**ESAIM: Control, Optimisation and Calculus of Variations** Will be set by the publisher
URL: http://www.emath.fr/cocv/

# PARABOLIC CONTROL PROBLEMS IN SPACE-TIME MEASURE SPACES *

Eduardo Casas[1] and Karl Kunisch[2]

**Abstract**. Optimal control problems in measure spaces governed by parabolic equations with are considered. The controls appear as spatial measure in the initial condition and as space-time measures as forcing functions. First order optimality conditions are derived and certain structural properties, in particular sparsity, are discussed. An framework for approximation if these highly irregular problems is also proposed.

**1991 Mathematics Subject Classification.** 90C48, 49J52,49K20.

.

## 1. Introduction

In this paper we study optimal control problems for parabolic equations where the space-time controls appear as volume sources and also as measure valued initial conditions. In particular, we consider the following problem:

$$\text{(P)} \quad \min_{(u,u_0)\in\mathcal{M}(Q_c)\times\mathcal{M}(\Omega)} J(u,u_0) = \frac{1}{q}\|y-y_d\|^q_{L^q(Q)} + \alpha\|u\|_{\mathcal{M}(Q_c)} + \beta\|u_0\|_{\mathcal{M}(\Omega)}, \tag{1.1}$$

where $y$ is the solution of the problem

$$\left\{\begin{array}{rcll} \dfrac{\partial y}{\partial t} - \Delta y &=& u & \text{in } Q = \Omega\times(0,T)\\ y(x,0) &=& u_0 & \text{in } \Omega\\ y(x,t) &=& 0 & \text{on } \Sigma = \Gamma\times(0,T). \end{array}\right. \tag{1.2}$$

Here $Q_c = \omega \times I$, where $\omega \subset \Omega$ is the control domain, the subinterval $I$ of $(0,T)$, is the control horizon, and $\mathcal{M}(Q_c)$ and $\mathcal{M}(\Omega)$ denote measure spaces. More details on the notation and the variational solution concept to (1.2) will be given in the following section. For results of this paper the Laplacian can be replaced by a second order elliptic operator with regular coefficients.

*Keywords and phrases:* space-time measure controls, optimal control, sparsity, parabolic equations, first order optimality conditions, numerical approximation

* *The first author was supported by Spanish Ministerio de Economía y Competitividad under project MTM2011-22711. The second was supported by the Austrian Science Fund (FWF) under grant SFB F32 (SFB "Mathematical Optimization and Applications in Biomedical Sciences")*

[1] Eduardo Casas. 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).

[2] Karl Kunisch. Institute for Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, A-8010 Graz, Austria (karl.kunisch@uni-graz.at), and Radon Institute, Austrian Academy of Sciences.

The importance of measure valued controls is by now well-established. The solution of measure-valued optimal control problems have the structural property of sparsity. This property can be used for formulating the problem of optimal actuator placement or for source identification problems. Formally, these properties could also be achieved by an $L^1$ approach. This space, however, does not allow an appropriate topology for compactness arguments to guarantee existence of solutions to (1.1). The papers [4, 9] may have been the first ones to address measure valued control problems with the goal of sparsity for linear elliptic control systems. The approach was extended to semi-linear elliptic equations in [6]. A-priori error estimates for finite element approximation of linear elliptic optimal control problems with measure valued controls were investigated in [16]. The parabolic case was considered in [5] and [13] with different measure valued topologies to enhance directional spatial sparsity. The terminology directional sparsity was introduces in [12]. In the present paper we succeed in establishing an analytical framework which allows to consider measure-valued controls on the space-time cylinder. Let us mention that the difficulty of dealing with existence in the presence of $L^1$ controls can also be addressed by utilizing either constraints on the controls or regularization terms in a finer norm than the $L^1$ norm which then allows to use weak or weak* convergence arguments. Finally we mention the recent paper [7] which also uses measure-valued controls in the context of approximate controllability into an $L^2(\Omega)$ ball. To compensate for the lack of sufficient regularity of the trajectories the controls only act on a subset of the full time horizon.

The plan of the paper is as follows. In the following section we address well-posedness of the state equation (1.2) and existence of solutions to (1.1). The optimality system and sparsity properties of the solution are analyzed in Section 3. An approximation framework to these highly irregular problems is developed in Section 4, where strong (subsequential) convergence of the discrete optimal trajectories and weak* convergence of the discrete optimal controls to optimal trajectories and optimal controls of the continuous problem is proved. In the final section we consider the case when the observation is only assumed to be available in a sub-cylinder $\Omega_o \times I_0$ of $\Omega \times (0, T)$ and analyze the support of the optimal controls relative to the location of $\Omega_o \times I_0$ and $\Omega \times I$.

## 2. Assumptions and existence of solutions

The following notation will be utilized through this paper. By $\Omega$ we denote an open bounded domain in $\mathbb{R}^d$, for $d \in \{1, 2, 3\}$, with a Lipschitz boundary $\Gamma$. Let us set $Q_c = \omega \times I$, where $\omega$ is a relatively closed domain of $\Omega$ and $I$ is an interval relatively closed in $(0, T)$ for some $T > 0$ given. With $\mathcal{M}(Q_c)$ and $\mathcal{M}(\Omega)$ we denote the spaces of real and regular Borel measures in $Q_c$ and $\Omega$, respectively. The appearance of $u \in \mathcal{M}(Q_c)$ in the state equation (1.2) is understood as an extension to $Q$ by zero outside $Q_c$. We assume that $\alpha > 0$, $\beta > 0$ and $1 \le q < \min\{2, \frac{d+2}{d}\}$. We also assume that $y_d \in L^q(Q)$. Under these assumptions we shall prove that (P) is well defined and it has a unique solution. To this end, we first analyze the state equation (1.2).

**Definition 2.1.** *We say that a function $y \in L^1(Q)$ is a solution of (1.2) if the following identity holds*

$$\int_Q -(\frac{\partial \phi}{\partial t} + \Delta\phi) y \, dxdt = \int_{Q_c} \phi \, du + \int_\Omega \phi(0) \, du_0, \quad \forall \phi \in \Phi, \tag{2.1}$$

*where*

$$\Phi = \{\phi \in L^2(0, T; H_0^1(\Omega)) : \frac{\partial \phi}{\partial t} + \Delta\phi \in L^\infty(Q) \quad and \quad \phi(x, T) = 0 \ in \ \Omega\}.$$

Let us observe that the problem

$$\left\{ \begin{array}{rcll} \dfrac{\partial \phi}{\partial t} + \Delta\phi & = & f & \text{in } Q \\ \phi(x, T) & = & 0 & \text{in } \Omega \\ \phi(x, t) & = & 0 & \text{on } \Sigma \end{array} \right. \tag{2.2}$$

has a unique solution $\phi \in L^2(0,T;H_0^1(\Omega)) \cap C([0,T];L^2(\Omega))$ for every $f \in L^\infty(Q)$. Moreover, the regularity $\phi \in C(\bar{Q})$ holds. This continuity property follows from the results in [1]; see [3, Theorem 5.1]. The reader is also referred to [17].

**Theorem 2.2.** *There exists a unique solution $y$ of (1.2). Moreover, $y \in L^r(0,T;W_0^{1,p}(\Omega))$ for all $p, r \in [1,2)$ with $(2/r) + (d/p) > d + 1$, and the following estimate holds*

$$\|y\|_{L^r(0,T;W_0^{1,p}(\Omega))} \leq C_{r,p}(\|u\|_{\mathcal{M}(Q_c)} + \|u_0\|_{\mathcal{M}(\Omega)}). \tag{2.3}$$

*Proof.* The uniqueness is an immediate consequence of (2.1) and the fact that the mapping $\frac{\partial}{\partial t} + \Delta : \Phi \longrightarrow L^\infty(Q)$ is surjective; see (2.2). To prove the existence and the regularity we choose two sequences $\{u_k\}_k \subset C(\bar{Q}_c)$ and $\{u_{0k}\}_k \subset C(\bar{\Omega})$ such that $u_k \stackrel{*}{\rightharpoonup} u$ in $\mathcal{M}(Q_c)$ and $u_{0k} \stackrel{*}{\rightharpoonup} u_0$ in $\mathcal{M}(\Omega)$, and $\|u_k\|_{L^1(Q_c)} \leq \|u\|_{\mathcal{M}(Q_c)}$ and $\|u_{0k}\|_{L^1(\Omega)} \leq \|u_0\|_{\mathcal{M}(\Omega)}$. This can be achieved by taking the convolution with sequences of mollifiers. Associated with $(u_k, u_{0k})$ we define the sequence of solutions $\{y_k\}_k \subset L^2(0,T;H_0^1(\Omega))$ of (1.2). Then, using the regularity of $y_k$ we can make integration by parts to obtain for every $\phi \in \Phi$

$$\int_Q -(\frac{\partial \phi}{\partial t} + \Delta\phi) y_k \, dxdt = \int_Q (\frac{\partial y_k}{\partial t} + \Delta y_k)\phi \, dxdt + \int_\Omega \phi(0) u_{0k} \, dx$$

$$= \int_{Q_c} \phi u_k \, dx \, dt + \int_\Omega \phi(0) u_{0k} \, dx. \tag{2.4}$$

Let us obtain the estimates (2.3) for $y_k$. To this end, we take $\{\psi_j\}_{j=0}^d \subset \mathcal{D}(Q)$ and take $\phi \in \Phi$ satisfying

$$\left\{\begin{array}{rcll} \dfrac{\partial \phi}{\partial t} + \Delta \phi & = & \psi_0 - \dfrac{\partial \psi_j}{\partial x_j} & \text{in } Q \\ \phi(x,T) & = & 0 & \text{in } \Omega \\ \phi(x,t) & = & 0 & \text{on } \Sigma. \end{array}\right. \tag{2.5}$$

Following [1] and [3, Theorem 5.1], we know that there exists a constant $C$ such that

$$\|\phi\|_{C(\bar{Q})} \leq C \sum_{j=0}^d \|\psi_j\|_{L^{r'}(0,T;L^{p'}(\Omega))}. \tag{2.6}$$

Now, using distributional derivatives we obtain from (2.4)-(2.6)

$$\langle y_k, \psi_0 \rangle + \sum_{j=1}^d \langle \partial_{x_j} y_k, \psi_j \rangle = \int_Q y_k(\psi_0 - \sum_{j=0}^d \partial_{x_j} \psi_j) \, dx \, dt$$

$$= \int_Q (\frac{\partial \phi}{\partial t} + \Delta\phi) y_k \, dxdt = -\int_{Q_c} \phi u_k \, dx \, dt - \int_\Omega \phi(0) u_{0k} \, dx$$

$$\leq C(\|u_k\|_{L^1(Q_c)} + \|u_{0k}\|_{L^1(\Omega)}) \sum_{j=0}^d \|\psi_j\|_{L^{r'}(0,T;L^{p'}(\Omega))}$$

$$\leq C(\|u\|_{\mathcal{M}(Q_c)} + \|u_0\|_{\mathcal{M}(\Omega)}) \sum_{j=0}^d \|\psi_j\|_{L^{r'}(0,T;L^{p'}(\Omega))}.$$

This proves that $\{y_k\}_k \subset L^r(0,T;W_0^{1,p}(\Omega))$ and every $y_k$ satisfies (2.3). Finally, by taking a subsequence, we deduce the existence of $y \in L^r(0,T;W_0^{1,p}(\Omega))$ such that $y_k \rightharpoonup y$ in this space. Then, passing to the limit in (2.4) we obtain that $y$ satisfies (2.1), as well as (2.3). □

**Remark 2.3.** *The notion of solution given here differs from definitions used in [3] or [7]; see also [2]. In those papers, the solution was assumed to belong to $L^r(0,T;W_0^{1,p}(\Omega))$ from the beginning. However, here our solution is supposed to belong just to $L^1(Q)$. This is convenient for the numerical analysis that will be carried out later. In the case of parabolic equations with regular coefficients both definitions coincide.*

**Remark 2.4.** *For $d=2$ or 3 the condition $1 \le q < \frac{d+2}{d}$ was required in the formulation of the cost-functional. If we take*

$$\tilde{p} = \frac{dq}{dq+q-2},$$

*then $(2/q)+(d/\tilde{p}) = d+1$ and $\tilde{p} \in (1, \frac{d}{d-1}]$ are satisfied. From Sobolev's embedding we also have $W_0^{1,\tilde{p}}(\Omega) \subset L^{(d+2)/d}(\Omega)$. Therefore, for any $q < \min(2, \frac{2+d}{d})$ there exists some $p$ such that*

$$1 \le p < \tilde{p}, \quad \frac{2}{q} + \frac{d}{p} > d+1, \quad \textit{and} \quad W_0^{1,p}(\Omega) \subset L^q(\Omega) \textit{ compactly.} \tag{2.7}$$

*Hence, Theorem 2.2 states that the solution $y$ of (1.2) belongs to $L^q(0,T;W_0^{1,p}(\Omega)) \subset L^q(Q)$. This motivates the choice of $q$ in the cost functional.*

*In dimension $d=1$, the condition on $q$ is $1 \le q < 2$, and there always exists $p > 1$ such that $(2/q)+(1/p) > 2$. For any such $p$ we have $L^q(0,T;W_0^{1,p}(\Omega)) \subset L^q(Q)$ and the compact embedding $W_0^{1,p}(\Omega) \subset L^q(\Omega)$.*

**Remark 2.5.** *Since the solutions of (1.2) belong to $L^q(Q)$, the density of $L^\infty(Q)$ in $L^{q'}(Q)$ implies that the identity (2.1) is valid for every $\phi$ in the space*

$$\Phi_q = \{\phi \in L^2(0,T;H_0^1(\Omega)) : \frac{\partial \phi}{\partial t} + \Delta\phi \in L^{q'}(Q) \quad \textit{and} \quad \phi(x,T) = 0 \textit{ in } \Omega\}.$$

*Indeed, first we observe that $q' > 2$ and $q' > 1 + \frac{d}{2}$, which follows from the inequality $q < \min\{2, \frac{d+2}{d}\}$. Then for any $g \in L^{q'}(Q)$, there exists a sequence $\{g_n\} \subset L^\infty(Q)$ with $g_n \to g$ in $L^{q'}(Q)$. As a consequence of the regularity results in [1] and [14] there exists a sequence $\{\phi_n\} \subset \Phi_q$ and $\phi \in \Phi$ with $\frac{\partial \phi_n}{\partial t} + \Delta\phi_n = g_n$, $\frac{\partial \phi}{\partial t} + \Delta\phi = g$, and $\lim_n \phi_n = \phi$ in $C(\bar{Q})$. Passing to the limit in (2.2) with $\phi$ replaced by $\phi_n$ implies that $\phi \in \Phi_g$.*

Before proving the existence of an optimal control for (P), let us establish a technical lemma that will be useful later.

**Lemma 2.6.** *Let $\{(u_k, u_{0k})\}_k \subset L^1(Q_c) \times L^1(\Omega)$ be a weakly* convergence sequence in $\mathcal{M}(Q_c) \times \mathcal{M}(\Omega)$ to an element $(u, u_0)$. Then, the associated states $\{y_k\}_k$ converge strongly in $L^q(Q)$ to the state $y$ corresponding to $(u, u_0)$ for every $1 \le q < \frac{d+2}{d}$.*

*Proof.* Let us take $p$ as in the previous remark. Then, from (2.3) we get that $y_k \rightharpoonup y$ in $L^q(0,T;W_0^{1,p}(\Omega))$ and $\{\partial_t y_k\}_k$ is bounded in $L^1(0,T;W^{-1,p}(\Omega)) = L^1(0,T;W_0^{1,p'}(\Omega)^*)$. Indeed, obviously $\{\Delta y_k\}_k$ is bounded in $L^q(0,T;W^{-1,p}(\Omega))$. In addition, we observe that the weak$^*$ convergence of $\{u_k\}_k$ implies its boundedness in $L^1(Q)$. Now, we check that $L^1(Q) \subset L^1(0,T;W^{-1,p}(\Omega))$. For this purpose it is enough to show that $W_0^{1,p'}(\Omega) \subset C(\bar{\Omega})$ which holds if $p' > d$. The latter is is obvious for $d=1$. For dimensions $d=2$ and $d=3$ we have that $p < 2$ and $p < 3/2$, respectively, which leads to $p' > d$. Finally, we have that $W_0^{1,p}(\Omega) \subset L^q(\Omega) \subset W^{-1,p}(\Omega)$, where the first inclusion is compact and the second is continuous. Then, from [19, Corollary 4], we deduce the strong convergence $y_k \to y$ in $L^q(Q)$. □

We conclude this section by studying the existence of solutions for the control problem (P).

**Theorem 2.7.** *Problem* (P) *has at least one solution* $(\bar{u}, \bar{u}_0) \in \mathcal{M}(Q_c) \times \mathcal{M}(\Omega)$ *for every* $1 \le q < \min\{2, \frac{d+2}{d}\}$*. Furthermore, if* $q > 1$ *then the solution is unique.*

*Proof.* Let $\{(u_k, u_{0k})\}_k$ be a minimizing sequence. By the coercivity of $J$ we can obtain a weak$^*$ convergence subsequence, denoted in the same way, with limit $(\bar{u}, \bar{u}_0)$. By Theorem 2.2 we get that the sequence of states $\{y_k\}_k$ associated to $\{(u_k, u_{0k})\}_k$ converges to $\bar{y} = y(\bar{u}, \bar{u}_0)$ weakly in $L^r(0, T; W_0^{1,p}(\Omega))$ for every $(2/r) + (d/p) > d + 1$. Now, (2.7) implies that $y_k \rightharpoonup \bar{y}$ in $L^q(Q)$. Hence, $J(\bar{u}, \bar{u}_0) \le \liminf_{k\to\infty} J(u_k, u_{0k}) = \inf\,(\mathrm{P})$ holds. The uniqueness follows from the strict convexity of $J$ for $q > 1$ and the injectivity of the mapping $(u, u_0) \to y$ from $\mathcal{M}(Q_c) \times \mathcal{M}(\Omega)$ to $L^q(Q)$. Here we use that the observation is taken on the whole domain. □

**Remark 2.8.** *Let us observe that the existence of a solution to* (P) *can also be obtained for arbitrary* $q \ge \min\{2, \frac{d+2}{d}\}$*. The only change in the argument of the above proof is that* $y_k \rightharpoonup \bar{y}$ *in* $L^q(Q)$ *is obtained from the boundedness of* $\{J(u_k, u_{0k})\}_k$ *and the fact that* $\frac{1}{q}\|y_k - y_d\|^q_{L^q(Q)} \le J(u_k, u_{0k})$*. Our assumption on the parameter* $q$ *will be needed in the following section devoted the necessary optimality conditions.*

## 3. Optimality Conditions

In this section, we state the optimality conditions satisfied by a solution $(\bar{u}, \bar{u}_0)$ of (P) and discuss its sparsity structure. Let us fix some notation. By $S : \mathcal{M}(Q_c) \times \mathcal{M}(\Omega) \longrightarrow L^q(Q)$ we denote the solution operator of (1.2). We write the cost functional in the form

$$J(u, u_0) = (F \circ S)(u, u_0) + \alpha j_Q(u) + \beta j_\Omega(u_0),$$

where

$$\begin{aligned} &F : L^q(Q) \longrightarrow \mathbb{R}, \quad F(y) = \frac{1}{q}\|y - y_d\|^q_{L^q(Q)} \\ &j_Q : \mathcal{M}(Q_c) \longrightarrow \mathbb{R}, \quad j_Q(u) = \|u\|_{\mathcal{M}(Q_c)} \\ &j_\Omega : \mathcal{M}(\Omega) \longrightarrow \mathbb{R}, \quad j_\Omega(u_0) = \|u_0\|_{\mathcal{M}(\Omega)}. \end{aligned}$$

We set as usual

$$\mathrm{sign}(s) = \begin{cases} \{+1\} & \text{if } s > 0 \\ \{-1\} & \text{if } s < 0 \\ [-1, +1] & \text{if } s = 0. \end{cases}$$

Further, we observe that for $q > 1$ the mapping $F$ is of class $C^1$ and for $q = 1$ the subdifferential of $F$ is given by

$$\partial F(y) = \{g \in L^\infty(Q) : g(x,t) \in \mathrm{sign}(y(x,t) - y_d(x,t)) \ \text{ a.e.}\}. \tag{3.1}$$

**Theorem 3.1.** *Let* $(\bar{u}, \bar{u}_0)$ *denote a solution to* (P) *with associated state* $\bar{y}$. *Then, there exists an element* $\bar{\varphi} \in L^2(0,T;H_0^1(\Omega)) \cap C(\bar{Q})$ *satisfying*

$$\begin{cases} -\dfrac{\partial\bar{\varphi}}{\partial t} - \Delta\bar{\varphi} &= \bar{g} \quad \text{in } Q \\ \bar{\varphi}(x,T) &= 0 \quad \text{in } \Omega \\ \bar{\varphi}(x,t) &= 0 \quad \text{on } \Sigma, \end{cases} \tag{3.2}$$

$$\begin{cases} \displaystyle\int_{Q_c} \bar{\varphi}\, d\bar{u} + \alpha\|\bar{u}\|_{\mathcal{M}(Q_c)} = 0 \\ \|\bar{\varphi}\|_{C(\bar{Q}_c)} \begin{cases} = \alpha & \text{if } \bar{u} \neq 0 \\ \leq \alpha & \text{if } \bar{u} = 0, \end{cases} \end{cases} \tag{3.3}$$

$$\begin{cases} \displaystyle\int_{\Omega} \bar{\varphi}(0)\, d\bar{u}_0 + \beta\|\bar{u}_0\|_{\mathcal{M}(\Omega)} = 0 \\ \|\bar{\varphi}(0)\|_{C(\bar{\Omega})} \begin{cases} = \beta & \text{if } \bar{u}_0 \neq 0 \\ \leq \beta & \text{if } \bar{u}_0 = 0, \end{cases} \end{cases} \tag{3.4}$$

*where*

$$\bar{g}(x,t) \begin{cases} = |\bar{y}(x,t) - y_d(x,t)|^{q-2}(\bar{y}(x,t) - y_d(x,t)) & \text{if } 1 < q < \min\{2, \frac{d+2}{d}\} \\ \in \operatorname{sign}(\bar{y}(x,t) - y_d(x,t)) & \text{if } q = 1. \end{cases} \tag{3.5}$$

*Furthermore,* $\bar{\varphi}$ *is unique if* $q > 1$.

*Proof.* First we consider the case $q > 1$. In this case we can compute the derivative of the mapping $F \circ S$. Given $(u, u_0) \in \mathcal{M}(Q_c) \times \mathcal{M}(\Omega)$, we denote $y = S(u, u_0)$. Then, we have

$$\langle (F \circ S)'(\bar{u}, \bar{u}_0), (u, u_0)\rangle = \langle S^* F'(\bar{y}), (u, u_0)\rangle = \langle F'(\bar{y}), S(u, u_0)\rangle = \int_Q \bar{g} y\, dx\, dt, \tag{3.6}$$

where $\bar{g}$ is given by (3.5). Since $\bar{g} \in L^{q'}(Q)$, to follows from Remark 2.5 that there exists a unique solution $\bar{\varphi}$ of (3.2), that additionally satisfies the identity (2.1). From there and (3.6) we get

$$\langle (F \circ S)'(\bar{u}, \bar{u}_0), (u, u_0)\rangle = \int_{Q_c} \bar{\varphi}\, du + \int_{\Omega} \bar{\varphi}(0)\, du_0. \tag{3.7}$$

Now, using the optimality of $(\bar{u}, \bar{u}_0)$, the convexity of $j_Q$ and $j_\Omega$, and the differentiability of $F \circ S$, we get

$$0 \leq \limsup_{\rho \to 0} \frac{1}{\rho}[J(\bar{u} + \rho(u - \bar{u}), \bar{u}_0 + \rho(u_0 - \bar{u}_0))) - J(\bar{u}, \bar{u}_0)]$$

$$\leq \langle (F \circ S)'(\bar{u}, \bar{u}_0), (u - \bar{u}, u_0 - \bar{u}_0)\rangle + \alpha[j_Q(u) - j_Q(\bar{u})] + \beta[j_\Omega(u_0) - j_\Omega(\bar{u}_0)].$$

By inserting (3.7) in this expression we infer $\forall (u, u_0) \in \mathcal{M}(Q_c) \times \mathcal{M}(\Omega)$

$$-\int_{Q_c} \bar{\varphi}\, d(u - \bar{u}) - \int_{\Omega} \bar{\varphi}\, d(u_0 - \bar{u}_0) + \alpha j_Q(\bar{u}) + \beta j_\Omega(\bar{u}_0) \leq \alpha j_Q(u) + \beta j_\Omega(u_0). \tag{3.8}$$

In the case $q = 1$, we use the convexity and continuity of the three functionals defining $J$ and the rules of the subdifferential calculus to get

$$0 \in \partial J(\bar{u}, \bar{u}_0) = \partial(F \circ S)(\bar{u}, \bar{u}_0) + \partial[\alpha j_Q(\bar{u}) + \beta j_\Omega(\bar{u}_0)]$$

$$\subset S^* \partial F(\bar{y}) + \partial[\alpha j_Q(\bar{u}) + \beta j_\Omega(\bar{u}_0)].$$

Hence, we deduce the existence of $\bar g \in \partial F(\bar y)$ such that $-S^*\bar g \in \partial[\alpha j_Q(\bar u) + \beta j_\Omega(\bar u_0)]$, or equivalently $\forall (u, u_0) \in \mathcal{M}(Q_c) \times \mathcal{M}(\Omega)$

$$-\langle \bar g, S(u - \bar u, u_0 - \bar u_0)\rangle + \alpha j_Q(\bar u) + \beta j_\Omega(\bar u_0) \le \alpha j_Q(u) + \beta j_\Omega(u_0).$$

From (3.1) we have that $\bar g \in L^\infty(Q)$ and it is given by (3.5). Taking $\bar\varphi$ in the space $L^2(0,T;H^1_0(\Omega)) \cap C(\bar Q)$ solution of (3.2) and using (2.1), we get again (3.8) from the above inequality.

Finally, (3.3) and (3.4) follow from (3.8). We will prove (3.3), the proof of (3.4) being analogous. Let us take $u_0 = \bar u_0$ in (3.8), then we have

$$-\int_{Q_c} \bar\varphi\, d(u - \bar u) + \alpha j_Q(\bar u) \le \alpha j_Q(u) \quad \forall u \in \mathcal{M}(Q_c).$$

Taking in these inequalities $u = 0$ and $u = 2\bar u$, respectively, we deduce the first identity of (3.3). Hence, we get

$$\int_{Q_c} \bar\varphi\, du \le \alpha j_Q(u) \quad \forall u \in \mathcal{M}(Q_c).$$

This implies that $\|\bar\varphi\|_{C(\bar Q_c)} \le \alpha$. But the first identity of (3.3) leads to the equality $\|\bar\varphi\|_{C(\bar Q_c)} = \alpha$ if $\bar u \neq 0$.

We conclude the proof noting that the uniqueness of $\bar\varphi$ for $q > 1$ is an immediate consequence of (3.2) and the definition of $\bar g$.

□

From (3.3) and (3.4), and [5, Lemma 3.4] we deduce the following corollary which shows the sparsity structure of $(\bar u, \bar u_0)$.

**Corollary 3.2.** *Under the assumptions and notations of Theorem 3.1 we have that*

$$\left\{\begin{array}{l} \mathrm{Supp}(\bar u^+) \subset \{(x,t) \in \bar Q_c : \bar\varphi(x,t) = -\alpha\} \\ \mathrm{Supp}(\bar u^-) \subset \{(x,t) \in \bar Q_c : \bar\varphi(x,t) = +\alpha\} \end{array}\right. \tag{3.9}$$

$$\left\{\begin{array}{l} \mathrm{Supp}(\bar u_0^+) \subset \{(x,t) \in \Omega : \bar\varphi(x,t) = -\beta\} \\ \mathrm{Supp}(\bar u_0^-) \subset \{(x,t) \in \bar\Omega : \bar\varphi(x,t) = +\beta\} \end{array}\right. \tag{3.10}$$

*where $\bar u = \bar u^+ - \bar u^-$ and $\bar u_0 = \bar u_0^+ - \bar u_0^-$ are the Jordan decompositions of $\bar u$ and $\bar u_0$, respectively.*

**Remark 3.3.** *Let us observe that $\|\bar\varphi\|_{C(\bar Q)} \le M$ for some constant $M$ independently of $\alpha$ and $\beta$. Indeed, this is obvious for $q = 1$ because $\|\bar\varphi\|_{C(\bar Q)} \le M\|\bar g\|_{L^\infty(Q)} = M$. For $q > 1$, we have that*

$$\|\bar\varphi\|_{C(\bar Q)} \le C\|\bar g\|_{L^{q'}(Q)} = C\|\bar y - y_d\|^{q-1}_{L^q(Q)}$$

$$\le C[q\, J(\bar u, \bar u_0)]^{\frac{q-1}{q}} \le C[q\, J(0,0)]^{\frac{q-1}{q}} = C\|y_d\|^{q-1}_{L^q(Q)} =: M.$$

*As a consequence, if $\alpha > M$ or $\beta > M$, then $\bar u \equiv 0$, respectively $\bar u_0 \equiv 0$.*

Let us mention some additional consequences of the optimality conditions. If $\alpha < \beta$ and $\bar Q_c \supset \Omega \times \{0\}$, then $\|\bar\varphi(0)\|_{C(\bar\Omega)} \le \alpha < \beta$, with (3.4) implies that $\bar u_0 \equiv 0$. Conversely, if $\alpha > \beta$, then by uniform continuity of $\bar\varphi$ there exists $\varepsilon > 0$ such that $|\bar\varphi(x,t)| < \alpha$ for every $(x,t) \in \Omega \times [0, \varepsilon]$. Hence, $\mathrm{supp}(\bar u) \subset \bar Q_c \cap (\Omega \times [\varepsilon, T])$ holds.

**Remark 3.4.** *The results of Sections 2 and 3, in particular Theorem 2.7, Theorem 3.1, and Corollary 3.2 can be extended to the case where $F$ is a convex, weakly lower semi-continuous functional on $L^q(Q)$, which in addition is bounded from below, and which is strictly convex if $q > 1$. In this case* (3.5) *needs to be replaced by $\bar g \in \partial F(\bar y) \in L^{q'}(Q)$, where $q'$ is the conjugate exponent to $q$.*

## 4. Numerical Approximation of (P)

In this section, $\Omega$ is supposed to be convex. To avoid technicalities in the presentation we also assume that $Q_c = Q$. We consider a dG(0)cG(1) discontinuous Galerkin approximation of the state equation (1.2) (i.e., piecewise constant in time and linear nodal basis finite elements in space; see, e.g., [20]). Associated with a parameter $h$ we consider a family of triangulations $\{\mathcal{K}_h\}_{h>0}$ of $\bar{\Omega}$. To every element $K \in \mathcal{K}_h$ we assign two parameters $\rho(K)$ and $\vartheta(K)$, where $\rho(K)$ denotes the diameter of $K$ and $\vartheta(K)$ is the diameter of the biggest ball contained in $K$. The size of the grid is given by $h = \max_{K\in\mathcal{K}_h} \rho(K)$. We will denote by $\{x_j\}_{j=1}^{N_h}$ the interior nodes of the triangulation $\mathcal{K}_h$. In addition, the following usual regularity assumptions on the triangulation are assumed.

(i) There exist two positive constants $\rho_\Omega$ and $\vartheta_\Omega$ such that

$$\frac{h}{\rho(K)} \le \rho_\Omega \quad \text{and} \quad \frac{\rho(K)}{\vartheta(K)} \le \vartheta_\Omega$$

hold for every $K \in \mathcal{K}_h$ and all $h > 0$.

(ii) Let us set $\overline{\Omega}_h = \cup_{K\in\mathcal{K}_h} K$ with $\Omega_h$ and $\Gamma_h$ being its interior and boundary, respectively. We assume that the vertices of $\mathcal{K}_h$ placed on the boundary $\Gamma_h$ are also points of $\Gamma$.

We also introduce a temporal grid $0 = t_0 < t_1 < \ldots < t_{N_\tau} = T$ with $\tau_k = t_k - t_{k-1}$ and set $\tau = \max_{1\le k\le N_\tau} \tau_k$. We assume that there exist $\rho_T > 0$, $C_{\Omega,T} > 0$ and $c_{\Omega,T} > 0$ independent of $h$ and $\tau$ such that

$$\tau \le \rho_T \tau_k, \ \text{for } 1 \le k \le N_\tau. \tag{4.1}$$

We will use the notation $\sigma = (\tau, h)$ and $Q_h = \Omega_h \times (0, T)$.

### 4.1. Discretization of the controls and states

We first discuss the spatial discretization, which follows [4]. Associated to the interior nodes $\{x_j\}_{j=1}^{N_h}$ of $\mathcal{K}_h$ we consider the spaces

$$U_h = \left\{ u_h \in \mathcal{M}(\Omega) : u_h = \sum_{j=1}^{N_h} u_j \delta_{x_j}, \ \text{where } \{u_j\}_{j=1}^{N_h} \subset \mathbb{R} \right\}$$

and

$$Y_h = \left\{ y_h \in C_0(\Omega) : y_h = \sum_{j=1}^{N_h} y_j e_j, \ \text{where } \{y_j\}_{j=1}^{N_h} \subset \mathbb{R} \right\},$$

where $\{e_j\}_{j=1}^{N_h}$ is the nodal basis formed by the continuous piecewise linear functions such that $e_j(x_i) = \delta_{ij}$ for every $1 \le i, j \le N_h$.

For every $\sigma$ we define the space of discrete controls and states by

$$\mathcal{U}_\sigma = \{u_\sigma \in L^1(I, U_h) : u_\sigma|_{I_k} \in U_h, \ 1 \le k \le N_\tau\}$$

and

$$\mathcal{Y}_\sigma = \{y_\sigma \in L^2(I, Y_h) : y_\sigma|_{I_k} \in Y_h, \ 1 \le k \le N_\tau\},$$

where $I_k = (t_{k-1}, t_k]$. The elements $u_\sigma \in \mathcal{U}_\sigma$ and $y_\sigma \in \mathcal{Y}_\sigma$ can be represented in the form

$$u_\sigma = \sum_{k=1}^{N_\tau} u_{k,h} \chi_k \quad \text{and} \quad y_\sigma = \sum_{k=1}^{N_\tau} y_{k,h} \chi_k,$$

where $\chi_k$ is the indicator function of $I_k$, $u_{k,h} \in U_h$ and $y_{k,h} \in Y_h$. Moreover, by definition of $U_h$ and $Y_h$, we can write

$$u_\sigma = \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} u_{kj} \chi_k \delta_{x_j} \quad \text{and} \quad y_\sigma = \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} y_{kj} \chi_k e_j.$$

Thus $\mathcal{U}_\sigma$ and $\mathcal{Y}_\sigma$ are finite dimensional spaces of dimension $N_\tau \times N_h$, and bases are given by $\{\chi_k \delta_{x_j}\}_{k,j}$ and $\{\chi_k e_j\}_{k,j}$.

As in [5], associated to the triangulation of $\Omega$ we define the linear operators $\Lambda_h : \mathcal{M}(\Omega) \longrightarrow U_h \subset \mathcal{M}(\Omega)$ and $\Pi_h : C_0(\Omega) \longrightarrow Y_h \subset C_0(\Omega)$ by

$$\Lambda_h u_0 = \sum_{j=1}^{N_h} \langle u_0, e_j \rangle \delta_{x_j} \quad \text{and} \quad \Pi_h y = \sum_{j=1}^{N_h} y(x_j) e_j. \tag{4.2}$$

The operator $\Pi_h$ is the nodal interpolation operator for $Y_h$. Concerning the operator $\Lambda_h$ we have the following result.

**Proposition 4.1** ( [4, Theorem 3.1])**.** *The following properties hold.*

(i) *For every $u_0 \in \mathcal{M}(\Omega)$ and every $y \in C_0(\Omega)$ and $y_h \in Y_h$ we have*

$$\langle u_0, y_h \rangle = \langle \Lambda_h u_0, y_h \rangle, \tag{4.3}$$
$$\langle u_0, \Pi_h y \rangle = \langle \Lambda_h u_0, y \rangle. \tag{4.4}$$

(ii) *For every $u_0 \in \mathcal{M}(\Omega)$ we have*

$$\|\Lambda_h u_0\|_{\mathcal{M}(\Omega)} \leq \|u_0\|_{\mathcal{M}(\Omega)}, \tag{4.5}$$
$$\Lambda_h u_0 \stackrel{*}{\rightharpoonup} u_0 \text{ in } \mathcal{M}(\Omega) \text{ and } \|\Lambda_h u_0\|_{\mathcal{M}(\Omega)} \to \|u_0\|_{\mathcal{M}(\Omega)} \text{ as } h \to 0. \tag{4.6}$$

Analogously, we define for every $\sigma$ the operators

$$\begin{aligned} \Upsilon_\sigma : \mathcal{M}(Q) \longrightarrow \mathcal{U}_\sigma \qquad & \Upsilon_\sigma u = \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} \frac{1}{\tau_k} \int_{I_k} \int_\Omega e_j \, du \, \delta_{x_j} \chi_k, \\ \Psi_\sigma : C([0,T], C_0(\Omega)) \longrightarrow \mathcal{Y}_\sigma \quad & \Psi_\sigma y = \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} \frac{1}{\tau_k} \int_{I_k} y(x_j, t) \, dt \, e_j \chi_k. \end{aligned} \tag{4.7}$$

Now, we prove the proposition analogous to 4.1.

**Proposition 4.2.** *The following properties hold.*

(i) *For every $u \in \mathcal{M}(Q)$ and every $y \in C([0,T], C_0(\Omega))$ and $y_\sigma \in \mathcal{Y}_\sigma$ we have*

$$\langle u, y_\sigma \rangle = \langle \Upsilon_\sigma u, y_\sigma \rangle, \tag{4.8}$$
$$\langle u, \Psi_\sigma y \rangle = \langle \Upsilon_\sigma u, y \rangle. \tag{4.9}$$

(ii) *For every $u \in \mathcal{M}(Q)$ we have*

$$\|\Upsilon_\sigma u\|_{\mathcal{M}(Q)} \leq \|u\|_{\mathcal{M}(Q)}, \tag{4.10}$$
$$\Upsilon_\sigma u \stackrel{*}{\rightharpoonup} u \text{ in } \mathcal{M}(Q) \text{ and } \|\Upsilon_\sigma u\|_{\mathcal{M}(Q)} \to \|u\|_{\mathcal{M}(Q)} \text{ as } |\sigma| \to 0. \tag{4.11}$$

*Proof.* Using the representation of $y_\sigma$ in the base $\{e_j\chi_k\}_{j,k}$ we have

$$\langle u, y_\sigma\rangle = \sum_{k=1}^{N_\tau}\sum_{j=1}^{N_h} y_{kj}\int_{I_k}\int_\Omega e_j\,du.$$

On the other hand,

$$\langle \Upsilon_\sigma u, y_\sigma\rangle = \sum_{k=1}^{N_\tau}\sum_{j=1}^{N_h}\frac{1}{\tau_k}\int_{I_k}\int_\Omega e_j\,du\,\langle\delta_{x_j}\chi_k, y_\sigma\rangle$$

$$= \sum_{k=1}^{N_\tau}\sum_{j=1}^{N_h}\frac{1}{\tau_k}\int_{I_k}\int_\Omega e_j\,du\int_{I_k} y_\sigma(x_j,t)\,dt = \sum_{k=1}^{N_\tau}\sum_{j=1}^{N_h} y_{kj}\int_{I_k}\int_\Omega e_j\,du,$$

which implies (4.8). Turning to (4.9), we get

$$\langle u, \Psi_\sigma y\rangle = \sum_{k=1}^{N_\tau}\sum_{j=1}^{N_h}\frac{1}{\tau_k}\int_{I_k} y(x_j,t)\,dt\langle u, e_j\chi_k\rangle = \sum_{k=1}^{N_\tau}\sum_{j=1}^{N_h}\frac{1}{\tau_k}\int_{I_k} y(x_j,t)\,dt\int_{I_k}\int_\Omega e_j\,du,$$

and

$$\langle \Upsilon_\sigma u, y\rangle = \sum_{k=1}^{N_\tau}\sum_{j=1}^{N_h}\frac{1}{\tau_k}\int_{I_k}\int_\Omega e_j\,du\,\langle\delta_{x_j}\chi_k, y\rangle = \sum_{k=1}^{N_\tau}\sum_{j=1}^{N_h}\frac{1}{\tau_k}\int_{I_k}\int_\Omega e_j\,du\int_{I_k} y(x_j,t)\,dt.$$

As claimed, these two expressions coincide. Inequality (4.10) is obtained as follows

$$\|\Upsilon_\sigma u\|_{\mathcal{M}(Q)} \le \sum_{k=1}^{N_\tau}\sum_{j=1}^{N_h}\frac{1}{\tau_k}\int_{I_k}\int_\Omega e_j\,d|u|\,\|\delta_{x_j}\chi_k\|_{\mathcal{M}(Q)}$$

$$= \sum_{k=1}^{N_\tau}\sum_{j=1}^{N_h}\int_{I_k}\int_\Omega e_j\,d|u| \le \int_Q d|u| = \|u\|_{\mathcal{M}(Q)}.$$

From this estimate we deduce the existence of a subsequence, denoted in the same way, such that $\Upsilon_\sigma u \stackrel{*}{\rightharpoonup} \tilde u$ as $|\sigma|\to 0$ for some $\tilde u\in\mathcal{M}(Q)$. For any function $y\in C([0,T],C_0(\Omega))$ we know that $\Psi_\sigma y\to y$ in $C([0,T],C_0(\Omega))$ as $|\sigma|\to 0$. Hence, using (4.9) we obtain

$$\langle\tilde u, y\rangle = \lim_{|\sigma|\to 0}\langle\Upsilon_\sigma u, y\rangle = \lim_{|\sigma|\to 0}\langle u, \Psi_\sigma y\rangle = \langle u, y\rangle.$$

Therefore $\tilde u = u$ and consequently the whole sequence $\{\Upsilon_\sigma u\}_\sigma$ converges weakly$^*$ to $u$. This convergence and (4.10) imply that

$$\|u\|_{\mathcal{M}(Q)} \le \liminf_{|\sigma|\to 0}\|\Upsilon u\|_{\mathcal{M}(Q)} \le \|u\|_{\mathcal{M}(Q)},$$

which concludes the proof of (4.11) □

## 4.2. Discrete state equation

In this section we approximate the state equation. We recall that $I_k$ was defined as $(t_{k-1},t_k]$ and consequently $y_{k,h} = y_\sigma(t_k) = y_\sigma|_{I_k}$, $1\le k\le N_\tau$. To approximate the state equation in time we use a dG(0) discontinuous

Galerkin method, which can be formulated as an implicit Euler time stepping scheme. Given a control $(u, u_0) \in \mathcal{M}(Q) \times \mathcal{M}(\Omega)$, for $k = 1, \ldots, N_\tau$ and $z_h \in Y_h$ we set

$$\left\{ \begin{array}{l} \left(\dfrac{y_{k,h} - y_{k-1,h}}{\tau_k}, z_h\right) + a(y_{k,h}, z_h) = \dfrac{1}{\tau_k} \displaystyle\int_{I_k} \int_\Omega z_h \, du \\ y_{0,h} = y_{0h}, \end{array} \right. \tag{4.12}$$

where $(\cdot, \cdot)$ denotes the scalar product in $L^2(\Omega)$, $a$ is the bilinear form associated to the operator $-\Delta$, i.e.,

$$a(y, z) = \int_\Omega \nabla y \nabla z \, dx,$$

and $y_{0h}$ is the unique element of $Y_h$ satisfying

$$(y_{0h}, z_h) = \int_\Omega z_h \, du_0 \quad \forall z_h \in Y_h. \tag{4.13}$$

Obviously, the discrete state $y_\sigma$ associated to $u$ is uniquely defined by (4.12). The strong convergence of $y_\sigma$ to $y = y(u)$ in $L^2(0, T; H_0^1(\Omega))$ for regular functions $(u, u_0)$, for instance $(u, u_0) \in L^2(Q) \times L^2(\Omega)$, is well known. Indeed, the proof of the weak convergence is standard. The strong convergence follows from the compactness result [21, Theorem 3.1].

## 4.3. Definition of the discrete problem $(\mathrm{P}_\sigma)$ and convergence analysis

The approximation of the optimal control problem (P) is defined as

$$(\mathrm{P}_\sigma) \qquad \min_{(u_\sigma, u_{0\sigma}) \in \mathcal{U}_\sigma \times U_h} J_\sigma(u_\sigma, u_{0h}) = \frac{1}{q} \|y_\sigma - y_d\|^q_{L^q(Q_h)} + \alpha \|u_\sigma\|_{\mathcal{M}(Q)} + \|u_{0h}\|_{\mathcal{M}(\Omega)},$$

where $y_\sigma$ is the discrete state associated to $(u_\sigma, u_{0h})$, i.e., the solution to (4.12).

Let us recall that $1 \le q < \min\{2, \frac{d+2}{d}\}$. Hence, its conjugate $q' = q/(q-1)$ satisfies $\max\{2, \frac{d+2}{2}\} < q' \le \infty$. We make the following assumption.

**(A)** Given $q$ such that $1 < q < \min\{2, \frac{d+2}{d}\}$, $\forall f \in L^{q'}(\Omega)$ there exists a unique solution $\varphi$ of (2.2) belonging to $L^{q'}(0, T; W^{2,q'}(\Omega)) \cap W_0^{1,q'}(\Omega)$.

Under assumption (A), we also deduce from the equation (2.2) that $\varphi \in H^1(Q)$. In addition, since $q' > 1 + \frac{d}{2}$ and $f \in L^{q'}(Q)$, then using again [1] we get that $\varphi \in C(\bar{Q})$.

It is well known that assumption (A) holds if $\Gamma$ is of class $C^{1,1}$; see for instance [14, Theorem 9.1]. In the case of a convex polygonal domain $\Omega \subset \mathbb{R}^2$, assumption (A) also holds for $q' < \frac{2}{2 - (\pi/\theta)}$, where $\theta$ is the biggest angle of the polygon. This regularity can be proved by standard arguments and using the $W^{2,p}(\Omega)$ regularity for elliptic problems in polygonal domains; see [11].

Now, we state the main result of this section.

**Theorem 4.3.** $(\mathrm{P}_\sigma)$ *has at least one solution* $(\bar{u}_\sigma, \bar{u}_{0h})$. *Furthermore, if assumption (A) holds,* $1 < q < \min\{2, \frac{d+2}{d}\}$, *and* $\{(\bar{u}_\sigma, \bar{u}_{0h})\}_\sigma$ *denotes a sequence of such solutions with associated states* $\{\bar{y}_\sigma\}_\sigma$, *then the following convergence properties hold*

$$\lim_{|\sigma| \to 0} \|\bar{y} - \bar{y}_\sigma\|_{L^q(Q)} = 0, \tag{4.14}$$

$$(\bar{u}_\sigma, \bar{u}_{0h}) \stackrel{*}{\rightharpoonup} (\bar{u}, \bar{u}_0) \quad \text{as } |\sigma| \to 0 \quad \text{in } \mathcal{M}(Q) \times \mathcal{M}(\Omega), \tag{4.15}$$

$$\lim_{|\sigma| \to 0} \left(\|\bar{u}_\sigma\|_{\mathcal{M}(Q)}, \|\bar{u}_{0h}\|_{\mathcal{M}(\Omega)}\right) = \left(\|\bar{u}\|_{\mathcal{M}(Q)}, \|\bar{u}_0\|_{\mathcal{M}(\Omega)}\right), \tag{4.16}$$

*where* $(\bar{u}, \bar{u}_0)$ *is the unique solution of* (P) *and* $\bar{y}$ *its associated state.*

*Proof.* The existence of a solution of $(\mathrm{P}_\sigma)$ is an immediate consequence of the finite dimension of $\mathcal{U}_\sigma \times U_h$, and the continuity and coercivity of $J_\sigma$. Let us prove (4.14)-(4.16). First, we observe that $J_\sigma(\bar u_\sigma), \bar u_{0h}) \le J(0,0) \le \frac{1}{q}\|y_d\|^q_{L^q(Q)}$. Consequently, the sequences $\{\bar u_\sigma, \bar u_{0h})\}_\sigma$ and $\{\bar y_\sigma\}_\sigma$ are bounded in $\mathcal{M}(Q)\times\mathcal{M}(\Omega)$ and $L^q(Q)$, respectively. Hence, we can take subsequences, denoted in the same way such that

$$(\bar u_\sigma, \bar u_{0h}) \stackrel{*}{\rightharpoonup} (\tilde u, \tilde u_0) \ \text{ in } \mathcal{M}(Q)\times\mathcal{M}(\Omega) \ \text{ and } \ \bar y_\sigma \rightharpoonup \tilde y \ \text{ in } L^q(Q). \tag{4.17}$$

Let us split the rest of the proof into several steps.

*I - $\tilde y$ is the solution of* (1.2) *corresponding to* $(\tilde u, \tilde u_0)$. Let us take $\xi \in C^1[0,T]$ with $\xi(T) = 0$, and $\psi \in W^{2,q'}(\Omega)\cap W_0^{1,q'}(\Omega)$. We approximate $\psi$ by $\psi_h \in Y_h$ satisfying

$$a(\psi_h, z_h) = a(\psi, z_h) \quad \forall z_h \in Y_h, \quad \|\psi - \psi_h\|_{C(\bar\Omega)} \to 0 \ \text{ as } h \to 0. \tag{4.18}$$

Many papers are devoted to prove error estimates for $\|\psi - \psi_h\|_\infty$; see, for instance, [8, Theorem 19.3, pages 143-144] for a simple proof or [18] and the references therein for improved error estimates. Using (4.12) we have

$$\int_0^T (\bar y_\sigma(t), \psi_h)\xi'(t)\,dt = \sum_{k=1}^{N_\tau}\int_{I_k}(y_{k,h},\psi_h)\xi'(t)\,dt = \sum_{k=1}^{N_\tau}(y_{k,h},\psi_h)(\xi(t_k)-\xi(t_{k-1}))$$

$$= -\sum_{k=1}^{N_\tau}(y_{k,h}-y_{k-1,h},\psi_h)\xi(t_{k-1}) - (y_{0h},\psi_h)\xi(0)$$

$$= \sum_{k=1}^{N_\tau}\left\{\tau_k a(y_{k,h},\psi_h) - \int_{I_k}\int_\Omega \psi_h\,d\bar u_\sigma\right\}\xi(t_{k-1}) - (y_{0h},\psi_h)\xi(0)$$

$$= \int_0^T a(\bar y_\sigma(t),\psi_h)\xi(t)\,dt - \int_0^T\int_\Omega \psi_h\xi(t)\,d\bar u_\sigma - (y_{0h},\psi_h)\xi(0)$$

$$+\sum_{k=1}^{N_\tau}\left\{\int_{I_k} a(y_{k,h},\psi_h)(\xi(t_{k-1})-\xi(t))\,dt - \int_{I_k}\int_\Omega \psi_h(\xi(t_{k-1})-\xi(t))\,d\bar u_\sigma\right\}$$

and with (4.13), (4.18) and $\psi \in W^{2,q'}(\Omega)\cap W_0^{1,q'}(\Omega)$

$$= \int_0^T a(\bar y_\sigma(t),\psi)\xi(t)\,dt - \int_0^T\int_\Omega \psi_h\xi(t)\,d\bar u_\sigma - \int_\Omega \psi_h\,d\bar u_{0h}\xi(0)$$

$$+\sum_{k=1}^{N_\tau}\left\{\int_{I_k} a(y_\sigma(t),\psi)(\xi(t_{k-1})-\xi(t))\,dt - \int_{I_k}\int_\Omega \psi_h(\xi(t_{k-1})-\xi(t))\,d\bar u_\sigma\right\}$$

$$= \int_0^T (\bar y_\sigma(t),-\Delta\psi)\xi(t)\,dt - \int_0^T\int_\Omega \psi_h\xi(t)\,d\bar u_\sigma - \int_\Omega \psi_h\,d\bar u_{0h}\xi(0)$$

$$+\sum_{k=1}^{N_\tau}\left\{\int_{I_k} (y_\sigma(t),-\Delta\psi)(\xi(t_{k-1})-\xi(t))\,dt - \int_{I_k}\int_\Omega \psi_h(\xi(t_{k-1})-\xi(t))\,d\bar u_\sigma\right\}.$$

Using (4.18), it is immediate to pass to the limit and to obtain

$$\lim_{|\sigma|\to 0}\int_0^T (\bar y_\sigma(t),\psi_h)\xi'(t)\,dt = \int_0^T (\tilde y(t),\psi)\xi'(t)\,dt$$

and

$$\lim_{|\sigma|\to 0}\left\{\int_0^T(\bar y_\sigma(t),-\Delta\psi)\xi(t)\,dt-\int_0^T\int_\Omega\psi_h\xi(t)\,d\bar u_\sigma-\int_\Omega\psi_h\,d\bar u_{0h}\,\xi(0)\right\}$$
$$=\int_0^T(\tilde y(t),-\Delta\psi)\xi(t)\,dt-\int_0^T\int_\Omega\psi\xi(t)\,d\tilde u-\int_\Omega\psi\,d\tilde u_0\,\xi(0).$$

The remaining terms can be estimated as follows

$$\left|\sum_{k=1}^{N_\tau}\left\{\int_{I_k}(y_\sigma(t),-\Delta\psi)(\xi(t_{k-1})-\xi(t))\,dt-\int_{I_k}\int_\Omega\psi_h(\xi(t_{k-1})-\xi(t))\,d\bar u_\sigma\right\}\right|$$
$$\leq\|\bar y_\sigma\|_{L^q(Q)}\|\Delta\psi\|_{L^{q'}(\Omega)}\tau\|\xi'\|_\infty+\|\psi_h\|_{C(\bar\Omega)}\|\bar u_\sigma\|_{\mathcal{M}(Q)}\tau\|\xi'\|_\infty\to 0\quad\text{as } |\sigma|\to 0.$$

From the above equalities we infer that

$$\int_Q-\tilde y(\frac{\partial}{\partial t}+\Delta)(\psi\xi)\,dx\,dt=\int_Q(\psi\xi)\,d\tilde u+\int_\Omega(\psi\xi)(0)\,d\tilde u_0.$$

Since $\tilde y\in L^{q'}(Q)$, by density arguments, we have that the identity (2.1) is satisfied by $\tilde y$ for every $\phi\in L^{q'}(0,T;W^{2,q'}(\Omega)\cap W_0^{1,q'}(\Omega))\cap H^{1,q'}(Q)\cap C(\bar Q)$. Due to assumption (A), the solutions of (2.2) enjoy this regularity for every $f\in L^\infty(Q)$. Hence, we conclude that $\tilde y$ is the solution of (1.2) associated to $(\tilde u,\tilde u_0)$.

*II* - $J(\tilde u,\tilde u_0)\leq J(u,u_0)$ $\forall(u,u_0)\in C(\bar Q)\times C(\bar\Omega)$. Since $\Omega$ is convex, the solution $y$ of (1.2) associated to one of these regular controls $(u,u_0)$ belongs to $L^2(0,T;H^2(\Omega)\cap H_0^1(\Omega))\cap H^1(Q)$. As we mentioned above, the corresponding discrete solutions $y_\sigma$, of (4.12), converge strongly to $y$ in $L^2(0,T;H_0^1(\Omega))\subset L^q(Q)$ since $q<2$.

Now, set $(u_\sigma,u_{0h})=(\Upsilon_\sigma u,\Lambda_h u_0)$. From (4.3) and (4.8), we deduce that the discrete states associated to $(u,u_0)$ and $(u_\sigma,u_{0h})$ coincide. Indeed, first we observe that (4.3) implies that

$$\int_\Omega z_h\,du_0=\int_\Omega z_h\,du_{0,h},\quad\forall z_h\in Y_h.$$

Therefore, (4.13) shows that $y_{0h}$ coincides for both controls. Second, we use (4.8) replacing $y_\sigma$ by $z_h\chi_k\in\mathcal{Y}_\sigma$, for any $z_h\in Y_h$ and $1\leq k\leq N_\tau$, then we get

$$\frac{1}{\tau_k}\int_{I_k}\int_\Omega z_h\,du=\frac{1}{\tau_k}\int_{I_k}\int_\Omega z_h\,du_\sigma.$$

Hence, the effects of $(u,u_0)$ and $(u_\sigma,u_{0h})$ on the discretized equation (4.12) coincide and they provide the same solution. From (4.6) and (4.11), and $y_\sigma\to y$ in $L^q(Q)$, it follows that $J_\sigma(u_\sigma,u_{0h})\to J(u,u_0)$. Using that $(\bar u_\sigma,\bar u_{0h})$ is a solution of (P$_\sigma$) and (4.17) we infer that

$$J(\tilde u,\tilde u_0)\leq\liminf_{|\sigma|\to 0}J_\sigma(\bar u_\sigma,\bar u_{0h})\leq\limsup_{|\sigma|\to 0}J_\sigma(\bar u_\sigma,\bar u_{0h})$$
$$\leq\limsup_{|\sigma|\to 0}J_\sigma(u_\sigma,u_{0h})=J(u,u_0).\tag{4.19}$$

*III* - $(\tilde u,\tilde u_0)=(\bar u,\bar u_0)$. To prove this, it is enough to show that $(\tilde u,\tilde u_0)$ is a solution of (P). Then the uniqueness implies the desired equality. To this purpose, let us chose a sequence $(u_k,u_{0k})\in C(\bar Q)\times C(\bar\Omega)$ such that

$$\begin{aligned}&(u_k,u_{0k})\stackrel{*}{\rightharpoonup}(\bar u,\bar u_0)\text{ in }\mathcal{M}(Q)\times\mathcal{M}(\Omega),\\&\|u_k\|_{L^1(Q)}\leq\|u\|_{\mathcal{M}(Q)}\ \text{ and }\ \|u_{0k}\|_{L^1(\Omega)}\leq\|u\|_{\mathcal{M}(\Omega)}.\end{aligned}\tag{4.20}$$

From Lemma 2.6 we deduce the strong convergence $y_k \to \bar{y}$, where $\{y_k\}_k$ are the states associated to $\{(u_k, u_{0k})\}_k$. On the other hand, (4.20) implies that

$$\|\bar{u}\|_{\mathcal{M}(Q)} \le \liminf_{k\to\infty} \|u_k\|_{L^1(Q)} \le \limsup_{k\to\infty} \|u_k\|_{L^1(Q)} \le \|\bar{u}\|_{\mathcal{M}(Q)}.$$

Hence, $\|u_k\|_{\mathcal{M}(Q)} \to \|\bar{u}\|_{\mathcal{M}(Q)}$. Analogously we get the convergence $\|u_{0k}\|_{\mathcal{M}(Q)} \to \|\bar{u}_0\|_{\mathcal{M}(\Omega)}$. Altogether shows that $J(u_k, u_{0k}) \to J(\bar{u}, \bar{u}_0)$. Together with (4.19), this implies that $J(\tilde{u}, \tilde{u}_0) \le J(\bar{u}, \bar{u}_0) = \inf$ (P). Hence, we have that $(\tilde{u}, \tilde{u}_0) = (\bar{u}, \bar{u}_0)$, and using once again (4.19) and (4.17) we get

$$\lim_{|\sigma|\to 0} J_\sigma(\bar{u}_\sigma, \bar{u}_{0h}) = J(\bar{u}, \bar{u}_0) \ \text{ and } \ \bar{y}_\sigma \rightharpoonup \bar{y} \ \text{ in } L^q(Q). \tag{4.21}$$

*IV - Proof of* (4.14)-(4.16). We have proved that any subsequence of solutions of ($P_\sigma$) converges to the unique solution $(\bar{u}, \bar{u}_0)$ of (P). This gives (4.15). From (4.21) we deduce

$$\frac{1}{q}\|\bar{y} - y_d\|^q_{L^q(Q)} \le \liminf_{|\sigma|\to 0} \frac{1}{q}\|\bar{y}_\sigma - y_d\|^q_{L^q(Q)} \le \limsup_{|\sigma|\to 0} \frac{1}{q}\|\bar{y}_\sigma - y_d\|^q_{L^q(Q)}$$

$$= \limsup_{|\sigma|\to 0} \left\{J_\sigma(\bar{u}_\sigma, \bar{u}_{0h}) - \alpha\|\bar{u}_\sigma\|_{\mathcal{M}(Q)} - \beta\|\bar{u}_{0h}\|_{\mathcal{M}(\Omega)}\right\}$$

$$\le \limsup_{|\sigma|\to 0} J_\sigma(\bar{u}_\sigma, \bar{u}_{0h}) - \liminf_{|\sigma|\to 0} \left\{\alpha\|\bar{u}_\sigma\|_{\mathcal{M}(Q)} + \beta\|\bar{u}_{0h}\|_{\mathcal{M}(\Omega)}\right\}$$

$$\le J(\bar{u}, \bar{u}_0) - \left\{\alpha\|\bar{u}\|_{\mathcal{M}(Q)} + \beta\|\bar{u}_0\|_{\mathcal{M}(\Omega)}\right\} = \frac{1}{q}\|\bar{y} - y_d\|^q_{L^q(Q)}.$$

Together with the weak converge $\bar{y}_\sigma \rightharpoonup \bar{y}$ in $L^q(Q)$, this implies the strong convergence (4.14). To prove that $\|\bar{u}_\sigma\|_{\mathcal{M}(Q)} \to \|\bar{u}\|_{\mathcal{M}(Q)}$ we proceed in a similar way

$$\alpha\|\bar{u}\|_{\mathcal{M}(Q)} \le \liminf_{|\sigma|\to 0} \alpha\|\bar{u}_\sigma\|_{\mathcal{M}(Q)} \le \limsup_{|\sigma|\to 0} \alpha\|\bar{u}_\sigma\|_{\mathcal{M}(Q)}$$

$$= \limsup_{|\sigma|\to 0} \left\{J_\sigma(\bar{u}_\sigma, \bar{u}_{0h}) - \frac{1}{q}\|\bar{y}_\sigma - y_d\|^q_{L^q(Q_h)} - \beta\|u_{0h}\|_{\mathcal{M}(\Omega)}\right\}$$

$$\le \limsup_{|\sigma|\to 0} J_\sigma(\bar{u}_\sigma, \bar{u}_{0h}) - \liminf_{|\sigma|\to 0} \left\{\frac{1}{q}\|\bar{y}_\sigma - y_d\|^q_{L^q(Q_h)} + \beta\|u_{0h}\|_{\mathcal{M}(\Omega)}\right\}$$

$$\le J(\bar{u}, \bar{u}_0) - \left\{\frac{1}{q}\|\bar{y} - y_d\|^q_{L^q(Q)} + \beta\|u_0\|_{\mathcal{M}(\Omega)}\right\} = \alpha\|\bar{u}\|_{\mathcal{M}(Q)}.$$

Finally, $\|\bar{u}_{0h}\|_{\mathcal{M}(\Omega)} \to \|\bar{u}_0\|_{\mathcal{M}(\Omega)}$ is an immediate consequence of (4.14), $\|\bar{u}_\sigma\|_{\mathcal{M}(Q)} \to \|\bar{u}\|_{\mathcal{M}(Q)}$ and (4.21). $\square$

## 5. Extensions

In this section we analyze the situations where not both controls $u$ and $u_0$ are simultaneously present in the state equation. We also consider some cases where the observation domain is a strict subset of the physical domain $\Omega$ and temporal observation is not necessarily during the whole time $(0, T)$. We are especially interested in the consequences on the sparsity structure of the optimal controls.

### 5.1. **Separated control and observation domains**

Here, we consider where the observation takes places in a open set $\Omega_o$ and during an interval of time $I_o$. Let us denote $Q_o = \Omega_o \times I_o$. On the other hand, the distributed control $u$ is supported on a region $\omega$ such that $\bar\omega \cap \bar\Omega_o = \emptyset$. The cost functional is then given by

$$J(u, u_0) = \frac{1}{q}\|y - y_d\|^q_{L^q(Q_o)} + \alpha\|u\|_{\mathcal{M}(Q_c)} + \beta\|u_0\|_{\mathcal{M}(\Omega)}.$$

For this new cost functional, Theorem 2.7 is still valid except for the uniqueness of solutions. The difficulty arises from the lack of injectivity of the control to observation mapping, which excludes the strict convexity of $J$ even if $q > 1$. Of course, this effects that Theorem 4.3 on the numerical approximation in the sense that we can have different sequences of discrete optimal controls converging to different solutions of (P). Otherwise the convergence properties still hold along (4.14)-(4.16), now interpreted subsequentially. In the optimality system (3.2)-(3.4), the definition of $\bar g$ given by (3.5) is only correct in $Q_o$ and it should taken as zero outside.

Let us discuss the sparsity properties of the optimal controls $(\bar u, \bar u_0)$. From (3.2) we get that

$$\frac{\partial\bar\varphi}{\partial t} + \Delta\bar\varphi = 0 \ \text{ in } Q_1 = [(\Omega \setminus \bar\Omega_o) \times (0,T)] \cup [\Omega \times ((0,T) \setminus \bar I_0)\,].$$

From the properties of the heat operator we deduce that $\bar\varphi \in C^\infty(Q_1) \cap C(\bar Q)$. Let us verify that there exists $0 < T_0 < T$ such that the support of $\bar u$ is contained in $(\partial\omega \cap \Omega) \times [0, T_0]$. Indeed, according to (3.9), it is enough to show that $|\bar\varphi(x,t)| < \alpha$ for every $x \in \omega$ and all $t > T_0$. Since $\bar\varphi(x,T) = 0$ $\forall x \in \bar\Omega$ and $\bar\varphi$ is continuous in $\bar Q$, we deduce the existence of $0 < T_0 < T$ such that $|\bar\varphi(x,t)| < \alpha$ $\forall (x,t) \in \bar\Omega \times (T_0,T)$. Let us prove that $|\bar\varphi(x,t)| < \alpha$ for every $x \in \omega$. We argue by contradiction and let us assume that there exists a point $x_0 \in \omega$ and some $t_0 \in [0,T_0]$ such that $\bar\varphi(x_0,t_0) = \alpha$. From (3.3) this means that the maximum of $\bar\varphi$ is achieved at $(x_0,t_0)$. Then, from the parabolic strong maximum principle [10, Theorem 11, page 375] and the connectivity of $\omega$ we deduce that $\bar\varphi(x,t) = \alpha$ $\forall (x,t) \in \omega \times [t_0,T]$, which contradicts that $\bar\varphi(x,t) = 0$ whenever $t > T_0$. In the same manner we can exclude the possibility of achieving the value $-\alpha$ in $\omega$. In the case $d = 1$ and $\omega = (a,b)$ with $[a,b] \subset \Omega$, then $\partial\omega = \{a,b\}$, which implies that $\bar u = \bar u_a\delta_a + \bar u_b\delta_b$, with $\bar u_a, \bar u_b \in \mathcal{M}([0,T_0])$, and $\delta_a$ and $\delta_b$ denote the Dirac measures concentrated at $a$ and $b$, respectively. Moreover, since the maximum and minimum values of $\bar\varphi$ are achieved on the boundary of $\omega$ for every $t$ and since $|\bar\varphi(x,t)| \not\equiv \alpha$ in $\omega$ for all $t$, then if $\bar\varphi(a) = \alpha$, then $\bar\varphi(b) < \alpha$. We can argue in the same way with $b$. This shows that $\mathrm{supp}(u_a^+) \cap \mathrm{supp}(u_b^+) = \emptyset$ and $\mathrm{supp}(u_a^-) \cap \mathrm{supp}(u_b^-) = \emptyset$.

To deal with $\bar u_0$ we assume that $0 \notin \bar I_o$. Then, we prove that $\bar u_0$ is supported on a set in $\Omega$ with a zero Lebesgue measure. To this end, now we use (3.10). Since the mapping $x \in \Omega \longrightarrow \bar\varphi(x,0) \in \mathbb{R}$ is analytic, then either $|\bar\varphi(x,0)| = \beta$ in $\Omega$ or the set of points where $|\bar\varphi(x,0)| = \beta$ has a zero Lebesgue measure. But the boundary condition $\bar\varphi(x,0) = 0$ for $x \in \Gamma$ excludes the first possibility. Once again, we can get some extra information in the one-dimensional case, $d = 1$. Indeed, the analyticity of $x \in \Omega \longrightarrow \bar\varphi(x,0) \in \mathbb{R}$ implies that the set points where $|\bar\varphi(x,0)| = \beta$ in $\Omega$ is finite. Let us denote them by $\{x_k\}_{k=1}^m$. Hence, the equality $\bar u_0 = \sum_{k=1}^m \bar\lambda_k\delta_{x_k}$ holds for some real numbers $\{\bar\lambda_k\}_{k=1}^m$.

### 5.2. **Terminal observation with initial controls**

In this case, we consider the cost functional is given by

$$J(u_0) = \frac{1}{q}\|y(T) - y_d\|^q_{L^q(\Omega)} + \beta\|u_0\|_{\mathcal{M}(\Omega)},$$

where $y$ is the unique solution of the state equation (1.2) with $u = 0$, and $y_d \in L^q(\Omega)$ is given. From the state equation we deduce that any feasible state $y$ belongs to $C^\infty(\Omega \times (0,T])$. Hence, the control problem is well formulated for any $q \in [1,+\infty)$. In any of these cases, there exists at least one optimal control. Moreover, if

$q > 1$, then the solution is unique. Indeed, let us assume that $u_{01}$ and $u_{02}$ are solutions of the problem. Then, the convexity of the norm $\|\cdot\|_{\mathcal{M}(\Omega)}$ and the strict convexity of the functional $z \to \|z - y_d\|^q_{L^q(\Omega)}$ imply that $y_{u_{01}}(T) = y_{u_{02}}(T)$. Set $u_0 = u_{01} - u_{02}$ and let $y$ be the state associated to $u_0$. Then, $y$ is solution of the heat equation in $Q$, vanishing on the boundary $\Sigma$ and $y(T) = 0$. Now, the backward uniqueness of the heat equation implies that $y = 0$ and hence $u_0 = y(0) = 0$.

The optimality system satisfied by an optimal control $\bar{u}_o$ is formulated as follows

$$\left\{\begin{array}{rcll} -\dfrac{\partial\bar\varphi}{\partial t} - \Delta\bar\varphi & = & 0 & \text{in } Q \\ \bar\varphi(x,T) & = & \bar g & \text{in } \Omega \\ \bar\varphi(x,t) & = & 0 & \text{on } \Sigma, \end{array}\right. \tag{5.1}$$

$$\left\{\begin{array}{l} \displaystyle\int_\Omega \bar\varphi(0)\,d\bar u_0 + \beta\|\bar u_0\|_{\mathcal{M}(\Omega)} = 0 \\ \|\bar\varphi(0)\|_{C(\bar\Omega)} \left\{\begin{array}{ll} = \beta & \text{if } \bar u_0 \neq 0 \\ \leq \beta & \text{if } \bar u_0 = 0, \end{array}\right. \end{array}\right. \tag{5.2}$$

where

$$\bar g(x) \left\{\begin{array}{ll} = |\bar y(x,T) - y_d(x)|^{q-2}(\bar y(x,T) - y_d(x)) & \text{if } 1 < q < +\infty \\ \in \operatorname{sign}(\bar y(x,T) - y_d(x)) & \text{if } q = 1. \end{array}\right. \tag{5.3}$$

To prove this optimality system we proceed in an analogous way to Theorem 3.1 using the fact that $\bar\varphi \in C(\bar\Omega \times [0,T))$. From this optimality system we can deduce the same sparsity structure for $\bar u_0$ as obtained in the second paragraph of the previous sub-section.

This problem is related to applications for inverse problems of source identification studied in the literature. Under the above formulation of the control problem, we deduce for $d = 1$ that the optimal control has the structure $\bar u_0 = \sum_{k=1}^m \bar\lambda_k \delta_{x_k}$, as assumed in [15].

## References


[1] E. Di Benedetto. On the local behaviour of solutions of degenerate parabolic equations with measurable coefficients. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze*, 13(3):487–535, 1986.

[2] L. Boccardo and T. Gallouet. Non-linear elliptic and parabolic equations involving measure data. *Journal of Functional Analysis*, 87(1):149 – 169, 1989.

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

[4] E. Casas, C. Clason, and K. Kunisch. Approximation of elliptic control problems in measure spaces with sparse solutions. *SIAM Journal on Control and Optimization*, 50(4):1735–1752, 2012.

[5] E. Casas, C. Clason, and K. Kunisch. Parabolic control problems in measure spaces with sparse solutions. *SIAM J. Control Optim.*, 51(1):28–63, 2013.

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

[7] E. Casas and E. Zuazua. Spike controls for elliptic and parabolic pde. *Systems Control Lett.*, 62:311–318, 2013.

[8] P.G. Ciarlet and J.L. Lions. *Handbook of Numerical Analysis: II*. Handbook of Numerical Analysis. North-Holland, 1990.

[9] C. Clason and K. Kunisch. A duality-based approach to elliptic control problems in non-reflexive Banach spaces. *ESAIM: Control, Optimisation and Calculus of Variations*, 17(1):243–266, 2011.

[10] L.C. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 2010.

[11] P. Grisvard. *Elliptic Problems in Nonsmooth Domains*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1985.

[12] R. Herzog and G. Stadler und G. Wachsmuth. Directional sparsity in optimal control of partial differential equations. *SIAM Journal on Control and Optimization*, 50(2):943–963, 2012.

[13] K. Kunisch, K. Pieper, and B. Vexler. Measure valued directional sparsity for parabolic optimal control problems. Submitted, 2014.

[14] O.A. Ladyzhenskaia, V.A. Solonnikov, and N.N. Uralceva. *Linear and Quasi-linear Equations of Parabolic Type*. American Mathematical Society, translations of mathematical monographs. American Mathematical Society, 1968.

[15] Y. Li, S. Osher, and R. Tsai. Heat source identification based on $l_1$ constrained minimization. *Inverse Probl. and Imaging*, 8(1):199–221, 2014.

[16] K. Pieper and B. Vexler. A priori error analysis for discretization of sparse elliptic optimal control problems in measure space. *SIAM Journal on Control and Optimization*, 51(4):2788–2808, 2013.

[17] J.P. Raymond and H. Zidani. Hamiltonian pontryagin's principles for control problems governed by semilinear parabolic equations. *Applied Mathematics and Optimization*, 39(2):143–177, 1999.

[18] A.H. Schatz. Pointwise error estimates and asymptotic error expansion inequalities for the finite element method on irregular grids. *Part I. Global Estimates, Math. Comp*, pages 877–899, 1998.

[19] J. Simon. Compact sets in the space $L^p(O,T;B)$. *Annali di Matematica Pura ed Applicata*, 146(1):65–96, 1986.

[20] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems (Springer Series in Computational Mathematics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[21] N. Walkington. Compactness properties of the dg and cg time stepping schemes for parabolic equations. *SIAM Journal on Numerical Analysis*, 47(6):4680–4710, 2010.