

AN A POSTERIORI ERROR ANALYSIS OF AN ELLIPTIC OPTIMAL CONTROL PROBLEM IN MEASURE SPACE

FRANCISCO FUICA, ENRIQUE OTÁROLA, AND ABNER J. SALGADO

ABSTRACT. We propose an a posteriori error estimator for an optimal control problem with sparsity promoting cost: the control variable lies in the space of regular Borel measures. We consider a solution technique that relies on the discretization of the control variable as a linear combination of Dirac measures. The proposed a posteriori error estimator can be decomposed into the sum of two contributions: an error estimator in the maximum norm for the discretization of the adjoint equation and an estimator in the L^2 -norm that accounts for the approximation of the state equation. We prove that the designed error estimator is locally efficient and we explore its reliability properties. The analysis is valid for two and three-dimensional domains. We illustrate the theory with numerical examples.

1. INTRODUCTION

This work is dedicated to the design and analysis of an efficient a posteriori error estimator for an elliptic optimal control problem with sparsity promoting cost: the control variable is sought in the space of regular Borel measures. To make matters precise, for $d \in \{2, 3\}$, we let $\Omega \subset \mathbb{R}^d$ be an open, bounded and convex polytopal domain; the boundary of Ω is denoted by $\partial\Omega$. Given a desired state $y_d \in L^2(\Omega)$, and a sparsity parameter $\alpha > 0$, we introduce the cost functional

$$(1.1) \quad J(y, u) := \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \alpha \|u\|_{\mathcal{M}(\Omega)},$$

and thus define our optimal control problem with sparsity promoting cost as follows:
Find

$$(1.2) \quad \min J(y, u)$$

subject to the linear and elliptic PDE

$$(1.3) \quad -\Delta y = u \text{ in } \Omega, \quad y = 0 \text{ on } \partial\Omega.$$

Notice that the control variable u lies in the space of regular Borel measures $\mathcal{M}(\Omega)$.

The design and analysis of solution techniques for optimal control problems that induce a sparse structure in the control variable have been widely studied in the literature over the last decade. The first work that provides an analysis for this class of problems is [32]; the sparsity arises from the consideration of a $L^1(\Omega)$ -control cost term in the quadratic cost functional. The author of [32] studied a regularized problem, derived optimality conditions, proposed and analyzed a semismooth

Key words and phrases. sparse controls, a posteriori error analysis, adaptive finite elements.

FF is supported by USM through Programa de Incentivos a la Iniciación Científica (PIIC). EO is partially supported by CONICYT through FONDECYT project 11180193. AJS is partially supported by NSF grant DMS-1720213.

Newton method. Solution techniques based on finite element methods have been proposed and analyzed in [37] when the state equation is linear and in [7, 9, 16] when the state equation is a semilinear elliptic PDE. For an up-to-date overview of the theory we refer the interested reader to [5]. The extension of the theory to the evolutionary case has been recently explored in [6, 8, 12, 17].

Regarding the optimal control problem (1.2)–(1.3), and to the best of our knowledge, the first work that provided an analysis was [18]. In this work, the authors analyzed elliptic control problems with measures and functions of bounded variation as controls; existence and uniqueness of the corresponding predual problems were discussed together with the solution of the optimality systems by a semismooth Newton method. Subsequently, in [14], the authors address the feasibility of optimal source placement by optimal control in measure spaces: they extend [18] by including partial observation, control on subdomains and non-negativity properties of the controls. A numerical scheme based on finite element techniques was later proposed in [10]. The space $\mathcal{M}(\Omega)$ was discretized by the set of linear combinations of Dirac masses at interior mesh points. The authors proved convergence of the scheme and provided error estimates. On the basis of this discretization scheme, improved error estimates were obtained in [31]. However, these error estimates, are not optimal in terms of approximation. This is due to the fact that the state variable exhibits reduced regularity properties. It is thus just natural to propose adaptive finite element methods (AFEMs) to efficiently resolve the optimal control problem (1.2)–(1.3) and recover optimal rates of convergence for the state variable.

AFEMs are a fundamental numerical instrument in science and engineering that allows for the resolution of PDEs with relatively modest computational resources. They are known to outperform classical FEM in practice and deliver optimal convergence rates when the latter cannot. To extract the local errors incurred by FEM, and thus be able to equidistribute them, AFEMs rely on *a posteriori error estimators*, which are computable quantities, that depend on the discrete solution and data. In contrast to the well established theory for linear elliptic PDEs, the design and analysis of a posteriori error estimators for optimal control problems are being currently developed. In view of their inherent nonlinear feature the analysis involves more arguments and technicalities. To the best of our knowledge, the only work that provides an advance concerning the a posteriori error analysis for (1.2)–(1.3) is [15]. In this reference, the authors propose a functional error estimator, and prove that its square root yields an upper bound for the approximation error of the state variable and the error between the discrete and continuous cost functionals [15, Section 5]; an efficiency analysis, however, is not provided.

In light of the discussion given above, the main objective of this work is to propose and analyze an efficient a posteriori error estimator for the optimal control problem (1.2)–(1.3). We consider a solution technique for (1.2)–(1.3) that relies on the discretization of the state and adjoint variables with piecewise linear functions, whereas the control variable is discretized using the framework presented in [10, Section 3]. The proposed a posteriori error estimator only accounts for the discretization of the state and adjoint variables. We measure the error of the state variable in $L^2(\Omega)$ -norm and the error of the adjoint variable in $L^\infty(\Omega)$ -norm, and we derive local efficiency results. We also explore the reliability properties of the designed error estimator. On the basis of the constructed a posteriori error estimator, we also design simple adaptive strategies that yield optimal rates of convergence for

the numerical examples that we perform. We would like to mention that our error estimator is simpler than the one considered in [15, Section 5]. In addition, and in contrast to [15, Section 5], the error indicator that we consider for the adjoint variable in the $L^\infty(\Omega)$ -norm allows for unbounded forcing terms. This is of importance since, as it can be observed from (2.3), the adjoint equation has $\bar{y} - y_d$ as a forcing term and, in general, $\bar{y} - y_d \notin L^\infty(\Omega)$. Additional assumptions must be imposed on y_d in order to have that $\bar{y} - y_d \in L^\infty(\Omega)$ [31, Theorem 2.5].

The outline of this manuscript is as follows. In Section 2 we present existence and uniqueness results together with first-order optimality conditions. In Section 3 we present the finite element discretization of our optimal control problem; it relies on the discretization of the state and adjoint equations by using piecewise linear functions, whereas the control variable is approximated with Dirac deltas. The a posteriori error analysis of elliptic problems with delta sources together with maximum-norm a posteriori error estimation of elliptic problems are reviewed in Section 4. The core of our work is Section 5, where we design an a posteriori error estimator and study reliability and efficiency results. We conclude, in Section 6, with a series of numerical examples that illustrate our theory.

Throughout this work $d \in \{2, 3\}$. If \mathcal{X} and \mathcal{Y} are normed vector spaces, 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. If X is a function space over the domain $G \subset \mathbb{R}^d$, we denote by $\langle \cdot, \cdot \rangle_G$ the duality pairing between X' and X . The relation $a \lesssim b$ indicates that $a \leq Cb$, with a nonessential constant C that might change at each occurrence.

2. THE OPTIMAL CONTROL PROBLEM IN MEASURE SPACE

In this section we review some of the main results related to the existence and uniqueness of solutions for problem (1.2)–(1.3). In addition, we present first-order necessary and sufficient optimality conditions.

We recall that the space of regular Borel measures $\mathcal{M}(\Omega)$ can be identified, by the Riesz Theorem, with the dual of the space of continuous functions that vanish on the boundary $\partial\Omega$, which from now on we shall denote by $C_0(\Omega)$. Given a measure $\mu \in \mathcal{M}(\Omega)$, we have that

$$(2.1) \quad \|\mu\|_{\mathcal{M}(\Omega)} = \sup_{\substack{\varphi \in C_0(\Omega) \\ \|\varphi\|_{C_0(\Omega)} \leq 1}} \langle \mu, \varphi \rangle_\Omega = \sup_{\substack{\varphi \in C_0(\Omega) \\ \|\varphi\|_{C_0(\Omega)} \leq 1}} \int_{\Omega} \varphi \, d\mu.$$

Given $u \in \mathcal{M}(\Omega)$, we define the weak solution of problem (1.3) as follows:

$$(2.2) \quad y \in W_0^{1,r}(\Omega) : \quad (\nabla y, \nabla v)_{L^2(\Omega)} = \langle u, v \rangle_\Omega \quad \forall v \in W_0^{1,r'}(\Omega),$$

where $1 \leq r < d/(d-1)$ and r' denotes the conjugate exponent of r . Problem (2.2) has a unique solution $y \in W_0^{1,r}(\Omega)$ that satisfies [11, Theorem 4], [33, Théorème 9.1]

$$\|\nabla y\|_{L^r(\Omega)} \lesssim \|u\|_{\mathcal{M}(\Omega)}.$$

We notice that, in view of the fact that $W_0^{1,r}(\Omega) \hookrightarrow L^2(\Omega)$ for $2d/(d+2) \leq r < d/(d-1)$, the cost functional J , which is defined by (1.2), is well-defined.

The following existence result follows from [18, Proposition 2.2].

Theorem 2.1 (existence and uniqueness). *The sparse optimal control problem (1.2)–(1.3) has a unique solution $(\bar{y}, \bar{u}) \in W_0^{1,r}(\Omega) \times \mathcal{M}(\Omega)$ for $1 \leq r < d/(d-1)$.*

The next result establishes optimality conditions for problem (1.2)–(1.3); see [10, Theorem 2.1] and [31, Theorem 2.2].

Theorem 2.2 (optimality conditions). *Let $(\bar{y}, \bar{u}) \in W_0^{1,r}(\Omega) \times \mathcal{M}(\Omega)$ be the (unique) solution of (1.2)–(1.3). Then, there exists a unique element $\bar{p} \in H^2(\Omega) \cap H_0^1(\Omega)$ satisfying*

$$(2.3) \quad (\nabla w, \nabla \bar{p})_{L^2(\Omega)} = (\bar{y} - y_d, w)_{L^2(\Omega)} \quad \forall w \in H_0^1(\Omega),$$

such that, for all $u \in \mathcal{M}(\Omega)$,

$$(2.4) \quad -\langle u - \bar{u}, \bar{p} \rangle_\Omega + \alpha \|\bar{u}\|_{\mathcal{M}(\Omega)} \leq \alpha \|u\|_{\mathcal{M}(\Omega)}.$$

In addition,

$$(2.5) \quad \|\bar{p}\|_{C_0(\Omega)} = \alpha \text{ if } \bar{u} \neq 0, \quad \|\bar{p}\|_{C_0(\Omega)} \leq \alpha \text{ if } \bar{u} = 0.$$

As a conclusion, the pair $(\bar{y}, \bar{u}) \in W_0^{1,r}(\Omega) \times \mathcal{M}(\Omega)$ is optimal for (1.2)–(1.3) if and only if the triplet $(\bar{y}, \bar{p}, \bar{u}) \in W_0^{1,r}(\Omega) \times H_0^1(\Omega) \times \mathcal{M}(\Omega)$ satisfies the following optimality system:

$$(2.6) \quad \begin{cases} (\nabla \bar{y}, \nabla v)_{L^2(\Omega)} = \langle \bar{u}, v \rangle_\Omega & \forall v \in W_0^{1,r'}(\Omega), \\ (\nabla w, \nabla \bar{p})_{L^2(\Omega)} = (\bar{y} - y_d, w)_{L^2(\Omega)} & \forall w \in H_0^1(\Omega), \\ -\langle u - \bar{u}, \bar{p} \rangle_\Omega + \alpha \|\bar{u}\|_{\mathcal{M}(\Omega)} \leq \alpha \|u\|_{\mathcal{M}(\Omega)} & \forall u \in \mathcal{M}(\Omega). \end{cases}$$

3. FINITE ELEMENT DISCRETIZATION

We recall the finite element approximation of the sparse optimal control problem (1.2)–(1.3) developed in [10, Section 3]. In addition, we present convergence rates for such a discretization [31].

We begin by introducing some ingredients of standard finite element approximation [13, 24]. Let $\mathcal{T} = \{T\}$ be a conforming partition of $\bar{\Omega}$ into simplices T with size $h_T := \text{diam}(T)$, and set $h_{\mathcal{T}} := \max_{T \in \mathcal{T}} h_T$. Let us denote by \mathbb{T} the collection of conforming and shape regular meshes that are refinements of \mathcal{T}_0 , where \mathcal{T}_0 represents an initial mesh. Given $\mathcal{T} \in \mathbb{T}$, we denote by $N_{\mathcal{T}}$ the number of interior nodes of \mathcal{T} and by $\{x_i\}_{i=1}^{N_{\mathcal{T}}}$ the set of interior nodes of \mathcal{T} .

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

$$(3.1) \quad \mathbb{V}(\mathcal{T}) = \{v_{\mathcal{T}} \in C_0(\Omega) : v_{\mathcal{T}}|_T \in \mathbb{P}_1(T) \forall T \in \mathcal{T}\}.$$

Given a node x_i in the mesh \mathcal{T} , we introduce the function $\phi_i \in \mathbb{V}(\mathcal{T})$, which is such that $\phi_i(x_j) = \delta_{ij}$ for all $j = 1, \dots, N_{\mathcal{T}}$. The set $\{\phi_i\}_{i=1}^{N_{\mathcal{T}}}$ is the so-called Courant basis of the space $\mathbb{V}(\mathcal{T})$.

With this setting at hand, we define the following discrete version of the optimal control problem (1.2)–(1.3): Find $(y_{\mathcal{T}}, u) \in \mathbb{V}(\mathcal{T}) \times \mathcal{M}(\Omega)$ that minimizes

$$(3.2) \quad J(y_{\mathcal{T}}, u)$$

subject to the discrete state equation

$$(3.3) \quad (\nabla y_{\mathcal{T}}, \nabla v_{\mathcal{T}})_{L^2(\Omega)} = \langle u, v_{\mathcal{T}} \rangle_\Omega \quad \forall v_{\mathcal{T}} \in \mathbb{V}(\mathcal{T}).$$

Notice that, since the control variable is not discretized, the solution technique (3.2)–(3.3) corresponds to an instance of the so-called variational discretization approach [26].

The discrete optimal control problem (3.2)–(3.3) has a solution. In contrast to the continuous case, the discrete version of the control-to-state map $S_{\mathcal{T}} : \mathcal{M}(\Omega) \ni u \mapsto y_{\mathcal{T}} \in \mathbb{V}(\mathcal{T})$ is not injective [10, 31]. This and the fact that the norm $\|\cdot\|_{\mathcal{M}(\Omega)}$ is not strictly convex imply that J is not strictly convex. As a consequence, the uniqueness of the control variable cannot be guaranteed. However, among all the possible optimal controls, a special solution can be explicitly characterized using a particular discrete space [10, Section 3]:

$$(3.4) \quad \mathbb{U}(\mathcal{T}) := \left\{ u_{\mathcal{T}} \in \mathcal{M}(\Omega) : u_{\mathcal{T}} = \sum_{i=1}^{N_{\mathcal{T}}} u_i \delta_{x_i}, \quad u_i \in \mathbb{R}, \quad 1 \leq i \leq N_{\mathcal{T}} \right\}.$$

Notice that this discrete space consists of linear combinations of Dirac measures associated to the interior nodes of the mesh \mathcal{T} . Finally, we introduce the following operator [10, Section 3]:

$$(3.5) \quad \Lambda_{\mathcal{T}} : \mathcal{M}(\Omega) \rightarrow \mathbb{U}(\mathcal{T}), \quad \Lambda_{\mathcal{T}}(u) = \sum_{i=1}^{N_{\mathcal{T}}} \langle u, \phi_i \rangle_{\Omega} \delta_{x_i}.$$

With the previous definitions at hand, we are in position to present the following results related to the convergence of the discrete solutions; see [10, Theorems 3.2 and 3.5] and [31, Theorem 3.1].

Theorem 3.1 (convergence). *Among all the optimal controls of problem (3.2)–(3.3), there exists a unique $\bar{u}_{\mathcal{T}} \in \mathbb{U}(\mathcal{T})$. Any other optimal control $\tilde{u}_{\mathcal{T}} \in \mathcal{M}(\Omega)$ of (3.2)–(3.3) satisfies that $\Lambda_{\mathcal{T}} \tilde{u}_{\mathcal{T}} = \bar{u}_{\mathcal{T}}$. In addition, we have the following convergence properties as $h_{\mathcal{T}} \rightarrow 0$:*

$$(3.6) \quad \bar{u}_{\mathcal{T}} \xrightarrow{*} \bar{u} \text{ in } \mathcal{M}(\Omega), \quad \|\bar{u}_{\mathcal{T}}\|_{\mathcal{M}(\Omega)} \rightarrow \|\bar{u}\|_{\mathcal{M}(\Omega)}, \quad \|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(\Omega)} \rightarrow 0.$$

The following results present optimality conditions for the discrete optimal control problem (3.2)–(3.3); see [31, Theorem 3.2].

Theorem 3.2 (discrete optimality conditions). *Let $(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}) \in \mathbb{V}(\mathcal{T}) \times \mathbb{U}(\mathcal{T})$ be the discrete solution, as in Theorem 3.1. Then there exists a unique discrete adjoint state $\bar{p}_{\mathcal{T}} \in \mathbb{V}(\mathcal{T})$ such that*

$$(3.7) \quad (\nabla w_{\mathcal{T}}, \nabla \bar{p}_{\mathcal{T}})_{L^2(\Omega)} = (\bar{y}_{\mathcal{T}} - y_d, w_{\mathcal{T}})_{L^2(\Omega)} \quad \forall w_{\mathcal{T}} \in \mathbb{V}(\mathcal{T}),$$

and that satisfies

$$(3.8) \quad -\langle u - \bar{u}_{\mathcal{T}}, \bar{p}_{\mathcal{T}} \rangle_{\Omega} + \alpha \|\bar{u}_{\mathcal{T}}\|_{\mathcal{M}(\Omega)} \leq \alpha \|u\|_{\mathcal{M}(\Omega)} \quad \forall u \in \mathcal{M}(\Omega).$$

Consequently, the discrete pair $(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}) \in \mathbb{V}(\mathcal{T}) \times \mathbb{U}(\mathcal{T})$ is optimal for (3.2)–(3.3) if and only if the triplet $(\bar{y}_{\mathcal{T}}, \bar{p}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}) \in \mathbb{V}(\mathcal{T}) \times \mathbb{V}(\mathcal{T}) \times \mathbb{U}(\mathcal{T})$ solves

$$(3.9) \quad \begin{cases} (\nabla \bar{y}_{\mathcal{T}}, \nabla v_{\mathcal{T}})_{L^2(\Omega)} = \langle \bar{u}_{\mathcal{T}}, v_{\mathcal{T}} \rangle_{\Omega} & \forall v_{\mathcal{T}} \in \mathbb{V}(\mathcal{T}), \\ (\nabla w_{\mathcal{T}}, \nabla \bar{p}_{\mathcal{T}})_{L^2(\Omega)} = (\bar{y}_{\mathcal{T}} - y_d, w_{\mathcal{T}})_{L^2(\Omega)} & \forall w_{\mathcal{T}} \in \mathbb{V}(\mathcal{T}), \\ -\langle u - \bar{u}_{\mathcal{T}}, \bar{p}_{\mathcal{T}} \rangle_{\Omega} + \alpha \|\bar{u}_{\mathcal{T}}\|_{\mathcal{M}(\Omega)} \leq \alpha \|u\|_{\mathcal{M}(\Omega)} & \forall u \in \mathcal{M}(\Omega). \end{cases}$$

To conclude this section, we present a priori error estimates for the approximation of the optimal state variable. To state it, we will need an extra assumption on the desired state y_d ; see [31, Section 4]. Let us assume that

$$(3.10) \quad y_d \in L^{\infty}(\Omega) \text{ for } d = 2, \quad y_d \in L^3(\Omega) \text{ for } d = 3.$$

Theorem 3.3 (a priori error estimates). *Let y_d satisfy (3.10). Let $(\bar{y}, \bar{u}) \in W_0^{1,r}(\Omega) \times \mathcal{M}(\Omega)$ be the solution of problem (1.2)–(1.3), and let $(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}) \in \mathbb{V}(\mathcal{T}) \times \mathbb{U}(\mathcal{T})$ the solution of problem (3.2)–(3.3), given as in Theorem 3.1. Then there holds*

$$(3.11) \quad \|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(\Omega)} \lesssim h_{\mathcal{T}}^{2-d/2} |\ln h_{\mathcal{T}}|^{\frac{\gamma}{2}},$$

$$(3.12) \quad \|\bar{u} - \bar{u}_{\mathcal{T}}\|_{H^{-2}(\Omega)} \lesssim h_{\mathcal{T}}^{2-d/2} |\ln h_{\mathcal{T}}|^{\frac{\gamma}{2}},$$

with $\gamma = \frac{7}{2}$ for $d = 2$, and $\gamma = 1$ for $d = 3$.

Proof. We refer the reader to [31, Theorem 4.4] for a proof of (3.11) and [31, Corollary 4.5] for (3.12). \square

Theorem 3.4 (a priori error estimates). *Let $y_d \in L^\infty(\Omega)$, which implies that the solution $(\bar{y}, \bar{u}) \in W_0^{1,r}(\Omega) \times \mathcal{M}(\Omega)$ of problem (1.2)–(1.3) satisfies that $\bar{y} \in H_0^1(\Omega) \cap L^\infty(\Omega)$. If $(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}) \in \mathbb{V}(\mathcal{T}) \times \mathbb{U}(\mathcal{T})$ denote the solution of problem (3.2)–(3.3), given as in Theorem 3.1, then*

$$(3.13) \quad \|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(\Omega)} \lesssim h_{\mathcal{T}} |\ln h_{\mathcal{T}}|^{\frac{\rho}{2}},$$

with $\rho = 2$ for $d = 2$ and $\rho = 11/4$ for $d = 3$.

Proof. We refer the reader to [31, Theorem 5.1]. \square

4. A POSTERIORI ERROR ANALYSIS FOR THE LAPLACIAN

In the next section we will construct an a posteriori error estimator for the optimal control problem with sparsity promoting cost (1.2)–(1.3) that will be based on two error contributions: one associated to the discretization of the state equation (2.2) and another one related to the discretization of the adjoint equation (2.3). In order to design these contributions, and in an effort to make the presentation of the material as clear as possible, in this section we briefly review a posteriori error estimates for Poisson problems. We first review the L^2 a posteriori error estimator, developed in [3], for a Poisson problem that involves a Dirac measure as a source term, and then the pointwise a posteriori error estimator of [2] for a Poisson problem involving an unbounded forcing term. The latter is needed because the adjoint problem (2.3) has the function $\bar{y} - y_d$ as a forcing term, which in general does not belong to $L^\infty(\Omega)$. In order to have that $\bar{y} - y_d \in L^\infty(\Omega)$, an additional assumption must be imposed: $y_d \in L^\infty(\Omega)$; see [31, Theorem 2.5].

4.1. A posteriori error estimates for the Laplacian with Dirac sources. Let ξ be an interior point of Ω and consider the following elliptic boundary value problem: Find z such that

$$(4.1) \quad -\Delta z = \delta_\xi \text{ in } \Omega, \quad z = 0 \text{ on } \partial\Omega.$$

Consider the following weak formulation of problem (4.1):

$$(4.2) \quad z \in W_0^{1,r}(\Omega) : \quad (\nabla z, \nabla v)_{L^2(\Omega)} = \delta_\xi(v) \quad \forall v \in W_0^{1,r'}(\Omega),$$

where $1 \leq r < d/(d-1)$ and r' denotes its conjugate exponent. We immediately notice that, since $r' > d$, we have that $W^{1,r'}(\Omega) \hookrightarrow C(\bar{\Omega})$ and, consequently, the term on the right-hand side of (4.2) is well-defined.

We now define the Galerkin approximation to (4.2) as the following problem:

$$(4.3) \quad z_{\mathcal{T}} \in \mathbb{V}(\mathcal{T}) : \quad (\nabla z_{\mathcal{T}}, \nabla v_{\mathcal{T}})_{L^2(\Omega)} = \delta_\xi(v_{\mathcal{T}}) \quad \forall v_{\mathcal{T}} \in \mathbb{V}(\mathcal{T}),$$

where the space $\mathbb{V}(\mathcal{T})$ is defined in (3.1).

In order to present the error estimator developed in [3], we introduce standard notation in a posteriori error analysis [36]. We define \mathcal{S} as the set of internal $(d-1)$ -dimensional interelement boundaries S of \mathcal{T} . For $T \in \mathcal{T}$, let \mathcal{S}_T denote the subset of \mathcal{S} that contains the sides in \mathcal{S} which are sides of T . We also denote by \mathcal{N}_S the subset of \mathcal{T} that contains the two elements that have S as a side. In addition, we define the following patches associated with an element $T \in \mathcal{T}$:

$$(4.4) \quad \mathcal{N}_T := \bigcup_{T' \in \mathcal{T}: T \cap T' \neq \emptyset} T',$$

and

$$(4.5) \quad \mathcal{N}_T^* := \bigcup_{T' \in \mathcal{T}: \mathcal{S}_T \cap \mathcal{S}_{T'} \neq \emptyset} T'.$$

Given a discrete function $z_{\mathcal{T}} \in \mathbb{V}(\mathcal{T})$, we define, for any internal side $S \in \mathcal{S}$, the jump or interelement residual $\llbracket \nabla z_{\mathcal{T}} \cdot \nu \rrbracket$ by

$$(4.6) \quad \llbracket \nabla z_{\mathcal{T}} \cdot \nu \rrbracket = \nu^+ \cdot \nabla z_{\mathcal{T}}|_{T^+} + \nu^- \cdot \nabla z_{\mathcal{T}}|_{T^-},$$

where $\mathcal{N}_S = \{T^+, T^-\}$ and ν^+, ν^- denote the unit normals to S pointing towards $T^+, T^- \in \mathcal{T}$, respectively.

With these ingredients at hand, we define the following a posteriori error indicators. If $\xi \notin \{x_i\}_{i=1}^{N_{\mathcal{T}}}$ but $\xi \in T$, then

$$(4.7) \quad \mathcal{E}_{\delta}^2(z_{\mathcal{T}}; T) := h_T^{4-d} \chi(T) + h_T^3 \|\llbracket \nabla z_{\mathcal{T}} \cdot \nu \rrbracket\|_{L^2(\partial T \setminus \partial \Omega)}^2,$$

otherwise,

$$(4.8) \quad \mathcal{E}_{\delta}^2(z_{\mathcal{T}}; T) := h_T^3 \|\llbracket \nabla z_{\mathcal{T}} \cdot \nu \rrbracket\|_{L^2(\partial T \setminus \partial \Omega)}^2.$$

With these indicators at hand, we thus define the global a posteriori error estimator

$$\mathcal{E}_{\delta}^2(z_{\mathcal{T}}; \mathcal{T}) := \left(\sum_{T \in \mathcal{T}} \mathcal{E}_{\delta}^2(z_{\mathcal{T}}; T) \right)^{\frac{1}{2}}.$$

The following result states the reliability and local efficiency of the global error estimator \mathcal{E}_{δ} . For a proof, we refer the reader to [3, Theorem 4.1].

Theorem 4.1. *Let $z \in W_0^{1,r}(\Omega)$ and $z_{\mathcal{T}} \in \mathbb{V}(\mathcal{T})$ be the solutions to problems (4.1) and (4.3), respectively. We thus have that*

$$(4.9) \quad \|z - z_{\mathcal{T}}\|_{L^2(\Omega)} \lesssim \mathcal{E}_{\delta}(z_{\mathcal{T}}; \mathcal{T}),$$

and

$$(4.10) \quad \mathcal{E}_{\delta}(z_{\mathcal{T}}, T) \lesssim \|z - z_{\mathcal{T}}\|_{L^2(\mathcal{N}_{\mathcal{T}})},$$

where the hidden constants are independent of z , $z_{\mathcal{T}}$, T and the cardinality of \mathcal{T} .

Remark 4.2 (convexity of Ω). Assuming convexity is customary when performing an a posteriori error analysis based on duality. Indeed, the convexity of the domain Ω is imposed so that the associated dual problem exhibits suitable regularity properties which are used to show the reliability of \mathcal{E}_{δ} in the L^2 -norm; see the proof of [3, Theorem 4.1] and [22].

We conclude this section by commenting on the alternative a posteriori error analyses developed in [1] and [25]. In [1], the authors introduce an a posteriori error estimator in weighted Sobolev spaces for an elliptic PDE with a Dirac delta as a forcing term. The authors prove reliability and locally efficiency results in two- and three-dimensional domains. A different approach is considered in [25], where the authors explore the use of fractional Sobolev spaces. The authors design an a posteriori error estimator and derive reliability and locally efficiency properties on polygonal domains.

4.2. Pointwise a posteriori error estimation. Since the variational inequality (2.4), that characterizes the optimal control, involves the duality pairing between the spaces $C_0(\Omega)$ and $\mathcal{M}(\Omega)$, it is thus imperative to consider a pointwise error estimator for the adjoint problem (2.3).

Pointwise a posteriori error estimates have been studied by several authors in the literature. To the best of our knowledge, the earliest two works that study L^∞ residual a posteriori error estimators for a Poisson problem with a bounded forcing term are [23, 28]. The analysis of [28] was subsequently extended to $d = 3$ in [19] and later improved in [20, 21]. The theory has also been extended to obstacle, monotone semilinear, and geometric problems [4, 21, 29, 30]. A standard requirement, in most of these works, is that the right hand side of the underlying PDE belongs to $L^\infty(\Omega)$. However, as it was previously mentioned, the adjoint equation (2.3), for the sparse optimal control problem, has the function $\bar{y} - y_d$ as a forcing term, which, in general is not bounded. For this reason, in what follows we will present an a posteriori error analysis in the maximum norm for a Poisson problem with an unbounded forcing term [2, 4, 21].

Let $f \in L^2(\Omega)$, and consider the following elliptic boundary value problem:

$$(4.11) \quad z \in H_0^1(\Omega) : \quad (\nabla z, \nabla v)_{L^2(\Omega)} = (f, v)_{L^2(\Omega)} \quad \forall v \in H_0^1(\Omega).$$

Notice that, since we are in a convex polytope, we conclude that $z \in H^2(\Omega)$ and that this in turn implies, via Sobolev embedding, that $z \in W^{1,t}(\Omega) \cap C^{0,\kappa}(\bar{\Omega})$ for some $t > d$ and $\kappa > 0$. In view of this, it is legitimate to study the a posteriori error estimation in $L^\infty(\Omega)$ of problem (4.11).

We begin by defining the Galerkin approximation to problem (4.11) as

$$(4.12) \quad z_{\mathcal{T}} \in \mathbb{V}(\mathcal{T}) : \quad (\nabla z_{\mathcal{T}}, \nabla v_{\mathcal{T}})_{L^2(\Omega)} = (f, v_{\mathcal{T}})_{L^2(\Omega)} \quad \forall v_{\mathcal{T}} \in \mathbb{V}(\mathcal{T}).$$

We thus introduce the following a posteriori local error indicators

$$(4.13) \quad \mathcal{E}_\infty(z_{\mathcal{T}}; T) := h_T^{2-d/2} \|f\|_{L^2(T)} + h_T \|\nabla z_{\mathcal{T}} \cdot \nu\|_{L^\infty(\partial T \setminus \partial\Omega)},$$

and the error estimator $\mathcal{E}_\infty(z_{\mathcal{T}}; \mathcal{T}) := \max_{T \in \mathcal{T}} \mathcal{E}_\infty(z_{\mathcal{T}}; T)$.

In order to present the reliability of the global error estimator \mathcal{E}_∞ , we define

$$(4.14) \quad \ell_{\mathcal{T}} := \left| \log \left(\max_{T \in \mathcal{T}} \frac{1}{h_T} \right) \right|.$$

The proof of the next result can be found in [2, Lemma 4.2].

Lemma 4.3 (global reliability). *Let $z \in H_0^1(\Omega) \cap L^\infty(\Omega)$ and $z_{\mathcal{T}} \in \mathbb{V}(\mathcal{T})$ be the solutions of (4.11) and (4.12), respectively. Then*

$$(4.15) \quad \|z - z_{\mathcal{T}}\|_{L^\infty(\Omega)} \lesssim \ell_{\mathcal{T}} \mathcal{E}_\infty(z_{\mathcal{T}}; \mathcal{T}),$$

where the hidden constant is independent of f , z , $z_{\mathcal{T}}$, the size of the elements in the mesh \mathcal{T} and $\#\mathcal{T}$.

To present the local efficiency of the indicators \mathcal{E}_∞ , we define for any $g \in L^2(\Omega)$, and $\mathcal{M} \subset \mathcal{T}$,

$$(4.16) \quad \text{osc}_{\mathcal{T}}(g; \mathcal{M}) := \left(\sum_{T \in \mathcal{M}} h_T^{2(2-d/2)} \|g - \Pi_{\mathcal{T}} g\|_{L^2(T)}^2 \right)^{\frac{1}{2}},$$

where $\Pi_{\mathcal{T}}$ is the L^2 -projection operator onto piecewise linear functions over \mathcal{T} .

The local efficiency of the indicators (4.13) is as follows. For a proof we refer the reader to [2, Lemma 4.3]

Lemma 4.4 (local efficiency). *Let $z \in H_0^1(\Omega) \cap L^\infty(\Omega)$ and $z_{\mathcal{T}} \in \mathbb{V}(\mathcal{T})$ be the solutions to problems (4.11) and (4.12), respectively. Then*

$$(4.17) \quad \mathcal{E}_\infty(z_{\mathcal{T}}; T) \lesssim \|z - z_{\mathcal{T}}\|_{L^\infty(\mathcal{N}_T^*)} + \text{osc}_{\mathcal{T}}(f; \mathcal{N}_T^*)$$

for all $T \in \mathcal{T}$, where \mathcal{N}_T^* is given by (4.5), and the hidden constant is independent of $f, z, z_{\mathcal{T}}$, the size of the elements in the mesh \mathcal{T} and $\#\mathcal{T}$.

5. A POSTERIORI ERROR ANALYSIS FOR THE OPTIMAL CONTROL PROBLEM WITH SPARSITY PROMOTING COST

On the basis of the error indicators and estimators presented in Sections 4.1 and 4.2, we proceed with the design of an a posteriori error estimator for the sparse optimal control problem (1.2)–(1.3). The error estimator can be decomposed as the sum of two contributions:

$$(5.1) \quad \mathcal{E}_{\text{ocp}}^2(\bar{y}_{\mathcal{T}}, \bar{p}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; \mathcal{T}) := \mathcal{E}_y^2(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; \mathcal{T}) + \mathcal{E}_p^2(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; \mathcal{T}),$$

where $\mathcal{T} \in \mathbb{T}$ and $\bar{y}_{\mathcal{T}}, \bar{p}_{\mathcal{T}}$ and $\bar{u}_{\mathcal{T}}$ denote the discrete optimal variables that solve the discrete optimality system (3.9).

Let us now describe each contribution to (5.1) separately. First, on the basis of the results presented in Section 4.1, we define, for $T \in \mathcal{T}$, the local error indicators

$$(5.2) \quad \mathcal{E}_y^2(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; T) := h_T^3 \|[\nabla \bar{y}_{\mathcal{T}} \cdot \nu]\|_{L^2(\partial T \setminus \partial \Omega)}^2.$$

The global error estimator $\mathcal{E}_y(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; \mathcal{T})$ is thus defined by

$$(5.3) \quad \mathcal{E}_y(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; \mathcal{T}) := \left(\sum_{T \in \mathcal{T}} \mathcal{E}_y^2(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; T) \right)^{\frac{1}{2}}.$$

We immediately notice that, since the optimal control $\bar{u}_{\mathcal{T}}$ is sought in the discrete space $\mathbb{U}(\mathcal{T})$, it can be thus written as a linear combination of Dirac measures supported on $\{x_i\}_{i=1}^{N_{\mathcal{T}}}$. As a consequence, the definition of the local indicators $\mathcal{E}_y(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; T)$ does not involve the additional term $h_T^{4-d} \chi(T)$ that appears in (4.7); see [3, Remark 4.1].

The second error contribution in (5.1) is based on the maximum-norm error estimator that we presented in Section 4.2. Locally, it is defined by

$$(5.4) \quad \mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; T) := h_T^{2-d/2} \|\bar{y}_{\mathcal{T}} - y_d\|_{L^2(T)} + h_T \|[\nabla \bar{p}_{\mathcal{T}} \cdot \nu]\|_{L^\infty(\partial T \setminus \partial \Omega)}.$$

The global error estimator $\mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; \mathcal{T})$ is then defined by

$$(5.5) \quad \mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; \mathcal{T}) := \max_{T \in \mathcal{T}} \mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; T).$$

Since, it will be useful in the analysis that we will perform, we introduce the following auxiliary variables. First, let $(\hat{y}, \hat{p}) \in W_0^{1,r}(\Omega) \times H_0^1(\Omega)$ be such that

$$(5.6) \quad \begin{aligned} (\nabla \hat{y}, \nabla v)_{L^2(\Omega)} &= \langle \bar{u}_{\mathcal{T}}, v \rangle_{\Omega} & \forall v \in W_0^{1,r'}(\Omega), \\ (\nabla w, \nabla \hat{p})_{L^2(\Omega)} &= (\bar{y}_{\mathcal{T}} - y_d, w)_{L^2(\Omega)} & \forall w \in H_0^1(\Omega), \end{aligned}$$

where $1 \leq r < d/(d-1)$ and r' denotes its conjugate exponent. We also define $\tilde{p} \in H_0^1(\Omega)$ to be the solution of

$$(5.7) \quad (\nabla w, \nabla \tilde{p})_{L^2(\Omega)} = (\hat{y} - y_d, w)_{L^2(\Omega)} \quad \forall w \in H_0^1(\Omega).$$

We notice that $(\bar{y}_{\mathcal{T}}, \bar{p}_{\mathcal{T}})$ can be understood as a finite element approximation of (\hat{y}, \hat{p}) . Consequently, on the basis of the results presented in Sections 4.1 and 4.2, the a posteriori error estimators defined in (5.3) and (5.5) satisfy the following reliability properties:

$$(5.8) \quad \|\hat{y} - \bar{y}_{\mathcal{T}}\|_{L^2(\Omega)} \lesssim \mathcal{E}_y(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; \mathcal{T}), \quad \|\hat{p} - \bar{p}_{\mathcal{T}}\|_{L^\infty(\Omega)} \lesssim \ell_{\mathcal{T}} \mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; \mathcal{T}).$$

We thus have all the ingredients at hand to develop our a posteriori error analysis for the sparse optimal control problem (1.2)–(1.3).

5.1. A posteriori error estimator: reliability.

Theorem 5.1 (global reliability). *Let $(\bar{y}, \bar{p}, \bar{u}) \in W_0^{1,r}(\Omega) \times H_0^1(\Omega) \times \mathcal{M}(\Omega)$, with $2d/(d+2) \leq r < d/(d-1)$, be the solution to (2.6), and let $(\bar{y}_{\mathcal{T}}, \bar{p}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}) \in \mathbb{V}(\mathcal{T}) \times \mathbb{V}(\mathcal{T}) \times \mathbb{U}(\mathcal{T})$ be its numerical approximation obtained as the solution to the discrete optimality system (3.9). Then*

$$(5.9) \quad \begin{aligned} \|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(\Omega)}^2 + \|\bar{p} - \bar{p}_{\mathcal{T}}\|_{L^\infty(\Omega)}^2 + \|\bar{u} - \bar{u}_{\mathcal{T}}\|_{H^{-2}(\Omega)}^2 \\ \lesssim \mathcal{E}_y^2(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; \mathcal{T}) + \ell_{\mathcal{T}}^2 \mathcal{E}_p^2(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; \mathcal{T}) + \ell_{\mathcal{T}} \mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; \mathcal{T}), \end{aligned}$$

where $\ell_{\mathcal{T}}$ is defined in (4.14), and the hidden constant is independent of the continuous and discrete optimal variables, the size of the elements of the mesh \mathcal{T} and its cardinality $\#\mathcal{T}$.

Proof. We proceed in four steps.

Step 1. The objective of this step is to bound the error $\|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(\Omega)}$. To accomplish this task, we invoke the auxiliary state variable \hat{y} , defined as the solution to (5.6), and apply the triangle inequality to arrive at the estimate

$$(5.10) \quad \|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(\Omega)}^2 \lesssim \|\bar{y} - \hat{y}\|_{L^2(\Omega)}^2 + \mathcal{E}_y^2(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; \mathcal{T}),$$

where we have also used (5.8). We now focus on controlling the term $\|\bar{y} - \hat{y}\|_{L^2(\Omega)}^2$. Set $u = \bar{u}_{\mathcal{T}}$ in (2.4) and $u = \bar{u}$ in (3.8), and obtain that

$$\begin{aligned} -\langle \bar{u}_{\mathcal{T}} - \bar{u}, \bar{p} \rangle_{\Omega} + \alpha \|\bar{u}\|_{\mathcal{M}(\Omega)} &\leq \alpha \|\bar{u}_{\mathcal{T}}\|_{\mathcal{M}(\Omega)}, \\ -\langle \bar{u} - \bar{u}_{\mathcal{T}}, \bar{p}_{\mathcal{T}} \rangle_{\Omega} + \alpha \|\bar{u}_{\mathcal{T}}\|_{\mathcal{M}(\Omega)} &\leq \alpha \|\bar{u}\|_{\mathcal{M}(\Omega)}. \end{aligned}$$

Adding the previous inequalities, we thus obtain that

$$(5.11) \quad \langle \bar{u} - \bar{u}_{\mathcal{T}}, \bar{p} - \tilde{p} \rangle_{\Omega} + \langle \bar{u} - \bar{u}_{\mathcal{T}}, \tilde{p} - \bar{p}_{\mathcal{T}} \rangle_{\Omega} \leq 0,$$

where we have invoked the auxiliary adjoint state \tilde{p} , defined in (5.7).

Let us concentrate now on the term $\langle \bar{u} - \bar{u}_{\mathcal{T}}, \bar{p} - \tilde{p} \rangle_{\Omega}$. First, we notice that the functions $\bar{y} - \hat{y}$ and $\bar{p} - \tilde{p}$ satisfy

$$(5.12) \quad \bar{y} - \hat{y} \in W_0^{1,r}(\Omega) : \quad (\nabla(\bar{y} - \hat{y}), \nabla v)_{L^2(\Omega)} = \langle \bar{u} - \bar{u}_{\mathcal{T}}, v \rangle_{\Omega} \quad \forall v \in W_0^{1,r'}(\Omega),$$

$$(5.13) \quad \bar{p} - \tilde{p} \in H_0^1(\Omega) : \quad (\nabla w, \nabla(\bar{p} - \tilde{p}))_{L^2(\Omega)} = (\bar{y} - \hat{y}, w)_{L^2(\Omega)} \quad \forall w \in H_0^1(\Omega),$$

where $2d/(d+2) \leq r < d/(d-1)$ and $r' > d$ is its conjugate exponent. Notice now that, since $\bar{y} - \hat{y} \in L^2(\Omega)$, there exists $r' > d$ for which $\bar{p} - \tilde{p} \in W^{1,r'}(\Omega)$. Consequently, we are able to set $v = \bar{p} - \tilde{p}$ in (5.12). This yields

$$(\nabla(\bar{y} - \hat{y}), \nabla(\bar{p} - \tilde{p}))_{L^2(\Omega)} = \langle \bar{u} - \bar{u}_{\mathcal{T}}, \bar{p} - \tilde{p} \rangle_{\Omega}.$$

Following a similar reasoning, we would like to set $w = \bar{y} - \hat{y}$ in (5.13). However, $\bar{y} - \hat{y} \notin H_0^1(\Omega)$. Exploiting the fact that $\bar{y} - \hat{y} \in W_0^{1,r}(\Omega)$ with $2d/(d+2) \leq r < d/(d-1)$ and that there exists $r' > d$ such that $\bar{p} - \tilde{p} \in W^{1,r'}(\Omega)$, a density argument allows us to conclude that

$$(\nabla(\bar{y} - \hat{y}), \nabla(\bar{p} - \tilde{p}))_{L^2(\Omega)} = (\bar{y} - \hat{y}, \bar{y} - \hat{y})_{L^2(\Omega)},$$

and thus that

$$(5.14) \quad \langle \bar{u} - \bar{u}_{\mathcal{T}}, \bar{p} - \tilde{p} \rangle_{\Omega} = \|\bar{y} - \hat{y}\|_{L^2(\Omega)}^2.$$

Replacing the previous term in (5.11), and inserting the auxiliary variable \hat{p} defined as the solution to (5.6), we obtain that

$$(5.15) \quad \|\bar{y} - \hat{y}\|_{L^2(\Omega)}^2 \leq \langle \bar{u}_{\mathcal{T}} - \bar{u}, \tilde{p} - \hat{p} \rangle_{\Omega} + \langle \bar{u}_{\mathcal{T}} - \bar{u}, \hat{p} - \bar{p}_{\mathcal{T}} \rangle_{\Omega}.$$

We now apply Theorem 3.1 to conclude that the norms $\|\bar{u}\|_{\mathcal{M}(\Omega)}$ and $\|\bar{u}_{\mathcal{T}}\|_{\mathcal{M}(\Omega)}$ are bounded. This, combined with the estimate (5.8), reveals that

$$(5.16) \quad \langle \bar{u}_{\mathcal{T}} - \bar{u}, \hat{p} - \bar{p}_{\mathcal{T}} \rangle_{\Omega} \lesssim \ell_{\mathcal{T}} \mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; \mathcal{T}).$$

To estimate the first term on the right hand side of (5.15), we notice that $\hat{y} - \bar{y}$ and $\tilde{p} - \hat{p}$ satisfy

$$(5.17) \quad \hat{y} - \bar{y} \in W_0^{1,r}(\Omega) : \quad (\nabla(\hat{y} - \bar{y}), \nabla v)_{L^2(\Omega)} = \langle \bar{u}_{\mathcal{T}} - \bar{u}, v \rangle_{\Omega} \quad \forall v \in W_0^{1,r'}(\Omega),$$

$$(5.18) \quad \tilde{p} - \hat{p} \in H_0^1(\Omega) : \quad (\nabla w, \nabla(\tilde{p} - \hat{p}))_{L^2(\Omega)} = (\hat{y} - \bar{y}_{\mathcal{T}}, w)_{L^2(\Omega)} \quad \forall w \in H_0^1(\Omega).$$

Similar density arguments to those used to obtain (5.14), allow us to set $v = \tilde{p} - \hat{p}$ and $w = \hat{y} - \bar{y}$. This yields

$$(5.19) \quad \langle \bar{u}_{\mathcal{T}} - \bar{u}, \tilde{p} - \hat{p} \rangle_{\Omega} = (\hat{y} - \bar{y}_{\mathcal{T}}, \hat{y} - \bar{y}).$$

Replacing (5.16) and the obtained result into (5.15) we thus obtain that

$$(5.20) \quad \|\bar{y} - \hat{y}\|_{L^2(\Omega)}^2 \lesssim \|\hat{y} - \bar{y}_{\mathcal{T}}\|_{L^2(\Omega)}^2 + (\hat{y} - \bar{y}_{\mathcal{T}}, \bar{y}_{\mathcal{T}} - \bar{y})_{L^2(\Omega)} + \ell_{\mathcal{T}} \mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; \mathcal{T}).$$

Now, on the basis of (5.8), an application of Young's inequality allows us to arrive at the estimate

$$(5.21) \quad \|\bar{y} - \hat{y}\|_{L^2(\Omega)}^2 \lesssim \mathcal{E}_y^2(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; \mathcal{T}) + \ell_{\mathcal{T}} \mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; \mathcal{T}) + \frac{1}{4} \|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(\Omega)}^2.$$

Finally, replacing (5.21) into (5.10), we obtain the estimate

$$(5.22) \quad \|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(\Omega)}^2 \lesssim \mathcal{E}_y^2(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; \mathcal{T}) + \ell_{\mathcal{T}} \mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; \mathcal{T}).$$

Step 2. The goal of this step is to estimate the term $\|\bar{u} - \bar{u}_{\mathcal{T}}\|_{H^{-2}(\Omega)}$. To accomplish this task, we follow the arguments elaborated in [31, Corollary 4.5] that guarantee

$$(5.23) \quad \|\bar{u} - \bar{u}_{\mathcal{T}}\|_{H^{-2}(\Omega)}^2 \lesssim \|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(\Omega)}^2 + \|\bar{y}_{\mathcal{T}} - \hat{y}\|_{L^2(\Omega)}^2,$$

where \hat{y} is defined as the solution to (5.6). We now invoke, once again, the estimate (5.8) together with (5.22) to conclude that

$$(5.24) \quad \|\bar{u} - \bar{u}_{\mathcal{T}}\|_{H^{-2}(\Omega)}^2 \lesssim \mathcal{E}_y^2(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; \mathcal{T}) + \ell_{\mathcal{T}} \mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; \mathcal{T}).$$

Step 3. In this step we bound the error $\|\bar{p} - \bar{p}_{\mathcal{T}}\|_{L^{\infty}(\Omega)}$. To accomplish this task, we use the triangle inequality and then the estimate (5.8) to obtain that

$$(5.25) \quad \|\bar{p} - \bar{p}_{\mathcal{T}}\|_{L^{\infty}(\Omega)}^2 \lesssim \|\bar{p} - \hat{p}\|_{L^{\infty}(\Omega)}^2 + \ell_{\mathcal{T}}^2 \mathcal{E}_p^2(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; \mathcal{T}).$$

To estimate the term $\|\bar{p} - \hat{p}\|_{L^{\infty}(\Omega)}^2$, we observe that $\bar{p} - \hat{p} \in H^2(\Omega) \hookrightarrow C(\bar{\Omega})$ so that this, together with (5.22), gives

$$(5.26) \quad \|\bar{p} - \hat{p}\|_{L^{\infty}(\Omega)}^2 \lesssim \|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(\Omega)}^2 \lesssim \mathcal{E}_y^2(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; \mathcal{T}) + \ell_{\mathcal{T}} \mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; \mathcal{T}).$$

Finally, inserting (5.26) into (5.25), we arrive at the estimate

$$(5.27) \quad \|\bar{p} - \bar{p}_{\mathcal{T}}\|_{L^{\infty}(\Omega)}^2 \lesssim \mathcal{E}_y^2(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; \mathcal{T}) + \ell_{\mathcal{T}}^2 \mathcal{E}_p^2(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; \mathcal{T}) + \ell_{\mathcal{T}} \mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; \mathcal{T}).$$

Step 4. The desired estimate (5.9) follows upon gathering the estimates (5.22), (5.24) and (5.27). \square

5.2. A posteriori error estimator: efficiency. In this section we analyze the efficiency properties of the local a posteriori error indicator

$$(5.28) \quad \mathcal{E}_{\text{ocp}}^2(\bar{y}_{\mathcal{T}}, \bar{p}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; T) = \mathcal{E}_y^2(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; T) + \mathcal{E}_p^2(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; T).$$

To accomplish this task, we study each of its contributions separately. We start with the indicator $\mathcal{E}_y(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; T)$ defined by (5.2). Before embarking ourselves with the efficiency analysis of $\mathcal{E}_y(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; T)$, we introduce the following notation: for an edge, triangle or tetrahedron G , let $\mathcal{V}(G)$ be the set of vertices of G .

Let $T \in \mathcal{T}$ and $S \in \mathcal{S}_T$. Recall that \mathcal{N}_S denotes the patch composed by the two elements T and T' sharing S . We introduce the following edge bubble function

$$(5.29) \quad \psi_S|_{\mathcal{N}_S} = d^{4d} \left(\prod_{v \in \mathcal{V}(S)} \phi_v^T \phi_v^{T'} \right)^2,$$

where, for $v \in \mathcal{V}(S)$, ϕ_v^T and $\phi_v^{T'}$ denote the barycentric coordinates of T and T' , respectively, which are understood as functions over \mathcal{N}_S . The following properties of the bubble function ψ_S follow immediately: $\psi_S \in \mathbb{P}_{4d}(\mathcal{N}_S)$, $\psi_S \in C^2(\mathcal{N}_S)$, and $\psi_S = 0$ on $\partial \mathcal{N}_S$. In addition, we have that

$$(5.30) \quad \nabla \psi_S = 0 \text{ on } \partial \mathcal{N}_S, \quad [\![\nabla \psi_S \cdot \nu]\!] = 0 \text{ on } S.$$

With all these ingredients at hand, we are ready to prove the local efficiency of $\mathcal{E}_y^2(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; T)$.

Lemma 5.2 (local efficiency of \mathcal{E}_y). *Let $(\bar{y}, \bar{p}, \bar{u}) \in W_0^{1,r}(\Omega) \times H_0^1(\Omega) \times \mathcal{M}(\Omega)$, with $2d/(d+2) \leq r < d/(d-1)$, be the solution to (2.6), and let $(\bar{y}_{\mathcal{T}}, \bar{p}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}) \in \mathbb{V}(\mathcal{T}) \times \mathbb{V}(\mathcal{T}) \times \mathbb{U}(\mathcal{T})$ be its numerical approximation obtained as the solution to the discrete optimality system (3.9). Then, for $T \in \mathcal{T}$, the local error indicator \mathcal{E}_y , defined as in (5.2), satisfies that*

$$(5.31) \quad \mathcal{E}_y^2(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; T) \lesssim \|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(\mathcal{N}_T^*)}^2 + \sum_{S \in \mathcal{S}_T} \|\bar{u} - \bar{u}_{\mathcal{T}}\|_{H^{-2}(\mathcal{N}_S)}^2,$$

where \mathcal{N}_T^* is defined as in (4.5) and the hidden constant is independent of the continuous and discrete optimal variables, the size of the elements in the mesh \mathcal{T} , and $\#\mathcal{T}$.

Proof. Let $v \in W_0^{1,r'}(\Omega)$, with $r' > d$, be such that $v|_T \in C^2(T)$ for all $T \in \mathcal{T}$. Consider v as a test function in the state equation (2.2) and apply integration by parts to arrive at

$$(5.32) \quad \int_{\Omega} \nabla(\bar{y} - \bar{y}_{\mathcal{T}}) \nabla v = \sum_{T \in \mathcal{T}} \langle \bar{u}, v \rangle_T + \sum_{S \in \mathcal{S}} \int_S [\nabla \bar{y}_{\mathcal{T}} \cdot \nu] v.$$

Since we have that $v \in C^2(T)$ on each $T \in \mathcal{T}$, we can integrate by parts, again, to conclude that

$$(5.33) \quad \int_{\Omega} \nabla(\bar{y} - \bar{y}_{\mathcal{T}}) \nabla v = - \sum_{S \in \mathcal{S}} \int_S [\nabla v \cdot \nu] (\bar{y} - \bar{y}_{\mathcal{T}}) - \sum_{T \in \mathcal{T}} \int_T (\bar{y} - \bar{y}_{\mathcal{T}}) \Delta v.$$

As a conclusion, from (5.32) and (5.33), we arrive at the following identity:

$$(5.34) \quad \sum_{T \in \mathcal{T}} (\langle \bar{u} - \bar{u}_{\mathcal{T}}, v \rangle_T + \langle \bar{u}_{\mathcal{T}}, v \rangle_T) + \sum_{S \in \mathcal{S}} \int_S [\nabla \bar{y}_{\mathcal{T}} \cdot \nu] v \\ = - \sum_{S \in \mathcal{S}} \int_S [\nabla v \cdot \nu] (\bar{y} - \bar{y}_{\mathcal{T}}) - \sum_{T \in \mathcal{T}} \int_T (\bar{y} - \bar{y}_{\mathcal{T}}) \Delta v,$$

which holds for every $v \in W_0^{1,r'}(\Omega)$, with $r' > d$, and is such that $v|_T \in C^2(T)$ for all $T \in \mathcal{T}$.

Let $S \in \mathcal{S}$, and set $v = \beta_S = [\nabla \bar{y}_{\mathcal{T}} \cdot \nu] \psi_S$ in (5.34). As it is customary in a posteriori error analysis, we extend the jump term defined on S to the patch \mathcal{N}_S following the arguments of [35, Section 3]. In what follows, we will make no distinction between the jump term defined on S and its extension onto \mathcal{N}_S . Now, notice that we have that $\beta_S \in H_0^2(\mathcal{N}_S)$ and, moreover, $[\nabla \beta_S \cdot \nu] = 0$ on S . In addition, since $\bar{u}_{\mathcal{T}} \in \mathbb{U}(\mathcal{T})$, we can conclude for every $T \in \mathcal{N}_S$ that

$$\langle \bar{u}_{\mathcal{T}}, \beta_S \rangle_T = \sum_{v \in \mathcal{V}(T)} u_v \langle \delta_v, \beta_S \rangle_T = 0.$$

Consequently, a simple application of the Cauchy–Schwarz inequality reveals that

$$(5.35) \quad \int_S [\nabla \bar{y}_{\mathcal{T}} \cdot \nu]^2 \psi_S = - \sum_{T' \in \mathcal{N}_S} \left(\int_{T'} (\bar{y} - \bar{y}_{\mathcal{T}}) \Delta \beta_S + \langle \bar{u} - \bar{u}_{\mathcal{T}}, \beta_S \rangle_{T'} \right) \\ \leq \sum_{T' \in \mathcal{N}_S} \left(\|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(T')} \|\Delta \beta_S\|_{L^2(T')} \right) + \|\bar{u} - \bar{u}_{\mathcal{T}}\|_{H^{-2}(\mathcal{N}_S)} \|\beta_S\|_{H^2(\mathcal{N}_S)}.$$

Now, since $[\nabla \bar{y}_{\mathcal{T}} \cdot \nu] \in \mathbb{R}$, we have that $\Delta \beta_S = [\nabla \bar{y}_{\mathcal{T}} \cdot \nu] \Delta \psi_S$. Standard arguments allow us to obtain the following bound

$$\|\Delta \beta_S\|_{L^2(T)}^2 \lesssim [\nabla \bar{y}_{\mathcal{T}} \cdot \nu]^2 \|\Delta \psi_S\|_{L^2(T)}^2 \\ \lesssim |T| |S|^{-1} h_T^{-4} \|[\nabla \bar{y}_{\mathcal{T}} \cdot \nu]\|_{L^2(S)}^2 \lesssim h_T^{-3} \|[\nabla \bar{y}_{\mathcal{T}} \cdot \nu]\|_{L^2(S)}^2.$$

With these estimates at hand, we thus use standard bubble functions arguments, the Poincaré inequality, and the shape regularity property of the family $\{\mathcal{T}\}$ to arrive at

$$(5.36) \quad h_T^3 \|\llbracket \nabla \bar{y}_{\mathcal{T}} \cdot \nu \rrbracket\|_{L^2(S)}^2 \lesssim \sum_{T' \in \mathcal{N}_S} \left\{ \|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(T')}^2 \right\} + \|\bar{u} - \bar{u}_{\mathcal{T}}\|_{H^{-2}(\mathcal{N}_S)}^2,$$

which concludes our proof. \square

We now continue with the study of the local efficiency properties of the indicators $\mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; T)$ defined by (5.4).

Lemma 5.3 (local efficiency of \mathcal{E}_p). *Let $(\bar{y}, \bar{p}, \bar{u}) \in W_0^{1,r}(\Omega) \times H_0^1(\Omega) \times \mathcal{M}(\Omega)$, where $2d/(d+2) \leq r < d/(d-1)$, be the solution to the optimality system (2.6), and $(\bar{y}_{\mathcal{T}}, \bar{p}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}) \in \mathbb{V}(\mathcal{T}) \times \mathbb{V}(\mathcal{T}) \times \mathbb{U}(\mathcal{T})$ be its numerical approximation obtained as the solution to the discrete optimality system (3.9). Then, for $T \in \mathcal{T}$, the local error indicator $\mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; T)$, defined in (5.4), satisfies that*

$$(5.37) \quad \mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; T) \lesssim \|\bar{p} - \bar{p}_{\mathcal{T}}\|_{L^\infty(\mathcal{N}_T^*)} + h_T^{2-d/2} \|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(\mathcal{N}_T^*)} + \text{osc}_{\mathcal{T}}(y_d; \mathcal{N}_T^*),$$

where \mathcal{N}_T^* is defined in (4.5) and the hidden constant is independent of the optimal variables, their approximations, the size of the elements in the mesh \mathcal{T} , and its cardinality $\#\mathcal{T}$.

Proof. We follow closely the arguments elaborated in [2, Lemma 5.4]. Let $w \in H_0^1(\Omega)$ be such that $w|_T \in C^2(\Omega)$ for $T \in \mathcal{T}$. Consider w as a test function in the adjoint equation (2.3). An argument based on integration by parts reveals that

$$(5.38) \quad \int_{\Omega} \nabla w \nabla (\bar{p} - \bar{p}_{\mathcal{T}}) = \sum_{T \in \mathcal{T}} \int_T (\bar{y} - y_d) w + \sum_{S \in \mathcal{S}} \int_S \llbracket \nabla \bar{p}_{\mathcal{T}} \cdot \nu \rrbracket w.$$

In view of the regularity properties of the function w , we can, again, integrate by parts to obtain that

$$(5.39) \quad \int_{\Omega} \nabla w \nabla (\bar{p} - \bar{p}_{\mathcal{T}}) = - \sum_{S \in \mathcal{S}} \int_S \llbracket \nabla w \cdot \nu \rrbracket (\bar{p} - \bar{p}_{\mathcal{T}}) - \sum_{T \in \mathcal{T}} \int_T (\bar{p} - \bar{p}_{\mathcal{T}}) \Delta w.$$

As a conclusion, we have thus obtained the following identity

$$(5.40) \quad \begin{aligned} & \sum_{T \in \mathcal{T}} \int_T (\bar{y} - y_d) w + \sum_{S \in \mathcal{S}} \int_S \llbracket \nabla \bar{p}_{\mathcal{T}} \cdot \nu \rrbracket w \\ &= - \sum_{S \in \mathcal{S}} \int_S \llbracket \nabla w \cdot \nu \rrbracket (\bar{p} - \bar{p}_{\mathcal{T}}) - \sum_{T \in \mathcal{T}} \int_T (\bar{p} - \bar{p}_{\mathcal{T}}) \Delta w, \end{aligned}$$

that holds for every $w \in H_0^1(\Omega)$ such that $w|_T \in C^2(T)$ for all $T \in \mathcal{T}$. With this error equation at hand, we proceed to derive the local efficiency properties of the error indicator $\mathcal{E}_p(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; T)$: the analysis involves two steps.

Step 1. In this step we bound the term $h_T^{2-d/2} \|\bar{y}_{\mathcal{T}} - y_d\|_{L^2(T)}$ in (5.4) for $T \in \mathcal{T}$. We begin with a simple application of the triangle inequality:

$$h_T^{2-d/2} \|\bar{y}_{\mathcal{T}} - y_d\|_{L^2(T)} \leq h_T^{2-d/2} \|\bar{y}_{\mathcal{T}} - \Pi_{\mathcal{T}} y_d\|_{L^2(T)} + h_T^{2-d/2} \|\Pi_{\mathcal{T}} y_d - y_d\|_{L^2(T)},$$

where $\Pi_{\mathcal{T}}$ denotes the L^2 -projection operator onto piecewise linear functions over \mathcal{T} . Now we set $w = \zeta_T := (\bar{y}_{\mathcal{T}} - \Pi_{\mathcal{T}} y_d) \varphi_T^2$ in (5.40), where φ_T is the standard

bubble function over T [34, 36]. This yields

$$(5.41) \quad \begin{aligned} \int_T (\bar{y}_\mathcal{T} - \Pi_\mathcal{T} y_d) \zeta_T &= \int_T [(\bar{y} - y_d) - (\bar{y} - \bar{y}_\mathcal{T}) + (y_d - \Pi_\mathcal{T} y_d)] \zeta_T \\ &= - \int_T (\bar{p} - \bar{p}_\mathcal{T}) \Delta \zeta_T - \int_T (\bar{y} - \bar{y}_\mathcal{T}) \zeta_T + \int_T (y_d - \Pi_\mathcal{T} y_d) \zeta_T =: \text{I} + \text{II} + \text{III}. \end{aligned}$$

To compute $\Delta \zeta_T$ on T , we first notice that $\Delta(\bar{y}_\mathcal{T} - \Pi_\mathcal{T} y_d)|_T = 0$. A basic computation thus reveals that

$$\Delta \zeta_T = 4\nabla(\bar{y}_\mathcal{T} - \Pi_\mathcal{T} y_d) \cdot \nabla \varphi_T \varphi_T + 2(\bar{y}_\mathcal{T} - \Pi_\mathcal{T} y_d)(\varphi_T \Delta \varphi_T + \nabla \varphi_T \cdot \nabla \varphi_T).$$

The control of the term I follows from the the previous identity, the properties of the bubble function φ_T and an inverse inequality. In fact, we have that

$$(5.42) \quad \begin{aligned} |\text{I}| &\lesssim \left(h_T^{\frac{d}{2}-1} \|\nabla(\bar{y}_\mathcal{T} - \Pi_\mathcal{T} y_d)\|_{L^2(T)} + h_T^{\frac{d}{2}-2} \|\bar{y}_\mathcal{T} - \Pi_\mathcal{T} y_d\|_{L^2(T)} \right) \\ &\quad \cdot \|\bar{p} - \bar{p}_\mathcal{T}\|_{L^\infty(T)} \lesssim h_T^{d/2-2} \|\bar{y}_\mathcal{T} - \Pi_\mathcal{T} y_d\|_{L^2(T)} \|\bar{p} - \bar{p}_\mathcal{T}\|_{L^\infty(T)}. \end{aligned}$$

The terms II and III in (5.41) are bounded in view of standard properties of the bubble function φ_T : we have that

$$(5.43) \quad |\text{II}| \lesssim \|\bar{y} - \bar{y}_\mathcal{T}\|_{L^2(T)} \|\bar{y}_\mathcal{T} - \Pi_\mathcal{T} y_d\|_{L^2(T)},$$

and that

$$(5.44) \quad |\text{III}| \lesssim \|y_d - \Pi_\mathcal{T} y_d\|_{L^2(T)} \|\bar{y}_\mathcal{T} - \Pi_\mathcal{T} y_d\|_{L^2(T)}.$$

Finally, since $\|\bar{y}_\mathcal{T} - \Pi_\mathcal{T} y_d\|_{L^2(T)}^2 \lesssim \int_T (\bar{y}_\mathcal{T} - \Pi_\mathcal{T} y_d) \zeta_T$, the identity (5.41), combined with the estimates (5.42), (5.43) and (5.44), allow us to conclude that

$$(5.45) \quad \begin{aligned} h_T^{2-\frac{d}{2}} \|\bar{y}_\mathcal{T} - y_d\|_{L^2(T)} &\lesssim \|\bar{p} - \bar{p}_\mathcal{T}\|_{L^\infty(T)} \\ &\quad + h_T^{2-\frac{d}{2}} \|\bar{y} - \bar{y}_\mathcal{T}\|_{L^2(T)} + \text{osc}_\mathcal{T}(y_d; T), \end{aligned}$$

where $\text{osc}_\mathcal{T}(y_d; T)$ is defined by (4.16).

Step 2. Let $T \in \mathcal{T}$ and $S \in \mathcal{S}_T$. The objective of this step is control the term $h_T \|\llbracket \nabla \bar{p}_\mathcal{T} \cdot \nu \rrbracket\|_{L^\infty(\partial T \setminus \partial \Omega)}$ in (5.4). To accomplish this task, we invoke the standard bubble function over S which we denote φ_S . Recall that, according to [34, 36] this function satisfies

$$|\nabla^l \varphi_S| \approx h_S^{-l}, \quad l = 0, 1, 2, \quad |S| \|\llbracket \nabla \bar{p}_\mathcal{T} \cdot \nu \rrbracket\|_{L^\infty(S)} \lesssim \left| \int_S \llbracket \nabla \bar{p}_\mathcal{T} \cdot \nu \rrbracket \varphi_S \right|.$$

To bound the term on the right-hand side of the previous inequality, we set $w = \varphi_S$ in (5.40) and obtain that

$$\begin{aligned} \left| \int_S [\nabla \bar{p}_{\mathcal{T}} \cdot \nu] \varphi_S \right| &\leq \sum_{T' \in \mathcal{N}_S} \int_{T'} |\bar{y} - y_d| \varphi_S + \sum_{T' \in \mathcal{N}_S} \int_{T'} |\bar{p} - \bar{p}_{\mathcal{T}}| |\Delta \varphi_S| \\ &\quad + \sum_{T' \in \mathcal{N}_S} \sum_{S' \in \mathcal{S}_{T'}} \int_{S'} |\bar{p} - \bar{p}_{\mathcal{T}}| |[\nabla \varphi_S \cdot \nu]| \\ &\lesssim \sum_{T' \in \mathcal{N}_S} |T'|^{\frac{1}{2}} (\|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(T')} + \|\bar{y}_{\mathcal{T}} - y_d\|_{L^2(T')}) \\ &\quad + \sum_{T' \in \mathcal{N}_S} \left(h_S^{-2} |T'| + h_S^{-1} \sum_{S' \in \mathcal{S}_{T'}} |S'| \right) \|\bar{p} - \bar{p}_{\mathcal{T}}\|_{L^\infty(T')}. \end{aligned}$$

Upon multiplying this inequality by $h_T |S|^{-1}$, using the shape regularity of the mesh \mathcal{T} , in addition with (5.45), yields

$$\begin{aligned} (5.46) \quad h_T \|[\nabla \bar{p}_{\mathcal{T}} \cdot \nu]\|_{L^\infty(S)} &\lesssim \|\bar{p} - \bar{p}_{\mathcal{T}}\|_{L^\infty(\mathcal{N}_S)} \\ &\quad + h_T^{2-d/2} \|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(\mathcal{N}_S)} + \text{osc}_{\mathcal{T}}(y_d; \mathcal{N}_S). \end{aligned}$$

Combining (5.45) and (5.46), we arrive at (5.37). \square

The results of Lemmas 5.2 and 5.3 immediately yield the following result.

Theorem 5.4 (local efficiency of \mathcal{E}_{ocp}). *Let $(\bar{y}, \bar{p}, \bar{u}) \in W_0^{1,r}(\Omega) \times H_0^1(\Omega) \times \mathcal{M}(\Omega)$, with $2d/(d+2) \leq r < d/(d-1)$, be the solution to the optimality system (2.6), and $(\bar{y}_{\mathcal{T}}, \bar{p}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}) \in \mathbb{V}(\mathcal{T}) \times \mathbb{V}(\mathcal{T}) \times \mathbb{U}(\mathcal{T})$ be its numerical approximation obtained as the solution to the discrete optimality system (3.9). Then, for $T \in \mathcal{T}$, it holds*

$$\begin{aligned} \mathcal{E}_y^2(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; T) + \mathcal{E}_p^2(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; T) &\lesssim \|\bar{p} - \bar{p}_{\mathcal{T}}\|_{L^\infty(\mathcal{N}_T^*)}^2 + (1 + h_T^{4-d}) \|\bar{y} - \bar{y}_{\mathcal{T}}\|_{L^2(\mathcal{N}_T^*)}^2 \\ &\quad + \sum_{S \in \mathcal{S}_T} \|\bar{u} - \bar{u}_{\mathcal{T}}\|_{H^{-2}(\mathcal{N}_S)}^2 + \text{osc}_{\mathcal{T}}^2(y_d; \mathcal{N}_T^*), \end{aligned}$$

where \mathcal{N}_T^* is defined in (4.5) and the hidden constant is independent of the optimal variables, their approximations, the size of the elements in the mesh \mathcal{T} , and its cardinality $\#\mathcal{T}$.

6. NUMERICAL EXAMPLES

In this section we conduct a series of numerical examples that illustrate the performance of the error estimator that we designed and analyzed in Section 5. In Section 6.3 below, we go beyond the presented analysis and perform a numerical experiment where the assumption that our domain is convex is violated. The presented numerical examples have been carried out with the help of a code that we implemented using MATLAB[®] (R2017a). When assembling all system matrices we have used exact integration. The right hand sides as well as the approximation errors are computed by a quadrature formula which is exact for polynomials of degree 19.

For a given partition \mathcal{T} , we seek $(\bar{y}_{\mathcal{T}}, \bar{p}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}) \in \mathbb{V}(\mathcal{T}) \times \mathbb{V}(\mathcal{T}) \times \mathbb{U}(\mathcal{T})$ that solves the discrete optimality system (3.9). We solve such a nonlinear system of equations by using the solution technique that was devised in [10, Section 6]. Such a

Algorithm 1: Adaptive semismooth Newton algorithm.**Input:** Initial mesh \mathcal{T}_0 , desired state y_d and sparsity parameter α .**Set:** $i = 0$.**Active set strategy:**

- 1 Compute $[\bar{y}_{\mathcal{T}_i}, \bar{p}_{\mathcal{T}_i}, \bar{u}_{\mathcal{T}_i}] = \text{Newton method}[\mathcal{T}_i, \alpha, y_d]$, which implements the semismooth Newton method of [10, Section 6].

A posteriori error estimation:

- 2 For each $T \in \mathcal{T}_i$ compute the local error indicator $\mathcal{E}_{\text{ocp}, T}^2$ given in (6.1).
- 3 Mark an element T for refinement if $\mathcal{E}_{\text{ocp}, T}^2 > \frac{1}{2} \max_{T' \in \mathcal{T}_i} \mathcal{E}_{\text{ocp}, T'}^2$.
- 4 From step 3, construct a new mesh, using a longest edge bisection algorithm. Set $i \leftarrow i + 1$, and go to step 1.

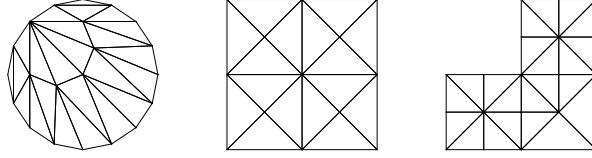


FIGURE 1. The initial meshes used when the domain Ω is a circle (Example 1), a square (Example 2) and a L-shaped domain (Example 3).

numerical scheme is based on an equivalent formulation of the optimality conditions (2.4) and (2.5):

$$\langle \bar{u}, \bar{p} - p \rangle \leq 0 \quad \forall \|p\|_{C_0(\Omega)} \leq \alpha.$$

In the discrete setting of Section 3 this variational inequality can be reformulated using a complementarity function as

$$\bar{u}_{\mathcal{T}} + \max\{0, -\bar{u}_{\mathcal{T}} + \bar{p}_{\mathcal{T}} - \alpha\} + \min\{0, -\bar{u}_{\mathcal{T}} + \bar{p}_{\mathcal{T}} + \alpha\} = 0,$$

which should be understood pointwise. Since the underlying map is locally Lipschitz, the reformulated discrete optimality system can be solved by a locally superlinearly convergent semismooth Newton method [27]. We refer the reader to [10, Section 6] for details. Once a discrete solution is obtained, on the basis of (5.9), we compute the error indicator

$$(6.1) \quad \mathcal{E}_{\text{ocp}, T} := \mathcal{E}_{\text{ocp}}(\bar{y}_{\mathcal{T}}, \bar{p}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; T) = (\mathcal{E}_y^2(\bar{y}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; T) + \mathcal{E}_p^2(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; T))^{\frac{1}{2}},$$

which is defined in terms of the local indicators given by (5.2) and (5.4), to drive the adaptive procedure described in Algorithm 1 and the global a posteriori error estimator \mathcal{E}_{ocp} in order to assess the accuracy of the approximation. For the numerical results, we define the total number of degrees of freedom $\text{Ndof} = 2 \dim(\mathbb{V}(\mathcal{T})) + \dim(\mathbb{U}(\mathcal{T}))$, and the errors $e_y := \bar{y} - \bar{y}_{\mathcal{T}}$ and $e_p := \bar{p} - \bar{p}_{\mathcal{T}}$. Since $\|\bar{u} - \bar{u}_{\mathcal{T}}\|_{H^{-2}(\Omega)}$ is not computable, we measure experimental rates of convergence for the error in the norm

$$\|(e_y, e_p)\|_{\Omega} := \left(\|e_y\|_{L^2(\Omega)}^2 + \|e_p\|_{L^{\infty}(\Omega)}^2 \right)^{\frac{1}{2}}.$$

The initial meshes for our numerical examples are shown in Figure 1.

We now provide three numerical experiments. First we consider a problem where an exact solution can be obtained: it is based on the solution constructed in [31, Section 8.1]. In the second example the exact solution is not known. Finally, in

the third example, we go beyond the presented analysis and perform a numerical experiment where we violate the assumption of the convexity of the domain.

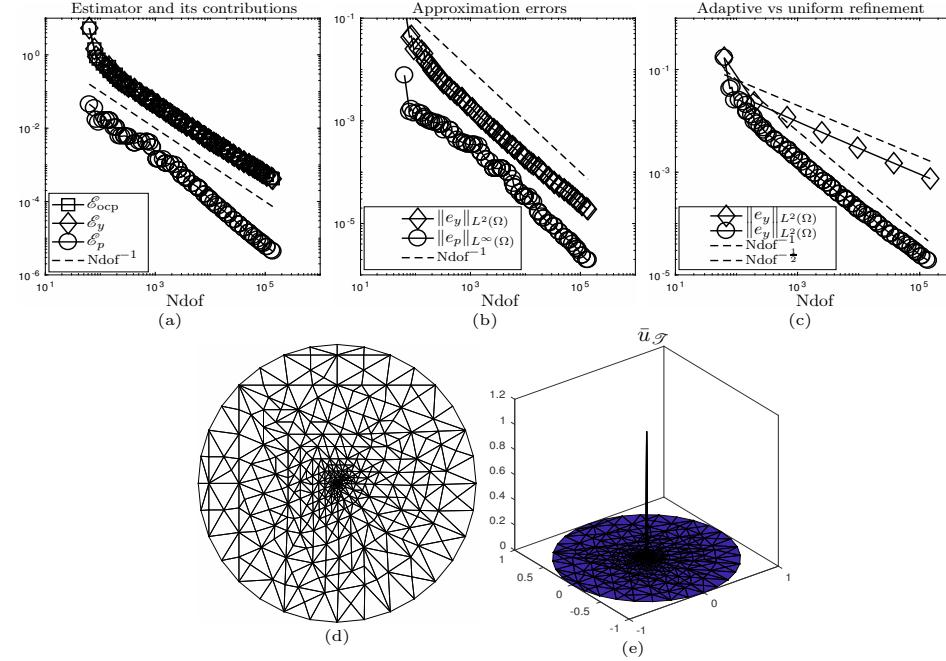


FIGURE 2. An example with known exact solution; see Section 6.1. (a) Experimental rates of convergence for $\mathcal{E}_{\text{ocp}}(\bar{y}_T, \bar{p}_T, \bar{u}_T; \mathcal{T})$, and each of his contributions. (b) Experimental rates of convergence for the errors e_y and e_p . (c) Experimental rates of convergence for e_y on uniform and adaptive refinement. (d) The 30th adaptively refined mesh. (e) The approximate discrete solution \bar{u}_T .

6.1. A problem with known exact solution. For this problem we set $\Omega = B_0(1) \subset \mathbb{R}^2$, that is, the unit two dimensional ball, and $\alpha = 10^{-2}$. Following [31, Section 8.1], the exact optimal state, adjoint state, and exact optimal control are given by

$$\bar{y} = -\frac{1}{2\pi} \ln(|x|), \quad \bar{p} = -0.02|x|^3 + 0.03|x|^2 - 0.01, \quad \bar{u} = \delta_0,$$

where for a vector $\xi \in \mathbb{R}^2$ we denote by $|\xi|$ its Euclidean norm. Note that the optimal state \bar{y} is simply a Green's function. On the basis of the solution, we compute the desired state via

$$y_d = \Delta \bar{p} + \bar{y} = -\alpha \frac{6(3|x|^2 - 2|x|)}{|x|} - \frac{1}{2\pi} \ln(|x|);$$

see [31, Section 8.1] for details. The results for this experiment are presented in Figure 2 and Table 1. In Figure 2 we show the experimental rates of convergence for the error estimator \mathcal{E}_{ocp} , as well as for each one of its contributions. Since, in this case, we have access to the exact solution, we also compute the decay rates of the errors $\|e_y\|_{L^2(\Omega)}$ and $\|e_p\|_{L^\infty(\Omega)}$. It can be observed that the AFEM described Algorithm 1 outperforms the FEM of [10] since it delivers optimal experimental rates of convergence when the latter cannot. We also present, in Figure 2, the

mesh obtained after 30 iterations of the designed AFEM and the discrete optimal variable obtained on such an adaptively refined mesh. We observe that most of the refinement occurs near $x = 0$ which corresponds to the support of $\bar{u} = \delta_0$. This attests to the efficiency of the devised estimator.

Ndofs	\mathcal{E}_{ocp}	\mathcal{E}_y	\mathcal{E}_p	$\ e_y\ _{L^2(\Omega)}$	$\ e_p\ _{L^\infty(\Omega)}$
111	5.66984e-01	5.66737e-01	1.67262e-02	2.55767e-02	1.46470e-03
1077	4.37467e-02	4.37232e-02	1.43511e-03	1.75347e-03	3.40616e-04
12801	3.66509e-03	3.66419e-03	8.12535e-05	1.50904e-04	2.82187e-05
33792	1.43789e-03	1.43769e-03	2.39289e-05	6.18704e-05	8.89136e-06
63876	8.38781e-04	8.38713e-04	1.07119e-05	3.76121e-05	4.70434e-06
91896	5.99521e-04	5.99480e-04	7.01987e-06	2.66131e-05	2.96516e-06
137298	4.17480e-04	4.17457e-04	4.39216e-06	1.89848e-05	1.96305e-06

TABLE 1. The total error estimator \mathcal{E}_{ocp} , its contributions \mathcal{E}_y and \mathcal{E}_p , and the errors $\|e_y\|_{L^2(\Omega)}$ and $\|e_p\|_{L^\infty(\Omega)}$ obtained when using the adaptive loop of Algorithm 1 to approximate the solution to the problem described in Section 6.1.

6.2. An example with no analytical solution. We now explore the performance of our devised a posteriori error estimators in problems where no analytical solution is available. Set $\Omega = (-1, 1)^2$ and for $(x_1, x_2) \in \Omega$:

$$(6.2) \quad y_d = 10 \left(\exp(-50\{(x_1 - 0.2)^2 + (x_2 + 0.1)^2\}) - \exp(-50\{(x_1 + 0.1)^2 + (x_2 - 0.2)^2\}) \right).$$

The purpose of this example is to investigate the effect of varying the sparsity parameter α . We consider

$$\alpha \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\}.$$

The results of this experiment are presented in Figure 3. We observe, in subfigures (a)–(f), that optimal experimental rates of convergence in terms of approximation are obtained for the error estimator \mathcal{E}_{ocp} and its individual contributions \mathcal{E}_y and \mathcal{E}_p for all the values of the parameter α that we considered. For $\alpha = 10^{-4}$, we display the adaptive mesh obtained after 25 iterations of the designed AFEM and the obtained discrete optimal control variable.

6.3. An example on a nonconvex domain. In this section we go beyond the presented analysis and perform a numerical experiment where the domain convexity assumption is violated. We let $\Omega = (-1, 1)^2 \setminus (-1, 0] \times [0, 1)$, i.e., a L shaped domain and $\alpha = 5 \cdot 10^{-3}$. The desired state y_d is given by

$$y_d = -\log(\sqrt{(x_1 - 0.2)^2 + (x_2 + 0.2)^2}), \quad (x_1, x_2) \in \Omega.$$

The results are shown in Figure 4 where we observe optimal experimental rates of convergence for the proposed a posteriori error estimator \mathcal{E}_{ocp} and each one of its contributions. We also observe that the refinement is being concentrated about the singularity exhibited by the optimal control and to a lesser extent the re-entrant corner. To appreciate the nature of the aforementioned singularity, we also display the computed optimal control variable.

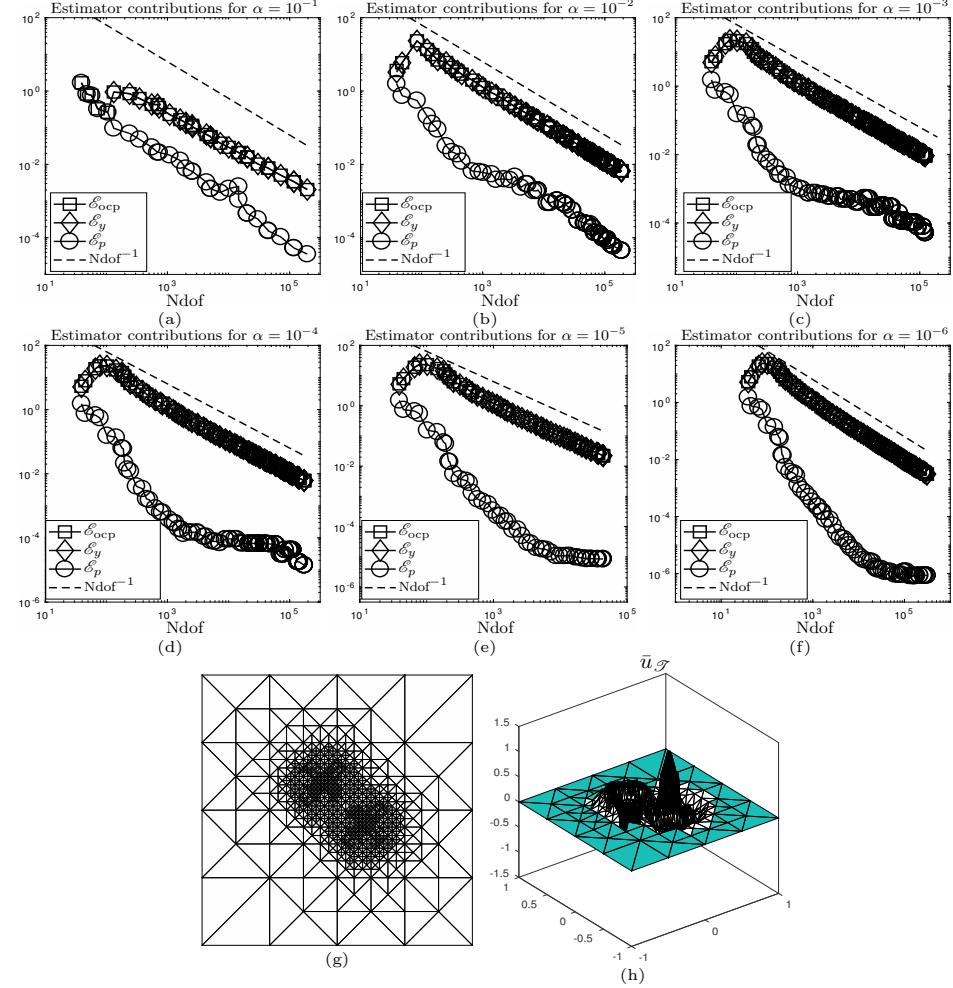


FIGURE 3. Results for the example of Section 6.2. (a)–(f) Experimental rates of convergence for $\mathcal{E}_{\text{ocp}}(\bar{y}_{\mathcal{T}}, \bar{p}_{\mathcal{T}}, \bar{u}_{\mathcal{T}}; \mathcal{T})$ and its contributions for $\alpha \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\}$, respectively. (g) The 25th adaptively refined mesh for $\alpha = 10^{-4}$. (h) The discrete control $\bar{u}_{\mathcal{T}}$ on the 25th adaptively refined mesh for $\alpha = 10^{-4}$.

6.4. Conclusions. From the presented numerical examples several general conclusions can be drawn:

- The error estimator \mathcal{E}_{ocp} , as well as each of his contributions, exhibit optimal experimental rates of convergence for the experiments that we perform.
- Most of the refinement occurs near the singularity points, which attests to the efficiency of the devised a posteriori error estimator.
- The contribution $\mathcal{E}_y(\bar{p}_{\mathcal{T}}, \bar{y}_{\mathcal{T}}; \mathcal{T})$ to the error estimator is the dominating one. We believe that this shows the very singular nature of the problem that defines the state variable.
- The third example shows that, even in the presence of a nonconvex domain, the error estimator \mathcal{E}_{ocp} exhibits optimal experimental rates of convergence with respect to approximation.

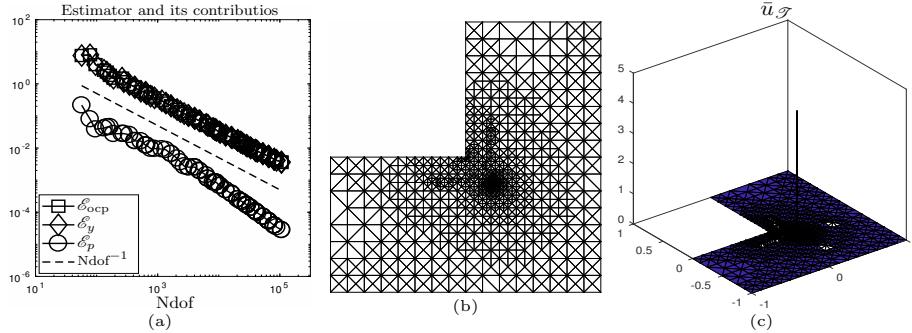


FIGURE 4. Results for the example of section 6.3. (a) Experimental rates of convergence for $\mathcal{E}_{\text{cp}}(\bar{y}_T, \bar{p}_T, \bar{u}_T; \mathcal{T})$, and each of his contributions. (b) The 25th adaptively refined mesh. (c) The discrete control \bar{u}_T on the 25th adaptively refined mesh.

- In spite of the very singular nature of the solution to the state equation, our proposed estimator is able to deliver optimal rates of convergence.

ACKNOWLEDGEMENT

The authors would like to thank Cesare Molinari (UTFSM) for insightful discussions.

REFERENCES

- [1] J. AGNELLI, E. GARAU, AND P. MORIN, *A posteriori error estimates for elliptic problems with Dirac measure terms in weighted spaces*, ESAIM Math. Model. Numer. Anal., 48 (2014), pp. 1557–1581.
- [2] A. ALLENDES, E. OTÁROLA, R. RANKIN, AND A. J. SALGADO, *An a posteriori error analysis for an optimal control problem with dirac measures*, ESAIM Math. Model. Numer. Anal., (2018). DOI: <https://doi.org/10.1051/m2an/m2an170055>.
- [3] R. ARAYA, E. BEHRENS, AND R. RODRÍGUEZ, *A posteriori error estimates for elliptic problems with Dirac delta source terms*, Numer. Math., 105 (2006), pp. 193–216.
- [4] F. CAMACHO AND A. DEMLOW, *L_2 and pointwise a posteriori error estimates for FEM for elliptic PDEs on surfaces*, IMA J. Numer. Anal., 35 (2015), pp. 1199–1227.
- [5] E. CASAS, *A review on sparse solutions in optimal control of partial differential equations*, SeMA J., 74 (2017), pp. 319–344.
- [6] E. CASAS, R. HERZOG, AND G. WACHSMUTH, *Analysis of spatio-temporally sparse optimal control problems of semilinear parabolic equations*, ESAIM Control Optim. Calc. Var., 23 (2017), pp. 263–295.
- [7] 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.
- [8] E. CASAS AND K. KUNISCH, *Stabilization by sparse controls for a class of semilinear parabolic equations*, SIAM J. Control Optim., 55 (2017), pp. 512–532.
- [9] 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.
- [10] E. CASAS, C. CLASON, AND K. KUNISCH, *Approximation of elliptic control problems in measure spaces with sparse solutions*, SIAM J. Control Optim., 50 (2012), pp. 1735–1752.
- [11] E. CASAS, *Control of an elliptic problem with pointwise state constraints*, SIAM J. Control Optim., 24 (1986), pp. 1309–1318.
- [12] E. CASAS, M. MATEOS, AND A. RÖSCH, *Improved approximation rates for a parabolic control problem with an objective promoting directional sparsity*, Comput. Optim. Appl., 70 (2018), pp. 239–266.

- [13] P. G. CIARLET, *The finite element method for elliptic problems*, SIAM, Philadelphia, PA, 2002.
- [14] C. CLASON AND K. KUNISCH, *A measure space approach to optimal source placement*, Comput. Optim. Appl., 53 (2012), pp. 155–171.
- [15] C. CLASON, B. KALTENBACHER, AND D. WACHSMUTH, *Functional error estimators for the adaptive discretization of inverse problems*, Inverse Problems, 32 (2016), pp. 104004, 25.
- [16] E. CASAS, R. HERZOG, AND G. WACHSMUTH, *Approximation of sparse controls in semilinear elliptic equations*, in Large-scale scientific computing, vol. 7116 of Lecture Notes in Comput. Sci., Springer, Heidelberg, 2012, pp. 16–27.
- [17] E. CASAS, M. MATEOS, AND A. RÖSCH, *Finite element approximation of sparse parabolic control problems*, Math. Control Relat. Fields, 7 (2017), pp. 393–417.
- [18] C. CLASON AND K. KUNISCH, *A duality-based approach to elliptic control problems in non-reflexive Banach spaces*, ESAIM Control Optim. Calc. Var., 17 (2011), pp. 243–266.
- [19] E. DARI, R. G. DURÁN, AND C. PADRA, *Maximum norm error estimators for three-dimensional elliptic problems*, SIAM J. Numer. Anal., 37 (2000), pp. 683–700.
- [20] A. DEMLOW AND E. H. GEORGULIS, *Pointwise a posteriori error control for discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 50 (2012), pp. 2159–2181.
- [21] A. DEMLOW AND N. KOPTEVA, *Maximum-norm a posteriori error estimates for singularly perturbed elliptic reaction-diffusion problems*, Numer. Math., 133 (2016), pp. 707–742.
- [22] A. DEMLOW AND R. STEVENSON, *Convergence and quasi-optimality of an adaptive finite element method for controlling L_2 errors*, Numer. Math., 117 (2011), pp. 185–218.
- [23] K. ERIKSSON, *An adaptive finite element method with efficient maximum norm error control for elliptic problems*, Math. Models Methods Appl. Sci., 4 (1994), pp. 313–329.
- [24] A. ERN AND J.-L. GUERMOND, *Theory and practice of finite elements*, vol. 159 of Applied Mathematical Sciences, Springer-Verlag, New York, 2004.
- [25] F. D. GASPOZ, P. MORIN, AND A. VEESER, *A posteriori error estimates with point sources in fractional Sobolev spaces*, Numer. Methods Partial Differential Equations, 33 (2017), pp. 1018–1042.
- [26] M. HINZE, *A variational discretization concept in control constrained optimization: the linear-quadratic case*, Comput. Optim. Appl., 30 (2005), pp. 45–61.
- [27] B. KUMMER, *Newton's method based on generalized derivatives for nonsmooth functions: convergence analysis*, in Advances in optimization (Lambrecht, 1991), vol. 382 of Lecture Notes in Econom. and Math. Systems, Springer, Berlin, 1992, pp. 171–194.
- [28] R. NOCHETTO, *Pointwise a posteriori error estimates for elliptic problems on highly graded meshes*, Math. Comp., 64 (1995), pp. 1–22.
- [29] R. H. NOCHETTO, A. SCHMIDT, K. G. SIEBERT, AND A. VEESER, *Pointwise a posteriori error estimates for monotone semi-linear equations*, Numer. Math., 104 (2006), pp. 515–538.
- [30] R. H. NOCHETTO, K. G. SIEBERT, AND A. VEESER, *Pointwise a posteriori error control for elliptic obstacle problems*, Numer. Math., 95 (2003), pp. 163–195.
- [31] K. PIEPER AND B. VEXLER, *A priori error analysis for discretization of sparse elliptic optimal control problems in measure space*, SIAM J. Control Optim., 51 (2013), pp. 2788–2808.
- [32] 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.
- [33] G. STAMPACCHIA, *Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus*, Ann. Inst. Fourier (Grenoble), 15 (1965), pp. 189–258.
- [34] R. VERFÜRTH, *A posteriori error estimators for the Stokes equations*, Numer. Math., 55 (1989), pp. 309–325.
- [35] R. VERFÜRTH, *A posteriori error estimators for convection-diffusion equations*, Numer. Math., 80 (1998), pp. 641–663.
- [36] R. VERFÜRTH, *A posteriori error estimation techniques for finite element methods*, Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford, 2013.
- [37] 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.

(F. Fuica) DEPARTAMENTO DE MATEMÁTICA, UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA,
VALPARAÍSO, CHILE.

E-mail address: francisco.fuica@sansano.usm.cl

(E. Otárola) DEPARTAMENTO DE MATEMÁTICA, UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA,
VALPARAÍSO, CHILE.

E-mail address: enrique.otorola@usm.cl

(A.J. Salgado) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TENNESSEE, KNOXVILLE, TN
37996, USA.

E-mail address: asalgad1@utk.edu