

FRACTIONAL, SEMILINEAR, AND SPARSE OPTIMAL CONTROL: A PRIORI ERROR BOUNDS*

FRANCISCO BERSETCHE[†], FRANCISCO FUICA[‡], ENRIQUE OTÁROLA[§], AND DANIEL QUERO[¶]

Abstract. In this work, we use the integral definition of the fractional Laplace operator and study a sparse optimal control problem involving a fractional, semilinear, and elliptic partial differential equation as state equation; control constraints are also considered. We establish the existence of optimal solutions and first and second order optimality conditions. We also analyze regularity properties for optimal variables. We propose and analyze two finite element strategies of discretization: a fully discrete scheme, where the control variable is discretized with piecewise constant functions, and a semidiscrete scheme, where the control variable is not discretized. For both discretization schemes, we analyze convergence properties and a priori error bounds.

Key words. sparse optimal control, fractional diffusion, first and second order optimality conditions, regularity estimates, finite elements, a priori error bounds.

AMS subject classifications. 35R11, 49J20, 49K20 49M25, 65K10, 65N15 65N30.

1. Introduction. Let $d \in \mathbb{N}$ be such that $d \geq 2$. Let $\Omega \subset \mathbb{R}^d$ be an open and bounded domain with Lipschitz boundary $\partial\Omega$. Define the cost functional

$$(1.1) \quad J(u, q) := \int_{\Omega} L(x, u) dx + \frac{\lambda}{2} \|q\|_{L^2(\Omega)}^2 + \mu \|q\|_{L^1(\Omega)},$$

where $L : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ denotes an appropriate Caratheodory function, $\lambda > 0$ corresponds to the so-called *regularization* parameter, and $\mu > 0$ denotes a *sparsity* parameter. The necessary assumptions on L are deferred to the section 2.4. In this paper, we are interested in the analysis and discretization of the following *nonconvex* and *nondifferentiable* optimal control problem for a *fractional, semilinear, and elliptic* partial differential equation (PDE): Find $\min J(u, q)$ subject to the *state equation*

$$(1.2) \quad (-\Delta)^s u + a(\cdot, u) = q \text{ in } \Omega, \quad u = 0 \text{ in } \Omega^c,$$

where $\Omega^c = \mathbb{R}^d \setminus \Omega$ and $s \in (0, 1)$, and the *control constraints*

$$(1.3) \quad q \in \mathbb{Q}_{ad}, \quad \mathbb{Q}_{ad} := \{v \in L^2(\Omega) : \alpha \leq v(x) \leq \beta \text{ a.e. } x \in \Omega\}.$$

We adopt the integral definition of the fractional Laplace operator $(-\Delta)^s$. For a smooth function $w : \mathbb{R}^d \rightarrow \mathbb{R}$, the operator $(-\Delta)^s$ is defined as follows:

$$(1.4) \quad (-\Delta)^s w(x) := C(d, s) \text{p.v.} \int_{\mathbb{R}^d} \frac{w(x) - w(y)}{|x - y|^{d+2s}} dy, \quad C(d, s) := \frac{2^{2s} s \Gamma(s + \frac{d}{2})}{\pi^{\frac{d}{2}} \Gamma(1 - s)},$$

*FB is supported by ANID through FONDECYT grant 3220254. FF is supported by ANID through FONDECYT grant 3230126. EO is partially supported by ANID through FONDECYT grant 1220156. DQ is supported by UTFSM through Programa de Incentivo a la Investigación Científica (PIIC) and by ANID/Subdirección del Capital Humano/Doctorado Nacional/2021–21210988.

[†]Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile. francisco.bersetche@usm.cl

[‡]Facultad de Matemáticas, Pontificia Universidad Católica de Chile, Santiago, Chile. francisco.fuica@mat.uc.cl

[§]Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile. enrique.otarola@usm.cl

[¶]Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso, Chile. daniel.quero@alumnos.usm.cl

where p.v. stands for the *Cauchy principal value* and $C(d, s)$ is a normalization constant. The necessary assumptions on the function a are deferred until section 2.4. The control bounds α and β are such that $\alpha < 0 < \beta$; see [6, Remark 2.1] for a discussion.

The optimal control problem under consideration involves a cost functional J that contains the $L^1(\Omega)$ -norm of the control variable. The study of this type of optimal control problems is mainly motivated by the following two observations: First, $\|\cdot\|_{L^1(\Omega)}$ is a natural measure of the control cost. Second, $\|\cdot\|_{L^1(\Omega)}$ leads to sparsely supported optimal controls, i.e., optimal controls that are non-zero only in a small region of the considered domain. This is a desirable property in applications, for example in the optimal placement of discrete actuators [19]. From a mathematical point of view, the analysis and discretization of the optimal control problem considered here is anything but trivial and interesting, especially due to the following considerations:

1. *Fractional diffusion*: The efficient approximation of problems involving the *integral* fractional Laplacian carries two main difficulties. The first and most important is that $(-\Delta)^s$ is a non-local operator [3, 9]. The second is the lack of boundary regularity, which leads to reduced convergence rates [3, 4].
2. *Non-linearity/Non-convexity*: Since the state equation is a *semilinear* PDE, the control problem is non-convex. Consequently, first order optimality conditions are necessary conditions for local optimality; sufficiency requires the investigation of second order optimality conditions [6, 21].
3. *Non-differentiability*: Due to the presence of $\|\cdot\|_{L^1(\Omega)}$ in the cost functional J , the optimal control problem becomes non-differentiable ($\alpha < 0 < \beta$). This leads to some difficulties that do not occur in the differential case $\lambda > 0$ and $\mu = 0$ [21], especially when analysing second order optimality conditions [6] and investigating finite element techniques [6, 22].

For the special case $s = 1$, there are several papers in the literature that provide error estimates for finite element discretizations of control problems related to ours. As far as we know, the first paper is [22], in which the authors consider a linear PDE and propose several finite element strategies to discretize the admissible control set. For all strategies considered, the authors obtain bounds for the error that occurs when approximating the optimal control variable in $L^2(\Omega)$. The semilinear scenario was later developed in [6]. In this paper, the authors develop two strategies for discretization: one based on the variational discretization approach and another where the admissible control set is discretized with piecewise constant functions. Based on a complete study of second order optimality conditions, the authors obtain error estimates in $L^\infty(\Omega)$ for the error that occurs when approximating all the optimal variables involved. We would also like to mention [5] for the piecewise linear approximation of the admissible control set and [2, 22] for a posteriori error analyses.

For the non-local and linear scenario, $s \in (0, 1)$ and $a \equiv 0$, there are several papers in which finite element strategies are analyzed. We refer the interested reader to [8] for the analysis of a priori error bounds for the differentiable case $\lambda > 0$ and $\mu = 0$ and to the more recent work [23] for the analysis of a posteriori error bounds in the non-differentiable scenario. We also mention [14, 17], where the authors consider the linear version of (1.1)–(1.3), but with the *spectral* definition of the fractional Laplacian. It is important to mention that the treatment of discretizations of problems involving the spectral and integral definitions of the fractional Laplacian is fundamentally different due to regularity and discretization properties. As far as we know, this work is the *first* to provide a complete analysis for the semilinear control problem (1.1)–(1.3), which also includes the development and analysis of finite element strategies.

The structure of this article is as follows. In section 2 we introduce the notation,

the functional framework and the assumptions that we will use in our work. In section 3 we give an overview of (1.2) and its discretization by finite elements. In section 4 we present a weak formulation of (1.1)–(1.3), analyze existence results, and derive first and second order optimality conditions; furthermore, regularity properties are also analyzed. In section 5, we introduce a fully discrete method and provide convergence properties and error bounds. In section 6, we develop a semidiscretization scheme and derive error bounds. We conclude our work with section 7 in which we perform a numerical experiment that illustrates the performance of the proposed methods.

2. Notation and preliminary remarks. Let us set the notation and recall some facts that will be useful later.

2.1. Notation. We denote by Ω^c the complement of Ω . For normed spaces \mathcal{X} and \mathcal{Y} , we write $\mathcal{X} \hookrightarrow \mathcal{Y}$ to denote that \mathcal{X} is continuously embedded in \mathcal{Y} . We denote by \mathcal{X}' and $\|\cdot\|_{\mathcal{X}'}$ the dual and the norm of \mathcal{X} , respectively. The duality pairing between \mathcal{X} and \mathcal{X}' is denoted by $\langle \cdot, \cdot \rangle_{\mathcal{X}, \mathcal{X}'}$. When the spaces \mathcal{X} and \mathcal{X}' are clear from the context, we simply write $\langle \cdot, \cdot \rangle$. Let $\{x_n\}_{n=1}^\infty$ be a sequence in \mathcal{X} . We denote by $x_n \rightarrow x$, $x_n \rightharpoonup x$, and $x_n \overset{*}{\rightharpoonup} x$ the strong, weak, and weak* convergence, respectively, of $\{x_n\}_{n=1}^\infty$ to x in \mathcal{X} as $n \uparrow \infty$. Finally, $\mathbf{a} \lesssim \mathbf{b}$ indicates that $\mathbf{a} \leq C\mathbf{b}$, where C is a positive constant that does not depend on either \mathbf{a} or \mathbf{b} . The value of C might change at each occurrence.

2.2. Subgradients and subdifferentials. We denote $\mathbb{R} \cup \{+\infty\}$ by \mathbb{R}_∞ . Let $j : Z \rightarrow \mathbb{R}_\infty$ be a given function, where Z is a real normed space, and let z be a point in $\text{Dom } j$, i.e., $z \in Z$ is such that $j(z) < \infty$. An element $\zeta \in Z'$ is called a *subgradient* of j at z if it satisfies the following *subgradient* inequality [7, Chapter 4.1]:

$$(2.1) \quad j(y) - j(z) \geq \langle \zeta, y - z \rangle_{Z', Z} \quad \forall y \in Z.$$

The set of all subgradients of j at z is denoted by $\partial j(z)$ and called the *subdifferential* of j at z . Of particular interest is the case where $Z = L^1(\Omega)$ and $j : Z \rightarrow \mathbb{R}_0^+$ is defined by $j(z) = \|z\|_{L^1(\Omega)}$. In this scenario, it follows that $\zeta \in \partial j(z)$ if and only if

$$\zeta(x) = 1 \text{ if } z(x) > 0, \quad \zeta(x) = -1 \text{ if } z(x) < 0, \quad \zeta(x) \in [-1, 1] \text{ if } z(x) = 0,$$

for a.e. $x \in \Omega$ [18, Proposition 4.6.2]. For $z, v \in L^1(\Omega)$, the directional derivative of j at z in the direction v is given by

$$(2.2) \quad j'(z; v) = \lim_{\rho \rightarrow 0} \frac{j(z + \rho v) - j(z)}{\rho} = \int_{\Omega_z^+} v dx - \int_{\Omega_z^-} v dx + \int_{\Omega_z^0} |v| dx,$$

where Ω_z^+, Ω_z^- , and Ω_z^0 denote the sets of points in Ω where z is positive, negative, and zero, respectively.

Let $M \subseteq Z$ be nonempty. The functional $\mathbb{I}_M : Z \rightarrow \mathbb{R}_\infty$ defined by $\mathbb{I}_M(z) = 0$ if $z \in M$ and $\mathbb{I}_M(z) = +\infty$ if $z \in Z \setminus M$ is called the *indicator functional* of M . We note that \mathbb{I}_M is proper and convex if and only if M is nonempty and convex. Let $z \in M$. It follows from (2.1) that $\zeta \in \partial \mathbb{I}_M(z)$ if and only if $\langle \zeta, y - z \rangle_{Z', Z} \leq 0$ for all $y \in M$.

2.3. Function spaces. Let $s \geq 0$ and let $\mathbb{R}^d \ni \xi \mapsto \iota(\xi) = (1 + |\xi|^2)^{\frac{s}{2}} \in \mathbb{R}$. With \mathcal{F} we denote the Fourier transform. We define the fractional Sobolev space $H^s(\mathbb{R}^d) := \{v \in L^2(\mathbb{R}^d) : \iota \mathcal{F}(v) \in L^2(\mathbb{R}^d)\}$, which is endowed with the norm $\|v\|_{H^s(\mathbb{R}^d)} := \|\iota \mathcal{F}(v)\|_{L^2(\mathbb{R}^d)}$; see [20, Definition 15.7] and [11, Chapter 1, Section 7].

We define $\tilde{H}^s(\Omega)$ as the closure of $C_0^\infty(\Omega)$ in $H^s(\mathbb{R}^d)$. According to [12, Theorem 3.29], we have the characterization $\tilde{H}^s(\Omega) = \{v \in H^s(\mathbb{R}^d) : \text{supp } v \subset \overline{\Omega}\}$. We endow the fractional Sobolev space $\tilde{H}^s(\Omega)$ with the following inner product and norm:

$$(u, v)_{\tilde{H}^s(\Omega)} := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{(u(x) - u(y))(v(x) - v(y))}{|x - y|^{d+2s}} dx dy, \quad \|v\|_{\tilde{H}^s(\Omega)} := (v, v)_{\tilde{H}^s(\Omega)}^{\frac{1}{2}}.$$

We denote by $H^{-s}(\Omega)$ the dual space of $\tilde{H}^s(\Omega)$. Finally, we introduce

$$\mathcal{A} : \tilde{H}^s(\Omega) \times \tilde{H}^s(\Omega) \rightarrow \mathbb{R}, \quad \mathcal{A}(u, v) = 2^{-1}C(d, s)(u, v)_{\tilde{H}^s(\Omega)}, \quad \|v\|_s := \mathcal{A}(v, v)^{\frac{1}{2}}.$$

We will use the following continuous and compact embedding results repeatedly. Let $s \in (0, 1)$. If $\mathfrak{r} \in [1, 2d/(d - 2s)]$, then $H^s(\Omega) \hookrightarrow L^\mathfrak{r}(\Omega)$ [1, Theorem 7.34]. If $\mathfrak{r} \in [1, 2d/(d - 2s))$, then the embedding is compact [9, Corollary 7.2].

2.4. Assumptions. We will operate under the following assumptions on a and L . However, we must mention right away that some of the results obtained in this paper are also valid under less restrictive conditions; when possible, we explicitly mention the assumptions on a and L that are required to obtain a particular result.

- (A.1) $a : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second variable and $a(\cdot, 0) \in L^2(\Omega) \cap L^r(\Omega)$ for $r > d/2s$.
- (A.2) $\frac{\partial a}{\partial u}(x, u) \geq 0$ for a.e. $x \in \Omega$ and for all $u \in \mathbb{R}$.
- (A.3) For all $\mathfrak{m} > 0$, there exists $C_{\mathfrak{m}} > 0$ such that

$$\sum_{i=1}^2 \left| \frac{\partial^i a}{\partial u^i}(x, u) \right| \leq C_{\mathfrak{m}}, \quad \left| \frac{\partial^2 a}{\partial u^2}(x, u) - \frac{\partial^2 a}{\partial u^2}(x, v) \right| \leq C_{\mathfrak{m}}|u - v|$$

for a.e. $x \in \Omega$ and u, v such that $|u|, |v| \leq \mathfrak{m}$.

We note that it follows directly from (A.3) and the mean value theorem that a and $\frac{\partial a}{\partial u}$ are locally Lipschitz with respect to the second variable.

- (B.1) $L : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ is a Carathéodory function of class C^2 with respect to the second variable and $L(\cdot, 0) \in L^1(\Omega)$.
- (B.2) For all $\mathfrak{m} > 0$, there exist $\psi_{\mathfrak{m}}, \phi_{\mathfrak{m}} \in L^r(\Omega)$ with $r > d/2s$ such that

$$\left| \frac{\partial L}{\partial u}(x, u) \right| \leq \psi_{\mathfrak{m}}(x), \quad \left| \frac{\partial^2 L}{\partial u^2}(x, u) \right| \leq \phi_{\mathfrak{m}}(x)$$

for a.e. $x \in \Omega$ and u such that $|u| \leq \mathfrak{m}$.

The following assumption is necessary to obtain further regularity properties for optimal control variables and to derive error estimates.

- (C.1) For all $\mathfrak{m} > 0$ and $u \in [-\mathfrak{m}, \mathfrak{m}]$, $\frac{\partial L}{\partial u}(\cdot, u) \in L^2(\Omega)$ and $\frac{\partial^2 L}{\partial u^2}(\cdot, u) \in L^{\frac{d}{s}}(\Omega)$.

3. Fractional semilinear PDEs. Let $s \in (0, 1)$ and let $q \in L^r(\Omega)$ with $r > d/2s$. We introduce the following weak formulation for the *fractional, semilinear, and elliptic* PDE (1.2): Find $u \in \tilde{H}^s(\Omega)$ such that

$$(3.1) \quad \mathcal{A}(u, v) + \int_{\Omega} a(x, u)v dx = \int_{\Omega} qv dx \quad \forall v \in \tilde{H}^s(\Omega).$$

Here, $a = a(x, u) : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ denotes a Carathéodory function which is monotone increasing in u . We assume that for every $\mathfrak{m} > 0$ there exists $\varphi_{\mathfrak{m}} \in L^t(\Omega)$ such that

$$(3.2) \quad |a(x, u)| \leq \varphi_{\mathfrak{m}}(x) \text{ a.e. } x \in \Omega, \quad u \in [-\mathfrak{m}, \mathfrak{m}], \quad \mathfrak{t} = 2d/(d + 2s).$$

If, in addition, $a(\cdot, 0) \in L^r(\Omega)$, then (3.1) admits a unique solution $u \in \tilde{H}^s(\Omega) \cap L^\infty(\Omega)$ satisfying the stability bound $\|u\|_s + \|u\|_{L^\infty(\Omega)} \lesssim \|q - a(\cdot, 0)\|_{L^r(\Omega)}$ [15, Theorem 3.1].

THEOREM 3.1 (Sobolev regularity). *Let $s \in (0, 1)$ and let $q \in L^2(\Omega) \cap L^r(\Omega)$. If $a(\cdot, 0) \in L^2(\Omega)$ and a is locally Lipschitz with respect to the second variable, then the solution u of problem (3.1) belongs to $H^{s+\kappa-\varepsilon}(\Omega)$ for all $0 < \varepsilon < s$, where $\kappa = \frac{1}{2}$ for $\frac{1}{2} < s < 1$ and $\kappa = s - \varepsilon$ for $0 < s \leq \frac{1}{2}$. Moreover, we have the bound*

$$\|u\|_{H^{s+\kappa-\varepsilon}(\Omega)} \lesssim C\varepsilon^{-\nu} \|q - a(\cdot, 0)\|_{L^2(\Omega)} \quad \forall \varepsilon \in (0, s),$$

where $\nu = \frac{1}{2}$ for $\frac{1}{2} < s < 1$ and $\nu = \frac{1}{2} + \nu_0$ for $0 < s \leq \frac{1}{2}$. Here, ν_0 and C denote positive constants that depend on Ω and d and Ω , d , and s , respectively.

Proof. The proof follows from a direct application of [4, Theorem 2.1 and inequality (2.6)] using the fact that a is locally Lipschitz with respect to the second variable so that $\|q - a(\cdot, u)\|_{L^2(\Omega)} \lesssim \|q - a(\cdot, 0)\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)} \lesssim \|q - a(\cdot, 0)\|_{L^2(\Omega)}$. \square

3.1. Finite element discretization. We now present a finite element approximation of problem (3.1) under the additional assumption that Ω is a Lipschitz *polytope*. Let $\{\mathcal{T}_h\}_{h>0}$ be a collection of conforming and quasi-uniform meshes \mathcal{T}_h made of closed simplices T , where $h = \max\{h_T : T \in \mathcal{T}_h\}$ and $h_T = \text{diam}(T)$. For each mesh \mathcal{T}_h , we introduce the following standard finite element space:

$$(3.3) \quad \mathbb{V}_h := \{v_h \in C(\bar{\Omega}) : v_h|_T \in \mathbb{P}_1(T) \ \forall T \in \mathcal{T}_h, v_h = 0 \text{ on } \partial\Omega\}.$$

The discrete approximation of (3.1) is as follows: Find $\mathbf{u}_h \in \mathbb{V}_h$ such that

$$(3.4) \quad \mathcal{A}(\mathbf{u}_h, v_h) + \int_{\Omega} a(x, \mathbf{u}_h) v_h dx = \int_{\Omega} q v_h dx \quad \forall v_h \in \mathbb{V}_h.$$

The existence of a discrete solution follows from Brouwer's fixed point theorem; uniqueness follows from the monotonicity of a . Moreover, we have $\|\mathbf{u}_h\|_s \lesssim \|q\|_{H^{-s}(\Omega)}$.

We now state a priori error estimates. For this purpose, we will assume that

$$(3.5) \quad |a(x, u) - a(x, v)| \leq |\phi(x)| |u - v|$$

for a.e. $x \in \Omega$ and $u, v \in \mathbb{R}$. The function ϕ belongs to $L^\eta(\Omega)$, where $\eta = d/2s$.

THEOREM 3.2 (a priori error estimates). *Let $s \in (0, 1)$ and let $q \in L^r(\Omega)$ with $r > d/2s$. Let $u \in \dot{H}^s(\Omega)$ be the solution to (3.1) and let $\mathbf{u}_h \in \mathbb{V}_h$ be its finite element approximation obtained as the solution to (3.4). If a satisfies (3.5), then we have*

$$(3.6) \quad \|u - \mathbf{u}_h\|_s \lesssim \|u - v_h\|_s \quad \forall v_h \in \mathbb{V}_h.$$

If, in addition, $q \in L^2(\Omega)$, a is locally Lipschitz with respect to the second variable, and $a(\cdot, 0) \in L^2(\Omega)$, then we have the error bound

$$(3.7) \quad \|u - \mathbf{u}_h\|_s \lesssim h^\gamma |\log h|^\varphi \|q - a(\cdot, 0)\|_{L^2(\Omega)}, \quad \gamma = \min\{s, \frac{1}{2}\},$$

If, in addition, a satisfies the condition (3.5) with $\eta = d/s$, then we have

$$(3.8) \quad \|u - \mathbf{u}_h\|_{L^2(\Omega)} \lesssim h^{2\gamma} |\log h|^{2\varphi} \|q - a(\cdot, 0)\|_{L^2(\Omega)}, \quad \gamma = \min\{s, \frac{1}{2}\}.$$

Here, $\varphi = \nu$ if $s \neq \frac{1}{2}$, $\varphi = 1 + \nu$ if $s = \frac{1}{2}$, and $\nu \geq \frac{1}{2}$ is the constant in Theorem 3.1.

Proof. The proof of (3.6) can be found in [15, Theorem 5.2]. The error estimates (3.7) and (3.8) can be found in [16, Theorem 5.1] and [16, Theorem 5.2]. \square

We conclude this section with the following convergence result.

LEMMA 3.3 (convergence). *Let $s \in (0, 1)$ and let $u_h \in \mathbb{V}_h$ be the solution to*

$$\mathcal{A}(u_h, v_h) + \int_{\Omega} a(x, u_h) v_h dx = \int_{\Omega} q_h v_h dx \quad \forall v_h \in \mathbb{V}_h,$$

where $q_h \in L^r(\Omega)$ with $r > d/2s$. If a satisfies the condition (3.5), then we have $q_h \rightharpoonup q$ in $L^r(\Omega) \implies u_h \rightarrow u$ in $L^r(\Omega)$ as $h \rightarrow 0$. Here, $\tau \leq 2d/(d-2s)$.

Proof. See [15, Proposition 5.3] for a proof. We note that in the statement of [15, Proposition 5.3] it is assumed that $\Omega \in C^2$. However, this assumption does not play any role and the same proof can be performed if Ω is a Lipschitz polytope. \square

4. The optimal control problem. In this section, we present the following weak formulation for the optimal control problem introduced in section 1: Find

$$(4.1) \quad \min\{J(u, q) : (u, q) \in \tilde{H}^s(\Omega) \times \mathbb{Q}_{ad}\}$$

subject to the *fractional semilinear*, and *elliptic* state equation

$$(4.2) \quad \mathcal{A}(u, v) + \int_{\Omega} a(x, u) v dx = \int_{\Omega} q v dx \quad \forall v \in \tilde{H}^s(\Omega).$$

Here, $a = a(x, u) : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ is a monotonically increasing in u Carathéodory function that satisfies (3.2) and $a(\cdot, 0) \in L^r(\Omega)$ with $r > d/2s$. As explained in section 3, problem (4.2) is well-posed under these assumptions on a . We therefore introduce the *control to state map* $\mathcal{S} : L^r(\Omega) \rightarrow \tilde{H}^s(\Omega) \cap L^\infty(\Omega)$ which, given a control q , associates to it the unique state u that solves (4.2).

4.1. Existence of an optimal solution. The existence of an optimal solution $(\bar{u}, \bar{q}) \in \tilde{H}^s(\Omega) \times \mathbb{Q}_{ad}$ is as follows.

THEOREM 4.1 (existence of an optimal solution). *Let $s \in (0, 1)$. Let $L = L(x, u) : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ be a Carathéodory function. Assume that, for every $\mathbf{m} > 0$, there exists $\varphi_{\mathbf{m}} \in L^r(\Omega)$ with $r > d/2s$ and $\sigma_{\mathbf{m}} \in L^1(\Omega)$ such that*

$$(4.3) \quad |a(x, u)| \leq \varphi_{\mathbf{m}}(x), \quad |L(x, u)| \leq \sigma_{\mathbf{m}}(x) \quad \text{a.e. } x \in \Omega, \quad u \in [-\mathbf{m}, \mathbf{m}].$$

Thus, (4.1)–(4.2) admits at least one solution $(\bar{u}, \bar{q}) \in \tilde{H}^s(\Omega) \cap L^\infty(\Omega) \times \mathbb{Q}_{ad}$.

Proof. Let $\{(u_k, q_k)\}_{k \in \mathbb{N}}$ be a minimizing sequence, i.e., for $k \in \mathbb{N}$, $q_k \in \mathbb{Q}_{ad}$ and $u_k = \mathcal{S}q_k \in \tilde{H}^s(\Omega)$ are such that $J(u_k, q_k) \rightarrow j := \inf\{J(\mathcal{S}q, q) : q \in \mathbb{Q}_{ad}\}$ as $k \uparrow \infty$. The arguments in the proof of [15, Theorem 4.1] show that, up to a nonrelabeled subsequence, $q_k \xrightarrow{*} \bar{q}$ in $L^\infty(\Omega)$ and $u_k \rightharpoonup \bar{u}$ in $\tilde{H}^s(\Omega)$ as $k \uparrow \infty$, where $\bar{u} = \mathcal{S}\bar{q}$. On the other hand, the Lebesgue dominated convergence theorem combined with (4.3) and the fact that $u_k \rightarrow \bar{u}$ in $L^\tau(\Omega)$ for every $\tau < 2d/(d-2s)$ show that $|\int_{\Omega} (L(x, u_k(x)) - L(x, \bar{u})) dx| \rightarrow 0$ as $k \uparrow \infty$. Since $\|\cdot\|_{L^1(\Omega)}$ and the square of $\|\cdot\|_{L^2(\Omega)}$ are continuous and convex in $L^1(\Omega)$ and $L^2(\Omega)$, respectively, we can arrive at $J(\bar{u}, \bar{q}) \leq j$. This completes the proof. \square

4.2. First order necessary optimality conditions. In this section, we develop necessary first order optimality conditions for (4.1)–(4.2). Since this problem is nonconvex, we discuss optimality conditions in the context of local solutions: We say that $\bar{q} \in \mathbb{Q}_{ad}$ is a local solution in $L^2(\Omega)$ for (4.1)–(4.2) if there exists $\varepsilon > 0$ such that

$$(4.4) \quad J(\mathcal{S}\bar{q}, \bar{q}) \leq J(\mathcal{S}q, q) \quad \forall q \in \mathbb{Q}_{ad} : \|q - \bar{q}\|_{L^2(\Omega)} \leq \varepsilon.$$

A solution \bar{q} is called *strict* if \bar{q} is the unique solution satisfying (4.4) for some $\varepsilon > 0$.

We now introduce $F : L^2(\Omega) \rightarrow \mathbb{R}$ and $j : L^1(\Omega) \rightarrow \mathbb{R}$ by

$$(4.5) \quad F(q) = \int_{\Omega} L(x, \mathcal{S}q) dx + \frac{\lambda}{2} \|q\|_{L^2(\Omega)}^2, \quad j(q) = \|q\|_{L^1(\Omega)}.$$

We also introduce the *reduced cost functional* $j : \mathbb{Q}_{ad} \rightarrow \mathbb{R}$ by $j(q) = F(q) + \mu j(q)$.

In what follows, we discuss differentiability properties for \mathcal{S} and F .

PROPOSITION 4.2 (differentiability properties of \mathcal{S}). *Let $s \in (0, 1)$ and let $r > d/2s$. Assume that (A.1)–(A.3) hold. Then, the control to state map $\mathcal{S} : L^r(\Omega) \rightarrow \tilde{H}^s(\Omega) \cap L^\infty(\Omega)$ is of class C^2 . In addition, if $q, w \in L^r(\Omega)$, then $\phi = \mathcal{S}'(q)w \in \tilde{H}^s(\Omega) \cap L^\infty(\Omega)$ corresponds to the unique solution to the problem*

$$(4.6) \quad \mathcal{A}(\phi, v) + \int_{\Omega} \frac{\partial a}{\partial u}(x, u) \phi v dx = \int_{\Omega} w v dx \quad \forall v \in \tilde{H}^s(\Omega),$$

where $u = \mathcal{S}q$. If $w_1, w_2 \in L^r(\Omega)$, then $\psi = \mathcal{S}''(q)(w_1, w_2) \in \tilde{H}^s(\Omega) \cap L^\infty(\Omega)$ corresponds to the unique solution to

$$(4.7) \quad \mathcal{A}(\psi, v) + \int_{\Omega} \frac{\partial a}{\partial u}(x, u) \psi v dx = - \int_{\Omega} \frac{\partial^2 a}{\partial u^2}(x, u) \phi_{w_1} \phi_{w_2} v dx \quad \forall v \in \tilde{H}^s(\Omega),$$

where $u = \mathcal{S}q$ and $\phi_{w_i} = \mathcal{S}'(q)w_i$, with $i \in \{1, 2\}$.

Proof. See [15, Theorem 4.3] for a proof. \square

To present the following result, we introduce the *adjoint state* p , which corresponds to the solution to the problem: Find $p \in \tilde{H}^s(\Omega) \cap L^\infty(\Omega)$ such that

$$(4.8) \quad \mathcal{A}(v, p) + \int_{\Omega} \frac{\partial a}{\partial u}(x, u) p v dx = \int_{\Omega} \frac{\partial L}{\partial u}(x, u) v dx \quad \forall v \in \tilde{H}^s(\Omega),$$

where $u = \mathcal{S}q \in \tilde{H}^s(\Omega) \cap L^\infty(\Omega)$. Since $\partial a / \partial u(x, u) \geq 0$ for a.e. $x \in \Omega$ and for all $u \in \mathbb{R}$ and $\partial L / \partial u(\cdot, u) \in L^r(\Omega)$ for every $m > 0$ and $u \in [-m, m]$, the adjoint problem (4.8) is well-posed.

PROPOSITION 4.3 (differentiability properties of F). *Let $s \in (0, 1)$ and let $r > d/2s$. Assume that (A.1)–(A.3) and (B.1)–(B.2) hold. Then, $F : L^2(\Omega) \cap L^r(\Omega) \rightarrow \mathbb{R}$ is of class C^2 . In addition, if $q, w \in L^2(\Omega) \cap L^r(\Omega)$, then*

$$(4.9) \quad F'(q)w = \int_{\Omega} (p + \lambda q) w dx,$$

where p solves (4.8). If $w_1, w_2 \in L^2(\Omega) \cap L^r(\Omega)$, then we have

$$(4.10) \quad F''(q)(w_1, w_2) = \int_{\Omega} \left(\frac{\partial^2 L}{\partial u^2}(x, u) \phi_{w_1} \phi_{w_2} + \lambda w_1 w_2 - p \frac{\partial^2 a}{\partial u^2}(x, u) \phi_{w_1} \phi_{w_2} \right) dx,$$

where $u = \mathcal{S}q$ and $\phi_{w_i} = \mathcal{S}'(q)w_i$, with $i \in \{1, 2\}$.

Proof. See [15, Proposition 4.5] for a proof. \square

The necessary first order optimality conditions are as follows.

THEOREM 4.4 (first order optimality conditions). *If $\bar{q} \in \mathbb{Q}_{ad}$ is a local solution to (4.1)–(4.2), then there exists $\bar{\eta} \in \partial j(\bar{q})$ such that*

$$(4.11) \quad \int_{\Omega} (\bar{p} + \lambda \bar{q} + \mu \bar{\eta})(q - \bar{q}) dx \geq 0 \quad \forall q \in \mathbb{Q}_{ad}.$$

Here, \bar{p} denotes the solution to (4.8) where u is replaced by $\bar{u} = \mathcal{S}\bar{q}$.

Proof. Since \bar{q} is a local solution, there exists $\varepsilon > 0$ such that $j(\bar{q}) \leq j(q)$ for every $q \in \mathbb{Q}_{ad}$ such that $\|\bar{q} - q\|_{L^2(\Omega)} \leq \varepsilon$. Let $q \in \mathbb{Q}_{ad}$ and let $\rho \in (0, 1)$ be sufficiently small so that $q_{\rho} := \bar{q} + \rho(q - \bar{q}) = \rho q + (1 - \rho)\bar{q} \in \mathbb{Q}_{ad}$ satisfies $\|\bar{q} - q_{\rho}\|_{L^2(\Omega)} \leq \varepsilon$. Thus

$$(4.12) \quad 0 \leq j(q_{\rho}) - j(\bar{q}) = [F(q_{\rho}) - F(\bar{q})] + \mu[j(q_{\rho}) - j(\bar{q})].$$

Note that $j(q_{\rho}) - j(\bar{q}) \leq \rho(j(q) - j(\bar{q}))$ because j is convex. Dividing (4.12) by ρ and taking the limit as $\rho \downarrow 0$ yield $0 \leq F'(\bar{q})(q - \bar{q}) + \mu(j(q) - j(\bar{q}))$ for every $q \in \mathbb{Q}_{ad}$. This inequality can be rewritten as follows: $\bar{q} \in \mathbb{Q}_{ad}$ is such that

$$\ell(\bar{q}) \leq \ell(q) \quad \forall q \in \mathbb{Q}_{ad}, \quad \ell : L^2(\Omega) \rightarrow \mathbb{R}, \quad \ell(q) := F'(\bar{q})q + \mu j(q),$$

i.e., \bar{q} is a global minimizer of ℓ over \mathbb{Q}_{ad} . We must thus have that $0 \in \partial(\ell + \mathbb{I}_{\mathbb{Q}_{ad}})(\bar{q}) = \partial\ell(\bar{q}) + \partial\mathbb{I}_{\mathbb{Q}_{ad}}(\bar{q})$ [18, page 134 and Proposition 4.5.1]. Here, $\mathbb{I}_{\mathbb{Q}_{ad}}$ corresponds to the indicator functional of \mathbb{Q}_{ad} ; see section 2.2. We now use that $F'(\bar{q}) = (\bar{p} + \lambda \bar{q})$ and let $\bar{\eta} \in \partial j(\bar{q})$ to arrive at $-(\bar{p} + \lambda \bar{q}) - \mu \bar{\eta} \in \partial\mathbb{I}_{\mathbb{Q}_{ad}}(\bar{q})$. This allows us to conclude. \square

Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ be such that $\mathbf{a} < \mathbf{b}$. We introduce the operator $\Pi_{[\mathbf{a}, \mathbf{b}]} : L^1(\Omega) \rightarrow \mathbb{Q}_{ad}$ by $\Pi_{[\mathbf{a}, \mathbf{b}]}(v) = \min\{\mathbf{b}, \max\{\mathbf{a}, v\}\}$ and present the following result.

THEOREM 4.5 (projection formulas). *If $\bar{q}, \bar{u}, \bar{p}$, and $\bar{\eta}$ are as in the statement of Theorem 4.4, then*

$$(4.13) \quad \bar{q}(x) = \Pi_{[\alpha, \beta]}(-\lambda^{-1}(\bar{p}(x) + \mu \bar{\eta}(x))), \quad \bar{q}(x) = 0 \Leftrightarrow |\bar{p}(x)| \leq \mu,$$

$$(4.14) \quad \bar{\eta}(x) = \Pi_{[-1, 1]}(-\mu^{-1}\bar{p}(x))$$

for a.e. $x \in \Omega$. In particular, the subgradient $\bar{\eta} \in \partial j(\bar{q})$ is uniquely determined and both \bar{q} and $\bar{\eta}$ belong to $\tilde{H}^s(\Omega) \cap L^\infty(\Omega)$.

Proof. The derivation of the projection formula for \bar{q} in (4.13) is standard in PDE-constrained optimization. The equivalence $\bar{q}(x) = 0 \Leftrightarrow |\bar{p}(x)| \leq \mu$ for a.e. $x \in \Omega$ can be found in [6, Corollary 3.2]. The projection formula for $\bar{\eta}$ in (4.14) can also be found in [6, Corollary 3.2]. This formula directly guarantees the uniqueness of $\bar{\eta}$. The desired regularity properties for \bar{q} and $\bar{\eta}$ follow from the projection formulas in (4.13) and (4.14), the fact that $\max\{0, \tau\} = (\tau + |\tau|)/2$ for all $\tau \in \mathbb{R}$, and [13, Theorem 1]. \square

4.3. Second order optimality conditions. Let $\bar{q} \in \mathbb{Q}_{ad}$ be a local minimum and let $\bar{\eta} \in \partial j(\bar{q})$ be the corresponding subgradient. We define

$$(4.15) \quad C_{\bar{q}} := \{w \in L^2(\Omega) : w \text{ satisfies (4.16) and } F'(\bar{q})w + \mu j'(\bar{q}; w) = 0\},$$

where $F'(\bar{q})(\cdot)$ and $j'(\bar{q}, \cdot)$ are described in Proposition 4.3 and (2.2), respectively, and

$$(4.16) \quad w(x) \geq 0 \text{ if } \bar{q}(x) = \alpha, \quad w(x) \leq 0 \text{ if } \bar{q}(x) = \beta.$$

The set $C_{\bar{q}}$ is a closed and convex cone in $L^2(\Omega)$ [6, Proposition 3.4].

We formulate necessary second order optimality conditions as follows.

THEOREM 4.6 (second order necessary optimality conditions). *If $\bar{q} \in \mathbb{Q}_{ad}$ is a local minimum of problem (4.1)–(4.2), then $F''(\bar{q})w^2 \geq 0$ for every $w \in C_{\bar{q}}$.*

Proof. The proof follows the same arguments as in [6, Theorem 3.7]. \square

To present the following result, for $\tau > 0$, we introduce the cone

$$(4.17) \quad C_{\bar{q}}^\tau := \{w \in L^2(\Omega) : w \text{ satisfies (4.16) and } F'(\bar{q})w + \mu j'(\bar{q}; w) \leq \tau \|w\|_{L^2(\Omega)}\}.$$

THEOREM 4.7 (equivalence). *Let $\bar{q} \in \mathbb{Q}_{ad}$ be a local minimum and let $\bar{\eta} \in \partial j(\bar{q})$ be the corresponding subgradient such that (4.11) holds. Then, the following statements are equivalent:*

- (i) $F''(\bar{q})w^2 > 0$ for all $w \in C_{\bar{q}} \setminus \{0\}$.
- (ii) There exist $\tau, \delta > 0$ such that $F''(\bar{q})w^2 \geq \delta \|w\|_{L^2(\Omega)}^2$ for all $w \in C_{\bar{q}}^\tau$.

Proof. Since $C_{\bar{q}} \subset C_{\bar{q}}^\tau$ for every $\tau > 0$, it is immediate that (ii) implies (i).

We now prove that (i) implies (ii). To do so, we follow the arguments of the proof of [6, Theorem 3.8] and proceed by contradiction. Indeed, we assume the existence of a sequence $\{v_k\}_{k \in \mathbb{N}}$ such that

$$v_k \in C_{\bar{q}}^{1/k}, \quad F''(\bar{q})v_k^2 < k^{-1} \|v_k\|_{L^2(\Omega)}^2, \quad k \in \mathbb{N}.$$

Define $w_k := \|v_k\|_{L^2(\Omega)}^{-1} v_k$. Note that $w_k \in C_{\bar{q}}^{1/k}$ because $C_{\bar{q}}^{1/k}$ is a cone. Moreover,

$$(4.18) \quad \|w_k\|_{L^2(\Omega)} = 1, \quad F''(\bar{q})w_k^2 < k^{-1}, \quad k \in \mathbb{N}.$$

Since $\{w_k\}_{k \in \mathbb{N}}$ is uniformly bounded in $L^2(\Omega)$, we can extract a nonrelabeled subsequence such that $w_k \rightharpoonup w$ in $L^2(\Omega)$ as $k \uparrow \infty$. Note that w satisfies (4.16). Apply [6, Lemma 3.5] to derive that $F'(\bar{q})w + \mu j'(\bar{q}; w) \geq 0$. On the other hand, since $j'(\bar{q}; \cdot)$ is weakly lower semicontinuous in $L^2(\Omega)$ and $F'(\bar{q})w_k \rightarrow F'(\bar{q})w$ as $k \uparrow \infty$ we obtain

$$F'(\bar{q})w + \mu j'(\bar{q}; w) \leq \liminf_{k \uparrow \infty} F'(\bar{q})w_k + \mu j'(\bar{q}; w_k) \leq \liminf_{k \uparrow \infty} k^{-1} = 0.$$

To obtain the last inequality we used that $w_k \in C_{\bar{q}}^{1/k}$. This proves that $F'(\bar{q})w + \mu j'(\bar{q}; w) = 0$ and thus that $w \in C_{\bar{q}}$.

We now prove that $w \equiv 0$. Since (i) holds, it suffices to prove that $F''(\bar{q})w^2 \leq 0$. To accomplish this task, we use the characterization (4.10) and write

$$(4.19) \quad F''(\bar{q})w_k^2 = \int_{\Omega} \left(\frac{\partial^2 L}{\partial u^2}(x, \bar{u}) \phi_k^2 + \lambda w_k^2 - \bar{p} \frac{\partial^2 a}{\partial u^2}(x, \bar{u}) \phi_k^2 \right) dx.$$

Here, $\phi_k = \mathcal{S}'(\bar{q})w_k \in \tilde{H}^s(\Omega) \cap L^\infty(\Omega)$ solves (4.6) with u and w replaced by $\bar{u} = \mathcal{S}(\bar{q})$ and w_k , respectively and $\bar{p} \in \tilde{H}^s(\Omega) \cap L^\infty(\Omega)$ solves (4.8) with u replaced by \bar{u} . Since $w_k \rightharpoonup w$ in $L^2(\Omega)$ as $k \uparrow \infty$, we deduce that $\phi_k \rightharpoonup \phi = \mathcal{S}'(\bar{q})w$ in $\tilde{H}^s(\Omega)$ as $k \uparrow \infty$. Note that ϕ solves (4.6) with u replaced by \bar{u} . This convergence result and the compact embedding of section 2.3 show that $\phi_k \rightarrow \phi$ in $L^t(\Omega)$ for every $t < 2d/(d-2s)$. Thus, we invoke the assumptions (A.3) and (B.2) and obtain

$$\begin{aligned} \left| \int_{\Omega} \frac{\partial^2 L}{\partial u^2}(x, \bar{u}) (\phi_k^2 - \phi^2) dx \right| &\leq \left\| \frac{\partial^2 L}{\partial u^2}(\cdot, \bar{u}) \right\|_{L^r(\Omega)} \|\phi_k - \phi\|_{L^t(\Omega)} \|\phi_k + \phi\|_{L^t(\Omega)}, \\ \left| \int_{\Omega} \frac{\partial^2 a}{\partial u^2}(x, \bar{u}) \bar{p} (\phi_k^2 - \phi^2) dx \right| &\leq \left\| \frac{\partial^2 a}{\partial u^2}(\cdot, \bar{u}) \right\|_{L^r(\Omega)} \|\bar{p}\|_{L^\infty(\Omega)} \|\phi_k - \phi\|_{L^t(\Omega)} \|\phi_k + \phi\|_{L^t(\Omega)}. \end{aligned}$$

Here $r^{-1} + 2t^{-1} = 1$. Note that, since $r > d/2s$, we have that $t < 2d/(d - 2s)$.

Since the square of $\|\cdot\|_{L^2(\Omega)}$ is continuous and convex in $L^2(\Omega)$, the aforementioned convergence results based on (4.19) and (4.18) yield

$$F''(\bar{q})w^2 \leq \liminf_{k \rightarrow \infty} F''(\bar{q})w_k^2 \leq 0.$$

The condition (i), which reads: $F''(\bar{q})w^2 > 0$ for all $w \in C_{\bar{q}} \setminus \{0\}$, therefore implies that $w = 0$ and that $F''(\bar{q})w_k^2 \rightarrow 0$ as $k \uparrow \infty$.

After proving that $w = 0$, we conclude the proof by arriving at a contradiction. We begin by noting that $w_k \rightarrow 0$ in $L^2(\Omega)$ implies that $\phi_k \rightarrow 0$ in $L^r(\Omega)$ for every $r < 2d/(d - 2s)$. We use the latter convergence result and $\|w_k\|_{L^2(\Omega)} = 1$ to obtain

$$F''(\bar{q})w_k^2 = \lambda + \int_{\Omega} \left(\frac{\partial^2 L}{\partial u^2}(x, \bar{u})\phi_k^2 - \bar{p} \frac{\partial^2 a}{\partial u^2}(x, \bar{u})\phi_k^2 \right) dx \rightarrow \lambda > 0, \quad k \uparrow \infty.$$

This is a contradiction because $F''(\bar{q})w_k^2 \rightarrow 0$ as $k \uparrow \infty$. This concludes the proof. \square

We conclude this section with second order sufficient optimality conditions.

THEOREM 4.8 (sufficient second order optimality conditions). *Let $\bar{q} \in \mathbb{Q}_{ad}$ be a local minimum and let $\bar{\eta} \in \partial j(\bar{q})$ be the corresponding subgradient such that (4.11) holds. If $F''(\bar{q})w^2 > 0$ for all $w \in C_{\bar{q}} \setminus \{0\}$, then there exist $\delta > 0$ and $\varepsilon > 0$ such that*

$$j(q) \geq j(\bar{q}) + \frac{\delta}{4} \|q - \bar{q}\|_{L^2(\Omega)}^2 \quad \forall q \in \mathbb{Q}_{ad} : \|q - \bar{q}\|_{L^2(\Omega)} \leq \varepsilon.$$

Proof. The proof follows the same arguments as in [6, Theorem 3.9]. \square

4.4. Regularity properties. Let \bar{q} be a local minimum of problem (4.1)–(4.2). In this section, we provide regularity results for all involved optimal control variables, i.e., \bar{q} , \bar{u} , \bar{p} , and $\bar{\eta}$. To accomplish this task, we assume that the conditions (A.1)–(A.3) and (B.1)–(B.2) hold together with the additional property (C.1).

To present the following result, we define

$$(4.20) \quad \Lambda(L, u) = \left\| \frac{\partial L}{\partial u}(\cdot, u) \right\|_{L^2(\Omega)}.$$

THEOREM 4.9 (regularity properties). *Let \bar{q} be a local minimum of (4.1)–(4.2) and let \bar{u} , \bar{p} , and $\bar{\eta}$ be the associated optimal variables. Then, $\bar{u}, \bar{p}, \bar{q}, \bar{\eta} \in H^{s+\kappa-\varepsilon}(\Omega)$ for all $\varepsilon \in (0, s)$, where $\kappa = \frac{1}{2}$ for $\frac{1}{2} < s < 1$ and $\kappa = s - \varepsilon$ for $0 < s \leq \frac{1}{2}$. Moreover,*

$$(4.21) \quad \|\bar{u}\|_{H^{s+\kappa-\varepsilon}(\Omega)} \lesssim \varepsilon^{-\nu} \|\bar{q} - a(\cdot, 0)\|_{L^2(\Omega)},$$

$$(4.22) \quad \|\bar{p}\|_{H^{s+\kappa-\varepsilon}(\Omega)} \lesssim \varepsilon^{-\nu} \Lambda(L, \bar{u}),$$

and

$$(4.23) \quad \|\bar{q}\|_{H^{s+\kappa-\varepsilon}(\Omega)} + \|\bar{\eta}\|_{H^{s+\kappa-\varepsilon}(\Omega)} \lesssim \varepsilon^{-\nu} [1 + \Lambda(L, \bar{u})],$$

where $\nu = \frac{1}{2}$ for $\frac{1}{2} < s < 1$ and $\nu = \frac{1}{2} + \nu_0$ for $0 < s \leq \frac{1}{2}$, with $\nu_0 = \nu_0(\Omega, d) > 0$. The hidden constant in all inequalities is independent of ε .

Proof. Since $\bar{q} \in \mathbb{Q}_{ad} \subset L^2(\Omega) \cap L^r(\Omega)$ and a is locally Lipschitz with respect to the second variable and satisfies (A.1), the desired regularity property for \bar{u} follows from Theorem 3.1. We recall that $\bar{u} \in L^\infty(\Omega)$ and that $\|\bar{u}\|_{L^\infty(\Omega)} \lesssim \|\bar{q} - a(\cdot, 0)\|_{L^r(\Omega)}$. We now focus on the optimal adjoint variable \bar{p} . We first note that since $\bar{u} \in L^\infty(\Omega)$

and, for every $\mathbf{m} > 0$, $|\partial L/\partial u(x, u)| \leq \psi_{\mathbf{m}}(x)$ for a.e. $x \in \Omega$ and $u \in [-\mathbf{m}, \mathbf{m}]$, with $\psi_{\mathbf{m}} \in L^r(\Omega)$, we can conclude that $\bar{p} \in L^\infty(\Omega)$ [15, Theorem 3.1]. With this regularity result, we invoke the assumptions (A.3) and (C.1) to arrive at

$$\frac{\partial L}{\partial u}(\cdot, \bar{u}) - \frac{\partial a}{\partial u}(\cdot, \bar{u})\bar{p} \in L^2(\Omega).$$

Thus, an application of [4, Theorem 2.1 and inequality (2.6)] show that $\bar{p} \in H^{s+\kappa-\varepsilon}(\Omega)$ together with the estimate

$$\|\bar{p}\|_{H^{s+\kappa-\varepsilon}(\Omega)} \lesssim \varepsilon^{-\nu} (\Lambda(\cdot, \bar{u}) + \|\bar{p}\|_{L^2(\Omega)}) \lesssim \varepsilon^{-\nu} \Lambda(\cdot, \bar{u}).$$

We note that since $\bar{u} \in L^\infty(\Omega)$ and the assumption (A.3) holds, $|\partial a/\partial u(x, \bar{u})| \leq C_{\mathbf{m}}$ for a.e. $x \in \Omega$. The desired regularity property for $\bar{\eta}$ and part of the estimate in (4.23) follow from the projection formula (4.14), the fact that $\max\{0, \tau\} = (\tau + |\tau|)/2$ for all $\tau \in \mathbb{R}$, and [13, Theorem 1], which applies because $s + \kappa - \varepsilon = s + \frac{1}{2} - \varepsilon < \frac{3}{2}$ if $\frac{1}{2} < s < 1$ and $s + \kappa - \varepsilon = 2s - 2\varepsilon \leq 1 - 2\varepsilon < \frac{3}{2}$ if $0 < s \leq \frac{1}{2}$. The desired regularity property for \bar{q} follows the same arguments. This concludes the proof. \square

5. A fully discrete scheme for the optimal control problem. Before we begin our analysis, we would like to mention that in this section we assume that Ω is a Lipschitz *polytope*. In the following, we propose and analyze the following fully discrete finite element approximation of our optimal control problem (4.1)–(4.2): Find $\min J(u_h, q_h)$ subject to the *discrete state equation*

$$(5.1) \quad \mathcal{A}(u_h, v_h) + \int_{\Omega} a(x, u_h)v_h dx = \int_{\Omega} q_h v_h dx \quad \forall v_h \in \mathbb{V}_h$$

and the *control constraints* $q_h \in \mathbb{Q}_{ad,h}$. Here, $\mathbb{Q}_{ad,h} := \mathbb{Q}_{ad} \cap \mathbb{Q}_h$, where $\mathbb{Q}_h := \{q_h \in L^\infty(\Omega) : q_h|_T \in \mathbb{P}_0(T) \forall T \in \mathcal{T}_h\}$. \mathbb{V}_h denotes the finite element space defined in (3.3).

The existence of at least one optimal solution is standard. To provide first order optimality conditions, we introduce $F_h : L^2(\Omega) \rightarrow \mathbb{R}$ and $j_h : \mathbb{Q}_h \rightarrow \mathbb{R}$ to be such that

$$F_h(q) = \int_{\Omega} L(x, \mathcal{S}_h q) dx + \frac{\lambda}{2} \|q\|_{L^2(\Omega)}^2, \quad j_h(q_h) = \|q_h\|_{L^1(\Omega)},$$

where $\mathcal{S}_h : \mathbb{Q}_{ad,h} \rightarrow \mathbb{V}_h$ denotes the *discrete control to state map*.

First order optimality conditions for the fully discrete scheme are as follows.

THEOREM 5.1 (first order optimality conditions). *If $\bar{q}_h \in \mathbb{Q}_{ad,h}$ is a local solution to the fully discrete scheme, then there exists $\bar{\eta}_h \in \partial j_h(\bar{q}_h)$ such that*

$$(5.2) \quad \int_{\Omega} (\bar{p}_h + \lambda \bar{q}_h + \mu \bar{\eta}_h)(q_h - \bar{q}_h) dx \geq 0 \quad \forall q_h \in \mathbb{Q}_{ad,h}.$$

Here, \bar{p}_h is the solution to the discrete adjoint equation: Find $\bar{p}_h \in \mathbb{V}_h$ such that

$$(5.3) \quad \mathcal{A}(v_h, \bar{p}_h) + \int_{\Omega} \frac{\partial a}{\partial u}(x, \bar{u}_h) \bar{p}_h v_h dx = \int_{\Omega} \frac{\partial L}{\partial u}(x, \bar{u}_h) v_h dx \quad \forall v_h \in \mathbb{V}_h,$$

where $\bar{u}_h = \mathcal{S}_h \bar{q}_h$.

Proof. The proof follows the same arguments as in Theorem 4.4. \square

Analogous to the continuous case, we obtain projection formulas, but now on a discrete level.

THEOREM 5.2 (discrete projection formulas). *If $\bar{q}_h \in \mathbb{Q}_{ad,h}$ and $\bar{\eta}_h \in \partial j_h(\bar{q}_h)$ are as in the statement of Theorem 5.1, then for every $T \in \mathcal{T}_h$, we have the formulas*

$$(5.4) \quad \bar{q}_h|_T = \Pi_{[\alpha,\beta]} \left[-\frac{1}{\lambda} \left(\frac{1}{|T|} \int_T \bar{p}_h dx + \mu \bar{\eta}_h|_T \right) \right],$$

$$(5.5) \quad \bar{q}_h|_T = 0 \Leftrightarrow \frac{1}{|T|} \left| \int_T \bar{p}_h dx \right| \leq \mu,$$

$$(5.6) \quad \bar{\eta}_h|_T = \Pi_{[-1,1]} \left(-\frac{1}{\mu|T|} \int_T \bar{p}_h dx \right).$$

In particular, the subgradient $\bar{\eta}_h \in \partial j_h(\bar{q}_h)$ is uniquely determined.

Proof. From the variational inequality (5.2) we have that, for every $q_T \in [\alpha, \beta]$,

$$\left(\int_T \bar{p}_h dx + [\lambda \bar{q}_h|_T + \mu \bar{\eta}_h|_T] |T| \right) (q_T - \bar{q}_h|_T) \geq 0.$$

The projection formula (5.4) can be derived from this inequality. The sparse property (5.5) and the projection formula (5.6) are derived in [6, equivalence (4.4b)] and [6, formula (4.4c)], respectively. \square

We conclude this section with the following error bounds for the discretization of the adjoint equation and the corresponding subdifferential variables by finite elements. These error bounds hold under the assumption that discrete solutions u_h of the problem (5.1) are uniformly bounded in $L^\infty(\Omega)$, i.e.,

$$(5.7) \quad \exists C > 0 : \quad \|u_h\|_{L^\infty(\Omega)} \leq C \quad \forall h > 0.$$

To present a proof of some of the aforementioned error bounds, we introduce $\mathcal{P}_h : L^2(\Omega) \rightarrow \mathbb{Q}_h$, the orthogonal projection operator onto piecewise constant functions.

THEOREM 5.3 (error bounds). *Let \bar{q} be a local solution to (4.1)–(4.2) and let $\bar{\eta} \in \partial j(\bar{q})$ be as in Theorems 4.4 and 4.5. Let \bar{p} be the solution to (4.8) with u replaced by $\bar{u} = S\bar{q}$. Let \bar{q}_h be a local solution to the fully discrete scheme and let $\bar{\eta}_h \in \partial j_h(\bar{q}_h)$ be as in Theorems 5.1 and 5.2. Let \bar{p}_h be the solution to (5.3). Let us assume that (A.1)–(A.3), (B.1)–(B.2), (C.1), and (5.7) hold. Then, we have*

$$(5.8) \quad \|\bar{p} - \bar{p}_h\|_s \lesssim h^\gamma |\log h|^\varphi + \|\bar{q} - \bar{q}_h\|_{L^2(\Omega)}, \quad \gamma = \min\{s, \frac{1}{2}\}$$

$$(5.9) \quad \|\bar{p} - \bar{p}_h\|_{L^2(\Omega)} + \|\bar{\eta} - \bar{\eta}_h\|_{L^2(\Omega)} \lesssim h^{2\gamma} |\log h|^{2\varphi} + \|\bar{q} - \bar{q}_h\|_{L^2(\Omega)},$$

where $\varphi = \nu$ if $s \neq \frac{1}{2}$, $\varphi = 1 + \nu$ if $s = \frac{1}{2}$, and $\nu \geq \frac{1}{2}$ is the constant in Theorem 3.1.

Proof. The error estimates for $\bar{p} - \bar{p}_h$ in $\tilde{H}^s(\Omega)$ and $L^2(\Omega)$ can be found in [16, Theorem 6.2, estimate (6.17)] and [16, Theorem 6.2, estimate (6.18)], respectively. It is worth noting that in the proof of [16, Theorem 6.2] it is assumed that $\partial L/\partial u$ is locally Lipschitz with respect to the second variable. This assumption can be removed at the expense of having the assumption on the second derivative in (C.1). It remains to prove the estimate for $\bar{\eta} - \bar{\eta}_h$ in $L^2(\Omega)$. To do this, we use the projection formulas for $\bar{\eta}$ and $\bar{\eta}_h$, which are given in (4.14) and (5.6), respectively, and arrive at

$$(5.10) \quad \|\bar{\eta} - \bar{\eta}_h\|_{L^2(\Omega)} \leq \mu^{-1} \|\bar{p} - \mathcal{P}_h \bar{p}_h\|_{L^2(\Omega)}.$$

A simple application of a triangle inequality thus results in $\|\bar{\eta} - \bar{\eta}_h\|_{L^2(\Omega)} \leq \mu^{-1} \|\bar{p} - \mathcal{P}_h \bar{p}\|_{L^2(\Omega)} + C\mu^{-1} \|\bar{p} - \bar{p}_h\|_{L^2(\Omega)}$ using the fact that \mathcal{P}_h is stable in $L^2(\Omega)$. The regularity property for \bar{p} derived in Theorem 4.9, namely, $\bar{p} \in H^{s+\kappa-\varepsilon}(\Omega)$, where κ and

ε are the same as in the statement of Theorem 4.9, and a basic error estimate for the orthogonal projection \mathcal{P}_h yield

$$(5.11) \quad \|\bar{p} - \mathcal{P}_h \bar{p}\|_{L^2(\Omega)} \lesssim h^{2\gamma} |\log h|^\nu \Lambda(L, \bar{u}), \quad \gamma = \min\{s, \frac{1}{2}\}.$$

This and the derived error estimate for $\bar{p} - \bar{p}_h$ in $L^2(\Omega)$ allow us to complete the error estimate (5.9). This completes the proof. \square

5.1. Convergence of discretizations. We begin this section with a convergence result that guarantees that a sequence of discrete global solutions $\{\bar{q}_h\}_{h>0}$ contains subsequences that converge to global solutions of problem (4.1)–(4.2) as $h \rightarrow 0$.

THEOREM 5.4 (convergence of discrete solutions). *Let us assume that (A.1)–(A.3), (B.1)–(B.2), and (C.1) hold. Let $h > 0$ and let $\bar{q}_h \in \mathbb{Q}_{ad,h}$ be a global solution to the fully discrete control problem. Then, there exist nonrelabeled subsequences of $\{\bar{q}_h\}_{h>0}$ such that $\bar{q}_h \overset{*}{\rightharpoonup} \bar{q}$ in the weak* topology of $L^\infty(\Omega)$ as $h \rightarrow 0$ and \bar{q} corresponds to a global solution of the continuous control problem (4.1)–(4.2). Furthermore,*

$$(5.12) \quad \lim_{h \rightarrow 0} \|\bar{q} - \bar{q}_h\|_{L^2(\Omega)} = 0, \quad \lim_{h \rightarrow 0} j_h(\bar{q}_h) = j(\bar{q}).$$

If, in addition, (5.7) holds, then

$$(5.13) \quad \|\bar{\eta} - \bar{\eta}_h\|_{L^2(\Omega)} \rightarrow 0$$

as $h \rightarrow 0$.

Proof. Since $\{\bar{q}_h\}_{h>0} \subset \mathbb{Q}_{ad,h}$ is uniformly bounded in $L^\infty(\Omega)$, we can extract a nonrelabeled subsequence such that $\bar{q}_h \overset{*}{\rightharpoonup} \bar{q}$ in the weak* topology of $L^\infty(\Omega)$ as $h \rightarrow 0$. In the following, we prove that the limit \bar{q} is a global solution to (4.1)–(4.2).

Let $\bar{q} \in \mathbb{Q}_{ad}$ be a global solution to (4.1)–(4.2) and define $\mathbf{q}_h = \mathcal{P}_h \bar{q} \in \mathbb{Q}_h$. We note that according to the definition of \mathcal{P}_h , \mathbf{q}_h is such that $\mathbf{q}_h|_T = \int_T \bar{q}(x) dx / |T|$ for each $T \in \mathcal{T}_h$. This immediately guarantees that $\mathbf{q}_h \in \mathbb{Q}_{ad,h}$, because $\bar{q} \in \mathbb{Q}_{ad}$. We now take advantage of the fact that $\bar{q} \in H^{s+\kappa-\varepsilon}(\Omega)$, where κ and ε are as in the statement of Theorem 4.9, and a basic error estimate for \mathcal{P}_h to deduce that $\|\bar{q} - \mathbf{q}_h\|_{L^2(\Omega)} \rightarrow 0$ as $h \rightarrow 0$. We now use the global optimality of $\bar{q} \in \mathbb{Q}_{ad}$ in conjunction with the fact that $\bar{q} \in \mathbb{Q}_{ad}$, a convergence result based on Lemma 3.3 and the fact that $\bar{q}_h \overset{*}{\rightharpoonup} \bar{q}$ in the weak* topology of $L^\infty(\Omega)$ as $h \rightarrow 0$, the global discrete optimality of $\bar{q}_h \in \mathbb{Q}_{ad,h}$ combined with the fact that $\mathbf{q}_h \in \mathbb{Q}_{ad,h}$, and another convergence result based on Lemma 3.3 and the fact that $\|\bar{q} - \mathbf{q}_h\|_{L^2(\Omega)} \rightarrow 0$ as $h \rightarrow 0$ to obtain

$$j(\bar{q}) \leq j(\bar{q}) \leq \liminf_{h \rightarrow 0} j_h(\bar{q}_h) \leq \limsup_{h \rightarrow 0} j_h(\bar{q}_h) \leq \limsup_{h \rightarrow 0} j_h(\mathbf{q}_h) = j(\bar{q}).$$

As a result, we have obtained that $j(\bar{q}) \leq j(\bar{q}) \leq j(\bar{q})$, which guarantees the global optimality of \bar{q} . We also proved that $j_h(\bar{q}_h) \rightarrow j(\bar{q})$ as $h \rightarrow 0$.

We now prove that $\|\bar{q} - \bar{q}_h\|_{L^2(\Omega)} \rightarrow 0$ as $h \rightarrow 0$. Due to Lemma 3.3, we have that $\bar{u}_h = \mathcal{S}_h \bar{q}_h \rightarrow \bar{u} = \mathcal{S} \bar{q}$ in $L^\tau(\Omega)$ for every $\tau \leq 2d/(d-2s)$. From this follows $|\int_\Omega (L(x, \bar{u}_h) - L(x, \bar{u})) dx| \rightarrow 0$ as $h \rightarrow 0$. Since $j_h(\bar{q}_h) \rightarrow j(\bar{q})$ as $h \rightarrow 0$, we obtain

$$\frac{\lambda}{2} \|\bar{q}_h\|_{L^2(\Omega)}^2 + \mu \|\bar{q}_h\|_{L^1(\Omega)} \rightarrow \frac{\lambda}{2} \|\bar{q}\|_{L^2(\Omega)}^2 + \mu \|\bar{q}\|_{L^1(\Omega)}, \quad h \rightarrow 0.$$

This and the weak convergence $\bar{q}_h \rightharpoonup \bar{q}$ in $L^2(\Omega)$ as $h \rightarrow 0$ guarantee the strong one $\|\bar{q} - \bar{q}_h\|_{L^2(\Omega)} \rightarrow 0$ as $h \rightarrow 0$. The convergence of the term $\|\bar{\eta} - \bar{\eta}_h\|_{L^2(\Omega)}$ follows immediately from Theorem 5.3. This concludes the proof. \square

To present the next result, we introduce $B_\epsilon(\bar{q}) := \{q \in L^2(\Omega) : \|\bar{q} - q\|_{L^2(\Omega)} \leq \epsilon\}$.

THEOREM 5.5. *Let the assumptions of Theorem 5.4 hold. If \bar{q} is a strict local minimum of the control problem (4.1)–(4.2), then there is a sequence $\{\bar{q}_h\}_{0 < h < h_\dagger}$ of local minima of the fully discrete control problems such that (5.12) and (5.13) hold.*

Proof. Since \bar{q} is a strict local minimum of (4.1)–(4.2), there exists $\epsilon > 0$ such that \bar{q} is the unique solution to the following problem: Find $\min\{j(q) : q \in \mathbb{Q}_{ad} \cap B_\epsilon(\bar{q})\}$.

We now introduce the following discrete problem for each $h > 0$: Find $\min\{j_h(q_h) : q_h \in \mathbb{Q}_{ad,h} \cap B_\epsilon(\bar{q})\}$. We note that there exists $h_\star > 0$ so that for each $h \in (0, h_\star)$ the discrete function $\mathcal{P}_h \bar{q}$ belongs to $\mathbb{Q}_{ad,h} \cap B_\epsilon(\bar{q})$. Consequently, $\mathbb{Q}_{ad,h} \cap B_\epsilon(\bar{q})$ is not empty and the previously introduced discrete problem admits a solution.

Let $h \in (0, h_\star)$ and let $\bar{q}_h \in \mathbb{Q}_{ad,h} \cap B_\epsilon(\bar{q})$ be a global solution of the aforementioned discrete problem. The arguments elaborated in the proof of Theorem 5.4 show the existence of a nonrelabeled subsequence of $\{\bar{q}_h\}_{0 < h < h_\star}$ such that it converges strongly in $L^2(\Omega)$ as $h \rightarrow 0$ to a global solution of the following problem: Find $\min\{j(q) : q \in \mathbb{Q}_{ad} \cap B_\epsilon(\bar{q})\}$. As mentioned at the beginning of the proof, this problem admits a unique solution \bar{q} . Consequently, the whole sequence $\{\bar{q}_h\}_{0 < h < h_\star}$ must converge to \bar{q} in $L^2(\Omega)$ as $h \rightarrow 0$. As a result, the constraint $\bar{q}_h \in \mathbb{Q}_{ad,h} \cap B_\epsilon(\bar{q})$ is not active for h sufficiently small so that \bar{q}_h solves the fully discrete scheme. \square

5.2. Error estimates. This section is dedicated to the derivation of error estimates. For this purpose, we let $\{\bar{q}_h\}_{h>0} \subset \mathbb{Q}_{ad,h}$ be a sequence of local minima of the fully discrete optimal control problems such that $\|\bar{q} - \bar{q}_h\|_{L^2(\Omega)} \rightarrow 0$ as $h \rightarrow 0$, where \bar{q} corresponds to a local solution of (4.1)–(4.2); see Theorems 5.4 and 5.5. The main goal of this section is to obtain an error estimate for $\bar{q} - \bar{q}_h$ in $L^2(\Omega)$, namely

$$(5.14) \quad \|\bar{q} - \bar{q}_h\|_{L^2(\Omega)} \lesssim h^{2\gamma} |\log h|^{2\varphi} \quad \forall h \leq h_\bullet, \quad \gamma = \min\left\{s, \frac{1}{2}\right\},$$

where $\varphi = \nu$ if $s \neq \frac{1}{2}$, $\varphi = 1 + \nu$ if $s = \frac{1}{2}$, and $\nu \geq \frac{1}{2}$ is the constant in Theorem 3.1.

The following result is helpful to obtain (5.14).

THEOREM 5.6 (instrumental error estimate). *Let us assume that (A.1)–(A.3), (B.1)–(B.2), (C.1), and (5.7) hold. Let $\bar{q} \in \mathbb{Q}_{ad}$ satisfy the second order optimality condition (i), or equivalently (ii) in Theorem 4.7. If (5.14) is false, then there exists $h_\square > 0$ such that*

$$(5.15) \quad \mathfrak{C} \|\bar{q} - \bar{q}_h\|_{L^2(\Omega)}^2 \leq [F'(\bar{q}_h) - F'(\bar{q})](\bar{q}_h - \bar{q}) \quad \forall h \in (0, h_\square],$$

where $\mathfrak{C} := 2^{-1}\delta$ and δ is the constant that appears in the item (ii) of Theorem 4.7.

Proof. In a first step, we invoke the C^2 regularity of F in $L^2(\Omega) \cap L^r(\Omega)$, with $r > d/2s$ (cf. Proposition 4.3), the $L^2(\Omega)$ -convergence of \bar{q}_h to \bar{q} as $h \rightarrow 0$, and the mean value theorem to conclude the existence of $\epsilon > 0$ and $h_\epsilon > 0$ such that

$$(5.16) \quad [F'(\bar{q}_h) - F'(\bar{q})](\bar{q}_h - \bar{q}) \geq F''(\bar{q})(\bar{q}_h - \bar{q})^2 - \frac{\delta}{2} \|\bar{q}_h - \bar{q}\|_{L^2(\Omega)}^2 \quad \forall h \leq h_\epsilon.$$

It therefore suffices to prove that $\bar{q}_h - \bar{q} \in C_\tau^r$ for some $\tau > 0$ and for every $h \leq h_\tau$, where $h_\tau > 0$, to apply item (ii) of Theorem 4.7 and deduce (5.15) with $h_\square := \min\{h_\tau, h_\epsilon\}$. Therefore, the rest of the proof is devoted to proving that $\bar{q}_h - \bar{q} \in C_\tau^r$.

For each $h > 0$, we define the function $v_h := (\bar{q}_h - \bar{q})/\|\bar{q}_h - \bar{q}\|_{L^2(\Omega)}$. Since $\{v_h\}_{h>0}$ is uniformly bounded in $L^2(\Omega)$, there exists a nonrelabeled subsequence such that $v_h \rightharpoonup v$ in $L^2(\Omega)$ as $h \rightarrow 0$. We now prove the existence of $\tau > 0$ and $h_\tau > 0$ such

that $v_h \in C_{\bar{q}}^\tau$ for every $h \leq h_\tau$. The fact that every v_h satisfies the sign conditions (4.16) is trivial. It is therefore sufficient to prove that

$$(5.17) \quad F'(\bar{q})v_h + \mu j'(\bar{q}; v_h) \leq \tau \quad \forall h \leq h_\tau.$$

Since [6, Lemma 3.5] guarantees that $F'(\bar{q})v_h + \mu j'(\bar{q}; v_h) \geq 0$, which holds for every $h > 0$ because $v_h \in L^2(\Omega)$ satisfies (4.16), we will obtain (5.17) with the help of the limit $F'(\bar{q})v_h + \mu j'(\bar{q}; v_h) \rightarrow 0$ as $h \rightarrow 0$. This will be now the focus of the proof.

Let us first note that the arguments developed in the proof of [15, Theorem 7.4] based on the weak convergence $v_h \rightharpoonup v$ in $L^2(\Omega)$ as $h \rightarrow 0$ show that

$$(5.18) \quad \left\{ F'(\bar{q})v + \mu \int_{\Omega} \bar{\eta} v dx \right\} = \lim_{h \rightarrow 0} \left\{ F'(\bar{q}_h)v_h + \mu \int_{\Omega} \bar{\eta}_h v_h dx \right\} \leq 0.$$

This is the place where we use that (5.14) is false. On the other hand, $F'(\bar{q})v + \mu \int_{\Omega} \bar{\eta} v dx = \int_{\Omega} [\bar{p} + \lambda \bar{q} + \mu \bar{\eta}] v dx \geq 0$ as a consequence of the variational inequality (4.11) and the fact that v satisfies (4.16). The relation $F'(\bar{q})v + \mu \int_{\Omega} \bar{\eta} v dx = 0$ can be derived from this.

We now prove that $j'(\bar{q}; v_h) \rightarrow \int_{\Omega} \bar{\eta} v dx$ as $h \rightarrow 0$; recall that $j'(\bar{q}; \cdot)$ is defined in (2.2). To do this, we first note that $v_h \rightharpoonup v$ in $L^2(\Omega)$ as $h \rightarrow 0$ yields

$$(5.19) \quad \lim_{h \rightarrow 0} \left\{ \int_{\Omega_{\bar{q}}^+} v_h dx - \int_{\Omega_{\bar{q}}^-} v_h dx \right\} = \int_{\Omega_{\bar{q}}^+} v dx - \int_{\Omega_{\bar{q}}^-} v dx.$$

We now study $\int_{\Omega_{\bar{q}}^0} |v_h| dx$, where $\Omega_{\bar{q}}^0 = \{x \in \Omega : \bar{q}(x) = 0\}$. In view of (4.13), the set $\Omega_{\bar{q}}^0$ can be rewritten as $\Omega_{\bar{q}}^0 = \{x \in \Omega : |\bar{p}(x)| \leq \mu\}$. We decompose this set as follows:

$$\begin{aligned} \Omega_{\bar{q}}^0 &= \Omega_{\mu}^+ \cup \Omega_{\mu}^- \cup \Omega_{\mu}^{\text{less}}, & \Omega_{\mu}^+ &:= \{x \in \Omega : \bar{p}(x) = \mu\}, \\ \Omega_{\mu}^- &:= \{x \in \Omega : \bar{p}(x) = -\mu\}, & \Omega_{\mu}^{\text{less}} &:= \{x \in \Omega : |\bar{p}(x)| < \mu\}. \end{aligned}$$

With these sets at hand, we can write the integral $\int_{\Omega_{\bar{q}}^0} |v_h| dx$ as follows:

$$(5.20) \quad \int_{\Omega_{\bar{q}}^0} |v_h| dx = \int_{\Omega_{\mu}^{\text{less}}} |v_h| dx + \int_{\Omega_{\mu}^+} |v_h| dx + \int_{\Omega_{\mu}^-} |v_h| dx := \mathfrak{J}_h + \mathfrak{K}_h + \mathfrak{L}_h.$$

In the following, we proceed in three steps to examine the three preceding terms.

Step 1. We study the limit value of \mathfrak{J}_h as $h \rightarrow 0$. For this purpose, for each $h > 0$ and $T \in \mathcal{T}_h$, we introduce the average $\bar{p}_{h,T} := \int_T \bar{p}_h dx / |T|$. We also define

$$(5.21) \quad \mathcal{T}_{1,h} := \{T \in \mathcal{T}_h : |\bar{p}_{h,T}| \leq \mu\}, \quad \Omega_{1,h} := \Omega_{\mu}^{\text{less}} \cap \mathcal{T}_{1,h},$$

$$(5.22) \quad \mathcal{T}_{2,h} := \{T \in \mathcal{T}_h : |\bar{p}_{h,T}| > \mu\}, \quad \Omega_{2,h} := \Omega_{\mu}^{\text{less}} \cap \mathcal{T}_{2,h}.$$

In an abuse of notation, in what follows, by $\mathcal{T}_{1,h}$ and $\mathcal{T}_{2,h}$, we will indistinctively denote either these sets or the union of the triangles that comprise them.

We now note that for each $T \in \mathcal{T}_{1,h}$, $\bar{q}_h|_T = 0$ as a consequence of (5.5). Thus,

$$v_h(x)|_T = \frac{\bar{q}_h(x) - \bar{q}(x)}{\|\bar{q}_h - \bar{q}\|_{L^2(\Omega)}} \Big|_T = \frac{-\bar{q}(x)}{\|\bar{q}_h - \bar{q}\|_{L^2(\Omega)}} \Big|_T = 0$$

for a.e. $x \in \Omega_{\mu}^{\text{less}} \cap T$ and for each $T \in \mathcal{T}_{1,h}$ due to (4.13). As a result, we obtain

$$\int_{\Omega_{1,h}} |v_h| dx = \sum_{T \in \mathcal{T}_{1,h}} \int_{T \cap \Omega_{\mu}^{\text{less}}} |v_h| dx = 0 \implies \mathfrak{J}_h = \int_{\Omega_{\mu}^{\text{less}}} |v_h| dx = \int_{\Omega_{2,h}} |v_h| dx.$$

We now bound $\int_{\Omega_{2,h}} |v_h| dx$. For this purpose we let $T \in \mathcal{T}_{2,h}$ and note that

$$0 < |\mu - |\bar{p}(x)|| < ||\bar{p}_{h,T} - |\bar{p}(x)||$$

for a.e. $x \in \Omega_\mu^{\text{less}} \cap T$. Integrating over $\Omega_\mu^{\text{less}} \cap T$ results in

$$0 \leq \int_{T \cap \Omega_\mu^{\text{less}}} |\mu - |\bar{p}||^2 dx < \int_T ||\mathcal{P}_h \bar{p}_h - |\bar{p}||^2 dx \leq \|\mathcal{P}_h \bar{p}_h - \bar{p}\|_{L^2(T)}^2.$$

Summing over all elements T in $\mathcal{T}_{2,h}$ and using the arguments from the proof of Theorem 5.3 which guarantee that $\|\mathcal{P}_h \bar{p}_h - \bar{p}\|_{L^2(\Omega)} \rightarrow 0$ as $h \rightarrow 0$, it follows that $|\Omega_{2,h}| \rightarrow 0$ as $h \rightarrow 0$. This implies that

$$(5.23) \quad \lim_{h \rightarrow 0} \mathfrak{J}_h = 0.$$

Step 2. We now study the limit value of \mathfrak{K}_h as $h \rightarrow 0$. To do this, we define

$$(5.24) \quad \mathcal{T}_{3,h} := \{T \in \mathcal{T}_h : \bar{p}_{h,T} < -\mu\}, \quad \Omega_{3,h} := \Omega_\mu^+ \cap \mathcal{T}_{3,h},$$

$$(5.25) \quad \mathcal{T}_{4,h} := \{T \in \mathcal{T}_h : \bar{p}_{h,T} > +\mu\}, \quad \Omega_{4,h} := \Omega_\mu^+ \cap \mathcal{T}_{4,h},$$

and $\Omega_{5,h} := \Omega_\mu^+ \cap \mathcal{T}_{1,h}$, where $\mathcal{T}_{1,h}$ is defined in (5.21). If we proceed as at the beginning of *Step 1*, we can deduce that $v_h(x) = 0$ for a.e. $x \in \Omega_{5,h}$. Consequently,

$$\mathfrak{K}_h = \int_{\Omega_\mu^+} |v_h| dx = \int_{\Omega_{3,h}} |v_h| dx + \int_{\Omega_{4,h}} |v_h| dx.$$

Let $T \in \mathcal{T}_{3,h}$. By definition, $\bar{p}_{h,T} < -\mu$. Recall that $\bar{p}(x) = \mu$ for a.e. $x \in \Omega_\mu^+$. Thus,

$$0 \leq 4\mu^2 |T \cap \Omega_\mu^+| = \int_{T \cap \Omega_\mu^+} (\mu - (-\mu))^2 dx < \int_T (\bar{p} - \mathcal{P}_h \bar{p}_h)^2 dx = \|\bar{p} - \mathcal{P}_h \bar{p}_h\|_{L^2(T)}^2.$$

Summing over all elements T in $\mathcal{T}_{3,h}$ we obtain $|\Omega_{3,h}| \rightarrow 0$ as $h \rightarrow 0$. To complete the proof of this step, we let $T \in \mathcal{T}_{4,h}$. Note that $\bar{p}_{h,T} > \mu$. Using the projection formulas (5.4) and (5.6), we can deduce that $\bar{q}_h|_T < 0$. This implies that $v_h(x) < 0$ for a.e. $x \in T \cap \Omega_\mu^+$. As a result, $|v_h(x)| = -v_h(x)$ for a.e. $x \in T \cap \Omega_\mu^+$. We now invoke the weak convergence $v_h \rightharpoonup v$ in $L^2(\Omega)$ as $h \rightarrow 0$ to obtain

$$(5.26) \quad \begin{aligned} \lim_{h \rightarrow 0} \mathfrak{K}_h &= - \lim_{h \rightarrow 0} \int_{\Omega_{4,h}} v_h dx = - \lim_{h \rightarrow 0} \int_{\Omega_{4,h}} v_h dx - \lim_{h \rightarrow 0} \int_{\Omega_{3,h}} v_h dx \\ &\quad - \lim_{h \rightarrow 0} \int_{\Omega_{5,h}} v_h dx = - \lim_{h \rightarrow 0} \int_{\Omega_\mu^+} v_h dx = - \int_{\Omega_\mu^+} v dx, \end{aligned}$$

because $\int_{\Omega_{3,h}} v_h dx \rightarrow 0$ as $h \rightarrow 0$ and $\int_{\Omega_{5,h}} v_h dx = 0$ for each $h > 0$.

Step 3. Using arguments similar to those in *Step 2* we arrive at

$$(5.27) \quad \lim_{h \rightarrow 0} \mathfrak{L}_h = \int_{\Omega_\mu^-} v dx.$$

Collecting (5.20), (5.23), (5.26), and (5.27) we conclude that

$$(5.28) \quad \lim_{h \rightarrow 0} \int_{\Omega_\mu^0} |v_h| dx = - \int_{\Omega_\mu^+} v dx + \int_{\Omega_\mu^-} v dx.$$

After we have proved (5.28) we use (5.19) and the fact that $\bar{\eta}(x) = 1$ for a.e. $x \in \Omega_{\bar{q}}^+$ and $x \in \Omega_{\mu}^-$ and that $\bar{\eta}(x) = -1$ for a.e. $x \in \Omega_{\bar{q}}^-$ and $x \in \Omega_{\mu}^+$ to deduce that

$$(5.29) \quad \begin{aligned} \lim_{h \rightarrow 0} j'(\bar{q}; v_h) &= \int_{\Omega_{\bar{q}}^+} v dx - \int_{\Omega_{\bar{q}}^-} v dx - \int_{\Omega_{\mu}^+} v dx + \int_{\Omega_{\mu}^-} v dx \\ &= \int_{\Omega_{\bar{q}}^+} \bar{\eta} v dx + \int_{\Omega_{\bar{q}}^-} \bar{\eta} v dx + \int_{\Omega_{\mu}^+} \bar{\eta} v dx + \int_{\Omega_{\mu}^-} \bar{\eta} v dx = \int_{\Omega} \bar{\eta} v dx. \end{aligned}$$

To obtain the last equality we have used that $\|v_h\|_{L^1(\Omega_{\mu}^{\text{less}})} \rightarrow 0$, which follows from (5.23), and that $v_h \rightharpoonup v$ in $L^2(\Omega)$ as $h \rightarrow 0$ to conclude that $v = 0$ a.e. in $\Omega_{\mu}^{\text{less}}$.

The limit (5.29) and the relation (5.18) allows us to conclude that

$$\lim_{h \rightarrow 0} (F'(\bar{q})v_h + \mu j'(\bar{q}; v_h)) = 0.$$

Since $F'(\bar{q})v_h + \mu j'(\bar{q}; v_h) \geq 0$ for each $h > 0$, we deduce the existence of $\tau > 0$ and $h_{\tau} > 0$ such that $F'(\bar{q})v_h + \mu j'(\bar{q}; v_h) \leq \tau$ for all $h \leq h_{\tau}$. This concludes the proof. \square

We now provide a proof for the main result of this section.

THEOREM 5.7 (error bound for the approximation of an optimal control). *Let us assume that (A.1)–(A.3), (B.1)–(B.2), (C.1), and (5.7) hold. Let $\bar{q} \in \mathbb{Q}_{ad}$ satisfy the second order optimality condition (i), or equivalently (ii) in Theorem 4.7. Then, there is $h_{\bullet} > 0$ so that the estimate (5.14) holds.*

Proof. We proceed by contradiction: If we assume that (5.14) is false, then there exists $h_{\square} > 0$ such that the estimate (5.15) of Theorem 5.6 holds for every $h \in (0, h_{\square}]$. Based on the instrumental error estimate (5.15), we use the continuous and discrete optimality conditions, (4.11) and (5.2), respectively, to obtain [6, ineq. (4.14)]

$$\begin{aligned} \mathfrak{C} \|\bar{q} - \bar{q}_h\|_{L^2(\Omega)}^2 &\leq [F'_h(\bar{q}_h) - F'(\bar{q}_h)](\bar{q} - \bar{q}_h) + [F'_h(\bar{q}_h) - F'(\bar{q})](\mathcal{P}_h \bar{q} - \bar{q}) \\ &\quad + \left[F'(\bar{q})(\mathcal{P}_h \bar{q} - \bar{q}) + \mu \int_{\Omega} \bar{\eta}(\mathcal{P}_h \bar{q} - \bar{q}) dx \right] + \mu \int_{\Omega} (\bar{\eta} - \bar{\eta}_h)(\bar{q}_h - \bar{q}) dx \\ &\quad + \mu \int_{\Omega} (\bar{\eta}_h - \bar{\eta})(\mathcal{P}_h \bar{q} - \bar{q}) dx := \mathbf{I}_h + \mathbf{II}_h + \mathbf{III}_h + \mathbf{IV}_h + \mathbf{V}_h, \quad \forall h \leq h_{\square}. \end{aligned}$$

In the following, we proceed in several steps and estimate each of the terms \mathbf{I}_h , \mathbf{II}_h , \mathbf{III}_h , \mathbf{IV}_h , and \mathbf{V}_h individually.

Step 1. We first control the term \mathbf{IV}_h . To do this, we first use that $\bar{\eta} \in \partial j(\bar{q})$ and the definition of the subgradient given in (2.1) to obtain

$$\int_{\Omega} \bar{\eta}(\bar{q}_h - \bar{q}) dx \leq \|\bar{q}_h\|_{L^1(\Omega)} - \|\bar{q}\|_{L^1(\Omega)}.$$

Since $\bar{\eta}_h \in \partial j_h(\bar{q}_h)$, the characterization in [6, eq. (4.3)] leads to the conclusion that

$$\int_{\Omega} \bar{\eta}_h(\bar{q} - \bar{q}_h) dx = \int_{\Omega} \bar{\eta}_h \bar{q} dx - \|\bar{q}_h\|_{L^1(\Omega)} \leq \|\bar{q}\|_{L^1(\Omega)} - \|\bar{q}_h\|_{L^1(\Omega)}.$$

Consequently, we can control the term \mathbf{IV}_h as follows:

$$(5.30) \quad \mathbf{IV}_h \leq \mu (\|\bar{q}_h\|_{L^1(\Omega)} - \|\bar{q}\|_{L^1(\Omega)} + \|\bar{q}\|_{L^1(\Omega)} - \|\bar{q}_h\|_{L^1(\Omega)}) = 0.$$

Step 2. We estimate \mathbf{III}_h . To do so, we use the characterization of $F'(\bar{q})$ described in (4.9) and standard properties for the orthogonal projection operator \mathcal{P}_h to obtain

$$(5.31) \quad \begin{aligned} \mathbf{III}_h &= \int_{\Omega} (\bar{p} + \lambda\bar{q} + \mu\bar{\eta} - \mathcal{P}_h(\bar{p} + \lambda\bar{q} + \mu\bar{\eta}))(\mathcal{P}_h\bar{q} - \bar{q})dx \\ &\leq (\|\bar{p} - \mathcal{P}_h\bar{p}\|_{L^2(\Omega)} + \mu\|\bar{\eta} - \mathcal{P}_h\bar{\eta}\|_{L^2(\Omega)}) \|\bar{q} - \mathcal{P}_h\bar{q}\|_{L^2(\Omega)}. \end{aligned}$$

Here, $\bar{p} \in \tilde{H}^s(\Omega) \cap L^\infty(\Omega)$ denotes the solution to (4.8) with u replaced by $\bar{u} = \mathcal{S}\bar{q}$. The control of the error $\|\bar{p} - \mathcal{P}_h\bar{p}\|_{L^2(\Omega)}$ follows from a standard error estimate for \mathcal{P}_h in conjunction with the regularity property $\bar{p} \in H^{s+\kappa-\varepsilon}(\Omega)$ for all $\varepsilon \in (0, s)$, where $\kappa = \frac{1}{2}$ for $\frac{1}{2} < s < 1$ and $\kappa = s - \varepsilon$ for $0 < s \leq \frac{1}{2}$ (see Theorem 4.9). In fact, we have

$$(5.32) \quad \|\bar{p} - \mathcal{P}_h\bar{p}\|_{L^2(\Omega)} \lesssim h^{2\gamma} |\log h|^\nu \Lambda(L, \bar{u}), \quad \gamma = \min\{s, \frac{1}{2}\},$$

where $\nu = \frac{1}{2}$ for $\frac{1}{2} < s < 1$ and $\nu = \frac{1}{2} + \nu_0$ for $0 < s \leq \frac{1}{2}$, with $\nu_0 = \nu_0(\Omega, d) > 0$. The term $\Lambda(L, \bar{u})$ is defined in (4.20). Similarly, we have

$$(5.33) \quad \|\bar{q} - \mathcal{P}_h\bar{q}\|_{L^2(\Omega)} + \|\bar{\eta} - \mathcal{P}_h\bar{\eta}\|_{L^2(\Omega)} \lesssim h^{2\gamma} |\log h|^\nu [1 + \Lambda(L, \bar{u})].$$

If we replace the estimates obtained in (5.32) and (5.33) into (5.31), we obtain

$$(5.34) \quad \mathbf{III}_h \lesssim h^{4\gamma} |\log h|^{2\nu} [1 + \Lambda(L, \bar{u})]^2.$$

Step 3. An estimate for the term \mathbf{V}_h follows easily from the definition of \mathcal{P}_h and the estimate (5.33) derived in the previous step:

$$(5.35) \quad \mathbf{V}_h = \mu \int_{\Omega} (\mathcal{P}_h\bar{\eta} - \bar{\eta})(\mathcal{P}_h\bar{q} - \bar{q})dx \lesssim h^{4\gamma} |\log h|^{2\nu} [1 + \Lambda(L, \bar{u})]^2.$$

Step 4. The aim of this step is to estimate the term \mathbf{I}_h . To accomplish this task, we introduce the variables $\hat{u} \in \tilde{H}^s(\Omega)$ and $\hat{p} \in \tilde{H}^s(\Omega)$, which solve, respectively,

$$\mathcal{A}(\hat{u}, v) + \int_{\Omega} a(x, \hat{u})vdx = \int_{\Omega} \bar{q}_h vdx \quad \forall v \in \tilde{H}^s(\Omega)$$

and

$$\mathcal{A}(\hat{p}, v) + \int_{\Omega} \frac{\partial a}{\partial u}(x, \hat{p})\hat{p}vdx = \int_{\Omega} \frac{\partial L}{\partial u}(x, \hat{p})vdx \quad \forall v \in \tilde{H}^s(\Omega).$$

With these variables, we can rewrite and estimate the term \mathbf{I}_h as follows:

$$(5.36) \quad \mathbf{I}_h = \int_{\Omega} [(\bar{p}_h + \lambda\bar{q}_h) - (\hat{p} + \lambda\bar{q}_h)] (\bar{q} - \bar{q}_h)dx \leq \frac{1}{\mathfrak{C}} \|\bar{p}_h - \hat{p}\|_{L^2(\Omega)}^2 + \frac{\mathfrak{C}}{4} \|\bar{q} - \bar{q}_h\|_{L^2(\Omega)}^2.$$

To obtain the last inequality, we used Young's inequality. Here, $\mathfrak{C} = 2^{-1}\delta$ is the constant that appears in Theorem 5.6.

The rest of this step is dedicated to bound the term $\|\bar{p}_h - \hat{p}\|_{L^2(\Omega)}$. For this purpose, we introduce \tilde{p} as the solution to: Find $\tilde{p} \in \tilde{H}^s(\Omega)$ such that

$$\mathcal{A}(\tilde{p}, v) + \int_{\Omega} \frac{\partial a}{\partial u}(x, \bar{u}_h)\tilde{p}vdx = \int_{\Omega} \frac{\partial L}{\partial u}(x, \bar{u}_h)vdx \quad \forall v \in \tilde{H}^s(\Omega).$$

We note that, in view of (5.7) and the assumptions on a and L all terms in this weak formulation are well-posed. With the variable \tilde{p} at hand, the triangle inequality in $L^2(\Omega)$ yields $\|\bar{p}_h - \hat{p}\|_{L^2(\Omega)} \leq \|\bar{p}_h - \tilde{p}\|_{L^2(\Omega)} + \|\tilde{p} - \hat{p}\|_{L^2(\Omega)}$. To bound $\|\bar{p}_h - \tilde{p}\|_{L^2(\Omega)}$, we use that \bar{p}_h corresponds to the finite element approximation of \tilde{p} within \mathbb{V}_h . Indeed, an application of a suitable modification of Theorem [16, Theorem 6.1] yields

$$(5.37) \quad \|\bar{p}_h - \hat{p}\|_{L^2(\Omega)} \lesssim h^{2\gamma} |\log h|^{2\varphi} \Lambda(L, \bar{u}_h).$$

We note that the same arguments we used to derive the regularity results for \bar{p} in Theorem 4.9 also apply to \tilde{p} with a similar estimate. It remains to bound the term $\|\tilde{p} - \hat{p}\|_{L^2(\Omega)}$. To do this, we first note that $\tilde{p} - \hat{p}$ is such that

$$(5.38) \quad \tilde{p} - \hat{p} \in \tilde{H}^s(\Omega) : \quad \mathcal{A}(\tilde{p} - \hat{p}, v) + \int_{\Omega} \frac{\partial a}{\partial u}(x, \hat{u})(\tilde{p} - \hat{p})v dx \\ = \int_{\Omega} \frac{\partial^2 L}{\partial u^2}(x, u_{\theta})(\bar{u}_h - \hat{u})v dx + \int_{\Omega} \frac{\partial^2 a}{\partial u^2}(x, u_{\theta})(\hat{u} - \bar{u}_h)\tilde{p}v dx$$

for all $v \in \tilde{H}^s(\Omega)$, where $u_{\theta} := \hat{u} + \theta_h(\bar{u}_h - \hat{u})$ and $u_{\vartheta} := \bar{u}_h + \vartheta_h(\hat{u} - \bar{u}_h)$ with $\theta_h, \vartheta_h \in (0, 1)$. Given the assumptions (C.1) and (5.7), the $L^\infty(\Omega)$ -regularity of \tilde{p} , and Hölder's inequality we have that all terms in (5.38) are well-defined. We now invoke a stability estimate for problem (5.38), assumptions (A.3) and (C.1), the $L^\infty(\Omega)$ -regularity of \tilde{p} , and the embedding $H^s(\Omega) \hookrightarrow L^r(\Omega)$, which holds for every $r \leq 2d/(d - 2s)$, to derive the estimates

$$(5.39) \quad \|\tilde{p} - \hat{p}\|_{L^2(\Omega)} \lesssim \|\tilde{p} - \hat{p}\|_s \leq \left\| \left(\frac{\partial^2 L}{\partial u^2}(\cdot, u_{\theta}) - \frac{\partial^2 a}{\partial u^2}(\cdot, u_{\vartheta})\tilde{p} \right) (\bar{u}_h - \hat{u}) \right\|_{H^{-s}(\Omega)} \\ \lesssim \left(\left\| \frac{\partial^2 L}{\partial u^2}(\cdot, u_{\theta}) \right\|_{L^{\frac{d}{s}}(\Omega)} + \|\tilde{p}\|_{L^\infty(\Omega)} \right) \|\bar{u}_h - \hat{u}\|_{L^2(\Omega)} \lesssim h^{2\gamma} |\log h|^{2\varphi} \|\bar{q}_h - a(\cdot, 0)\|_{L^2(\Omega)}.$$

To obtain the last estimate, we used the fact that \bar{u}_h corresponds to the finite element approximation of \hat{u} within \mathbb{V}_h and Theorem 3.2. If we replace the estimates (5.37) and (5.39) in (5.36), we obtain that

$$(5.40) \quad \mathbf{I}_h \leq Ch^{4\gamma} |\log h|^{4\varphi} (\|\bar{q}_h - a(\cdot, 0)\|_{L^2(\Omega)} + \Lambda(L, \bar{u}_h))^2 + \frac{\mathfrak{C}}{4} \|\bar{q} - \bar{q}_h\|_{L^2(\Omega)}^2.$$

Step 5. In this step, we bound \mathbf{II}_h . To do this, we use the definition of the orthogonal projection operator \mathcal{P}_h and the regularity properties of \bar{p} and \bar{q} to obtain

$$\mathbf{II}_h = \int_{\Omega} ((\bar{p}_h + \lambda\bar{q}_h - (\bar{p} + \lambda\bar{q}))(\mathcal{P}_h\bar{q} - \bar{q})) dx = \int_{\Omega} (\bar{p}_h - \bar{p})(\mathcal{P}_h\bar{q} - \bar{q}) dx \\ + \lambda \|\mathcal{P}_h\bar{q} - \bar{q}\|_{L^2(\Omega)}^2 \leq \|\bar{p}_h - \bar{p}\|_{L^2(\Omega)} \|\mathcal{P}_h\bar{q} - \bar{q}\|_{L^2(\Omega)} + \lambda \|\mathcal{P}_h\bar{q} - \bar{q}\|_{L^2(\Omega)}^2.$$

Use the estimates (5.33) and (5.9), as well as Young's inequality, to obtain

$$(5.41) \quad \mathbf{II}_h \leq Ch^{4\gamma} |\log h|^{4\varphi} [1 + \Lambda(L, \bar{u})]^2 + \frac{\mathfrak{C}}{4} \|\bar{q} - \bar{q}_h\|_{L^2(\Omega)}^2.$$

Step 6. Through the collection of (5.30), (5.34), (5.35), (5.40), and (5.41), we conclude that (5.14) holds, which is a contradiction. This concludes the proof. \square

As a corollary, we present the following estimate for $\bar{\eta} - \bar{\eta}_h$.

COROLLARY 5.8 (error bound for the approximation of an optimal subgradient). *In the framework of Theorem 5.7, we have the error bound*

$$(5.42) \quad \|\bar{\eta} - \bar{\eta}_h\|_{L^2(\Omega)} \lesssim h^{2\gamma} |\log h|^{2\varphi} \quad \forall h \leq h_\bullet, \quad \gamma = \min\left\{s, \frac{1}{2}\right\},$$

where $\varphi = \nu$ if $s \neq \frac{1}{2}$, $\varphi = 1 + \nu$ if $s = \frac{1}{2}$, and $\nu \geq \frac{1}{2}$ is the constant in Theorem 3.1.

Proof. The error bound is an immediate consequence of Theorems 5.3 and 5.7. \square

6. A semidiscrete scheme for the optimal control problem. In the following, we propose a semidiscretization strategy based on the variational discretization approach [10]. Here, only the state space is discretized (the control space is not discretized). The semidiscrete approach is as follows: Find $\min J(u_h, \mathbf{q})$ subject to

$$(6.1) \quad \mathcal{A}(u_h, v_h) + \int_{\Omega} a(x, u_h) v_h dx = \int_{\Omega} \mathbf{q} v_h dx \quad \forall v_h \in \mathbb{V}_h,$$

and the control constraints $\mathbf{q} \in \mathbb{Q}_{ad}$. Standard arguments show that there is at least one optimal solution to this problem. Furthermore, as in Theorem 4.4, it can be proved that if $\bar{\mathbf{q}} \in \mathbb{Q}_{ad}$ denotes a local solution, then there exists $\bar{\eta} \in \partial j(\bar{\mathbf{q}})$ such that

$$(6.2) \quad \int_{\Omega} (\bar{p}_h + \lambda \bar{\mathbf{q}} + \mu \bar{\eta})(q - \bar{\mathbf{q}}) dx \geq 0 \quad \forall q \in \mathbb{Q}_{ad}.$$

Here, \bar{p}_h solves the problem (5.3), where \bar{u}_h corresponds to the solution of (6.1) with \mathbf{q} replaced by $\bar{\mathbf{q}}$. As in Theorem 4.5, the following projection formulas can be derived for every $x \in \Omega$:

$$(6.3) \quad \bar{\mathbf{q}}(x) = \Pi_{[\alpha, \beta]}(-\lambda^{-1}(\bar{p}_h(x) + \mu \bar{\eta}(x))), \quad \bar{\mathbf{q}}(x) = 0 \Leftrightarrow |\bar{p}_h(x)| \leq \mu,$$

$$(6.4) \quad \bar{\eta}(x) = \Pi_{[-1, 1]}(-\mu^{-1} \bar{p}_h(x)).$$

Since $\bar{\mathbf{q}}$ and $\bar{\eta}$ implicitly depend on h , we will use the notation $\bar{\mathbf{q}}_h$ and $\bar{\eta}_h$ in the following. Assuming that discrete solutions \bar{u}_h of (6.1) are uniformly bounded in $L^\infty(\Omega)$ and that (A.1)–(A.3), (B.1)–(B.2), and (C.1) hold, we can provide error bounds for the approximation error of the adjoint state and subdifferential variables. The error estimate for the latter is simpler than that in Theorem 5.3 because of the bound $\|\bar{\eta} - \bar{\eta}_h\|_{L^2(\Omega)} \leq \mu^{-1} \|\bar{p} - \bar{p}_h\|_{L^2(\Omega)}$. Furthermore, minor adjustments in the proofs of the Theorems 5.4 and 5.5 lead to the following convergence results.

- Let $h > 0$ and let $\bar{\mathbf{q}}_h \in \mathbb{Q}_{ad}$ be a global solution of the semidiscrete scheme. Then, there exist nonrelabeled subsequences of $\{\bar{\mathbf{q}}_h\}_{h>0}$ such that $\bar{\mathbf{q}}_h \xrightarrow{*} \bar{q}$ in $L^\infty(\Omega)$ as $h \rightarrow 0$ and \bar{q} corresponds to a global solution to (4.1)–(4.2). In addition, the convergence results (5.12) and (5.13) hold.
- If $\bar{q} \in \mathbb{Q}_{ad}$ is a strict local minimum of the control problem (4.1)–(4.2), then there exists a sequence of local minima $\{\bar{\mathbf{q}}_h\}_{0 < h < h_\dagger}$ of the semidiscrete scheme such that (5.12) and (5.13) hold.

We now derive the error bound for the semidiscrete scheme given in (6.6). For this purpose, we let $\{\bar{\mathbf{q}}_h\} \subset \mathbb{Q}_{ad}$ be a sequence of local minima of such a scheme such that $\bar{\mathbf{q}}_h \rightarrow \bar{q}$ in $L^2(\Omega)$ as $h \rightarrow 0$, where $\bar{q} \in \mathbb{Q}_{ad}$ is a local solution to (4.1)–(4.2). In a first step, we provide an instrumental result that is analogous to Theorem 5.6.

THEOREM 6.1 (instrumental error estimate). *Let us assume that (A.1)–(A.3), (B.1)–(B.2), (C.1) and (5.7) hold. Let $\bar{q} \in \mathbb{Q}_{ad}$ satisfy the second order optimality condition (i), or equivalently (ii) in Theorem 4.7. Then, there exists $h_* > 0$ such that*

$$(6.5) \quad \mathcal{E}\|\bar{q} - \bar{\mathbf{q}}_h\|_{L^2(\Omega)}^2 \leq [F'(\bar{\mathbf{q}}_h) - F'(\bar{q})](\bar{\mathbf{q}}_h - \bar{q}) \quad \forall h \in (0, h_*],$$

where $\mathfrak{C} = 2^{-1}\delta$ and δ is the constant that appears in the item (ii) of Theorem 4.7.

Proof. The proof follows analogous arguments as in the proof of Theorem 5.6. The main modifications are the redefinitions of the sets

$$\begin{aligned}\Omega_{1,h} &:= \{x \in \Omega_\mu^{\text{less}} : |\bar{p}_h(x)| \leq \mu\}, & \Omega_{2,h} &:= \{x \in \Omega_\mu^{\text{less}} : |\bar{p}_h(x)| > +\mu\}, \\ \Omega_{3,h} &:= \{x \in \Omega_\mu^+ : \bar{p}_h(x) < -\mu\}, & \Omega_{4,h} &:= \{x \in \Omega_\mu^+ : \bar{p}_h(x) > +\mu\},\end{aligned}$$

and $\Omega_{5,h} := \{x \in \Omega_\mu^+ : |\bar{p}_h(x)| \leq \mu\}$. For the sake of simplicity, we skip the details. \square

We now derive the main result of this section.

THEOREM 6.2 (error bound for the approximation of an optimal control). *Let us assume that (A.1)–(A.3), (B.1)–(B.2), (C.1), and (5.7) hold. Let $\bar{q} \in \mathbb{Q}_{ad}$ satisfy the second order optimality condition (i), or equivalently (ii) in Theorem 4.7. Then, there exists $h_\bullet > 0$ such that*

$$(6.6) \quad \|\bar{q} - \bar{q}_h\|_{L^2(\Omega)} \lesssim h^{2\gamma} |\log h|^{2\varphi} \quad \forall h \leq h_\bullet, \quad \gamma = \min\left\{s, \frac{1}{2}\right\},$$

where $\varphi = \nu$ if $s \neq \frac{1}{2}$, $\varphi = 1 + \nu$ if $s = \frac{1}{2}$, and $\nu \geq \frac{1}{2}$ is the constant in Theorem 3.1.

Proof. We proceed as in the proof of Theorem 5.7 and use the instrumental error bound (6.5) and the variational inequalities (4.11) and (6.2) with $q = \bar{q}_h$ and $q = \bar{q}$, respectively, to obtain (cf. [6, Theorem 5.1])

$$\mathfrak{C} \|\bar{q} - \bar{q}_h\|_{L^2(\Omega)}^2 \leq [F'_h(\bar{q}_h) - F'(\bar{q}_h)](\bar{q} - \bar{q}_h) + \mu \int_\Omega (\bar{\eta} - \bar{\eta}_h)(\bar{q}_h - \bar{q}) dx \quad \forall h \leq h_*.$$

We immediately notice that the first and second terms on the right-hand side of the previous inequality correspond to \mathbf{I}_h and \mathbf{IV}_h , respectively, from the proof of Theorem 5.7. These terms, \mathbf{I}_h and \mathbf{IV}_h , are estimated in (5.40) and (5.30), respectively. \square

7. Numerical examples. We present a numerical experiment that illustrates the performance of the fully and semidiscrete methods presented Sections 5 and 6, respectively, when used to approximate a solution of the control problem (4.1)–(4.2). A MATLAB implementation is used for the experiment, and the methods are solved using a semi-smooth Newton method.

The setting of the experiment is as follows: we set $d = 2$, $\Omega = B(0, 1)$, and $\lambda = 1$, where $B(0, 1)$ denotes the unit disc. We let $a(\cdot, u) = u^3$ and $L(\cdot, u) = (u - u_\Omega)^2/2$, where u_Ω is such that the exact optimal state and the optimal adjoint state are

$$(7.1) \quad \bar{u}(x) = \bar{p}(x) = (2^{2s}\Gamma^2(1+s))^{-1}(1 - |x|^2)_+^s, \quad t_+ = \max\{0, t\}.$$

We also consider $a = -1$, $b = 1$, and $s \in \{0.2, 0.4, 0.6, 0.8\}$. Additionally, for $s \leq 0.5$, we set $\mu = 0.6$, and for $s > 0.5$, we choose $\mu = 0.25$.

Figures 7.1, 7.2, and 7.3 show the results for the fully discrete and semidiscrete schemes. Figure 7.1 shows the experimental convergence rates for $\|\bar{u} - \bar{u}_h\|_s$ and $\|\bar{p} - \bar{p}_h\|_s$ for $s \in \{0.2, 0.4, 0.6, 0.8\}$. The experimental convergence rates for $\|\bar{u} - \bar{u}_h\|_{L^2(\Omega)}$ and $\|\bar{p} - \bar{p}_h\|_{L^2(\Omega)}$ are shown in Figure 7.2, while the results for $\|\bar{q} - \bar{q}_h\|_{L^2(\Omega)}$ and $\|\bar{\eta} - \bar{\eta}_h\|_{L^2(\Omega)}$ are shown in Figure 7.3. It can be observed that when $s \geq 0.5$ the experimental convergence rates for all involved approximation errors are in agreement with the error estimates obtained in Sections 5.2 and 6. However, when $s < 0.5$, the reported experimental convergence rates exceed those predicted in our manuscript. We present a discussion in the following two Remarks.

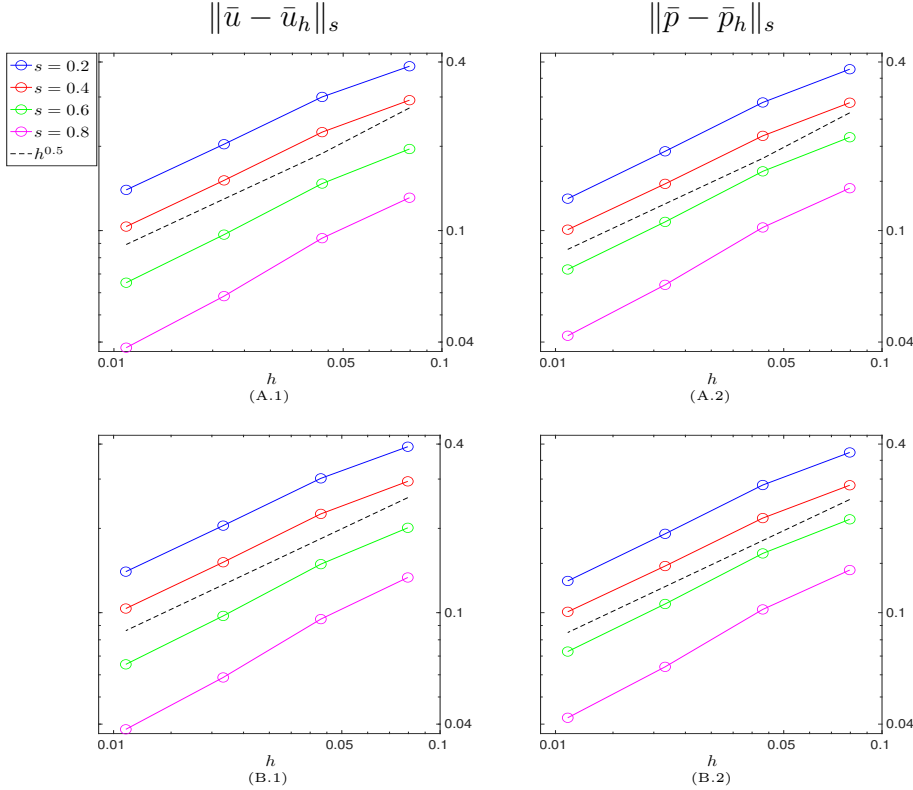


FIG. 7.1. Experimental rates of convergence for $\|\bar{u} - \bar{u}_h\|_s$ and $\|\bar{p} - \bar{p}_h\|_s$ considering the fully discrete (A.1)–(A.2) and semidiscrete schemes (B.1)–(B.2) for $s \in \{0.2, 0.4, 0.6, 0.8\}$.

Remark 7.1 (convergence rates: state and adjoint variables). The Figures 7.1 and 7.2 show that the experimental convergence rates for the approximation errors of the state and adjoint variables exceed the rates predicted by the combination of Theorems 5.3, 5.7, and 6.2 when $s < 0.5$, but agree in terms of the regularity

$$(7.2) \quad H^{s+\frac{1}{2}-\epsilon}(\Omega), \quad 0 < \epsilon < s + \frac{1}{2}.$$

The error bounds in Theorem 5.3 are based on the regularity results of Theorem 4.9, which in turn are inspired by the estimates in Theorem 3.1 ([4, Theorem 2.1]). If $s \in (0, 0.5]$, $\bar{u}, \bar{p} \in H^{2s-2\epsilon}(\Omega)$ for every $\epsilon \in (0, s)$, which is weaker than (7.2). As explained in [4, page 1921], one expects the solutions to be smoother than just $H^{2s}(\Omega)$ if the corresponding forcing term belongs to $H^r(\Omega)$, for some $r > 0$; however, such a result of higher regularity cannot be derived from [4, Theorem 2.1]. It is important to note that in our particular setting \bar{u} and \bar{p} defined in (7.1) satisfy (7.2). Regarding the approximation in the $L^2(\Omega)$ -norm, we note that it behaves as prescribed by the property of maximal regularity (7.2), but limited to the rate $\mathcal{O}(h)$; see [4, §6.1].

Remark 7.2 (convergence rates: control and subgradient variables). Figure 7.3 shows the experimental convergence rates for the approximation error of the control and subgradient variables for both the fully discrete and the semidiscrete schemes. As for the approximation of the control variable, the experimental convergence rate $\mathcal{O}(h)$ is observed for both methods for $s \in \{0.2, 0.4, 0.6, 0.8\}$, which exceeds the expected rate when $s < 1/2$. This behaviour is due to the sparse structure of our optimal

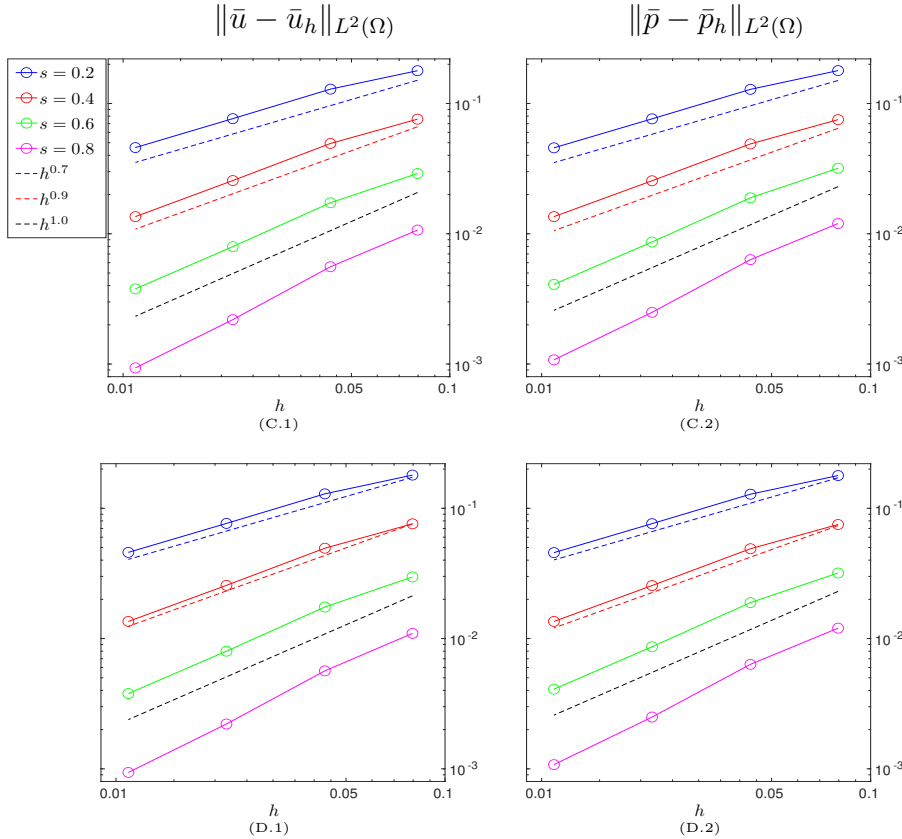


FIG. 7.2. Experimental rates of convergence for $\|\bar{u} - \bar{u}_h\|_{L^2(\Omega)}$ and $\|\bar{p} - \bar{p}_h\|_{L^2(\Omega)}$ considering the fully discrete (C.1)–(C.2) and semidiscrete schemes (D.1)–(D.2) for $s \in \{0.2, 0.4, 0.6, 0.8\}$.

control problem: the optimal control variable \bar{q} is supported in a small region $\omega \subset \Omega$; see Figure 7.4 for a graphical representation of this comment in a suitable finite element approximation. Consequently, the singularities near the boundary that \bar{q} could have as a consequence of the singular behaviour of \bar{u} and \bar{p} are not present. This is in contrast to the behaviour of $\bar{\eta}$; see Figure 7.4 for a graphical representation of a suitable finite element approximation. The latter also explains, in view of the comments in Remark 7.1, the experimentally determined convergence rates obtained for the error within the approximation of $\bar{\eta}$.

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] A. ALLENDES, F. FUICA, AND E. OTÁROLA, *Adaptive finite element methods for sparse PDE-constrained optimization*, IMA J. Numer. Anal., 40 (2020), pp. 2106–2142, <https://doi.org/10.1093/imanum/drz025>.
 [3] A. BONITO, J. P. BORTHAGARAY, R. H. NOCHETTO, E. OTÁROLA, AND A. J. SALGADO, *Numerical methods for fractional diffusion*, Comput. Vis. Sci., 19 (2018), pp. 19–46, <https://doi.org/10.1007/s00791-018-0289-y>.
 [4] J. P. BORTHAGARAY, D. LEYKEKHMAN, AND R. H. NOCHETTO, *Local energy estimates for the fractional Laplacian*, SIAM J. Numer. Anal., 59 (2021), pp. 1918–1947, <https://doi.org/>

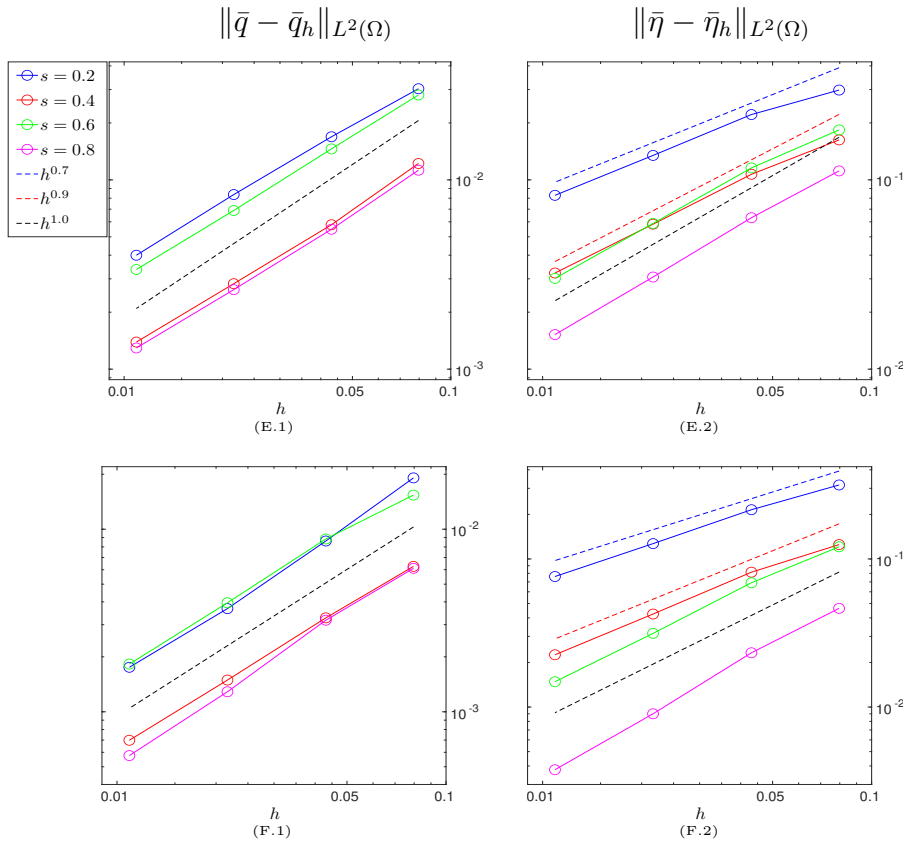


FIG. 7.3. Experimental rates of convergence for $\|\bar{q} - \bar{q}_h\|_{L^2(\Omega)}$ and $\|\bar{\eta} - \bar{\eta}_h\|_{L^2(\Omega)}$ considering the fully discrete (E.1)–(E.2) and semidiscrete schemes (F.1)–(F.2) for $s \in \{0.2, 0.4, 0.6, 0.8\}$.

10.1137/20M1335509.

- [5] E. CASAS, R. HERZOG, AND G. WACHSMUTH, *Approximation of sparse controls in semilinear equations by piecewise linear functions*, Numer. Math., 122 (2012), pp. 645–669, <https://doi.org/10.1007/s00211-012-0475-7>.
- [6] E. CASAS, R. HERZOG, AND G. WACHSMUTH, *Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional*, SIAM J. Optim., 22 (2012), pp. 795–820, <https://doi.org/10.1137/110834366>.
- [7] F. CLARKE, *Functional analysis, calculus of variations and optimal control*, vol. 264 of Graduate Texts in Mathematics, Springer, London, 2013, <https://doi.org/10.1007/978-1-4471-4820-3>.
- [8] M. D’ELIA, C. GLUSA, AND E. OTÁROLA, *A priori error estimates for the optimal control of the integral fractional Laplacian*, SIAM J. Control Optim., 57 (2019), pp. 2775–2798, <https://doi.org/10.1137/18M1219989>.
- [9] E. DI NEZZA, G. PALATUCCI, AND E. VALDINOCI, *Hitchhiker’s guide to the fractional Sobolev spaces*, Bull. Sci. Math., 136 (2012), pp. 521–573, <https://doi.org/10.1016/j.bulsci.2011.12.004>.
- [10] M. HINZE, *A variational discretization concept in control constrained optimization: the linear-quadratic case*, Comput. Optim. Appl., 30 (2005), pp. 45–61, <https://doi.org/10.1007/s10589-005-4559-5>.
- [11] J.-L. LIONS AND E. MAGENES, *Non-homogeneous boundary value problems and applications. Vol. I*, Die Grundlehren der mathematischen Wissenschaften, Band 182, Springer-Verlag, New York-Heidelberg, 1972. Translated from the French by P. Kenneth.
- [12] W. MCLEAN, *Strongly elliptic systems and boundary integral equations*, Cambridge University Press, Cambridge, 2000.

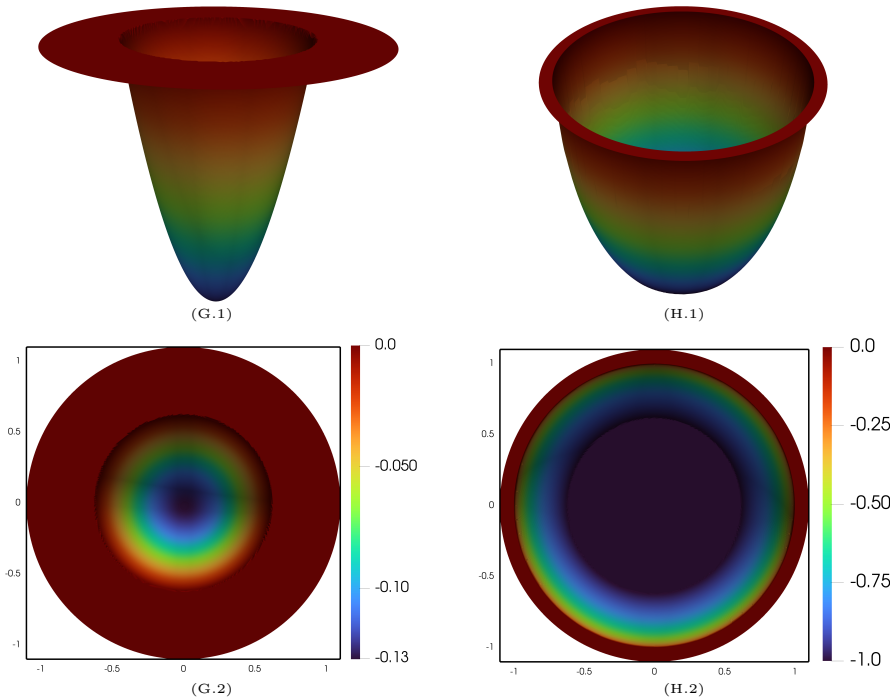


FIG. 7.4. Finite element solutions \bar{q}_h (left) and $\hat{\eta}_h$ (right), obtained by the semidiscrete scheme with $s = 0.4$. The sparse behavior in the control variable \bar{q}_h is evident. In addition, a singular behavior can be observed for $\hat{\eta}_h$ near the boundary $\partial\Omega$.

[13] P. OSWALD, *On the boundedness of the mapping $f \rightarrow |f|$ in Besov spaces*, Comment. Math. Univ. Carolin., 33 (1992), pp. 57–66.

[14] E. OTÁROLA, *An adaptive finite element method for the sparse optimal control of fractional diffusion*, Numer. Methods Partial Differential Equations, 36 (2020), pp. 302–328, <https://doi.org/10.1002/num.22429>.

[15] E. OTÁROLA, *Fractional semilinear optimal control: optimality conditions, convergence, and error analysis*, SIAM J. Numer. Anal., 60 (2022), pp. 1–27, <https://doi.org/10.1137/20M1356294>.

[16] E. OTÁROLA, *Error estimates for fractional semilinear optimal control on Lipschitz polytopes*, Appl. Math. Optim., 88 (2023), pp. Paper No. 40, 32, <https://doi.org/10.1007/s00245-023-10009-1>.

[17] E. OTÁROLA AND A. J. SALGADO, *Sparse optimal control for fractional diffusion*, Comput. Methods Appl. Math., 18 (2018), pp. 95–110, <https://doi.org/10.1515/cmam-2017-0030>.

[18] W. SCHIROTZEK, *Nonsmooth analysis*, Universitext, Springer, Berlin, 2007, <https://doi.org/10.1007/978-3-540-71333-3>.

[19] G. STADLER, *Elliptic optimal control problems with L^1 -control cost and applications for the placement of control devices*, Comput. Optim. Appl., 44 (2009), pp. 159–181, <https://doi.org/10.1007/s10589-007-9150-9>.

[20] L. TARTAR, *An introduction to Sobolev spaces and interpolation spaces*, vol. 3 of Lecture Notes of the Unione Matematica Italiana, Springer, Berlin; UMI, Bologna, 2007.

[21] F. TRÖLTZSCH, *Optimal control of partial differential equations*, American Mathematical Society, Providence, RI, 2010, <https://doi.org/10.1090/gsm/112>.

[22] G. WACHSMUTH AND D. WACHSMUTH, *Convergence and regularization results for optimal control problems with sparsity functional*, ESAIM Control Optim. Calc. Var., 17 (2011), pp. 858–886, <https://doi.org/10.1051/cocv/2010027>.

[23] F. WANG, Q. WANG, AND Z. ZHOU, *Adaptive finite element approximation of sparse optimal control with integral fractional laplacian*, 2023, <https://arxiv.org/abs/2309.08130>.