

## PARABOLIC CONTROL PROBLEMS IN MEASURE SPACES WITH SPARSE SOLUTIONS\*

EDUARDO CASAS<sup>†</sup>, CHRISTIAN CLASON<sup>‡</sup>, AND KARL KUNISCH<sup>‡</sup>

**Abstract.** Optimal control problems in measure spaces lead to controls that have small support, which is desirable, e.g., in the context of optimal actuator placement. For problems governed by parabolic partial differential equations, well-posedness is guaranteed in the space of square-integrable measure-valued functions, which leads to controls with a spatial sparsity structure. A conforming approximation framework allows one to derive numerically accessible optimality conditions as well as convergence rates. In particular, although the state is discretized, the control problem can still be formulated and solved in the measure space. Numerical examples illustrate the structural features of the optimal controls.

**Key words.** measure controls, optimal control, sparsity, parabolic partial differential equations, convergence estimates

**AMS subject classifications.** 90C48, 49K20, 65N30, 35K10

**DOI.** 10.1137/120872395

**1. Introduction.** This paper is concerned with the analysis and approximation of the optimal control problem

$$(P) \quad \min_{u \in L^2(I, \mathcal{M}(\Omega))} J(u) = \frac{1}{2} \|y - y_d\|_{L^2(\Omega_T)}^2 + \alpha \|u\|_{L^2(\mathcal{M})},$$

where  $I = [0, T]$  and  $y$  is the unique solution to the initial-boundary value problem

$$(1.1) \quad \begin{cases} \partial_t y - \Delta y = u & \text{in } \Omega_T = \Omega \times (0, T), \\ y = 0 & \text{on } \Sigma_T = \Gamma \times (0, T), \\ y(x, 0) = y_0 & \text{in } \Omega \end{cases}$$

for given  $y_0 \in L^2(\Omega)$ . We assume that  $\alpha > 0$ ,  $y_d \in L^2(\Omega_T)$ , and  $\Omega$  is a bounded domain in  $\mathbb{R}^n$ ,  $1 \leq n \leq 3$ , which is supposed to either be convex or have a  $C^{1,1}$  boundary  $\Gamma$ . Hereafter  $\mathcal{M}(\Omega)$  denotes the space of regular Borel measures in  $\Omega$ , and  $\|u\|_{L^2(\mathcal{M})}$  denotes the norm of  $u$  in the space  $L^2(I, \mathcal{M}(\Omega))$ ; see section 2 below for details.

Formulating the control problem in a measure space is motivated by the observation that the resulting optimal controls possess sparsity properties (i.e., have small support), which is desirable in many applications such as optimal sensor or actuator placement; see [5, 2] in the context of elliptic equations. Although similar features can be achieved using  $L^1$  control costs, the corresponding control problem in general

\*Received by the editors April 5, 2012; accepted for publication (in revised form) September 5, 2012; published electronically January 8, 2013.

<http://www.siam.org/journals/sicon/51-1/87239.html>

<sup>†</sup>Departamento de Matemática Aplicada y Ciencias de la Computación, E.T.S.I. Industriales y de Telecomunicación, Universidad de Cantabria, 39005 Santander, Spain (eduardo.casas@unican.es). This author was supported by the Spanish Ministerio de Ciencia e Innovación under project MTM2011-22711.

<sup>‡</sup>Institute for Mathematics and Scientific Computing, University of Graz, A-8010 Graz, Austria (christian.clason@uni-graz.at, karl.kunisch@uni-graz.at). These authors were supported by the Austrian Science Fund (FWF) under grant SFB F32 (SFB “Mathematical Optimization and Applications in Biomedical Sciences”).

does not admit a solution in the absence of further regularization because  $L^1$  spaces lack the necessary compactness properties. For parabolic problems, the situation is even more delicate since (1.1) is not well-posed for right-hand sides in  $\mathcal{M}(\Omega_T)$  (which would require  $C(\Omega_T)$  regularity for the adjoint equation; see Definition 2.1 below). This leads to considering controls in  $L^2(I, \mathcal{M}(\Omega))$ . The associated norm  $\|u\|_{L^2(\mathcal{M})}$  for the control is a natural one from the point of view of well-posedness of the state equation (1.1) and allows for sparsity in space. The numerical results will illustrate precisely this property of our formulation. The spatiotemporal coupling of the corresponding control cost, however, presents a challenge for deriving numerically useful optimality conditions.

Besides the analysis of the control problem (P), the main focus of this paper consists in providing an approximation framework which, in spite of the difficulties due to the measure space setting, leads to implementable schemes for which a priori error estimates can be provided. We show that the optimal measure controls can be approximated efficiently by linear combinations of Dirac measures in space which are piecewise constant in time. We point out that even after discretization, the control problem is formulated and solved in the measure space.

Let us mention some related works. A similar approximation framework for elliptic control problems in measure spaces was proposed in [2]. Differently from the elliptic case, parabolic control problems with sparsity-promoting constraints have received very little attention. In [3], the approximate control of  $y(T)$  by measures  $u \in \mathcal{M}([t_0, t_1] \times \Omega)$  with  $0 < t_0 < t_1 < T$  is discussed (using the smoothing property of the heat equation to ensure  $y(T) \in L^2(\Omega)$ ); finite-dimensional approximation and numerical solution are not addressed. Although not specifically concerned with parabolic equations, the approach of [9] covers control problems with  $L^1(\Omega, L^2([0, T]))$  control costs (together with additional pointwise control constraints). The resulting optimal controls have *directional sparsity*; i.e., their support is constant in time. In contrast, we will show that solutions to (P) have a nonseparable sparsity structure.

This paper is organized as follows. In the next section, we discuss the functional analytic setting of the control problem and analyze well-posedness of the state equation. Section 3 is concerned with existence of and optimality conditions for solutions to (P), the latter implying a sparsity property of the optimal controls. The proposed approximation framework is the subject of section 4, where we introduce the discretization (section 4.1) and show convergence of solutions to the discretized state equation (section 4.2) and to the discrete optimal control problem (section 4.3). Convergence rates are derived in section 5. Section 6 addresses the numerical solution of the discrete control problem, for which we derive a reformulated optimality system that is amenable to solution by a semismooth Newton method. (The continuous counterpart of this optimality system is sketched in the appendix.) Finally, section 7 illustrates the structure of the optimal controls with some numerical examples.

**2. Function spaces and well-posedness of the state equation.** In this section we first define the control space and give some of its properties. Then we turn to the analysis of the state equation.

**2.1. Control space.** We denote by  $C_0(\Omega)$  the space of continuous functions in  $\bar{\Omega}$  vanishing on  $\Gamma = \partial\Omega$ , endowed with the supremum norm  $\|\cdot\|_\infty$ . Its topological dual is identified with the space of regular Borel measures in  $\Omega$ , denoted by  $\mathcal{M}(\Omega)$ .

Moreover, we have

$$\|u\|_{\mathcal{M}} = \sup \left\{ \int_{\Omega} z \, du : z \in C_0(\Omega) \text{ and } \|z\|_{\infty} \leq 1 \right\} = |u|(\Omega),$$

where  $|u|$  denotes the total variation measure.

Associated to the interval  $I = [0, T]$  we define the spaces  $L^2(I, C_0(\Omega))$  and  $L^2(I, \mathcal{M}(\Omega))$ , where  $L^2(I, C_0(\Omega))$  is the space of measurable functions  $z : [0, T] \rightarrow C_0(\Omega)$  for which the associated norm given by

$$\|z\|_{L^2(C_0)} = \left( \int_0^T \|z(t)\|_{\infty}^2 \, dt \right)^{1/2}$$

is finite. Due to the fact that  $C_0(\Omega)$  is a separable Banach space,  $L^2(I, C_0(\Omega))$  is also a separable Banach space; see, e.g., [18, Theorem I.5.18].

As a consequence of the nonseparability of  $\mathcal{M}(\Omega)$ , the definition of the space  $L^2(I, \mathcal{M}(\Omega))$  is more delicate. Indeed, we need to distinguish between weakly and strongly measurable functions  $u : [0, T] \rightarrow \mathcal{M}(\Omega)$ . Hereafter we denote by  $L^2(I, \mathcal{M}(\Omega))$  the space of weakly measurable functions  $u$  for which the norm

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

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

$$\langle u, z \rangle_{L^2(\mathcal{M}), L^2(C_0)} = \int_0^T \langle u(t), z(t) \rangle \, dt,$$

with  $\langle \cdot, \cdot \rangle$  denoting the duality between  $\mathcal{M}(\Omega)$  and  $C_0(\Omega)$ . The reader is referred to [6, section 8.14.1 and Proposition 8.15.3] for the different notions of measurability and [6, Theorem 8.20.3] for the duality identification. (The distinction between weak and strong measurability is not required for the space  $L^2(I, C_0(\Omega))$  because  $C_0(\Omega)$  is separable, and hence both notions are equivalent; see [6, Theorem 8.15.2].)

**2.2. Analysis of the state equation.** Given  $1 < p < \infty$ , we denote by  $W_0^{1,p}(\Omega)$  the Sobolev space of functions of  $L^p(\Omega)$  with distributional derivatives in  $L^p(\Omega)$  and having a zero trace on  $\Gamma$ , and we set  $W^{-1,p'}(\Omega)$  to be the dual of  $W_0^{1,p}(\Omega)$ , where  $1/p' + 1/p = 1$ . These spaces are reflexive and separable, and hence the spaces  $L^2(I, W_0^{1,p}(\Omega))$  formed by the measurable functions  $y : [0, T] \rightarrow W_0^{1,p}(\Omega)$  for which the norm

$$\|y\|_{L^2(W_0^{1,p})} = \left( \int_0^T \|y(t)\|_{W_0^{1,p}}^2 \, dt \right)^{1/2}$$

is finite are separable and reflexive Banach spaces whose dual is identified with  $L^2(I, W^{-1,p'}(\Omega))$ ; see [6, Theorem 8.25.5].

The notion of solution to the state equation makes use of the following space of test functions:

$$\mathcal{Z} = \{z \in H^{2,1}(\Omega_T) : z = 0 \text{ on } \Sigma_T \text{ and } z(T) = 0 \text{ in } \Omega\},$$

where

$$H^{2,1}(\Omega_T) = \left\{ z \in L^2(\Omega_T) : \partial_t z, \frac{\partial^{|\beta|} z}{\partial x^\beta} \in L^2(\Omega_T), \text{ with } \beta \in \mathbb{N}^n, |\beta| \leq 2 \right\}$$

is endowed with the graph norm. By the Rellich–Kondrachov theorem,  $\mathcal{Z}$  embeds compactly into  $L^2(I, C_0(\Omega))$ .

DEFINITION 2.1. We say that  $y \in L^2(\Omega_T)$  is a solution to (1.1) if

$$(2.1) \quad \int_{\Omega_T} y(-\partial_t z - \Delta z) dx dt = \int_0^T \langle u(t), z(t) \rangle dt + \int_{\Omega} y_0(x)z(x, 0) dx \quad \forall z \in \mathcal{Z}.$$

THEOREM 2.2. For all  $(u, y_0) \in L^2(I, \mathcal{M}(\Omega)) \times L^2(\Omega)$ , (1.1) has a unique solution  $y$ . Moreover,  $y \in L^2(I, W_0^{1,p}(\Omega))$  for every  $p \in [1, \frac{n}{n-1})$ , and there exist constants  $C_p$  such that

$$(2.2) \quad \|y\|_{L^2(W_0^{1,p})} \leq C_p (\|u\|_{L^2(\mathcal{M})} + \|y_0\|_{L^2(\Omega)}).$$

Proof. We adapt the proof of [1]. Let  $\{u_k\}_k$  be a sequence in  $C(\bar{\Omega}_T)$  satisfying

$$(2.3) \quad u_k \xrightarrow{*} u \text{ in } L^2(I, \mathcal{M}(\Omega)) \text{ and } \|u_k\|_{L^2(L^1)} \leq \|u\|_{L^2(\mathcal{M})}.$$

Let  $y_k \in L^2(I, H_0^1(\Omega))$  denote the variational solution to

$$(2.4) \quad \begin{cases} \partial_t y_k - \Delta y_k = u_k & \text{in } \Omega_T, \\ y_k = 0 & \text{on } \Sigma_T, \\ y_k(x, 0) = y_0(x) & \text{in } \Omega. \end{cases}$$

For  $\psi = (\psi_0, \dots, \psi_n) \in \mathcal{D}(\Omega_T)^{n+1}$  we denote by  $z \in \mathcal{Z}$  the solution to

$$(2.5) \quad \begin{cases} -\partial_t z - \Delta z = \psi_0 - \sum_{j=1}^n \partial_{x_j} \psi_j & \text{in } \Omega_T, \\ z = 0 & \text{on } \Sigma_T, \\ z(x, T) = 0 & \text{in } \Omega. \end{cases}$$

From the last two equations we get for any  $1 < p < \frac{n}{n-1}$

$$\begin{aligned} \int_{\Omega_T} \left( \psi_0 y_k + \sum_{j=1}^n \psi_j \partial_{x_j} y_k \right) dx dt &= \int_{\Omega_T} u_k z dx dt + \int_{\Omega} y_0(x)z(x, 0) dx \\ &\leq \|u_k\|_{L^2(L^1)} \|z\|_{L^2(W_0^{1,p'})} + \|y_0\|_{L^2(\Omega)} \|z(0)\|_{L^2(\Omega)}. \end{aligned}$$

In the following estimate we use maximal regularity of the heat equation in an essential way. If  $\Omega$  is convex, its boundary is of Lipschitz class, and hence there exists a  $\hat{p}$  with  $\hat{p} > 4$  if  $n = 2$  and  $\hat{p} > 3$  when  $n = 3$  such that  $\Delta : W_0^{1,p}(\Omega) \rightarrow W^{-1,p}(\Omega)$  is an isomorphism for each  $\hat{p}' < p < \hat{p}$ , where  $1/\hat{p}' + 1/\hat{p} = 1$ ; see [10]. (If  $n = 1$  or if  $\Omega$  has a  $C^{1,1}$  boundary,  $\Delta : W_0^{1,p}(\Omega) \rightarrow W^{-1,p}(\Omega)$  is an isomorphism for every  $1 < p < +\infty$ .) In particular, combining [8, Theorem 5.4] and (2.3), we obtain for every  $\hat{p}' < p < \frac{n}{n-1} < \hat{p}$  the existence of a constant  $\hat{C}_p$  such that

$$\begin{aligned} \int_{\Omega_T} y_k \left( \psi_0 - \sum_{j=1}^n \partial_{x_j} \psi_j \right) dx dt &= \int_{\Omega_T} \left( \psi_0 y_k + \sum_{j=1}^n \psi_j \partial_{x_j} y_k \right) dx dt \\ &\leq \hat{C}_p (\|y_0\|_{L^2(\Omega)} + \|u\|_{L^2(\mathcal{M})}) \sum_{j=0}^n \|\psi_j\|_{L^2(L^{\hat{p}'})}. \end{aligned}$$

From the density of  $\{\psi_0 - \sum_{j=1}^n \partial_{x_j} \psi_j : \psi \in \mathcal{D}(\Omega_T)^{n+1}\}$  in  $L^2(I, W^{-1,p'}(\Omega))$  and the duality identification  $L^2(I, W_0^{1,p}(\Omega))^* = L^2(I, W^{-1,p'}(\Omega))$ , we deduce the boundedness of  $\{y_k\}_{k=1}^\infty$  in  $L^2(I, W_0^{1,p}(\Omega))$  and the existence of a constant  $C_p$  such that

$$(2.6) \quad \|y_k\|_{L^2(W_0^{1,p})} \leq C_p (\|u\|_{L^2(\mathcal{M})} + \|y_0\|_{L^2(\Omega)}).$$

Using the reflexivity of  $L^2(I, W_0^{1,p}(\Omega))$ , we can obtain a subsequence, denoted in the same way, and an element  $y \in L^2(I, W_0^{1,p}(\Omega))$  such that  $y_k \rightharpoonup y$  in  $L^2(I, W_0^{1,p}(\Omega))$ .

For  $\psi_0 \in L^2(\Omega_T)$  arbitrary and  $z \in \mathcal{Z}$  solution to (2.5) for  $\psi_j = 0$ ,  $1 \leq j \leq n$ , it follows from (2.4) and (2.5) that

$$\int_{\Omega_T} y_k (-\partial_t z - \Delta z) dx dt = \int_{\Omega_T} y_k \psi_0 dx dt = \int_{\Omega_T} u_k z dx dt + \int_{\Omega} y_0(x) z(x, 0) dx.$$

Passing to the limit in this identity and in (2.6), we obtain (2.1) and (2.2). Using the fact that  $\partial_t + \Delta$  is an isomorphism from  $\mathcal{Z}$  to  $L^2(\Omega_T)$  and (2.1), we conclude the uniqueness of  $y \in W_0^{1,p}(\Omega)$ .

Finally, independence of  $y$  with respect to  $p$  follows from the existence of a solution  $y$  in  $L^2(I, W_0^{1,p}(\Omega))$  for every  $\hat{p}' < p < \frac{n}{n-1}$  and its uniqueness in  $L^2(\Omega_T)$ , since  $W_0^{1,p_1}(\Omega) \subset W_0^{1,p_2}(\Omega)$  for  $p_1 > p_2$ .  $\square$

*Remark 2.3.* (i) The solution to (1.1) belongs to  $L^2(I, W_0^{1,p}(\Omega))$  for every  $\hat{p} \leq p < \frac{n}{n-1}$ , and from (1.1) we know that  $\partial_t y \in L^2(I, W^{-1,p}(\Omega))$ . Observe that  $W_0^{1,p}(\Omega) \subset L^2(\Omega)$  for  $p \geq p_0 := \max\{\hat{p}', \frac{2n}{n+2}\}$ , with  $\hat{p}$  as in the proof of Theorem 2.2, and hence  $y \in L^2(\Omega_T)$ . As a consequence, we deduce that  $y \in C(I, L^2(\Omega))$ ; see [16, Proposition III.1.2].

(ii) Under our regularity conditions, an equivalent definition for the solution to (1.1) is the following. A function  $y \in L^2(I, W_0^{1,p}(\Omega))$  with  $p_0 < p < \frac{n}{n-1}$  is called a solution to (1.1) if

$$\begin{aligned} - \int_0^T \langle y(t), \partial_t z(t) \rangle_{W_0^{1,p}, W^{-1,p'}} dt + \int_{\Omega_T} \nabla y \nabla z dx dt \\ = \int_0^T \langle u(t), z(t) \rangle dt + \int_{\Omega} y_0(x) z(x, 0) dx \end{aligned}$$

for all  $z \in L^2(I, W_0^{1,p'}(\Omega))$  such that  $\partial_t z \in L^2(I, W^{-1,p'}(\Omega))$  (which implies  $z(\cdot, 0) \in L^2(\Omega)$ ; see (i)) and  $z(T) = 0$ . This follows from (2.1) and the density of  $\mathcal{Z}$  in this new space of test functions. Theorem 2.2 remains valid with this definition if we only assume that  $\Omega$  has a Lipschitz boundary. This is the regularity of  $\Omega$  required to have the maximal parabolic regularity; see [8]. We have chosen the above definition because it is more convenient for the numerical analysis to be developed later in this paper.

(iii) The preceding theorem as well as the rest of the results given in this paper are valid if we replace the heat operator in (1.1) by a more general parabolic operator  $\partial_t + A$  that enjoys maximal parabolic regularity.

We finish this section by proving a continuity result of the states with respect to the controls.

**THEOREM 2.4.** *Let  $\{u_k\}_{k=1}^\infty \subset L^2(I, \mathcal{M}(\Omega))$  be a sequence such that  $u_k \xrightarrow{*} u$  in  $L^2(I, \mathcal{M}(\Omega))$ . If  $y_k$  and  $y$  denote the states associated to  $u_k$  and  $u$ , respectively, then  $\|y_k - y\|_{L^2(\Omega_T)} \rightarrow 0$ .*

*Proof.* For every  $k$ , let  $z_k \in \mathcal{Z}$  satisfy

$$\begin{cases} -\partial_t z_k - \Delta z_k = y - y_k & \text{in } \Omega_T, \\ z_k = 0 & \text{on } \Sigma_T, \\ z_k(x, T) = 0 & \text{in } \Omega. \end{cases}$$

Then, from Definition 2.1 and using the boundedness of  $\{u_k\}_{k=1}^\infty$  in  $L^2(I, \mathcal{M}(\Omega))$ , we have

$$\begin{aligned} \|y - y_k\|_{L^2(\Omega_T)}^2 &= \int_{\Omega_T} (y - y_k)(-\partial_t z_k - \Delta z_k) dx dt = \int_0^T \langle u(t) - u_k(t), z_k(t) \rangle dt \\ &\leq \|u - u_k\|_{L^2(\mathcal{M})} \|z_k\|_{L^2(C_0)} \leq C \|z_k\|_{L^2(C_0)}. \end{aligned}$$

From Theorem 2.2, we know that  $y_k \rightharpoonup y$  in  $L^2(\Omega_T)$ ; therefore  $z_k \rightharpoonup 0$  in  $H^{2,1}(\Omega_T)$ . Since the embedding  $H^{2,1}(\Omega_T) \subset L^2(I, C_0(\Omega))$  is compact, we get that  $\|z_k\|_{L^2(C_0)} \rightarrow 0$ . This convergence and the above inequality conclude the proof.  $\square$

**3. Analysis of the control problem.** In this section we establish existence of an optimal control and derive the optimality conditions.

PROPOSITION 3.1. *The control problem (P) has a unique solution  $\bar{u}$ .*

*Proof.* Let  $\{u_k\}_{k=1}^\infty$  be a minimizing sequence, which is thus bounded in the space  $L^2(I, \mathcal{M}(\Omega))$ . Since the predual  $L^2(I, C_0(\Omega))$  is separable, there exists a subsequence, denoted in the same way, converging weakly-\* to some  $\bar{u} \in L^2(I, \mathcal{M}(\Omega))$ . From Theorem 2.4 we get that  $y(u_k) \rightarrow y(\bar{u})$  strongly in  $L^2(\Omega_T)$ . Hence, the weakly-\* lower semicontinuity of the norm  $\|\cdot\|_{L^2(\mathcal{M})}$  implies that  $\bar{u}$  is a solution. The uniqueness is a consequence of the strict convexity of  $J$ , which follows from the injectivity of the control-to-state mapping.  $\square$

Hereafter  $\bar{u}$  will denote the solution to (P) 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).

THEOREM 3.2. *There exists a unique element  $\bar{\varphi} \in H^{2,1}(\Omega_T)$  satisfying*

$$(3.1) \quad \begin{cases} -\partial_t \bar{\varphi} - \Delta \bar{\varphi} = \bar{y} - y_d & \text{in } \Omega_T, \\ \bar{\varphi} = 0 & \text{on } \Sigma_T, \\ \bar{\varphi}(x, T) = 0 & \text{in } \Omega, \end{cases}$$

such that

$$(3.2) \quad \int_0^T \langle \bar{u}(t), \bar{\varphi}(t) \rangle dt + \alpha \|\bar{u}\|_{L^2(\mathcal{M})} = 0,$$

$$(3.3) \quad \|\bar{\varphi}\|_{L^2(C_0)} \begin{cases} = \alpha & \text{if } \bar{u} \neq 0, \\ \leq \alpha & \text{if } \bar{u} = 0. \end{cases}$$

*Proof.* Let us introduce  $j(u) = \|u\|_{L^2(\mathcal{M})}$  and  $F(u) = \frac{1}{2} \|y(u) - y_d\|_{L^2(\Omega_T)}^2$ , so that  $J(u) = F(u) + \alpha j(u)$ . By the differentiability of  $F$  and the convexity of  $j$  we obtain

$$F'(\bar{u})(u - \bar{u}) + \alpha j(u) - \alpha j(\bar{u}) \geq 0 \quad \forall u \in L^2(I, \mathcal{M}(\Omega)),$$

and hence

$$\int_{\Omega_T} (\bar{y} - y_d)(y(u) - \bar{y}) dx dt + \alpha j(u) - \alpha j(\bar{u}) \geq 0.$$

Utilizing the adjoint equation (3.1) and the state equation (2.1), we deduce from the above inequality

$$(3.4) \quad \int_0^T \langle u(t) - \bar{u}(t), \bar{\varphi}(t) \rangle dt + \alpha j(u) - \alpha j(\bar{u}) \geq 0 \quad \forall u \in L^2(I, \mathcal{M}(\Omega)).$$

Taking  $u = 2\bar{u}$  and  $u = \frac{1}{2}\bar{u}$ , respectively, in (3.4), we obtain (3.2). On the other hand, setting  $u = \bar{u} - v$  in (3.4), it follows that

$$(3.5) \quad \int_0^T \langle v(t), \bar{\varphi}(t) \rangle dt \leq \alpha(j(\bar{u} - v) - j(\bar{u})) \leq \alpha \|v\|_{L^2(\mathcal{M})} \quad \forall v \in L^2(I, \mathcal{M}(\Omega)).$$

By the duality  $L^2(I, \mathcal{M}(\Omega)) = L^2(I, C_0(\Omega))^*$  we have that

$$(3.6) \quad \|\bar{\varphi}\|_{L^2(C_0)} = \max_{\|v\|_{L^2(\mathcal{M})} \leq 1} \int_0^T \langle v(t), \bar{\varphi}(t) \rangle dt \leq \alpha.$$

Then (3.3) is an immediate consequence of (3.2) and (3.6).  $\square$

From now on, we will assume that the optimal control  $\bar{u} \neq 0$ . By using (3.2) and (3.3) we can prove some sparsity property for  $\bar{u}$ . 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.

**THEOREM 3.3.** *For almost every  $t \in I$  the following embeddings hold:*

$$(3.7) \quad \text{Supp}(\bar{u}^+(t)) \subset \{x \in \Omega : \bar{\varphi}(x, t) = -\|\bar{\varphi}(t)\|_\infty\},$$

$$(3.8) \quad \text{Supp}(\bar{u}^-(t)) \subset \{x \in \Omega : \bar{\varphi}(x, t) = +\|\bar{\varphi}(t)\|_\infty\}.$$

*Proof.* Since  $\bar{\varphi} : I \times \bar{\Omega} \rightarrow \mathbb{R}$  is a Carathéodory function, there exists a measurable selection  $t \in I \mapsto x_t \in \bar{\Omega}$  such that  $\bar{\varphi}(x_t, t) = \|\bar{\varphi}(t)\|_\infty$ ; see [7, Chapter 8, Theorem 1.2]. Now, we define the element  $v \in L^2(I, \mathcal{M}(\Omega))$  by  $v(t) = \text{sign}(\bar{\varphi}(x_t)) \|u(t)\|_{\mathcal{M}} \delta_{x_t}$ . We have to check that  $v : I \rightarrow \mathcal{M}(\Omega)$  is weakly measurable. To this end the only delicate point is the weak measurability of  $t \in I \mapsto \delta_{x_t} \in \mathcal{M}(\Omega)$ . This follows from the measurability of the mapping  $t \mapsto x_t$  and the continuity of  $x \in \bar{\Omega} \mapsto \delta_x \in \mathcal{M}(\Omega)$  when  $\mathcal{M}(\Omega)$  is endowed with the weak-\* topology. By definition of  $v$  we get

$$(3.9) \quad \langle v(t), \bar{\varphi}(t) \rangle = \|\bar{u}(t)\|_{\mathcal{M}} \|\bar{\varphi}(t)\|_\infty \geq -\langle \bar{u}(t), \bar{\varphi}(t) \rangle$$

and

$$(3.10) \quad \begin{aligned} \|v\|_{L^2(\mathcal{M})} &= \left( \int_0^T \|\bar{u}(t)\|_{\mathcal{M}}^2 \|\delta_{x_t}\|_{\mathcal{M}}^2 dt \right)^{1/2} = \left( \int_0^T \|\bar{u}(t)\|_{\mathcal{M}}^2 dt \right)^{1/2} \\ &= \|\bar{u}\|_{L^2(\mathcal{M})}. \end{aligned}$$

From (3.2), (3.9), (3.5), and (3.10) we obtain

$$\alpha \|\bar{u}\|_{L^2(\mathcal{M})} = - \int_0^T \langle \bar{u}(t), \bar{\varphi}(t) \rangle dt \leq \int_0^T \langle v(t), \bar{\varphi}(t) \rangle dt \leq \alpha \|v\|_{L^2(\mathcal{M})} = \alpha \|\bar{u}\|_{L^2(\mathcal{M})}.$$

As a consequence of these inequalities and (3.9) we conclude that

$$(3.11) \quad \|\bar{u}(t)\|_{\mathcal{M}} \|\bar{\varphi}(t)\|_\infty = -\langle \bar{u}(t), \bar{\varphi}(t) \rangle \quad \text{for a.e. } t \in I.$$

Finally, (3.7) and (3.8) follow from (3.11) and Lemma 3.4 below applied to  $\mu = -\bar{u}(t)$ .  $\square$

LEMMA 3.4. *Let  $\mu \in \mathcal{M}(\Omega)$  and  $z \in C_0(\Omega)$ , both of them not zero, be such that*

$$(3.12) \quad \langle \mu, z \rangle = \|\mu\|_{\mathcal{M}} \|z\|_{\infty},$$

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

$$(3.13) \quad \text{Supp}(\mu^+) \subset \Omega_+ = \{x \in \Omega : z(x) = +\|z\|_{\infty}\},$$

$$(3.14) \quad \text{Supp}(\mu^-) \subset \Omega_- = \{x \in \Omega : z(x) = -\|z\|_{\infty}\}.$$

*Proof.* We will prove (3.13), the proof of (3.14) being analogous. First we observe that due to (3.12) we obtain for all measures  $\nu \in \mathcal{M}(\Omega)$  with  $\|\nu\|_{\mathcal{M}} \leq \|\mu\|_{\mathcal{M}}$  that

$$(3.15) \quad \langle \nu, z \rangle \leq \|\nu\|_{\mathcal{M}} \|z\|_{\infty} \leq \|\mu\|_{\mathcal{M}} \|z\|_{\infty} = \langle \mu, z \rangle.$$

We have as well that

$$\langle \mu, z \rangle = \langle \mu^+, z^+ \rangle + \langle \mu^-, z^- \rangle - \langle \mu^+, z^- \rangle - \langle \mu^-, z^+ \rangle \leq \langle \mu^+, z^+ \rangle + \langle \mu^-, z^- \rangle.$$

Moreover, the inequality is strict unless  $\mu^+$  and  $\mu^-$  are concentrated at the set of points  $x \in \Omega$  where  $z(x) \geq 0$  and  $z(x) \leq 0$ , respectively. Let us define the sets

$$A_+ = \{x \in \Omega : z(x) \geq 0\} \quad \text{and} \quad A_- = \{x \in \Omega : z(x) \leq 0\}$$

and the measures  $\nu^+ = \mu^+|_{A_+}$ ,  $\nu^- = \mu^-|_{A_-}$ , and  $\nu = \nu^+ - \nu^-$ . Then we have that  $\|\nu\|_{\mathcal{M}} \leq \|\mu\|_{\mathcal{M}}$  and  $\langle \nu, z \rangle > \langle \mu, z \rangle$  if  $\text{Supp}(\mu^+) \not\subset A_+$  or  $\text{Supp}(\mu^-) \not\subset A_-$ . Because of (3.15) we conclude that  $\text{Supp}(\mu^+) \subset A_+$  and  $\text{Supp}(\mu^-) \subset A_-$ . Now we distinguish two cases in the proof of (3.13) depending on whether the norm bound is attained from above.

*Case 1.*  $\max_{x \in \bar{\Omega}} z(x) < \|z\|_{\infty}$ . In this case we prove that  $\mu^+ = 0$ . Indeed, let  $x_0 \in \Omega$  such that  $z(x_0) = -\|z\|_{\infty}$  and define  $\nu = -\mu^+(\Omega)\delta_{x_0} - \mu^-$ . Then it is obvious that  $\|\nu\|_{\mathcal{M}} = \|\mu\|_{\mathcal{M}}$ . If  $\mu^+ \neq 0$ , since the support of  $\mu^+$  is in  $A_+$  and  $\max_{x \in \bar{\Omega}} z(x) < \|z\|_{\infty}$ , we have that

$$\langle \nu, z \rangle = \|z\|_{\infty} \mu^+(\Omega) - \langle \mu^-, z \rangle > \langle \mu^+, z \rangle - \langle \mu^-, z \rangle = \langle \mu, z \rangle,$$

which contradicts (3.15). Then (3.13) holds.

*Case 2.*  $\max_{x \in \bar{\Omega}} z(x) = \|z\|_{\infty}$ . Let  $x_0 \in \Omega$  be such that  $z(x_0) = \|z\|_{\infty}$ . We argue by contradiction and assume that  $\mu^+(S) > 0$ , where

$$S = \{x \in \Omega : 0 \leq z(x) < \|z\|_{\infty}\}.$$

We take  $\nu = \mu^+(\Omega)\delta_{x_0} - \mu^-$  and once again

$$\|\nu\|_{\mathcal{M}} = \|\mu\|_{\mathcal{M}} \quad \text{and} \quad \langle \nu, z \rangle = \mu^+(\Omega)\|z\|_{\infty} - \langle \mu^-, z \rangle > \langle \mu, z \rangle,$$

since  $\mu^+(S) > 0$ . Again this contradicts (3.15). Therefore,  $\mu^+(S) = 0$ , and hence (3.13) follows from the inclusion  $\text{Supp}(\mu^+) \subset A_+$ .  $\square$

COROLLARY 3.5. *There exists  $\bar{\alpha} > 0$  such that  $\bar{u} = 0$  for every  $\alpha > \bar{\alpha}$ .*

*Proof.* Let us denote by  $J_{\alpha}$  the cost functional associated to the parameter  $\alpha$ . Similarly, let  $(u_{\alpha}, y_{\alpha}, \varphi_{\alpha})$  denote the solution to the corresponding optimality system. For each  $\alpha > 0$  we have the inequalities

$$\frac{1}{2} \|y_{\alpha} - y_d\|_{L^2(\Omega_T)}^2 \leq J_{\alpha}(u_{\alpha}) \leq J_{\alpha}(0) = \frac{1}{2} \|\hat{y}_0 - y_d\|_{L^2(\Omega_T)}^2,$$



where  $\hat{y}_0$  denotes the uncontrolled state, i.e., the solution to (1.1) with  $u = 0$ . Consequently,  $\|y_\alpha - y_d\|_{L^2(\Omega_T)} \leq \|\hat{y}_0 - y_d\|_{L^2(\Omega_T)}$  holds for every  $\alpha > 0$ . From the adjoint state equation (3.1) and the embedding of  $H^{2,1}(\Omega_T) \hookrightarrow L^2(I, C(\bar{\Omega}))$ , we deduce the existence of a constant  $C > 0$  such that

$$\|\varphi_\alpha\|_{L^2(C_0)} \leq C' \|\bar{\varphi}\|_{H^{2,1}} \leq C \|y_\alpha - y_d\|_{L^2(\Omega_T)} \leq C \|\hat{y}_0 - y_d\|_{L^2(\Omega_T)}.$$

Setting  $\bar{\alpha} = C \|\hat{y}_0 - y_d\|_{L^2(\Omega_T)}$ , we obtain from the above inequality and (3.3) that  $u_\alpha = 0$  for every  $\alpha > \bar{\alpha}$ .  $\square$

**4. Approximation of the control problem.** We consider a dG(0)cG(1) discontinuous Galerkin approximation of the state equation (1.1) (i.e., piecewise constant in time and linear nodal basis finite elements in space; see, e.g., [17]). Associated with a parameter  $h$  we consider a family of triangulations  $\{\mathcal{K}_h\}_{h>0}$  of  $\bar{\Omega}$ . To every element  $K \in \mathcal{K}_h$  we assign two parameters  $\rho(K)$  and  $\vartheta(K)$ , where  $\rho(K)$  denotes the diameter of  $K$  and  $\vartheta(K)$  is the diameter of the biggest ball contained in  $K$ . The size of the grid is given by  $h = \max_{K \in \mathcal{K}_h} \rho(K)$ . We will denote by  $\{x_j\}_{j=1}^{N_h}$  the interior nodes of the triangulation  $\mathcal{K}_h$ . In this section  $\Omega$  will be assumed to be convex. In addition, the following usual regularity assumptions on the triangulation are assumed:

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

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

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

- (ii) Let us set  $\bar{\Omega}_h = \cup_{K \in \mathcal{K}_h} K$  with  $\Omega_h$  and  $\Gamma_h$  being its interior and boundary, respectively. We assume that the vertices of  $\mathcal{K}_h$  placed on the boundary  $\Gamma_h$  are also points of  $\Gamma$  and that there exists a constant  $C_\Gamma > 0$  such that  $\text{dist}(x, \Gamma) \leq C_\Gamma h^2$  for every  $x \in \Gamma_h$ . This always holds if  $\Gamma$  is a  $C^2$  boundary. In the case of polygonal or polyhedral domains, it is reasonable to assume that the triangulation satisfies that  $\Gamma_h = \Gamma$ . From this assumption we know [14, section 5.2] that

$$(4.1) \quad |\Omega \setminus \Omega_h| \leq Ch^2,$$

where  $|\cdot|$  denotes the Lebesgue measure.

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

$$(4.2) \quad \tau \leq \rho_T \tau_k \text{ for } 1 \leq k \leq N_\tau \quad \text{and} \quad c_{\Omega, T} h^{\max\{n, 2\}} \leq \tau \leq C_{\Omega, T} h^{\max\{n, 2\}}.$$

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

**4.1. Discretization of the controls and states.** We first discuss the spatial discretization, which follows [2]. Associated to the interior nodes  $\{x_j\}_{j=1}^{N_h}$  of  $\mathcal{K}_h$  we consider the spaces

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

and

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

where  $\{e_j\}_{j=1}^{N_h}$  is the nodal basis formed by the continuous piecewise linear functions such that  $e_j(x_i) = \delta_{ij}$  for every  $1 \leq i, j \leq N_h$ . Such functions attain their maximum and minimum at one of the nodes, and thus for all  $y_h \in Y_h$ ,

$$\|y_h\|_\infty = \max_{1 \leq j \leq N_h} |y_j| = |\vec{y}_h|_\infty,$$

where we have identified  $y_h$  with the vector  $\vec{y}_h = (y_1, \dots, y_{N_h})^T \in \mathbb{R}^{N_h}$  of its expansion coefficients, and  $|\cdot|_p$  denotes the usual  $p$ -norm in  $\mathbb{R}^{N_h}$ . Similarly, we have for all  $u_h \in U_h$  that

$$\|u_h\|_{\mathcal{M}} = \sup_{\|v\|_\infty=1} \sum_{j=1}^{N_h} u_j \langle \delta_{x_j}, v \rangle = \sum_{j=1}^{N_h} |u_j| = |\vec{u}_h|_1 \quad \forall u_h \in U_h.$$

Hence endowed with these norms,  $U_h$  is the topological dual of  $Y_h$  with respect to the duality pairing

$$\langle u_h, y_h \rangle = \sum_{j=1}^{N_h} u_j y_j = \vec{u}_h^T \vec{y}_h.$$

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

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

and

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

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

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

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

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

Thus  $\mathcal{U}_\sigma$  and  $\mathcal{Y}_\sigma$  are finite-dimensional spaces of dimension  $N_\tau \times N_h$ , and bases are given by  $\{\chi_k \delta_{x_j}\}_{k,j}$  and  $\{\chi_k e_j\}_{k,j}$ . Identifying again  $u_\sigma$  with the vector  $\vec{u}_\sigma$  of expansion coefficients  $u_{kj}$ , we have for all  $u_\sigma \in \mathcal{U}_\sigma$  that

$$\begin{aligned} \|u_\sigma\|_{L^2(\mathcal{M})}^2 &= \int_0^T \left\| \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} u_{kj} \chi_k \delta_{x_j} \right\|_{\mathcal{M}}^2 dt = \sum_{k=1}^{N_\tau} \int_{I_k} \left\| \sum_{j=1}^{N_h} u_{kj} \delta_{x_j} \right\|_{\mathcal{M}}^2 dt \\ &= \sum_{k=1}^{N_\tau} \tau_k \left( \sum_{j=1}^{N_h} |u_{kj}| \right)^2 = \sum_{k=1}^{N_\tau} \tau_k |\vec{u}_k|_1^2 \end{aligned}$$

for  $\vec{u}_k = (u_{k1}, \dots, u_{kN_h})^T$ , and similarly for all  $y_\sigma \in \mathcal{Y}_\sigma$  that

$$\|y_\sigma\|_{L^2(C_0)}^2 = \sum_{k=1}^{N_\tau} \tau_k \left( \max_{1 \leq j \leq N_h} |y_{kj}| \right)^2 = \sum_{k=1}^{N_\tau} \tau_k |\vec{y}_k|_\infty^2.$$

It is thus straightforward to verify that, endowed with these norms,  $\mathcal{U}_\sigma$  is the topological dual of  $\mathcal{Y}_\sigma$  with respect to the duality pairing

$$(4.3) \quad \langle u_\sigma, y_\sigma \rangle = \sum_{k=1}^{N_\tau} \tau_k \sum_{j=1}^{N_h} u_{kj} y_{kj} = \sum_{k=1}^{N_\tau} \tau_k (\vec{u}_k^T \vec{y}_k).$$

Next we define the linear operators  $\Lambda_h : \mathcal{M}(\Omega) \rightarrow U_h \subset \mathcal{M}(\Omega)$  and  $\Pi_h : C_0(\Omega) \rightarrow Y_h \subset C_0(\Omega)$  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  $\Pi_h$  is the nodal interpolation operator for  $Y_h$ . Concerning the operator  $\Lambda_h$  we have the following result.

**THEOREM 4.1** (see [2, Theorem 3.1]). *The following properties hold:*

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

$$\begin{aligned} \langle u, y_h \rangle &= \langle \Lambda_h u, y_h \rangle, \\ \langle u, \Pi_h y \rangle &= \langle \Lambda_h u, y \rangle. \end{aligned}$$

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

$$\begin{aligned} \|\Lambda_h u\|_{\mathcal{M}} &\leq \|u\|_{\mathcal{M}}, \\ \Lambda_h u &\xrightarrow{*} u \text{ in } \mathcal{M}(\Omega) \text{ and } \|\Lambda_h u\|_{\mathcal{M}} \rightarrow \|u\|_{\mathcal{M}} \text{ as } h \rightarrow 0. \end{aligned}$$

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

$$\begin{aligned} \|u - \Lambda_h u\|_{W^{-1,p}(\Omega)} &\leq Ch^{1-n/p'} \|u\|_{\mathcal{M}}, \quad 1 < p < \frac{n}{n-1}, \\ \|u - \Lambda_h u\|_{(W_0^{1,\infty}(\Omega))^*} &\leq Ch \|u\|_{\mathcal{M}}, \end{aligned}$$

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

Similarly to  $\Lambda_h$  and  $\Pi_h$  we define the linear operators

$$\Phi_\sigma : L^2(I, \mathcal{M}(\Omega)) \rightarrow \mathcal{U}_\sigma \subset L^2(I, \mathcal{M}(\Omega))$$

and

$$\Psi_\sigma : L^2(I, C_0(\Omega)) \rightarrow \mathcal{Y}_\sigma \subset L^2(I, C_0(\Omega))$$

by

$$\begin{aligned} \Phi_\sigma u &= \sum_{k=1}^{N_\tau} \frac{1}{\tau_k} \int_{I_k} \Lambda_h(u(t)) dt \chi_k = \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} \frac{1}{\tau_k} \int_{I_k} \langle u(t), e_j \rangle dt \chi_k \delta_{x_j}, \\ \Psi_\sigma y &= \sum_{k=1}^{N_\tau} \frac{1}{\tau_k} \int_{I_k} \Pi_h(y(t)) dt \chi_k = \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} \frac{1}{\tau_k} \int_{I_k} y(x_j, t) dt \chi_k e_j. \end{aligned}$$

Analogously to Theorem 4.1 we obtain the following result concerning  $\Phi_\sigma$  and  $\Psi_\sigma$ .

**THEOREM 4.2.** *The following properties hold:*

(i) *For every  $u_\sigma \in \mathcal{U}_\sigma$  and every  $y_\sigma \in \mathcal{Y}_\sigma$  we have*

$$(4.4) \quad \Phi_\sigma u_\sigma = u_\sigma \quad \text{and} \quad \Psi_\sigma y_\sigma = y_\sigma.$$

(ii) *For every  $u \in L^2(I, \mathcal{M}(\Omega))$  and every  $y \in L^2(I, C_0(\Omega))$  and  $y_\sigma \in \mathcal{Y}_\sigma$  we have*

$$(4.5) \quad \langle u, y_\sigma \rangle = \langle \Phi_\sigma u, y_\sigma \rangle,$$

$$(4.6) \quad \langle u, \Psi_\sigma y \rangle = \langle \Phi_\sigma u, y \rangle.$$

(iii) *For every  $u \in L^2(I, \mathcal{M}(\Omega))$  and  $y \in L^2(I, C_0(\Omega))$  we have*

$$(4.7) \quad \|\Phi_\sigma u\|_{L^2(\mathcal{M})} \leq \|u\|_{L^2(\mathcal{M})},$$

$$(4.8) \quad \|\Psi_\sigma y\|_{L^2(C_0)} \leq \|y\|_{L^2(C_0)}.$$

(iv) *For every  $u \in L^2(I, \mathcal{M}(\Omega))$  and  $y \in L^2(I, C_0(\Omega))$  we have*

$$(4.9) \quad \Phi_\sigma u \xrightarrow{*} u \text{ in } L^2(I, \mathcal{M}(\Omega)) \text{ and } \|\Phi_\sigma u\|_{L^2(\mathcal{M})} \rightarrow \|u\|_{L^2(\mathcal{M})},$$

$$(4.10) \quad \Psi_\sigma y \rightarrow y \text{ in } L^2(I, C_0(\Omega)).$$

*Proof.* The formulas of (4.4) follow from the linearity of the operators and the identities  $\Phi_\sigma(\chi_l \delta_{x_i}) = \chi_l \delta_{x_i}$  and  $\Psi_\sigma(\chi_l e_i) = \chi_l e_i$  for all  $1 \leq l \leq N_\tau$  and  $1 \leq i \leq N_h$ .

Identity (4.5) is a consequence of (4.4) and (4.6). Let us prove the latter. First we observe that

$$(4.11) \quad \Phi_\sigma u = \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} u_{kj} \chi_k \delta_{x_j}, \quad \text{with } u_{kj} = \frac{1}{\tau_k} \int_{I_k} \langle u(t), e_j \rangle dt,$$

$$(4.12) \quad \Psi_\sigma y = \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} y_{kj} \chi_k e_j, \quad \text{with } y_{kj} = \frac{1}{\tau_k} \int_{I_k} y(x_j, t) dt.$$

From (4.11) and (4.12) we have

$$\begin{aligned} \langle \Phi_\sigma u, y \rangle &= \int_0^T \langle (\Phi_\sigma u)(t), y(t) \rangle dt = \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} u_{kj} \int_0^T \langle \chi_k \delta_{x_j}, y(t) \rangle dt \\ &= \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} u_{kj} \int_{I_k} y(x_j, t) dt = \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} \tau_k u_{kj} y_{kj}. \end{aligned}$$

Analogously we get

$$\begin{aligned} \langle u, \Psi_\sigma y \rangle &= \int_0^T \langle u(t), (\Psi_\sigma y)(t) \rangle dt = \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} y_{kj} \int_0^T \langle u(t), \chi_k e_j \rangle dt \\ &= \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} y_{kj} \int_{I_k} \langle u(t), e_j \rangle dt = \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} \tau_k u_{kj} y_{kj}, \end{aligned}$$

as desired.

We turn to (4.7). First we recall that the norm of  $\Phi_\sigma u$  is given by

$$\|\Phi_\sigma u\|_{L^2(\mathcal{M})} = \left( \sum_{k=1}^{N_\tau} \tau_k \left( \sum_{j=1}^{N_h} |u_{kj}| \right)^2 \right)^{1/2}.$$

Next we define  $y_\sigma \in \mathcal{Y}_\sigma$  by

$$y_{kj} = \left( \sum_{i=1}^{N_h} |u_{ki}| \right) \text{sign}(u_{kj}),$$

where we set  $\text{sign}(0) = 0$ . For  $y_\sigma$  we compute the expressions

$$\begin{aligned} (4.13) \quad \langle u, y_\sigma \rangle &= \int_0^T \langle u(t), y_\sigma(t) \rangle dt = \sum_{k=1}^{N_\tau} \int_{I_k} \sum_{j=1}^{N_h} y_{kj} \langle u(t), e_j \rangle dt \\ &= \sum_{k=1}^{N_\tau} \tau_k \sum_{j=1}^{N_h} y_{kj} u_{kj} = \sum_{k=1}^{N_\tau} \tau_k \left( \sum_{j=1}^{N_h} |u_{kj}| \right)^2 \\ &= \|\Phi_\sigma u\|_{L^2(\mathcal{M})}^2 \end{aligned}$$

and

$$\begin{aligned} (4.14) \quad \|y_\sigma\|_{L^2(C_0)} &= \left( \int_0^T \|y_\sigma(t)\|_\infty^2 dt \right)^{1/2} = \left( \sum_{k=1}^{N_\tau} \int_{I_k} \left\| \sum_{j=1}^{N_h} y_{kj} e_j \right\|_\infty^2 dt \right)^{1/2} \\ &= \left( \sum_{k=1}^{N_\tau} \tau_k \left( \sum_{j=1}^{N_h} |u_{kj}| \right)^2 \right)^{1/2} \\ &= \|\Phi_\sigma u\|_{L^2(\mathcal{M})}. \end{aligned}$$

From (4.13) and (4.14) we deduce

$$\|\Phi_\sigma u\|_{L^2(\mathcal{M})}^2 = \langle u, y_\sigma \rangle \leq \|u\|_{L^2(\mathcal{M})} \|y_\sigma\|_{L^2(C_0)} = \|u\|_{L^2(\mathcal{M})} \|\Phi_\sigma u\|_{L^2(\mathcal{M})},$$

which implies (4.7).

To establish (4.8) we choose  $y \in L^2(I, C_0(\Omega))$  and estimate

$$\begin{aligned} \|\Psi_\sigma y\|_{L^2(C_0)} &= \left( \sum_{k=1}^{N_\tau} \int_{I_k} \|(\Psi_\sigma y)(t)\|_\infty^2 dt \right)^{1/2} \\ &= \left( \sum_{k=1}^{N_\tau} \frac{1}{\tau_k} \left\| \sum_{j=1}^{N_h} \left( \int_{I_k} y(x_j, t) dt \right) e_j \right\|_\infty^2 \right)^{1/2} \\ &\leq \left( \sum_{k=1}^{N_\tau} \frac{1}{\tau_k} \left( \int_{I_k} \|y(t)\|_\infty dt \right)^2 \right)^{1/2} \leq \left( \sum_{k=1}^{N_\tau} \int_{I_k} \|y(t)\|_\infty^2 dt \right)^{1/2} \\ &= \|y\|_{L^2(C_0)}. \end{aligned}$$

Before proving (4.9), we will consider (4.10). It is well known that (4.10) holds for functions in  $C^\infty(\bar{\Omega}_T)$  vanishing on  $\Sigma_T$ . From the density of these functions in  $L^2(I, C_0(\Omega))$  and from inequality (4.8) we deduce (4.10).

Finally, we prove (4.9). From (4.7) we know that  $\{\Phi_\sigma u\}_\sigma$  is bounded in the space  $L^2(I, \mathcal{M}(\Omega))$ . Then there exists a subsequence, denoted in the same way, and an element  $\tilde{u} \in L^2(I, \mathcal{M}(\Omega))$  such that  $\Phi_\sigma u \overset{*}{\rightharpoonup} \tilde{u}$  in  $L^2(I, \mathcal{M}(\Omega))$ . Then for every  $y \in L^2(I, C_0(\Omega))$  it holds that

$$\lim_{\sigma \rightarrow 0} \int_0^T \langle (\Phi_\sigma u)(t), y(t) \rangle dt = \int_0^T \langle \tilde{u}(t), y(t) \rangle dt.$$

Using (4.6) and (4.10), we find

$$\lim_{\sigma \rightarrow 0} \int_0^T \langle (\Phi_\sigma u)(t), y(t) \rangle dt = \lim_{\sigma \rightarrow 0} \int_0^T \langle u(t), (\Psi_\sigma y)(t) \rangle dt = \int_0^T \langle u(t), y(t) \rangle dt.$$

Combining these two equalities, we have that

$$\int_0^T \langle \tilde{u}(t), y(t) \rangle dt = \int_0^T \langle u(t), y(t) \rangle dt \quad \forall y \in L^2(I, C_0(\Omega)),$$

and therefore  $u = \tilde{u}$  and the whole sequence  $\{\Phi_\sigma u\}_\sigma$  converges weakly- $*$  to  $u$ .

By the convergence  $\Phi_\sigma u \overset{*}{\rightharpoonup} u$  and (4.7) we obtain

$$\|u\|_{L^2(\mathcal{M})} \leq \liminf_{\sigma \rightarrow 0} \|\Phi_\sigma u\|_{L^2(\mathcal{M})} \leq \limsup_{\sigma \rightarrow 0} \|\Phi_\sigma u\|_{L^2(\mathcal{M})} \leq \|u\|_{L^2(\mathcal{M})},$$

which concludes the proof of (4.9).  $\square$

We finish this section by proving the following approximation result.

**THEOREM 4.3.** *Let  $y$  and  $y^\sigma$  be the solutions to (1.1) corresponding to  $u$  and  $\Phi_\sigma u$ , respectively. Then there exists a constant  $C > 0$  independent of  $u$  and  $\sigma$  such that*

$$(4.15) \quad \|y - y^\sigma\|_{L^2(\Omega_T)} \leq Ch^{2-\frac{n}{2}} \|u\|_{L^2(\mathcal{M})} \quad \forall u \in L^2(I, \mathcal{M}(\Omega)).$$

*Proof.* Let  $f \in L^2(\Omega_T)$  be arbitrary and take  $z \in \mathcal{Z}$  satisfying

$$(4.16) \quad \begin{cases} -\partial_t z - \Delta z = f & \text{in } \Omega_T, \\ z = 0 & \text{on } \Sigma_T, \\ z(x, T) = 0 & \text{in } \Omega. \end{cases}$$

Due to the convexity of  $\Omega$ , there exists a constant  $\tilde{C}$  independent of  $f$  such that  $\|z\|_{H^{2,1}(\Omega_T)} \leq \tilde{C} \|f\|_{L^2(\Omega_T)}$ . By (2.1) and (4.6) we get

$$(4.17) \quad \begin{aligned} \int_{\Omega_T} (y - y^\sigma) f \, dx \, dt &= \int_0^T \langle u(t) - (\Phi_\sigma u)(t), z(t) \rangle dt \\ &= \int_0^T \langle u(t), z(t) - (\Psi_\sigma z)(t) \rangle dt \\ &\leq \|u\|_{L^2(\mathcal{M})} \|z - \Psi_\sigma z\|_{L^2(C_0)}. \end{aligned}$$

Now, we will prove that

$$(4.18) \quad \|z - \Psi_\sigma z\|_{L^2(C_0)} \leq Ch^{2-\frac{n}{2}} \|z\|_{H^{2,1}(\Omega_T)}.$$

From the error estimates of the interpolation in Sobolev spaces [4, Chapter 3] we get (4.19)

$$\begin{aligned} \|z - \Pi_h z\|_{L^2(C_0)} &= \left( \int_0^T \|z(t) - \Pi_h z(t)\|_\infty^2 dt \right)^{1/2} \leq Ch^{2-\frac{n}{2}} \left( \int_0^T \|z(t)\|_{H^2(\Omega)}^2 dt \right)^{1/2} \\ &\leq Ch^{2-\frac{n}{2}} \|z\|_{H^{2,1}(\Omega_T)}. \end{aligned}$$

Here and below,  $C$  denotes a constant independent of  $\sigma$ . By an inverse inequality (see [4, Theorem 17.2]) and using (4.2) for the last inequality in the following estimate, we obtain

$$\begin{aligned} (4.20) \quad \|\Pi_h z - \Psi_\sigma z\|_{L^2(C_0)} &= \left( \sum_{k=1}^{N_\tau} \int_{I_k} \left\| \Pi_h z(t) - \frac{1}{\tau_k} \int_{I_k} \Pi_h z(s) ds \right\|_\infty^2 dt \right)^{1/2} \\ &\leq \left( \sum_{k=1}^{N_\tau} \frac{1}{\tau_k} \int_{I_k} \int_{I_k} \|\Pi_h z(t) - \Pi_h z(s)\|_\infty^2 ds dt \right)^{1/2} \\ &\leq \frac{C}{h^{n/2} \sqrt{\tau}} \left( \sum_{k=1}^{N_\tau} \int_{I_k} \int_{I_k} \|\Pi_h z(t) - \Pi_h z(s)\|_{L^2(\Omega)}^2 ds dt \right)^{1/2} \\ &\leq \frac{C}{h^{n/2} \sqrt{\tau}} \left( \sum_{k=1}^{N_\tau} \int_{I_k} \int_{I_k} \|\Pi_h z(t) - z(t)\|_{L^2(\Omega)}^2 ds dt \right)^{1/2} \\ &\quad + \frac{C}{h^{n/2} \sqrt{\tau}} \left( \sum_{k=1}^{N_\tau} \int_{I_k} \int_{I_k} \|\Pi_h z(s) - z(s)\|_{L^2(\Omega)}^2 ds dt \right)^{1/2} \\ &\quad + \frac{C}{h^{n/2} \sqrt{\tau}} \left( \sum_{k=1}^{N_\tau} \int_{I_k} \int_{I_k} \|z(t) - z(s)\|_{L^2(\Omega)}^2 ds dt \right)^{1/2} \\ &\leq \frac{Ch^2}{h^{n/2}} \|z\|_{H^{2,1}(\Omega_T)} + \frac{C}{h^{n/2} \sqrt{\tau}} \left( \sum_{k=1}^{N_\tau} \int_{I_k} \int_{I_k} \left\| \int_{I_k} \partial_t z(\theta) d\theta \right\|_{L^2(\Omega)}^2 ds dt \right)^{1/2} \\ &\leq C \frac{h^2 + \tau}{h^{n/2}} \|z\|_{H^{2,1}(\Omega_T)} \leq Ch^{2-\frac{n}{2}} \|z\|_{H^{2,1}(\Omega_T)}. \end{aligned}$$

Inequality (4.18) follows from (4.19) and (4.20). Finally, (4.17) and (4.18) lead to

$$\int_{\Omega_T} (y - y^\sigma) f dx dt \leq Ch^{2-\frac{n}{2}} \|z\|_{H^{2,1}(\Omega_T)} \leq Ch^{2-\frac{n}{2}} \|f\|_{L^2(\Omega_T)} \quad \forall f \in L^2(\Omega_T),$$

which implies (4.15).  $\square$

**4.2. Discrete state equation.** In this section we approximate the state equation and provide error estimates. 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  $u \in L^2(I, \mathcal{M}(\Omega))$ , for  $k = 1, \dots, N_\tau$  and  $z_h \in Y_h$  we set

$$(4.21) \quad \begin{cases} \left( \frac{y_{k,h} - y_{k-1,h}}{\tau_k}, z_h \right) + a(y_{k,h}, z_h) = \frac{1}{\tau_k} \int_{I_k} \langle u(t), z_h \rangle dt, \\ y_{0,h} = y_{0h}, \end{cases}$$

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 an element of  $Y_h$  satisfying for some  $C_0 > 0$

$$(4.22) \quad \|y_0 - y_{0h}\|_{H^{-1}(\Omega)} \leq C_0 h \|y_0\|_{L^2(\Omega)}.$$

For instance, we can choose for  $y_{0h}$  the projection  $P_h y_0$  of  $y_0$  on  $Y_h$  given by the variational equation

$$(P_h y_0, z_h) = (y_0, z_h) \quad \forall z_h \in Y_h.$$

For any such choice of  $y_{0h}$ , the estimate (4.22) implies that there exists a constant  $C_1 > 0$  independent of  $h$  such that

$$(4.23) \quad \|y_{0h}\|_{L^2(\Omega)} \leq C_1 \|y_0\|_{L^2(\Omega)}.$$

Indeed, by using an inverse inequality and the well-known estimates for the projection operator  $P_h : L^2(\Omega) \rightarrow Y_h$ , we obtain

$$\begin{aligned} \|y_{0h}\|_{L^2(\Omega)} &\leq \|y_{0h} - P_h y_0\|_{L^2(\Omega)} + \|P_h y_0\|_{L^2(\Omega)} \\ &\leq \frac{C}{h} \|y_{0h} - P_h y_0\|_{H^{-1}(\Omega)} + \|y_0\|_{L^2(\Omega)} \\ &\leq \frac{C}{h} (\|y_{0h} - y_0\|_{H^{-1}(\Omega)} + \|y_0 - P_h y_0\|_{H^{-1}(\Omega)}) + \|y_0\|_{L^2(\Omega)} \\ &\leq (C + 1) \|y_0\|_{L^2(\Omega)}. \end{aligned}$$

Obviously (4.21) defines a unique solution  $y_{\sigma}$ . Let us observe that from (4.5) we have the following important consequence.

LEMMA 4.4. *Let  $y_{\sigma}$  and  $\tilde{y}_{\sigma}$  denote the solutions to (4.21) associated to the controls  $u$  and  $\Phi_{\sigma} u$ , respectively. Then the identity  $y_{\sigma} = \tilde{y}_{\sigma}$  holds.*

The rest of the section is devoted to the proof of the stability of the scheme (4.21) and to the derivation of error estimates for  $\|y - y_{\sigma}\|_{L^2(\Omega_T)}$ , where  $y$  and  $y_{\sigma}$  are the solutions to (1.1) and (4.21) associated to a given control  $u \in L^2(I, \mathcal{M}(\Omega))$ . To this end, we introduce some operators that will be used in the proof of the theorems. For every  $h$  we consider the Ritz projection  $R_h : H_0^1(\Omega) \rightarrow Y_h$  given by

$$a(y_h, R_h z) = a(y_h, z) \quad \forall y_h \in Y_h.$$

From the theory of finite elements we know that for all  $z \in H^2(\Omega) \cap H_0^1(\Omega)$ ,

$$(4.24) \quad \begin{cases} \|z - R_h z\|_{L^2(\Omega)} + h \|z - R_h z\|_{H^1(\Omega)} \leq Ch^2 \|z\|_{H^2(\Omega)}, \\ \|z - R_h z\|_{L^{\infty}(\Omega)} \leq Ch^{2-\frac{n}{2}} \|z\|_{H^2(\Omega)}. \end{cases}$$

Now, for every  $\sigma = (\tau, h)$  we define  $\mathcal{R}_{\sigma} : L^2(I, H_0^1(\Omega)) \rightarrow \mathcal{Y}_{\sigma}$  by

$$\mathcal{R}_{\sigma} z = \sum_{k=1}^{N_{\tau}} \frac{1}{\tau_k} \int_{I_k} R_h z(t) \, dt \chi_k = \sum_{k=1}^{N_{\tau}} z_{k,h} \chi_k.$$



The operator  $\mathcal{R}_\sigma$  enjoys for all  $z \in L^2(I, H_0^1(\Omega))$  and  $y_\sigma \in \mathcal{Y}_\sigma$  the property

$$(4.25) \quad \int_0^T a(y_\sigma(t), z(t) - \mathcal{R}_\sigma z(t)) dt = \sum_{k=1}^{N_\tau} \int_{I_k} a(y_{k,h}, z(t) - z_{k,h}) dt = 0.$$

Indeed, for every  $k = 1, \dots, N_\tau$  we have

$$\begin{aligned} \int_{I_k} a(y_{k,h}, z(t)) dt &= \int_{I_k} a(y_{k,h}, R_h z(t)) dt \\ &= \tau_k a\left(y_{k,h}, \frac{1}{\tau_k} \int_{I_k} R_h z(t) dt\right) \\ &= \int_{I_k} a(y_{h,k}, z_{h,k}) dt. \end{aligned}$$

**THEOREM 4.5.** *Given a control  $u \in L^2(I, \mathcal{M}(\Omega))$ , let  $y_\sigma$  be the solution to (4.21) corresponding to  $u$ . Then there exist constants  $C_i > 0$ ,  $i = 1, 2$ , independent of  $u$  and  $\sigma$  such that*

$$(4.26) \quad \sum_{k=1}^{N_\tau} \|y_{k,h} - y_{k-1,h}\|_{L^2(\Omega)}^2 + \tau \max_{1 \leq k \leq N_\tau} \|\nabla y_{k,h}\|_{L^2(\Omega)}^2 \leq C_1 \left( \|y_0\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\mathcal{M})}^2 \right),$$

$$(4.27) \quad \|y_\sigma\|_{L^2(\Omega_T)} \leq C_2 \left( \|y_0\|_{L^2(\Omega)} + \|u\|_{L^2(\mathcal{M})} \right).$$

*Proof.* Let us set  $z_h = y_{k,h} - y_{k-1,h}$  in (4.21). Then we obtain for  $1 \leq k \leq N_\tau$  that

$$\frac{1}{\tau_k} \|y_{k,h} - y_{k-1,h}\|_{L^2(\Omega)}^2 + a(y_{k,h}, y_{k,h} - y_{k-1,h}) = \frac{1}{\tau_k} \int_{I_k} \langle u(t), y_{k,h} - y_{k-1,h} \rangle dt.$$

From here we get with the aid of an inverse estimate [4, Theorem 17.2]

$$\begin{aligned} &\frac{1}{\tau} \|y_{k,h} - y_{k-1,h}\|_{L^2(\Omega)}^2 + \frac{1}{2} [a(y_{k,h}, y_{k,h}) - a(y_{k-1,h}, y_{k-1,h})] \\ &\leq \frac{1}{\tau_k} \|y_{k,h} - y_{k-1,h}\|_{L^2(\Omega)}^2 + \frac{1}{2} [a(y_{k,h}, y_{k,h}) - a(y_{k-1,h}, y_{k-1,h}) \\ &\quad + a(y_{k,h} - y_{k-1,h}, y_{k,h} - y_{k-1,h})] \\ &= \frac{1}{\tau_k} \|y_{k,h} - y_{k-1,h}\|_{L^2(\Omega)}^2 + a(y_{k,h}, y_{k,h} - y_{k-1,h}) \\ &= \frac{1}{\tau_k} \int_{I_k} \langle u(t), y_{k,h} - y_{k-1,h} \rangle dt \\ &\leq \frac{1}{\sqrt{\tau_k}} \|u\|_{L^2(I_k, \mathcal{M})} \|y_{k,h} - y_{k-1,h}\|_\infty \\ &\leq \frac{Ch^{-n/2}}{\sqrt{\tau_k}} \|u\|_{L^2(I_k, \mathcal{M})} \|y_{k,h} - y_{k-1,h}\|_{L^2(\Omega)} \\ &\leq \frac{C^2 h^{-n} \tau}{2\tau_k} \|u\|_{L^2(I_k, \mathcal{M})}^2 + \frac{1}{2\tau} \|y_{k,h} - y_{k-1,h}\|_{L^2(\Omega)}^2 \\ &\leq \frac{C^2 \rho_T C_{\Omega, T}}{2\tau} \|u\|_{L^2(I_k, \mathcal{M})}^2 + \frac{1}{2\tau} \|y_{k,h} - y_{k-1,h}\|_{L^2(\Omega)}^2. \end{aligned}$$

In the last inequality we used (4.2). Summing from  $k = 1$  to  $m$  and using (4.2), it follows that

$$\frac{1}{\tau} \sum_{k=1}^m \|y_{k,h} - y_{k-1,h}\|_{L^2(\Omega)}^2 + a(y_{m,h}, y_{m,h}) - a(y_{0h}, y_{0h}) \leq \frac{C^2 \rho_T C_{\Omega,T}}{\tau} \|u\|_{L^2(\mathcal{M})}^2.$$

Hence

$$(4.28) \quad \sum_{k=1}^m \|y_{k,h} - y_{k-1,h}\|_{L^2(\Omega)}^2 + \tau \|\nabla y_{m,h}\|_{L^2(\Omega)}^2 \leq C(\|y_0\|_{L^2(\Omega)}^2 + \|u\|_{L^2(\mathcal{M})}^2).$$

Here we have used an inverse inequality, (4.2), and (4.23) to get

$$\tau \|y_{0h}\|_{H^1(\Omega)}^2 \leq \frac{C\tau}{h^2} \|y_{0h}\|_{L^2(\Omega)}^2 \leq C \|y_0\|_{L^2(\Omega)}^2.$$

Finally, since  $1 \leq m \leq N_\tau$  is arbitrary, (4.26) follows from (4.28).

Now we prove (4.27). Given  $f \in L^2(\Omega_T)$ , we take  $z \in \mathcal{Z}$  satisfying (4.16). Integrating by parts, we get

$$\begin{aligned} \int_{\Omega_T} y_\sigma f \, dxdt &= \sum_{k=1}^{N_\tau} \int_{I_k} \int_{\Omega} y_{k,h}(x) f(x, t) \, dxdt \\ &= \sum_{k=1}^{N_\tau} \int_{I_k} \{-\partial_t(y_{k,h}, z(t)) + a(y_{k,h}, z(t))\} \, dt \\ &= \sum_{k=1}^{N_\tau} \left\{ (y_{k,h}, z(t_{k-1}) - z(t_k)) + \int_{I_k} a(y_{k,h}, z(t)) \, dt \right\} \\ &= \sum_{k=1}^{N_\tau} \left\{ (y_{k,h} - y_{k-1,h}, z(t_{k-1})) + \int_{I_k} a(y_{k,h}, z(t)) \, dt \right\} + (y_{0h}, z(0)). \end{aligned}$$

Taking  $z_\sigma = \mathcal{R}_\sigma z$ , we get from the above identity and (4.25) that

$$(4.29) \quad \begin{aligned} \int_{\Omega_T} y_\sigma f \, dxdt &= \sum_{k=1}^{N_\tau} \{(y_{k,h} - y_{k-1,h}, z_{k,h}) + \tau_k a(y_{k,h}, z_{k,h})\} + (y_{0h}, z(0)) \\ &+ \sum_{k=1}^{N_\tau} \left\{ (y_{k,h} - y_{k-1,h}, z(t_{k-1}) - z_{k,h}) + \int_{I_k} a(y_{k,h}, z(t) - z_{k,h}) \, dt \right\} \\ &= \int_0^T \langle u(t), z_\sigma(t) \rangle \, dt + (y_{0h}, z(0)) + \sum_{k=1}^{N_\tau} (y_{k,h} - y_{k-1,h}, z(t_{k-1}) - z_{k,h}). \end{aligned}$$

Let us estimate each of these terms. From the definition of  $z_\sigma$  and (4.23) we obtain

$$(4.30) \quad \begin{aligned} \int_0^T \langle u(t), z_\sigma(t) \rangle \, dt + (y_{0h}, z(0)) &\leq \|u\|_{L^2(\mathcal{M})} \|z_\sigma\|_{L^2(C_0)} + \|y_{0h}\|_{L^2(\Omega)} \|z(0)\|_{L^2(\Omega)} \\ &\leq C \|z\|_{H^{2,1}(\Omega_T)} (\|u\|_{L^2(\mathcal{M})} + \|y_0\|_{L^2(\Omega)}), \end{aligned}$$

where we have used that there exists a constant  $C > 0$  independent of  $\sigma$  such that

$$(4.31) \quad \|\mathcal{R}_\sigma v\|_{L^2(C_0)} \leq C \|v\|_{H^{2,1}(\Omega_T)} \quad \forall v \in H^{2,1}(\Omega_T).$$

Indeed,

$$\begin{aligned} \|\mathcal{R}_\sigma v\|_{L^2(C_0)} &= \left( \sum_{k=1}^{N_\tau} \int_{I_k} \|\mathcal{R}_\sigma v(t)\|_\infty^2 dt \right)^{1/2} = \left( \sum_{k=1}^{N_\tau} \int_{I_k} \left\| \frac{1}{\tau_k} \int_{I_k} R_h v(s) ds \right\|_\infty^2 dt \right)^{1/2} \\ &\leq \left( \sum_{k=1}^{N_\tau} \int_{I_k} \|R_h v(s)\|_\infty^2 ds \right)^{1/2}. \end{aligned}$$

Using (4.24), we deduce that

$$\|R_h w\|_\infty \leq \|R_h w - w\|_\infty + \|w\|_\infty \leq Ch^\kappa \|w\|_{H^2(\Omega)} + \|w\|_\infty \leq C \|w\|_{H^2(\Omega)}$$

for every  $w \in H^2(\Omega) \cap H_0^1(\Omega)$ , with  $\kappa = 1$  if  $n \leq 2$  and  $\kappa = 1/2$  if  $n = 3$ . Then (4.31) follows from the above inequalities.

Concerning the last term of (4.29), we will prove

$$(4.32) \quad \sum_{k=1}^{N_\tau} (y_{k,h} - y_{k-1,h}, z(t_{k-1}) - z_{k,h}) \leq Ch^\kappa \|z\|_{H^{2,1}(\Omega_T)} (\|u\|_{L^2(\mathcal{M})} + \|y_0\|_{L^2(\Omega)}),$$

where  $\kappa$  is defined as above. First we observe that (4.26) implies

$$\begin{aligned} (4.33) \quad &\sum_{k=1}^{N_\tau} |(y_{k,h} - y_{k-1,h}, z(t_{k-1}) - z_{k,h})| \\ &\leq \left( \sum_{k=1}^{N_\tau} \|y_{k,h} - y_{k-1,h}\|_{L^2(\Omega_h)}^2 \right)^{1/2} \left( \sum_{k=1}^{N_\tau} \|z(t_{k-1}) - z_{k,h}\|_{L^2(\Omega_h)}^2 \right)^{1/2} \\ &\leq C (\|u\|_{L^2(\mathcal{M})} + \|y_0\|_{L^2(\Omega)}) \left( \sum_{k=1}^{N_\tau} \|z(t_{k-1}) - z_{k,h}\|_{L^2(\Omega_h)}^2 \right)^{1/2}. \end{aligned}$$

From the definition of  $z_\sigma$  and (4.24) we deduce

$$\begin{aligned} \|z(t_{k-1}) - z_{k,h}\|_{L^2(\Omega_h)} &= \left( \int_{\Omega_h} \left| \frac{1}{\tau_k} \int_{I_k} \{z(t_{k-1}) - R_h z(s)\} ds \right|^2 dx \right)^{1/2} \\ &\leq \left( \frac{1}{\tau_k} \int_{\Omega_h} \int_{I_k} |z(t_{k-1}) - R_h z(s)|^2 ds dx \right)^{1/2} \\ &\leq \left( \frac{1}{\tau_k} \int_{\Omega_h} \int_{I_k} |z(t_{k-1}) - z(s)|^2 ds dx \right)^{1/2} \\ &\quad + \left( \frac{1}{\tau_k} \int_{I_k} \|z(s) - R_h z(s)\|_{L^2(\Omega_h)}^2 ds \right)^{1/2} \\ &\leq \left( \int_{\Omega_h} \int_{I_k} \int_{I_k} |\partial_t z(\theta)|^2 d\theta ds dx \right)^{1/2} \\ &\quad + Ch^2 \left( \frac{1}{\tau_k} \int_{I_k} \|z(s)\|_{H^2(\Omega)}^2 ds \right)^{1/2} \\ &\leq \sqrt{\tau} \|\partial_t z\|_{L^2(I_k, L^2(\Omega))} + \frac{Ch^2 \sqrt{\rho_T}}{\sqrt{\tau}} \|z\|_{L^2(I_k, H^2(\Omega))} \\ &\leq Ch^\kappa (\|\partial_t z\|_{L^2(I_k, L^2(\Omega))} + \|z\|_{L^2(I_k, H^2(\Omega))}). \end{aligned}$$

Inserting this estimate into (4.33), we infer (4.32). Finally, (4.29), (4.30), and (4.32) imply that

$$\int_{\Omega_T} y_\sigma f \, dx dt \leq C \|f\|_{L^2(\Omega_T)} (\|u\|_{L^2(\mathcal{M})} + \|y_0\|_{L^2(\Omega)}) \quad \forall f \in L^2(\Omega_T),$$

which is equivalent to (4.27).  $\square$

In the next theorem we show error estimates for the discretization of the state equation.

**THEOREM 4.6.** *Given  $u \in L^2(I, \mathcal{M}(\Omega))$ , let  $y$  and  $y_\sigma$  be the solutions to (1.1) and (4.21). Then there exists a constant  $C$  independent of  $u \in L^2(I, \mathcal{M}(\Omega))$ ,  $y_0 \in L^2(\Omega)$ , and  $\sigma$  such that*

$$(4.34) \quad \|y - y_\sigma\|_{L^2(\Omega_T)} \leq Ch^\kappa (\|u\|_{L^2(\mathcal{M})} + \|y_0\|_{L^2(\Omega)}),$$

where  $\kappa = 1$  if  $n \leq 2$  and  $\kappa = 1/2$  if  $n = 3$ .

*Proof.* As in the proof of Theorem 4.5, we take an arbitrary element  $f \in L^2(\Omega_T)$ ,  $z \in \mathcal{Z}$  solution to (4.16), and  $z_\sigma = \mathcal{R}_\sigma z$ . Then from (2.1) we obtain

$$(4.35) \quad \int_{\Omega_T} (y - y_\sigma) f \, dx dt = \int_0^T \langle u(t), z(t) \rangle dt + \int_\Omega y_0(x) z(x, 0) dx - \sum_{k=1}^{N_\tau} \int_{I_k} \{-(y_{k,h}, \partial_t z(t)) + a(y_{k,h}, z(t))\} dt.$$

Integrating by parts, we get

$$\begin{aligned} \sum_{k=1}^{N_\tau} \int_{I_k} -(y_{k,h}, \partial_t z(t)) dt &= \sum_{k=1}^{N_\tau} (y_{k,h}, z(t_{k-1}) - z(t_k)) \\ &= \sum_{k=1}^{N_\tau} (y_{k,h} - y_{k-1,h}, z(t_{k-1})) + (y_{0h}, z(0)). \end{aligned}$$

From this identity, (4.21), and (4.25) we deduce

$$\begin{aligned} &\sum_{k=1}^{N_\tau} \int_{I_k} \{-(y_{k,h}, \partial_t z(t)) + a(y_{k,h}, z(t))\} dt \\ &= \sum_{k=1}^{N_\tau} \int_{I_k} \left\{ (y_{k,h} - y_{k-1,h}, z(t_{k-1})) + \int_{I_k} a(y_{k,h}, z(t)) \right\} dt + (y_{0h}, z(0)) \\ &= \sum_{k=1}^{N_\tau} \int_{I_k} \left\{ (y_{k,h} - y_{k-1,h}, z_{k,h}) + \int_{I_k} a(y_{k,h}, z_{k,h}) \right\} dt \\ &\quad + \sum_{k=1}^{N_\tau} \int_{I_k} (y_{k,h} - y_{k-1,h}, z(t_{k-1}) - z_{k,h}) + (y_{0h}, z(0)) \\ &= \int_0^T \langle u(t), z_\sigma(t) \rangle dt + \int_\Omega y_{0h}(x) z(x, 0) dx \\ &\quad + \sum_{k=1}^{N_\tau} \int_{I_k} (y_{k,h} - y_{k-1,h}, z(t_{k-1}) - z_{k,h}). \end{aligned}$$

Inserting this identity in (4.35), we infer

$$(4.36) \quad \int_{\Omega_T} (y - y_\sigma) f \, dx \, dt = \int_0^T \langle u(t), z(t) - \mathcal{R}_\sigma z(t) \rangle \, dt + \int_\Omega (y_0(x) - y_{0h}(x)) z(x, 0) \, dx \\ - \sum_{k=1}^{N_\tau} \int_{I_k} (y_{k,h} - y_{k-1,h}, z(t_{k-1}) - z_{k,h}).$$

Let us estimate each of these three terms. For the first term we observe that

$$\|z - \mathcal{R}_\sigma z\|_{L^2(C_0)} \leq Ch^\kappa \|z\|_{H^{2,1}(\Omega_T)}.$$

The proof of this inequality is the same as that of (4.18); it is enough to replace  $\Pi_h$  by  $R_h$  and to use (4.24). Using this inequality, we obtain the first estimate as follows:

$$(4.37) \quad \left| \int_0^T \langle u(t), z(t) - \mathcal{R}_\sigma z(t) \rangle \, dt \right| \leq \|u\|_{L^2(\mathcal{M})} \|z - \mathcal{R}_\sigma z\|_{L^2(C_0)} \\ \leq ch^\kappa \|u\|_{L^2(\mathcal{M})} \|z\|_{H^{2,1}(\Omega_T)}.$$

For the second term we proceed with the aid of (4.23):

$$(4.38) \quad \left| \int_\Omega (y_0(x) - y_{0h}(x)) z(x, 0) \, dx \right| \leq \|y_0 - y_{0h}\|_{H^{-1}(\Omega)} \|z(0)\|_{H_0^1(\Omega)} \\ \leq Ch \|y_0\|_{L^2(\Omega)} \|z\|_{H^{2,1}(\Omega_T)}.$$

Finally, the third term of (4.36) was estimated in (4.32). Thus, using (4.37), (4.38), and (4.32) in (4.36), the inequality

$$\int_{\Omega_T} (y - y_\sigma) f \, dx \, dt \leq Ch^\kappa (\|u\|_{L^2(\mathcal{M})} + \|y_0\|_{L^2(\Omega)}) \|z\|_{H^{2,1}(\Omega_T)} \\ \leq Ch^\kappa (\|u\|_{L^2(\mathcal{M})} + \|y_0\|_{L^2(\Omega)}) \|f\|_{L^2(\Omega_T)}$$

is obtained, which leads to (4.34).  $\square$

**4.3. Discrete optimal control problem.** The approximation of the optimal control problem (P) is defined as

$$(P_\sigma) \quad \min_{u \in L^2(I, \mathcal{M}(\Omega))} J_\sigma(u) = \frac{1}{2} \|y_\sigma - y_d\|_{L^2(\Omega_{hT})}^2 + \alpha \|u\|_{L^2(\mathcal{M})},$$

where  $y_\sigma$  is the discrete state associated to  $u$ , i.e., the solution to (4.21). Let us observe that, analogously to  $J$ , the functional  $J_\sigma$  is convex. However, it is not strictly convex due to the noninjectivity of the control-to-discrete-state mapping and the nonstrict convexity of the norm of  $L^2(I, \mathcal{M}(\Omega))$ . Although the existence of a solution can be shown in the same way as for problem (P), we therefore cannot deduce its uniqueness. On the other hand, if  $\tilde{u}_\sigma$  is a solution to  $(P_\sigma)$  and if we take  $\bar{u}_\sigma = \Phi_\sigma \tilde{u}_\sigma$ , then Lemma 4.4 and the inequality (4.7) imply that  $J_\sigma(\bar{u}_\sigma) \leq J_\sigma(\tilde{u}_\sigma)$ , and hence  $\bar{u}_\sigma$  is also a solution to  $(P_\sigma)$ . Since for  $u_\sigma \in \mathcal{U}_\sigma$  the mapping  $u_\sigma \mapsto y_\sigma(u_\sigma)$ , with  $y_\sigma(u_\sigma)$  the solution to (4.21) for  $u = u_\sigma$ , is linear, injective, and  $\dim \mathcal{U}_\sigma = \dim \mathcal{Y}_\sigma$ , this mapping is bijective. Therefore, the cost functional  $J_\sigma$  is strictly convex on  $\mathcal{U}_\sigma$ , and hence  $(P_\sigma)$

has a unique solution in  $\mathcal{U}_\sigma$ , which will be denoted by  $\bar{u}_\sigma$  hereafter. We summarize this discussion in the following theorem.

**THEOREM 4.7.** *Problem  $(P_\sigma)$  admits at least one solution. Among all solutions, there exists a unique solution  $\bar{u}_\sigma$  belonging to  $\mathcal{U}_\sigma$ . Moreover, any other solution  $\tilde{u} \in L^2(I, \mathcal{M}(\Omega))$  to  $(P_\sigma)$  satisfies  $\Phi_\sigma \tilde{u} = \bar{u}_\sigma$ .*

*Remark 4.8.* The fact that  $(P_\sigma)$  has exactly one solution in  $\mathcal{U}_\sigma$  is of practical interest. Indeed, recall that  $\bar{u}_\sigma$ , as an element of  $\mathcal{U}_\sigma$ , can be uniquely represented as

$$\bar{u}_\sigma = \sum_{k=1}^{N_\tau} \sum_{j=1}^{N_h} \bar{u}_{kj} \chi_k \delta_{x_j}.$$

The numerical computation of  $\bar{u}_\sigma$  therefore is equivalent to the computation of the coefficients  $\{\bar{u}_{kj} : 1 \leq k \leq N_\tau, 1 \leq j \leq N_h\}$ ; see section 6.

We finish this section by analyzing the convergence of the solution in  $\mathcal{U}_\sigma$  to  $(P)$  to the solution to  $(P)$ .

**THEOREM 4.9.** *For every  $\sigma$ , let  $\bar{u}_\sigma$  be the unique solution to  $(P_\sigma)$  belonging to  $\mathcal{U}_\sigma$  and let  $\bar{u}$  be the solution to  $(P)$ . Then the following convergence properties hold for  $\sigma \rightarrow 0^+$ :*

(4.39)  $\bar{u}_\sigma \overset{*}{\rightharpoonup} \bar{u}$  in  $L^2(I, \mathcal{M}(\Omega))$ ,

(4.40)  $\|\bar{u}_\sigma\|_{L^2(\mathcal{M})} \rightarrow \|\bar{u}\|_{L^2(\mathcal{M})}$ ,

(4.41)  $\|\bar{y} - \bar{y}_\sigma\|_{L^2(\Omega_T)} \rightarrow 0$ ,

(4.42)  $J_\sigma(\bar{u}_\sigma) \rightarrow J(\bar{u})$ ,

where  $\bar{y}$  and  $\bar{y}_\sigma$  are the continuous and discrete states associated to  $\bar{u}$  and  $\bar{u}_\sigma$ , respectively.

*Proof.* First of all, let us show that

(4.43)  $u_\sigma \overset{*}{\rightharpoonup} u$  in  $L^2(I, \mathcal{M}(\Omega))$  implies  $\|y_\sigma - y\|_{L^2(\Omega_T)} \rightarrow 0$ ,

where  $y_\sigma$  and  $y$  are the discrete and continuous states associated to the controls  $u_\sigma$  and  $u$ , respectively. Indeed, let us write  $y - y_\sigma = (y - y^\sigma) + (y^\sigma - y_\sigma)$ , where  $y^\sigma$  is the continuous state associated to  $u_\sigma$ . Then by Theorems 2.4 and 4.6 we deduce (4.43).

Turning to the verification of (4.39), we observe that

$$\alpha \|\bar{u}_\sigma\|_{L^2(\mathcal{M})} \leq J_\sigma(\bar{u}_\sigma) \leq J_\sigma(0) = \frac{1}{2} \|\hat{y}_{\sigma 0} - y_d\|_{L^2(\Omega_{hT})}^2 \leq \frac{1}{2} \|\hat{y}_{\sigma 0} - y_d\|_{L^2(\Omega_T)}^2,$$

with  $\hat{y}_{\sigma 0}$  denoting the uncontrolled discrete state, which implies the boundedness of  $\{\bar{u}_\sigma\}_\sigma$  in  $L^2(I, \mathcal{M}(\Omega))$ . By taking a subsequence, we have that  $\bar{u}_\sigma \overset{*}{\rightharpoonup} u$  in  $L^2(I, \mathcal{M}(\Omega))$ . Then using (4.1), (4.43), lower semicontinuity of the norm  $\|\cdot\|_{L^2(\mathcal{M})}$ , and (4.9), we obtain

$$J(u) \leq \liminf_{\sigma \rightarrow 0} J_\sigma(\bar{u}_\sigma) \leq \limsup_{\sigma \rightarrow 0} J_\sigma(\bar{u}_\sigma) \leq \limsup_{\sigma \rightarrow 0} J_\sigma(\Psi_\sigma \bar{u}) = J(\bar{u}).$$

Hence  $u = \bar{u}$  by the uniqueness of the solution to  $(P)$ , and the whole sequence  $\{\bar{u}_\sigma\}_\sigma$  converges weakly- $*$  to  $\bar{u}$ . In addition, the above inequality implies (4.42). Using again (4.43), we deduce (4.41). Finally, (4.40) follows immediately from (4.41) and (4.42).  $\square$

**5. Error estimates.** We now turn to the proof of error estimates for the optimal costs and for the optimal states. We still require  $\Omega$  to be convex and assume in addition

$$(5.1) \quad y_d \in L^2(I, L^r(\Omega)) \quad \text{with } r = \begin{cases} 2 & \text{if } n = 1, \\ 4 & \text{if } n = 2, \\ \frac{8}{3} & \text{if } n = 3. \end{cases}$$

Recall that  $\bar{y}$  and  $\bar{y}_\sigma$  denote the continuous and discrete states associated to the optimal controls  $\bar{u}$  and  $\bar{u}_\sigma$ , respectively.

**THEOREM 5.1.** *There exists a constant  $C > 0$  independent of  $\sigma$  such that*

$$(5.2) \quad |J(\bar{u}) - J_\sigma(\bar{u}_\sigma)| \leq Ch^\kappa,$$

where  $\kappa = 1$  if  $n \leq 2$  and  $\kappa = 1/2$  if  $n = 3$ .

*Proof.* Taking  $r$  as in (5.1) and using Hölder's inequality and (4.1), we deduce that for all  $\phi \in L^2(I, L^r(\Omega))$  and  $n = 2$  or  $3$ ,

$$(5.3) \quad \|\phi\|_{L^2(I, L^2(\Omega \setminus \Omega_h))} \leq \|\phi\|_{L^2(I, L^r(\Omega \setminus \Omega_h))} |\Omega \setminus \Omega_h|^{\frac{r-2}{2r}} \leq C \|\phi\|_{L^2(I, L^r(\Omega \setminus \Omega_h))} h^{\frac{\kappa}{2}}$$

holds. Observe that  $\Omega = \Omega_h$  for  $n = 1$ ; consequently (5.3) holds with  $C = 0$ .

Let  $y$  and  $y_\sigma$  be the continuous and discrete states associated to a given control  $u$ . As a consequence of (4.34) and (5.3), with  $\phi = y - y_d$ , we obtain

$$(5.4) \quad \begin{aligned} \left| \|y - y_d\|_{L^2(\Omega_T)}^2 - \|y_\sigma - y_d\|_{L^2(\Omega_{hT})}^2 \right| &\leq \|y - y_d\|_{L^2(I, L^2(\Omega \setminus \Omega_h))}^2 \\ &+ (\|y - y_d\|_{L^2(\Omega_{hT})} + \|y_\sigma - y_d\|_{L^2(\Omega_{hT})}) \|y - y_\sigma\|_{L^2(\Omega_{hT})} \\ &\leq C \left( \|y - y_d\|_{L^2(I, L^r(\Omega \setminus \Omega_h))}^2 + \|u\|_{L^2(\mathcal{M})} + \|y_0\|_{L^2(\Omega)} \right) h^\kappa. \end{aligned}$$

Now, by the optimality of  $\bar{u}$  and  $\bar{u}_\sigma$  we have

$$J(\bar{u}) - J_\sigma(\bar{u}) \leq J(\bar{u}) - J_\sigma(\bar{u}_\sigma) \leq J(\bar{u}_\sigma) - J_\sigma(\bar{u}_\sigma),$$

and hence

$$(5.5) \quad |J(\bar{u}) - J_\sigma(\bar{u}_\sigma)| \leq \max \{ |J(\bar{u}) - J_\sigma(\bar{u})|, |J(\bar{u}_\sigma) - J_\sigma(\bar{u}_\sigma)| \}.$$

From (4.40) we deduce that  $\{\bar{u}_\sigma\}_\sigma$  is bounded in  $L^2(I, \mathcal{M}(\Omega))$ . Therefore, (2.2) implies that the continuous associated states  $\{y_{\bar{u}_\sigma}\}_\sigma$  are bounded in  $L^2(I, W_0^{1,p}(\Omega))$  for every  $1 \leq p < \frac{n}{n-1}$ , and therefore in  $L^2(I, L^r(\Omega))$  as well. We now apply (5.4) with  $u = \bar{u}_\sigma$  and  $u = \bar{u}$ , respectively. Together with (5.5) this establishes (5.2).  $\square$

In the following theorem we establish a rate of convergence for the states.

**THEOREM 5.2.** *There exists a constant  $C > 0$  independent of  $h$  such that*

$$(5.6) \quad \|\bar{y} - \bar{y}_\sigma\|_{L^2(\Omega_T)} \leq Ch^{\frac{\kappa}{2}},$$

with  $\kappa$  as defined in Theorem 4.1.

*Proof.* Let  $S : L^2(I, \mathcal{M}(\Omega)) \rightarrow L^2(\Omega_T)$  and  $S_\sigma : L^2(I, \mathcal{M}(\Omega)) \rightarrow L^2(\Omega_T)$  be the solution operators associated to (1.1) and (4.21), respectively. From (4.34) it follows that

$$(5.7) \quad \|Su - S_\sigma u\|_{L^2(\Omega_T)} \leq Ch^\kappa (\|u\|_{L^2(\mathcal{M})} + \|y_0\|_{L^2(\Omega)}).$$

By the optimality of  $\bar{u}$  we have for all  $u \in L^2(I, \mathcal{M}(\Omega))$  that

$$(S\bar{u} - y_d, Su - S\bar{u}) + \alpha[\|u\|_{L^2(\mathcal{M})} - \|\bar{u}\|_{L^2(\mathcal{M})}] \geq 0,$$

where  $(\cdot, \cdot)$  now denotes the scalar product in  $L^2(\Omega_T)$ . In particular, taking  $u = \bar{u}_\sigma$ , we get

$$(5.8) \quad (S\bar{u} - y_d, S\bar{u}_\sigma - S\bar{u}) + \alpha[\|\bar{u}_\sigma\|_{L^2(\mathcal{M})} - \|\bar{u}\|_{L^2(\mathcal{M})}] \geq 0.$$

Analogously, the optimality of  $\bar{u}_\sigma$  implies that

$$(5.9) \quad (S_\sigma \bar{u}_\sigma - y_d, S_\sigma \bar{u} - S_\sigma \bar{u}_\sigma) + \alpha[\|\bar{u}\|_{L^2(\mathcal{M})} - \|\bar{u}_\sigma\|_{L^2(\mathcal{M})}] \geq 0.$$

We point out that by definition of  $Y_h$ , we have  $S_\sigma u = 0$  in  $I \times (\Omega \setminus \Omega_h)$ . Then the scalar product above in  $L^2(\Omega_T)$  coincides with that in  $L^2(\Omega_{hT})$ . Now, we rearrange terms in (5.9) as follows:

$$(5.10) \quad (S\bar{u}_\sigma - y_d, S\bar{u} - S\bar{u}_\sigma) + (S_\sigma \bar{u}_\sigma - S\bar{u}_\sigma, S_\sigma \bar{u} - S_\sigma \bar{u}_\sigma) \\ + (y_d, S\bar{u} - S_\sigma \bar{u} + S_\sigma \bar{u}_\sigma - S\bar{u}_\sigma) + (S\bar{u}_\sigma, S_\sigma \bar{u} - S\bar{u} + S\bar{u}_\sigma - S_\sigma \bar{u}_\sigma) \\ + \alpha[\|\bar{u}\|_{L^2(\mathcal{M})} - \|\bar{u}_\sigma\|_{L^2(\mathcal{M})}] \geq 0.$$

Adding (5.8) and (5.10), we obtain

$$(5.11) \quad \|S\bar{u} - S_\sigma \bar{u}_\sigma\|_{L^2(\Omega_T)}^2 = (S\bar{u} - S_\sigma \bar{u}_\sigma, S\bar{u} - S_\sigma \bar{u}_\sigma) \\ \leq (S_\sigma \bar{u}_\sigma - S\bar{u}_\sigma, S_\sigma \bar{u} - S_\sigma \bar{u}_\sigma) \\ + (y_d - S\bar{u}_\sigma, S\bar{u} - S_\sigma \bar{u} + S_\sigma \bar{u}_\sigma - S\bar{u}_\sigma).$$

Let us estimate the right-hand terms. For the first one we apply the Cauchy–Schwarz inequality and use (5.7) to deduce

$$(5.12) \quad (S_\sigma \bar{u}_\sigma - S\bar{u}_\sigma, S_\sigma \bar{u} - S_\sigma \bar{u}_\sigma) \leq \|S_\sigma \bar{u}_\sigma - S\bar{u}_\sigma\|_{L^2(\Omega_T)} \|S_\sigma \bar{u} - S_\sigma \bar{u}_\sigma\|_{L^2(\Omega_T)} \leq Ch^\kappa,$$

where we have used that  $\{\bar{u}_\sigma\}_\sigma$ ,  $\{S_\sigma \bar{u}\}_\sigma$ , and  $\{S_\sigma \bar{u}_\sigma\}_\sigma$  are bounded due to (4.40) and (4.27). For the second term we use once again (5.7) to obtain

$$(5.13) \quad (y_d - S\bar{u}_\sigma, S\bar{u} - S_\sigma \bar{u} + S_\sigma \bar{u}_\sigma - S\bar{u}_\sigma) \\ + \|y_d - S\bar{u}_\sigma\|_{L^2(\Omega_T)} \|(S - S_\sigma)(\bar{u} - \bar{u}_\sigma)\|_{L^2(\Omega_T)} \\ + C(\|\bar{u} - \bar{u}_\sigma\|_{L^2(\mathcal{M})} + \|y_0\|_{L^2(\Omega)})h^\kappa \leq Ch^\kappa,$$

where we have also used that  $y_d \in L^2(I, L^r(\Omega))$  and (2.2). Finally, (5.11), (5.12), and (5.13) prove (5.6).  $\square$

*Remark 5.3.* Let us observe that (5.2) and (5.6) imply that

$$\left| \|\bar{u}\|_{L^2(\mathcal{M})} - \|\bar{u}_\sigma\|_{L^2(\mathcal{M})} \right| \leq Ch^{\frac{\kappa}{2}}$$

for some constant  $C > 0$  independent of  $\sigma$ .

**6. Numerical solution.** We now address the computation of minimizers  $\bar{u}_\sigma$  of problem  $(P_\sigma)$ . First of all, we note that if we define  $y_{d,\sigma}$  as the  $L^2(\Omega_{hT})$  projection of  $y_d$  on  $\mathcal{Y}_\sigma$ , then

$$J_\sigma(u) = \frac{1}{2} \|y_\sigma - y_{d,\sigma}\|_{L^2(\Omega_{hT})}^2 + \alpha \|u\|_{L^2(\mathcal{M})} + \frac{1}{2} \|y_d - y_{d,\sigma}\|_{L^2(\Omega_{hT})}^2.$$



Therefore, problems  $(P_\sigma)$  and

$$(Q_\sigma) \quad \min_{u \in L^2(I, \mathcal{M}(\Omega))} \tilde{J}_\sigma(u) = \frac{1}{2} \|y_\sigma - y_{d,\sigma}\|_{L^2(\Omega_{hT})}^2 + \alpha \|u\|_{L^2(\mathcal{M})}$$

are equivalent. In this section we present a numerical algorithm to solve  $(Q_\sigma)$  as an alternative formulation to  $(P_\sigma)$ .

Due to the spatiotemporal coupling of the norm in  $L^2(I, \mathcal{M}(\Omega))$ , its subdifferential is difficult to characterize. However, using Fenchel duality combined with an equivalent reformulation that decouples the spatiotemporal structure, we can obtain optimality conditions that can be solved using a semismooth Newton method.

For the reader's convenience, we recall the Fenchel duality theory, e.g., from [7, Chapter 4]. Let  $V$  and  $Y$  be Banach spaces with topological duals  $V^*$  and  $Y^*$ , respectively, and let  $\Lambda : V \rightarrow Y$  be a continuous linear operator. Setting  $\bar{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ , let  $\mathcal{F} : V \rightarrow \bar{\mathbb{R}}$ ,  $\mathcal{G} : Y \rightarrow \bar{\mathbb{R}}$  be convex lower semicontinuous functionals which are not identically equal to  $\infty$  and for which there exists a  $v_0 \in V$  such that  $\mathcal{F}(v_0) < \infty$ ,  $\mathcal{G}(\Lambda v_0) < \infty$ , and  $\mathcal{G}$  is continuous at  $\Lambda v_0$ . Let  $\mathcal{F}^* : V^* \rightarrow \bar{\mathbb{R}}$  denote the Fenchel conjugate of  $\mathcal{F}$  defined by

$$\mathcal{F}^*(q) = \sup_{v \in V} \langle q, v \rangle_{V^*, V} - \mathcal{F}(v),$$

which we can calculate using the fact that

$$(6.1) \quad \mathcal{F}^*(q) = \langle q, v \rangle_{V^*, V} - \mathcal{F}(v) \quad \text{if and only if} \quad q \in \partial \mathcal{F}(v).$$

Here,  $\partial \mathcal{F}$  denotes the subdifferential of the convex function  $\mathcal{F}$ , which reduces to the Gâteaux derivative if it exists, and the left-hand side arises from differentiating the duality pairing.

The Fenchel duality theorem states that under the assumptions given above,

$$(6.2) \quad \inf_{v \in V} \mathcal{F}(v) + \mathcal{G}(\Lambda v) = \sup_{q \in Y^*} -\mathcal{F}^*(\Lambda^* q) - \mathcal{G}^*(-q)$$

holds, and that the right-hand side of (6.2) has at least one solution. Furthermore, the equality in (6.2) is attained at  $(\bar{v}, \bar{q})$  if and only if

$$(6.3) \quad \begin{cases} \Lambda^* \bar{q} \in \partial \mathcal{F}(\bar{v}), \\ -\bar{q} \in \partial \mathcal{G}(\Lambda \bar{v}), \end{cases}$$

where the derivative of the duality pairing again enters the left-hand side.

We now apply the Fenchel duality theorem to  $(Q_\sigma)$ , which we express in terms of the expansion coefficients  $\bar{u}_{kj}$ . Let  $N_\sigma = N_\tau \times N_h$  and identify as above  $u_\sigma \in \mathcal{U}_\sigma$  with the vector  $\bar{u}_\sigma = (u_{11}, \dots, u_{1N_h}, \dots, u_{N_\tau N_h})^T \in \mathbb{R}^{N_\sigma}$  of coefficients, and similarly  $y_{d,\sigma} \in \mathcal{Y}_\sigma$ ; see section 4.1. To keep the notation simple, we will omit the vector arrows from here on. Denote by  $M_h = ((e_j, e_k))_{j,k=1}^{N_h}$  the mass matrix and by  $A_h = (a(e_j, e_k))_{j,k=1}^{N_h}$  the stiffness matrix corresponding to  $Y_h$ . For the sake of presentation, we fix  $y_0 = 0$ . Then the discrete state equation (4.21) can be expressed as  $L_\sigma y_\sigma = u_\sigma$  with

$$L_\sigma = \begin{pmatrix} \tau_1^{-1} M_h + A_h & 0 & 0 \\ -\tau_1^{-1} M_h & \tau_2^{-1} M_h + A_h & 0 \\ 0 & \ddots & \ddots \end{pmatrix} \in \mathbb{R}^{N_\sigma \times N_\sigma}.$$

(Note that the “mass matrix” corresponding to  $(\langle \delta_{x_j}, e_k \rangle)_{j,k=1}^{N_h}$  is the identity.) Introducing for  $v_\sigma \in \mathbb{R}^{N_\sigma}$  the vectors  $v_k = (v_{k1}, \dots, v_{kN_h})^T \in \mathbb{R}^{N_h}$ ,  $1 \leq k \leq N_\tau$ , the discrete optimal control problem (Q $_\sigma$ ) can be stated in reduced form as

$$\min_{u_\sigma \in \mathbb{R}^{N_\sigma}} \frac{1}{2} \sum_{k=1}^{N_\tau} \tau_k [L_\sigma^{-1} u_\sigma - y_{d,\sigma}]_k^T M_h [L_\sigma^{-1} u_\sigma - y_{d,\sigma}]_k + \alpha \left( \sum_{k=1}^{N_\tau} \tau_k |u_k|_1^2 \right)^{1/2}.$$

We now set  $\Lambda : \mathbb{R}^{N_\sigma} \rightarrow \mathbb{R}^{N_\sigma}$ ,  $\Lambda v = L_\sigma^{-1} v$ ,

$$\begin{aligned} \mathcal{F} : \mathbb{R}^{N_\sigma} &\rightarrow \mathbb{R}, & \mathcal{F}(v) &= \alpha \left( \sum_{k=1}^{N_\tau} \tau_k |v_k|_1^2 \right)^{1/2}, \\ \mathcal{G} : \mathbb{R}^{N_\sigma} &\rightarrow \mathbb{R}, & \mathcal{G}(v) &= \frac{1}{2} \sum_{k=1}^{N_\tau} \tau_k (v_k - y_{d,k})^T M_h (v_k - y_{d,k}), \end{aligned}$$

and calculate the Fenchel conjugates with respect to the topology induced by the duality pairing (4.3). For  $\mathcal{G}$ , we have by direct calculation that

$$\begin{aligned} \mathcal{G}^*(q) &= \sup_{v \in \mathbb{R}^{N_\sigma}} \sum_{k=1}^{N_\tau} \tau_k q_k^T v_k - \frac{1}{2} \sum_{k=1}^{N_\tau} \tau_k (v_k - y_{d,k})^T M_h (v_k - y_{d,k}) \\ &= \frac{1}{2} \sum_{k=1}^{N_\tau} \tau_k \left( (q_k + M_h y_{d,k})^T M_h^{-1} (q_k + M_h y_{d,k}) - y_{d,k}^T M_h y_{d,k} \right) \end{aligned}$$

since the supremum is attained if and only if  $q_k = M_h (v_k - y_{d,k})$  for each  $1 \leq k \leq N_\tau$  due to (6.1) and the definition of the duality pairing. For  $\mathcal{F}$ , we appeal to the fact that in any Banach space the Fenchel conjugate (with respect to the weak-\* topology) of a norm is the indicator function of the unit ball with respect to the dual norm (see, e.g., [15, Example 2.2.6]), and to the duality between  $\mathcal{U}_\sigma$  and  $\mathcal{Y}_\sigma$ , to obtain

$$\mathcal{F}^*(q) = \iota_\alpha(q) := \begin{cases} 0 & \text{if } \left( \sum_{k=1}^{N_\tau} \tau_k |q_k|_\infty^2 \right)^{1/2} \leq \alpha, \\ \infty & \text{otherwise.} \end{cases}$$

The adjoint  $\Lambda^* : \mathbb{R}^{N_\sigma} \rightarrow \mathbb{R}^{N_\sigma}$  (with respect to the above duality pairing) is given by  $L_\sigma^{-T}$ . Dropping the constant term in  $\mathcal{G}^*$  and substituting  $p_\sigma = \Lambda^* q_\sigma$ , i.e.,  $q_\sigma = L_\sigma^T p_\sigma$ , we obtain the dual problem

$$(6.4) \quad \min_{p_\sigma \in \mathbb{R}^{N_\sigma}} \frac{1}{2} \sum_{k=1}^{N_\tau} \tau_k \left( [L_\sigma^T p_\sigma]_k - M_h y_{d,k} \right)^T M_h^{-1} \left( [L_\sigma^T p_\sigma]_k - M_h y_{d,k} \right) + \iota_\alpha(p_\sigma).$$

Since  $v_0 = 0 = \Lambda v_0$  satisfies the regular point condition, the Fenchel duality theorem is applicable, implying the existence of a solution  $\bar{p}_\sigma$  which is unique due to the strict convexity in (6.4).

While the second relation of (6.3),

$$(6.5) \quad \tau_k (L_\sigma^T \bar{p}_\sigma)_k = \tau_k M_h (L_\sigma^{-1} \bar{u}_\sigma - y_{d,\sigma})_k \quad \forall 1 \leq k \leq N_\tau,$$

can in principle be used to obtain  $\bar{u}_\sigma$  from  $\bar{p}_\sigma$ , the first relation remains impractical for numerical computation. We thus consider the following equivalent reformulation

of (6.4), which decouples the spatiotemporal constraint given by the term  $\iota_\alpha(p_\sigma)$ :

$$\begin{cases} \min_{p_\sigma \in \mathbb{R}^{N_\sigma}, c_\sigma \in \mathbb{R}^{N_\tau}} \frac{1}{2} \sum_{k=1}^{N_\tau} \tau_k ([L_\sigma^T p_\sigma]_k - M_h y_{d,k})^T M_h^{-1} ([L_\sigma^T p_\sigma]_k - M_h y_{d,k}) \\ \text{s.t. } |p_k|_\infty \leq c_k \quad \forall 1 \leq k \leq N_\tau \quad \text{and} \quad \sum_{k=1}^{N_\tau} \tau_k c_k^2 = \alpha^2, \end{cases}$$

where  $c_\sigma = (c_1, \dots, c_{N_\tau})^T \in \mathbb{R}^{N_\tau}$ . Since the constraints satisfy a Slater condition (take  $p_\sigma = 0$  and  $c_k = T^{-1/2}\alpha$ ,  $1 \leq k \leq N_\tau$ ), we obtain (e.g., from [12]) existence of Lagrange multipliers  $\mu_k^1, \mu_k^2 \in \mathbb{R}^{N_h}$ ,  $1 \leq k \leq N_\tau$ , and  $\lambda \in \mathbb{R}$  such that the (unique) solution  $(\bar{p}_\sigma, \bar{c}_\sigma)$  satisfies the optimality conditions

$$(6.6) \quad \begin{cases} \tau_k [L_\sigma M_\sigma^{-1} (L_\sigma^T \bar{p}_\sigma - M_\sigma y_{d,\sigma})]_k = \mu_k^1 + \mu_k^2, \quad 1 \leq k \leq N_\tau, \\ \sum_{j=1}^{N_h} (-\mu_{kj}^1 + \mu_{kj}^2) + 2\lambda \tau_k \bar{c}_k = 0, \quad 1 \leq k \leq N_\tau, \\ (\mu_k^1)^T (\bar{p}_k - \bar{c}_k) = 0, \quad (\mu_k^2)^T (\bar{p}_k + \bar{c}_k) = 0, \quad \mu_k^1 \leq 0, \quad \mu_k^2 \geq 0, \quad 1 \leq k \leq N_\tau, \\ \sum_{k=1}^{N_\tau} \tau_k \bar{c}_k^2 - \alpha^2 = 0, \end{cases}$$

where  $M_\sigma \in \mathbb{R}^{N_\sigma \times N_\sigma}$  is a block diagonal matrix containing  $N_\tau$  copies of  $M_h$ .

We now rewrite the optimality system in a form amenable to the numerical solution using a semismooth Newton method. First,  $\mu_k^1$  and  $\mu_k^2$  are scaled by  $\tau_k > 0$  to eliminate this factor from the first and second relations (which does not affect the complementarity conditions). Using the componentwise max and min functions, the complementarity conditions for  $\mu_k^1, \mu_k^2$  and  $\bar{p}_k$  can be expressed equivalently for any  $\gamma > 0$  as

$$\mu_k^1 + \max(0, -\mu_k^1 + \gamma(\bar{p}_k - \bar{c}_k)) = 0, \quad \mu_k^2 + \min(0, -\mu_k^2 + \gamma(\bar{p}_k + \bar{c}_k)) = 0.$$

Since  $\mu_k^2 = 0$  if  $\bar{p}_k > -\bar{c}_k$  and  $\mu_k^1 = 0$  if  $\bar{p}_k < \bar{c}_k$ , we have by componentwise inspection

$$\max(0, -\mu_k^1 + \gamma(\bar{p}_k - \bar{c}_k)) = \max(0, -\mu_k^1 - \mu_k^2 + \gamma(\bar{p}_k - \bar{c}_k)).$$

We argue similarly for the min term. Furthermore, comparing the first relation of (6.6) with (6.5), we deduce that  $\bar{u}_k = \mu_k^1 + \mu_k^2$  for all  $1 \leq k \leq N_\tau$ . Finally, to avoid having to form  $M_\sigma^{-1}$ , we introduce  $\bar{y}_\sigma \in \mathbb{R}^{N_\sigma}$  satisfying

$$L_\sigma^T \bar{p}_\sigma = M_\sigma (\bar{y}_\sigma - y_{d,\sigma}).$$

Inserting these relations into (6.6), we obtain for every  $\gamma > 0$  the optimality system

$$(6.7) \quad \begin{cases} L_\sigma \bar{y}_\sigma - \bar{u}_\sigma = 0, \\ L_\sigma^T \bar{p}_\sigma - M_\sigma (\bar{y}_\sigma - y_{d,\sigma}) = 0, \\ \bar{u}_k + \max(0, -\bar{u}_k + \gamma(\bar{p}_k - \bar{c}_k)) + \min(0, -\bar{u}_k + \gamma(\bar{p}_k + \bar{c}_k)) = 0, \quad 1 \leq k \leq N_\tau, \\ \sum_{j=1}^{N_h} [-\max(0, -\bar{u}_k + \gamma(\bar{p}_k - \bar{c}_k)) + \min(0, -\bar{u}_k + \gamma(\bar{p}_k + \bar{c}_k))]_j + 2\lambda \bar{c}_k = 0, \\ \hspace{15em} 1 \leq k \leq N_\tau, \\ \sum_{k=1}^{N_\tau} \tau_k \bar{c}_k^2 - \alpha^2 = 0. \end{cases}$$

Since the max and min functions are globally Lipschitz mappings in finite dimensions, this defines a semismooth equation which can be solved using a generalized Newton method; see, e.g., [13, 11]. Here we recall that the Newton derivative of  $\max(0, v)$  with respect to  $v$  is given componentwise by

$$[D_N \max(0, v)h]_k = \begin{cases} h_k & \text{if } v_k \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

and similarly that of  $\min(0, v)$ . In practice, we have to account for the possibly local convergence of the Newton method. To compute a suitable starting point, as an initialization step we successively solve a sequence of approximating problems that are obtained from (6.7) by replacing the max and min terms with

$$\max(0, \gamma(\bar{p}_k - \bar{c}_k)) \quad \text{and} \quad \min(0, \gamma(\bar{p}_k + \bar{c}_k)),$$

respectively, and letting  $\gamma$  tend to infinity. (This can be interpreted as a Moreau–Yosida regularization of the complementarity conditions.) Since now  $u_k$  no longer appears in the argument of the max and min functions, it can be eliminated from the optimality system using the third equation (which also allows computing  $\bar{u}_k$  given  $(\bar{p}_k, \bar{c}_k)$ ), yielding

$$(6.8) \quad \begin{cases} L_\sigma^T p_\gamma - M_\sigma(y_\gamma - y_{d,\sigma}) = 0, \\ L_\sigma y_\gamma + \gamma[\max(0, p_\gamma - c_\gamma) + \min(0, p_\gamma + c_\gamma)] = 0, \\ \sum_{j=1}^{N_h} \gamma [-\max(0, p_{\gamma,k} - c_{\gamma,k}) + \min(0, p_{\gamma,k} + c_{\gamma,k})]_j + 2\lambda_\gamma c_{\gamma,k} = 0, \quad 1 \leq k \leq N_\tau, \\ \sum_{k=1}^{N_\tau} \tau_k c_{\gamma,k}^2 - \alpha^2 = 0. \end{cases}$$

Starting with  $\gamma = 1$  and  $p^0 = y^0 = 0$ ,  $c^0 = T^{-1/2}\alpha$ , and  $\lambda^0 = 1$ , we solve (6.8) using a semismooth Newton method, increase  $\gamma$ , and compute a new solution for increased  $\gamma$  with the previous solution as starting point. Once a solution satisfies the constraints (or a stopping value  $\gamma^*$  is reached), we use it as a starting point for the solution of (6.7) with  $\gamma = 1$ .

*Remark 6.1.* By virtue of the chosen discretization (specifically, the adjoint consistency of discontinuous Galerkin methods and the discrete topology mirroring the continuous one), the discrete optimality system (6.8) coincides with the discretization of the continuous optimality system obtained by applying Fenchel duality, the relaxation approach, and a Moreau–Yosida approximation to problem (P). Since the continuous optimality system may be of independent interest, the derivation is sketched in the appendix.

**7. Numerical examples.** We illustrate the structure of the optimal controls with some one-dimensional examples. For this purpose we set  $\Omega = (-1, 1)$ ,  $T = 2$ ,  $\nu = 10^{-1}$  and consider the state equation

$$\begin{cases} y_t - \nu \Delta y = u, \\ y(0) = 0, \end{cases}$$

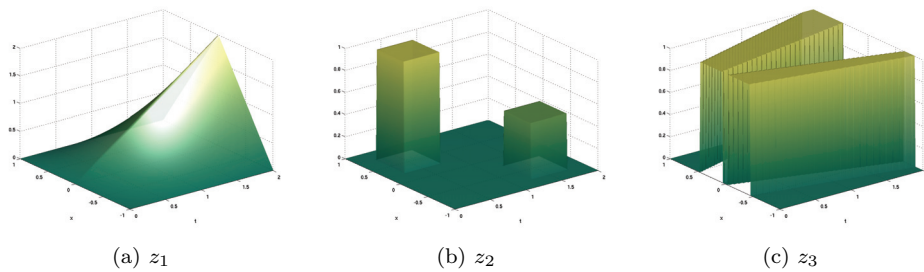


FIG. 7.1. Targets for numerical experiments.

with homogeneous Dirichlet conditions. The spatial domain is discretized using  $N_h = 128$  uniformly distributed nodes (which corresponds to  $h \approx 0.0156$ ). Following (4.2), we take  $N_\tau = 1024$  time steps (which corresponds to  $\tau \approx 0.00195$ ). The targets are chosen as (see Figure 7.1)

$$z_1 = t(1 - |x|),$$

$$z_2 = \begin{cases} 1 & \text{if } 0.25 \leq t \leq 0.75 \text{ and } 0.25 \leq x \leq 0.75, \\ \frac{1}{2} & \text{if } 1.25 \leq t \leq 1.75 \text{ and } -0.25 \geq x \geq -0.75, \\ 0 & \text{otherwise,} \end{cases}$$

$$z_3 = \begin{cases} 1 & \text{if } |x - 0.25 - t/4| < (0.2 + t/20), \\ 1 & \text{if } |x + 0.25 + t/4| < (0.2 - t/20), \\ 0 & \text{otherwise.} \end{cases}$$

The semismooth Newton method for the solution of the optimality system (6.7) is implemented in MATLAB, where the initialization is calculated as discussed in section 6 with  $\gamma_{k+1} = 10\gamma_k$  and  $\gamma^* = 10^{12}$ . For each target the optimal control is computed for  $\alpha = 10^{-3}$  and  $\alpha = 10^{-1}$ . In every case, the discrete optimality system is solved to an accuracy below  $10^{-12}$ , and the bounds on  $p_\sigma$  and on  $c_\sigma$  are attained within machine precision.

The respective optimal controls  $u_\sigma$  (in the form of linearly interpolated expansion coefficients  $u_{kj}$ ), optimal states  $y_\sigma$ , and bounds  $c_\sigma$  are shown in Figures 7.2–7.4. The predicted sparsity structure of the optimal controls can be observed clearly: The spatiotemporal coupling of the control cost predominantly promotes spatial sparsity; see Figure 7.3(b) in particular. The structural features of the norm  $\|u\|_{L^2(\mathcal{M})}$  are further illustrated by the fact that larger values of  $\alpha$  lead to both increased sparsity in space and increased smoothness in time. It is instructive to compare the optimal controls obtained with our  $\|u\|_{L^2(\mathcal{M})}$  regularization to those obtained numerically using a (Moreau–Yosida approximation of an)  $\mathcal{M}(\Omega_T)$ -norm penalty term. Figure 7.5 shows the latter for all considered targets and values of  $\alpha$ . While for  $\alpha = 10^{-3}$  both types of control have comparable structure, for  $\alpha = 10^{-1}$  the controls in  $\mathcal{M}(\Omega_T)$  demonstrate strong temporal sparsity, which is absent in the case of controls in  $L^2(I, \mathcal{M}(\Omega))$ .

We now investigate the convergence behavior as  $h \rightarrow 0$ . In the absence of a known exact solution, we take as a reference solution the computed optimal discrete control and optimal discrete state on the finest grid with  $N_{h^*} = 256$  and  $N_{\tau^*} = 4096$ , corresponding to  $h^* \approx 0.00781$  and  $\tau^* \approx 0.000488$ . As a representative example,

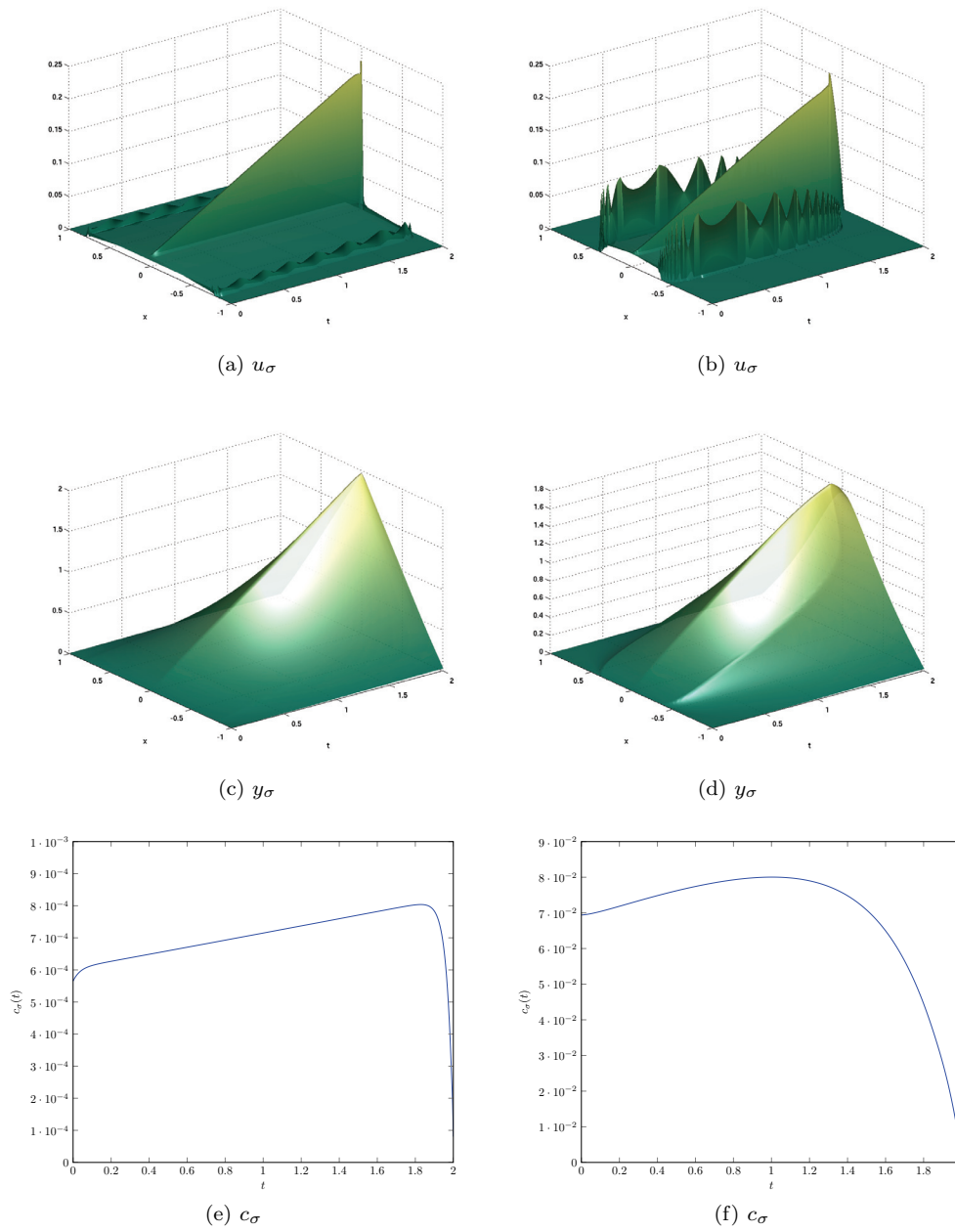


FIG. 7.2. Optimal control  $u_\sigma$ , state  $y_\sigma$ , and bound  $c_\sigma$  for target  $z_1$  and  $\alpha = 10^{-3}$  (left),  $\alpha = 10^{-1}$  (right).

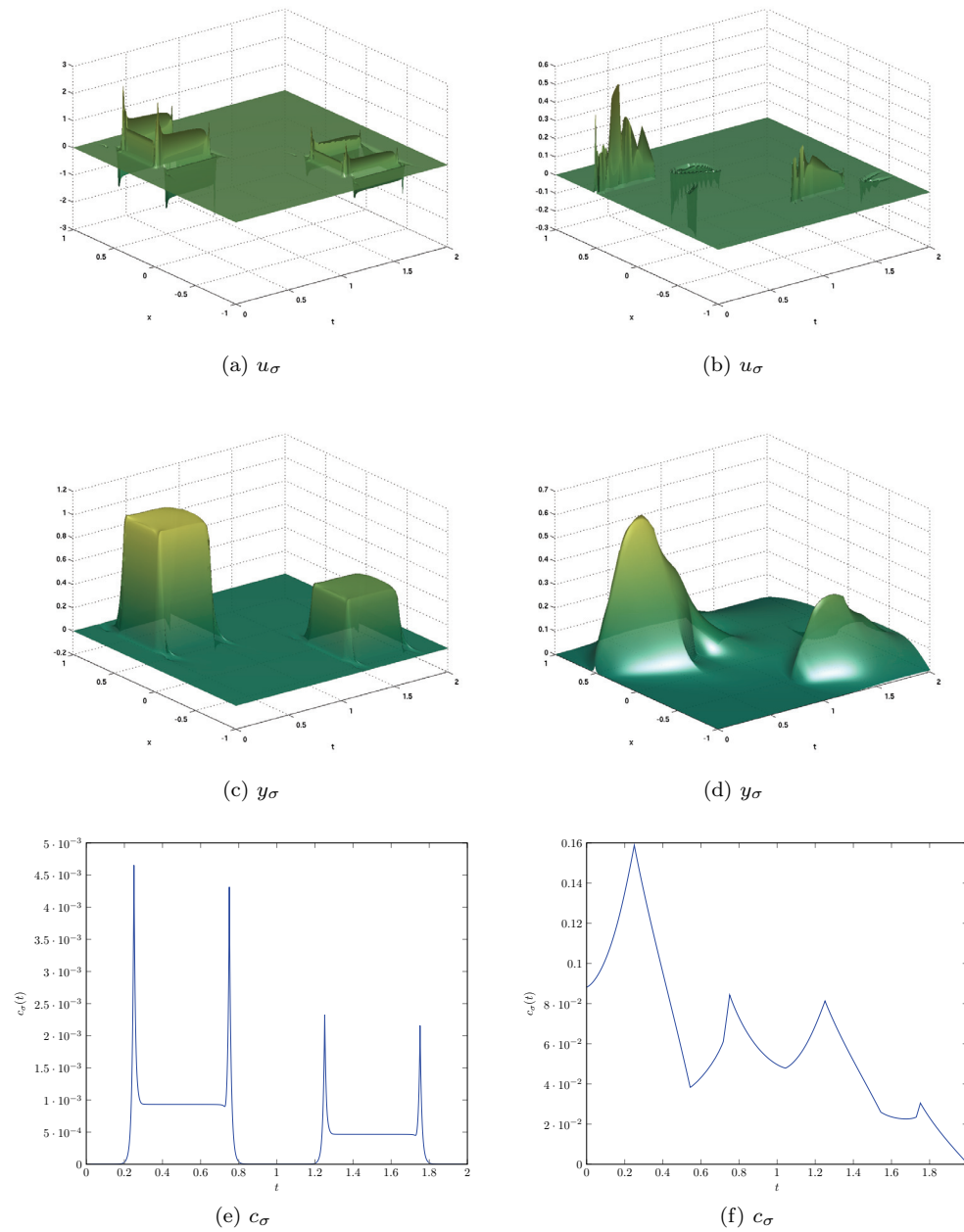
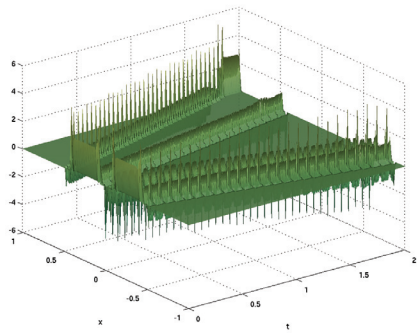
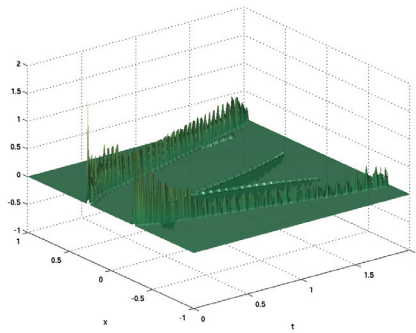


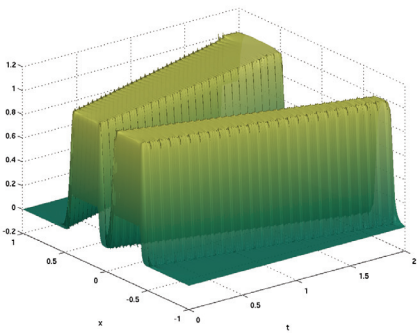
FIG. 7.3. Optimal control  $u_\sigma$ , state  $y_\sigma$ , and bound  $c_\sigma$  for target  $z_2$  and  $\alpha = 10^{-3}$  (left),  $\alpha = 10^{-1}$  (right).



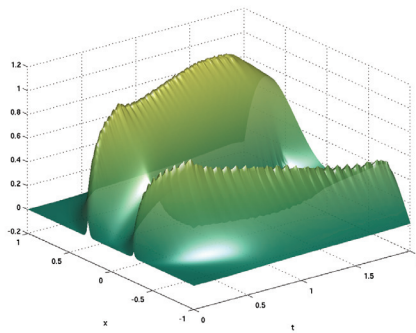
(a)  $u_\sigma$



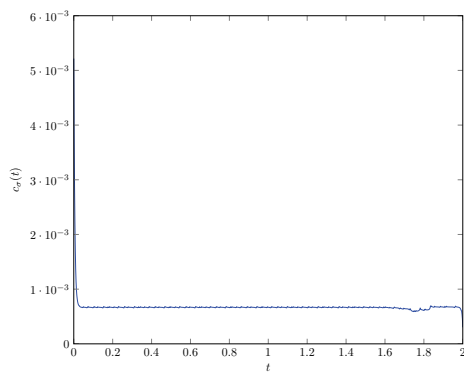
(b)  $u_\sigma$



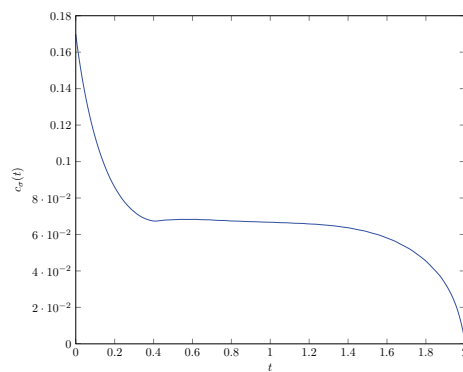
(c)  $y_\sigma$



(d)  $y_\sigma$



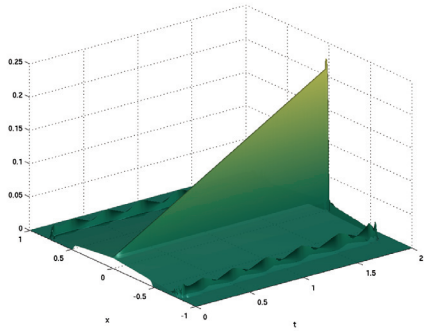
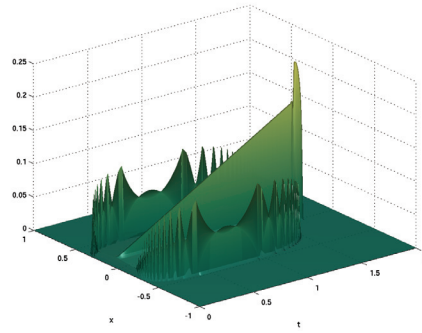
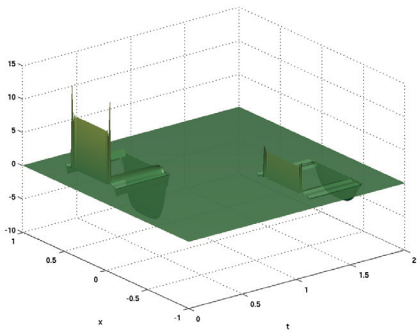
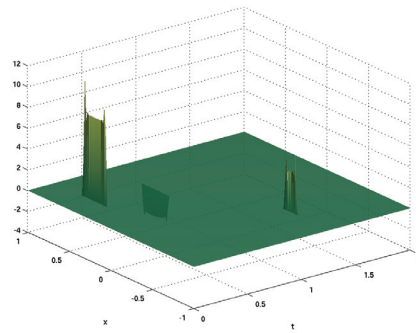
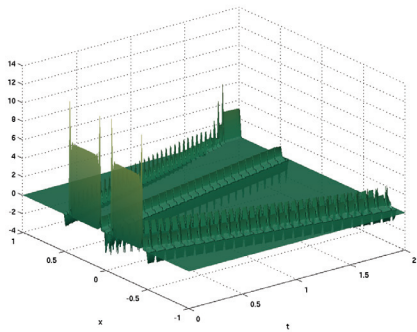
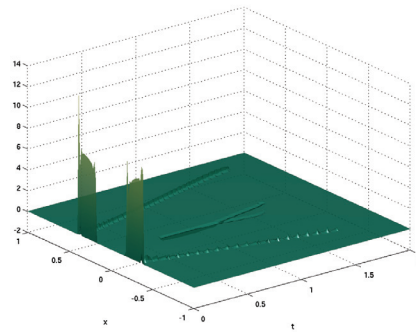
(e)  $c_\sigma$



(f)  $c_\sigma$

FIG. 7.4. Optimal control  $u_\sigma$ , state  $y_\sigma$ , and bound  $c_\sigma$  for target  $z_3$  and  $\alpha = 10^{-3}$  (left),  $\alpha = 10^{-1}$  (right).



(a) target  $z_1$ ,  $\alpha = 10^{-3}$ (b) target  $z_1$ ,  $\alpha = 10^{-1}$ (c) target  $z_2$ ,  $\alpha = 10^{-3}$ (d) target  $z_2$ ,  $\alpha = 10^{-1}$ (e) target  $z_3$ ,  $\alpha = 10^{-3}$ (f) target  $z_3$ ,  $\alpha = 10^{-1}$ FIG. 7.5. Optimal controls with  $\mathcal{M}(\Omega_T)$  penalty.

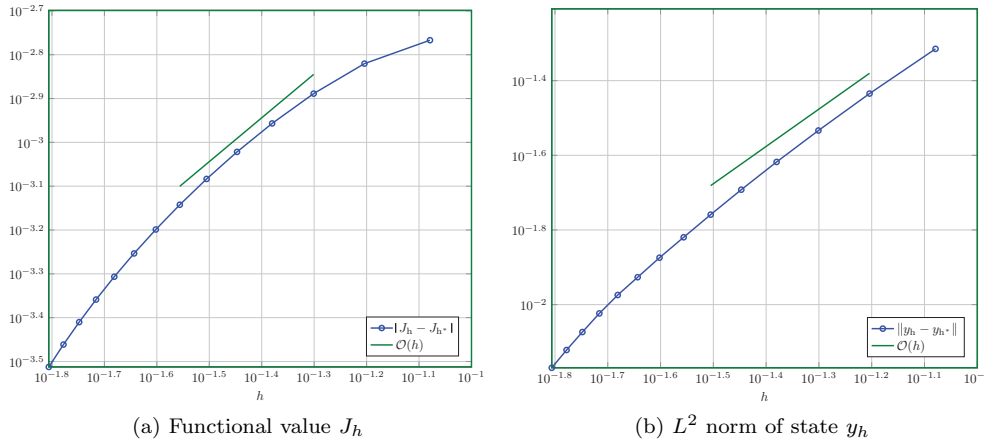


FIG. 7.6. Illustration of convergence order for target  $z_1$  and  $\alpha = 0.1$ .

we consider the target  $z_1$  and  $\alpha = 0.1$ . Figure 7.6(a) shows the difference  $|J_h - J_{h^*}|$  for a series of successively refined grids with  $N_h = 32, 40, \dots, 128$  and  $N_{\tau(h)} = \frac{1}{16} N_h^2$ . The observed approximately linear convergence rate agrees with the rate obtained in Theorem 5.1. The corresponding  $L^2$  error  $\|y_h - y_{h^*}\|_{L^2}$  of the discrete states also decays with a linear rate, which is faster than predicted by Theorem 5.2. A similar behavior was observed in the elliptic case; see [2].

**8. Conclusion.** For the appropriate functional-analytic setting of parabolic optimal control problems in measure spaces, there exists a straightforward approximation framework that retains the structural properties of the norm in the measure-valued Banach space and allows deriving numerically accessible optimality conditions as well as convergence rates. In particular, although the state is discretized, the control problem is still formulated and solved in measure space. The numerical results demonstrate that the optimal controls exhibit the expected sparsity pattern.

**Appendix. Continuous optimality system.** In this section we sketch the derivation of the continuous optimality system using Fenchel duality and the relaxation approach. Let  $S : L^2(I, \mathcal{M}(\Omega)) \rightarrow L^2(\Omega_T)$  denote the solution operator corresponding to the state equation (1.1) with homogeneous initial conditions. It will be convenient to introduce the parabolic differential operator  $L$  such that the solution  $y$  to (1.1) satisfies  $Ly = u$ . Then we can express problem (P) in reduced form as

$$\min_{u \in L^2(I, \mathcal{M}(\Omega))} \frac{1}{2} \|Su - y_d\|_{L^2(\Omega_T)}^2 + \alpha \|u\|_{L^2(I, \mathcal{M}(\Omega))}.$$

To apply Fenchel duality, we set

$$\begin{aligned} \mathcal{F} : L^2(I, \mathcal{M}(\Omega)) &\rightarrow \mathbb{R}, & \mathcal{F}(v) &= \alpha \|v\|_{L^2(I, \mathcal{M}(\Omega))}, \\ \mathcal{G} : L^2(\Omega_T) &\rightarrow \mathbb{R}, & \mathcal{G}(v) &= \frac{1}{2} \|v - y_d\|_{L^2(\Omega_T)}^2, \\ \Lambda : L^2(I, \mathcal{M}(\Omega)) &\rightarrow L^2(\Omega_T), & \Lambda u &= Su. \end{aligned}$$

Similarly to the discrete case, the Fenchel conjugates (with respect to the weak-\* topology) are given by

$$\begin{aligned}\mathcal{F}^* : L^2(I, C_0(\Omega)) &\rightarrow \mathbb{R}, & \mathcal{F}^*(q) &= \iota_\alpha(q), \\ \mathcal{G}^* : L^2(\Omega_T) &\rightarrow \mathbb{R}, & \mathcal{G}^*(q) &= \frac{1}{2}\|q + y_d\|_{L^2(\Omega_T)}^2 - \frac{1}{2}\|y_d\|_{L^2(\Omega_T)}^2,\end{aligned}$$

where

$$\iota_\alpha(q) = \begin{cases} 0 & \text{if } \|q\|_{L^2(I, C_0(\Omega))} \leq \alpha, \\ \infty & \text{otherwise.} \end{cases}$$

Due to the definition of the solution to (1.1) via duality (see Definition 2.1), we obtain the existence of a weak-\* adjoint operator  $\Lambda^* := S^* : L^2(\Omega_T) \rightarrow L^2(I, C_0(\Omega))$  defined via the solution to (2.5). Furthermore, there exists a weak-\* adjoint  $L^*$  of  $L$  such that, for given  $\psi_0 \in L^2(\Omega_T)$ , the solution  $z \in L^2(I, C_0(\Omega))$  of (2.5) satisfies  $L^*z = \psi_0$ . The dual problem is then found to be

$$\min_{q \in L^2(\Omega_T)} \frac{1}{2}\|q - y_d\|_{L^2(\Omega_T)}^2 + \iota_\alpha(S^*q).$$

We again substitute  $p = S^*q \in L^2(I, C_0(\Omega))$ , i.e.,  $q = L^*p$ , introduce  $c \in L^2(I)$  by

$$c(t) := \|p(t)\|_\infty \quad \text{for a.e. } 0 \leq t \leq T,$$

and consider

$$(A.1) \quad \begin{cases} \min_{p \in L^2(I, C_0(\Omega)), c \in L^2(I)} \frac{1}{2}\|L^*p - y_d\|_{L^2(\Omega_T)}^2 \\ \text{s.t. } \|p(t)\|_\infty \leq c(t) \text{ for a.e. } 0 \leq t \leq T \\ \text{and } \int_0^T c(t)^2 dt = \alpha^2. \end{cases}$$

The Moreau–Yosida regularization of (A.1) is given by

$$\begin{cases} \min_{p \in L^2(I, C_0(\Omega)), c \in L^2(I)} \frac{1}{2}\|L^*p - y_d\|_{L^2(\Omega_T)}^2 + \frac{\gamma}{2} \left[ \|\max(0, p - c)\|_{L^2(\Omega_T)}^2 \right. \\ \qquad \qquad \qquad \left. + \|\min(0, p + c)\|_{L^2(\Omega_T)}^2 \right] \\ \text{s.t. } \int_0^T c(t)^2 dt = \alpha^2, \end{cases}$$

where the max and min functions should be understood pointwise in  $\Omega$  for almost every  $0 \leq t \leq T$ . Its solution is denoted by  $(p_\gamma, c_\gamma) \in L^2(I, C_0(\Omega)) \times L^2(I)$ . Since the cost functional is Fréchet differentiable and a Slater condition is again satisfied for the constraint on  $c$  (take  $c = T^{-1/2}\alpha$ ), we obtain existence of a Lagrange multiplier  $\lambda_\gamma \in \mathbb{R}$ . Introducing once more  $y_\gamma$  satisfying  $L^*p_\gamma = y_\gamma - y_d$ , this yields the continuous optimality system

$$\begin{cases} L^*p_\gamma - (y_\gamma - y_d) = 0, \\ Ly_\gamma + \gamma \max(0, p_\gamma - c_\gamma) + \gamma \min(0, p_\gamma + c_\gamma) = 0, \\ \gamma \int_\Omega -\max(0, p_\gamma - c_\gamma) + \min(0, p_\gamma + c_\gamma) dx + 2\lambda_\gamma c_\gamma = 0, \\ \int_0^T c_\gamma^2 dt - \alpha^2 = 0. \end{cases}$$

By approximating  $p_\gamma$  and  $y_\gamma$  in  $\mathcal{Y}_\sigma$ , using the fact that for linear finite elements the pointwise maximum and minimum is attained at the nodes, and the adjoint consistency of discontinuous Galerkin methods (i.e.,  $(L^*)_\sigma = L_\sigma^T$ ), we recover (6.8).

## REFERENCES

- [1] E. CASAS, *Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations*, SIAM J. Control Optim., 35 (1997), pp. 1297–1327.
- [2] E. CASAS, C. CLASON, AND K. KUNISCH, *Approximation of elliptic control problems in measure spaces with sparse solutions*, SIAM J. Control Optim., 50 (2012), pp. 1735–1752.
- [3] E. CASAS AND E. ZUAZUA, *Spike Controls for Elliptic and Parabolic PDE*, tech. rep., Basque Center for Applied Mathematics, 2012.
- [4] P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.
- [5] C. CLASON AND K. KUNISCH, *A measure space approach to optimal source placement*, Comput. Optim. Appl., 53 (2012), pp. 155–171.
- [6] R. E. EDWARDS, *Functional Analysis. Theory and Applications*, Holt, Rinehart and Winston, New York, 1965.
- [7] I. EKELAND AND R. TÉMAM, *Convex Analysis and Variational Problems*, Classics Appl. Math. 28, SIAM, Philadelphia, 1999.
- [8] R. HALLER-DINTELMANN AND J. REHBERG, *Maximal parabolic regularity for divergence operators including mixed boundary conditions*, J. Differential Equations, 247 (2009), pp. 1354–1396.
- [9] R. HERZOG, G. STADLER, AND G. WACHSMUTH, *Directional sparsity in optimal control of partial differential equations*, SIAM J. Control Optim., 50 (2012), pp. 943–963.
- [10] D. JERISON AND C. E. KENIG, *The inhomogeneous Dirichlet problem in Lipschitz domains*, J. Funct. Anal., 130 (1995), pp. 161–219.
- [11] B. KUMMER, *Newton's method based on generalized derivatives for nonsmooth functions: Convergence analysis*, in Advances in Optimization (Lambrecht, 1991), Lecture Notes in Econom. and Math. Systems 382, Springer, Berlin, 1992, pp. 171–194.
- [12] H. MAURER AND J. ZOWE, *First and second order necessary and sufficient optimality conditions for infinite-dimensional programming problems*, Math. Programming, 16 (1979), pp. 98–110.
- [13] L. QI AND J. SUN, *A nonsmooth version of Newton's method*, Math. Programming, 58 (1993), pp. 353–367.
- [14] P.-A. RAVIART AND J.-M. THOMAS, *Introduction à l'Analyse Numérique des Equations aux Dérivées Partielles*, Masson, Paris, 1983.
- [15] W. SCHIROTZEK, *Nonsmooth Analysis*, Universitext, Springer, Berlin, 2007.
- [16] R. E. SHOWALTER, *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*, Math. Surveys Monogr. 49, AMS, Providence, RI, 1997.
- [17] V. THOMÉE, *Galerkin Finite Element Methods for Parabolic Problems*, 2nd ed., Springer Ser. Comput. Math. 25, Springer, Berlin, 2006.
- [18] J. WARGA, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.