ERROR ESTIMATES FOR A POINTWISE TRACKING OPTIMAL CONTROL PROBLEM OF A SEMILINEAR ELLIPTIC EQUATION*

ALEJANDRO ALLENDES†, FRANCISCO FUICA‡, AND ENRIQUE OTÁROLA§

Abstract. We consider a pointwise tracking optimal control problem for a semilinear elliptic partial differential equation. We derive the existence of optimal solutions and obtain first order optimality conditions. We also obtain necessary and sufficient second order optimality conditions. To approximate the solution of the aforementioned optimal control problem we devise a finite element technique that approximates the solution to the state and adjoint equations with piecewise linear functions and the control variable with piecewise constant functions. We analyze convergence properties and prove that the error approximation of the control variable converges with rate $\mathcal{O}(h|\log h|)$ when measured in the L^2-norm.

Key words. optimal control problem, semilinear elliptic PDE, Dirac measures, finite element approximations, maximum-norm estimates, a priori error estimates.

AMS subject classifications. 35J61, 35R06, 49J20, 49M25, 65N15, 65N30.

1. Introduction. In this work we shall be interested in the analysis and discretization of a pointwise tracking optimal control problem for a semilinear elliptic partial differential equation (PDE). This PDE-constrained optimization problem entails the minimization of a cost functional that involves point evaluations of the state; control constrains are also considered. Let us make this discussion precise. Let $\Omega \subset \mathbb{R}^d$, with $d \in \{2, 3\}$, be an open, bounded, and convex polytope with boundary $\partial\Omega$ and $\mathcal{D}$ be a finite ordered subset of Ω with cardinality $\#\mathcal{D} < \infty$. Given a set of desired states $\{y_t\}_{t\in\mathcal{D}} \subset \mathbb{R}$, a regularization parameter $\alpha > 0$, and the cost functional

$$J(y, u) := \frac{1}{2}\sum_{t\in\mathcal{D}} (y(t) - y_t)^2 + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)}, \tag{1.1}$$

the problem under consideration reads as follows: Find $\min J(y, u)$ subject to the *monotone, semilinear, and elliptic PDE*

$$-\Delta y + a(\cdot, y) = u \text{ in } \Omega, \qquad y = 0 \text{ on } \partial\Omega, \tag{1.2}$$

and the *control constraints*

$$u \in \mathbb{U}_{ad}, \quad \mathbb{U}_{ad} := \{v \in L^2(\Omega) : \mathtt{a} \le v(x) \le \mathtt{b} \text{ a.e. } x \in \Omega\}. \tag{1.3}$$

The control bounds $\mathtt{a}, \mathtt{b} \in \mathbb{R}$ are such that $\mathtt{a} < \mathtt{b}$. Assumptions on the function a will be deferred until section 2.2.

The analysis of a priori error estimates for finite element approximations of distributed semilinear optimal control problems has previously been considered in a number of works. To the best of our knowledge, the work [5] appears to be the

*AA is partially supported by CONICYT through FONDECYT project 1170579. FF is supported by UTFSM through Beca de Mantención. EO is partially supported by CONICYT through FONDECYT Project 11180193.

†Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile. (alejandro.allendes@usm.cl).

‡Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile. (francisco.fuica@sansano.usm.cl).

§Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile. (enrique.otarola@usm.cl, http://eotarola.mat.utfsm.cl/).

first to provide error estimates for a such class of problems; control constraints are also considered. Within a general setting, the authors consider the cost functional $J(y,u) := \int_\Omega L(x,y,u)\mathrm{d}x$, where L satisfies the conditions stated in [5, assumption **A2**], and devise finite element techniques to solve the underlying optimal control problem. To be precise, the authors propose a fully discrete scheme on quasi–uniform meshes that utilize piecewise constant functions to approximate the control variable and piecewise linear functions to approximate the state and adjoint variables. Assuming that $\Omega \subset \mathbb{R}^d$, with $d \in \{2,3\}$, is a convex domain with a boundary $\partial\Omega$ of class $C^{1,1}$ and the mesh–size is sufficiently small, the authors derive a priori error estimates for the approximation of the optimal control variable in the $L^2(\Omega)$-norm [5, Theorem 5.1] and the $L^\infty(\Omega)$-norm [5, Theorem 5.2]; the one derived in the $L^2(\Omega)$-norm being optimal in terms of approximation. Since the publication of [5], several additional studies have enriched our understanding within such a scenario. We refer the reader to [14] for references and also for an up-to-date discussion including linear approximation of the optimal control, the so-called variational discretization approach, superconvergence and postprocessing step, and time dependent problems.

For the particular case $a \equiv 0$, there are several works available in the literature that provide a priori error estimates for finite element discretizations of (1.1)–(1.3). In two and three dimensions and utilizing that the associated adjoint variable belongs to $W_0^{1,r}(\Omega)$, for every $r < d/(d-1)$, the authors of [19] obtain a priori and a posteriori error estimates for the so-called variational discretization of (1.1)–(1.3); the state and adjoint equations are discretized with continuous piecewise linear finite elements. The following rates of convergence for the error approximation of the control variable are derived [19, Theorem 3.2]: $\mathcal{O}(h)$ in two dimensions and $\mathcal{O}(h^{1/2})$ in three dimensions. Later, the authors of [9] analyze a fully discrete scheme that approximates the optimal state, adjoint, and control variables with piecewise linear functions and obtain a $\mathcal{O}(h)$ rate of convergence for the error approximation of the control variable in two dimensions [9, Theorem 5.1]. The authors of [9] also analyze the variational discretization scheme and derive a priori error estimates for the error approximation of the control variable in [9, Theorem 5.2]. In [4], the authors invoke the theory of Muckenhoupt weights and weighted Sobolev spaces to provide error estimates for a numerical scheme that discretizes the control variable with piecewise constant functions; the state and adjoint equations are discretized with continuous piecewise linear finite elements. In two and three dimenions, the authors derive a priori error estimates for the error approximation of the optimal control variable; the one in two dimensions being nearly-optimal in terms of approximation [4, Theorem 4.3]. In three dimensions the estimate behaves as $\mathcal{O}(h^{1/2}|\log h|)$; it is suboptimal in terms of approximation. This has been recently improved in [7, Theorem 6.6]. We finally mention the works [23] and [6] for extensions of the aforementioned results to the Stokes system.

In contrast to the aforementioned advances and to the best of our knowledge, this exposition is the first one that studies approximation techniques for a pointwise tracking optimal control problem involving a semilinear elliptic PDE. In what follows, we list, what we believe are, the main contributions of our work:

- *Existence of an optimal control:* Assuming that $a = a(x,y)$ is a Carathéodory function that is monotone increasing and locally Lipschitz in y with $a(\cdot,0) \in L^2(\Omega)$, we show that our control problem admits at least a solution; see Theorem 3.1.
- *Optimality conditions:* We obtain first order optimality conditions in Theorem 3.3. Under additional assumptions on a, we derive second order necessary

and sufficient optimality conditions with a minimal gap; see section 3.3. Since the cost functional of our problem involves point evaluations of the state, we have that $\bar{p} \in W_0^{1,r}(\Omega) \notin H_0^1(\Omega) \cap C(\bar{\Omega})$, where $r < d/(d-1)$. This requires a suitable adaption of the arguments available in the literature [14, section 6], [32]. The arguments in [14, section 6] utilize that $\bar{p} \in W^{2,p}(\Omega)$ with $p > d$.

- *Convergence of discretization and error estimates:* We prove that the sequence $\{\bar{u}_h\}_{h>0}$ of global solutions of suitable discrete control problems converge to a solution of the continuous optimal control problem. We also derive a nearly-optimal local error estimate in maximum–norm for semilinear PDEs in Theorem 4.1, which is instrumental for proving that the error approximation of the control variable converges with rate $\mathcal{O}(h|\log h|)$, when measured in the L^2-norm. The analysis involves estimate in L^∞-norm and $W^{1,p}$-spaces, combined with having to deal with the variational inequality that characterizes the optimal control and suitable second order optimality conditions. This subtle intertwining of ideas is one of the highlights of this contribution.

The outline of this manuscript is as follows. In section 2 we introduce the notation and functional framework we shall work with and briefly review basic results for semilinear elliptic PDEs. In section 3 we analyze a weak version of the optimal control problem (1.1)–(1.3); we show existence of solutions and obtain first and second order optimality conditions. In section 4 we present a finite element discretization of (1.1)–(1.3) and review some results related to the discretization of the state and adjoint equations. In section 5 we derive a nearly-optimal estimate for the error approximation of the control variable. We conclude in section 6 by presenting a numerical example that confirms our theoretical results.

2. Notation and assumptions.

Let us set notation and describe the setting we shall operate with.

2.1. Notation.

Throughout this work $d \in \{2, 3\}$ and $\Omega \subset \mathbb{R}^d$ is an open, bounded, and convex polytopal domain. If $\mathscr{X}$ and $\mathscr{Y}$ are Banach function spaces, we write $\mathscr{X} \hookrightarrow \mathscr{Y}$ to denote that $\mathscr{X}$ is continuously embedded in $\mathscr{Y}$. We denote by $\|\cdot\|_{\mathscr{X}}$ the norm of $\mathscr{X}$. Given $r \in (1, \infty)$, we denote by r' its Hölder conjugate, i.e., the real number such that $1/r + 1/r' = 1$. The relation $\mathfrak{a} \lesssim \mathfrak{b}$ indicates that $\mathfrak{a} \leq C\mathfrak{b}$, with a positive constant that depends neither on $\mathfrak{a}$, $\mathfrak{b}$ nor on the discretization parameter. The value of C might change at each occurrence.

2.2. Assumptions.

We will consider the following assumptions on the nonlinear function a. We notice, however, that some of the results that we present in this work hold under less restrictive requirements. When possible we explicitly mention the assumptions on a that are needed to obtain a particular result.

(A.1) $a : \Omega \times \mathbb{R} \to \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second variable and $a(\cdot, 0) \in L^2(\Omega)$.

(A.2) $\frac{\partial a}{\partial y}(x, y) \geq 0$ for a.e. $x \in \Omega$ and for all $y \in \mathbb{R}$.

(A.3) For all $\mathscr{m} > 0$, there exists a positive constant $C_{\mathscr{m}}$ such that

$$\sum_{i=1}^{2} \left| \frac{\partial^i a}{\partial y^i}(x, y) \right| \leq C_{\mathscr{m}}, \qquad \left| \frac{\partial^2 a}{\partial y^2}(x, v) - \frac{\partial^2 a}{\partial y^2}(x, w) \right| \leq C_{\mathscr{m}} |v - w|$$

for a.e. $x \in \Omega$ and $y, v, w \in [-\mathscr{m}, \mathscr{m}]$.

2.3. State equation.

Here, we collect some facts on problem (1.2) that are well-known and will be used repeatedly. Given $f \in L^q(\Omega)$, with $q > d/2$, we introduce the following weak problem: Find $\mathsf{y} \in H_0^1(\Omega)$ such that

$$(2.1) \qquad (\nabla \mathsf{y}, \nabla v)_{L^2(\Omega)} + (a(\cdot, \mathsf{y}), v)_{L^2(\Omega)} = (f, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega).$$

We begin with the following result that states the well–posedness of problem (2.1) and further regularity properties for its solution y.

THEOREM 2.1 (well–posedness and regularity). *Let $f \in L^q(\Omega)$ with $q > d/2$. Let $a = a(x, y) : \Omega \times \mathbb{R} \to \mathbb{R}$ be a Carathéodory function that is monotone increasing and locally Lipschitz in y a.e. in Ω. If Ω denotes an open and bounded domain with Lipschitz boundary and $a(\cdot, 0) \in L^q(\Omega)$, with $q > d/2$, then problem (2.1) has a unique solution $\mathsf{y} \in H_0^1(\Omega) \cap L^\infty(\Omega)$. If, in addition, Ω is convex and $f, a(\cdot, 0) \in L^2(\Omega)$, then*

$$\|\mathsf{y}\|_{H^2(\Omega)} \lesssim \|f - a(\cdot, 0)\|_{L^2(\Omega)}.$$

The hidden constant is independent of a and f.

Proof. The existence of a unique solution $\mathsf{y} \in H_0^1(\Omega) \cap L^\infty(\Omega)$ follows from the main theorem on monotone operators [33, Theorem 26.A], [28, Theorem 2.18] combined with an argument due to Stampacchia [31], [25, Theorem B.2]. The $H^2(\Omega)$-regularity of y follows from the fact that $f, a(\cdot, 0) \in L^2(\Omega)$ and that Ω is convex; see [24, Theorems 3.2.1.2 and 4.3.1.4] when $d = 2$ and [24, Theorems 3.2.1.2] and [26, section 4.3.1] when $d = 3$. ∎

The following result is contained in [32, Theorem 4.16].

THEOREM 2.2. *Let $f_1, f_2 \in L^q(\Omega)$ with $q > d/2$. Let $a = a(x, y) : \Omega \times \mathbb{R} \to \mathbb{R}$ be a Carathéodory function of class C^1 with respect to y such that (A.2) holds. Assume that $|\partial a/\partial y(x, y)| \leq C_{\mathfrak{m}}$ for a.e. $x \in \Omega$ and $y \in [-\mathfrak{m}, \mathfrak{m}]$. If Ω denotes an open and bounded domain with Lipschitz boundary and $a(\cdot, 0) \in L^q(\Omega)$, with $q > d/2$, then*

$$(2.2) \qquad \|\nabla(\mathsf{y}_1 - \mathsf{y}_2)\|_{L^2(\Omega)} + \|\mathsf{y}_1 - \mathsf{y}_2\|_{L^\infty(\Omega)} \lesssim \|f_1 - f_2\|_{L^q(\Omega)},$$

where $i \in \{1, 2\}$ and y_i solves problem (2.1) with f replaced by f_i.

3. The pointwise tracking optimal control problem. In this section, we analyze the following weak version of the pointwise tracking optimal control problem (1.1)–(1.3): Find

$$(3.1) \qquad \min\{J(y, u) : (y, u) \in H_0^1(\Omega) \cap L^\infty(\Omega) \times \mathbb{U}_{ad}\}$$

subject to the *the state equation*

$$(3.2) \qquad (\nabla y, \nabla v)_{L^2(\Omega)} + (a(\cdot, y), v)_{L^2(\Omega)} = (u, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega).$$

Let $a = a(x, y) : \Omega \times \mathbb{R} \to \mathbb{R}$ be a Carathéodory function that is monotone increasing and locally Lipschitz in y with $a(\cdot, 0) \in L^2(\Omega)$. Since Ω is convex, Theorem 2.1 yields the existence of a unique solution $y \in H^2(\Omega) \cap H_0^1(\Omega)$ of problem (3.2). We immediately notice that, in view of the continuous embedding $H^2(\Omega) \hookrightarrow C(\bar{\Omega})$, point evaluations of y in (1.1) are well-defined.

3.1. Existence of optimal controls. As it is customary in optimal control theory, to analyze (3.1)–(3.2), we introduce the so-called control to state operator $\mathcal{S} : L^2(\Omega) \to H^2(\Omega) \cap H_0^1(\Omega)$ which, given a control u, associates to it the unique state y that solves (3.2). With this operator at hand, we define the reduced cost functional $j : L^2(\Omega) \to \mathbb{R}$ by $j(u) := J(\mathcal{S}u, u)$.

Since the optimal control problem (3.1)–(3.2) is not convex, we discuss existence results and optimality conditions in the context of local solutions. A control $\bar{u} \in \mathbb{U}_{ad}$ is said to be locally optimal in $L^2(\Omega)$ for (3.1)–(3.2) if there exists $\delta > 0$ such that $J(\bar{y}, \bar{u}) \leq J(y, u)$ for all $u \in \mathbb{U}_{ad}$ such that $\|u - \bar{u}\|_{L^2(\Omega)} \leq \delta$. Here, $\bar{y} = \mathcal{S}\bar{u}$ and $y = \mathcal{S}u$. Since the set $\mathbb{U}_{ad}$ is bounded in $L^\infty(\Omega)$, it can be proved that local optimality in $L^2(\Omega)$ is equivalent to local optimality in $L^q(\Omega)$ for $q \in (1, \infty)$; see [14, section 5] for details.

The existence of an optimal state-control pair $(\bar{y}, \bar{u})$ is as follows.

THEOREM 3.1 (existence of an optimal pair). *Let Ω be an open, bounded, and convex domain. Let $a = a(x, y) : \Omega \times \mathbb{R} \to \mathbb{R}$ be a Carathéodory function that is monotone increasing and locally Lipschitz in y with $a(\cdot, 0) \in L^2(\Omega)$. Thus, the optimal control problem (3.1)–(3.2) admits at least one solution $(\bar{y}, \bar{u}) \in H_0^1(\Omega) \cap H^2(\Omega) \times \mathbb{U}_{ad}$.*

Proof. Define $\Phi : L^2(\Omega) \to \mathbb{R}$ and $\Psi : H^2(\Omega) \cap H_0^1(\Omega) \to \mathbb{R}$ by

$$\Phi(v) := \alpha \|v\|_{L^2(\Omega)}^2, \qquad \Psi(y) := \sum_{t \in \mathcal{D}} |y(t) - y_t|^2.$$

It is immediate that Φ is continuous and convex in $L^2(\Omega)$. It is thus weakly lower semicontinuous in $L^2(\Omega)$. On the other hand, Ψ is continuous as a map from $H_0^1(\Omega) \cap H^2(\Omega)$ to $\mathbb{R}$. The fact that $\mathbb{U}_{ad}$ is weakly sequentially compact allows us to conclude; see [32, Theorem 4.15] for details. ∎

3.2. First order necessary optimality conditions. In this section, we formulate first order necessary optimality conditions. To accomplish this task, we begin by analyzing differentiability properties of the control to state operator $\mathcal{S}$.

THEOREM 3.2 (differentiability properties of $\mathcal{S}$). *Assume that (A.1), (A.2), and (A.3) hold. Then, the control to state map $\mathcal{S} : L^2(\Omega) \to H^2(\Omega) \cap H_0^1(\Omega)$ is of class C^2. In addition, if $u, v \in L^2(\Omega)$, then $z = \mathcal{S}'(u)v \in H^2(\Omega) \cap H_0^1(\Omega)$ corresponds to the unique solution to*

$$(\nabla z, \nabla w)_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, y) z, w\right)_{L^2(\Omega)} = (v, w)_{L^2(\Omega)} \quad \forall w \in H_0^1(\Omega), \tag{3.3}$$

where $y = \mathcal{S}u$. If $v_1, v_2 \in L^2(\Omega)$, then $\mathfrak{z} = \mathcal{S}''(u)(v_1, v_2) \in H^2(\Omega) \cap H_0^1(\Omega)$ is the unique solution to

$$(\nabla \mathfrak{z}, \nabla w)_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, y)\mathfrak{z}, w\right)_{L^2(\Omega)} = -\left(\tfrac{\partial^2 a}{\partial y^2}(\cdot, y) z_{v_1} z_{v_2}, w\right)_{L^2(\Omega)} \tag{3.4}$$

for all $w \in H_0^1(\Omega)$, where $z_{v_i} = \mathcal{S}'(u)v_i$, with $i = 1, 2$, and $y = \mathcal{S}u$.

Proof. The first order Fréchet differentiability of $\mathcal{S}$ from $L^2(\Omega)$ into $H^2(\Omega) \cap H_0^1(\Omega)$ follows from a slight modification of the arguments of [32, Theorem 4.17] that basically entails replacing $H^1(\Omega) \cap C(\bar{\Omega})$ by $H^2(\Omega) \cap H_0^1(\Omega)$ and $L^r(\Omega)$ by $L^2(\Omega)$. [32, Theorem 4.17] also yields that $z = \mathcal{S}'(u)v \in H^2(\Omega) \cap H_0^1(\Omega)$ corresponds to the unique solution to (3.3). The second order Fréchet differentiability of $\mathcal{S}$ can be obtained by using the implicit function theorem; see, for instance, the proof of [32, Theorem 4.24] and [14, Proposition 16] for details. ∎

We begin the analysis of optimality conditions with a classical result. If $\bar{u} \in \mathbb{U}_{ad}$ denotes a locally optimal control for problem (3.1)–(3.2), then we have the variational inequality [32, Lemma 4.18]

$$j'(\bar{u})(u - \bar{u}) \geq 0 \quad \forall u \in \mathbb{U}_{ad}. \tag{3.5}$$

We recall that, for $u \in \mathbb{U}_{ad}$, the reduced cost functional is defined as $j(u) = J(\mathcal{S}u, u)$. In (3.5), $j'(\bar{u})$ denotes the Gateâux derivative of j at $\bar{u}$. To explore the variational inequality (3.5), we introduce the adjoint variable $p \in W_0^{1,r}(\Omega)$, with $r \in (1, d/(d-1))$, as the unique solution to the *adjoint equation*

$$(3.6)\quad (\nabla w, \nabla p)_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, y)p, w\right)_{L^2(\Omega)} = \sum_{t\in\mathcal{D}} \langle (y(t) - y_t)\delta_t, w\rangle \quad \forall w \in W_0^{1,r'}(\Omega).$$

Here, $r' > d$ denotes the Hölder conjugate of r and $y = \mathcal{S}u$ corresponds to the solution to (3.2). We immediately notice that, in view of assumptions (A.1)–(A.3), problem (3.6) is well–posed; see [10, Theorem 1].

We are now in position to present first order necessary optimality conditions for our PDE–constrained optimization problem.

THEOREM 3.3 (first order necessary optimality conditions). *Assume that (A.1), (A.2), and (A.3) hold. Then every locally optimal control $\bar{u} \in \mathbb{U}_{ad}$ for problem (3.1)–(3.2) satisfies the variational inequality*

$$(3.7)\quad (\bar{p} + \alpha\bar{u}, u - \bar{u})_{L^2(\Omega)} \geq 0 \quad \forall u \in \mathbb{U}_{ad},$$

where $\bar{p} \in W_0^{1,r}(\Omega)$, with $r < d/(d-1)$, denotes the unique solution to problem (3.6) with y replaced by $\bar{y} = \mathcal{S}\bar{u}$.

Proof. A simple computation reveals that the first order optimality condition (3.5) can be written as follows:

$$(3.8)\quad \sum_{t\in\mathcal{D}} (\mathcal{S}\bar{u}(t) - y_t)\cdot \mathcal{S}'(\bar{u})(u-\bar{u})(t) + \alpha(\bar{u}, u-\bar{u})_{L^2(\Omega)} \geq 0 \qquad \forall u \in \mathbb{U}_{ad}.$$

Let us concentrate on the first term of the left hand side of (3.8). To accomplish this task, we begin by defining $z := \mathcal{S}'(\bar{u})(u - \bar{u})$. Since $u, \bar{u} \in L^2(\Omega)$, the results of Theorem 3.2 guarantees that $z \in H^2(\Omega) \cap H_0^1(\Omega) \hookrightarrow W_0^{1,\mathfrak{q}}(\Omega)$ for every $\mathfrak{q} < 2d/(d-2)$ [1, Theorem 4.12]. In particular, since $2d/(d-2) > d$, we have that $z \in W_0^{1,\mathfrak{q}}(\Omega)$ for every $\mathfrak{q} \in (d, 2d/(d-2))$. We are thus able to set $w = z$ as a test function in the adjoint problem (3.6) to obtain

$$(3.9)\quad (\nabla z, \nabla \bar{p})_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, \bar{y})\bar{p}, z\right)_{L^2(\Omega)} = \sum_{t\in\mathcal{D}} (\bar{y}(t) - y_t)\cdot z(t).$$

On the other hand, we would like to set $w = \bar{p}$ in the problem that $z = \mathcal{S}'(\bar{u})(u-\bar{u})$ solves. If that were possible, we would obtain

$$(3.10)\quad (\nabla z, \nabla \bar{p})_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, \bar{y})z, \bar{p}\right)_{L^2(\Omega)} = (u - \bar{u}, \bar{p})_{L^2(\Omega)}.$$

However, since $\bar{p} \in W_0^{1,r}(\Omega)$ with $r < d/(d-1)$, we have that $\bar{p} \notin H_0^1(\Omega)$ so that (3.10) must be justified by different means. Let $\{p_n\}_{n\in\mathbb{N}} \subset C_0^\infty(\Omega)$ be such that $p_n \to \bar{p}$ in $W_0^{1,r}(\Omega)$ for every $r < d/(d-1)$. Setting, $w = p_n$, with $n \in \mathbb{N}$, in the problem that $z = \mathcal{S}'(\bar{u})(u - \bar{u})$ solves yields

$$(\nabla z, \nabla p_n)_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, \bar{y})z, p_n\right)_{L^2(\Omega)} = (u - \bar{u}, p_n)_{L^2(\Omega)}.$$

The right hand side of this expression converges to $(u - \bar{u}, \bar{p})_{L^2(\Omega)}$. In fact,

$$|(u-\bar{u}, \bar{p})_{L^2(\Omega)} - (u - \bar{u}, p_n)_{L^2(\Omega)}| \leq \|u - \bar{u}\|_{L^\infty(\Omega)} \|\bar{p} - p_n\|_{L^1(\Omega)} \to 0, \quad n \uparrow \infty.$$

Since there is $m > 0$ such that $|\bar{y}(x)| \leq m$ for a.e. $x \in \Omega$, (A.3) reveals that

$$\left| \left(\frac{\partial a}{\partial y}(\cdot, \bar{y}) z, \bar{p} \right)_{L^2(\Omega)} - \left(\frac{\partial a}{\partial y}(\cdot, \bar{y}) z, p_n \right)_{L^2(\Omega)} \right| \leq C_m \|z\|_{L^\infty(\Omega)} \|\bar{p} - p_n\|_{L^1(\Omega)} \to 0$$

as $n \uparrow \infty$; $\|z\|_{L^\infty(\Omega)}$ is uniformly bounded because $z \in H^2(\Omega) \cap H_0^1(\Omega)$. Finally,

$$|(\nabla z, \nabla(\bar{p} - p_n))_{L^2(\Omega)}| \leq \|\nabla z\|_{L^{r'}(\Omega)} \|\nabla(\bar{p} - p_n)\|_{L^r(\Omega)} \to 0, \quad n \uparrow \infty,$$

for every $r < d/(d-1)$.

The desired variational inequality (3.7) follows from (3.8), (3.9), and (3.10). ∎

We present the following projection formula for $\bar{u}$. The local optimal control $\bar{u}$ satisfies (3.7) if and only if [32, section 4.6]

$$\bar{u}(x) := \Pi_{[\mathtt{a},\mathtt{b}]}(-\alpha^{-1}\bar{p}(x)) \text{ a.e. } x \in \Omega, \tag{3.11}$$

where $\Pi_{[\mathtt{a},\mathtt{b}]} : L^1(\Omega) \to \mathbb{U}_{ad}$ is defined by $\Pi_{[\mathtt{a},\mathtt{b}]}(v) := \min\{\mathtt{b}, \max\{v, \mathtt{a}\}\}$ a.e. in Ω. We can thus immediately conclude that $\bar{u} \in W^{1,r}(\Omega)$ for every $r < d/(d-1)$.

We now present the following regularity result, which will be of importance to derive the error estimate of Theorem 5.1.

THEOREM 3.4 (extra regularity of $\bar{u}$). *Suppose that assumptions (A.1), (A.2), and (A.3) hold. Then, every locally optimal control $\bar{u} \in H^1(\Omega) \cap C^{0,1}(\bar{\Omega})$.*

Proof. The proof relies on the projection formula (3.11) and on the local regularity of the locally optimal adjoint state $\bar{p}$. For a detailed proof we refer the reader to [16, Lemma 3.3] and [12, Theorem 3.4]; see also [18, Theorem 4.2]. ∎

3.3. Second order sufficient optimality condition. In this section, we derive second order optimality conditions. To be precise, we formulate second order necessary optimality conditions in Theorem 3.6 and derive, in Theorem 3.7, sufficient optimality conditions with a minimal gap with respect to the necessary ones derived in Theorem 3.6.

We begin our analysis with the following result.

THEOREM 3.5 (j is of class C^2 and j'' is locally Lipschitz). *Assume that (A.1), (A.2), and (A.3) hold. Then the reduced cost functional $j : L^2(\Omega) \to \mathbb{R}$ is of class C^2. Moreover, for every $u, v_1, v_2 \in L^2(\Omega)$, we have*

$$j''(u)(v_1, v_2) = \alpha(v_1, v_2)_{L^2(\Omega)} - \left(\frac{\partial^2 a}{\partial y^2}(\cdot, y) z_{v_1} z_{v_2}, p \right)_{L^2(\Omega)} + \sum_{t \in \mathcal{D}} z_{v_1}(t) z_{v_2}(t), \tag{3.12}$$

where p solves (3.6) and $z_{v_i} = \mathcal{S}'(u) v_i$, with $i \in \{1, 2\}$. In addition, if $v, u_1, u_2 \in L^2(\Omega)$ and there exists $m > 0$ is such that $\max\{\|u_1\|_{L^2(\Omega)}, \|u_2\|_{L^2(\Omega)}\} \leq m$, then there exists $C_m > 0$ such that

$$|j''(u_1)v^2 - j''(u_2)v^2| \leq C_m \|u_1 - u_2\|_{L^2(\Omega)} \|v\|^2_{L^2(\Omega)}. \tag{3.13}$$

Proof. The fact that j is of class C^2 is an immediate consequence of the differentiability properties of the control to state map $\mathcal{S}$ given in Theorem 3.2. It thus suffices to derive (3.12) and (3.13). To accomplish this task, we begin with a basic computation, which reveals that, for every $u, v_1, v_2 \in L^2(\Omega)$, we have

$$j''(u)(v_1, v_2) = \alpha(v_1, v_2)_{L^2(\Omega)} + \sum_{t \in \mathcal{D}} \left[(\mathfrak{z}(t) \cdot (\mathcal{S}u(t) - y_t) + z_{v_1}(t) z_{v_2}(t) \right], \tag{3.14}$$

where $\mathfrak{z}, z_{v_1}, z_{v_2} \in H^2(\Omega) \cap H_0^1(\Omega)$ are as in the statement of Theorem 3.2. Set $w = \mathfrak{z}$ in (3.6) and invoke a similar approximation argument to that used in the proof of Theorem 3.3, that essentially allows us to set $w = p$ in (3.4), to obtain

$$\sum_{t\in\mathcal{D}} \mathfrak{z}(t) \cdot (\mathcal{S}u(t) - y_t) = -\left(\tfrac{\partial^2 a}{\partial y^2}(\cdot, y) z_{v_1} z_{v_2}, p\right)_{L^2(\Omega)}.$$

Replacing the previous identity into (3.14) yields (3.12).

Let $u_1, u_2, v \in L^2(\Omega)$ and $\mathscr{m} > 0$ be such that $\max\{\|u_1\|_{L^\infty(\Omega)}, \|u_2\|_{L^\infty(\Omega)}\} \leq \mathscr{m}$. Define $\chi = \mathcal{S}'(u_1)v$ and $\psi = \mathcal{S}'(u_2)v$. Notice that χ and ψ correspond to the unique solutions to (3.3) with $y = y_{u_1} := \mathcal{S}u_1$ and $y = y_{u_2} := \mathcal{S}u_2$, respectively. In view of the identity (3.12) we obtain

$$\begin{aligned} j''(u_1)v^2 - j''(u_2)v^2 &= \left(\tfrac{\partial^2 a}{\partial y^2}(\cdot, y_{u_2})\psi^2, p_{u_2}\right)_{L^2(\Omega)} - \left(\tfrac{\partial^2 a}{\partial y^2}(\cdot, y_{u_1})\chi^2, p_{u_1}\right)_{L^2(\Omega)} \\ &\quad + \sum_{t\in\mathcal{D}} (\chi^2(t) - \psi^2(t)) =: \mathbf{I} + \sum_{t\in\mathcal{D}} \mathbf{II}_t. \end{aligned} \tag{3.15}$$

Here, $i = \{1, 2\}$ and $p_{u_i} \in W_0^{1,r}(\Omega)$, with $r < d/(d-1)$, denotes the unique solution to the adjoint equation (3.6) with y replaced by y_{u_i}. In what follows we estimate $\mathbf{I}$ and $\mathbf{II}_t$ for every $t \in \mathcal{D}$. To estimate $\mathbf{I}$, we first rewrite it as follows:

$$\begin{aligned} \mathbf{I} = \left(\left[\tfrac{\partial^2 a}{\partial y^2}(\cdot, y_{u_2}) - \tfrac{\partial^2 a}{\partial y^2}(\cdot, y_{u_1})\right]\psi^2, p_{u_2}\right)_{L^2(\Omega)} + \left(\tfrac{\partial^2 a}{\partial y^2}(\cdot, y_{u_1})\psi^2, p_{u_2} - p_{u_1}\right)_{L^2(\Omega)} \\ + \left(\tfrac{\partial^2 a}{\partial y^2}(\cdot, y_{u_1})[\psi^2 - \chi^2], p_{u_1}\right)_{L^2(\Omega)} =: \mathbf{I}_1 + \mathbf{I}_2 + \mathbf{I}_3. \end{aligned}$$

Invoke (A.3), a generalized Hölder inequality, the Sobolev embedding $H_0^1(\Omega) \hookrightarrow L^4(\Omega)$, the well–posedness of problem (3.3), and the Lipschitz property (2.2), to obtain

$$\mathbf{I}_1 \lesssim \|y_{u_1} - y_{u_2}\|_{L^\infty(\Omega)} \|\nabla\psi\|^2_{L^2(\Omega)} \|p_{u_2}\|_{L^2(\Omega)} \lesssim \|u_1 - u_2\|_{L^2(\Omega)} \|v\|^2_{L^2(\Omega)}, \tag{3.16}$$

where we have also used the stability estimate

$$\|p_{u_2}\|_{L^2(\Omega)} \lesssim \|\nabla p_{u_2}\|_{L^r(\Omega)} \lesssim \|y_{u_2}\|_{L^\infty(\Omega)} + \sum_{t\in\mathcal{D}} |y_t| \lesssim \mathscr{M} + \sum_{t\in\mathcal{D}} |y_t|. \tag{3.17}$$

Notice that Theorem 2.1 and the assumption on u_2 yields $\|y_{u_2}\|_{L^\infty(\Omega)} \leq C\|u_2\|_{L^2(\Omega)} \leq C\mathscr{m}$, where $C > 0$. To guarantee that $p_{u_2} \in L^2(\Omega)$ and the first estimate in (3.17) we further restrict the exponent r to belong to $[2d/(d+2), d/(d-1))$ [1, Theorem 4.12]. To control $\mathbf{I}_2$, we invoke similar arguments to the ones that lead to (3.16). We obtain

$$\begin{aligned} \mathbf{I}_2 \leq C_{\mathscr{m}} \|\psi\|^2_{L^4(\Omega)} \|p_{u_1} - p_{u_2}\|_{L^2(\Omega)} &\lesssim \|\nabla\psi\|^2_{L^2(\Omega)} \|\nabla(p_{u_1} - p_{u_2})\|_{L^r(\Omega)} \\ &\lesssim \|v\|^2_{L^2(\Omega)} \|y_{u_1} - y_{u_2}\|_{L^\infty(\Omega)} \lesssim \|v\|^2_{L^2(\Omega)} \|u_1 - u_2\|_{L^2(\Omega)}. \end{aligned}$$

Finally, to estimate $\mathbf{I}_3$, we notice that $\psi - \chi \in H_0^1(\Omega) \cap L^\infty(\Omega)$ solves

$$(\nabla(\psi - \chi), \nabla w) + \left(\tfrac{\partial a}{\partial y}(\cdot, y_{u_2})(\psi - \chi), w\right)_{L^2(\Omega)} = \left(\left[\tfrac{\partial a}{\partial y}(\cdot, y_{u_1}) - \tfrac{\partial a}{\partial y}(\cdot, y_{u_2})\right]\chi, w\right)_{L^2(\Omega)}$$

for all $w \in H_0^1(\Omega)$. The stability estimate

$$\|\psi - \chi\|_{L^\infty(\Omega)} \lesssim \left\|\left[\tfrac{\partial a}{\partial y}(\cdot, y_{u_1}) - \tfrac{\partial a}{\partial y}(\cdot, y_{u_2})\right]\chi\right\|_{L^2(\Omega)},$$

combined with (A.3) and the Lipschitz property (2.2), allows us to conclude that

$$\|\psi - \chi\|_{L^\infty(\Omega)} \lesssim \|y_{u_1} - y_{u_2}\|_{L^\infty(\Omega)}\|\chi\|_{L^2(\Omega)} \lesssim \|u_1 - u_2\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)}. \tag{3.18}$$

Therefore, utilizing (A.3), the well-posedness of problem (3.3) and (3.18) we obtain

$$\mathbf{I}_3 \lesssim \|p_{u_1}\|_{L^2(\Omega)}\|\psi - \chi\|_{L^\infty(\Omega)}(\|\psi\|_{L^2(\Omega)} + \|\chi\|_{L^2(\Omega)}) \lesssim \|u_1 - u_2\|_{L^2(\Omega)}\|v\|^2_{L^2(\Omega)},$$

where we have also used the stability estimate $\|\psi\|_{L^2(\Omega)} + \|\chi\|_{L^2(\Omega)} \lesssim \|v\|_{L^2(\Omega)}$ and an estimate for $\|p_{u_1}\|_{L^2(\Omega)}$ which is similar to the one derived in (3.17).

The collection of the previous estimates allows us to arrive at

$$\mathbf{I} = \mathbf{I}_1 + \mathbf{I}_2 + \mathbf{I}_3 \lesssim \|u_1 - u_2\|_{L^2(\Omega)}\|v\|^2_{L^2(\Omega)}. \tag{3.19}$$

Let $t \in \mathcal{D}$. We now estimate $\mathbf{II}_t$ in (3.15). Combining the estimate (3.18) with an stability estimate for (3.3), it immediately follows that

$$\mathbf{II}_t \lesssim \|\psi - \chi\|_{L^\infty(\Omega)}(\|\psi\|_{L^\infty(\Omega)} + \|\chi\|_{L^\infty(\Omega)}) \lesssim \|u_1 - u_2\|_{L^2(\Omega)}\|v\|^2_{L^2(\Omega)}. \tag{3.20}$$

We conclude the desired estimate (3.13) by replacing estimates (3.19) and (3.20) into (3.15). This concludes the proof. ∎

Let $\bar{u} \in \mathbb{U}_{ad}$ satisfy the first order optimality conditions (3.2), (3.6), and (3.7). Define $\bar{\mathfrak{p}} := \bar{p} + \alpha\bar{u}$. The variational inequality (3.7) immediately yields

$$\bar{\mathfrak{p}}(x) \begin{cases} = 0 & \text{a.e. } x \in \Omega \text{ if } \mathtt{a} < \bar{u} < \mathtt{b}, \\ \geq 0 & \text{a.e. } x \in \Omega \text{ if } \bar{u} = \mathtt{a}, \\ \leq 0 & \text{a.e. } x \in \Omega \text{ if } \bar{u} = \mathtt{b}. \end{cases} \tag{3.21}$$

To formulate second order optimality conditions we introduce the *cone of critical directions*

$$C_{\bar{u}} := \{v \in L^2(\Omega) \text{ satisfying (3.23) and } v(x) = 0 \text{ if } \bar{\mathfrak{p}}(x) \neq 0\}, \tag{3.22}$$

where condition (3.23) reads as follows:

$$v(x) \begin{cases} \geq 0 \text{ a.e. } x \in \Omega \text{ if } \bar{u}(x) = \mathtt{a}, \\ \leq 0 \text{ a.e. } x \in \Omega \text{ if } \bar{u}(x) = \mathtt{b}. \end{cases} \tag{3.23}$$

From now on, we will restrict the exponent r to belong to $[2d/(d+2), d/(d-1))$ so that p, the solution to (3.6), belongs to $L^2(\Omega)$ [1, Theorem 4.12]. This immediately implies that $\bar{\mathfrak{p}} \in L^2(\Omega)$.

We are now in position to present second order necessary and sufficient optimality conditions. While it is fair to say that for distributed and semilinear optimal control problems such a theory is well-understood, our main source of difficulty here is that the solution to the adjoint problem does not belong to $H_0^1(\Omega) \cap C(\bar{\Omega})$: $\bar{p} \in W_0^{1,r}(\Omega) \setminus H_0^1(\Omega)$ with $r \in [2d/(d+2), d/(d-1))$.

THEOREM 3.6 (second order necessary optimality conditions). *If $\bar{u} \in \mathbb{U}_{ad}$ denotes a locally optimal control for problem (3.1)–(3.2), then*

$$j''(\bar{u})v \geq 0 \quad \forall v \in C_{\bar{u}}.$$

Proof. Let $v \in C_{\bar{u}}$. Define, for every $k \in \mathbb{N}$, the function

$$v_k(x) := \begin{cases} 0 & \text{if} \quad x : \mathtt{a} < \bar{u}(x) < \mathtt{a} + \frac{1}{k}, \quad \mathtt{b} - \frac{1}{k} < \bar{u}(x) < \mathtt{b}, \\ \Pi_{[-k,k]}(v(x)) & \text{otherwise.} \end{cases}$$

Notice that, since $v \in C_{\bar{u}}$, it immediately follows that $v_k \in C_{\bar{u}}$. In fact, a.e. $x \in \Omega$, we have $v(x) = 0 \implies v_k(x) = 0$, $v(x) \geq 0 \implies v_k(x) \geq 0$, and $v(x) \leq 0 \implies v_k(x) \leq 0$. In addition, $|v_k(x)| \leq |v(x)|$ and $v_k(x) \to v(x)$ as $k \uparrow \infty$ for a.e. $x \in \Omega$. Consequently, $v_k \to v$ in $L^2(\Omega)$ as $k \uparrow \infty$. On the other hand, simple computations reveal that for every $0 < \rho \leq k^{-2}$, we have $\bar{u} + \rho v_k \in \mathbb{U}_{ad}$. We can thus invoke the fact that $\bar{u}$ is a local minimum to conclude that $j(\bar{u}) \leq j(\bar{u} + \rho v_k)$ for ρ small enough. We now apply Taylor's theorem for j at $\bar{u}$ and utilize that $j'(\bar{u})v_k = 0$, which follows from the fact that $v_k \in C_{\bar{u}}$, to conclude that, for ρ sufficiently small, we have

$$0 \leq j(\bar{u} + \rho v_k) - j(\bar{u}) = \rho j'(\bar{u})v_k + \tfrac{\rho^2}{2} j''(\bar{u} + \rho\theta_k v_k)v_k^2 = \tfrac{\rho^2}{2} j''(\bar{u} + \rho\theta_k v_k)v_k^2,$$

with $\theta_k \in (0,1)$. Divide by ρ^2 and let $\rho \downarrow 0$ to arrive at $j''(\bar{u})v_k^2 \geq 0$. Let now $k \uparrow \infty$ and recall that $v_k \to v$ in $L^2(\Omega)$ to conclude, in view of (3.12), that $j''(\bar{u})v^2 \geq 0$. This concludes the proof. ∎

We now derive a sufficient condition with a minimal gap with respect to the necessary one obtained in Theorem 3.6.

THEOREM 3.7 (second order sufficient optimality conditions). *Let $(\bar{y}, \bar{p}, \bar{u})$ be a local minimum of (3.1)–(3.2) satisfying the first order optimality conditions (3.2), (3.6), and (3.7). If $j''(\bar{u})v^2 > 0$ for all $v \in C_{\bar{u}} \setminus \{0\}$, then there exist $\mu > 0$ and $\sigma > 0$ such that*

$$j(u) \geq j(\bar{u}) + \tfrac{\mu}{2}\|u - \bar{u}\|^2_{L^2(\Omega)} \quad \forall u \in \mathbb{U}_{ad} : \|u - \bar{u}\|_{L^2(\Omega)} \leq \sigma. \tag{3.24}$$

In particular, $\bar{u}$ is a locally optimal control in the sense of $L^2(\Omega)$.

Proof. We will proceed by contradiction. Assume that (3.24) does not hold. Hence, for any $k \in \mathbb{N}$ we are able to find an element $u_k \in \mathbb{U}_{ad}$ such that

$$\|\bar{u} - u_k\|_{L^2(\Omega)} < \tfrac{1}{k}, \qquad j(u_k) < j(\bar{u}) + \tfrac{1}{2k}\|\bar{u} - u_k\|^2_{L^2(\Omega)}. \tag{3.25}$$

Define

$$\rho_k := \|u_k - \bar{u}\|_{L^2(\Omega)}, \qquad v_k := \rho_k^{-1}(u_k - \bar{u}). \tag{3.26}$$

Taking a subsequence if necessary we can assume that $v_k \rightharpoonup v$ in $L^2(\Omega)$. In what follows we will prove, first, that the limit $v \in C_{\bar{u}}$ and thus that $v = 0$.

Since the set of elements satisfying condition (3.23) is closed and convex in $L^2(\Omega)$, it is weakly closed. Consequently, v satisfies (3.23). To verify the remaining condition in (3.22), we invoke the mean value theorem, (3.26), and (3.25) to arrive at

$$j'(\tilde{u}_k)v_k = \tfrac{1}{\rho_k}(j(u_k) - j(\bar{u})) < \tfrac{\rho_k}{2k} \to 0, \quad k \uparrow \infty, \tag{3.27}$$

where $\tilde{u}_k = \bar{u} + \theta_k(u_k - \bar{u})$ and $\theta_k \in (0,1)$. Define $\tilde{y}_k := \mathcal{S}\tilde{u}_k$ and $\tilde{p}_k$ as the unique solution to (3.6) with $y = \tilde{y}_k$. Since $\tilde{u}_k \to \bar{u}$ in $L^2(\Omega)$ as $k \uparrow \infty$, an application of Theorem 2.2 yields $\tilde{y}_k \to \bar{y}$ in $H_0^1(\Omega) \cap C(\bar{\Omega})$ as $k \uparrow \infty$. This, in view of [10, Theorem 1], implies that $\tilde{p}_k \to \bar{p}$ in $W_0^{1,r}(\Omega)$, for every $r < d/(d-1)$, as $k \uparrow \infty$. In particular, we have $\tilde{p}_k \to \bar{p}$ in $L^2(\Omega)$ as $k \uparrow \infty$. Consequently, since $\tilde{\mathfrak{p}}_k := \tilde{p}_k + \alpha \tilde{u}_k \to \bar{\mathfrak{p}} = \bar{p} + \alpha \bar{u}$ and $v_k \rightharpoonup v$ in $L^2(\Omega)$, as $k \uparrow \infty$, we invoke (3.27) to obtain

$$j'(\bar{u})v = \int_\Omega \bar{\mathfrak{p}}(x)v(x)\mathrm{d}x = \lim_{k\uparrow\infty} \int_\Omega \tilde{\mathfrak{p}}_k(x)v_k(x)\mathrm{d}x = \lim_{k\uparrow\infty} j'(\tilde{u}_k)v_k \leq 0.$$

On the other hand, in view of (3.7) we obtain $\int_\Omega \bar{\mathfrak{p}}(x)v_k(x) = \rho_k^{-1}\int_\Omega \bar{\mathfrak{p}}(x)(u_k(x) - \bar{u}(x))\mathrm{d}x \geq 0$. This implies $\int_\Omega \bar{\mathfrak{p}}(x)v(x)\mathrm{d}x \geq 0$. Consequently, $\int_\Omega \bar{\mathfrak{p}}(x)v(x)\mathrm{d}x = 0$. Since v satisfies the sign condition (3.23), the previous inequalities and (3.21) allow us to conclude that $\int_\Omega |\mathfrak{p}(x)v(x)|\mathrm{d}x = \int_\Omega \mathfrak{p}(x)v(x)\mathrm{d}x = 0$. This proves that, a.e. in Ω, $\bar{\mathfrak{p}}(x) \neq 0$ implies that $v(x) = 0$. We can thus conclude that $v \in C_{\bar{u}}$.

We now prove that $v = 0$. To accomplish this task, we invoke Taylor's theorem, the inequality in (3.25), and $j'(\bar{u})(u_k - \bar{u}) \geq 0$, for every $k \in \mathbb{N}$, to arrive at

$$\tfrac{\rho_k^2}{2} j''(\hat{u}_k)v_k^2 = j(u_k) - j(\bar{u}) - j'(\bar{u})(u_k - \bar{u}) \leq j(u_k) - j(\bar{u}) < \tfrac{\rho_k^2}{2k},$$

where, for $k \in \mathbb{N}$, $\hat{u}_k = \bar{u} + \theta_k(u_k - \bar{u})$ with $\theta_k \in (0,1)$. Thus, $\lim_k j''(\hat{u}_k)v_k^2 \leq 0$. We now prove that $j''(\bar{u})v^2 \leq \liminf_k j''(\hat{u}_k)v_k^2$. Let $\hat{z}_{v_k}$ and z_v be the solutions to (3.3) with forcing terms v_k and v, respectively. Invoke (3.12) and write

$$j''(\hat{u}_k)v_k^2 = \alpha\|v_k\|^2_{L^2(\Omega)} - \left(\tfrac{\partial^2 a}{\partial y^2}(\cdot,\hat{y}_k)\hat{z}^2_{v_k}, \hat{p}_k\right)_{L^2(\Omega)} + \sum_{t\in\mathcal{D}} \hat{z}^2_{v_k}(t).$$

Observe that, $\sum_{t\in\mathcal{D}} \hat{z}^2_{v_k}(t) \to \sum_{t\in\mathcal{D}} z^2_v(t)$. This is a consequence of the fact that $v_k \rightharpoonup v$ in $L^2(\Omega)$ implies that $\hat{z}_{v_k} \rightharpoonup z_v$ in $H^2(\Omega) \cap H^1_0(\Omega)$ as $k \uparrow \infty$ and the compact embedding $H^2(\Omega) \hookrightarrow C(\bar{\Omega})$. In addition, we have

$$(3.28)\quad \left|\int_\Omega \left(\tfrac{\partial^2 a}{\partial y^2}(x,\bar{y})z_v^2\bar{p} - \tfrac{\partial^2 a}{\partial y^2}(x,\hat{y}_k)\hat{z}^2_{v_k}\hat{p}_k\right)\mathrm{d}x\right| \leq C_m\|z_v\|^2_{L^\infty(\Omega)}\|\bar{p} - \hat{p}_k\|^2_{L^1(\Omega)}$$
$$+ C_m\|\hat{p}_k\|_{L^1(\Omega)}\left(\|\bar{y} - \hat{y}_k\|_{L^\infty(\Omega)}\|z_v\|^2_{L^\infty(\Omega)} + \|z_v + \hat{z}_{v_k}\|_{L^\infty(\Omega)}\|z_v - \hat{z}_{v_k}\|_{L^\infty(\Omega)}\right) \to 0$$

as $k \uparrow \infty$. To obtain (3.28), we used (A.3), $\tilde{p}_k \to \bar{p}$ in $W^{1,r}_0(\Omega)$, for every $r < d/(d-1)$, and $\tilde{y}_k \to \bar{y}$ in $H^1_0(\Omega) \cap C(\bar{\Omega})$ as $k \uparrow \infty$. Since the square of $\|v\|_{L^2(\Omega)}$ is continuous and convex, it is thus weakly lower semicontinuous in $L^2(\Omega)$. We have thus proved that $j''(\bar{u})v^2 \leq \liminf_k j''(\hat{u}_k)v_k^2$.

Therefore, since $j''(\bar{u})v^2 \leq \lim_k j''(\hat{u}_k)v_k^2 \leq 0$ and $v \in C_{\bar{u}}$, the second order optimality condition $j''(\bar{u})v^2 > 0$ for all $v \in C_{\bar{u}} \setminus \{0\}$ immediately yields $v = 0$.

Finally, since $v = 0$ we have $\hat{z}_{v_k} \rightharpoonup 0$ in $H^2(\Omega) \cap H^1_0(\Omega)$ as $k \uparrow \infty$. Consequently, from the identity

$$\alpha = \alpha\|v_k\|^2_{L^2(\Omega)} = j''(\hat{u}_k)v_k^2 + \left(\tfrac{\partial^2 a}{\partial y^2}(\cdot,\hat{y}_k)\hat{z}^2_{v_k}, \hat{p}_k\right)_{L^2(\Omega)} - \sum_{t\in\mathcal{D}} \hat{z}^2_{v_k}(t),$$

and the fact that $\liminf_k j''(\hat{u}_k)v_k^2 \leq 0$, we obtain that $\alpha \leq 0$, which is a contradiction. This concludes the proof. ◻

To present the following result, we define

$$C^\tau_{\bar{u}} := \{v \in L^2(\Omega) \text{ satisfying (3.23) and } v(x) = 0 \text{ if } |\bar{\mathfrak{p}}(x)| > \tau\}.$$

THEOREM 3.8 (equivalent optimality conditions). *If $(\bar{y}, \bar{p}, \bar{u})$ denotes a local minimum of* (3.1)–(3.2) *satisfying the first order optimality conditions* (3.2), (3.6), *and*

(3.7), *then the following statements are equivalent:*

$$j''(\bar{u})v^2 > 0 \quad \forall v \in C_{\bar{u}} \setminus \{0\} \tag{3.29}$$

and

$$\exists \mu, \tau > 0: \quad j''(\bar{u})v^2 \geq \mu \|v\|^2_{L^2(\Omega)} \quad \forall v \in C^{\tau}_{\bar{u}}. \tag{3.30}$$

Proof. Since, for every $\tau > 0$, we have $C_{\bar{u}} = C^0_{\bar{u}} \subset C^{\tau}_{\bar{u}}$, it follows immediately that (3.30) implies (3.29).

To prove that (3.29) implies (3.30) we proceed by contradiction. Assume that, for every $\tau > 0$, there exists $w_\tau \in C^{\tau}_{\bar{u}}$ such that $j''(\bar{u})w_\tau^2 < \tau \|w_\tau\|^2_{L^2(\Omega)}$. Define $v_\tau := w_\tau / \|w_\tau\|_{L^2(\Omega)}$. Upon taking a subsequence, if necessary, we have

$$v_\tau \in C^{\tau}_{\bar{u}}, \qquad \|v_\tau\|_{L^2(\Omega)} = 1, \qquad j''(\bar{u})v_\tau^2 < \tau, \qquad v_\tau \rightharpoonup v \text{ in } L^2(\Omega) \tag{3.31}$$

as $\tau \downarrow 0$. Since the set of elements satisfying condition (3.23) is closed and convex in $L^2(\Omega)$, it is weakly closed. Consequently, v satisfies (3.23). It suffices to prove that $\bar{\mathfrak{p}}(x) \neq 0$ implies $v(x) = 0$ for a.e. $x \in \Omega$ to conclude that $v \in C_{\bar{u}}$. To do so, we invoke property (3.21), the fact that $\bar{\mathfrak{p}} \in L^2(\Omega)$, $v_\tau \in C^{\tau}_{\bar{u}}$, and (3.31) to conclude that

$$0 \leq \int_\Omega \bar{\mathfrak{p}}(x)v(x)\mathrm{d}x = \lim_{\tau \downarrow 0} \int_\Omega \bar{\mathfrak{p}}(x)v_\tau(x)\mathrm{d}x = \lim_{\tau \downarrow 0} \int_{|\bar{\mathfrak{p}}| \leq \tau} \bar{\mathfrak{p}}(x)v_\tau(x)\mathrm{d}x \leq \lim_{\tau \downarrow 0} \tau \sqrt{|\Omega|} = 0.$$

Thus, $\int_\Omega |\bar{\mathfrak{p}}(x)v(x)|\mathrm{d}x = \int_\Omega \bar{\mathfrak{p}}(x)v(x)\mathrm{d}x = 0$. This proves that, a.e. in Ω, if $\bar{\mathfrak{p}} \neq 0$ then $v = 0$. Consequently, $v \in C_{\bar{u}}$. On the other hand, on the basis of the arguments developed in the proof of Theorem 3.7 we invoke (3.31) and obtain

$$j''(\bar{u})v^2 \leq \liminf_{\tau \downarrow 0} j''(\bar{u})v_\tau^2 \leq \limsup_{\tau \downarrow 0} j''(\bar{u})v_\tau^2 \leq 0.$$

Since $v \in C_{\bar{u}}$, (3.29) allows us to conclude that $v = 0$ and $j''(\bar{u})v_\tau^2 \to 0$ as $\tau \downarrow 0$. Now, since $v_\tau \rightharpoonup 0$ in $L^2(\Omega)$ implies $z_{v_\tau} \to 0$ in $C(\bar{\Omega})$, (3.12) yields

$$\liminf_{\tau \downarrow 0} j''(\bar{u})v_\tau = \liminf_{\tau \downarrow 0} \left[\alpha - \left(\tfrac{\partial^2 a}{\partial y^2}(\cdot, \bar{y}) z^2_{v_\tau}, \bar{p} \right)_{L^2(\Omega)} + \sum_{t \in \mathcal{D}} z^2_{v_\tau}(t) \right] = \alpha > 0.$$

This contradicts the fact that $j''(\bar{u})v_\tau^2 \to 0$ as $\tau \downarrow 0$. ∎

4. Finite element approximation.

We now introduce the discrete setting in which we will operate. We first introduce some terminology and a few basic ingredients [8, 20, 21] that will be common to all of our discretizations. We denote by $\mathscr{T}_h = \{T\}$ a conforming partition, or mesh, of $\bar{\Omega}$ into closed simplices T with size $h_T = \text{diam}(T)$. Define $h := \max_{T \in \mathscr{T}_h} h_T$. We denote by $\mathbb{T} = \{\mathscr{T}_h\}_{h>0}$ a collection of conforming and quasi–uniform meshes $\mathscr{T}_h$.

Given a mesh $\mathscr{T}_h \in \mathbb{T}$, we define the finite element space of continuous piecewise polynomials of degree one as

$$\mathbb{V}_h := \{v_h \in C(\bar{\Omega}) : v_h|_T \in \mathbb{P}_1(T) \ \forall T \in \mathscr{T}_h\} \cap H^1_0(\Omega). \tag{4.1}$$

In the following sections we will present convergence properties and suitable error estimates for finite element approximations of the state equation, the adjoint equation, and the optimal control problem (3.1)–(3.2), respectively.

4.1. Discrete state equation. Let $f \in L^2(\Omega)$. We define the Galerkin approximation of the solution y to problem (2.1) by

$$\mathsf{y}_h \in \mathbb{V}_h: \quad (\nabla \mathsf{y}_h, \nabla v_h)_{L^2(\Omega)} + (a(\cdot, \mathsf{y}_h), v_h)_{L^2(\Omega)} = (f, v_h)_{L^2(\Omega)} \quad \forall v_h \in \mathbb{V}_h. \tag{4.2}$$

Let $a = a(x, y) : \Omega \times \mathbb{R} \to \mathbb{R}$ be a Carathéodory function that is monotone increasing and locally Lipschitz in y, a.e. in Ω, with $a(\cdot, 0) \in L^2(\Omega)$. An application of Brouwer's fixed point theorem yields the existence of a unique solution to (4.2). In addition, we have $\|\nabla \mathsf{y}_h\|_{L^2(\Omega)} \lesssim \|f - a(\cdot, 0)\|_{L^2(\Omega)}$; see [27, Theorem 3.2] and [14, Section 7].

We now provide a local regularity result for the solution y of problem (2.1) that will be of importance to derive error estimates. Let $\Omega_1 \Subset \Omega_0 \Subset \Omega$ with Ω_0 smooth. Let $f \in L^2(\Omega) \cap L^t(\Omega_0)$, where $t \in [2, \infty)$. Since y can be seen as the solution to

$$\mathsf{y} \in H_0^1(\Omega): \quad (\nabla \mathsf{y}, \nabla v)_{L^2(\Omega)} = (f - a(\cdot, \mathsf{y}), v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega),$$

we can invoke [7, Lemma 4.2] to deduce that

$$\|\mathsf{y}\|_{W^{2,t}(\Omega_1)} \leq C_t \left(\|f - a(\cdot, \mathsf{y})\|_{L^t(\Omega_0)} + \|f - a(\cdot, \mathsf{y})\|_{L^2(\Omega)} \right), \tag{4.3}$$

where C_t behaves as Ct, with $C > 0$, as $t \uparrow \infty$. Notice that we further assume that a satisfies $a(\cdot, 0) \in L^t(\Omega)$, which, since $a = a(x, y)$ is locally Lipschitz in y, implies that $\|a(\cdot, \mathsf{y})\|_{L^t(\Omega)} \lesssim \|f\|_{L^2(\Omega)} + \|a(\cdot, 0)\|_{L^t(\Omega)}$.

THEOREM 4.1 (a priori error estimates). *Let $\Omega \subset \mathbb{R}^d$ be an open, bounded, and convex polytope. Let $a = a(x, y) : \Omega \times \mathbb{R} \to \mathbb{R}$ be a Carathéodory function that is monotone increasing and locally Lipschitz in y with $a(\cdot, 0) \in L^2(\Omega)$. Let $\mathsf{y} \in H_0^1(\Omega) \cap H^2(\Omega)$ and $\mathsf{y}_h \in \mathbb{V}_h$ be the solutions to (2.1) and (4.2), respectively, with $f \in L^2(\Omega)$. If h is sufficiently small, we thus have the following error estimates:*

$$\|\mathsf{y} - \mathsf{y}_h\|_{L^2(\Omega)} \lesssim h^2 \|f - a(\cdot, 0)\|_{L^2(\Omega)}, \tag{4.4}$$

and

$$\|\mathsf{y} - \mathsf{y}_h\|_{L^\infty(\Omega)} \lesssim h^{2 - \frac{d}{2}} \|f - a(\cdot, 0)\|_{L^2(\Omega)}. \tag{4.5}$$

Let $\Omega_1 \Subset \Omega_0 \Subset \Omega$ with Ω_0 smooth. If, in addition, $f \in L^\infty(\Omega_0)$ and $a(\cdot, 0) \in L^\infty(\Omega)$, we thus have the following local error estimate in maximum–norm:

$$\|\mathsf{y} - \mathsf{y}_h\|_{L^\infty(\Omega_1)} \lesssim h^2 |\log h|^2. \tag{4.6}$$

In all estimates the hidden constant is independent of h.

Proof. We refer the reader to [13, Lemma 4] and [13, Theorem 1] for a proof of the estimates (4.4) and (4.5), respectively; see also [13, Theorem 2]. We provide a proof of (4.6) that is inspired in the arguments developed in [27, Theorem 3.5] and [7, Lemma 4.4]. We begin with a simple application of the triangle inequality and write

$$\|\mathsf{y} - \mathsf{y}_h\|_{L^\infty(\Omega_1)} \leq \|\mathsf{y} - \mathfrak{y}_h\|_{L^\infty(\Omega_1)} + \|\mathfrak{y}_h - \mathsf{y}_h\|_{L^\infty(\Omega_1)},$$

where $\mathfrak{y}_h$ solves (4.2) with $a(\cdot, \mathsf{y}_h)$ replaced by $a(\cdot, \mathsf{y})$. Let Λ_1 be a smooth domain such that $\Omega_1 \Subset \Lambda_1 \Subset \Omega_0$. Since $(\nabla(\mathsf{y} - \mathfrak{y}_h), \nabla v_h)_{L^2(\Omega)} = 0$ for all $v_h \in \mathbb{V}_h$, we invoke [29, Corollary 5.1] to obtain the existence of $h_0 \in (0, 1)$ such that

$$\|\mathsf{y} - \mathfrak{y}_h\|_{L^\infty(\Omega_1)} \lesssim |\log h| \|\mathsf{y} - v_h\|_{L^\infty(\Lambda_1)} + \ell^{-d/2} \|\mathsf{y} - \mathfrak{y}_h\|_{L^2(\Omega)}, \quad v_h \in \mathbb{V}_h,$$

for every $h \le h_0$. Here, ℓ is such that $\mathrm{dist}(\Omega_1, \partial\Lambda_1) \ge \ell$, $\mathrm{dist}(\Lambda_1, \partial\Omega) \ge \ell$, and $Ch \le \ell$, where $C > 0$. Since $f \in L^\infty(\Omega_0) \cap L^2(\Omega)$ and $a(\cdot, 0) \in L^\infty(\Omega)$, the regularity estimate (4.3) implies that $\mathsf{y} \in W^{2,t}(\Lambda_1) \cap H^2(\Omega)$ for every $t < \infty$. Thus

$$\|\mathsf{y} - \mathfrak{y}_h\|_{L^\infty(\Omega_1)} \le C_1 |\log h| t h^{2-\frac{d}{t}} \left[\|f - a(\cdot, \mathsf{y})\|_{L^\infty(\Omega_0)} + \|f - a(\cdot, \mathsf{y})\|_{L^2(\Omega)}\right] + C_2 h^2 \|f - a(\cdot, \mathsf{y})\|_{L^2(\Omega)}, \quad C_1, C_2 > 0.$$

Inspired by [30, page 3], we thus set $t = |\log h|$ to arrive at the local estimate $\|\mathsf{y} - \mathfrak{y}_h\|_{L^\infty(\Omega_1)} \lesssim h^2 |\log h|^2 (\|f - a(\cdot, \mathsf{y})\|_{L^\infty(\Omega_0)} + \|f - a(\cdot, \mathsf{y})\|_{L^2(\Omega)})$. To control $\|\mathfrak{y}_h - \mathsf{y}_h\|_{L^\infty(\Omega_1)}$ we observe that

$$\mathfrak{y}_h - \mathsf{y}_h \in \mathbb{V}_h : \quad (\nabla(\mathfrak{y}_h - \mathsf{y}_h), \nabla v_h)_{L^2(\Omega)} = (a(\cdot, \mathsf{y}_h) - a(\cdot, \mathsf{y}), v_h)_{L^2(\Omega)} \quad \forall v_h \in \mathbb{V}_h.$$

Define $\mathfrak{y} \in H_0^1(\Omega)$ as the solution to $(\nabla \mathfrak{y}, \nabla v)_{L^2(\Omega)} = (a(\cdot, \mathsf{y}_h) - a(\cdot, \mathsf{y}), v)_{L^2(\Omega)}$ for all $v \in H_0^1(\Omega)$. Observe that $\mathfrak{y}_h - \mathsf{y}_h$ can be seen as the finite element approximation of $\mathfrak{y}$ within $\mathbb{V}_h$. We thus proceed as follows: $\|\mathfrak{y}_h - \mathsf{y}_h\|_{L^\infty(\Omega_1)} \le \|\mathfrak{y} - (\mathfrak{y}_h - \mathsf{y}_h)\|_{L^\infty(\Omega_1)} + \|\mathfrak{y}\|_{L^\infty(\Omega_1)}$. Invoke a stability estimate for the problem that $\mathfrak{y}$ solves, a basic error estimate, and the Lipschitz property of $a = a(x, y)$ in y to obtain

$$\|\mathfrak{y}_h - \mathsf{y}_h\|_{L^\infty(\Omega_1)} \lesssim \|a(\cdot, \mathsf{y}_h) - a(\cdot, \mathsf{y})\|_{L^2(\Omega)} \lesssim \|\mathsf{y}_h - \mathsf{y}\|_{L^2(\Omega)} \lesssim h^2 \left(\|f\|_{L^2(\Omega)} + \|a(\cdot, \mathsf{y})\|_{L^2(\Omega)}\right),$$

upon using (4.4). This concludes the proof. ∎

4.2. Discrete adjoint equation. Let $u \in \mathbb{U}_{ad}$ and $\{y_t\}_{t \in \mathcal{D}} \subset \mathbb{R}$. We define the Galerkin approximation to the adjoint equation (3.6) by

$$(4.7) \quad q_h \in \mathbb{V}_h : \; (\nabla w_h, \nabla q_h)_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, y) q_h, w_h\right)_{L^2(\Omega)} = \sum_{t \in \mathcal{D}} \langle (y(t) - y_t)\delta_t, w_h \rangle$$

for all $w_h \in \mathbb{V}_h$. In (4.7) the variable $y \in H_0^1(\Omega) \cap L^\infty(\Omega)$ denotes the unique solution to problem (3.2) with $u \in \mathbb{U}_{ad}$, i.e., $y = \mathcal{S}u$. Standard results yield the existence and uniqueness of a discrete solution.

We present the following error estimates.

THEOREM 4.2 (error estimates). *Let $a = a(x, y) : \Omega \times \mathbb{R} \to \mathbb{R}$ be a Carathéodory function of class C^1 with respect to the second variable such that $a(\cdot, 0) \in L^2(\Omega)$. Assume that (A.2) holds and that, for all $\mathfrak{m} > 0$, $|\partial a/\partial y(x, y)| \le C_{\mathfrak{m}}$ for a.e. $x \in \Omega$ and $y \in [-\mathfrak{m}, \mathfrak{m}]$. Let $p \in W_0^{1,r}(\Omega)$, with $r \in [2d/(d+2), d/(d-1))$, and $q_h \in \mathbb{V}_h$ be the solutions to (3.6) and (4.7), respectively. Then*

$$(4.8) \quad \|p - q_h\|_{L^2(\Omega)} \lesssim h^{2-\frac{d}{2}} \sum_{t \in \mathcal{D}} |y(t) - y_t| \lesssim h^{2-\frac{d}{2}} \left[\|u - a(\cdot, 0)\|_{L^2(\Omega)} + \sum_{t \in \mathcal{D}} |y_t|\right]$$

and

$$(4.9) \quad \|p - q_h\|_{L^1(\Omega)} \lesssim h^2 |\log h|^2.$$

In both estimates the hidden constants are independent of h.

Proof. Define $\mathfrak{a}(x) = \partial a/\partial y(x, y(x))$, where $y = \mathcal{S}u$ and $u \in \mathbb{U}_{ad}$. Since $\mathfrak{a} \in L^\infty(\Omega)$ and $\mathfrak{a}(x) \ge 0$ a.e. in Ω, we can apply [10, Theorem 3] in combination with

Theorem 2.1 to deduce (4.8). The proof of the estimate (4.9) follows similar arguments as the ones developed in [7, Lemma 5.3] and [22]. Let $\mathfrak{w}$ be the solution to

$$\mathfrak{B}(\mathfrak{w}, v) := (\nabla \mathfrak{w}, \nabla v)_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, y)\mathfrak{w}, v\right)_{L^2(\Omega)} = (\mathfrak{f}, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega)$$

and let $\mathfrak{w}_h$ be the Ritz projection of $\mathfrak{w}$ within $\mathbb{V}_h$, i.e., $\mathfrak{w}_h \in \mathbb{V}_h$ is such that $\mathfrak{B}(\mathfrak{w}_h, v_h) = \mathfrak{B}(\mathfrak{w}, v_h)$ for all $v_h \in \mathbb{V}_h$. Let $\mathfrak{f} = \operatorname{sgn}(p - q_h)$. Thus,

$$\|p - q_h\|_{L^1(\Omega)} = \int_\Omega \mathfrak{f}(p - q_h)\mathrm{d}x = \mathfrak{B}(\mathfrak{w}, p) - \mathfrak{B}(\mathfrak{w}_h, q_h) = \sum_{t \in \mathcal{D}} (y(t) - y_t)(\mathfrak{m} - \mathfrak{m}_h)(t),$$

where we have used that p and q_h solve (3.6) and (4.7), respectively. Since $\mathcal{D} \Subset \Omega$, similar arguments to the ones used in the proof of (4.6) yield

$$\|p - q_h\|_{L^1(\Omega)} \lesssim h^2 |\log h|^2 \left(\|y\|_{L^\infty(\Omega)} + \sum_{t \in \mathcal{D}} |y_t| \right) \|\mathfrak{f}\|_{L^\infty(\Omega)}.$$

This concludes the proof. ∎

Let $y_h \in \mathbb{V}_h$ be the unique solution to the discrete problem (4.2) with $f = u_h \in \mathbb{U}_{ad} \subset L^\infty(\Omega)$. Define now the discrete variable $p_h \in \mathbb{V}_h$ as the unique solution to

$$(4.10)\quad (\nabla w_h, \nabla p_h)_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, y_h)p_h, w_h\right)_{L^2(\Omega)} = \sum_{t \in \mathcal{D}} \langle (y_h(t) - y_t)\delta_t, w_h \rangle \quad \forall w_h \in \mathbb{V}_h.$$

We present the following error estimate, which will be of importance to perform an a priori error analysis for a suitable discretization of our optimal control problem.

THEOREM 4.3 (auxiliary error estimate). *Let $a = a(x, y) : \Omega \times \mathbb{R} \to \mathbb{R}$ be a Carathéodory function of class C^1 with respect to the second variable such that $a(\cdot, 0) \in L^2(\Omega)$. Assume that (A.2) holds and that, for all $\mathfrak{m} > 0$, $|\partial a/\partial y(x, y)| \leq C_{\mathfrak{m}}$ for a.e. $x \in \Omega$ and $y \in [-\mathfrak{m}, \mathfrak{m}]$. Let $u, u_h \in L^2(\Omega)$ be such that $\|u\|_{L^2(\Omega)} \leq C$ and $\|u_h\|_{L^2(\Omega)} \leq C$ for every $h > 0$, where $C > 0$. Let p and p_h be the solutions to (3.6) and (4.10) with $y = y(u)$ and $y_h = y_h(u_h)$, respectively. Then, we have*

$$(4.11)\quad \|p - p_h\|_{L^2(\Omega)} \lesssim \|u - u_h\|_{L^2(\Omega)} + h^{2 - \frac{d}{2}},$$

with a hidden constant that is independent of h.

Proof. We begin by introducing the auxiliary variable $\hat{p}$ as the unique solution to the problem: Find $\hat{p} \in W_0^{1,r}(\Omega)$, with $r \in [2d/(d+2), d/(d-1))$, such that

$$(\nabla w, \nabla \hat{p})_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, y_h)\hat{p}, w\right)_{L^2(\Omega)} = \sum_{t \in \mathcal{D}} \langle (y_h(t) - y_t)\delta_t, w \rangle \quad \forall w \in W_0^{1,r'}(\Omega).$$

With the variable $\hat{p}$ at hand, a trivial application of the triangle inequality yields

$$(4.12)\quad \|p - p_h\|_{L^2(\Omega)} \leq \|p - \hat{p}\|_{L^2(\Omega)} + \|\hat{p} - p_h\|_{L^2(\Omega)}.$$

We first estimate $\|\hat{p} - p_h\|_{L^2(\Omega)}$. Since p_h corresponds to the finite element approximation of the auxiliary variable $\hat{p}$, within $\mathbb{V}_h$, estimate (4.8) yields

$$(4.13)\quad \|\hat{p} - p_h\|_{L^2(\Omega)} \lesssim h^{2 - \frac{d}{2}} \sum_{t \in \mathcal{D}} |y_h(t) - y_t| \lesssim h^{2 - \frac{d}{2}} \left(\|y_h\|_{L^\infty(\Omega)} + \sum_{t \in \mathcal{D}} |y_t| \right).$$

Observe that $\|y_h\|_{L^\infty(\Omega)}$ is uniformly bounded. In fact, let us introduce the variable $\hat{y} \in H_0^1(\Omega) \cap L^\infty(\Omega)$ as the unique solution to problem (3.2) with $u = u_h$. Notice that y_h corresponds to the finite element approximation of $\hat{y}$ within $\mathbb{V}_h$. We thus invoke the error estimate (4.5), Theorem 2.1, and the assumption on u_h, to obtain

$$(4.14)\quad \|y_h\|_{L^\infty(\Omega)} \leq \|\hat{y} - y_h\|_{L^\infty(\Omega)} + \|\hat{y}\|_{L^\infty(\Omega)} \lesssim (h^{2-\frac{d}{2}} + 1)\|u_h - a(\cdot, 0)\|_{L^2(\Omega)} \leq C,$$

where C denotes a positive constant that is independent of the involved continuous and discrete variables and h. Replacing estimate (4.14) into (4.13), and the obtained one into (4.12), we conclude the error estimate

$$(4.15)\qquad \|p - p_h\|_{L^2(\Omega)} \lesssim \|p - \hat{p}\|_{L^2(\Omega)} + h^{2-\frac{d}{2}}.$$

We now bound $\|p - \hat{p}\|_{L^2(\Omega)}$ in (4.15). To accomplish this task, we introduce

$$\begin{aligned}\phi := p - \hat{p} \in W_0^{1,r}(\Omega): \quad & (\nabla w, \nabla \phi)_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, y)\phi, w\right)_{L^2(\Omega)} \\ & = \sum_{t\in\mathcal{D}} \langle (y(t) - y_h(t))\delta_t, w\rangle + \left(\left[\tfrac{\partial a}{\partial y}(\cdot, y_h) - \tfrac{\partial a}{\partial y}(\cdot, y)\right]\hat{p}, w\right)_{L^2(\Omega)}\end{aligned}$$

for all $w \in W_0^{1,r'}(\Omega)$. Here, $r \in [2d/(d+2), d/(d-1))$. An inf-sup condition that follows from [10, Theorem 1] yields the stability estimate

$$(4.16)\qquad \begin{aligned}\|\nabla\phi\|_{L^r(\Omega)} & \lesssim \sum_{t\in\mathcal{D}} |y(t) - y_h(t)| + \left\|\left[\tfrac{\partial a}{\partial y}(\cdot, y_h) - \tfrac{\partial a}{\partial y}(\cdot, y)\right]\hat{p}\right\|_{L^2(\Omega)} \\ & \lesssim \|y - y_h\|_{L^\infty(\Omega)}(1 + \|\hat{p}\|_{L^2(\Omega)}).\end{aligned}$$

To obtain the last estimate we have used that $\partial a/\partial y = \partial a/\partial y(x, y)$ is locally Lipschitz in y. We now bound the term $\|\hat{p}\|_{L^2(\Omega)}$. Since $W_0^{1,r}(\Omega) \hookrightarrow L^2(\Omega)$ for $r \in [2d/(d+2), d/(d-1))$, an stability estimate for the problem that $\hat{p}$ solves yields

$$\|\hat{p}\|_{L^2(\Omega)} \lesssim \|\nabla\hat{p}\|_{L^r(\Omega)} \lesssim \|y_h\|_{L^\infty(\Omega)} + \sum_{t\in\mathcal{D}} |y_t|.$$

This estimate, (4.16), and (4.14) yields $\|\nabla\phi\|_{L^r(\Omega)} \lesssim \|y - y_h\|_{L^\infty(\Omega)}$. We thus invoke that $\phi = p - \hat{p}$, $r \in [2d/(d+2), d/(d-1))$, and $W_0^{1,r}(\Omega) \hookrightarrow L^2(\Omega)$ to arrive at

$$(4.17)\qquad \|p - \hat{p}\|_{L^2(\Omega)} = \|\phi\|_{L^2(\Omega)} \lesssim \|y - y_h\|_{L^\infty(\Omega)},$$

with a hidden constant that is independent of the involved continuous and discrete variables and h.

Our final goal now is to bound $\|y - y_h\|_{L^\infty(\Omega)}$. Invoke the variable $\hat{y}$ and write

$$\|y - y_h\|_{L^\infty(\Omega)} \leq \|y - \hat{y}\|_{L^\infty(\Omega)} + \|\hat{y} - y_h\|_{L^\infty(\Omega)}.$$

In view of the Lipschitz property (2.2) and the estimate (4.5), we derive

$$(4.18)\qquad \|y - y_h\|_{L^\infty(\Omega)} \lesssim \|u - u_h\|_{L^2(\Omega)} + h^{2-\frac{d}{2}}\|u_h - a(\cdot, 0)\|_{L^2(\Omega)}.$$

Replacing the estimate (4.18) into (4.17) and the obtained one into (4.15), and taking into account the assumption on u_h, we conclude the desired estimate (4.11). ∎

4.3. Discretization of the control problem. Let us introduce the finite element space of piecewise constant functions over $\mathscr{T}_h$, $\mathbb{U}_h = \{u_h \in L^\infty(\Omega) : u_h|_T \in \mathbb{P}_0(T)\ \forall T \in \mathscr{T}_h\}$, and the space of discrete admissible controls, $\mathbb{U}_{ad,h} := \mathbb{U}_h \cap \mathbb{U}_{ad}$. With this discrete setting at hand, we propose the following finite element discretization of the optimal control problem (3.1)–(3.2): Find $\min J(y_h, u_h)$ subject to

$$(4.19)\quad y_h \in \mathbb{V}_h : \quad (\nabla y_h, \nabla v_h)_{L^2(\Omega)} + (a(\cdot, y_h), v_h)_{L^2(\Omega)} = (u_h, v_h)_{L^2(\Omega)} \quad \forall v_h \in \mathbb{V}_h,$$

and the discrete constraints $u_h \in \mathbb{U}_{ad,h}$. We recall that $\mathbb{V}_h$ is defined as in (4.1).

The existence of at least one solution for the previously defined discrete optimal control problem follows immediately from the compactness of $\mathbb{U}_{ad,h}$ and the continuity of the cost functional J. Let us introduce the discrete control to state map $\mathcal{S}_h$: $\mathbb{U}_h \ni u_h \mapsto y_h \in \mathbb{V}_h$, where y_h solves (4.19), and define the reduced cost functional $j_h(u_h) := J(\mathcal{S}_h u_h, u_h)$. With these ingredients at hand, as in the continuous case, we can derive first order optimality conditions for the discrete optimal control problem. In particular, if $\bar{u}_h$ denotes a local solution, then

$$(4.20)\qquad j_h'(\bar{u}_h)(u_h - \bar{u}_h) = (\bar{p}_h + \alpha\bar{u}_h, u_h - \bar{u}_h)_{L^2(\Omega)} \geq 0 \quad \forall u_h \in \mathbb{U}_{ad,h},$$

where $\bar{p}_h \in \mathbb{V}_h$ solves the discrete problem (4.10) with $y_h = \bar{y}_h := \mathcal{S}_h \bar{u}_h$.

The following error estimates can be found in [14, Lemmas 37 and 38].

THEOREM 4.4 (auxiliary error estimate). *Let Ω be a convex polytope. Assume that (A.1) and (A.2) hold. Let $u \in \mathbb{U}_{ad}$ and $u_h \in \mathbb{U}_{ad,h} \subset \mathbb{U}_{ad}$. Let $y = y(u)$ be the solution to (3.2) and let $y_h = y_h(u_h)$ be the solution to (4.19). Then,*

$$\|\nabla(y - y_h)\|_{L^2(\Omega)} \lesssim h + \|u - u_h\|_{L^2(\Omega)}, \quad \|y - y_h\|_{L^\infty(\Omega)} \lesssim h^{2-\frac{d}{2}} + \|u - u_h\|_{L^2(\Omega)}.$$

In addition, if $u_h \rightharpoonup u$ in $L^s(\Omega)$ as $h \downarrow 0$, with $s > d/2$, then $y_h \to y$ in $H_0^1(\Omega) \cap C(\bar{\Omega})$ as $h \downarrow 0$ and $j(u) \leq \liminf_{h\downarrow 0} j_h(u_h)$.

In what follows we provide a convergence result that, in essence, guarantees that the sequence of global solutions $\{\bar{u}_h\}$ of the discrete optimal control problems converge, as $h \downarrow 0$, to a solution of the continuous optimal control problem.

THEOREM 4.5 (convergence of the discrete solutions). *Assume that (A.1), (A.2), and (A.3) hold. Let $h > 0$ and $\bar{u}_h \in \mathbb{U}_{ad,h}$ be a global solution of the discrete optimal control problem. Then, there exist nonrelabeled subsequences of $\{\bar{u}_h\}_{h>0}$ such that $\bar{u}_h \overset{*}{\rightharpoonup} \bar{u}$ in the weak* topology of $L^\infty(\Omega)$, as $h \downarrow 0$, where $\bar{u}$ corresponds to a local solution of the optimal control problem (3.1)–(3.2). In addition, it follows that*

$$(4.21)\qquad \lim_{h\to 0} \|\bar{u} - \bar{u}_h\|_{L^2(\Omega)} = 0, \qquad \lim_{h\to 0} j_h(\bar{u}_h) = j(\bar{u}).$$

Proof. We begin the proof by noticing that, since $\bar{u}_h \in \mathbb{U}_{ad,h} \subset \mathbb{U}_{ad}$ for every $h > 0$, the sequence $\{\bar{u}_h\}_{h>0}$ is uniformly bounded in $L^\infty(\Omega)$. Then, there exists a nonrelabeled subsequence such that $\bar{u}_h \overset{*}{\rightharpoonup} \bar{u}$ in $L^\infty(\Omega)$ as $h \downarrow 0$. In what follows, we prove that $\bar{u} \in \mathbb{U}_{ad}$ is a solution to the optimal control problem (3.1)–(3.2) and that the convergence results in (4.21) hold.

Let $\tilde{u} \in \mathbb{U}_{ad}$ be a solution to (3.1)–(3.2). Define $\tilde{u}_h := \Pi_{L^2}\tilde{u} \in \mathbb{U}_{ad,h}$, the orthogonal projection of $\tilde{u}$ into piecewise constant functions over $\mathscr{T}_h$. We recall that

$$\Pi_{L^2} : L^2(\Omega) \to \mathbb{U}_h, \quad \Pi_{L^2} v|_T := \frac{1}{|T|}\int_T v \mathrm{d}x, \quad T \in \mathscr{T}, \quad v \in L^2(\Omega).$$

Since Theorem 3.4 guarantees that $\tilde{u} \in H^1(\Omega)$, we immediately conclude that $\|\tilde{u} - \tilde{u}_h\|_{L^2(\Omega)} \to 0$ as $h \downarrow 0$. We thus invoke, the local optimality of $\tilde{u}$, Theorem 4.4, the global optimality of $\bar{u}_h$, and the convergence result $\tilde{u}_h \to \tilde{u}$ in $L^2(\Omega)$ to obtain

$$j(\tilde{u}) \leq j(\bar{u}) \leq \liminf_{h\downarrow 0} j_h(\bar{u}_h) \leq \limsup_{h\downarrow 0} j_h(\bar{u}_h) \leq \limsup_{h\downarrow 0} j_h(\tilde{u}_h) = j(\tilde{u}).$$

This proves that $\bar{u}$ is a solution to problem (3.1)–(3.2) and that $\lim_{h\downarrow 0} j_h(\bar{u}_h) = j(\bar{u})$.

The strong convergence $\bar{u}_h \to \bar{u}$ in $L^2(\Omega)$ follows from $\lim_{h\downarrow 0} j_h(\bar{u}_h) = j(\bar{u})$ and $\bar{y}_h \to \bar{y}$ in $C(\bar{\Omega})$; see Theorem 4.4. In fact, the latter converge result implies that

$$\sum_{t\in\mathcal{D}} (y_h(t) - y_t)^2 \to \sum_{t\in\mathcal{D}} (y(t) - y_t)^2, \quad h \downarrow 0.$$

Since $\lim_{h\downarrow 0} j_h(\bar{u}_h) = j(\bar{u})$, we thus conclude that $\|\bar{u}_h\|^2_{L^2(\Omega)} \to \|\bar{u}\|^2_{L^2(\Omega)}$ as $h \downarrow 0$. The weak convergence $\bar{u}_h \rightharpoonup \bar{u}$ in $L^2(\Omega)$, as $h \downarrow 0$, allows us to conclude. ∎

5. Error estimates.

Let $\{\bar{u}_h\}_{h>0} \subset \mathbb{U}_{ad,h}$ be a sequence of local minima of the discrete optimal control problems such that $\bar{u}_h \to \bar{u}$ in $L^2(\Omega)$, as $h \downarrow 0$, where $\bar{u} \in \mathbb{U}_{ad}$ is a local solution of (3.1)–(3.2); see Theorem 4.5. The main goal of this section is to derive the following a priori error estimate for $\bar{u} - \bar{u}_h$ in $L^2(\Omega)$:

THEOREM 5.1 (error estimate). *Assume that (A.1), (A.2), and (A.3) hold, and that $a(\cdot, 0) \in L^\infty(\Omega)$. Let $\bar{u} \in \mathbb{U}_{ad}$ satisfies the sufficient second order optimality condition (3.29). Then there exists $h_\ddagger > 0$ such that the following inequality holds:*

$$\|\bar{u} - \bar{u}_h\|_{L^2(\Omega)} \lesssim h|\log h| \quad \forall h < h_\ddagger, \tag{5.1}$$

with a hidden constant that is independent of h.

To prove this result we will proceed by contradiction following [17, 11]. We will assume that $\{\bar{u}_h\}_{h>0}$ converges to $\bar{u}$ as $h \downarrow 0$ and (5.1) does not hold. If we assume that (5.1) is false, we can thus find, for every $k \in \mathbb{N}$, $h_k > 0$ such that $\|\bar{u} - \bar{u}_{h_k}\|_{L^2(\Omega)} > kh_k|\log h_k|$, and thus a sequence $\{h_k\}_{k\in\mathbb{N}} \subset \mathbb{R}^+$ such that

$$\lim_{h_k\downarrow 0} \|\bar{u} - \bar{u}_{h_k}\|_{L^2(\Omega)} \to 0, \qquad \lim_{h_k\downarrow 0} \frac{\|\bar{u} - \bar{u}_{h_k}\|_{L^2(\Omega)}}{h_k|\log h_k|} = +\infty. \tag{5.2}$$

To prove the estimate in Theorem 5.1, we need some preparatory lemmas.

LEMMA 5.2 (auxiliary result). *Assume that (A.1), (A.2), and (A.3) hold. Let $\bar{u} \in \mathbb{U}_{ad}$ satisfies the second order optimality condition (3.29). Let us assume, in addition, that (5.1) is false. Then there exists $h_\dagger > 0$ such that*

$$\mathfrak{C}\|\bar{u} - \bar{u}_h\|^2_{L^2(\Omega)} \leq [j'(\bar{u}_h) - j'(\bar{u})](\bar{u}_h - \bar{u}) \quad \forall h < h_\dagger, \tag{5.3}$$

where $\mathfrak{C} = 2^{-1}\min\{\mu, \alpha\}$, with α being the regularization parameter and μ the constant appearing in estimate (3.30).

Proof. Since (5.1) is false, there exists a sequence $\{h_k\}_{k\in\mathbb{N}}$ such that the limits in (5.2) hold. In an attempt to simplify the exposition of the material, in what follows, we will omit the subindex k, i.e., we denote $u_{h_k} = u_h$. Observe that $h \downarrow 0$ as $k \uparrow \infty$.

Define $v_h := (\bar{u}_h - \bar{u})/\|\bar{u}_h - \bar{u}\|_{L^2(\Omega)}$. Upon taking a subsequence, if necessary, we can assume that $v_h \rightharpoonup v$ in $L^2(\Omega)$ as $h \downarrow 0$. In what follows, we prove that $v \in C_{\bar{u}}$, with $C_{\bar{u}}$ defined as in (3.22). Since $\bar{u}_h \in \mathbb{U}_{ad,h} \subset \mathbb{U}_{ad}$, it is clear that v_h satisfies the sign conditions in (3.23). The fact that $v_h \rightharpoonup v$ in $L^2(\Omega)$ as $h \downarrow 0$ implies that v satisfies (3.23) as well. To show that $v(x) = 0$ if $\bar{\mathfrak{p}}(x) \neq 0$ for a.e. $x \in \Omega$, we introduce

$$\bar{\mathfrak{p}}_h := \bar{p}_h + \alpha \bar{u}_h. \tag{5.4}$$

Since $\|\bar{u} - \bar{u}_h\|_{L^2(\Omega)} \to 0$ as $h \downarrow 0$, Theorem 4.3 yields $\bar{\mathfrak{p}}_h \to \bar{\mathfrak{p}}$ in $L^2(\Omega)$ as $h \downarrow 0$. Thus,

$$\int_\Omega \bar{\mathfrak{p}}(x)v(x)\mathrm{d}x = \lim_{h\to 0} \int_\Omega \bar{\mathfrak{p}}_h(x)v_h(x)\mathrm{d}x = \lim_{h\to 0} \frac{1}{\|\bar{u}_h - \bar{u}\|_{L^2(\Omega)}}$$
$$\cdot \left(\int_\Omega \bar{\mathfrak{p}}_h(\Pi_{L^2}\bar{u} - \bar{u})\mathrm{d}x + \int_\Omega \bar{\mathfrak{p}}_h(\bar{u}_h - \Pi_{L^2}\bar{u})\mathrm{d}x \right) =: \lim_{h\to 0} \frac{\mathbf{I} + \mathbf{II}}{\|\bar{u}_h - \bar{u}\|_{L^2(\Omega)}}.$$

We recall that Π_{L^2} denotes the L^2-orthogonal projection into piecewise constant functions over $\mathscr{T}_h$. The discrete variational inequality (4.20) immediately yields $\mathbf{II} \leq 0$. On the other hand, $|\mathbf{I}| \leq \|\bar{\mathfrak{p}}_h\|_{L^2(\Omega)}\|\bar{u} - \Pi_{L^2}\bar{u}\|_{L^2(\Omega)} \lesssim h\|\nabla \bar{u}\|_{L^2(\Omega)}$, upon noticing that $\|\bar{\mathfrak{p}}_h\|_{L^2(\Omega)} \leq \|\bar{\mathfrak{p}}_h - \bar{\mathfrak{p}}\|_{L^2(\Omega)} + \|\bar{\mathfrak{p}}\|_{L^2(\Omega)} \leq C$, where $C > 0$. On the basis of (5.2) the previous inequalities yield

$$\int_\Omega \bar{\mathfrak{p}}(x)v(x)\mathrm{d}x \lesssim \lim_{h\to 0} \frac{h}{\|\bar{u}_h - \bar{u}\|_{L^2(\Omega)}} \lesssim \lim_{h\to 0} \frac{h|\log h|}{\|\bar{u}_h - \bar{u}\|_{L^2(\Omega)}} = 0.$$

Since v satisfies the sign condition (3.23), then $\bar{\mathfrak{p}}(x)v(x) \geq 0$. Therefore the previous inequality yields $\int_\Omega |\bar{\mathfrak{p}}(x)v(x)|\mathrm{d}x = \int_\Omega \bar{\mathfrak{p}}(x)v(x)\mathrm{d}x \leq 0$. Consequently, if $\bar{\mathfrak{p}}(x) \neq 0$, then $v(x) = 0$ for a.e. $x \in \Omega$. This allows us to conclude that $v \in C_{\bar{u}}$.

We now invoke the mean value theorem to deduce that

$$[j'(\bar{u}_h) - j'(\bar{u})](\bar{u}_h - \bar{u}) = j''(\hat{u}_h)(\bar{u}_h - \bar{u})^2, \quad \hat{u}_h = \bar{u} + \theta_h(\bar{u}_h - \bar{u}), \tag{5.5}$$

where $\theta_h \in (0,1)$. Let $y_{\hat{u}_h}$ be unique solution to (3.2) with $u = \hat{u}_h$ and $p_{\hat{u}_h}$ be the unique solution to (3.6) with $y = y_{\hat{u}_h}$. Since $\bar{u}_h \to \bar{u}$ in $L^2(\Omega)$ as $h \downarrow 0$, we have $y_{\hat{u}_h} \to \bar{y}$ in $H_0^1(\Omega) \cap C(\bar{\Omega})$ and $p_{\hat{u}_h} \to \bar{p}$ in $W_0^{1,r}(\Omega)$ as $h \downarrow 0$. Here $r < d/(d-1)$. Similarly, $v_h \rightharpoonup v$ in $L^2(\Omega)$ implies that $z_{v_h} \rightharpoonup z_v$ in $H^2(\Omega) \cap H_0^1(\Omega)$ as $h \downarrow 0$. Hence, invoke (3.12), the definition of v_h, and the second order condition (3.30) to obtain

$$\lim_{h\downarrow 0} j''(\hat{u}_h)v_h^2 = \lim_{h\downarrow 0} \left(\alpha - \left(\tfrac{\partial^2 a}{\partial y^2}(\cdot, y_{\hat{u}_h}) z_{v_h}^2, p_{\hat{u}_h} \right)_{L^2(\Omega)} + \sum_{t\in\mathcal{D}} z_{v_h}^2(t) \right)$$
$$= \alpha - \left(\tfrac{\partial^2 a}{\partial y^2}(\cdot, \bar{y}) z_v^2, \bar{p} \right)_{L^2(\Omega)} + \sum_{t\in\mathcal{D}} z_v^2(t)$$
$$= \alpha + j''(\bar{u})v^2 - \alpha\|v\|_{L^2(\Omega)}^2 \geq \alpha + (\mu - \alpha)\|v\|_{L^2(\Omega)}^2.$$

Therefore, since $\|v\|_{L^2(\Omega)} \leq 1$, we arrive at $\lim_{h\downarrow 0} j''(\hat{u}_h)v_h^2 \geq \min\{\mu, \alpha\} > 0$, which proves the existence of $h_\dagger > 0$ such that

$$j''(\hat{u}_h)v_h^2 \geq 2^{-1}\min\{\mu, \alpha\} \quad \forall h < h_\dagger.$$

This, in light of the definition of v_h and the identity (5.5), allows us to conclude. ◻

LEMMA 5.3 (auxiliary result). *Assume that (A.1), (A.2), and (A.3) hold, and that $a(\cdot, 0) \in L^\infty(\Omega)$. Let $u_1, u_2 \in \mathbb{U}_{ad}$ and $v \in L^\infty(\Omega)$. Thus, we have the estimates*

$$|j'(u_1)v - j_h'(u_1)v| \lesssim h^2|\log h|^2\|v\|_{L^\infty(\Omega)}, \tag{5.6}$$

and

$$|j_h'(u_1)v - j_h'(u_2)v| \lesssim \|u_1 - u_2\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)}. \tag{5.7}$$

Proof. We proceed on the basis of two steps.

Step 1. The goal of this step is to derive (5.6). To accomplish this task, we begin with a basic computation which reveals that $j'(u_1)v = (p_{u_1} + \alpha u_1, v)_{L^2(\Omega)}$, where

$$p_{u_1} \in W_0^{1,r}(\Omega): \quad (\nabla w, \nabla p_{u_1})_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, y_{u_1}) p_{u_1}, w\right)_{L^2(\Omega)} = \sum_{t \in \mathcal{D}} \langle (y_{u_1}(t) - y_t)\delta_t, w\rangle$$

for all $w \in W_0^{1,r'}(\Omega)$. Here, $r \in [2d/(d+2), d/(d-1))$, y_{u_1} denotes the unique solution to the state equation (3.2) with $u = u_1$, and r' denotes the Hölder's conjugate of r. A similar argument yields $j_h'(u_1)v = (\hat{p}_h + \alpha u_1, v)_{L^2(\Omega)}$, where $\hat{p}_h$ is such that

$$\hat{p}_h \in \mathbb{V}_h: \quad (\nabla w_h, \nabla \hat{p}_h)_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, \hat{y}_h)\hat{p}_h, w_h\right)_{L^2(\Omega)} = \sum_{t \in \mathcal{D}} \langle (\hat{y}_h(t) - y_t)\delta_t, w_h\rangle \tag{5.8}$$

for all $w_h \in \mathbb{V}_h$. In (5.8) the variable $\hat{y}_h \in \mathbb{V}_h$ corresponds to the solution to (4.19) with u_h replaced by u_1. Define $\hat{p} \in W_0^{1,r}(\Omega)$ as the unique solution to

$$(\nabla w, \nabla \hat{p})_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, \hat{y}_h)\hat{p}, w\right)_{L^2(\Omega)} = \sum_{t \in \mathcal{D}} \langle (\hat{y}_h(t) - y_t)\delta_t, w\rangle \quad \forall w \in W_0^{1,r'}(\Omega).$$

Here, $r \in [2d/(d+2), d/(d-1))$. Notice that $\hat{p}_h \in \mathbb{V}_h$ corresponds to the finite element approximation of $\hat{p}$ within $\mathbb{V}_h$. We also notice the following stability estimate for $\hat{p}$:

$$\|\nabla \hat{p}\|_{L^r(\Omega)} \lesssim \sum_{t \in \mathcal{D}} |\hat{y}_h(t) - y_t|. \tag{5.9}$$

With all these continuous and discrete variables at hand, we can write

$$j'(u_1)v - j_h'(u_1)v = (p_{u_1} - \hat{p}, v)_{L^2(\Omega)} + (\hat{p} - \hat{p}_h, v)_{L^2(\Omega)} := \mathbf{I} + \mathbf{II}. \tag{5.10}$$

To estimate the term $\mathbf{I}$ we define $\zeta := p_{u_1} - \hat{p} \in W_0^{1,r}(\Omega)$ and observe that

$$\begin{aligned}\zeta \in W_0^{1,r}(\Omega): \quad & (\nabla w, \nabla \zeta)_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, y_{u_1})\zeta, w\right)_{L^2(\Omega)} \\ & = \sum_{t \in \mathcal{D}} \langle (y_{u_1}(t) - \hat{y}_h(t))\delta_t, w\rangle + \left(\left[\tfrac{\partial a}{\partial y}(\cdot, \hat{y}_h) - \tfrac{\partial a}{\partial y}(\cdot, y_{u_1})\right]\hat{p}, w\right)_{L^2(\Omega)}\end{aligned}$$

for all $w \in W_0^{1,r'}(\Omega)$. An inf-sup condition that follows from [10, Theorem 1] yields

$$\|\nabla \zeta\|_{L^r(\Omega)} \lesssim \sum_{t \in \mathcal{D}} |y_{u_1}(t) - \hat{y}_h(t)| + \left\| \left[\tfrac{\partial a}{\partial y}(\cdot, \hat{y}_h) - \tfrac{\partial a}{\partial y}(\cdot, y_{u_1})\right]\hat{p}\right\|_{L^2(\Omega)}. \tag{5.11}$$

Let us concentrate on the second term of the right hand side of (5.11). Let Λ_1, Ω_0 be smooth domains such that $\Omega_1 \Subset \Lambda_1 \Subset \Omega_0 \Subset \Omega$ and $\mathcal{D} \subset \Omega_1$. Observe that

$$\begin{aligned}\mathfrak{I}^2 := \left\| \left[\tfrac{\partial a}{\partial y}(\cdot, \hat{y}_h) - \tfrac{\partial a}{\partial y}(\cdot, y_{u_1})\right]\hat{p}\right\|_{L^2(\Omega)}^2 &= \left\| \left[\tfrac{\partial a}{\partial y}(\cdot, \hat{y}_h) - \tfrac{\partial a}{\partial y}(\cdot, y_{u_1})\right]\hat{p}\right\|_{L^2(\Lambda_1)}^2 \\ &+ \left\| \left[\tfrac{\partial a}{\partial y}(\cdot, \hat{y}_h) - \tfrac{\partial a}{\partial y}(\cdot, y_{u_1})\right]\hat{p}\right\|_{L^2(\Omega\setminus\Lambda_1)}^2.\end{aligned} \tag{5.12}$$

In view of the estimates of Theorem 4.1 we can thus arrive at

$$\begin{aligned}\mathfrak{I}^2 &\lesssim \|\hat{p}\|_{L^2(\Lambda_1)}^2 \|\hat{y}_h - y_{u_1}\|_{L^\infty(\Lambda_1)}^2 + \|\hat{p}\|_{L^\infty(\Omega\setminus\Lambda_1)}^2 \|\hat{y}_h - y_{u_1}\|_{L^2(\Omega)}^2 \\ &\lesssim h^4 |\log h|^4 \|\hat{p}\|_{L^2(\Omega)}^2 + h^4 \|\hat{p}\|_{L^\infty(\Omega\setminus\Lambda_1)}^2 \|u_1 - a(\cdot, 0)\|_{L^2(\Omega)}^2.\end{aligned}$$

Invoke the Sobolev embedding $W_0^{1,r}(\Omega) \hookrightarrow L^2(\Omega)$, with $r \in [2d/(d+2), d/(d-1))$, and the fact that $u_1 \in \mathbb{U}_{ad}$ to obtain

$$\mathfrak{I}^2 \lesssim h^4 |\log h|^4 \|\nabla \hat{p}\|^2_{L^r(\Omega)} + h^4 \|\hat{p}\|^2_{L^\infty(\Omega\setminus\Lambda_1)}, \tag{5.13}$$

where the hidden constant is independent of the involved continuous and discrete variables but depends on the continuous optimal control problem data. We now utilize [12, Theorem 3.4] to conclude that $\|\hat{p}\|_{L^\infty(\Omega\setminus\Lambda_1)}$ is uniformly bounded. On the other hand, the stability estimate (5.9) and analogous arguments to the ones that lead to (4.14) allows us to conclude that $\|\nabla \hat{p}\|_{L^r(\Omega)} \leq C$, where C depends on $\{y_t\}_{t\in\mathcal{D}}$, a, $\mathtt{a}$, and $\mathtt{b}$. We thus invoke (5.13) to arrive at $\mathfrak{I} \lesssim h^2|\log h|^2$. This bound, estimate (5.11), and the local estimate (4.6), yield

$$\mathbf{I} \leq \|\zeta\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)} \lesssim \|\nabla\zeta\|_{L^r(\Omega)}\|v\|_{L^2(\Omega)} \lesssim h^2|\log h|^2\|v\|_{L^2(\Omega)}. \tag{5.14}$$

The control of $\mathbf{II}$ in (5.10) follows immediately from the error estimate (4.9):

$$\mathbf{II} \leq \|\hat{p} - \hat{p}_h\|_{L^1(\Omega)}\|v\|_{L^\infty(\Omega)} \lesssim h^2|\log h|^2\|v\|_{L^\infty(\Omega)}. \tag{5.15}$$

Upon combining (5.10), (5.14), and (5.15), we obtain the estimate (5.6).

Step 2. In this step we derive (5.7). From the previous step, we have that $j_h'(u_1)v = (\hat{p}_h + \alpha u_1, v)_{L^2(\Omega)}$, where $\hat{p}_h \in \mathbb{V}_h$ is the unique solution to problem (5.8). On the other hand, similar arguments yield $j_h'(u_2)v = (\tilde{p}_h + \alpha u_2, v)_{L^2(\Omega)}$, where $\tilde{p}_h \in \mathbb{V}_h$ is the unique solution to the discrete problem

$$(\nabla w_h, \nabla \tilde{p}_h)_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, \tilde{y}_h)\tilde{p}_h, w_h\right)_{L^2(\Omega)} = \sum_{t\in\mathcal{D}}\langle(\tilde{y}_h(t) - y_t)\delta_t, w_h\rangle \quad \forall w_h \in \mathbb{V}_h,$$

and $\tilde{y}_h \in \mathbb{V}_h$ corresponds to the solution to (4.19) with u_h replaced by u_2. Therefore,

$$|j_h'(u_1)v - j_h'(u_2)v| \leq \left(\|\hat{p}_h - \tilde{p}_h\|_{L^2(\Omega)} + \alpha\|u_1 - u_2\|_{L^2(\Omega)}\right)\|v\|_{L^2(\Omega)}. \tag{5.16}$$

The rest of the proof is dedicated to bound $\|\hat{p}_h - \tilde{p}_h\|_{L^2(\Omega)}$. To accomplish this task, we define

$$\begin{aligned} \xi \in W_0^{1,r}(\Omega): \quad & (\nabla w, \nabla \xi)_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot, \hat{y}_h)\xi, w\right)_{L^2(\Omega)} \\ = \sum_{t\in\mathcal{D}}\langle(\hat{y}_h(t) - \tilde{y}_h(t))\delta_t, w\rangle & + \left(\left[\tfrac{\partial a}{\partial y}(\cdot, \tilde{y}_h) - \tfrac{\partial a}{\partial y}(\cdot, \hat{y}_h)\right]\tilde{p}_h, w\right)_{L^2(\Omega)} \quad \forall w \in W_0^{1,r'}(\Omega). \end{aligned}$$

We also define $\xi_h := \hat{p}_h - \tilde{p}_h \in \mathbb{V}_h$ and immediately observe that ξ_h corresponds to the finite element approximation of ξ within $\mathbb{V}_h$. We thus invoke basic estimates, (4.8), and an stability estimate for the problem that ξ solves to arrive at

$$\begin{aligned} \|\xi_h\|_{L^2(\Omega)} &\leq \|\xi - \xi_h\|_{L^2(\Omega)} + \|\xi\|_{L^2(\Omega)} \lesssim \|\xi - \xi_h\|_{L^2(\Omega)} + \|\nabla\xi\|_{L^r(\Omega)} \\ &\lesssim (h^{2-\frac{d}{2}} + 1)\left(\|\hat{y}_h - \tilde{y}_h\|_{L^\infty(\Omega)} + \left\|\left[\tfrac{\partial a}{\partial y}(\cdot, \tilde{y}_h) - \tfrac{\partial a}{\partial y}(\cdot, \hat{y}_h)\right]\tilde{p}_h\right\|_{L^2(\Omega)}\right). \end{aligned}$$

The previous estimate, in light of assumption (A.3), immediately yields

$$\|\hat{p}_h - \tilde{p}_h\|_{L^2(\Omega)} = \|\xi_h\|_{L^2(\Omega)} \lesssim (1 + \|\tilde{p}_h\|_{L^2(\Omega)})\|\hat{y}_h - \tilde{y}_h\|_{L^\infty(\Omega)}. \tag{5.17}$$

We now bound $\|\hat{y}_h - \tilde{y}_h\|_{L^\infty(\Omega)}$. Before proceeding with such an estimation, we recall that y_{u_i} solves (3.2) with $u = u_i$, where $i \in \{1, 2\}$, and that $\hat{y}_h$ and $\tilde{y}_h$

correspond to the finite element approximations of y_{u_1} and y_{u_2}, respectively. Since $a(\cdot,\hat{y}_h) - a(\cdot,\tilde{y}_h) = \frac{\partial a}{\partial y}(\cdot,\mathsf{y}_h)(\hat{y}_h - \tilde{y}_h)$, where $\mathsf{y}_h = \tilde{y}_h + \theta_h(\hat{y}_h - \tilde{y}_h)$ and $\theta_h \in (0,1)$, we deduce that $\hat{y}_h - \tilde{y}_h \in \mathbb{V}_h$ solves the problem

$$(\nabla(\hat{y}_h - \tilde{y}_h), \nabla v_h)_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot,\mathsf{y}_h)(\hat{y}_h - \tilde{y}_h), v_h\right)_{L^2(\Omega)} = (u_1 - u_2, v_h)_{L^2(\Omega)} \quad \forall v_h \in \mathbb{V}_h.$$

Define $\eta \in H_0^1(\Omega)$ as the solution to

$$(\nabla \eta, \nabla v)_{L^2(\Omega)} + \left(\tfrac{\partial a}{\partial y}(\cdot,\mathsf{y}_h)\eta, v\right)_{L^2(\Omega)} = (u_1 - u_2, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega).$$

Noticing that $\hat{y}_h - \tilde{y}_h \in \mathbb{V}_h$ corresponds to the finite element approximation of η within $\mathbb{V}_h$, we conclude, in view of estimate (4.5), that

$$\|\hat{y}_h - \tilde{y}_h\|_{L^\infty(\Omega)} \leq \|(\hat{y}_h - \tilde{y}_h) - \eta\|_{L^\infty(\Omega)} + \|\eta\|_{L^\infty(\Omega)} \lesssim (h^{2-\frac{d}{2}} + 1)\|u_1 - u_2\|_{L^2(\Omega)}.$$

Replace this bound into (5.17) and the obtained one into (5.16) to conclude

$$(5.18) \qquad |j_h'(u_1)v - j_h'(u_2)v| \lesssim (1 + \|\tilde{p}_h\|_{L^2(\Omega)} + \alpha)\|u_1 - u_2\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)}.$$

We finally observe that similar arguments to the ones used to derive (4.14) yield

$$\|\tilde{p}_h\|_{L^2(\Omega)} \lesssim \|u_2 - a(\cdot,0)\|_{L^2(\Omega)} + \sum_{t\in\mathcal{D}} |y_t| \lesssim C,$$

where $C > 0$. This concludes the proof. ◻

Inspired by [17, Lemma 7.5] and [15, Lemma 4.17], we now introduce a suitable auxiliary variable and provide an error estimate.

LEMMA 5.4 (error estimate for an auxiliary variable). *There exists $h_\star > 0$ such that for $h < h_\star$ there exists $u_h^* \in \mathbb{U}_{ad,h}$ satisfying $j'(\bar{u})(\bar{u} - u_h^*) = 0$ and*

$$\|\bar{u} - u_h^*\|_{L^2(\Omega)} \leq Ch \quad \forall h < h_\star, \quad C > 0.$$

Proof. Define, for each $T \in \mathscr{T}_h$, $I_T := \int_T \bar{\mathfrak{p}}(x)\mathrm{d}x$ and $u_h^* \in \mathbb{U}_h$ by

$$(5.19) \qquad u_h^*|_T := \frac{1}{I_T}\int_T \bar{\mathfrak{p}}(x)\bar{u}(x)\mathrm{d}x \text{ if } I_T \neq 0, \qquad u_h^*|_T := \frac{1}{|T|}\int_T \bar{u}(x)\mathrm{d}x \text{ if } I_T = 0.$$

We recall that $\bar{\mathfrak{p}} = \bar{p} + \alpha\bar{u}$. In view of the fact that $\bar{u} \in C^{0,1}(\bar{\Omega})$, which follows from Theorem 3.4, there exists $h_\star > 0$ such that

$$|\bar{u}(x_1) - \bar{u}(x_2)| \leq (\mathtt{b} - \mathtt{a})/2 \quad \forall h < h_\star \ \forall x_1, x_2 \in T.$$

This implies, in particular, that, for each $T \in \mathscr{T}_h$, $\bar{u}$ do not take both values $\mathtt{a}$ and $\mathtt{b}$ in T. Therefore, with (3.21) at hand, we deduce that, for a.e $x \in T$, $\bar{\mathfrak{p}}(x) \geq 0$ or $\bar{\mathfrak{p}}(x) \leq 0$. Consequently, we have that $I_T = 0$ if and only if $\bar{\mathfrak{p}}(x) = 0$ for a.e. $x \in T$, and that, if $I_T \neq 0$, $\bar{\mathfrak{p}}(x)/I_T \geq 0$ for a.e. $x \in T$. From this fact, and in view of the generalized mean value theorem, we conclude the existence of $x_T \in T$ such that $u_h^*|_T = \bar{u}(x_T)$. Since $u_h^* \in \mathbb{U}_h$, we have thus obtained that $u_h^* \in \mathbb{U}_{ad,h}$. Now, let $T \in \mathscr{T}_h$. We estimate $\|\bar{u} - u_h^*\|_{L^2(T)}$ as follows:

$$\|\bar{u} - u_h^*\|_{L^2(T)} \leq \|\bar{u} - \Pi_{L^2}\bar{u}\|_{L^2(T)} + \|\Pi_{L^2}\bar{u} - u_h^*\|_{L^2(T)} \lesssim h\|\nabla\bar{u}\|_{L^2(T)} + h^{\frac{d}{2}}\|\bar{u}\|_{L^\infty(T)}.$$

We finally observe that (5.19) immediately yields

$$j'(\bar{u})u_h^* = (\bar{\mathfrak{p}}, u_h^*)_{L^2(\Omega)} = \sum_{T\in\mathscr{T}_h} (\bar{\mathfrak{p}}, u_h^*)_{L^2(T)} = \sum_{T\in\mathscr{T}_h} (\bar{\mathfrak{p}}, \bar{u})_{L^2(T)} = j'(\bar{u})\bar{u}.$$

This concludes the proof. ∎

Proof of Theorem 5.1. Adding and subtracting the term $j_h'(\bar{u}_h)(\bar{u} - \bar{u}_h)$ in the right hand side of inequality (5.3) we obtain

$$\|\bar{u} - \bar{u}_h\|^2_{L^2(\Omega)} \lesssim [j'(\bar{u}) - j_h'(\bar{u}_h)](\bar{u} - \bar{u}_h) + [j_h'(\bar{u}_h) - j'(\bar{u}_h)](\bar{u} - \bar{u}_h)$$

for every $h < h_\dagger$. Invoke inequality (5.6) in conjunction with that fact that $\bar{u}, \bar{u}_h \in \mathbb{U}_{ad}$ to immediately arrive at

$$\|\bar{u} - \bar{u}_h\|^2_{L^2(\Omega)} \lesssim [j'(\bar{u}) - j_h'(\bar{u}_h)](\bar{u} - \bar{u}_h) + h^2|\log h|^2. \tag{5.20}$$

We now estimate $[j'(\bar{u}) - j_h'(\bar{u}_h)](\bar{u} - \bar{u}_h)$. To accomplish this task, we set $u = \bar{u}_h$ in (3.5) and $u_h = u_h^*$ in (4.20) to obtain

$$0 \leq j'(\bar{u})(\bar{u}_h - \bar{u}), \qquad 0 \leq j_h'(\bar{u}_h)(u_h^* - \bar{u}_h) = j_h'(\bar{u}_h)(u_h^* - \bar{u}) + j_h'(\bar{u}_h)(\bar{u} - \bar{u}_h).$$

Adding these inequalities we arrive at $[j'(\bar{u}) - j_h'(\bar{u}_h)](\bar{u} - \bar{u}_h) \leq j_h'(\bar{u}_h)(u_h^* - \bar{u})$. We utilize that u_h^* is such that $j'(\bar{u})(u_h^* - \bar{u}) = 0$, which follows from Lemma 5.4, to obtain

$$\begin{aligned}
[j'(\bar{u}) - j_h'(\bar{u}_h)](\bar{u} - \bar{u}_h) &\leq [j_h'(\bar{u}_h) - j'(\bar{u})](u_h^* - \bar{u}) \\
&= [j_h'(\bar{u}_h) - j_h'(\bar{u})](u_h^* - \bar{u}) + [j_h'(\bar{u}) - j'(\bar{u})](u_h^* - \bar{u}).
\end{aligned}$$

We thus apply estimates (5.6) and (5.7) to obtain

$$[j'(\bar{u}) - j_h'(\bar{u}_h)](\bar{u} - \bar{u}_h) \lesssim \|\bar{u}_h - \bar{u}\|_{L^2(\Omega)}\|u_h^* - \bar{u}\|_{L^2(\Omega)} + h^2|\log h|^2.$$

Invoke Young's inequality and the estimate of Lemma 5.4 to arrive at

$$[j'(\bar{u}) - j_h'(\bar{u}_h)](\bar{u} - \bar{u}_h) \leq \tfrac{1}{2}\|\bar{u}_h - \bar{u}\|^2_{L^2(\Omega)} + Ch^2(1 + |\log h|^2) \tag{5.21}$$

for every $h < h_\star$. Here, $C > 0$. Finally, replacing estimate (5.21) into (5.20) we conclude (5.1). This, which contradicts (5.2), concludes the proof. ∎

6. Numerical example. In this section we conduct a numerical experiment that illustrates the performance of the scheme of section 4.3 when is used to approximate the solution to (1.1)–(1.3). The numerical experiment has been carried out with the help of a code that was implemented using `C++`. All matrices have been assembled exactly and global linear systems were solved using the multifrontal massively parallel sparse direct solver (MUMPS) [2, 3]. The right hand sides and the approximation errors were computed by a quadrature formula which is exact for polynomials of degree nineteen (19).

For a given partition $\mathscr{T}_h$, we seek $(\bar{y}_h, \bar{p}_h, \bar{u}_h) \in \mathbb{V}_h \times \mathbb{V}_h \times \mathbb{U}_{ad,h}$ that solves the discrete optimality problem presented in section 4.3. This problem is solved by using a primal–dual active set strategy [32, section 2.12.4] combined with a fixed point strategy.

Example. We set $\Omega = (0,1)^2$, $a(\cdot, y) = y^3$, $\mathtt{b} = -\mathtt{a} = 10$, $\alpha = 0.1$,

$$\mathcal{D} = \{(0.25, 0.25), (0.75, 0.25), (0.75, 0.75), (0.25, 0.75)\},$$

and

$$y_{(0.25,0.25)} = 3, \quad y_{(0.75,0.25)} = -3, \quad y_{(0.75,0.75)} = 3, \quad y_{(0.25,0.75)} = -3.$$

In the absence of an exact solution, we calculate the error committed in the approximation of the optimal control variable, by taking as a reference solution the discrete optimal control obtained on a fine triangulation $\mathscr{T}_h$: the mesh $\mathscr{T}_h$ is such that $h \approx 9 \cdot 10^{-4}$. In Figure 6.1, we observe that an optimal experimental order of convergence, in terms of approximation, is attained: $\mathcal{O}(h)$.

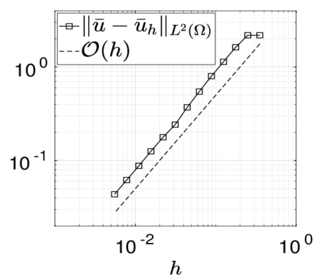

Fig. 6.1. *Experimental rate of convergence for the error* $\|\bar{u} - \bar{u}_h\|_{L^2(\Omega)}$.

REFERENCES

[1] R. A. Adams and J. J. F. Fournier, *Sobolev spaces*, vol. 140 of Pure and Applied Mathematics (Amsterdam), Elsevier/Academic Press, Amsterdam, second ed., 2003.

[2] P. Amestoy, I. Duff, and J.-Y. L'Excellent, *Multifrontal parallel distributed symmetric and unsymmetric solvers*, Computer Methods in Applied Mechanics and Engineering, 184 (2000), pp. 501 – 520, https://doi.org/10.1016/S0045-7825(99)00242-X.

[3] P. R. Amestoy, I. S. Duff, J.-Y. L'Excellent, and J. Koster, *A fully asynchronous multifrontal solver using distributed dynamic scheduling*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 15–41 (electronic), https://doi.org/10.1137/S0895479899358194.

[4] H. Antil, E. Otárola, and A. J. Salgado, *Some applications of weighted norm inequalities to the error analysis of PDE-constrained optimization problems*, IMA J. Numer. Anal., 38 (2018), pp. 852–883, https://doi.org/10.1093/imanum/drx018.

[5] N. Arada, E. Casas, and F. Tröltzsch, *Error estimates for the numerical approximation of a semilinear elliptic control problem*, Comput. Optim. Appl., 23 (2002), pp. 201–229, https://doi.org/10.1023/A:1020576801966.

[6] N. Behringer, *Improved error estimates for optimal control of the stokes problem with pointwise tracking in three dimensions*, Math. Control Relat. Fields, (2020), https://doi.org/10.3934/mcrf.2020038.

[7] N. Behringer, D. Meidner, and B. Vexler, *Finite element error estimates for optimal control problems with pointwise tracking*, Pure Appl. Funct. Anal., 4 (2019), pp. 177–204.

[8] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, vol. 15 of Texts in Applied Mathematics, Springer, New York, third ed., 2008, https://doi.org/10.1007/978-0-387-75934-0.

[9] C. Brett, A. Dedner, and C. Elliott, *Optimal control of elliptic PDEs at points*, IMA J. Numer. Anal., 36 (2016), pp. 1015–1050, https://doi.org/10.1093/imanum/drv040.

[10] E. Casas, *L^2 estimates for the finite element method for the Dirichlet problem with singular data*, Numer. Math., 47 (1985), pp. 627–632, https://doi.org/10.1007/BF01389461.

[11] E. Casas, *Using piecewise linear functions in the numerical approximation of semilinear elliptic control problems*, Adv. Comput. Math., 26 (2007), pp. 137–153, https://doi.org/10.1007/s10444-004-4142-0.

[12] E. Casas, *Necessary and sufficient optimality conditions for elliptic control problems with finitely many pointwise state constraints*, ESAIM Control Optim. Calc. Var., 14 (2008), pp. 575–589, https://doi.org/10.1051/cocv:2007063.

[13] E. Casas and M. Mateos, *Uniform convergence of the FEM. Applications to state constrained control problems*, vol. 21, 2002, pp. 67–100. Special issue in memory of Jacques-Louis Lions.

[14] E. Casas and M. Mateos, *Optimal control of partial differential equations*, in Computational mathematics, numerical analysis and applications, vol. 13 of SEMA SIMAI Springer Ser., Springer, Cham, 2017, pp. 3–59.

[15] E. Casas, M. Mateos, and J.-P. Raymond, *Error estimates for the numerical approximation of a distributed control problem for the steady-state Navier-Stokes equations*, SIAM J. Control Optim., 46 (2007), pp. 952–982, https://doi.org/10.1137/060649999.

[16] E. Casas, M. Mateos, and B. Vexler, *New regularity results and improved error estimates for optimal control problems with state constraints*, ESAIM Control Optim. Calc. Var., 20 (2014), pp. 803–822, https://doi.org/10.1051/cocv/2013084.

[17] E. Casas and J.-P. Raymond, *Error estimates for the numerical approximation of Dirichlet boundary control for semilinear elliptic equations*, SIAM J. Control Optim., 45 (2006), pp. 1586–1611, https://doi.org/10.1137/050626600.

[18] E. Casas and F. Tröltzsch, *Recent advances in the analysis of pointwise state-constrained elliptic optimal control problems*, ESAIM Control Optim. Calc. Var., 16 (2010), pp. 581–600, https://doi.org/10.1051/cocv/2009010.

[19] L. Chang, W. Gong, and N. Yan, *Numerical analysis for the approximation of optimal control problems with pointwise observations*, Math. Methods Appl. Sci., 38 (2015), pp. 4502–4520, https://doi.org/10.1002/mma.2861.

[20] P. G. Ciarlet, *The finite element method for elliptic problems*, SIAM, Philadelphia, PA, 2002, https://doi.org/10.1137/1.9780898719208.

[21] A. Ern and J.-L. Guermond, *Theory and practice of finite elements*, vol. 159 of Applied Mathematical Sciences, Springer-Verlag, New York, 2004, https://doi.org/10.1007/978-1-4757-4355-5.

[22] J. Frehse and R. Rannacher, *Eine L^1-Fehlerabschätzung für diskrete Grundlösungen in der Methode der finiten Elemente*, in Finite Elemente (Tagung, Univ. Bonn, Bonn, 1975), 1976, pp. 92–114. Bonn. Math. Schrift., No. 89.

[23] F. Fuica, E. Otárola, and D. Quero, *Error estimates for optimal control problems involving the Stokes system and Dirac measures*, Appl. Math. Optim., (2020), https://doi.org/10.1007/s00245-020-09693-0.

[24] P. Grisvard, *Elliptic problems in nonsmooth domains*, vol. 24 of Monographs and Studies in Mathematics, Pitman (Advanced Publishing Program), Boston, MA, 1985.

[25] D. Kinderlehrer and G. Stampacchia, *An introduction to variational inequalities and their applications*, vol. 31 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000, https://doi.org/10.1137/1.9780898719451. Reprint of the 1980 original.

[26] V. Maz'ya and J. Rossmann, *Elliptic equations in polyhedral domains*, vol. 162 of Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 2010, https://doi.org/10.1090/surv/162.

[27] I. Neitzel, J. Pfefferer, and A. Rösch, *Finite element discretization of state-constrained elliptic optimal control problems with semilinear state equation*, SIAM J. Control Optim., 53 (2015), pp. 874–904, https://doi.org/10.1137/140960645.

[28] T. Roubíček, *Nonlinear partial differential equations with applications*, vol. 153 of International Series of Numerical Mathematics, Birkhäuser/Springer Basel AG, Basel, second ed., 2013, https://doi.org/10.1007/978-3-0348-0513-1.

[29] A. H. Schatz and L. B. Wahlbin, *Interior maximum norm estimates for finite element methods*, Math. Comp., 31 (1977), pp. 414–442, https://doi.org/10.2307/2006424.

[30] A. H. Schatz and L. B. Wahlbin, *On the quasi-optimality in L_∞ of the $\dot{H}^1$-projection into finite element spaces*, Math. Comp., 38 (1982), pp. 1–22, https://doi.org/10.2307/2007461.

[31] G. Stampacchia, *Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus*, Ann. Inst. Fourier (Grenoble), 15 (1965), pp. 189–258, http://www.numdam.org/item?id=AIF_1965__15_1_189_0.

[32] F. Tröltzsch, *Optimal control of partial differential equations*, vol. 112 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2010, https://doi.org/10.1090/gsm/112. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.

[33] E. Zeidler, *Nonlinear functional analysis and its applications. II/B*, Springer-Verlag, New York, 1990, https://doi.org/10.1007/978-1-4612-0985-0. Nonlinear monotone operators, Translated from the German by the author and Leo F. Boron.