MATHEMATICAL CONTROL AND
RELATED FIELDS
Volume **5**, Number **3**, September **2015**

doi:10.3934/mcrf.2015.5.377

pp. **377–399**

SPARSE INITIAL DATA IDENTIFICATION FOR PARABOLIC PDE AND ITS FINITE ELEMENT APPROXIMATIONS

Eduardo Casas

Departamento de Matemática Aplicada y Ciencias de la Computación
E.T.S.I. Industriales y de Telecomunicación, Universidad de Cantabria
39005 Santander, Spain

Boris Vexler

Centre for Mathematical Sciences, Technische Universität München
Bolzmannstrasse 3, D-85747 Garching b. München, Germany

Enrique Zuazua(1),(2)

(1) Basque Center for Applied Mathematics (BCAM)
Alameda Mazarredo 14, E-48009 Bilbao, Basque Country, Spain
(2) IKERBASQUE, Basque Foundation for Science
E-48013 Bilbao, Basque Country, Spain

Abstract. We address the problem of inverse source identification for parabolic equations from the optimal control viewpoint employing measures of minimal norm as initial data. We adopt the point of view of approximate controllability so that the target is not required to be achieved exactly but only in an approximate sense. We prove an approximate inversion result and derive a characterization of the optimal initial measures by means of duality and the minimization of a suitable quadratic functional on the solutions of the adjoint system. We prove the sparsity of the optimal initial measures showing that they are supported in sets of null Lebesgue measure. As a consequence, approximate controllability can be achieved efficiently by means of controls that are activated in a finite number of pointwise locations. Moreover, we discuss the finite element numerical approximation of the control problem providing a convergence result of the corresponding optimal measures and states as the discretization parameters tend to zero.

1. **Introduction.** In this paper we address the issue of the backward resolution of parabolic equations formulated as a control problem, the control being the initial datum aiming to steer the solution to a given final value in a given time horizon.

We adopt the point of view of approximate controllability as in [13] where this problem was formulated and solved for initial data in L^p with $1 < p \leq \infty$. In the present paper we consider the case $p = 1$ or, to be more precise, the one in which

2010 *Mathematics Subject Classification.* 35K15, 49K20, 93B05, 93C20.

Key words and phrases. Parabolic equations, approximate controllability, sparse controls, Borel measures.

The first author was supported by Spanish Ministerio de Economía y Competitividad under project MTM2011-22711. The third author was supported by the Advanced Grant NUMERIWAVES/FP7-246775 of the European Research Council Executive Agency, FA9550-14-1-0214 of the EOARD-AFOSR, FA9550-15-1-0027 of AFOSR, the BERC 2014-2017 program of the Basque Government, the MTM2011-29306 and SEV-2013-0323 Grants of the MINECO, the CIMI-Toulouse Excellence Chair in PDEs, Control and Numerics and a Humboldt Award at the University of Erlangen-Nürnberg.

the initial data to be optimized are Borel measures. The range of the semigroup departing from this class of initial data is dense, since that is the case when $p > 1$ too.

Here we focus on the characterization of the initial measures of minimal norm ensuring that the final target is reached in an approximate sense. As we shall see, this can be done minimizing a suitable quadratic functional on the solutions of the adjoint system, following the arguments in [10].

Our analysis leads to sparsity results showing that the optimal initial measures have as support a set of zero Lebesgue measure. This also leads to effective inversion results within the class of initial data constituted by finite combinations of Dirac measures.

Our results apply in a broad class of parabolic equations, possibly semilinear with globally Lipschitz nonlinearities, as in [13]. But we shall focus on the classical heat equation with constant coefficients

$$\left\{\begin{array}{rcll} \dfrac{\partial y}{\partial t} - \Delta y & = & 0 & \text{in } Q = \Omega \times (0,T), \\ y & = & 0 & \text{on } \Sigma = \Gamma \times (0,T), \\ y(0) & = & u & \text{in } \Omega, \end{array}\right. \tag{1}$$

where the initial datum u, that plays the role of the control, is assumed to be a Borel measure, $\Omega \subset \mathbb{R}^n$ is an open connected bounded set and Γ is the boundary of Ω, that we will assume to be Lipschitz.

We view u as the control that we would like to choose such that the associated state y_u at time T, $y_u(T)$, is in the $L^2(\Omega)$-ball $\bar{B}_\varepsilon(y_d)$, where y_d represents the desired final state and $\epsilon > 0$ the admissible distance to the target.

It is well known that for any $\epsilon > 0$ it is possible to find $u \in L^2(\Omega)$ such that $y_u(T) \in \bar{B}_\varepsilon(y_d)$. This is a consequence of the fact that the range of the semigroup generated by the heat equation is dense (see [13]), which is equivalent to the well known and classical backward uniqueness property for the heat equation, see [24]. In fact the same holds when the control u has its support in a subset ω of Ω of positive measure. The interested reader is referred to [14] where this issue is investigated in a more general frame in order to obtain sharp bounds on the cost of approximate control.

We are interested on analyzing the structure of the initial data u of minimal norm. In the L^2-setting this can be done by considering the following minimization problem

$$\min_{y_u(T) \in \bar{B}_\varepsilon(y_d)} J(u) = \frac{1}{2}\|u\|^2_{L^2(\Omega)}. \tag{2}$$

It can be checked that this problem has a unique solution that is given by

$$\bar{u} = -\bar{\varphi}(0),$$

where $\bar{\varphi}$ is the unique solution of the adjoint equation

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

for some $\bar{g} \in L^2(\Omega)$ satisfying

$$\int_\Omega \bar{g}(x)(y(x) - \bar{y}(x,T))\,dx \le 0 \quad \forall y \in \bar{B}_\varepsilon(y_d).$$

Above, $\bar{y}$ denotes the state associated to the optimal initial datum $\bar{u}$. In this way, the obtained initial datum, being the trace at $t = 0$ of a solution of the adjoint system in the time interval $0 \leq t \leq T$, is smooth but non-zero at almost every point of Ω. This makes these initial data to be of little practical use in applications where one looks for initial data with small support.

This issue was recently addressed in [23], the goal being to develop efficient numerical algorithms to compute initial data constituted by a finite combination of Dirac measures. This was achieved by means of a fast Bregman iterative algorithm for ℓ_1 optimization. In that frame, of course, identifying the initial datum consists in determining a finite number of points for the support of the measure and the weight given to each of them.

In the present paper we develop the theory showing that sparse initial data exist and are characterized by means of a minimization principle over the class of solutions of the adjoint system. The proof requires two ingredients that are by now well known in the literature. On one hand, as mentioned above, the backward uniqueness for parabolic problems and, on the other, analyticity properties of solutions of the heat equation and, more precisely, the analyticity of solutions in Ω with respect to the space variable at the final time.

We will develop this program following closely the previous work [10], where the control was assumed to be a space-time dependent measure, acting as an exterior source. There it was shown that, by replacing the L^2-norm of the control, by its measure-norm turns out to be an efficient way to obtain sparse optimal controls with support in small regions.

To be more precise, we shall consider the problem above but by replacing in (2) the L^2-norm of the initial datum u, that plays the role of the control, by its total measure.

The main difficulty in the problem under consideration is the very strong irreversibility of the heat equation. The backward heat equation is ill-posed and, due to parabolic regularizing effects, the range of the generated semigroup at time $t = T$ is only constituted by very smooth functions. Thus, of course, it is impossible to reach exactly any given target in $L^2(\Omega)$. On the other hand, the range of the semigroup is dense and accordingly, the target can be approximated at any distance ε. This ensures the existence of solutions of the optimal control problem above. Note, however, that this existence result requires the regularizing effect ensuring that solutions of the heat equation departing from a measure belongs to L^2 at the final time. As we shall see, at the level of the adjoint equation this is reflected by the fact that the solutions of the adjoint system departing (at time $t = T$) from an L^2-datum, belong to the space of continuous functions at the initial time $t = 0$.

Using this adjoint methodology we shall prove the existence and uniqueness of the optimal measure.

The second issue we discuss is the sparsity of the obtained optimal initial measures. As we shall see, these initial measures, by duality, are supported on the set where the adjoint solution at time $t = 0$ reaches its L^∞-norm that, by the space analyticity turns out to be a set of null Lebesgue measure.

As proved in [4], a measure in Ω can be efficiently approximated by a combination of Dirac measures. As a consequence, we deduce that the approximate reachability can be achieved by activating the initial datum in some finite number of pointwise locations at the time $t = 0$. These pointwise locations are all of them placed in a very small region of Ω. Even more, in the one dimensional case, we prove that the optimal measure is a finite combination of Dirac measures, which provides a rigorous justification of the existence of the objects that are numerically approximated in [23].

Moreover, we provide a systematic approach to the discretization of the sparse optimal control problem under consideration with finite elements. To this end we discretize the state equation with a discontinuous Galerkin method in time, which is here a variant of the implicit Euler scheme, and using (conforming) linear finite elements in space. The discretiztaion of the optimal control problem is done in the spirit of [4]. For the resulting finite dimensional optimization problem we provide a convergence result for discretization parameters tending to zero, see Theorem 4.10 for details.

Let us briefly comment on some other related papers. Sparse optimal control problems in measure spaces are analyzed in [11, 26, 7] for elliptic and in [5, 21] for parabolic equations, where in both last papers also numerical analysis of space-time discretizations is performed, see also [15, 22] for related works on numerical analysis of pointwise control. Sparse optimal control problems with controls which are functions instead of measures are considered, e.g., in [6, 9, 19, 30, 32].

The plan of the paper is as follows. In the next section, we formulate the control problem in a precise way and, by means of backward uniqueness and standard regularizing effects for the heat equation (from L^2 into C_0) we infer that it has a unique solution. Later, in section §3, we consider the adjoint system and formulate the dual problem. In Section §4 we present a discrete version of the optimal control problem under consideration and prove the convergence as the discretization parameters tend to zero. Some final remarks are given in §5.

2. The control problem: Main results.

In this paper, we consider the following control problem

$$\text{(P)} \quad \left\{ \begin{array}{l} \min J(u) = \|u\|_{M(\Omega)}, \\ (u, y_u(T)) \in M(\Omega) \times \bar{B}_\varepsilon(y_d), \end{array} \right.$$

where y_u is the solution of the equation (1) associated to u, and $M(\Omega) = C_0(\Omega)^*$ denotes the Banach space of real and regular Borel measures in Ω, $C_0(\Omega)$ being the space of continuous functions in $\bar{\Omega}$ vanishing on Γ. In this space, the norm is defined by

$$\|u\|_{M(\Omega)} = |u|(\Omega) = \sup_{z \in C_0(\Omega),\ \|z\|_\infty \le 1} \int_\Omega z\, du,$$

with $|u|$ being the total variation measure associated to u; see, for instance, [28, Chapter 6].

The function y_d is fixed in $L^2(\Omega)$ and $\varepsilon > 0$ is given. To avoid trivial situations, we assume that $\|y_d\|_{L^2(\Omega)} > \varepsilon$. The case $\|y_d\|_{L^2(\Omega)} \le \varepsilon$ obviously leads to $\bar{u} = 0$ as the unique solution of the problem (P).

Before proving any property of (P), let us study the state equation (1). To this end let us first observe that the problem

$$\left\{ \begin{array}{rcll} \dfrac{\partial \varphi}{\partial t} + \Delta\varphi & = & f & \text{in } Q \\ \varphi(x,t) & = & 0 & \text{on } \Sigma \\ \varphi(x,T) & = & 0 & \text{in } \Omega \end{array} \right. \tag{3}$$

has a unique solution $\varphi \in L^2(0,T;H_0^1(\Omega)) \cap C([0,T];L^2(\Omega))$ for every $f \in L^\infty(Q)$. Moreover, the regularity $\varphi \in C(\bar{Q})$ holds. This continuity property follows from [16, Theorem 6.8] on general Lipschitz domains, see also [1] and [3, Theorem 5.1] or [27].

This allows to introduce the spaces

$$\Phi = \left\{ \varphi \in L^2(0,T;H_0^1(\Omega)) : \frac{\partial \varphi}{\partial t} + \Delta\varphi \in L^\infty(Q) \right\}$$

and

$$\Phi_T = \{\varphi \in \Phi : \varphi(x,T) = 0 \text{ in } \Omega\}.$$

Using the space Φ_T we define a solution concept for the state equation (1).

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

$$\int_Q -\left(\frac{\partial \varphi}{\partial t} + \Delta\varphi\right) y\, dxdt = \int_\Omega \varphi(0)\, du, \quad \forall \varphi \in \Phi_T. \tag{4}$$

As a consequence we have the following:

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

$$\|y\|_{L^r(0,T;W_0^{1,p}(\Omega))} + \|y(T)\|_{L^2(\Omega)} \le C_{r,p}\|u\|_{M(\Omega)}. \tag{5}$$

Furthermore, y satisfies

$$\int_\Omega \varphi(T)y(T)\, dx - \int_Q \left(\frac{\partial \varphi}{\partial t} + \Delta\varphi\right) y\, dxdt = \int_\Omega \varphi(0)\, du, \quad \forall \varphi \in \Phi. \tag{6}$$

Proof. For the proof of the first part of the Lemma, the reader is referred to [8, Theorem 2.2]. We only need to prove the estimate for $y(T)$ and (6). The identity (6) is obvious for regular data $u \in L^2(\Omega)$. Then, it is enough to take a sequence $\{u_k\}_{k=1}^\infty \subset L^2(\Omega)$ such that $u_k \stackrel{*}{\rightharpoonup} u$ in $M(\Omega)$, to write the equation for (u_k, y_k) and to pass to the limit as $k \to \infty$. In this limit, we only need to pay attention to the fact that $\varphi(0) \in C_0(\Omega)$ due to the regularizing effect of the heat equation. To prove the estimate for $y(T)$ let us define $\varphi \in L^2(0,T;H_0^1(\Omega)) \cap C([0,T], L^2(\Omega))$ solution of

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

Classical results on the gain of integrability and smoothing for the heat equation imply that

$$\|\varphi(0)\|_{C_0(\Omega)} \le c\|y(T)\|_{L^2(\Omega)}.$$

Then, using (6) we get

$$\|y(T)\|_{L^2(\Omega)}^2 = \int_\Omega \varphi(T)y(T)\, dx = \int_\Omega \varphi(0)\, du$$
$$\le \|u\|_{M(\Omega)}\|\varphi(0)\|_{C_0(\Omega)} \le C\|u\|_{M(\Omega)}\|y(T)\|_{L^2(\Omega)},$$

which implies (6). ∎

Remark 1. Choosing $r = 1$ in this lemma, we get $y \in L^1(0,T;W^{1,p}(\Omega))$ for all $1 \leq p < \frac{n}{n-1}$. This implies $y_t \in L^1(0,T;W^{-1,p}(\Omega))$ and therefore we conclude $y \in C([0,T];W^{-1,p}(\Omega))$.

The next lemma makes again use of the smoothing property of the heat equations.

Lemma 2.3. *Let $\{u_k\}_{k=1}^\infty \subset M(\Omega)$ be a sequence converging to u weakly* in $M(\Omega)$. If $\{y_k\}_{k=1}^\infty$ and y denote the corresponding states, solutions of (1), then the convergence $y_k(T) \to y(T)$ holds strongly in $L^2(\Omega)$.*

Proof. We consider the linear mapping $A_t\colon M(\Omega) \to L^2(\Omega)$, which maps a control $u \in M(\Omega)$ to the solution $y(t)$ of the (1) at time $0 < t \leq T$. By the previous lemma this mapping is linear and continuous. Moreover we consider an operator $B_{t_1,t_2}\colon L^2(\Omega) \to L^2(\Omega)$ mapping an initial condition $z_1 \in L^2(\Omega)$ of the following equation

$$\left\{\begin{array}{rcll} \dfrac{\partial z}{\partial t} - \Delta z &=& 0 & \text{in } \Omega \times (t_1,t_2), \\ z &=& 0 & \text{on } \Gamma \times (t_1,t_2), \\ z(t_1) &=& z_1 & \text{in } \Omega, \end{array}\right. \tag{7}$$

to the solution $z(t_2)$ at $t = t_2$ for $0 < t_1 < t_2 \leq T$. By a standard result we have $z \in L^2(t_1,t_2;H_0^1(\Omega)) \cap C([t_1,t_2],L^2(\Omega))$ and

$$\int_{t_1}^{t_2} \|\nabla z(t)\|^2_{L^2(\Omega)}\,dt \leq \|z_1\|^2_{L^2(\Omega)}. \tag{8}$$

For a given z_1 we consider a sequence $\{z_{1,m}\} \subset H_0^1(\Omega)$ with $z_{1,m} \to z_1$ in $L^2(\Omega)$. Let z_m be the solution of (7) with $z_{1,m}$ instead of z_1. By standard regularity we have that $z_m \in H^1(t_1,t_2;L^2(\Omega)) \cap C([t_1,t_2];H_0^1(\Omega))$. Then we test the weak formulation for z_m with $(t-t_1)\frac{\partial z_m}{\partial t}$ and obtain

$$(t-t_1)\left\|\frac{\partial z_m}{\partial t}(t)\right\|^2_{L^2(\Omega)} + (t-t_1)\int_\Omega \nabla z_m(t)\nabla\frac{\partial z_m}{\partial t}(t)\,dx = 0.$$

This results in

$$(t-t_1)\left\|\frac{\partial z_m}{\partial t}(t)\right\|^2_{L^2(\Omega)} + \frac{1}{2}\frac{d}{dt}\left((t-t_1)\|\nabla z_m(t)\|^2\right) = \frac{1}{2}\|\nabla z_m(t)\|^2.$$

Integrating in time we get

$$\int_{t_1}^{t_2}(t-t_1)\left\|\frac{\partial z_m}{\partial t}(t)\right\|^2_{L^2(\Omega)}dt + \frac{1}{2}(t_2-t_1)\|\nabla z_m(t_2)\|^2_{L^2(\Omega)} = \frac{1}{2}\int_{t_1}^{t_2}\|\nabla z_m(t)\|^2\,dt.$$

The last term can be estimated as (8) resulting in

$$(t_2-t_1)\|\nabla z_m(t_2)\|^2_{L^2(\Omega)} \leq \|z_{1,m}\|^2_{L^2(\Omega)}.$$

By the linearity of (7) we have for all $m, m' \in \mathbb{N}$

$$(t_2-t_1)\|\nabla\left(z_m(t_2) - z_{m'}(t_2)\right)\|^2_{L^2(\Omega)} \leq \|z_{1,m} - z_{1,m'}\|^2_{L^2(\Omega)}$$

and therefore $\{z_m(t_2)\}$ is a Cauchy sequence in $H_0^1(\Omega)$ converging strongly to $z(t_2)$. This implies $z(t_2) \in H_0^1(\Omega)$ and the well-known smoothing estimate of the heat equation

$$\|\nabla z(t_2)\|_{L^2(\Omega)} \leq \frac{1}{\sqrt{t_2-t_1}}\|z_1\|_{L^2(\Omega)}.$$

Due to the compact embedding of $H_0^1(\Omega)$ into $L^2(\Omega)$ the operator $B_{t_1,t_2}\colon L^2(\Omega) \to L^2(\Omega)$ is compact. Moreover, there holds

$$A_T u = B_{t_1,T} A_{t_1} u$$

for any $0 < t_1 < T$. Therefore, the operator $A_T\colon M(\Omega) \to L^2(\Omega)$ is also compact. This implies the statement of the lemma. □

Now we prove the existence and uniqueness of solution of (P). Let us observe that the uniqueness is proved despite that the cost functional J is not strictly convex.

Theorem 2.4. *There exists a unique solution $\bar u$ of the control problem* (P)*. Moreover, if $\bar y$ denotes the state associated to $\bar u$, then the identity $\|\bar y(T) - y_d\|_{L^2(\Omega)} = \varepsilon$ holds.*

Proof. First, from the approximate controllability properties of the heat equation, we know that the set of feasible controls is not empty; see [13].

The existence of a solution can be easily proved by taking a minimizing sequence and using the compactness the mapping of $u \in M(\Omega) \mapsto y_u(T) \in L^2(\Omega)$ established in Lemma 2.3.

Let us prove the uniqueness. To this end, we first observe that if $u \in M(\Omega)$, $u \neq 0$ and $\|y_u(T) - y_d\|_{L^2(\Omega)} < \varepsilon$, then we can take $0 < \lambda < 1$ such that

$$\|y_u(T) - y_d\|_{L^2(\Omega)} + \lambda \|y_u(T)\|_{L^2(\Omega)} \le \varepsilon.$$

Setting $u_\lambda = (1-\lambda)u$, then $J(u_\lambda) = (1-\lambda)J(u) < J(u)$ and $\|y_{u_\lambda}(T)-y_d\|_{L^2(\Omega)} \le \varepsilon$, hence u is not a solution of (P). Therefore, to any solution $\bar u$ of (P) corresponds an optimal state $\bar y$ such that $\bar y(T)$ is on the boundary of the ball $\bar B_\varepsilon(y_d)$.

Let us assume that u_1 and u_2 are two solutions of (P). We will prove that $u_1 = u_2$. First we note that the identity $y_{u_1}(T) = y_{u_2}(T)$ holds. Indeed, if $y_{u_1}(T) \neq y_{u_2}(T)$, then we take $u = (u_1 + u_2)/2$. Using the convexity of J, the strict convexity of the $L^2(\Omega)$-norm and the identity

$$\|y_{u_1}(T) - y_d\|_{L^2(\Omega)} = \|y_{u_2}(T) - y_d\|_{L^2(\Omega)} = \varepsilon,$$

we infer

$$J(u) \le \frac{1}{2}(J(u_1) + J(u_2)) = \inf\,(\mathrm{P}) \quad \text{and} \quad \|y_u(T) - y_d\|_{L^2(\Omega)} < \varepsilon,$$

which is not possible as we proved above. Finally, we take $u = u_1 - u_2 \neq 0$. For this control we obtain $y_u(T) = y_{u_1}(T) - y_{u_2}(T) = 0$. Then, for every $g \in L^2(\Omega)$ let φ_g be the solution of the problem

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

Then, since y_u is a regular function in $[\frac{1}{k}, T] \times \Omega$ for every $k \ge 1$, we have that

$$\int_\Omega \varphi_g\Big(\frac{1}{k}\Big) y_u\Big(\frac{1}{k}\Big)\,dx = \int_\Omega g y_u(T)\,dx = 0 \quad \forall g \in L^2(\Omega).$$

Since the space $S = \{\varphi_g(\frac{1}{k}) : g \in L^2(\Omega)\}$ is dense in $L^2(\Omega)$ due to the approximate controllability properties of the heat equation, we conclude that $y_u(\frac{1}{k}) = 0$. By Remark 1, we have that $y_u \in C([0,T], W^{-1,p}(\Omega))$ for $1 \le p < \frac{n}{n-1}$, and hence

$$u = y_u(0) = \lim_{k\to\infty} y\left(\frac{1}{k}\right) = 0,$$

which proves that $u_1 = u_2$. □

The next theorem characterizes the solution $\bar{u}$ of (P).

Theorem 2.5. *Let $\bar{u} \in M(\Omega)$ such that $\bar{y}(T) \in \bar{B}_\varepsilon(y_d)$, where $\bar{y}$ is the state associated to $\bar{u}$. Then, $\bar{u}$ is the solution of problem* (P) *if and only if there exist two elements $\bar{g} \in L^2(\Omega)$ and $\bar{\varphi} \in L^2([0,T], H_0^1(\Omega)) \cap C([0,T], L^2(\Omega))$ such that*

$$\int_\Omega \bar{g}(x)(y(x) - \bar{y}(x,T))\,dx \le 0 \quad \forall y \in \bar{B}_\varepsilon(y_d), \tag{9}$$

$$\begin{cases} \dfrac{\partial \bar{\varphi}}{\partial t} + \Delta \bar{\varphi} &= 0 \quad \text{in } Q, \\ \bar{\varphi} &= 0 \quad \text{on } \Sigma, \\ \bar{\varphi}(T) &= \bar{g} \quad \text{in } \Omega, \end{cases} \tag{10}$$

$$\|\bar{u}\|_{M(\Omega)} = -\int_\Omega \bar{\varphi}(x,0)\,d\bar{u}, \tag{11}$$

$$\|\bar{\varphi}(0)\|_{C_0(\Omega)} = 1. \tag{12}$$

Furthermore, $\bar{\varphi}$ and $\bar{g}$ are unique, and there exists a real number $\bar{\lambda} > 0$ such that $\bar{g} = \bar{\lambda}(\bar{y}(T) - y_d)$.

Proof. Let us consider the linear mapping $A \in \mathcal{L}(M(\Omega), L^2(\Omega))$, defined by $Au = y_u(T)$. The continuity of A follows from (5). We formulate (P) in an equivalent way. To this end we define the functional $\mathcal{J} : M(\Omega) \longrightarrow (-\infty, +\infty]$ by

$$\mathcal{J}(u) = J(u) + I_{\bar{B}_\varepsilon(y_d)}(Au),$$

where $I_{\bar{B}_\varepsilon(y_d)}$ denotes the indicator function of the ball $\bar{B}_\varepsilon(y_d)$; which means that it vanishes in $\bar{B}_\varepsilon(y_d)$ and takes the value $+\infty$ outside. The problem (P) can be reformulated as the minimization of the convex functional $\mathcal{J}$. Then, $\bar{u}$ is a solution of (P) if and only if $0 \in \partial\mathcal{J}(\bar{u})$. Now, we apply the rules of the sub-differential calculus of convex functions; see, for instance, [12, Chapter 1, §5.3]. In particular, we can apply the chain rule because according to the proof of Theorem 2.4, there exists $u_0 \in M(\Omega)$ such that $Au_0 \in B_\varepsilon(y_d)$, which means that the Slater condition is fulfilled, consequently

$$0 \in \partial\mathcal{J}(\bar{u}) \subset \partial J(\bar{u}) + A^*\partial I_{\bar{B}_\varepsilon(y_d)}(\bar{y}(T)).$$

This implies that there exists $\bar{g} \in \partial I_{\bar{B}_\varepsilon(y_d)}(\bar{y}(T))$ such that $-A^*\bar{g} \in \partial J(\bar{u})$. Relation (9) is precisely the definition of $\bar{g} \in \partial I_{\bar{B}_\varepsilon(y_d)}(\bar{y}(T))$. Now, we take $\bar{\varphi}$ as the solution of (10). Then, from (6) we deduce

$$\langle A^*\bar{g}, u\rangle = \langle \bar{g}, Au\rangle = \int_\Omega \bar{g}(x) y_u(x,T)\,dx = \int_\Omega \bar{\varphi}(x,0)\,du \quad \forall u \in M(\Omega).$$

Combining this identity with the definition of $-A^*\bar{g} \in \partial J(\bar{u})$

$$\langle -A^*\bar{g}, u - \bar{u}\rangle + J(\bar{u}) \le J(u) \quad \forall u \in M(\Omega),$$

we obtain

$$\int_\Omega \bar{\varphi}(x,0)\,d\bar{u} - \int_\Omega \bar{\varphi}(x,0)\,du + \|\bar{u}\|_{M(\Omega)} \le \|u\|_{M(\Omega)} \quad \forall u \in M(\Omega).$$

Taking $u = 2\bar{u}$ and $\bar{u}/2$, respectively, in the above inequality, we get (11). Therefore, the above inequality and (11) imply

$$-\int_\Omega \bar{\varphi}(x,0)\,du \le \|u\|_{M(\Omega)} \quad \forall u \in M(\Omega).$$

For any point $x_0 \in \Omega$ we select $u = \pm\delta_{x_0}$ in the above inequality, which shows that $\pm\bar\varphi(x_0, 0) \le 1$. Since x_0 is arbitrary in Ω, we get that $\|\bar\varphi(0)\|_{C_0(\Omega)} \le 1$. This inequality along with (11) and the fact that $\bar u \ne 0$ imply (12).

Finally, we prove the uniqueness of $\bar g$, the corresponding uniqueness for $\bar\varphi$ being an immediate consequence. From (9) it follows

$$\int_\Omega \bar g(y - y_d)\,dx \le \int_\Omega \bar g(\bar y(T) - y_d)\,dx \quad \forall y \in \bar B_\varepsilon(y_d),$$

or equivalently

$$\int_\Omega \bar g y\,dx \le \int_\Omega \bar g(\bar y(T) - y_d)\,dx \quad \forall y \in \bar B_\varepsilon(0),$$

which implies the existence of some positive number $\bar\lambda$ such that $\bar g = \bar\lambda(\bar y(T) - y_d)$. Observe that $\bar g \ne 0$ because $\|\bar\varphi(0)\|_{C_0(\Omega)} = 1$. Moreover, the last identity implies that $\bar\lambda$ is uniquely determined, which concludes the proof. □

As a consequence of the previous theorem we get the desired sparsity of the optimal measure. In the sequel $|A|$ will denote the Lebesgue measure of a measurable set $A \subset \Omega$.

Corollary 1. *Let $\bar u$ be the solution of* (P) *and consider the Jordan decomposition of the measure $\bar u$: $\bar u = \bar u^+ - \bar u^-$. Then, the following inclusions hold*

$$\operatorname{supp}(\bar u^+) \subset \Omega^- = \{x \in \Omega : \bar\varphi(x,0) = -1\}, \tag{13}$$

$$\operatorname{supp}(\bar u^-) \subset \Omega^+ = \{x \in \Omega : \bar\varphi(x,0) = +1\}. \tag{14}$$

Furthermore, we have that $|\Omega^+| = |\Omega^-| = 0$. In addition, in dimension $n = 1$, there exist finitely many points $\{x_j\}_{j=1}^m \subset \Omega$ and real numbers $\{\bar\lambda_j\}_{j=1}^m$ such that

$$\bar u = \sum_{j=1}^m \bar\lambda_j \delta_{x_j} \quad \text{with} \quad \bar\lambda_j \begin{cases} \le 0 & \text{if } \bar\varphi(x_j,0) = +1, \\ \ge 0 & \text{if } \bar\varphi(x_j,0) = -1. \end{cases} \tag{15}$$

Proof. Let us denote $\Omega_{\bar u^+} = \operatorname{supp}(\bar u^+)$ and $\Omega_{\bar u^-} = \operatorname{supp}(\bar u^-)$. From (11) and (12) we deduce

$$\begin{aligned} \|\bar u\|_{M(\Omega)} = |\bar u|(\Omega) &= \int_{\Omega_{\bar u^+}} d\bar u^+ + \int_{\Omega_{\bar u^-}} d\bar u^- \\ &\ge -\int_{\Omega_{\bar u^+}} \bar\varphi(x,0)\,d\bar u^+ + \int_{\Omega_{\bar u^-}} \bar\varphi(x,0)\,d\bar u^- \\ &= -\int_\Omega \bar\varphi(x,0)\,d\bar u = \|\bar u\|_{M(\Omega)}. \end{aligned}$$

Hence,

$$\int_{\Omega_{\bar u^+}} (1 + \bar\varphi(x,0))d\bar u^+ + \int_{\Omega_{\bar u^-}} (1 - \bar\varphi(x,0))d\bar u^- = 0,$$

therefore

$$\int_{\Omega_{\bar u^+}} (1 + \bar\varphi(x,0))d\bar u^+ = \int_{\Omega_{\bar u^-}} (1 - \bar\varphi(x,0))d\bar u^- = 0.$$

These identities imply (13) and (14). On the other hand, because of the properties of the heat equation, we know that the function $x \in \Omega \longrightarrow \bar\varphi(x,0) \in \mathbb{R}$ is analytic; see, for instance, [20]. Hence, the maximum and minimum values of this function are achieved in a set of points having zero Lebesgue measure, unless it is a constant function. This is not the case because $\bar\varphi(x,0) = 0$ for $x \in \Gamma$ and $\|\bar\varphi\|_{C_0(\Omega)} = 1$.

This proves that $|\Omega^+| = |\Omega^-| = 0$. Finally, if $n = 1$, then the set of points where a non constant analytic function achieves the extreme values is finite. Using (13) and (14) this directly implies (15). $\square$

3. **The adjoint formulation.** So far we have proved that the system (1) can be approximately controlled by using Borel measures with sparse support. To do this we have followed a direct approach, just looking for the minimum of problem (P). However, the analysis of the approximate controllability of the heat equation has traditionally followed a different approach. Indeed, the approximate controllability of (1) by using L^2 initial controls has been obtained by studying the adjoint optimization problem

$$(\mathrm{P}_\varepsilon) \quad \min_{g\in L^2(\Omega)} J_\varepsilon(g) = \frac{1}{2}\int_\Omega |\varphi_g(x,0)|^2\,dxdt + \varepsilon\|g\|_{L^2(\Omega)} - \int_\Omega y_d g\,dx,$$

where φ_g is the solution of the adjoint system

$$\left\{\begin{array}{rcll} \dfrac{\partial\varphi}{\partial t} + \Delta\varphi & = & 0 & \text{in } Q,\\ \varphi & = & 0 & \text{on } \Sigma,\\ \varphi(T) & = & g & \text{in } \Omega. \end{array}\right. \tag{16}$$

This problem has a unique solution in $L^2(0,T;H_0^1(\Omega))\cap C([0,T];L^2(\Omega))$ for every element $g \in L^2(\Omega)$. The functional J_ε is well defined, continuous and strictly convex. The coercivity is more delicate to prove but, as pointed out in [13] it is a consequence of the backward uniqueness property ensuring that $\varphi(x,0)\equiv 0$ implies that $g\equiv 0$.

In this way, in [13], it was proved that (P_ε) has a unique solution $\bar g$. Then, the initial datum $\bar u \in L^2(\Omega)$ of minimum norm, with associated state $\bar y$, satisfying that $\|\bar y(T) - y_d\|_{L^2(\Omega)} \le \varepsilon$ is given by $\bar u = \bar\varphi(x,0)$, where $\bar\varphi$ is the solution of (9) corresponding to $\bar g$.

The analysis in [13] covered also the case where the norm of $\varphi_g(x,0)$ in $L^2(\Omega)$ was replaced by the norm in any other space $L^q(\Omega)$ with $1\le q<\infty$, thus leading to optimal initial data in any L^p-setting with $1<p\le\infty$.

In the present paper however, we are interested in optimal data in the sense of measures. The functional above has to be then modified so to replace the $L^2(\Omega)$-norm of $\varphi_g(x,0)$ by the L^∞-one. We shall do this following the arguments in [10].

The minimization problem in the adjoint system to be considered with that purpose is the following:

$$(\mathrm{P}_{\infty,\varepsilon}) \quad \min_{g\in L^2(\Omega)} J_{\infty,\varepsilon}(g) = \frac{1}{2}\|\varphi_g(x,0)\|^2_{C_0(\Omega)} + \varepsilon\|g\|_{L^2(\Omega)} - \int_\Omega y_d g\,dx.$$

Let us analyze this control problem. First of all, it is obvious that $J_{\infty,\varepsilon}$ is strictly convex and continuous. Moreover, it is coercive. Indeed, let $\{g_k\}_{k=1}^\infty \subset L^2(\Omega)$ such that $\|g_k\|_{L^2(\Omega)}\to+\infty$. We claim that

$$\liminf_{k\to\infty}\frac{J_{\infty,\varepsilon}(g_k)}{\|g_k\|_{L^2(\Omega)}} \ge \varepsilon.$$

To this end, we set $\tilde g_k = g_k/\|g_k\|_{L^2(\Omega)}$ and, by taking a subsequence, we can assume that $\tilde g_k \rightharpoonup g$ weakly in $L^2(\Omega)$. Denote $\varphi_k = \varphi_{g_k}$ and $\tilde\varphi_k = \varphi_{\tilde g_k}$. Then,

$$\frac{J_{\infty,\varepsilon}(g_k)}{\|g_k\|_{L^2(\Omega)}} = \frac{1}{2}\|g_k\|_{L^2(\Omega)}\|\tilde\varphi_k(x,0)\|^2_{C_0(\Omega)} + \varepsilon - \int_\Omega y_d\tilde g_k\,dx.$$

The following two cases may occur:

1.- $\liminf_{k\to\infty} \|\tilde\varphi_k(x,0)\|_{C_0(\Omega)} > 0$. In this case we obtain immediately that

$$\frac{J_{\infty,\varepsilon}(g_k)}{\|g_k\|_{L^2(\Omega)}} \to \infty.$$

2.- $\liminf_{k\to\infty} \|\tilde\varphi_k(x,0)\|_{C_0(\Omega)} = 0$. Now, using the weak* convergence $\tilde\varphi_k(x,0) \rightharpoonup \varphi_g(x,0)$ in $L^\infty(\Omega)$, we get by the lower semi-continuity

$$\|\varphi_g(x,0)\|_{L^\infty(\Omega)} \le \liminf_{k\to\infty} \|\tilde\varphi_k(x,0)\|_{L^\infty(\Omega)} = 0.$$

Hence, $\varphi_g(x,0) = 0$ in Ω. Now, the classical backward uniqueness property of the heat equation implies that $\varphi_g \equiv 0$ in $\Omega\times(0,T)$ and consequently $g = 0$. Therefore, $\tilde g_k \rightharpoonup 0$ weakly in $L^2(\Omega)$ and $\int_\Omega y_d\tilde g_k\,dx \to 0$ as well. Finally, we have

$$\liminf_{k\to\infty} \frac{J_{\infty,\varepsilon}(g_k)}{\|g_k\|_{L^2(\Omega)}} \ge \liminf_{k\to\infty}[\varepsilon - \int_\Omega y_d\tilde g_k\,dx] = \varepsilon,$$

which concludes the proof.

Let us denote by $\bar g_\infty$ the solution of $(\mathrm{P}_{\infty,\varepsilon})$ and by $\bar\varphi_\infty = \varphi_{\bar g_\infty}$ the associated state. Since we have assumed that $\|y_d\|_{L^2(\Omega)} > \varepsilon$, setting $g_\lambda = \lambda y_d$, it is easy to check that $J_{\infty,\varepsilon}(g_\lambda) < 0$ if $\lambda > 0$ is small enough. Consequently, we have that $\bar g_\infty \neq 0$. To write the optimality conditions satisfied by $\bar g_\infty$, we will use the linear operator $A^* \in \mathcal{L}(L^2(\Omega), C_0(\Omega))$, pre-adjoint of the linear continuous operator $Au = y(T)$ from the space of measures $M(\Omega)$ into $L^2(\Omega)$, and given by $A^*g = \varphi_g(0)$. Then, $J_{\infty,\varepsilon}$ can be expressed in the form

$$J_{\infty,\varepsilon}(g) = \frac{1}{2}\|A^*g\|^2_{C_0(\Omega)} + \varepsilon\|g\|_{L^2(\Omega)} - \int_\Omega y_d g\,dx.$$

Thus, it holds $0 \in \partial J_{\infty,\varepsilon}(\bar g_\infty)$, which implies the existence of an element $\bar u_\infty \in \partial\|\cdot\|_{C_0(\Omega)}(\bar\varphi_\infty(0))$, where $\bar\varphi_\infty = \varphi_{\bar g_\infty}$, such that

$$0 = \|\bar\varphi_\infty(0)\|_{C_0(\Omega)} A\bar u_\infty + \frac{\varepsilon}{\|\bar g_\infty\|_{L^2(\Omega)}}\bar g_\infty - y_d.$$

If $\bar y_\infty$ denotes the solution of (1) corresponding to $\bar u_\infty$, then $A\bar u_\infty = \bar y_\infty(T)$. Then the above equality can be rewritten

$$\|\bar\varphi_\infty(0)\|_{C_0(\Omega)}\bar y_\infty(T) - y_d = -\frac{\varepsilon}{\|\bar g_\infty\|_{L^2(\Omega)}}\bar g_\infty. \tag{17}$$

On the other hand, $\bar u_\infty \in \partial\|\cdot\|_{C_0(\Omega)}(\bar\varphi_\infty(0))$ implies by definition

$$\int_\Omega (z - \bar\varphi_\infty(0))\,d\bar u_\infty + \|\bar\varphi_\infty(0)\|_{C_0(\Omega)} \le \|z\|_{C_0(\Omega)} \quad \forall z \in C_0(\Omega).$$

These inequalities are equivalent to

$$\|\bar\varphi_\infty(0)\|_{C_0(\Omega)} = \int_\Omega \bar\varphi_\infty(0)\,d\bar u_\infty \quad \text{and} \quad \|\bar u_\infty\|_{M(\Omega)} = 1. \tag{18}$$

Finally, if we set

$$\begin{pmatrix}\bar g\\ \bar\varphi\end{pmatrix} = -\frac{1}{\|\bar\varphi_\infty(0)\|_{C_0(\Omega)}}\begin{pmatrix}\bar g_\infty\\ \bar\varphi_\infty\end{pmatrix} \text{ and } \begin{pmatrix}\bar u\\ \bar y\end{pmatrix} = \|\bar\varphi_\infty(0)\|_{C_0(\Omega)}\begin{pmatrix}\bar u_\infty\\ \bar y_\infty\end{pmatrix},$$

then we get from (17) and (18) that $(\bar g, \bar\varphi, \bar u, \bar y)$ satisfies (9)-(12), and hence $\bar u$ is the solution of (P).

4. **Numerical approximation of problem** (P).

4.1. **Motivation.** In this section, for the sake of simplicity, we assume $\Omega \subset \mathbb{R}^n$ to be convex and $n \leq 3$. The results of this section will probably hold without these restrictions. But such an extension would require further technical developments.

We consider a family of triangulations $\{\mathcal{K}_h\}_{h>0}$ of $\bar{\Omega}$, defined in the standard way. To each element $K \in \mathcal{K}_h$ we associate two parameters h_K and ϱ_K, where h_K denotes the diameter of the set K and ϱ_K is the diameter of the largest ball contained in K. The size of the mesh is defined by $h = \max_{K\in\mathcal{K}_h} h_K$. We also assume the standard regularity assumptions on the triangulation:

(i) – There exist two positive constants $\varrho_{\mathcal{K}}$ and $\delta_{\mathcal{K}}$ such that $\frac{h_K}{\varrho_K} \leq \varrho_{\mathcal{K}}$ and $\frac{h}{h_K} \leq \delta_{\mathcal{K}}$ $\forall K \in \mathcal{K}_h$ and $\forall h > 0$.
(ii) – Define $\overline{\Omega}_h = \cup_{K\in\mathcal{K}_h} K$, and let Ω_h and Γ_h denote its interior and its boundary, respectively. We assume that the vertices of $\mathcal{K}_h$ placed on the boundary Γ_h are points of Γ.

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 $\tau = \max_{1\leq k\leq N_\tau} \tau_k$. We assume that there exists a constant κ such that

$$\kappa^{-1} \leq \frac{\tau_k}{\tau_{k+1}} \leq \kappa,$$

cf. the corresponding assumption in [25].

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

To each triangulation K_h we associate the usual space of linear finite elements

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

where $\{x_j\}_{j=1}^{N_h}$ are the interior nodes of $\mathcal{K}_h$ and $\{e_j\}_{j=1}^{N_h}$ is the nodal basis formed by the continuous piecewise linear functions such that $e_j(x_i) = \delta_{ij}$ for every $1 \leq i, j \leq N_h$.

Following [4] we define the space of discrete controls by

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

where δ_{x_j} denotes the Dirac measure centered at the point x_j. For every σ we introduce the discrete state space

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

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

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

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

$$y_\sigma = \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} y_{kj} \chi_k e_j.$$

Thus U_h and $\mathcal{Y}_\sigma$ are finite dimensional spaces of dimension N_h and $N_\tau \times N_h$, respectively, and bases are given by $\{\delta_{x_j}\}_{j=1}^{N_h}$ and $\{\chi_k e_j\}_{k,j}$.

Next we define the linear operators $\Lambda_h : M(\Omega) \to U_h \subset M(\Omega)$ and $\Pi_h : C_0(\Omega) \to Y_h \subset C_0(\Omega)$ by

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

The operator Π_h is the nodal interpolation operator for Y_h. Concerning the operator Λ_h we have the following result.

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

(i) For every $u \in M(\Omega)$ and every $y \in C_0(\Omega)$ and $y_h \in Y_h$ we have

$$\langle u, y_h \rangle = \langle \Lambda_h u, y_h \rangle, \tag{19}$$

$$\langle u, \Pi_h y \rangle = \langle \Lambda_h u, y \rangle. \tag{20}$$

(ii) For every $u \in M(\Omega)$ we have

$$\|\Lambda_h u\|_{M(\Omega)} \leq \|u\|_{M(\Omega)}, \tag{21}$$

$$\Lambda_h u \overset{*}{\rightharpoonup} u \text{ in } M(\Omega) \text{ and } \|\Lambda_h u\|_{M(\Omega)} \to \|u\|_{M(\Omega)} \text{ as } h \to 0. \tag{22}$$

(iii) There exists a constant $C > 0$ such that for every $u \in M(\Omega)$ we have

$$\|u - \Lambda_h u\|_{W^{-1,p}(\Omega)} \leq C h^{1-n/p'} \|u\|_{M(\Omega)}, \quad 1 < p < \frac{n}{n-1}, \tag{23}$$

$$\|u - \Lambda_h u\|_{(W_0^{1,\infty}(\Omega))^*} \leq C h \|u\|_{M(\Omega)}, \tag{24}$$

with $1/p' + 1/p = 1$.

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 $y_{k,h} = y_\sigma(t_k) = y_\sigma|_{I_k}$, $1 \leq k \leq N_\tau$. 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 $u \in M(\Omega)$, for $k = 1, \ldots, N_\tau$ we set

$$\begin{cases} \left(\dfrac{y_{k,h} - y_{k-1,h}}{\tau_k}, z_h \right) + a(y_{k,h}, z_h) = 0 & \forall z_h \in Y_h \\ y_{0,h} = y_{0h}, \end{cases} \tag{25}$$

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

$$\int_\Omega y_{0h} z_h \, dx = \int_\Omega z_h \, du \quad \forall z_h \in Y_h. \tag{26}$$

The existence and uniqueness of the solution of (25) is obvious. Let us reformulate (25) in an equivalent form in terms of $y_\sigma \in \mathcal{Y}_\sigma$ with $y_\sigma|_{I_k} = y_{k,h}$. We define the bilinear form $B_\sigma : \mathcal{Y}_\sigma \times \mathcal{Y}_\sigma \longrightarrow \mathbb{R}$ by

$$B_\sigma(y_\sigma, z_\sigma) = \sum_{k=2}^{N_\tau} (y_{k,h} - y_{k-1,h}, z_{k,h}) + \int_0^T a(y_\sigma(t), z_\sigma(t)) \, dt + (y_{1,h}, z_{1,h}). \tag{27}$$

Then, it is immediate to check that y_σ is the solution of (25) if and only if

$$B_\sigma(y_\sigma, z_\sigma) = \int_\Omega z_{1,h} \, du \quad \forall z_\sigma \in \mathcal{Y}_\sigma. \tag{28}$$

The following theorem establishes the convergence of these approximations.

Theorem 4.2. *Let us assume that Γ is of class $C^{1,1}$ and let $\{u_\sigma\}_\sigma \subset M(\Omega)$ be a sequence such that $u_\sigma \stackrel{*}{\rightharpoonup} u$ in $M(\Omega)$. If y is the state associated to u, solution of (1), and $\{y_\sigma\}_\sigma$ denote the discrete states associated to $\{u_\sigma\}_\sigma$, then $y_\sigma \rightharpoonup y$ in $L^r(Q)$ for every $r \in [1, \frac{4}{3})$ and $y_\sigma(T) \to y(T)$ in $L^\infty(\Omega)$.*

First, we will prove the boundedness of $\{(y_\sigma, y_\sigma(T))\}_\sigma$ in $L^r(Q) \times L^2(\Omega)$. To do it we need some technical lemmas. We recall that given $y_h \in Y_h$, its discrete Laplacian $\Delta_h y_h \in Y_h$ is defined through the identity

$$(\Delta_h y_h, z_h) = -(\nabla y_h, \nabla z_h) \quad \forall z_h \in Y_h; \tag{29}$$

see, for instance, [31, Chapter 2]. The statement of the next lemma is similar to the discrete Gagliardo–Nirenberg inequality, see [18, Lemma 3.3].

Lemma 4.3. *For every $\delta \in (0, \frac{1}{2}]$ there exists $C_\delta > 0$ such that*

$$\|y_h\|_{L^\infty(\Omega)} \le C_\delta \|\nabla y_h\|_{L^2(\Omega)}^{\frac{1}{2}-\delta} \|\Delta_h y_h\|_{L^2(\Omega)}^{\frac{1}{2}+\delta} \quad \forall y_h \in Y_h. \tag{30}$$

Proof. Let us fix $0 < \delta \le 1/2$ and set $p_\delta = \frac{3}{1-\delta}$. Therefore, $3 < p_\delta \le 6$ holds. Then, from the Sobolev embedding $W^{1,p_\delta}(\Omega) \subset C(\bar{\Omega})$ (recall that $n \le 3$) and the interpolation inequality in L^p spaces we have

$$\|y_h\|_{L^\infty(\Omega)} \le c_\delta \|\nabla y_h\|_{L^{p_\delta}(\Omega)} \le c'_\delta \|\nabla y_h\|_{L^2(\Omega)}^{\frac{1}{2}-\delta} \|\nabla y_h\|_{L^6(\Omega)}^{\frac{1}{2}+\delta}. \tag{31}$$

Now, we consider the Dirichlet problem

$$\begin{cases} -\Delta y^h &= -\Delta_h y_h & \text{in } \Omega, \\ y^h &= 0 & \text{on } \Gamma. \end{cases}$$

Since Γ is of class $C^{1,1}$ we know that $y^h \in H^2(\Omega) \cap H_0^1(\Omega)$; see [17, Chapter 2]. Moreover, the following identity holds

$$(\nabla(y^h - y_h), \nabla z_h) = 0 \quad \forall z_h \in Y_h.$$

Hence, y_h is the Ritz projection of y^h. Using again the regularity of Γ, so to be in the assumptions of [2, Theorem 8.5.3], we infer

$$\|\nabla y_h\|_{L^6(\Omega)} \le C\|\nabla y^h\|_{L^6(\Omega)} \le C\|y^h\|_{H^2(\Omega)} \le C'\|\Delta_h y_h\|_{L^2(\Omega)}.$$

Combining this inequality with (31) we deduce (30). □

The next lemma provides a discrete analog of the compact embedding of $H^2(\Omega)$ into $L^\infty(\Omega)$.

Lemma 4.4. *Let $\{y_h\}_h$, $y_h \in Y_h$, be a sequence for $h \to 0$ with*

$$\|\Delta_h y_h\|_{L^2(\Omega)} \le C.$$

Then, there exists a subsequence (denoted again by $\{y_h\}$) with $y_h \to y$ strongly in $L^\infty(\Omega)$.

Proof. As in the proof of the previous lemma we consider for each $y_h \in Y_h$ the corresponding element $y^h \in H^2(\Omega) \cap H_0^1(\Omega)$ as the solution of

$$\begin{cases} -\Delta y^h &= -\Delta_h y_h & \text{in } \Omega, \\ y^h &= 0 & \text{on } \Gamma. \end{cases}$$

As noted above y_h is the Ritz projection of y^h. There holds

$$\|y^h\|_{H^2(\Omega)} \le c\|\Delta y^h\|_{L^2(\Omega)} = c\|\Delta_h y_h\|_{L^2(\Omega)} \le C.$$

Due to the compact embedding of $H^2(\Omega)$ into $L^\infty(\Omega)$, there is a subsequence (denoted in the same way) with $y^h \to y$ strongly in $L^\infty(\Omega)$. For this subsequence we obtain

$$\begin{aligned}\|y - y_h\|_{L^\infty(\Omega)} &\le \|y - y^h\|_{L^\infty(\Omega)} + \|y^h - y_h\|_{L^\infty(\Omega)}\\ &\le \|y - y^h\|_{L^\infty(\Omega)} + ch^{2-\frac{n}{2}}\|y^h\|_{H^2(\Omega)}\\ &\le \|y - y^h\|_{L^\infty(\Omega)} + ch^{2-\frac{n}{2}}\|\Delta_h y_h\|_{L^2(\Omega)}.\end{aligned}$$

For $h \to 0$ within this subsequence we obtain $y_h \to y$ strongly in $L^\infty(\Omega)$. □

In the next lemma we provide a result on discrete smoothing, which is similar to [18], cf. also [29].

Lemma 4.5. *Let $g \in L^2(\Omega)$ and $\delta \in (0, \frac{1}{2}]$ be given. For every σ we take $\varphi_\sigma \in \mathcal{Y}_\sigma$ as the unique element of Y_σ satisfying*

$$B_\sigma(z_\sigma, \varphi_\sigma) = (g, z_{N_\tau,h}) \quad \forall z_\sigma \in Y_\sigma. \tag{32}$$

Then, there exists $C'_\delta > 0$ independent of g such that the following inequalities hold

$$\|\varphi_{k,h}\|_{L^\infty(\Omega)} \le C'_\delta (T - t_{k-1})^{-(\frac{3}{4}+\frac{\delta}{2})}\|g\|_{L^2(\Omega)}, \tag{33}$$

for all $1 \le k \le N_\tau$.

Proof. First, we observe that the inequalities

$$(T - t_{k-1})\|\Delta_h \varphi_{N_\tau+1-k,h}\|_{L^2(\Omega)} + \sqrt{T - t_{k-1}}\|\nabla \varphi_{N_\tau+1-k,h}\|_{L^2(\Omega)} \le C\|g\|_{L^2(\Omega)}, \tag{34}$$

for all $1 \le k \le N_\tau$, can be established by repeating the arguments of [25, Theorem 4.5-Inequality (4.14)]. Now, it is enough to combine these inequalities with (30) to get (33). □

Lemma 4.6. *Let $y_\sigma \in \mathcal{Y}_\sigma$ be the solution of (25) (or equivalently (28)), and let δ and C'_δ be as in Lemma 4.5. Then, the following inequalities hold*

$$\|y_{k,h}\|_{L^2(\Omega)} \le C'_\delta t_k^{-(\frac{3}{4}+\frac{\delta}{2})}\|u_\sigma\|_{M(\Omega)} \quad \text{for } 1 \le k \le N_\tau \tag{35}$$

and

$$\|\Delta_h y_\sigma(T)\|_{L^2(\Omega)} \le C_\delta T^{-(\frac{7}{4}+\frac{\delta}{2})}\|u_\sigma\|_{M(\Omega)}. \tag{36}$$

Proof. Obviously it is enough to prove (35) for $k = N_\tau$, hence $t_k = T$. Indeed, for any other k we can replace the interval $[0, T]$ for $[0, t_k]$ in the above estimates and proceed in the same way as we do for $k = N_\tau$. Let $\varphi_\sigma \in \mathcal{Y}_\sigma$ be as in Lemma 4.5 with $g = y_{N_\tau,h}$. Taking $z_\sigma = y_\sigma$ in (32), we get with (28) and (33)

$$\begin{aligned}\|y_{N_\tau,h}\|^2_{L^2(\Omega)} &= B_\sigma(y_\sigma, \varphi_\sigma) = \int_\Omega \varphi_{1,h}\, du_\sigma\\ \le&\|u_\sigma\|_{M(\Omega)}\|\varphi_{1,h}\|_{L^\infty(\Omega)} \le C'_\delta T^{-(\frac{3}{4}+\frac{\delta}{2})}\|u_\sigma\|_{M(\Omega)}\|y_{N_\tau,h}\|_{L^2(\Omega)},\end{aligned}$$

which implies (35). To prove (36) we combine (35) for $\tilde{k} = [N_\tau/2]$ and again the estimate

$$\frac{T}{2}\|\Delta_h y_\sigma(T)\|_{L^2(\Omega)} \le C\|y_{\tilde{k},h}\|_{L^2(\Omega)}$$

from [25, Theorem 4.5-Inequality (4.14)]. □

Lemma 4.7. *The sequence $\{y_\sigma\}_\sigma$ of solutions of problems (25) is bounded in $L^r(Q)$ for every $r \in [1, \frac{4}{3})$.*

Proof. Let $\delta \in (0, \frac{1}{2})$ arbitrary and take $r \in [1, \frac{4}{3+2\delta})$. Using (35) we infer

$$\begin{aligned}
\|y_\sigma\|^r_{L^r(Q)} &= \int_0^T \|y_\sigma(t)\|^r_{L^r(\Omega)}\,dt = \sum_{k=1}^{N_\tau} \tau_k \|y_{k,h}\|^r_{L^r(\Omega)} \\
&\le \left(\sum_{k=1}^{N_\tau} t_k^{-r(\frac{3}{4}+\frac{\delta}{2})}\tau_k\right)\left(C'_\delta\|u\|_{M(\Omega)}\right)^r \\
&\le \left(\sum_{k=1}^{N_\tau} \int_{t_{k-1}}^{t_k} t^{-r(\frac{3}{4}+\frac{\delta}{2})}\,dt\right)\left(C'_\delta\|u\|_{M(\Omega)}\right)^r \\
&= \left(\int_0^T t^{-r(\frac{3}{4}+\frac{\delta}{2})}\,dt\right)\left(C'_\delta\|u\|_{M(\Omega)}\right)^r.
\end{aligned}$$

Since $r(\frac{3}{4}+\frac{\delta}{2}) < 1$, then the above integral is finite and the boundedness of $\{y_\sigma\}_\sigma$ in $L^r(Q)$ follows. Finally, for every $r < \frac{4}{3}$ we can take $\delta > 0$ sufficiently close to 0 such that $r < \frac{4}{3+2\delta}$. This concludes the proof. ☐

Proof of Theorem 4.2. Taking $\delta = \frac{1}{2}$ in (36) we get

$$\|\Delta_h y_\sigma(T)\|_{L^2(\Omega)} = \|\Delta_h y_{N_\tau,h}\|_{L^2(\Omega)} \le \frac{C'_{1/2}}{T^2}\|u_\sigma\|_{M(\Omega)}. \tag{37}$$

Since $\{u_\sigma\}_\sigma$ is bounded in $M(\Omega)$, we deduce with Lemma 4.7, Lemma 4.6 and Lemma 4.4 the existence of a subsequence, denoted in the same way, such that

$$y_\sigma \rightharpoonup y \text{ in } L^r(Q) \text{ and } y_\sigma(T) \to y_T \in L^\infty(\Omega) \tag{38}$$

for some $y \in L^r(Q)$ and $y_T \in L^\infty(\Omega)$. We will prove that y is the solution of (1) associated to u and $y_T = y(T)$. Therefore, since every convergent subsequence has the same limit, the whole sequence convergence to this limit. Without loss of generality we can assume that $r > 1$. Let $\psi \in C^\infty(\bar\Omega) \cap C_0(\Omega)$ and $\xi \in C^1[0,T]$. Let ψ_h be the Ritz projection of ψ on Y_h, then $\psi_h \to \psi$ in $W_0^{1,r'}(\Omega)$. Then, we have

$$\begin{aligned}
&\int_0^T (y_\sigma(t), \psi_h)\xi'(t)\,dt \\
&= \sum_{k=1}^{N_\tau}\int_{t_{k-1}}^{t_k} (y_{k,h}, \psi_h)\xi'(t)\,dt = \sum_{k=1}^{N_\tau}(y_{k,h}, \psi_h)(\xi(t_k) - \xi(t_{k-1})) \\
&= -\sum_{k=1}^{N_\tau}(y_{k,h} - y_{k-1,h}, \psi_h)\xi(t_{k-1}) + (y_{N_\tau,h}, \psi_h)\xi(T) - (y_{0,h}, \psi_h)\xi(0) \\
&= \sum_{k=1}^{N_\tau}\tau_k a(y_{k,h}, \psi_h)\xi(t_{k-1}) + (y_\sigma(T), \psi_h)\xi(T) - \int_\Omega \psi_h\,du_\sigma\xi(0) \\
&= \sum_{k=1}^{N_\tau}\tau_k a(y_{k,h}, \psi)\xi(t_{k-1}) + (y_\sigma(T), \psi_h)\xi(T) - \int_\Omega \psi_h\,du_\sigma\xi(0)
\end{aligned}$$

$$= -\sum_{k=1}^{N_\tau} \tau_k (y_{k,h}, \Delta\psi)\xi(t_{k-1}) + (y_\sigma(T), \psi_h)\xi(T) - \int_\Omega \psi_h \, du_\sigma \xi(0)$$

$$= -\int_0^T (y_\sigma(t), \Delta\psi)\xi(t)\, dt + (y_\sigma(T), \psi_h)\xi(T) - \int_\Omega \psi_h \, du_\sigma \xi(0)$$

$$+ \sum_{k=1}^{N_\tau} \int_{t_{k-1}}^{t_k} (y_\sigma(t), \Delta\psi)(\xi(t) - \xi(t_{k-1}))\, dt.$$

Since $r' > 4$, we have that $W_0^{1,r'}(\Omega) \subset C_0(\Omega)$ and hence $\psi_h \to \psi$ in $C_0(\Omega)$. Moreover, $|\xi(t) - \xi(t_{k-1})| \leq \tau \|\xi'\|_{C[0,1]}$ for $1 \leq k \leq N_\tau$, therefore we can pass to the limit in the above identity and obtain

$$(y_T, \psi)\xi(T) - \int_Q y \left(\frac{\partial(\xi\psi)}{\partial t} + \Delta(\xi\psi) \right) dx\, dt = \int_\Omega \psi \, du\, \xi(0). \tag{39}$$

Since the linear combination of functions $\xi\psi$ are dense in $L^{r'}(0,T; W^{2,r'}(\Omega) \cap W_0^{1,r'}(\Omega)) \cap W^{1,r'}(Q)$, Φ is contained in this space, and $y \in L^r(Q)$, we deduce that (4) holds. Hence, y is the solution of (1). Moreover, taking $\xi \in C^1[0,T]$ with $\xi(T) = 1$ and $\xi(0) = 0$, we get from (6) and (39) that $(y(T), \psi) = (y_T, \psi)$ $\forall \psi \in W^{2,r'}(\Omega) \cap W_0^{1,r'}(\Omega)$, which implies that $y(T) = y_T$. □

4.3. **Discrete control problem.** Now, we define the discrete control problem as follows

$$(\mathrm{P}_\sigma) \begin{cases} \min J(u) = \|u\|_{M(\Omega)}, \\ (u, y_{\sigma,u}(T)) \in M(\Omega) \times \bar{B}_\varepsilon(y_d), \end{cases}$$

where $y_{\sigma,u} \in \mathcal{Y}_\sigma$ is the discrete state associated to u, i.e. the solution of (25) or equivalently (28)

Let us study the existence of solutions for problems (P_σ).

Theorem 4.8. *There exists $\sigma_0 > 0$ such that for every $\sigma = (\tau, h)$, with $|\sigma| = \tau + h \leq \sigma_0$, the discrete problem (P_σ) has at least one solution. Among them, there exists a unique solution $\bar{u}_\sigma \in U_h$. Moreover, any other solution $\tilde{u}_\sigma \in M(\Omega)$ of (P_σ) satisfies $\Lambda_h \tilde{u}_\sigma = \bar{u}_\sigma$. Finally, the identity $\|\bar{y}_\sigma(T) - y_d\|_{L^2(\Omega)} = \varepsilon$ holds.*

Proof. First, we have to prove that the set of controls $u \in M(\Omega)$ for which $y_{\sigma,u}(T) \in \bar{B}_\varepsilon(y_d)$ is non empty. From the approximate controllability property of the heat equation, we know that there exists a regular elements u_0, for instance $u_0 \in H_0^2(\Omega) \subset M(\Omega)$ such that its associated state y_{u_0} belongs to the open ball $B_\varepsilon(y_d)$. Due to the regularity of u_0 we have that $\|y_{u_0}(T) - y_{\sigma,u_0}(T)\|_{L^2(\Omega)} \to 0$; see [31, Chapter 9]. Hence, there exists $\sigma_0 > 0$ such that $y_{\sigma,u_0} \in B_\varepsilon(y_d)$ for all $|\sigma| \leq \sigma_0$. Now, the proof of the existence of a solution $\tilde{u}_\sigma \in M(\Omega)$ is like in Theorem 2.4.

Given a solution $\tilde{u}_\sigma \in M(\Omega)$ of (P_σ), let us define $\bar{u}_\sigma = \Lambda_h \tilde{u}_\sigma \in U_h$. Then, (19) implies that the initial states $y_{0,h}$ corresponding to $\tilde{u}_\sigma$ and $\bar{u}_\sigma$, defined by (26), are equal. Hence, the identity $y_{\sigma,\bar{u}_\sigma} = y_{\sigma,\tilde{u}_\sigma}$ holds. In particular we have that $y_{\sigma,\bar{u}_\sigma}(T) = y_{\sigma,\tilde{u}_\sigma}(T) \in B_\varepsilon(y_d)$, which proves that $\bar{u}_\sigma$ is a feasible control for (P_σ). Additionally, (21) shows that $J(\bar{u}_\sigma) \leq J(\tilde{u}_\sigma)$, consequently $\bar{u}_\sigma$ is also a solution of (P_σ).

For any solution $\bar{u}_\sigma$ of (P_σ) with the corresponding state $\bar{y}_\sigma$ the identity

$$\|\bar{y}_\sigma(T) - y_d\|_{L^2(\Omega)} = \varepsilon$$

is proved like in Theorem 2.2.

Let us prove the uniqueness of solutions in U_h. Let $\bar u_\sigma, \bar v_\sigma \in U_h$ be two solutions of (P_σ) with associated discrete states $\bar y_\sigma$ and $\bar z_\sigma$, respectively. Arguing as in the proof of Theorem 2.4, we obtain that $\bar y_\sigma(T) = \bar z_\sigma(T)$. Now, subtracting the first equations in (25) satisfied by $\bar y_\sigma$ and $\bar z_\sigma$, and computing $\bar y_{k-1,h} - \bar z_{k-1,h}$ from $\bar y_{k,h} - \bar z_{k,h}$, starting at $\bar y_{N_\tau,h} - \bar z_{N_\tau,h} = 0$, we conclude that $\bar y_\sigma - \bar z_\sigma = 0$. By (26) this results in

$$\int_\Omega \psi_h d(\bar u_\sigma - \bar v_\sigma) = 0 \quad \text{for all} \quad \psi_h \in Y_h.$$

Hence, $\bar u_\sigma - \bar v_\sigma = 0$ by the structure of U_h. □

The next theorem characterizes the solutions of (P_σ).

Theorem 4.9. *Let σ_0 be as in Theorem 4.8 and assume that $|\sigma| \le \sigma_0$. Let $\bar u_\sigma \in M(\Omega)$ such that $\bar y_\sigma(T) \in \bar B_\varepsilon(y_d)$, where $\bar y_\sigma$ is the state discrete associated to $\bar u_\sigma$. Then, $\bar u_\sigma$ is the solution of problem (P_σ) if and only if there exist two elements $\bar g_\sigma \in Y_h$ and $\bar\varphi_\sigma \in \mathcal{Y}_\sigma$ such that*

$$\int_\Omega \bar g_\sigma(x)(y_h(x) - \bar y_\sigma(x,T))\,dx \le 0 \quad \forall y_h \in \bar B_\varepsilon(y_d) \cap Y_h, \tag{40}$$

$$\begin{cases} \text{For } k = N_\tau, \dots, 1 \\ \left(\dfrac{\bar\varphi_{k,h} - \bar\varphi_{k+1,h}}{\tau_k}, z_h\right) + a(z_h, \bar\varphi_{k,h}) = 0 \qquad \forall z_h \in Y_h, \\ \bar\varphi_{N_\tau+1,h} = \bar g_\sigma, \end{cases} \tag{41}$$

$$\|\bar u_\sigma\|_{M(\Omega)} = -\int_\Omega \bar\varphi_{1,h}\,d\bar u_\sigma, \tag{42}$$

$$\|\bar\varphi_{1,h}\|_{C_0(\Omega)} = 1. \tag{43}$$

Furthermore, $\bar\varphi_\sigma$ and $\bar g_\sigma$ are unique, and there exists a real number $\bar\lambda_\sigma > 0$ such that $\bar g_\sigma = \bar\lambda_\sigma(\bar y_\sigma(T) - y_{d,\sigma})$, where $y_{d,\sigma}$ is the $L^2(\Omega)$ projection of y_d on Y_h.

Proof. We proceed as in the proof of Theorem 2.5. To this end, we define the operator $A_\sigma \in \mathcal{L}(M(\Omega), L^2(\Omega))$ by $A_\sigma u = y_{\sigma,u}(T)$ and $\mathcal{J}_\sigma(u) = J(u) + I_{\bar B_\varepsilon(y_d)}(A_\sigma u)$. In the proof of Theorem 4.8 it was established the existence of an element $u_0 \in M(\Omega)$ such that $y_{\sigma,u_0}(T) \in B_\varepsilon(y_d)$. Hence, we can use the chain rule and to deduce that $\bar u_\sigma$ is a solution of (P_σ) if and only if there exits $\tilde g_\sigma \in \partial I_{\bar B_\varepsilon(y_d)}(\bar y_\sigma(T))$ such that $0 \in \partial J(\bar u_\sigma) + A_\sigma^* \tilde g_\sigma$ with A_σ^* being the pre-adjoint operator of A_σ. Now, we take $\bar g_\sigma$ as the $L^2(\Omega)$ projection of $\tilde g_\sigma$ on Y_h and define $\bar\varphi_\sigma$ as the solution of (41). Then, for every $u \in M(\Omega)$ we have

$$\begin{aligned} \langle A_\sigma^* \tilde g_\sigma, u\rangle = \langle \tilde g_\sigma, A_\sigma u\rangle &= \int_\Omega \tilde g_\sigma y_{\sigma,u}(T)\,dx = \int_\Omega \bar g_\sigma y_{\sigma,u}(T)\,dx \\ &= \int_\Omega \bar\varphi_\sigma(T) y_{\sigma,u}(T) = \int_\Omega y_{0,h}\bar\varphi_{1,h}\,dx = \int_\Omega \bar\varphi_{1,h}\,du, \end{aligned}$$

where we have used the fact that (41) is the adjoint state equation of (25), and the definition of $y_{0,h}$ given by (26). Using that $-A_\sigma^*\tilde g \in \partial J(\bar u_\sigma)$ and the above identity we infer

$$\int_\Omega \bar\varphi_{1,h}\,du - \int_\Omega \bar\varphi_{1,h}\,d\bar u_\sigma + \|\bar u_\sigma\|_{M(\Omega)} \le \|u\|_{M(\Omega)} \quad \forall u \in M(\Omega).$$

As in the proof of Theorem 2.5, (42) and (43) follow from these inequalities. Using that $\bar{g}_\sigma$ is the projection of $\tilde{g}_\sigma$ and $\tilde{g}_\sigma \in \partial I_{\bar{B}_\varepsilon(y_d)}(\bar{y}_\sigma(T))$, we obtain

$$\int_\Omega \bar{g}_\sigma(y_h - \bar{y}_\sigma(T))\,dx = \int_\Omega \tilde{g}_\sigma(y_h - \bar{y}_\sigma(T))\,dx \le 0 \quad \forall y_h \in \bar{B}_\varepsilon(y_d) \cap Y_h,$$

which proves (40). Furthermore, from the last inequality it follows the existence of $\bar{\lambda}_\sigma > 0$ such that $\tilde{g}_\sigma = \bar{\lambda}_\sigma(\bar{y}_\sigma(T) - y_d)$, therefore $\bar{g}_\sigma = \bar{\lambda}_\sigma(\bar{y}_\sigma(T) - y_{d,h})$. The uniqueness of $\bar{\lambda}_\sigma$ follows from (43), which implies the uniqueness of $\bar{\varphi}_\sigma$. □

We have the following result analogous to Corollary 1.

Corollary 2. *Let $\bar{u}_\sigma \in U_h$ be the solution of (P_σ) for $|\sigma| \le \sigma_0$. Then, we have*

$$\bar{u}_\sigma = \sum_{j=1}^{N_h} \bar{\lambda}_j \delta_{x_j}, \quad \text{where } \bar{\lambda}_j \begin{cases} = 0 & \text{if } |\bar{\varphi}_{1,h}| < 1, \\ \le 0 & \text{if } \bar{\varphi}_{1,h} = +1, \\ \ge 0 & \text{if } \bar{\varphi}_{1,h} = -1. \end{cases} \tag{44}$$

This corollary is an immediate consequence of (42), (43), and the definition of U_h. We finish this section by proving the convergence of (P_σ).

Theorem 4.10. *Let $\{\bar{u}_\sigma\}_{|\sigma|\le\sigma_0}$ be a sequence formed by solutions of problems (P_σ), and let $\{\bar{y}_\sigma\}_{|\sigma|\le\sigma_0}$ be the associate discrete states. Then, the following convergence properties are fulfilled*

$$\bar{u}_\sigma \stackrel{*}{\rightharpoonup} \bar{u} \quad \text{in } M(\Omega) \quad \text{and} \quad \lim_{\sigma\to 0} \|\bar{u}_\sigma\|_{M(\Omega)} = \|\bar{u}\|_{M(\Omega)}, \tag{45}$$

$$\bar{y}_\sigma \rightharpoonup \bar{y} \text{ in } L^r(Q)\ \forall r \in [1, \frac{4}{3}) \text{ and } \bar{y}_\sigma(T) \to \bar{y}(T) \text{ in } L^\infty(\Omega), \tag{46}$$

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

Proof. Let $u_0 \in M(\Omega)$ be the control introduced in the proof of Theorem 4.8. Then, we know that u_0 is an admissible control for the problem (P_σ) for every $|\sigma| \le \sigma_0$. Hence, $J(\bar{u}_\sigma) \le J(u_0)$ for all $|\sigma| \le \sigma_0$, and therefore $\{\bar{u}_\sigma\}_{|\sigma|\le\sigma_0}$ is bounded in $M(\Omega)$. Let us take a subsequence, denoted in the same way, such that $\bar{u}_\sigma \stackrel{*}{\rightharpoonup} u$ in $M(\Omega)$, and let y be the associated continuous state. We will prove that u is the solution $\bar{u}$ of (P) and $y = \bar{y}$. From Theorem 4.2 we have that

$$\bar{y}_\sigma \rightharpoonup y \text{ in } L^r(Q)\ \forall r \in [1, \frac{4}{3}) \text{ and } \bar{y}_\sigma(T) \to y(T) \text{ in } L^\infty(\Omega).$$

Since $\bar{y}_\sigma(T) \in \bar{B}_\varepsilon(y_d)$ for every $|\sigma| \le \sigma_0$, we deduce that $y(T) \in \bar{B}_\varepsilon(y_d)$ as well. Hence, u is an admissible control for problem (P). For every $\rho \in (0,1)$ we set $u_\rho = \bar{u} + \rho(u_0 - \bar{u})$. If we denote by y_ρ the state associated to u_ρ and by y_0 the state associated to u_0, then we have

$$\|y_\rho(T) - y_d\|_{L^2(\Omega)} \le (1-\rho)\|\bar{y}(T) - y_d\|_{L^2(\Omega)} + \rho\|y_0(T) - y_d\|_{L^2(\Omega)} < \varepsilon.$$

We fix ρ and consider $y_{\rho,\sigma} \in \mathcal{Y}_\sigma$ being the discrete solution associated to u_ρ. By Theorem 4.2 we have $y_{\rho,\sigma}(T) \to y_\rho(T)$ in $L^\infty(\Omega)$. Hence, $y_{\rho,\sigma}(T) \in B_\varepsilon(y_d)$ and u_ρ is an admissible control for (P_σ) for every σ sufficiently small. Therefore, using that $\bar{u}_\sigma$ is a solution of (P_σ), that $\bar{u}$ is the solution of (P) and u is an admissible control for (P), we infer

$$J(\bar{u}) \le J(u) \le \liminf_{\sigma\to 0} J(\bar{u}_\sigma) \le \limsup_{\sigma\to 0} J(\bar{u}_\sigma) \le J(u_\rho).$$

Finally, passing to the limit as $\rho \to 0$ we conclude

$$J(\bar u) \le J(u) \le \liminf_{\sigma\to 0} J(\bar u_\sigma) \le \limsup_{\sigma\to 0} J(\bar u_\sigma) \le J(\bar u).$$

This proves that $u = \bar u$ and consequently $y = \bar y$. Moreover, the above inequality implies (45) as well. $\square$

Corollary 3. *Under the conditions of Theorem 4.10 there holds*

$$\bar\lambda_\sigma \to \bar\lambda, \quad \bar g_\sigma \to \bar g \text{ in } L^2(\Omega), \quad \text{and} \quad \bar\varphi_{1,h} \to \bar\varphi(0) \text{ in } L^\infty(\Omega).$$

Proof. Due to the fact that $\bar\lambda \neq 0$ and $\bar\lambda_\sigma \neq 0$ we can define

$$\hat\varphi = \frac{1}{\bar\lambda}\bar\varphi \text{ and } \hat\varphi_\sigma = \frac{1}{\bar\lambda_\sigma}\bar\varphi_\sigma.$$

Therefore $\hat\varphi$ fulfills the equation

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

and $\hat\varphi_\sigma$ fulfills the corresponding discrete equation with the initial condition $y_\sigma(T) - y_{d,\sigma}$. Using the convergence

$$\bar y_\sigma(T) - y_{d,\sigma} \to \bar y(T) - y_d \quad \text{in } L^2(\Omega)$$

and applying Theorem 4.2 to the sequence $\hat\varphi_\sigma$ we obtain $\hat\varphi_{1,h} \to \hat\varphi(0)$ in $L^\infty(\Omega)$. From Theorem 2.5 and Theorem 4.9 we get

$$\bar\lambda = -\frac{\|\bar u\|_{M(\Omega)}}{\int_\Omega \hat\varphi(0)\,d\bar u} \quad \text{and} \quad \bar\lambda_\sigma = -\frac{\|\bar u_\sigma\|_{M(\Omega)}}{\int_\Omega \hat\varphi_{1,h}\,d\bar u_\sigma}$$

and therefore by Theorem 4.10 we obtain $\bar\lambda_\sigma \to \bar\lambda$. The statements

$$\bar g_\sigma \to \bar g \text{ in } L^2(\Omega), \quad \bar\varphi_{1,h} \to \bar\varphi(0) \text{ in } L^\infty(\Omega)$$

follow then directly by the definitions of $\bar g$, $\bar g_\sigma$, $\hat\varphi$ and $\hat\varphi_\sigma$. $\square$

5. **Final remarks.** In this section, we address two final issues. First, we see that the control problem (P) is the limit, as $p \to 1$, of the corresponding control problems where the measures are replaced by functions of $L^p(\Omega)$. In the second part, we consider the case where the control u is supported in a given small set $\omega \subset \Omega$ with possibly an empty interior.

5.1. **Problem** (P) **as limit of control problems with functions.** Given $1 < p < +\infty$, we define the control problem

$$(\mathrm{P}_p) \left\{\begin{array}{l} \min J_p(u) = \frac{1}{p}\|u\|_{L^p(\Omega)},\\ (u, y_u(T)) \in L^p(\Omega) \times \bar B_\varepsilon(y_d). \end{array}\right.$$

It is immediate that this problem has a unique solution. Let us check that (P) is the limit of problems (P_p) as $p \to 1$. First, we observe that the case $p = 1$ is not a well posed problem because of the lack of compactness properties of $L^1(\Omega)$. Consequently, there is a mathematical reason to consider the space of measures in the definition of (P). Now, let us denote by u_p the solution of (P_p) and fix $u_\infty \in L^\infty(\Omega)$ such that its associated state y_∞, solution of (1) for $u = u_\infty$, belongs to the ball $\bar{B}_\varepsilon(y_d)$. Therefore, u_∞ is an admissible control for any problem (P$_p$). We consider a sequence u_p of solutions to (P$_p$) for $p \to 1$ and obtain:

$$\|u_p\|_{L^1(\Omega)} \leq |\Omega|^{1-\frac{1}{p}} \|u_p\|_{L^p(\Omega)} = p|\Omega|^{1-\frac{1}{p}} J_p(u_p) \leq p|\Omega|^{1-\frac{1}{p}} J_p(u_\infty) \leq |\Omega| \, \|u_\infty\|_\infty.$$

Then, taking a subsequence, denoted in the same way, we have that $u_p \stackrel{*}{\rightharpoonup} \tilde{u}$ in $M(\Omega)$. Let us prove that $\tilde{u} = \bar{u}$, the solution of (P). First, we observe

$$\begin{aligned} J(\bar{u}) &\leq J(\tilde{u}) \leq \liminf_{p\to 1} \|u_p\|_{L^1(\Omega)} \leq \limsup_{p\to 1} \|u_p\|_{L^1(\Omega)} \\ &\leq \limsup_{p\to 1} p|\Omega|^{1-\frac{1}{p}} J_p(u_p) \leq \limsup_{p\to 1} p|\Omega|^{1-\frac{1}{p}} J_p(u_\infty) = \|u_\infty\|_{L^1(\Omega)}. \end{aligned}$$

Since u_∞ was taken arbitrarily in $L^\infty(\Omega)$, the above inequalities prove that

$$\begin{aligned} J(\bar{u}) &\leq J(\tilde{u}) \leq \liminf_{p\to 1} \|u_p\|_{L^1(\Omega)} \leq \limsup_{p\to 1} \|u_p\|_{L^1(\Omega)} \\ &\leq \inf\Big\{\|u\|_{L^1(\Omega)} : \ u \in L^\infty(\Omega), \ \text{ and } \ y_u(T) \in \bar{B}_\varepsilon(y_d)\Big\}. \end{aligned} \tag{47}$$

Let us prove that the last infimum coincides with $J(\bar{u})$. Of course it is bigger or equal to $J(\bar{u})$, we establish the opposite inequality. To this end, let us fix again $u_\infty \in L^\infty(\Omega)$ such that $\|y_\infty(T) - y_d\|_{L^2(\Omega)} < \varepsilon$. Define $u_\rho = (1-\rho)\bar{u} + \rho u_\infty \in M(\Omega)$ for every $\rho \in (0,1)$. Then, the state y_ρ associated with the control u_ρ satisfies $\|y_\rho(T) - y_d\|_{L^2(\Omega)} < \varepsilon$. Now, we take a sequence $\{u_k\}_{k=1}^\infty \subset C(\bar{\Omega})$ such that $u_k \stackrel{*}{\rightharpoonup} u_\rho$ in $M(\Omega)$ and $\lim_{k\to\infty} \|u_k\|_{L^1(\Omega)} = \|u_\rho\|_{M(\Omega)}$. Since $y_k(T) \to y_\rho(T)$ in $L^2(\Omega)$ as $k \to \infty$ (see Lemma 2.3), we deduce the existence of k_ε such that $\|y_k(T) - y_d\|_{L^2(\Omega)} < \varepsilon$ for every $k \geq k_\varepsilon$. Hence, we get

$$\begin{aligned} &\inf\Big\{\|u\|_{L^1(\Omega)} : \ u \in L^\infty(\Omega), \ \text{ and } \ y_u(T) \in \bar{B}_\varepsilon(y_d)\Big\} \\ &\leq \lim_{k\to\infty} \|u_k\|_{L^1(\Omega)} = \|u_\rho\|_{M(\Omega)} \to \|\bar{u}\|_{M(\Omega)} \text{ as } \rho \to 0. \end{aligned}$$

This along with (47) imply that $J(\bar{u}) = J(\tilde{u})$, hence the uniqueness of a solution of (P) shows that $\bar{u} = \tilde{u}$. Moreover, the above inequalities also prove that $\|u_p\|_{L^1(\Omega)} \to \|\bar{u}\|_{M(\Omega)}$ as $p \to 1$. Finally, from the uniqueness of the limit $\bar{u}$, we deduce that these properties of convergence are valid for the whole sequence $\{u_p\}_{p>1}$, not only for a subsequence.

5.2. **Controls supported in little domains ω.** Let ω be a relatively closed subset of Ω with a Lebesgue measure $|\omega| > 0$. Instead of taking $M(\Omega)$ as space of controls, we can consider the controls $u \in M(\omega)$. Then, the control problem is formulated in the form

$$\text{(P)} \ \begin{cases} \min J(u) = \|u\|_{M(\omega)}, \\ (u, y_u(T)) \in M(\omega) \times \bar{B}_\varepsilon(y_d), \end{cases}$$

If we denote

$$C_0(\omega) = \{z \in C(\bar{\omega}) : z(x) = 0 \ \ \forall x \in \partial\omega \cap \Gamma\},$$

then $M(\omega) = C_0(\omega)^*$. All the previous results remain valid. First we observe, that for any $\epsilon > 0$, (P) has feasible controls. Indeed, first we observe that the space $\{y_u(T) : u \in M(\omega)\}$ is dense in $L^2(\Omega)$. To check this we proceed as usual: if $g \in L^2(\Omega)$ satisfies that $\int_\Omega g y_u(T)\, dx = 0$ for every $u \in M(\omega)$, then we prove that

$g = 0$. To this end, we take $\varphi_g \in L^2(0,T;H_0^1(\Omega)) \cap C([0,T], L^2(\Omega))$ solution of (16). Then, according to (6) we have

$$\int_\omega \varphi_g(0)\,du = \int_\Omega g y_u(T)\,dx = 0 \quad \forall u \in M(\omega),$$

hence $\varphi_g(x,0) = 0\ \forall x \in \omega$. But, $x \in \Omega \to \varphi_g(x,0)$ is analytic and $|\omega| > 0$, therefore $\varphi_g(x,0) = 0\ \forall x \in \Omega$ holds. Finally, the backward uniqueness property of the heat equation implies that $\varphi_g = 0$ in $\Omega \times (0,T)$, and consequently $g = 0$.

Now, the problem $(\mathrm{P}_{\infty,\varepsilon})$ defined in §3 should be modified as follows

$$(\mathrm{P}_{\infty,\varepsilon}) \quad \min_{g \in L^2(\Omega)} J_{\infty,\varepsilon}(g) = \frac{1}{2}\|\varphi_g(x,0)\|^2_{C_0(\omega)} + \varepsilon\|g\|_{L^2(\Omega)} - \int_\Omega y_d g\,dx.$$

The proof of the coercivity of $J_{\infty,\varepsilon}$ follows the same lines as in §3, and we deduce that $\varphi_g(x,0) = 0$ for every $x \in \omega$. Arguing as above, we get again the identity $g = 0$.

REFERENCES

[1] E. Di Benedetto, On the local behaviour of solutions of degenerate parabolic equations with measurable coefficients, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4),* **13** (1986), 487–535.

[2] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, Berlin, Heidelberg, 2008, Third edition.

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

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

[5] ______, Parabolic control problems in measure spaces with sparse solutions, *SIAM J. Control Optim.,* **51** (2013), 28–63.

[6] E. Casas, R. Herzog and G. Wachsmuth, Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional, *SIAM J. Optim.,* **22** (2012), 795–820.

[7] E. Casas and K. Kunisch, Optimal control of semilinear elliptic equations in measure spaces, *SIAM J. Control Optim.,* **52** (2014), 339–364.

[8] ______, Parabolic control problems in space-time measure spaces, To appear in ESAIM Control Optim. Calc. Var.

[9] 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), 1010–1033.

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

[11] 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.

[12] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems*, North-Holland-Elsevier, New York, 1976.

[13] C. Fabre, J. P. Puel and E. Zuazua, On the density of the range of the semigroup for semilinear heat equations, Control and optimal design of distributed parameter systems (Minneapolis, MN, 1992), *IMA Vol. Math. Appl.,* Springer, New York, **70** (1995), 73–91.

[14] E. Fernández-Cara and E. Zuazua, The cost of approximate controllability for heat equations: The linear case, *Advances in Differential Equations,* **5** (2000), 465–514.

[15] W. Gong, M. Hinze and Z. Zhou, A Priori Error Analysis for Finite Element Approximation of Parabolic Optimal Control Problems with Pointwise Control, *SIAM J. Control Optim.,* **52** (2014), 97–119.

[16] J. A. Griepentrog, Maximal regularity for nonsmooth parabolic problems in Sobolev-Morrey spaces, *Adv. Differential Equations,* **12** (2007), 1031–1078.

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

[18] A. Hansbo, Strong stability and non-smooth data error estimates for discretizations of linear parabolic problems, *BIT,* **42** (2002), 351–379.

[19] R. Herzog, G. Stadler and G. Wachsmuth, Directional sparsity in optimal control of partial differential equations, *SIAM J. Control Optim.,* **50** (2012), 943–963.

[20] J. Kovats, Real analytic solutions of parabolic equations with time-measurable coefficients, *Proc. Amer. Math. Soc.,* **130** (2002), 1055–1064.

[21] K. Kunisch, K. Pieper and B. Vexler, Measure valued directional sparsity for parabolic optimal control problems, *SIAM J. Control Optim.,* **52** (2014), 3078–3108.

[22] D. Leykekhman and B. Vexler, Optimal a priori error estimates of parabolic optimal control problems with pointwise control, *SIAM J. Numer. Anal.,* **51** (2013), 2797–2821.

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

[24] J.-L. Lions and B. Malgrange, Sur l'unicité rétrograde des équations paraboliques, *Math. Scand.,* **8** (1960), 277–286.

[25] D. Meidner, R. Rannacher and B. Vexler, A priori error estimates for finite element discretizations of parabolic optimization problems with pointwise state constraints in time, *SIAM J. Control Optim.,* **49** (2011), 1961–1997.

[26] 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), 2788–2808.

[27] J. P. Raymond and H. Zidani, Hamiltonian Pontryagin's principles for control problems governed by semilinear parabolic equations, *App. Math. Optim.,* **39** (1999), 143–177.

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

[29] A. H. Schatz, V. C. Thomée and L. B. Wahlbin, Maximum norm stability and error estimates in parabolic finite element equations, *Comm. Pure Appl. Math.,* **33** (1980), 265–304.

[30] G. Stadler, Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices, *Comput. Optim. Appl.,* **44** (2009), 159–181.

[31] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Second edition, Spinger-Verlag, Berlin, 2006.

[32] G. Wachsmuth and D. Wachsmuth, Convergence and regularisation results for optimal control problems with sparsity functional, *ESAIM Control Optim. Calc. Var.,* **17** (2011), 858–886.

Received August 2014; revised November 2014.

E-mail address: eduardo.casas@unican.es
E-mail address: vexler@ma.tum.de
E-mail address: zuazua@bcamath.org