Measure Control of a Semilinear Parabolic Equation with a Nonlocal Time Delay *

Eduardo Casas[†]   Mariano Mateos[‡]   Fredi Tröltzsch[§]

September 13, 2018

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 non-differentiable 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.

Keywords: optimal control, parabolic equation, nonlocal time delay, measure control

AMS Subject classification: 49K20, 35K58, 49M25

1 Introduction

We consider optimal control problems for the parabolic equation

$$
\begin{cases}
\dfrac{\partial y}{\partial t} - \Delta y + R(y) & = \displaystyle\int_0^T y(x, t-s)\,\mathrm{d}u(s) & \text{in } Q = \Omega \times (0, T), \\[2mm]
\partial_n y & = 0 & \text{on } \Sigma = \Gamma \times (0, T), \\[1mm]
y(x,t) & = y_0(x,t) & \text{in } Q_- = \Omega \times [-T, 0],
\end{cases}
\tag{1.1}
$$

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

[†]Departmento 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.

1

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 [24]. 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. Main ideas for the associated analysis were prepared in [11] and we are able to proceed similarly, at 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 exemplarily the papers [2, 3, 12, 17, 18], that have some relation to distributed parameter systems, or the surveys [4, 27]. More recent contributions are e.g. [16, 22, 23].

However, to our best knowledge, the optimal control of parabolic equations with nonlocal time delay was only investigated in [24]. The case of measures as controls is new for this type of equations. 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 [34] and [35].

The equation above generalizes different models of Pyragas type feedback that are very popular in theoretical physics. We mention the seminal paper by Pyragas [25], where the feedback of the form (1.2) was introduced to stabilize periodic orbits; see also [26]. We also refer to [29, 31, 32], where nonlocal Pyragas feedback operators of the type (1.5) are discussed for different kernels u.

Let us also mention [19], where the implementation of nonlocal feedback controllers is investigated. In particular, these equations have applications in Laser technology; we refer to associated contributions in [29].

Let us also mention a few examples for the equation (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*

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

is obtained as particular case of (1.1) with $u = \kappa \left(\delta_\tau - \delta_0 \right)$. Equations of this type with fixed time delay τ are known in the context of the so-called Pyragas type feedback control, [25, 26, 29].

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

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

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

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

is obtained. Here, the control $(u_1, \ldots, u_m) \in \mathbb{R}^m$ 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*

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

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

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

for any real numbers $a \leq b$. Here, $|u|$ denotes the total variation measure of u; see [28, 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 the sequel. Thus, we distinguish

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

for real numbers $a \leq b$.

We recall our general state equation (1.1),

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y + R(y) &=& \displaystyle\int_0^T y(x, t - s) \, du(s) & \text{in } Q, \\[2mm] \partial_n y &=& 0 & \text{on } \Sigma, \\[2mm] y(x, t) &=& y_0(x, t) & \text{in } Q_-. \end{cases}$$

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 $R : \mathbb{R} \longrightarrow \mathbb{R}$ is a function of class C^1 such that

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

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 [21].

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, $i = 0, \ldots, k$, 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 $R : \Omega \times \mathbb{R} \to \mathbb{R}$, 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*

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

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

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

The theory remains also 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 $R : \mathbb{R} \to \mathbb{R}$ of class C^1 and satisfying condition (2.1).

In the sequel, 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 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 exemplarily to the expositions [7, 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*

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

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

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 $y^\lambda(x,t) = \mathrm{e}^{-\lambda t}y(x,t)$ with arbitrary $\lambda > 0$. Hence equation (1.1) is transformed to

$$
\begin{cases}
\dfrac{\partial y^\lambda}{\partial t} - \Delta y^\lambda + \mathrm{e}^{-\lambda t}R(\mathrm{e}^{\lambda t}y^\lambda) + \lambda y^\lambda &=\quad \displaystyle\int_0^T \mathrm{e}^{-\lambda s}y^\lambda(x,t-s)\,\mathrm{d}u(s) \quad &\text{in } Q, \\[2mm]
\partial_n y^\lambda &=\quad 0 &\text{on } \Sigma, \\[2mm]
y^\lambda(x,t) &=\quad \mathrm{e}^{-\lambda t}y_0(x,t) &\text{in } Q_-.
\end{cases}
$$

To simplify the notation we introduce, for every $\lambda \geq 0$, the following family of operators $K_\lambda[u], K_\lambda^+[u] :$ $C(\bar{Q}) \longrightarrow L^\infty(Q)$, defined for $(x,t) \in \bar{Q}$ by

$$
(K_\lambda[u]z)(x,t) = \int_{[0,t)} \mathrm{e}^{-\lambda s}z(x,t-s)\,\mathrm{d}u(s),
$$

$$
(K_\lambda^+[u]z)(x,t) = \int_{(0,t)} \mathrm{e}^{-\lambda s}z(x,t-s)\,\mathrm{d}u(s).
$$

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

Moreover, we define the family of functions

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

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

$$
\begin{cases}
\dfrac{\partial y^\lambda}{\partial t} - \Delta y^\lambda + \mathrm{e}^{-\lambda t}R(\mathrm{e}^{\lambda t}y^\lambda) + \lambda y^\lambda &=\quad K_\lambda[u]y^\lambda + g_{\lambda,u} \quad &\text{in } Q, \\[2mm]
\partial_n y^\lambda &=\quad 0 &\text{on } \Sigma, \\[2mm]
y^\lambda(x,0) &=\quad y_0(x,0) &\text{in } \Omega,
\end{cases} \tag{2.4}
$$

with the additional extension $y^\lambda(x,t) = y_0(x,t)$ 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

$$
\int_0^T e^{-\lambda s}y^\lambda(x,t-s)\,du(s) = \int_{[0,t)} e^{-\lambda s}y^\lambda(x,t-s)\,du(s) + \int_t^T e^{-\lambda s}y_0(x,t-s)\,du(s)
$$

$$
= (K_\lambda[u]y^\lambda + g_{\lambda,u})(x,t),
$$

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

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

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

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

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

$$\begin{cases} \|K_\lambda^+[u]z\|_{L^2(Q)} & \leq \varepsilon\big(1 + \|u\|_{\mathcal{M}[0,T]}\big)\|z\|_{L^2(Q)} \quad \forall z \in C(\bar{Q}), \\ \|K_\lambda^+[u]z\|_{L^\infty(Q)} & \leq \varepsilon\big(1 + \|u\|_{\mathcal{M}[0,T]}\big)\|z\|_{C(\bar{Q})} \quad \forall z \in C(\bar{Q}), \\ \|g_{\lambda,u}\|_{L^\infty(Q)} & \leq \varepsilon\big(1 + \|u\|_{\mathcal{M}[0,T]}\big)\|y_0\|_{C(\bar{Q}_-)}. \end{cases} \tag{2.7}$$

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

$$\|K_\lambda[u]z\|_{L^2(Q)}^2 = \int_Q \left(\int_{[0,t)} e^{-\lambda s} z(x, t - s) \, du(s) \right)^2 dx \, dt$$

$$\leq \int_Q \left(\int_{[0,t)} z^2(x, t - s) \, d|u|(s) \right) \left(\int_{[0,t)} e^{-2\lambda s} \, d|u|(s) \right) dx \, dt$$

$$\leq \left(\int_\Omega \int_0^T \int_s^T z^2(x, t - s) \, dt \, d|u|(s) \, dx \right) \left(\int_0^T e^{-2\lambda s} \, d|u|(s) \right) = I_1 I_2.$$

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

$$I_1 = \int_\Omega \int_0^T \int_0^{T-s} z^2(x, \sigma) \, d\sigma \, d|u|(s) \, dx \leq \int_\Omega \int_0^T \int_0^T z^2(x, \sigma) \, d\sigma \, d|u|(s) \, dx$$

$$= \|z\|_{L^2(Q)}^2 \|u\|_{\mathcal{M}[0,T]}.$$

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: $\forall t \in [0, T]$

$$\|(K_\lambda[u]z)(\cdot, t)\|_{L^2(\Omega)}^2 = \int_\Omega \left(\int_{[0,t)} e^{-\lambda s} z(x, t - s) \, du(s) \right)^2 dx \, dt$$

$$\leq \left(\int_\Omega \int_{[0,t)} z^2(x, t - s) \, d|u|(s) \, dx \right) \left(\int_0^T e^{-2\lambda s} \, d|u|(s) \right)$$

$$= \left(\int_{[0,t)} \|z(\cdot, t - s)\|_{L^2(\Omega)}^2 \, d|u|(s) \right) I_2 \leq \|z\|_{C([0,T], L^2(\Omega))}^2 \|u\|_{\mathcal{M}[0,T]}^2.$$

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

$$\|K_\lambda[u]z\|_{L^\infty(Q)} \leq \|z\|_{C(\bar{Q})} I_2 \leq \|z\|_{C(\bar{Q})} \|u\|_{\mathcal{M}[0,T]}.$$

In a completely analogous way we prove (2.6). To establish the first two inequalities of (2.7) we proceed exactly as above replacing I_2 by I_2^+

$$I_2^+ = \int_{(0,T]} e^{-2\lambda s} \, d|u|(s) \leq \int_{(0, \frac{1}{\sqrt{\lambda}})} e^{-2\lambda s} \, d|u|(s) + \int_{\frac{1}{\sqrt{\lambda}}}^T e^{-2\lambda s} \, d|u|(s)$$

$$\leq |u|\big(0, \frac{1}{\sqrt{\lambda}}\big) + e^{-2\sqrt{\lambda}} \|u\|_{\mathcal{M}[0,T]}. \tag{2.8}$$

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

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

holds to conclude the desired inequalities. Finally, we prove the last inequality of (2.7). To this end, we observe that

$$|g_{\lambda,u}(x,t)| \leq \int_{(0,T]} \mathrm{e}^{-\lambda s} |y_0(x,t-s)| \; \mathrm{d}|u|(s) \leq I_1^+ \|y_0\|_{C(\bar{Q}_-)} \quad \forall (x,t) \in Q,$$

where

$$I_1^+ = \int_{(0,T]} \mathrm{e}^{-\lambda s} \; \mathrm{d}|u|(s) \leq |u|\big(0, \frac{1}{\sqrt{\lambda}}\big) + \mathrm{e}^{-\sqrt{\lambda}} \|u\|_{\mathcal{M}[0,T]} \leq \varepsilon\big(1 + \|u\|_{\mathcal{M}[0,T]}\big).$$

Combining this estimate with (2.8) we deduce the desired inequality. □

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

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y + R_\lambda^+(t,y) & = & K_\lambda^+[u]z + g_{\lambda,u} & \text{in } Q, \\[2mm] \partial_n y & = & 0 & \text{on } \Sigma, \\[2mm] y(x,0) & = & y_0(x,0) & \text{in } \Omega, \end{cases} \tag{2.9}$$

where $R_\lambda^+(t,y) = \mathrm{e}^{-\lambda t} R(\mathrm{e}^{\lambda t} y) + (\lambda - u(\{0\}))y$. 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

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

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 [8] or [33], 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 equality $R_\lambda^+(t,0) = R(0)$ we know the estimate

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

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

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

Let B_M be the closed ball of $C(\bar{Q})$ with center at 0 and radius M. We define the continuous mapping $F : B_M \longrightarrow B_M$ that associates to 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 (2.11), (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

$$\|y\|_{C^{0,\mu_0}(\bar{Q})} \le C_\mu \Big(\|K_\lambda^+[u]z + g_{\lambda,u}\|_{L^\infty(Q)} + \|y_0(\cdot, 0)\|_{C^{0,\mu}(\bar{\Omega})} + \max_{|\rho| \le M} |R(\rho)| + 1 \Big);$$

see [20, §III-10]. 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 $y_{0k} \to y_0(\cdot, 0)$ in $C(\bar{\Omega})$. Hence, for every $k \ge 1$, (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

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

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

$$\frac{1}{2}\|(y_k^\lambda - y_m^\lambda)(T)\|_{L^2(\Omega)} + \int_Q |\nabla(y_k^\lambda - y_m^\lambda)|^2 \, \mathrm{d}x \, \mathrm{d}t + \int_Q |y_k^\lambda - y_m^\lambda|^2 \, \mathrm{d}x \, \mathrm{d}t$$

$$\le \varepsilon\big(1 + \|u\|_{\mathcal{M}[0,T]}\big)\|y_k^\lambda - y_m^\lambda\|_{L^2(Q)}^2 + \frac{1}{2}\|y_{0k} - y_{0m}\|_{L^2(\Omega)}^2$$

$$\le C\varepsilon\big(1 + \|u\|_{\mathcal{M}[0,T]}\big)\|y_k^\lambda - y_m^\lambda\|_{L^2(0,T;H^1(\Omega))}^2 + \frac{1}{2}\|y_{0k} - y_{0m}\|_{L^2(\Omega)}^2.$$

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

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

Notice that the left hand side of the above chain of inequalities absorbs the term appearing with the factor $C\varepsilon\big(1 + \|u\|_{\mathcal{M}[0,T]}\big)$ 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

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

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

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

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 $y_k^\lambda \to y^\lambda$ in Y. It is easy to see that y^λ is a solution of (2.4). Now, we re-substitute $y_u(x, t) = \mathrm{e}^{\lambda t} y^\lambda(x, t)$ 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

$$
\begin{cases}
\dfrac{\partial y^\lambda}{\partial t} - \Delta y^\lambda + \dfrac{\partial R_\lambda^+}{\partial y}(t, \hat{y}^\lambda)y^\lambda & = & K_\lambda^+[u]y^\lambda & \text{in } Q, \\[2mm]
\partial_n y^\lambda & = & 0 & \text{on } \Sigma, \\[2mm]
y^\lambda(x, 0) & = & 0 & \text{in } \Omega,
\end{cases}
$$

where $\hat{y}^\lambda = y_1^\lambda + \theta(y_2^\lambda - y_1^\lambda)$ is some intermediate state with $0 \le \theta(x,t) \le 1$. Multiplying this equation by y^λ and invoking again (2.10) along with (2.5), we obtain

$$
\frac{1}{2}\|y^\lambda(T)\|_{L^2(\Omega)}^2 + \int_Q |\nabla y^\lambda|^2 \, \mathrm{d}x \, \mathrm{d}t + \int_Q |y^\lambda|^2 \, \mathrm{d}x \, \mathrm{d}t
$$
$$
\le \varepsilon\big(1 + \|u\|_{\mathcal{M}[0,T]}\big)\|y^\lambda\|_{L^2(Q)}^2.
$$

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

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

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

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

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

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

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

Then, multiplying equation (2.4) by y^λ and using this inequality along with (2.5) and (2.6), we obtain for every $0 < T' < T$ and $Q_{T'} = \Omega \times (0, T')$

$$
\frac{1}{2}\|y^\lambda(T')\|_{L^2(\Omega)}^2 + \int_{Q_{T'}} |\nabla y^\lambda|^2 \, \mathrm{d}x \, \mathrm{d}t + 2(1 + \|u\|_{\mathcal{M}[0,T]})\int_{Q_{T'}} |y^\lambda|^2 \, \mathrm{d}x \, \mathrm{d}t
$$
$$
\le \int_{Q_{T'}} (K_\lambda[u]y^\lambda + g_{\lambda,u})y^\lambda \, \mathrm{d}x \, \mathrm{d}t + |R(0)|\int_{Q_{T'}} |y^\lambda| \, \mathrm{d}x \, \mathrm{d}t + \frac{1}{2}\|y_0(\cdot, 0)\|_{L^2(\Omega)}^2
$$
$$
\le \|u\|_{\mathcal{M}[0,T]}\big(\|y^\lambda\|_{L^2(Q_{T'})} + \|y_0\|_{L^2(Q_-)}\big)\|y^\lambda\|_{L^2(Q_{T'})}
$$
$$
+ |R(0)||Q|^{\frac{1}{2}}\|y^\lambda\|_{L^2(Q_{T'})} + \frac{1}{2}\|y_0(\cdot, 0)\|_{L^2(\Omega)}^2
$$
$$
\le \big(\frac{1}{2} + \|u\|_{\mathcal{M}[0,T]}\big)\|y^\lambda\|_{L^2(Q_{T'})}^2 + \frac{1}{2}\|u\|_{\mathcal{M}[0,T]}^2\|y_0\|_{L^2(Q_-)}^2
$$
$$
+ \frac{|Q|}{2}|R(0)|^2 + \frac{1}{2}\|y^\lambda\|_{L^2(Q_{T'})}^2 + \frac{1}{2}\|y_0(\cdot, 0)\|_{L^2(\Omega)}^2
$$
$$
= \big(1 + \|u\|_{\mathcal{M}[0,T]}\big)\|y^\lambda\|_{L^2(Q_{T'})}^2 + \frac{1}{2}\|u\|_{\mathcal{M}[0,T]}^2\|y_0\|_{L^2(Q_-)}^2
$$
$$
+ \frac{|Q|}{2}|R(0)|^2 + \frac{1}{2}\|y_0(\cdot, 0)\|_{L^2(\Omega)}^2.
$$

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

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

To prove (2.3) we use the second inequality of (2.5), (2.6) and the results of [20, §III-7] applied to (2.4) to obtain

$$\|y^\lambda\|_{C(\bar{Q})} \leq C\Big(\|K^+_\lambda[u]y^\lambda\|_{L^\infty((0,T),L^2(\Omega))} + \|g_{\lambda,u}\|_{L^\infty(Q)} + \|y_0(\cdot,0)\|_{C(\bar\Omega)} + |R(0)|\Big)$$
$$\leq C\Big(\|u\|_{\mathcal{M}[0,T]}\|y^\lambda\|_{C([0,T],L^2(\Omega))} + \|u\|_{\mathcal{M}[0,T]}\|y_0\|_{C(Q_-)} + \|y_0(\cdot,0)\|_{C(\bar\Omega)} + |R(0)|\Big).$$

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).

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*

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

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

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

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, [30, §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 [30, §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 $G : \mathcal{M}[0,T] \longrightarrow Y$ that associates to $u \in \mathcal{M}[0,T]$ 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*

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

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

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

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

$$\mathcal{F} : \mathcal{Y} \times \mathcal{M}[0,T] \longrightarrow L^\infty(Q) \times C(\bar{\Omega}),$$

$$\mathcal{F}(y,u) = \left(\frac{\partial y}{\partial t} - \Delta y + R(y) - K[u]y - g_u, y(\cdot,0) - y_0(\cdot,0) \right).$$

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 $\frac{\partial \mathcal{F}}{\partial y}(y,u) : \mathcal{Y} \longrightarrow L^\infty(Q) \times C(\bar{\Omega})$ 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

$$\begin{cases} \dfrac{\partial z}{\partial t} - \Delta z + R'(y)z & = & K[u]z + f & \text{in } Q, \\[2mm] \partial_n z & = & 0 & \text{on } \Sigma, \\[2mm] z(x,0) & = & z_0 & \text{in } \Omega. \end{cases}$$

The existence and uniqueness of such a solution is proved in the same way as for the problem (1.1). Hence, an application of the implicit function theorem implies that G is of class C^1. The equation (2.16) follows easily by differentiating the identity $\mathcal{F}(G(u),u) = 0$ with respect to u. $\qquad\square$

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 equation (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 prerequisites to study our optimal control problem, namely

$$\text{(P)} \min_{u \in \mathcal{M}[0,T]} J(u) = \frac{1}{2} \int_Q |y_u - y_d|^2 \, \mathrm{d}x \, \mathrm{d}t + \nu \|u\|_{\mathcal{M}[0,T]},$$

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 $u_k \overset{*}{\rightharpoonup} \bar{u}$ in $\mathcal{M}[0,T]$ for $k \to \infty$ and let y_k and $\bar{y}$ be the states associated with u_k and $\bar{u}$, respectively; then $y_k \to \bar{y}$ in Y.*

Proof. Since $u_k \overset{*}{\rightharpoonup} \bar{u}$ in $\mathcal{M}[0,T]$ as $k \to \infty$, 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 $y_k^\lambda(x,t) = \mathrm{e}^{-\lambda t} y_k(x,t)$ and $\bar{y}^\lambda(x,t) = \mathrm{e}^{-\lambda t}\bar{y}(x,t)$. Then y_k^λ and $\bar{y}^\lambda$ 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 $R_\lambda(t,y) = \mathrm{e}^{-\lambda t} R(\mathrm{e}^{\lambda t} y) + \lambda y$ with $\lambda \geq 2(M+1) - \min\{0, C_R\}$, we get

$$
\begin{cases}
\dfrac{\partial w_k^\lambda}{\partial t} - \Delta w_k^\lambda + \dfrac{\partial R_\lambda}{\partial y}(t, \hat{y}_k^\lambda) w_k^\lambda = K_\lambda[u_k]y_k^\lambda - K_\lambda[\bar{u}]\bar{y}^\lambda + g_{\lambda, u_k} - g_{\lambda, \bar{u}} & \text{in } Q, \\[2mm]
\partial_n w_k^\lambda = 0 & \text{on } \Sigma, \\[2mm]
w_k^\lambda(x,0) = 0 & \text{in } \Omega.
\end{cases}
\tag{3.1}
$$

with intermediate states $\hat{y}_k^\lambda$.

Testing this equation by w_k^λ and invoking (2.7), we get for every $0 < T' < T$

$$
\frac{1}{2}\|w_k^\lambda(T')\|_{L^2(\Omega)}^2 + \int_{Q_{T'}} \left[|\nabla w_k^\lambda|^2 + 2(M+1)|w_k^\lambda|^2\right] \,\mathrm{d}x\,\mathrm{d}t
$$

$$
\leq \int_{Q_{T'}} \left[K_\lambda[u_k]y_k^\lambda - K_\lambda[\bar{u}]\bar{y} + g_{\lambda, u_k} - g_{\lambda, \bar{u}}\right] w_k^\lambda \,\mathrm{d}x\,\mathrm{d}t
$$

$$
= \int_{Q_{T'}} \left[K_\lambda[u_k]w_k^\lambda + K_\lambda[u_k - \bar{u}]\bar{y} + g_{\lambda, u_k - \bar{u}}\right] w_k^\lambda \,\mathrm{d}x\,\mathrm{d}t
$$

$$
\leq \|u_k\|_{\mathcal{M}[0,T]}\|w_k^\lambda\|_{L^2(Q_{T'})}^2 + \|K_\lambda[u_k - \bar{u}]\bar{y} + g_{\lambda, u_k - \bar{u}}\|_{L^2(Q)}\|w_k^\lambda\|_{L^2(Q_{T'})}
$$

$$
\leq (M+1)\|w_k^\lambda\|_{L^2(Q_{T'})} + \frac{1}{2}\|K_\lambda[u_k - \bar{u}]\bar{y} + g_{\lambda, u_k - \bar{u}}\|_{L^2(Q)}^2.
$$

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

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

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

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

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

$$
\|K_\lambda[u_k - \bar{u}]\bar{y} + g_{\lambda, u_k - \bar{u}}\|_{L^\infty(Q)} \leq \left(\|\bar{y}\|_{C(\bar{Q})} + \|y_0\|_{C(\bar{Q}_-)}\right)\|u_k - \bar{u}\|_{\mathcal{M}[0,T]}
$$

$$
\leq 2M\left(\|\bar{y}\|_{C(\bar{Q})} + \|y_0\|_{C(\bar{Q}_-)}\right) \quad \forall k.
$$

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

Let us show show the uniform convergence. From equation (3.1), using the estimates of [20, §III-8] and (2.5) we infer for $p > 1 + \frac{n}{2}$

$$
\|w_k^\lambda\|_{C(\bar{Q})} \leq C_1\|K_\lambda[u_k]w_k^\lambda\|_{L^\infty((0,T),L^2(\Omega))} + C_2\|K_\lambda[u_k - \bar{u}]\bar{y} + g_{\lambda, u_k - \bar{u}}\|_{L^p(\Omega))}
$$

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

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

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

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

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 $u_k \stackrel{*}{\rightharpoonup} \bar{u}$ 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 $y_k \to \bar{y}$ in $L^2(Q)$. This convergence, along with (2.3), implies that $J(\bar{u}) \le \liminf_{k\to\infty} J(u_k) = \inf(\text{P})$, and hence $\bar{u}$ is a solution of (P).

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) *in the sense of $\mathcal{M}[0,T]$ (respectively $H^1(0,T)^*$) if there exists a ball $B_\varepsilon(\bar{u})$ in the associated space such that $J(\bar{u}) \le J(u) \ \forall u \in \mathcal{M}[0,T] \cap B_\varepsilon(\bar{u})$. We will say that $\bar{u}$ is a local solution if it is a local solution in some of the two notions defined above.*

Due to the continuity of the above embeddings, it follows immediately that, if $\bar{u}$ is a local solution in the $H^1(0,T)^*$ sense, then it is also a local solution in the $\mathcal{M}[0,T]$ sense. The converse implication is not true, in general.

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

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

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

$$\begin{cases} -\dfrac{\partial\varphi}{\partial t} - \Delta\varphi + R'(y_u)\varphi &= K^*[u]\varphi + y_u - y_d & \text{in } Q, \\[2mm] \partial_n\varphi &= 0 & \text{on } \Sigma \\[2mm] \varphi(x,T) &= 0 & \text{in } \Omega, \end{cases} \tag{3.4}$$

and the operator K^* is defined by

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

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*

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

$$\|\varphi\|_{C(\bar{Q})} \le M_\infty\|y_u - y_d\|_{L^{\bar{p}}(Q)}. \tag{3.7}$$

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

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

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 $(K_\lambda[u]\psi)(x,t) = (e^{-\lambda t}K^*[u]\varphi)(x, T - t)$, and (3.4) is transformed to the forward equation

$$\begin{cases} \dfrac{\partial \psi^\lambda}{\partial t} - \Delta \psi^\lambda + R'(\hat{y}_u)\psi^\lambda + \lambda \psi^\lambda &=& K_\lambda[u]\psi + f & \text{in } Q, \\[2mm] \partial_n \psi^\lambda &=& 0 & \text{on } \Sigma \\[2mm] \psi^\lambda(x, 0) &=& 0 & \text{in } \Omega, \end{cases} \tag{3.9}$$

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 [20, §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

$$\int_Q (K^*[u]w)(x,t)z(x,t)\,\mathrm{d}x\,\mathrm{d}t = \int_Q \left(\int_{[0,T-t)} w(x, t+s)\,\mathrm{d}u(s) \right) z(x,t)\,\mathrm{d}x\,\mathrm{d}t$$

$$= \int_\Omega \int_{[0,T)} \left(\int_0^{T-s} w(x, t+s)z(x,t)\,\mathrm{d}t \right) \mathrm{d}u(s)\,\mathrm{d}x$$

$$= \int_\Omega \int_{[0,T)} \left(\int_s^T w(x, \tau)z(x, \tau - s)\,\mathrm{d}\tau \right) \mathrm{d}u(s)\,\mathrm{d}x$$

$$= \int_\Omega \int_0^T \left(\int_{[0,\tau)} z(x, \tau - s)\,\mathrm{d}u(s) \right) w(x, \tau)\,\mathrm{d}\tau\,\mathrm{d}x$$

$$= \int_Q (K[u]z)(x, \tau)w(x, \tau)\,\mathrm{d}x\,\mathrm{d}\tau. \tag{3.10}$$

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 equation (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

$$F'(u)v = \int_Q (y_u - y_d)z_v\,\mathrm{d}x\,\mathrm{d}t$$

$$= \int_Q \left[-\frac{\partial \varphi_u}{\partial t}z_v + \nabla\varphi_u\nabla z_v + R'(y_u)\varphi_u z_v - (K^*[u]\varphi_u)z_v \right] \mathrm{d}x\,\mathrm{d}t$$

$$= \int_Q \left[\frac{\partial z_v}{\partial t}\varphi_u + \nabla\varphi_u\nabla z_v + R'(y_u)z_v\varphi_u - (K[u]z_v)\varphi_u \right] \mathrm{d}x\,\mathrm{d}t$$

$$= \int_Q \left(K[v]y_u + g_v \right)\varphi_u\,\mathrm{d}x\,\mathrm{d}t,$$

which proves (3.3).

We continue by studying the function $j : \mathcal{M}[0,T] \longrightarrow \mathbb{R}$. Since j is Lipschitz and convex, we know that it has a nonempty subdifferential and possesses directional derivatives at every point $u \in \mathcal{M}[0,T]$ and in any direction $v \in \mathcal{M}[0,T]$. They will be denoted by $\partial j(u)$ and $j'(u;v)$, respectively.

Let us recall some properties of $\partial j(u)$ and $j'(u;v)$; 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*

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

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

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

Proof. By definition of the subdifferential, we have

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

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

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

By the established properties, we find

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

therefore

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

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

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

Hence, we get

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

which proves (3.12). $\qquad\square$

Now we study the directional derivatives of j. Following [10], we introduce another notation. Given $u, v \in \mathcal{M}[0, T]$, we consider the Lebesgue decomposition of v with respect to $|u|$: $v = v_a + v_s$, where v_a is the absolutely continuous part of v with respect to $|u|$ and v_s is the singular part; see, for instance, [28, Chapter 6]. We denote by h_v the Radon-Nikodym derivative of v_a with respect to $|u|$, i.e. $dv_a = h_v d|u|$. Then we have

$$\|v\|_{\mathcal{M}[0,T]} = \|v_a\|_{\mathcal{M}[0,T]} + \|v_s\|_{\mathcal{M}[0,T]} = \int_0^T |h_v(s)| \, \mathrm{d}|u|(s) + \|v_s\|_{\mathcal{M}[0,T]}. \qquad (3.15)$$

Moreover, it is obvious that u is absolutely continuous with respect to $|u|$. We have $du = h \, d|u|$, $du^+ = h^+ \, d|u|$, and $du^- = h^- \, d|u|$, where $|h(s)| = 1$ for every $s \in [0, T]$.

In the next statement, we derive the expression for the directional derivatives of j.

Proposition 3.7. *For every $u, v \in \mathcal{M}[0, T]$, we have*

$$j'(u; v) = \int_0^T h_v(s) \, \mathrm{d}u(s) + \int_0^T \mathrm{d}|v_s|(s). \qquad (3.16)$$

We refer to [10, Proposition 3.3] for the proof.

Theorem 3.8. *Let $\bar{u}$ be a local solution of* (P). *Then there exist $\bar{y} \in Y \cap C(\bar{Q} \cup \bar{Q}_-)$, $\bar{\varphi} \in H^1(Q) \cap C(\bar{Q} \cup \bar{Q}_+)$ and $\bar{\lambda} \in C[0, T] \cap \partial j(\bar{u})$ such that*

$$\begin{cases} \dfrac{\partial \bar{y}}{\partial t} - \Delta \bar{y} + R(\bar{y}) & = & K[\bar{u}]\bar{y} + g_{\bar{u}} & in \ Q, \\[2mm] \partial_n \bar{y} & = & 0 & on \ \Sigma, \\[2mm] \bar{y}(x, t) & = & y_0(x, t) & in \ Q_-, \end{cases} \qquad (3.17)$$

$$\begin{cases} -\dfrac{\partial \bar{\varphi}}{\partial t} - \Delta \bar{\varphi} + R'(\bar{y})\bar{\varphi} & = & K^*[\bar{u}]\bar{\varphi} + \bar{y} - y_d & in \ Q, \\[2mm] \partial_n \bar{\varphi} & = & 0 & on \ \Sigma, \\[2mm] \bar{\varphi} & = & 0 & in \ Q_+, \end{cases} \qquad (3.18)$$

$$\bar{\lambda}(s) = -\frac{1}{\nu} \int_{-s}^T \int_\Omega \bar{y}(x, t)\bar{\varphi}(x, t + s) \, \mathrm{d}x \, \mathrm{d}t \quad \forall s \in [0, T], \qquad (3.19)$$

where $Q_+ = \Omega \times [T, 2T]$.

Proof. The existence and uniqueness of solutions to (3.17) and (3.18) have already been discussed in Theorem 2.2 and Proposition 3.5. Notice that the condition $\bar{\varphi}(x, T) = 0$ in Ω has been extended to $\bar{\varphi}(x, t) = 0$ in $Q_+ = \Omega \times [T, 2T]$. It is obvious that this extension by 0 defines a continuous function in $\bar{Q} \cup \bar{Q}_+$. Now, we define $\bar{\lambda}$ by (3.19). 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

$$0 \le \lim_{\rho \searrow 0} \frac{J(\bar{u} + \rho(u - \bar{u})) - J(\bar{u})}{\rho} \le F'(\bar{u})(u - \bar{u}) + \nu j(u) - \nu j(\bar{u})$$

$$= \int_Q \bar{\varphi}(K[u - \bar{u}]\bar{y} + g_{u - \bar{u}}) \, \mathrm{d}x \, \mathrm{d}t + \nu j(u) - \nu j(\bar{u}). \qquad (3.20)$$

$$\int_Q \bar{\varphi}\left(K[u-\bar{u}]\bar{y} + g_{u-\bar{u}}\right) \mathrm{d}x\, \mathrm{d}t$$

$$= \int_\Omega \int_0^T \bar{\varphi}(x,t)\left(\int_0^T \bar{y}(x,t-s)\,\mathrm{d}(u-\bar{u})(s)\right) \mathrm{d}t\, \mathrm{d}x$$

$$= \int_0^T \int_\Omega \int_0^T \bar{y}(x,t-s)\bar{\varphi}(x,t)\,\mathrm{d}t\, \mathrm{d}x\, \mathrm{d}(u-\bar{u})(s)$$

$$= \int_0^T \int_\Omega \int_{-s}^{T-s} \bar{y}(x,\tau)\bar{\varphi}(x,\tau+s)\,\mathrm{d}\tau\, \mathrm{d}x\, \mathrm{d}(u-\bar{u})(s)$$

$$= \int_0^T \left(\int_{-s}^T \int_\Omega \bar{y}(x,\tau)\bar{\varphi}(x,\tau+s)\,\mathrm{d}x\, \mathrm{d}\tau\right) \mathrm{d}(u-\bar{u})(s)$$

$$= -\nu \int_0^T \bar{\lambda}(s)\,\mathrm{d}(u-\bar{u})(s).$$

Combining this with (3.20), we find

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

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

From Proposition 3.6 and Theorem 3.8 we deduce the following sparsity structure of the optimal control $\bar{u}$.

Corollary 3.9. *Let $\bar{u}$ be a local minimum of (P) and let $\bar{y}$, $\bar{\varphi}$ and $\bar{\lambda}$ satisfy the optimality system (3.17)-(3.19), then if $\bar{u} \not\equiv 0$*

$$\|\bar{\lambda}\|_{C[0,T]} = 1, \tag{3.21}$$

$$\begin{aligned} supp(\bar{u}^+) &\subset \{t \in [0,T] : \bar{\lambda}(t) = +1\}, \\ supp(\bar{u}^-) &\subset \{t \in [0,T] : \bar{\lambda}(t) = -1\}, \end{aligned} \tag{3.22}$$

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

Proposition 3.10. *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

$$\|\bar{y}\|_{L^2(Q)} \leq C_1 < \infty \quad \text{and} \quad \|\bar{u}\|_{\mathcal{M}[0,T]} \leq \frac{C_2}{\nu} \tag{3.23}$$

for some constants independent of ν. Arguing similarly as in the proof of inequality (2.13), we get from equation (3.18) that

$$\|\bar{\varphi}\|_{C([0,T],L^2(\Omega))} \leq C_3\|\bar{y} - y_d\|_{L^2(Q)} \leq C_4. \tag{3.24}$$

According to (3.23), C_3 and C_4 are independent of $\nu \geq 1$. Now, from (3.19), (3.23) and (3.24) we

get

$$|\bar{\lambda}(s)| \leq \frac{1}{\nu} \int_{-s}^{T} \|\bar{y}(t)\|_{L^2(\Omega)} \|\bar{\varphi}(t+s)\|_{L^2(\Omega)} \, \mathrm{d}t$$

$$\leq \frac{1}{\nu} \Big(\int_{-T}^{T} \|\bar{y}(t)\|_{L^2(\Omega)} \, \mathrm{d}t \Big) \|\bar{\varphi}\|_{C([0,T],L^2(\Omega))}$$

$$= \frac{1}{\nu} \Big(\int_{-T}^{0} \|y_0(t)\|_{L^2(\Omega)} \, \mathrm{d}t + \int_{0}^{T} \|\bar{y}(t)\|_{L^2(\Omega)} \, \mathrm{d}t \Big) \|\bar{\varphi}\|_{C([0,T],L^2(\Omega))}$$

$$\leq \frac{\sqrt{T}}{\nu} \big(\|y_0\|_{L^2(Q_-)} + \|\bar{y}\|_{L^2(Q)} \big) \|\bar{\varphi}\|_{C([0,T],L^2(\Omega))} \leq \frac{\sqrt{T}}{\nu} \big(\|y_0\|_{L^2(Q_-)} + C_1 \big) C_4.$$

If we take $\bar{\nu} > \max \big\{ 1, \sqrt{T} \big(\|y_0\|_{L^2(Q_-)} + C_1 \big) C_4 \big\}$ we infer that $|\bar{\lambda}(s)| < 1 \ \forall s \in [0,T]$. Then, (3.22) implies that $\bar{u} \equiv 0$. □

4 Discretization of the Control Space

In this section we are going to consider the approximation of $\mathcal{M}[0,T]$ by finite dimensional subspaces $\mathcal{U}_\tau$. Associated to each space $\mathcal{U}_\tau$ we define a new problem (P_τ). Then, we analyze the convergence of the solutions of (P_τ). First we consider a grid of points $0 = t_0 < t_1 < \ldots < t_{N_\tau} = T$. 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

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

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

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

The following proposition states some properties of this mapping.

Proposition 4.1. *The following statements hold*

1. *$\|\Lambda_\tau u\|_{\mathcal{M}[0,T]} \leq \|u\|_{\mathcal{M}[0,T]} \ \forall u \in \mathcal{M}[0,T]$.*

2. *$\Lambda_\tau u \overset{*}{\rightharpoonup} u$ in $\mathcal{M}[0,T] \ \forall u \in \mathcal{M}[0,T]$.*

3. *$\lim_{\tau \to 0} \|\Lambda_\tau u\|_{\mathcal{M}[0,T]} = \|u\|_{\mathcal{M}[0,T]} \ \forall u \in \mathcal{M}[0,T]$.*

Proof. 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

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

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

$$|\langle u - \Lambda_\tau u, y \rangle| = \Big| \int_0^T y(s)\, \mathrm{d}u(s) - \sum_{k=0}^{N_\tau} y(t_k) u(I_k) \Big| = \Big| \sum_{k=0}^{N_\tau} \int_{I_k} [y(s) - y(t_k)]\, \mathrm{d}u(s) \Big|$$

$$\leq \sum_{k=0}^{N_\tau} \int_{I_k} |y(s) - y(t_k)|\, \mathrm{d}|u|(s) \leq \varepsilon \|u\|_{\mathcal{M}[0,T]}.$$

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

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

which concludes the proof. $\qquad\square$

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

$$(\mathrm{P}_\tau) \min_{u_\tau \in \mathcal{U}_\tau} J(u_\tau) = \frac{1}{2} \int_Q |y_{u_\tau} - y_d|^2\, \mathrm{d}x\, \mathrm{d}t + \nu \sum_{k=0}^{N_\tau} |u_k|,$$

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

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

We identify the dual of $\mathcal{U}_\tau$ 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 4.2. *With the above notation, we have $\lambda_\tau \in \partial j_\tau(\bar{u}_\tau)$ if and only if the following identity holds*

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

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. $\qquad\square$

Using this proposition, the following theorem can be proved as Theorem 3.8.

Theorem 4.3. *Let $\bar{u}_\tau$ be a local solution of* (P_τ)*. Then there exist $\bar{y}_\tau \in Y \cap C(\bar{Q} \cup \bar{Q}_-)$, $\bar{\varphi}_\tau \in H^1(Q) \cap C(\bar{Q} \cup \bar{Q}_+)$ and $\bar{\lambda}_\tau \in \partial j_\tau(\bar{u}_\tau)$ such that*

$$\begin{cases} \dfrac{\partial \bar{y}_\tau}{\partial t} - \Delta \bar{y}_\tau + R(\bar{y}_\tau) & = & K[\bar{u}_\tau]\bar{y}_\tau + g_{\bar{u}_\tau} & in \ Q, \\[2mm] \partial_n \bar{y}_\tau & = & 0 & on \ \Sigma, \\[2mm] \bar{y}_\tau(x,t) & = & y_0(x,t) & in \ Q_-, \end{cases} \tag{4.3}$$

$$\begin{cases} -\dfrac{\partial \bar{\varphi}_\tau}{\partial t} - \Delta \bar{\varphi}_\tau + R'(\bar{y}_\tau)\bar{\varphi}_\tau & = & K^*[\bar{u}_\tau]\bar{\varphi}_\tau + \bar{y}_\tau - y_d & in \ Q, \\[2mm] \partial_n \bar{\varphi}_\tau & = & 0 & on \ \Sigma, \\[2mm] \bar{\varphi}_\tau & = & 0 & in \ Q_+, \end{cases} \tag{4.4}$$

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

Combining Proposition 4.2 and (4.5) we deduce the following corollary.

Corollary 4.4. *Let $\bar{u}_\tau$ be a local minimum of* (P_τ) *with $\bar{u}_\tau \not\equiv 0$ and let $\bar{y}_\tau$, $\bar{\varphi}_\tau$ and $\bar{\lambda}_\tau$ satisfy* (4.3)-(4.5), *then*

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

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

Finally, we analyze the convergence of the above discretization.

Theorem 4.5. *Let $\{\bar{u}_\tau\}_{\tau \searrow 0}$ be a sequence of discrete controls such that every control $\bar{u}_\tau$ 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 $J(\bar{u}_\tau) \to \inf (P)$ as $\tau \searrow 0$. In addition, if $\tau_j \to 0$ and $u_{\tau_j} \overset{*}{\rightharpoonup} \bar{u}$ in $\mathcal{M}[0,T]$, then $\lim_{j\to\infty} \|\bar{u}_{\tau_j}\|_{\mathcal{M}[0,T]} = \|\bar{u}\|_{\mathcal{M}[0,T]}$ and $\lim_{j\to\infty} \|\bar{y}_{\tau_j} - \bar{y}\|_Y = 0$, where $\bar{y}_{\tau_j}$ and $\bar{y}$ denote the states associated to $\bar{u}_{\tau_j}$ 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 $j \to \infty$. From Lemma 3.2 we have $\bar{y}_{\tau_j} \to \bar{y}$ in Y. Let $\tilde{u}$ be a solution of (P). Then, we get with Proposition 4.1 and Lemma 3.2

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

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 $J(\bar{u}_{\tau_j}) \to J(\bar{u})$. This convergence along with $y_{\bar{u}_{\tau_j}} \to \bar{y}$ in $L^2(Q)$ implies that $\|\bar{u}_{\tau_j}\|_{\mathcal{M}[0,T]} \to \|\bar{u}\|_{\mathcal{M}[0,T]}$. $\qquad\square$

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

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

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 $y_\tau^c \in Y \cap C(\bar{Q} \cup \bar{Q}_-)$, $\varphi_\tau^c \in H^1(Q) \cap C(\bar{Q} \cup \bar{Q}_+)$ and $\lambda_\tau^c \in \partial j_\tau(u_\tau^c)$ such that*

$$\begin{cases} \dfrac{\partial y_\tau^c}{\partial t} - \Delta y_\tau^c + R(y_\tau^c) &= K[u_\tau^c] y_\tau^c + g_{u_\tau^c} & in\ Q, \\[2mm] \partial_n y_\tau^c &= 0 & on\ \Sigma, \\[1mm] y_\tau^c(x,t) &= y_0(x,t) & in\ Q_-, \end{cases} \tag{5.1}$$

$$\begin{cases} -\dfrac{\partial \varphi_\tau^c}{\partial t} - \Delta \varphi_\tau^c + R'(y_\tau^c)\varphi_\tau^c &= K^*[u_\tau^c]\varphi_\tau^c + y_\tau^c - y_d & in\ Q, \\[2mm] \partial_n \varphi_\tau^c &= 0 & on\ \Sigma, \\[1mm] \varphi_\tau^c &= 0 & in\ Q_+, \end{cases} \tag{5.2}$$

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

Taking into account Proposition 4.2, the condition on the subgradient $\lambda_\tau^c \in \partial j_\tau(u_\tau^c)$ can be written as

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

We solve the system (5.1)–(5.4) by a semi-smooth Newton method. 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. To solve the delay linear parabolic partial differential equations that appear in the process we consider the standard continuous piecewise linear finite elements in space and piecewise constant discontinuous Galerkin method in time, i.e., 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 $c = c_k$, $k \ge 0$, it is taken as the initial guess to solve the optimality system for $c = c_{k+1} > c_k$. The process stops when further changes of c_k do not alter the solution in a significant way. We choose $C = c\nu$ at every iteration.

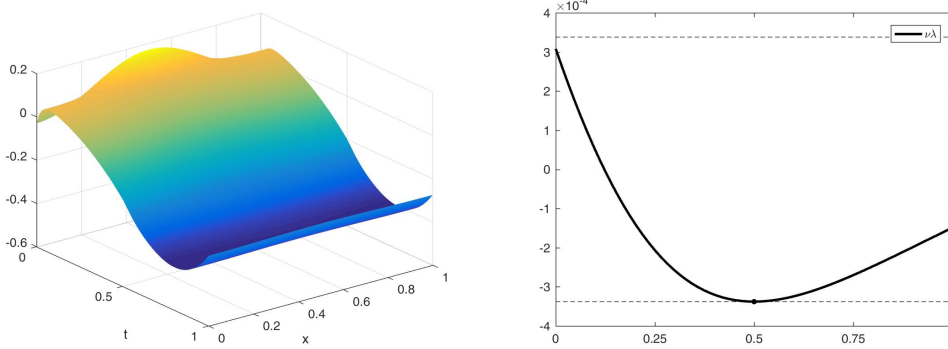In all our examples, the reaction term is of the form

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

Example 5.1 (Example with known critical point).

To test the discretization and the optimization algorithm, we first construct an example with known solution of the optimality system given in Theorem 3.8.

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

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

Figure 1: Target (left) and $\nu\bar\lambda(t)$ for Example 5.1

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 prehistory

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

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

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

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

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

Taking into account that $-\Delta\bar\varphi = 0$, we have that $\bar\varphi$ satisfies the adjoint state equation; see Figure 1 for a picture of the computed target.

Finally, we compute

$$\nu\bar\lambda(s) = -\int_0^T \int_\Omega \bar y(x,t-s)\bar\varphi(x,t)\,\mathrm{d}x\,\mathrm{d}t.$$

With our choices of t^*, u^*, and $\bar\varphi(x,t)$, we have that $s \mapsto \nu\bar\lambda(s)$ is a strictly convex function in $[0,1]$ that has a minimum at t^* such that $\nu\bar\lambda(t^*) = -3.39817 \times 10^{-4}$ (see Figure 1). If we define $\nu = 3.39817 \times 10^{-4}$, we have that $|\bar\lambda(s)| < 1$ for all $s \ne t^*$ and $\bar\lambda(t^*) = -1$, and therefore $(\bar u, \bar y, \bar\varphi, \bar\lambda)$ satisfies the first order optimality conditions for problem (P) with data ν and y_d.

The values of the differentiable and non-differentiable parts of the functional are

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

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 1. Notice that we are able to confirm all the results of Theorem 4.5. 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.

τ	$\bar{u}_\tau$	$\|\bar{y} - \bar{y}_\tau\|_Y$	$\|\bar{u}_\tau\|_{\mathcal{M}[0,T]}$	$F(\bar{u}_\tau)$	$\nu j(\bar{u}_\tau)$
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	$9.067e-5$	$2.617e-3$

Table 1: Results of Example with known solution

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 $\bar{u}_\tau$ increase. As it was proved in Proposition 3.10, there 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 2.

ν	$F_\tau(\bar{u}_\tau)$	$\|\bar{u}_\tau\|_{\mathcal{M}[0,T]}$	$\sharp\mathrm{supp}\bar{u}_\tau$
$1e-1$	$8.85e-2$	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

Table 2: Sensitivity of the norm and the support of the optimal control

Example 5.3 (Recovering the solution of a system with non-local Pyragas feedback control).

We consider the data of Example 1 in [24], namely $\Omega = (-20, 20)$, $T = 40$, $\rho = 1$, $y_1 = 0$, $y_2 = 0.25$, $y_3 = 1$. The prehistory is 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

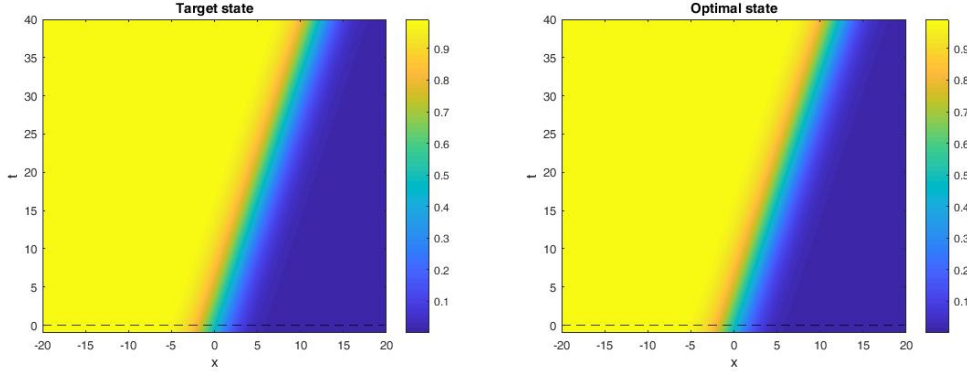$$\int_{[0,T]} y(x, t - s)du_d(s) = \kappa\left(\frac{1}{t_b - t_a}\int_{t_a}^{t_b} y(x, t - s)ds - y(x, t)\right),$$

Figure 2: Target (left) obtained with a non-local Pyragas feedback control and optimal state (right)

with parameters $\kappa = 0.5$, $t_a = 0.456$, $t_b = 0.541$. We fix the parameter $\nu = 1e-2$ –this is big 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

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

The associated values of the objective are $F(\bar{u}) = 5.3e-5$ and $\nu j(\bar{u}) = 4.87e-3$, hence $J(\bar{u}) = 0.49\nu$. The target and the optimal state are illustrated in Figure 2. To solve the equations, we have used a space mesh with 513 evenly spaced nodes and a time grid with a variable stepsize and 778 nodes. The discretization of the control has been done with 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 prehistory is 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 $[T/2, T]$. Therefore, here we redefine the differentiable part of the functional $J(u)$ by

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

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

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

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

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

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

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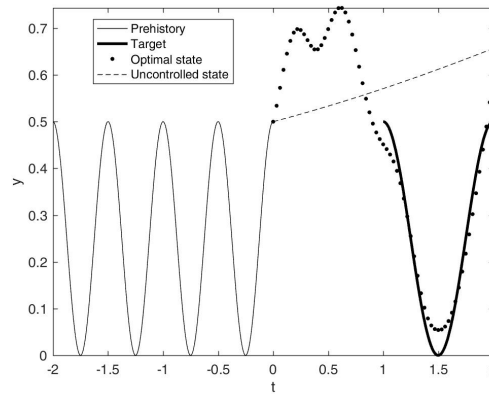

Figure 3: Data and solution for Example 5.5

References

[1] N. U. Ahmed. Existence of optimal controls for a general class of impulsive systems on Banach spaces. *SIAM J. Control Optim.*, 42(2):669–685, 2003. URL: `http://dx.doi.org/10.1137/S0363012901391299`, `doi:10.1137/S0363012901391299`.

[2] H. T. Banks and J. A. Burns. Hereditary control problems: numerical methods based on averaging approximations. *SIAM J. Control Optimization*, 16(2):169–208, 1978. URL: `http://dx.doi.org/10.1137/0316013`, `doi:10.1137/0316013`.

[3] H. T. Banks, J. A. Burns, and E. M. Cliff. Parameter estimation and identification for systems with delays. *SIAM J. Control Optim.*, 19(6):791–828, 1981. URL: `http://dx.doi.org/10.1137/0319051`, `doi:10.1137/0319051`.

[4] H. T. Banks and Andrzej Manitius. Application of abstract variational theory to hereditary systems–a survey. *IEEE Trans. Automatic Control*, AC-19:524–533, 1974.

[5] D. D. Baĭnov and D. P. Mishev. *Oscillation theory for neutral differential equations with delay.* Adam Hilger, Ltd., Bristol, 1991.

[6] Alfredo Bellen and Marino Zennaro. *Numerical methods for delay differential equations.* Numerical Mathematics and Scientific Computation. The Clarendon Press, Oxford University Press, New York, 2003. URL: `http://dx.doi.org/10.1093/acprof:oso/9780198506546.001.0001`, `doi:10.1093/acprof:oso/9780198506546.001.0001`.

[7] Richard Bellman. *A survey of the mathematical theory of time-lag, retarded control, and hereditary processes.* The Rand Corporation, Santa Monica, Calif., 1954. With the assistance of John M. Danskin, Jr.

[8] E. Casas. Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations. *SIAM J. Control Optim.*, 35(4):1297–1327, 1997.

[9] E. Casas, F. Kruse, and K. Kunisch. Optimal control of semilinear parabolic equations by BV-functions. *SIAM J. Control Optim.*, to appear, 2016.

[10] E. Casas and K. Kunisch. Optimal control of semilinear elliptic equations in measure spaces. *SIAM J. Control Optim.*, 52(1):339–364, 2013.

[11] E. Casas, C. Ryll, and F. Tröltzsch. Sparse optimal control of the Schlögl and FitzHugh-Nagumo systems. *Computational Methods in Applied Mathematics*, 13:415–442, 2014. `doi: 10.1515/cmam-2013-0016`.

[12] F. Colonius and D. Hinrichsen. Optimal control of hereditary differential systems. In *Recent theoretical developments in control (Proc. Conf., Univ. Leicester, Leicester, 1976)*, pages 215–239. Academic Press, London-New York, 1978. With discussion.

[13] T. Erneux. *Applied delay differential equations*, volume 3 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer, New York, 2009.

[14] P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. Pitman, Boston-London-Melbourne, 1985.

[15] Jack K. Hale and Sjoerd M. Verduyn Lunel. *Introduction to functional-differential equations*, volume 99 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1993. URL: `http://dx.doi.org/10.1007/978-1-4612-4342-7`, `doi:10.1007/978-1-4612-4342-7`.

[16] Jin-Mun Jeong and Hae-Jun Hwang. Optimal control problems for semilinear retarded functional differential equations. *J. Optim. Theory Appl.*, 167(1):49–67, 2015. URL: `http://dx.doi.org/10.1007/s10957-015-0726-8`, `doi:10.1007/s10957-015-0726-8`.

[17] F. Kappel and K. Kunisch. Spline approximations for neutral functional-differential equations. *SIAM J. Numer. Anal.*, 18(6):1058–1080, 1981. URL: `http://dx.doi.org/10.1137/0718072`, `doi:10.1137/0718072`.

[18] Karl Kunisch. Approximation schemes for nonlinear neutral optimal control systems. *J. Math. Anal. Appl.*, 82(1):112–143, 1981. URL: `http://dx.doi.org/10.1016/0022-247X(81)90228-6`, `doi:10.1016/0022-247X(81)90228-6`.

[19] Y. N. Kyrychko, K. B. Blyuss, and E. Schöll. Amplitude death in systems of coupled oscillators with distributed-delay coupling. *Eur. Physi. J. B*, 84:307–315, 2011. `doi:10.1140/epjb/e2011-20677-8`.

[20] O.A. Ladyzhenskaya, V.A. Solonnikov, and N.N. Ural'tseva. *Linear and Quasilinear Equations of Parabolic Type*. American Mathematical Society, 1988.

[21] J. Löber, R. Coles, J. Siebert, H. Engel, and E. Schöll. Control of chemical wave propagation. *arXiv*, 1403:3363, 2014.

[22] Boris S. Mordukhovich, Dong Wang, and Lianwen Wang. Optimal control of delay-differential inclusions with functional endpoint constraints in infinite dimensions. *Nonlinear Anal.*, 71(12):e2740–e2749, 2009. URL: `http://dx.doi.org/10.1016/j.na.2009.06.022`, `doi:10.1016/j.na.2009.06.022`.

[23] Boris S. Mordukhovich, Dong Wang, and Lianwen Wang. Optimization of delay-differential inclusions in infinite dimensions. *Pac. J. Optim.*, 6(2):353–374, 2010.

[24] P. Nestler, E. Schöll, and F. Tröltzsch. Optimization of nonlocal time-delayed feedback controllers. *Computational Optimization and Applications*, DOI 10.1007/s10589-015-9809-6, published online 2015.

[25] K. Pyragas. Continuous control of chaos by self-controlling feedback. *Phys. Rev. Lett.*, A 170:421, 1992.

[26] K. Pyragas. Delayed feedback control of chaos. *Phil. Trans. R. Soc*, A 364:2309, 2006.

[27] Jean-Pierre Richard. Time-delay systems: an overview of some recent advances and open problems. *Automatica J. IFAC*, 39(10):1667–1694, 2003. URL: `http://dx.doi.org/10.1016/S0005-1098(03)00167-5`, `doi:10.1016/S0005-1098(03)00167-5`.

[28] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, London, 1970.

[29] E. Schöll and H.G. Schuster. *Handbook of Chaos Control*. Wiley-VCH, Weinheim, 2008.

[30] R. E. Showalter. *Monotone operators in Banach space and nonlinear partial differential equations*, volume 49 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1997.

[31] 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. *Physical Review E*, 89, 052909, 2014. `doi:10.1103/PhysRevE.89.052909`.

[32] J. Siebert and E. Schöll. Front and turing patterns induced by mexican-hat-like nonlocal feedback. *Europhys. Lett.*, 109, 40014, 2015.

[33] F. Tröltzsch. *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Philadelphia, 2010.

[34] F. Unger and L. von Wolfersdorf. On a control problem for memory kernels in heat conduction. *Z. Angew. Math. Mech.*, 75(5):365–370, 1995. URL: `http://dx.doi.org/10.1002/zamm.19950750505`, `doi:10.1002/zamm.19950750505`.

[35] L. von Wolfersdorf. On optimality conditions in some control problems for memory kernels in viscoelasticity. *Z. Anal. Anwendungen*, 12(4):745–750, 1993.