; then $y_k \rightarrow \bar{y}$ in Y .*

Proof. Since $u_k \xrightarrow{*} \bar{u}$ in $\mathcal{M}[0, T]$ as $k \rightarrow \infty$, we know that there exists a constant $M > 0$ such that $\|u_k\|_{\mathcal{M}[0,T]} \leq M \, \forall k \geq 1$. Hence, $\|\bar{u}\|_{\mathcal{M}[0,T]} \leq M$ holds as well. Set $y_k^\lambda(x, t) = e^{-\lambda t} y_k(x, t)$ and $\bar{y}^\lambda(x, t) = e^{-\lambda t} \bar{y}(x, t)$. Then y_k^λ and \bar{y}^λ satisfy (2.4) for controls $u := u_k$ and $u := \bar{u}$, respectively. Let us define $w_k^\lambda = y_k^\lambda - \bar{y}^\lambda$. Then, subtracting these two equations and taking again $R_\lambda(t, y) = e^{-\lambda t} R(e^{\lambda t} y) + \lambda y$ with $\lambda \geq 2(M + 1) - \min\{0, C_R\}$, we get

$$(3.1) \quad \begin{cases} \frac{\partial w_k^\lambda}{\partial t} - \Delta w_k^\lambda + \frac{\partial R_\lambda}{\partial y}(t, \hat{y}_k^\lambda) w_k^\lambda = K_\lambda[u_k] y_k^\lambda - K_\lambda[\bar{u}] \bar{y}^\lambda + g_{\lambda, u_k} - g_{\lambda, \bar{u}} & \text{in } Q, \\ \partial_n w_k^\lambda = 0 & \text{on } \Sigma, \\ w_k^\lambda(x, 0) = 0 & \text{in } \Omega. \end{cases}$$

with intermediate states \hat{y}_k^λ .

Testing this equation by w_k^λ and invoking (2.5), we get for every $0 < T' < T$

$$\begin{aligned} & \frac{1}{2} \|w_k^\lambda(T')\|_{L^2(\Omega)}^2 + \int_{Q_{T'}} [|\nabla w_k^\lambda|^2 + 2(M + 1)|w_k^\lambda|^2] \, dx \, dt \\ & \leq \int_{Q_{T'}} [K_\lambda[u_k] y_k^\lambda - K_\lambda[\bar{u}] \bar{y}^\lambda + g_{\lambda, u_k} - g_{\lambda, \bar{u}}] w_k^\lambda \, dx \, dt \\ & = \int_{Q_{T'}} [K_\lambda[u_k] w_k^\lambda + K_\lambda[u_k - \bar{u}] \bar{y}^\lambda + g_{\lambda, u_k - \bar{u}}] w_k^\lambda \, dx \, dt \\ & \leq \|u_k\|_{\mathcal{M}[0,T]} \|w_k^\lambda\|_{L^2(Q_{T'})}^2 + \|K_\lambda[u_k - \bar{u}] \bar{y}^\lambda + g_{\lambda, u_k - \bar{u}}\|_{L^2(Q)} \|w_k^\lambda\|_{L^2(Q_{T'})} \\ & \leq (M + 1) \|w_k^\lambda\|_{L^2(Q_{T'})}^2 + \frac{1}{2} \|K_\lambda[u_k - \bar{u}] \bar{y}^\lambda + g_{\lambda, u_k - \bar{u}}\|_{L^2(Q)}^2. \end{aligned}$$

The first term of the right-hand side can be absorbed by the left-hand side and we infer

$$(3.2) \quad \|w_k^\lambda\|_{L^2(0,T;H^1(\Omega))} + \|w_k^\lambda\|_{C([0,T],L^2(\Omega))} \leq C(\|K_\lambda[u_k - \bar{u}] \bar{y}^\lambda + g_{\lambda, u_k - \bar{u}}\|_{L^2(Q)}).$$

Let us prove that the right-hand side of the inequality converges to zero. From the convergence $u_k \xrightarrow{*} \bar{u}$ in $\mathcal{M}[0, T]$ and by the continuity of \bar{y} we get for $k \rightarrow \infty$

$$(K_\lambda[u_k - \bar{u}] \bar{y}^\lambda + g_{\lambda, u_k - \bar{u}})(x, t) = \int_0^T e^{-\lambda s} \bar{y}^\lambda(x, t - s) d(u_k - \bar{u})(s) \rightarrow 0 \quad \forall (x, t) \in Q,$$

i.e., pointwise convergence. Moreover, from (2.5) and (2.6) we have

$$\begin{aligned} & \|K_\lambda[u_k - \bar{u}] \bar{y}^\lambda + g_{\lambda, u_k - \bar{u}}\|_{L^\infty(Q)} \leq (\|\bar{y}^\lambda\|_{C(\bar{Q})} + \|y_0\|_{C(\bar{Q}_-)}) \|u_k - \bar{u}\|_{\mathcal{M}[0,T]} \\ & \leq 2M(\|\bar{y}^\lambda\|_{C(\bar{Q})} + \|y_0\|_{C(\bar{Q}_-)}) \quad \forall k. \end{aligned}$$

From the Lebesgue dominated convergence theorem we conclude that $K_\lambda[u_k - \bar{u}]\bar{y}^\lambda + g_{\lambda, u_k - \bar{u}} \rightarrow 0$ in $L^p(Q)$ for every $p < \infty$. Therefore, we infer from (3.2) the convergence $w_k^\lambda \rightarrow 0$ in $L^2(0, T; H^1(\Omega)) \cap C([0, T], L^2(\Omega))$ and, hence, $y_k^\lambda \rightarrow \bar{y}^\lambda$ in $L^2(0, T; H^1(\Omega)) \cap C([0, T], L^2(\Omega))$.

Let us show the uniform convergence. From (3.1), using the estimates of [21, section III-8] and (2.5), we infer for $p > 1 + \frac{n}{2}$

$$\begin{aligned} \|w_k^\lambda\|_{C(\bar{Q})} &\leq C_1 \|K_\lambda[u_k]w_k^\lambda\|_{L^\infty((0, T), L^2(\Omega))} + C_2 \|K_\lambda[u_k - \bar{u}]\bar{y}^\lambda + g_{\lambda, u_k - \bar{u}}\|_{L^p(\Omega)} \\ &\leq C_1 M \|w_k^\lambda\|_{C([0, T], L^2(\Omega))} + C_2 \|K_\lambda[u_k - \bar{u}]\bar{y}^\lambda + g_{\lambda, u_k - \bar{u}}\|_{L^p(Q)} \rightarrow 0 \text{ as } k \rightarrow \infty. \end{aligned}$$

We have proved that $y_k^\lambda \rightarrow \bar{y}^\lambda$ in Y . Transforming y_k^λ and \bar{y}^λ back to y_k and \bar{y} , this leads to $\|y_k - \bar{y}\|_Y \rightarrow 0$. \square

Proof of Theorem 3.1. Let $\{u_k\}_{k=1}^\infty \subset \mathcal{M}[0, T]$ be a minimizing sequence of (P). Since

$$\nu \|u_k\|_{\mathcal{M}[0, T]} \leq J(u_k) \leq J(0) < +\infty,$$

we deduce that $\{u_k\}_{k=1}^\infty$ is bounded in $\mathcal{M}[0, T]$. Hence, we can extract a subsequence, denoted in the same way, such that $u_k \xrightarrow{*} \bar{u}$ in $\mathcal{M}[0, T]$. Denote by y_k and \bar{y} the states associated with u_k and \bar{u} , respectively. From Lemma 3.2 we know that $y_k \rightarrow \bar{y}$ in $L^2(Q)$. This convergence implies that $J(\bar{u}) \leq \liminf_{k \rightarrow \infty} J(u_k) = \inf(P)$ and, hence, \bar{u} is a solution of (P). \square

Next we derive the first order optimality conditions that have to be satisfied by any local solution of the problem (P). We distinguish between two different types of local solutions. To this end, we recall that $\mathcal{M}[0, T] \subset H^1(0, T)^*$, the embedding being continuous and compact. Notice that $H^1(0, T)$ is compactly embedded in $C[0, T]$ and then by transposition we deduce the compactness of $\mathcal{M}[0, T] \subset H^1(0, T)^*$.

DEFINITION 3.3. *A control \bar{u} is called a local solution or local minimum of (P) if there exists a ball $B_\varepsilon(\bar{u}) \subset \mathcal{M}[0, T]$ such that $J(\bar{u}) \leq J(u) \forall u \in B_\varepsilon(\bar{u})$.*

Let us define the two different functionals forming $J(u) = F(u) + \nu j(u)$ by

$$F(u) = \frac{1}{2} \int_Q |y_u - y_d|^2 dx dt \quad \text{and} \quad j(u) = \|u\|_{\mathcal{M}[0, T]}.$$

THEOREM 3.4. *The functional $F : \mathcal{M}[0, T] \rightarrow \mathbb{R}$ is of class C^1 . Its derivative is given by*

$$(3.3) \quad F'(u)v = \int_Q \varphi_u (K[v]y_u + g_v) dx dt \quad \forall u, v \in \mathcal{M}[0, T],$$

where $\varphi_u \in H^1(Q) \cap C(\bar{Q})$ is the solution of the adjoint state equation

$$(3.4) \quad \begin{cases} -\frac{\partial \varphi}{\partial t} - \Delta \varphi + R'(y_u)\varphi = K^*[u]\varphi + y_u - y_d & \text{in } Q, \\ \partial_n \varphi = 0 & \text{on } \Sigma, \\ \varphi(x, T) = 0 & \text{in } \Omega, \end{cases}$$

and the operator K^* is defined by

$$(3.5) \quad (K^*[u]w)(x, t) = \int_{[0, T-t]} w(x, t+s) du(s) \quad \forall w \in C(\bar{Q}).$$

Before proving this theorem we analyze the adjoint state equation (3.4).

PROPOSITION 3.5. *For all $u \in \mathcal{M}[0, T]$, there exists a unique solution $\varphi \in H^1(Q) \cap C(\bar{Q})$ of (3.4) and it holds*

$$(3.6) \quad \|\varphi\|_{H^1(Q)} \leq M_{1,1} \|y_u - y_d\|_{L^2(Q)},$$

$$(3.7) \quad \|\varphi\|_{C(\bar{Q})} \leq M_\infty \|y_u - y_d\|_{L^{\bar{p}}(Q)},$$

where \bar{p} is the order of integrability of y_d introduced in (P). Moreover, if either Γ is of class $C^{1,1}$ or Ω is convex, then $\varphi \in H^{2,1}(Q)$ and

$$(3.8) \quad \|\varphi\|_{H^{2,1}(Q)} \leq M_{2,1} \|y_u - y_d\|_{L^2(Q)}.$$

The constants $M_{1,1}$, M_∞ , and $M_{2,1}$ depend on u , but they can be taken fixed on bounded subsets of $\mathcal{M}[0, T]$.

Proof. Given $\lambda > 0$, we set $\psi^\lambda(x, t) = e^{-\lambda t} \varphi(x, T - t)$ in Q . Then we have $(K_\lambda[u] \psi^\lambda)(x, t) = (e^{-\lambda t} K^*[u] \varphi)(x, T - t)$, and (3.4) is transformed to the forward equation

$$(3.9) \quad \begin{cases} \frac{\partial \psi^\lambda}{\partial t} - \Delta \psi^\lambda + R'(\hat{y}_u) \psi^\lambda + \lambda \psi^\lambda = K_\lambda[u] \psi + f & \text{in } Q, \\ \partial_n \psi^\lambda = 0 & \text{on } \Sigma, \\ \psi^\lambda(x, 0) = 0 & \text{in } \Omega, \end{cases}$$

where $\hat{y}_u(x, t) = y_u(x, T - t)$ and $f(x, t) = (y_u - y_d)(x, T - t)$. Now, we can argue as in Theorems 2.2 and 2.5 to get the existence, uniqueness, and regularity. The only difference is that $f \in L^{\bar{p}}(Q)$ with $\bar{p} > 1 + \frac{n}{2}$, which is enough to deduce the Hölder regularity of the solution of (3.9); see [21, section III-10]. \square

Let us observe that, in some sense, the operator $K^*[u]$ is the adjoint of $K[u]$ with respect to the $L^2(Q)$ scalar product. Indeed, given $w, z \in C(\bar{Q})$, applying Fubini's theorem, and making the change of variables $\tau = t + s$ we get

$$(3.10) \quad \begin{aligned} & \int_Q (K^*[u]w)(x, t) z(x, t) \, dx \, dt = \int_Q \left(\int_{[0, T-t]} w(x, t + s) \, du(s) \right) z(x, t) \, dx \, dt \\ & = \int_\Omega \int_{[0, T)} \left(\int_0^{T-s} w(x, t + s) z(x, t) \, dt \right) \, du(s) \, dx \\ & = \int_\Omega \int_{[0, T)} \left(\int_s^T w(x, \tau) z(x, \tau - s) \, d\tau \right) \, du(s) \, dx \\ & = \int_\Omega \int_0^T \left(\int_{[0, \tau)} z(x, \tau - s) \, du(s) \right) w(x, \tau) \, d\tau \, dx \\ & = \int_Q (K[u]z)(x, \tau) w(x, \tau) \, dx \, d\tau. \end{aligned}$$

Therefore, $K^*[u]$ is the transposition of $K[u]$ with respect to the scalar product of $L^2(Q)$, but restricted to the space $C(\bar{Q})$.

Proof of Theorem 3.4. Let us set $z_v = G'(u)v$. Thanks to Remark 2.7 and Proposition 3.5, we have that $z_v, \varphi_u \in H^1(Q)$. Hence, we can multiply (3.4) by

z_v and perform an integration by parts. Using (3.10), (2.16), and the fact that $\varphi_u(x, T) = z_v(x, 0) = 0$ in Ω , we get

$$\begin{aligned} F'(u)v &= \int_Q (y_u - y_d)z_v \, dx \, dt \\ &= \int_Q \left[-\frac{\partial \varphi_u}{\partial t} z_v + \nabla \varphi_u \nabla z_v + R'(y_u)\varphi_u z_v - (K^*[u]\varphi_u)z_v \right] \, dx \, dt \\ &= \int_Q \left[\frac{\partial z_v}{\partial t} \varphi_u + \nabla \varphi_u \nabla z_v + R'(y_u)z_v \varphi_u - (K[u]z_v)\varphi_u \right] \, dx \, dt \\ &= \int_Q \left(K[v]y_u + g_v \right) \varphi_u \, dx \, dt, \end{aligned}$$

which proves (3.3). \square

We continue by studying the function $j : \mathcal{M}[0, T] \rightarrow \mathbb{R}$. Since j is Lipschitz and convex, we know that it has a nonempty subdifferential at every point $u \in \mathcal{M}[0, T]$. As usual, it will be denoted by $\partial j(u)$. Let us prove an interesting property of $\partial j(u)$ for the sparsity analysis of (P); see [9] and [10] for similar results.

PROPOSITION 3.6. *If $\lambda \in \partial j(\bar{u})$ with $\bar{u} \neq 0$ and $\lambda \in C[0, T]$, then the following properties hold:*

$$(3.11) \quad \|\lambda\|_{C[0, T]} = 1,$$

$$(3.12) \quad \begin{aligned} \text{supp}(\bar{u}^+) &\subset \{t \in [0, T] : \lambda(t) = +1\}, \\ \text{supp}(\bar{u}^-) &\subset \{t \in [0, T] : \lambda(t) = -1\}, \end{aligned}$$

where $\bar{u} = \bar{u}^+ - \bar{u}^-$ is the Jordan decomposition of the measure \bar{u} .

Proof. By definition of the subdifferential, we have

$$(3.13) \quad \langle u - \bar{u}, \lambda \rangle_{\mathcal{M}[0, T], C[0, T]} + j(\bar{u}) \leq j(u) \quad \forall u \in \mathcal{M}[0, T].$$

Taking $u = 0$ and $u = 2\bar{u}$, respectively, in (3.13) we deduce that $\langle \bar{u}, \lambda \rangle_{\mathcal{M}[0, T], C[0, T]} = j(\bar{u})$. Hence (3.13) implies that

$$\langle u, \lambda \rangle_{\mathcal{M}[0, T], C[0, T]} \leq j(u) = \|u\|_{\mathcal{M}[0, T]} \quad \forall u \in \mathcal{M}[0, T].$$

Now, for every $s \in [0, T]$ we take $u = \pm \delta_s$ in the above inequality. This leads to

$$(3.14) \quad |\lambda(s)| \leq 1 \quad \forall s \in [0, T].$$

By the established properties, we find

$$\|\bar{u}\|_{\mathcal{M}[0, T]} = j(\bar{u}) = \int_0^T \lambda(s) \, d\bar{u}(s) \leq \int_0^T |\lambda(s)| \, d|\bar{u}|(s) \leq \int_0^T d|\bar{u}|(s) = \|\bar{u}\|_{\mathcal{M}[0, T]};$$

therefore,

$$\int_0^T [1 - |\lambda(s)|] \, d|\bar{u}|(s) = 0 \quad \text{and} \quad \int_0^T \lambda(s) \, d\bar{u}(s) = \int_0^T |\lambda(s)| \, d|\bar{u}|(s).$$

The second identity and (3.14) imply (3.11). Let us prove (3.12). From (3.11) we infer

$$\begin{aligned} 0 &\leq \int_0^T (1 - \lambda(s)) \, d\bar{u}^+(s) + \int_0^T (1 + \lambda(s)) \, d\bar{u}^-(s) \\ &= \int_0^T d|\bar{u}|(s) - \int_0^T \lambda(s) \, d\bar{u}(s) = \int_0^T (1 - |\lambda(s)|) \, d|\bar{u}|(s) = 0. \end{aligned}$$

Hence, we get

$$\int_0^T (1 - \lambda(s)) d\bar{u}^+(s) = \int_0^T (1 + \lambda(s)) d\bar{u}^-(s) = 0,$$

which proves (3.12). □

THEOREM 3.7. *Let \bar{u} be a local solution of (P). Then there exist $\bar{y} \in Y \cap C(\bar{Q} \cup \bar{Q}_-)$, $\bar{\varphi} \in H^1(Q) \cap C(\bar{Q})$, and $\bar{\lambda} \in C[0, T] \cap \partial j(\bar{u})$ such that*

$$(3.15) \quad \begin{cases} \frac{\partial \bar{y}}{\partial t} - \Delta \bar{y} + R(\bar{y}) = K[\bar{u}]\bar{y} + g_{\bar{u}} & \text{in } Q, \\ \partial_n \bar{y} = 0 & \text{on } \Sigma, \\ \bar{y}(x, t) = y_0(x, t) & \text{in } Q_-, \end{cases}$$

$$(3.16) \quad \begin{cases} -\frac{\partial \bar{\varphi}}{\partial t} - \Delta \bar{\varphi} + R'(\bar{y})\bar{\varphi} = K^*[\bar{u}]\bar{\varphi} + \bar{y} - y_d & \text{in } Q, \\ \partial_n \bar{\varphi} = 0 & \text{on } \Sigma, \\ \bar{\varphi}(x, T) = 0 & \text{in } \Omega, \end{cases}$$

$$(3.17) \quad \bar{\lambda}(s) = -\frac{1}{\nu} \int_{-s}^{T-s} \int_{\Omega} \bar{y}(x, t) \bar{\varphi}(x, t+s) dx dt \quad \forall s \in [0, T].$$

Proof. The existence and uniqueness of solutions to (3.15) and (3.16) have already been discussed in Theorem 2.2 and Proposition 3.5. Now, we define $\bar{\lambda}$ by (3.17). The continuity of \bar{y} and $\bar{\varphi}$ implies that $\bar{\lambda} \in C[0, T]$. It remains to prove that $\bar{\lambda} \in \partial j(\bar{u})$. To this end, we use that \bar{u} is a local minimizer of (P). Hence, for any $u \in \mathcal{M}[0, T]$, we get from the convexity of j and (3.3) that

$$(3.18) \quad \begin{aligned} 0 &\leq \lim_{\rho \searrow 0} \frac{J(\bar{u} + \rho(u - \bar{u})) - J(\bar{u})}{\rho} \leq F'(\bar{u})(u - \bar{u}) + \nu j(u) - \nu j(\bar{u}) \\ &= \int_Q \bar{\varphi} (K[u - \bar{u}]\bar{y} + g_{u-\bar{u}}) dx dt + \nu j(u) - \nu j(\bar{u}). \end{aligned}$$

The last integral can be expressed in terms of $\bar{\lambda}$ and $u - \bar{u}$ as follows:

$$\begin{aligned} &\int_Q \bar{\varphi} (K[u - \bar{u}]\bar{y} + g_{u-\bar{u}}) dx dt \\ &= \int_{\Omega} \int_0^T \bar{\varphi}(x, t) \left(\int_0^T \bar{y}(x, t-s) d(u - \bar{u})(s) \right) dt dx \\ &= \int_0^T \int_{\Omega} \int_0^T \bar{y}(x, t-s) \bar{\varphi}(x, t) dt dx d(u - \bar{u})(s) \\ &= \int_0^T \int_{\Omega} \int_{-s}^{T-s} \bar{y}(x, \tau) \bar{\varphi}(x, \tau+s) d\tau dx d(u - \bar{u})(s) \\ &= -\nu \int_0^T \bar{\lambda}(s) d(u - \bar{u})(s). \end{aligned}$$

Combining this with (3.18), we find

$$\int_0^T \bar{\lambda}(s) d(u - \bar{u})(s) + j(\bar{u}) \leq j(u) \quad \forall u \in \mathcal{M}[0, T].$$

This is the definition of $\bar{\lambda} \in \partial j(\bar{u})$. □

From Proposition 3.6 and Theorem 3.7 we deduce the following sparsity structure of the optimal control \bar{u} .

COROLLARY 3.8. *Let \bar{u} be a local minimum of (P) and let \bar{y} , $\bar{\varphi}$, and $\bar{\lambda}$ satisfy the optimality system (3.15)–(3.17), then if $\bar{u} \neq 0$*

$$(3.19) \quad \|\bar{\lambda}\|_{C[0,T]} = 1,$$

$$(3.20) \quad \begin{aligned} \text{supp}(\bar{u}^+) &\subset \{t \in [0, T] : \bar{\lambda}(t) = +1\}, \\ \text{supp}(\bar{u}^-) &\subset \{t \in [0, T] : \bar{\lambda}(t) = -1\}, \end{aligned}$$

where $\bar{u} = \bar{u}^+ - \bar{u}^-$ is the Jordan decomposition of the measure \bar{u} .

PROPOSITION 3.9. *There exists $\bar{\nu} > 0$ such that 0 is the only solution of (P) for every $\nu \geq \bar{\nu}$.*

Proof. Let \bar{u} be a solution of (P). From the inequality $J(\bar{u}) \leq J(0)$ we deduce that

$$(3.21) \quad \|\bar{y}\|_{L^2(Q)} \leq C_1 < \infty \quad \text{and} \quad \|\bar{u}\|_{\mathcal{M}[0,T]} \leq \frac{C_2}{\nu}$$

for some constants independent of ν . Arguing similarly as in the proof of inequality (2.13), we get from (3.16) that

$$(3.22) \quad \|\bar{\varphi}\|_{C([0,T],L^2(\Omega))} \leq C_3 \|\bar{y} - y_d\|_{L^2(Q)} \leq C_4.$$

According to (3.21), C_3 and C_4 are independent of $\nu \geq 1$. Now, from (3.17), (3.21), and (3.22) we get

$$\begin{aligned} |\bar{\lambda}(s)| &\leq \frac{1}{\nu} \int_{-s}^T \|\bar{y}(t)\|_{L^2(\Omega)} \|\bar{\varphi}(t+s)\|_{L^2(\Omega)} dt \\ &\leq \frac{1}{\nu} \left(\int_{-T}^T \|\bar{y}(t)\|_{L^2(\Omega)} dt \right) \|\bar{\varphi}\|_{C([0,T],L^2(\Omega))} \\ &= \frac{1}{\nu} \left(\int_{-T}^0 \|y_0(t)\|_{L^2(\Omega)} dt + \int_0^T \|\bar{y}(t)\|_{L^2(\Omega)} dt \right) \|\bar{\varphi}\|_{C([0,T],L^2(\Omega))} \\ &\leq \frac{\sqrt{T}}{\nu} (\|y_0\|_{L^2(Q_-)} + \|\bar{y}\|_{L^2(Q)}) \|\bar{\varphi}\|_{C([0,T],L^2(\Omega))} \leq \frac{\sqrt{T}}{\nu} (\|y_0\|_{L^2(Q_-)} + C_1) C_4. \end{aligned}$$

If we take $\bar{\nu} > \max\{1, \sqrt{T}(\|y_0\|_{L^2(Q_-)} + C_1)C_4\}$ we infer that $|\bar{\lambda}(s)| < 1 \forall s \in [0, T]$. Then, (3.20) implies that $\bar{u} \equiv 0$. \square

To finish this section, we prove the following result concerning the dependence of the measure norm of the (globally) optimal controls with respect to the sparsity parameter $\nu > 0$.

PROPOSITION 3.10. *Let $\{\nu_k\}$ be a sequence of positive real numbers such that $\nu_k \rightarrow \nu > 0$ and let, for each k , u_k be a solution of*

$$\min_{u \in \mathcal{M}[0,T]} J_k(u) = \frac{1}{2} \int_Q |y_u - y_d|^2 dx dt + \nu_k \|u\|_{\mathcal{M}[0,T]},$$

where $y_u \in Y$ is the solution of (1.1). Then, the sequence $\{u_k\}$ is bounded in $\mathcal{M}[0, T]$ and for every subsequence that converges weakly* to some \bar{u} in $\mathcal{M}[0, T]$, we have that \bar{u} is a solution of (P), and $\|u_k\|_{\mathcal{M}[0,T]} \rightarrow \|\bar{u}\|_{\mathcal{M}[0,T]}$.

Proof. Let \tilde{u} be a global solution of (P). Notice that $\nu_k \|u_k\|_{\mathcal{M}[0,T]} \leq J_k(u_k) \leq J_k(0) = J(0)$. Since ν_k is convergent to $\nu > 0$, we have that $\{u_k\}$ is a bounded sequence in $\mathcal{M}[0, T]$. Hence, we can extract a subsequence, denoted in the same way, such that $u_k \overset{*}{\rightharpoonup} \tilde{u}$ in $\mathcal{M}[0, T]$. From Lemma 3.2, we have that $y_{u_k} \rightarrow y_{\tilde{u}}$ in Y . This convergence and the optimality of every u_k lead to

$$J(\tilde{u}) \leq \liminf_{k \rightarrow \infty} J_k(u_k) \leq \limsup_{k \rightarrow \infty} J_k(u_k) \leq \limsup_{k \rightarrow \infty} J_k(\tilde{u}) = J(\tilde{u}) = \inf (P).$$

Therefore \tilde{u} is a solution of (P) and $J_k(u_k) \rightarrow J(\tilde{u})$. Finally, the last convergence, $y_{u_k} \rightarrow y_{\tilde{u}}$ in Y , and $\nu_k \rightarrow \nu$ imply $\|u_k\|_{\mathcal{M}[0,T]} \rightarrow \|\tilde{u}\|_{\mathcal{M}[0,T]}$. \square

4. Discretization of the control space. In this section, we first consider the approximation of $\mathcal{M}[0, T]$ by finite dimensional subspaces \mathcal{U}_τ . Then, associated with each space \mathcal{U}_τ we define a new problem (P_τ) and analyze the convergence of the solutions of (P_τ) to solutions of (P). We do not discuss a convergence analysis of the discretization of the state and adjoint state equation, since this is beyond the scope of the paper.

Let us consider a grid of points $0 = t_0 < t_1 < \dots < t_{N_\tau} = T$. We set $\tau_k = t_k - t_{k-1}$ for $1 \leq k \leq N_\tau$ and $\tau = \max_{1 \leq k \leq N_\tau} \tau_k$. We also set $I_k = (t_{k-1}, t_k]$ for $1 \leq k \leq N_\tau$, and $I_0 = \{0\}$. Associated with this grid we define the space

$$\mathcal{U}_\tau = \left\{ u_\tau = \sum_{k=0}^{N_\tau} u_k \delta_{t_k} : (u_k)_{k=0}^{N_\tau} \in \mathbb{R}^{N_\tau+1} \right\},$$

where δ_{t_k} denotes the Dirac measure centered at t_k . Thus, \mathcal{U}_τ has dimension $N_\tau + 1$ and \mathcal{U}_τ is a vector subspace of $\mathcal{M}[0, T]$. Now, we introduce the linear mapping

$$\Lambda_\tau : \mathcal{M}[0, T] \longrightarrow \mathcal{U}_\tau \text{ defined by } \Lambda_\tau u = \sum_{k=0}^{N_\tau} u(I_k) \delta_{t_k}.$$

The following proposition states some properties of this mapping.

PROPOSITION 4.1. *The following statements hold:*

1. $\|\Lambda_\tau u\|_{\mathcal{M}[0,T]} \leq \|u\|_{\mathcal{M}[0,T]} \quad \forall u \in \mathcal{M}[0, T]$.
2. $\Lambda_\tau u \overset{*}{\rightharpoonup} u$ in $\mathcal{M}[0, T] \quad \forall u \in \mathcal{M}[0, T]$.
3. $\lim_{\tau \rightarrow 0} \|\Lambda_\tau u\|_{\mathcal{M}[0,T]} = \|u\|_{\mathcal{M}[0,T]} \quad \forall u \in \mathcal{M}[0, T]$.

Remark 4.2. Since $\mathcal{M}[0, T]$ is a nonseparable Banach space, it cannot be approximated in the norm by means of finite dimensional spaces. In this sense, statements 2 and 3 in Proposition 4.1 mean that \mathcal{U}_τ is a good approximation of $\mathcal{M}[0, T]$. For a different approach in the definition of Λ_τ , see [8, Theorem 3.1].

Proof of Proposition 4.1. 1. It is obtained as follows:

$$\|\Lambda_\tau u\|_{\mathcal{M}[0,T]} = \sum_{k=0}^{N_\tau} |u(I_k)| \leq \sum_{k=0}^{N_\tau} |u(I_k)| = |u|([0, T]) = \|u\|_{\mathcal{M}[0,T]}.$$

2. Let us take $y \in C[0, T]$. Given an arbitrary $\varepsilon > 0$, the continuity of y implies that there exists $\tau_\varepsilon > 0$ such that

$$(4.1) \quad |y(t) - y(s)| < \varepsilon \quad \forall s, t \in [0, T] \text{ such that } |t - s| < \tau_\varepsilon.$$

Then for every $\tau < \tau_\varepsilon$ we have

$$\begin{aligned} |\langle u - \Lambda_\tau u, y \rangle| &= \left| \int_0^T y(s) \, du(s) - \sum_{k=0}^{N_\tau} y(t_k) u(I_k) \right| = \left| \sum_{k=0}^{N_\tau} \int_{I_k} [y(s) - y(t_k)] \, du(s) \right| \\ &\leq \sum_{k=0}^{N_\tau} \int_{I_k} |y(s) - y(t_k)| \, |du(s)| \leq \varepsilon \|u\|_{\mathcal{M}[0,T]}. \end{aligned}$$

Since y is an arbitrary element of $C[0, T]$, this proves that $\Lambda_\tau u \overset{*}{\rightharpoonup} u$ in $\mathcal{M}[0, T]$.

3. Combining 2 and 1 we get

$$\|u\|_{\mathcal{M}[0,T]} \leq \liminf_{\tau \rightarrow 0} \|\Lambda_\tau u\|_{\mathcal{M}[0,T]} \leq \limsup_{\tau \rightarrow 0} \|\Lambda_\tau u\|_{\mathcal{M}[0,T]} \leq \|u\|_{\mathcal{M}[0,T]},$$

which concludes the proof. □

Now, for every $\tau > 0$ we consider the control problem with discretized controls

$$(P_\tau) \quad \min_{u_\tau \in \mathcal{U}_\tau} J(u_\tau) = \frac{1}{2} \int_Q |y_{u_\tau} - y_d|^2 \, dx \, dt + \nu \sum_{k=0}^{N_\tau} |u_k|.$$

From Lemma 3.2 we deduce the continuity of the functional $J : \mathcal{U}_\tau \rightarrow \mathbb{R}$. Therefore, taking into account that \mathcal{U}_τ is a finite dimensional vector space and J is coercive on \mathcal{U}_τ , we deduce the existence of at least one global solution \bar{u}_τ of (P_τ) . Let us study the sparse structure of the solutions \bar{u}_τ of (P_τ) . We denote by $j_\tau : \mathcal{U}_\tau \rightarrow \mathbb{R}$ the restriction of j to \mathcal{U}_τ :

$$j_\tau(u_\tau) = j(u_\tau) = \sum_{k=0}^{N_\tau} |u_k|.$$

We identify the dual of \mathcal{U}_τ with $\mathbb{R}^{N_\tau+1}$ as follows:

$$\forall \lambda_\tau = (\lambda_k)_{k=0}^{N_\tau} \in \mathbb{R}^{N_\tau+1} \text{ and } \forall u_\tau = \sum_{k=0}^{N_\tau} u_k \delta_{t_k} \text{ we set } \langle \lambda_\tau, u_\tau \rangle = \sum_{k=0}^{N_\tau} \lambda_k u_k.$$

Then Proposition 3.6 is reformulated as follows.

PROPOSITION 4.3. *With the above notation, we have $\lambda_\tau \in \partial j_\tau(\bar{u}_\tau)$ if and only if the following identity holds:*

$$(4.2) \quad \lambda_k \begin{cases} = +1 & \text{if } \bar{u}_k > 0, \\ = -1 & \text{if } \bar{u}_k < 0, \\ \in [-1, +1] & \text{if } \bar{u}_k = 0. \end{cases} \quad 0 \leq k \leq N_\tau,$$

Proof. By definition of the subdifferential we have that $\lambda_\tau \in \partial j_\tau(\bar{u}_\tau)$ if and only if

$$\sum_{k=0}^{N_\tau} \lambda_k (u_k - \bar{u}_k) + \sum_{k=0}^{N_\tau} |\bar{u}_k| \leq \sum_{k=0}^{N_\tau} |u_k| \quad \forall u_\tau \in \mathcal{U}_\tau.$$

The above relation is equivalent to

$$\lambda_k (u_k - \bar{u}_k) + |\bar{u}_k| \leq |u_k| \quad \forall 0 \leq k \leq N_\tau \text{ and } \forall u_\tau \in \mathcal{U}_\tau.$$

Obviously, the above inequalities are equivalent to (4.2). □

Using this proposition, the following theorem can be proved like Theorem 3.7.

THEOREM 4.4. *Let \bar{u}_τ be a local solution of (P_τ) . Then there exist $\bar{y}_\tau \in Y \cap C(\bar{Q} \cup \bar{Q}_-)$, $\bar{\varphi}_\tau \in H^1(Q) \cap C(\bar{Q})$, and $\bar{\lambda}_\tau \in \partial j_\tau(\bar{u}_\tau)$ such that*

$$(4.3) \quad \begin{cases} \frac{\partial \bar{y}_\tau}{\partial t} - \Delta \bar{y}_\tau + R(\bar{y}_\tau) = K[\bar{u}_\tau] \bar{y}_\tau + g_{\bar{u}_\tau} & \text{in } Q, \\ \partial_n \bar{y}_\tau = 0 & \text{on } \Sigma, \\ \bar{y}_\tau(x, t) = y_0(x, t) & \text{in } Q_-, \end{cases}$$

$$(4.4) \quad \begin{cases} -\frac{\partial \bar{\varphi}_\tau}{\partial t} - \Delta \bar{\varphi}_\tau + R'(\bar{y}_\tau) \bar{\varphi}_\tau = K^*[\bar{u}_\tau] \bar{\varphi}_\tau + \bar{y}_\tau - y_d & \text{in } Q, \\ \partial_n \bar{\varphi}_\tau = 0 & \text{on } \Sigma, \\ \bar{\varphi}_\tau(\cdot, T) = 0 & \text{in } \Omega, \end{cases}$$

$$(4.5) \quad \bar{\lambda}_k = -\frac{1}{\nu} \int_{-t_k}^{T-t_k} \int_{\Omega} \bar{y}_\tau(x, t) \bar{\varphi}_\tau(x, t + t_k) \, dx \, dt \quad \forall 0 \leq k \leq N_\tau.$$

Combining Proposition 4.3 and (4.5) we deduce the following corollary.

COROLLARY 4.5. *Let \bar{u}_τ be a local minimum of (P_τ) with $\bar{u}_\tau \neq 0$ and let $\bar{y}_\tau, \bar{\varphi}_\tau$, and $\bar{\lambda}_\tau$ satisfy (4.3)–(4.5), then*

$$(4.6) \quad \max_{0 \leq k \leq N_\tau} |\bar{\lambda}_k| = 1,$$

$$(4.7) \quad \begin{cases} \bar{u}_k > 0 \Rightarrow \bar{\lambda}_k = +1, \\ \bar{u}_k < 0 \Rightarrow \bar{\lambda}_k = -1. \end{cases}$$

Finally, we analyze the convergence of the above discretization.

THEOREM 4.6. *Let $\{\bar{u}_\tau\}_{\tau \searrow 0}$ be a sequence of discrete controls such that every control \bar{u}_τ is a solution of (P_τ) . This sequence is bounded in $\mathcal{M}[0, T]$. Any weak* limit point of a subsequence is a solution of (P) , and $J(\bar{u}_\tau) \rightarrow \inf (P)$ as $\tau \searrow 0$. In addition, if $\tau_j \rightarrow 0$ and $u_{\tau_j} \overset{*}{\rightharpoonup} \bar{u}$ in $\mathcal{M}[0, T]$, then $\lim_{j \rightarrow \infty} \|\bar{u}_{\tau_j}\|_{\mathcal{M}[0, T]} = \|\bar{u}\|_{\mathcal{M}[0, T]}$ and $\lim_{j \rightarrow \infty} \|\bar{y}_{\tau_j} - \bar{y}\|_Y = 0$, where \bar{y}_{τ_j} and \bar{y} denote the states associated with \bar{u}_{τ_j} and \bar{u} , respectively.*

Proof. The boundedness in $\mathcal{M}[0, T]$ is an immediate consequence of the inequalities $J(\bar{u}_\tau) \leq J(0) \, \forall \tau$. Assume that $u_{\tau_j} \overset{*}{\rightharpoonup} \bar{u}$ in $\mathcal{M}[0, T]$ as $j \rightarrow \infty$. From Lemma 3.2 we have $\bar{y}_{\tau_j} \rightarrow \bar{y}$ in Y . Let \tilde{u} be a solution of (P) . Then, we get, with Proposition 4.1 and Lemma 3.2,

$$J(\tilde{u}) \leq \liminf_{j \rightarrow \infty} J(\bar{u}_{\tau_j}) \leq \limsup_{j \rightarrow \infty} J(\bar{u}_{\tau_j}) \leq \limsup_{j \rightarrow \infty} J(\Lambda_{\tau_j} \tilde{u}) = J(\tilde{u}) = \inf (P).$$

From these inequalities we deduce that \bar{u} is a solution of (P) . Consequently, we have that $J(\bar{u}) = J(\tilde{u})$. Then, using again the above inequalities we get that $J(\bar{u}_{\tau_j}) \rightarrow J(\bar{u})$. This convergence along with $y_{\bar{u}_{\tau_j}} \rightarrow \bar{y}$ in $L^2(Q)$ implies that $\|\bar{u}_{\tau_j}\|_{\mathcal{M}[0, T]} \rightarrow \|\bar{u}\|_{\mathcal{M}[0, T]}$. \square

5. Numerical examples. Next we present some test examples to illustrate our results. To solve the problem, we have used a Tikhonov regularization of (P_τ) . For

$c > 0$, we consider the problem

$$(P_\tau^c) \quad \min_{u_\tau \in \mathcal{U}_\tau} J(u_\tau) + \frac{1}{2c} \sum_{k=0}^{N_\tau} u_k^2.$$

The first order optimality conditions for (P_τ^c) read as follows.

THEOREM 5.1. *Let u_τ^c be a local solution of (P_τ^c) . Then there exist $y_\tau^c \in Y \cap C(\bar{Q} \cup \bar{Q}_-)$, $\varphi_\tau^c \in H^1(Q) \cap C(\bar{Q})$, and $\lambda_\tau^c \in \partial j_\tau(u_\tau^c)$ such that*

$$(5.1) \quad \begin{cases} \frac{\partial y_\tau^c}{\partial t} - \Delta y_\tau^c + R(y_\tau^c) = K[u_\tau^c]y_\tau^c + g_{u_\tau^c} & \text{in } Q, \\ \partial_n y_\tau^c = 0 & \text{on } \Sigma, \\ y_\tau^c(x, t) = y_0(x, t) & \text{in } Q_-, \end{cases}$$

$$(5.2) \quad \begin{cases} -\frac{\partial \varphi_\tau^c}{\partial t} - \Delta \varphi_\tau^c + R'(y_\tau^c)\varphi_\tau^c = K^*[u_\tau^c]\varphi_\tau^c + y_\tau^c - y_d & \text{in } Q, \\ \partial_n \varphi_\tau^c = 0 & \text{on } \Sigma, \\ \varphi_\tau^c(\cdot, T) = 0 & \text{in } \Omega, \end{cases}$$

$$(5.3) \quad -\nu \lambda_k^c = \int_{-t_k}^{T-t_k} \int_\Omega y_\tau^c(x, t) \varphi_\tau^c(x, t + t_k) \, dx \, dt + \frac{1}{c} u_k^c \quad \forall 0 \leq k \leq N_\tau.$$

Taking into account Proposition 4.3, the condition on the subgradient $\lambda_\tau^c \in \partial j_\tau(u_\tau^c)$ can be written as

$$(5.4) \quad u_k^c = \max\{0, u_k^c + C(\lambda_k^c - 1)\} + \min\{0, u_k^c + C(\lambda_k^c + 1)\} \quad \forall C > 0.$$

We solve the system (5.1)–(5.4) with a semismooth Newton method. To that aim, we consider the function $G : \mathbb{R}^{2N_\tau+2} \rightarrow \mathbb{R}^{2N_\tau+2}$:

$$G(u, \lambda) = \begin{pmatrix} \left(\nu \lambda_k + \int_{-t_k}^{T-t_k} \int_\Omega y_{u_\tau}(x, t) \varphi_{u_\tau}(x, t + t_k) \, dx \, dt + \frac{1}{c} u_k \right)_{k=0}^{N_\tau} \\ \left(u_k - \max\{0, u_k - C(\lambda_k - 1)\} - \min\{0, u_k + C(\lambda_k + 1)\} \right)_{k=0}^{N_\tau} \end{pmatrix},$$

where $u = (u_k)_{k=0}^{N_\tau}$, $\lambda = (\lambda_k)_{k=0}^{N_\tau}$, and we denote y_{u_τ} and φ_{u_τ} the state and adjoint state, respectively, associated with the control $u_\tau = \sum_{k=0}^{N_\tau} u_k \delta_{t_k}$. Since this is a finite dimensional optimization problem, the $\max(0, \cdot)$ and $\min(0, \cdot)$ functions are semismooth—see, e.g., [16, Lemma 3.1]—and the method converges locally super-linearly to a critical point $(u_\tau^c, \lambda_\tau^c)$ such that $(y_{u_\tau^c}, \varphi_{u_\tau^c}, u_\tau^c, \lambda_\tau^c)$ is a solution of the optimality system (5.1)–(5.4), provided that the slant derivative of G has a uniformly bounded inverse in a neighborhood of $(u_\tau^c, \lambda_\tau^c)$; cf. [16, Theorem 1.1]. In practice, this means that we can solve the linear system arising at each iteration.

Let us comment on the practical details. To solve the delay linear parabolic partial differential equations that appear in the process we consider standard contin-

uous piecewise linear finite elements in space and a piecewise constant discontinuous Galerkin method in time, i.e., a dG(0)cG(1) discretization. The problem is solved first for some small value $c = c_0 > 0$ with initial guess $u \equiv 0$. Given the solution for some value $c = c_n, n \geq 0$, it is taken as the initial guess to solve the optimality system for $c = c_{n+1} > c_n$. The process stops when $\|u_\tau^{c_{n+1}} - u_\tau^{c_n}\|_{\mathcal{M}[0,T]} < \epsilon$ for some prescribed tolerance $\epsilon > 0$. We choose $C = c\nu$ at every iteration. The linear system arising at each iteration is reduced to a linear system for the active part of the control variable that is solved using GMRES.

In all our examples, the reaction term is of the form

$$R(y) = \rho(y - y_1)(y - y_2)(y - y_3).$$

This cubic polynomial is characteristic for certain bistable reactions and is often considered in the literature. For instance, it appears in the FitzHugh–Nagumo or Schlögl model of mathematical physics.

Example 5.1 (example with known critical point). To test the discretization and the optimization algorithm, we first construct an example with a known solution of the optimality system given in Theorem 3.7.

Consider $\Omega = (0, 1) \subset \mathbb{R}, T = 1, \rho = 1/3, -y_1 = y_3 = \sqrt{3}, y_2 = 0$. Define

$$\bar{u} = u^* \delta_{t^*},$$

where $t^* = 0.5$ and $u^* = -7.7$. With this control, we compute (an approximation of) its related state \bar{y} solving the state equation with initial data

$$y_0(x, t) = \frac{1+t}{5} \sin^2(\pi x).$$

For this example, we use a discretization of 257 evenly spaced nodes both in space and time to solve the parabolic partial differential equations.

Next we define

$$\bar{\varphi}(x, t) = \begin{cases} \cos^2(\frac{\pi}{2}t) & \text{if } 0 \leq t \leq T, \\ 0 & \text{if } T \leq t. \end{cases}$$

This function satisfies the boundary and final conditions of the adjoint state equation. Moreover, we define

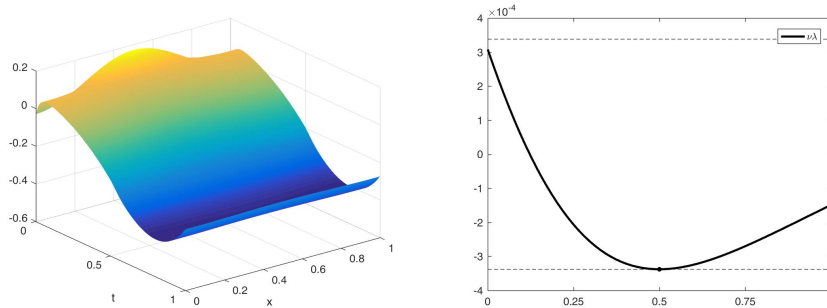
$$y_a(x, t) = \partial_t \bar{\varphi}(x, t) - R'(\bar{y})\bar{\varphi} + \bar{y} + \int_0^T \bar{\varphi}(x, t+s) d\bar{u}(s).$$

Taking into account that $-\Delta \bar{\varphi} = 0$, we have that $\bar{\varphi}$ satisfies the adjoint state equation; see Figure 5.1 for a picture of the computed target.

Finally, we compute $\Lambda(s)$, an approximation at the time nodes of $\nu \bar{\lambda}(s)$:

$$\Lambda(s) = - \int_0^T \int_\Omega \bar{y}(x, t-s) \bar{\varphi}(x, t) dx dt.$$

With our selections of t^*, u^* , and $\bar{\varphi}$, the computed function $s \mapsto \Lambda(s)$ is strictly convex

FIG. 5.1. Target (left) and $\nu\bar{\lambda}(t)$ for Example 5.1.TABLE 5.1
Results of Example 5.1 with known solution.

τ	\bar{u}_τ	$\ \bar{y} - \bar{y}_\tau\ _Y$	$\ \bar{u}_\tau\ _{\mathcal{M}[0,T]}$	$F(\bar{u}_\tau)$	$\nu j(\bar{u}_\tau)$
3^{-2}	$-3.33\delta_{t_\tau^-} - 4.50\delta_{t_\tau^+}$	$3.15e-5$	7.831	$9.132e-5$	$2.661e-3$
3^{-3}	$-3.30\delta_{t_\tau^-} - 4.43\delta_{t_\tau^+}$	$9.40e-7$	7.730	$9.107e-5$	$2.627e-3$
3^{-4}	$-2.55\delta_{t_\tau^-} - 5.17\delta_{t_\tau^+}$	$2.87e-8$	7.720	$9.067e-5$	$2.623e-3$
3^{-5}	$-0.11\delta_{t_\tau^-} - 7.61\delta_{t_\tau^+}$	$9.72e-9$	7.719	$9.066e-5$	$2.623e-3$
exact	$-7.7\delta_{t^*}$	0	7.7	$9.067e-5$	$2.617e-3$

in $[0, 1]$, has a minimum at t^* such that $\Lambda(t^*) \approx -3.39817 \times 10^{-4}$, and $|\Lambda(s)| < 3.39817 \times 10^{-4}$ for all $s \neq t^*$; see Figure 5.1, right. We should mention that the point $t^* = 0.5$ is a node of the grid underlying the piecewise linear approximation Λ .

If we define $\nu = 3.39817 \times 10^{-4}$, and $\bar{\lambda}(s) = \Lambda(s)/\nu$, we have that $|\bar{\lambda}(s)| < 1$ for all $s \neq t^*$ and $\bar{\lambda}(t^*) = -1$ and, therefore, (the approximation of) $(\bar{u}, \bar{y}, \bar{\varphi}, \bar{\lambda})$ satisfies the first order optimality conditions for problem (P) with data ν and y_d up to finite element precision.

Here and in what follows, the sign \approx means that the unknown exact values of the associated values are approximated by the given numbers. We expect but do not claim that these values are close to the exact ones.

The values of the differentiable and nondifferentiable parts of the functional are

$$F(\bar{u}) \approx 9.067 \times 10^{-5} \text{ and } \nu j(\bar{u}) \approx 2.617 \times 10^{-3}.$$

If we solve the problem with a discretization of the control space such that $\bar{u} \in \mathcal{U}_\tau$, we recover the original solution with four digits of accuracy. To set a more realistic scenario, we test our software in grids with constant time step $\tau = 3^{-k}$, $k = 2 : 5$, so that $\bar{u} \notin \mathcal{U}_\tau$.

The numerical results are displayed in Table 5.1. Notice that we are able to confirm all the results of Theorem 4.6. We write $t_\tau^- = t^* - \tau/2$ and $t_\tau^+ = t^* + \tau/2$ for the closest points to t^* in the control mesh. The values corresponding to the original solution \bar{u} are included in the last row of the table.

Example 5.2 (sensitivity to the regularization parameter ν). For the same problem as above, we illustrate how the solution changes as ν varies. It can be expected that the value of F decreases as ν decreases, and both $\|\bar{u}_\tau\|_{\mathcal{M}[0,T]}$ as well as the number of points in the support of \bar{u}_τ increase. As was proved in Proposition 3.9, there

TABLE 5.2

Example 5.2. Sensitivity of the norm and the support of the optimal control.

ν	$F_\tau(\bar{u}_\tau)$	$\ \bar{u}_\tau\ _{\mathcal{M}[0,T]}$	$\#\text{supp}\bar{u}_\tau$
$1e-1$	$8.85e-2$	0	0
$6e-2$	$8.85e-2$	0	0
$5e-2$	$8.74e-2$	0.02	1
$1e-2$	$1.95e-2$	3.54	1
$1e-3$	$3.79e-3$	7.36	2
$1e-4$	$5.97e-5$	7.95	2
$1e-5$	$4.31e-5$	8.83	4
$1e-6$	$3.38e-5$	13.3	8
$1e-7$	$2.12e-5$	59.4	31
$1e-8$	$1.85e-5$	170.0	56

is a $\bar{\nu} > 0$ such that the optimal control is zero for $\nu \geq \bar{\nu}$. In this example, we use a discretization of 65 equidistant nodes both in space and time. We use the same time grid for the control discretization. Our results are shown in Table 5.2.

Example 5.3 (recovering the solution of a system with nonlocal Pyragas feedback control). We consider the data of Example 1 in [25], namely, $\Omega = (-20, 20)$, $T = 40$, $\rho = 1$, $y_1 = 0$, $y_2 = 0.25$, $y_3 = 1$. The initial data are given by

$$y_0(x, t) = \frac{1}{2}(y_1 + y_3) + \frac{1}{2}(y_1 - y_3) \tanh\left(\frac{(y_3 - y_1)}{2\sqrt{2}}(x - ct)\right),$$

where

$$c = \frac{y_1 + y_3 - 2y_2}{\sqrt{2}}.$$

The desired state is the solution of the state equation with delay term given by the measure $u_d \in \mathcal{M}[0, T]$ defined as

$$\int_0^T y(x, t - s) du_d(s) = \kappa \left(\frac{1}{t_b - t_a} \int_{t_a}^{t_b} y(x, t - s) ds - y(x, t) \right)$$

with parameters $\kappa = 0.5$, $t_a = 0.456$, $t_b = 0.541$. This measure is a linear combination of a continuous measure with support in $[a, b]$ and a Dirac measure concentrated at $t = 0$.

We fix the parameter $\nu = 1e - 2$ —this is large enough to obtain a combination of Dirac measures—and we will look for solutions in $\mathcal{M}[0, 1]$. Since for the given delay term $\|u_d\|_{\mathcal{M}[0,1]} = 2\kappa = 1$ and $F(u_d) = 0$, it holds $J(u_d) = \nu$. Numerically, we obtain the solution

$$\bar{u} \approx -0.240821\delta_{s=0} + 0.246667\delta_{s=1}.$$

The associated values of the objective are $F(\bar{u}) \approx 5.3e - 5$ and $\nu j(\bar{u}) \approx 4.87e - 3$, hence $J(\bar{u}) \approx 0.49\nu$. The target and the optimal state are illustrated in Figure 5.2.

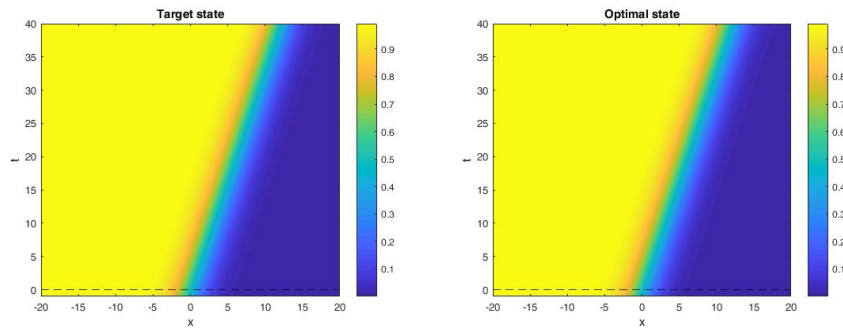


FIG. 5.2. Target (left) obtained with a nonlocal Pyragas feedback control and optimal state (right).

To solve the equations, we have used a space mesh with 513 evenly spaced nodes and a time grid with a variable step size and 778 nodes. For the discretization of the control, we selected a grid of 101 equidistant nodes in $[0, 1]$.

Example 5.4 (steering the system to an unstable equilibrium point). Here, our data are $\Omega = (0, 1)$, $T = 2$, $\rho = 1$, $y_1 = 0$, $y_2 = 0.25$, $y_3 = 1$. The initial data are given by $y_0(x, t) \equiv y_3$, which is a stable equilibrium point and the target is $y_d(x, t) \equiv y_2$, which is an unstable equilibrium point. Since the data do not depend on x and the boundary conditions are satisfied, the problem is equivalent to controlling a nonlinear delay ODE. We fix $\nu = 1e - 3$ and consider the tracking only on $[T/2, T]$. The reason is that the desired periodicity will need some time to be developed.

Therefore, here we redefine the differentiable part of the functional $J(u)$ by

$$(5.5) \quad F(u) = \frac{1}{2} \int_{T/2}^T \int_{\Omega} (y_u(x, t) - y_d(x, t))^2 dx dt.$$

This redefinition causes only minor modifications of our theory. The functional (5.5) is covered by multiplying the square $(y_u(x, t) - y_d(x, t))^2$ with the characteristic function of $(T/2, T) \times \Omega$ in the objective functional of (P). This causes only an obvious modification in the adjoint equation. However, to reduce the complexity of the notation, we avoided the general consideration of such modified functionals.

With $N_\tau = 512$ time steps, we obtain

$$\bar{u} \approx -1.304\delta_{0.418} + 0.134\delta_{1.977} + 0.220\delta_{1.978},$$

$$F(\bar{u}) \approx 1.29e - 4, \text{ and } J(\bar{u}) \approx 1.787e - 3.$$

Example 5.5 (changing the period of an incoming wave). We use the same data as in the previous example, but $y_0(x, t) = \cos^2(2\pi t)/2$ and $y_d(x, t) = \cos^2(\pi t)/2$. We fix $\nu = 1e - 3$ and take $F(u)$ as defined in (5.5). By a discretization with $N_\tau = 256$ time steps, we obtain

$$\bar{u} \approx 0.3188\delta_0 - 1.5499\delta_{0.4219} - 0.9964\delta_{0.8047} + 2.7233\delta_{1.5234}$$

and the objective values $F(\bar{u}) \approx 6.57e - 4$ and $J(\bar{u}) \approx 6.25e - 3$. The initial data, target, uncontrolled state, and the state associated with the computed optimal delay control are illustrated in Figure 5.3, e plot the functions for $x = 0$.

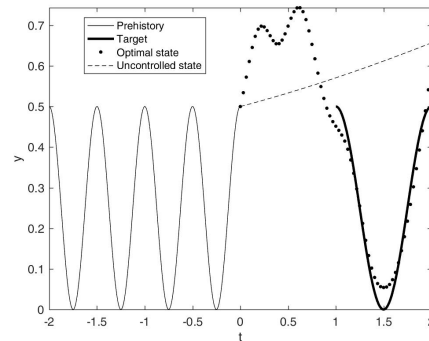


FIG. 5.3. Data and solution for Example 5.5.

REFERENCES

- [1] N. U. AHMED, *Existence of optimal controls for a general class of impulsive systems on Banach spaces*, SIAM J. Control Optim., 42 (2003), pp. 669–685.
- [2] H. T. BANKS AND J. A. BURNS, *Hereditary control problems: Numerical methods based on averaging approximations*, SIAM J. Control Optim., 16 (1978), pp. 169–208.
- [3] H. T. BANKS, J. A. BURNS, AND E. M. CLIFF, *Parameter estimation and identification for systems with delays*, SIAM J. Control Optim., 19 (1981), pp. 791–828.
- [4] H. T. BANKS AND A. MANITIUS, *Application of abstract variational theory to hereditary systems—a survey*, IEEE Trans. Automat. Control, AC-19 (1974), pp. 524–533.
- [5] D. D. BAĬ NOV AND D. P. MISHEV, *Oscillation Theory for Neutral Differential Equations with Delay*, Adam Hilger, Bristol, England, 1991.
- [6] R. BELLMAN, *A survey of the mathematical theory of time-lag, retarded control, and hereditary processes*, The Rand Corporation, Santa Monica, CA, 1954.
- [7] E. CASAS, *Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations*, SIAM J. Control Optim., 35 (1997), pp. 1297–1327.
- [8] E. CASAS, C. CLASON, AND K. KUNISCH, *Approximation of elliptic control problems in measure spaces with sparse solutions*, SIAM J. Control Optim., 50 (2012), pp. 1735–1752.
- [9] E. CASAS, F. KRUSE, AND K. KUNISCH, *Optimal control of semilinear parabolic equations by BV-functions*, SIAM J. Control Optim., 55 (2017), pp. 1752–1788.
- [10] E. CASAS AND K. KUNISCH, *Optimal control of semilinear elliptic equations in measure spaces*, SIAM J. Control Optim., 52 (2014), pp. 339–364.
- [11] E. CASAS, C. RYLL, AND F. TRÖLTZSCH, *Sparse optimal control of the Schlögl and FitzHugh-Nagumo systems*, Comput. Methods Appl. Math., 13 (2014), pp. 415–442.
- [12] F. COLONIUS AND D. HINRICHSSEN, *Optimal control of hereditary differential systems*, in Recent Theoretical Developments in Control, University of Leicester, Leicester, 1976, Academic Press, London, 1978, pp. 215–239.
- [13] T. ERNEUX, *Applied Delay Differential Equations*, Surv. Tutor. Appl. Math. Sci. 3, Springer, New York, 2009.
- [14] P. GRISVARD, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, 1985.
- [15] J. K. HALE AND S. M. VERDUYN LUNEL, *Introduction to functional-differential equations*, Appl. Math. Sci. 99, Springer, New York, 1993.
- [16] M. HINTERMÜLLER, K. ITO, AND K. KUNISCH, *The primal-dual active set strategy as a semi-smooth Newton method*, SIAM J. Optim., 13 (2002), pp. 865–888.
- [17] J.-M. JEONG AND H.-J. HWANG, *Optimal control problems for semilinear retarded functional differential equations*, J. Optim. Theory Appl., 167 (2015), pp. 49–67.
- [18] F. KAPPEL AND K. KUNISCH, *Spline approximations for neutral functional differential equations*, SIAM J. Numer. Anal., 18 (1981), pp. 1058–1080.
- [19] K. KUNISCH, *Approximation schemes for nonlinear neutral optimal control systems*, J. Math. Anal. Appl., 82 (1981), pp. 112–143.
- [20] Y. N. KYRYCHKO, K. B. BLYUSS, AND E. SCHÖLL, *Amplitude death in systems of coupled oscillators with distributed-delay coupling*, Eur. Phys. J. B, 84 (2011), pp. 307–315.

- [21] O. LADYZHENSKAYA, V. SOLONNIKOV, AND N. URAL'TSEVA, *Linear and Quasilinear Equations of Parabolic Type*, AMS, Providence, RI, 1968.
- [22] J. LÖBER, R. COLES, J. SIEBERT, H. ENGEL, AND E. SCHÖLL, *Control of a chemical wave propagation*, in World Scientific Lecture Notes in Complex Systems: Vol. 12: Engineering of Chemical Complexity II, Chap. 11, pp. 185–207, 2014.
- [23] B. S. MORDUKHOVICH, D. WANG, AND L. WANG, *Optimal control of delay-differential inclusions with functional endpoint constraints in infinite dimensions*, *Nonlinear Anal.*, 71 (2009), pp. e2740–e2749.
- [24] B. S. MORDUKHOVICH, D. WANG, AND L. WANG, *Optimization of delay-differential inclusions in infinite dimensions*, *Pacific J. Optim.*, 6 (2010), pp. 353–374.
- [25] P. NESTLER, E. SCHÖLL, AND F. TRÖLTZSCH, *Optimization of nonlocal time-delayed feedback controllers*, *Comput. Optim. Appl.*, 64 (2016), pp. 265–294.
- [26] K. PYRAGAS, *Continuous control of chaos by self-controlling feedback*, *Phys. Lett. A*, 170 (1992), pp. 421–428.
- [27] K. PYRAGAS, *Delayed feedback control of chaos*, *Philos. Trans. Roy. Soc. A*, 364 (2006), pp. 2309–2334.
- [28] J.-P. RICHARD, *Time-delay systems: An overview of some recent advances and open problems*, *Automatica J. IFAC*, 39 (2003), pp. 1667–1694.
- [29] W. RUDIN, *Real and Complex Analysis*, McGraw-Hill, London, 1970.
- [30] E. SCHÖLL AND H. SCHUSTER, *Handbook of Chaos Control*, Wiley-VCH, Weinheim, Germany, 2008.
- [31] R. E. SHOWALTER, *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*, *Math. Surveys Monogr.* 49, AMS, Providence, RI, 1997.
- [32] J. SIEBERT, S. ALONSO, M. BÄR, AND E. SCHÖLL, *Dynamics of reaction-diffusion patterns controlled by asymmetric nonlocal coupling as a limiting case of differential advection*, *Phys. Rev. E* (3), 89 (2014), 052909.
- [33] J. SIEBERT AND E. SCHÖLL, *Front and Turing patterns induced by Mexican-hat-like nonlocal feedback*, *EPL*, 109 (2015), 40014.
- [34] F. TRÖLTZSCH, *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*, *Grad. Stud. Math.* 112, AMS, Providence, RI, 2010.
- [35] F. UNGER AND L. VON WOLFERSDORF, *On a control problem for memory kernels in heat conduction*, *Z. Angew. Math. Mech.*, 75 (1995), pp. 365–370.
- [36] L. VON WOLFERSDORF, *On optimality conditions in some control problems for memory kernels in viscoelasticity*, *Z. Anal. Anwend.*, 12 (1993), pp. 745–750.