SeMa manuscript No.
(will be inserted by the editor)

A review on sparse solutions in optimal control of partial differential equations

Eduardo Casas

Received: date / Revised version: date

Abstract In this paper a review of the results on sparse controls for partial differential equations is presented. There are two different approaches to the sparsity study of control problems. One approach consists of taking functions to control the system, putting in the cost functional a convenient term that promotes the sparsity of the optimal control. A second approach deals with controls that are Borel measures and the norm of the measure is involved in the cost functional. The use of measures as controls allows to obtain optimal controls supported on a zero Lebesgue measure set, which is very interesting for practical implementation. If the state equation is linear, then we can carry out a complete analysis of the control problem with measures. However, if the equation is nonlinear the use of measures to control the system is still an open problem, in general, and the use of functions to control the system seems to be more appropriate.

Keywords sparse controls · measure controls · elliptic equations · parabolic equations

Mathematics Subject Classification (2000) 35J15 · 35K10 · 49K20 · 49M25 · 65N30

1 Introduction

In the control of distributed parameter systems, those formulated by partial differential equations, usually we cannot put control devices at every point of the domain. Actually, we are allowed to use small regions to put the controllers. Then the big issue

This work was partially supported by the Spanish Ministerio de Economía y Competitividad under project MTM2014-57531-P

Eduardo Casas
Dpto. de Matemática Aplicada y Ciencias de la Computación, E.T.S.I. Industriales y de Telecomunicación, Universidad de Cantabria, Av. de los Castros s/n, 39005 Santander, Spain. E-mail: eduardo.casas@unican.es

is which region is the most convenient to localize them. Of course, we have to determine the power of the controllers as well. These controls are called *sparse* because they are not zero only in a small region of the domain. In the last few years, some researchers have focused their investigation in this direction. First, it was observed that the use of the L^1 norm of the control in the cost functional leads to the sparsity of the solution. Of course, this introduces some mathematical difficulties in the problem due to the lack of differentiability of this functional. However, despite this difficulty, a lot of progress has been done, and the numerical computations show the interest and applicability of this approach.

Taking a further step in this direction, we find that many times it is even desirable to put the controllers only in finitely many points of the domain, or along a line (in two dimensions), or on a surface (in three dimensions). In these cases we need to use controllers that are localized in a zero Lebesgue measure set. These controllers cannot be identified with functions, they are measures. This is the starting point of a new type of control problems where the controls are Borel measures. Adding the norm of the measure to the cost functional, we obtain optimal controls having the desired sparsity property.

In this paper we present the results obtained in the analysis of sparse control problems, both taking functions or measures as controls. The case of elliptic and parabolic control problems is considered. The paper is organized as follows. In Sect. 2 the sparse control of semilinear elliptic equations is studied. In Sect. 3 we present the parabolic case. In Sects. 4 and 5 the elliptic and parabolic cases corresponding to measure controls are analyzed. Though some results are indicated and references are provided, we have not considered the numerical approximation of the control problems because this would lead to a very long paper.

2 Sparse control of smilinear elliptic equations

As far as the author knows, the first paper devoted to the study of sparse controls of elliptic systems is due to Stadler [29]. In this paper the author considers a distributed control problem associated to a linear elliptic equation. The cost functional is the usual quadratic cost that inludes the Tikhonov term, and the L^1 norm of the control is added to the functional. Control constraints are included in the formulation of the problem. The resulting problem is strictly convex, then it has a unique solution. Therefore, the first order optimality conditions are necessary and sufficient for global optimality. He derives these conditions and deduces from them the sparsity of the optimal control. Finally, he uses a semismooth Newton method to compute a discrete version of the control problem. For some additional analysis of the same problem and the proof of error estimates for the numerical discretization, the reader is referred to [31,32]. Later the control problem associated to a semilinear elliptic equation and a more general cost functional was investigated in [9,10]. The material presented in this section is based on the theoretical part of [10].

2.1 Setting of the control problem and preliminary results.

In this section Ω will denote an open bounded subset of $\mathbb{R}^n$, $n = 2$ or 3, with a Lipschitz boundary Γ.

We consider the following control problem

$$\text{(P)} \begin{cases} \min J(u) \\ \alpha \le u(x) \le \beta \text{ for a.a. } x \in \Omega \end{cases}$$

where $J(u) = F(u) + \gamma j(u)$, with $F : L^2(\Omega) \longrightarrow \mathbb{R}$ and $j : L^1(\Omega) \longrightarrow \mathbb{R}$ defined by

$$F(u) = \int_\Omega L(x, y_u(x))\, dx + \frac{\nu}{2} \int_\Omega u^2(x)\, dx \ \text{ and } \ j(u) = \int_\Omega |u(x)|\, dx,$$

y_u being the solution of the state equation

$$\begin{cases} Ay + a(x,y) = u & \text{in } \Omega, \\ \qquad\qquad y = 0 & \text{on } \Gamma. \end{cases} \tag{2.1}$$

A is the linear operator

$$Ay = -\sum_{i,j=1}^{n} \partial_{x_j}[a_{ij}(x)\, \partial_{x_i} y] + a_0(x)\, y.$$

We make the following assumptions on the functions and parameters involved in the control problem (P).

Assumption 1.— The coefficients of A have the following properties: $a_0, a_{ij} \in L^\infty(\Omega)$,

$$a_0(x) \ge 0 \ \text{ and } \ \sum_{i,j=1}^{n} a_{ij}(x)\, \xi_i\, \xi_j \ge \Lambda\, |\xi|^2 \text{ for a.a. } x \in \Omega \ \text{ and } \ \forall \xi \in \mathbb{R}^n, \tag{2.2}$$

for some $\Lambda > 0$.

Assumption 2.— $a : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second variable, with $a(\cdot, 0) \in L^{\bar{p}}(\Omega)$ for some $\bar{p} > \frac{n}{2}$, and satisfying

$$\begin{cases} \dfrac{\partial a}{\partial y}(x,y) \ge 0 \ \text{ for a.a. } \ x \in \Omega \ \text{ and } \ \forall y \in \mathbb{R} \\ \forall M > 0\ \exists C_M > 0 : \displaystyle\sum_{j=1}^{2} \left| \frac{\partial^j a}{\partial y^j}(x,y) \right| \le C_M \ \text{ for a.a. } x \in \Omega \ \text{ and } |y| \le M, \\ \forall \varepsilon > 0 \text{ and } \forall M > 0\ \exists \delta : \text{ if and } |y_2 - y_1| < \delta, |y_1|, |y_2| \le M \text{ then} \\ \qquad\qquad \left| \dfrac{\partial^2 a}{\partial y^2}(x, y_2) - \dfrac{\partial^2 a}{\partial y^2}(x, y_1) \right| \le \varepsilon \text{ for a.a. } x \in \Omega. \end{cases} \tag{2.3}$$

Assumption 3.— We also assume that $-\infty < \alpha < 0 < \beta < +\infty$, $\gamma > 0$, $\nu \ge 0$, and $L : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second

variable such that $L(\cdot,0)\in L^1(\Omega)$ and for every $M>0$ there exists a function $\psi_M\in L^{\bar p}(\Omega)$ satisfying

$$\left|\frac{\partial L}{\partial y}(x,y)\right|\le \psi_M(x)\ \ \forall |y|\le M \ \text{ and for a.a. } x\in\Omega. \tag{2.4}$$

Additionally we assume

$$\begin{cases} \forall M>0\ \exists C_M>0: \left|\dfrac{\partial^2 L}{\partial y^2}(x,y)\right|\le C_M \ \text{ for a.a. } x\in\Omega \ \text{ and } |y|\le M, \\ \forall \varepsilon>0 \text{ and } \forall M>0\ \exists\delta: \text{ if and } |y_2-y_1|<\delta, |y_1|,|y_2|\le M \text{ then} \\ \qquad\qquad \left|\dfrac{\partial^2 L}{\partial y^2}(x,y_2)-\dfrac{\partial^2 L}{\partial y^2}(x,y_1)\right|\le\varepsilon \text{ for a.a. } x\in\Omega. \end{cases} \tag{2.5}$$

In the sequel, we will denote the set of feasible controls by

$$\mathbb{K}=\{u\in L^\infty(\Omega): \alpha\le u(x)\le\beta \ \text{ for a.a. } x\in\Omega\}.$$

Let us notice that the usual function $L(x,y)=\frac12(y-y_d(x))^2$ satisfies Assumption 3 if $y_d\in L^2(\Omega)$.

Remark 1 In *Assumption 3* we made the hypothesis $\alpha<0<\beta$. Since we are looking for sparsity of the optimal control, it does not make sense to consider $0<\alpha$ or $\beta<0$. However, the cases $\alpha=0$ or $\beta=0$ are frequent in practice. In these situations, the sparsity of the optimal control is also induced by the presence of the term $\gamma\|u\|_{L^1(\Omega)}$, see Remark 2. Nevertheless, in this case the L^1 norm is linear on $\mathbb{K}$, hence the cost functional J is differentiable and the control problem (P) falls into the framework of well studied optimal control problems.

The first step in the analysis of (P) is the study of the state equation and the relation control-to-state. This is established in the next theorem.

Theorem 1 *The following statements hold.*

1. *For any $u\in L^p(\Omega)$, with $n/2<p\le\bar p$, there exists a unique solution of* (2.1) $y_u\in Y=H_0^1(\Omega)\cap C(\bar\Omega)$.
2. *The mapping $G:L^p(\Omega)\longrightarrow Y$ defined by $G(u)=y_u$ is of class C^2. Moreover, for $v\in L^p(\Omega)$, $z_v=G'(u)v$ is the unique solution of*

$$\begin{cases} Az+\dfrac{\partial a}{\partial y}(x,y)z=v & \text{in } \Omega, \\ z=0 & \text{on } \Gamma, \end{cases} \tag{2.6}$$

and given $v_1,v_2\in L^p(\Omega)$, $w_{v_1,v_2}=G''(u)(v_1,v_2)$ is the unique solution of

$$\begin{cases} Aw+\dfrac{\partial a}{\partial y}(x,y)w+\dfrac{\partial^2 a}{\partial y^2}(x,y)z_{v_1}z_{v_2}=0 & \text{in } \Omega, \\ w=0 & \text{on } \Gamma, \end{cases} \tag{2.7}$$

where $z_{v_i}=G'(u)v_i$, $i=1,2$.

The existence and uniqueness of a solution of (2.1) in Y is obtained by classical arguments; see, for instance, [2] and [23, Chapter 8]. The differentiability of G can be obtained from the implicit function theorem as follows. We considered the space

$$V = \{y \in Y : Ay \in L^p(\Omega)\}$$

endowed with the graph norm. Now we define the nonlinear operator

$$\mathscr{F} : V \times L^p(\Omega) \longrightarrow L^p(\Omega), \quad \mathscr{F}(y,u) = Ay + a(\cdot,y) - u.$$

Then, it is immediate to check that $\mathscr{F}$ is of class C^2 and $\mathscr{F}(G(u),u) = 0$ for every $u \in L^p(\Omega)$. Using the results of [23, Chapter 8] again, we deduce that

$$\frac{\partial \mathscr{F}}{\partial y}(G(u),u) : V \longrightarrow L^p(\Omega)$$

is an isomorphism. Thus, the assumptions of the implicit function theorem are fulfilled and some simple calculations prove (2.6) and (2.7).

By using the chain rule and the previous theorem we infer the differentiability of F.

Theorem 2 *Functional $F : L^2(\Omega) \longrightarrow \mathbb{R}$ is of class C^2 and the first and second derivatives are given by*

$$F'(u)\,v = \int_\Omega (\varphi_u + \nu\, u)\, v\, dx \tag{2.8}$$

$$F''(u)(v_1,v_2) = \int_\Omega \left\{ \left[\frac{\partial^2 L}{\partial y^2}(x,y_u) - \frac{\partial^2 a}{\partial y^2}(x,y_u)\,\varphi_u \right] z_{v_1} z_{v_2} + \nu\, v_1\, v_2 \right\} dx, \tag{2.9}$$

where $z_{v_i} = G'(u)\,v_i$, $i = 1,2$, and $\varphi_u \in Y$ is the adjoint state defined as the unique solution of

$$\begin{cases} A^*\varphi + \dfrac{\partial a}{\partial y}(x,y_u)\,\varphi = \dfrac{\partial L}{\partial y}(x,y_u) & \text{in } \Omega, \\ \varphi = 0 & \text{on } \Gamma, \end{cases} \tag{2.10}$$

A^* being the adjoint operator of A.

The existence of a solution of (P) can be proved by standard arguments. However, due to the nonconvexity of (P), we have to distinguish between local and global solutions.

Definition 1 We will say that $\bar{u}$ is alocal minimum of (P) in the $L^p(\Omega)$ sense, $1 \le p \le +\infty$, if there exists a ball $B_\varepsilon(\bar{u}) \subset L^p(\Omega)$ such that $J(\bar{u}) \le J(u)$ $\forall u \in \mathbb{K} \cap B_\varepsilon(\bar{u})$. The element $\bar{u}$ will be said a strict local minimum if the inequality $J(\bar{u}) < J(u)$ holds $\forall u \in \mathbb{K} \cap B_\varepsilon(\bar{u})$ with $\bar{u} \ne u$.

Since $\mathbb{K}$ is a bounded subset of $L^\infty(\Omega)$, if $\bar{u}$ is a (strict) local minimum of (P) in the $L^p(Q)$ sense, for some $1 \le p < +\infty$, then $\bar{u}$ is a (strict) local minimum of (P) in the $L^q(Q)$ sense for every $q \in [1,+\infty]$. However, if $\bar{u}$ is a local minimum in the $L^\infty(\Omega)$ sense, it is not necessarily a local minimum in the $L^p(\Omega)$ sense for any $p \in [1,+\infty)$. In the sequel, if nothing is precised, when we say that $\bar{u}$ is a local minimum of (P), it should be intended in the $L^p(\Omega)$ sense for some $p \in [1,+\infty]$.

2.2 First and second order optimality conditions.

To derive the first order optimality conditions we need to say something about the nondifferentiable term j of the cost functional. Since j is convex and Lipschitz, the subdifferential in the sense of convex analysis and the generalized gradients introduced by Clarke coincide. Moreover, a simple computation shows that $\lambda \in \partial j(u)$ if and only if

$$\begin{cases} \lambda(x) = +1 & \text{if } u(x) > 0, \\ \lambda(x) = -1 & \text{if } u(x) < 0, \\ \lambda(x) \in [-1,+1] & \text{if } u(x) = 0. \end{cases} \tag{2.11}$$

As any Lipschitz and convex funcional, j has directional derivatives that can be easily computed

$$j'(u;v) = \lim_{\rho \searrow 0} \frac{j(u+\rho\, v) - j(u)}{\rho} = \int_{\Omega_u^+} v\, dx - \int_{\Omega_u^-} v\, dx + \int_{\Omega_u^0} |v|\, dx, \tag{2.12}$$

for $u, v \in L^1(\Omega)$, where Ω_u^+, Ω_u^- and Ω_u^0 represent the sets of points where u is positive, negative or zero, respectively. Now, we have the following result.

Theorem 3 *If $\bar{u}$ is a local minimum of* (P)*, then there exist $\bar{y}, \bar{\varphi} \in Y$ and $\bar{\lambda} \in \partial j(\bar{u})$ such that*

$$\begin{cases} A\bar{y} + a(x,\bar{y}) = \bar{u} & \text{in } \Omega, \\ \bar{y} = 0 & \text{on } \Gamma, \end{cases} \tag{2.13a}$$

$$\begin{cases} A^*\bar{\varphi} + \dfrac{\partial a}{\partial y}(x,\bar{y})\,\bar{\varphi} = \dfrac{\partial L}{\partial y}(x,\bar{y}) & \text{in } \Omega, \\ \bar{\varphi} = 0 & \text{on } \Gamma, \end{cases} \tag{2.13b}$$

$$\int_\Omega (\bar{\varphi} + \nu\,\bar{u} + \gamma\bar{\lambda})(u - \bar{u})\, dx \geq 0 \quad \forall u \in \mathbb{K}. \tag{2.13c}$$

Sketch of Proof. Since $\mathbb{K}$ is convex, for any $u \in \mathbb{K}$ and $0 \leq \rho \leq 1$, we have that $\bar{u} + \rho(u - \bar{u}) \in \mathbb{K}$. Hence, the local optimality of $\bar{u}$ implies that $J(\bar{u}) \leq J(\bar{u} + \rho(u - \bar{u}))$ for every $\rho > 0$ small enough. Now, from the convexity of j we deduce for every $\rho > 0$ sufficiently small

$$0 \leq \frac{J(\bar{u} + \rho(u - \bar{u})) - J(\bar{u})}{\rho} \leq \frac{F(\bar{u} + \rho(u - \bar{u})) - F(\bar{u})}{\rho} + \gamma j(u) - \gamma j(\bar{u}).$$

Taking the limit as $\rho \to 0$ and using the differentiability of F and (2.8) we get

$$\begin{aligned} 0 &\leq F'(\bar{u})(u - \bar{u}) + \gamma j(u) - \gamma j(\bar{u}) \\ &= \int_\Omega (\bar{\varphi} + \nu\bar{u})(u - \bar{u})\, dx + \gamma j(u) - \gamma j(\bar{u}) \;\; \forall u \in \mathbb{K}, \end{aligned}$$

where $\bar{\varphi}$ satisfies (2.13b). Therefore $\bar{u}$ is the solution of the convex problem

$$\min_{u\in L^\infty(\Omega)} I(u) = \int_\Omega (\bar{\varphi}+\nu\bar{u})u\,dx + \gamma j(u) + I_{\mathbb{K}}(u),$$

where $I_{\mathbb{K}}$ denotes the indicator function of $\mathbb{K}$, taking the value 0 if $u\in\mathbb{K}$ and $+\infty$ if $u\notin\mathbb{K}$. Finally, since I is convex, we can use the subdifferential calculus to obtain

$$0\in\partial I(\bar{u}) = (\bar{\varphi}+\nu\bar{u}) + \gamma\partial j(\bar{u}) + \partial I_{\mathbb{K}}(\bar{u}).$$

Hence, there exists an element $\bar{\lambda}\in\partial j(\bar{u})$ such that $-(\bar{\varphi}+\nu\bar{u}+\gamma\bar{\lambda})\in\partial I_{\mathbb{K}}(\bar{u})$, which is equivalent to (2.13c). □

From (2.13c) we deduce the following corollary.

Corollary 1 *Let $\bar{u}$ be a local minimum of* (P). *Then the following properties hold. If $\nu>0$:*

$$\bar{u}(x) = \mathrm{Proj}_{[\alpha,\beta]}\Big(-\frac{1}{\nu}\big(\bar{\varphi}(x)+\gamma\bar{\lambda}(x)\big)\Big), \tag{2.14a}$$

$$\bar{u}(x) = 0 \quad\Leftrightarrow\quad |\bar{\varphi}(x)|\le\gamma. \tag{2.14b}$$

If $\nu=0$:

$$|\bar{\varphi}(x)|<\gamma \Rightarrow \bar{u}(x)=0, \tag{2.15a}$$

$$\bar{\varphi}(x)>+\gamma \Rightarrow \bar{u}(x)=\alpha, \tag{2.15b}$$

$$\bar{\varphi}(x)<-\gamma \Rightarrow \bar{u}(x)=\beta, \tag{2.15c}$$

$$\bar{\varphi}(x)=+\gamma \Rightarrow \bar{u}(x)\le 0, \tag{2.15d}$$

$$\bar{\varphi}(x)=-\gamma \Rightarrow \bar{u}(x)\ge 0. \tag{2.15e}$$

Moreover, for any $\nu\ge 0$ the function $\bar{\lambda}\in Y$, it is unique and it satisfies

$$\bar{\lambda}(x) = \mathrm{Proj}_{[-1,+1]}\Big(-\frac{1}{\gamma}\bar{\varphi}(x)\Big). \tag{2.16}$$

Finally, if $\nu>0$ then $\bar{u}\in Y$ holds.

The proof of the result for $\nu>0$ can be found in [10] The case $\nu=0$ is analyzed in [4]. Let us point out that the relations (2.14b) and (2.15a) imply the sparsity of local optimal controls. Since $\mathbb{K}$ is a bounded subset of $L^\infty(\Omega)$, it is easy to prove that there exists a constant M independent of γ and ν such that $\|\bar{\varphi}\|_{L^\infty(\Omega)}\le M$. Therefore, if $\gamma>M$, then (2.14b) and (2.15a) prove that $\bar{u}\equiv 0$. However, for $\gamma=0$, then (2.14a) and (2.15b)-(2.15c) prove that $\bar{u}$ is not zero whenever $\bar{\varphi}$ is not zero. Tipically in this case, $\bar{u}(x)\neq 0$ almost everywhere in Ω. Now, we can monitor γ in the interval $[0,M]$ to get an optimal control with a small support.

Remark 2 The sparsity of the local minima of (P) continues to hold in the cases $\alpha = 0$ or $\beta = 0$. Indeed, if $\nu > 0$ and $\alpha = 0$, it is easy to deduce from (2.14a) that $\bar{u}(x) = 0$ if and only if $\bar{\varphi}(x) \geq -\gamma$, which also implies the sparsity. For $\nu > 0$ and $\beta = 0$, we have that $\bar{u}(x) = 0$ if and only if $\bar{\varphi}(x) \leq +\gamma$. In the case $\nu = 0$ and $\alpha = 0$, we have that $\bar{u}(x) = 0$ if $\bar{\varphi}(x) > -\gamma$, there is no information about $\bar{u}(x)$ when $\bar{\varphi}(x) = -\gamma$. Analogously we have that for $\nu = 0$ and $\beta = 0$, then $\bar{u}(x) = 0$ whenever $\bar{\varphi}(x) < +\gamma$.

When $\nu = 0$, if the set of points where $|\bar{\varphi}(x)| = \gamma$ has a zero Lebesgue measure, then $\bar{u}(x)$ only take the values $\{\alpha, 0, \beta\}$ for almost every point $x \in \Omega$. In this case, we can say that $\bar{u}$ is a bang-bang-bang control.

The following example shows how the sparsity of the optimal control can be monitored through the parameter γ. The domain $\Omega = B_1(0) \subset \mathbb{R}^2$ is the unit circle and the state equation (2.1) is given by

$$\begin{aligned} -\Delta y + y^3 &= u \quad \text{in } \Omega, \\ y &= 0 \quad \text{on } \Gamma. \end{aligned}$$

The part L of the objective is the standard tracking functional, i.e. $L(x,y) = \frac{1}{2}(y - y_d(x))^2$ with $y_d(x_1,x_2) = 4\sin(2\pi x_1)\sin(\pi x_2)\mathrm{e}^{x_1}$, We have taken the parameters $\nu = 0.002$, $\alpha = -12$ and $\beta = +12$. The pictures in Fig. 1 show the solution of (P) for the values of $\gamma = 0$ and $\gamma = 2^i \cdot 10^{-3}$ for $i = 0, \ldots, 7$. The pictures are taken from [8].

To formulate the sufficient second order optimality conditions we need to introduce the cone of critical directions. Let $\bar{u} \in \mathbb{K}$ satisfy the first order optimality conditions (2.13). Then, the following cone is a straightforward extension of the corresponding cone for a finite dimensional optimization problem.

$$C_{\bar{u}} = \{v \in L^2(\Omega) : v(x) \begin{cases} \geq 0 \text{ if } \bar{u}(x) = \alpha, \\ \leq 0 \text{ if } \bar{u}(x) = \beta, \end{cases} \text{ and } F'(\bar{u})v + \gamma j'(\bar{u};v) = 0\}. \tag{2.17}$$

The following result was proved in [10].

Theorem 4 *Let us assume that $\bar{u} \in \mathbb{K}$ satisfy the first order optimality conditions* (2.13)*. Then, $C_{\bar{u}}$ is a closed convex cone of $L^2(\Omega)$. Additionally, if $\bar{u}$ is a local solution of* (P)*, then $F''(\bar{u})v^2 \geq 0$ $\forall v \in C_{\bar{u}}$.*

Let us observe that there is no a second order contribution of the functional j, it only appears in the definition of the cone $C_{\bar{u}}$. Next we give a sufficient condition for local optimality. First we recall that in finite dimension a sufficient condition for a local minimum $\bar{u}$ of a functional J is the following: $J'(\bar{u}) = 0$ and $J''(\bar{u})v^2 > 0$ $\forall v \neq 0$. This second order condition is equivalent to the existence of some $\delta > 0$ such that $J''(\bar{u})v^2 \geq \delta\|v\|^2$. However, both second order conditions are not equivalent, in general, in infinite dimensional optimization problems. Therefore, the issue is whether the first second order condition is sufficient for a local minimum in infinite dimensional problems or we need to assume the second condition, or none of both is sufficient. The following example shows that it is not enough to assume, in general, that $J''(\bar{u}) > 0$ for $v \neq 0$.

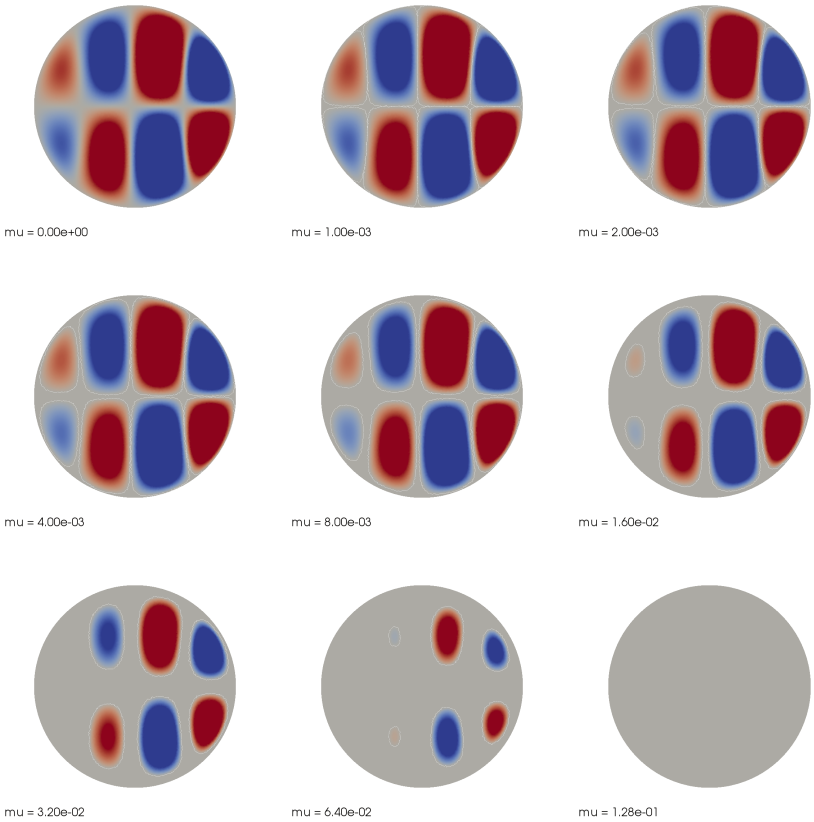

Fig. 1 Solution of an example problem for different values of $\gamma = \mu = 0$ and $\gamma = \mu = 2^i \cdot 10^{-3}$, $i = 0, \ldots, 7$, by piecewise linear controls.

Example 1 Consider the optimization problem

$$\min_{u \in L^\infty(0,1)} J(u) = \int_0^1 [tu^2(t) - u^3(t)]\, dt.$$

The function $\bar{u}(t) \equiv 0$ satisfies the first-order necessary condition $J'(\bar{u}) = 0$ and

$$J''(\bar{u})v^2 = \int_0^1 2tv^2(t)\, dt > 0 \;\; \forall v \in L^\infty(0,1) \setminus \{0\}.$$

However, $\bar{u}$ is not a local minimum of J. Indeed, if we define

$$u_k(t) = \begin{cases} 2t \text{ if } t \in (0, \dfrac{1}{k}), \\ 0 \;\text{ otherwise,} \end{cases}$$

then it holds $J(u_k) = -\frac{1}{k^4} < J(\bar{u})$, and $\|u_k - \bar{u}\|_{L^\infty(0,1)} = \frac{2}{k}$.

It is a classical result that the condition $J''(\bar{u})v^2 \geq \delta\|v\|^2$ is sufficient. Surprisingly, in the control problems, when the Tikhonov term appears in the cost functional ($\nu > 0$), the condition $J''(\bar{u})v^2 > 0$ if $v \neq 0$ is a sufficient condition as well. This results is also valid for the case of control constraints if we restrict v to the cone $C_{\bar{u}}$.

Theorem 5 ([10, Theorem 3.9]) *Assume that $\nu > 0$. Let $\bar{u} \in \mathbb{K}$ and $\bar{\lambda} \in \partial j(\bar{u})$ be such that* (2.13) *hold. Furthermore, let us assume that $F''(\bar{u})\,v^2 > 0$ for all $v \in C_{\bar{u}} \setminus \{0\}$, then there exist $\kappa > 0$ and $\varepsilon > 0$ such that*

$$J(\bar{u}) + \frac{\kappa}{4}\|u-\bar{u}\|^2_{L^2(\Omega)} \leq J(u) \;\; \forall u \in \mathbb{K} \cap B_\varepsilon(\bar{u}), \tag{2.18}$$

where $B_\varepsilon(\bar{u})$ denotes the $L^2(\Omega)$ ball of center $\bar{u}$ and radius ε.

Actually, the situation when $\nu > 0$ is similar to the finite dimensional case. The following theorem is an immediate consequence of [10, Theorm 3.8].

Theorem 6 *Assume that $\nu > 0$. Let $\bar{u} \in \mathbb{K}$ and $\bar{\lambda} \in \partial j(\bar{u})$ be such that* (2.13) *hold. Then the following statements are equivalent.*

1. *$F''(\bar{u})\,v^2 > 0$ for all $v \in C_{\bar{u}} \setminus \{0\}$.*
2. *There exists $\delta > 0$ such that $F''(\bar{u})\,v^2 \geq \delta\,\|v\|^2_{L^2(\Omega)}$ for all $v \in C_{\bar{u}}$.*
3. *There exist $\tau > 0$ and $\delta > 0$ such that $F''(\bar{u})\,v^2 \geq \delta\,\|v\|^2_{L^2(\Omega)}$ for all $v \in C^\tau_{\bar{u}}$, where*

$$C^\tau_{\bar{u}} = \{v \in L^2(\Omega) : v(x) \begin{cases} \geq 0 \text{ if } \bar{u}(x) = \alpha, \\ \leq 0 \text{ if } \bar{u}(x) = \beta, \end{cases} \text{ and } F'(\bar{u})\,v + \gamma\, j'(\bar{u};v) \leq \tau\,\|v\|_{L^2(\Omega)}\}.$$

The assumption $\nu > 0$ is essential in the theorems 5 and 6. When $\nu = 0$, the situation is completely different. Since we have pointwise control constraints, the critical cone $C_{\bar{u}}$ is too small to formulate the sufficient second order conditions. The following example due to Dunn [21] demonstrates that $J''(\bar{u})v^2 \geq \delta\|v\|^2_{L^2(\Omega)}$ for every $v \in C_{\bar{u}}$ is not in general sufficient for local optimality.

Example 2 We define $J : L^2(0,1) \to \mathbb{R}$ by

$$J(u) = \int_0^1 [2a(x)u(x) - \operatorname{sign}(a(x))u(x)^2]\,dx,$$

where $a(x) = 1 - 2x$. The set of admissible functions u is defined by

$$\mathbb{K} := \{u \in L^\infty(0,1) : \; 0 \leq u(x) \leq 2 \text{ for a.a. } x \in [0,1]\},$$

and the optimization problem is

$$\min_{u \in \mathbb{K}} J(u).$$

Let us set $\bar{u}(x) = \max\{0, -a(x)\}$; then $\bar{u}(x) = 0$ holds on $[0,1/2]$ and $0 < \bar{u}(x) < 2$ on $(1/2,2)$. We have

$$\begin{aligned} J'(\bar{u})v &= \int_0^1 2[a(x) - \operatorname{sign}(a(x))\bar{u}(x)]v(x)\,dx = \int_0^1 \bar{d}(x)v(x)dx \\ &= \int_0^{1/2} 2a(x)v(x)dx \geq 0 \end{aligned}$$

for all $v \in L^2(0,1)$ with $v(x) \geq 0$ on $[0,1/2]$. Since $u - \bar{u}$ is nonnegative for all $u \in \mathbb{K}$, $\bar{u}$ satisfies the first order necessary optimality conditions.

Since $\bar{d}(x) > 0$ on $[0,1/2)$, the critical cone for this example is

$$C_{\bar{u}} = \{v \in L^2(0,1) : v(x) = 0 \text{ on } [0,1/2)\}.$$

For all $v \in C_{\bar{u}}$, we obtain

$$\begin{aligned} J''(\bar{u})v^2 =& -\int_0^1 2\,\mathrm{sign}(a(x))\,v^2(x)\,dx = 2\int_{1/2}^1 v^2(x)\,dx - 2\int_0^{1/2} v^2(x)\,dx \\ =& 2\int_{1/2}^1 v^2(x)\,dx = 2\,\|v\|^2_{L^2(0,1)}, \end{aligned}$$

Therefore $J''(\bar{u})v^2 \geq \delta\|v\|^2_{L^2(0,1)}$ $\forall v \in C_{\bar{u}}$ is fulfilled with $\delta = 2$. However, $\bar{u}$ is not a local minimum in $L^2(0,1)$. Indeed, take for $0 < \varepsilon < 1/2$

$$u_\varepsilon(x) = \begin{cases} 3\varepsilon, & \text{if } x \in [\frac{1}{2} - \varepsilon, \frac{1}{2}] \\ \bar{u}(x), & \text{else.} \end{cases}$$

Then we have

$$J(u_\varepsilon) - J(\bar{u}) = \int_{\frac{1}{2}-\varepsilon}^{\frac{1}{2}} [6\varepsilon(1-2x) - 9\varepsilon^2]dx = -3\varepsilon^3 < 0.$$

This example shows that, in general, it is necessary to extend $C_{\bar{u}}$ to a bigger cone to formulate the second order condition. To this end, the cone $C^\tau_{\bar{u}}$ was introduced in [4] to deal with the case $\nu = 0$. There it was proved that a second order condition can be formulated on this extended cone, but it is not the condition that the reader is maybe thinking of. It was also proved in [4] that $J''(\bar{u})v^2 \geq \delta\|v\|^2_{L^2(\Omega)}$ $\forall v \in C^\tau_{\bar{u}}$ never holds. A different condition was given:

Theorem 7 *Assume that* $\nu = 0$ *and* $\bar{u} \in \mathbb{K}$ *satisfies* (2.13). *We also suppose that there exist* $\delta > 0$ *and* $\tau > 0$ *such that*

$$J''(\bar{u})v^2 \geq \delta\|z_v\|^2_{L^2(\Omega)} \quad \forall v \in C^\tau_{\bar{u}},$$

where $z_v = G'(\bar{u})v$. *Then, there exist* $\varepsilon > 0$ *and* $\kappa > 0$ *such that*

$$J(\bar{u}) + \frac{\kappa}{2}\|y_u - \bar{y}\|^2_{L^2(\Omega)} \leq J(u) \quad \forall u \in \mathbb{K} \cap B_\varepsilon(\bar{u}), \tag{2.19}$$

where $B_\varepsilon(\bar{u})$ *is the* $L^2(\Omega)$ *ball centered at* $\bar{u}$ *and radius* ε.

Let us compare (2.18) and (2.19). The inequality (2.18) is crutial to analyze the stability fo the solution $\bar{u}$ of the control problem with respect to perturbations in the data of the problem; see, for instance, [16]. In this paper the authors analyze the stability for a solution of a control problem associated to a semilinear parabolic equation, but the resuts and the methods of proof are identical to the elliptic case. However, (2.19) only allows, in general, to prove the stability of the optimal states

with respect to perturbations of the data of (P). In particular, using (2.19) we can prove that the difference between the optimal states of (P) for $\nu > 0$ and the optimal ones for $\nu = 0$ are of order $o(\sqrt{\nu})$; see [16].

When $\nu > 0$, the sufficient second order condition is a crucial tool for the proof of error estimates among continuous and discrete optimal controls; see [9, 10]. However, for $\nu = 0$, we can get only error estimates for the corresponding optimal states by using (2.19). Under some aditional assumption on the optimal adjoint state, some estimates can be deduced as well for the optimal controls in some cases. This has been done for linear-quadratic control problems in [31].

Finally, let us mention that the analysis for state-constrained control problems of semilinear elliptic equations with sparse controls can be found in [15].

3 Sparse control of semilinear parabolic equations

In this section, we analyze some optimal control problems governed by semilinear parabolic equations where the cost functional involves a functional j acting on the control which promotes the sparsity of the optimal control. The reader is referred to [11] for the proofs of the results of this section. Related references are [5] and [14]. These papers are devoted to the control of the Navier-Stokes and FitzHugh-Nagumo systems respectively by sparse controls.

The control problem is formulated as follows

$$\text{(P)} \begin{cases} \text{Minimize } \ J(u) = F(u) + \gamma j(u) \\ \alpha \le u(x,t) \le \beta \quad \text{for a.a. } (x,t) \in \Omega_T = \Omega \times (0,T) \end{cases}$$

where

$$F(u) = \int_{\Omega_T} L(x,t,y_u(x,t))\,dx\,dt + \frac{\nu}{2}\int_{\Omega_T} u^2(x,t)\,dx\,dt,$$

$j : L^2(\Omega_T) \to \mathbb{R}$ is a Lipschitz continuous and convex but not Frchet differentiable function, $\nu \ge 0$, $\gamma > 0$. The state y_u is the solution of the semilinear, parabolic equation

$$\begin{cases} \partial_t y + Ay + a(x,t,y) = u \ \text{ in } \ \Omega_T, \\ \qquad y = 0 \ \text{ on } \ \Sigma_T = \Gamma \times (0,T), \\ \qquad\qquad y(0) = y_0 \ \text{ in } \ \Omega. \end{cases} \tag{3.1}$$

Here, A is the same linear elliptic operator considered in §2. We mention that it is possible to replace the Dirichlet boundary condition $y_u = 0$ by a Neumann boundary condition $\partial_{n_A} y_u = g$ with $g \in L^p(\Sigma)$, provided that $p > n-1$, then $L^\infty(\Omega_T)$ estimates for the solution of the boundary value problem are obtained.

The analysis of (P) will be done for each of the three following choices for the functional j

$$j_1 : L^1(\Omega_T) \to \mathbb{R},\ \ j_1(u) = \|u\|_{L^1(\Omega_T)} = \int_{\Omega_T} |u(x,t)|\,dx\,dt,$$

$$j_2 : L^2(0,T;L^1(\Omega)) \to \mathbb{R},\ \ j_2(u) = \|u\|_{L^2(0,T;L^1(\Omega))} = \left[\int_0^T \|u(t)\|^2_{L^1(\Omega)}\,dt\right]^{1/2},$$

$$j_3 : L^1(\Omega;L^2(0,T)) \to \mathbb{R},\ \ j_3(u) = \|u\|_{L^1(\Omega;L^2(0,T))} = \int_\Omega \|u(x)\|_{L^2(0,T)}\,dx.$$

When we take $j = j_1$, the corresponding problem (P) will be denoted by (P_1). Analogously, we define the the control problems (P_2) and (P_3) corresponding to the other two functionals j_2 and j_3.

Each of these choices for j induces a different spatio-temporal sparsity pattern for the optimal control, all of them being interesting. The functional j_3 with linear elliptic and parabolic equations has been studied in [24], and the term *directionally sparse* controls was coined there. Due to the linearity of the equation and the convexity of the cost functional, no second-order analysis was necessary in [24].

Besides the first-order necessary conditions, we derive in this section second-order necessary and sufficient conditions for the non-convex problems (P_1)–(P_3), which, in case $\nu > 0$, both use the same cone of critical directions and thus provide the minimal gap between second-order necessary and sufficient conditions. Note that the second-order directional derivatives $v \mapsto j''(u;v^2)$ of the above functionals do not exist in all directions.

3.1 Assumptions and preliminary results.

In this section Ω denotes an open, bounded subset of $\mathbb{R}^n$, $1 \le n \le 3$, with a Lipschitz boundary Γ. The final time $T > 0$ is given. We make the following assumptions on the functions and parameters involved in the control problem (P).

Assumption 1.— The assumptions on A are the same than in §2.

Assumption 2.— The initial datum $y_0 \in L^\infty(\Omega)$ and $a : \Omega_T \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the last variable, satisfying the following assumptions

$$\begin{cases} \exists \psi_a \in L^{\hat p}(0,T;L^{\hat q}(\Omega)) \text{ and } C_a \in \mathbb{R} \text{ such that} \\ a(x,t,y)y \ge \psi_a(x,t) + C_a y^2 \ \text{ for a.a. } (x,t) \in \Omega_T \text{ and } \forall y \in \mathbb{R}, \end{cases} \tag{3.2}$$

$$\begin{cases} a(\cdot,\cdot,0) \in L^{\hat p}(0,T;L^{\hat q}(\Omega)) \text{ and } \forall M > 0\ \exists C_M > 0 \text{ such that} \\ \left|\dfrac{\partial^j a}{\partial y^j}(x,t,y)\right| \le C_M \text{ for a.a. } (x,t) \in \Omega_T, \forall |y| \le M, \text{ with } j = 1,2 \end{cases} \tag{3.3}$$

$$\begin{cases} \forall \rho > 0 \text{ and } \forall M > 0\ \exists \varepsilon_{M,\rho} > 0 \text{ such that for a.a. } (x,t) \in \Omega_T \\ \left|\dfrac{\partial^2 a}{\partial y^2}(x,t,y_2) - \dfrac{\partial^2 a}{\partial y^2}(x,t,y_1)\right| \le \rho, \forall |y_i| \le M \text{ with } |y_2 - y_1| \le \varepsilon_{M,\rho}, \end{cases} \tag{3.4}$$

where $\hat p, \hat q \in [2,+\infty]$, with $\frac{1}{\hat p} + \frac{n}{2\hat q} < 1$.

Assumption 3.— We also assume $-\infty < \alpha < 0 < \beta < +\infty$, $\gamma > 0$, $\nu \geq 0$, and $L : \Omega_T \times \mathbb{R} \longrightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the last variable such that $L(\cdot,\cdot,0) \in L^1(\Omega_T)$. Furthermore, for every $M > 0$ there exists a function $\psi_M \in L^{\hat{p}}(0,T;L^{\hat{q}}(\Omega))$ satisfying

$$\begin{cases} \forall M > 0\ \exists \psi_M \in L^{\hat{p}}(0,T;L^{\hat{q}}(\Omega)) \text{ and } C_M > 0 \text{ such that} \\ \left|\dfrac{\partial L}{\partial y}(x,t,y)\right| \leq \psi_M(x,t)\ \forall |y| \leq M \text{ and for a.a. } (x,t) \in \Omega_T, \\ \left|\dfrac{\partial^2 L}{\partial y^2}(x,t,y)\right| \leq C_M\ \forall |y| \leq M \text{ and for a.a. } (x,t) \in \Omega_T, \end{cases} \tag{3.5}$$

$$\begin{cases} \forall \rho > 0 \text{ and } \forall M > 0\ \exists \varepsilon_{M,\rho} > 0 \text{ such that for a.a. } (x,t) \in \Omega_T \\ \left|\dfrac{\partial^2 L}{\partial y^2}(x,t,y_2) - \dfrac{\partial^2 L}{\partial y^2}(x,t,y_1)\right| \leq \rho, \forall |y_i| \leq M, \text{ and } |y_2 - y_1| \leq \varepsilon_{M,\rho}. \end{cases} \tag{3.6}$$

In the sequel, we will denote the set of feasible controls by

$$\mathbb{K} = \left\{ u \in L^\infty(\Omega_T) : \alpha \leq u(x,t) \leq \beta \ \text{ for a.a. } (x,t) \in \Omega_T \right\}.$$

As usual, we denote by $W(0,T)$ the space of functions $y \in L^2(0,T;H_0^1(\Omega))$ such that $\partial_t y \in L^2(0,T;H^{-1}(\Omega))$. It is well known that every function $y \in W(0,T)$ belongs, after a modification on a set of zero Lebesgue measure, to $C([0,T],L^2(\Omega))$. Now, we analyze the existence, uniqueness and regularity of a solution of (3.1), as well as the differentiability of the relation control-to-state.

Theorem 8 *Under the assumptions 1 and 2, for ever $u \in L^{\hat{p}}(0,T;L^{\hat{q}}(\Omega))$ the equation* (3.1) *has a unique solution $y_u \in W(0,T) \cap L^\infty(\Omega_T)$. Furthermore, the mapping $G : L^{\hat{p}}(0,T;L^{\hat{q}}(\Omega)) \longrightarrow W(0,T) \cap L^\infty(\Omega_T)$, defined by $G(u) = y_u$, is of class C^2. For all elements u, v, v_1 and v_2 of $L^{\hat{p}}(0,T;L^{\hat{q}}(\Omega))$, the functions $z_v = G'(u)v$ and $z_{v_1 v_2} = G''(u)(v_1,v_2)$ are the solutions of the problems*

$$\begin{cases} \dfrac{\partial z}{\partial t} + Az + \dfrac{\partial a}{\partial y}(x,t,y_u)z = v & \text{in } \Omega_T, \\ z = 0 & \text{on } \Sigma_T, \\ z(0) = 0 & \text{in } \Omega, \end{cases} \tag{3.7}$$

and

$$\begin{cases} \dfrac{\partial z}{\partial t} + Az + \dfrac{\partial a}{\partial y}(x,t,y_u)z + \dfrac{\partial^2 a}{\partial y^2}(x,t,y_u)z_{v_1}z_{v_2} = 0 & \text{in } \Omega_T, \\ z = 0 & \text{on } \Sigma_T, \\ z(0) = 0 & \text{in } \Omega, \end{cases} \tag{3.8}$$

respectively.

The proof of the existence and uniqueness of a solution of (3.1) in $W(0,T) \cap L^\infty(\Omega_T)$ is standard. The reader is referred, for instance, to [3] where the arguments

used for a Robin boundary condition can be easily adapted to the Dirichlet case. For the proof of the differentiability we can proceed as follows. We set

$$Y = \Big\{ y \in W(0,T) \cap L^\infty(\Omega_T) : \frac{\partial y}{\partial t} + Ay \in L^{\hat{p}}(0,T;L^{\hat{q}}(\Omega)) \Big\},$$

endowed with the graph norm. Y is a Banach space and $Y \subset C([0,T],L^2(\Omega))$. Observe that $L^{\hat{p}}(0,T;L^{\hat{q}}(\Omega)) \subset L^2(\Omega_T) \subset L^2(0,T;H^{-1}(\Omega))$. Moreover, we have $y(0) = \lim_{t\to 0} y(t) \in L^\infty(\Omega)$ for every $y \in Y$. Now, we define

$$\mathscr{F} : Y \times L^{\hat{p}}(0,T;L^{\hat{q}}(\Omega)) \longrightarrow L^{\hat{p}}(0,T;L^{\hat{q}}(\Omega)) \times L^\infty(\Omega),$$

$$\mathscr{F}(y,u) = (\frac{\partial y}{\partial t} + Ay + a(x,t,y) - u,\ y(0) - y_0).$$

Using that $y \in L^\infty(\Omega_T)$ and (3.3) we deduce that $a(\cdot,\cdot,y) \in L^{\hat{p}}(0,T;L^{\hat{q}}(\Omega))$. Hence, $\mathscr{F}$ is well defined and we can apply the implicit function theorem to deduce that G is of class C^2 and to show that (3.7) and (3.8) represent its first and second derivatives, respectively.

Now, we have the following differentiability result of F.

Theorem 9 *Under the Assumptions 1–3, $F : L^{\hat{p}}(0,T;L^{\hat{q}}(\Omega)) \longrightarrow \mathbb{R}$ is of class C^2. Moreover, for all u, v, v_1 and v_2 of $L^{\hat{p}}(0,T;L^{\hat{q}}(\Omega))$ we have*

$$F'(u)\,v = \int_{\Omega_T} \Big(\frac{\partial L}{\partial y}(x,t,y_u)\, z_v + \nu\, u\, v \Big)\, dx\, dt = \int_{\Omega_T} (\varphi_u + \nu\, u)\, v\, dx\, dt, \tag{3.9}$$

$$F''(u)(v_1,v_2) = \int_{\Omega_T} \Big\{ \Big(\frac{\partial^2 L}{\partial y^2}(x,t,y_u) - \varphi_u \frac{\partial^2 a}{\partial y^2}(x,t,y_u) \Big) z_{v_1} z_{v_2} + \nu\, v_1 v_2 \Big\}\, dx\, dt, \tag{3.10}$$

where $z_{v_i} = G'(u)\, v_i$, $i = 1,2$, and $\varphi_u \in W(0,T) \cap L^\infty(\Omega_T)$ is the solution of

$$\begin{cases} -\dfrac{\partial \varphi}{\partial t} + A^*\varphi + \dfrac{\partial a}{\partial y}(x,t,y_u)\varphi = \dfrac{\partial L}{\partial y}(x,t,y_u) & \text{in } \Omega_T, \\ \varphi = 0 & \text{on } \Sigma_T, \\ \varphi(T) = 0 & \text{in } \Omega, \end{cases} \tag{3.11}$$

A^* being the adjoint operator of A.

The fact that F is of class C^2 is an immediate consequence of Theorem 9 and the chain rule. On the other hand, since $y_u \in L^\infty(\Omega_T)$, we deduce from (3.5) that $\frac{\partial L}{\partial y}(\cdot,\cdot,y_u) \in L^{\hat{p}}(0,T;L^{\hat{q}}(\Omega))$, which implies that φ_u is well defined and enjoys the indicated regularity. The formulas (3.9) and (3.10) follow from standard computations.

Analogously to Y, we define the space

$$\Phi = \big\{ \varphi \in W(0,T) \cap L^\infty(\Omega_T) : -\frac{\partial \varphi}{\partial t} + A^* y \in L^{\hat{p}}(0,T;L^{\hat{q}}(\Omega)) \big\},$$

endowed with the graph norm. As established for Y, we also have the embedding $\Phi \subset C([0,T],L^2(\Omega))$.

We conclude this section with the following theorem, whose proof follows from classical arguments by taking a minimizing sequence.

Theorem 10 *Problem* (P) *has at least one solution* $\bar{u}$.

3.2 First order optimality conditions.

The next theorem states the first-order optimality conditions satisfied by a local minimum of (P); see Definition 1 and the subsequent comments. The proof follows the same steps than the one of Theorem 3, using (3.9) instead of (2.8).

Theorem 11 *If* $\bar{u}$ *is a local minimum of* (P)*, then there exist* $\bar{y} \in Y$*,* $\bar{\varphi} \in \Phi$*, and* $\bar{\lambda} \in \partial j(\bar{u})$ *such that*

$$\begin{cases} \partial_t \bar{y} + A\bar{y} + a(x,t,\bar{y}) = \bar{u} \text{ in } \Omega_T, \\ \bar{y} = 0 \text{ on } \Sigma_T, \\ \bar{y}(0) = y_0 \text{ in } \Omega, \end{cases} \tag{3.12a}$$

$$\begin{cases} -\partial_t \bar{\varphi} + A^* \bar{\varphi} + \dfrac{\partial a}{\partial y}(x,t,\bar{y})\,\bar{\varphi} = \dfrac{\partial L}{\partial y}(x,t,\bar{y}) \text{ in } \Omega_T, \\ \bar{\varphi} = 0 \text{ on } \Sigma_T, \\ \bar{\varphi}(T) = 0 \text{ in } \Omega, \end{cases} \tag{3.12b}$$

$$\int_{\Omega_T} (\bar{\varphi} + \nu\bar{u} + \gamma\bar{\lambda})(u - \bar{u})\,dx\,dt \geq 0 \quad \forall u \in \mathbb{K}. \tag{3.12c}$$

Now, we use the optimality system (3.12) to deduce the sparse structure of $\bar{u}$ for the three choices of j.

Problem (P_1)**.** The subdiferential and the directional derivatives of the functional $j_1 : L^1(\Omega_T) \longrightarrow \mathbb{R}$ were given in (2.11) and (2.12), where it is enough to change Ω by Ω_T.

Corollary 2 *Let* $\bar{u}$*,* $\bar{\varphi}$ *and* $\bar{\lambda}$ *be as in Theorem 11 with* $j = j_1$*, then the following relations hold for almost all* $(x,t) \in \Omega_T$

$$\text{if } \nu > 0 \text{ then} \quad \begin{cases} \bar{u}(x,t) = 0 \Leftrightarrow |\bar{\varphi}(x,t)| \leq \gamma, \\ \bar{u}(x,t) = \mathrm{Proj}_{[\alpha,\beta]}\left(-\dfrac{1}{\nu}[\bar{\varphi}(x,t) + \gamma\bar{\lambda}(x,t)]\right), \end{cases} \tag{3.13}$$

$$\text{if } \nu = 0 \text{ then} \quad \begin{cases} |\bar{\varphi}(x,t)| < \gamma \Rightarrow \bar{u}(x,t) = 0, \\ \bar{\varphi}(x,t) > \quad \gamma \Rightarrow \bar{u}(x,t) = \alpha, \\ \bar{\varphi}(x,t) < -\gamma \Rightarrow \bar{u}(x,t) = \beta, \end{cases} \tag{3.14}$$

$$\bar{\lambda}(x,t) = \mathrm{Proj}_{[-1,+1]}\left(-\frac{1}{\gamma}\bar{\varphi}(x,t)\right). \tag{3.15}$$

Moreover, $\bar{\lambda} \in L^2(0,T;H_0^1(\Omega))$ *holds and it is unique. Finally, if* $\nu > 0$*, we also have that* $\bar{u} \in L^2(0,T;H_0^1(\Omega))$.

In the case $\nu = 0$, (3.14) implies that if the set of points $(x,t) \in \Omega_T$ where $|\bar{\varphi}(x,t)| = \gamma$ has a zero Lebesgue measure (which is expected in many cases), then $\bar{u}(x,t) \in \{\alpha, 0, \beta\}$ for almost all $(x,t) \in \Omega_T$, which means that the optimal control has a bang-bang-bang structure.

From the state equation (3.12a) we can get an $L^\infty(\Omega_T)$ estimate for $\bar{y}$ depending on α and β, but independent of $\bar{u}$. Now using (3.12b), we deduce an estimate $\|\bar{\varphi}\|_{L^\infty(\Omega)} \leq M$ with M independent of $\bar{u}$. Hence, from (3.13) and (3.14) we conclude that $\bar{u} \equiv 0$ if $\gamma > M$. Therefore, we may influence the size of an optimal control's support by adjusting γ in the interval $[0, M]$.

Problem (P_2)**.** The characterization of the subdifferential of j_2 and its directional derivatives are given in the following propositions.

Proposition 1 *For $u \neq 0$, $\lambda \in \partial j_2(u)$ is equivalent to $\lambda \in L^2(0,T;L^\infty(\Omega))$ and*

$$\lambda(x,t) \in \mathrm{Sign}(u(x,t)) \frac{\|u(t)\|_{L^1(\Omega)}}{\|u\|_{L^2(L^1)}} \quad \textit{a.e. in } \Omega_T, \tag{3.16}$$

In case $u = 0$, we have $\partial j_2(u) = \{\lambda \in L^2(0,T;L^\infty(\Omega)) : \|\lambda\|_{L^2(L^\infty)} \leq 1\}$. Moreover, for $\lambda \in \partial j_2(u)$ we have

$$\begin{cases} \operatorname{supp} u^+(t) \subset \left\{x \in \Omega : \lambda(x,t) = +\|\lambda(t)\|_{L^\infty(\Omega)}\right\} \\ \operatorname{supp} u^-(t) \subset \left\{x \in \Omega : \lambda(x,t) = -\|\lambda(t)\|_{L^\infty(\Omega)}\right\} \end{cases} \textit{a.e. in } (0,T). \tag{3.17}$$

Proposition 2 *For every $u, v \in L^2(0,T;L^1(\Omega))$ we have either*

$$j_2'(u;v) = \frac{1}{\|u\|_{L^2(L^1)}} \int_0^T j_\Omega'(u(t);v(t)) \|u(t)\|_{L^1(\Omega)}\, dt \tag{3.18}$$

in case $u \neq 0$ or $j_2'(u;v) = j_2(v)$ in case $u = 0$.

From the previous propositions and Theorem 11 we deduce the following corollary.

Corollary 3 *Let $\bar{u}$, $\bar{\varphi}$ and $\bar{\lambda}$ be as in Theorem 11 and $j = j_2$, then the following relations hold for almost all $(x,t) \in \Omega_T$*

$$\textit{if } \nu > 0 \quad \begin{cases} \bar{u}(x,t) = \mathrm{Proj}_{[\alpha,\beta]}\Big(-\dfrac{1}{\nu}[\bar{\varphi}(x,t) + \gamma\bar{\lambda}(x,t)]\Big), \\ \bar{u}(x,t) = 0 \Leftrightarrow |\bar{\varphi}(x,t)| \leq \gamma\bar{\sigma}(t), \end{cases} \tag{3.19}$$

$$\textit{if } \nu = 0 \quad \begin{cases} |\bar{\varphi}(x,t)| < \gamma\bar{\sigma}(t) \Rightarrow \bar{u}(x,t) = 0, \\ \bar{\varphi}(x,t) > +\gamma\bar{\sigma}(t) \Rightarrow \bar{u}(x,t) = \alpha, \\ \bar{\varphi}(x,t) < -\gamma\bar{\sigma}(t) \Rightarrow \bar{u}(x,t) = \beta, \end{cases} \tag{3.20}$$

$$\bar{\lambda}(x,t) = \mathrm{Proj}_{[-\bar{\sigma}(t),+\bar{\sigma}(t)]}\Big(-\frac{1}{\gamma}\bar{\varphi}(x,t)\Big), \tag{3.21}$$

where

$$\bar{\sigma}(t) = \begin{cases} \dfrac{1}{\|\bar{u}\|_{L^2(L^1)}} \|\bar{u}(t)\|_{L^1(\Omega)} & \text{if } \bar{u} \neq 0 \\ 1 & \text{if } \bar{u} = 0. \end{cases}$$

Moreover, $\bar{\lambda} \in L^2(0,T;H_0^1(\Omega))$ *holds and it is unique. Finally, if* $\nu > 0$*, we also have that* $\bar{u} \in L^2(0,T;H_0^1(\Omega))$.

Analogously to the problem (P_1), we can prove that there exists $M > 0$ independent of γ such that $\bar{u} \equiv 0$ if $\gamma > M$. Hence, we can monitor the parameter γ in the interval $[0,M]$ to get a suitable support for $\bar{u}$.

Problem (P_3)**.** The expression for the subdifferential and the directional derivatives of j_3 are given in the following proposition.

Proposition 3 *The following statements hold.*

1. $\lambda \in \partial j_3(u)$ *is equivalent to* $\lambda \in L^\infty(\Omega;L^2(0,T))$ *and*

$$\begin{cases} \|\lambda(x)\|_{L^2(0,T)} \leq 1 & \text{for a.a. } x \in \Omega_u^0, \\ \lambda(x,t) = \dfrac{u(x,t)}{\|u(x)\|_{L^2(0,T)}} & \text{for a.a. } x \in \Omega_u \text{ and } t \in (0,T), \end{cases} \tag{3.22}$$

where

$$\Omega_u = \{x \in \Omega : \|u(x)\|_{L^2(0,T)} \neq 0\} \ \text{ and } \ \Omega_u^0 = \Omega \setminus \Omega_u.$$

2. *For every* $u,v \in L^1(\Omega;L^2(0,T))$

$$j_3'(u;v) = \int_{\Omega_u^0} \|v(x)\|_{L^2(0,T)}\,dx + \int_{\Omega_u} \frac{1}{\|u(x)\|_{L^2(0,T)}} \int_0^T u\,v\,dt\,dx. \tag{3.23}$$

As a consequence of the above proposition and Theorem 11 we get the following corollary.

Corollary 4 *Let* $\bar{u}$*,* $\bar{\varphi}$ *and* $\bar{\lambda}$ *be as in Theorem 11, then the following relations hold for almost all* $(x,t) \in \Omega_T$

$$\text{if } \nu > 0 \quad \begin{cases} \bar{u}(x,t) = \mathrm{Proj}_{[\alpha,\beta]}\Big(-\dfrac{1}{\nu}[\bar{\varphi}(x,t) + \gamma\bar{\lambda}(x,t)]\Big), \\ \|\bar{u}(x)\|_{L^2(0,T)} = 0 \Leftrightarrow \|\bar{\varphi}(x)\|_{L^2(0,T)} \leq \gamma, \end{cases} \tag{3.24}$$

$$\text{if } \nu = 0 \quad \begin{cases} \|\bar{\varphi}(x)\|_{L^2(0,T)} < \gamma \Rightarrow \|\bar{u}(x)\|_{L^2(0,T)} = 0, \\ \|\bar{u}(x)\|_{L^2(0,T)} = 0 \Rightarrow \|\bar{\varphi}(x)\|_{L^2(0,T)} \leq \gamma, \end{cases} \tag{3.25}$$

$$\bar{\lambda}(x,t) = \begin{cases} -\dfrac{1}{\gamma}\bar{\varphi}(x,t) & \text{if } x \in \Omega_{\bar{u}}^0, \\ \dfrac{\bar{u}(x,t)}{\|\bar{u}(x)\|_{L^2(0,T)}} & \text{if } x \in \Omega_{\bar{u}}. \end{cases} \tag{3.26}$$

Moreover, $\bar{\lambda}$ *is unique.*

Once again, we can prove the existence of some $M > 0$ independent of γ such that $\bar{u} \equiv 0$ if $\gamma > M$.

Remark 3 It is interesting to compare the sparsity properties of the local solutions $\bar{u}$ corresponding to the studied problems. From (3.13) and (3.14) we obtain that the local solutions $\bar{u}$ of (P_1) are sparse in space and time. However, the solutions of (P_3) are only sparse in space as proved by (3.24) and (3.25), the sparsity region remaining constant throughout time. When we look at (3.19) and (3.20), we observe that the sparsity region of the solutions of (P_2) can change with the time. Any of the three formulations can be interesting with different possible applications.

3.3 Second order optimality conditions.

To derive the second order necessary optimality conditions we introduce the cone of critical directions analogously as it was done in §2.2

$$C_{\bar{u}} = \{v \in L^2(\Omega_T) : v(x,t) \begin{cases} \geq 0 \text{ if } \bar{u}(x,t) = \alpha, \\ \leq 0 \text{ if } \bar{u}(x,t) = \beta, \end{cases} \text{ and } F'(\bar{u})v + \gamma j'(\bar{u};v) = 0\}. \quad (3.27)$$

The set $C_{\bar{u}}$ is a closed, convex cone in $L^2(\Omega_T)$. On this cone we formulate the second order necessary conditions. Now we introduce the following notation for the three different functionals under investigation

$$j_1''(u;v^2) = 0, \quad (3.28)$$

$$j_2''(u;v^2) = \frac{1}{\|u\|_{L^2(L^1)}} \left\{ \int_0^T j'_\Omega(u(t);v(t))^2\,dt - j'_2(u;v)^2 \right\} \qquad \text{if } u \neq 0$$
$$j_2''(u;v^2) = 0 \qquad \text{if } u = 0 \quad (3.29)$$

$$j_3''(u;v^2) = \int_{\Omega_u} \frac{1}{\|u(x)\|_{L^2(0,T)}} \left[\int_0^T v^2(x,t)\,dt - \left(\int_0^T \frac{u(x,t)\,v(x,t)}{\|u(x)\|_{L^2(0,T)}}\,dt \right)^2 \right] dx. \quad (3.30)$$

In (3.29), $j_\Omega : L^1(\Omega) \longrightarrow \mathbb{R}$ is defined by $j(u) = \|u\|_{L^1(\Omega)}$ and $j'_\Omega(u;v)$ is given by (2.12). In (3.30), Ω_u is defined as in Proposition 3.

The above expressions don't mean that there exist the second derivatives, they are simply notation. When the second directional derivative exists for some u and some v, then it is given by the above expressions. Concerning $j_3''(u;v^2)$, the reader should observe that a simple application of Schwarz inequality implies that the integrand in Ω_u is nonnegative. Hence, $j_3''(u;v^2)$ is always well defined but it can take the value $+\infty$ in some cases. In the sequel we set $J''(u;v^2) = F''(u)v^2 + \gamma j''(u;v^2)$, where j is one of the three functionals above defined.

Theorem 12 *Let $\nu \geq 0$ and $\bar{u}$ be a local minimum of* (P). *Then $J''(\bar{u};v^2) \geq 0$ for every $v \in C_{\bar{u}}$.*

For the sufficient second order conditions we introduce an extended cone, analogously as we did in §2.2.

$$C_{\bar u}^{\tau}=\{v\in L^2(\Omega_T): v(x)\begin{cases}\geq 0 \text{ if } \bar u(x,t)=\alpha,\\ \leq 0 \text{ if } \bar u(x,t)=\beta,\end{cases} \text{ and } F'(\bar u)\,v+\gamma\, j'(\bar u;v)\leq \tau\,\|v\|_{L^2(\Omega_T)}\}.$$

The second order optimality conditions can be formulated indistinctly on $C_{\bar u}$ or $C_{\bar u}^{\tau}$ if $\nu>0$. However for $\nu=0$, then the cone $C_{\bar u}^{\tau}$ is the correct one. Analogously to Theorem 6 we have the following result.

Theorem 13 *Let us assume that $\nu>0$. Then the following statements are equivalent*

$$J''(\bar u)\,v^2>0 \quad \forall v\in C_{\bar u}\setminus\{0\}, \tag{3.31}$$

$$\exists\delta>0 \text{ and } \tau>0 \text{ such that } J''(\bar u)\,v^2\geq \delta\,\|v\|^2_{L^2(\Omega_T)} \quad \forall v\in C_{\bar u}^{\tau}, \tag{3.32}$$

$$\exists\delta>0 \text{ and } \tau>0 \text{ such that } J''(\bar u)\,v^2\geq \delta\,\|z_v\|^2_{L^2(\Omega_T)} \ \forall v\in C_{\bar u}^{\tau}. \tag{3.33}$$

where $z_v=G'(\bar u)\,v$ is the solution of the linearized parabolic equation (2.8) *corresponding to $y_u=\bar y$.*

Theorem 14 *Let $\nu>0$ and assume that $\bar u\in\mathbb{K}$ satisfies* (3.12) *and* (3.31)*. Then there exist $\varepsilon>0$ and $\kappa>0$ such that*

$$J(\bar u)+\frac{\kappa}{2}\|u-\bar u\|^2_{L^2(\Omega_T)}\leq J(u) \quad \forall u\in\mathbb{K}\cap B_\varepsilon(\bar u), \tag{3.34}$$

where $B_\varepsilon(\bar u)$ denotes the ball of $L^2(\Omega_T)$ centered at $\bar u$ and with radius ε.

For the case $\nu=0$ we have the following result.

Theorem 15 *Let $\nu=0$ and assume that $\bar u\in\mathbb{K}$ satisfies* (3.12) *and* (3.33)*. We also assume that $j=j_1$ or j_2. Then there exist $\varepsilon>0$ and $\kappa>0$ such that*

$$J(\bar u)+\frac{\kappa}{2}\|y_u-\bar y\|^2_{L^2(\Omega_T)}\leq J(u) \quad \forall u\in\mathbb{K}\cap B_\varepsilon(\bar u), \tag{3.35}$$

where $B_\varepsilon(\bar u)$ denotes the ball of $L^2(\Omega_T)$ centered at $\bar u$ and with radius ε.

In the case $\nu=0$ and $j=j_3$, the second order sufficient conditions for local optimality is an open question for the moment.

4 Elliptic control problems in measure spaces with sparse solutions

This section is dedicated to the analysis of the optimal control problem

$$\text{(P)} \qquad \min_{u\in\mathscr{M}(\Omega)} J(u)=\frac{1}{2}\|y_u-y_d\|^2_{L^2(\Omega)}+\gamma\|u\|_{\mathscr{M}(\Omega)},$$

where y_u is the unique solution to the Dirichlet problem

$$\begin{cases} Ay=u \text{ in } \Omega,\\ \ \ y=0 \text{ on } \Gamma,\end{cases} \tag{4.1}$$

A being the operator introduced in §2, satisfying the Assumption 1. We assume that $\gamma > 0$, $y_d \in L^2(\Omega)$ and Ω is a bounded domain in $\mathbb{R}^n$, $n = 2$ or 3 with a Lipschitz boundary Γ. The controls are taken in the space of real regular Borel measures $\mathscr{M}(\Omega)$. As usual, $\mathscr{M}(\Omega)$ is identified by the Riesz theorem with the dual space of $C_0(\Omega)$ – consisting of the continuous functions in $\bar{\Omega}$ vanishing on Γ – endowed with the norm

$$\|u\|_{\mathscr{M}(\Omega)} = \sup_{\|z\|_{C_0(\Omega)} \le 1} \langle u, z \rangle = \sup_{\|z\|_{C_0(\Omega)} \le 1} \int_\Omega z(x)\, du,$$

which is equivalent to the total variation of u.

We will see that this formulation leads to optimal controls which are sparse. This is relevant for many applications in distributed parameter control; see [19]. Moreover, the support of the optimal control provides information on the optimal placements of control actuators. The main advantage of the use of measures as controls is that the support of the optimal controls can be much smaller than the corresponding to the formulation with $L^1(\Omega)$ functions, as in §2. Actually, this supports can have a zero Lebesgue measure. Most of the results presented in this section are proved in [6].

As usual, given a measure $u \in \mathscr{M}(\Omega)$, we say that $y \in L^1(\Omega)$ is a solution to (4.1) if

$$\int_\Omega y A^* \phi \, dx = \int_\Omega \phi \, du \qquad \text{for all } \phi \in \Phi, \tag{4.2}$$

where A^* is the adjoint operator of A and

$$\Phi = \{\phi \in Y : A^*\phi \in L^\infty(\Omega)\}.$$

It is well known, see for instance [30], that there exists a unique solution to (4.1) in the sense of (4.2). Moreover, $y \in W_0^{1,p}(\Omega)$ for every $1 \le p < \frac{n}{n-1}$ and

$$\|y\|_{W_0^{1,p}(\Omega)} \le C_p \|u\|_{\mathscr{M}(\Omega)}. \tag{4.3}$$

Since $W_0^{1,p}(\Omega) \subset L^2(\Omega)$ for very $\frac{2n}{n+2} \le p < \frac{n}{n-1}$, the cost functional is well defined on $\mathscr{M}(\Omega)$. Furthermore, the control-to-state mapping is injective, and therefore the cost functional J is strictly convex. Then, it can be obtained by the standard approach that (P) has a unique solution; see [19] for details. Hereafter, this optimal solution will be denoted by $\bar{u}$ with an associated state $\bar{y}$. By using subdifferential calculus of convex functions and introducing the adjoint state we get the following result (see also [19, 20]).

Theorem 16 *There exists a unique element $\bar{\varphi} \in C(\bar{\Omega}) \cap H_0^1(\Omega)$ satisfying*

$$\begin{cases} A^*\bar{\varphi} = \bar{y} - y_d & \text{in } \Omega, \\ \bar{\varphi} = \quad 0 & \text{on } \Gamma, \end{cases} \tag{4.4}$$

such that

$$\gamma \|\bar{u}\|_{\mathscr{M}(\Omega)} + \int_\Omega \bar{\varphi}\, d\bar{u} = 0, \tag{4.5}$$

$$\|\bar{\varphi}\|_{C_0(\Omega)} \begin{cases} = \gamma & \text{if } \bar{u} \ne 0, \\ \le \gamma & \text{if } \bar{u} = 0. \end{cases} \tag{4.6}$$

Since the control problem is convex, the conditions (4.4)-(4.6) are necessary and sufficient for optimality of $\bar{u}$. Relations (4.5) and (4.6) imply the sparsity of $\bar{u}$. Indeed, this is an immediate consequence of the following lemma whose proof can be found in [7]

Lemma 1 *Let $\mu \in \mathscr{M}(\Omega)$ and $z \in C_0(\Omega)$, both of them not zero, be such that*

$$\langle \mu, z \rangle = \|\mu\|_{\mathscr{M}(\Omega)} \|z\|_\infty,$$

and let $\mu = \mu^+ - \mu^-$ be the Jordan decomposition of μ. Then we have

$$\begin{aligned} \operatorname{supp}(\mu^+) &\subset \Omega_+ = \{x \in \Omega : z(x) = +\|z\|_\infty\}, \\ \operatorname{supp}(\mu^-) &\subset \Omega_- = \{x \in \Omega : z(x) = -\|z\|_\infty\}. \end{aligned}$$

Combining Lemma 1 and (4.5)-(4.6) we deduce

$$\begin{cases} \operatorname{supp}(\bar{u}^+) \subset \{x \in \Omega : \bar{\varphi}(x) = -\gamma\}, \\ \operatorname{supp}(\bar{u}^-) \subset \{x \in \Omega : \bar{\varphi}(x) = +\gamma\}. \end{cases} \tag{4.7}$$

As the numerical results show, the set $\{x \in \Omega : |\bar{\varphi}(x)| = \gamma\}$ is small, which yields the sparsity of $\bar{u}$. Moreover, it can be proved that tere exists a real number $M > 0$ independent of γ such that $\bar{u} \equiv 0$ if $\gamma > M$.

Remark 4 There is a very interesting paper by Pieper and Vexler [28] where the following regularity result is proved: under the assumption $y_d \in L^\infty(\Omega)$, then $\bar{y} \in H_0^1(\Omega) \cap L^\infty(\Omega)$, and consequently $\bar{u} \in \mathscr{M}(\Omega) \cap H^{-1}(\Omega)$. Hence, one can imagine that the optimal control is supported on curves in the two-dimensional case, or on surfaces in the three-dimensional case. Dirac measures are excluded and measures concentrated on curves are also excluded for $n = 3$.

In [6], the numerical analysis of (P) is carried out by using a convenient finite element approximation of $\mathscr{M}(\Omega)$. In particular, convergence of the optimal controls and error estimates for the difference between the continuous and discrete optimal states are stablished; see [28] for some improved error estimates. Since $\mathscr{M}(\Omega)$ is not a separable Banach space, we cannot get strong convergence of the discretizations, but we can prove

$$\bar{u}_h \overset{*}{\rightharpoonup} \bar{u} \text{ in } \mathscr{M}(\Omega) \text{ and } \|\bar{u}_h\|_{\mathscr{M}(\Omega)} \to \|\bar{u}\|_{\mathscr{M}(\Omega)} \text{ as } h \to 0.$$

The discrete problem can be solved by a semismooth Newton method. Let us illustrate the theoretical results presented above with an example taken from [6]. In the computations we have taken $\Omega = (-1,+1)^2$, the operator is $A = -\Delta$, $\gamma = 10^{-2}$, and $y_d(x) = 10\exp(-50|x|^2)$. The optimal control is shown in Figure 2.

Before finishing this section let us address a natural question that the reader is probably wondering. What happen if the state equation is semilinear? Let us consider the equation (2.1) with $u \in \mathscr{M}(\Omega)$. The associated control problem was studied in [12]. The main difficulty is the solvability of the state equation. Indeed, it is well

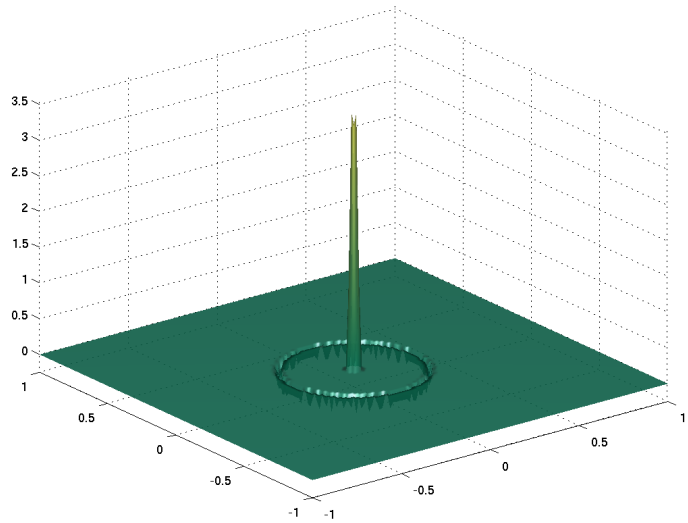

Fig. 2 Optimal Control

known that (2.1) has no solution if we take $A = -\Delta$, $a(x,y) = y^3$ and $u = \delta_{x_0}$ with $x_0 \in \Omega$; see [1]. This difficulty is overcome in [12] in two different ways. First we assume that the growth order of $a(x,y)$ with respect to y is polynomial of arbitrary degree if $n = 2$ and degree < 3 if $n = 3$. Under this assumption we have existence and uniqueness of a solution of the state equation. A second approach consists of taking the controls in a subspace of $\mathscr{M}(\Omega)$ for which (2.1) has a unique solution for arbitrary nonlinearities of $a(x,y)$ with respect to y. Of course, as in §2, we require the function a to be monotone nondecreasing with respect to y. This subspace, denoted by $\mathscr{M}_\infty(\Omega)$, is formed by the measures $u \in \mathscr{M}(\Omega)$ such that the solution of the Dirichlet problem

$$\begin{cases} -\Delta y = u & \text{in } \Omega, \\ \quad y = 0 & \text{on } \Gamma \end{cases}$$

belongs to $L^\infty(\Omega)$. For every $u \in \mathscr{M}_\infty(\Omega)$, (2.1) has a unique solution $y_u \in H_0^1(\Omega) \cap L^\infty(\Omega)$. Actually, the embedding $\mathscr{M}_\infty(\Omega) \subset H^{-1}(\Omega)$ holds. If we recall Remark 4, we can conclude that $\mathscr{M}_\infty(\Omega)$ is the natural space for the controls. However, since $\mathscr{M}_\infty(\Omega)$ is not a closed subspace of $\mathscr{M}(\Omega)$, the proof of the existence of a solution for the associated control problem is not obvious, but it was established in [12]. In both approaches we get the optimality system satisfied by the locally optimal controls, and we prove the sparsity given by (4.7).

5 Parabolic control problems in measure spaces with sparse solutions

In this section we consider control problems for parabolic equations where the controls are measures. As we mentioned in §3, the possibilities to formulate a control problem leading to sparse optimal controls are higher for the parabolic case than for the elliptic one. Here we will show three different formulations that are the natural extensions of the three problems studied in §3. We restrict the formulation to linear equations because of the difficulty of dealing with measures for nonlinear parabolic equations. Additionally the cost functional will involve a tracking type term for the state plus the norm of the correponding control. In all these formulations the state equation will be

$$\begin{cases} \partial_t y + Ay = u \ \text{ in } \ \Omega_T, \\ \qquad\quad y = 0 \ \text{ on } \ \Sigma_T, \\ \quad\ y(0) = y_0 \ \text{ in } \ \Omega, \end{cases} \tag{5.1}$$

where $y_0 \in L^2(\Omega)$, and A is defined as in §2, enjoying the same assumptions. We also assume that $\Omega \subset \mathbb{R}^n$ with $1 \le n \le 3$ and Γ is Lipschitz. We denote the solution of (5.1) by y_u. Now the three problems are formulated as follows:

$$(\mathrm{P}_1) \qquad \min_{u \in \mathscr{M}(\Omega_T)} \quad J(u) = \tfrac{1}{q}\|y_u - y_d\|^q_{L^q(\Omega_T)} + \gamma\|u\|_{\mathscr{M}(\Omega_T)},$$

$$(\mathrm{P}_2) \qquad \min_{u \in L^2((0,T),\mathscr{M}(\Omega))} \quad J(u) = \tfrac{1}{2}\|y_u - y_d\|^2_{L^2(\Omega_T)} + \gamma\|u\|_{L^2((0,T),\mathscr{M}(\Omega))},$$

$$(\mathrm{P}_3) \qquad \min_{u \in \mathscr{M}(\Omega,L^2(0,T))} \quad J(u) = \tfrac{1}{2}\|y_u - y_d\|^2_{L^2(\Omega_T)} + \gamma\|u\|_{\mathscr{M}(\Omega,L^2(0,T))}.$$

We analyze each of these problems in the next subsections.

5.1 Analysis of (P_1).

The problem (P_1) has been studied in [13]. In this problem the controls are measures in Ω_T. If u is a function, then its norm in the measure space coincides with the $L^1(\Omega_T)$-norm. Hence, this can be considered as an extension of the functional j_1 introduced in §3 to the space of measures. The term involving the state is different of the corresponding term for the problems (P_2) and (P_3). The reason for that is the lack of regularity of the solution of (5.1) when $u \in \mathscr{M}(\Omega_T)$. We say that a function $y \in L^1(Q)$ is a solution of (5.1) if the following identity holds

$$\int_{\Omega_T} (-\frac{\partial \phi}{\partial t} + A^*\phi) y \, dx \, dt = \int_{\Omega_T} \phi \, du + \int_{\Omega} y_0 \phi(0) \, dx, \quad \forall \phi \in \Phi, \tag{5.2}$$

where

$$\Phi = \{\phi \in L^2(0,T;H^1_0(\Omega)) : -\frac{\partial \phi}{\partial t} + A^*\phi \in L^\infty(\Omega_T) \ \text{ and } \ \phi(x,T) = 0 \text{ in } \Omega\}.$$

Let us observe that the problem

$$\begin{cases} -\dfrac{\partial \phi}{\partial t} + A^*\phi = f \text{ in } Q \\ \phi(x,T) = 0 \text{ in } \Omega \\ \phi(x,t) = 0 \text{ on } \Sigma \end{cases} \tag{5.3}$$

has a unique solution $\phi \in L^2(0,T;H_0^1(\Omega)) \cap C([0,T];L^2(\Omega))$ for every $f \in L^\infty(\Omega_T)$. Moreover, the regularity $\phi \in C(\bar{\Omega}_T)$ holds; see [26, Chapter 3].

Then we have the following result of existence, uniqueness and regularity; see, for instance, [13].

Theorem 17 *There exists a unique solution y of* (5.1). *Moreover,* $y \in L^r(0,T;W_0^{1,p}(\Omega))$ *for all* $p,r \in [1,2)$ *with* $(2/r)+(d/p) > n+1$, *and the following estimate holds*

$$\|y\|_{L^r(0,T;W_0^{1,p}(\Omega))} \le C_{r,p}(\|u\|_{\mathscr{M}(\Omega_T)}. \tag{5.4}$$

From this theorem we deduce that $y_u \in L^q(\Omega_T)$ if $1 \le q < \min\{2, \frac{n+2}{n}\}$. This is the reason to replace the $L^2(\Omega_T)$-norm in the cost functional by the $L^q(\Omega_T)$-norm. Next we will assume that $1 < q < \min\{2, \frac{n+2}{n}\}$. The solution of (P_1) is characterized for the first order optimality conditions expressed in the following theorem.

Theorem 18 *Let* $\bar{u}$ *denote a solution to* (P_1) *with associated state* $\bar{y}$. *Then, there exists an element* $\bar{\varphi} \in L^2(0,T;H_0^1(\Omega)) \cap C(\bar{\Omega}_T)$ *satisfying*

$$\begin{cases} -\dfrac{\partial \bar{\varphi}}{\partial t} + A^*\bar{\varphi} = \bar{g} \text{ in } \Omega_T \\ \bar{\varphi}(x,T) = 0 \text{ in } \Omega \\ \bar{\varphi}(x,t) = 0 \text{ on } \Sigma, \end{cases} \tag{5.5}$$

$$\begin{cases} \displaystyle\int_{\Omega_T} \bar{\varphi}\, d\bar{u} + \gamma\|\bar{u}\|_{\mathscr{M}(\Omega_T)} = 0 \\ \|\bar{\varphi}\|_{C(\bar{\Omega}_T)} \begin{cases} = \gamma \text{ if } \bar{u} \neq 0 \\ \le \gamma \text{ if } \bar{u} = 0, \end{cases} \end{cases} \tag{5.6}$$

where $\bar{g}(x,t) = |\bar{y}(x,t) - y_d(x,t)|^{q-2}(\bar{y}(x,t) - y_d(x,t))$.

The continuity of $\bar{\varphi}$ follows from the fact that $g \in L^{q'}(\Omega_T)$ and $q' > 5/2$. From Lemma 1 and (5.6) we deduce the following corollary that shows the sparse character of the optimal control.

Corollary 5 *Under the assumptions and notations of Theorem 18 we have that*

$$\begin{cases} \operatorname{supp}(\bar{u}^+) \subset \{(x,t) \in \bar{\Omega}_T : \bar{\varphi}(x,t) = -\gamma\} \\ \operatorname{supp}(\bar{u}^-) \subset \{(x,t) \in \bar{\Omega}_T : \bar{\varphi}(x,t) = +\gamma\} \end{cases} \tag{5.7}$$

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

5.2 Analysis of (P_2).

This section is based on the paper [7]. Hereafter we denote by $L^2((0,T),\mathscr{M}(\Omega))$ the space of weakly measurable functions $u:[0,T]\to\mathscr{M}(\Omega)$ for which the norm

$$\|u\|_{L^2((0,T),\mathscr{M}(\Omega))}=\left(\int_0^T\|u(t)\|^2_{\mathscr{M}(\Omega)}\,dt\right)^{1/2}$$

is finite. This choice makes $L^2((0,T),\mathscr{M}(\Omega))$ a Banach space that can be identified with the dual of $L^2((0,T),C_0(\Omega))$, where the duality relation is given by

$$\langle u,z\rangle=\int_0^T\langle u(t),z(t)\rangle_{\mathscr{M}(\Omega),C_0(\Omega)}\,dt,$$

The reader is referred to [22, section 8.14.1 and Proposition 8.15.3] for the different notions of measurability and [22, Theorem 8.20.3] for the duality identification.

Given $u\in L^2((0,T),\mathscr{M}(\Omega))$, we say that $y\in L^1(\Omega_T)$ is a solution of (5.1) if the following identity holds

$$\int_{\Omega_T}(-\frac{\partial\phi}{\partial t}+A^*\phi)y\,dx\,dt=\int_0^T\langle u(t),\phi(t)\rangle\,dt+\int_\Omega y_0\phi(0)\,dx,\quad\forall\phi\in\Phi,\tag{5.8}$$

where

$$\Phi=\{\phi\in L^2(0,T;H_0^1(\Omega)):-\frac{\partial\phi}{\partial t}+A^*\phi\in L^\infty(\Omega_T)\ \text{ and }\ \phi(x,T)=0\text{ in }\Omega\}.$$

Then we have the following theorem.

Theorem 19 *For all $(u,y_0)\in L^2((0,T),\mathscr{M}(\Omega))\times L^2(\Omega)$ the equation* (5.1) *has a unique solution y. Moreover, $y\in L^2((0,T),W_0^{1,p}(\Omega))$ for every $p\in[1,\frac{n}{n-1})$ and there exist constants C_p such that*

$$\|y\|_{L^2((0,T),W_0^{1,p}(\Omega))}\le C_p\left(\|u\|_{L^2((0,T),\mathscr{M}(\Omega))}+\|y_0\|_{L^2(\Omega)}\right).\tag{5.9}$$

As a consequence of this theorem is immediate that (P_2) has at least a solution. Moreover, due to its strict convexity we get the uniqueness of such a solution. Hereafter $\bar u$ will denote the solution to (P_2) and $\bar y$ the associated state. Now, we give the first order optimality conditions, which are necessary and sufficient due to the convexity of (P_2).

Theorem 20 *There exists a unique element $\bar\varphi\in W(0,T)\cap L^2((0,T),C_0(\Omega))$ satisfying*

$$\begin{cases}-\partial_t\bar\varphi+A^*\bar\varphi=\bar y-y_d & \text{in }\Omega_T,\\ \bar\varphi=0 & \text{on }\Sigma_T,\\ \bar\varphi(x,T)=0 & \text{in }\Omega,\end{cases}\tag{5.10}$$

such that

$$\int_0^T \langle \bar{u}(t), \bar{\varphi}(t)\rangle\, dt + \gamma \|\bar{u}\|_{L^2((0,T),\mathscr{M}(\Omega))} = 0, \tag{5.11}$$

$$\|\bar{\varphi}\|_{L^2((0,T),C_0(\Omega))} \begin{cases} = \gamma \text{ if } \bar{u} \neq 0, \\ \leq \gamma \text{ if } \bar{u} = 0. \end{cases} \tag{5.12}$$

From (5.11) and (5.12) we deduce the sparse property of the optimal control $\bar{u}$. Let us consider the Jordan decomposition $\bar{u}(t) = \bar{u}^+(t) - \bar{u}^-(t)$ for almost every $t \in I$. Then we have the following theorem.

Corollary 6 *For almost every $t \in (0,T)$ the following embeddings hold*

$$\operatorname{supp}(\bar{u}^+(t)) \subset \{x \in \Omega : \bar{\varphi}(x,t) = -\|\bar{\varphi}(t)\|_\infty\}, \tag{5.13}$$

$$\operatorname{supp}(\bar{u}^-(t)) \subset \{x \in \Omega : \bar{\varphi}(x,t) = +\|\bar{\varphi}(t)\|_\infty\}. \tag{5.14}$$

5.3 Analysis of (P_3).

The reader is referred to [25] for a detailed study of this control problem. First of all we make some comments on the control space $\mathscr{M}(\Omega, L^2(0,T))$. Let us denote the σ-algebra of Borel sets in Ω by $\mathscr{B}(\Omega)$. The space $\mathscr{M}(\Omega, L^2(0,T))$ is formed by the set of all countably additive vector measures $u : \mathscr{B}(\Omega) \longrightarrow L^2(0,T)$ of bounded total variation. For $u \in \mathscr{M}(\Omega, L^2(0,T))$ the total variation measure $|u| \in \mathscr{M}(\Omega)$ is defined by

$$|u|(B) = \sup\Big\{ \sum_{k=1}^{\infty} \|u(B_k)\|_{L^2(0,T)} : \{B_k\}_{k=1}^{\infty} \subset \mathscr{B}(\Omega) \text{ is a disjoint partition of } B \Big\}.$$

It is easy to check that

$$\|u(B)\|_{L^2(0,T)} \leq |u|(B) \quad \forall B \in \mathscr{B}(\Omega).$$

$\mathscr{M}(\Omega, L^2(0,T))$ is a Banach space when endowed with the norm

$$\|u\|_{\mathscr{M}(\Omega, L^2(0,T))} = \||u|\|_{\mathscr{M}(\Omega)} = |u|(\Omega);$$

see [27, Chapter 12.3]. Given $u \in \mathscr{M}(\Omega, L^2(0,T))$, the mapping $t \in (0,T) \to u(t) \in \mathscr{M}(\Omega)$ is well defined for almost all $t \in (0,T)$. In fact we have that $\mathscr{M}(\Omega, L^2(0,T)) \subsetneq L^2((0,T), \mathscr{M}(\Omega))$, the embedding being continuous. The identity is false. Indeed, if we take $u(t) = g(t)\delta_t$ with $g \in L^2(0,T)$, then $u \in L^2((0,T), \mathscr{M}(\Omega))$, but $u \notin \mathscr{M}(\Omega, L^2(0,T))$. From the embedding $\mathscr{M}(\Omega, L^2(0,T)) \subsetneq L^2((0,T), \mathscr{M}(\Omega))$ we deduce that Theorem 19 holds for the controls $u \in \mathscr{M}(\Omega, L^2(0,T))$. We also have the existence and uniqueness of an optimal control $\bar{u}$, solution of (P_3).

If we denote by $\bar\varphi$ the adjoint state associated with $\bar u$, which is the solution of (5.10), then we have [25, Theorem 2.11]

$$-\frac{1}{\gamma}\int_0^T \langle \bar u(t), \bar\varphi(t)\rangle\, dt + \|\bar u\|_{\mathscr{M}(\Omega,L^2(0,T))} \le \|u\|_{\mathscr{M}(\Omega,L^2(0,T))} \quad \forall u \in \mathscr{M}(\Omega, L^2(0,T)).$$

This leads to the following sparse property of $\bar u$

$$\operatorname{supp}(|\bar u|) \subset \{x \in \Omega : \|\bar\varphi(x)\|_{L^2(0,T)} = \gamma\}. \tag{5.15}$$

If we compare the sparse structure of $\bar u$ defined through (5.13)-(5.14) and (5.15) we can observe the same differences that we already found in the comparison of (3.19)-(3.20) and (3.24)-(3.25); see Remark 3.

Related with the material presented in this section is the controllability of parabolic equations by sparse measures; see [17] and [18].

References

1. Ph. Benilan and H. Brezis, *Nonlinear problems related to the Thomas-Fermi equation, dedicated to Philippe Benilan*, J. Evol. Equ. **3** (2004), 673–770.
2. E. Casas, *Boundary control of semilinear elliptic equations with pointwise state constraints*, SIAM J. Control Optim. **31** (1993), no. 4, 993–1006.
3. ———, *Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations*, SIAM J. Control Optim. **35** (1997), no. 4, 1297–1327.
4. ———, *Second order analysis for bang-bang control problems of PDEs*, SIAM J. Control Optim. **50** (2012), no. 4, 2355–2372.
5. E. Casas and K. Chrysafinos, *Analysis of the velocity tracking control problem for the 3d evolutionary navier-stokes equations*, ESAIM:COCV, **23** (2017), 263-295.
6. E. Casas, C. Clason, and K. Kunisch, *Approximation of elliptic control problems in measure spaces with sparse solutions*, SIAM J. Control Optim. **50** (2012), no. 4, 1735–1752.
7. ———, *Parabolic control problems in measure spaces with sparse solutions*, SIAM J. Control Optim. **51** (2013), no. 1, 28–63.
8. E. Casas, R. Herzog, and G. Wachsmuth, *Approximation of sparse controls in semilinear elliptic equations*, Large-scale scientific computing, Lecture Notes in Comput. Sci., vol. 7116, Springer, Heidelberg, 2012, pp. 16–27.
9. E. Casas, R. Herzog, and G. Wachsmuth, *Approximation of sparse controls in semilinear equations by piecewise linear functions*, Numer. Math. **122** (2012), 645–669.
10. ———, *Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional*, SIAM J. Optim. **22** (2012), no. 3, 795–820.
11. ———, *Analysis of spatio-temporally sparse optimal control problems of semilinear parabolic equations*, ESAIM Control Optim. Calc. Var. **23** (2017), 263–295.
12. E. Casas and K. Kunisch, *Optimal control of semilinear elliptic equations in measure spaces*, SIAM J. Control Optim. **52** (2013), no. 1, 339–364.
13. ———, *Parabolic control problems in space-time measure spaces*, ESAIM Control Optim. Calc. Var. **22** (2016), no. 2, 355–370.
14. E. Casas, C. Ryll, and F. Tröltzsch, *Sparse optimal control of the Schlögl and FitzHugh-Nagumo systems*, Comput. Methods Appl. Math. **13** (2014), 415–442.
15. E. Casas and F. Tröltzsch, *Second-order and stability analysis for state-constrained elliptic optimal control problems with sparse controls*, SIAM J. Control Optim. **52** (2014), no. 2, 1010–1033.
16. ———, *Second order optimality conditions for weak and strong local solutions of parabolic optimal control problems*, Vietnam J. Math. **44** (2016), no. 1, 181–202.
17. E. Casas, B. Vexler, and E. Zuazua, *Sparse initial data identification for parabolic PDE and its finite element approximations*, Math. Control Relat. Fields **5** (2015), no. 3, 377–399.
18. E. Casas and E. Zuazua, *Spike controls for elliptic and parabolic pde*, Systems Control Lett. **62** (2013), 311–318.

19. C. Clason and K. Kunisch, *A duality-based approach to elliptic control problems in non-reflexive Banach spaces*, ESAIM Control Optim. Calc. Var. **17** (2011), 243–266, DOI: 10.1051/cocv/2010003.
20. C. Clason and K. Kunisch, *A measure space approach to optimal source placement*, Comput. Optim. Appl. **53** (2012), no. 1, 155–171.
21. J.C. Dunn, *On second order sufficient optimality conditions for structured nonlinear programs in infinite-dimensional function spaces*. In Fiacco, A. (ed.) Mathematical Programming with Data Perturbations, pp. 83–107, Marcel Dekker, New York (1998).
22. R.E. Edwards, *Functional analysis*, Holt, Rinehart and Winston, New York, 1965.
23. D. Gilbarg and N.S. Trudinger, *Elliptic partial differential equations of second order*, Springer-Verlag, Berlin, 1977.
24. R. Herzog, G. Stadler, and G. Wachsmuth, *Directional sparsity in optimal control of partial differential equations*, SIAM J. Control Optim. **50** (2012), no. 2, 943–963.
25. K: Kunisch, K. Pieper, and B. Vexler, *Measure valued directional sparsity for parabolic optimal control problems*, SIA: J. Control Optim. **52** (2014), no. 5, 3078–3108.
26. O.A. Ladyzhenskaya, V.A. Solonnikov, and N.N. Ural'tseva, *Linear and quasilinear equations of parabolic type*, American Mathematical Society, Providence (1988).
27. S. Lang, *Real analysis, 2nd ed.*, Addison Wesley, Redison, MA, 1983.
28. K. Pieper and B. Vexler, *A priori error analysis for discretization of sparse elliptic optimal control problems in measure space*, SIAM J. Control Optim. **51** (2013), no. 4, 2788–2808.
29. G. Stadler, *Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices*, Comput. Optim. Appl. **44** (2009), no. 2, 159–181.
30. G. Stampacchia, *Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus*, Ann. Inst. Fourier (Grenoble) **15** (1965), 189–258.
31. D. Wachsmuth and G. Wachsmuth, *Convergence and regularisation results for optimal control problems with sparsity functional*, ESAIM Control Optim. Calc. Var. **17** (2011), no. 3, 858–886.
32. _____, *Regularization error estimates and discrepancy principle for optimal control problems with inequality constraints*, Control Cybernet. **40** (2011), no. 4, 1125–1158.