ESAIM: COCV 29 (2023) 23
https://doi.org/10.1051/cocv/2023013

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

OPTIMAL CONTROL OF A PARABOLIC EQUATION WITH MEMORY

Eduardo Casas[1,*,**] and Jiongmin Yong[2,***]

Abstract. An optimal control problem for a semilinear parabolic partial differential equation with memory is considered. The well-posedness as well as the first and the second order differentiability of the state equation is established by means of Schauder fixed point theorem and the implicity function theorem. For the corresponding optimal control problem with the quadratic cost functional, the existence of optimal control is proved. The first and the second order necessary conditions are presented, including the investigation of the adjoint equations which are linear parabolic equations with a measure as a coefficient of the operator. Finally, the sufficiency of the second order optimality condition for the local optimal control is proved.

Mathematics Subject Classification. 49J20, 49K20, 35K58, 45K05.

Received July 12, 2022. Accepted March 1, 2023.

1. Introduction

In this paper, we study the following optimal control problem

$$\text{(P)} \qquad \min_{u\in L^\infty(0,T;L^2(\Omega))} J(u) := \frac{1}{2}\int_Q (y_u - y_d)^2 \,\mathrm{d}x\,\mathrm{d}t + \frac{\kappa}{2}\int_{Q_\omega} u^2 \,\mathrm{d}x\,\mathrm{d}t,$$

where $Q = \Omega \times (0,T)$ with Ω a bounded domain of $\mathbb{R}^n$, $1 \le n \le 3$, $0 < T < \infty$, $Q_\omega = \omega \times (0,T)$, ω a measurable subset of Ω with positive measure, and y_u is the solution of the following Neumann initial-boundary value problem

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y + f(x,t,y) + K[y] = g + \chi_\omega u & \text{in } Q,\\ \partial_\nu y = 0 \quad \text{on } \Sigma,\ \ y(0) = y_0 & \text{in } \Omega. \end{cases} \tag{1.1}$$

** This author was supported by MCIN/ AEI/10.13039/501100011033/ under research project PID2020-114837GB-I00.
*** This author is partially supported by NSF Grant DMS-1812921. The main part of this work was completed while this author was visiting the first author at Universidad de Cantabria. This author would like to thank the first author for his hospitality.

Keywords and phrases: Parabolic partial differential equation with memory, optimal control, optimality conditions.

[1] Departmento de Matemática Aplicada y Ciencias de la Computación, E.T.S.I. Industriales y de Telecomunicación, Universidad de Cantabria, 39005 Santander, Spain.
[2] Department of Mathematics, University of Central Florida, Orlando, FL 32816, USA.

* Corresponding author: eduardo.casas@unican.es

In the above equation, $K[y]$ denotes the function

$$K[y](x,t) = \int_{[0,t]} a(x,t,s,y(x,s))\,\mathrm{d}\mu(s) \quad \text{for } (x,t) \in Q, \tag{1.2}$$

where μ is a real Borel measure in $[0,T]$. Such a term represent the memory. It is easy to understand that memory exists in almost all applications, in particular, in the diffusion processes described by parabolic partial differential equations. There are at least two commonly accepted situations that this will happen: (i) As it is known, the classical heat equation is derived from the Fourier's law. In the derivation, for simplicity, people neglect the memory/time delay effect. If one takes this into account, then the memory appears. See [15], as well as [20, 27]. We see that the situation is actually more difficult since the memory can appear in the highest derivative terms. Here, we only consider a much simpler version. But it is still meaningful since we may regard it as an external heat source/sink depending on the past and up to current temperature. (ii) In diffusion of population/epidemic models, it is easy to understand that the current diffusion situation heavily depends on the past and up to current concentration of the spices. See [22], and references cited therein. By using the term $K[y]$ is one way to describe such a situation. If we consider the Lebesgue decomposition of μ: $\mathrm{d}\mu = h\mathrm{d}t + \sigma$, a typical form of the measure μ corresponds with the case where σ is a combination of Dirac measures, $\sigma = \sum_{i=1}^{\infty} c_i\delta_{t_i}$, with $\{t_i\}_{i=1}^{\infty} \subset [0,T]$ being an increasing sequence, $\{c_i\}_{i=1}^{\infty} \subset \mathbb{R}$, and $\sum_{i=1}^{\infty}|c_i| < \infty$. Hence, we have

$$\int_{[0,t]} \phi(t)\,\mathrm{d}\mu(t) = \int_0^t \phi(t)h(t)\,\mathrm{d}t + \sum_{t_i \le t} c_i\phi(t_i) \quad \forall \phi \in C[0,T].$$

In the above, the two terms on the right-hand side represent the continuous and the discrete memories, respectively. We point out that the diffusion process under consideration could have some special memory at some specific time moments. For example, suppose we are considering a heating process starting from certain initial temperature distribution $y_0(x)$. Then the changing of the temperature distribution at time t might be affected by the action made by the programmed machine at the previous moments $t_1, t_2, \cdots$. One could easily cook up some other similar examples.

The so-called *fading memory* is a main feature of the *memory kernel* $a(\cdot,\cdot,\cdot,\cdot)$. It plays an essential role for infinite horizon problems, and such a feature could be also interesting for finite horizon problems. The fading memory can be characterized by the following:

$$t \mapsto a(x,t,s,y) \text{ is decreasing on } [s,T], \qquad s \mapsto a(x,t,s,y) \text{ is increasing on } [0,t].$$

Without the term $K[y]$, the semi-linear parabolic equations have been extensively studied. See [19] for the standard classical theory, and [10, 16, 17] for some further/recent developments. The corresponding optimal control problems can be found, for instance, in [4, 7, 9] and the references therein. Parabolic equations with memory have been investigated by a number of authors for various situations [2, 13, 24, 25]. There were some optimal control problems studied for the abstract evolution equations and some PDEs with memory, see [1, 26]. However, it seems to us that the equation of form (1.1) has not been discussed and of course, the corresponding optimal control problem has not been touched. The purpose of the current paper is to analyze equations of the form (1.1), and carry out the corresponding optimal control theory.

The rest of the paper is organized as follows. In Section 2, we present a careful analysis on the state equation. It turns out that due to the appearance of the memory term $K[y]$ governed by a general memory kernel and the general signed measure, together with the possibly super-linear growth of the nonlinear term f, the well-posedness of (1.1) becomes a little technically subtle. Optimal control problem is investigated in Section 3. It includes the existence of optimal controls, the first and the second order necessary conditions, and the sufficiency of the second order optimality condition for the local optimal control. We indicate that because of the memory term involves the real valued measure μ, the adjoint equation has a term of unknown function with μ as a part of the coefficient. This bring us a proper Bochner integral interpretation of the term, which makes the first and the second order necessary conditions very interesting and attractive. Some concluding remarks are collected in Section 4.

2. Analysis of the parabolic equation with memory

In this section, we perform the analysis of the following semilinear parabolic equation with memory:

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y + f(x,t,y) + K[y] = g & \text{in } Q, \\ \partial_n y = 0 \quad \text{on } \Sigma, \ y(0) = y_0 & \text{in } \Omega. \end{cases} \tag{2.1}$$

We make the following assumptions on the data of this equation.

(A1) Ω is a bounded domain of $\mathbb{R}^n$, $1 \le n \le 3$, with a Lipschitz boundary Γ; $T \in (0,\infty)$ is a finite horizon; $Q = \Omega \times (0,T)$; and $\Sigma = \Gamma \times (0,T)$.

(A2) The map g belongs to $L^r(0,T;L^p(\Omega))$ with $\frac{1}{r} + \frac{n}{2p} < 1$ and $p, r \in [2,\infty]$, and $y_0 \in C(\bar{\Omega})$.

(A3) The map $f : Q \times \mathbb{R} \longrightarrow \mathbb{R}$ is Carathéodory function and of class C^2 with respect to the third variable, satisfying

$$f(x,t,0) = 0, \tag{2.2}$$

$$\exists\, \Lambda_f \ge 0 \text{ such that } \frac{\partial f}{\partial y}(x,t,y) \ge -\Lambda_f \quad \forall y \in \mathbb{R}, \tag{2.3}$$

$$\forall M > 0\, \exists\, C_{f,M} \text{ such that } \Big|\frac{\partial f}{\partial y}(x,t,y)\Big| + \Big|\frac{\partial^2 f}{\partial y^2}(x,t,y)\Big| \le C_{f,M} \quad \text{if } |y| \le M, \tag{2.4}$$

for almost all $(x,t) \in Q$.

(A4) The function $a : Q \times [0,T] \times \mathbb{R} \longrightarrow \mathbb{R}$ is measurable and of class C^2 with respect to the last variable and it satisfies:

$$a(x,t,s,0) = 0, \tag{2.5}$$

$$\exists\, C_a \text{ such that } \Big|\frac{\partial a}{\partial y}(x,t,s,y)\Big| + \Big|\frac{\partial^2 a}{\partial y^2}(x,t,s,y)\Big| \le C_a, \tag{2.6}$$

for almost all $(x,t) \in Q$. Furthermore, we assume that a and $\frac{\partial^j a}{\partial y^j}$, $j = 1,2$, are continuous with respect to the third variable.

(A5) We assume that μ belongs to $M[0,T]$, the space of real valued regular Borel measures in $[0,T]$.

With $C[0,T]$ we denote the space of continuous real functions defined in $[0,T]$. This is a Banach space when endowed with the supremum norm. It is well-known that $M[0,T]$ is the dual space of $C[0,T]$ and

$$\|\mu\|_{C[0,T]^*} = \|\mu\|_{M[0,T]} = |\mu|([0,T]),$$

where $|\mu|$ denotes the total variation measure of μ; see, for instance, Chapter 6 of [21]. In the sequel, we will simply write $\|\mu\|$. Given a function $y : \Omega \times [0,T] \longrightarrow \mathbb{R}$ continuous with respect to the second variable, we set

$$K[y](x,t) = \int_{[0,t]} a(x,t,s,y(x,s))\, \mathrm{d}\mu(s) \qquad \text{for } (x,t) \in Q.$$

We denote $W(0,T) = \{y \in L^2(0,T;H^1(\Omega)) : \frac{\partial y}{\partial t} \in L^2(0,T;H^1(\Omega)^*\}$ endowed with the norm

$$\|y\|_{W(0,T)} = \|y\|_{L^2(0,T;H^1(\Omega))} + \Big\|\frac{\partial y}{\partial t}\Big\|_{L^2(0,T;H^1(\Omega)^*)}.$$

Remark 2.1. We observe that the assumption (2.2) can be replaced by the more general hypothesis $f(\cdot\,,\cdot\,,0) \in L^r(0,T;L^p(\Omega))$. Indeed, it is enough to rename f and g as $f - f(\cdot\,,\cdot\,,0)$ and $g - f(\cdot\,,\cdot\,,0)$, respectively. Analogously, we can relax the assumption (2.5). If we denote by $\hat{a} : Q \longrightarrow \mathbb{R}$ the function defined by

$$\hat{a}(x,t) = \int_0^t a(x,t,s,0)\,\mathrm{d}\mu(s),$$

it is enough to assume that $\hat{a} \in L^r(0,T;L^p(\Omega))$ and to replace a and g by $a - \hat{a}$ and $g - \hat{a}$, and to define $K[y]$ accordingly. The condition on $\hat{a}$ holds if $\sup_{s\in[0,T]} \|a(\cdot\,,\cdot\,,s,0)\|_{L^r(0,T;L^p(\Omega))} < \infty$ is satisfied.

Now, we address the issue of existence, uniqueness, and regularity of a solution to (2.2).

Theorem 2.2. *Under the assumptions (A1)–(A5), (2.1) has a unique solution $y \in W(0,T) \cap C(\bar{Q})$. In addition, there exist constants C_W and C_∞ independent of (g, y_0) such that the following estimates are satisfied:*

$$\|y\|_{W(0,T)} \le C_W\Big(\|g\|_{L^2(Q)} + \|y_0\|_{L^2(\Omega)}\Big), \tag{2.7}$$

$$\|y\|_{C(\bar{Q})} \le C_\infty\Big(\|g\|_{L^r(0,T;L^p(\Omega))} + \|y_0\|_{L^\infty(\Omega)}\Big). \tag{2.8}$$

Finally, if the weak convergence $g_k \rightharpoonup g$ in $L^r(0,T;L^p(\Omega))$ holds, then $y_k \rightharpoonup y$ in $W(0,T)$ and $\|y_k - y\|_{C(\bar{Q})} \to 0$ as $k \to \infty$, where y_k and y are the states associated with g_k and g, respectively.

Proof. Let $\{y_{0,k}\}_{k=1}^\infty$ be a sequence of Lipschitz functions in $\bar{\Omega}$ such that $\|y_{0,k}\|_{C(\bar{\Omega})} \le \|y_0\|_{C(\bar{\Omega})}$ and $\|y_{0,k} - y_0\|_{C(\bar{\Omega})} \to 0$ as $k \to \infty$. Associated with k, we also define the functions $a_k(x,t,s,y) = a(x,t,s,\mathrm{Proj}_{[-k,+k]}(y))$, where $\mathrm{Proj}_{[-k,+k]}(y) = \max\{-k, \min\{y, +k\}\}$, and

$$K_k[w](x,t) = \int_{[0,t]} a_k(x,t,s,w(x,s))\,\mathrm{d}\mu(s).$$

From (2.5) and (2.6) we get $|a_k(x,t,s,y)| \le C_a k$ and, consequently,

$$\|K_k[w]\|_{L^\infty(Q)} \le C_a k\|\mu\| \tag{2.9}$$

for every Carathéodory function function $w : \Omega \times [0,T] \longrightarrow \mathbb{R}$. Now, we define the function $F_k : C(\bar{Q}) \longrightarrow C(\bar{Q})$ by $y_{k,w} = F_k(w)$ solution of the equation

$$\left\{\begin{array}{l} \dfrac{\partial y}{\partial t} - \Delta y + f(x,t,y) = g - K_k[w] \quad \text{in } Q,\\ \partial_n y = 0 \quad \text{on } \Sigma,\ y(0) = y_{0,k} \quad \text{in } \Omega. \end{array}\right. \tag{2.10}$$

Due to (2.9), we get that $g - K_k[w] \in L^r(0,T;L^p(\Omega))$. Hence, (2.10) has a unique solution $y_{k,w} \in W(0,T) \cap C(\bar{Q})$ and it satisfies, for some $\alpha \in (0,1]$ independent of w

$$\|y_{k,w}\|_{W(0,T)} \le C\Big(\|g\|_{L^2(Q)} + C_k\|\mu\| + \|y_{0,k}\|_{L^2(\Omega)}\Big),$$

$$\|y_{k,w}\|_{C^{0,\alpha}(\bar{Q})} \le C\big(\|g\|_{L^r(0,T;L^p(\Omega))} + C_k\|\mu\| + \|y_{0,k}\|_{C^{0,1}(\bar{\Omega})}\Big);$$

see [4] or [7] for the existence and uniqueness of solutions and [11] for the Hölder estimate. Above, as along the proofs in this paper, C will denote a generic constant that could be different from line to line.

Therefore, the image of F_k is a bounded and closed subset of $C^{0,\alpha}(\bar{Q})$, hence it is a compact subset of $C(\bar{Q})$. Then, applying Schauder's fixed point theorem we infer the existence of a function $y_k \in W(0,T) \cap C^{0,\alpha}(\bar{Q})$ satisfying

$$\begin{cases} \dfrac{\partial y_k}{\partial t} - \Delta y_k + f(x,t,y_k) + K_k[y_k] = g \quad \text{in } Q, \\ \partial_n y_k = 0 \quad \text{on } \Sigma, \; y_k(0) = y_{0,k} \quad \text{in } \Omega. \end{cases} \tag{2.11}$$

Moreover, since $y_{0,k} \in C^{0,1}(\bar{\Omega}) \subset H^1(\Omega)$ and $g - f(\cdot,\cdot,y_{k,w}) - K_k[w] \in L^2(Q)$, we deduce from Proposition III-2.5 of [23] that $y_{k,w} \in H^1(Q)$. We prove an estimate for y_k in $L^2(0,T;H^1(\Omega)) \cap L^\infty(0,T;L^2(\Omega))$ independent of k. First, we observe that with (2.5)–(2.6) we obtain

$$\begin{aligned} &\|K_k[y_k](t)\|_{L^2(\Omega)} = \Big(\int_\Omega \Big[\int_{[0,t]} a_k(x,t,s,y_k(x,s)) \,\mathrm{d}\mu(s) \Big]^2 \mathrm{d}x \Big)^{\frac{1}{2}} \\ &\le C_a \Big(\int_\Omega \Big[\int_{[0,t]} |y_k(x,s)| \,\mathrm{d}|\mu|(s) \Big]^2 \mathrm{d}x \Big)^{\frac{1}{2}} \\ &\le C_a \int_{[0,t]} \|y_k(s)\|_{L^2(\Omega)} \,\mathrm{d}|\mu|(s) \le C_a \|\mu\| \max_{0\le s\le t} \|y_k(s)\|_{L^2(\Omega)}. \end{aligned} \tag{2.12}$$

Now, testing (2.11) with $\mathrm{e}^{-4\Lambda_f t} y_k$, where $\Lambda_f \ge 0$ is given satisfying (2.3). We infer for every $t \in (0,T]$

$$\begin{aligned} &\frac{\mathrm{e}^{-4\Lambda_f t}}{2} \|y_k(t)\|^2_{L^2(\Omega)} + \int_0^t \mathrm{e}^{-4\Lambda_f s} \int_\Omega [|\nabla y_k|^2 + \Lambda_f |y_k|^2] \,\mathrm{d}x \,\mathrm{d}s \\ &+ \int_0^t \mathrm{e}^{-4\Lambda_f s} \int_\Omega [\Lambda_f y_k^2 + f(x,s,y_k) y_k] \,\mathrm{d}x \,\mathrm{d}s = \frac{1}{2} \|y_{0,k}\|^2_{L^2(\Omega)} \\ &+ \int_0^t \mathrm{e}^{-4\Lambda_f t} \int_\Omega g y_k \,\mathrm{d}x \,\mathrm{d}s - \int_0^t \mathrm{e}^{-4\Lambda_f t} \int_\Omega K[y_k](x,s) y_k(x,s) \,\mathrm{d}x \,\mathrm{d}s \\ &\le \frac{1}{2} \|y_{0,k}\|^2_{L^2(\Omega)} + \frac{1}{2} \|g\|^2_{L^2(Q)} + \frac{1}{2} \int_0^t \int_\Omega y_k^2 \,\mathrm{d}x \,\mathrm{d}s + C_a^2 \|\mu\|^2 \int_0^t \max_{0\le\tau\le s} \|y_k(\tau)\|^2_{L^2(\Omega)} \mathrm{d}s \\ &\le \frac{1}{2} \|y_{0,k}\|^2_{L^2(\Omega)} + \frac{1}{2} \|g\|^2_{L^2(Q)} + (\frac{1}{2} + C_a^2 \|\mu\|^2) \int_0^t \max_{0\le\tau\le s} \|y_k(\tau)\|^2_{L^2(\Omega)} \mathrm{d}s. \end{aligned}$$

From (2.2)–(2.3) we get that $\Lambda_f y_k^2 + f(x,t,y_k) y_k \ge 0$. Hence, we infer from the above inequalities and the fact that t was arbitrarily selected in $(0,T]$

$$\begin{aligned} &\max_{0\le s\le t} \|y_k(s)\|^2_{L^2(\Omega)} + \int_0^t \int_\Omega [|\nabla y_k|^2 + \Lambda_f y_k^2] \,\mathrm{d}x \,\mathrm{d}s \\ &\le C \Big(\|y_{0,k}\|^2_{L^2(\Omega)} + \|g\|^2_{L^2(Q)} + \int_0^t \max_{0\le\tau\le s} \|y_k(\tau)\|^2_{L^2(\Omega)} \mathrm{d}s \Big) \end{aligned}$$

Applying Gronwall's inequality to the function $h(t) = \max_{0\le s\le t} \|y_k(s)\|^2_{L^2(\Omega)}$ we get

$$\max_{0\le s\le t} \|y_k(s)\|^2_{L^2(\Omega)} \le C \Big(\|y_{0,k}\|^2_{L^2(\Omega)} + \|g\|^2_{L^2(Q)} \Big).$$

Therefore, the last estimates and the fact that $y_{0,k} \to y_0$ in $L^2(\Omega)$ yield

$$\|y_k\|_{L^\infty(0,T;L^2(\Omega))} + \|y_k\|_{L^2(0,T;H^1(\Omega))} \le C\Big(\|y_0\|_{L^2(\Omega)} + \|g\|_{L^2(Q)}\Big) \quad \forall k \ge 1. \tag{2.13}$$

Combining this estimate with (2.12), using again [4] or [7], and the fact that $\|y_{0,k}\|_{L^\infty(\Omega)} \le \|y_0\|_{L^\infty(\Omega)}$, we deduce

$$\|y_k\|_{C(\bar Q)} \le C\Big(\|g\|_{L^r(0,T;L^p(\Omega))} + \|y_0\|_{L^\infty(\Omega)}\Big) \quad \forall k \ge 1. \tag{2.14}$$

Hence, we have that $K_k[y_k] = K[y_k]$ for every k large enough. Using the above estimates and the fact that $y_{0,k} \to y_0$, it is easy to pass to the limit in (2.11) and to deduce that $y_k \rightharpoonup y$ in $W(0,T)$ and y is a solution of (2.1). Moreover, the estimates (2.7) and (2.8) are straightforward consequences of the estimates proved for $\{y_k\}_{k=1}^\infty$.

Let us prove the uniqueness of solution. Let $y_1, y_2 \in W(0,T) \cap C(\bar Q)$ be solutions of (2.1) and set $y = y_2 - y_1$. Then, subtracting the equations satisfied by y_2 and y_1, we obtain with the mean value theorem

$$\left\{\begin{array}{l} \dfrac{\partial y}{\partial t} - \Delta y + \dfrac{\partial f}{\partial y}(x,t,\hat y)y + \displaystyle\int_{[0,t]} \dfrac{\partial a}{\partial y}(x,t,s,\tilde y(x,s))y(x,s)\,\mathrm{d}\mu(s) = 0 \quad \text{in } Q,\\ \partial_n y = 0 \quad \text{on } \Sigma,\ y(0) = 0 \quad \text{in } \Omega, \end{array}\right. \tag{2.15}$$

where $\hat y = y_2 + \hat\theta(y_1 - y_2)$ and $\tilde y = y_2 + \tilde\theta(y_1 - y_2)$ with $\hat\theta$ and $\tilde\theta$ measurable functions from Q to $[0,1]$. Testing (2.15) with $\mathrm{e}^{-4\Lambda_f t}y$, taking into account (2.6), and arguing similarly as above, we infer from the Gronwall inequality that y satisfies the inequality (2.13) with $g = 0$ and $y_0 = 0$ in the right hand side. Then, the equality $y = 0$ follows.

Finally, we prove the continuity of the solution with respect to the right-hand-side of the equation. Let $g_k \rightharpoonup g$ in $L^r(0,T;L^p(\Omega))$ and denote by y_k and y the solutions of (2.1) corresponding to g_k and g, respectively. From (2.7) and (2.8) we know that $\{y_k\}_{k=1}^\infty$ is bounded in $W(0,T) \cap C(\bar Q)$. Therefore, for a subsequence, $y_k \rightharpoonup \tilde y$ in $W(0,T)$. Since the embedding $W(0,T) \subset L^2(Q)$ is compact and $\{y_k\}_{k=1}^\infty$ is bounded in $C(\bar Q)$, we can assume, taking a new subsequence if necessary, that $y_k(x,t) \to \tilde y(x,t)$ almost everywhere in Q and $y_k \to \tilde y$ strongly in $L^q(Q)$ for every $q < \infty$. Hence, it is easy to pass to the limit in the equation satisfied by y_k and to deduce that $\tilde y$ is the state associated to g. Therefore, the identity $\tilde y = y$ follows and the whole sequence $\{y_k\}_{k=1}^\infty$ converges weakly to y in $W(0,T)$. Now, setting $z_k = y_k - y$ we have

$$\left\{\begin{array}{l} \dfrac{\partial z_k}{\partial t} - \Delta z_k + \dfrac{\partial f}{\partial y}(x,t,\hat y_k)z_k + \displaystyle\int_{[0,t]} \dfrac{\partial a}{\partial y}(x,t,s,\tilde y_k(x,s))z_k(x,s)\,\mathrm{d}\mu(s) = g_k - g \quad \text{in } Q,\\ \partial_n z_k = 0 \quad \text{on } \Sigma,\ z_k(0) = 0 \quad \text{in } \Omega, \end{array}\right.$$

where $\hat y_k = y + \hat\theta_k(y_k - y)$ and $\tilde y_k = y + \tilde\theta_k(y_k - y)$ with $\hat\theta_k$ and $\tilde\theta_k$ measurable functions from Q to $[0,1]$. From this equation we get that $\{z_k\}_{k=1}^\infty$ is bounded in a Hölder space $C^{0,\alpha}(\bar Q)$; see [11]. Therefore, the convergence $z_k \to 0$ strongly in $C(\bar Q)$ holds. □

Now, we define the mapping $F : L^r(0,T;L^p(\Omega)) \longrightarrow W(0,T) \cap C(\bar Q)$ by $F(g) = y_g$ solution of (2.1) associated with g. The next theorem analyzes the differentiability of F. First, we introduce the following notation. Given functions $y, z, z_1, z_2 : \Omega \times [0,T] \longrightarrow \mathbb{R}$ continuous with respect to the second variable, we denote for $(x,t) \in \Omega \times [0,T]$

$$(K'[y]z)(x,t) = \int_{[0,t]} \frac{\partial a}{\partial y}(x,t,s,y(x,s))z(x,s)\,\mathrm{d}\mu(s),$$

$$(K''[y](z_1,z_2))(x,t) = \int_{[0,t]} \frac{\partial^2 a}{\partial y^2}(x,t,s,y(x,s))z_1(x,s)z_2(x,s)\,\mathrm{d}\mu(s).$$

Theorem 2.3. *Under the assumptions (A1)–(A5), the mapping F is of class C^2. Moreover, given $g,h,h_1,h_2 \in L^r(0,T;L^p(\Omega))$, the function $z_h = F'(g)h$ satisfies the equation*

$$\left\{\begin{array}{l} \dfrac{\partial z}{\partial t} - \Delta z + \dfrac{\partial f}{\partial y}(x,t,y_g)z + K'[y_g]z = h \quad \text{in } Q,\\ \partial_n z = 0 \quad \text{on } \Sigma,\ z(0) = 0 \quad \text{in } \Omega, \end{array}\right. \tag{2.16}$$

and $z_{h_1,h_2} = F''(g)(h_1,h_2)$ satisfies equation

$$\left\{\begin{array}{l} \dfrac{\partial z}{\partial t} - \Delta z + \dfrac{\partial f}{\partial y}(x,t,y_g)z + K'[y_g]z \\ \qquad -\dfrac{\partial^2 f}{\partial y^2}(x,t,y_g)z_{h_1}z_{h_2} - K''[y_g](z_{h_1},z_{h_2}) \quad \text{in } Q,\\ \partial_n z = 0 \quad \text{on } \Sigma,\ z(0) = 0 \quad \text{in } \Omega, \end{array}\right. \tag{2.17}$$

where $z_{h_i} = F'(y_g)h_i$, $i = 1,2$.

Proof. We are going to apply the implicit function theorem. To this end, we define the space

$$Y = \left\{y \in W(0,T) \cap C(\bar{Q}) : \frac{\partial y}{\partial t} - \Delta y \in L^r(0,T;L^p(\Omega))\right\}$$

and endowed it with the norm

$$\|y\|_Y = \|y\|_{W(0,T)} + \|y\|_{C(\bar{Q})} + \left\|\frac{\partial y}{\partial t} - \Delta y\right\|_{L^r(0,T;L^p(\Omega))}.$$

Then, Y is a Banach space. We also define the function

$$\begin{aligned} &\mathcal{F} : Y \times L^r(0,T;L^p(\Omega)) \longrightarrow L^r(0,T;L^p(\Omega)) \times C(\bar{\Omega}),\\ &\mathcal{F}(y,g) = \Big(\frac{\partial y}{\partial t} - \Delta y + f(\cdot,\cdot,y) + K[y] - g, y(0) - y_0\Big). \end{aligned}$$

It is immediate to check that $\mathcal{F}$ is of class C^2 and $\mathcal{F}(F(g),g) = (0,0)$ for every $g \in L^r(0,T;L^p(\Omega))$. Moreover, for $y,z \in Y$ we have

$$\frac{\partial \mathcal{F}}{\partial y}(y,g)z = \Big(\frac{\partial z}{\partial t} - \Delta z + \frac{\partial f}{\partial y}(\cdot,\cdot,y)z + K'[y]z, z(0)\Big).$$

Since $\frac{\partial \mathcal{F}}{\partial y}(y,g) : Y \longrightarrow L^r(0,T;L^p(\Omega)) \times C(\bar{\Omega})$ is a linear continuous mapping, it is an isomorphism if and only if the equation

$$\left\{\begin{array}{l} \dfrac{\partial z}{\partial t} - \Delta z + \dfrac{\partial f}{\partial y}(x,t,y_g)z + K'[y_g]z = h \quad \text{in } Q,\\ \partial_n z = 0 \quad \text{on } \Sigma,\ z(0) = z_0 \quad \text{in } \Omega \end{array}\right. \tag{2.18}$$

has a unique solution $z \in Y$ for every $(h, z_0) \in L^r(0,T;L^p(\Omega)) \times C(\bar\Omega)$. This property follows from Theorem 2.2. Indeed, if we define $\hat f(x,t,z) = \frac{\partial f}{\partial y}(x,t,y_g(x,t))z$, $b(x,t,s,z) = \frac{\partial a}{\partial y}(x,t,s,y_g(x,s))z$, and

$$\hat K[z](x,t) = \int_{[0,t]} b(x,t,s,z(x,s))\,\mathrm{d}\mu(s),$$

taking into account that $y_g \in C(\bar Q)$, we infer that $\hat f$ and b satisfy the assumptions (A3) and (A4), respectively. Then, Theorem 2.2 applies to the equation

$$\left\{\begin{array}{l} \dfrac{\partial z}{\partial t} - \Delta z + \hat f(x,t,z) + \hat K[z] = h \quad \text{in } Q,\\ \partial_n z = 0 \quad \text{on } \Sigma,\ z(0) = z_0 \quad \text{in } \Omega. \end{array}\right.$$

Therefore, from the implicit function theorem we deduce that F is of class C^2 and (2.16) and (2.17) follow by diferentiation of the identity $\mathcal{F}(F(g), g) = 0$. □

Before finishing this section we are going to carry out a more detailed study of the linearized equation (2.16).

Theorem 2.4. *For every $g \in L^r(0,T;L^p(\Omega))$ and $h \in L^2(Q)$, the equation (2.16) has a unique solution $z \in H^1(Q) \cap C([0,T];H^1(\Omega))$. Further, there exists a constant C_g depending of g, but independent of h, such that*

$$\|z\|_{H^1(Q)} + \|z\|_{C([0,T];H^1(\Omega))} \le C_g\|h\|_{L^2(Q)}. \tag{2.19}$$

Let us mention that, given a function $z \in C([0,T];L^2(\Omega))$, the integral defining $K'[y_g]z$ is a Bochner integral and actually we have that $K'[y_g]z \in L^\infty(0,T;L^2(\Omega))$. Indeed, for every $t \in [0,T]$, we get with (2.6)

$$\begin{aligned}\|(K'[y_g]z)(t)\|_{L^2(\Omega)} &\le \int_{[0,t]} \Big(\int_\Omega \Big|\frac{\partial a}{\partial y}(x,t,s,y_g(x,s))z(x,s)\Big|^2\,\mathrm{d}x\Big)^{\frac12}\,\mathrm{d}|\mu|(s)\\ &\le C_a\int_{[0,T]} \|z(s)\|_{L^2(\Omega)}\,\mathrm{d}|\mu|(s) \le C_a\|\mu\|\|z\|_{C([0,T];L^2(\Omega))},\end{aligned}$$

which proves that

$$\|K'[y_g]z\|_{L^\infty(0,T;L^2(\Omega))} \le C_a\|\mu\|\|z\|_{C([0,T];L^2(\Omega))}. \tag{2.20}$$

Proof of Theorem 2.4. Let $\{h_k\}_{k=1}^\infty \subset L^r(0,T;L^p(\Omega))$ be a sequence converging strongly to h in $L^2(Q)$. Denote by z_k the solution of (2.16) corresponding to h_k. Then, defining $\hat K[z]$ and $\hat f$ as we did at the end of the proof of Theorem 2.3 and applying Theorem 2.2 we infer the existence of a constant independent of k such that $\|z_k\|_{W(0,T)} \le C\|h_k\|_{L^2(Q)}$. Hence, $\{z_k\}_{k=1}^\infty$ is bounded in $W(0,T)$. By taking a subsequence, we obtain $z_k \rightharpoonup z$ in $W(0,T)$. Then, it is easy to pass to the limit in the equations satisfied by z_k and to deduce that z is a solution of (2.16). Observe that $W(0,T) \subset C([0,T];L^2(\Omega))$ and, hence, $z \in C([0,T];L^2(\Omega))$. Now, from the boundedness of y_g and (2.20), the regularity $z \in H^1(Q) \cap C([0,T];H^1(\Omega))$ follows; see Section III-2 of [23]. Moreover, using that $\|z\|_{W(0,T))} \le C\|h\|_{L^2(Q)}$ and the estimates of [23], the inequality (2.19) is obtained. □

3. Optimal control problem

In this section, we analyze the control problem (P). We prove existence of a solution and derive first and second order optimality conditions. For this purpose we make the following assumptions:

(A6) The target state y_d belongs to $L^2(Q)$ and the coefficient κ in the cost functional is strictly positive.

(A7) In the state equation (1.1), g is an element of $L^\infty(0,T;L^2(\Omega))$ and the controls u belong to $L^\infty(0,T;L^2(\omega))$ with $Q_\omega = \omega\times(0,T)$. By ω we denote a measurable subset of Ω with positive Lebesgue measure. We denote by χ_ω the characteristic function of ω. Hence, we have that $(u\chi_\omega)(x,t)=0$ if $(x,t)\notin Q_\omega$ and equal to $u(x,t)$ if $(x,t)\in Q_\omega$.

Under the assumption (A7) we have that $g+u\chi_\omega\in L^\infty(0,T;L^2(\Omega))$. Then, we can use Theorem 2.2 with $r=\infty$ and $p=2$ to deduce the existence and uniqueness of a solution $y_u\in W(0,T)\cap C(\bar Q)$ for every control $u\in L^\infty(0,T;L^2(\omega))$. Actually, the mapping $G: L^r(0,T;L^2(\omega))\longrightarrow W(0,T)\cap C(\bar Q)$ associating to each control its corresponding state $G(u)=y_u$ is well defined if $r>\frac{4}{4-n}$. Moreover, from Theorem 2.3 we get that $G(u)=F(g+u\chi_\omega)$ is of class C^2. We observe that $z_v=G'(u)v$ is the solution of (2.16) with $h=v\chi_\omega$. By the chain rule we infer that the cost functional $J: L^r(0,T;L^2(\omega))\longrightarrow\mathbb{R}$ is also of class C^2. The following theorem provides the expressions for the first and second derivatives of J.

Theorem 3.1. *For every $u,v,v_1,v_2\in L^r(0,T;L^2(\omega))$ with $r>\frac{4}{4-n}$ the following identities hold*

$$J'(u)v=\int_{Q_\omega}(\varphi_u+\kappa u)v\,\mathrm{d}x\,\mathrm{d}t, \tag{3.1}$$

$$\begin{aligned}&J''(u)(v_1,v_2)\\&=\int_Q\Big[\Big(1-\varphi_u\frac{\partial^2 f}{\partial y^2}(x,t,y_u)\Big)z_{v_1}z_{v_2}-\varphi_u K''[y_u](z_{v_1},z_{v_2})\Big]\,\mathrm{d}x\,\mathrm{d}t+\kappa\int_{Q_\omega}v_1v_2\,\mathrm{d}x\,\mathrm{d}t,\end{aligned} \tag{3.2}$$

where $z_{v_i}=G'(u)v_i$, $i=1,2$, and φ_u is the unique solution in $L^2(0,T;H^1(\Omega))\cap L^\infty(0,T;L^2(\Omega))$ of the adjoint equation

$$\left\{\begin{array}{l}-\dfrac{\partial\varphi}{\partial t}-\Delta\varphi+\dfrac{\partial f}{\partial y}(x,t,y_u)\varphi+K'[y_u]^*\varphi\mu=y_u-y_d\quad\text{in } Q,\\ \partial_n\varphi=0\quad\text{on }\Sigma,\ \ \varphi(T)=0\quad\text{in }\Omega\end{array}\right. \tag{3.3}$$

with

$$(K'[y_u]^*\varphi)(x,t)=\int_t^T\frac{\partial a}{\partial y}(x,s,t,y_u(x,t))\varphi(x,s)\,\mathrm{d}s. \tag{3.4}$$

Before proving this theorem let us comment about the expression $K'[y_u]^*\varphi\mu$. First of all, we observe that for any function $h\in C[0,T]$ and any real valued measure $\mu\in M[0,T]$ the product $h\mu$ is defined as an element of $M[0,T]$ by the identity

$$\langle h\mu,\phi\rangle=\int_{[0,T]}h(t)\phi(t)\,\mathrm{d}\mu(t).$$

Now, we have that for $x\in\Omega$ the mapping $h(t)=\int_t^T\frac{\partial a}{\partial y}(x,s,t,y_u(x,t))\varphi(x,s)\,\mathrm{d}s$ is continuous in $[0,T]$ due to the continuity of y_u and the continuity of $\frac{\partial a}{\partial y}$ on the last two variables. Hence, for every function $z\in C(\bar Q)$ the following identities are fulfilled:

$$\begin{aligned}\langle K'[y_u]^*\varphi\mu,z\rangle_Q&=\int_\Omega\int_{[0,T]}\Big(\int_0^t\frac{\partial a}{\partial y}(x,s,t,y_u(x,t))\varphi(x,s)\,\mathrm{d}s\Big)\,z(x,t)\,\mathrm{d}\mu(t)\,\mathrm{d}x\\&=\int_\Omega\int_0^T\Big(\int_{[0,t]}\frac{\partial a}{\partial y}(x,t,s,y_u(x,s))z(x,s)\,\mathrm{d}\mu(s)\Big)\,\varphi(x,t)\,\mathrm{d}t\,\mathrm{d}x\end{aligned}$$

$$= \int_Q (K'[y_u]z)(x,t)\varphi(x,t)\,\mathrm{d}x\,\mathrm{d}t = \langle K'[y_u]z, \varphi\rangle_Q. \tag{3.5}$$

Regarding equation (3.3), we have to explain what we mean by a solution.

Definition 3.2. We say that $\varphi \in L^1(Q)$ is a solution of (3.3) if

$$\int_Q \left\{\frac{\partial z}{\partial t} - \Delta z + \frac{\partial f}{\partial y}(x,t,y_u)z + K'[y_u]z\right\}\varphi\,\mathrm{d}x\,\mathrm{d}t = \int_Q (y_u - y_d)z\,\mathrm{d}x\,\mathrm{d}t \quad \forall z \in Z, \tag{3.6}$$

where

$$Z = \{z \in H^1(Q) : \frac{\partial z}{\partial t} - \Delta z \in L^\infty(Q),\ z(0) = 0,\ \partial_n z = 0\}.$$

Remark 3.3. Let us observe that $H^1(Q) \subset C([0,T];L^2(\Omega))$, hence the initial condition $z(0) = 0$ makes sense for very $z \in Z$. Moreover, given $z \in H^1(Q)$ with $\frac{\partial z}{\partial t} - \Delta z \in L^\infty(Q)$ there exists $\phi \in L^\infty(Q)$ such that $-\Delta z(t) = \phi(t) - \frac{\partial z}{\partial t}(t) \in L^2(\Omega)$ for almost every $t \in (0,T)$. As a consequence, the existence of $\partial_n z(t) \in H^{-\frac{1}{2}}(\Gamma)$ follows for almost every t; see, for instance, ([14], Cor. I.2.6). In addition, we have that Z is continuously embedded in $C(\bar{Q})$ [11]. Hence, the integrals in (3.6) are well defined.

Lemma 3.4. *Equation (3.3) has a unique solution $\varphi_u \in L^1(Q)$ for all $u \in L^\infty(0,T;L^2(\omega))$. Moreover, φ_u belongs to $L^2(0,T;H^1(\Omega)) \cap L^\infty(0,T;L^2(\Omega))$ and there exists a constant M depending of $\|\mu\|$, but independent of u, such that the following estimate is fulfilled:*

$$\|\varphi_u\|_{L^2(0,T;H^1(\Omega))} + \|\varphi_u\|_{L^\infty(0,T;L^2(\Omega))} \le M\|y_u - y_d\|_{L^2(Q)}. \tag{3.7}$$

Proof. To prove the uniqueness is enough to show that $\varphi = 0$ is the unique solution of the homogeneous equation. To establish this we select z as the solution of (2.18) with $y_g = y_u$ and $h(x,t) = \operatorname{sign}(\varphi(x,t))$. Then, $z \in Z$ and (3.6) holds with right hand side equal to 0, which implies that $\varphi = 0$.

To prove the existence of a solution we firstly consider the case where $\mu \in L^\infty(0,T)$. For every integer $k \ge 1$ we consider the equation

$$\left\{\begin{array}{l} -\dfrac{\partial \varphi}{\partial t} - \Delta\varphi + \dfrac{\partial f}{\partial y}(x,t,y_u)\varphi + K'[y_u]^*\operatorname{Proj}_{[-k,+k]}(\varphi)\mu = y_u - y_d \quad \text{in } Q,\\ \partial_n\varphi = 0 \quad \text{on } \Sigma,\ \varphi(T) = 0 \quad \text{in } \Omega \end{array}\right.$$

Then following the lines of the proof of Theorem 2.2 it is easy to deduce the existence of a solution $\varphi_k \in W(0,T) \cap C(\bar{Q})$ to this equation. Further, the following inequality is satisfied:

$$\|\varphi_k\|_{L^\infty(0,T;L^2(\Omega))} + \|\varphi_k\|_{L^2(0,T;H^1(\Omega))} \le C\|y_u - y_d\|_{L^2(Q)} \quad \forall k \ge 1. \tag{3.8}$$

As in inequality (2.13), the constant C depends on $\|\mu\|_{L^1(0,T)}$. Using this estimate in (3.3) we infer that $\frac{\partial f}{\partial y}(x,t,y_u)\varphi_k + K'[y_u]^*\varphi_k\mu$ is uniformly bounded in $L^2(Q)$. Hence, we have that $\{\varphi_k\}_{k=1}^\infty$ is bounded in $W(0,T)$. Therefore, for a subsequence denoted by itself, we have that $\varphi_k \rightharpoonup \varphi$ in $W(0,T)$ and $\varphi_k \to \varphi$ strongly in $L^2(Q)$. Whence, we can pass to the limits as $k \to \infty$ in the equation satisfied by φ_k and to get that φ solves (3.3). Moreover, the estimate (3.7) follows from (3.8).

Finally, we get rid of the assumption $\mu \in L^\infty(0,T)$. For this purpose, given $\mu \in M[0,T]$ we consider a sequence $\{\mu_k\}_{k=1}^\infty \subset L^\infty(0,T)$ such that $\mu_k \stackrel{*}{\rightharpoonup} \mu$ in $M[0,T]$ and $\|\mu_k\|_{L^1(0,T)} \le \|\mu\|$. Then, we get solutions y_k to (3.3) corresponding to the functions μ_k. For every k, the estimate (3.7) holds due to the boundedness of $\|\mu_k\|_{L^1(0,T)}$. Hence, taking a subsequence, we infer that $\varphi_k \stackrel{*}{\rightharpoonup} \varphi$ in $L^2(0,T;H^1(\Omega)) \cap L^\infty(0,T;L^2(\Omega))$. To prove that φ solves (3.3) it is enough to pass to the limit in the identities

$$\int_Q \left\{ \frac{\partial z}{\partial t} - \Delta z + \frac{\partial f}{\partial y}(x,t,y_u)z + K_k'[y_u]z \right\} \varphi_k \,\mathrm{d}x\,\mathrm{d}t = \int_Q (y_u - y_d)z \,\mathrm{d}x\,\mathrm{d}t \quad \forall z \in Z,$$

where

$$(K_k'[y_u]z)(x,t) = \int_{[0,t]} \frac{\partial a}{\partial y}(x,t,s,y_u(x,s))z(x,s)\mu_k(s)\,\mathrm{d}s.$$

Precisely, the unique delicate point to pass to the limit is in the integral

$$\int_Q (K_k'[y_u]z)(x,t)\varphi_k(x,t)\,\mathrm{d}x\,\mathrm{d}t \to \int_Q (K'[y_u]z)(x,t)\varphi(x,t)\,\mathrm{d}x\,\mathrm{d}t. \tag{3.9}$$

To prove this we observe that the convergence $\mu_k \stackrel{*}{\rightharpoonup} \mu$ in $M[0,T]$ implies the pointwise convergence $(K_k'[y_u]z)(x,t) \to (K'[y_u]z)(x,t)$. Moreover, we have $|(K_k'[y_u]z)(x,t)| \le C_a\|\mu_k\|_{L^1(0,T)}\|z\|_{C(\bar Q)} \le C_a\|\mu\|\|z\|_{C(\bar Q)}$. Applying the Lebesgue dominated convergence theorem we infer that $K_k'[y_u]z \to K'[y_u]z$ in $L^2(Q)$. This combined with the weak convergence $\varphi_k \rightharpoonup \varphi$ in $L^2(Q)$ proves (3.9). Therefore, φ is solution of (3.3) and satisfies (3.7). $\square$

Remark 3.5. Since $L^\infty(Q)$ is dense in $L^\infty(0,\infty;L^2(\Omega))$ and $\varphi_u \in L^\infty(0,T;L^2(\Omega))$, the identity (3.6) also holds if we assume that $\frac{\partial z}{\partial t} - \Delta z + \frac{\partial f}{\partial y}(x,t,y_u)z + K'[y_u]z \in L^\infty(0,T;L^2(\Omega))$. In particular, this is true if we take $z = z_v = G'(u)v$ for $v \in L^r(0,T;L^2(\Omega))$.

Proof of Theorem 3.1. Let us show the formulas (3.1) and (3.2). Given $u, v \in L^r(0,T;L^2(\omega))$ and setting $z_v = G'(u)v$, the chain rule yields

$$J'(u)v = \int_Q (y_u - y_d)z_v\,\mathrm{d}x\,\mathrm{d}t + \kappa\int_{Q_\omega} uv\,\mathrm{d}x\,\mathrm{d}t. \tag{3.10}$$

From Definition 3.2 and Remark 3.5, we get

$$\begin{aligned}&\int_Q (y_u - y_d)z_v\,\mathrm{d}x\,\mathrm{d}t\\ &= \int_Q \left\{ \frac{\partial z_v}{\partial t} - \Delta z_v + \frac{\partial f}{\partial y}(x,t,y_u)z_v + K'[y_u]z_v \right\}\varphi_u\,\mathrm{d}x\,\mathrm{d}t = \int_{Q_\omega} \varphi_u v\,\mathrm{d}x\,\mathrm{d}t.\end{aligned}$$

Identity (3.1) is a straightforward consequence of this identity and (3.10). Let us prove (3.2). For $u, v_1, v_2 \in L^r(0,T;L^2(\omega))$, differentiating the expression (3.10) with $v = v_1$ we obtain

$$J''(u)(v_1,v_2) = \int_Q [(y_u - y_d)z_{v_1,v_2} + z_{v_1}z_{v_2}]\,\mathrm{d}x\,\mathrm{d}t + \kappa\int_{Q_\omega} v_1v_2\,\mathrm{d}x\,\mathrm{d}t, \tag{3.11}$$

where $z_{v_1,v_2} = G''(u)(v_1,v_2)$ and $z_{v_i} = G'(u)v_i$, $i = 1,2$. Invoking again Remark 3.5 and using (2.17), we deduce

$$\begin{aligned}
&\int_Q (y_u - y_d) z_{v_1,v_2}\,\mathrm{d}x\,\mathrm{d}t \\
&= \int_Q \left\{ \frac{\partial z_{v_1,v_2}}{\partial t} - \Delta z_{v_1,v_2} + \frac{\partial f}{\partial y}(x,t,y_g) z_{v_1,v_2} + K'[y_u] z_{v_1,v_2} \right\} \varphi_u\,\mathrm{d}x\,\mathrm{d}t \\
&= -\int_Q \left\{ \frac{\partial^2 f}{\partial y^2}(x,t,y_u) z_{v_1} z_{v_2} + K''[y_u](z_{v_1}, z_{v_2}) \right\} \mathrm{d}x\,\mathrm{d}t.
\end{aligned}$$

Combining (3.11) and the above identity, (3.2) follows. □

Remark 3.6. We observe that the linear form $J'(\bar u) : L^r(0,T;L^2(\omega)) \longrightarrow \mathbb{R}$ can be extended to a continuous linear form on $L^2(Q_\omega)$ by the same expression (3.1). It is an obvious consequence of the fact that $\varphi_{u|_\omega} + \kappa u \in L^2(Q_\omega)$. The same extension is possible for the bilinear form $J''(\bar u) : L^r(0,T;L^2(\omega)) \times L^r(0,T;L^2(\omega)) \longrightarrow \mathbb{R}$. Indeed, this is consequence of the estimate (2.19) that implies

$$\begin{aligned}
&\int_Q \left| \varphi_u \frac{\partial f}{\partial y}(x,t,y_u) z_{v_1} z_{v_2} \right| \mathrm{d}x\,\mathrm{d}t \\
&\le \left\| \frac{\partial f}{\partial y}(x,t,y_u) \right\|_{L^\infty(Q)} \|\varphi_u\|_{L^\infty(0,T;L^2(\Omega))} \|z_{v_1}\|_{L^2(0;T;L^4(\Omega))} \|z_{v_1}\|_{L^2(0;T;L^4(\Omega))} \\
&\le C_u \|v_1\|_{L^2(Q_\omega)} \|v_2\|_{L^2(Q_\omega)}
\end{aligned}$$

and

$$\begin{aligned}
&\int_Q \left| \varphi_u K''[y_u](z_{v_1}, z_{v_2}) \right| \mathrm{d}x\,\mathrm{d}t \le C_a \int_0^T \int_{[0,t]} \int_\Omega |\varphi_u(x,t) z_{v_1}(x,s) z_{v_2}(x,s)|\,\mathrm{d}x\,\mathrm{d}|\mu|(s)\,\mathrm{d}t \\
&\le \|\mu\| \|\varphi_u\|_{L^1(0,T;L^2(\Omega))} \|z_{v_1}\|_{L^\infty(0;T;L^4(\Omega))} \|z_{v_2}\|_{L^\infty(0;T;L^4(\Omega))} \le C_u \|v_1\|_{L^2(Q_\omega)} \|v_2\|_{L^2(Q_\omega)}.
\end{aligned}$$

Now, we address the issue of existence of a solution for control problem (P). Since the cost functional J is not coercive on $L^\infty(0,T;L^2(\omega))$, the classical approach based on a minimizing sequence does not work to establish the existence of a solution. An alternative idea is used for the proof.

Theorem 3.7. (P) *has at least one solution* $\bar u$.

Proof. For every integer $k \ge 1$ we define the control problem

$$(\mathrm{P}_k) \qquad \min_{u \in U_k} J(u) := \frac{1}{2} \int_Q (y_u - y_d)^2\,\mathrm{d}x\,\mathrm{d}t + \frac{\kappa}{2} \int_{Q_\omega} u^2\,\mathrm{d}x\,\mathrm{d}t,$$

where $U_k = \{u \in L^\infty(0,T;L^2(\omega)) : \|u\|_{L^\infty(0,T;L^2(\omega))} \le k\}$. Using Theorem 2.2 and the fact that U_k is weakly* closed and bounded in $L^\infty(0,T;L^2(\omega))$, the existence of a solution u_k to (P_k) follows. Since the control $u \equiv 0$ belongs to U_k for every k, we have that $\frac{1}{2}\|y_{u_k} - y_d\|^2_{L^2(Q)} \le J(u_k) \le J(0)$ for every $k \ge 1$. Hence, $\{y_{u_k} - y_d\}_{k=1}^\infty$ is a bounded sequence in $L^2(Q)$. Then, from (3.7) the boundedness of $\{\varphi_{u_k}\}_{k=1}^\infty$ in $L^\infty(0,T;L^2(\Omega))$ follows. Moreover, u_k satisfies the first order optimality condition: $J'(u_k)(u - u_k) \ge 0$ for all $u \in U_k$. According to (3.1), this implies that $\int_{Q_\omega} (\varphi_k + \kappa u_k)(u - u_k)\,\mathrm{d}x\,\mathrm{d}t \ge 0$ for all $u \in U_k$ or equivalently $u_k = \mathrm{Proj}_{U_k}\big(-\frac{1}{\kappa}\varphi_{k|_\omega}\big)$. Consequently, we have that $\|u_k\|_{L^\infty(0,T;L^2(\omega))} \le \frac{1}{\kappa}\|\varphi_{u_k}\|_{L^\infty(0,T;L^2(\Omega))} \le C$ for every $k \ge 1$.

Select $k_0 > C$ and take $\bar u = u_{k_0}$. Then, $\bar u$ is a solution of (P). Indeed, let u be an arbitrary control in $L^\infty(0,T;L^2(\omega))$. If $\|u\|_{L^\infty(0,T;L^2(\omega))} \le k_0$, then we obviously have $J(\bar u) = J(u_{k_0}) \le J(u)$. If to the contrary, $\|u\|_{L^\infty(0,T;L^2(\omega))} > k_0$, we take an integer k such that $k > \|u\|_{L^\infty(0,T;L^2(\omega))}$. Let u_k be a solution of (P$_k$). Then, as proved above, we have that $\|u_k\|_{L^\infty(0,T;L^2(\omega))} \le C < k_0$. Therefore, $u_k \in U_{k_0}$ and $u \in U_k$ hold. Using the optimality of $\bar u = u_{k_0}$ and u_k we obtain: $J(\bar u) \le J(u_k) \le J(u)$. Hence, $J(\bar u) \le J(u)$ for every $u \in L^\infty(0,T;L^2(\omega))$. ☐

We continue this section by deriving the optimality conditions. Since (P) is not a convex problem, it is convenient to deal not only with global minimizers, but also with local minimizers. We will say that $\bar u$ is a local minimizer or local solution of (P) in the $L^r(0,T;L^2(\omega))$ sense with $r > \frac{4}{4-n}$ if there exists $\varepsilon > 0$ such that $J(\bar u) \le J(u)$ whenever $\|u - \bar u\|_{L^r(0,T;L^2(\omega))} \le \varepsilon$.

Theorem 3.8. *Let $\bar u$ be a local solution of* (P) *in the $L^r(0,T;L^2(\omega))$ sense with $r > \frac{4}{4-n}$. Then, there exist $\bar y \in W(0,T) \cap C(\bar Q)$ and $\bar\varphi \in L^2(0,T;H^1(\Omega)) \cap L^\infty(0,T;L^2(\Omega))$ such that*

$$\left\{\begin{array}{l} \dfrac{\partial \bar y}{\partial t} - \Delta \bar y + f(x,t,\bar y) + K[\bar y] = g + \chi_\omega \bar u \quad \text{in } Q,\\ \partial_\nu \bar y = 0 \quad \text{on } \Sigma,\ \bar y(0) = y_0 \quad \text{in } \Omega.\end{array}\right. \tag{3.12}$$

$$\left\{\begin{array}{l} -\dfrac{\partial \bar\varphi}{\partial t} - \Delta \bar\varphi + \dfrac{\partial f}{\partial y}(x,t,\bar y)\varphi + K'[\bar y]^*\bar\varphi\mu = \bar y - y_d \quad \text{in } Q,\\ \partial_n \bar\varphi = 0 \quad \text{on } \Sigma,\ \bar\varphi(T) = 0 \quad \text{in } \Omega\end{array}\right. \tag{3.13}$$

$$\bar\varphi_{|_\omega} + \kappa \bar u = 0. \tag{3.14}$$

Moreover, the inequality $J''(\bar u)v^2 \ge 0$ is fulfilled for every $v \in L^2(Q_\omega)$.

The optimality system (3.12)–(3.14) is a straightforward consequence of (3.1) and the necessary optimality conditions $J'(\bar u) = 0$. It is also well-known that a local solution must satisfy $J''(\bar u)v^2 \ge 0$ for every $v \in L^\infty(0,T;L^2(\omega))$. However, as established in Remark 3.6, $J''(\bar u)$ is a continuous bilinear form on $L^2(Q_\omega)$ and $L^\infty(0,T;L^2(\omega))$ is dense in $L^2(Q_\omega)$. Hence, the inequality $J''(\bar u)v^2 \ge 0$ also holds for every $v \in L^2(Q_\omega)$.

The next theorem establishes a sufficient condition for local optimality.

Theorem 3.9. *Let $\bar u \in L^\infty(0,T;L^2(\omega))$ satisfy the first order optimality conditions* (3.12)–(3.14) *and the second order condition $J''(\bar u)v^2 > 0$ for every $v \in L^2(Q_\omega) \setminus \{0\}$. Then, for every $r > \frac{4}{4-n}$ there exist $\varepsilon > 0$ and $\delta > 0$ such that*

$$J(\bar u) + \frac{\delta}{2}\|u - \bar u\|^2_{L^2(Q_\omega)} \le J(u) \quad \text{if} \quad \|u - \bar u\|_{L^r(0,T;L^2(\omega))} \le \varepsilon. \tag{3.15}$$

Proof. We argue by contradiction. If the statement of the theorem is false, then for every integer $k \ge 1$ there exists a control u_k such that

$$\|u_k - \bar u\|_{L^r(0,T;L^2(\omega))} < \frac{1}{k} \quad \text{and} \quad J(u_k) < J(\bar u) + \frac{1}{2k}\|u_k - \bar u\|^2_{L^2(Q_\omega)}. \tag{3.16}$$

Let us set

$$\rho_k = \|u_k - \bar u\|_{L^2(Q_\omega)} \quad \text{and} \quad v_k = \frac{1}{\rho_k}(u_k - \bar u). \tag{3.17}$$

Since $\|v_k\|_{L^2(Q_\omega)} = 1$, we can take a subsequence, still denoted by itself, such that $v_k \rightharpoonup v$ in $L^2(Q_\omega)$. Then we can perform a Taylor expansion and use that $J'(\bar u) = 0$ to get

$$J(u_k) = J(\bar u) + \frac{1}{2} J''(\bar u + \theta_k(u_k - \bar u))(u_k - \bar u)^2.$$

This equality along with (3.16) and (3.17) leads to $J''(\bar u + \theta_k(u_k - \bar u))v_k^2 < \frac{1}{k}$. We denote $\hat u_k = \bar u + \theta_k(u_k - \bar u)$, $\hat y_k = G(\hat u_k)$, $z_k = G'(\hat u_k)v_k$, and φ_k the adjoint state corresponding to $\hat u_k$. Then, recalling (3.2), the above inequality can be written

$$\int_Q \left\{ \Big[1 - \varphi_k \frac{\partial^2 f}{\partial y^2}(x,t,\hat y_k)\Big] z_k^2 - \varphi_k K''[\hat y_k](z_k, z_k) \right\} \mathrm{d}x\,\mathrm{d}t + \kappa \int_{Q_\omega} v_k^2 \,\mathrm{d}x\,\mathrm{d}t < \frac{1}{k}. \tag{3.18}$$

From (3.16) we get that $\hat u_k \to \bar u$ in $L^r(0,T;L^2(\omega))$, therefore $\hat y_k = G(\hat u_k) \to G(\bar u) = \bar y$ in $W(0,T) \cap C(\bar Q)$. This convergence implies that $\varphi_k \to \bar\varphi$ in $L^2(0,T;H^1(\Omega)) \cap L^\infty(0,T;L^2(\Omega))$. Using the compactness of the embedding $H^1(Q) \subset L^2(0,T;L^4(\Omega))$, see Proposition III-1.3 of [23], and Theorem 2.4 we infer that $z_k \to z_v = G'(\bar u)v$ in $L^2(0,T;L^4(\Omega))$. Using all these convergence properties it is easy to pass to the limits in (3.18) and to get

$$\begin{aligned} J''(\bar u)v^2 \le \lim_{k\to\infty} \int_Q \left\{ \Big[1 - \varphi_k \frac{\partial^2 f}{\partial y^2}(x,t,\hat y_k)\Big] z_k^2 - \varphi_k K''[\hat y_k](z_k, z_k) \right\} \mathrm{d}x\,\mathrm{d}t \\ + \kappa \liminf_{k\to\infty} \int_{Q_\omega} v_k^2 \,\mathrm{d}x\,\mathrm{d}t \le 0. \end{aligned}$$

Due to the assumption $J''(\bar u)v^2 > 0$ if $v \in L^2(Q_\omega) \setminus \{0\}$, we infer that $v = 0$. Hence, we have $z_k \to 0$ in $L^2(0,T;L^4(\Omega))$. Then, passing to the limits in (3.18) and using that $\|v_k\|_{L^2(Q_\omega)} = 1$ we obtain $\kappa \le 0$, which contradicts our assumption on κ. ☐

The presence of the Tikhonov regularizing term $\frac{\kappa}{2}\|u\|^2_{L^2(Q_\omega)}$ is crucial in the proof of the above theorem. When $\kappa = 0$, the second order analysis is more complicate; see [3, 6].

4. Concluding remarks

We have presented a general theory of optimal control problem for a class of semilinear parabolic equations with a possibly super-linear nonlinearity and with a memory term governed by a general memory kernel and general real valued Borel measure. Here are some remarks that we would like to collect.

– The appearance of a memory term of the form (1.2) makes the well-posedness of the state equation, as well as that of the adjoint equations, technically difficult. A careful analysis involving the Bochner integral and some delicate regularity results for parabolic equations help us to overcome the difficult.

– Due to the density of $L^r(0,T;L^2(\omega))$ with $r > \frac{4}{4-n}$ in $L^2(Q_\omega)$ and recalling Remark 3.6, the sufficient second order condition for local optimality $J''(\bar u)v^2 > 0$ for all $v \in L^2(Q_\omega) \setminus \{0\}$ is equivalent to $J''(\bar u)v^2 > 0$ for all $v \in L^r(0,T;L^2(\omega)) \setminus \{0\}$. Moreover, this condition is still equivalent to

$$\exists \delta > 0 \ \text{ such that } \ J''(\bar u)v^2 \ge \delta \|v\|^2_{L^2(Q_\omega)} \quad \forall v \in L^2(Q_\omega).$$

Indeed, it is obvious that this condition implies that $J''(\bar u)v^2 > 0$ for all $v \in L^2(Q_\omega) \setminus \{0\}$. To prove the converse implication we use (3.15). We define $I : L^r(0,T;L^2(\omega)) \longrightarrow \mathbb{R}$ by $I(u) = J(u) - \frac{\delta}{2}\|u - \bar u\|^2_{L^2(Q_\omega)}$. From (3.15) we infer that $I(\bar u) \le I(u)$ if $\|u - \bar u\|_{L^r(0,T;L^2(\omega))} \le \varepsilon$. Hence, $\bar u$ is a local solution of I and, consequently, as in Theorem 3.8 we have that $J''(\bar u)v^2 - \delta\|v\|^2_{L^2(Q_\omega)} = I''(\bar u)v^2 \ge 0$ for every $v \in L^2(Q_\omega)$. Therefore, (P) enjoys the so-called two-norm discrepancy: the functional J is differentiable with respect to the (stronger) norm of

$L^r(0,T;L^2(\omega))$, but the sufficient condition $J''(\bar{u})v^2 \geq \delta\|v\|^2_{L^2(Q_\omega)}$ holds for a different (weaker) norm. The reader is referred to [8] for additional comments on this issue.

– It is possible to include control constraints in the control problem such as

$$U_{ad} = \{u \in L^2(0,T;L^2(\omega)) : u(t) \in K_{ad}\},$$

where K_{ad} is a closed, convex, and bounded subset of $L^2(\omega)$. In this case, the existence of an optimal control and the first order optimality conditions can be easily obtained following the approach of this paper with obvious modifications. For the choices

$$K_{ad} = \{v \in L^2(\omega) : \|v\|_{L^2(\omega)} \leq \gamma\},\ 0 < \gamma < \infty,$$
$$K_{ad} = \{v \in L^2(\omega) : \alpha \leq v(x) \leq \beta \text{ for a.a. } x \in \omega\},\ -\infty < \alpha < \beta < \infty,$$

the second order analysis can be performed by using the techniques of [5].

References

[1] P. Cannarsa, H. Frankowska and E.M. Marchini, Optimal control for evolution equations with memory. *J. Evol. Equ.* **13** (2013) 197–227.

[2] P. Cannarsa and D. Sforza, Global solutions of abstract semilinear parabolic equations with memory terms. *Nonlinear Diff. Equ. Appl.* **10** (2003) 399–430.

[3] E. Casas, Second order analysis for bang-bang control problems of PDEs. *SIAM J. Control Optim.* **50** (2022) 2355–2372.

[4] E. Casas and K. Kunisch, Optimal control of semilinear parabolic equations with non-smooth pointwise-integral control constraints in time-space. *Appl. Math. Optim.* **85** (2022).

[5] E. Casas and K. Kunisch, Infinite horizon optimal control for a general class of semilinear parabolic equations (2022) Submitted.

[6] E. Casas and M. Mateos, Critical cones for sufficient second order conditions in PDE constrained optimization. *SIAM J. Optim.* **30** (2020) 585–603.

[7] E. Casas, C. Ryll and F. Tröltzsch, Optimal control of a class of reaction-diffusion systems. *Comput. Optim. Appl.* **70** (2018) 677–707.

[8] E. Casas and F. Tröltzsch, Second order analysis for optimal control problems: improving results expected from abstract theory. *SIAM J. Optim.* **22** (2012) 261–279.

[9] E. Casas and F. Tröltzsch, Second order optimality conditions for weak and strong local solutions of parabolic optimal control problems. *Vietnam J. Math.* **44** (2016) 181–202.

[10] A. Constantin and S. Peszat, Global existence of soutions of semilinear parabolic evolution equations. *Diff. Int. Equ.* **13** (2000) 99–114.

[11] K. Disser, A.F.M. ter Elst and J. Rehberg, Hölder estimates for parabolic operators on domains with rough boundary. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **17** (2017) 65–79.

[12] K. Ezzinbi and P. Ndambomve, Solvability of some partial functional integrodifferential equations with finite delay and optimal controls in Banach spaces. *SpringerPlus* **5** (2016) 1264.

[13] Z.B. Fang and L Qiu, Global existence and uniform energy decay rates for the semilinear parabolic equation with a memory term and mixed boundary condition. *Abst. Appl. Anal.* (2013) 532935 https://doi-org.unican.idm.oclc.org/10.1155/2013/532935.

[14] P. Girault and P.A. Raviart, *Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms.* Springer-Verlag, Berlin, Heidelberg, New York, Tokio (1986).

[15] M.E. Gurtin and A.C. Pipkin, A general theory of heat conduction with finite wave speeds. *Arch. Ratl. Mech. Anal.* **31** (1968) 113–126.

[16] D. Henry, Geometric Theory of Semilinear Parabolic Equations. Lecture Notes in Math., Vol. 840, Springer-Verlag (1981).

[17] B. Hu, Blow-up Theories for Semilinear Parabolic Equation. Springer (2011).

[18] S. Larsson, Semilinear parabolic partial differential equations: theory, approximation, and application. *New Trends in the Mathematical and Computer Sciences,* 153-194, Publ. ICMCS, 3, Int. Cent. Math. Comp. Sci. (ICMCS), Lagos (2006).

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

[20] J.W. Nunziato, On heat conduction in materials with memory. *Quart. Appl. Math.* (1971) 187–204.

[21] W. Rudin, Real and Complex Analysis. McGraw-Hill Book Co., London (1970).

[22] J. Shi, C. Wang and H. Wang, Diffusive spatial movement with memory and maturation delays. *Nonlinearuy* **32** (2019) 3188–3208.

[23] R.E. Showalter, Monotone Operators in Banach Space and Nonlinear Partial Differential Equations, volume 49 of *Math. Surv. and Monogr.* American Mathematical Society, Providence, RI (1997).
[24] D. Sofiane, B. Abdelfatah and O. Taki-Eddine, Study of solution for a parabolic intergodifferential equation with the second kind integral condition. *Int. J. Anal. Appl.* **16** (2018) 569–593.
[25] P.G. Surkov, Tracking of solutions to parabolic equations with memory in a general class of controls. *Russian Math.* **60** (2016) 44–54.
[26] N.D. Toan, Optimal control of nonclassical diffusion equations with memory. *Acta Appl. Math.* **169** (2020) 533–558.
[27] J. Yong and X. Zhang, Heat equations with memory. *Nonlinear Anal.* **63** (2005) e99–e108.